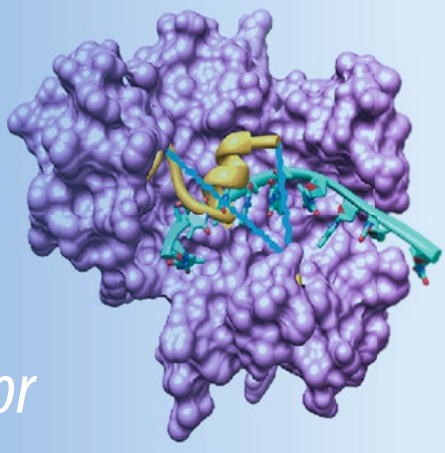
Methods in Molecular Biology 2444

Springer Protocols



Nima Mosammaparast *Editor*

DNA Damage Responses

Methods and Protocols

Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor
John M. Walker
School of Life and Medical Sciences
University of Hertfordshire
Hatfield, Hertfordshire, UK

For further volumes:
http://www.springer.com/series/7651

For over 35 years, biological scientists have come to rely on the research protocols and methodologies in the critically acclaimed *Methods in Molecular Biology* series. The series was the first to introduce the step-by-step protocols approach that has become the standard in all biomedical protocol publishing. Each protocol is provided in readily-reproducible step-by-step fashion, opening with an introductory overview, a list of the materials and reagents needed to complete the experiment, and followed by a detailed procedure that is supported with a helpful notes section offering tips and tricks of the trade as well as troubleshooting advice. These hallmark features were introduced by series editor Dr. John Walker and constitute the key ingredient in each and every volume of the *Methods in Molecular Biology* series. Tested and trusted, comprehensive and reliable, all protocols from the series are indexed in PubMed.

DNA Damage Responses

Methods and Protocols

Edited by

Nima Mosammaparast

Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA

Humana Press

Editor
Nima Mosammaparast
Department of Pathology
and Immunology
Washington University in St. Louis School of Medicine
St. Louis, MO, USA

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-0716-2062-5 ISBN 978-1-0716-2063-2 (eBook)
https://doi.org/10.1007/978-1-0716-2063-2

This Humana imprint is published by the registered company Springer Science+Business Media, LLC part of Springer Nature.
The registered company address is: 1 New York Plaza, New York, NY 10004, U.S.A.

Preface

For those of us studying nucleic acid damage responses, the field has been extraordinarily exciting in recent years. From expanding our understanding of the fundamental chemistry and biochemistry of nucleic acid modifications, to uncovering the precise mechanisms of DNA damage signaling and repair, to the clear connections these processes have to DNA replication and transcription, not to mention affecting human health, the field continues to provide ample opportunities for those interested in a wide berth of both basic and translational biomedical sciences. At the same time, despite the fact that these positive motivators make it an appealing area to study, for nonexperts to move into this field can be a daunting task. There is a tendency to believe that DNA repair is "hyper-specialized" and that the assays used are not accessible to individuals who did not do a significant portion of their training in this field.

This edition of *Methods in Molecular Biology* is a testament that the above is simply not true. While many contributors here have been studying DNA repair for a long time, many of us, myself included, started focusing on DNA repair towards the beginning of their independent careers and had to learn and adapt these methods from our colleagues or from the literature. This reflects the fact that much of this field relies on more general established methods, which are then modified as needed to reflect the question at hand related to DNA damage repair.

While far from exhaustive, the purpose of these *Methods in Molecular Biology* chapters is to sample some of the key approaches used to study DNA repair, with the occasional "dabble" into DNA replication, which is intimately associated with many repair processes. We begin with more systems-based approaches to identify factors and pathways that are involved in the repair process and their links to cancer (Chapters 1 and 2). Subsequently, we then move to proteomics and biophysical approaches that have been used successfully to understand the function and mechanisms of proteins involved in the repair process (Chapters 3 and 4). More specialized approaches to study physiological DNA repair in immune cells are covered in Chapter 5. In Chapters 6 and 7, methods related to the analysis of DNA replication dynamics are described. Following this, specific pathways and systems used to monitor nucleic acid base damage and modifications are discussed in Chapters 8–10. We end on several chapters related to the biochemical reconstitution of several key pathways involved in DNA double-stranded break repair and DNA damage signaling. Even the more specialized chapters will hopefully yield insights to improve established methodologies in a modern molecular biology laboratory.

I would like to sincerely thank all of the authors here who took much of their time and contributed to this edition. It is their generosity and collegiality that make this field of study so fun.

Saint Louis, MO, USA — *Nima Mosammaparast*
June 2021

Contents

Contributors

Dauren Alimbetov • *Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA*

Rita Anoh • *Department of Science, Mount St. Mary's University, Emmitsburg, MD, USA*

Albino Bacolla • *Departments of Cancer Biology and of Molecular and Cellular Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA*

Ryan P. Barnes • *Department of Environmental and Occupational Health, University of Pittsburgh Graduate School of Public Health, and UPMC Hillman Cancer Center, Pittsburgh, PA, USA*

Jeffrey J. Bednarski • *Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA*

Kate A. Burke • *Department of Science, Mount St. Mary's University, Emmitsburg, MD, USA*

Kathryn H. Burnett • *Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

Bo-Ruei Chen • *Division of Hematology and Oncology, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA; O'Neal Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL, USA*

Naga Babu Chinnam • *Department of Molecular and Cellular Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA*

Rajashree A. Deshpande • *The Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX, USA*

James M. Dewar • *Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN, USA*

Rachel DeWeerd • *Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA*

Elise Fouquerel • *Department of Biochemistry and Molecular Biology, Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA*

Adit Ganguly • *Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA*

Abby M. Green • *Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA; Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA*

Aye Su Hlaing • *Biological Systems and Engineering Division, Department of BioEngineering & BioMedical Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

Greg L. Hura • *Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA, USA; Chemistry and Biochemistry Department, University of California Santa Cruz, Santa Cruz, CA, USA*

Jessica Jackson • *Division of Oncology, Department of Medicine, Washington University in St. Louis, St. Louis, MO, USA*

Rachel Johnston • *Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA*

Tamar Kavlashvili • *Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, TN, USA*
Linda Lee • *Department of Biochemistry and Molecular Biology, Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada*
Susan P. Lees-Miller • *Department of Biochemistry and Molecular Biology, Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada*
Judy G. Lin • *Biological Systems and Engineering Division, Department of BioEngineering & BioMedical Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
Patrick M. Lombardi • *Department of Science, Mount St. Mary's University, Emmitsburg, MD, USA*
Su S. Maw • *Biological Systems and Engineering Division, Department of BioEngineering & BioMedical Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
Nima Mosammaparast • *Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA*
Patricia L. Opresko • *Department of Environmental and Occupational Health, University of Pittsburgh Graduate School of Public Health, and UPMC Hillman Cancer Center, Pittsburgh, PA, USA*
John M. Pascal • *Department of Biochemistry and Molecular Medicine, Université de Montréal, Montreal, QC, Canada*
Tanya T. Paull • *The Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX, USA*
Ishtiaque Rashid • *Departments of Internal Medicine, Molecular Genetics and Microbiology and the University of New Mexico Comprehensive Cancer Center, University of New Mexico, Albuquerque, NM, USA*
Rebecca Rodell • *Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA*
Cody M. Rogers • *Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA*
Daniel J. Rosenberg • *Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA, USA; Graduate Group in Biophysics, University of California, Berkeley, Berkeley, CA, USA*
Dhane P. Schmelyun • *Department of Science, Mount St. Mary's University, Emmitsburg, MD, USA*
Brian Shih • *Biological Systems and Engineering Division, Department of BioEngineering & BioMedical Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
Barry P. Sleckman • *Division of Hematology and Oncology, Department of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA; O'Neal Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, AL, USA*
Jennifer M. Soll • *Department of Pathology & Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA*
Aleksandr Sverzhinsky • *Department of Biochemistry and Molecular Medicine, Université de Montréal, Montreal, QC, Canada*
Aleem Syed • *Department of Molecular and Cellular Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX, USA*
John A. Tainer • *Department of Molecular and Cellular Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA; Molecular Biophysics and Integrated*

Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA, USA; Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

Sanjana A. Thosar • *Department of Environmental and Occupational Health, University of Pittsburgh Graduate School of Public Health, and UPMC Hillman Cancer Center, Pittsburgh, PA, USA*

Aye C. Thwin • *Biological Systems and Engineering Division, Department of BioEngineering & BioMedical Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

Alan E. Tomkinson • *Departments of Internal Medicine, Molecular Genetics and Microbiology and the University of New Mexico Comprehensive Cancer Center, University of New Mexico, Albuquerque, NM, USA*

Brittany A. Townley • *Department of Pathology & Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA*

Miaw-Sheue Tsai • *Biological Systems and Engineering Division, Department of BioEngineering & BioMedical Sciences, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

Ning Tsao • *Department of Pathology and Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO, USA*

Susan E. Tsutakawa • *Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

Alessandro Vindigni • *Division of Oncology, Department of Medicine, Washington University in St. Louis, St. Louis, MO, USA*

Meiling Wang • *Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA*

Lynn S. White • *Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA*

Yaping Yu • *Department of Biochemistry and Molecular Biology, Charbonneau Cancer Institute, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada*

Weixing Zhao • *Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA*

Chapter 1

Robust Computational Approaches to Defining Insights on the Interface of DNA Repair with Replication and Transcription in Cancer

Albino Bacolla and John A. Tainer

Abstract

The massive amount of experimental DNA and RNA sequence information provides an encyclopedia for cell biology that requires computational tools for efficient interpretation. The ability to write and apply simple computing scripts propels the investigator beyond the boundaries of online analysis tools to more broadly interrogate laboratory experimental data and to integrate them with all available datasets to test and challenge hypotheses. Here we describe robust prototypic bash and C++ scripts with metrics and methods for validation that we have made publicly available to address the roles of non-B DNA-forming motifs in eliciting genetic instability and to query The Cancer Genome Atlas. Importantly, the methods presented provide practical data interpretation tools to examine fundamental relationships and to enable insights and correlations between alterations in gene expression patterns and patient outcome. The exemplary source codes described are simple and can be efficiently modified, elaborated, and applied to other relationships and areas of investigation.

Key words Non-B DNA, Cancer genome, TCGA analyses, Parallel computing, Bash, Kaplan-Meier survival, Tumor normal pair, Gene expression correlation analysis, Custom scripts

1 Introduction

From searching for reagents to analyzing data, computers have become an integral part of a molecular biology laboratory, and the integration of "wet" laboratory data with bioinformatics and metadata analyses has become a powerful means for casting experimental information into a broader spectrum of knowledge to test, challenge, and validate novel paradigms [1–4]. The massive amount of experimental data generated by The Cancer Genome Atlas effort may in practical terms only be harnessed through computational means [5, 6]. Likewise, both the processing and biological interpretation of next-generation sequencing and RNA-Seq data are achieved through computational programming [7–10]. The

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_1,

widespread use of computing resources is also generating an increasing need for the molecular biology laboratory not only to acquire the skills necessary for installing and running off-the-shelf software or use online analytical tools but also to devise ad hoc computational scripts to interrogate database information in a manner suited to specific needs that emerge during laboratory projects. Exemplary evidence for this need was captured by a recent seminar, where the speaker commented that "an investigator who knows even just the basics of a computer language owns the data": this insight makes a qualitative distinction between the "button pusher" and the active interrogator and interpreter of the data that are being queried. Indeed, it is often of clear advantage to be able to write even simple codes to address specific questions or to process large data files in batch mode, so as to avoid manual mistakes and ensure reproducibility.

Here we present few enabling robust scripts in bash and C++ tailored to support the molecular biologist to address two types of questions: (1) the roles of non-B DNA in genomic instability and (2) RNA-Seq analyses of TCGA data, which we have used to help elucidate the interface between DNA repair (and its deficiencies) and DNA replication and transcription. Although distinct, these two areas of investigation are intimately interconnected, since DNA secondary structures are arising as powerful partners with the DNA repair arsenal to thrive a normal cell towards a malignant state [11–14]. There are many helpful tools available to search for non-B DNA-forming sequences [15–19] and to explore TCGA data online, including among others cBioPortal (https://www.cbioportal.org), the Xena browser (https://xena.ucsc.edu), Mexpress (https://mexpress.be), and TCGA Wanderer (http://maplab.imppc.org/wanderer/). The source codes presented here are available at https://github.com/abacolla. They are not a substitute for other available resources; rather, their goal is to stimulate and enable investigator interest in using and writing ad hoc codes to adapt the output of a search to specific needs, in other words, to "own," interrogate, and integrate experimental laboratory data with public datasets.

2 Materials

We assume a basic knowledge of command line syntax and, preferably, access to a High Performance Computing (HPC) cluster. The intent of this chapter is to provide access to scripts and selected sample files on which to run the scripts. A flow-through of the steps described here may also be found in the README.md files at the GitHub site. The scripts used in this article are available at https://github.com/abacolla. Files containing the TCGA RNA-Seq normalized Rsem, and patient clinical data may be downloaded from

https://www.researchgate.net/profile/Albino_Bacolla, under the TCGA Analyses project. These files have been obtained using the TCGA Assembler-2 (https://github.com/compgenome365/TCGA-Assembler-2). Other utilities include twoBitToFa from http://hgdownload.cse.ucsc.edu/admin/exe and the human reference genome sequence contained in the hg38.2bit file, from http://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/.

3 Methods

3.1 Compile C++ Programs

1. Inspect the Makefile associated with each cpp file, edit if needed, and compile running "make." This will generate a "build" directory containing the executable (*see* **Note 1**).

3.2 Search for Non-B DNA-Forming Sequences in Fasta Files

This method aids at determining whether non-B DNA-forming repeats may be enriched at particular genomic loci. It searches for non-B DNA-forming motifs in fasta files and reports their total number. It also returns the distance between the center of each non-B DNA-forming motif and the middle position of the DNA fasta sequence, which is meant to assess whether non-B DNA-forming motifs occur more often near junctions (breakpoints) of genomic rearrangements, such as deletions, duplications, inversions, and translocations than expected by chance [14, 20–22]. Scripts are contained in directories non-B-DB and submitMpi at https://github.com/abacolla.

1. Generate file fastaList.fa containing a list of 31 fasta records on which to perform a search of non-B DNA-forming motifs by executing "twoBitToFa -noMask -seqList=fastaList hg38.2bit fastaList.fa" (*see* **Note 2**).
2. Run each non-B DNA search script on fastaList.fa, i.e., "./dr_get.sh fastaList.fa," to search for direct repeats (dr), "./h_get.sh fastaList.fa" to search for triplex-forming repeats (h), "./g4_get.sh fastaList.fa" to retrieve G4-forming repeats (g4), "./z_get.sh fastaList.fa" to obtain Z DNA-forming repeats (z), and "./ir_get.sh fastaList.fa" for inverted repeats, which can form cruciforms (ir). Running the searches on a single processor is inefficient; to speed up the process, use parallel computing, as follows.
3. Split fastaList.fa into individual fasta records by executing the "csplit" command reported in Step 2 of the README.md file at https://github.com/abacolla/nonB-DNA; delete file bin_000 ("rm bin_000"). This generates 31 fasta files named bin_001 to bin_031.
4. Make bash files to process bin_001 to bin_031 for all non-B DNA search scripts by executing "./makeFile.sh dr";

"./makeFile.sh z"; "./makeFile.sh g4"; "./makeFile.sh h"; and "./makeFile.sh ir." This generates 5 sets of 31 bash files each (drdna_1.sh to drdna_31.sh; zdna_1.sh to zdna_31.sh; g4dna_1.sh to g4dna_31.sh; hdna_1.sh to hdna_31.sh; and irdna_1.sh to irdna_31.sh).

5. Launch the job using the directive "ibrun (or mpirun depending on HPC instructions) vga_submitMpiJob drdna_," where drdna_ is the prefix of the 31 bash files generated in **step 4** (see launchJob.sh for an example template) (*see* **Note 3**). The time required to complete a job will depend on the type of script used, the total number of tasks (fasta files to process), and the number of processors that have been requested per node (Fig. 1).
6. Process the output. The output consists of a number of files containing chromosome number, hg38 coordinate, length, distance from start and end of the sequence to the center of the fasta sequence, and DNA sequence (for dr, g4, z) and chromosome number, hg38 coordinate, length of stem, length of loop, distance from the center loop to the center of the fasta sequence, and the sequence of both stems (for h, ir). Use nonB_getRes.sh to extract the number of tracts. Its usage is: nonB_getRes.sh file_suffix(dr | ir | q1k | z1k | triplex); i.e., "nonB_getRes.sh triplex." A comparison of all results should show that the number of inverted repeats is significantly higher in bin_22 than in the other bins. For studies aimed at assessing whether non-B DNA-forming repeats are enriched at rearrangement junctions, such as in cancer genomes, plot the distance of the tracts (or all the non-B DNA-forming bases) from the center positions of the fasta files (Fig. 2).

3.3 Assess TCGA Gene Expression Levels Between Tumor and Normal Controls

RNA-Seq gene expression data for the TCGA repository are widely available, and several online tools exist to visualize the data for a specific gene of interest. However, it is often desirable to create custom-made graphs where one can control aesthetic features for publication-quality figures. In addition, it may be necessary to compare the data for large sets of genes, and this can be accomplished most easily from custom-generated flat text files. Here we present vga_makeBoxPlotRsem.sh, a bash script that accomplishes these tasks easily. Scripts are contained in the directory tcgaAnalyses at https://github.com/abacolla.

1. vga_makeBoxPlotRsem.sh is a script that generates a high-quality png box plot using mRNA expression of a given gene for 15 TCGA tumor and normal matched controls suitable for publication with minimal editing. The number of tumor/normal pairs is limited to those cancer sets with at least ten normal controls. Its usage is: vga_makeBoxPlotRsem.sh

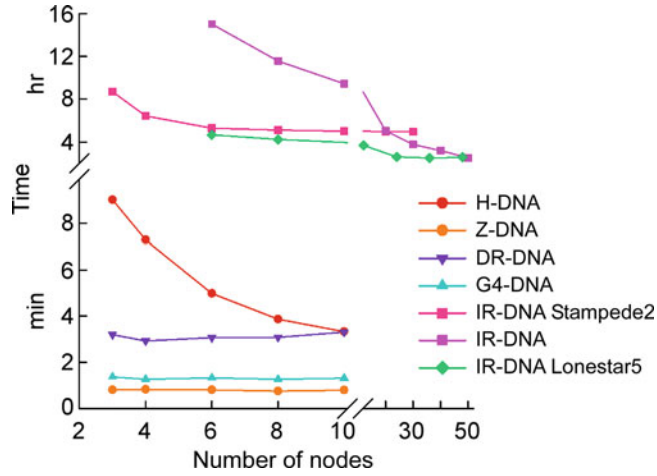


Fig. 1 HPC performance for processing non-B DNA-related batch jobs. A total of 125 individual fasta files, 10 kb of DNA sequence each, were used to search for non-B DNA-forming repeats using vga_submitMpi on variable number of nodes. Except where indicated, jobs were run on Bridges2 (RM nodes, 64 cores/node) at the Pittsburgh Supercomputing Center, Pittsburgh, PA. Stampede2 (Knights Landing compute nodes, 68 cores/node) and Lonestar5 (24 cores/node) were from the Texas Advanced Computing Center, Austin, TX

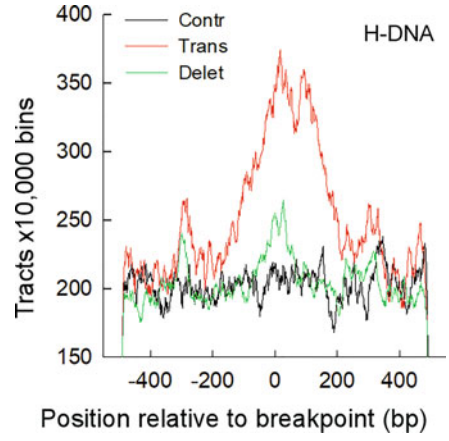


Fig. 2 Triplex DNA-forming repeats are enriched at translocation junctions in cancer genomes. Line plots of number of triplex-forming repeats flanking the junction breakpoints of translocations, deletions, and control sets from the Catalogue Of Somatic Mutations In Cancer (COSMIC, https:/cancer.sanger.ac.uk/cosmic). (Reproduced from [14] by permission of Oxford University Press)

<geneName>, where <geneName> is an official gene name in capital letters, i.e., vga_makeBoxPlotRsem.sh BLM. Before running the script, edit lines 8–11 to load any module required for the R language and edit DIR0 on line 13 to point to the RNA gene expression files. A copy of these files may be found in

directory TCGA Analyses at https://www.researchgate.net/profile/Albino_Bacolla, which were obtained using the TCGA-Assembler v.2.0 (https://github.com/compgenome365/TCGA-Assembler-2). Box plots are drawn according to the list on lines 63–77; to change the ranking, such as plotting according to p-values, change the order of tumor/normal pairs on lines 63–77.

2. vga_makeBoxPlotRsem.sh calls automatically the R program vga_pngBoxPlotRsem.R. Options in vga_pngBoxPlotRsem.R that control main aesthetic features include the y-axis range on line 52 (ylim), p-values (on, off) on line 53 (stats_compare_means), colors for the plots on line 73 (scale_fill_manual), the x-axis line (axis.line.x) on line 71, and notch (true, false) on line 46.
3. vga_makeBoxPlotRsem.sh generates two files: a box plot and a text file named after the input file, i.e., "boxPlotGeneExprBLM.png" and "pngBoxPlotBLM.out," containing the statistical data (*see* **Note 4**). For *BLM*, which encodes the BLM helicase with functions in DNA replication and repair and whose deficiency is associated with the autosomal recessive Bloom syndrome, the data show that the gene is significantly upregulated in all types of cancer (Fig. 3) (*see* **Note 5**).
4. vga_makeBoxPlotRsem.sh can be scaled up using vga_submitMpiJob, which is detailed in directory submitMpi and in Subheading 3.1, and the associated .out files, which can be queried in batch mode to assess statistical trends.

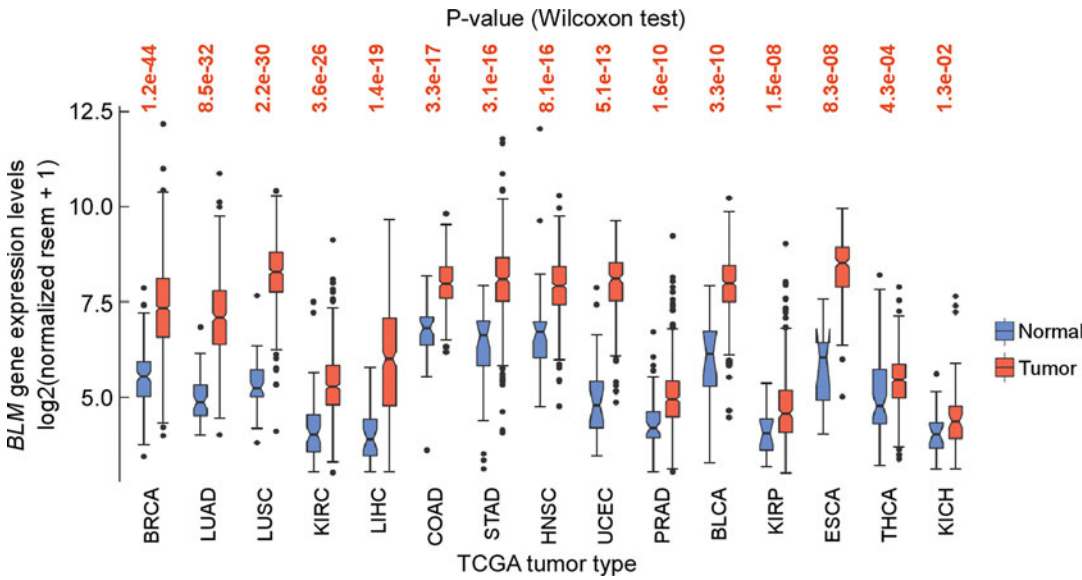


Fig. 3 Box plot of TCGA *BLM* mRNA levels between tumor and matched controls. The core plot generated with vga_makeBoxPlotRsem.sh was edited with Canvas (https:/www.canvasgfx.com)

3.4 Kaplan-Meier Survival Curves in TCGA Patients with High and Low mRNA Levels

It can be critically enabling to create custom graphs for Kaplan-Meier survival curve analyses of gene expression and to use the associated flat text files for large-scale analyses. Script "vga_survivalCurve.sh" in the https://github.com/abacolla/tcgaAnalyses repository serves this purpose.

1. vga_survivalCurve.sh generates a Kaplan-Meier survival curve comparing patients with high (above mean) versus low (below mean) expression levels for a given gene. Its usage is: vga_survivalCurve.sh <TCGA_TUMOR> <GENE_NAME>, where TCGA_TUMOR is the TCGA tumor code and GENE_NAME an official gene name, both in capital letters: i.e., vga_survivalCurve.sh KIRC ERCC1. Edit lines 8–11 to load any module required in R. Line 19 launches the vga_spotLight binary; its path needs to be specified. The --optFdat option points to the TCGA gene expression files; edit the path. Edit line 21 to point to the TCGA clinical data files, which can be extracted from https://www.researchgate.net/profile/Albino_Bacolla. Line 32 calls vga_survival.R; verify its path. The example above will generate a graphic file named kirc_ercc1.png and a text file named survival_ercc1_kirc.out.
2. Use vga_submitMpiJob, which is detailed in directory https://github.com/abacolla/submitMpi, to scale up vga_survivalCurve.sh to process more (or all) genes for the 33 TCGA tumor types, and use the associated .out files for statistical comparisons [23] (Fig. 4).

3.5 Gene Correlation Expression Analyses (GCEA)

The program vga_geneExprMain.cpp at https://github.com/abacolla/tcgaAnalyses generates the vga_spotLight binary, a utility for performing gene expression correlation analyses (GCEA) and other analyses using TCGA RNA-Seq gene expression data.

1. Edit Makefile to point to the BOOST library and preload any module required for MPI. Edit lines 96 and 97 of vga_geneExprUsage.hpp to point to the directories containing the gene expression and mutation data (files for the mutation data are not included and are not required for gene expression analyses). File testStart.sh may be used as a guide to test the vga_spotLight compiled binary.
2. Use option A to find a correlation between the expression of two genes. The usage is: vga_spotLight --optAdat <dataset> --optAgene1 <GENE1> --optAgene2 <GENE2>, where <dataset> is the TCGA gene expression file and <GENE1> and <GENE2> are official gene names. For example, launching "ibrun -n 1 vga_spotLight --optAdat ACC__geneExprT.txt --optAgene1 GRB2 --optAgene2 FGFR2" will generate an output file named "ACC_GRB2_FGFR2_expr.txt" containing

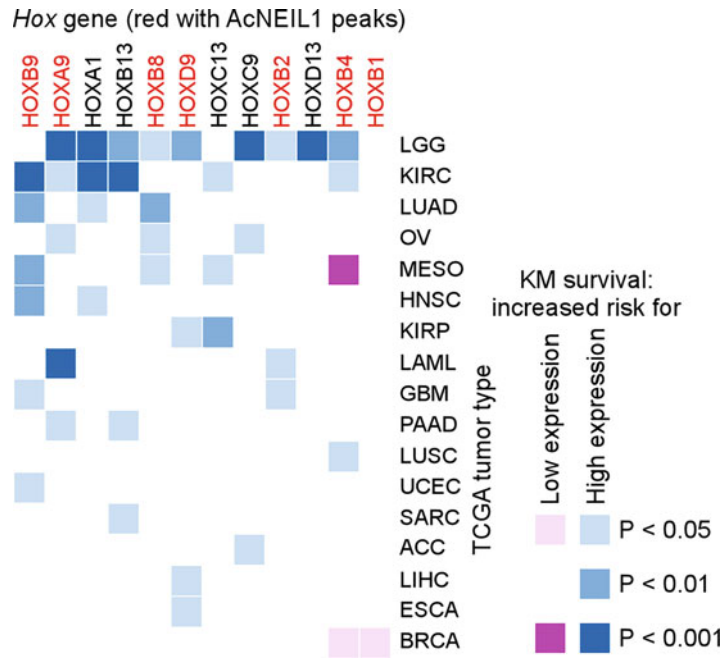


Fig. 4 Heat map of the effect of *HOX* gene expression on survival in TCGA patients. In adult tissues *HOX* genes display tissue-restricted expression, and hyper- or re-activation of *HOX* gene expression in tumors is generally associated with poor survival in most cases. *Red*, *HOX* genes actively repaired from oxidative DNA damage by acetylated base excision repair DNA glycosylase NEIL1. (Reproduced from [23] by permission of Oxford University Press)

the gene symbols; patient codes; log2 of normalized Rsem values for the genes, i.e., log2(norm rsem + 1); the regression coefficient; and *p*-value (*see* **Note 6**).

3. Use option B to compute the correlation between one gene and all other genes (~20,500) of a given dataset. The usage is: vga_spotLight --optBdat <dataset> --optBgene <GENE>, where <dataset> is a TCGA gene expression file and <GENE> an official gene name. For example, "ibrun -n 1 vga_spotLight --optBdat ACC__geneExprT.txt --optBgene GRB2" will generate file "GRB2_toAll_ACC_T.txt," returning input gene name, test gene name, entrez gene record, number of observations, linear regression coefficient, and *p*-value. The program returns 2 and 1 in place of the linear regression coefficient and *p*-value when there is an insufficient number of observations (*see* **Note 7**). Use this option to explore co-expression patterns between a test gene and potential members of its pathways, as exemplified for *KAT8* encoding MOF, a member of the MYST histone acetyltransferase protein family, in which the co-expression data support a function in chromosome segregation [24](Fig. 5a, b).

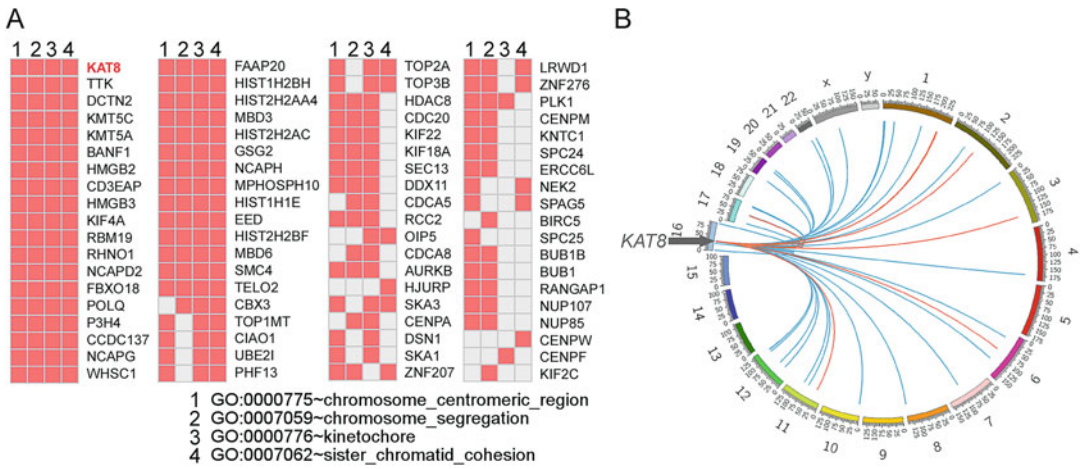


Fig. 5 Using GCEA to explore gene pathways. (**a**) List of genes highly co-expressed with *KAT8/MOF* in TCGA comprising Gene Ontology terms related to "chromosome segregation." (**b**) Circos plot of chromosome location of genes significantly ($P < 0.01$) co-expressed (blue links) or anticorrelated (orange links) with *KAT8/MOF* from panel **a** in KIRC patients. (Adapted from [24] by permission of the American Society for Microbiology)

4. Option F returns the gene expression results for one gene in a given dataset, and its use is: vga_spotLight --optFdat <dataset> --optFgene <GENE>. Using "ibrun -n 1 vga_spotLight --optFdat ACC__geneExprT.txt --optFgene GRB2" will generate file "ACC_GRB2_exprOne.txt" with patient code and log2(norm rsem + 1) values.
5. Options C–E were implemented to explore correlations between gene expression and mutation loads; we found correlations among common sets of genes in various types of tumor, including cell cycle and DNA repair in four tumors (KICH, LUAD, PRAD, and LGG), mitochondrial respiration in three tumors (STAD, THCA, and CHOL), antigen processing and presentation in CESC, reactivation of olfactory receptor genes in SKCM, and the unfolded protein response in BRCA [20] (*see* **Note 8**).

3.6 Basic Utilities

When working with custom scripts on repetitive tasks, it is helpful to create utilities that perform routine operations, such as using the information from file 1 to extract matching information in file 2 or running t-tests or linear regressions, without the need for user intervention. Here we present few scripts that may help automating these tasks.

1. Intersect two files. Let's assume we have RNA-Seq data, and we wish to assess whether there is a correlation between differentially up- and downregulated transcripts and gene length. Because gene length information is not contained in the

analyses of RNA-Seq output files, we need to obtain it from additional sources, such as in file knownGenes.txt from http://hgdownload.soe.ucsc.edu/goldenPath/hg38/database/. We can simplify our gene length file by executing: "awk '{ print $1 "\t" $5-$4 }' knownGene.txt | sed 's/\.[0-9]*//' > knownGenes_ens_length.txt" (*see* **Note 9**) to only contain ENS number and gene length:

```
ENST00000619216        68
ENST00000473358        1544
ENST00000469289        843
```

("knownGenes_ens_length.txt," first 3 lines)

Given our files "rnaSeq_up.txt" and "rnaSeq_down.txt" containing upregulated and downregulated transcripts, respectively, each with the list of significantly differentially expressed transcripts containing three fields, ESN number, geneID, and gene name:

```
ENST00000225964      1277      COL1A1-201
ENST00000358171       871      SERPINH1-201
ENST00000250383     10202      DHRS2-201
```

("rnaSeq_up.txt," first 3 lines)

we use the first field of "rnaSeq_up.txt" to extract its matching ENS number from "knownGenes_ens_length.txt." We could use Unix "grep"; however, "grep" is slow, particularly if the -w option is used and files are large. Script vga_intersect2x3.cpp at https://github.com/abacolla/intersect is a convenient alternative; its usage is: vga_intersect2x3 file1 file2 file3, where file1 is in our case "knownGenes_ens_length.txt"; file2 is, for example, "rnaSeq_up.txt"; and file3 is the output file (e.g., up_output). The vga_intersect2x3 script will be run separately on both up- and downregulated RNA-Seq files. The output file returned by vga_intersect2x3 contains all five fields from the two input files so as to verify that the match is correct. The two output files may be used to run a t-test (or a Wilcoxon test) to assess for statistical significance in gene (transcript) length as follows.

2. Run a t-test on the two output files obtained from vga_intersect2x3 above. Take the log10 of the second field from up_output and down_output (awk '{ print log($2)/log(10) }' up_output > logf2_up; awk '{print log($2)/log [10]}'

down_output > logf2_down), and run script vga_tTest as "vga_tTest logf2_up logf2_down." The t-test script may be found at https://github.com/abacolla/tTest.

3. Use vga_linearRegression to embed a linear regression analysis from an input file containing paired x and y coordinates and vga_fisherTest to compute Fisher exact tests from an input file containing one or more sets of data. The scripts may be found at https://github.com/abacolla/linearRegression and https://github.com/abacolla/fisherTest, respectively. All the commands above can be inserted directly into a single bash or other script and run without user intervention, thereby fulfilling our goal of simplifying the workflow and maintaining a record for reproducibility purposes.

4 Notes

1. We suggest moving the binaries to ~/bin and other scripts such as bash and R to ~/sbin and to modify PATH in .bash_profile, i.e., "export PATH=$HOME/bin:$PATH"; "export PATH=$HOME/sbin:$PATH." To compile, use the appropriate compiler type and its associated commands: Intel and icpc, AMD (AOCC) and clang, Gnu and g++, and PGI and pgc++.
2. The -noMask flag returns the DNA sequences in capital letters.
3. When scaling up jobs, it is useful to use a utility, such as remora, to track CPU usage and optimize the number of nodes requested for a job.
4. It is generally most convenient to crop the box frame of the original .png file, import it into a graphic program, and fill in with any additional statistical data using the associated .out file.
5. In the R ggplot2 package, notches for the box plots extend 1.58 * IQR / sqrt(n), where IQR is the interquartile range. Since these are asymmetrical with respect to the median, they can recurve when the data are asymmetrical, as seen in Fig. 2 for the ESCA data in normal tissues. In these cases, the notch option on line 46 may be turned off.
6. Consult the user manual for the HPC system in use to compile and run MPI applications in C++.
7. It is common to find strong co-expression for genes located near each other on the same chromosome in tumors.
8. It is essential to be mindful that whereas all scripts reported here return results, the results obtained ought to be cast onto broader contexts, and extensive controls need to be used, to avoid incorrect biological interpretation. For example, we found that in cancer genomes several *HOX* genes display

significant co-expression with two transcription factors, *FOXM1* and *MYBL2*. However, when compared with the co-expression levels of all genes, which revealed −log10 *P*-values for the linear regressions down to $\sim 10^{-250}$ and *r*-values up to 0.94, the results for the *HOX* genes were comparatively weak, and indeed the comprehensive comparison suggested indirect transactivation or perhaps non-casual relationships [23]. Likewise, the range of Kaplan-Meier *p*-values is strongly dependent upon tumor types, and therefore a "significant" *p*-value of 0.01, for example, may need to be interpreted with extreme caution. It is for this reasons that we think it is important to conduct broad-based investigations using in-house scripts rather than relying on limited analyses using exclusively online tools.

9. The alterative "sed" command: "awk '{ print $1 "\t" $5-$4 }' knownGene.txt | sed 's/\.[0-9]\+//' > knownGenes_ens_length.txt" may also work. Verify that fields are separated by a tab without additional white spaces; in "vim" you can use ":set list" to verify.

Acknowledgments

This work was supported by grants from the National Institutes of Health P01 CA092584, R35 CA220430, by the Cancer Prevention and Research Institute of Texas RP180813 and by a Robert A. Welch Chemistry Chair to J.A.T. The research used the Bridges/Bridges2 Pittsburgh Supercomputing Center through the Extreme Science and Engineering Discovery Environment (XSEDE), which are supported by the National Science Foundation grants ACI-1445606 and ACI-1548562, and the Texas Advanced Computing Center, supported by National Science Foundation grant ACI-1134872.

References

1. Pucker B, Schilbert HM, Schumacher SF (2019) Integrating molecular biology and bioinformatics education. J Integr Bioinform 16: 20190005
2. Houl JH, Ye Z, Brosey CA, Balapiti-Modarage LPF, Namjoshi S, Bacolla A, Laverty D, Walker BL, Pourfarjam Y, Warden LS et al (2019) Selective small molecule PARG inhibitor causes replication fork stalling and cancer cell death. Nat Commun 10:5654
3. Eckelmann BJ, Bacolla A, Wang H, Ye Z, Guerrero EN, Jiang W, El-Zein R, Hegde ML, Tomkinson AE, Tainer JA et al (2020) XRCC1 promotes replication restart, nascent fork degradation and mutagenic DNA repair in BRCA2-deficient cells. NAR. Cancer 2: zcaa013
4. Lees-Miller JP, Cobban A, Katsonis P, Bacolla A, Tsutakawa SE, Hammel M, Meek K, Anderson DW, Lichtarge O, Tainer JA et al (2020) Uncovering DNA-PKcs ancient phylogeny, unique sequence motifs and insights for human disease. Prog Biophys Mol Biol 163:87–108
5. Consortium ITP-CAoWG (2020) Pan-cancer analysis of whole genomes. Nature 578:82–93

6. Gerstung M, Jolly C, Leshchiner I, Dentro SC, Gonzalez S, Rosebrock D, Mitchell TJ, Rubanova Y, Anur P, Yu K et al (2020) The evolutionary history of 2,658 cancers. Nature 578:122–128
7. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J et al (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 43:11 10 11–11 10 33
8. Franke KR, Crowgey EL (2020) Accelerating next generation sequencing data analysis: an evaluation of optimized best practices for genome analysis toolkit algorithms. Genomics Inform 18:e10
9. Hesketh AR (2019) RNA sequencing best practices: experimental protocol and data analysis. Methods Mol Biol 2049:113–129
10. Vieth B, Parekh S, Ziegenhain C, Enard W, Hellmann I (2019) A systematic evaluation of single cell RNA-seq analysis pipelines. Nat Commun 10:4667
11. van Wietmarschen N, Sridharan S, Nathan WJ, Tubbs A, Chan EM, Callen E, Wu W, Belinky F, Tripathi V, Wong N et al (2020) Repeat expansions confer WRN dependence in microsatellite-unstable cancers. Nature 586:292–298
12. McKinney JA, Wang G, Vasquez KM (2020) Distinct mechanisms of mutagenic processing of alternative DNA structures by repair proteins. Mol Cell Oncol 7:1743807
13. Berroyer A, Kim N (2020) The functional consequences of eukaryotic topoisomerase 1 interaction with G-quadruplex DNA. Genes 11:193
14. Bacolla A, Tainer JA, Vasquez KM, Cooper DN (2016) Translocation and deletion breakpoints in cancer genomes are associated with potential non-B DNA-forming sequences. Nucleic Acids Res 44:5673–5688
15. Puig Lombardi E, Londono-Vallejo A (2020) A guide to computational methods for G-quadruplex prediction. Nucleic Acids Res 48:1–15
16. Cer RZ, Donohue DE, Mudunuri US, Temiz NA, Loss MA, Starner NJ, Halusa GN, Volfovsky N, Yi M, Luke BT et al (2013) Non-B DB v2.0: a database of predicted non-B DNA-forming motifs and its associated tools. Nucleic Acids Res 41:D94–D100
17. Brazda V, Kolomaznik J, Lysek J, Haronikova L, Coufal J, St'astny J (2016) Palindrome analyser - a new web-based server for predicting and evaluating inverted repeats in nucleotide sequences. Biochem Biophys Res Commun 478:1739–1745
18. Buske FA, Bauer DC, Mattick JS, Bailey TL (2012) Triplexator: detecting nucleic acid triple helices in genomic and transcriptomic data. Genome Res 22:1372–1381
19. Hon J, Martinek T, Rajdl K, Lexa M (2013) Triplex: an R/Bioconductor package for identification and visualization of potential intramolecular triplex patterns in DNA sequences. Bioinformatics 29:1900–1901
20. Bacolla A, Ye Z, Ahmed Z, Tainer JA (2019) Cancer mutational burden is shaped by G4 DNA, replication stress and mitochondrial dysfunction. Prog Biophys Mol Biol 147:47–61
21. Zhao J, Wang G, Del Mundo IM, McKinney JA, Lu X, Bacolla A, Boulware SB, Zhang C, Zhang H, Ren P et al (2018) Distinct mechanisms of nuclease-directed DNA-structure-induced genetic instability in cancer genomes. Cell Rep 22:1200–1210
22. Seo SH, Bacolla A, Yoo D, Koo YJ, Cho SI, Kim MJ, Seong MW, Kim HJ, Kim JM, Tainer JA et al (2020) Replication-based rearrangements are a common mechanism for SNCA duplication in Parkinson's disease. Mov Disord 35:868–876
23. Bacolla A, Sengupta S, Ye Z, Yang C, Mitra J, De-Paula RB, Hegde ML, Ahmed Z, Mort M, Cooper DN et al (2021) Heritable pattern of oxidized DNA base repair coincides with pre-targeting of repair complexes to open chromatin. Nucleic Acids Res 49:221–243
24. Singh M, Bacolla A, Chaudhary S, Hunt CR, Pandita S, Chauhan R, Gupta A, Tainer JA, Pandita TK (2020) Histone acetyltransferase MOF orchestrates outcomes at the crossroad of oncogenesis, DNA damage response, proliferation, and stem cell development. Mol Cell Biol 40

Chapter 2

A Whole Genome CRISPR/Cas9 Screening Approach for Identifying Genes Encoding DNA End-Processing Proteins

Bo-Ruei Chen and Barry P. Sleckman

Abstract

DNA double-strand breaks (DSBs) are mainly repaired by homologous recombination (HR) and non-homologous end joining (NHEJ). The choice of HR or NHEJ is dictated in part by whether the broken DNA ends are resected to generate extended single-stranded DNA (ssDNA) overhangs, which are quickly bound by the trimeric ssDNA binding complex RPA, the first step of HR. Here we describe a series of protocols for generating Abelson murine leukemia virus-transformed pre-B cells (abl pre-B cells) with stably integrated inducible Cas9 that can be used to identify and study novel pathways regulating DNA end processing. These approaches involve gene inactivation by CRISPR/Cas9, whole genome guide RNA (gRNA) library-mediated screen, and flow cytometry-based detection of chromatin-bound RPA after DNA damage.

Key words CRISPR/Cas9, Genome-wide screen, NHEJ, HR, DNA end resection, RPA

1 Introduction

DNA double-strand breaks (DSBs) are among the most deleterious DNA lesions that can occur upon exposure to genotoxic agents such as ionizing radiation and chemotherapeutics or during physiological processes including replication, transcription, and meiosis. To prevent genome instability arising from un- or mis-repaired DNA DSBs, two major DNA DSB repair pathways exist in cells. For cells in S and G_2 phases of the cell cycle, homologous recombination (HR) is utilized to accurately repair DSBs using sister chromatids as the repair templates [1]. Non-homologous end joining (NHEJ), on the other hand, can function in all phases of the cell cycle phases to directly join broken DNA ends that have not been processed to generate long single-stranded DNA (ssDNA) overhangs [2]. A critical bifurcation of HR and NHEJ is the resection of broken DNA ends leading to the generation of single-stranded DNA (ssDNA) overhangs. DSBs to be repaired by HR is first

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_2,

resected to generate ssDNA overhangs that require nucleases and non-nuclease proteins such as MRE11, CtIP, and BRCA1 and are quickly bound by the trimeric ssDNA binding RPA complex [3, 4]. In contrast, NHEJ only joins blunt DNA ends or ends with minimal processing and relies on DNA end protection proteins such 53BP1, RIF1, and the shieldin complex to prevent ssDNA generation at DSB ends [2, 5–7]. Therefore, HR requires a concerted action of nuclease activities to generate ssDNA, while NHEJ depends on DNA end protection proteins to limit nucleolytic activity, especially in G_0/G_1 cells, where NHEJ must repair DNA DSBs due to the lack of sister chromatids.

The CRISPR/Cas system was originally discovered as a unique RNA-guided DNA endonuclease that functions as a form of bacterial adaptive immunity [8, 9]. Not long after the discovery of CRISPR/Cas9 in bacteria, it was quickly adopted in engineering the genomes of cultured human cells and has revolutionized genome editing across many species [10]. CRISPR/Cas9 has also been adopted in genome-wide genetic screens and more widely used now than short hairpin RNA (shRNA)- or short interfering RNA (siRNA)-based approaches [11–13]. The biggest distinction between CRISPR/Cas9 technology and RNA interference is that CRISPR/Cas9 generates mutations that inactivate genes or makes other genetic alterations in cells, thereby not requiring continuous Cas9 or guide RNA (gRNA) expression. In contrast, gene inactivation by RNA interference will quickly subside when cells lose expression of the shRNA or siRNA.

Abelson murine leukemia virus (A-MuLV)-transformed pre-B cell lines, hereafter abl pre-B cells, were first developed by transducing pre-B cells from spleen or bone marrow with non-replicative A-MuLV and have been instrumental to studies on the regulation of B lymphocyte development [14, 15]. Transformation by A-MuLV halts B lymphocyte differentiation primarily at the large pre-B cell stage, and abl pre-B cells proliferate in culture in the absence of growth factors such as interleukin-7 (IL-7). In addition, ectopic expression Bcl2 transgene facilitates the generation of abl pre-B cells and promotes their survival in culture [16]. However, inhibition of the viral tyrosine kinase v-abl by imatinib arrests abl pre-B cells in G_0/G_1 and promotes their differentiation into small pre-B cells [17]. One of the genetic programs activated in imatinib-treated abl pre-B cells is V(D)J recombination, the process that drives antigen receptor gene assembly in lymphocytes. During V(D)J recombination, the endonuclease complex RAG1/RAG2 generates DSBs at antigen receptor gene segments participating in the recombination, and NHEJ is required to repair the DSB intermediates to generate genes encoding antibodies in B cells or T cell receptors in T cells [18, 19]. The requirement of NHEJ during V(D)J recombination and the fact that imatinib causes abl pre-B

cells arrest in G_0/G_1 phases of the cell cycle, make abl pre-B cells an excellent model to study genetic pathways that regulate NHEJ.

Abl pre-B cells have been used to demonstrate several key features of NHEJ [16, 20–24]. We have recently applied the CRISPR/Cas9 technology to generate abl pre-B cell lines with stably integrated doxycycline-inducible FLAG-Cas9. Upon doxycycline treatment, homogenous induction leads to almost all cells in the culture expressing the FLAG-Cas9 protein. The homogenous expression of Cas9 in these cells upon exposure to doxycycline allows highly efficient gene inactivation and depletion of target proteins with single-guide RNAs, which we term bulk gene inactivation. By modulating the duration of doxycycline treatment, this approach can be used to efficiently deplete proteins essential for cell proliferation (3–4 day doxycycline treatment) as well as proteins with high abundance and stability that require many cell divisions for their loss (7-day doxycycline treatment). The high efficiency of CRISPR/Cas9-mediated gene inactivation in abl pre-B cells and the ease of scaling up the size of the culture make these cells ideal for genome-scale gRNA screening. Here we describe protocols for generating abl pre-B cell lines, isolating abl pre-B cells with stably integrated inducible Cas9 using lentivirus, and preparing cellular gRNA library in abl pre-B cells with inducible Cas9. We also include protocols for a flow cytometry-based RPA assay for monitor resection [25] and for using this assay to conduct genome-wide screen for genes regulating DNA end processing.

2 Materials

2.1 Cell Culture

1. DMEM media + 100 U/ml penicillin–streptomycin, 2 mM L-glutamine, 1× non-essential amino acid, 1 mM sodium pyruvate, 55 μM 2-mercaptoethanol, 10% fetal bovine serum (FBS).
2. Plat E media: DMEM media + 1 μg/ml puromycin, 10 μg/ml blasticidin, 100 U/ml penicillin-streptomycin, 10% FBS.
3. Opti-MEM.
4. 293T cells.
5. Platinum-E cells.
6. Abl pre-B cells: *see* Subheading 3.1 for preparation details.

2.2 Plasmids

1. pMSCV-v-abl.
2. pMX-Bcl2-hCD271.
3. pCMV-dR8.2 dvpr.
4. pCMV-VSV-G.
5. pCW-Cas9.
6. Mouse genome-wide lentiviral CRISPR gRNA library version 2 (Addgene #67988).

2.3 Antibodies

1. Rat anti-RPA32 (4E4) (Cell Signaling Technology, 2208S).
2. Alexa Fluor 488 goat anti-rat IgG (BioLegend, 405,418).
3. Mouse anti-FLAG M2 (Sigma Aldrich, F1804).
4. Alexa Fluor 488 goat anti-mouse IgG (BioLegend, 405319).

2.4 Other Reagents

1. Puromycin.
2. Blasticidin.
3. Lipofectamine 2000.
4. Sequa-brene.
5. Doxycycline hydrochloride.
6. FACS wash (2% FBS in 1× PBS).
7. BD Permeabilization Buffer Plus.
8. BD Perm/Wash Buffer.
9. BD Cytofix/Cytoperm.
10. Rapid lysis buffer (100 mM Tris pH 8.5, 5 mM EDTA, 0.2% SDS, 200 mM NaCl, 10 μg/ml proteinase K).
11. Platinum SuperFi DNA Polymerase.
12. QIAquick Gel Extraction Kit.

2.5 PCR Primers

Primers for generating NextSeq library are listed in Table 1 (*see* **Note 1**).

3 Methods

3.1 Generation of abl Pre-B Cell Lines

1. Plate 2 × 10^6 Plat-E cells in a 6-cm tissue culture dish 1 day before transfection in 5 ml of antibiotics-free DMEM with 10% fetal bovine serum (FBS).
2. Prepare transfection mix in 1.5 ml microfuge tubes. For each transfection (a 5-cm dish): (a) tube 1: mix 250 μl of Opti-MEM + 20 μl of Lipofectamine 2000 and incubate at room temperature for 5 min. (b) tube 2: mix 250 μl of Opti-MEM + 9 μg of pMSCV-v-abl. (c) Mix tube 2 to tube 1 and incubate the mixture at room temperature for 20 min.
3. Add transfection mix to a plate of Plat-E cells drop-wise across the entire plate, and gently shake the plate to mix.
4. Replace the transfection media with transfection mix with 5 ml of fresh DMEM + 10% FBS media 6–8 h after transfection.
5. Return the plates to the incubator, and collect and filter retroviral supernatant with 0.45 μm filter 2 days after transfection.
6. Isolate femurs from a mouse and snip off both ends of the bones, and force bone marrow out into a 15 ml conical using a

Table 1
Sequences of oligos/primers used in this protocol

Oligonucleotide name	Oligonucleotide sequences
pKLV lib330F:	AATGGACTATCATATGCTTACCGT
pKLV lib490R:	CCTACCGGT GGATGTGGAATG
PE.P5_pKLV lib195 Fwd	AATGATACGGCGACCACCGAGATCTGGCTTTATATATCTTGTGGAAAGGAC
P7 index180 Rev	CAAGCAGAAGACGGCATACGAGAT*INDEX*GTGACTGGAGTTCAGACGTGTGCTCTTCC GATCCAGACTGCCTTGGGAAAAGC

syringe with a 25-gauge needle and 3 ml (1.5 ml per femur) of DMEM + 20% FBS.

7. Break marrow into single cell suspension by repeatedly pipetting.
8. Plate 2×10^6 cells in 1 ml of DMEM + 20% FBS in a six-well plate. Set up 3–6 wells per mouse.
9. Add 1 ml of pMSCV-v-abl viral supernatant to each well of marrow suspension.
10. Five days later, feed each well with 1 ml of DMEM + 20% FBS, and continue adding 1 ml of DMEM + 20% FBS two more times when the color of the medium turn orange (*see* **Note 2**). The transformed cells will initially appear as clusters of large cells on the bottom of the well and will divide rapidly from there.
11. Split cells after they are fairly dense in the well—as they will be sensitive to being too sparse. Begin by splitting them 1:2 (e.g., 2 ml of cell suspension with 2 ml of DMEM + 20% FBS), and increase the fold of dilution after every five splits (*see* **Note 3**). At the first time of splitting, cells should be moved into a new plate to get them away from the carryover fibroblasts.
12. Once cells continuously grow well at 1:10 split, reduce the concentration of FBS to 10%, the concentration of FBS for established abl pre-B cell lines.

3.2 Establishing abl Pre-B Cell Line with Inducible Cas9 Using Lentiviral pCW-Cas9

1. Plate 2×10^6 293T cells in a 6-cm tissue culture dish 1 day before transfection in 5 ml of antibiotics-free DMEM + 10% FBS media.
2. Prepare transfection mix in 1.5 ml microfuge tubes. For each transfection (a 5-cm dish): (a) tube 1: mix 250 μl of Opti-MEM + 20 μl of Lipofectamine 2000, and incubate at room temperature for 5 min. (b) tube 2: mix 200 μl of Opti-MEM + 4 μg of pCMV-dR8.2 dvpr, 4 μg of pCW-Cas9, and 0.8 μg of pCMV-VSV-G. (c) Mix tube 2 to tube 1, and incubate the mixture at room temperature for 20 min.
3. Add transfection mix to a plate of 293T cells drop-wise across the entire plate, and gently shake the plate to mix.
4. Replace the transfection media with transfection mix with 5 ml of fresh DMEM + 10% FBS 6–8 h after transfection.
5. Return the plates to the incubator, and collect and filter lentiviral supernatant with 0.45 μm filter 2 days after transfection.
6. Transduce abl pre-B cells with lentiviral pCW-Cas9 by mixing 2×10^6 cells (in 1 ml of medium) with 1.5 ml of pCW-Cas9 lentiviral supernatant and sequa-brene at 5 μg/ml in a six-well plate, and spin the cell-virus mixture at 1800 rpm for 1.5 h.

7. Add 1 ml of medium to cells 1 day after transduction.
8. Two days after transduction, select transduced cells with 2 μg/ml of puromycin at 2×10^5 cells (*see* **Note 4**) along with un-transduced cells as the control.
9. Subclone puromycin-resistant cells 5–7 days after treating with puromycin in 96-well plates by serial dilution.
10. Upon expansion of single clones, treat a small portion of cells from these clones with 3 μg/ml of doxycycline for 2 days.
11. Collect doxycycline-treated cells, and wash with 1 ml of FACS wash. Use cells not treated with doxycycline or not transduced with pCW-Cas9 as the control.
12. Fix cells in 100 μl of BD Cytofix/Cytoperm at room temperature for 20 min.
13. Wash cells with 1 ml of FACS wash, and permeabilize cells with 100 μl of BD Permeabilization Buffer Plus on ice for 10 min.
14. Wash cells with 1 ml of FACS wash, and fix cells again with 100 μl of BD Cytofix/Cytoperm at room temperature for 5 min.
15. Wash cells with 1 ml of FACS wash, and stain cells with anti-FLAG M2 antibody at 1:500 dilution in 100 μl of 1× BD Perm/Wash buffer at room temperature for 2 h.
16. Wash cells with 1 ml of 1× BD Perm/Wash buffer, and stain cells with Alexa Fluor 488 Goat anti-mouse IgG antibody at 1:500 dilution in 100 μl of 1× BD Perm/Wash buffer in the dark at room temperature for 1 h.
17. Wash cells with 1 ml of 1× BD Perm/Wash buffer, and resuspend cells in 300 μl of 1× PBS for analysis on a flow cytometer.
18. Identify clones in which doxycycline treatment promotes homogenous expression of FLAG-Cas9 in all cells (Fig. 1) (*see* **Note 5**).

3.3 Preparation of Lentiviral gRNA Library

1. Split 293T cells from one confluent 10-cm culture dish to three 10-cm dishes in 10 ml of antibiotics-free growth media 1 day before transfection (*see* **Note 6**).
2. Prepare transfection mix in 1.5 ml microfuge tubes. For each transfection (a 10-cm dish): (a) tube 1: mix 400 μl of Opti-MEM + 30 μl of Lipofectamine 2000, and incubate at room temperature for 5 min. (b) tube 2: mix 400 μl of Opti-MEM + 6 μg of pCMV-dR8.2 dvpr, 6 μg of lentiviral gRNA library DNA, and 1.2 μg of pCMV-VSV-G. (c) Mix tube 2 to tube 1 and incubate the mixture at room temperature for 20 min.
3. Add transfection mix to a plate of 293T cells drop-wise across the entire plate, and gently shake the plate to mix.

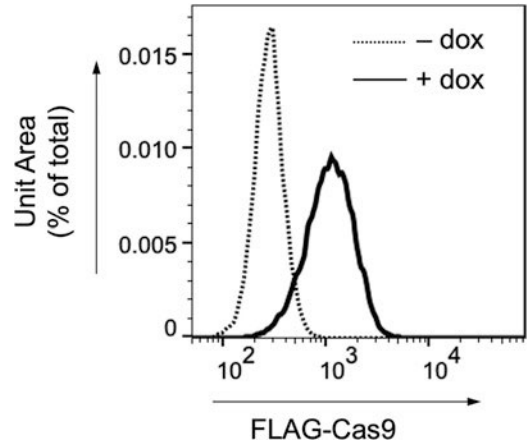


Fig. 1 Flow cytometric analysis of FLAG-Cas9 in *Lig4*$^{-/-}$ cells with (+dox) and without (−dox) doxycycline to induce expression of FLAG-Cas9

4. Replace the transfection media with transfection mix with 10 ml of fresh media 6–8 h after transfection.
5. Return the plates to the incubator, and collect lentiviral supernatant 2 days after transfection.
6. Filter the viral supernatant with 0.45 μm filter, and store at −80 °C if not used immediately (*see* **Note 7**).

3.4 Preparation of Genome-Scale abl Pre-B Cell gRNA Libraries

1. Mix one million cells (in 1 ml of growth media) with 1 ml of filtered and diluted lentiviral gRNA library supernatant and 2 μl of 5 mg/ml Sequa-brene (*see* **Note 7**).
2. Place the cell-virus mixture in a well of six-well plates. For 1000-fold coverage at 50% transduction efficiency, set up 30 six-well plates (180 million cells) (*see* **Note 8**).
3. Spin plates at 1800 rpm for 90 min, and return plate to the incubator.
4. Add 1 ml of fresh media 1 day after transduction.
5. Three days after transduction, sort BFP-positive cells using BD FACS Aria Cell Sorter.
6. One day after cell sorting, treat cells with doxycycline at 3 μg/ml for 3–7 days (*see* **Note 9**).

3.5 Genome-Scale Screen for Genes Regulating DNA DSB End Resection in G0/G1-Arrested Abl Pre-B Cells Using Flow Cytometry-Based RPA Assay

1. Treat lentiviral gRNA library-transduced, BFP-sorted abl pre-B cells with 3 μM imatinib at 2×10^6 cells/ml for 48 h to arrest cells in G_0/G_1.
2. Expose cells to 15 Gy IR (*see* **Note 10**).
3. 18 h after IR, spin down 100 million cells, and wash cells with 20 ml of FACS wash, and spin cells at 1200 rpm for 5 min.
4. Extract cells with 0.6 ml of 0.05% Triton X-100 in 1× PBS on ice for 10 min (*see* **Note 11**).

5. Wash with 20 ml of cold 1× FACS wash, and spin cells at 1200 rpm for 5 min.
6. Fix cells with 0.6 ml of BD Cytofix/Cytoperm at room temperature for 5 min.
7. Wash with 20 ml cold 1× FACS wash, and spin cells at 1200 rpm for 5 min.
8. Stain cells with 1 ml of anti-RPA32 antibody at 1:500 dilution in 1× BD Perm/Wash at room temperature for 2 h.
9. Wash with 2 ml 1× BD Perm/Wash, and spin cells at 1200 rpm for 5 min.
10. Stain cells with 1 ml of Alexa Fluor 488 goat anti-rat IgG at 1:500 dilution in 1× BD Perm/Wash in the dark at room temperature for 1 h.
11. Sort cells on BD FACS Aria Cell Sorter for cells with higher (gRNAs targeting end protection genes) or lower (gRNAs targeting resection promoting genes) RPA antibody staining intensity (*see* **Note 12**).
12. Lyse sorted cells with rapid lysis buffer at 55 °C overnight.
13. Extract DNA by adding equal volume of isopropanol, mix thoroughly, and spin at 3000 rpm for 5 min.
14. Wash the DNA pellet with 1 ml 70% ethanol, and dissolve DNA in 1× TE.

3.6 PCR of gRNAs from Genomic DNA Isolated from Sorted Cells

1. first round PCR: Use 3 μg gDNA per reaction (~0.5 million cells) in a 50 μl reaction (10 μl of 5× SuperFi buffer, 10 μl of GC Enhancer, 2 μl of 5 mM dNTP mix, 2.5 μl of 10 μM pKLV lib330F primer, 2.5 μl of 10 μM pKLV lib 490R primer, 0.5 μl of SuperFi DNA Polymerase, adding Milli-Q H_2O and DNA to a final of 50 μl). The thermocycle program is as the following: 98 °C/5 min—[98 °C/15 s—60 °C/15 s—72 °C/1 min] × 15–72 °C/5 min. Amplicon size is 329 bp (*see* **Note 13**).
2. second round PCR (to add Illumina adaptors and indexes): Pool together all first PCR products (no need to purify), and use 5 μl of first PCR product as the template for each second PCR reaction (10 μl of 5× SuperFi buffer, 10 μl of GC Enhancer, 2 μl of 5 mM dNTP mix, 2.5 μl of 10 μM PE. P5_pKLV lib195 Fwd primer, 2.5 μl of 10 μM P7 index180 Rev. primer, 0.5 μl of SuperFi DNA Polymerase, adding 5 μl of DNA template and 17.5 μl of Milli-Q H_2O). Set up at least ten reactions using the following program: 98 °C/5 min—[98 °C/15 s—60 °C/30 s—72 °C/20 s] × 10–72 °C/5 min. Amplicon size is 283 bp (*see* **Note 13**).

3. Resolve the PCR product in 1.5% agarose gel, excise gel slices containing PCR product DNA, and purify the resulting PCR product using QIAUICK Gel Extraction Kit for NextSeq (*see* **Note 14**).

4 Notes

1. The oligonucleotide sequences are designed based on the pKLV vector sequence used in the mouse genome-wide lentiviral CRISPR gRNA library version 2 (addgene, #67988), the adaptor sequences for Illumina NextSeq. If using a different gRNA library and sequencing platform, the sequences need to be changed accordingly.
2. Avoid splitting cells at this stage (the first 7–9 days) as during the initial phase of transformation, cells are sensitive to being sparse. The time for feeding cells with fresh media varies from cell line to cell line. The color indicator in the media turning orange gives indication of good proliferation and metabolic activities of cultured cells.
3. If dead cells start to increase after starting to split the culture at higher dilution, return to splitting cells at lower dilution until cells can proliferate normally and maintain good viability after each split.
4. Puromycin selection on abl pre-B cells will not be effective at high cell density. The control cells (no puromycin-resistant gene) should never grow out in the culture. One can stop the selection when the transduced cells start to grow robustly, making the culture look turbid, while the control culture remains nearly clear.
5. We have found that the expression levels of Cas9 after doxycycline treatment are not as critical as the homogeneity of the expression in the entire culture. If all cells in a culture express Cas9 homogenously, even at a low level, the gene inactivation efficiency is usually still very high.
6. We normally plate 10–12 10-cm plates for each batch of viral library preparation to collect ~100–120 ml of viral library supernatant. This is usually enough for four to six screens in abl pre-B cells as described in this chapter.
7. The titer/transduction efficiency of frozen viral supernatant should be determined every time before each use.
8. The Yusa mouse guide RNA library contains 90,230 gRNAs. Assuming one cell only receives one gRNA, to achieve 1000-fold coverage, 9×10^7 transduced cells are required. If transduction efficiency is pre-determined at 50%, ~1.8×10^8 will be required for viral transduction.

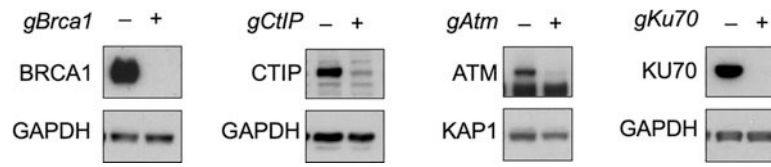


Fig. 2 Western blot on abl pre-B cells with inducible Cas9 and with (+) or without (−) indicated gRNAs (+) after doxycycline treatment for 4 (*gBrca1* and *gCtIP*) or 7 days (*gATM* and *gKu70*) using specified antibodies

9. We have routinely been able to use abl pre-B cells with inducible Cas9 to conduct bulk knockout of genes essential and not essential (e.g., *Atm*, *Ku70*) for cell proliferation with gRNAs by changing the duration of doxycycline treatment. For genes critical for survival or proliferation (e.g., *Brca1*, *CtIP*), we only treat cells with doxycycline for 3–4 days. For bulk knockout of genes not required for proliferation, we generally treat cells with doxycycline for 7 days to achieve maximal protein depletion (Fig. 2). If the screen values the contribution of essential genes, the time of doxycycline treatment can be shorter (4 days) to increase the inclusion of cells with inactivating mutations in the screen. On the other hand, treating doxycycline for a longer period of time ensures the cells with inactivating mutations can efficiently eliminate proteins encoded by the targeted genes during cellular doublings.
10. We typically irradiate 100 million in 50 ml of culture in a T75 flask in a XRAD 320 (Precision X-ray Inc) irradiator. The number of cells and type of vessels could be adjusted according to the available irradiator.
11. The concentration of Triton X-100, 0.05%, is empirically determined for imatinib-treated abl pre-B cells. If using proliferating abl pre-B cells, 0.2% Triton X-100 in 1× PBS should be used for the pre-extraction step. For other cell types, concentration of Triton X-100 should be determined empirically for optimal extraction efficiency.
12. *See* Fig. 3 for examples of minimal and elevated chromatin-bound RPA after irradiation in imatinib-treated *Lig4*$^{-/-}$ and *Lig4*$^{-/-}$*:53bp1*$^{-/-}$ abl pre-B cells, respectively, reflecting limited and extended resection in these cells.
13. PCR steps are based on Invitrogen Platinum SuperFi DNA Polymerase. A distinct band of PCR product should not or only weakly be visible in gel after the first PCR. A product of ~300 bp will be readily detectable after the second PCR amplification using the product from the first round PCR as the template.

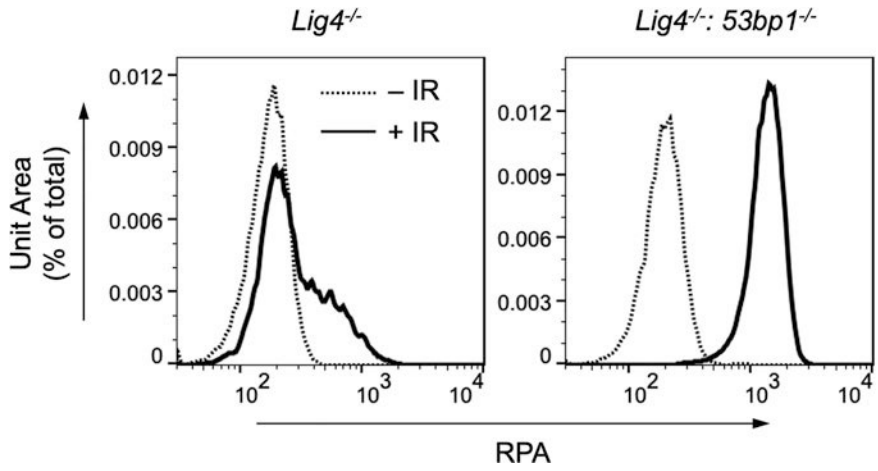


Fig. 3 Flow cytometric analysis of chromatin-bound RPA in imatinib-treated *Lig4*$^{-/-}$ and *Lig4*$^{-/-}$*: 53 bp1*$^{-/-}$ abl pre-B cells 18 h after treated with (+IR) and without (-IR) 15 Gy IR

14. An example of how to analyze and interpret the NextSeq data of a screening can be found at https://www.biorxiv.org/content/10.1101/2021.04.21.440786v1.

Acknowledgments

B.P.S. is supported by National Institutes of Health grants R01 AI047829 and R01 AI074953.

References

1. Prakash R, Zhang Y, Feng W, Jasin M (2015) Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. Cold Spring Harb Perspect Biol 7(4):a016600. https://doi.org/10.1101/cshperspect.a016600
2. Chang HHY, Pannunzio NR, Adachi N, Lieber MR (2017) Non-homologous DNA end joining and alternative pathways to double-strand break repair. Nat Rev Mol Cell Biol 18(8):495–506. https://doi.org/10.1038/nrm.2017.48
3. Symington LS, Gautier J (2011) Double-strand break end resection and repair pathway choice. Annu Rev Genet 45:247–271. https://doi.org/10.1146/annurev-genet-110410-132435
4. Ciccia A, Elledge SJ (2010) The DNA damage response: making it safe to play with knives. Mol Cell 40(2):179–204. https://doi.org/10.1016/j.molcel.2010.09.019
5. Mirman Z, de Lange T (2020) 53BP1: a DSB escort. Genes Dev 34(1–2):7–23. https://doi.org/10.1101/gad.333237.119
6. Setiaputra D, Durocher D (2019) Shieldin—the protector of DNA ends. EMBO Rep 20(5):e47560. https://doi.org/10.15252/embr.201847560
7. Bunting SF, Callen E, Wong N, Chen HT, Polato F, Gunn A, Bothmer A, Feldhahn N, Fernandez-Capetillo O, Cao L, Xu X, Deng CX, Finkel T, Nussenzweig M, Stark JM, Nussenzweig A (2010) 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. Cell 141(2):243–254. https://doi.org/10.1016/j.cell.2010.03.012
8. Wiedenheft B, Sternberg SH, Doudna JA (2012) RNA-guided genetic silencing systems in bacteria and archaea. Nature 482(7385):331–338. https://doi.org/10.1038/nature10886

9. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337(6096):816–821. https://doi.org/10.1126/science.1225829
10. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. Science 339(6121):823–826. https://doi.org/10.1126/science.1232033
11. Unniyampurath U, Pilankatta R, Krishnan MN (2016) RNA interference in the age of CRISPR: will CRISPR interfere with RNAi? Int J Mol Sci 17(3):291. https://doi.org/10.3390/ijms17030291
12. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root DE, Doench JG, Zhang F (2014) Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 343(6166):84–87. https://doi.org/10.1126/science.1247005
13. Mohr SE, Smith JA, Shamu CE, Neumuller RA, Perrimon N (2014) RNAi screening comes of age: improved techniques and complementary approaches. Nat Rev Mol Cell Biol 15(9):591–600. https://doi.org/10.1038/nrm3860
14. Rosenberg N, Kincade PW (1994) B-lineage differentiation in normal and transformed cells and the microenvironment that supports it. Curr Opin Immunol 6(2):203–211. https://doi.org/10.1016/0952-7915(94)90093-0
15. Rosenberg N, Baltimore D, Scher CD (1975) In vitro transformation of lymphoid cells by Abelson murine leukemia virus. Proc Natl Acad Sci U S A 72(5):1932–1936. https://doi.org/10.1073/pnas.72.5.1932
16. Bredemeyer AL, Sharma GG, Huang CY, Helmink BA, Walker LM, Khor KC, Nuskey B, Sullivan KE, Pandita TK, Bassing CH, Sleckman BP (2006) ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. Nature 442(7101):466–470. https://doi.org/10.1038/nature04866
17. Muljo SA, Schlissel MS (2003) A small molecule Abl kinase inhibitor induces differentiation of Abelson virus-transformed pre-B cell lines. Nat Immunol 4(1):31–37. https://doi.org/10.1038/ni870
18. Helmink BA, Sleckman BP (2012) The response to and repair of RAG-mediated DNA double-strand breaks. Annu Rev Immunol 30:175–202. https://doi.org/10.1146/annurev-immunol-030409-101320
19. Schatz DG, Swanson PC (2011) V(D)J recombination: mechanisms of initiation. Annu Rev Genet 45:167–202. https://doi.org/10.1146/annurev-genet-110410-132552
20. Hung PJ, Johnson B, Chen BR, Byrum AK, Bredemeyer AL, Yewdell WT, Johnson TE, Lee BJ, Deivasigamani S, Hindi I, Amatya P, Gross ML, Paull TT, Pisapia DJ, Chaudhuri J, Petrini JJH, Mosammaparast N, Amarasinghe GK, Zha S, Tyler JK, Sleckman BP (2018) MRI is a DNA damage response adaptor during classical non-homologous end joining. Mol Cell 71(2):332–342 e338. https://doi.org/10.1016/j.molcel.2018.06.018
21. Hung PJ, Chen BR, George R, Liberman C, Morales AJ, Colon-Ortiz P, Tyler JK, Sleckman BP, Bredemeyer AL (2017) Deficiency of XLF and PAXX prevents DNA double-strand break repair by non-homologous end joining in lymphocytes. Cell Cycle 16(3):286–295. https://doi.org/10.1080/15384101.2016.1253640
22. Tubbs AT, Dorsett Y, Chan E, Helmink B, Lee BS, Hung P, George R, Bredemeyer AL, Mittal A, Pappu RV, Chowdhury D, Mosammaparast N, Krangel MS, Sleckman BP (2014) KAP-1 promotes resection of broken DNA ends not protected by gamma-H2AX and 53BP1 in G(1)-phase lymphocytes. Mol Cell Biol 34(15):2811–2821. https://doi.org/10.1128/MCB.00441-14
23. Dorsett Y, Zhou Y, Tubbs AT, Chen BR, Purman C, Lee BS, George R, Bredemeyer AL, Zhao JY, Sodergen E, Weinstock GM, Han ND, Reyes A, Oltz EM, Dorsett D, Misulovin Z, Payton JE, Sleckman BP (2014) HCoDES reveals chromosomal DNA end structures with single-nucleotide resolution. Mol Cell 56(6):808–818. https://doi.org/10.1016/j.molcel.2014.10.024
24. Helmink BA, Tubbs AT, Dorsett Y, Bednarski JJ, Walker LM, Feng Z, Sharma GG, McKinnon PJ, Zhang J, Bassing CH, Sleckman BP (2011) H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes. Nature 469(7329):245–249. https://doi.org/10.1038/nature09585
25. Forment JV, Walker RV, Jackson SP (2012) A high-throughput, flow cytometry-based method to quantify DNA-end resection in mammalian cells. Cytometry A 81(10):922–928. https://doi.org/10.1002/cyto.a.22155

Chapter 3

Immunoaffinity Purification of Epitope-Tagged DNA Repair Complexes from Human Cells

Brittany A. Townley, Jennifer M. Soll, and Nima Mosammaparast

Abstract

Immunoaffinity purification allows for the purification of epitope-tagged proteins and their associated multisubunit complexes from mammalian cells. Subsequent identification of the proteins by proteomic analysis enables unbiased biochemical characterization of their associated partners, potentially revealing the physiological or functional context of any given protein. Here, we use immunoaffinity isolation of the Activating Signal Co-integrator Complex (ASCC) from human cells as an example, demonstrating the utility of the approach in revealing protein complexes involved in genotoxic stress responses.

Key words Protein purification, Mass spectrometry, ASCC

1 Introduction

One approach to determining the function of an uncharacterized protein is to identify its interaction partners, as proteins are often integrated into stable multimeric protein complexes. Traditionally, this can be accomplished through utilizing either the yeast two-hybrid system or immunoaffinity purification coupled with mass spectrometry. While the yeast two-hybrid system has been used to successfully discover direct physical interaction between proteins, the method characteristically features high levels of false positives and false negatives, in addition to having multiple protocol-specific biases [1–3]. Affinity purification, however, enables the discovery of intramolecular protein interactors, thus widening the potential to identify both multi-protein complexes to which the uncharacterized protein belongs, as well as members of distinct interacting protein complexes [4].

Immunoaffinity purification techniques rely on recognition of the protein of interest by an antibody bound to a resin (e.g., agarose or Sepharose beads). After cell extract is passed over the antibody/resin conjugate, specific interacting proteins remain bound, while

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_3,

unbound proteins are washed away. Elution of the remaining bound proteins is accomplished through addition of competing epitope peptides, or by addition of high salt or acidic pH. Affinity purification techniques encompass native complex purification, purification using epitope tagging, and various chromatography-based purification strategies that utilize differential interactions on a resin during the purification process. Native complex purification is reliant upon the availability of high-affinity, specific antibodies against the protein of interest, whereas the various forms of conventional chromatography can decrease the stability of multi-protein complexes [5].

Epitope-tagging and immunoaffinity purification of protein complexes allow for a simplified purification strategy using well-characterized antibodies. Epitope tags commonly used include Flag, HA, and cMyc, among others [6]. Here, we outline a simplified strategy for the isolation of epitope-tagged protein complexes from mammalian cell extracts. After generation and expansion of a stable cell line expressing an epitope-tagged protein of interest, the tagged protein and its associated proteins are purified from fractionated extracts using immobilized anti-epitope antibody, coupled with competitive elution using epitope peptides (Fig. 1). We demonstrate using this protocol to purify the Activating Signal Co-integrator Complex (ASCC) from nuclear extracts of cells stably expressing Flag-tagged ASCC subunits. Although this method is focused on nuclear extracts due to our focus on DNA repair complexes, any subcellular fractionation method can be coupled to this affinity purification technique [7].

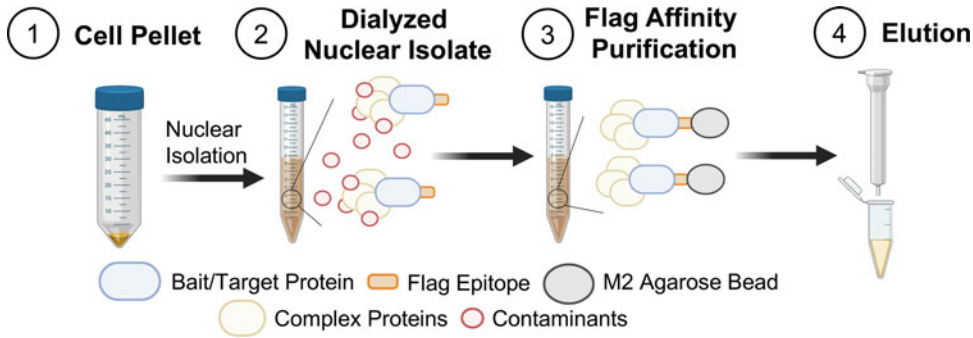


Fig. 1 Schematic for epitope-tagged protein complex purification from human cell nuclear extracts. (**a**) Cells expressing the epitope-tagged protein of interest are harvested and pelleted. (**b**) Nuclear extract is harvested, then dialyzed, and subsequently used for (**c**) flag affinity purification of the protein of interest and associated complex proteins. Contaminating proteins are removed via a series of washes, and the protein of interest and associated complex proteins are eluted (**d**) through addition of competing Flag peptide. The process may be continued for HA affinity purification to improve specificity

2 Materials

2.1 Production of Stably Expressing Proteins in Mammalian Cell Lines

1. 10 cm and 15 cm cell culture dishes.
2. 1.5 mL microcentrifuge tubes.
3. Virus packaging (HEK293T) cells.
4. Complete DMEM Media: DMEM plus 10% fetal bovine serum and 1× penicillin–streptomycin.
5. Trypsin-EDTA.
6. Target cell lines for expression. We often use HeLa-S3 cells due to their ability to grow in suspension. Adherent HeLa, HEK293T, or U2OS cells may also be used for this purpose.
7. 3 L and 10 L spinner flasks with magnetic stirrer.
8. Spinner flask incubator with magnetic stirrer plate.
9. Retroviral expression vector encoding epitope-tagged protein of interest.
10. Retroviral packaging vectors (*see* **Note 1**).
11. Mirus Transit293 reagent (Mirus Bio Cat# MIR 2700).
12. Opti-MEM (Thermo Fisher #31985070).
13. Polybrene solution (100 mg/mL initial stock: prepare in 1× phosphate buffered saline (PBS), and filter with 0.2 μm filter and keep sterile. 4 mg/mL [1000×] stock: dilute 100 mg/mL stock 1:25 with sterile 1× PBS. Keep sterile and store at −20 °C until ready for use).
14. Puromycin (2 mg/mL 1× PBS, filter-sterilized).
15. Blasticidin (5 mg/mL in 1× PBS, filter-sterilized).

2.2 Cell Extract Preparation

1. 40 mL and 15 mL Dounce homogenizers with "tight" pestle.
2. 1.5 mL microcentrifuge tubes.
3. 25 or 50 mL glass beakers.
4. Magnetic stir plate and small magnetic stir bars.
5. 3.0 mL syringes.
6. 23-gauge needles.
7. 15 mL and 50 mL plastic conical tubes.
8. 1× PBS.
9. PMSF solution: 100 mM in isopropanol.
10. Aprotinin, leupeptin, and pepstatin A at 1 mg/mL each.
11. β-Mercaptoethanol.
12. Hypotonic buffer: 10 mM Tris–HCl pH 7.3, 10 mM KCl, 1.5 mM $MgCl_2$.

13. Low salt buffer: 20 mM Tris–HCl pH 7.3, 20 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 25% glycerol.
14. High salt buffer: 20 mM Tris, pH 7.3, 1.2 M KCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 25% Glycerol.
15. BC100 dialysis buffer: 20 mM Tris, pH 7.3, 100 mM KCl, 0.2 mM EDTA, 20% Glycerol.
16. Swinging bucket centrifuge (for 15 mL and 50 mL tubes).
17. Fixed rotor centrifuge (for harvesting large volumes of cells).
18. Dialysis membrane tubing, 12,000 to 14,000 Dalton Molecular Weight Cut-Off.

2.3 Immunoaffinity Purification

1. 15 mL Falcon tubes.
2. 15 mL Falcon tube rotator.
3. 1.5 mL microcentrifuge tubes.
4. Plastic disposable chromatography columns.
5. Parafilm.
6. Anti-Flag resin: Anti-Flag M2 affinity gel (Sigma A2220).
7. Flag peptide: 5 mg/mL (Sigma F3290) dissolved in TAP wash buffer (see below).
8. Anti-HA resin.
9. HA peptide.
10. 100 mM glycine, pH 2.5.
11. 1.0 M Tris–HCl pH 7.9.
12. TAP wash buffer: 50 mM Tris–HCl pH 7.9, 100 mM KCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 10% Glycerol, 0.1% NP-40.

3 Methods

3.1 Generation of Cell Lines Expressing Epitope-Tagged Protein of Interest

Our preferred vector for constitutive protein expression is the pMSCV-Flag-HA plasmid with either puromycin or blasticidin selectable marker (*see* **Note 2**). The gene of interest is cloned in-frame into this vector, and retrovirus is made using a packaging cell line. The cell line of choice to be used for expression (e.g., HeLa-S3 cells; ATCC CCL-2.2) is then transduced and selected, and expression of the protein of interest is analyzed by Western blotting. Once confirmed, the cell culture is expanded and grown. While the choice of cell line is dependent on experimental parameters, we have had success using the pMSCV-TAP vector system to generate a variety of stable cell lines suitable for immunoaffinity purification on a large scale (10^7–10^8 cells). We recommend performing parallel mock purification from cells transduced with

either an empty pMSCV-TAP plasmid or with pMSCV-TAP-eGFP (*see* **Note 3**).

3.1.1 Transfection of Retroviral Vectors into Packaging Cells

Transient transfection of 293T cells is used to produce retrovirus for subsequent stable protein expression.

1. Plate HEK293T cells in P10 tissue culture dish using complete DMEM media.
2. When cells are 60–75% confluent, they are ready to transfect.
3. For each transfection, prepare a sterile 1.5 mL microfuge tube, and add 900 μL of OptiMEM media. To this, add 30 μL of Mirus Transit293 reagent drop-by-drop and mix by inversion. Incubate at room temperature for 5–10 min.
4. For each transfection, prepare a sterile 1.5 mL microfuge tube, and add in order: 100 μL of OptiMEM media, 1.5 μg of pVSV-G DNA, 1.5 μg of pGag-Pol DNA, and 7.0 μg of pMSCV-TAP retroviral vector containing the gene of interest.
5. Add the diluted Transit293/OptiMEM mix to the DNA mixture. Incubate at room temperature for at least 15–30 min.
6. Add the above transfection mixture drop-by-drop to the plate of HEK293T cells. Do not change the media during the transfection.
7. Recover the retroviral supernatant 48–72 h after transfection. Western blotting of whole cell lysates from transfected cells can be used to first confirm the desired tagged protein is being expressed, using antibodies directed against the Flag or HA epitope.
8. Filter supernatant with 0.45 μm filter, and proceed to retroviral transduction. Collected virus can be stored for up to 1 week at 4 °C or aliquoted and frozen at −80 °C.

3.1.2 Retroviral Transduction and Antibiotic Selection

1. Grow target cells of interest (e.g., HeLa-S or U2OS cells) in complete DMEM media. We typically plate ~5×10^5 cells per transduction in a P10 plate.
2. Incubate overnight at 37 °C.
3. Remove cell culture media from the plate (*see* **Note 4**).
4. Add 2 mL of the viral supernatant plus 10 μL of polybrene solution. Bring total volume of the plate to 10 mL using complete DMEM media.
5. Incubate for 24 h and then replace the viral supernatant and media with 10 mL of complete DMEM media.
6. Incubate for an additional 48 h at 37 °C.
7. Trypsinize the transduced cells into a new P10, and plate ~$1–2 \times 10^6$ cells.

8. Add the appropriate selection antibiotic. Puromycin is used at a final concentration of 1.0 μg/mL and blasticidin at 5 μg/mL (*see* **Note 5**).
9. Continue to select cells until they reach 80% confluency.
10. Before proceeding to large-scale growth, Western blotting of whole cell lysates from selected cells should be used to confirm the desired tagged protein is being expressed, using antibodies directed against the appropriate epitope (*see* **Note 6**).

3.1.3 Large-Scale Growth of Transduced Cells

1. Expand transduced and selected cells into 15 cm culture plates using complete DMEM media. Maintain appropriate antibiotic selection. Use 20–25 mL of media per plate.
2. When 90–95% confluent, expand to 4–5 × 15 cm plates. Continue growth until 90–95% confluent.
3. For adherent cell lines, we recommend continuing to expand using 15–20 × 15 cm plates. This ensures sufficient initial material (approaching 10^8 cells) for the subsequent affinity purification. For HeLa-S3 cells, expand the 4–5 × 15 cm plates into a 3 L spinner flask with 500 mL of media. We typically reduce serum to 5% at this point and maintain antibiotic selection. Put the spinner bottle on a magnetic stirrer in a 37 °C incubator, and stir at 50 rpm (*see* **Note 7**).
4. Monitor cell density daily. When the media start to change color to slightly yellow (usually in 3–4 days), count the cells. At approximately 4 × 10^5 per mL, add media to 1 L total volume. Additional antibiotic selection is not necessary at this stage.
5. Continue to grow to the desired volume as follows (with media containing 5% serum and no antibiotic), and collect at a maximum cell density of 1 × 10^6 cells/mL:
   (a) For 3 L total volume: once media change color, add media to total 3 L.
   (b) If the protein complex is relatively low in abundance, the suspension culture can be expanded further to improve purification (e.g., total of 16 L). Split the 3 L cell culture into 2 × 10 L spinner flasks, with 1.5 L initial culture in each flask. Add media to total 8 L per flask, and continue incubation.
6. Collect and spin cells at 4000 × *g* for 15 min. Discard supernatant.
7. Transfer cell pellet into 50 mL conical tubes with ice cold 1× PBS. Centrifuge and remove supernatant. Repeat the PBS wash to remove traces of media. For HeLa-S3 cells, the pellet volume will be approximately 1.5 mL per L of cultured cells.

8. Carefully remove all PBS, and proceed to nuclear extract preparation (*see* **Note 8**).

3.2 Preparation of Nuclear Extracts

We have found that initial cellular fractionation significantly reduces background contaminants in subsequent analysis of protein complexes. If a cytoplasmic fraction is needed for a particular protein complex, we refer the reader to an established protocol [8].

1. Add PMSF (to 0.2 mM final concentration), protease inhibitors (to 1 μg/mL final concentration), and β-ME (to 1.0 mM final concentration; *see* **Note 9**) immediately prior to use.
2. To the cell pellet, add at least 5× volume of ice-cold hypotonic buffer (HB). Very gently resuspend using a pipette aid.
3. Incubate on ice for 10–15 min, with inversion of the suspension every 2–3 min.
4. In a swinging bucket rotor, centrifuge at 4100 × *g* at 4 °C for 10 min. Carefully remove supernatant, and check cells under cell culture microscope. Most cells will appear swollen; some will be lysed.
5. Add 1× pellet volume HB to pellet, and resuspend with pipette aid.
6. Dounce 10–12 times with 15 mL Dounce homogenizer, with tight pestle. If a larger culture is used, use the 40 mL Dounce.
7. Check lysis under microscope. At least 90% of the cells should be lysed. If lysis appears lower, dounce five more times.
8. In a swinging bucket rotor, centrifuge at 5750 × *g* at 4 °C for 15 min. The pellet contains the nuclei, while the supernatant is pre-S100 cytoplasmic extract. The latter can be kept at −80 °C for later use, if desired.
9. Resuspend nuclear pellet in precisely 0.5× pellet volume low-salt buffer.
10. Transfer to 15 mL dounce homogenizer, and dounce seven to ten times to resuspend the nuclei.
11. Transfer the dounced nuclear material to a small beaker containing a magnetic stirrer in a cold room. Set up the stirrer next to a clamp.
12. Attach a 23-gauge needle to the 3.0 mL syringe, and remove syringe plunger. Place the syringe/needle directly above the nuclear extract in the beaker, with the needle pointing downwards. Add 0.5× pellet volume high-salt buffer to the syringe (*see* **Note 10**).
13. With the magnetic stirrer on (at 100–120 rpm), drip the high-salt buffer into the beaker using the syringe/needle at 4 °C. Continue stirring beaker to promote prompt mixing.

14. Once all the high-salt buffer has mixed in, use the plunger to ensure any of the buffer remaining in the needle has been transferred into the extract.
15. Stir for an additional 30–45 min.
16. Transfer the nuclear extract to 1.5 mL centrifuge tubes on ice; it will take five to eight of these to accommodate the volume (*see* **Note 11**). For larger volumes, transfer to a 50 mL conical tube instead.
17. Spin at 20,000 × *g* for 30 min at 4 °C.
18. The supernatant is the soluble nuclear extract, with the pellet being the remaining chromatin fraction. If desired, the pellet can be flash frozen at −80 °C if the chromatin-associated material is desired at this stage.
19. Dialyze overnight against 3 L of BC100 buffer (containing protease inhibitors and 1 mM β-ME) at 4 °C.

3.3 Immunoaffinity Purification

For Flag-HA tagged proteins, a one-step anti-Flag affinity purification usually suffices to purify the protein of interest and its associated partners with reasonable yield, although some contaminants are likely [9]. If desired, a subsequent anti-HA affinity purification will reduce associated contaminants but at a cost of yield.

3.3.1 Anti-Flag Immunoaffinity Purification

1. Retrieve dialyzed material. For smaller volumes (<10 mL), aliquot into 1.5 mL tubes. For larger volumes, transfer into 50 mL conical tubes. Spin at 20,000 × *g* for 30 min at 4 °C.
2. While spinning, prepare M2-agarose beads. Resuspend 100–250 μL of beads (bed volume) in 1.0 mL TAP wash buffer. Spin at 800 × *g* at 4 °C, and carefully remove supernatant.
3. Repeat the wash two times with 1.0 mL 100 mM glycine (pH 2.5).
4. Neutralize the pH by washing once with 1.0 mL 1.0 M Tris–HCl (pH 7.9).
5. Repeat the wash once with 1.0 mL TAP-wash buffer (*see* **Note 12**).
6. Transfer spun nuclear extract to 15 mL Falcon tube, reserving 1–5% of the extract volume as input. Add the washed beads to the nuclear extract.
7. Rotate the extract/beads at 4 °C for up to 4 h (*see* **Note 13**).
8. Spin down beads 800 × *g* at 4 °C, and carefully remove supernatant. Supernatant may be kept as unbound material.
9. Wash beads by resuspending in 10 mL of ice-cold TAP wash buffer (with protease inhibitors and β-ME). Spin down beads 800 × *g* at 4 °C, and carefully remove supernatant.

10. Repeat wash step two additional times.
11. In a cold room, transfer beads to 3 mL disposable column using a pipette tip. Use additional ice-cold TAP wash buffer to transfer any beads remaining in the original 15 mL tube.
12. Wash three times with 3 mL ice-cold TAP wash buffer. Allow buffer to go through column by gravity flow.
13. Transfer column to 15 mL Falcon tube. Spin briefly at 200 × *g* at 4 °C to remove traces of wash buffer. Keep the Falcon tube for use in **step 11** below.
14. Cap the bottom of the column. Add 0.5–1.0 mL of TAP-Wash Buffer plus 0.4 mg/mL Flag peptide.
15. Parafilm the top of the column. For extra safety, double seal with Parafilm. Transfer to 15 mL Falcon tube.
16. Rotate at 4 °C for 1 h, and retrieve eluate as your Flag eluate. Because repeated freeze–thaw cycles can decrease protein complex stability, we recommend immediately proceeding to HA-immunoaffinity purification or dividing elution into small aliquots (20–50 μL) prior to flash freezing and storage at −80 °C.

3.3.2 Anti-HA Immunoaffinity Purification

1. Prepare anti-HA (12CA5) antibody-conjugated beads as described earlier for the anti-Flag beads (*see* **steps 2–5** in the previous section). Use 20–50 μL (bed volume) of anti-HA antibody-conjugated beads (*see* **Note 14**).
2. Spin the beads at 800 × *g* for 1 min, and carefully remove the supernatant. Load 200–500 μL of the Flag-immunopurified material from above. Save 10–20 μL of the Flag-immunopurified for SDS-PAGE analysis.
3. Rotate beads at 4 °C for 2–4 h.
4. Spin down at 800 × *g* for 1 min, and transfer the supernatant to a new 1.5 mL tube. Save this unbound material for SDS-PAGE analysis.
5. Wash beads three times with 1.0 mL of TAP wash buffer (with protease inhibitors and β-ME).
6. Resuspend in 200 μL TAP wash buffer, and transfer to a micro-spin column (Bio-Rad Bio-Spin or equivalent) set up over a 1.5 mL tube.
7. Spin at 800 × *g* for 1 min.
8. Wash again with 200 μL TAP wash buffer by centrifugation at 800 × *g* for 1 min.
9. After centrifugation, close the bottom with the cap. Elute using 40 μL (or 2 bed volumes) of TAP wash buffer with 0.2 mg/mL HA peptide. Incubate at 4 °C for 1 hr, mixing occasionally.

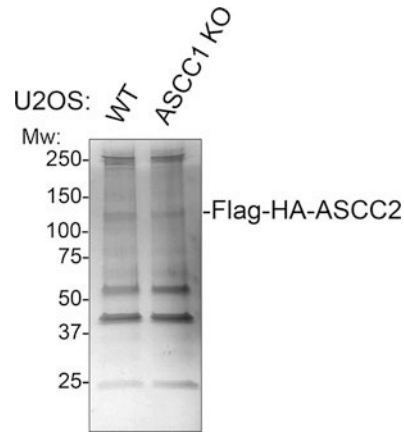


Fig. 2 Flag-HA-tagged ASCC2 complexes were purified from U2OS nuclear extract (WT and ASCC1 knockout). The Flag-eluted material was separated on 4–12% SDS-PAGE gel and silver stained

10. Recover the eluant by centrifugation at 800 × *g* for 1 min.
11. Because repeated freeze–thaw cycles can decrease protein complex stability, we recommend immediately dividing elution into 10–20 μL aliquots prior to freezing at −80 °C.

3.3.3 Analysis of Affinity Purified Complexes

The eluted material can be analyzed by SDS-PAGE and Western blotting to detect the presence of the bait protein as well as potential interacting partners. Prior to precipitating the eluted material for proteomic analysis by mass spectrometry, we typically analyze 2–10 μL by silver staining to determine relative abundance of the bait and co-purifying proteins (*see* example in Fig. 2). We prefer trichloroacetic acid (TCA) precipitation as a method for simultaneously concentrating and removing contaminating salts from our protein samples [10]. Such samples may then be analyzed by mass spectrometry, with pairwise cell lines or genetic backgrounds being used to compare interaction partners (*see* example in Fig. 3).

4 Notes

1. Our lab uses the pCMV-VSV-G (Addgene #8454) and pGAG/Pol packaging vectors for retrovirus production. These are also available upon request from the corresponding author. For a list of commonly used retroviral packaging vectors, we refer the reader to https://www.addgene.org/viral-vectors/retrovirus/.
2. The empty pMSCV-TAP vector is available on Addgene (#12570). Our laboratory uses Gateway-based pMSCV-TAP vectors, with puromycin or blasticidin selection markers. These are available upon request from the corresponding author.

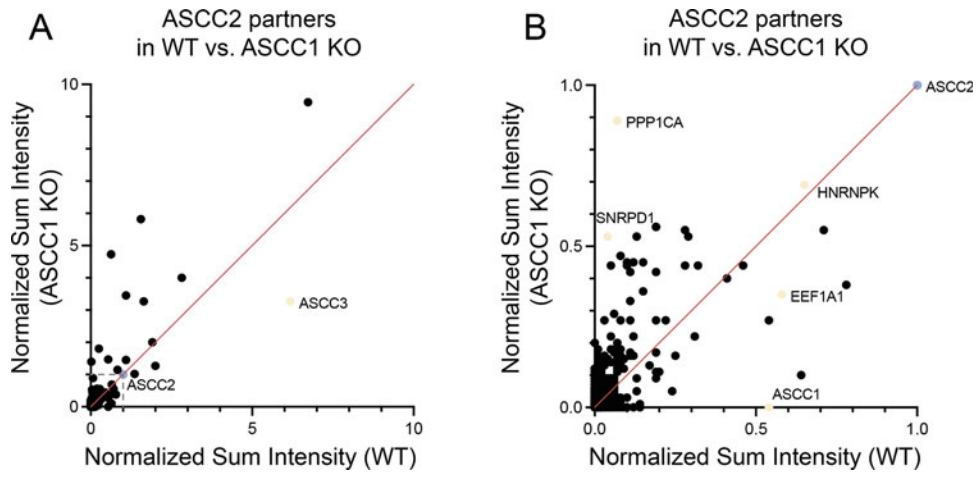


Fig. 3 Flag-HA-tagged ASCC2 was purified from WT or ASCC1 knockout U2OS cells and analyzed by LC-MS/MS. Ion intensities for identified proteins were normalized to ASCC2 bait ion intensity. Expanded view is shown on the right. Note complete loss of ASCC1 in the ASCC1 knockout cells relative to WT cells, and altered ion intensities for selected proteins, including ASCC3, PPP1CA, SNRPD1, and EEF1A1, with HNRNPK appearing unchanged

3. The pMSCV-eGFP vector is available upon request from the corresponding author.
4. HeLa-S3 cells grow as a mixed population of floating and attached cells. For these cells, collect media from the overnight cultured plate, add 5 mL fresh media to plate, and centrifuge collected cells at 200 × *g* for 2 min. Resuspend the centrifuged cells in 3 mL complete DMEM media, and transfer back to plate prior to resuming transduction.
5. We have noticed that different cell lines and different lots of antibiotics may require higher or lower concentrations for proper selection. Therefore, it is important to establish the proper amount of antibiotic necessary for each cell line a priori. At the appropriate concentration, puromycin and blasticidin will kill >90% of non-transduced cells within 3–4 days. Higher concentrations beyond this point are discouraged because this may affect growth of transduced cells.
6. If available, Western blotting should also be used to determine the expression of the tagged protein in comparison to the endogenous protein of interest. Ideally, the expression of the tagged protein should be no more than two- to threefold higher than the endogenous protein. Higher expression may prevent appropriate assembly of the target protein into physiological macromolecular complexes. If higher expression levels are seen, it is possible that the multiplicity of infection of the retrovirus needs to be reduced.

7. HeLa-S3 cells do not require CO_2 for their growth. However, they will grow at a moderately reduced rate under these conditions.
8. We routinely flash freeze these large cell pellets using dry ice/ethanol, followed by storage at −80 °C. Yields are not noticeably reduced from frozen cell pellets.
9. Higher amounts of reducing agent may be detrimental to immunoaffinity purification. This is due to the potential uncoupling of the antibody conjugate from the bead.
10. The precise volume of low salt buffer and high salt buffer are critical, as this determines the final concentration of KCl present in the nuclear extract. The protocol is designed such that the final concentration of KCl in this step is 300 mM, which we have found to be appropriate for most complexes of interest. At this concentration of KCl, there is at least partial dissociation of many chromatin-associated complexes from chromatin, without complete disassembly of the complexes themselves.
11. At this stage, the nuclear extract may have noticeably increased viscosity. This is normal. Measures to reduce viscosity (e.g., by sonication) should be avoided as this will fragment the chromatin and may lead to artifacts in complex purification.
12. These bead preparatory steps will dissociate any material bound to the antibody and may increase yield.
13. Overnight incubation with anti-Flag beads will significantly increase nonspecific binding without much increase in yield.
14. In general, we find that 20 μL of the HA beads provides sufficient capacity to pull down from the previous Flag affinity purification. However, we often find protein complexes have poor binding to the anti-HA antibody, perhaps due to poor accessibility to the epitope. If a significant amount of the unbound complex is found, increase the amount of the antibody beads.

Acknowledgments

Work in our laboratory is supported by the NIH (R01 CA193318, R01 CA227001, and P01 CA092584), the American Cancer Society research scholar program (RSG-18-156-01-DMC), and the Alvin J. Siteman Cancer Center Investment Program, which is supported by the Foundation for Barnes-Jewish Hospital Cancer Frontier Fund and the National Cancer Institute, Cancer Support Grant P30 CA091842. B.A.T. is supported by the NCI Ruth L. Kirschstein National Research Service Award (NCI F31CA254143).

References

1. Yu H, Braun P, Yildirim MA, Lemmens I, Venkatesan K, Sahalie J, Hirozane-Kishikawa T, Gebreab F, Li N, Simonis N, Hao T, Rual JF, Dricot A, Vazquez A, Murray RR, Simon C, Tardivo L, Tam S, Svrzikapa N, Fan C, de Smet AS, Motyl A, Hudson ME, Park J, Xin X, Cusick ME, Moore T, Boone C, Snyder M, Roth FP, Barabási AL, Tavernier J, Hill DE, Vidal M (2008) High-quality binary protein interaction map of the yeast interactome network. Science 322(5898):104–110
2. Deane CM, Salwiński Ł, Xenarios I, Eisenberg D (2002) Protein interactions: two methods for assessment of the reliability of high throughput observations. Mol Cell Proteomics 1(5):349–356
3. Kuchaiev O, Rasajski M, Higham DJ, Przulj N (2009) Geometric de-noising of protein-protein interaction networks. PLoS Comput Biol 5(8):1000454
4. Dunham WH, Mullin M, Gingras AC (2012) Affinity-purification coupled to mass spectrometry: basic principles and strategies. Proteomics 12(10):1576–1590
5. Yoshihiro N, Vasily O (2003) Immunoaffinity purification of mammalian protein complexes. Methods Enzymol 370:430–444
6. Kimple ME, Brill AL, Pasker RL (2013) Overview of affinity tags for protein purification. Curr Protoc Protein Sci 73:9.9.1–9.9.23
7. Brickner JR, Soll JM, Lombardi PM, Vågbø CB, Mudge MC, Oyeniran C, Rabe R, Jackson J, Sullender ME, Blazosky E, Byrum AK, Zhao Y, Corbett MA, Gécz J, Field M, Vindigni A, Slupphaug G, Wolberger C, Mosammaparast N (2017) A ubiquitin-dependent signalling axis specific for ALKBH-mediated DNA dealkylation repair. Nature 551:389–393
8. Tomomori-Sato C, Sato S, Conaway RC, Conaway JW (2013) Immunoaffinity purification of protein complexes from mammalian cells. Methods Mol Biol 977:273–287
9. Trinkle-Mulcahy L, Boulon S, Lam YW, Urcia R, Boisvert FM, Vandermoere F, Morrice NA, Swift S, Rothbauer U, Leonhardt H, Lamond A (2008) Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. J Cell Biol 183(2):223–239
10. Link AJ, LaBaer J (2011) Trichloroacetic acid (TCA) precipitation of proteins. Cold Spring Harb Protoc 2011:993–994

Chapter 4

Universally Accessible Structural Data on Macromolecular Conformation, Assembly, and Dynamics by Small Angle X-Ray Scattering for DNA Repair Insights

Naga Babu Chinnam, Aleem Syed, Kathryn H. Burnett, Greg L. Hura, John A. Tainer, and Susan E. Tsutakawa

Abstract

Structures provide a critical breakthrough step for biological analyses, and small angle X-ray scattering (SAXS) is a powerful structural technique to study dynamic DNA repair proteins. As toxic and mutagenic repair intermediates need to be prevented from inadvertently harming the cell, DNA repair proteins often chaperone these intermediates through dynamic conformations, coordinated assemblies, and allosteric regulation. By measuring structural conformations in solution for both proteins, DNA, RNA, and their complexes, SAXS provides insight into initial DNA damage recognition, mechanisms for validation of their substrate, and pathway regulation. Here, we describe exemplary SAXS analyses of a DNA damage response protein spanning from what can be derived directly from the data to obtaining super resolution through the use of SAXS selection of atomic models. We outline strategies and tactics for practical SAXS data collection and analysis. Making these structural experiments in reach of any basic and clinical researchers who have protein, SAXS data can readily be collected at government-funded synchrotrons, typically at no cost for academic researchers. In addition to discussing how SAXS complements and enhances cryo-electron microscopy, X-ray crystallography, NMR, and computational modeling, we furthermore discuss taking advantage of recent advances in protein structure prediction in combination with SAXS analysis.

Key words Protein structure, SAXS analysis, Conformational flexibility, DNA repair, RNA, Endonuclease

1 Introduction

DNA repair pathways are highly regulated and coordinated, often through dynamic conformations and assemblies, as seen for TFIIH that adopts distinct functional conformations depending on the assembly context in transcription and DNA repair [1]. As their repair intermediates are often more toxic than the damage itself, many DNA repair enzymes chaperone the intermediate, through product inhibition and required coordination with the next enzyme

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_4,

in the pathway to release their product [2–5]. For example, nicks formed by repair pathways to cut out the DNA damage [1, 6, 7] create a risk for the cell, as fraying can lead to 3′ overhangs that can invade sister strands [1, 6, 7]. Thus, DNA repair nucleases that process ends or nick the DNA represent the committed step, as seen for the MRE11 nuclease in DNA double-strand break repair (DSBR) or WRN exonuclease in processing ends [8, 9]. As a result, they strictly validate their substrates and need to be licensed for incision [10].

This type of coordination is often enabled through the DNA repair protein's intrinsic ability to adopt multiple conformations, protein-induced DNA conformations, and DNA-induced protein conformations, making small angle X-ray scattering (SAXS) an effective tool for structure-based analyses and for decoding the structural mechanisms of DNA damage responses and DNA repair. Although X-ray crystallography and NMR can provide precise atomic structures, systematic analyses of their accuracy show that X-ray structures are too rigid and NMR structures are too flexible [11, 12]. Yet, for DNA repair and damage responses ranging from oxidized base repair to DSBR, we have found that accurate measures of flexibility and conformational change by SAXS helped decipher functional mechanisms, as exemplified by NEIL1 intrinsically disordered tail [13–15] and ATP-driven RAD50 states [16, 17]. SAXS provides an accurate measure of the solution ensemble and the means to assess conformational changes and states critical to DNA repair activities and valuable for enhancing X-ray, cryo-electron microscopy (EM), and NMR structural analysis [18–20].

As an important component to their function, DNA repair proteins face the difficulty of differentiating their target DNA damage from the much more populated undamaged DNA [10, 15]. With rigid protein structural features, they often sterically mold and distort the DNA to check for the presence of damage or other specific characteristics of their substrate. For example, glycosylases and apurinic/apyrimidinic endonucleases use phosphate backbone pinching to test for disrupted base stacking that allows for flipping out of the nucleotide or phosphodiester into damage-binding rigid structural features [21–24]. Indeed, stable binding to flipped out alkylated DNA bases can mark the base damage and enable a handoff from base to nucleotide excision repair (NER) for efficient damage removal [25, 26]. For excision enzymes, only if the flipped out DNA can be retained is activity enabled. As a more elaborate example, flap endonuclease (FEN) is a structure-specific endonuclease that uses DNA distortion and DNA-induced protein conformational changes to validate the presence of a 5′ flap (varying length) and a one nucleotide 3′ flap within double-stranded (ds)DNA before licensing incision [27–29]. MRE11/RAD50 also undergoes dramatic global conformational changes that

include validation of dsDNA ends by RAD50 coiled regions [30, 31].

To coordinate repair and reduce the risk of toxic intermediates, repair enzymes often chaperone the intermediate to the next enzyme present. Indeed, there is growing appreciation for the multi-protein assemblies of DNA repair enzymes. In DSBR and replication restart, there is a coordinated assembly of proteins at DNA ends [32–38]. DNA end joining requires the DNA ends to be brought together, processed, and positioned for ligation, which, in turn, requires flexibility in the proteins bound to the DNA ends and in the coordinating, scaffold proteins such as XRCC1 [39, 40]. In NER, the extreme precision of the excised oligonucleotide supports the presence of a TFIIH-centered NERsome-based ruler that strictly dictates where and when the incision sites occur relative to the lesion [1, 41].

Given that dynamic features and assemblies are essential elements in DNA repair functions, knowledge of solution structures and states of DNA repair proteins helps in deciphering their mechanisms [42]. SAXS is an enabling technique to structurally characterize proteins in solution under near physiological conditions and with super resolution [43]. *Enabling integrative structural biology,* SAXS results complement and enhance structural results from cryo-EM, X-ray crystallography, NMR, and computational modeling [20, 43–46]. A major reason for obtaining SAXS data is that collecting SAXS data is straightforward and essentially available to any scientist who has protein, RNA, or DNA [47–50]. No labeling or crystallization is required. The size limitation at our SIBYLS beamline ranges from a 8 kD to 600 kD [20]. Expensive in-house SAXS equipment is not necessary, as data collection for research is typically provided for free by all biological SAXS beamlines, with one at almost every synchrotron [20, 50].

With potential for clinical implications and actionable knowledge, these considerations make SAXS experiments important for DNA repair complexes and readily accessible and relevant to both clinical and basic researchers. For example, SAXS provided insights into a super responder mutation identified in a cancer study; a time course SAXS study showed that the RAD50 super-responder mutation kinetically slowed the conformational change without significantly altering the beginning and end conformational states [17]. At a more advanced level, Gold-SAXS was useful for measuring MutS DNA conformations and dynamics [51], and the current approach to X-ray Scattering Interferometry with gold nanoparticle-conjugated DNA is provided elsewhere in this issue [52].

For SAXS theory, readers are directed to excellent reviews [44, 45, 53], as here we describe prototypic and practical methods for investigators interested in using SAXS to explore DNA repair proteins. Simplistically, the SAXS from proteins in solution contains

information on a statistical distribution of electron pair distances referred to as a pair-wise distribution or real space function, $P(r)$, analogous to the Patterson function in X-ray crystallography [54]. This distribution provides information on molecular mass of particle in solution (e.g., stoichiometry), protein density (e.g., flexibility), shape, and whether or not an atomic model is consistent with the experimental solution data. If vectors are drawn from every electron to every other electron and are sorted based on vector length, this is analogous to the real space $P(r)$ curve representation of the SAXS data. A misconception is that SAXS only detects the protein surface; like crystallography, SAXS measures the coherent scattering from all electron pairs (surface and internal) [44, 47, 55].

Key structural information can be derived from SAXS data [19]. First, SAXS provides information on the stoichiometry of proteins in solution; over half of all bacterial proteins form multimers in our structural genomics study [47]. For eukaryotic DNA repair systems, these proteins act in complex assemblies [35, 56–58]. Second, SAXS can test and define ligand-induced changes in assembly [46, 59, 60]. Third, SAXS quantitatively measures protein density, which can be a reflection of protein flexibility [19]. Fourth, distances can be derived from atomic models and used to predict SAXS curves that can be compared to the experimental data [18, 61, 62]. If the atomic model reflects the true conformations and assemblies occurring in solution, the predicted SAXS curve will match the experimental SAXS curve. If multiple conformations are occurring in solution, an ensemble of conformations can be put together, and their predicted scatter can be compared to the experimental data. If two domains are moving relative to each other, the distances of many electron pairs will change, so SAXS may provide experimental data for computational analyses of transient electrostatic orientation and interactions [63]. A flexible region which alters many distances is detectable in SAXS. A caveat is that if the proportion of well-folded region of the protein is large relative to the disordered region, then detection of flexibility is difficult.

Due to the powerful Fourier transform relationship of X-ray scattering to structure, scattering data can be represented in two ways. In reciprocal space, scattering curves of intensities I are plotted as a function of low momentum transfer q, or scattering angle. In real space, SAXS is shown as histograms of relative proportion P of electron pairs at distance r, i.e., $P(r)$. Thus, investigators can do analyses in either reciprocal space or real space.

At the SIBYLS beamline, SAXS data can be collected in high-throughput (HT) SAXS mode or Size-Exclusion Chromatography-coupled (SEC) SAXS. HT-SAXS requires stoichiometrically monodisperse protein solution (20 μl of a 0.5–2 mg/ml) and the corresponding buffer. Users mail in their samples in a sealed 96-well plate for analyses. The advantage of HT-SAXS is that proteins can be collected at higher concentration and more samples

can be collected at a time. SAXS data is collected as consecutive .dat files that are averaged to obtain the highest signal-to-noise possible. For well-behaved monodisperse samples, HT-SAXS works well [47]. For difficult proteins or in vitro-folded RNA samples that aggregate or for complexes that are equilibrating between the bound and unbound state, SEC-SAXS is best as SAXS data is collected as the sample is eluted from the SEC [20, 64, 65]. Although there is at least fourfold dilution, separation of aggregates and stoichiometrically heterogeneous samples is possible and is the primary advantage of SEC-SAXS. When there is slow conformational change on the minute time-scale, researchers are even able to separate conformationally heterogeneous but stoichiometrically homogenous samples [40, 66]. SAXS data is collected as sequential .dat files that are averaged over the elution to obtain the highest signal-to-noise possible. If there are overlapping peaks indicating a mixture of assemblies or conformations, the SAXS signal from stoichiometrically monodisperse populations can be extracted from SEC-SAXS data [67, 68]. Thus, SEC-SAXS is particularly powerful for probing dynamic DNA repair assemblies [20, 40, 66].

Here we provide an exemplary SAXS analysis for a well-behaved alkylation damage response protein, based on what we have found to be an efficient and robust approach. We provide directions for both HT-SAXS and SEC-SAXS. We describe three ways to prepare samples and equivalent buffer for HT-SAXS. Additional instructions are available at the SIBYLS website for HT-SAXS and SEC-SAXS (https://bl1231.als.lbl.gov/saxs_protocols/). Once a single SAXS curve has been calculated, both HT-SAXS and SEC-SAXS data can be processed similarly.

2 Materials

1. Protein(s). The concentration of protein required to get a good signal-to-noise depends on the size of the protein, as proteins scatter to the square of the mass. Our beamline has found that the rule of 100 gives an optimal signal (Molecular weight (kD) × concentration (mg/ml) = 100). Thus, for a 50 kD protein, 2 mg/ml concentration at the beam gives a good signal. It is possible to go lower in protein concentration, but there is the risk of lower signal-to-noise and confidence in the conclusions. For complexes, it is the aggregate mass of the complex that determines the scattering.
   (a) For HT SAXS where the protein is directly loaded into the SAXS sample cell, the rule of 100 works fine. The SIBYLS HT-SAXS sample cell requirement is 30 μl at a concentration of 0.5–2 mg/ml.

(b) For SEC-SAXS, depending on how many peaks your protein will elute in, expect about a four- to sixfold dilution if there is only one primary peak. At SIBYLS, SEC-SAXS requires 50–100 μl of 2–10 mg/ml protein.

(c) If analyzing complexes, collect data on individual proteins, in addition to data on the complex. Comparison of the SAXS signal from individual components and the complex can be used to confirm the presence of the complex. It's best to analyze complexes by SEC-SAXS, to obtain stoichiometrically homogenous population of the complex and separate out unbound protein. Smaller proteins or DNA can be added in 1.2 molar excess to increase the complex concentration. For complexes with low affinities and not stable over SEC (*see* **Note 1**).

2. Buffer. The most important condition of the buffer is that your protein is stable. Salt is fine up to 0.5 M. We recommend 1% glycerol to prevent radiation damage. If buffer contains >2% glycerol, it will be hard for the sample loading in HT-SAXS or raise column pressure in SEC-SAXS. Use detergent over the critical micelle concentration or sucrose with caution. For detergent, 0.1% OG is under the CMC and works well [7]. *See* **Note 2** for DNA repair complexes.
3. For SIBYLS HT-SAXS, Corning Axygen® 96-well Polypropylene PCR Microplate, Full Skirt, Clear, Nonsterile, Product Number PCR-96-FS-C. This plate has been calibrated at our beamline. Other plates may break the sample loading needle.
4. To protect samples in your HT-SAXS plate, Corning Axygen® AxyMats™ 96 Round Well Sealing Mat for PCR Microplates, Nonsterile works sufficiently.
5. For HT-SAXS alternative dialysis sample preparation, 50 μl dialysis buttons, Hampton HR3-326. Spectra/Por dialysis membrane (varying pore sizes).
6. For HT-SAXS alternative SEC or concentrator-based sample preparation, 4 ml protein concentrator, one per protein sample.
7. HT SAXS Sample Preparation.

(a) For SIBYLS HT-SAXS, Corning Axygen® 96-well Polypropylene PCR Microplate, Full Skirt, Clear, Nonsterile, Product Number PCR-96-FS-C. This plate has been calibrated at our beamline. Other plates may break the sample loading needle.

(b) To protect samples in your HT-SAXS plate, Corning Axygen® AxyMats™ 96 Round Well Sealing Mat for PCR Microplates, Nonsterile works sufficiently.

3 Method

3.1 Obtaining SAXS Data Collection Time at the SIBYLS Beamline

We utilize an easy-to-use web server. Each SAXS beamline has its own application process. The beamline scientists can be contacted with any questions or concerns about setting up the SAXS experiment. Here, we provide instructions on how to obtain beamtime at our beamline. The SIBYLS beamline mail-in SAXS page has information on how to obtain beamtime.

1. Go to the SIBYLS beamline website for applying for beamtime. https://bl1231.als.lbl.gov/htsaxs
2. Register for a SIBYLS mail-in SAXS account. https://bl1231.als.lbl.gov/htsaxs/request_account/new
3. Register for an ALSHub account with the Advanced Light Source at https://alshub.als.lbl.gov/.
4. Log in to ALSHub to submit a RAPIDD proposal any time before sending samples.
5. Log into SIBYLS beamline account to book sample slots at https://bl1231.als.lbl.gov/htsaxs/users/sign_in.
6. Click on "Book Slot."
7. The page will show the available beamtimes, for HT or SEC-SAXS.
8. Follow sample preparation and shipping instructions on the HT-SAXS page https://bl1231.als.lbl.gov/htsaxs/instructions/htsaxs or the SEC-SAXS page https://bl1231.als.lbl.gov/htsaxs/instructions/secsaxs.

3.2 HT-SAXS Sample Preparation

There are three ways that HT-SAXS samples can be prepared, all to obtain the best possible buffer for subtraction. Samples need to be loaded onto 96-well plates and mailed to the beamline.

3.2.1 HT-SAXS Sample Preparation by Dialysis

1. Prepare Spectra Por dialysis membranes according to instructions.
2. Assemble sample in 50 μl Hampton dialysis button (HR3–326), per instructions.
3. Equilibrate overnight in buffer at 4 °C.
4. Load 96-well plate with buffer from dialysis, protein sample, buffer from dialysis.

3.2.2 HT-SAXS Sample Preparation by Concentrator

1. Wash protein concentrator (e.g., Amicon Ultra 4) three to five times with similar buffer. Chemicals used for long-term storage of the filters in concentrators will throw off buffer subtraction.
2. Load concentrator with protein.
3. Spin for 1 min. Throw away first eluant.

4. Spin until obtaining enough flow through (FT) to load 30 μl into sample plate.
5. Multiple protein concentrations are recommended (e.g., 1, 2, 5 mg/ml) to check for concentration-dependent multimerization (*see* **Note 3**).
6. Spin 1–2 min more until ~2× concentrated. Mix sample, and pull out 30 μl aliquot for loading onto SAXS plate.
7. Spin 10–15 min more (~4× concentrated). Mix sample, and pull out 30 μl aliquot for loading onto SAXS plate.
8. Use concentrator FT for buffer subtraction.
9. Load 96-well plate with buffers and then original and concentrated samples at increasing concentration and then another buffer. If the original starting protein concentration is low, loading three buffer wells before the sample provides better signal to noise, based on our experience (e.g., buffer, buffer, buffer, low, medium, high, buffer).

3.2.3 HT-SAXS Sample Preparation by SEC

1. Use a 24 ml SEC column that will best separate out the target peak from larger aggregates. Particles scatter X-rays to the square of the mass, so larger particles will scatter disproportionately greater than smaller particles. E.g. A dimer will scatter four times more than a monomer.
2. Equilibrate column with three column volumes (CV) (*see* Note 4).
3. Load 300 μl of 10–20 mg/ml, and collect 300 μl fractions throughout elution.
4. Collect and keep separate 30 μl aliquot of the peak fraction (n) and the two fractions following (n + 1 and n + 2) for loading onto SAXS plate to get truly monodisperse sample. Even at low concentration, they are important as a standard to validate lack of aggregation in the concentrated samples. Often the peak fraction may be contaminated with a small amount of aggregated sample, so save some to compare n with n + 1. If there is aggregation, n + 1 can be compared with n + 2. After that, it is generally too dilute to be useful.
5. Save buffer from the column eluant before void volume for SAXS buffer subtraction (generally, 7–8 ml range) (*see* **Note 2**).
6. If the fractions are low protein concentration (e.g., <2–5 mg/ml) after going through the SEC column, concentrate each individual n and n + 1 gel filtration fractions separately, as described above in Subheading 3.2, **step 3** (*see* **Note 3**).
7. Spin 1 min. Throw away first FT.
8. Spin 1–2 min more until ~2× concentrated. Mix sample, and pull out 30 μl aliquot for loading onto SAXS plate.

9. Spin 10–15 min more (~4× concentrated). Mix sample, and pull out 30 μl aliquot for loading onto SAXS plate.
10. Use concentrator FT for buffer subtraction.
11. Load 96-well plate with three buffers and then the SEC fractions in reverse. For original SEC fractions, load seven wells with buffer, buffer, buffer, $n + 2$, $n + 1$, n, and buffer. For concentrated fractions, load six wells with FT, FT, FT, 2× conc, 4× conc, and FT. If there isn't enough FT, use the SEC eluant for the first two wells.

3.3 SEC-SAXS Sample Preparation

1. Mail one to seven proteins in Eppendorf tubes to SIBYLS beamline the week of the scheduled data collection. Send 50–100 μl of 1–5 mg/ml protein.
2. Mail 50 ml 10× buffer filtered with 0.22 μm filter. If sending frozen, separate out 50 ml into two tubes to prevent tubes breaking because buffer volume increased after freezing.
3. Select one of three Shodex columns for the SEC. 802.5 has size exclusion limit of 150 kDa; 803, 700 kDa; and 804, 1000 kDa. The most important thing is to separate out the protein peak from aggregate in the void, so when the expected protein mass is close to the size exclusion limit, use the column with the next bigger size exclusion limit.
4. Fill out SEC-SAXS form with column preference, buffer, and samples.

3.4 HT-SAXS Data Averaging

1. The SIBYLS beamline will collect ~30 consecutive 0.3 s exposures from each well and process the data with information on the wavelength, sample-to-detector distance, and other parameters based on a scattering control. The beamline will subtract the sample well with the closest buffer before and after the sample. Individual folders will be created with the exposures/frames for the buffer before subtraction, the buffer after subtraction, and an average.
2. The beamline will do a preliminary analysis for data quality (aggregation, low signal to noise) and technical errors (e.g., bubbles in the sample cell) and a recommendation on which data subtraction to use and send out information on how to access the HT-SAXS data.
3. Download the data to your hard drive.
4. Open the webserver FrameSlice (https://bl1231.als.lbl.gov/ran) on your browser. The SIBYLS beamline has a YouTube video tutorial on how to analyze HT-SAXS data and use FrameSlice (https://www.youtube.com/watch?v=c4-a7dEMAeY).

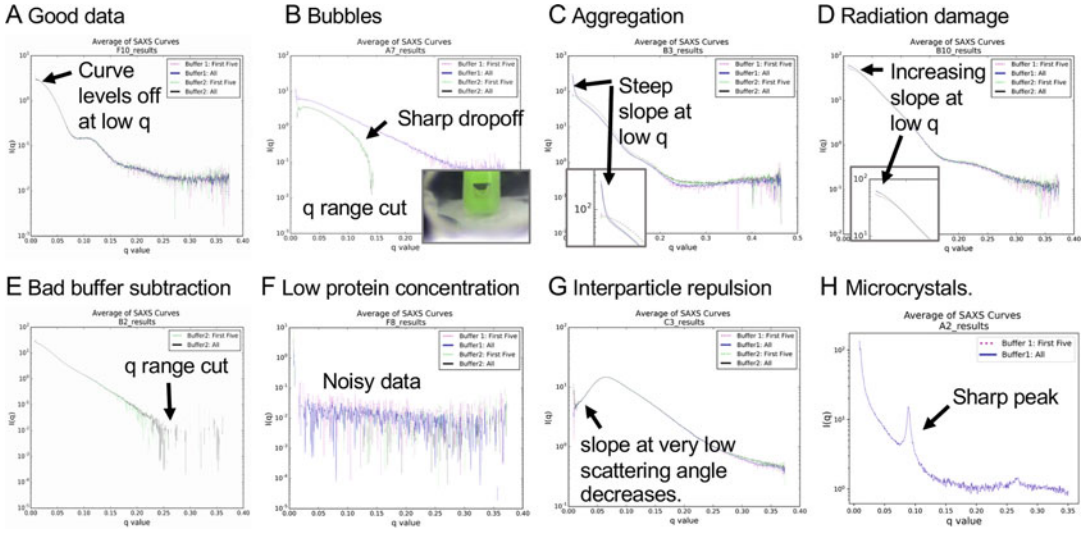


Fig. 1 Examples of good and bad SAXS data. (**a**) Good data shows that intensity levels off as *q* goes toward 0. (**b**) Bubbles in the sample or buffer will cause a sharp drop-off of data. (**c**) Aggregation is detected when the intensity increases as *q* goes toward 0. (**d**) Samples often aggregate from radiation damage, and aggregation appears and becomes worse over multiple exposures. Merge frames before the first appearance of aggregation. (**e**) Mismatched buffer will cause bad buffer subtraction and is indicated when the SAXS data is cut off at a lower *q*. (**f**) When the protein concentration is too low, even the low *q* region (0–0.2. $Å^{-1}$) which is where the signal-to-noise of the data should be high. (**g**) Interparticle repulsion occurs when the molecules in solution are not randomly oriented relative to each other. Intensity decreases as *q* goes toward 0. (**h**) Microcrystals are rare, but can provide useful information. Typically features in the SAXS data are smooth, but microcrystal diffraction will look like sharp peaks in the curve

5. Look for the following experimental errors [50] in the intensity vs. scattering angle (q) on FrameSlice (Fig. 1).
   (a) Good data will go the full width, and the intensity will show a leveling off as the scattering angle q goes to 0. If the SAXS data has curves, the protein is more spherical and globular. If the SAXS data is fairly straight, the protein is likely extended and/or rod shaped.
   (b) Bubbles in the sample cell or buffer cell gives data with a sharp drop-off of signal, and data does not extend to the full scattering angle q range. The SIBYLS beamline provides images of the sample cell for each exposure, so that users can check them manually.
   (c) Aggregation in the Guinier region is revealed by a sharp increase in Intensity as the scattering angle goes to zero. Monodisperse samples will level off close to zero.
   (d) Radiation-induced aggregation is evident from aggregation (*see* **step c**) that increases after multiple consecutive exposures. Sometimes, there are also changes in other parts of the scattering curve, indicating a change in

conformation. Select frames that match the first exposure, before radiation damage occurs.

(e) Bad buffer subtraction is suggested when data does not extend to the full q range. It may be possible to adjust the buffer subtraction, but this requires manual intervention by the beamline scientist.

(f) Too low protein concentration is evident from significant noise in the sample even at low scattering angle q, where the signal should be the best.

(g) Interparticle repulsion is observed when the intensity decreases as the scattering angle q goes to 0. If this occurs, the current data set is not usable. Data should be recollected with higher salt in the buffer and/or lower protein concentration (*see* **Note 5**).

(h) Microcrystals are unusual but do occur. Instead of a smooth line in the reciprocal SAXS curve, there will be small, sharp peaks. This information can provide interesting distance information for your systems [69].

6. If data quality is fine, average the file as demonstrated, and go to analysis Subheading 3.8.

3.5 SEC-SAXS Data Averaging

SEC-SAXS data analysis should take into account any additional information collected, such as in-line Multi-Angle Light Scattering (MALS). Most SEC-SAXS beamlines, including the SIBYLS beamline, have inline MALS.

1. The MALS elution profile should be examined for monodispersity. A single peak well-separated from aggregate or other peaks indicates monodispersity of the protein sample and is ideal for SAXS analysis. The molecular mass estimate through the peak ideally should be flat. If the molecular mass decreases, particularly for a complex, it indicates that the protein sample may be stoichiometrically heterogeneous.
2. The MALS elution profile should also be examined for places of high light scattering fractions, even without strong UV signal. These regions should be avoided during buffer subtraction.
3. Select buffer frames and merge SAXS exposures, following program tutorials and/or instructions. The SIBYLS beamline collects ~600 consecutive exposures over the course of the SEC elution. There are several programs for merging SEC-SAXS frames. For well-behaved samples, we use SCÅTTER IV (the program is open source and downloadable free at https://bl1231.als.lbl.gov/scatter/) [68], which is good for well-behaved monodisperse peaks. For elution profiles with overlapping peaks, the program RAW has an excellent option for separating out SAXS signal from different populations (e.g.,

monomer and dimer) using singular value decomposition (SVD) analysis and evolving factor analysis (EFA) (https://bioxtas-raw.readthedocs.io/en/latest/index.html) [31, 68]. For additional SEC-SAXS reduction issues (*see* **Note 6**).

3.6 Calculation of Metrics from HT-SAXS or SEC-SAXS Data

There are multiple programs for carrying out the SAXS data analysis, including SCÅTTER [68], RAW [67], and the ATSAS suite [70]. All programs are free for academic users.

1. Guinier analysis provides the R_g and is a quality control against aggregation and can be done in all three programs (Fig. 2). The Guinier plot should be linear, such that the residuals are unbiased on either side of the line and do not smile or frown. Points can be removed or added in this region. Sometimes, there is noise from the scatter of the beamstop that needs to be removed. Aggregation is indicated by points that fall above the line at the low q end of the Guinier plot. For SEC-SAXS, there should be no aggregation as long as the peak is not close to the void. For HT-SAXS, aggregation can be an issue. If aggregation is mild, some points can be removed. Select the maximum q used in the Guinier plot when maximum $q \times R_g < 1.3$ (*see* Fig. 2a).
2. The molecular mass in solution can be determined based on Volume of Correlation and R_g [71], implemented in both SCÅTTER and RAW. This estimate is typically within 10% of the true molecular mass in solution. For some SAXS data, the curve should be reduced down to $q < 0.32$, as this range was used to develop the constants to calculate the molecular mass.
3. Flexibility by the normalized Kratky analysis, $(I(q)/I(0)) \times (q \times R_g)^2$ vs. $q \times R_g$, can be done in SCÅTTER. A well-folded molecule will be above the crosshairs, which indicate the Guinier-Kratky point of $(\sqrt{3}, 1.1)$ [68, 72].
4. Quantitative flexibility metric based on Porod-Debye exponent can be determined using flexibility and volume analyses in SCÅTTER, as described [68]. The Porod-Debye exponent (P_x) reflects the protein density and provides a quantitative metric of flexibility [19, 45]. For well-folded proteins, P_x is 4 and the Porod-Debye plot shows a plateau (Fig. 2). For proteins with flexible domains, P_x is between 3 and below 4, depending on the level of flexibility. For denatured proteins, P_x is 2. Please *see* **Note 7** for additional information.
5. Checking for complex in SCÅTTER. If SAXS data is available for individual components and the complex, it is possible to test if the components are truly forming a complex or not interacting. If the components are not interacting, the SAXS data for the individual components can be added together to

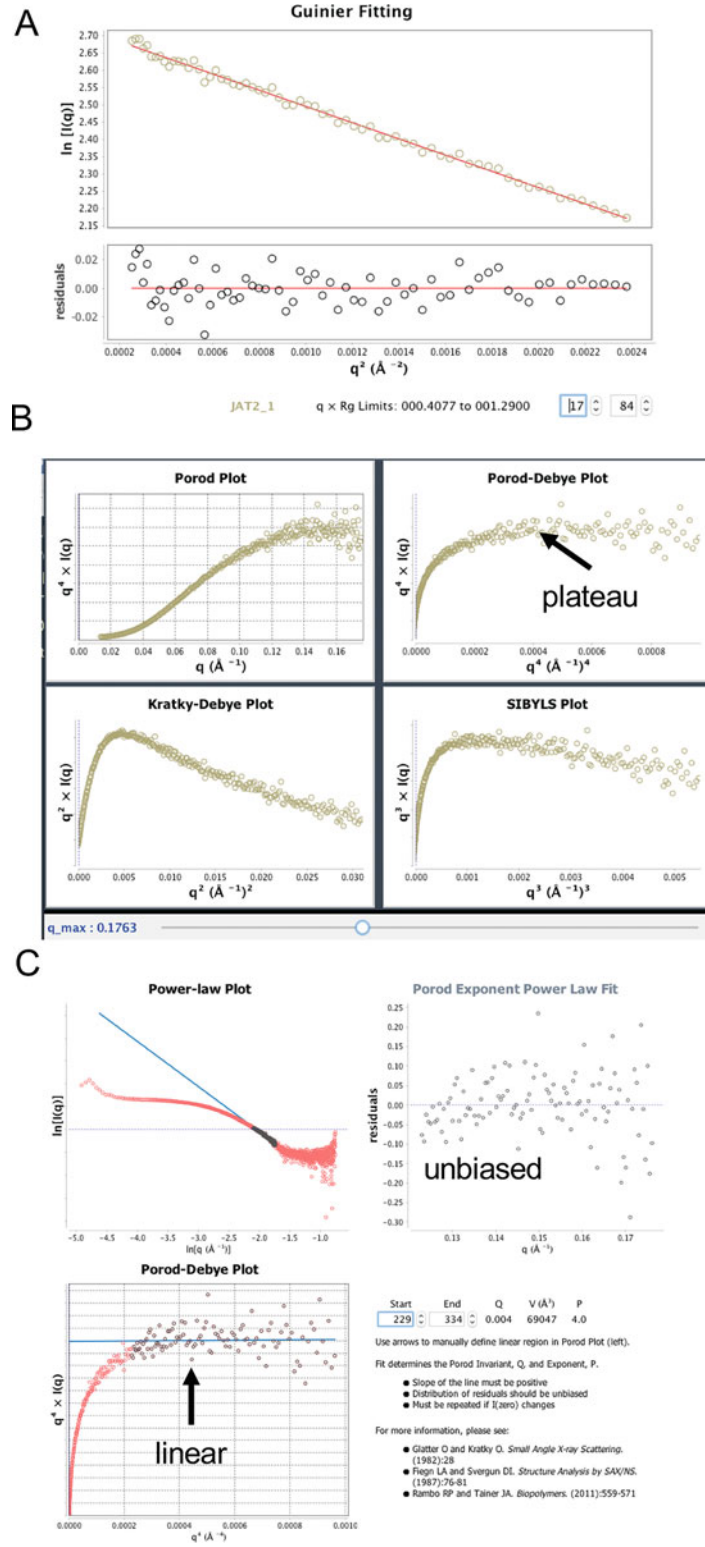


Fig. 2 SAXS metrics in SCÅTTER. Screenshots from SCÅTTER. (**a**) Reciprocal Space Guinier Plot. (**b**) Flexibility plot with slider placed so that one curve

obtain the "complex" data. If the components are truly forming a complex, this would not be true. Unbiased residuals indicate that the components are forming a complex.

6. Both SCÅTTER and GNOM implemented in PRIMUS program in the ATSAS program suite can generate the real space $P(r)$ plot [68, 73] (Fig. 3). Selection of the D_{max} should be manually judged. For most proteins, the curve should come down to the baseline smoothly. For ring-like PCNA, the P (r) will approach the baseline more abruptly at Dmax. Significant flexibility will be indicated by a "tail" near D_{max}. Real space estimates are shown for R_g. For the conformationally heterogeneous proteins in solution, the reciprocal R_g and the real space R_g do not necessarily match.

3.7 Comparison of Atomic Models to Experimental SAXS Data by FOXS

There are multiple programs for predicting SAXS curves from atomic models and comparing them to experimental SAXS data [74–79]. *See* **Note 8** for additional information and caveats. Here, we present how to do this in FOXS server (https://modbase.compbio.ucsf.edu/foxs/).

1. Input PDB name or upload a PDB file or multiple PDB files compressed together as a zip file. SAXS is sensitive to all electrons that belong to the molecule in solution and thus will not match pdb files missing regions in the protein, including any purification tags [47]. It is important to make a homology model with the missing regions modeled in. MODELLER implemented in CHIMERA graphics program is useful to create a full-length model of the protein [80, 81].
2. FOXS outputs the fit of the atomic model to the experimental data. If the atomic model is the conformation in solution, the ideal χ^2 should be close to 1 (*see* **Note 9** for caveats to the use of χ^2). Two things should be considered. The noise of the data is a denominator to χ^2, which means that the noisier the data, the lower the χ^2. Indeed, χ^2 can drop below 1, as shown in this example (Fig. 4). Because of the large contribution of the noise in determining χ^2, the χ^2's of a model against different experimental data with varying noise should not be compared with each other.
3. Make sure that the fitting constant, c_2, is not close to 4. A c_2 close to 4 indicates overfitting of the model curve to the

Fig. 2 (continued) plateaus. In this case, plateau is occurring in Porod-Debye plot. (**c**) Volume analysis to obtain Porod exponent (P_x). Start and end points are set such that the line in the Porod-Debye plots represents a linear portion of the curve and the residuals are unbiased. In this example, the protein is well-folded. If the protein is flexible, the line would have a positive slope

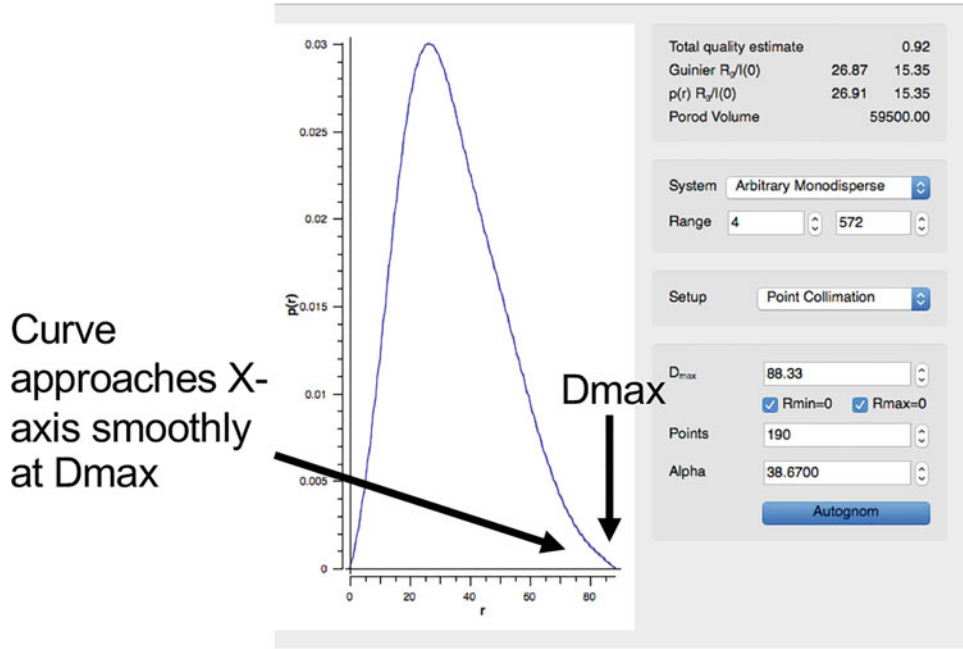
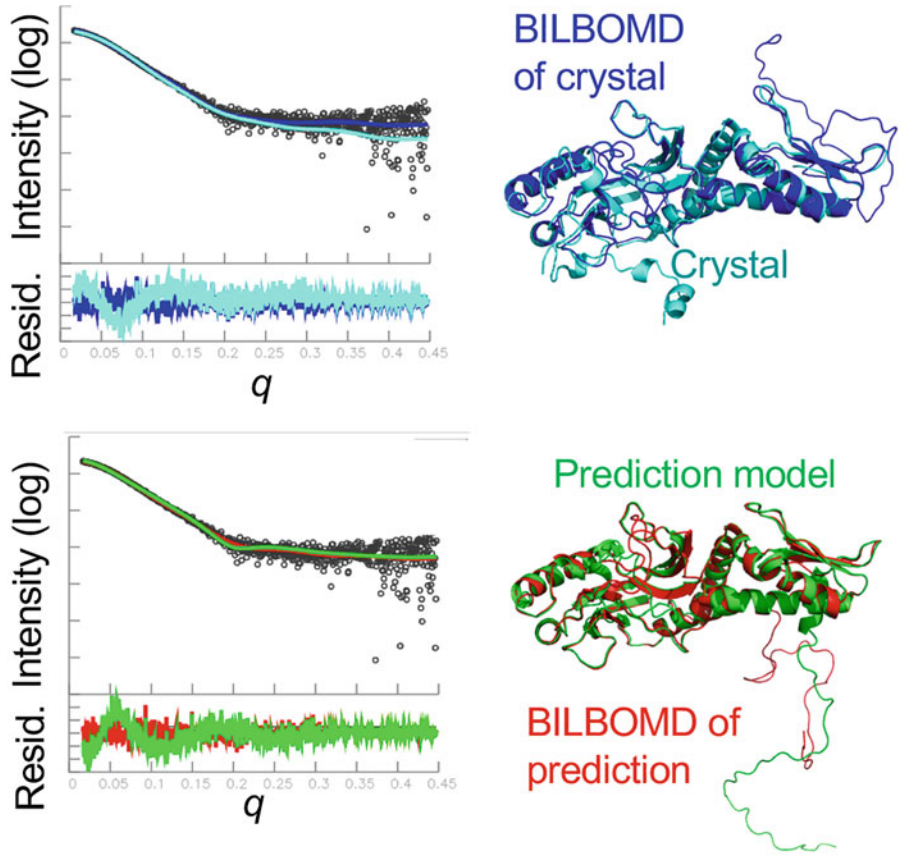


Fig. 3 Generation of *P*(*r*) using Gnom. Screenshot from ATSAS PRIMUS/GNOM. Set Dmax so that *P*(*r*) curve approaches the *X*-axis smoothly

experimental data. In the example, the crystal structure missing regions that could not be modeled has c_2 close to 4.

4. Examine the low scattering angle q region. A slope near $q = 0$ in the experimental data higher than the prediction indicates either that the conformation of the protein in solution is more expanded or, not infrequently, there is contribution of multimeric species in the experimental data.
5. Examine the features in the SAXS curves. If the experimental data has more features and is curvier, then the conformation in solution is more spherical. If it is more linear, then the conformation in solution is more expanded. DsDNA will have a linear SAXS curve.
6. After the FOXS server has run, multiple models as an ensemble can be assessed by "Multi-state models by MultiFOXS." If the protein sample has multiple stoichiometric or conformation states (indicated by the P_x), additional models will significantly improve the model. If there is little conformational heterogeneity contributing to the experimental SAXS data, then there will be no significant improvement with additional models or the solution state is not found in the models given to the program.
7. To double check the fit of the SAXS curve predicted atomic model(s) to the experimental data, it is informative to compare



PDB file	show all/hide all	χ^2	c_1	c_2	R_g	# atoms
predictionfullbilbo	☑	0.79	1.02	-0.70	26.35	2777
crystalfullbilbo	☑	0.82	1.03	-0.28	26.08	2777
predictionfullmodel	☑	2.12	1.03	0.22	28.16	2777
crystal	☑	2.20	0.99	3.85	24.17	2599
crystalfullmodeller	☐	2.50	1.00	1.41	24.35	2777

Fig. 4 FOXS and BILBOMD. Screenshots from FOXS webpage. FOXS is useful to generate and compare atomic models to the experimental SAXS data. BILBOMD is useful for generating diverse population of conformations that can be compared as single models or an ensemble of models to the experimental data. Here, the crystal structure missing flexible regions has a higher χ^2 and poor c_2 close to 4

the $P(r)$ curves. Taking the FOXS- or MultiFOXS-generated .fit file, the scattering angle q is in the first column; the intensity for the experimental data is in the second column, while the intensity for the atomic model is in the fourth column. To get the proper input of the model SAXS curve for gnom, use the linux command "awk '{print \$1,\$4}' fox.fit >model.dat" where foxs.fit is the FOXS-generated .fit file.

3.8 Generation of Flexible Atomic Models to Compare Ensembles with Experimental SAXS Data by BILBOMD

Proteins require motion to perform their functions and can adopt multiple conformation in solution, but crystal structures offer only static conformations trapped in the crystallographic lattice and may not represent conformations in solution. Indeed, studies of crystal structures reveal over half have different conformations in solution [82, 83]. BILBOMD is a useful program to generate models with different conformations, keeping domains as rigid bodies. *See* **Note 10** for additional BILBOMD information.

1. Go to BILBOMD server (https://bl1231.als.lbl.gov/bilbomd).
2. Follow the directions for uploading each segment or chain separately and the experimental SAXS data file. Keep file names short and with no special characters.
3. Define which regions are kept rigid, and input a R_g range, typically 10–20 Å below and above the experimentally defined R_g.
4. Analyze the output as described for FOXS.

3.9 Future Prospects

Promising to revolutionize the study of biology and medicine, an artificial intelligence algorithm in 2020 made an astounding breakthrough in solving the Fifty-Year Protein Folding Grand Challenge posed by Christian Anfinsen in his Nobel prize acceptance speech. The Deepmind AlphaFold algorithm accurately predicted structures for 90% of single chain protein structures, as assessed by Critical Assessment of protein Structure Prediction (CASP) 14 [84]. It is likely that when these more reliable protein structure predictions become available, DNA repair scientists will begin to use them. However, there are limitations even for AlphaFold with regard to certain proteins that are difficult or currently impossible to predict with high reliability and accuracy—in particular, multimeric, multi-domain, or flexible proteins. SAXS helps to fill this gap by providing experimental validation and information on conformations occurring in solution.

We tested this hypothesis using the AlphaFold model predicted for our protein in the last CASP14 in 2020, obtained from the CASP website (https://predictioncenter.org/index.cgi). Our protein has two domains, and the AlphaFold predicted the Cα backbone to 0.8 Å RMSD, including the relative orientation of the two domains (Fig. 4). Based on the AlphaFold model, we added in missing regions using MODELLER and created multiple models allowing certain regions to move. Using BILBOMD to allow flexible regions to move, we were able to match our experimental SAXS data with a χ^2 fit of 0.79. If a crystal structure was not available, this result suggests that SAXS could be used to validate an atomic model. Further work will be needed to improve prediction algorithms for multimeric protein/protein and protein/nucleic acid

complexes, but protein structure prediction, combined with SAXS-based experimental validation, has promise for generating atomic models and testable mechanistic hypotheses for the DNA repair field.

4 Notes

1. For complexes with low affinities and not stable over SEC, HT-SAXS may be required to preserve concentrations sufficient for complex formation. As exemplified by a successful study of a MutSα and PCNA complex [28], it is possible to obtain information from HT-SAXS as the signal of the complex will be greater than the individual components, depending on the details of the protein system and with the significant caveat that the SAXS signal comes from a stoichiometrically heterogeneous population.
2. Buffer Note. DNA/protein interactions are often based on electrostatics and specificity relies on salt concentrations. Many if not most DNA binding proteins will bind DNA non-specifically at low salt, and in such conditions, proteins will accumulate on the DNA ends or any available DNA footprint leading to non-informative stoichiometries. If salt is increased too high, all binding may be wiped out. Optimal salt concentration may be determined using HT-SAXS with SEC-SAXS later run with a running buffer that maximizes the specific interaction of interest. Importantly, the buffer must be an exact match to the protein sample. This is critical for HT-SAXS of eukaryotic DNA repair proteins that typically have disordered regions and can only be concentrated to relatively low concentrations.
3. Note for concentrated samples. We recommend to take aliquots out during concentration, instead of concentrating and then diluting to obtain protein samples at multiple concentrations. If aggregation occurs during concentration, it is most common that the proteins will remain irreversibly aggregated even when diluted.
4. HT SAXS Sample preparation by SEC. Buffer subtraction is still off after 1.5 CV, although conductivity may look fine. During equilibration, divert the buffer to the fraction collector to thoroughly clean out tubing. Dried particulates could cause artifacts in the SAXS sample.
5. Interparticle repulsion in DNA repair proteins is caused by molecules in solution affecting each other and is detectable in the low angle SAXS scattering pattern. PCNA, with its negatively charged outer surface, is particularly sensitive to

interparticle repulsion. Keeping protein concentration below 5 mg/ml and with 150 mM salt eliminates this problem [18].

6. SEC-SAXS data reduction programs. RAW is also useful for SEC-SAXS data reduction when X-ray radiation induces protein precipitation on the sample window, as detected as an increase in background from before and after the protein peak (https://bioxtas-raw.readthedocs.io/en/latest/index.html) [67].
7. The Porod exponent metric (P_x) arises from analysis of different plots, originating with the Kratky plot ($q^2 \times I(q)$ plotted against q^2) and expanding to the SIBYLS ($q^3 \times I(q)$ plotted against q^3) and Porod-Debye ($q^4 \times I(q)$ plotted against q^4) plots. Further explanation on the theory can be found on the BIOSIS.net website and reviews [19, 45, 85]. Correlating to flexibility, these plots reflect the protein density or volume occupied by the protein of a certain mass.
8. CRYSOL in the Svergun ATSAS program suite was one of the first programs for calculating SAXS data from explicit, atomic models of the protein or complex [74]. FOXS has an easy-to-use server with links to MultiFOXS for predicting SAXS data from ensembles of protein models [75–78]. This latter program is useful for looking at both conformationally diverse ensembles of the same protein [18, 62] or adding in a mixture of individual proteins and complexes [61]. The caveat to these two programs may seem minor but is important to remember. SAXS includes the hydration layer around the protein. Both programs do not use explicit modeling of waters but estimate the water layer. When experimental SAXS data is included in the prediction analysis, the hydration layer is adjusted to improve the fit to the experimental data through two fitting constants, c_1 and c_2, in FOXS. In our experience, a c_2 variable close to 4 suggests that there is likely overfitting to the experimental SAXS data and the atomic model does not represent what is occurring in solution and is not as good as suggested from the χ^2 value.
9. Comparison of experimental data to atomic models requires a quantitative metric on the similarity of their predicted atomic model to the SAXS data. Yet, the most common metrics that score the agreement of a model with experiment are often misunderstood in SAXS analysis. *First,* SAXS is an exponentially decaying function with 100- to 1000-fold difference in signal over the typical curve and strong bias toward low q regions. This is why we display the SAXS curve on log scale. Thus, any algorithm that fits data based on the residual difference between experiment and calculated scattering intensities ($I_{exp} - I_{calc}$) is dominated by information at very low q. For example,

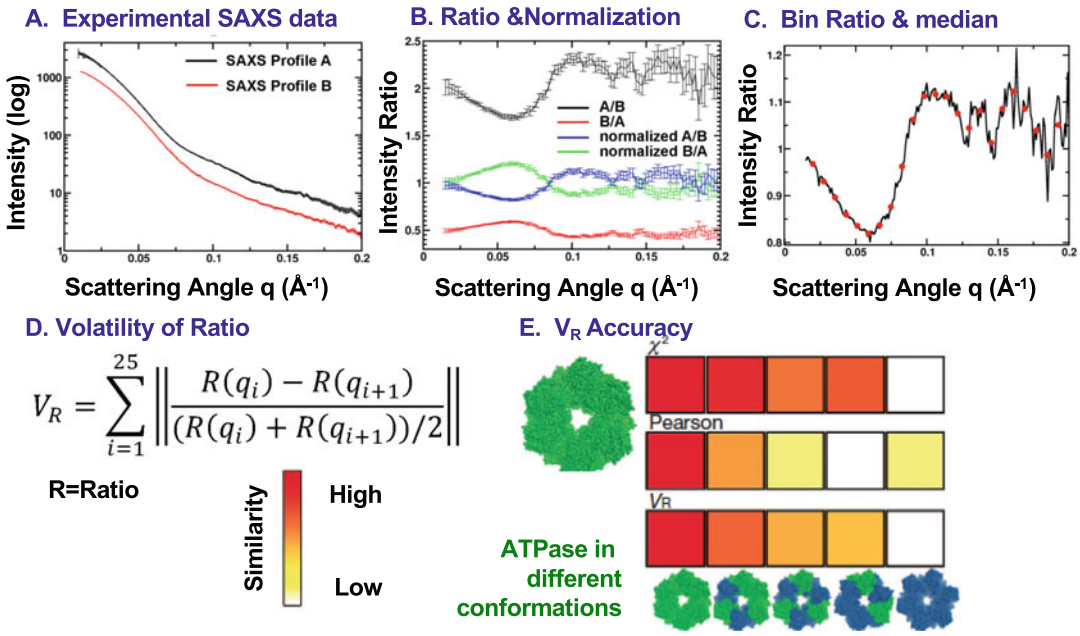


Fig. 5 Volatility ratio (V_R) for comparison of atomic models to experimental data or more generally, SAXS curve to another. (**a**) Two SAXS curves for comparison to be compared. Curves are not scaled. (**b**) The ratios of one SAXS curve over the other. (**c**) V_R divides the SAXS curve into bins and calculates each bin median. (**d**) A volatility measure, adopted from stock market analysis, assesses if a difference is real or noise. (**e**) χ^2, Pearson, and V_R are calculated for a hexameric ATPase in open, closed, or mixed populations, all of which have the same general shape. Only V_R accurately predicts the relative similarity

a 1% difference at low q (100 vs. 99) would be weighed the same as a 100-fold difference at higher q (1.01 vs. 0.01); thus, small differences in low q data outweigh all differences over the rest of the curve. This residual bias is inherent in the traditionally used χ^2 and simple RMSD. Second, χ^2 and RMSD require that curves be appropriately scaled which adds another parameter for fitting. Third, χ^2 assumes that all data points are independent, which is not true based on Shannon information theory [86]. Another metric is Volatility Ratio (V_R) [87] (Fig. 5). V_R has three advantages. (1) It uses a ratio of one SAXS curve over the other, so that the entire q region contributes equally to the comparison and is independent of scaling error (Fig. 5b). A subtractive method would depend on proper scaling of the experimental and predicted curve. (2) It exploits Shannon information theory such that a SAXS curve can be recapitulated by a smaller number of independent points and that the SAXS curve is vastly oversampled [86]. V_R divides the SAXS curve into bins and calculates each bin median, thereby reducing bias from experimental error (Fig. 5c). (3) It uses a volatility measure, adopted from stock market analysis (Fig. 5d). This equation assesses if a difference is real

or noise. The accuracy of V_R is evident in comparison of a hexameric ATPase in open, closed, or mixed populations, all of which have the same general shape (Fig. 5e) [87].

10. BILBOMD is a useful server to generate a diverse population of conformations by defining certain region(s) as rigid bodies. Promoting a broad sampling of conformations, undefined regions are then "boiled" in silico and allowed to move but keeping bond distances constant. Bond angles however are non-biological, and conclusions should keep this caveat in mind. BILBOMD is best for beads on a string and tends to move domains away from each other. It is not as useful for generating conformations where multiple domains pack against each other. FOXSDOCK is useful if there are only two domains [76], but more advanced computational docking programs are needed for docking mutiple domains.

Acknowledgments

We thank long-time collaborators and colleagues in the SAXS field, including Dmitry Svergun, Michal Hammel, Robert P. Rambo, Andrej Sali, Dina Schneidman, Jesse Hopkins, and Daniel Rosenberg, for their many insights and contributions including useful programs for SAXS data collection and analysis. Work is supported by NCI P01 CA092584 (to S.E.T., G.H., J.A.T.), NIGMS R01GM110387 (to S.E.T.), R35 CA220430 (to J.A.T.), and 1R01GM137021 (To S.E.T. and G.H.). JAT effort is also supported by Cancer Prevention Research Institute of Texas (CPRIT) grant RP180813 and a Robert A Welch Chemistry Chair. The SIBYLS beamline's efforts are supported by DOE-BER IDAT under contract DE-AC02-05CH11231.

References

1. Tsutakawa SE, Tsai CL, Yan C, Bralic A, Chazin WJ, Hamdan SM, Scharer OD, Ivanov I, Tainer JA (2020) Envisioning how the prototypic molecular machine TFIIH functions in transcription initiation and DNA repair. DNA Repair (Amst) 96:102972. https://doi.org/10.1016/j.dnarep.2020.102972
2. Wilson SH, Kunkel TA (2000) Passing the baton in base excision repair. Nat Struct Biol 7(3):176–178. https://doi.org/10.1038/73260
3. Parikh SS, Mol CD, Hosfield DJ, Tainer JA (1999) Envisioning the molecular choreography of DNA base excision repair. Curr Opin Struct Biol 9(1):37–47. https://doi.org/10.1016/s0959-440x(99)80006-2
4. Mol CD, Izumi T, Mitra S, Tainer JA (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. Nature 403(6768):451–456. https://doi.org/10.1038/35000249
5. Hitomi K, Iwai S, Tainer JA (2007) The intricate structural chemistry of base excision repair machinery: implications for DNA damage recognition, removal, and repair. DNA Repair 6(4):410–428. https://doi.org/10.1016/j.dnarep.2006.10.004
6. Krokan HE, Bjoras M (2013) Base excision repair. Cold Spring Harb Perspect Biol 5(4):a012583. https://doi.org/10.1101/cshperspect.a012583

7. Tsutakawa SE, Sarker AH, Ng C, Arvai AS, Shin DS, Shih B, Jiang S, Thwin AC, Tsai MS, Willcox A, Her MZ, Trego KS, Raetz AG, Rosenberg D, Bacolla A, Hammel M, Griffith JD, Cooper PK, Tainer JA (2020) Human XPG nuclease structure, assembly, and activities with insights for neurodegeneration and cancer from pathogenic mutations. Proc Natl Acad Sci U S A 117(25): 14127–14138. https://doi.org/10.1073/pnas.1921311117
8. Shibata A, Moiani D, Arvai AS, Perry J, Harding SM, Genois MM, Maity R, van Rossum-Fikkert S, Kertokalio A, Romoli F, Ismail A, Ismalaj E, Petricci E, Neale MJ, Bristow RG, Masson JY, Wyman C, Jeggo PA, Tainer JA (2014) DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. Mol Cell 53(1):7–18. https://doi.org/10.1016/j.molcel.2013.11.003
9. Perry JJ, Yannone SM, Holden LG, Hitomi C, Asaithamby A, Han S, Cooper PK, Chen DJ, Tainer JA (2006) WRN exonuclease structure and molecular mechanism imply an editing role in DNA end processing. Nat Struct Mol Biol 13(5):414–422. https://doi.org/10.1038/nsmb1088
10. Tsutakawa SE, Lafrance-Vanasse J, Tainer JA (2014) The cutting edges in DNA repair, licensing, and fidelity: DNA and RNA repair nucleases sculpt DNA to measure twice, cut once. DNA Repair (Amst) 19:95–107. https://doi.org/10.1016/j.dnarep.2014.03.022
11. Holton JM, Classen S, Frankel KA, Tainer JA (2014) The R-factor gap in macromolecular crystallography: an untapped potential for insights on accurate structures. FEBS J 281(18):4046–4060. https://doi.org/10.1111/febs.12922
12. Fowler NJ, Sljoka A, Williamson MP (2020) A method for validating the accuracy of NMR protein structures. Nat Commun 11(1):6321. https://doi.org/10.1038/s41467-020-20177-1
13. Hegde PM, Dutta A, Sengupta S, Mitra J, Adhikari S, Tomkinson AE, Li GM, Boldogh I, Hazra TK, Mitra S, Hegde ML (2015) THE C-terminal domain (CTD) of human DNA glycosylase NEIL1 is required for forming BERosome repair complex with DNA replication proteins at THE replicating genome dominant negative function of the CTD. J Biol Chem 290(34):20919–20933. https://doi.org/10.1074/jbc.M115.642918
14. Hegde ML, Tsutakawa SE, Hegde PM, Holthauzen LM, Li J, Oezguen N, Hilser VJ, Tainer JA, Mitra S (2013) The disordered C-terminal domain of human DNA glycosylase NEIL1 contributes to its stability via intramolecular interactions. J Mol Biol 425(13): 2359–2371. https://doi.org/10.1016/j.jmb.2013.03.030
15. Bacolla A, Sengupta S, Ye Z, Yang C, Mitra J, De-Paula RB, Hegde ML, Ahmed Z, Mort M, Cooper DN, Mitra S, Tainer JA (2021) Heritable pattern of oxidized DNA base repair coincides with pre-targeting of repair complexes to open chromatin. Nucleic Acids Res 49(1): 221–243. https://doi.org/10.1093/nar/gkaa1120
16. Williams GJ, Williams RS, Williams JS, Moncalian G, Arvai AS, Limbo O, Guenther G, SilDas S, Hammel M, Russell P, Tainer JA (2011) ABC ATPase signature helices in Rad50 link nucleotide state to Mre11 interface for DNA repair. Nat Struct Mol Biol 18(4):423–431. https://doi.org/10.1038/nsmb.2038
17. Deshpande RA, Williams GJ, Limbo O, Williams RS, Kuhnlein J, Lee JH, Classen S, Guenther G, Russell P, Tainer JA, Paull TT (2014) ATP-driven Rad50 conformations regulate DNA tethering, end resection, and ATM checkpoint signaling. EMBO J 33(5): 482–500. https://doi.org/10.1002/embj.201386100
18. Tsutakawa SE, Van Wynsberghe AW, Freudenthal BD, Weinacht CP, Gakhar L, Washington MT, Zhuang Z, Tainer JA, Ivanov I (2011) Solution X-ray scattering combined with computational modeling reveals multiple conformations of covalently bound ubiquitin on PCNA. Proc Natl Acad Sci U S A 108(43):17672–17677. https://doi.org/10.1073/pnas.1110480108
19. Rambo RP, Tainer JA (2011) Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. Biopolymers 95(8): 559–571. https://doi.org/10.1002/bip.21638
20. Hammel M, Rosenberg DJ, Bierma J, Hura GL, Thapar R, Lees-Miller SP, Tainer JA (2020) Visualizing functional dynamicity in the DNA-dependent protein kinase holoenzyme DNA-PK complex by integrating SAXS with cryo-EM. Prog Biophys Mol Biol 163:74–86. https://doi.org/10.1016/j.pbiomolbio.2020.09.003
21. Tsutakawa SE, Shin DS, Mol CD, Izumi T, Arvai AS, Mantha AK, Szczesny B, Ivanov IN, Hosfield DJ, Maiti B, Pique ME, Frankel KA, Hitomi K, Cunningham RP, Mitra S, Tainer JA (2013) Conserved structural chemistry for

incision activity in structurally non-homologous apurinic/apyrimidinic endonuclease APE1 and endonuclease IV DNA repair enzymes. J Biol Chem 288(12): 8445–8455. https://doi.org/10.1074/jbc.M112.422774

22. Parikh SS, Mol CD, Slupphaug G, Bharati S, Krokan HE, Tainer JA (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. EMBO J 17(17): 5214–5226. https://doi.org/10.1093/emboj/17.17.5214
23. Mullins EA, Rodriguez AA, Bradley NP, Eichman BF (2019) Emerging roles of DNA glycosylases and the base excision repair pathway. Trends Biochem Sci 44(9):765–781. https://doi.org/10.1016/j.tibs.2019.04.006
24. Tsutakawa SE, Sarker AH (2021) Breaking the rules: protein sculpting in NEIL2 regulation. Structure 29(1):1–2. https://doi.org/10.1016/j.str.2020.12.011
25. Tubbs JL, Latypov V, Kanugula S, Butt A, Melikishvili M, Kraehenbuehl R, Fleck O, Marriott A, Watson AJ, Verbeek B, McGown G, Thorncroft M, Santibanez-Koref MF, Millington C, Arvai AS, Kroeger MD, Peterson LA, Williams DM, Fried MG, Margison GP, Pegg AE, Tainer JA (2009) Flipping of alkylated DNA damage bridges base and nucleotide excision repair. Nature 459(7248): 808–813. https://doi.org/10.1038/nature08076
26. Latypov VF, Tubbs JL, Watson AJ, Marriott AS, McGown G, Thorncroft M, Wilkinson OJ, Senthong P, Butt A, Arvai AS, Millington CL, Povey AC, Williams DM, Santibanez-Koref MF, Tainer JA, Margison GP (2012) Atl1 regulates choice between global genome and transcription-coupled repair of O(6)-alkylguanines. Mol Cell 47(1):50–60. https://doi.org/10.1016/j.molcel.2012.04.028
27. Grasby JA, Finger LD, Tsutakawa SE, Atack JM, Tainer JA (2012) Unpairing and gating: sequence-independent substrate recognition by FEN superfamily nucleases. Trends Biochem Sci 37(2):74–84. https://doi.org/10.1016/j.tibs.2011.10.003
28. Rashid F, Harris PD, Zaher MS, Sobhy MA, Joudeh LI, Yan C, Piwonski H, Tsutakawa SE, Ivanov I, Tainer JA, Habuchi S, Hamdan SM (2017) Single-molecule FRET unveils induced-fit mechanism for substrate selectivity in flap endonuclease 1. Elife 6:e21884. https://doi.org/10.7554/eLife.21884
29. Tsutakawa SE, Classen S, Chapados BR, Arvai AS, Finger LD, Guenther G, Tomlinson CG, Thompson P, Sarker AH, Shen B, Cooper PK, Grasby JA, Tainer JA (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. Cell 145(2):198–211. https://doi.org/10.1016/j.cell.2011.03.004
30. Syed A, Tainer JA (2018) The MRE11-RAD50-NBS1 complex conducts the orchestration of damage signaling and outcomes to stress in DNA replication and repair. Annu Rev Biochem 87:263–294. https://doi.org/10.1146/annurev-biochem-062917-012415
31. Kashammer L, Saathoff JH, Lammens K, Gut F, Bartho J, Alt A, Kessler B, Hopfner KP (2019) Mechanism of DNA end sensing and processing by the Mre11-Rad50 complex. Mol Cell 76(3):382–394. e386. https://doi.org/10.1016/j.molcel.2019.07.035
32. Wang JL, Duboc C, Wu Q, Ochi T, Liang S, Tsutakawa SE, Lees-Miller SP, Nadal M, Tainer JA, Blundell TL, Strick TR (2018) Dissection of DNA double-strand-break repair using novel single-molecule forceps. Nat Struct Mol Biol 25(6):482–487. https://doi.org/10.1038/s41594-018-0065-1
33. Carney SM, Moreno AT, Piatt SC, Cisneros-Aguirre M, Lopezcolorado FW, Stark JM, Loparo JJ (2020) XLF acts as a flexible connector during non-homologous end joining. Elife 9:e61920. https://doi.org/10.7554/eLife.61920
34. Conlin MP, Reid DA, Small GW, Chang HH, Watanabe G, Lieber MR, Ramsden DA, Rothenberg E (2017) DNA ligase IV guides end-processing choice during nonhomologous end joining. Cell Rep 20(12):2810–2819. https://doi.org/10.1016/j.celrep.2017.08.091
35. Hammel M, Rey M, Yu Y, Mani RS, Classen S, Liu M, Pique ME, Fang S, Mahaney BL, Weinfeld M, Schriemer DC, Lees-Miller SP, Tainer JA (2011) XRCC4 protein interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double strand break repair. J Biol Chem 286(37):32638–32650. https://doi.org/10.1074/jbc.M111.272641
36. Mahaney BL, Hammel M, Meek K, Tainer JA, Lees-Miller SP (2013) XRCC4 and XLF form long helical protein filaments suitable for DNA end protection and alignment to facilitate DNA double strand break repair. Biochem Cell Biol 91(1):31–41. https://doi.org/10.1139/bcb-2012-0058
37. Williams GJ, Hammel M, Radhakrishnan SK, Ramsden D, Lees-Miller SP, Tainer JA (2014) Structural insights into NHEJ: building up an integrated picture of the dynamic DSB repair super complex, one component and interaction

at a time. DNA Repair (Amst) 17:110–120. https://doi.org/10.1016/j.dnarep.2014.02.009

38. Zhao B, Watanabe G, Morten MJ, Reid DA, Rothenberg E, Lieber MR (2019) The essential elements for the noncovalent association of two DNA ends during NHEJ synapsis. Nat Commun 10(1):3588. https://doi.org/10.1038/s41467-019-11507-z
39. Eckelmann BJ, Bacolla A, Wang H, Ye Z, Guerrero EN, Jiang W, El-Zein R, Hegde ML, Tomkinson AE, Tainer JA, Mitra S (2020) XRCC1 promotes replication restart, nascent fork degradation and mutagenic DNA repair in BRCA2-deficient cells. NAR. Cancer 2(3): zcaa013. https://doi.org/10.1093/narcan/zcaa013
40. Hammel M, Rashid I, Sverzhinsky A, Pourfarjam Y, Tsai MS, Ellenberger T, Pascal JM, Kim IK, Tainer JA, Tomkinson AE (2021) An atypical BRCT-BRCT interaction with the XRCC1 scaffold protein compacts human DNA ligase IIIalpha within a flexible DNA repair complex. Nucleic Acids Res 49(1): 306–321. https://doi.org/10.1093/nar/gkaa1188
41. Staresincic L, Fagbemi AF, Enzlin JH, Gourdin AM, Wijgers N, Dunand-Sauthier I, Giglia-Mari G, Clarkson SG, Vermeulen W, Scharer OD (2009) Coordination of dual incision and repair synthesis in human nucleotide excision repair. EMBO J 28(8):1111–1120. https://doi.org/10.1038/emboj.2009.49
42. Rambo RP, Tainer JA (2010) Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. Curr Opin Struct Biol 20(1):128–137. https://doi.org/10.1016/j.sbi.2009.12.015
43. Rambo RP, Tainer JA (2013) Super-resolution in solution X-ray scattering and its applications to structural systems biology. Annu Rev Biophys 42:415–441. https://doi.org/10.1146/annurev-biophys-083012-130301
44. Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q Rev Biophys 40(3):191–285. https://doi.org/10.1017/S0033583507004635
45. Brosey CA, Tainer JA (2019) Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology. Curr Opin Struct Biol 58:197–213. https://doi.org/10.1016/j.sbi.2019.04.004
46. Brosey CA, Ho C, Long WZ, Singh S, Burnett K, Hura GL, Nix JC, Bowman GR, Ellenberger T, Tainer JA (2016) Defining NADH-driven Allostery regulating apoptosis-inducing factor. Structure 24(12):2067–2079. https://doi.org/10.1016/j.str.2016.09.012
47. Hura GL, Menon AL, Hammel M, Rambo RP, Poole FL 2nd, Tsutakawa SE, Jenney FE Jr, Classen S, Frankel KA, Hopkins RC, Yang SJ, Scott JW, Dillard BD, Adams MW, Tainer JA (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat Methods 6(8):606–612. https://doi.org/10.1038/nmeth.1353
48. Classen S, Rodic I, Holton J, Hura GL, Hammel M, Tainer JA (2010) Software for the high-throughput collection of SAXS data using an enhanced Blu-ice/DCS control system. J Synchrotron Radiat 17(6):774–781. https://doi.org/10.1107/S0909049510028566
49. Classen S, Hura GL, Holton JM, Rambo RP, Rodic I, McGuire PJ, Dyer K, Hammel M, Meigs G, Frankel KA, Tainer JA (2013) Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the advanced light source. J Appl Crystallogr 46(Pt 1):1–13. https://doi.org/10.1107/S0021889812048698
50. Dyer KN, Hammel M, Rambo RP, Tsutakawa SE, Rodic I, Classen S, Tainer JA, Hura GL (2014) High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. Methods Mol Biol 1091:245–258. https://doi.org/10.1007/978-1-62703-691-7_18
51. Hura GL, Tsai CL, Claridge SA, Mendillo ML, Smith JM, Williams GJ, Mastroianni AJ, Alivisatos AP, Putnam CD, Kolodner RD, Tainer JA (2013) DNA conformations in mismatch repair probed in solution by X-ray scattering from gold nanocrystals. Proc Natl Acad Sci U S A 110(43):17308–17313. https://doi.org/10.1073/pnas.1308595110
52. Rosenberg DJ, Syed A, Tainer JA, Hura GL Monitoring nuclease activity by X-ray scattering interferometry using gold nanoparticle conjugated DNA. Methods Mol Biol. (In this issue)
53. Mertens HD, Svergun DI (2010) Structural characterization of proteins and complexes using small-angle X-ray solution scattering. J Struct Biol 172(1):128–141. https://doi.org/10.1016/j.jsb.2010.06.012
54. Tang HYH, Tainer JA, Hura GL (2017) High resolution distance distributions determined by X-ray and Neutron scattering. Adv Exp Med

Biol 1009:167–181. https://doi.org/10.1007/978-981-10-6038-0_10

55. Tsutakawa SE, Hura GL, Frankel KA, Cooper PK, Tainer JA (2007) Structural analysis of flexible proteins in solution by small angle X-ray scattering combined with crystallography. J Struct Biol 158(2):214–223. https://doi.org/10.1016/j.jsb.2006.09.008
56. Hammel M, Yu Y, Fang S, Lees-Miller SP, Tainer JA (2010) XLF regulates filament architecture of the XRCC4.Ligase IV complex. Structure 18(11):1431–1442. https://doi.org/10.1016/j.str.2010.09.009
57. Hammel M, Yu Y, Mahaney BL, Cai B, Ye R, Phipps BM, Rambo RP, Hura GL, Pelikan M, So S, Abolfath RM, Chen DJ, Lees-Miller SP, Tainer JA (2010) Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. J Biol Chem 285(2):1414–1423. https://doi.org/10.1074/jbc.M109.065615
58. Hammel M, Yu Y, Radhakrishnan SK, Chokshi C, Tsai MS, Matsumoto Y, Kuzdovich M, Remesh SG, Fang S, Tomkinson AE, Lees-Miller SP, Tainer JA (2016) An intrinsically disordered APLF links Ku, DNA-PKcs, and XRCC4-DNA ligase IV in an extended flexible non-homologous end joining complex. J Biol Chem 291(53):26987–27006. https://doi.org/10.1074/jbc.M116.751867
59. Guo HF, Tsai CL, Terajima M, Tan X, Banerjee P, Miller MD, Liu X, Yu J, Byemerwa J, Alvarado S, Kaoud TS, Dalby KN, Bota-Rabassedas N, Chen Y, Yamauchi M, Tainer JA, Phillips GN Jr, Kurie JM (2018) Pro-metastatic collagen lysyl hydroxylase dimer assemblies stabilized by Fe (2+)-binding. Nat Commun 9(1):512. https://doi.org/10.1038/s41467-018-02859-z
60. Walker RG, Kattamuri C, Goebel EJ, Zhang F, Hammel M, Tainer JA, Linhardt RJ, Thompson TB (2021) Heparin-mediated dimerization of follistatin. Exp Biol Med (Maywood) 246(4):467–482. https://doi.org/10.1177/1535370220966296
61. Cotner-Gohara E, Kim IK, Hammel M, Tainer JA, Tomkinson AE, Ellenberger T (2010) Human DNA ligase III recognizes DNA ends by dynamic switching between two DNA-bound states. Biochemistry 49(29):6165–6176. https://doi.org/10.1021/bi100503w
62. Tsutakawa SE, Yan C, Xu X, Weinacht CP, Freudenthal BD, Yang K, Zhuang Z, Washington MT, Tainer JA, Ivanov I (2015) Structurally distinct ubiquitin- and sumo-modified PCNA: implications for their distinct roles in the DNA damage response. Structure 23(4):724–733. https://doi.org/10.1016/j.str.2015.02.008
63. Roberts VA, Freeman HC, Olson AJ, Tainer JA, Getzoff ED (1991) Electrostatic orientation of the electron-transfer complex between plastocyanin and cytochrome c. J Biol Chem 266(20):13431–13441
64. Rambo RP, Tainer JA (2010) Improving small-angle X-ray scattering data for structural analyses of the RNA world. RNA 16(3):638–646. https://doi.org/10.1261/rna.1946310
65. Thapar R, Wang JL, Hammel M, Ye R, Liang K, Sun C, Hnizda A, Liang S, Maw SS, Lee L, Villarreal H, Forrester I, Fang S, Tsai MS, Blundell TL, Davis AJ, Lin C, Lees-Miller SP, Strick TR, Tainer JA (2021) Mechanism of efficient double-strand break repair by a long non-coding RNA. Nucleic Acids Res 49(2):1199–1200. https://doi.org/10.1093/nar/gkaa1233
66. Zhou Y, Millott R, Kim HJ, Peng S, Edwards RA, Skene-Arnold T, Hammel M, Lees-Miller SP, Tainer JA, Holmes CFB, Glover JNM (2019) Flexible tethering of ASPP proteins facilitates PP-1c catalysis. Structure 27:1485–1496.e4. https://doi.org/10.1016/j.str.2019.07.012
67. Hopkins JB, Gillilan RE, Skou S (2017) BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. J Appl Crystallogr 50(Pt 5):1545–1553. https://doi.org/10.1107/S1600576717011438
68. Tully MD, Tarbouriech N, Rambo RP, Hutin S (2021) Analysis of SEC-SAXS data via EFA deconvolution and scatter. J Vis Exp 167:10.3791/61578
69. Hammel M, Amlanjyoti D, Reyes FE, Chen JH, Parpana R, Tang HY, Larabell CA, Tainer JA, Adhya S (2016) HU multimerization shift controls nucleoid compaction. Sci Adv 2(7):e1600650. https://doi.org/10.1126/sciadv.1600650
70. Manalastas-Cantos K, Konarev PV, Hajizadeh NR, Kikhney AG, Petoukhov MV, Molodenskiy DS, Panjkovich A, Mertens HDT, Gruzinov A, Borges C, Jeffries CM, Svergun DI, Franke D (2021) ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. J Appl Crystallogr 54(1):343–355. https://doi.org/10.1107/S1600576720013412
71. Rambo RP, Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. Nature 496(7446):477–481. https://doi.org/10.1038/nature12070

72. Durand D, Vives C, Cannella D, Perez J, Pebay-Peyroula E, Vachette P, Fieschi F (2010) NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers. J Struct Biol 169(1): 45–53. https://doi.org/10.1016/j.jsb.2009.08.009
73. Svergun DI (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J Appl Crystallogr 25:495–503. https://doi.org/10.1107/S0021889892001663
74. Svergun D, Barberato C, Koch MHJ (1995) CRYSOL– a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. J Appl Crystallogr 28(6):768–773. https://doi.org/10.1107/S0021889895007047
75. Schneidman-Duhovny D, Kim SJ, Sali A (2012) Integrative structural modeling with small angle X-ray scattering profiles. BMC Struct Biol 12:17. https://doi.org/10.1186/1472-6807-12-17
76. Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A (2016) FoXS, FoXSDock and MultiFoXS: single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. Nucleic Acids Res 44 (W1):W424–W429. https://doi.org/10.1093/nar/gkw389
77. Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A (2013) Accurate SAXS profile computation and its assessment by contrast variation experiments. Biophys J 105(4):962–974. https://doi.org/10.1016/j.bpj.2013.07.020
78. Schneidman-Duhovny D, Hammel M, Sali A (2010) FoXS: a web server for rapid computation and fitting of SAXS profiles. Nucleic Acids Res 38(Web Server issue):W540–W544. https://doi.org/10.1093/nar/gkq461
79. Forster F, Webb B, Krukenberg KA, Tsuruta H, Agard DA, Sali A (2008) Integration of small-angle X-ray scattering data into structural modeling of proteins and their assemblies. J Mol Biol 382(4):1089–1106. https://doi.org/10.1016/j.jmb.2008.07.074
80. Yang Z, Lasker K, Schneidman-Duhovny D, Webb B, Huang CC, Pettersen EF, Goddard TD, Meng EC, Sali A, Ferrin TE (2012) UCSF Chimera, MODELLER, and IMP: an integrated modeling system. J Struct Biol 179(3):269–278. https://doi.org/10.1016/j.jsb.2011.09.006
81. Webb B, Sali A (2016) Comparative protein structure modeling using modeller. Curr Protoc Bioinformatics 54:5 6 1–5 6 37. https://doi.org/10.1002/cpbi.3
82. Hura GL, Hodge CD, Rosenberg D, Guzenko D, Duarte JM, Monastyrskyy B, Grudinin S, Kryshtafovych A, Tainer JA, Fidelis K, Tsutakawa SE (2019) Small angle X-ray scattering-assisted protein structure prediction in CASP13 and emergence of solution structure differences. Proteins 87(12): 1298–1314. https://doi.org/10.1002/prot.25827
83. Andrec M, Snyder DA, Zhou Z, Young J, Montelione GT, Levy RM (2007) A large data set comparison of protein structures determined by crystallography and NMR: statistical test for structural differences and the effect of crystal packing. Proteins 69(3):449–465. https://doi.org/10.1002/prot.21507
84. Callaway E (2020) It will change everything': DeepMind's AI makes gigantic leap in solving protein structures. Nature 588(7837): 203–204. https://doi.org/10.1038/d41586-020-03348-4
85. Reyes FE, Schwartz CR, Tainer JA, Rambo RP (2014) Methods for using new conceptual tools and parameters to assess RNA structure by small-angle X-ray scattering. Methods Enzymol 549:235–263. https://doi.org/10.1016/B978-0-12-801122-5.00011-8
86. Moore P (1980) Small-angle scattering. Information content and error analysis. J Appl Crystallogr 13(2):168–175. https://doi.org/10.1107/S002188988001179X
87. Hura GL, Budworth H, Dyer KN, Rambo RP, Hammel M, McMurray CT, Tainer JA (2013) Comprehensive macromolecular conformations mapped by quantitative SAXS analyses. Nat Methods 10(6):453–454. https://doi.org/10.1038/nmeth.2453

Chapter 5

Assessing DNA Damage Responses Using B Lymphocyte Cultures

Rachel Johnston, Lynn S. White, and Jeffrey J. Bednarski

Abstract

Development of B cells requires the programmed generation and repair of double-stranded DNA breaks in antigen receptor genes. Investigation of the cellular responses to these DNA breaks has established important insights into B cell development and, more broadly, has provided fundamental advances into the molecular mechanisms of DNA damage response pathways. Abelson transformed pre-B cell lines and primary pre-B cell cultures are malleable experimental systems with diverse applications for studying DNA damage responses. This chapter describes methods for generating these cellular systems, inducing and quantifying DSBs, and assessing DNA damage programs.

Key words B cells, DNA double-stranded breaks, V(D)J recombination, DNA damage response

1 Introduction

Chromosomal double-stranded DNA breaks (DSBs) are generated during normal physiologic processes, such as antigen receptor gene recombination in immune cells, DNA replication, and meiosis, or by exposure of cells to exogenous genotoxins, such as radiation or chemotherapy. Regardless of mechanism of injury, DSBs activate a highly conserved signaling network that coordinates diverse cellular responses, including DSB repair, cell cycle checkpoint, and cell death [1–3].

Developing B cells (and T cells) must recombine antigen receptor genes in order to create a diverse receptor repertoire to recognize pathogens. The assembly of antigen receptor genes occurs through a process termed V(D)J recombination, which depends on the successful joining of two distant DNA segments through generation and repair of DSBs [4–7]. The DSBs necessary for this process are made by the RAG endonuclease, which is comprised of the RAG1 and RAG2 proteins [4–7]. RAG DSBs are repaired by

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_5,

nonhomologous end joining to ensure proper assembly of the genes and ongoing lymphocyte development [4–7].

Development of two experimental systems established methods for studying the mechanisms and cellular factors that coordinate the generation, repair, and cellular response to RAG DSBs. In one method, immortalized pre-B cell lines are generated by retroviral expression of the viral Abelson kinase (v-abl) in bone marrow pre-B cells [8–10]. In a second approach, primary (non-transformed) pre-B cells are expanded in vitro by culture with the cytokine interleukin-7 (IL-7) [11–16]. In both models, the proliferating pre-B cells do not express RAG or generate DSBs. RAG DSBs are induced by treatment of abl pre-B cells with imatinib, which inhibits the v-abl kinase, or by removal of IL-7 from primary pre-B cell cultures [8–16]. Both approaches result in cell cycle arrest (in G1 phase), which leads to expression of RAG and subsequent generation of RAG DSBs [17]. Transformed abl pre-B cells and primary pre-B cell cultures can be established from genetically modified mice and are amenable to RNAi-mediated inhibition of gene expression as well as CRISPR-mediated gene deletion [9, 10, 12, 13, 18–20]. Consequently, these pre-B cell systems are useful tools for evaluating the genetic factors that regulate DNA damage responses, including DSB generation and repair [5, 21–30]. Effects of DNA damage responses on gene transcription and developmental programs can also be investigated in these systems [12, 13, 16, 18–21]. Studies using these pre-B cell models have revealed insights into molecular mechanisms of DNA damage responses that are broadly applicable to other cell types [19, 20].

This chapter describes protocols for establishing abl pre-B cells and primary pre-B cell cultures as well as approaches for using these experimental systems to evaluate DNA damage responses.

2 Materials

2.1 Cell Lines and Plasmids

1. Platinum-E (Plat-E) Retroviral Packaging Cell Line (Cell Biolabs, Inc).
2. pMSCV-v-abl retroviral plasmid [8–10].

2.2 Culture Media and Solutions

1. Tissue culture media: DMEM supplemented with 10% FBS ("DMEM media plus 10% FBS"), 100 mg/mL penicillin, 100 mg/mL streptomycin, 2 mM L-glutamine, 1 mM sodium pyruvate, 1 mM non-essential amino acids, and 0.05 mM 2-mercaptoethanol.
2. Tissue culture media: DMEM supplemented with 20% FBS ("DMEM media plus 20% FBS"), 100 mg/mL penicillin, 100 mg/mL streptomycin, 2 mM L-glutamine, 1 mM sodium

pyruvate, 1 mM non-essential amino acids, and 0.05 mM 2-mercaptoethanol.

3. Murine recombinant interleukin-7 (IL-7; Miltenyi Biotec).
4. Imatinib (STI-571; Millipore Sigma).
5. Rapid lysis buffer with Proteinase K: 100 mM Tris pH 8.5, 5 mM ethylenediaminetetraacetic acid (EDTA), 0.2% sodium dodecyl sulfate (SDS), 200 mM sodium chloride, and 67 μg/mL Proteinase K.
6. Lipofectamine 2000 (Invitrogen).
7. Polybrene Infection/Transfection Reagent (Sequabrene; Sigma).
8. Cell-Tak Cell and Tissue Adhesive (Corning).
9. NEBNext dsDNA Fragmentase (New England Biolabs).
10. Brilliant II SYBR Green QPCR Master Mix (Agilent Technologies).
11. PCR primers (examples):
    (a) *CD19* Forward – TGTTCTCCTTCCTCCTCTTTCT.
    (b) *CD19* Reverse – CTCAACTCAGAACCCAGACTTT.
    (c) *Igk* J1 Forward – CTGTTCCTCTTCAGTGAG GAGGGT.
    (d) *Igk* J1 Reverse – GACATAGAAGCCACAGACATAGA CAACGG.

3 Methods

3.1 Generating Abelson Pre-B Cell Lines

3.1.1 Generation of V-abl-Expressing Retrovirus

1. (Day 1) Culture 3–5 × 10^6 PlatE cells in 10 mL of tissue culture media +10% FBS in a T75 tissue culture flask at 37 ° C with 5% CO_2 (*see* **Note 1**).
2. (Day 2) When cells reach ~60–80% confluence, aspirate media. Add 10 mL of DMEM media plus 10% FBS without antibiotics (no penicillin or streptomycin). In both steps, use caution to not disrupt the adherent cell monolayer on the bottom of the flask. Place in tissue culture incubator at 37 ° C with 5% CO_2.
3. (Day 2) Prepare DNA/Lipofectamine mixture. Dilute 60 mL Lipofectamine 2000 in 1.5 mL DMEM media (no additives). Gently mix and incubate at room temperature for 5 min (*see* **Note 2**).
4. In a separate tube, dilute 24 μg pMSCV-v-abl retroviral plasmid in 1.5 mL DMEM media (no additives). Mix and then add to diluted Lipofectamine 2000 (from **step 3** above). Gently mix and incubate for 20 min at room temperature.

5. (Day 2) Add 3 mL of DNA/Lipofectamine mix (from **step 4**) dropwise to media on cultured PlatE cells (from **step 2**). Place in tissue culture incubator at 37 °C with 5% CO_2.
6. (Day 3) 24 h later, carefully aspirate media, and add 10 mL of fresh tissue culture media with 10% FBS (and all antibiotics). Use caution to not disrupt the adherent cell monolayer on the bottom of the flask. Return to tissue culture incubator at 37 °C with 5% CO_2.
7. (Day 5) 48 h later, remove supernatant and save. Add 10 mL of fresh DMEM media plus 10% FBS (and all antibiotics). Use caution to not disrupt the adherent cell monolayer on the bottom of the flask. Return to tissue culture incubator at 37 °C with 5% CO_2. Filter collected viral supernatant through 0.45 μm syringe filter to remove any cell debris. This will also remove bacterial contaminants. Do not use 0.22 μm filter as this can shear viral particles. Viral supernatant can be used immediately to transduce cells (below), can be stored at 4 °C for up to 1 week, or can be frozen at −80 °C for long-term storage (up to 1 year).
8. (Days 6–7) **Step 6** above can be repeated for 2 more days to collect additional viral supernatant for use or storage.

3.1.2 Isolation of Hematopoietic Cells from Murine Bone Marrow

1. Euthanize 4–6-week-old mouse using inhaled carbon dioxide. Isolate femurs from both hind legs. Cut off both ends of each bone, and use a syringe with 21-gauge needle to flush out the bone marrow from each femur with 3 mL of tissue culture media into a 15 mL conical tube [31].
2. Generate a single cell suspension by mixing several times with 5 mL serological pipette.
3. Add tissue culture media to bring to a total of 10 mL.
4. Count cells. Typical cell yield from 2 femurs is approximately 16–20 × 10^6 total cells.

3.1.3 Transduction of Murine Hematopoietic Cells with V-abl Retrovirus

1. Centrifuge cell suspension from Subheading 3.1.2 at 300 × *g* for 5 min at room temperature. Aspirate supernatant, and resuspend at 2 × 10^6 cells/mL in DMEM media +20% FBS.
2. Plate 1 mL (2 × 10^6 cells) per well in a six-well plate. Typically, plate six wells per mouse.
3. Add 1 mL of viral supernatant containing 10 μg/mL polybrene to each well of cells. Final volume will be 2 mL/well with final polybrene concentration of 5 μg/mL.
4. Incubate in tissue culture incubator at 37 ° C with 5% CO_2 untouched for 5 days.
5. On fifth day of culture, add 1 mL of DMEM media +20% FBS to each well.

6. Transformed cells will initially appear as clusters of large cells on the bottom of the well and will divide rapidly from there. Monitor the cells every 2–3 days. Add 1 mL of media on day 9 or 10. Once cells reach a density that covers bottom of well, they can be split 1:2 with fresh DMEM media +20% FBS. The first split should be into a new six-well plate to move away from adherent fibroblasts on the bottom of the well that may also be transformed by the v-abl virus (*see* **Note 3**).
7. Continue to split cells 1:2 when they reach a cell density that covers the bottom of the well. Once cells are being split every 2–3 days, increase to 1:3 split within the six-well plate for approximately 3–4 weeks. Cells can then be moved to T25 flask and split 1:5 into 10 mL of DMEM media +20% FBS.
8. Once in the T25 flask, and cells are being split every 2–3 days at 1:5, they can be transitioned to tissue culture media with 10% FBS. Cell growth will initially slow down with this decrease in serum concentration. Once growth recovers and cells are being split 1:5 every 2–3 days, they are ready to be used for experiments.
9. Cells can be maintained by splitting every 2–3 days at 1:5 or 1:10.
10. Cells can be frozen in aliquots of 20–30 $\times$ 10^6 cells in 1 mL of FBS + 10% DMSO at −80 °C or in vapor phase of liquid nitrogen tank.
11. To thaw cells, run vial under hot water or place in 37 °C water bath until thawed. Add the 1 mL of cell suspension to 10 mL of DMEM media +10% FBS, and centrifuge at 300 $\times$ *g* at room temperature to pellet cells. Aspirate supernatant. Resuspend cells in 10 mL of DMEM media +10% FBS, and incubate in tissue culture incubator at 37 °C with 5% CO_2. Cells will start expanding in 2–5 days and can be maintained as in **step 8** above.

3.2 Establishing Primary Pre-B Cell Cultures

1. Centrifuge cell suspension from Subheading 3.1.2 at 300 $\times$ *g* for 5 min at room temperature. Aspirate supernatant. Resuspend cells at 4 $\times$ 10^6 cells/3 mL (~1.3 $\times$ 10^6 cells/mL) in media with 10% FBS and 5 ng/mL of recombinant murine IL-7.
2. Place 3 mL (4 $\times$ 10^6 cells/mL) into four wells of a six-well plate.
3. Incubate in tissue culture incubator at 37 ° C with 5% CO_2.
4. Leave the cells untouched in the six-well plates for 4–5 days. Pre-B cells will selectively expand, and after 5–7 days, cultures will be comprised only of pre-B cells. Pre-B cells grow in

suspension. Numerous adherent stromal cells will be present on the bottom of the wells at the end of this culture period.

5. On days 4–5 of culture, combine all wells from the same mouse, and transfer to a T75 flask. Add an additional 8 mL of fresh DMEM media containing 10% FBS and 5 ng/mL of IL-7. Final volume will be approximately 20 mL (*see* **Note 4**). Incubate in tissue culture incubator at 37 °C with 5% CO_2.
6. Cells can be used for experiments beginning on day 7 of culture.
7. To maintain cell cultures, add fresh DMEM media with 10% FBS and 5 ng/mL IL-7 every 2–3 days to keep concentration of the cells between 3–5 × 10^6 cells/mL (*see* **Note 5**). Cell cultures should be continually expanded by addition of media and not split as with traditional immortalized cell lines. Additionally, culture media should not be replaced but rather additional media added as needed. As cells continue in culture, a monolayer of adherent stromal cells will accumulate. Transfer the pre-B cells (in suspension) to a new flask when stromal cells occupy more than 30–50% of surface area of the flask. Cultures can be maintained for 6–8 weeks until cells undergo senescence.

3.3 Induction of RAG-Mediated DSBs

3.3.1 RAG DSBs in Abl Pre-B Cells

1. Prepare imatinib stock solution at 30 mM in DMSO. Aliquots can be stored at −20 ° C.
2. Count abl pre-B cells, and resuspend at 1 × 10^6 cells/mL in tissue culture media with 10% FBS and 3 μM imatinib. Place in appropriate tissue culture vessel (*see* **Note 6**).
3. After incubation with imatinib, abl pre-B cells undergo cell cycle arrest, induce RAG1 and RAG2 expression, and generate RAG-mediated DSBs. Cell cycle arrest can be confirmed by visual change to smaller cell size by microscopy or by flow cytometry.
4. RAG DSBs are generated beginning at ~24 h after addition of imatinib, and maximal generation is at 48 h post-imatinib.
5. Cells can be collected at desired timepoints after addition of imatinib, for example, at 24, 48, and 72 h post-imatinib. Genomic DNA is collected for measurement of DNA breaks. Other samples can be collected as needed for desired endpoints, i.e., protein, RNA, and cells for immunofluorescence to evaluate DNA damage foci.
6. Abl pre-B cells will undergo cell death after treatment with imatinib. Cell death will occur gradually over 72 h (*see* **Note 7**).

3.3.2 RAG DSBs in Primary Pre-B Cells

1. Count primary pre-B cells, and resuspend cells at 2×10^6 cells per mL in tissue culture media with 10% FBS but *without* IL-7.
2. After removal of IL-7, primary pre-B cells undergo cell cycle arrest, induce RAG1 and RAG2 expression, and generate RAG-mediated DSBs. Cell cycle arrest can be confirmed by visual change to smaller cell size by microscopy or by flow cytometry.
3. RAG DSBs are generated beginning at ~24 h after IL-7 withdrawal, and maximal generation is at 48 h post IL-7 withdrawal.
4. Cells can be collected at desired timepoints after IL-7 withdrawal, for example, at 24, 48, and 72 h post-withdrawal. Genomic DNA is collected for measurement of DNA breaks. Other endpoints can include protein, RNA, and cells for immunofluorescence to evaluate DNA damage foci.
5. Primary pre-B cells will undergo cell death after withdrawal of IL-7. Cell death will occur gradually over 72 h (*see* **Note 7**).

3.4 Quantitation of DNA Breaks

3.4.1 Isolation of Genomic DNA

1. Collect cells (between $2–5 \times 10^6$ cells) at 24–72 h after addition of imatinib or withdrawal of IL-7. Centrifuge at $300 \times g$ for 5 min at room temperature to pellet cells. Aspirate supernatant.
2. Resuspend cells in 500 μL Rapid Lysis Buffer with Proteinase K (*see* **Note 8**). Buffer is stored at 4 °C and should be warmed to 37 °C before use to dissolve SDS. To ensure the pellet is resuspended, rake tube across a tube rack rather than mixing by pipetting as the suspension becomes too viscous.
3. Incubate at 55 °C for 4 h to overnight. Mix well with wide bore 1000 mL pipette tip. Transfer to 1.5 mL Eppendorf tube.
4. Add 500 μL isopropanol. Invert tubes several times to precipitate DNA. Centrifuge at $12{,}000 \times g$ for 1 min to pellet DNA. Gently decant supernatant into a small beaker, being careful not to lose DNA pellet.
5. Wash pellet with 100 μL 70% ethanol. Ensure the pellet is resuspended and doesn't remain fixed to the bottom of the tube. Centrifuge at $12{,}000 \times g$ for 1 min. Remove supernatant with pipette. Do not decant as pellet may be lost.
6. Wash pellet with 100 μL 100% ethanol. Centrifuge as above, and remove supernatant with pipette. Air dry for 5 min to allow any residual ethanol in the tube to evaporate.

7. Add 400 μL TE, and incubate at 55 °C for ≥1 h to resuspend DNA pellet. Mix with wide bore 1000 mL pipette and then with 200 μL pipette to ensure uniform resuspension. Solution will be very viscous and will be easier to pipette as DNA is resuspended.
8. Quantitate DNA. DNA can be stored at 4 °C.

3.4.2 Quantitative PCR Protocol for Measuring DNA Breaks

1. To quantify DSBs, quantitative PCR (qPCR) is run using primers that span the DSB site. Reduction in PCR product relative to a control sample without RAG DSBs measures DSB generation (controls can be from pre-B cells prior to imatinib or IL-7 withdrawal or can be from *Rag*-deficient pre-B cells, which do not generate RAG DSBs). qPCR can be conducted with SYBR Green or with TaqMan probes. We describe SYBR green approach here. This method for quantitating DSBs is applicable to cells with DNA repair defects where the DSB is not repaired. For repair-sufficient pre-B cells, the method is not appropriate as the DSB is repaired and PCR will not be able to detect whether a DSB break was generated.
2. Following genomic DNA isolation, digest 2 μg of DNA with NEBNext dsDNA Fragmentase for 10 min following manufacturer's protocol.
3. Use a commercial PCR cleanup kit to purify the DNA from above.
4. Quantitate DNA. Dilute DNA to 12.5 ng/μL.
5. Prepare 96-well plate for qPCR. In each well, combine 8.5 μL Brilliant II SYBR Green QPCR Master Mix with forward and reverse primers at final concentrations of 0.1 μM each. Bring final volume up to 15 μL with H_2O. Add 2 μL diluted DNA (25 ng total DNA) to each well. Each experimental sample and control are run in triplicate with primers spanning the DSB site as well as primers for a control region of uncut genomic DNA, such as *CD19* (*see* **Note 9**).
6. qPCR reaction: 95 °C × 10 min followed by 40 cycles of 95 °C × 30 s, 55 °C × 1 min (data collected here), 72 °C × 1 min. Melt curve can be run at the end of the 40 cycles, if desired: 95 °C × 1 min, 55 °C × 30 s, 95 °C 30 s.
7. Data analysis uses ΔΔCT calculation. First, calculate the ΔCt (sample) by subtracting the mean Ct (cycle threshold) value for the three replicates of qPCR across the control region (termed "Ct(sample, uncut site)") from the mean Ct for the three replicates of qPCR across the DSB (termed "Ct(sample, DSB site)"). Repeat this calculation to obtain the ΔCt(control) for the control sample that does not have any DSBs. To obtain

ΔΔCT, subtract ΔCt(control) from ΔCt(sample). Relative fold change is calculated as $2^{-\Delta\Delta CT}$. Calculations are as follows:

(a) ΔCt(sample) = Ct(sample, DSB site) − Ct(sample, uncut site).

(b) ΔCt(control) = Ct(control, DSB site) − Ct(control, uncut site).

(c) ΔΔCT = ΔCt(sample) − ΔCt(control).

(d) Relative fold change = $2^{-\Delta\Delta CT}$.

3.5 Measurement of DNA Damage Responses

3.5.1 Protein and RNA Collection from Abl Pre-B Cells and Primary Pre-B Cells

1. Collect cells (between 2–10 × 10^6 cells) at 24–72 h after addition of imatinib or withdrawal of IL-7. Centrifuge at 300 × *g* for 5 min at room temperature to pellet cells. Aspirate supernatant.
2. Prepare protein and RNA samples per standard protocols.

3.5.2 Immunofluorescence of DNA Damage Foci in Abl Pre-B Cells and Primary Pre-B Cells

1. Prepare coverslips. Mix Cell-Tak and 2 M sodium carbonate in a 4:1 ratio, and then add isopropanol to a final concentration of 3%. Place 8 μL of this solution on 18 mm circular coverslip, and spread using the edge of a rectangular coverslip to evenly coat the surface. Allow to air dry and then rinse with ddH_2O. Air dry and store at 4 °C until use (can be stored for up to 2 months).
2. Collect 1 × 10^6 cells at 24–72 h after addition of imatinib or withdrawal of IL-7. Centrifuge at 300 × *g* for 5 min at room temperature to pellet cells. Aspirate supernatant.
3. Resuspend cells in 1 mL phosphate buffered saline (PBS). Centrifuge at 300 × *g* for 5 min at room temperature. Aspirate supernatant.
4. Resuspend cells in 100 μL PBS, and add cell suspension dropwise to the prepared coverslip (from **step 1**). Allow to attach for 20 min at 37 °C.
5. Wash coverslip with 1 mL of PBS.
6. Standard immunofluorescence protocols can be used for evaluating DNA damage foci.

4 Notes

1. For smaller quantities of viral supernatant, PlatE cells can be cultured in six-well plates. Culture 1 × 10^6 cells in 3 mL tissue culture media with 10% FBS per well. DNA/Lipofectamine is generated using 20 μL Lipofectamine 2000 in 0.5 mL of

DMEM (no additives) and 8 μg plasmid in 0.5 mL of DMEM (no additives).

2. Inclusion of FBS in the DMEM in **steps 3** and **4** will interfere with transfection efficiency. Inclusion of antibiotics in **steps 2–4** will be toxic to the PlatE cells and will result in cell death.
3. Do not split the cells unless they are fairly dense in the well as, at this point, they are sensitive to being too sparse.
4. Flow cytometry with anti-B220 antibody can be done to confirm culture is comprised only of B cells.
5. Lower concentrations can stress the cells.
6. Typically, we include a *Rag1*$^{-/-}$ abl pre-B cell line or primary pre-B cell culture, which does not generate DSBs, as a negative control and an *Artemis*$^{-/-}$ abl pre-B cell line or primary pre-B cell culture, which generates but cannot repair RAG DSBs, as a positive control.
7. BCL2 transgene can be expressed in abl pre-B cells or primary pre-B cell cultures to block cell death and promote survival after imatinib treatment. This can be done by either transduction of BCL2-expressing retrovirus or generation of abl pre-B cells or primary pre-B cell cultures from BCL2 transgenic mice (Subheadings 3.1 and 3.2).
8. Commercial DNA isolation kits can be used in place of **steps 2–7**.
9. We typically use CD19 as a normalization control. Any genomic DNA region without a DNA break can be used as a normalization control.

Acknowledgments

RJ is supported by a training grant through the Alvin J. Siteman Cancer Center. JJB is supported by National Institutes of Health grant R56 AI153234, an American Society of Hematology Scholar Award, the Children's Discovery Institute at St. Louis Children's Hospital and Washington University School of Medicine, the St. Louis Children's Hospital Foundation, and the Gabrielle's Angel Foundation.

References

1. Ciccia A, Elledge SJ (2010) The DNA damage response: making it safe to play with knives. Mol Cell 40(2):179–204. https://doi.org/10.1016/j.molcel.2010.09.019
2. Matsuoka S, Ballif BA, Smogorzewska A, McDonald ER 3rd, Hurov KE, Luo J et al (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. Science 316(5828):1160–1166. https://doi.org/10.1126/science.1140321
3. Shiloh Y, Ziv Y (2013) The ATM protein kinase: regulating the cellular response to

genotoxic stress, and more. Nat Rev Mol Cell Biol 14(4):197–210. https://doi.org/10.1038/nrm3546

4. Alt FW, Zhang Y, Meng FL, Guo C, Schwer B (2013) Mechanisms of programmed DNA lesions and genomic instability in the immune system. Cell 152(3):417–429. https://doi.org/10.1016/j.cell.2013.01.007
5. Helmink BA, Sleckman BP (2012) The response to and repair of RAG-mediated DNA double-strand breaks. Annu Rev Immunol 30:175–202. https://doi.org/10.1146/annurev-immunol-030409-101320
6. Bassing CH, Swat W, Alt FW (2002) The mechanism and regulation of chromosomal V(D)J recombination. Cell 109(Suppl):S45–S55. https://doi.org/10.1016/s0092-8674(02)00675-x
7. Fugmann SD, Lee AI, Shockett PE, Villey IJ, Schatz DG (2000) The RAG proteins and V(D)J recombination: complexes, ends, and transposition. Annu Rev Immunol 18:495–527. https://doi.org/10.1146/annurev.immunol.18.1.495
8. Muljo SA, Schlissel MS (2003) A small molecule Abl kinase inhibitor induces differentiation of Abelson virus-transformed pre-B cell lines. Nat Immunol 4(1):31–37. https://doi.org/10.1038/ni870
9. Bredemeyer AL, Sharma GG, Huang CY, Helmink BA, Walker LM, Khor KC et al (2006) ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. Nature 442(7101):466–470. https://doi.org/10.1038/nature04866
10. Bredemeyer AL, Helmink BA, Innes CL, Calderon B, McGinnis LM, Mahowald GK et al (2008) DNA double-strand breaks activate a multi-functional genetic program in developing lymphocytes. Nature 456(7223):819–823. https://doi.org/10.1038/nature07392
11. Rolink A, Kudo A, Karasuyama H, Kikuchi Y, Melchers F (1991) Long-term proliferating early pre B cell lines and clones with the potential to develop to surface Ig-positive, mitogen reactive B cells in vitro and in vivo. EMBO J 10(2):327–336
12. Bednarski JJ, Nickless A, Bhattacharya D, Amin RH, Schlissel MS, Sleckman BP (2012) RAG-induced DNA double-strand breaks signal through Pim2 to promote pre-B cell survival and limit proliferation. J Exp Med 209(1):11–17. https://doi.org/10.1084/jem.20112078
13. Bednarski JJ, Pandey R, Schulte E, White LS, Chen BR, Sandoval GJ et al (2016) RAG-mediated DNA double-strand breaks activate a cell type-specific checkpoint to inhibit pre-B cell receptor signals. J Exp Med 213(2):209–223. https://doi.org/10.1084/jem.20151048
14. Johnson K, Hashimshony T, Sawai CM, Pongubala JM, Skok JA, Aifantis I et al (2008) Regulation of immunoglobulin light-chain recombination by the transcription factor IRF-4 and the attenuation of interleukin-7 signaling. Immunity 28(3):335–345. https://doi.org/10.1016/j.immuni.2007.12.019
15. Ochiai K, Maienschein-Cline M, Mandal M, Triggs JR, Bertolino E, Sciammas R et al (2012) A self-reinforcing regulatory network triggered by limiting IL-7 activates pre-BCR signaling and differentiation. Nat Immunol 13(3):300–307. https://doi.org/10.1038/ni.2210
16. Steinel NC, Lee BS, Tubbs AT, Bednarski JJ, Schulte E, Yang-Iott KS et al (2013) The ataxia telangiectasia mutated kinase controls Igkappa allelic exclusion by inhibiting secondary Vkappa-to-Jkappa rearrangements. J Exp Med 210(2):233–239. https://doi.org/10.1084/jem.20121605
17. Lin WC, Desiderio S (1994) Cell cycle regulation of V(D)J recombination-activating protein RAG-2. Proc Natl Acad Sci U S A 91(7):2733–2737. https://doi.org/10.1073/pnas.91.7.2733
18. Soodgupta D, White LS, Yang W, Johnston R, Andrews JM, Kohyama M et al (2019) RAG-mediated DNA breaks attenuate PU.1 activity in early B cells through activation of a SPIC-BCLAF1 complex. Cell Rep 29(4):829–43 e5. https://doi.org/10.1016/j.celrep.2019.09.026
19. Collins PL, Purman C, Porter SI, Nganga V, Saini A, Hayer KE et al (2020) DNA double-strand breaks induce H2Ax phosphorylation domains in a contact-dependent manner. Nat Commun 11(1):3158. https://doi.org/10.1038/s41467-020-16926-x
20. Purman CE, Collins PL, Porter SI, Saini A, Gupta H, Sleckman BP et al (2019) Regional gene repression by DNA double-strand breaks in G1 phase cells. Mol Cell Biol 39(24):e00181-19. https://doi.org/10.1128/MCB.00181-19
21. Bednarski JJ, Sleckman BP (2019) At the intersection of DNA damage and immune responses. Nat Rev Immunol 19(4):231–242. https://doi.org/10.1038/s41577-019-0135-6
22. Difilippantonio S, Gapud E, Wong N, Huang CY, Mahowald G, Chen HT et al (2008) 53BP1 facilitates long-range DNA

end-joining during V(D)J recombination. Nature 456(7221):529–533. https://doi.org/10.1038/nature07476

23. Gapud EJ, Sleckman BP (2011) Unique and redundant functions of ATM and DNA-PKcs during V(D)J recombination. Cell Cycle 10(12):1928–1935. https://doi.org/10.4161/cc.10.12.16011
24. Tubbs AT, Dorsett Y, Chan E, Helmink B, Lee BS, Hung P et al (2014) KAP-1 promotes resection of broken DNA ends not protected by gamma-H2AX and 53BP1 in G(1)-phase lymphocytes. Mol Cell Biol 34(15):2811–2821. https://doi.org/10.1128/MCB.00441-14
25. Helmink BA, Tubbs AT, Dorsett Y, Bednarski JJ, Walker LM, Feng Z et al (2011) H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes. Nature 469(7329):245–249. https://doi.org/10.1038/nature09585
26. Lescale C, Abramowski V, Bedora-Faure M, Murigneux V, Vera G, Roth DB et al (2016) RAG2 and XLF/Cernunnos interplay reveals a novel role for the RAG complex in DNA repair. Nat Commun 7:10529. https://doi.org/10.1038/ncomms10529
27. Lescale C, Lenden Hasse H, Blackford AN, Balmus G, Bianchi JJ, Yu W et al (2016) Specific roles of XRCC4 paralogs PAXX and XLF during V(D)J recombination. Cell Rep 16(11):2967–2979. https://doi.org/10.1016/j.celrep.2016.08.069
28. Hung PJ, Johnson B, Chen BR, Byrum AK, Bredemeyer AL, Yewdell WT et al (2018) MRI is a DNA damage response adaptor during classical non-homologous end joining. Mol Cell 71(2):332–42 e8. https://doi.org/10.1016/j.molcel.2018.06.018
29. Savic V, Yin B, Maas NL, Bredemeyer AL, Carpenter AC, Helmink BA et al (2009) Formation of dynamic gamma-H2AX domains along broken DNA strands is distinctly regulated by ATM and MDC1 and dependent upon H2AX densities in chromatin. Mol Cell 34(3):298–310. https://doi.org/10.1016/j.molcel.2009.04.012
30. Ba Z, Lou J, Ye AY, Dai HQ, Dring EW, Lin SG et al (2020) CTCF orchestrates long-range cohesin-driven V(D)J recombinational scanning. Nature 586(7828):305–310. https://doi.org/10.1038/s41586-020-2578-0
31. Liu X, Quan N (2015) Immune cell isolation from mouse femur bone marrow. Bio Protoc 5(20):e1631. https://doi.org/10.21769/bioprotoc.1631

Chapter 6

Studying Single-Stranded DNA Gaps at Replication Intermediates by Electron Microscopy

Jessica Jackson and Alessandro Vindigni

Abstract

Single-stranded DNA gaps are frequent structures that accumulate on newly synthesized DNA under conditions of replication stress. The identification of these single-stranded DNA gaps has been instrumental to uncover the mechanisms that allow the DNA replication machinery to skip intrinsic replication obstacles or DNA lesions. DNA fiber assays provide an essential tool for detecting perturbations in DNA replication fork dynamics genome-wide at single molecule resolution along with identifying the presence of single-stranded gaps when used in combination with S1 nuclease. However, electron microscopy is the only technique allowing the actual visualization and localization of single-stranded DNA gaps on replication forks. This chapter provides a detailed method for visualizing single-stranded DNA gaps at the replication fork by electron microscopy including psoralen cross-linking of cultured mammalian cells, extraction of genomic DNA, and finally enrichment of replication intermediates followed by spreading and platinum rotary shadowing of the DNA onto grids. Discussion on identification and analysis of these gaps as well as on the advantages and disadvantages of electron microscopy relative to the DNA fiber technique is also included.

Key words Electron microscopy, DNA replication, Replication structures, ssDNA gaps, DNA replication stress

1 Introduction

Single-stranded DNA (ssDNA) discontinuities (or gaps) are genome-destabilizing structures that need to be quickly repaired (or filled) to prevent DNA breakage and genome instability. ssDNA gaps accumulate both on leading and lagging strands of DNA replication forks after treatment with a wide range of DNA-damaging agents [1–7]. Lagging strand gaps can form because of the discontinuous nature of Okazaki fragment synthesis. On the other hand, leading strand gaps form when DNA synthesis resumes downstream of a replication-blocking lesion through a process called fork repriming. Repriming involves re-initiation of DNA synthesis beyond a DNA lesion, leaving unreplicated ssDNA

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_6,

gaps to be filled post-replicatively, and is mediated by human Primase and DNA-directed Polymerase (PRIMPOL) in mammalian cells [8–11].

The DNA fiber approach exploits the ability of many organisms to incorporate halogenated pyrimidine nucleoside analogs into replicating DNA and provides a powerful tool to monitor genome-wide replication perturbations at single-molecule resolution [12–17]. Ongoing replication events are typically labeled with two thymidine analogs—e.g., iododeoxyuridine (IdU) and chlorodeoxyuridine (CldU)—and individual two-color labeled DNA tracts are visualized on stretched DNA fibers by immunofluorescence. This approach can be adapted to detect ssDNA gaps on the labeled strand by taking advantage of the unique enzymatic cleavage properties of the S1 nuclease. The S1 nuclease from *Aspergillus oryzae* [18] is used to nick the ssDNA and convert the ssDNA gap into a double-stranded break [19]. In the presence of ssDNA gaps, treatment with the S1 nuclease leads to shorter thymidine labeled tracts, which can be detected by DNA fiber analysis [19]. This approach can detect ssDNA gaps as short as 1–3 nucleotides. However, it also has some important limitations including the inability to determine the actual size of the ssDNA gaps as well as their exact location on the newly synthesized DNA.

Electron microscopy (EM) is the only technique that allows direct visualization and quantification of replication intermediates [20–22]. EM has been applied to study ssDNA gaps on replication forks as it can provide unique structural insight including the size and location of the gaps on the replication forks. This chapter focuses on the experimental procedures related to the in-cell psoralen cross-linking of mammalian cell cultures, DNA extraction, enrichment, spreading, platinum shadowing, and finally viewing of the actual DNA replication intermediates and ssDNA gaps via EM (Fig. 1). Details on the carbon coating of grids for use in DNA spreading are also noted. These methods have been adapted from previously published protocols [20, 23] with emphasis on identification and analysis of ssDNA gaps within the newly synthesized DNA daughter strands. In addition, we discuss how the S1 nuclease DNA fiber assay can be used in conjunction with EM, as well as the advantages and disadvantages of using EM versus DNA fiber.

2 Materials

2.1 In-Cell Psoralen Cross-Linking and Lysis

1. 150 mm tissue culture dish.
2. Cold 1× PBS (Phosphate Buffer Saline).
3. 60 mm tissue culture dish.
4. 200 μg/ml 4,5′,8-Trimethylpsoralen (TMP) in 100% ethanol. Care should be taken when handling TMP-containing

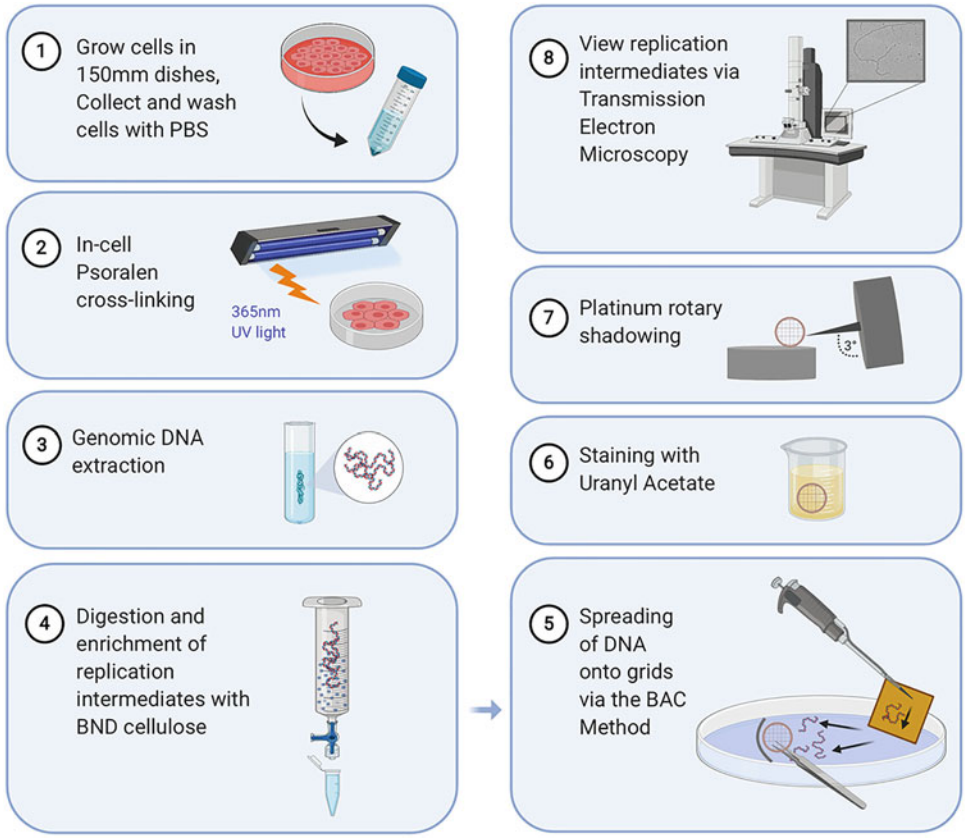


Fig. 1 Illustrative depiction of the electron microscopy protocol summarizing the main steps

solutions, including appropriate personal protection equipment along with use in conjunction with a certified fume hood and proper disposal, due to its potential DNA-damaging properties.

5. UV Cross-linker, 365 nm.
6. Flat frozen ice pack(s).
7. Flat metal plate cooled to −20 °C, no larger than the width of the interior of the cross-linker. The cold metal plate will be placed on top of the frozen ice packs and used to cool the samples, in the 60 mm petri dishes, in order to prevent heating during the incubation and irradiation times during cross-linking.
8. Cold Lysis Buffer stock solution: 1.28 M Sucrose, 40 mM Tris–HCL pH 7.5, 20 mM $MgCl_2$, 4% Triton X-100.

2.2 Genomic DNA Extraction

1. Cold 1× PBS.
2. Digestion Buffer: 800 mM guanidine-HCl, 30 mM Tris–HCl pH 8.0, 30 mM EDTA pH 8.0, 5% Tween-20, 0.5% Triton X-100. Store at RT.

3. 20 mg/ml proteinase K: made up fresh in digestion buffer.
4. Chloroform/Isoamyl Alcohol 24:1. Chloroform is highly toxic. Any handling of chloroform-containing solutions should be used within a certified fume hood.
5. 100% Isopropanol.
6. 70% Ethanol.
7. Nalgene Oak Ridge High-Speed Centrifuge Tubes, 50 ml, specifically designed for chloroform extractions (Thermo Fisher).
8. 1× TE Buffer.

2.3 DNA Digestion and Enrichment of Replication Intermediates

1. PvuII-HF Restriction Enzyme.
2. 10 mg/ml RNase A.
3. Low Salt Buffer: 10 mM Tris–HCl pH 8.0, 300 mM NaCl. Store at RT.
4. High Salt Buffer: 10 mM Tris–HCl pH 8.0, 1 M NaCl. Store at RT.
5. 1.8% Caffeine Solution: made up in High Salt Buffer, incubated at 50 °C, and requires thorough mixing and vortexing in order to completely dissolve into solution. Store at RT for up to 4 months. Make new if precipitate has formed.
6. 100 mg/ml BND Cellulose: made up in Low Salt Buffer; mix thoroughly by vortexing before use. Store at 4 °C for up to 2 months.
7. Poly-Prep Chromatography Columns, 2 ml.
8. Amicon Ultra 0.5 ml Centrifugal Filter Tubes.
9. 1× TE Buffer.

2.4 Carbon Coating of the Grids

1. 25 × 25 mm mica sheet (Electron Microscopy Sciences) pre-coated with carbon 12 nm in thickness. Pre-coated sheets may be obtained through any departmental core equipped with a coating machine with a carbon gun. Carbon-coated mica sheets may be stored for up to 5 months.
2. Scotch tape solution: 30 cm of clear Scotch brand tape in 100 ml of chloroform. After thoroughly mixing, the adhesive on the tape will dissolve into the chloroform, leaving just the backing. Store at RT for up to 1 year.
3. Copper 400 mesh square grids (Electron Microscopy Sciences).

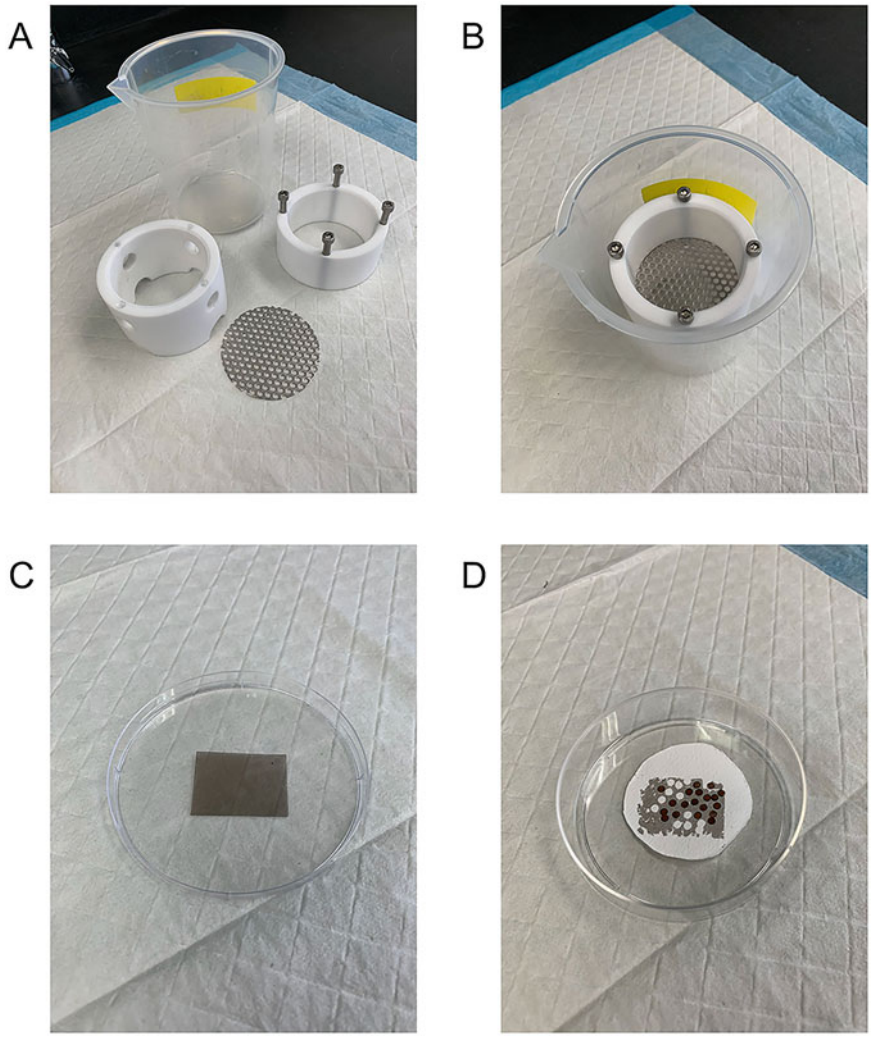


Fig. 2 (**a**, **b**) Apparatus used to apply a 12 nm carbon layer on top of the grids. (**c**) Piece of carbon-coated mica sheet from which the carbon layer is taken. (**d**) Completed carbon-coated grids

4. Filter paper circles: diameter 45 mm.
5. Filter paper circles: diameter 90 mm.
6. Apparatus capable of holding approximately 100 ml of double distilled water in an area 90 mm wide and 180 mm deep. These may be custom produced in any machine shop and can be made up of two Teflon or plastic rings that can be separated with a removable metal wire mesh middle to allow for water flow and placement of grids (Fig. 2a,b).

2.5 Spreading of the DNA

1. Double distilled water.
2. 33.3 μg/ml ethidium bromide: made up fresh in sterile double distilled water. Sterile water should be used in conjunction with

highly purified ethidium bromide in order to eliminate any particulates in solution that could become adhered to the grids and interfere with the visualization of DNA molecules. Ethidium bromide is highly toxic and should be used within a certified fume hood due to its DNA-damaging properties.

3. 100% formamide.
4. 0.2% w/v benzyl-dimethyl-alkylammonium chloride (BAC): made up in formamide. Store at RT for up to 1 year.
5. 1:10 BAC diluted in 1× tris EDTA (TE) buffer: made up fresh before use.
6. 100% ethanol.
7. 5 mM uranyl acetate (UrAc): made up in 5 mM HCl. Store at 4 °C for up to 1 year. UrAc has some sensitivity to light and should be stored in the dark or thoroughly covered. UrAc has a mild radioactivity level of 0.37–0.51 μCi/g, and appropriate personal protective equipment should be worn along with careful handling of the powdered form in a certified fume hood.
8. 1:10 UrAc diluted in 100% ethanol.
9. 150 mm tissue culture dish.
10. Graphite powder.
11. Mica sheet: cut to approximately 12.5 mm × 10 mm, freshly cleaved before use.
12. Super fine tip tweezers.

2.6 Platinum Shadowing of Grids

1. High vacuum coater: capable of low angle (3 degrees) rotary shadowing, at least one platinum gun, quartz crystal monitor, rotary stage and holder for the copper grids.
2. Platinum rod inserts for gun.

2.7 Visualization of DNA Via Transmission Electron Microscope

1. Transmission electron microscope (JEOL 1400 or equivalent) with preferably bottom mounted camera system (AMT XR401 High Sensitivity sCMOS Camera for TEM or equivalent).
2. Camera software to allow for saving images as a .tiff file format.

2.8 Analysis of Single-Stranded DNA Gaps at the Replication Fork

1. ImageJ or similar software to measure the length of ssDNA gaps.
2. Excel software program to measure the presence of ssDNA gaps in a given sample.

3 Methods

3.1 In-Cell Psoralen Cross-Linking and Lysis

1. Asynchronous mammalian cells are grown in three 150 mm tissue culture plates, in 10–20 ml growth media, reaching a confluency of approximately 60–80% for each sample. Amount of growth media may vary due to treatment conditions; care should be taken to collect cells into one 50 ml tube to minimize loss of cells.
2. Scrape each plate, and collect cells into 50 ml tube in their growth media.
3. Add 5 ml 1× PBS, scraping each plate again then collecting cells into their same respective 50 ml tubes.
4. Spin cells down for 5 min, at 900 × *g* at 4 °C.
5. Remove media via vacuum, and add 10 ml cold 1× PBS to wash (*see* **Note 1**).
6. Remove PBS via vacuum, and add 10.5 ml cold 1×PBS, resuspend cells, and transfer to a 60 mm tissue culture dish. At this point, 1 ml of cells+PBS may be removed to a 1.5 ml tube and spun down, and the pellet can be frozen at −80 °C for future processing (e.g., protein or RNA extraction). If further processing is not necessary, 9.5 ml cold 1× PBS may be used to resuspend the cell pellet.
7. Place 60 mm dish on the pre-cooled metal block, and place on top of the frozen ice pack(s).
8. In hood, add 500 μl TMP stock solution to each dish in a circular pattern to evenly spread throughout sample. Cover with the dish top.
9. Place entire setup in dark for 5 min.
10. Take ice packs and metal block with samples on top, and place in the cross-linker. Program cross-linker to highest setting, 9999, and press start.
11. Repeat **step 8** an additional two times for a total of 3 TMP additions and cross-linking cycles (*see* **Note 2**).
12. Transfer cross-linked cells to fresh 15 ml tube. Add 1 ml 1× PBS to each 60 mm dish to wash and collect any cells left behind to their respective tubes. Repeat two more times for a total of three 1 ml PBS washes of the 60 mm dish.
13. Spin cells down for 5 min, at 900 × *g* at 4 °C.
14. Remove PBS + TMP, and collect for appropriate hazardous waste disposal. Resuspend cell pellet in 10 ml cold 1× PBS to wash.
15. Remove PBS via vacuum, and add another 10 ml cold 1× PBS to wash. Repeat once more for a total of three PBS washes.

16. Remove PBS via vacuum, and thoroughly resuspend pellet in 2 ml cold 1× PBS.
17. Dilute lysis buffer stock solution 1:4 with cold double distilled water.
18. Add 8 ml cold lysis buffer to 2 ml cell pellet+PBS, and invert several times to mix (do not vortex), and incubate on ice for 10 min.
19. Spin down lysed cells for 15 min at 300 × *g* at 4 °C.
20. Remove lysis+PBS via vacuum, and resuspend pellet in 4 ml cold lysis buffer to wash.
21. Spin down cells for an additional 15 min, at 300 × *g* at 4 °C.
22. Remove lysis buffer completely via vacuum, and freeze pellet at −80 °C overnight.

3.2 Genomic DNA Extraction

1. On ice, completely resuspend frozen nuclei pellet in 200 µl cold 1× PBS, using a cutoff pipet tip.
2. Add 5 ml digestion buffer, and mix via pipet (do not vortex).
3. Add 200 µl 20 mg/ml proteinase K solution, and incubate at 50 °C for 1.5 h in a water bath.
4. Remove samples and let cool to RT.
5. In hood, add 5 ml chloroform/isoamyl alcohol, vortex immediately four times to thoroughly mix the two layers, and pour into a 50 ml Oak Ridge centrifuge tube.
6. In hood, balance tubes with digestion buffer by using a scale to achieve exact measurements.
7. Spin for 30 min at 9000 × *g* at 4 °C.
8. Promptly remove tubes from centrifuge. There should now be a clear separation of both layers with a milky interface in the middle. In the hood, carefully transfer the upper layer to a new oak ridge centrifuge tube, using a cutoff pipet tip. Care should be taken to not disturb nor take up the white, cloudy interface. Additional chloroform extractions may be performed until interface is clear (*see* **Note 3**). Volume should be approximately 4 ml. Collect and appropriately discard the bottom layer as chloroform hazardous waste.
9. Add same volume, approximately 4 ml, of 100% isopropanol. Cap tube, and vigorously swirl to thoroughly mix in order to precipitate the DNA. Long white strands of DNA should be visible at this point and be concentrated into a single clump.
10. Balance tubes with isopropanol by using a scale to achieve exact measurements, and spin for 10 min at 9000 × *g* at 4 °C.

11. Carefully pipet off the isopropanol, without disturbing the DNA pellet. Pellet may be semi-transparent and could be difficult to see.
12. Wash pellet with 5 ml 70% ethanol.
13. Spin for 5 min at 9000 × *g* at 4 °C.
14. Pipet off most of the ethanol, without disturbing the pellet, and place in 50 °C water bath for 15 min or until dry. Keep tube uncapped and water bath top removed to facilitate drying. Pellet will be clear when dry and difficult to see.
15. Add 200 μl 1× TE buffer to pellet, cap tube, place in 50 ml tube holder, and gently rock overnight at RT to completely dissolve pellet.
16. Transfer dissolved DNA into a 1.7 ml tube using a cutoff pipet tip. Measure concentration with 1 μl on a NanoDrop. Measurement should be taken at least three times and average used as the final concentration.

3.3 DNA Digestion and Enrichment of Replication Intermediates

1. Digest 20 μg of genomic DNA with the PvuII HF restriction enzyme, along with appropriate buffer (*see* **Note 4**). Place at 37 °C for 4 h. Add 1 μl 10 mg/ml RNase to the reaction mix when 1 h remaining in the reaction. Volume of reaction mix should total 250 μl (*see* **Note 5**).
2. Add 2.0 ml 100 mg/ml BND cellulose resin (0.1 ml BND suspension/1 μg digested DNA) to a chromatography column, with a cutoff pipet tip, and allow liquid to flow through. Washing of the resin takes approximately 1 h and can begin 1 h before completion of the digestion.
3. Wash the resin 6× with 1 ml high salt buffer. Allow all liquid to run through column before adding next wash. Gently resuspend resin each time after the addition of buffer. Take care not to allow resin to dry as it may interfere with its interaction with the DNA when applied.
4. Wash column 6× with 1 ml low salt buffer. Gently resuspend resin each time after the addition of buffer (*see* **Note 6**).
5. After digestion is complete, adjust reaction mix to 300 mM NaCl final concentration by adding 5 M NaCl. Adjust to final volume of 600 μl by adding low salt buffer. Salt adjustment for PvuII HF restriction enzyme: 250 μl sample, 14.7 μl 5 M NaCl, and 335.3 μl low salt buffer (*see* **Note 7**).
6. Close bottom of column, and add the 600 μl NaCl adjusted sample to resin; incubate for 30 min at RT to allow for full binding of DNA to the BND cellulose. Gently resuspend the resin bed with a cutoff pipet tip every 10 min.

7. Remove cap and collect sample in a 1.5 ml tube. Save as "flowthrough DNA" in case of defective DNA binding to resin.
8. Wash DNA bound resin 2× with 1 ml high salt buffer. Resuspend resin after each addition. Collect flowthrough in a 2.5 ml tube, and save as "linear dsDNA."
9. Close bottom of column, and add 600 μl of pre-warmed 50 °C 1.8% caffeine/high salt buffer in order to elute the remaining DNA molecules from the BND cellulose. Incubate for 10 min at RT. Gently resuspend resin bed after 5 min.
10. Remove cap and collect flowthrough into a 1.5 ml tube as the "enriched replication intermediate DNA."
11. The DNA sample may now be purified and concentrated using Amicon Ultra size exclusion columns. Place the collection 600 μl enriched DNA sample into the column, which is placed into the collection tube, and spin for 8 min at 7600 × *g*.
12. Discard flowthrough and wash column with 200 μl 1× TE. Spin for 5 min at 7600 × *g*.
13. Discard flowthrough, and wash column one final time with 200 μl 1× TE. Spin for 3.5 min at 7600 × *g*.
14. Discard flowthrough, turn column upside down into a new collection tube, and use the short spin cycle to force the bound DNA from the column and into the collection tube.
15. Remove column, and the remaining 20–30 μl of 1× TE now contains the purified DNA enriched for replication intermediates.
16. Optional: 1 μl purified, enriched DNA may be loaded onto a 0.8% agarose gel to determine DNA quantity. 1 μl of the flowthrough and linear dsDNA can be loaded as well, as a control to confirm majority of DNA is obtained from the final elution (*see* **Note 8**).
17. Concentrate DNA samples using a standard speed vac for approximately 10 min on high. Final volume should be around 15–20 μl. Seal tubes containing DNA with parafilm to avoid evaporation of sample.

3.4 Carbon Coating of the Grids

1. Arrange 30 of the 400 mesh copper grids, shiny side down, on a piece of filter paper, and place in a glass petri dish.
2. In a certified fume hood, cover grids with the chloroform/tape mixture, using a dropper, and let dry. Repeat for a total of three chloroform/tape evaporation treatments (*see* **Note 9**). Remove filter paper containing sticky grids to a new covered petri dish.
3. Place a pre-coated 25 × 25 mm carbon-coated mica sheet (Fig. 2c) in a petri dish lined with a moist piece of filter paper at 37 °C for 30 min.

4. Set up grid coating apparatus (Fig. 2a,b) by placing the 90 mm × 180 mm chamber in a 500 ml beaker. Place a new 45 mm piece of filter paper inside the bottom of the chamber, and weigh down. Fill entire beaker with double distilled water to the top of the chamber. Remove weight on filter paper.
5. Place grids onto the filter paper, located inside the water-filled chamber, sticky side up, taking care to arrange the grids in a tight-packed area, the size of the mica sheet, without overlapping. Use super fine tip tweezers to handle the grids.
6. Take the carbon-coated mica sheet, carbon side up, parallel to the water, and carefully lower, starting with one side, onto the top of the water, inside the center of the chamber, where the grids are located below. Use a hooked super fine tip tweezers in order to angle the carbon-coated mica sheet into the water.
7. Proceed to slowly turn the mica sheet perpendicular to the surface of the water, angling down, in order to separate the carbon from the mica sheet. As the mica sheet is moved in a downward fashion, under the water, the carbon will remain at the surface. When the carbon is completely free from the mica sheet, slowly raise the mica sheet out of the water, being careful to not break or bring up the carbon at the surface.
8. With the carbon now floating at the surface of the water, directly above the grids, use a vacuum to slowly remove the water from the beaker, guiding the carbon over the grids as the water lowers, taking care not to break the fragile carbon sheet.
9. As the water lowers and empties, guide the carbon on top of the grids, making sure all are covered with carbon (Fig. 2d).
10. After the water is gone, remove the filter paper, containing the carbon-coated grids, onto a dry piece of filter paper, located in a petri dish.
11. Slightly cover the still wet carbon-coated grids, with the top of the dish, and allow to dry at least overnight before using. Carbon-coated grids may be used for up to 5 months.

3.5 Spreading of DNA

1. In a certified fume hood, place 20 μl of the working ethidium bromide solution onto a piece of parafilm, creating a drop onto which the carbon-coated grid, carbon side down, is placed on top (Fig. 3a). Incubate for 20 min at RT. Cover with a large petri dish to avoid evaporation.
2. Take grids from drop, gently remove the excess ethidium bromide by dabbing the grid on a piece of filter paper, and place carbon side down on top of a new piece of filter paper to dry.
3. Utilizing an enclosed space free of air currents (*see* **Note 10**), set up all components required for spreading.

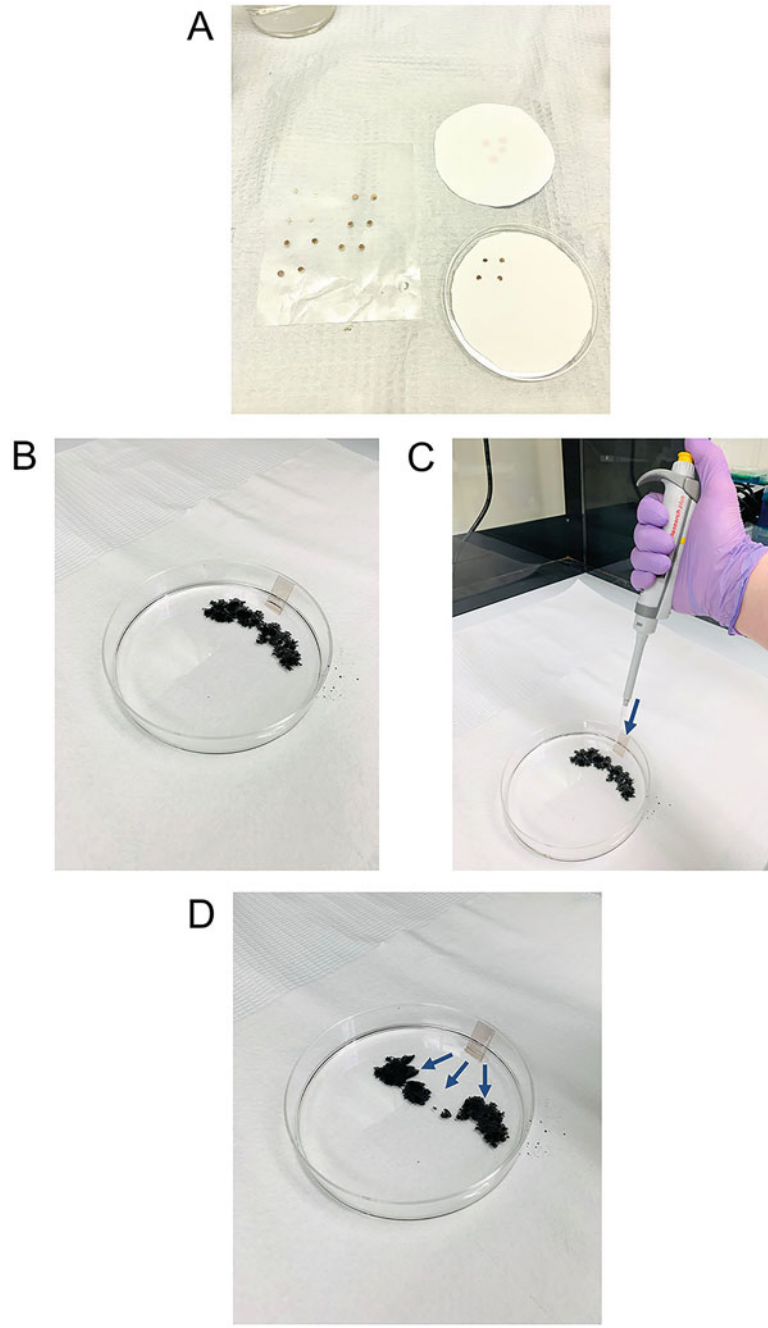


Fig. 3 Series of images representing important steps of the spreading process. (**a**) Carbon-coated grids are incubated with ethidium bromide, excess is removed by dabbing onto a piece of filter paper, and the grid is dried carbon side down onto a fresh piece of filter paper. (**b**) A 150 mm cell-culture dish is prepared for spreading by placing a mica sheet, at an angle, against the side of the dish, containing 20 ml of distilled water and then sprinkling granite powder into the water. (**c**) A drop of DNA is pipetted onto the mica sheet, very near to the surface of the water, and allowed to slide down into the water. (**d**) The DNA is then spread onto the surface of the water (blue arrows) accumulating at the edge of the granite powder wall

4. Pour 20 ml double distilled water into a 150 mm tissue culture dish.
5. Mix an amount of formamide (2–3.5 μl) equal to the amount of DNA being spread and 0.5 μl BAC working solution in the bottom of a 1.5 ml tube. Add an equal amount of DNA to the side of the 1.5 ml tube, keeping the two mixes separate until immediately before spreading (*see* **Note 11**). Volume of DNA used depends on the concentration of sample. Typical amount used is 3 μl of enriched replication intermediate DNA, but can be adjusted if concentrations are too high/low.
6. Position a freshly cleaved 12.5 mm × 10 mm piece of mica, with tweezers, into the 150 mm dish containing 20 ml water at a 45 degree angle against the side of the dish, partially submerged into the water.
7. Sprinkle a small amount of graphite powder close to the mica sheet, forming a wall around it (Fig. 3b).
8. Shortly spin sample containing the formamide, BAC, and DNA in a minifuge to mix all components together (spin for 5–10 s).
9. Pipet up the mix, and place droplet directly above the mica sheet, right above the water line, to allow DNA mix to slide down the mica sheet and spread out over the surface of the water (Fig. 3c), concentrating along the edge of the graphite powder wall (*see* **Note 12**). As the DNA mix hits the water, it will force the graphite powder wall to expand outwards as it travels across the surface of the water. Thus, the DNA will be most prominently located near the graphite's edge (Fig. 3d).
10. Using a fine tip tweezers, position the grid as parallel as possible to the water, and lightly touch the surface, nearest to the graphite wall, taking care not to take up the graphite itself.
11. Immediately remove grid from the water, and, using the tweezers, hold the grid into the diluted uranyl acetate/ethanol solution for 15 s, and then dip the grid into 100% ethanol, to wash away any excess uranyl acetate. Place grid carbon side up on a piece of filter paper to dry.
12. Wipe down tweezers to remove excess ethanol (*see* **Note 13**). A second grid can now be obtained starting from **step 10**, taking care to collect DNA from a different area, near the graphite wall, in order to collect a maximum amount of DNA (*see* **Note 14**).
13. Grids are now ready for platinum shadowing.

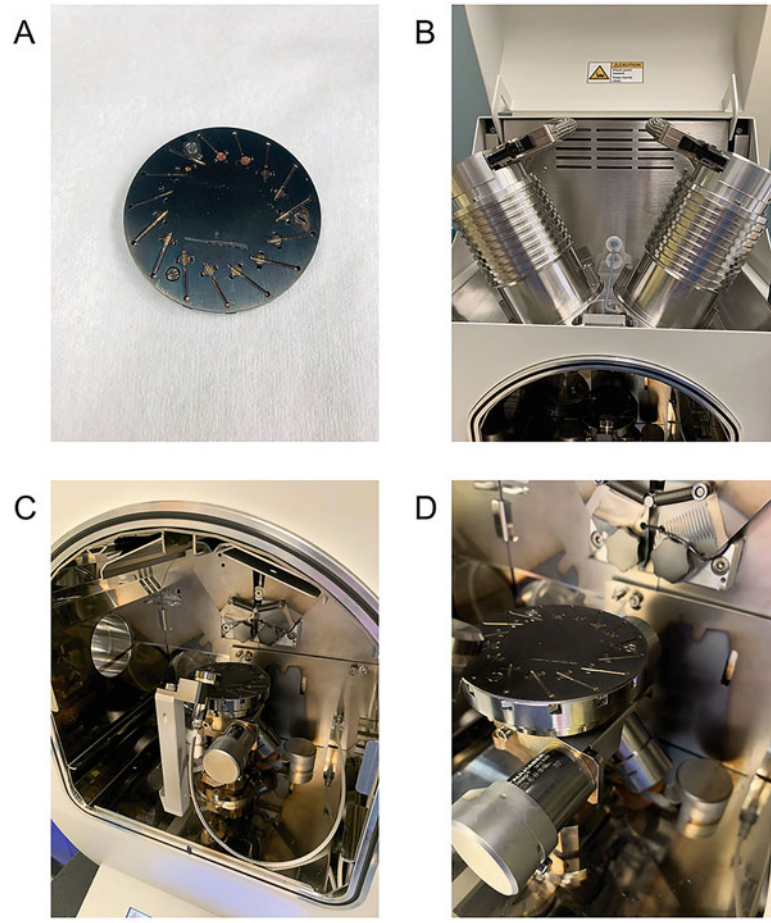


Fig. 4 Crucial components of the Leica EM ACE600 coater used in the rotary platinum shadowing. (**a**) Rotary stage capable of securing up to 15 grids. (**b**) Carbon gun (left) and essential platinum gun (right). (**c**, **d**) Positioning of the rotary stage inside the coater right alongside the quartz monitor, located directly in front of stage, responsible for calculating the specific amount of platinum output

3.6 Platinum Shadowing of Grids

1. Secure grids to a rotary stage of a high vacuum coater (Fig. 4a).
2. Obtain a vacuum pressure of approximately 4.5 × 10-6 mbar, and set shadowing at an angle of 3 degrees, with a platinum deposition of 3.5 nm. Grids should be rotated during shadowing to obtain even distribution of platinum.
3. Shadowed grids are now ready to be viewed by TEM and may be stored in an appropriate EM specimen holder at RT indefinitely.

3.7 Visualization of DNA Via Transmission Electron Microscope

1. Ideal TEM parameters for the visualization of DNA include a lanthanum boride (LaB6) filament necessary for maximum brightness and sharpness of the image, along with a preferably bottom mount high-resolution camera for high-quality images. Imaging of the DNA begins at approximately 10,000× with higher magnification images being visualized at 100,000×.
2. Due to the BAC spreading method, in conjunction with uranyl acetate staining and platinum shadowing, DNA will appear as a long darkened, cylindrical fiber in high contrast to the lighter carbon granular background. Double-stranded DNA molecules average a thickness of around 7 nm, whereas lighter, thinner single-stranded DNA average a thickness in the order of 2–3 nm [24]. Due to various circumstances, DNA staining or amount may be less than ideal. In these cases, spreading and shadowing can be repeated in order to obtain better quality grids. If success is not achieved, re-enrichment of the genomic DNA may be performed followed by new spreadings and shadowing.
3. Quantity of DNA on the grid is of great importance and depends on the concentration of the sample, size of the spreading surface, and quality of the carbon grid. Spreading of the DNA may need to be repeated in order to achieve an ideal amount of DNA on a grid. Certain parts of the same grid may be better than others. Optimal concentration of DNA on the grid should allow a good amount of empty space between each DNA molecule in order to clearly distinguish replication intermediates from random crossings of DNA fibers. Having an ideal concentration of DNA reduces the likelihood of crossover events, ensuring that any meeting of three or four-stranded structures represent a true junction.
4. Most of the DNA present in the sample is linear, and only a small percentage can be classified as a replication intermediate. Correct identification of these intermediates requires several established criteria. Replication forks are initially identified by visualizing a three-way junction or the meeting of three DNA fibers at one contact point (Fig. 5). These fibers represent a moment in time when the DNA is actively being replicated, one arm being the parental strand and the other two arms the daughter strands. If the identified three-way junction structure corresponds to a replication fork, the daughter strands should be of equal or similar length because the PvuII restriction enzyme will cut the two daughter strands in the same location (*see* Subheading 3.8 for more details). To confirm that the junction is a seamless joining of three strands, the junction needs also to be imaged at high magnifications of 100,000× or greater. ssDNA gaps may also be visualized at this point as

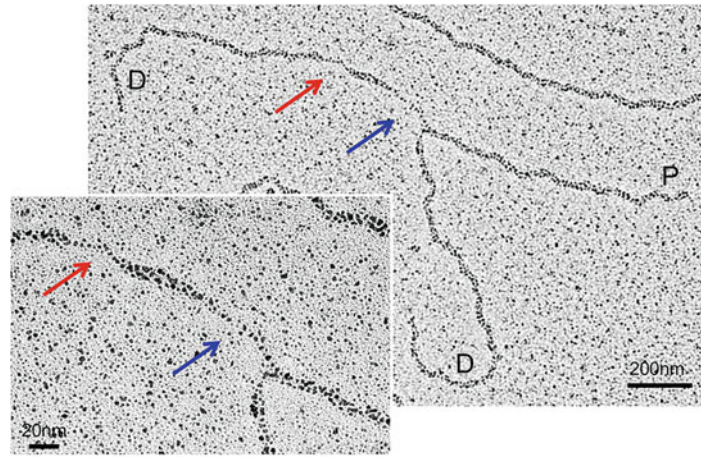


Fig. 5 Representative image of a replication fork containing two daughter strands (D) of equal length and a single parental strand (P). The red arrow indicates the ssDNA gap within the daughter strand, and the blue arrow is indicating a ssDNA gap at the junction

lighter, thinner sections located within the DNA fibers. These gaps are predominately found on the daughter strands as well as at the junction itself.

5. A smaller percentage of these replication intermediates can be classified as a reversed replication fork, characterized by a four-way junction. As with three-way junctions, reversed forks contain one parental strand and two daughter strands of similar length, along with an additional, typically shorter, fourth strand, indicative of the reversed arm. Particular attention is paid to the junction in order to confirm the presence of an open junction and rule out a possible crossover of DNA fragments. Often times the junction may be collapsed, and other indicators such as daughter strand symmetry, presence of single-stranded DNA at the junction, or the entire structure itself must be taken into account. For more details on the EM analysis of reversed fork structures, please *see* [20].
6. Approximately 80–120 images should be taken per sample. This allows for a small percentage of DNA molecules that may be discarded because of the lack of appropriate characteristics of a true replication intermediate upon later analysis. These images should be saved as .tiff files for the computational analysis performed with analytical software such as ImageJ. Standard statistical analysis using Microsoft Excel can be applied along with graphical representations of chosen parameters to visualize differences between samples.

3.8 Analysis of ssDNA Gaps at the Replication Fork

1. Once images are acquired, detailed analysis can commence using ImageJ and Microsoft Excel. Classification of a three-way junction, representing a replication fork, is made by looking at the symmetry of the daughter strands and quality of the junction. Symmetry is determined by measuring the lengths of each strand. At least two strands should be of equal length due to DNA digestion with the PvuII restriction enzyme. This symmetry is due to the same cut site being equidistant on both daughter strands. Certain circumstances may affect the symmetrical length of the daughter strands including incomplete digestion, DNA breakage due to handling, and responses to DNA damage. These instances should only affect a small percentage of three-way junctions. The meeting point at which the three strands join should be carefully analyzed using high magnification images. Visualization of the high magnification image allows to confirm the cohesive connection of all three strands. Some replication forks may contain areas of single-stranded DNA at the junction. ssDNA appears much lighter and thinner than the duplex strand. The presence of a single-stranded region at the junction is a further indicator of a replication fork.
2. Once a replication fork is established, careful inspection of the thickness of the daughter strand filaments allow for the identification of regions of single-stranded DNA. These ssDNA gaps can be pinpointed by monitoring changes in the filament thickness (double-stranded DNA molecules have an average thickness of approximately 7 nm, whereas single-stranded DNA has an average thickness of approximately 2–3 nm). Therefore, ssDNA gaps can be easily located as they appear thinner and lighter when compared to the surrounding double-stranded DNA. The presence or absence of these gaps, along with their quantity, can be statistically analyzed in Microsoft Excel. For example, the frequency of ssDNA gap accumulation can be calculated as a function of the specific treatment conditions or of any genetic modification made to the cell. The lengths of the gap may also be measured using ImageJ.

4 Notes

1. All growth media should be thoroughly removed via washing, so it does not absorb any of the monochromatic light during cross-linking.
2. For the cross-linking step, a 365 nm bulb should be used. The power should be set at 6.2 mW/cm^2 and confirmed with a UV meter. Several factors, including the sample distance from bulbs or bulb life, may interfere with this total irradiation power, so

irradiation cycle times may be adjusted to obtain ideal cross-links.

3. Care should be taken not to take up any of the cloudy white interphase where extracted proteins are located. Subsequent chloroform extractions might be performed in order to eliminate this unwanted middle layer. However, loss of DNA, located in the upper layer, may occur during transports to new tubes. Additional extractions should only be considered if it is too difficult to remove the top layer without taking up too much of the milky middle layer. Additional spin time may also be used to more clearly separate the two layers.
4. DNA is digested before enrichment in order to isolate the replication forks for easy viewing and identification in the field of vision of electron microscopy. The PvuII restriction enzyme is used because of its ideal cutting frequency of genomic DNA, resulting in DNA fragments in the order of 30–50 base pairs, allowing for high-resolution imaging of a single fork. In addition, PvuII cleavage leaves blunt ends, in contrast to ssDNA overhangs, and these blunt ends will not compete for binding to the BND cellulose.
5. It is important to remove any RNA contained in the sample before loading onto the BND cellulose, so that it does not interfere with the enrichment of the replication intermediate DNA.
6. Extensive washings of the BND cellulose must be performed in order to remove any fine particulate material from the BND suspension that may coat the grids and interfere with detection of the DNA molecules.
7. Salt adjustments are different for each restriction enzyme. The listed amounts are specific to PvuII. Refer to the restriction enzyme insert for respective adjustments.
8. If a specific quantity of DNA is desired, it is necessary to use an agarose gel to run a small amount of the acquired purified DNA alongside an appropriate marker. Quantification of DNA via NanoDrop is not recommended due to caffeine's interference with standard spectrometry readings.
9. In the chloroform/tape mixture, the chloroform will dissolve the sticky residue from the backing of the tape. When this mixture is applied to the grids, the sticky residue will adhere to the surface of the grids, while the chloroform will evaporate away. This provides a sticky base to which the carbon may adhere. If the correct tape is not used or too much tape is used in the chloroform solution, it can cause the grids to acquire an unwanted cloudy coating on top. If there is not enough sticky residue, the carbon will not properly adhere to the grids and slough away when spreading the DNA.

10. Spreading of the DNA onto the surface of the water is highly sensitive to air currents and vibrations in the environment. Therefore, it is recommended to perform this procedure in an isolated location or in an enclosed area, in order to avoid any disruptions to the DNA layer on top of the water.
11. Formamide is a partially denaturing reagent and helps to unfold the DNA molecules during spreading. This is essential to reveal a clear linear image of the dsDNA and subsequent replication intermediates. However, the DNA and formamide should remain separate until right before spreading, because too much exposure to the formamide could cause uncoupling of the two DNA strands.
12. It is crucial for the mica sheet to be freshly cleaved in order to provide a charged surface for the DNA to repel against, providing the necessary momentum into the water and across the surface. Because of this momentum, DNA will be concentrated at the edge of the graphite wall where the grid should be placed to pick up the highest amount.
13. Care should be taken to remove excess ethanol on the tweezers as it can upset the surface in which the DNA is located and decrease uptake on the grid.
14. It is recommended that no more than two grids be used per one spreading. The DNA spread on the surface will slowly sink to the bottom, and very little will remain to pick up on a third grid.
15. EM allows direct visualization of ssDNA gaps behind replication forks and can be combined with the modified DNA fiber assay utilizing the S1 nuclease to further characterize ssDNA gap accumulation throughout the genome (Figs. 5 and 6). For a general comparison of the EM and DNA fiber techniques, *see* [22]. Briefly, the DNA fiber assay starts with the incorporation of thymidine analogs into the replicating DNA, followed by spreading of the DNA onto charged microscope slides and staining with fluorescent antibodies in order to visualize the thymidine-labeled DNA tracts on a microscope [12–17]. The resolution of a DNA fiber experiment is typically limited to a few kilobases of the stretched DNA. The modified DNA fiber approach used for ssDNA gap detection takes advantage of the unique ssDNA cutting properties of the S1 nuclease. The S1 nuclease is added to the reaction after pulse labeling with the thymidine analogs to cleave the thymidine-labeled DNA containing ssDNA gaps. If gaps are present in the DNA, S1 cleavage will lead to shorter tract lengths relative to S1-untreated controls [19]. The S1 DNA fiber approach can detect ssDNA gaps as short as a few nucleotides or even nicks due to the ability of S1 nuclease to cleave substrates that contain single

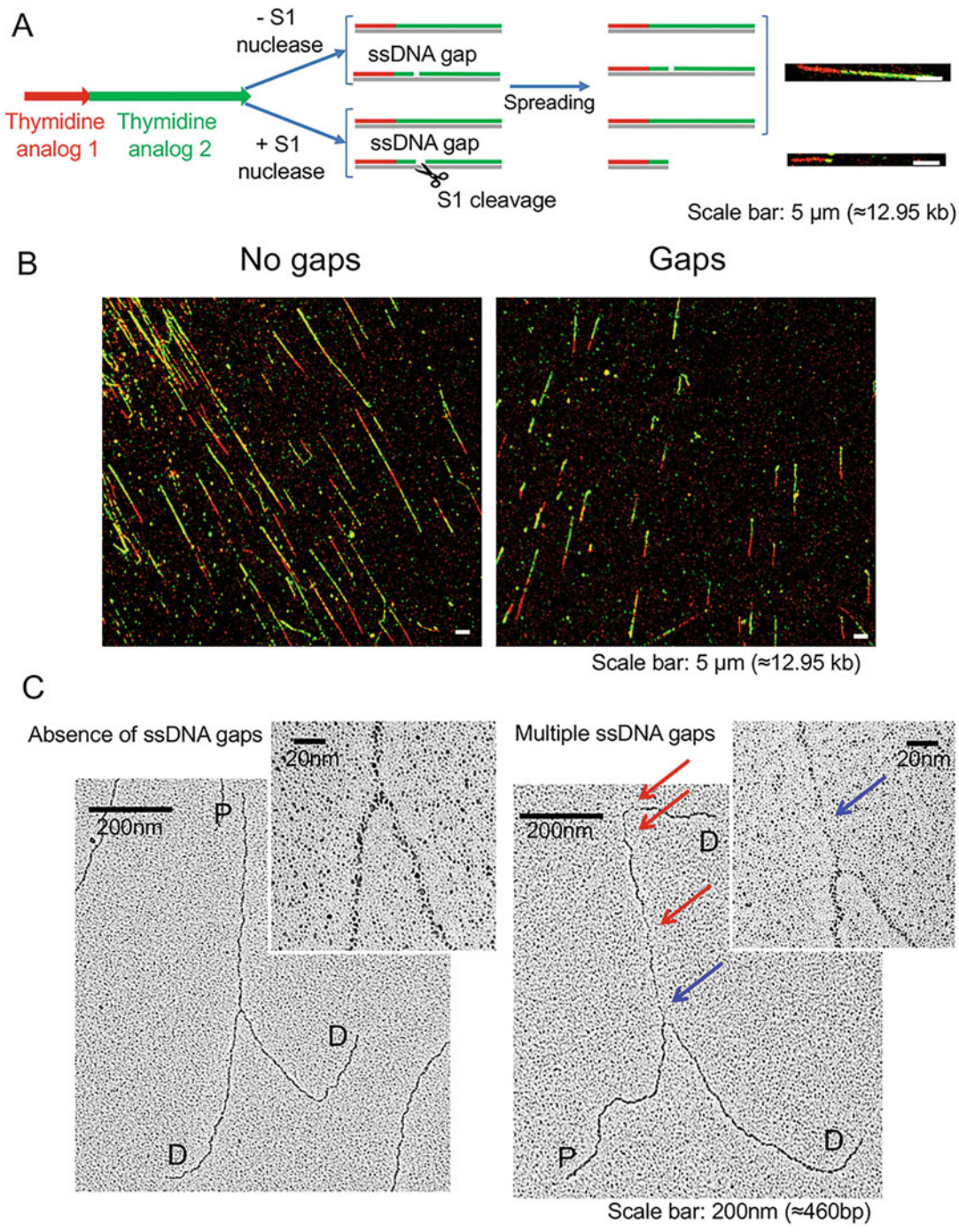


Fig. 6 (**a**) Schematic representation of the DNA fiber assay utilizing the S1 nuclease. (**b**) Fluorescent microscopy images of red and green labeled ongoing forks, using the S1 nuclease, resulting in either a longer tract length due to no gaps (left) or short tract lengths due to gaps (right). (**c**) Electron microscopy images detailing a replication fork that has no gaps (left) or multiple gaps, indicated by the red arrows, and a single gap at the junction indicated by the blue arrow (right)

nucleotide nicks [18]. The EM technique has a higher resolution compared to the DNA fiber approach. However, unlike the S1 DNA fiber approach, it cannot detect ssDNA gaps shorter than 40–60 nucleotides [25]. A major advantage of the EM technique is that it shows the actual location of the ssDNA gaps on the replication forks and it allows for the measurement of the size of the ssDNA gaps [26]. This information is extremely useful to determine whether ssDNA gaps are equally distributed on the two parental strands of a replication fork and whether the size of the gaps might change under different experimental conditions. Moreover, the EM technique allows quantification of the number of gaps on an individual fork and to distinguish between ssDNA gaps that are present at replication fork junctions versus ssDNA gaps located behind the junctions. These distinctions cannot be made using the DNA fiber approach. On the other hand, an important drawback of the EM technique is that the ssDNA gaps must be located within approximately 500 nucleotides from the fork junction in order to be detected due to the relatively short size of the DNA fragments analyzed by EM (1–2 kilobases). Conversely, the S1 DNA fiber approach can potentially detect ssDNA gaps that are several kilobases away from the fork junction given that a typical thymidine-labeling scheme allows the incorporation of the thymidine analog for 20–30 kilobases. Table 1 summarizes the advantages and disadvantages of the EM and DNA fiber techniques. Taken together, both these techniques complement each other and can each contribute to a more thorough detection and ultimately confirmation of the presence of ssDNA gaps in the genome.

Acknowledgments

We thank Alice Meroni and Annabel Quinet for their help with the preparation of the figures and for their comments on the manuscript. This work was supported by the NCI under grant numbers R01CA237263 and R01CA248526 and by the DOD BCRP Expansion Award BC191374 to A.V. This research was supported by the Alvin J. Siteman Cancer Center Siteman Investment Program (supported by The Foundation for Barnes-Jewish Hospital, Cancer Frontier Fund) and the Barnard Foundation to A.V. Figures were created with BioRender.com.

Table 1
Summary of the advantages and disadvantages of the electron microscopy and S1 nuclease DNA fiber assays

Assay	Advantages	Disadvantages
Electron Microscopy	▪ High resolution imaging of DNA as low as 40nt ▪ Actual visualization of the ssDNA gaps at the replication fork ▪ Quantification of the number of gaps per replication fork on one or both strands of DNA as well as length ▪ Able to pinpoint gaps on daughter strands specifically	▪ Gaps must be near the replication fork, within 0.5-1kb, due to restriction enzyme digestion of the DNA ▪ ssDNA gaps must be at least 40nt in length to accurately detect ▪ Time intensive ▪ Specialized equipment and reagents are required ▪ Use of an electron microscope can be costly and requires specific training
S1 Nuclease DNA Fiber Assay	▪ Single molecule resolution ▪ Recognition of gaps 1-3nt in length ▪ Results achieved in a relatively short time frame ▪ Less cells and reagents needed ▪ Access to only a fluorescent microscope	▪ Unknown amount of gaps present ▪ Inability to determine gap length ▪ Unable to identify newly synthesized strands

References

1. Diamant N, Hendel A, Vered I, Carell T, Reissner T, de Wind N, Geacinov N, Livneh Z (2012) DNA damage bypass operates in the S and G2 phases of the cell cycle and exhibits differential mutagenicity. Nucleic Acids Res 40(1):170–180. https://doi.org/10.1093/nar/gkr596
2. Elvers I, Johansson F, Groth P, Erixon K, Helleday T (2011) UV stalled replication forks restart by re-priming in human fibroblasts. Nucleic Acids Res 39(16):7049–7057. https://doi.org/10.1093/nar/gkr420
3. Jansen JG, Tsaalbi-Shtylik A, Hendriks G, Verspuy J, Gali H, Haracska L, de Wind N (2009) Mammalian polymerase zeta is essential for post-replication repair of UV-induced DNA lesions. DNA Repair (Amst) 8(12):1444–1451. https://doi.org/10.1016/j.dnarep.2009.09.006
4. Lehmann AR (1972) Post-replication repair of DNA in ultraviolet-irradiated mammalian cells. No gaps in DNA synthesized late after ultraviolet irradiation. Eur J Biochem 31(3):438–445. https://doi.org/10.1111/j.1432-1033.1972.tb02550.x
5. Lopes M, Foiani M, Sogo JM (2006) Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. Mol Cell 21(1):15–27. https://doi.org/10.1016/j.molcel.2005.11.015
6. Meneghini R (1976) Gaps in DNA synthesized by ultraviolet light-irradiated WI38 human cells. Biochim Biophys Acta 425(4):419–427. https://doi.org/10.1016/0005-2787(76)90006-x
7. Quinet A, Vessoni AT, Rocha CR, Gottifredi V, Biard D, Sarasin A, Menck CF, Stary A (2014) Gap-filling and bypass at the replication fork are both active mechanisms for tolerance of low-dose ultraviolet-induced DNA damage in the human genome. DNA Repair (Amst) 14:27–38. https://doi.org/10.1016/j.dnarep.2013.12.005
8. Bianchi J, Rudd SG, Jozwiakowski SK, Bailey LJ, Soura V, Taylor E, Stevanovic I, Green AJ, Stracker TH, Lindsay HD, Doherty AJ (2013)

PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. Mol Cell 52(4):566–573. https://doi.org/10.1016/j.molcel.2013.10.035

9. García-Gómez S, Reyes A, Martínez-Jiménez MI, Chocrón ES, Mourón S, Terrados G, Powell C, Salido E, Méndez J, Holt IJ, Blanco L (2013) PrimPol, an archaic primase/polymerase operating in human cells. Mol Cell 52(4):541–553. https://doi.org/10.1016/j.molcel.2013.09.025
10. Mourón S, Rodriguez-Acebes S, Martínez-Jiménez MI, García-Gómez S, Chocrón S, Blanco L, Méndez J (2013) Repriming of DNA synthesis at stalled replication forks by human PrimPol. Nat Struct Mol Biol 20(12):1383–1389. https://doi.org/10.1038/nsmb.2719
11. Wan L, Lou J, Xia Y, Su B, Liu T, Cui J, Sun Y, Lou H, Huang J (2013) hPrimpol1/CCDC111 is a human DNA primase-polymerase required for the maintenance of genome integrity. EMBO Rep 14(12):1104–1112. https://doi.org/10.1038/embor.2013.159
12. Bensimon A, Simon A, Chiffaudel A, Croquette V, Heslot F, Bensimon D (1994) Alignment and sensitive detection of DNA by a moving interface. Science 265(5181):2096–2098. https://doi.org/10.1126/science.7522347
13. Jackson DA, Pombo A (1998) Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. J Cell Biol 140(6):1285–1295. https://doi.org/10.1083/jcb.140.6.1285
14. Merrick CJ, Jackson D, Diffley JF (2004) Visualization of altered replication dynamics after DNA damage in human cells. J Biol Chem 279(19):20067–20075. https://doi.org/10.1074/jbc.M400022200
15. Michalet X, Ekong R, Fougerousse F, Rousseaux S, Schurra C, Hornigold N, van Slegtenhorst M, Wolfe J, Povey S, Beckmann JS, Bensimon A (1997) Dynamic molecular combing: stretching the whole human genome for high-resolution studies. Science 277(5331):1518–1523. https://doi.org/10.1126/science.277.5331.1518
16. Parra I, Windle B (1993) High resolution visual mapping of stretched DNA by fluorescent hybridization. Nat Genet 5(1):17–21. https://doi.org/10.1038/ng0993-17
17. Techer H, Koundrioukoff S, Azar D, Wilhelm T, Carignon S, Brison O, Debatisse M, Le Tallec B (2013) Replication dynamics: biases and robustness of DNA fiber analysis. J Mol Biol 425(23):4845–4855. https://doi.org/10.1016/j.jmb.2013.03.040
18. Vogt VM (1973) Purification and further properties of single-strand-specific nuclease from Aspergillus oryzae. Eur J Biochem 33(1):192–200. https://doi.org/10.1111/j.1432-1033.1973.tb02669.x
19. Quinet A, Carvajal-Maldonado D, Lemacon D, Vindigni A (2017) DNA fiber analysis: mind the gap! Methods Enzymol 591:55–82. https://doi.org/10.1016/bs.mie.2017.03.019
20. Neelsen KJ, Chaudhuri AR, Follonier C, Herrador R, Lopes M (2014) Visualization and interpretation of eukaryotic DNA replication intermediates in vivo by electron microscopy. Methods Mol Biol 1094:177–208. https://doi.org/10.1007/978-1-62703-706-8_15
21. Neelsen KJ, Lopes M (2015) Replication fork reversal in eukaryotes: from dead end to dynamic response. Nat Rev Mol Cell Biol 16(4):207–220. https://doi.org/10.1038/nrm3935
22. Vindigni A, Lopes M (2017) Combining electron microscopy with single molecule DNA fiber approaches to study DNA replication dynamics. Biophys Chem 225:3–9. https://doi.org/10.1016/j.bpc.2016.11.014
23. Lopes M (2009) Electron microscopy methods for studying in vivo DNA replication intermediates. Methods Mol Biol 521:605–631. https://doi.org/10.1007/978-1-60327-815-7_34
24. Vollenweider HJ, Sogo JM, Koller T (1975) A routine method for protein-free spreading of double- and single-stranded nucleic acid molecules. Proc Natl Acad Sci U S A 72(1):83–87. https://doi.org/10.1073/pnas.72.1.83
25. Mijic S, Zellweger R, Chappidi N, Berti M, Jacobs K, Mutreja K, Ursich S, Ray Chaudhuri A, Nussenzweig A, Janscak P, Lopes M (2017) Replication fork reversal triggers fork degradation in BRCA2-defective cells. Nat Commun 8(1):859. https://doi.org/10.1038/s41467-017-01164-5
26. Hashimoto Y, Ray Chaudhuri A, Lopes M, Costanzo V (2010) Rad51 protects nascent DNA from Mre11-dependent degradation and promotes continuous DNA synthesis. Nat Struct Mol Biol 17(11):1305–1311. https://doi.org/10.1038/nsmb.1927

Chapter 7

Approaches to Monitor Termination of DNA Replication Using *Xenopus* Egg Extracts

Tamar Kavlashvili and James M. Dewar

Abstract

DNA replication is crucial for cell viability and genome integrity. Despite its crucial role in genome duplication, the final stage of DNA replication, which is termed termination, is relatively unexplored. Our knowledge of termination is limited by cellular approaches to study DNA replication, which cannot readily detect termination. In contrast, the *Xenopus laevis* egg extract system allows for all of DNA replication to be readily detected. Here we describe the use of this system and assays to monitor replication termination.

Key words DNA replication, *Xenopus* egg extracts, DNA synthesis, Fork merger, Ligation, Decatenation, Chromatin capture

1 Introduction

Complete and accurate DNA replication is essential for cell viability and genome integrity. The final stages of DNA replication, which are collectively called "termination," are under-studied compared to the earlier stages, despite being equally important for cell viability and genome integrity [1]. This paucity of data arises because cellular approaches cannot readily detect termination events, although it is possible to inactivate key events during termination once the genetic requirements have been determined [2, 3]. While termination can be reconstituted using purified yeast proteins [4], termination in yeast and vertebrates differs in important ways [2–6], which necessitates alternate approaches. The *Xenopus laevis* egg extract system supports in vitro DNA replication using the full set of vertebrate proteins [7]. This approach allows vertebrate

The original version of this chapter was revised. The correction to this chapter is available at https://doi.org/10.1007/978-1-0716-2063-2_17

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_7,

termination to be monitored in detail on custom DNA templates and carefully manipulated [5, 6, 8–10]. In this chapter, we describe the use of *Xenopus laevis* egg extracts ("egg extracts") to replicate plasmid DNA and assays to monitor termination.

The egg extract system described here [7] involves sequential addition of two different extracts to plasmid DNA. Initially, plasmid DNA is incubated in a high-speed supernatant (HSS) of interphase cytosol, consisting of the soluble cellular proteins. HSS stimulates loading of the MCM2-7 complex onto DNA by ORC, CDC6, and CDT1 ("licensing") [11]. Once MCM2-7 complexes are loaded, they do not initiate replication because HSS contains low concentrations of the CDK and DDK kinases that are necessary for replication initiation. To initiate replication, reactions are supplemented with NucleoPlasmic Extract (NPE), which comprises highly purified nucleoplasm. NPE supplies a high concentration of CDK and DDK [12, 13], resulting in activation of replicative CMG helicase (CDC45-MCM2-7-GINS) and establishment of replication forks ("initiation"). Once replication forks are established, they duplicate the majority of the template ("elongation") before encountering each other head-on on the same stretch of DNA ("termination"). This process results in a single round of DNA replication, which is enforced by Cdt1 destruction at the onset of DNA synthesis [14], as is also the case in vivo [15].

Termination involves completion of DNA synthesis, separation of daughter molecules, and removal of replication proteins from the DNA [1]. Although many questions about termination remain, current data support the following model for termination in vertebrates. First, replication forks converge on the same stretch of DNA and advance towards each other (Fig. 1i, ii"fork convergence"). Replication forks do not slow or stall as they converge [8, 10] due to the activity of topoisomerase II, which removes pre-catenanes to prevent accumulation of topological stress that would otherwise cause converging forks to stall [6]. Once replication forks encounter each other, the CMG helicases pass each other on opposite DNA strands, which unwinds the final stretch of parental duplex (Fig. 1ii, iii "fork merger") [8, 10]. CMGs then translocate over the replicated lagging strands from the opposing fork, which allows for ligation of the daughter strands (Fig. 1iii, iv "ligation") [8, 10]. Topoisomerase II then removes any remaining DNA intertwines to unlink chromosomes (Fig. 1iv, v "decatenation") [8]. In parallel, the CMG helicases are ubiquitylated by the $CRL2^{Lrr1}$ ubiquitin ligase and extracted from DNA by the p97 ATPase (Fig. 1iv, v "unloading") [3, 9, 16] due to loss of contact with the lagging strand template [10].

In this chapter, we describe assays to monitor replication termination. Initially, we describe how to replicate plasmid DNA using HSS and NPE. We also describe how immunodepletion can be used to interrogate protein function during replication. We then explain how to assay for defects in initiation or elongation. Finally, we describe assays to monitor fork merger, ligation, decatenation,

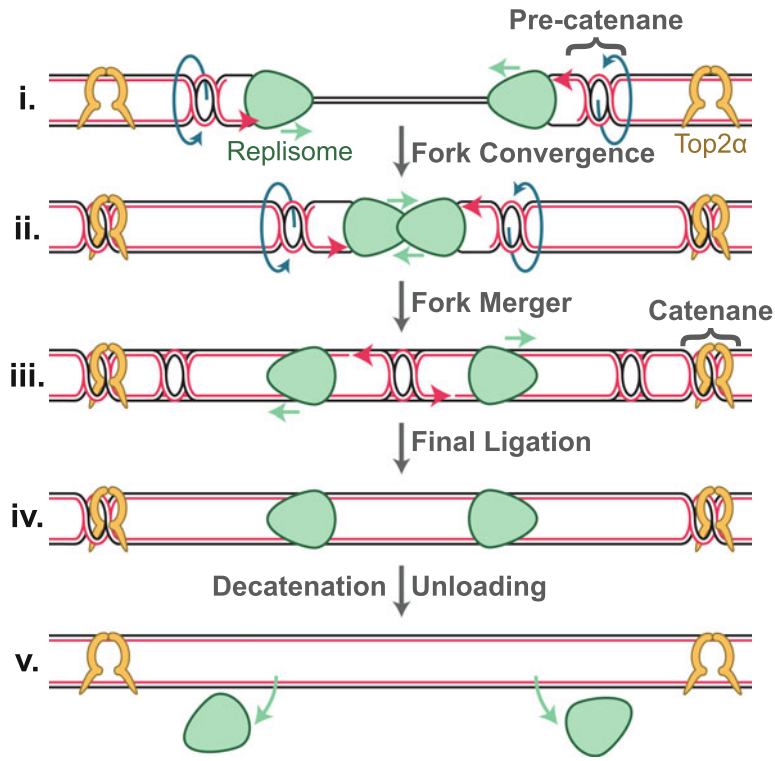


Fig. 1 Model for DNA replication termination in vertebrates. (**a**) (i) Termination begins when two replication forks emanating from adjacent origins converge upon the same stretch of DNA ("fork convergence"). (ii) When converging replication forks meet, the replisomes pass each other on opposite parental strands to unwind the final stretch of DNA ("fork merger") which allows the nascent strands to be extended past each other. (iii) Replisomes travel over duplex DNA from the opposing fork, which allows ligation of leading strands to downstream lagging strands ("final ligation"). (iv) In parallel, topoisomerase II removes any remaining catenanes ("decatenation"), and the replisomes are removed from DNA ("unloading")

and unloading with topoisomerase IIα-depleted extracts as an exemplar for defective termination [6]. We do not cover preparation of *Xenopus* egg extracts, which has been described elsewhere [17]. We also do not cover other *Xenopus* egg extract systems that can also provide valuable insights into DNA replication [18]. We also note that the *Xenopus* egg extract system is highly flexible and the approaches outlined in this chapter can readily be adapted or extended. For example, plasmid templates can be chemically modified to include DNA lesions so that replication-coupled repair processes can be examined [19–21]. Additionally, replication forks can be reversibly stalled using an array of *lac* repressor molecules, which can be used to synchronize and localize termination [8].

2 Materials

2.1 Immunodepletion of Xenopus Egg Extracts

1. Dynabeads Protein A (Invitrogen).
2. PBS (Phosphate-Buffered Saline) 10×: 1.37 M NaCl, 27 mM KCl, 101.4 mM Na_2HPO_4, 17.6 mM KH_2PO_4, pH 6.6.
3. PBS 1×: Dilute PBS 10× (*see* **item 2**) 1:10 with Milli-Q water.
4. PBST (PBS-Tween) 1×: 1× PBS supplemented with 0.01% Tween.
5. 5× ELB-sucrose: 5× ELB salts (from a 10× stock), 1.25 M sucrose (from Sucrose, for molecular biology, ≥99.5% (GC)). Make fresh once a month, and store 600 μL aliquots at +4 °C.
6. 20% Tween.
7. 1× ELB: Dilute 5× ELB-sucrose (*see* **item 5**) 1:5 with Milli-Q water.
8. 1× ELB + Salts: 1× ELB supplemented with 0.5 M NaCl.
9. 1× ELBT: 1× ELB supplemented with 0.01% Tween.
10. 1× ELBT + Salts: 1× ELBT + Salts supplemented with 0.01% Tween.
11. Magnetic rack for microcentrifuge tubes.
12. Low protein binding microcentrifuge tubes (Costar).

2.2 Replication of Plasmid DNA in Xenopus Egg Extracts

1. [α-P 32]-dATP (3000 Ci/mmol).
2. 1 M PC (phosphocreatine disodium salt): 10 mM potassium phosphate, pH 7.0; store 50 μL aliquots at −20 °C.
3. 0.2 M ATP (adenosine 5′-triphosphate disodium salt hydrate): dissolve in sterile water; adjust the pH to 7.0 with 10 M NaOH. Store 50 μL aliquots at −20 °C.
4. 5 mg/mL CPK (creatine phosphokinase): 50 mM NaCl, 50% glycerol, and 10 mM HEPES-KOH, pH 7.5. Store 250 μL aliquots at −20 °C. These aliquots are stable for 2–6 months.
5. ARS (ATP-regenerating system): Combine 10 μL 1 M PC (*see* **item 2**), 5 μL 0.2 M ATP (*see* **item 3**), and 0.5 μL 5 mg/mL CPK (*see* **item 4**) immediately before use. Store on ice.
6. 1 M DTT (dithiothreitol): dissolve in sterile Milli-Q water; store 20 μL aliquots at −20 °C.
7. 10× ELB salts: 25 mM $MgCl_2$, 500 mM KCl, 100 mM HEPES-KOH, pH 7.7. Filter-sterilize and store at 4 °C.
8. 5× ELB: 5× ELB salts (from a 10x stock), 1.25 M sucrose (from Sucrose, for molecular biology, ≥99.5% (GC)). Make fresh once a month, and store 600 μL aliquots at +4 °C.
9. 0.5 mg/mL nocodazole: Dissolve in DMSO; store 50 μL aliquots at −20 °C.

10. 300 ng/μL DNA plasmid in 10 mM Tris–HCL (pH 8.0) (*see* **Note 1**).
11. High-speed supernatant (HSS) [6].
12. NucleoPlasmic Extract (NPE) [6].
13. 0.5 mL Safe-Lock tubes.
14. Extraction Stop: 0.5% SDS, 25 mM EDTA pH 8.0 in 50 mM Tris–HCl, pH 7.5.
15. 20 mg/mL Proteinase K: Dissolve in Milli-Q water; store 50 μL aliquots at −20 °C.
16. 2 mg/mL RNase.
17. Phosphorimager such as the GE Typhoon Imager.

2.3 Purification of DNA Replication Intermediates

1. 10 mM Tris–HCl, pH 8.0.
2. Phenol/chloroform/isoamyl alcohol (25:24:1).
3. Chloroform/isoamyl alcohol 24:1 (v/v).
4. 3 M sodium acetate, pH 5.5. Dissolve in Milli-Q water, and adjust pH with glacial acetic acid.
5. 20 mg/mL glycogen from mussels. Store in 20–50 μL aliquots at −20 °C.
6. Ice-cold 100% ethanol.
7. Ice-cold 70% ethanol.
8. Clear 0.6 mL Tubes (Axygen).
9. epTIPS (Eppendorf).

2.4 Analysis of DNA Synthesis Using Native Agarose Gels

1. SDS DNA loading buffer (6×): 15% Ficoll-400, 66.5 mM EDTA, 20 mM Tris–HCl pH 8.0, 0.1% SDS, 0.09% Bromophenol blue.
2. TBE (10×): 0.89 M Tris Base, 0.89 M boric acid, 25.5 mM EDTA, pH 8.0.
3. Gel migration buffer: 1× TBE (*see* **item 2**).
4. GTG Agarose (SeaKem).
5. High voltage power supply.
6. Whatman paper.
7. Amersham Hybond XL Roll.
8. Paper towels.
9. Gel dryer attached to a vacuum pump.

2.5 Analysis of Fork Merger Using Native Agarose Gels

1. XmnI restriction enzyme, 20,000 units/mL.
2. CutSmart Buffer 10×.
3. *See* **items 1–9** in Subheading 2.4.

2.6 Analysis of Ligation Using Denaturing Agarose Gels

1. AlwnI restriction enzyme, 20,000 units/mL.
2. CutSmart Buffer 10.
3. Alkaline loading buffer (6×): 18.5% Ficoll, 6.19 mM EDTA, 0.26% xylene cyanol, 0.16% bromophenol green.
4. Alkaline buffer (10×): 500 mM NaOH, 10 mM EDTA.
5. TCA 7% (w/v).
6. *See* **items 4–9** in Subheading 2.4.

2.7 Analysis of Replication Protein Binding by Chromatin Capture

1. Biotinylated *lac* repressor (LacR) [8].
2. Streptavidin magnetic beads.
3. Low protein binding microcentrifuge tubes (Costar).
4. Bovine serum albumin (BSA) protein.
5. 1 M Tris–HCl, pH 8.0.
6. 5 M NaCl.
7. 0.5 M EDTA.
8. Tween (20%).
9. ELB-sucrose: 5X ELB salts (from a 10× stock), 1.25 M sucrose (from Sucrose, for molecular biology, ≥99.5% (GC)). Make fresh once a month, and store 600 µL aliquots at +4 °C.
10. 10× ELB salts: 25 mM $MgCl_2$, 500 mM KCl, 100 mM HEPES-KOH, pH 7.7. Filter-sterilize and store at 4 °C.
11. Binding buffer: 50 mM Tris–HCl (pH 8.0), 1 mM EDTA, 0.02% Tween.
12. ELB-BT: 2× ELB, 0.5 mg/mL BSA, 0.04% Tween.
13. ELB Salts/Tween: 2× ELB salts, 0.25 mg/mL BSA, 0.03% Tween.
14. 1 M DTT.
15. Protein sample buffer (2.2×): 6% SDS, 125 mM Tris–HCl (pH 6.8), 20% glycerol, 0.01% bromophenol blue.
16. Protein sample buffer (1×): 3% SDS, 62.5 mM Tris–HCl (pH 6.8), 10% glycerol, 0.005% bromophenol blue, 50 mM DTT.
17. Magnetic rack for microcentrifuge tubes.

3 Methods

3.1 Immunodepletion of Xenopus Egg Extracts

1. Aliquot 40 µL Dynabeads Protein A in low protein binding tubes.
2. Wash beads 4× with 400 µL 1× PBST.

3. Resuspend in 40 μL 1× PBST. Transfer 24 μL to a new tube for the immunodepletion ("depletion") and 12 μL to a new tube for the control ("mock") (*see* **Note 2**).
4. Remove supernatant, and add 12 μg antibodies for the depletion and 12 μg antibodies for the mock (*see* **Note 3**).
5. Resuspend beads thoroughly by pipetting, and incubate overnight at 4 °C with rotation.
6. Wash beads 2× with 1× PBST.
7. Wash beads 2×with 1× ELB-Tween + Salts.
8. Wash beads 1× with 1× ELB-Tween.
9. Transfer beads to new tubes. Resuspend depletion beads to 20 μL using 1× ELB, and mock beads to 10 μL using 1× ELB.
10. Aliquot 2 × 5 μL mock beads for mock depletion of NPE and 2 × 5 μL depletion beads for depletion of NPE. Additionally aliquot 2 × 5 μL depletion beads for HSS. Place all beads on ice (*see* **Note 4**).
11. Activate HSS (*see* Subheading 3.2).
12. Deplete 10 μL HSS with antibody bound beads. To do this, remove the ELB supernatant and then add 10 μL HSS to the beads, and mix by gently pipetting (*see* **Note 5**). Rotate at room temperature (RT) for 20 min.
13. Place the depleted HSS in a magnetic rack to pellet the beads. Transfer the supernatant to a new tube to perform an additional round of depletion, as in **step 12**.
14. Activate NPE (*see* Subheading 3.2) 5 min after initiating the second round of HSS depletion in **step 13**.
15. After NPE is activated, deplete 10 μL NPE using mock beads and 10 μL NPE using depletion beads, as in **step 12**.
16. After the second round of HSS depletion is complete, recover the HSS, as in **step 13**, and then prepare a licensing mix as in Subheading 3.2, **step 5**.
17. After the first round of NPE depletion is complete, the depleted NPE and perform an additional round of NPE depletion, as in **step 13**.
18. After the second round of NPE depletion is complete, transfer NPE to new tubes, and save 1 μL of depleted NPE to validate the extent of depletion by Western blotting (*see* **Note 6**).
19. Initiate replication (*see* Subheading 3.2).

3.2 Replication of Plasmid DNA in Xenopus *Egg Extracts*

1. Assemble ARS (*see* Subheading 2.2), and store on ice.
2. Thaw a 33 μL aliquot of HSS, and then add 1 μL ARS and 0.2 μL nocodazole (*see* **Note 7**).
3. Centrifuge HSS at 16,000 × *g* for 5 min at RT.

4. Transfer 24 μL of activated HSS to a new tube, and pipette up and down 8–10 times to mix.
5. Add 19 μL of activated HSS to 1 μL of plasmid DNA (300 ng/μL) ("licensing mix"). Supplement with 0.3–1 μL of [α-P32]-dATP (depending on activity) and mix thoroughly. Incubate for 30 min prior to addition to NPE (*see* **Note 8**).
6. At $T = 8$–12 min, thaw a tube of 20 μL NPE. Activate by addition of 0.6 μL ARS, 0.858 μL DTT, and 4 μL ELB (5×). Add water to bring the mix to 40 μL, and then pipette up and down eight to ten times to mix.
7. At $T = 30$ min, initiate replication by mixing 2 volumes of diluted NPE with 1 volume of licensing mix (*see* **Note 9**).
8. Sample the reaction at desired timepoints by adding 2 μL of the replication reaction to 40 μL Extraction Stop.
9. Add 1.9 μL RNAse to the sample, and incubate at 37 °C for >30 min.
10. Add 1.83 μL of proteinase K to the sample, and incubate for 1 h at 37 °C or overnight at RT.

3.3 Purification of DNA Replication Intermediates

1. Transfer 10 μL of each sample (from Subheading 3.2 **step 10**) to new tubes to be analyzed directly (*see* Subheading 3.4, below). Dilute the remainder of each sample to 110 μL with Extraction Stop, and add 110 μL Tris 10 mM, pH 8.0. Mix by inversion.
2. Add 220 μL phenol/chloroform/isoamyl alcohol (25:24:1), and mix by carefully inverting tubes 10–15 times (*see* **Note 10**).
3. Centrifuge samples at 16,000 × *g* for 4 min at RT.
4. Transfer 180 μL of the aqueous phase to Safe-Lock tubes containing 180 μL chloroform/isoamyl alcohol (24:1). Invert tubes 10–15 times to mix. Centrifuge at 16,000 × *g* for 4 min at RT.
5. During centrifugation, prepare microcentrifuge tubes (Axygen) each containing 14 μL NaOAc (3 M) and 1 μL glycogen, one per sample (*see* **Note 11**).
6. After centrifugation, carefully remove 140 μL of the upper aqueous phase, and add it to the NaOAc and glycogen. Pipette up and down three times to mix.
7. Place samples on ice, and add 2.5 volumes of ice-cold 100% EtOH. Mix tubes thoroughly by inversion (*see* **Note 12**).
8. Incubate samples on ice for 15 min, and then centrifuge at 24,000 × *g* for >30 min at 4 °C.
9. Remove all tubes from the centrifuge, and keep upright on ice. Aspirate all but ~50 μL of the supernatant from each tube.

10. Add 400 μL of ice-cold 70% EtOH, and centrifuge at 24,000 × *g* for 5–10 min at 4 °C.
11. Aspirate all but ~50 μL of the supernatant from each tube again, and then briefly (>1 min) centrifuge at 24,000 × *g* at 4 °C.
12. Using an epTIPS pipette tip, aspirate all traces of supernatant from around the glycogen pellet. Cap the tubes, and return them to ice until all supernatants are aspirated.
13. Once ethanol is aspirated from all pellets, uncap the tubes, and put them on a rack at RT. Monitor each tube, and note when the pellet turns translucent. Wait ~4 min after this, and then add 10 mM Tris, pH 8.0 (1 μL per 5–10 ng DNA).
14. Resuspend pellets at RT for >30 min. After 15 min and 30 min, flick the tube and centrifuge briefly. Store at 4 °C for up to 1 week. Store at −20 °C for longer periods.

3.4 Separation of Replication Intermediates Using Native Agarose Gels

1. Add 2 μL DNA SDS loading dye (6×) to 10 μL sample (from Subheading 3.3 **step 1**, above). Briefly centrifuge, and flick the tubes to mix.
2. Prepare a solution of agarose 1% (w/v) and water by microwaving thoroughly until all agarose is melted (*see* **Note 13**).
3. Stir constantly while the mixture cools.
4. Once the gel has cooled to ~55 °C, pour the gel into the casting tray.
5. After the gel has cooled and solidified, submerge the gel in an electrophoresis tank containing 1× TBE.
6. Load approximately 1.5 μL of each sample per mm of lane width on a 1% agarose gel, and perform electrophoresis at 5 V/cm (*see* **Note 14**).
7. After the dye front has migrated ~10 cm, stop the electrophoresis, and then cut the gel ~1 cm above the dye front. Discard the bottom half of the gel in a radioactive waste container as this contains most of the unincorporated radionucleotides.
8. Place the gel face down on plastic wrap. Cover it with a piece of Hybond and 2 pieces of Whatman paper cut to the same size of the agarose gel.
9. Place a 5″ stack of paper towels on top of the gel, and apply weight of ~1 kg for >1 h (*see* **Note 15**).
10. Remove the paper towels and the top Whatman paper, and then carefully flip the gel and replace the top Whatman paper, now in direct contact with the gel. Place the gel in this same orientation, with the Hybond facing down, in a gel dryer attached to a vacuum pump. The gel dryer should be heated to 80 °C to increase the speed of drying.

11. After the gel is fully dried, remove the bottom Whatman paper, wrap the gel in plastic wrap, and visualize the gel by phosphorimaging (*see* **Note 16**).

3.5 Analysis of DNA Synthesis and Decatenation Using Native Agarose Gels

1. The intermediates observed during unperturbed replication are depicted in Fig. 2a, and Fig. 2b shows an autoradiogram of replication intermediates from mock- and topoisomerase IIα (Top2α)-depleted extracts. In mock-depleted extracts, the replication fork-containing structures (Fig. 2a, θ) form a distinct band (Fig. 2b, θ) and are then rapidly resolved to nicked circular monomers (nCMs) and supercoiled circular monomers (NCMs), which exist as an equilibrium. In Top2α-depleted extracts, θs persist because replication forks stall during termination. Fork merger eventually occurs independently of Top2α activity, resulting in highly catenated catenanes (Cats+) that are slowly converted to less-catenated species (Cats-) by residual Top2 activity [6].
2. To measure total DNA synthesis, total incorporation of radioactive dATP is determined. Open the image in ImageJ, and draw a rectangle around each lane to measure the mean signal ($\bar{x}$[lane]) and total area (A[lane]). Measure the mean background signal for each lane ($\bar{x}$[bg]) from an area of the gel not containing replication intermediates. Total signal is calculated as ($\bar{x}$[lane] − $\bar{x}$[bg]) × (A[lane]). Plot total signal in each lane over max against time (Fig. 2c) (*see* **Note 17**). In the example shown, Top2α depletion has a negligible impact on DNA synthesis, which shows that neither initiation nor elongation was appreciably impacted. If a defect in DNA synthesis was detected, it would suggest a defect in either initiation, elongation, or both, and additional assays would have to be performed to address this [22, 23].
3. To measure decatenation, the percentage of scCMs is determined. Draw a rectangle around the scCMs in each lane to measure the mean signal ($\bar{x}$[CMs]) and total area (A[CMs]). Total scCM signal is calculated as ($\bar{x}$[CMs] − $\bar{x}$[bg]) × (A[CMs]). The percentage of scCMs is calculated as [scCM signal]/[total lane signal] × 100% (Fig. 2d). In the example shown, Top2α depletion completely blocks decatenation (Fig. 2d) despite having negligible impact on DNA synthesis (Fig. 2c) [6].

3.6 Preparation of Replication Fork Structures

1. Digest 1 μL purified replication intermediates with one unit of XmnI enzyme and CutSmart Buffer in a reaction volume of 10 μL.
2. Incubate the reaction at 37 °C for 1 h.

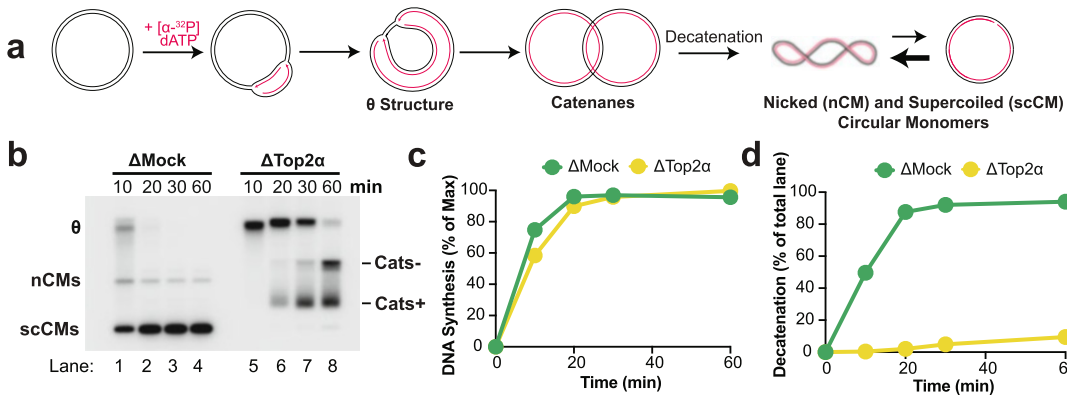


Fig. 2 Analysis of DNA synthesis and decatenation in *Xenopus* egg extracts. (**a**) Schematic of plasmid DNA replication intermediates and products. (**b**) Plasmid DNA was replicated in mock- and Top2α-immunodepleted extracts. Products were separated on an agarose gel and visualized by autoradiography. (**c**) Quantification of total DNA synthesis from (b). (**d**) Quantification of decatenation from (b)

3.7 Separation of Replication Fork Structures Using a Native Agarose Gel

1. Add 2 μL DNA SDS loading dye (6×) to 10 μL sample (from Subheading 3.6 **step 2**, above). Briefly centrifuge and flick the tubes to mix.
2. *See* **steps 2**–7 Subheading 3.4.

3.8 Analysis of Fork Merger Using Native Agarose Gels

1. The digested DNA structures that result from unperturbed replication are depicted in Fig. 3a, and Fig. 3b shows an autoradiogram of DNA structures from mock- and topoisomerase IIα(Top2α)-depleted extracts. In mock-depleted extracts, the replication fork-containing structures migrate as high molecular weight double-Ys (Fig. 3a, DYs) that are converted to much smaller linear molecules (Fig. 3b, lins) once the remaining parental duplex is unwound. In Top2α-depleted extracts, DYs persist, and accumulation of linears is delayed because replication forks stall during termination. Linears eventually accumulate due to fork merger independent of Top2 activity.
2. To measure fork merger, the percentage of linears is determined. Draw a rectangle around the lins each lane to measure the mean signal ($\bar{x}$[Lins]) and total area (A[Lins]). *See* Subheading 3.5 for measurement of background and total lane signal. Total linear signal is calculated as ($\bar{x}$[Lins] − $\bar{x}$[bg]) × (A[Lins]). The percentage of linears is calculated as [linear signal]/[total lane signal] × 100% (Fig. 3c). In the example shown, Top2α depletion delays fork merger (Fig. 3c) [6].

3.9 Preparation of Nascent Strands

1. Digest 1 μL of the purified replication intermediates with one unit of AlwnI enzyme in a reaction volume of 10 μL.
2. Incubate the reaction at 37 °C for 1 h.

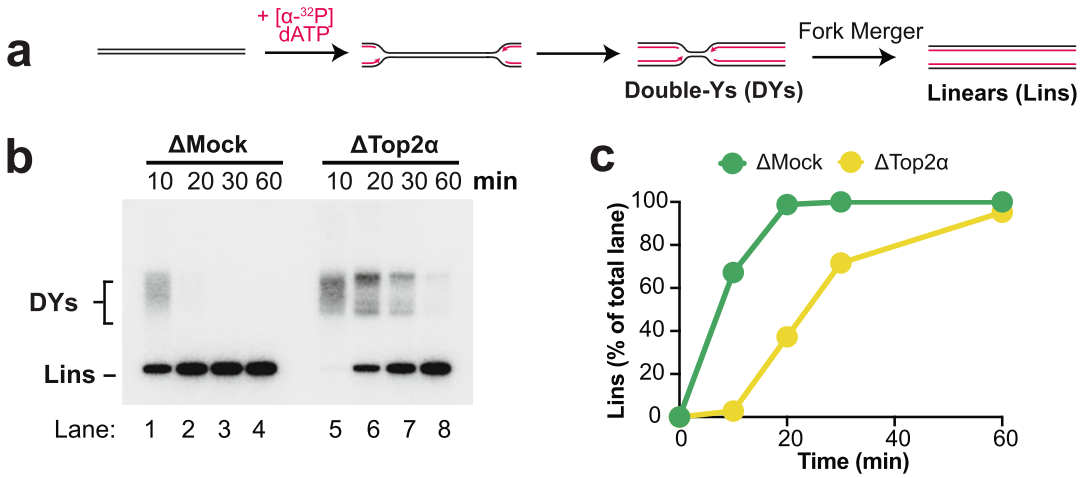


Fig. 3 Analysis of fork merger in *Xenopus* egg extracts. (**a**) Schematic of replication intermediates (DYs) and products (Lins) formed by restriction digest of plasmid intermediates. (**b**) Samples from Fig. 1b were digested with XmnI and then separated on an agarose gel and visualized by autoradiography. (**c**) Quantification of fork merger from (b)

3. Briefly centrifuge the samples, and then add 1 μL of 330 mM EDTA to each sample, and mix thoroughly (*see* **Note 18**).

3.10 Separation of Nascent Strands Using a Denaturing Agarose Gel

1. Add 2.2 μL (0.2 volumes) alkaline loading buffer to each sample and mix. Samples are ready for loading on an alkaline gel.
2. Prepare a solution of agarose 1.67% (w/v) and water by microwaving thoroughly until all agarose is melted (*see* **Note 13**).
3. Stir constantly while the mixture cools.
4. Once the gel has cooled to ~55 °C, add 1/ninth volumes of alkaline buffer (10×) to yield a 1.5% agarose gel in 1× alkaline buffer.
5. Continue stirring the gel with a stir bar for another minute, and then slowly pour the gel into the casting tray.
6. After the gel has cooled and solidified, submerge the gel in an electrophoresis tank containing 1× alkaline buffer (*see* **Note 19**).
7. Load approximately 1.5 μL of each sample per mm of lane, and perform electrophoresis at 1.5 V/cm until dye front has migrated ~15 cm (*see* **Note 20**).
8. Neutralize the gel with 7% TCA for 30 min (*see* **Note 21**).
9. Rinse the gel with Milli-Q water, and trim the edges to ensure it lays flat.

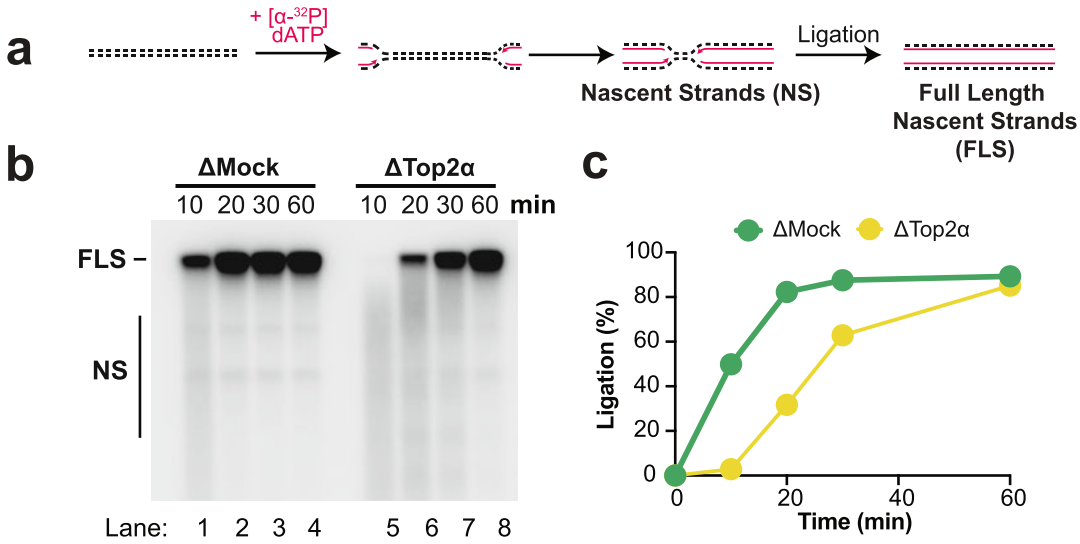


Fig. 4 Analysis of ligation in *Xenopus* egg extracts. (**a**) Schematic of nascent strands (NS) and full-length strands (FLS) formed by digestion and denaturation of plasmid intermediates. (**b**) Samples from Fig. 1b were digested with AlwNI and then separated on a denaturing agarose gel and visualized by autoradiography. (**c**) Quantification of ligation from (b)

10. Place the gel face down on plastic wrap. Cover it with a piece of Hybond and 2 pieces of Whatman paper cut to the exact size of the agarose gel.
11. Place 5″ stack of paper towels on top of the gel, and apply weight of ~1 kg for >2 h.
12. Remove the paper towels and the top Whatman paper, and then carefully flip the gel and replace the top Whatman paper, now in direct contact with the gel. Place the gel in this same orientation, with the Hybond facing down, in a gel dryer. The gel dryer should be heated to 60 °C to increase the speed of drying. Once the gel is almost completely dry, the temperature should be increased to 80 °C for 10 min to completely dry the gel (*see* **Note 22**).
13. After the gel is fully dried, remove the bottom Whatman paper, wrap the gel in plastic wrap, and visualize the gel by phosphorimaging.

3.11 Analysis of Ligation

1. The nascent strands that result from unperturbed replication are depicted in Fig. 4a, and Fig. 4b shows an autoradiogram of DNA structures from mock- and topoisomerase IIα(Top2α)-depleted extracts. In mock-depleted extracts, the nascent strands migrate as a smear (Fig. 4b, NS) that are converted to the much larger full-length strands (Fig. 4b, FLS) once the daughter strands are ligated. In Top2α-depleted extracts, NS persist and accumulation of FLS is delayed because replication

forks stall during termination. FLS eventually accumulate due to fork merger independent of Top2 activity.

2. To measure ligation, the percentage of full length strands is determined. Draw a rectangle around the FLS each lane to measure the mean signal ($\bar{x}$[FLS]) and total area (A[FLS]). *See* Subheading 3.5 for measurement of background and total lane signal. Total full length strands signal is calculated as ($\bar{x}$[FLS] − $\bar{x}$[bg]) × (A[FLS]). The percentage of full length strands is calculated as [full length strands signal]/[total lane signal] × 100% (Fig. 3c). In the example shown, Top2α depletion delays ligation (Fig. 3c) [6].

3.12 Recovery of Chromatin-Associated Proteins

1. Prepare binding buffer, ELB-BT, and ELB salts. Store ELB salts on ice.
2. Aliquot 22.5 μL streptavidin Dynabeads in a Costar microcentrifuge tube.
3. Wash beads 3× with binding buffer using a magnetic rack.
4. Resuspend beads in 135 μL binding buffer.
5. Add 22.5 pmol of LacR, and mix by vigorously inverting tube four to six times.
6. Incubate Dynabeads and LacR for 40 min at RT with end-over-end rotation.
7. Wash beads 2× with 400 μL binding buffer and 2× with 400 μL ELB-BT.
8. Resuspend beads in 112.5 μL ELB-BT, and aliquot 8 × 12.5 μL tubes on ice.
9. Replicate plasmid DNA according to Subheadings 3.1 and 3.2.
10. Sample 3 μL of reaction in ice-cold Dynabeads + LacR. Mix by vigorously pipetting up and down 40 times, and then store on ice until the end of the experiment.
11. In the cold room, resuspend all beads by pipetting, and then incubate for 30 min with end-over-end rotation.
12. After 30 min, place the samples on a magnetic rack until beads form a pellet against the side of the tube, and remove all supernatant.
13. Wash each tube 2× with 400 μL ELB salts/Tween (*see* **Note 23**).
14. After the second wash, remove supernatant, and centrifuge samples at 400 × *g* for 1 min.
15. Place tubes on a magnetic rack, and remove the last droplets of the supernatant by pipetting.
16. Add 25 μL protein sample buffer (1×) to each tube, bring them to RT, and vortex until beads are resuspended (*see* **Note 24**).

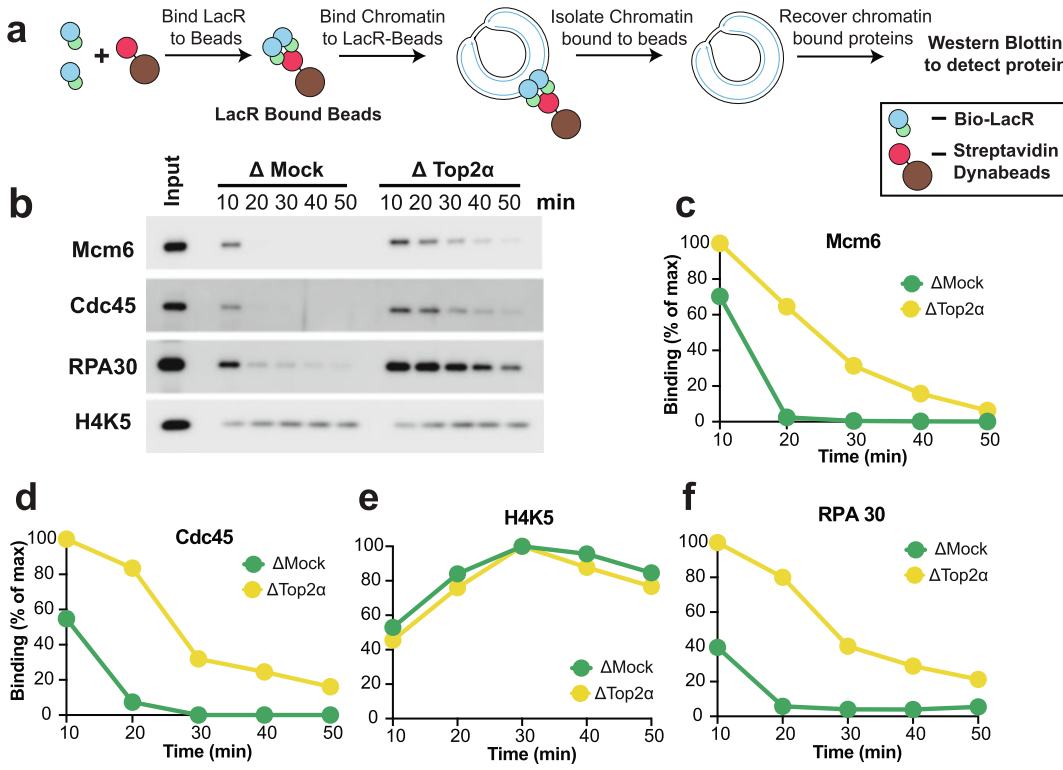


Fig. 5 Analysis of replisome unloading in *Xenopus* egg extracts. (**a**) Schematic of chromatin capture approach. (**b**) Replisome proteins were visualized by Western blotting. (**c–f**) Quantification of protein binding from (b)

17. Incubate tubes in a heat block at 95 °C for 5 min.
18. Allow tubes to cool to RT.
19. Centrifuge briefly, vortex, and then centrifuge once more to bring all droplets to the bottom.
20. Place tubes in a magnetic rack to pellet the beads, and remove supernatant to Costar microcentrifuge tubes.
21. Samples are now ready for analysis by Western blotting. Alternatively, store samples at −20 °C (*see* **Note 25**).

3.13 Detection of Chromatin Capture Samples by Western Blotting

1. Separate ~5 μL of sample per well using standard SDS-PAGE migration conditions.
2. To measure replisome unloading by plasmid pulldowns, perform Western blotting using antibodies that recognize CMG components, such as Cdc45 and Mcm6 (*see* **Note 26**).

3.14 Analysis of Replisome Unloading

1. The chromatin capture approach is outlined in Fig. 5a. Figure 4b shows Western blots of chromatin-bound MCM6, CDC45, RPA30, and H4K5 from mock- and topoisomerase IIα(Top2α)-depleted extracts. In mock-depleted extracts,

MCM6, CDC45, and RPA30 readily dissociated at the end of replication. In Top2α-depleted extracts, binding persists because replication forks stall during termination.

2. To measure replisome unloading, the relative abundance of replication proteins is determined. Draw a rectangle around the band in each lane to measure the mean protein signal ($\bar{x}$[PROT]) and total area (A[PROT]). *See* Subheading 3.5 for measurement of background. Total protein signal is calculated as ($\bar{x}$[PROT] − $\bar{x}$[bg]) × (A[PROT]). The data is normalized to peak signal for each protein max[PROT] by multiplying total protein signal by 100/(max[PROT]). In the example shown, Top2α depletion delays dissociation of MCM6, CDC45, and RPA30 (Fig. 3c–e) [6]. Dissociation of CDC45 and MCM6 serves as a measure of replisome unloading. Note that inhibition of replisome unloading only would not impact RPA dissociation but would cause retention of MCM6 and CDC45 [3, 9]. H4K5 is present at similar levels and serves as a loading control (Fig. 3f).

4 Notes

1. Any 3–5 kilobase pair plasmid can be used for replication assays, and, in principle, any restriction enzymes that cut only once can be used in the place of XmnI and AlwNI. XmnI and AlwNI are used in this chapter because they have been most extensively employed and validated.
2. Each time immunodepletion is performed against the antibody of interest, parallel set of extracts should be depleted with mock IgG control. To confirm specificity of the immunodepletion, the effect should ideally be either rescued by addition of purified recombinant protein or reproduced with a second independent antibody.
3. Concentrations >1 mg/mL may interfere with binding. Antibodies more concentrated than this should be diluted prior to adding to beads.
4. Typically, depleted HSS is used for all reactions to help ensure consistency of replication kinetics between conditions. Depleted HSS does not typically impact the mock condition due to the high abundance of target protein present in the mock-depleted NPE, but this should be verified for each antibody.
5. Allow each tube of beads to warm to RT for at least 5 min before addition of extract.
6. To validate the efficiency depletion, depleted extracts should be analyzed alongside undepleted extracts by Western blotting.

7. Thaw HSS by placing in RT water for 15–20 s. Add ARS and nocodazole to the sides of the HSS tube, and mix in rapidly by repeatedly pipetting HSS against the side of the tube.
8. The volume of the licensing mix can be scaled up or down as needed.
9. Licensing mix and NPE should each comprise 1/3 of the final reaction. Typically, this is achieved by diluting the NPE to 50% and mixing 2:1 with licensing mix. If additional reagents need to be added (e.g., small molecule inhibitors), then the NPE should be diluted less to account for the volume of additional reagents.
10. Phenol/chloroform/isoamyl alcohol must be warmed to RT to ensure SDS in the Extraction Stop does not precipitate. If SDS precipitates are formed, then large white pellets that do not readily resuspend in aqueous buffer will be obtained during the final step of DNA purification.
11. Make a master mix of NaOAC 3 M and glycogen by combining the appropriate amount of each reagent. Ensure proper mixing by vigorously pipetting up and down five to ten times and then vortexing for 30–60 s. Improper mixing will result in inconsistent recovery of DNA.
12. Ensure that each sample is mixed thoroughly by inverting vigorously. Do this by slightly shaking each tube during inversion to ensure no droplets are trapped by surface tension at either end. Improper mixing will result in inconsistent recovery of DNA.
13. When microwaving agarose in water, evaporation will occur. Weigh the solution before and after microwaving to ensure that gel percentage is accurate.
14. V/cm is the standard unit for electrophoresis. This number should be multiplied by the distance (cm) between electrodes of the electrophoresis tank to determine the appropriate voltage for electrophoresis.
15. Hard-back catalogs from vendors of lab supplies work best. Two can usually be balanced on top of the stack of paper towels.
16. When the gel has fully dried in a gel dryer, the bottom Whatman paper (the one contacting the membrane) will curl up and "pop" off. This is a sign that the gel is ready to be imaged.
17. Background counts are usually taken from underneath the smallest band but can also be taken from above the largest band.

18. EDTA is added to chelate all the Mg^{2+} which would otherwise precipitate out in the alkaline sample/migration buffer and disrupt migration of DNA out of the wells.
19. Best results are typically achieved if the denaturing gel sets for ~2 h prior to electrophoresis.
20. Much of the dye will diffuse out of the gel during long periods of electrophoresis. To increase the visibility of the dye front, load undiluted loading dye in one of the outer lanes. Alternatively, place a glass plate on top of the gel during the migration to prevent the dye diffusing out.
21. To determine how much TCA is needed, remove a small volume of migration buffer (e.g., 5 mL), and empirically determine the amount of 7% TCA needed to reach pH 6.5. This can readily be achieved using pH strips. Scale up the amount of TCA according to the volume of the gel.
22. Denaturing gels are sensitive to high heat and can only tolerate 80 °C for short amounts of time (~10–15 min).
23. After each set of beads is resuspended, place all tubes back on the rack, and wait until 1 min after all tubes have formed a stable pellet before proceeding with the next wash step.
24. Protein sample buffer (1×) must be kept at room temperature because SDS will precipitate out at 4 °C. A simple way to do this is to transport it into the cold room using a "cool rack" that has been warmed to RT.
25. Each time samples are thawed, they should be heated at 70 °C for 1 min to ensure all SDS redissolves. Samples can typically be freeze-thawed two to three times without impacting protein stability, but this must be empirically determined and varies based on the protein.
26. To monitor replisome unloading,any of the core replisome proteins can be analyzed as the entire replisome progression complex is coordinately unloaded during termination [3, 9].

Acknowledgments

JMD is supported by NIH grant R35GM128696.

References

1. Dewar JM, Walter JC (2017) Mechanisms of DNA replication termination. Nat Rev Mol Cell Biol 18(8):507–516. https://doi.org/10.1038/nrm.2017.42
2. Maric M, Maculins T, De Piccoli G, Labib K (2014) Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. Science (New York, NY) 346(6208):1253596. https://doi.org/10.1126/science.1253596
3. Sonneville R, Moreno SP, Knebel A, Johnson C, Hastie CJ, Gartner A, Gambus A,

Labib K (2017) CUL-2(LRR-1) and UBXN-3 drive replisome disassembly during DNA replication termination and mitosis. Nat Cell Biol 19(5):468–479. https://doi.org/10.1038/ncb3500

4. Deegan TD, Baxter J, Ortiz Bazan MA, Yeeles JTP, Labib KPM (2019) Pif1-family helicases support fork convergence during DNA replication termination in eukaryotes. Mol Cell 74(2):231–244.e9. https://doi.org/10.1016/j.molcel.2019.01.040
5. Deng L, Wu RA, Sonneville R, Kochenova OV, Labib K, Pellman D, Walter JC (2019) Mitotic CDK promotes replisome disassembly, fork breakage, and complex DNA rearrangements. Mol Cell 73(5):915–929.e6. https://doi.org/10.1016/j.molcel.2018.12.021
6. Heintzman DR, Campos LV, Byl JAW, Osheroff N, Dewar JM (2019) Topoisomerase II is crucial for fork convergence during vertebrate replication termination. Cell Rep 29(2):422–436.e425. https://doi.org/10.1016/j.celrep.2019.08.097
7. Walter J, Sun L, Newport J (1998) Regulated chromosomal DNA replication in the absence of a nucleus. Mol Cell 1(4):519–529
8. Dewar JM, Budzowska M, Walter JC (2015) The mechanism of DNA replication termination in vertebrates. Nature 525(7569):345–350. https://doi.org/10.1038/nature14887
9. Dewar JM, Low E, Mann M, Raschle M, Walter JC (2017) CRL2(Lrr1) promotes unloading of the vertebrate replisome from chromatin during replication termination. Genes Dev 31(3):275–290. https://doi.org/10.1101/gad.291799.116
10. Low E, Chistol G, Zaher MS, Kochenova OV, Walter JC (2020) The DNA replication fork suppresses CMG unloading from chromatin before termination. Genes Dev 34(21–22):1534–1545. https://doi.org/10.1101/gad.339739.120
11. Blow JJ (2001) Control of chromosomal DNA replication in the early Xenopus embryo. EMBO J 20(13):3293–3297. https://doi.org/10.1093/emboj/20.13.3293
12. Prokhorova TA, Mowrer K, Gilbert CH, Walter JC (2003) DNA replication of mitotic chromatin in Xenopus egg extracts. Proc Natl Acad Sci U S A 100(23):13241–13246. https://doi.org/10.1073/pnas.2336104100
13. Walter JC (2000) Evidence for sequential action of cdc7 and cdk2 protein kinases during initiation of DNA replication in Xenopus egg extracts. J Biol Chem 275(50):39773–39778. https://doi.org/10.1074/jbc.M008107200
14. Arias EE, Walter JC (2005) Replication-dependent destruction of Cdt1 limits DNA replication to a single round per cell cycle in Xenopus egg extracts. Genes Dev 19(1):114–126. https://doi.org/10.1101/gad.1255805
15. Pozo PN, Cook JG (2016) Regulation and function of Cdt1; a key factor in cell proliferation and genome stability. Genes (Basel) 8(1):2. https://doi.org/10.3390/genes8010002
16. Moreno SP, Bailey R, Campion N, Herron S, Gambus A (2014) Polyubiquitylation drives replisome disassembly at the termination of DNA replication. Science (New York, NY) 346(6208):477–481. https://doi.org/10.1126/science.1253585
17. Lebofsky R, Takahashi T, Walter JC (2009) DNA replication in nucleus-free Xenopus egg extracts. Methods Mol Biol 521:229–252. https://doi.org/10.1007/978-1-60327-815-7_13
18. Gillespie PJ, Gambus A, Blow JJ (2012) Preparation and use of Xenopus egg extracts to study DNA replication and chromatin associated proteins. Methods 57(2):203–213. https://doi.org/10.1016/j.ymeth.2012.03.029
19. Duxin JP, Dewar JM, Yardimci H, Walter JC (2014) Repair of a DNA-protein crosslink by replication-coupled proteolysis. Cell 159(2):346–357. https://doi.org/10.1016/j.cell.2014.09.024
20. Räschle M, Knipscheer P, Enoiu M, Angelov T, Sun J, Griffith JD, Ellenberger TE, Schärer OD, Walter JC (2008) Mechanism of replication-coupled DNA interstrand crosslink repair. Cell 134(6):969–980. https://doi.org/10.1016/j.cell.2008.08.030
21. Vrtis KB, Dewar JM, Chistol G, Wu RA, Graham TGW, Walter JC (2021) Single-strand DNA breaks cause replisome disassembly. Mol Cell 81(6):1309–1318.e1306. https://doi.org/10.1016/j.molcel.2020.12.039
22. Walter J, Newport J (2000) Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. Mol Cell 5(4):617–627
23. Yardimci H, Loveland AB, Habuchi S, van Oijen AM, Walter JC (2010) Uncoupling of sister replisomes during eukaryotic DNA replication. Mol Cell 40(5):834–840. https://doi.org/10.1016/j.molcel.2010.11.027

Chapter 8

Use of High-Performance Liquid Chromatography-Mass Spectrometry (HPLC-MS) to Quantify Modified Nucleosides

Rebecca Rodell, Ning Tsao, Adit Ganguly, and Nima Mosammaparast

Abstract

Physiological and chemically induced modifications to nucleosides are common in both DNA and RNA. Physiological forms of these modifications play critical roles in gene expression, yet aberrant marks, if left unrepaired, may be associated with increased genome instability. Due to the low prevalence of these marks in most samples of interest, a highly sensitive method is needed for their detection and quantitation. High-performance liquid chromatography, coupled to mass spectrometry (HPLC-MS), provides this high degree of sensitivity while also being adaptable to nearly any modified nucleoside of interest and still maintaining exquisite specificity. In this chapter, we demonstrate how to use HPLC-MS to analyze the catalytic activity of a nucleic acid demethylase, to quantify the prevalence of N^6-methyladenosine from RNA, and to determine the kinetics of alkylation damage repair.

Key words Mass spectrometry, Nucleoside, Methylation, AlkB, Epitranscriptomics

1 Introduction

Physiological modifications on nucleic acids, such as methylation, play a critical role in regulating gene expression and are tightly regulated by complex networks of proteins. These proteins perform three functions: they catalyze the addition of the modifications, they trigger downstream processes by reading the modifications, and they catalyze the removal of modifications—i.e., "writers," "readers," and "erasers," respectively. Dysregulation of these highly regulated gene regulation networks can lead to multiple diseases, including cancer [1]. Methylation marks are deposited on nucleic acids using S-adenosylmethionine (SAM)-dependent methyltransferases, and these modifications on DNA and RNA are important components of epigenetic and epitranscriptomic gene regulation [2]. A well-characterized methylation adduct associated with CpG islands in DNA is 5-methylcytosine—deposited by writers Dnmt1, Dnmt3A, and Dnmt3B [3]. The demethylation process occurs

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_8,

through TET dioxygenase-mediated oxidation of 5-methylcytosine, followed by replication-dependent dilution or removal of the oxidized adduct through base excision repair [4]. In mRNA, one such physiological modification is N^6-methyladenosine, which modulates mRNA translation and degradation and is deposited by the METTL3-METTL14 methyltransferase complex [5]. In turn, N^6-methyladenosine is reversed by erasers FTO [6] and ALKBH5 [7]—which are both human AlkB homologs. While the physiological functions of methylation modifications on mRNA are still being clarified, corresponding physiological tRNA methylation has been well characterized. Post-transcriptional modifications found in cytoplasmic tRNAs affect their folding, stability, and identity while regulating protein translation and cell growth [8].

In contrast to physiological modifications that are highly regulated, aberrant methylation damage induction can occur via treatment with chemical alkylating agents or by off-target effects of endogenous methyltransferases [9]. Alkylating agents are found naturally in low concentrations in the environment, while synthetic alkylators are commonly used anti-cancer drugs [10]. Methylmethane sulfonate (MMS) and *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine (MNNG) are examples of alkylating agents that primarily induce 7-methylguanosine and 3-methyladenine adducts while also catalyzing minor lesions such as 1-methyladenine and 3-methylcytosine on nucleic acids [10]. Similarly, the endogenous methyltransferase TRMT6/61A deposits 1-methyladenine within a GUUCRA tRNA-like motif [11], while another methyltransferase METTL8 is thought to catalyze 3-methylcytosine deposition on mRNA [12, 13]. 1-Methyladenine and 3-methylcytosine (Fig. 1) aberrant modifications on nucleic acids can be reversed by the AlkB family of dioxygenases in an oxygen, alpha-ketoglutarate and Fe (II)-dependent manner [14]. The *E. coli* AlkB protein is a well-studied oxidative dealkylation DNA repair enzyme, which has nine mammalian homologs, of which two—ALKBH2 and ALKBH3—have been shown to have DNA repair activity [15]. ALKBH3 catalyzes reversal of 1-methyladenine and 3-methylcytosine marks in ssDNA and RNA [16], and this catalytic activity is important for alkylation damage resistance in some cancer cell lines [17].

In order to determine the presence of these modifications on nucleic acids, methods such as 2D thin layer chromatography [6, 18] and dot blot analysis [6, 9, 17] have been used in the past. While these methods do provide a read on the presence or absence of the modification, they are at most semi-quantitative, not allowing for precise determination of the concentration of modified nucleotide present. Additionally, these methods tend to be dependent upon antibody recognition, which makes it difficult to quickly develop the ability to detect and read newly recognized modifications, in addition to the normal challenges associated with antibodies, such as specificity. In contrast, HPLC-MS allows for precise

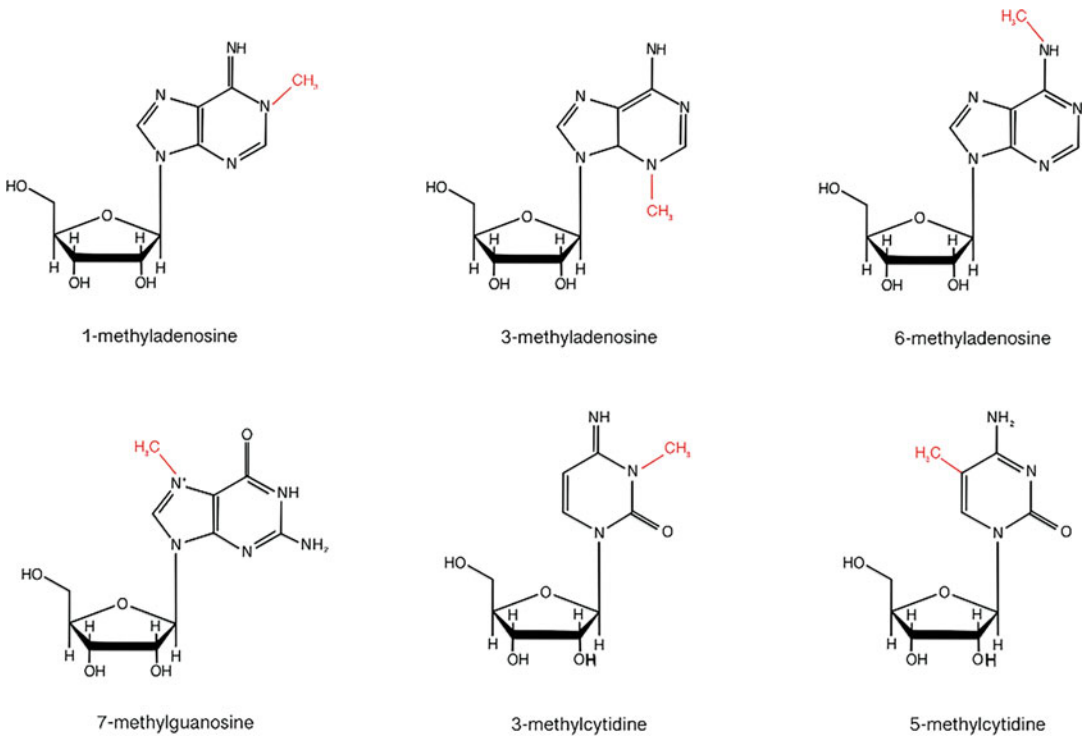


Fig. 1 Selected physiological and chemically induced base methylation adducts. The modifications are shown in the context of a ribonucleoside

quantification of the concentration of a lesion by comparing the spectral counts of a sample to pre-defined standard material. The high sensitivity of HPLC-MS allows for detection of compounds on the picogram to femtogram level. Furthermore, the HPLC-MS is also easily adaptable to quantify any lesion, given the availability of a pure standard of the modified nucleoside, allowing for substantial versatility in the application of the method. Additionally, one sample can be analyzed for multiple nucleosides in one run. This multiplexing capability can greatly increase efficiency of analysis for multiple different nucleosides, including modified and unmodified counterparts, in one sample.

Here, we describe a set of methods to quantify a prototypical methylated nucleoside, which can be used to assess its removal from a defined substrate by an AlkB demethylase in vitro, or quantify it in a biological sample. For the latter, the nucleic acid of interest is isolated from a sample using traditional nucleic acid purification methods. If needed, additional cell fractionation may be included in the purification protocol if the nucleic acid of interest is located in a specific subcellular compartment, such as mitochondria [19]. Once the nucleic acid is isolated, it is then digested down to individual nucleosides, as nucleotides do not behave well in a QQQ-type mass spectrometer. For this purpose, a combination of a nuclease and a

phosphatase is used. Thereafter, the sample is injected into the HPLC-MS system, where it is run in a solvent under high pressure through the HPLC column. Compounds then interact with the stationary phase of the column, eluting from the column at different time points, known as retention time, depending on their biochemical properties. The type of column and type of solvent impact the retention time of the compound; adjusting these measures adjusts the time at which the compound leaves the column. This allows for the possibility of manipulating the system to provide greater clarity between lesions which are difficult to differentiate by alternate methods. Subsequently, compounds are ionized and analyzed through a triple quadrupole mass spectrometer. Multiple reaction monitoring (MRM) is used, in which precursor ions of a specified mass to charge ratio (m/z) are selected in the first quadrupole and fragmented in the second quadrupole. The fragmented ions are once again selected through the third quadrupole, before finally reaching the ion detector where the quantity of the predefined product ion is recorded. The unique signature resulting from the combination of precursor and product ion allows for precise identification of the compound in question. When the data on the quantity of compounds detected is compared to a linear standard curve for the same compound, it is possible to precisely quantify the amount of a nucleoside present in a sample.

2 Materials

We describe three specific applications of nucleoside quantification using HPLC-MS. The first is quantitative demethylation of a 1-methyladenosine (m1A) containing DNA oligonucleotide using a prototypical human AlkB enzyme, ALKBH3 [15]. The second is the quantification of 6-methyladenosine (m6A) from total RNA isolated from a human cell line. Finally, we use the HPLC-MS approach to quantify the kinetics of 1-methyladenosine (1medA) loss from the genomic DNA of cultured cells, demonstrating the broad applicability of the approach. While the focus of this chapter is on specific methylated marks (see examples in Fig. 1), the same basic steps can be used for any modified nucleoside, with the only requirement being that a purified standard be available.

2.1 Setup of HPLC-MS/MS and Optimization of Nucleoside Standards

Every nucleoside has unique properties that determine the optimal collision energy and fragmentation values used by the mass spectrometer for its identification and quantification (Table 1). In our laboratory we utilize an Agilent LC-MS/MS system, which utilizes the Optimizer program for this purpose. Following this optimization, a standard curve can be created which in turn will be used for downstream quantification in experimental samples.

Table 1
Sample nucleoside compounds and associated values used for MS/MS

Compound	Precursor ion	Product ion	Fragment	Collision energy
2-Deoxyadenosine	252.11	136	149	20
2-Deoxycytosine	228.1	112	81	8
Adenosine	268.1	136	75	18
Cytosine	244.1	112	75	14
Guanosine	284.2	152	75	16
1-Methyl-2-deoxyadensine	266.13	150	91	16
N3-Methyl-2-deoxyadensine	242.1	109	81	48
	242.1	126.1	81	12
1-Methyladenosine	282	150	75	16
3-Methylcytosine	258	128	81	12
7-Methylguanosine	298	166	75	10

2.1.1 HPLC-MS/MS

1. Agilent 1290 Infinity II UHPLC system.
2. Agilent 6470 Triple Quadrupole MS.
3. MassHunter Optimizer software.
4. ZORBAX RRHD Eclipse Plus C18 2.1 × 50 mm (1.8 μm) column.
5. LC-MS grade methanol plus 0.1% formic acid (*see* **Note 1**).
6. LC-MS grade water plus 0.1% formic acid.
7. Agilent Optimizer Software Quick Start Guide.

2.1.2 Validating with Standard Curve

1. Modified nucleoside, such as N^6-methyladenosine, or other standard of interest.
2. Unmodified counterpart of modified nucleoside (i.e., adenosine).
3. LC-MS grade water.

2.2 ALKBH3-Mediated Demethylation Assay

The catalytic activity of any putative demethylase can be measured through a demethylation assay using a methylated oligonucleotide as a substrate. For this example, we will be demonstrating human ALKBH3 demethylation activity on a methylated RNA substrate.

2.2.1 Demethylation Assay

1. 0.8 M HEPES pH 7.5.
2. 100 mM ascorbate.
3. 100 mM 2-oxoglutarate.
4. 100 mM $FeSO_4$.

5. LC-MS grade water.
6. Recombinant 6X-His tagged ALKBH3. For the purification of ALKBH3 from bacteria, the reader is referred to previously published studies for additional details [17].
7. 1-Methyladenosine (m1A) containing RNA oligonucleotide.

2.2.2 Nucleic Acid Digestion

1. Nuclease S1 from *Aspergillus oryzae* (Sigma cat #N5661).
2. FastAP Thermosensitive Alkaline Phosphatase (Thermo Scientific cat #EF6051).
3. LC-MS grade water.
4. Incubator at 37 °C.
5. Incubator at 95 °C.
6. 11 mm Plastic Crimp/Snap Top Autosampler Vials (Thermo Scientific cat #C4011-13).
7. 11 mm Autosampler Snap-It Caps (Thermo Scientific cat #C4011-50).
8. 0.22 μm Millex syringe filter (Sigma cat #SLGVR04NK).
9. 10 mL syringe.

2.2.3 Quantitative Data Analysis

1. Agilent Quantitative Analysis software.

2.3 Quantification of N^6-Methyladenosine from RNA

A common epitranscriptomic modification in RNA is N^6-methyladenosine (m^6A). This modification can be quantified relative to its unmodified counterpart through LC-MS. After RNA purification, the subsequent steps utilize the same approach of sample digestion and HPLC-MS analysis.

2.3.1 RNA Extraction

1. Commercially available RNA purification kit. We use the miRNeasy Mini Kit (Qiagen #217004).
2. 1.7 mL microcentrifuge tubes.
3. LC-MS grade water.

2.4 Analysis of DNA Methylation Adduct Repair in Cultured Cells

Cells treated with alkylating agents generate various types of methylation adducts in nucleic acids including N^1-methyldeoxyadenosine (1medA). Here, we provide a method to analyze this lesion in genomic DNAs from methyl methanesulfonate (MMS)-treated cells. This approach allows the measurement of repair kinetics by monitoring the level of the lesion after recovery from acute MMS exposure. This modification can be quantified relative to its unmodified counterpart (dA) through HPLC-MS.

2.4.1 Induction of Methylation Damage

1. Adherent cell line of choice. We routinely use HeLa or other easily cultured cells.

2. 100 mm petri dishes.
3. CO_2 tissue culture incubator at 37 °C.
4. Methyl methanesulfonate (MMS) (Sigma #129925).
5. DMEM supplemented with 10% fetal bovine serum and penicillin/streptomycin, referred to as "complete medium" in the Methods sections that follow.
6. 1× PBS.

2.4.2 DNA Extraction and Purification

1. 1× PBS.
2. Cell lifter (Celltreat #229305).
3. 1.7 mL microcentrifuge tubes.
4. Centrifuge at 4 °C.
5. Commercially available genomic DNA purification kit. We use the DNeasy Blood & Tissue Kit (Qiagen #69506).

2.4.3 DNA Digestion

1. Nucleoside Digestion Mix (NEB # M0649S).
2. LC-MS grade water.
3. Incubator at 37 °C.
4. 11 mm Plastic Crimp/Snap Top Autosampler Vials.
5. 11 mm Autosampler Snap-It Caps.
6. 0.22 μm Millex syringe filter.
7. 10 mL syringe.

3 Methods

3.1 Optimization of a Standard

3.1.1 Using Optimizer

1. Prepare a nucleoside standard with a concentration of 1 ng/μL (*see* **Note 2**).
2. Select a suitable mobile phase with a 5–75% gradient, such as water and methanol (*see* **Note 1**). For the purpose of optimizing m^6A, we used 98% water and 2% methanol transitioning over 3 min to 92% water and 8% methanol, followed by 4 min of 2% water and 98% methanol.
3. Acquire data on your standard in MS2 Scan mode, injecting 10 μL of sample at a time. Check for the presence of your precursor ion of interest, and confirm by running the standard in MS2 SIM mode. Save the MS2 SIM method.
4. Using the Agilent Optimizer Software Quick Start Guide for assistance, create a Project in Optimizer using the MS2 SIM method created in **step 3**.
5. Run Optimizer. Once finished, it will provide you with a report of conditions for each target ion generated. Select target ions

with the highest prevalence for each compound; use these values for QQQ settings.

3.1.2 Validating with Standard Curve

1. Create a standard curve of the nucleoside consisting of five to six concentrations ranging from 2 nM to 500 nM.
2. Use the LC parameters and QQQ settings from the optimization of the standard above to create a new run method. *See* Table 1 for sample compounds and correspondent values. Run the standard curve using this method.
3. Open Agilent Quantitative Analysis software, and create a "New Batch."
4. Open all data files from run into this batch.
5. Under Methods, select New Method, and then Create Method from Acquired MRM Data. Select a single standard curve sample with a mid-range concentration value. Select *Create Levels from Calibration Samples*. Change CF Weight to 1/x. Exit the method creator, indicating that the method should be applied to all samples in the batch.
6. Check the created standard curve to ensure it is linear and has an R^2 value greater than 0.98 (see examples in Fig. 2). A very low concentration of a standard may be below the limit of linear quantification and may need to be removed for this analysis; if so, right click on the data point to remove it from the standard curve calculations, and then press "Analyze Batch" with the new standard curve.

3.2 ALKBH3-Mediated Demethylation Assay

3.2.1 Demethylation Assay

1. In one tube, make fresh 10× reaction buffer by combining HEPES pH 7.5 (to a final concentration of 500 mM), ascorbate (to a final concentration of 20 mM), 2-oxoglutarate (final concentration of 1 mM), and $FeSO_4$ (final concentration of 0.40 mM).
2. Add 5 μL of the 10× reaction buffer, methylated oligonucleotide (final concentration of 2 μM), and 0.1 μg of purified, recombinant ALKBH3. Bring to 50 μL using LC-MS grade water on ice.
3. Incubate the reaction at 37 ° C for the time required for the time-course assay (e.g., 0, 5, or 15 min).
4. Inactivate the reaction at 95 °C for 5 min.

3.2.2 Nucleic Acid Digestion

1. Add 1 μL of nuclease S1 to the sample. Gently vortex, spin down, and incubate overnight at 37 °C.
2. Add 1 μL of Fast AP and 5 μL of Fast AP 10× Buffer to the sample. Bring to 50 μL with LC-MS grade water.
3. Gently vortex, spin down, and incubate at 37 °C for 1 h.

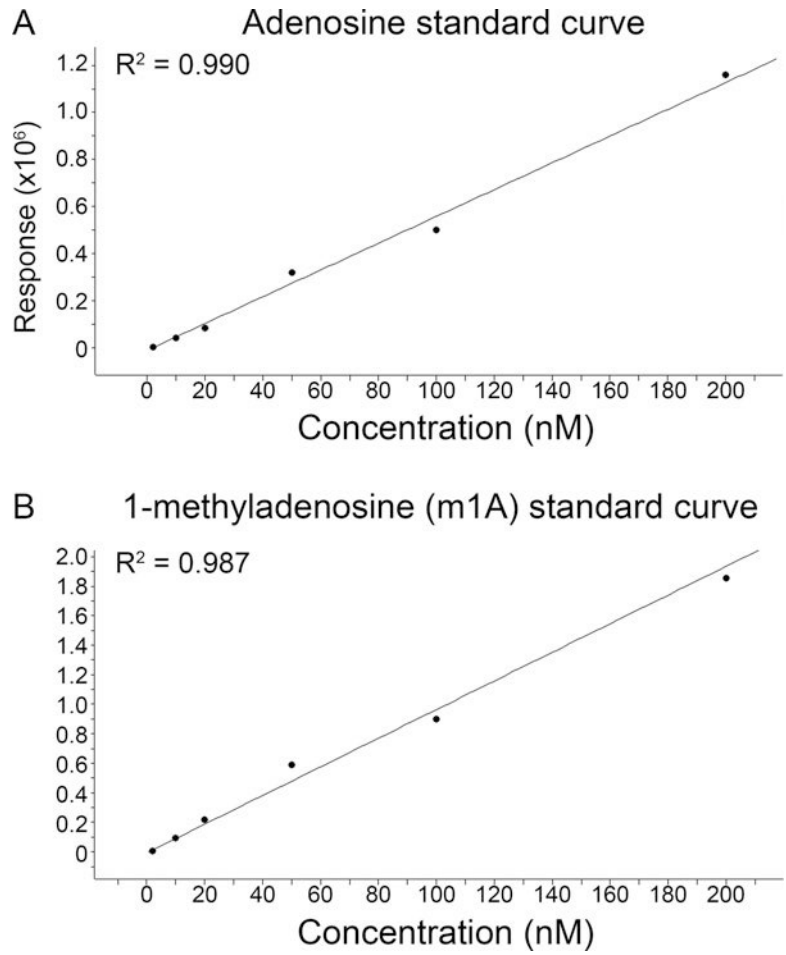


Fig. 2 HPLC-MS standard curves for a modified and unmodified ribonucleoside pair. After optimization, a standard curve was run for adenosine (**a**) and 1-methyladenosine (**b**). Concentrations analyzed were 2, 10, 20, 50, 100, and 200 nM. Calculated R^2 values are shown for each

4. Inactivate enzymes by incubating samples at 95 °C for 5 min.
5. Filter 25 μL of each sample and standard through a 0.22 μm syringe filter into a 200 μL autosampler vial (*see* **Note 3**). Cap vial, and shake tube to move all of sample to the bottom of the vial. Ensure no bubbles are in the sample. Filter at least 150 μL of water into a vial for use as a blank.

3.2.3 HPLC-MS

1. Use the MassHunter acquisition software to create a worklist for the run. Include vial position, the method created above, and data file location. For standard curve samples, make sure the sample type is "Calibration," and indicate the level of the sample, which should be equivalent to the concentration of the standard (*see* **Note 4**). Include at least two water samples at the

start of the run to flush the system, and one water sample indicated as "Blank" (*see* **Note 5**). Include script at the end of the program putting the machine on standby.

2. Load vials into autosampler, adhering to placement indicated in the method.
3. Turn on system, and allow all components to come to "Ready" status. Run the worklist.

3.2.4 Quantitative Data Analysis

1. Open Agilent Quantitative Analysis software, and create a "New Batch."
2. Open all data files from run into this batch.
3. Under Methods, select New Method, and then Create Method from Acquired MRM Data. Select a single standard curve sample with a mid-range concentration value. Select *Create Levels from Calibration Samples*. Change CF Weight to 1/x. Exit the method creator, indicating that the method should be applied to all samples in the batch.
4. Check the created standard curve to ensure it is linear and has an R^2 value greater than 0.98. If one value of the standard curve appears to be inhibiting this, right click on the dot to remove it from calculations, and then press "Analyze Batch" with the new standard curve. Do this for every standard used (*see* **Note 6**).
5. Export table to an Excel file for further analysis. To eliminate variability in nucleic acid amount injected in each sample, analysis should be done on the ratio of modified nucleoside to unmodified nucleoside (ex. m1A/A; *see* Fig. 3).

3.3 Quantification of m6A in Total RNA

3.3.1 Nucleic Acid Extraction, Digestion, and Analysis

1. Follow protocol from the miRNeasy extraction kit. Resuspend sample in 40 μL of LC-MS grade water in a 1.7 mL tube.
2. For digestion, add 1 μL of nuclease S1 to the sample. Gently vortex, spin down, and incubate overnight at 37 °C.
3. Add 1 μL of Fast AP and 5 μL of Fast AP 10× Buffer to the sample. Bring to 50 μL with LC-MS grade water. Gently vortex, spin down, and incubate at 37 °C for 1 h.
4. Inactivate enzymes by incubating samples at 95 °C for 5 min.
5. Filter 25 μL of each sample and standard through a 0.22 μm syringe filter into a 200 μL autosampler vial. Cap vial, and shake tube to move all of sample to the bottom of the vial. Ensure no bubbles are in the sample.
6. Filter at least 150 μL of water into a vial for use as a blank.
7. Use the MassHunter acquisition software to create a worklist for the run. Include vial position, the method created above, and data file location. For standard curve samples, make sure

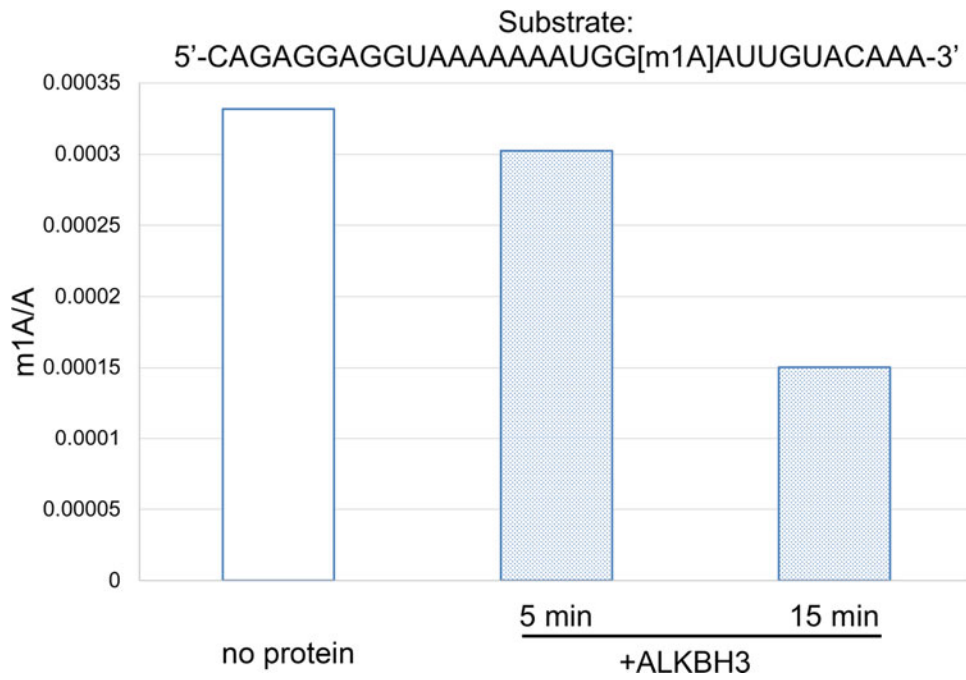


Fig. 3 In vitro analysis of ALKBH3-mediated demethylation of the indicated oligonucleotide. Quantified ratio of m1A/A in samples treated with or without ALKBH3 for the indicated period of time

the sample type is "Calibration," and indicate the level of the sample, which should be equivalent to the concentration of the standard (*see* **Note 4**). Include at least two water samples at the start of the run to flush the system, and one water sample indicated as "Blank" (*see* **Note 5**). Include script at the end of the program to put the machine on standby.

8. Load vials into autosampler, adhering to placement indicated in the method.
9. Turn on system, and allow all components to come to "Ready" status. Run the worklist.
10. As an alternative to the quantitative method, a qualitative analysis is also possible. For this, open Agilent Qualitative Analysis software.
11. Select the samples of interest, and open their respective chromatograms.
12. Under Method Editor, select the transition you are interested in (for m^6A, 282.1 to 150.1), and add it to the method. Then press "Extract Additional Chromatograms." This function will integrate the area under the curve of each peak identified by the program.

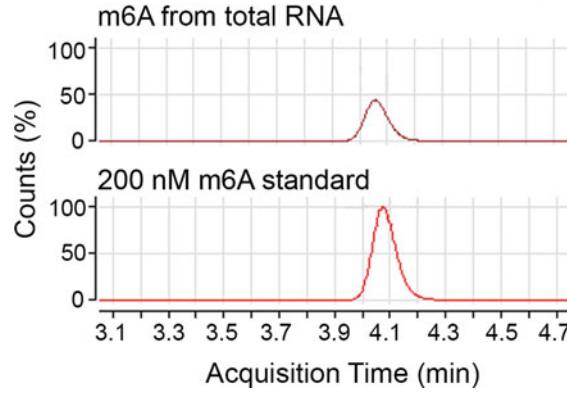


Fig. 4 The analyzed HPLC-MS chromatograms for m6A from a total RNA sample (top) or a 200 nM m6A standard (bottom)

13. Compare the peaks on the chromatograms of the samples by retention time to the peaks on the standard chromatograms to determine which peak is representing the nucleoside of interest (*see* **Note 7**). The integrated area under the curve can be compared to determine the relative amounts of m^6A in samples (Fig. 4).

3.4 Analysis of DNA Methylation Adduct Repair in Cultured Cells

3.4.1 Induction of Methylation Damage

1. Cell lines are plated in 100 mm petri dishes and grown to 60–70% confluency.
2. Make a 100 mM MMS solution by diluting stock MMS into the complete medium, and then make a 2 mM MMS-containing medium. Remove media from dishes, and add 8 mL 2 mM MMS-containing medium into each dish. Incubate cells for 1 h at 37 °C.
3. Recover cells from MMS exposure by washing with 5 mL 1× PBS. Leave one dish without recovery as the "0 hour" time point.

3.4.2 DNA Extraction and Purification

1. Harvest cells at the desired time points (e.g., 0, 2, 4, 8, and 24 h) after MMS exposure. Aspirate media, and wash cells once with 5 mL 1× PBS.
2. Aspirate PBS, add another 1 mL 1× PBS into each dish, scrape cells by using cell lifter, and collect the cell suspensions into 1.7 mL microcentrifuge tubes.
3. Spin down cells at 1000 × *g* for 2 min at 4 °C.
4. Aspirate all remaining liquid from tubes. Extract genomic DNA by following the "cultured cells" protocol provided by the DNeasy Blood & Tissue Kit.

3.4.3 DNA Digestion and HPLC-MS/MS

1. Digest 2 μg genomic DNA from each sample with 2 μL Nucleoside Digestion Mix in a total of 50 μL reaction volume, and incubate at 37 °C for at least 2 h.
2. Filter 25 μL of each sample and standard (1medA and dA) through a 0.22 μm syringe filter into a 200 μL autosampler vial (*see* **Note 3**). Cap vial, and shake tube to move all of sample to the bottom of the vial. Ensure no bubbles are in the sample. Filter at least 150 μL of water into a vial for use as a blank.
3. Dilute samples 100-fold by mixing 2 μL with 198 μL LC-MS grade water. Filter each sample as above. These diluted samples are used for acquiring the results of dA (*see* **Note 8**).
4. Use the MassHunter acquisition software to create a worklist for the run. Include vial position, the method created above, and data file location. Include at least two water samples at the start of the run to flush the system, and one water sample indicated as "Blank" (*see* **Note 5**). Include script at the end of the program putting the machine on standby.
5. Load vials into autosampler, adhering to placement indicated in the method.
6. Turn on system, and allow all components to come to "Ready" status. Run the worklist.

3.4.4 Data Analysis

1. Open Agilent Qualitative Analysis software. Select the samples of interest to open chromatograms for them.
2. Under Method Editor, select the transitions you are interested in (266.13 to 150 for 1medA, 252.11 to 136 for dA), and add it to the method. Then press "Extract Additional Chromatograms." This function will integrate the area under the curve of each peak identified by the program.
3. Compare the peaks on the chromatograms of the samples by retention time to the peaks on the standards chromatograms to determine which peak is representing the nucleoside of interest (*see* **Note 7**). The integrated area under the curve can be compared to determine relative amounts of 1medA in samples. The amount of 1medA is then normalized with its unmodified counterpart (dA) amount (Fig. 5).

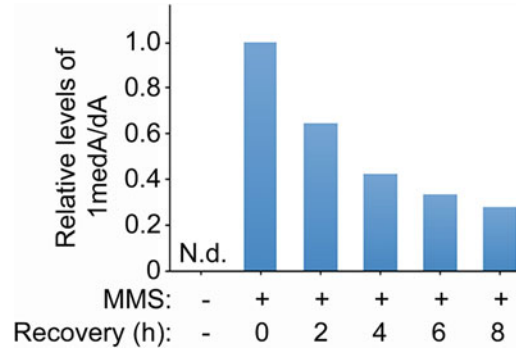


Fig. 5 Analysis of 1medA levels in genomic DNA from MMS-treated HeLa cells. Cells were treated with 2 mM MMS for 1 h, then recovered in fresh media, and harvested at the indicated time points after recovery. Genomic DNA was extracted from each sample and digested for LC-MS/MS analysis of 1medA and dA levels. The level of 1medA was normalized with its unmodified counterpart (dA). The relative 1medA/dA level of each sample is shown as compared to the "0 h" sample. N.d., not detected

4 Notes

1. In order to prevent the growth of organic material within the HPLC system, all solvent must contain 0.1% formic acid. To do so, add 1 mL of formic acid to 1 L of solvent.
2. Some standards may be only available in the nucleotide form. For these, alkaline phosphatase can be used to remove the phosphate moieties, converting them to nucleosides which will facilitate HPLC-MS analysis.
3. In order to allow for accurate quantitation of sample concentration, all samples must have a concentration which falls within the linear range of the standard curve. This means some samples may need to be diluted, especially when quantitating unmodified nucleosides. To do so, dilute the sample, and then proceed with filtering 25 μL as outlined above.
4. A typical standard curve should range between 10 nM and 1000 nM with five to six concentrations evenly spaced throughout. The standard curve should be run from the lowest concentration to the highest concentration. Due to the variability which may be found between runs of the system, it is best practice to include a standard curve with every run and to use that standard curve only for calculations of the samples with which it was run.
5. Water samples should be run at the beginning of every worklist to ensure the system is clear of remnants of any previous samples. Additionally, when first optimizing a method, it can be helpful to run a water sample between every sample to

ensure the column is completely clear and no extra runoff is present that would skew the determined concentration of subsequent samples. If excessive runoff is observed in these water samples, it may be helpful to extend the run time of the solvent for each sample.

6. After a certain concentration, the standard curve will no longer produce a linear relationship, particularly at higher concentrations. These points should not be considered during analysis. Any sample that gives a concentration in the non-linear region of the standard curve should be diluted and re-run to ensure accuracy of the determined concentration.
7. Some mass transitions can be the same for different compounds, such as m6A and m1A, due to structural similarities. Therefore, it is important to compare the peak of the standard by retention time to the samples to determine which of closely spaced peaks represents the compound of interest.
8. The amount of dA is usually >10,000-fold higher than 1medA in a genomic DNA sample extracted from mammalian cell lines. Given that, it is important to dilute a digested sample 100- to 1000-fold to make sure the detected amount of dA is in the linear range.

Acknowledgments

Work in our laboratory is supported the by the NIH (R01 CA193318, R01 CA227001, and P01 CA092584), the American Cancer Society research scholar program (RSG-18-156-01-DMC), and the Alvin J. Siteman Cancer Center Investment Program, which is supported by the Foundation for Barnes-Jewish Hospital Cancer Frontier Fund and the National Cancer Institute, Cancer Support Grant P30 CA091842.

References

1. Baylin SB, Jones PA (2011) A decade of exploring the cancer epigenome—biological and translational implications. Nat Rev Cancer 11(10):726–734
2. Shen L et al (2014) Mechanism and function of oxidative reversal of DNA and RNA methylation. Annu Rev Biochem 83:585–614
3. Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. Annu Rev Biochem 74:481–514
4. Wu X, Zhang Y (2017) TET-mediated active DNA demethylation: mechanism, function and beyond. Nat Rev Genet 18(9):517
5. Liu J et al (2014) A METTL3–METTL14 complex mediates mammalian nuclear RNA N 6-adenosine methylation. Nat Chem Biol 10(2):93–95
6. Jia G, Fu Y, Zhao X, Dai Q, Zheng G et al (2011) N6-Methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7:885–887
7. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM et al (2013) ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49:18–29

8. Phizicky EM, Hopper AK (2010) tRNA biology charges to the front. Genes Dev 24(17): 1832–1860
9. Rošić S et al (2018) Evolutionary analysis indicates that DNA alkylation damage is a byproduct of cytosine DNA methyltransferase activity. Nat Genet 50(3):452–459
10. Strauss B, Scudiero D, Henderson E (1975) The nature of the alkylation lesion in mammalian cells. In: Molecular mechanisms for repair of DNA. Springer, Boston, MA, pp 13–24
11. Li X et al (2017) Base-resolution mapping reveals distinct m1A methylome in nuclear- and mitochondrial-encoded transcripts. Mol Cell 68(5):993–1005
12. Xu L et al (2017) Three distinct 3-methylcytidine (m3C) methyltransferases modify tRNA and mRNA in mice and humans. J Biol Chem 292(35):14695–14703
13. Zhang L-H et al (2020) The SUMOylated METTL8 induces R-loop and tumorigenesis via m3C. iScience 23(3):100968
14. Fu D, Calvo J, Samson L (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. Nat Rev Cancer 12:104–120. https://doi.org/10.1038/nrc3185
15. Aas P, Otterlei M, Falnes P et al (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature 421:859–863. https://doi.org/10.1038/nature01363
16. Drabløs F et al (2004) Alkylation damage in DNA and RNA—repair mechanisms and medical significance. DNA Repair 3(11): 1389–1407
17. Dango S, Mosammaparast N, Sowa ME, Xiong L-J, Wu F, Park K, Rubin M, Gygi S, Harper JW, Shi Y (2011) DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation. Mol Cell 44:373–384
18. Barciszewska AM, Murawa D et al (2007) Analysis of 5-methylcytosine in DNA of breast and colon cancer tissues. IUBMB Life 59(12): 765–770. https://doi.org/10.1080/15216540701697412
19. Michalak M, Plitta-Michalak BP et al (2015) Global 5-methylcytosine alterations in DNA during ageing of Quercus robur seeds. Ann Bot 116(3):369–376. https://doi.org/10.1093/aob/mcv104
20. Azimzadeh P, Asadzadeh Aghdaei H, Tarban P, Akhondi MM, Shirazi A, Khorram Khorshid HR (2016) Comparison of three methods for mitochondria isolation from human liver cell line (*HepG2*). Gastroenterol Hepatol Bed Bench 9(2):105–113

Chapter 9

Targeted Formation of 8-Oxoguanine in Telomeres

Ryan P. Barnes, Sanjana A. Thosar, Elise Fouquerel, and Patricia L. Opresko

Abstract

Mammalian telomeres are guanine-rich sequences which cap the ends of linear chromosomes. While recognized as sites sensitive to oxidative stress, studies on the consequences of oxidative damage to telomeres have been primarily limited to experimental conditions which cause oxidative damage throughout the whole genome and cell. We developed a chemoptogenetic tool (FAP-mCER-TRF1) to specifically induce singlet oxygen at telomeres, resulting in the formation of the common oxidative lesion 8-oxo-guanine. Here, we describe this tool and detail how to generate cell lines which express FAP-mCER-TRF1 at telomeres and verify the formation of 8-oxo-guanine.

Key words Telomeres, Oxidative stress, DNA damage, Chemoptogenetic tool, 8-oxo-guanine, TRF1

1 Introduction

Linear chromosomes require specialized structures to prevent them from being recognized as DNA double-stranded breaks (DSBs) in the cell [1]. Mammalian telomeres consist of 5′ TTAGGG 3′ repeats which adopt a protective T-loop formation, shielding the exposed DNA ends. T-loops are stabilized with the assistance of the shelterin complex, which consists of TRF1, TRF2, TPP1, TIN2, POT1, and RAP1 [2]. TRF1 and TRF2 bind the duplex telomere sequence with high affinity and are localized specifically to telomeres. Several studies have shown telomeres are hotspots for DNA damage from free radicals during conditions of oxidative stress (reviewed in [3–6]). In these reports, telomeres from cells exposed to high oxygen tension, ionizing radiation, hydrogen peroxide, etc. show accelerated telomere shortening that is coincident with cellular senescence and/or apoptosis. However, all of these stressors have the important caveat that they damage other cellular components, not just the telomere, and they generate several types of DNA adducts as well as strand breaks [7]. Therefore,

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_9,

understanding mechanistically the contribution of a specific oxidative DNA adduct or lesion to the cellular outcome is not possible with general oxidant treatments.

To understand the specific contribution of the common oxidative DNA adduct, 8-oxo-guanine (8oxoG), we have employed a chemoptogenetic tool which can produce singlet oxygen with high spatial and temporal control [8, 9]. Singlet oxygen reacts rapidly and specifically with guanine to primarily form 8oxoG [10]. In this system, FAP (fluorogen activated peptide; dL5**) is a benign protein that binds malachite green (MG) dyes with nanomolar affinity. When bound, the MG-ester dye can be excited with 630 nm light, causing it to fluoresce at 670 nm and allowing for imaging in live or fixed cells. Di-iodination (MG-2I) of the MG-ester dye results in a bathochromic shift, moving the excitation maximum to a near-infrared wavelength (660 nm), which is not well absorbed by biological molecules, thereby limiting off-target effects. Moreover, this dramatically increases the yield of singlet oxygen (>100-fold) when MG-2I binds FAP and is exposed to 660 nm light, while there is no increase in singlet oxygen generation by free unbound dye [9]. Given the short half-life and radius of diffusion of singlet oxygen, the FAP coupled with the photosensitizer MG-2I dye allow for the targeted production of singlet oxygen at discreet locations within the cell.

The FAP and MG-2I system is a versatile tool that allows for the targeted production of singlet oxygen in a variety of cellular compartments [9, 11]. The dL5** sequence can be cloned into any gene of interest and expressed in cells as a fusion protein. To target the FAP to telomeres, we fused it to telomere binding protein TRF1 and validated this tool as a method to selectively produce 8oxoG exclusively at telomeres (Fig. 1a) [8]. The construct includes mCerulean fluorescent protein to enable visualization of the fusion protein in cells. In this chapter we detail how we generated stable cell lines expressing FAP-TRF1 homogeneously at telomeres, verified expression of the fusion protein, and confirmed induction of telomeric 8oxoG at the telomeres following dye and light treatment.

2 Materials

2.1 Generating FAP-mCER-TRF1 Expressing Cell Lines by Lentiviral Transduction and Single Cell Cloning

1. pLVX-FAP-mCER-TRF1 plasmid (Addgene # 168176).
2. Plasmid DNA purification kit (QIAGEN, or suitable alternative).
3. HEK-293T cells (ATCC).
4. Dulbecco's Modified Eagle Medium, high glucose (DMEM) supplemented with and without 50 units/mL penicillin, 50 μg/mL streptomycin, and 10% (v/v) fetal bovine serum.

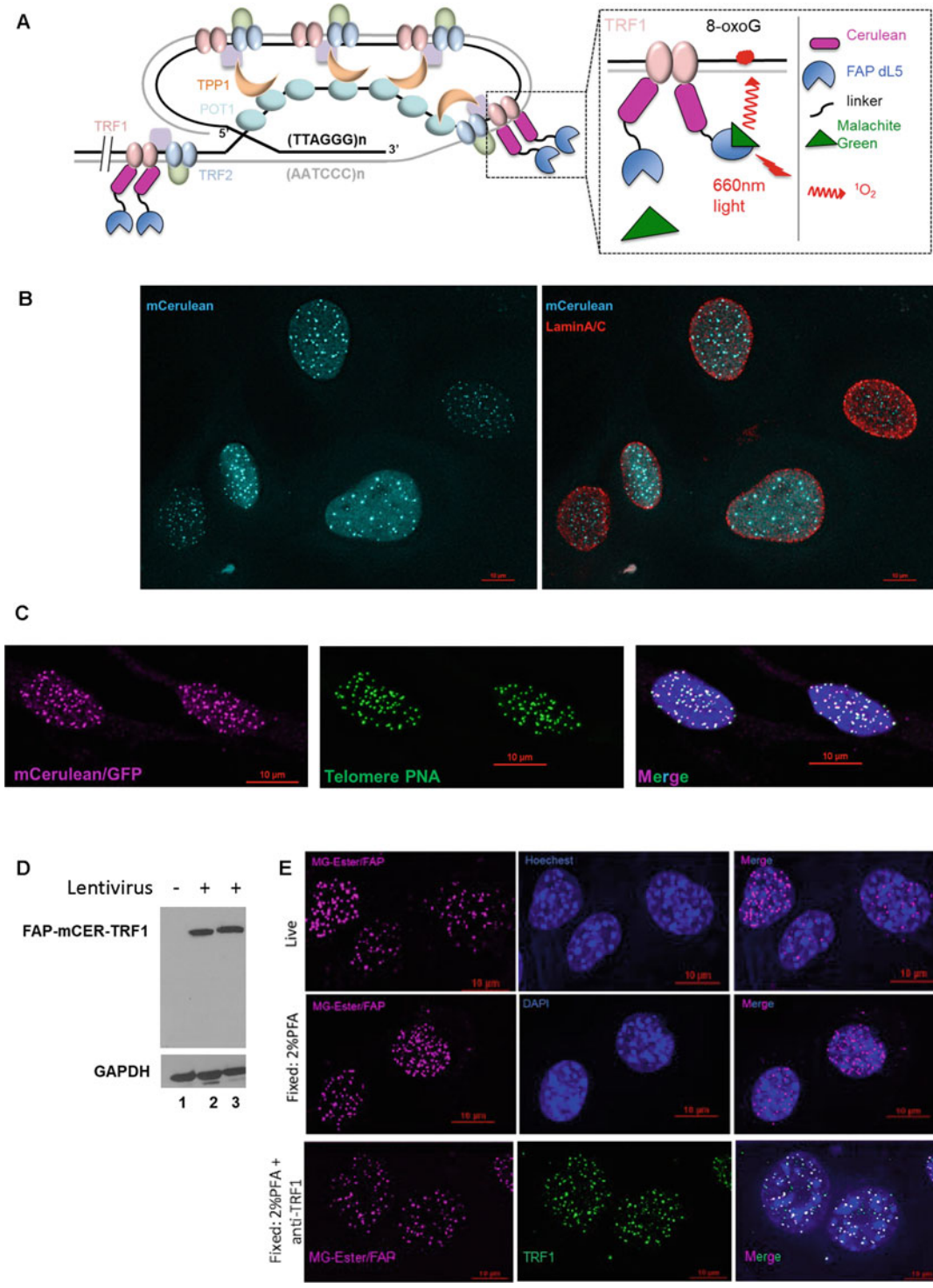


Fig. 1 (**a**) Schematic depicting human telomeres with shelterin proteins and the FAP-mCER-TRF1 system for targeted 8oxoG formation. (**b**) Direct visualization of mCerulean (cyan). Cell nuclei are identified with Lamin A/C immunofluorescence (red). (**c**) Immunofluorescence/FISH for FAP-mCER-TRF1 (purple) and telomeres (green). (**d**) Example of FAP-mCERT-TRF1 immunoblot before and after lentiviral infection to deliver the transgene. Lane 1 = uninfected cells; 2 = infected cells before cloning; 3 = infected cells after cloning. (**e**)

5. Opti-MEM media (Gibco).
6. FuGENE 6 Transfection Reagent (Promega #E2691).
7. Mission Lentivirus Packaging Mix (Sigma).
8. Polybrene 10 mg/mL (Sigma).
9. 0.22 μm syringe filter (Millipore) and syringe.
10. Any mammalian cell line of interest suitable for lentiviral transduction can be used. This protocol has been used successfully with HeLa, U2OS, BJ-5ta, and hTERT RPE-1 cells (ATCC).
11. G418 (Gibco).
12. 0.5% trypsin-EDTA.

2.2 Screening and Verification of FAP-mCER-TRF1 Expression and Localization by Fluorescence Microscopy

1. Coverslips 22 mm × 22 mm.
2. Formaldehyde (37%, Sigma #252549).
3. 1× PBS.
4. 1% Bovine Serum Albumin (BSA) in PBS.
5. PBS-T, 0.2% Triton X-100 in PBS.
6. IF Blocking buffer (10% normal Goat Serum, 1% BSA, 0.1% Triton X-100 in PBS).
7. DAPI.
8. Prolong Diamond Anti-Fade (Thermo Fisher).
9. Nikon Ti2E inverted fluorescent microscope, or equivalent instrument.
10. Apochromatic (APO) 60× oil immersion objective N.A. 1.4.
11. DAPI/Hoechst/Alexa Fluor 350 filter set.
12. EGFP/FITC/Cy2/Alexa Fluor 488 filter set.
13. Cy5 filter set.
14. Nikon NIS Elements software.
15. Nikon NIS Viewer software.

2.3 IF-FISH

1. Antibody that recognizes mCerulean (Anti-GFP Abcam #6556).
2. Secondary antibody Goat anti-Rabbit IgG, Cy5 conjugated (Invitrogen).
3. Ethanol 70, 90, and 100%.
4. Hybridization mix (30 μL per coverslip).

Fig. 1 (continued) MG-ester imaging. Cells were incubated with MG-ester (purple) 15 min with or without Hoechst (blue), and then imaged live, or fixed as indicated in the text. Bottom panel shows co-localization with TRF1 (green) as detected by immunofluorescence in fixed cells

(a) 70% deionized formamide (Millipore #S4117).

(b) 1× blocking reagent (Roche #11096176001). Prepare a 10% stock by dissolving in maleic acid buffer (100 mM maleic acid; 150 mM NaCl, pH 7.5).

(c) 10 mM Tris HCl, pH 7.5.

(d) 1× $MgCl_2$ buffer (82 mM Na_2HPO_4; 9 mM citric acid; 20 mM $MgCl_2$; pH 7.4).

(e) 1× telomeric PNA probe (PNA Bio #1002 for Alexa 488 conjugated $CCCTAA)_4$ or other fluorochrome.

(f) ddH_2O.

5. Hybridization wash (70% deionized formamide in Milli-Q water, 10 mM Tris pH 7.5).

2.4 FAP-Bound MG-Ester Imaging

1. 35 mm MatTek glass bottom dish for live cell imaging.
2. MG-ester dye (*see* **Note 1**).
3. Hoechst 33358 dye.
4. 2% paraformaldehyde in PBS.
5. Anti-TRF1 antibody (Abcam: ab1423).

2.5 Inducing Telomeric 8oxoG Formation

1. MG-2I photosensitizer dye (*see* **Note 1**).
2. 660 nm LED light source ~100 mW/cm^2. Construction of light boxes for FAP-MG-2I activation that provide illumination from above or below the culture dish are described in [9, 12].
3. FieldMaxII laser power meter with 7.9 mm MDL#OP-2 VIS sensor to measure light intensity from 660 nm light source.

2.6 Genomic DNA Preparation for 8oxoG Detection

1. Qiagen 100/G or 20/G DNA isolation kit.
2. Antioxidants.
   (a) Butylated hydroxytoluene (Sigma W218405)—dissolve in DMSO for a 100 mM solution, aliquot, and store at −20 °C.
   (b) Deferoxamine mesylate salt (Sigma D9533)—dissolve in water day of use for a 100 mM solution. This reagent is not stable and cannot be stored.
3. Cold PBS and media.
4. Tris EDTA pH 8 (TE).

2.7 Enzymatic Treatments for 8oxoG Detection

1. Telomere restriction enzymes *Alu*I, *Hph*I, *Mnl*I, *Hinf*I, and CutSmart buffer for human DNA (NEB). For mouse, use *Alu*I and *Mbo*I.
2. S1 Nuclease (Thermo EN0321).

3. S1 Diluent (20 mM Tris pH 7.5; 50 mM NaCl; 50% glycerol) store at -20 °C.
4. FPG, 8 U/μL (NEB).
5. OGG1, 125 nM final concentration in reaction (Novus).
6. APE1, 3.3 U/μg DNA final concentration in reaction *(NEB)*.

2.8 S Blot Analysis of Cleaved Telomere Restriction Fragments

1. 0.5× TBE, store at 4 °C.
2. Certified Megabase Agarose (Bio-Rad 1613109).
3. 2.5 kb Molecular Ruler DNA ladder (Bio-Rad 170-8205).
4. SYBR Green Dye (Thermo S7567).
5. Denaturation buffer (0.5 M NaOH, 1.5 M NaCl).
6. Neutralization buffer (0.5 M Tris pH 8, 1.5 M NaCl).
7. Church buffer (0.25 M sodium phosphate buffer pH 7.2, 1 mM EDTA, 1% BSA, 7% SDS).
   (a) 0.5 M Sodium phosphate pH 7.2 0.5 M—500 mL.
      - Prepare sodium phosphate solutions "A" and "B":
        - A = 1 M Na_2HPO_4, $2H_2O$: 177.9 g per liter.
        - B = 1 M NaH_2PO_4, H_2O: 138 g per liter.
      - For 1 L of 0.5 M sodium phosphate pH 7.2 0, mix 342 mL A, 158 mL B, with 500 mL water.

   (b) 0.5 M EDTA—2 mL.

   (c) BSA Fraction V—10 g.

   (d) SDS pellets—70 g.

   Dissolve the BSA in ~50 mL water *without heating*, separately from the SDS, which should be heated. Then add BSA, SDS, and EDTA to the sodium phosphate in order. Bring to 1 L with ddH_2O, and filter sterilize.
8. C-strand telomeric oligonucleotide probe 5′ $(CCCTAA)_4$ 3′ (10 pmol/μL).
9. G-strand telomeric oligonucleotide probe 5′ $(TTAGGG)_4$ 3′ (10 pmol/μL).
10. T4 Polynucleotide Kinase 10 U/μL and 10× T4 Kinase buffer.
11. ATP, [γ-^{32}P]-3000 Ci/mmol.
12. MicroSpin G-25 Sephadex columns (GE Healthcare).
13. 2× SSC, diluted from 20× stock in ddH_2O (Fisher BP-1325).
14. 0.1% SDS, 0.1× SSC.

3 Methods

3.1 Generating FAP-mCer-TRF1 Expressing Cell Lines by Lentiviral Transduction

The FAP-mCER-TRF1 transgene can be introduced into cell lines using standard plasmid transfection protocols, or by transduction with lentivirus, depending on the cell line (*see* **Note 2**).

1. Propagate and purify the pLVX-FAP-mCER-TRF1 lentiviral vector according to QIAGEN plasmid purification kit protocol, or suitable alternative (*see* **Note 3**).
2. Day 1: Seed and culture HEK293T cells according to the ATCC guidelines. Incubate at 37 °C and 5% CO_2 in a humidified cell culture incubator.
3. Day 2: Re-fresh the media, and examine cells under a light microscope for dish adherence and normal morphology.
4. Day 3: Split cells and seed 500,000 cells in a 6 cm dish, and incubate overnight. The cells should be ~60% confluent the next day.
5. Day 4: Warm Opti-MEM and DMEM media without antibiotics to 37 °C and the FuGENE reagent and lentivirus packaging mix to room temperature (RT). For each 6 cm dish, combine 1 μg of pLVX-mCER-TRF1 plasmid DNA, 10 μL lentivirus packaging mix, and 6 μL FuGENE in 70 μL of Opti-MEM. Mix well at RT for 15 min. Remove media from cells, add transfection mix dropwise, and then add fresh DMEM media without antibiotics. Incubate for 24 h.
6. Day 5: Discard old media, and add 4 mL fresh media with antibiotics and incubate overnight.
7. Day 6: Collect virus-containing media for harvesting, add fresh media to transfected HEK293T cells, and return to the incubator for 24 h. Filter the lentivirus media with a 22 μm syringe filter into a 15 mL falcon tube, and add fresh media for 6 mL total volume. Add polybrene at 1:1000 dilution (or 10 μg/mL final) to the falcon tube, and then add mixture directly to cells of interest for transgene introduction (3 mL per 6 cm dish and 1.5 mL per well of a six-well plate). Incubate cells of interest overnight (*see* **Notes 4** and **5**).
8. Day 7: Repeat **step 7** for a second lentiviral harvest and infection of the cells of interest. Incubate cells overnight. Discard transfected HEK293T cells.
9. Day 8: Add fresh media containing 1000 μg/mL G418 to lentiviral infected cell culture to select for cells that were successfully infected. The FAP-mCER-TRF1 vector contains a neomycin resistance gene. Culture the cells for 2–3 days in media containing 1000 μg/mL G418, or until uninfected control cells lacking the lentivirus die (*see* **Note 6**).
10. After uninfected cells have died, maintain the cell culture in media containing 500 μg/mL G418 (*see* **Note 7**).

3.2 Propagating FAP-mCER-TRF1 Expressing Cells and Single Cell Cloning

Single cell cloning ensures the expression of FAP-mCER-TRF1 at telomeres is fairly homogeneous. We found even with lentiviral infection, and especially standard transfection, the expression can vary among cells in a population. To expand clonal populations from single cells, any conventional approach is appropriate such as cloning disks or serial dilution and may require optimization depending on the cell line. The following protocol is for serial dilution.

1. Starting from a near confluent 10 cm dish, trypsinize and count the cells. Depending on the count, ~20% of the culture can be frozen for storage, ~70% can be harvested for Western blot confirmation of FAP-mCER-TRF1 expression (see below), and the remaining 10% can be used for clonal expansion.
2. Resuspend the cells in 10 mL of media. Based on the concentration (e.g., 200,000 cells/mL), dilute to 0.01 cells *per μL* so that addition of 100 μL to a well in a 96-well plate equates to 1 cell per well (*see* **Note 8**).
3. Place 50 mL of the final cell dilution in a sterile trough. For example, 200,000 cells/mL = 200 cells/μL, so three serial dilutions of 1:10 each (total 1:1000 dilution) = 0.2 cells/μL. Add 2.5 mL of 1:1000 dilution to 50 mL total volume for a final dilution of 0.01 cells/μL. Use a multichannel pipet to add 100 μL to each well for 1 cell per well.
4. Culture cells in media with 500 μg/mL G418, and expand clones to at least a confluent 10 cm dish.

3.3 Initial Screening of Clones with mCerulean Imaging

1. After clonal expansion, harvest the cells and seed ~75,000 on glass coverslips, and incubate overnight. Freeze the rest in cryovials for storage.
2. Fix the coverslips using 4% formaldehyde in PBS for 10 min at RT. Then, rinse with 1% BSA in PBS, and wash with 0.2% Triton X-100 in PBS.
3. Mount the coverslips on slides with Prolong Diamond Anti-Fade without DAPI (*see* **Note 9**).
4. Examine clones under a fluorescence microscope for FAP-mCER-TRF1 expression with a CFP filter set suitable for mCerulean imaging. A good clone displays distinct mCerulean foci and relatively homogeneous expression among different cells (Fig. 1b).

3.4 Verifications of FAP-mCER-TRF1 Expression and Localization

After the initial screen for mCerulean fluorescence, combined immunofluorescence and telomere fluorescence in situ hybridization (IF-FISH) is the most rigorous method to verify FAP-mCER-TRF1 expression and localization at telomeres.

3.4.1 Verification by IF-FISH

1. Stain fixed cells grown on coverslips with an anti-GFP antibody that recognizes mCerulean, and add a secondary antibody conjugated to Cy5 or another fluorophore using standard IF protocols (*see* **Note 10**).
2. After IF and washing with 0.2% Triton X-100 in PBS, fix again with 4% formaldehyde in PBS for 10 min at RT, and then rinse with 1% BSA in PBS.
3. Dehydrate the coverslips in subsequent solutions of 70%, 90%, and 100% ethanol incubations for 5 min each.
4. Prepare the hybridization mix containing Alexa 488 conjugated telomeric PNA probe during this time. Heat for 5 min on a heat block at 80 °C, and then place on ice.
5. Remove coverslips from 100% ethanol, and let air dry until ethanol is fully evaporated.
6. Place ~30 μL of hybridization mix on a glass slide. Place the coverslip face down on the hybridization mix while avoiding bubbles.
7. Hybridize for 10 min on the heat block. Turn heat block over so the flat side faces up.
8. Incubate for 2 h in a humid chamber with a wet Kim-wipe at RT or overnight at 4 °C.
9. Wash the coverslips two times 15 min each with 70% formamide, 10 mM Tris pH 7.5.
10. Wash three times with PBS-T 5 min each.
11. Rinse with water.
12. Stain with 1 μg/mL DAPI diluted in water for 10 min at RT.
13. Remove coverslips, and mount on slides with Prolong Diamond Anti-Fade.

    Image with FITC and Cy5 filter sets. A good clone will exhibit FAP-mCER-TRF1 localization exclusively at telomeres and show relatively homogeneous expression throughout the population of cells (Fig. 1c).

3.4.2 Verification by FAP-Bound MG-Ester Imaging

FAP-mCER-TRF1 expression can be verified by immunoblotting with either TRF1 or GFP (that recognizes mCerulean) antibodies (fusion protein ~120 kDa) (Fig. 1d). Alternatively, expression can be verified in live cells by imaging FAP bound to MG-ester dye with Hoechst dye staining the nuclear DNA or in fixed cells (Fig. 1e). This is useful if expression of FAP-TRF1 lacking the mCerulean protein is desired and can be coupled with IF against a telomeric protein to confirm localization to telomeres (Fig. 1e).

1. Seed cells on a coverslip (sterilized in ethanol) in a 35 mm dish or a well of a six-well plate. Seed cells in a glass bottom 35 mm dish for live cell imaging. Incubate overnight.

2. When cells are ~60% confluent, replace cell culture media with 2 mL Opti-MEM containing 100 nM MG-ester, and incubate for 15 min at 37 °C (*see* **Note 11**).
3. If performing live-cell imaging, add Hoechst (1:20000 dilution) with the MG-ester, and incubate for 15 min.
4. To fix the cells, wash twice with 1× PBS, and then add 2% paraformaldehyde in PBS for 10 min at RT.
5. Rinse with 1% BSA in PBS.
6. Wash three times with PBS-T for 5 min each, and then rinse with 1× PBS and water.
7. Samples can be stained with an antibody against a telomeric protein to confirm telomere localization using standard IF protocols. We successfully used anti-TRF1 antibody (Abcam # ab1423) (*see* **Note 12**).
8. Stain fixed cells with 1 μg/mL DAPI in water and mount on slides.
9. Examine live or fixed cells using the Cy5 filter set to visualize MG-ester.

3.5 Inducing and Verifying 8oxoG Formation at Telomeres

This approach is based on the enzymatic conversion of 8oxoG to a DSB so that cleaved telomeres can then be detected by changes in telomere length using gel electrophoresis and Southern blotting (Fig. 2). 8oxoG is first converted into a single-strand break using either FPG alone or OGG1 with APE1 and then converted into a DSB with S1 nuclease treatment. This approach works well for cells with long telomeres (≥20 kb), but is more difficult in cells with shorter telomere lengths (~10 kb). OGG1 specifically removes 8oxoG and then APE1 cleaves the abasic site. FPG recognizes 8oxoG and other oxidized purines, but is both an effective glycosylase and AP lyase (Fig. 3a).

3.5.1 Inducing Telomeric 8oxoG Formation

1. To induce singlet oxygen at telomeres for 8oxoG production, culture cells in a 10 cm dish until approximately 90% confluent (*see* **Notes 4** and **13**).
2. Change the media to Opti-MEM, and equilibrate for 15 min in the incubator.
3. Add MG-2I to 100 nM final concentration, and incubate for 15 min.
4. Measure the light intensity of 660 nm light source with the light meter. We use a light box fitted with 660 nm LED at 100 mW/cm^2.
5. Expose cells to 660 nm light starting at 5 min, and increase depending on the desired amount of damage since longer light

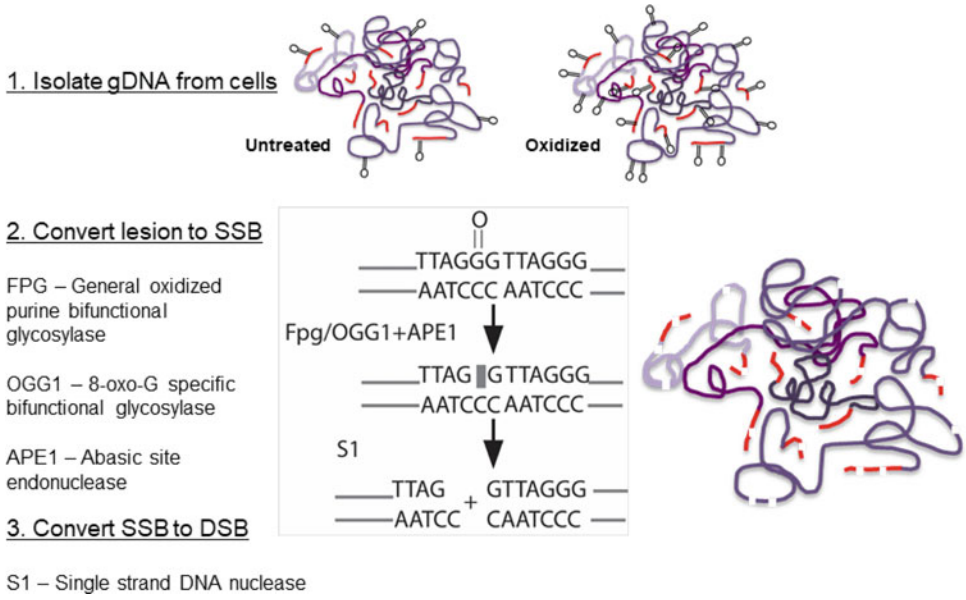


Fig. 2 Schematic for 8oxoG detection at telomeres. Genomic DNA (gDNA) is isolated from cells and then treated with DNA repair enzymes to convert 8oxoG into a DSB, before digestion to release telomere restriction fragments, and PFGE to separate intact telomeres from cleaved telomeres

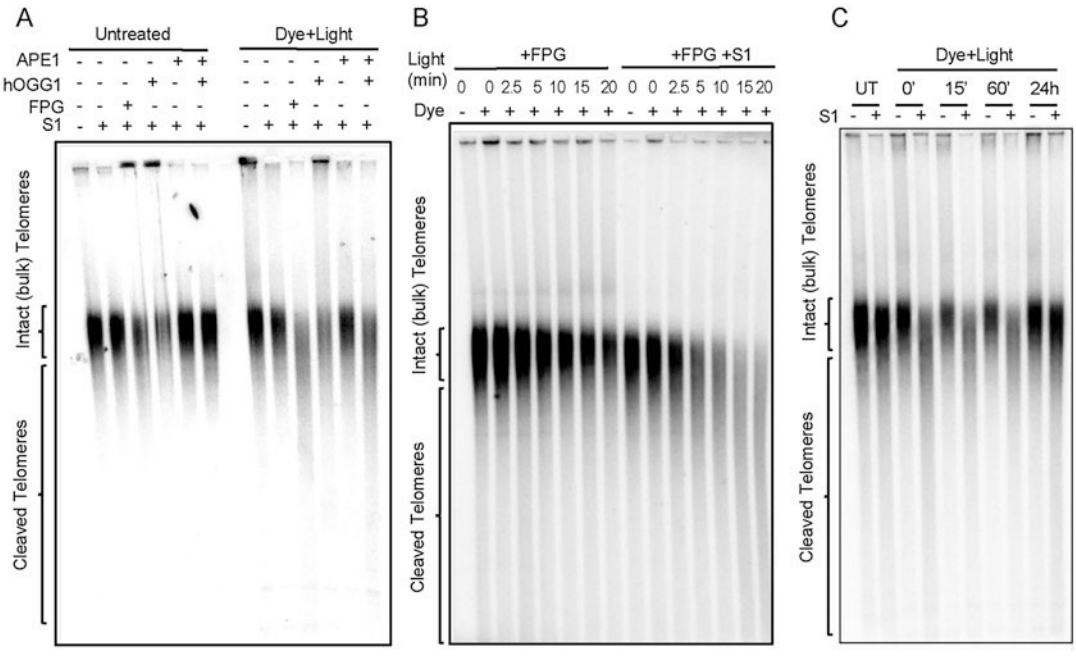


Fig. 3 Examples of S Blot analysis 8oxoG detection by conversion to DSBs in telomere restriction fragments. (**a**) DNA from untreated cells or 100 nM dye +5 min light-treated cells was treated with repair enzymes APE1, OGG1, APE1 + OGG1, or FPG and then digested with S1 nuclease. (**b**) Dose-response analysis. Cells were treated with 100 nM dye and increasing amounts of light exposure times as indicated. DNA was digested with FPG with and without S1 nuclease. (**c**) Recovery repair analysis. Cells were treated with dye and light for 5 min and then harvested immediately or allowed to recover for the indicated times. DNA was digested with FPG enzyme with and without S1 nuclease

exposure times increase the amount of telomeric 8oxoG formation (Fig. 3b) (*see* **Note 14**).

3.5.2 Genomic DNA Preparation

1. Harvest cells immediately after treatment, especially repair proficient cells, if the desire is to detect 8oxoG formation. Harvest after recovery periods if examining DNA repair (Fig. 3c).
2. Wash cells with 1XPBS, add trypsin, and then harvest by centrifugation. Resuspend with ~5 mL *cold* media. Alternatively, cells can be scraped off the dish on ice with cold PBS and collected.
3. Centrifuge again, and wash with *cold* PBS.
4. Proceed with Qiagen 100/G kit according to the manufacturer's instructions with the additions described below. Prepare a master mix of Qiagen buffer C1 and sterile ddH_2O at 1:3 ratio (3 mL C1 and 9 mL ddH_2O per sample). Add freshly dissolved deferoxamine and BHT to 100 μM final concentration (1:1000). Mix well and keep on ice (*see* **Note 15**).
5. Aspirate PBS wash, and resuspend in 2 mL cold PBS (use a serological pipet to add to multiple samples, and resuspend with P1000 pipetman). Then add 8 mL C1 master mix. Mix by inverting the tube, and incubate on ice 10 min.
6. Pellet nuclei by centrifugating 1300 RFC for 10 min at 4 °C.
7. Aspirate, wash with 4 mL C1 master mix (resuspension is not critical), and centrifuge again. Samples can be stored at −20 °C or further processed.
8. Add 5 mL G2 buffer containing 100 μM each Deferoxamine and BHT. Vortex 5–10 s, add 95 μL Qiagen Protease, mix by inverting, and incubate 50 °C for 30 min (*see* **Note 16**).
9. While incubating, equilibrate Qiagen Tip 100 with 4 mL Qiagen QBT buffer.
10. Vortex samples (2–5 s) before adding to columns (*see* **Note 17**).
11. Load column, and allow to empty by gravity flow. Then wash twice with 7.5 mL Qiagen QC wash buffer.
12. Move columns to 15 mL falcon tubes (secure with a piece of tape), and elute with 5 mL Qiagen QF buffer.
13. Precipitate DNA with 3.5 mL isopropanol. *Mix well in tube until DNA precipitate is visible.*
14. Pellet DNA by centrifugation at 5250 RFC for 30 min at 4 °C.
15. Remove supernatant by inverting the tube; do not aspirate as this may disturb the pellet.
16. Wash with 3 mL 70% ethanol, and centrifuge again for 5 min.

17. Pour off ethanol, invert tubes over absorbent surface, and let air dry ~15–20 min (*see* **Note 18**).
18. Resuspend DNA with 100–150 μL TE, based on pellet size. Place sample in microfuge tubes, and incubate in a shaker overnight at 22 °C 900 rpm for ~16 h.
19. Place samples on ice or at −20 °C the next morning. Measure the concentration with a NanoDrop, and record the average from two to three readings (*see* **Note 19**).

3.5.3 Enzymatic Treatments for 8oxoG Detection

This procedure describes using FPG with and without S1 nuclease, and example Southern blots are shown in Fig. 3b, c. If using the OGG1 and APE1 combination, the process is similar, but OGG1 alone and APE1 alone reactions are included as controls (Fig. 3a, lanes 4, 5, 10, and 11).

1. For each sample we recommend loading 2–3 μg of total genomic DNA per well. Determine the DNA concentration each time after thawing.
2. Prepare a 40 μL reaction containing 7 μg genomic DNA in 1× NEB CutSmart buffer with 2.5 μL FPG (8 U FPG per 3 μg genomic DNA) on ice. Mix well by pipetting up and down and then centrifuge briefly to spin down. Incubate for 37 °C for 2 h (water bath recommended) (*see* **Note 20**).
3. Prepare a restriction enzyme master mix (RE MM) depending on the number of samples. Each sample requires 0.5 μL of each telomere restriction enzymes (*Alu*I, *Hph*I, *MnlI*I, *Hinf*I). Add 2 μL RE MM to the sample, mix well, and spin down. Incubate at 37 °C overnight for ~16 h (*see* **Note 21**).
4. Place samples on ice, or samples can be stored at −20 °C.
5. Split samples in half, and place ~20 μL each in two separate tubes.
6. Treat one half with S1 nuclease and the other mock treat as a control. Dilute the stock S1 (100 U/μL) with S1 diluent to 1 U/μL. Per reaction combine 2 U S1 (2 μL), 6 μL 5× S1 buffer, and 2 μL ddH_2O (30 μL total). For mock replace the 2 μL S1 nuclease with S1 diluent.
7. Mix well, spin down, and incubate at 37 °C for 1 h. Then ice and spin down. Add 6 μL of 6× loading dye (1× final concentration).
8. To determine if the genomic DNA digestion is complete, run 2–4 μL on a 0.8% agarose gel in 1× TAE containing 1:20000 dilution of ethidium bromide in gel and running buffer. Run at 90 V for approximately 1 h. Digested genomic fragments should migrate at ~100 bp and can be compared to a DNA ladder.

3.5.4 S Blot Analysis of Cleaved Telomere Restriction Fragments

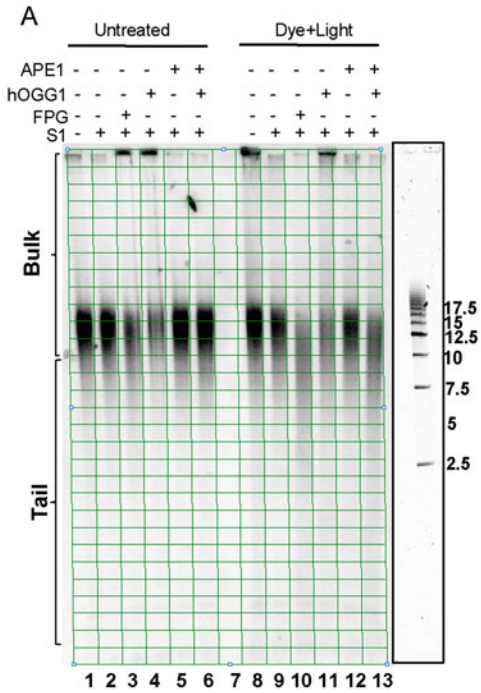
Any standard electrophoresis apparatus should be suitable for separating the intact and cleaved telomere restriction fragments. This protocol describes analysis by Pulse Field Gel Electrophoresis (PFGE), which is required for telomere restriction fragments in cell lines with long telomere lengths (≥20 kb) such as HeLa LT.

1. At least 3 h prior to PFGE, prepare 3 L of 0.5× TBE, and store at 4 °C after removing 200 mL to cast the gel. This is sufficient for the long (21 cm) gel casting tray.
2. Cast a 1% gel with megabase agarose lacking ethidium bromide.
3. Mix 2 μL DNA ladder with 1 μL 6× loading dye for 6 μL total.
4. Mix 30 μL of sample with 6 μL 6× loading dye for 36 μL total. We load less than the maximum reaction volume to ensure the same amount is loaded per well.
5. Run gel at 6 V, setting for 1 s initial switch and 6 s final switch, for 16 h. These settings are optimized for cells with average long telomere lengths. For cells with shorter telomeres, run for a shorter time (~8–10 h).
6. After ~40 min, turn on the pump first (to 60) and then the cooling module (set to 14 °C). Run overnight (*see* **Note 22**).
7. The next morning, preheat the vacuum drier to about 50 °C for ~10 min, before gel is finished running.
8. Place gel on a folded piece of Whatman paper, and cover with plastic wrap.
9. Dry gel under vacuum for 2 h at 50 °C. Monitor the temperature.
10. Gently remove gel from paper and plastic wrap by submerging in a water bath (*see* **Note 23**). The gel can be stored in a sealed plastic bag at 4 °C.
11. Place dried gel on nylon mesh, and roll up and place into hybridization bottle.
12. Dilute SYBR green 1:2000 in 50 mL 5× SSC, and add to hybridization bottle to stain gel in hybridization oven for 30–60 min at 42 °C.
13. Remove gel, and image SYBR green on a Typhoon imaging system using the blue/FAM setting at 375 PMT to visualize ladder.
14. Denature gel in a glass dish with 100 mL denaturation buffer for 15 min with gentle shaking at RT.
15. Wash with water for 10 min.
16. Neutralize for 15 min in neutralization buffer with gentle shaking at RT.

17. Place gel back on nylon mesh and in hybridization tube. Pre-hybridize by incubating with ~20 mL Church buffer for 20–30 min at 42 °C in hybridization oven.
18. Prepare Telomere probe. Combine 2 μL telomere G-strand probe (10 pmol/μL), 2 μL telomere C-strand probe (10 pmol/μL), 5 μL 10× Polynucleotide Kinase buffer, 2 μL Polynucleotide Kinase, 3 μL ^{32}P-γATP (10 μCi/μL), and 36 μL water. Mix well, spin down, and incubate 1 h at 37 °C. Inactivate at 65 °C for 20 min. Remove unincorporated ^{32}P-γ ATP with a G-25 MicroSpin column according to the manufacturer's instructions (*see* **Note 24**).
19. Add 20 μL of Telomere probe (see below) to 25 mL Church buffer. Remove pre-hybridization buffer and buffer containing the probe. Hybridize overnight in hybridization oven at 42 °C. Place a folded piece of filter paper in the oven in case of leaks.
20. The next day remove the buffer containing the probe (*see* **Note 25**).
21. Wash the gel in the hybridization bottle with the oven set to ~15 °C or RT. Wash for 10 min each with 2× SSC, 0.1× SSC containing 0.1% SDS, and 2× SSC.
22. Place gel in a plastic bag and seal. Expose for 2 h to overnight to a phosphor screen.

3.5.5 Analysis of S1 Blot

1. Visualize gel using a Typhoon phosphorimager.
2. Using ImageQuant (GE) or ImageJ, construct a grid with 30 rows and a column for each sample (Fig. 4a). A grid ensures all of the boxes are the same size and aligned properly, and 30 rows is standard for telomere length analyses [13]. Draw the top row around the wells, and extend until the bottom row reaches to the end of the gel.
3. Based on the image and the control sample lacking repair enzymes and S1 nuclease (lane 1 of Fig. 4a), identify rows that contain the intact telomere restriction fragments (bulk) (typically row 14 or 15). The boxes below contain cleaved telomere fragments (tail).
4. Include a column in a lane lacking sample to use for background subtraction from each row (see column 7 in Fig. 4a).
5. Use ImageQuant or ImageJ to quantify signal intensities in each box, and subtract background values. The percent of cleaved DNA (or percent in tail) is equal to the sum of tail rows (15–30) divided by the sum of all 30 rows. See example in Fig. 4b.



B

Dye+Light No Enzyme			Dye+Light +S1			Dye+Light FPG		
	Volume	-Background		Volume	-Background		Volume	-Background
1	904353.6	570564.08	1	661129.(	327340.3	1	399037.(	65247.77
2	565828.7	310629.65	2	469261.5	214062.8	2	310909.(	55710.15
3	496648.0	229172.42	3	415692.	148216.5	3	317915.(	50440.24
4	487544.9	201961.89	4	384138.(	98554.99	4	315796.(	30213.06
5	541604.3	247573.43	5	379504.(	85473.28	5	329577.(	35546.08
6	589342.8	278711.76	6	408366.(	97735.39	6	352317.(	41686.26
7	726108.9	378915.25	7	498123.(	150930	7	409638.(	62444.75
8	792433.5	427287.64	8	548774.(	183628.7	8	436426.(	71280.46
9	1510578.	1131148.3	9	1040753	661323	9	568170.(	188739.53
10	2993775.	2622453.1	10	2269889	1898567	10	1048200	676878.31
11	2592347.	2230212	11	2237556	1875421	11	1475214	1113078.7
12	1674908.	1382144	12	1613081	1320318	12	1513306	1220542.3
13	1059832.	794431.95	13	1082077	816677.2	13	1307218	1041817.5
14	748323.4	485564.49	14	769384.(	506625.5	14	1076914	814155.81
15	541753.2	304929.43	15	569749.(	332925.2	15	863661.(	626837.51
16	422389.1	212282.24	16	409877.(	199770.6	16	666616.(	456510.05
17	362586.3	157148.6	17	363821.(	158383.8	17	550336.(	344899.07
18	308696.5	112789.3	18	315559.(	119652.1	18	539509.(	343602.07
19	309681.0	133161.69	19	302684.(	126165.6	19	461100.(	284580.71
20	270628.8	96996.41	20	287145.(	113512.8	20	407341.(	233709.55
21	278172.1	86975.95	21	261766.(	70570.2	21	342199.(	151002.85
22	241854.0	59067.51	22	233743.(	50957.03	22	315121.(	132334.52
23	248885.8	65167.6	23	228796.(	45078.66	23	317856.(	134138.13
24	243034.9	59216.44	24	226080.(	42262.09	24	309481.(	125662.6
25	237353.0	37564.67	25	231602.(	31814.04	25	317303.(	117515.2
26	253428.4	36179.45	26	245408.(	28159.7	26	325091.(	107842.63
27	307180.3	93421.97	27	301250.(	87492.24	27	362642.(	148883.92
28	266605.4	67811.61	28	275019.(	76225.32	28	316520.(	117727.01
29	282979.2	88007.92	29	304100.(	109129.2	29	331410.(	136439.54
30	312442.7	102608.55	30	297617.(	87783.22	30	316779.(	106944.86
	Total	13004099		Total	10064755		Total	9036411.2
	Bulk	11290770		Bulk	8384874		Bulk	5467781
	Tail	1713329.3		Tail	1679882		Tail	3568630.2
	%Bulk	0.868247		%Bulk	0.833093		%Bulk	0.6050832
	%Tail	0.131753		%Tail	0.166907		%Tail	0.3949168

Fig. 4 Example of S blot quantification using ImageQuant. A grid of 13 columns and 30 rows was applied to the blot shown in Fig. 3a to measure radioactivity signal intensity in each box. The Sybr Green-stained DNA ladder is shown next to the blot assists in alignment. The spreadsheet shows an example of the data from lanes 8–10 showing the uncorrected signal intensity (volume) and corrected for background (– Background, subtracted values from lane 7). The "bulk" intact telomeres is the sum of rows 1–14 (yellow), and the "tail" (cleaved telomeres) is the sum of rows 15–30 (red)

4 Notes

1. Synthesis of the MG-ester and MG-2I dyes are described in [9, 11]. Alternatively, you can request these dyes from Dr. Marcel Bruchez at Carnegie Mellon University upon completion of an MTA.
2. In our hands, delivery of the FAP-mCER-TRF1 transgene by transduction with lentivirus appears to yield more homogeneous transgene expression in the cell population, compared to transfection reagents, and is needed for difficult-to-transfect cell lines.
3. Plasmid concentrations of $\geq$500 μg/mL are best for transfection into mammalian cells.
4. We culture "cells of interest" for infection with lentivirus to obtain stable expression of FAP-mCER-TRF1, in a low oxygen incubator set to 5% O_2 to reduce background oxidative damage.

5. The media containing lentivirus can also be stored at −4 °C overnight or −80 °C for future use. The transfected HEK293T cells can be discarded at this time. Alternatively, harvest the HEK293T cells and Western blot for expression of FAP-mCER-TRF1. This will confirm transfection of the HEK293T cells was successful.
6. Cell lines differ in sensitivity to G418. Therefore, a dose response needs to be determined for the cell line of interest to identify the concentration that kills cells that were not infected with the lentivirus.
7. Even though the lentiviral construct should integrate into the genome, we have noticed reduced expression of the transgene over time when selection pressure is not maintained by growing in G418 containing media.
8. We also seed a 96-well plate with 200 μL (i.e., 2 cells) per well to ensure we recover wells with an expanded clone from a surviving cell, and for harder to clone cells (i.e., BJ-5ta), we have added up to 16 cells per well.
9. For this experiment, DAPI cannot be used to stain the DNA for visualization of the nuclei because it can interfere with mCerulean imaging since they both emit in the blue spectral region. To visualize the nuclei, perform IF for nuclear proteins such as Lamin A/C. We have noticed mCerulean can bleach rapidly, so IF should be performed in the dark. Store slides in the dark.
10. Co-localization with a telomeric protein by IF is also possible. When not performing FISH, it is best to use a secondary antibody against anti-GFP (which binds mCerulean) that is conjugated to a green fluorophore since the mCerulean spectrum overlaps partially with the FITC and Alexa 488 spectra. Then the telomeric protein can be detected with a red fluorophore to avoid cross-talk between channels. When performing FISH, the denaturation steps destroy the mCerulean signal so cross-talk is not a concern.
11. Protect from light by covering dishes with foil, and work in the dark for rest of the protocol to preserve the MG-ester dye and prevent bleaching.
12. If combining MG-ester imaging in fixed cells with IF to confirm telomere localization, we recommend using a secondary antibody conjugated to a green fluorophore such as Alexa 488.
13. Expose a dish with less than 90% confluent cells if a longer recovery period is desired.
14. Light sources that illuminate from above or below are both effective. However, when illuminating from above, we remove the cell culture dish lid and expose to light in the cell culture

hood to maintain sterile conditions. It is important to ensure that the light exposures do not heat the cells. This can be tested for with a thermometer and can be mitigated using fans or limiting the duration of exposure. Dr. Marcel Bruchez at Carnegie Mellon University can be contacted for advice on light box construction.

15. Antioxidants Deferoxamine and BHT are included to prevent spontaneous oxidation of guanine during DNA isolation and processing.
16. Resuspension and homogenization are critical at this point. Otherwise digestion with the protease will not be complete and will potentially clog the purification column.
17. A pipet tip can be used to remove any "junk" floating in the solution to prevent clogging the column.
18. Do not overdry the pellet; this will make it very difficult to resuspend. Gauge dryness by smelling the tube, and if it still smells like ethanol, then let dry a bit longer.
19. Do not allow the samples to remain at RT for more than a day. It is best to measure the DNA concentrations on the day of enzymatic treatments after vortexing and to process the DNA for the experiment within a couple weeks due to concerns about 8oxoG stability.
20. If using OGG1 and APE1, instead of FPG, conduct 20 μL reactions each containing 125 nM OGG1, 3.3 U/μg DNA APE1, or both 125 nM OGG1 and 3.3 U/μg DNA APE1. Then proceed with S1 nuclease treatment.
21. Digestion times can be optimized and may be as short as 6 h. These restriction enzymes work well for releasing telomere restriction fragments from human genomic DNA. Mouse DNA can be digested with *Mbo*I and *Alu*I.
22. This prevents the pump from displacing any DNA in the wells before the DNA migrates into the gel.
23. The gel usually sticks to the plastic wrap. Therefore, lift the gel from the paper and then resubmerge in the water. Tap the plastic wrap and gel so the gel detaches. As the gel detaches, let the water flow in between the gel and the wrap to facilitate the separation. If the gel sticks to the paper instead, it will easily detach upon immersion in water. The gel can be stored in a sealed plastic bag at 4 °C.
24. Special protocols and safety procedures should be followed when using radioactive material and when disposing of solid and liquid radioactive waste. G-25 spin columns can occasionally dry out. If after vortexing the resin and spinning down the buffer by centrifugation, the collection tube contains less than ~50 μL buffer, then the column has probably dried out.

25. Probe can be saved in the refrigerator and reused once. Store at RT if reusing the same day; otherwise, store in a refrigerator designated for radioactive reagents. When reusing the probe, incubate it in the hybridization oven at 42 °C for at least 30 min to redissolve the SDS.

References

1. d'Adda di Fagagna F, Reaper PM, Clay-Farrace L, Fiegler H, Carr P, Von Zglinicki T et al (2003) A DNA damage checkpoint response in telomere-initiated senescence. Nature 426(6963):194–198. https://doi.org/10.1038/nature02118
2. de Lange T (2018) Shelterin-mediated telomere protection. Annu Rev Genet 52:223–247. https://doi.org/10.1146/annurev-genet-032918-021921
3. von Zglinicki T (2002) Oxidative stress shortens telomeres. Trends Biochem Sci 27(7):339–344. https://doi.org/10.1016/s0968-0004(02)02110-2
4. Barnes RP, Fouquerel E, Opresko PL (2019) The impact of oxidative DNA damage and stress on telomere homeostasis. Mech Ageing Dev 177:37–45. https://doi.org/10.1016/j.mad.2018.03.013
5. Ahmed W, Lingner J (2017) Impact of oxidative stress on telomere biology. Differentiation 99:21–27. https://doi.org/10.1016/j.diff.2017.12.002
6. Zhang J, Rane G, Dai X, Shanmugam MK, Arfuso F, Samy RP et al (2016) Ageing and the telomere connection: an intimate relationship with inflammation. Ageing Res Rev 25:55–69. https://doi.org/10.1016/j.arr.2015.11.006
7. Cooke MS, Evans MD, Dizdaroglu M, Lunec J (2003) Oxidative DNA damage: mechanisms, mutation, and disease. FASEB J 17(10):1195–1214. https://doi.org/10.1096/fj.02-0752rev
8. Fouquerel E, Barnes RP, Uttam S, Watkins SC, Bruchez MP, Opresko PL (2019) Targeted and persistent 8-oxoguanine base damage at telomeres promotes telomere loss and crisis. Mol Cell 75(1):117–30 e6. https://doi.org/10.1016/j.molcel.2019.04.024
9. He J, Wang Y, Missinato MA, Onuoha E, Perkins LA, Watkins SC et al (2016) A genetically targetable near-infrared photosensitizer. Nat Methods 13(3):263–268. https://doi.org/10.1038/nmeth.3735
10. Agnez-Lima LF, Melo JT, Silva AE, Oliveira AH, Timoteo AR, Lima-Bessa KM et al (2012) DNA damage by singlet oxygen and cellular protective mechanisms. Mutat Res Rev Mutat Res 751(1):15–28. https://doi.org/10.1016/j.mrrev.2011.12.005
11. Liang P, Kolodieznyi D, Creeger Y, Ballou B, Bruchez MP (2020) Subcellular singlet oxygen and cell death: location matters. Front Chem 8:592941. https://doi.org/10.3389/fchem.2020.592941
12. Xie W, Jiao B, Bai Q, Ilin VA, Sun M, Burton CE et al (2020) Chemoptogenetic ablation of neuronal mitochondria in vivo with spatiotemporal precision and controllable severity. elife 9:e51845. https://doi.org/10.7554/eLife.51845
13. Jenkins FJ, Kerr CM, Fouquerel E, Bovbjerg DH, Opresko PL (2017) Modified terminal restriction fragment analysis for quantifying telomere length using in-gel hybridization. J Vis Exp (125):56001. https://doi.org/10.3791/56001

Chapter 10

Qualitative and Quantitative Analysis of DNA Cytidine Deaminase Activity

Rachel DeWeerd and Abby M. Green

Abstract

The human genome encodes eleven DNA cytidine deaminases in the AID/APOBEC family, which encompass endogenous roles ranging from genetic diversification of the immunoglobulin locus to virus restriction. All AID/APOBEC functions are enabled by their catalyzation of cytidine deamination resulting in mutations and DNA damage. When acting aberrantly, deaminases can cause off-target mutations in the cellular genome resulting in somatic mutations, DNA damage, and genome instability. An association between cytidine deaminase-induced mutations and human cancers has been recognized over the last decade, necessitating assays for investigation of intracellular deaminase activity. Here we present two assays for deamination activity which enable in vitro evaluation of in vivo events. We define both a qualitative assay to confirm deaminase activity within cells as well as a quantitative assay for granular evaluation and comparisons of deamination activity across different cell populations or experimental conditions. The two procedures are customizable assays which can easily be adapted to individual labs and experiments.

Key words Deaminase, Cytidine deaminase, APOBEC, AID, DNA base editors

1 Introduction

Deaminase activity on the genome can result in base mutations, DNA breaks, and translocations [1–3]. The AID/APOBEC family of cytidine deaminases act on single-stranded DNA (ssDNA) or RNA to restrict viral infection and retrotransposition and, in the case of AID, promote antibody diversity and class switch recombination [4–8]. AID/APOBEC enzymes catalyze the hydrolysis of the NH_2 group of cytidine bases, causing a cytidine to uracil transition (Fig. 1). The uracil base can be removed through the base excision repair pathway or can be used as a template for replication; both mechanisms of uracil processing can cause mutations in viral or genomic DNA. Mutational patterns characteristic of deamination events have been identified in many subsets of cancer, suggesting that APOBEC activity plays a significant role in the mutational landscape of human cancers [9–11]. Because of the

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_10,

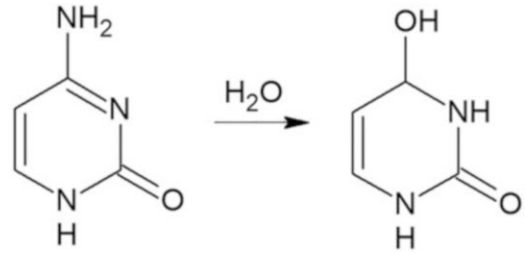


Fig. 1 Cytidine deamination to uracil. Cytidine deaminases catalyze a hydrolytic reaction, removing the amine group from cytidine, leaving a hydroxyl group in its place. The resulting nucleotide is a uracil

activity of cytidine deaminases on the genome, assays to confirm and quantify catalytic activity in laboratory models are required [12, 13].

Here we describe two assays which measure deaminase activity in cell lysates using ssDNA oligonucleotides that contain a single cytidine base and a fluorophore. In both assays, incubation of cell lysates with the oligonucleotide results in deamination of the cytidine base by active deaminases. Addition of uracil DNA glycosylase (UDG) removes the resulting uracil base, leaving an abasic site, which is then cleaved by sodium hydroxide (NaOH). The qualitative assay relies on a FAM-containing oligonucleotide which, after deamination and abasic site cleavage, reveals two distinct bands by gel electrophoresis (Fig. 2, left). In the quantitative assay, the oligonucleotide contains both a 5′ FAM fluorophore and a 3′ TAMRA quencher. Unmodified substrate remains quenched, but a deaminated and cleaved product emits fluorescence that can be detected by fluorimetry (Fig. 2, right). The qualitative assay offers a simple gel-based analysis and is ideal for determination of deamination activity in cell lysates (Fig. 3). The quantitative assay allows for more precise measurements of relative deaminase activity between samples (Fig. 4). Together, these two experimental procedures provide flexible systems which can be utilized for investigation of DNA deaminase activity in live cells.

2 Materials

Unless otherwise indicated, all stock components of buffers are prepared in deionized water.

2.1 Cell Lysis

1. Lysis buffer: 50 mM Tris HCl pH 7.4, 150 mM NaCl, 0.1% Triton X-100, 0.5% sodium deoxycholate, 1 mM sodium orthovanadate, 20 mM sodium fluoride. Add water to total desired volume.

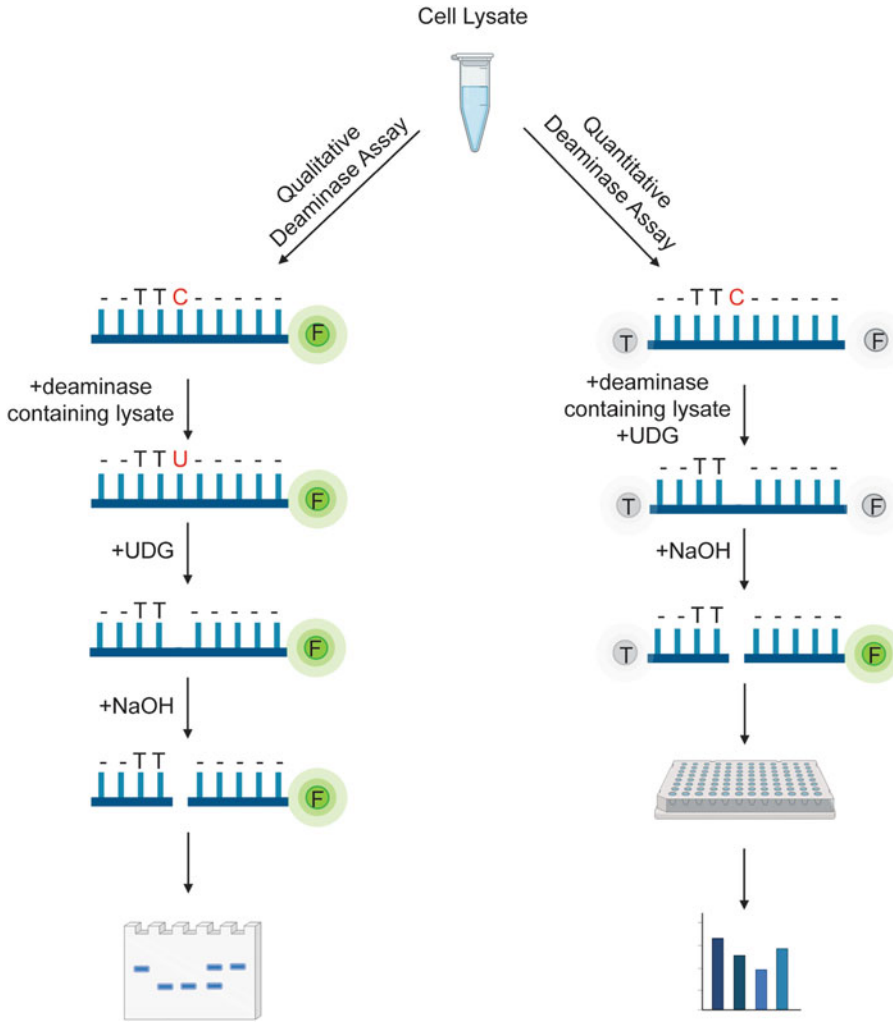


Fig. 2 Qualitative and quantitative deaminase assay work flow. Cells are lysed in the same manner for both assays. Qualitative (left): cell lysate is incubated with an oligonucleotide containing a single cytidine base and a FAM fluorophore, followed by UDG and NaOH. The resulting product is visualized by gel electrophoresis. Quantitative (right): cell lysate is incubated with UDG and an oligonucleotide containing a single cytidine base and both a FAM fluorophore (F) and TAMRA quencher (T), followed by NaOH and HCl. Fluorescence is measured using a fluorometric plate reader. Results generated are numerical and can be displayed as a graph or chart

2. Protease inhibitors: 0.1 M phenylmethylsulfonyl fluoride (PMSF) and protease inhibitor cocktail (PI).

2.2 Qualitative Deaminase Reaction

1. 10× Incubation buffer: 20 mM MES and 0.1% Tween 20 prepared in water.
2. Formamide loading dye: 5 mL formamide, 0.6 g NaOH, 100 µL 0.5 M EDTA, and enough bromophenol blue to create a visible dye.

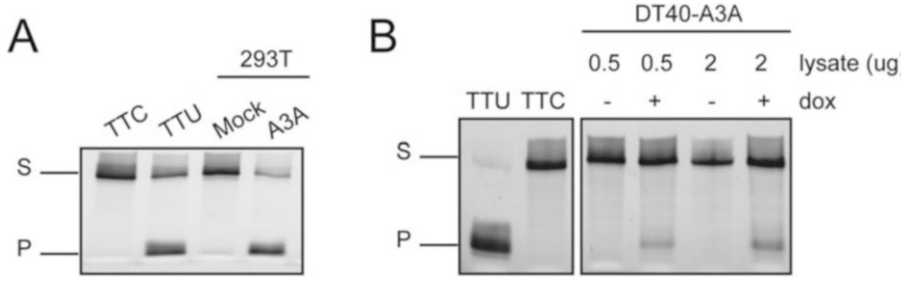


Fig. 3 Qualitative deaminase gels. (**a**) 293T cells were transfected with transfection reagent alone (mock) or with the cytidine deaminase APOBEC3A (A3A). TTC denotes the negative control for the full length substrate (S), TTU the positive control for the deaminated, cleaved product (P). (**b**) DT40-A3A cells were treated with doxycycline to induce A3A expression 72 h prior to lysis. Lysate containing 0.5 μg and 2 μg of protein was used

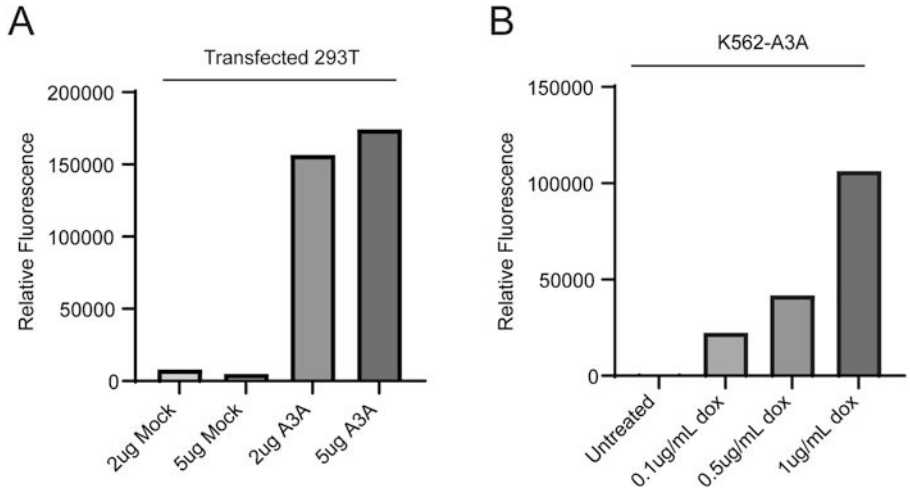


Fig. 4 Quantitative fluorescence deaminase. (**a**) 293T cells were transfected as described in panel 3a. Lysate containing 2 and 5 μg of lysate is shown. (**b**) K562-A3A cells were treated with increasing amounts of doxycycline to induce A3A expression 24 h prior to lysis

3. 25 μM 5′ FAM substrate oligonucleotide: an oligonucleotide with a 5′ FAM tag should be obtained. Individual oligonucleotide sequences can be adjusted for the context preferences of different deaminases [4, 14], but there should only be one cytidine base in the sequence. This will allow for consistent and predictable product band sizes. To prevent exonuclease degradation, the 3′ end should be modified for example with an inverted deoxythymidine (dT).

   Example: 5′-FAM TGAGGAATGAAGTTGATT**C**AAATGTG ATGAGGTGA.

4. 25 μM 5′ FAM positive control oligonucleotide: an oligonucleotide with a 5′ FAM tag should be created with an identical sequence to the substrate oligonucleotide, with the exception of a uracil base in place of the cytidine.

   Example: 5′-FAM TGAGGAATGAAGTTGATT**U**AAATGTG ATGAGGTGA.

5. Uracil DNA glycosylase (5000 U/mL).

2.3 Running Buffer and Urea-Acrylamide Gel

1. 10× TBE: 1 M Tris base, 1 M boric acid, and 0.02 M EDTA pH 8.
2. 1× TBE: 1:10 dilution of 10× TBE in water.
3. Urea-acrylamide master mix: 7 M urea and 20% acrylamide (19:1) prepared in 1× TBE.
4. Urea-acrylamide gel: 1000:10:1 urea-acrylamide master mix, ammonium persulfate (APS), and TEMED.

   Example: 12 mL urea-acrylamide master mix, 120 μL APS, and 12 μL TEMED.

2.4 Quantitative Deaminase Reaction

1. 50 mM Tris HCl pH 7.4.
2. 10 mM EDTA pH 8.
3. 25 μM 5′ FAM 3′ TAMRA substrate oligonucleotide: an oligonucleotide with a 5′ FAM and a 3′ TAMRA quencher. Individual oligonucleotide sequences can be adjusted for the context preferences of different deaminases, but there should only be one cytidine base in the sequence.

   Example: 5′-FAM AAATT**C**AGAGAGAGAATGTGA TAMRA-3′.

4. 25 μM 5′ FAM 3′ TAMRA positive control oligonucleotide: an oligonucleotide with a 5′ FAM and a 3′ TAMRA quencher should be created with an identical sequence to the substrate oligo, with the exception of a uracil base in place of the cytidine.

   Example: 5′-FAM AAATT**U**AGAGAGAGAATGTGA TAMRA-3′.

5. Uracil DNA glycosylase (5000 U/mL).
6. 4 N NaOH.
7. 4 N HCl.
8. 2 M Tris HCl pH 7.4.

2.5 Additional Equipment

1. Sonicator.
2. Fluorescent gel imager.
3. Fluorescent plate reader.

3 Methods

The assays presented here can be performed in any type of cultured or primary cell under a variety of conditions. Plating and treatment schemes should be adjusted for individual experiments.

3.1 Lysis

1. Pellet cells and remove residual supernatant.
2. Add protease inhibitors to lysis buffer immediately prior to lysis.
3. Resuspend cell pellet in a volume of lysis buffer that is equal to twice the volume of the cell pellet (*see* **Note 1**). Mix by pipette until homogenous.
4. Incubate lysate on ice for 10 min.
5. Sonicate 30 s (or more to minimize viscosity).
6. Centrifuge at 16,000 × *g* at 4 °C for 10 min. Save the supernatant, and discard the pellet.
7. Determine protein concentration (*see* **Note 2**).

3.2 Qualitative Deaminase Reactions

Aim to use 2 μg of cell lysate per reaction mixture (*see* **Note 3**).

1. In 1.5 mL tubes, prepare test reactions as follows (component; volume):
   - ~2 μg lysate; volume varies but must be 7.5 μL or less.
   - 25 μM 5′ FAM substrate oligonucleotide; 1 μL (final concentration 2.5 μM).
   - 10× Incubation buffer; 1 μL.
   - 50 mM EDTA pH 8; 0.5 μL.
   - Water; amount to total volume of 10 μL.
2. In 1.5 mL tubes, prepare separate positive and negative control reactions for full length substrate and cleaved product bands as follows (component; volume):
   - Lysis buffer; 1 μL.
   - 25 μM 5′ FAM substrate oligonucleotide OR 25 μM 5′ FAM positive control nucleotide; 1 μL (final concentration 2.5 μM)
   - 10× Incubation buffer; 1 μL
   - 50 mM EDTA pH 8; 0.5 μL
   - Water; 6.5 μL (total reaction volume of 10 μL).
3. Incubate test and control reactions at 37 °C for 2 h.
4. Add 0.5 μL (2.5 U) uracil DNA glycosylase.
5. Incubate at 37 °C for 15 min.

6. Add 10 μL formamide loading dye containing sodium hydroxide.
7. Incubate at 95 °C for 15 min.
8. Hold samples at 55 °C until use (*see* **Note 4**).

3.3 Qualitative Assay Gel Electrophoresis

1. Place gel in electrophoresis chamber. Pour 1× TBE buffer directly over the prepared 1.5 mm urea-acrylamide gel. Pipette 1× TBE directly into the wells to flush the wells before loading samples.
2. Fill gel chamber with 1× TBE as running buffer.
3. Load 10 μL of sample per well. Add 10 μL of formamide loading dye to all wells which do not contain sample (*see* **Note 5**).
4. Run at 40 watts until loading dye reaches the bottom of the gel (*see* **Note 6**).
5. Image the gel using the fluorescein channel on a fluorescent gel imager (excitation of FAM ~493 nm, emission ~517 nm).

3.4 Quantitative Deaminase Reactions

Following lysis as described in Subheading 3.1, reaction mixtures should be prepared in a clear 96-well, flat bottom plate. Aim to use 2 μg of cell lysate per reaction mixture (*see* **Note 3**).

1. Reaction master mix—create a stock of 50 mM Tris HCl pH 7.4, 10 mM EDTA pH 8, and 2.5 U of UDG per well. Prepare enough reaction mixture for 49 μL/well, including controls.
2. Prepare each sample well as follows (component; volume):
   - ~2 μg lysate; volume varies, must be 10 μL or less.
   - 25 μM 5′ FAM, 3′ TAMRA substrate oligonucleotide; 1 μL (final concentration 420 nM).
   - Reaction master mix; 49 μL.
   - Water; amount to total volume of 60 μL.
3. Prepare separate positive and negative control wells as follows (component; volume):
   - RIPA buffer; 1 μL.
   - 25 μM 5′ FAM, 3′ TAMRA substrate oligonucleotide OR 5′ FAM 3′ TAMRA positive control oligonucleotide; 1 μL (final concentration 420 nM)
   - Reaction master mix; 49 μL.
   - Water; 9 μL (total volume of 60 μL).
4. Incubate at 37 °C for 1.5 h.
5. Add 3 μL 4 N NaOH per well.
6. Incubate at 37 °C for 30 min.

7. Neutralize with 3 μL 4 N HCl and 27 μL 2 M Tris pH 7.9 per well.
8. Cool plate at 4 °C for several minutes before proceeding.

3.5 Quantitative Deaminase Fluorescence Measurements

1. Using a plate reader, read fluorescence with excitation set to ~493 and emission ~517 to detect the FAM fluorophore.
2. Subtract background fluorescence of the negative control wells from each sample (*see* **Note 7**).

4 Notes

1. Estimate pellet volume by pipetting liquid into an empty tube of the same size until the liquid volume is equal to the pellet size.
2. In all experiments shown here, protein concentration was determined by Bradford, although any sensitive assay for protein concentration measurement will suffice.
3. The protein concentration that is added to deamination assay reactions can be adjusted based on cell type, deaminase expression levels, or individual detectability. For the gels shown here, 2 μg per sample was sufficient for activity detection.
4. Holding at 55 °C prevents precipitation of the reaction. If not used on the same day, samples can be maintained at 4 °C for several days. Avoid precipitation by analyzing soon after the reaction has been completed.
5. To prevent uneven sample running during electrophoresis, load samples toward the middle of the gel, and fill all outer wells with an equal volume of formamide loading dye.
6. Depending on the size of the oligonucleotide, the deaminated product band may run close to the loading dye. It is important to run the samples toward the bottom of the gel to ensure separation between substrate and product bands, being careful not to run the samples off the gel.
7. The uracil containing 5′ FAM 3′ TAMRA oligonucleotide is not necessary for completing any calculations in quantitative assay, but does provide a positive control for assay efficiency.

Acknowledgments

Thanks to all members of the Green lab for advice and manuscript review. Special thanks to Rahul Kohli and Emily Schutsky for their work developing reagents for the qualitative deaminase assay. Figure 1 was created through ChemSketch, and the work flow in Fig. 2 was created in BioRender.

References

1. Casellas R, Basu U, Yewdell WT, Chaudhuri J, Robbiani DF, Di Noia JM (2016) Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. Nat Rev Immunol 16(3):164–176. https://doi.org/10.1038/nri.2016.2
2. Green AM, Weitzman MD (2019) The spectrum of APOBEC3 activity: from anti-viral agents to anti-cancer opportunities. DNA Repair (Amst) 83:102700. https://doi.org/10.1016/j.dnarep.2019.102700
3. Taylor BJ, Nik-Zainal S, Wu YL, Stebbings LA, Raine K, Campbell PJ, Rada C, Stratton MR, Neuberger MS (2013) DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. elife 2:e00534. https://doi.org/10.7554/eLife.00534
4. Chen H, Lilley CE, Yu Q, Lee DV, Chou J, Narvaiza I, Landau NR, Weitzman MD (2006) APOBEC3A is a potent inhibitor of adeno-associated virus and retrotransposons. Curr Biol 16(5):480–485. https://doi.org/10.1016/j.cub.2006.01.031
5. Okazaki I, Yoshikawa K, Kinoshita K, Muramatsu M, Nagaoka H, Honjo T (2003) Activation-induced cytidine deaminase links class switch recombination and somatic hypermutation. Ann N Y Acad Sci 987:1–8. https://doi.org/10.1111/j.1749-6632.2003.tb06027.x
6. Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D (2003) Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature 424(6944):99–103. https://doi.org/10.1038/nature01709
7. Harris RS, Bishop KN, Sheehy AM, Craig HM, Petersen-Mahrt SK, Watt IN, Neuberger MS, Malim MH (2003) DNA deamination mediates innate immunity to retroviral infection. Cell 113(6):803–809. https://doi.org/10.1016/s0092-8674(03)00423-9
8. Narvaiza I, Linfesty DC, Greener BN, Hakata Y, Pintel DJ, Logue E, Landau NR, Weitzman MD (2009) Deaminase-independent inhibition of parvoviruses by the APOBEC3A cytidine deaminase. PLoS Pathog 5(5):e1000439. https://doi.org/10.1371/journal.ppat.1000439
9. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, DJ MB, Bignell GR, Cooke SL, Shlien A, Gamble J, Whitmore I, Maddison M, Tarpey PS, Davies HIR, Papaemmanuil E, Stephens PJ, McLaren S, Butler AP, Teague JW, Jonsson G, Garber JE, Silver D, Miron P, Fatima A, Boyault S, Langerod A, Tutt A, Martens JW, Aparicio SA, Borg A, Salomon AV, Thomas G, Borresen-Dale AL, Richardson AL, Neuberger MS, Futreal PA, Campbell PJ, Stratton MR, Breast Cancer Working Group of the International Cancer Genome Consortium (2012) Mutational processes molding the genomes of 21 breast cancers. Cell 149(5):979–993. https://doi.org/10.1016/j.cell.2012.04.024
10. Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, Kiezun A, Kryukov GV, Carter SL, Saksena G, Harris S, Shah RR, Resnick MA, Getz G, Gordenin DA (2013) An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet 45(9):970–976. https://doi.org/10.1038/ng.2702
11. Leonard B, Hart SN, Burns MB, Carpenter MA, Temiz NA, Rathore A, Vogel RI, Nikas JB, Law EK, Brown WL, Li Y, Zhang Y, Maurer MJ, Oberg AL, Cunningham JM, Shridhar V, Bell DA, April C, Bentley D, Bibikova M, Cheetham RK, Fan JB, Grocock R, Humphray S, Kingsbury Z, Peden J, Chien J, Swisher EM, Hartmann LC, Kalli KR, Goode EL, Sicotte H, Kaufmann SH, Harris RS (2013) APOBEC3B upregulation and genomic mutation patterns in serous ovarian carcinoma. Cancer Res 73(24):7222–7231. https://doi.org/10.1158/0008-5472.CAN-13-1753
12. Schutsky EK, Nabel CS, Davis AKF, DeNizio JE, Kohli RM (2017) APOBEC3A efficiently deaminates methylated, but not TET-oxidized, cytosine bases in DNA. Nucleic Acids Res 45(13):7655–7665. https://doi.org/10.1093/nar/gkx345
13. Thielen BK, McNevin JP, McElrath MJ, Hunt BV, Klein KC, Lingappa JR (2010) Innate immune signaling induces high levels of TC-specific deaminase activity in primary monocyte-derived cells through expression of APOBEC3A isoforms. J Biol Chem 285(36):27753–27766. https://doi.org/10.1074/jbc.M110.102822
14. Chan K, Roberts SA, Klimczak LJ, Sterling JF, Saini N, Malc EP, Kim J, Kwiatkowski DJ, Fargo DC, Mieczkowski PA, Getz G, Gordenin DA (2015) An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat Genet 47(9):1067–1072. https://doi.org/10.1038/ng.3378

Chapter 11

Characterization of DNA-PK-Bound End Fragments Using GLASS-ChIP

Rajashree A. Deshpande and Tanya T. Paull

Abstract

Endonucleolytic cleavage of DNA ends by the human Mre11-Rad50-Nbs1 (MRN) complex occurs in a manner that is promoted by DNA-dependent protein kinase (DNA-PK). A method is described to isolate DNA-PK-bound fragments released from chromatin in human cells using a modified Gentle Lysis and Size Selection chromatin immunoprecipitation (GLASS-ChIP) protocol. This method, combined with real-time PCR or next-generation sequencing, can identify sites of MRN endonucleolytic cutting adjacent to DNA-PK binding sites in human cells.

Key words Double-strand breaks, DNA-PK, MRN complex, DNA repair

1 Introduction

In eukaryotes, double-strand breaks in DNA are repaired through two pathways, non-homologous end-joining (NHEJ) and homologous recombination (HR), each of which requires a dedicated set of repair factors [1, 2]. The Mre11/Rad50/Nbs1 (MRN) complex and the CtIP protein are central to the HR pathway because of their importance in 5′ resection: the removal of a few hundred basepairs (or more) from the 5′ strand at a double-strand break, a process required for subsequent Rad51-mediated strand invasion into an intact template for replication-based repair [3–5]. In contrast, NHEJ does not rely on a separate template but instead resolves breaks by direct ligation of the broken ends, often with small deletions or occasional insertions at the break points. The DNA-dependent protein kinase catalytic subunit (DNA-PKcs) is an important regulator of this process along with the Ku70/Ku80 (Ku) heterodimeric complex that recruits the kinase to DNA ends [6]. The fate of double-strand breaks with respect to these pathways has often been presented as a competition between the MRN and DNA-PK complexes; however, recent work suggests that the

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_11,

presence of DNA-PK on DNA ends actually promotes the initial endonucleolytic processing of DNA by MRN [7]. This result as well as other observations showing that NHEJ proteins arrive at double-strand break sites earlier than other factors [8–12] and results showing a fast NHEJ phase of repair preceding HR [13, 14] suggest that the decision process involves NHEJ factor recruitment followed by HR factors if the NHEJ pathway is unsuccessful.

We previously demonstrated CtIP-dependent MRN endonucleolytic cutting of DNA-PK-bound ends in vitro with purified recombinant components using ensemble and single-molecule approaches [7]. In that work we observed MRN/CtIP-dependent removal of DNA-PK from DNA ends and showed that this requires all components of MRN, CtIP, and DNA-PK (DNA-PKcs and Ku). We also demonstrated that fragments of DNA bound by DNA-PK are generated by MRN in human cells using an inducible ER-AsiSI system [15] to induce DSBs at specific sites in genomic DNA. In this DSB inducible via AsiSI (DIvA) system, the restriction enzyme AsiSI is fused to the estrogen receptor hormone-binding domain and can be induced to enter the nucleus where it generates DSBs at sequence-specific sites (5′-GCGATCGC-3′) upon 4-hydroxytamoxifen (4-OHT) treatment. There are approximately 1000 AsiSI restriction sites in the human genome although approximately 100 to 200 DSBs are actually generated by the AsiSI enzyme in human cells [16]. Because DNA-PK is bound to the ends that are cleaved by MRN, it is possible to recover these ends by cross-linking and immunoprecipitation of DNA-PK [7]. In contrast to standard chromatin immunoprecipitation (ChIP), these fragments are released from the chromatin and thus can be recovered from the supernatant rather than the pellet fraction that contains the vast majority of the chromatin (Fig. 1). This protocol describes the preparation of cells for this procedure, cross-linking, immunoprecipitation, and size selection of fragments to efficiently purify the DNA-PK-bound DNA fragments, in a protocol called Gentle Lysis and Size Selection ChIP (GLASS-ChIP). This assay, excluding the time for making appropriate cell lines and cell treatment and harvesting, can be completed within 3 to 4 days.

2 Materials

1. NU7441 (resuspend at 5 mM in DMSO; store at 4 °C in 1 ml aliquots or at −20 °C for long-term storage).
2. 4OHT (resuspend in methanol or ethanol; dilute to 600 μM for a 1000× stock).
3. Phosphate-buffered saline (PBS), 4 °C.
4. Formaldehyde (37% stock): (*see* **Note 1**).

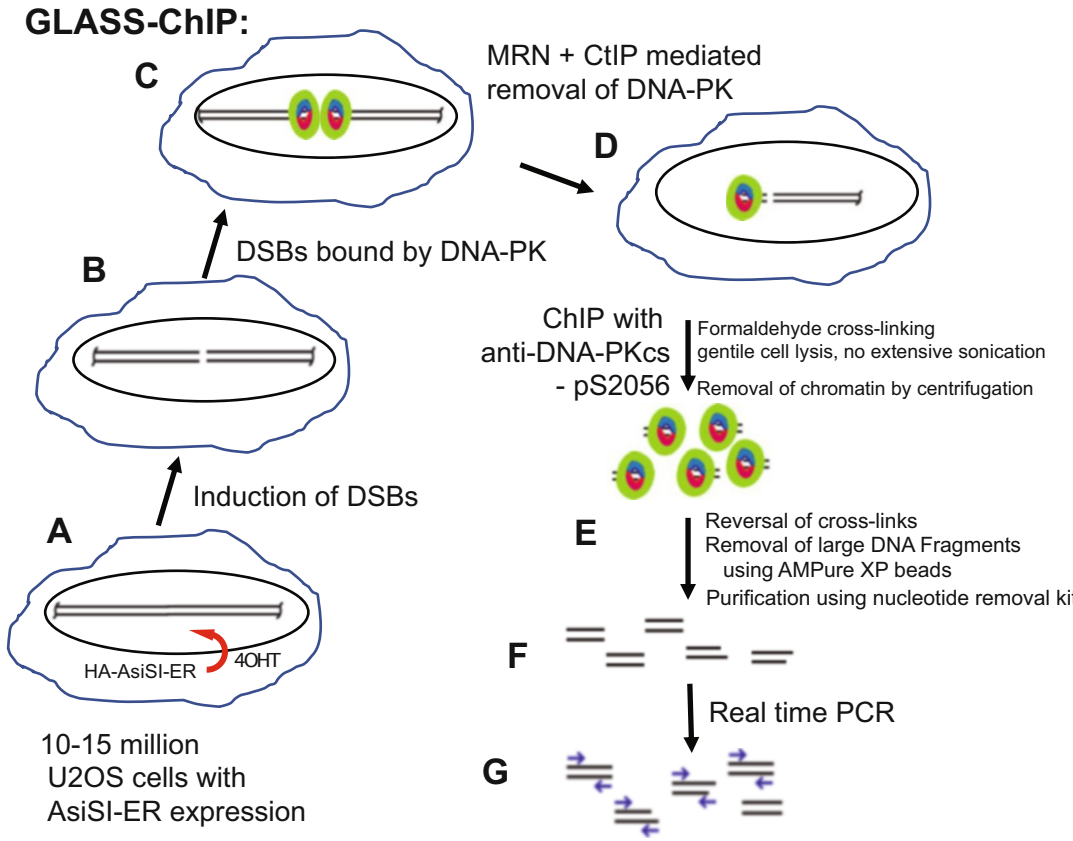


Fig. 1 Overview of GLASS-ChIP method. (**a**) U2OS cells with HA-AsiSI-ER expression [15], 10–15 million cells per treatment. 4OHT addition causes nuclear localization of AsiSI. (**b**) AsiSI cuts (DSBs) are induced by addition of 4OHT. (**c**) DSBs are bound by the Ku heterodimer (red and blue) and DNA-PKcs (green). (**d**) MRN, in cooperation with CtIP, cleaves the DNA bound by the DNA-PK complex. (**e**, **f**) GLASS-ChIP is carried out to collect these small fragments using anti-DNA-PK pS2056 antibodies. (**g**) The isolated DNA is quantified using qPCR

5. 1.25 M glycine (dissolve in water).
6. RIPA buffer: 50 mM Tris–HCl pH 8.0, 150 mM NaCl, 2 mM EDTA pH 8.0, 1% NP-40, 0.5% Sodium Deoxycholate, 0.1% SDS. Keep buffers cold until use.
7. Protease inhibitor tablets (Fisher A32955).
8. Anti-DNA-PKcs pS2056 rabbit antibody (Abcam ab124918).
9. Protein A/G magnetic beads (Fisher 88803).
10. Low salt wash buffer (TSE150): 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8.0, 150 mM NaCl.
11. High salt wash buffer (TSE500): 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8.0, 500 mM NaCl.
12. LiCl wash buffer: 0.25 M LiCl, 1% NP-40, 1% Sodium Deoxycholate, 1 mM EDTA, 10 mM Tris–HCl pH 8.0.

13. TE buffer: 10 mM Tris pH 8.0, 0.1 mM EDTA.
14. AMPure XP beads (Beckman Coulter A63880).
15. Nucleotide removal kit (Qiagen).
16. SYBR green PCR mix (Applied Biosystems PowerUp).
17. qPCR primers specific for DSB site of interest (*see* **Note 2**).
18. Instruments required: cell incubator, centrifuge, sonicator, rotator for dishes, rotator for tubes, 30 °C shaker, magnetic rack, 65 °C incubator/oven, qPCR machine.

2.1 Cell Culture Materials (Here Specified for Standard Growth of U2OS)

1. DMEM media, high glucose with pyruvate and glutamine (Millipore Sigma D6429).
2. Fetal bovine serum.
3. Trypsin-EDTA solution, 0.25%.
4. Penicillin/streptomycin (pen/strep) solution, 100×.

3 Methods (Fig. 2)

3.1 Cell Growth

1. Grow target cells in appropriate media to 60–70% confluency, and perform any treatment required for the experiment. For example, use U2OS cells grown in DMEM with 10% FBS and pen/strep, containing a virus encoding HA-AsiSI-ER to express the restriction enzyme AsiSI (DIvA system) [15]. The DNA-PKcs kinase inhibitor NU7441 can also be used to inhibit the NHEJ pathway and promote higher efficiency MRN cutting at DNA-PK-bound ends [7]. To treat the cells with DNA-PKcs inhibitor, add NU7441 to 10 μM final concentration in a 150 mm dish in 15 ml, and then incubate at 37 °C. After 1 h, add 4OHT (600 nM final concentration), and incubate at 37 °C for an additional 4 h.

3.2 Harvesting Cells

1. Place dishes on a horizontal rotator (90–110 rpm) at room temperature (RT). Add formaldehyde (37% stock) dropwise to 1% final concentration (400 μl for a 150 mm dish) while shaking, and continue incubation for 7 min.
2. Add 1.25 M glycine to 0.139 M final concentration (1.7 ml for 15 ml in a 150 mm dish) dropwise while shaking. Rotate for 5 min at RT.
3. Remove the media, and add 10 ml of cold PBS. Rinse the cells and discard the PBS. Add another 10 ml PBS. Leave cells in PBS until harvesting.
4. To harvest cells, aspirate PBS, and add 5 ml of fresh cold PBS. Scrape the cells from the plate using a cell scraper, and move to a 15 ml conical tube on ice. Add another 5 ml PBS to the dish,

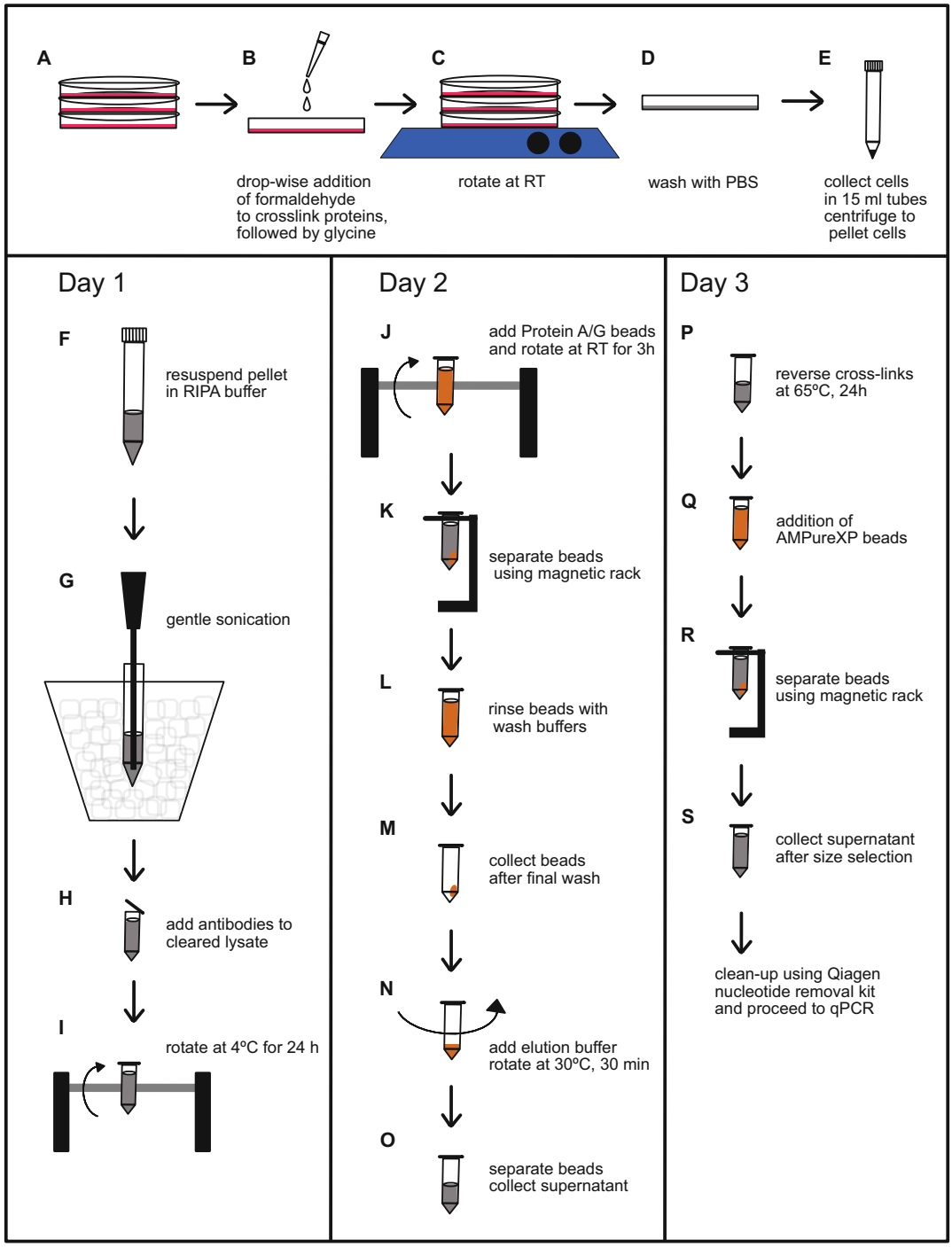


Fig. 2 Workflow of cell harvesting and GLASS-ChIP. (**a**–**e**) depict steps during harvesting of cells. (**a**) Dishes with cells to be used for GLASS-ChIP. (**b**) Cells are cross-linked with drop-wise addition of formaldehyde followed by neutralization with glycine. (**c**) Dishes are placed on a rotator during the cross-linking step. (**d**) Cells are rinsed with PBS after cross-linking. (**e**) Cell pellets in 15 ml conical tubes after harvesting of cells and centrifugation. (**f**–**s**) depict steps during GLASS-ChIP. (**f**) On the first day, cell pellets are thawed and resuspended in RIPA buffer. (**g**) Cell suspensions are placed on ice and sonicated gently. (**h**) Cell suspensions are transferred to microfuge tubes and centrifuged, and the supernatant is used for immunoprecipitation with

and scrape cells again to collect all cells. Combine with 5 ml in the tube. Keep the tubes cold during harvesting of all dishes.

5. Centrifuge the harvested cells at 1865 × *g* for 15 min at 4 °C. Aspirate the supernatant as much possible without losing cells. Freeze pellets in liquid nitrogen, and store at −70 °C. Cells from one 150 mm dish constitute one pellet in our experiment.

3.3 GLASS-ChIP

Day 1:

1. Prepare RIPA buffer with protease inhibitors sufficient for the samples. Use 1 tablet of protease inhibitors for 10 ml of cold RIPA buffer. Dissolve the inhibitors completely by brief vortexing, crush the tablet for easier dissolving, and keep buffers at 4 °C.
2. Thaw cell pellets at RT for 5 min, vortex gently for 2–3 s, and then transfer to ice bucket once thawed.
3. Add 2.3 ml RIPA buffer with protease inhibitors to the pellet. Mix with pipetting to resuspend the pellet completely; keep on ice for 15 min.
4. Keeping the tube on ice, sonicate at low power (Qsonica 55 with 1/8 in. microprobe at 15 setting) for 10 s, wait for 20 s, and then sonicate with 10–12 pulses of 1 s each. (Note: This sonication is mild compared to the standard ChIP sonication.)
5. Transfer the lysate to two 1.5 ml microfuge tubes.
6. Centrifuge the tubes at 845 × *g* (3000 rpm in microfuge) for 3 min. There will be a pellet at the bottom.
7. Label 1.5 ml microfuge tubes, one each for input, minus antibody and plus antibody pull-down. Transfer 1 ml each of the supernatant to the tubes labeled minus and plus antibody. Transfer 50 μl of the supernatant to the tube labeled as input. Keep all tubes in an ice bucket during these transfers.

Fig. 2 (continued) anti-DNA-PKcs antibody or other antibodies as appropriate. (**i**) Microfuge tubes are rotated overnight at 4 °C. (**j**) On the second day, Protein A/G magnetic beads are added to the tubes and rotated at RT for 3 h. (**k**) A magnetic rack is used to pull down the beads and remove supernatant. (**l**) Wash buffer is added, and beads are resuspended. Tubes are rotated at RT for 15–20 min similar to step **j**. Repeat steps **k** and **l** for all wash buffers. (**m**) Beads are collected after the final wash. (**n**) Elution buffer is added; tubes are incubated on shaker at 30 °C for 30 min. (**o**) Beads are separated using a magnetic rack, and the supernatant is collected. (**p**) The elutions are incubated at 65 °C for 24 h to reverse the cross-links. (**q**) On Day 3, size selection is carried out by adding AMPure XP beads. (**r**) A magnetic rack is used to pull down the beads. (**s**) The supernatant is collected after size selection and DNA purified using a Qiagen nucleotide removal kit. This DNA is then quantified using qPCR

8. Add 4 μl (1.6 μg) anti-DNA-PKcs-pS2056 antibody (Abcam, ab124918) to the plus antibody sample. The minus antibody tube serves as a control; a non-specific antibody should be added to this sample.
9. Incubate all (input, minus and plus antibody) tubes at 4 °C with rotation for 20 h (overnight).

Day 2:

1. Spin tubes briefly at 845 × *g* for 5 s. Keep the input tubes separately at 4 °C.
2. Mix Protein A/G magnetic beads (Pierce) thoroughly, and add 25 μl beads to the tubes labeled minus and plus antibody. Incubate these samples at RT for 3 h, with rotation.
3. Keep the buffers to be used for washing beads in the following steps at RT for warming. The SDS in elution buffer precipitates in cold temperatures, keep the elution buffer in 37 °C incubator till the precipitate dissolves completely, and then keep it at RT.
4. Centrifuge the tubes briefly at 845 × *g* for 5 s, and then isolate the beads using a magnetic rack.
5. Keeping the tubes on the magnetic rack, transfer supernatant to a new microfuge tube. Reserve for optional later use.
6. Add 1 ml of TSE150 buffer to the tubes with magnetic beads. Mix the beads by tapping the tubes, and rotate at RT for 15–20 min.
7. Transfer the tubes to the magnetic rack. Aspirate the supernatant, and add 1 ml of buffer TSE500. Mix the beads by tapping the tubes, and rotate at RT for 15–20 min.
8. Transfer the tubes to the magnetic rack. Aspirate the supernatant, and add 1 ml of LiCl buffer. Mix the beads by tapping the tubes, and rotate at RT for 15–20 min.
9. Transfer the tubes to the magnetic rack. Aspirate the supernatant, and add 1 ml of TE buffer. Mix the beads by pipetting, and transfer to a new microfuge tube and rotate at RT for 15–20 min.
10. Transfer the tubes to the magnetic rack. Aspirate the supernatant, and add 0.1 ml of elution buffer. Mix the beads by pipetting, and incubate at 30 °C for 30 min with shaking at 200 rpm (*see* **Note 3**).
11. Transfer the tubes to the magnetic rack. Aspirate the supernatant, and transfer to a new tube (*see* **Note 4**).
12. Incubate the elutions, and input tubes at 65 °C for 24 h (not less) to reverse the cross-links.

3.3.1 Size Selection

Day 3:

1. Cool down the sample tubes to room temperature for 30 min. At the same time, warm up AMPure XP beads to RT for 30 min. Keep the input tubes aside.
2. Mix the AMPure XP beads thoroughly, and add 65 μl beads to the minus antibody and plus antibody elutions. Mix well by pipetting up and down 10–15 times.
3. Incubate at RT for exactly 10 min. Separate beads using a magnetic rack.
4. Transfer the supernatant into a fresh PCR or DNA lo-bind tube.
5. Add 25 μl AMPure XP beads to this supernatant, mix well by pipetting up-down 10–15 times. Incubate at RT for exactly 10 min.
6. Transfer the supernatant to a fresh 1.5 ml vial. This is the size selected sample where larger fragments are removed for our experiment (*see* **Note 5**).
7. Purify the size selected supernatant as well as the input using Qiagen nucleotide removal kit. Elute in 60 μl elution buffer (Qiagen).

3.4 Quantitation of GLASS-ChIP DNA by qPCR

1. Use 3 μl input or size selected minus antibody or plus antibody samples as template in a 20 μl qPCR reaction containing 10 μl 2× PowerUp SYBR green PCR Master Mix and 0.5 μM of each primer. Each 96-well contains (a) 0.4 μl of 25 μM forward primer, (b) 0.4 μl of 25 μM reverse primer, (c) 10 μl of 2× PowerUp SYBR green PCR Master Mix (ABI), (d) 6.2 μl ddH_2O, and (e) 3 μl sample. Use PCR cycling as recommended by Thermo or other master mix suppliers. *See* Table 1 for examples of primer sets to use for AsiSI sites.
2. Calculate the percentage of DNA at selected sites. For each sample, a ΔC_t is calculated by subtracting the C_t value of the sample from the C_t value of the input. The percentage of DNA in the pull-down is calculated with the following equation: $2^{(Ct(\text{input}) - Ct(\text{test}))} \times 100\%$. To compensate for the difference in volume used for input (50 μl) and pull-down (1000 μl), the values are divided by 20. DNA values obtained for minus antibody samples are subtracted from the values obtained in presence of antibody. An example of the qPCR output of an experiment is shown in Fig. 3.

Table 1
Primers used for qPCR in AsiSI-expressing U2OS cells [7]

AsiSI site	Location as in hg18	Primer	Forward	Reverse	Amplicon size (bp)
DSB1	chr18: 7556705	U1	TCGGGGCCAGCGGCG TGTA	CGCCAGCCCGCTCCC	52
		D1	CGCGGGGCTCGGCGC	GGGAGA TGGCGCGGGAGC	40
		U2	GTGCTGGCTCAATG TGCTTATT	ACGATTTTGGGTC TGAGTGAA	132
		D2	CGCAGCCTC TTCCACAGTCA	GCCAC TACCGCCGCCGAA	139
DSB2	chr21: 32167382	U1	GGGAGCGGCCGCCCAG	GCTCCTAGCCG TGCGCT	40
		D1	CGGGAGCCCGACCCAA	CGCCGTC TGGCCCGCA	40
		U2	CGAAAGG TCCAGAAAACCCAA	GAAGCCACC TGAGCGCCAGA	132
		D2	TTGTCTACGCGCC TCGCT	CGGCTTCCCCGGC TTCT	119
DSB3	chr9: 129732985	U1	GACTGCGGCTGCA TCCAA	CGCCAGCGCC TCCCGC	41
		D1	CGCCTGCGGGTCCCGC	CTGAAGGATGC TGCAGCCGT	40
		U2	CCGCACTGGA TGAGAGCTT	CCTGGCGGATATCCC TCAA	112
		D2	GGACATCCATTCA TTGAACACA	GATCACGCGGGCAGC TGA	113
DSB4	chr22: 37194040	U1	CCCGGCCAAGAGTGCG T	CGCACCCGCGCGCCG	40
		D1	CGCGGAGCTGTGAGGC	GTCTCTAGG TGCCCCAGA	45
		U2	AAGATGAGGACAA TAGCAGGAA	AAGCCCCAATCTC TGCCTCA	125
		D2	CAGGGCGCTCCAGGTG T	GGGTCCTCCTCCTC TGAT	118

DSB sites from Aymard et al. [17]

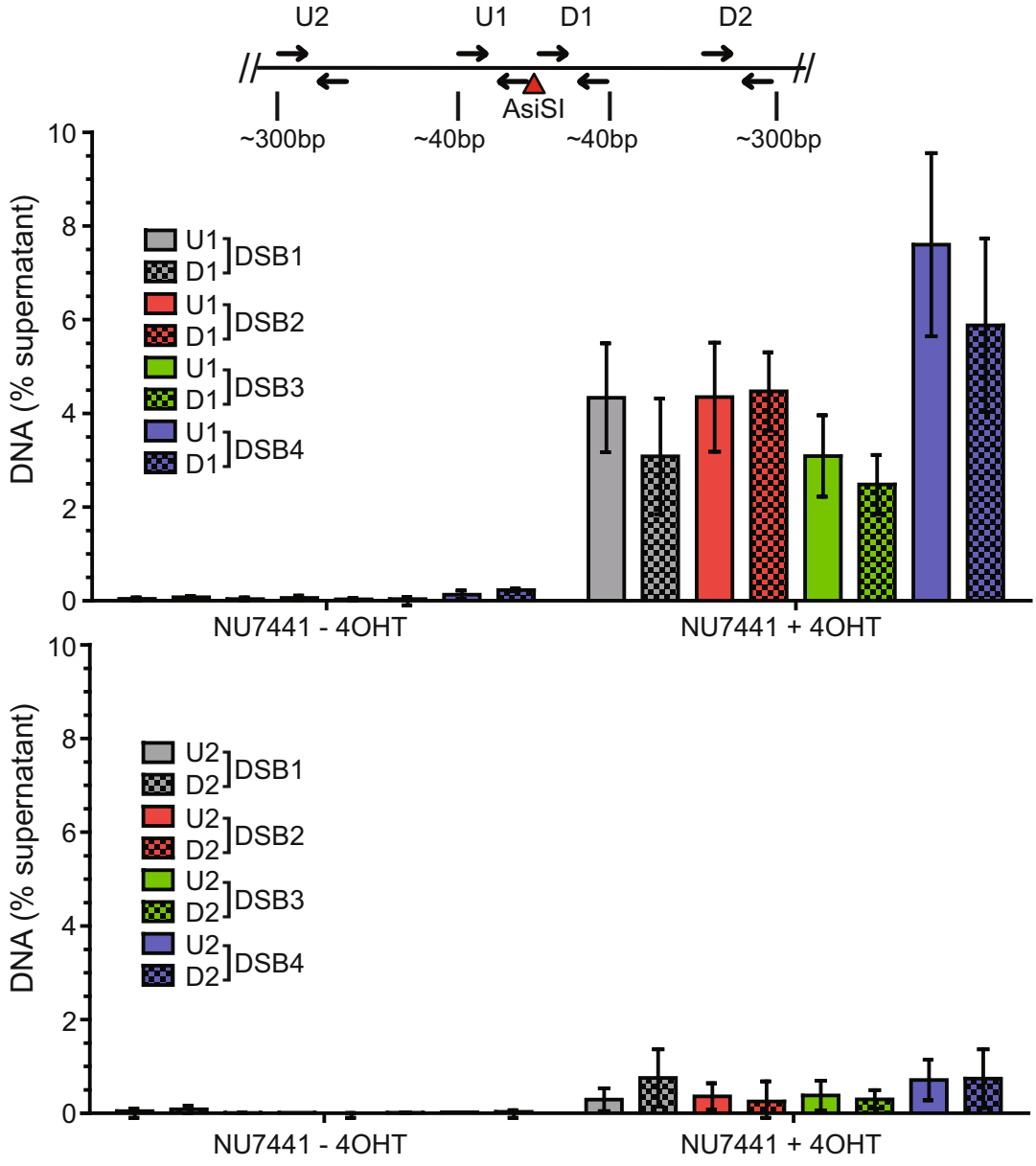


Fig. 3 Example of GLASS-ChIP qPCR result monitoring four AsiSI sites and upstream and downstream primers located either 40 bp or 300 bp from the DSB. Using the GLASS-ChIP protocol, small dsDNA products resulting from nucleolytic cleavage of DNA-PK-bound AsiSI-generated DNA ends were isolated from U2OS cells treated with 4-OHT or vehicle for 4 h as indicated. Cells were treated with NU7441 (10 μM) as indicated for 5 h starting at 1 h before 4-OHT addition. The isolated DNA was quantified by qPCR using primers located ~40 bp (primers U1 and D1) or ~300 bp (primers U2 and D2) from the AsiSI cut site. For each site, primer sets U1 and U2 (solid) are upstream, whereas D1 and D2 (checkered) are downstream of the AsiSI cut sites. At all four sites, DNA close to the AsiSI cut was detected with U1 and D1 primer sets upon induction of DSB with 4OHT addition. In comparison, the yield of DNA with U2 and D2 primer sets ~300 bp away from the AsiSI cut was low

4 Notes

1. Sometimes formaldehyde forms crystals at bottom, so it is best to use a fresh stock of formaldehyde or make sure the stock is clear.
2. qPCR primers: Use previously published primer sets close to AsiSI DSBs (Table 1) [7, 17] or any primer sets within approximately 100 bp of DSB location.
3. Since the elution volume is low, a shaker with horizontal circular motion was used. Tubes were kept at a 45° angle.
4. Optional: Use DNA lo-bind tubes (Eppendorf) for elution.
5. The amounts of AMPure beads used will remove DNA larger than 300 bp in size.

References

1. Pannunzio NR, Watanabe G, Lieber MR (2018) Nonhomologous DNA end-joining for repair of DNA double-strand breaks. J Biol Chem 293:10512–10523. https://doi.org/10.1074/jbc.TM117.000374
2. Symington LS (2016) Mechanism and regulation of DNA end resection in eukaryotes. Crit Rev Biochem Mol Biol 51:195–212. https://doi.org/10.3109/10409238.2016.1172552
3. Jasin M, Rothstein R (2013) Repair of strand breaks by homologous recombination. Cold Spring Harb Perspect Biol 5:a012740. https://doi.org/10.1101/cshperspect.a012740
4. Paull TT (2018) 20 years of Mre11 biology: no end in sight. Mol Cell 71:419–427. https://doi.org/10.1016/j.molcel.2018.06.033
5. Wright WD, Shah SS, Heyer W-D (2018) Homologous recombination and the repair of DNA double-strand breaks. J Biol Chem 293:10524–10535. https://doi.org/10.1074/jbc.TM118.000372
6. Blackford AN, Jackson SP (2017) ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response. Mol Cell 66:801–817. https://doi.org/10.1016/j.molcel.2017.05.015
7. Deshpande RA, Myler LR, Soniat MM et al (2020) DNA-dependent protein kinase promotes DNA end processing by MRN and CtIP. Sci Adv 6:eaay0922. https://doi.org/10.1126/sciadv.aay0922
8. Jones CE, Forsburg SL (2021) Monitoring *Schizosaccharomyces pombe* genome stress by visualizing end-binding protein Ku. Biology Open 10:bio054346. https://doi.org/10.1242/bio.054346
9. Kim J-S, Krasieva TB, Kurumizaka H et al (2005) Independent and sequential recruitment of NHEJ and HR factors to DNA damage sites in mammalian cells. J Cell Biol 170:341–347. https://doi.org/10.1083/jcb.200411083
10. Kochan JA, Desclos ECB, Bosch R et al (2017) Meta-analysis of DNA double-strand break response kinetics. Nucleic Acids Res 45:12625–12637. https://doi.org/10.1093/nar/gkx1128
11. Wu D, Topper LM, Wilson TE (2008) Recruitment and dissociation of nonhomologous end joining proteins at a DNA double-strand break in Saccharomyces cerevisiae. Genetics 178:1237–1249. https://doi.org/10.1534/genetics.107.083535
12. Yang G, Liu C, Chen S-H et al (2018) Super-resolution imaging identifies PARP1 and the Ku complex acting as DNA double-strand break sensors. Nucleic Acids Res 46:3446–3457. https://doi.org/10.1093/nar/gky088
13. Lobrich M, Shibata A, Beucher A et al (2010) gammaH2AX foci analysis for monitoring DNA double-strand break repair: strengths, limitations and optimization. Cell Cycle 9:662–669
14. Mao Z, Bozzella M, Seluanov A et al (2008) Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair (Amst) 7:1765–1771. https://doi.org/10.1016/j.dnarep.2008.06.018

15. Iacovoni JS, Caron P, Lassadi I et al (2010) High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. EMBO J 29:1446–1457. https://doi.org/10.1038/emboj.2010.38
16. Vítor AC, Huertas P, Legube G et al (2020) Studying DNA double-Strand break repair: an ever-growing toolbox. Front Mol Biosci 7:24. https://doi.org/10.3389/fmolb.2020.00024
17. Aymard F, Bugler B, Schmidt CK, Guillou E, Caron P, Briois S, Iacovoni JS, Daburon V, Miller KM, Jackson SP, Legube G. (2014) Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. Nat Struct Mol Biol. 21(4): 366–374. PMCID: PMC4300393
18. Zhou Y, Paull TT (2021) Quantifying DNA end resection in human cells. Methods Mol Biol 2153:59–69. https://doi.org/10.1007/978-1-0716-0644-5_5

Chapter 12

Monitoring Nuclease Activity by X-Ray Scattering Interferometry Using Gold Nanoparticle-Conjugated DNA

Daniel J. Rosenberg, Aleem Syed, John A. Tainer, and Greg L. Hura

Abstract

The biologically critical, exquisite specificity and efficiency of nucleases, such as those acting in DNA repair and replication, often emerge in the context of multiple other macromolecules. The evolved complexity also makes biologically relevant nuclease assays challenging and low-throughput. Meiotic recombination 11 homolog 1 (MRE11) is an exemplary nuclease that initiates DNA double-strand break (DSB) repair and processes stalled DNA replication forks. Thus, DNA resection by MRE11 nuclease activity is critical for multiple DSB repair pathways as well as in replication. Traditionally, in vitro nuclease activity of purified enzymes is studied either through gel-based assays or fluorescence-based assays like fluorescence resonance energy transfer (FRET). However, adapting these methods for a high-throughput application such as inhibitor screening can be challenging. Gel-based approaches are slow, and FRET assays can suffer from interference and distance limitations. Here we describe an alternative methodology to monitor nuclease activity by measuring the small-angle X-ray scattering (SAXS) interference pattern from gold nanoparticles (Au NPs) conjugated to 5′-ends of dsDNA using X-ray scattering interferometry (XSI). In addition to reporting on the enzyme activity, XSI can provide insight into DNA-protein interactions, aiding in the development of inhibitors that trap enzymes on the DNA substrate. Enabled by efficient access to synchrotron beamlines, sample preparation, and the feasibility of high-throughput XSI data collection and processing pipelines, this method allows for far greater speeds with less sample consumption than conventional SAXS techniques. The reported metrics and methods can be generalized to monitor not only other nucleases but also most other DNA-protein interactions.

Key words DNA repair, Small-angle X-ray scattering, Gold nanoparticles, Nuclease assay, MRE11A, X-ray scattering interferometry

1 Introduction

Meiotic recombination 11 homolog 1 (MRE11) is a critical nuclease that initiates DNA double-strand break (DSB) repair and processes stalled DNA replication forks [1–3]. Thus, DNA resection by MRE11 nuclease activity is important for multiple DSB repair

Daniel J. Rosenberg and Aleem Syed contributed equally to the work described in this chapter.

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_12,

pathways as well as in replication. The crystal structures of MRE11 show that the active site of MRE11 contains two Mn^{2+} ions with the protein forming a dimer both in an apo state and in DNA-bound states [4–6]. MRE11 is one of the first proteins to respond to DNA damage causing DSBs [1–3]. The DSBs are mainly repaired through either homologous-recombination (HR) that repairs the DSB in an error-free fashion or through non-homologous end-joining (NHEJ) pathway which may result in deletions or insertions in the repaired DNA. The first step in the HR repair pathway is MRE11-mediated resection (in the 3′–5′ direction) of the DSB leading to 3′ ssDNA overhangs [1]. These ssDNA overhangs inhibit NHEJ which requires very little processing of the broken DNA ends. Thus, MRE11 activity is the key determinant of whether DSBs are resected through HR or NHEJ [7].

In general, in vitro nuclease activity of purified enzymes is studied either through gel-based assays or fluorescence-based assays like fluorescence resonance energy transfer (FRET). However, adapting these methods for a high-throughput application such as inhibitor screening can be challenging. Gel-based approaches are slow, and FRET assays can suffer from interference and distance limitations (~1–10 nm) [8]. In the current method, we are combining our expertise in small-angle X-ray scattering (SAXS) with the scattering power of gold nanoparticles (Au NPs) conjugated to dsDNA substrates (Au-dsDNA). As will be demonstrated, when Au NPs are held at fixed distances on the 5′-ends of dsDNA, they act as molecular rulers through X-ray scattering interferometry (XSI) and can be used as a high-throughput technique to measure the binding and nuclease activity of MRE11 or other proteins that interact with DNA.

In biological research, SAXS is empowering for structural characterization of biomacromolecules at near physiological conditions [9]. Molecular assemblies, conformational changes, and flexibility can be robustly analyzed from a properly performed SAXS experiment [10]. SAXS is generally performed in solution with modest sample requirements, probing sub-nm distances and microsecond time-scales at many synchrotrons around the world. Coupled with sample handling robotics or microfluidics, impactful measurements can be made in high throughput. Although directional information is lost due to orientational averaging of the macromolecules relative to the probing beam, SAXS provides critical dynamics information to complement atomic resolution techniques like macromolecule X-ray crystallography (MX), nuclear magnetic resonance (NMR), and cryogenic-electron microscopy (cryo-EM) where both the distances and direction between atoms can be recorded [11]. Yet, high-resolution structures from MX or cryo-EM can typically only be attained on specific constructs that are sufficiently homogeneous in conformation and assembly. Fortunately, using these

experimental models and SAXS data, models of the full-length or alternate conformations can be determined along with information on flexibility, assembly, and conformational states [12]. SAXS is the right balance between information and throughput for many biological systems [9].

In a typical SAXS experiment, X-ray scattering from the biomacromolecule is measured in a buffer solution. The particle scattering intensity $I(q)$ is a function of momentum transfer $q = (4\pi \sin \theta)/\lambda$, where 2θ is the scattering angle and λ is the wavelength of the incident X-ray beam. $I(q)$ can be derived from the electron distribution within the biomacromolecule as:

$$I(q) = 4\pi \int_0^{D_{\max}} P(r) \frac{\sin(qr)}{qr} \mathrm{d}r$$

where r is the distance between election pairs within the macromolecule which leads to a statistical distribution of electron pair distances, or pair-distribution function, $P(r)$, where the maximal dimension, $D_{\max}$, of a molecule is found as the function goes to zero [13, 14].

SAXS is inherently a contrast measurement technique where the signal is derived from differences in electron density $\Delta\rho(r)$ between biomolecule $\rho(r)$ and that of the bulk solvent $\rho(s)$ [14] as:

$$\Delta\rho(r) = \rho(r) - \rho(s)$$

The approximate values for electron density of protein, DNA, and bulk solvent (pure water) are 0.43, 0.55, and 0.33 $e^-/Å^3$, respectively [15]. Given that the differences in the electron density between the biomolecules of interest and the buffer are already small, very minor fluctuations in buffer composition used for subtraction can greatly affect the results. Thus, a reasonable concentration of the analyte and a careful buffer subtraction are essential for obtaining useful information in SAXS experiments. To overcome the challenges in producing the large amounts of protein required for large-scale assays and the sensitivity of buffer fluctuations, high-throughput XSI can be used. This technique expands upon all of the same physical phenomenon of conventional SAXS by utilizing the interference pattern generated not between atom pairs but rather between heavy atom clusters (e.g., Au NPs) held at fixed distances by biomolecules (e.g., dsDNA). These Au NPs function as slits in reciprocal space to the atomic scale wavelengths of hard X-rays in an analogous way as the physical slits in Young's experiments with visible light from classical physics [16]. Importantly, Au NPs having significantly more electron density (5-nm NP ~4.6 $e^-/Å^3$) scatter X-rays with ~200-fold greater intensity as compared to a 172-kDa protein or 5400-fold higher than that of a 31-bp dsDNA since scattering intensity on an absolute scale, $I(0)$, (where $q = 0$) is

proportional to the square of the number of electrons (m) in a particle [17] as:

$$I(0) = N m^2 (1 - \rho(s)\,\psi)^2$$

where N is the number of particles and ψ is ratio between the particle volume and its number of electrons. The original idea of measuring scattering from heavy metals in a biomolecule was proposed in as early as the late 1940s and successfully performed in 1980 [18]. In 2008, Mathew-Fenn et al. were the first group to use Au-dsDNA as molecular rulers via XSI [19], applying this technique to measuring the double helix with exceptional accuracy [20]. The $P(r)$ functions derived from these experiments can be divided into two major peak regions. One corresponding to intra-Au and another for the inter-Au distances [17, 20, 21]. Using this approach, inter-particle distances between 2 Au NPs separated by up to 100-bp have been accurately measured [21], and greater distances are presumed possible. Others have followed this technique, studying Au NPs conjugated to DNA, RNA, and even proteins using XSI [22–27].

Our group has applied the XSI technique to probe the mismatch repair of MutS/L, demonstrating that the technique can be used to study damage-specific structural changes in the DNA caused by MutS/L [17]. This study focused on the qualitative changes in the inter-particle distances providing information on the DNA-protein interactions. In our study of MRE11 nuclease activity, we sought to observe such DNA-protein interactions as well as develop a more quantitative assay towards the future of high-throughput XSI experiments. As such we have designed Au-DNA substrates of two different lengths (37-bp and 57-bp) conjugated to 10 nm Au NPs via a Trithiol (TrT) (Letsinger's type) linker on the 5′-end annealed to a shorter (9-bp) duplex forming ssDNA oligo leaving a long stretch of ssDNA available for MRE11 binding. In both substrate cases, the inter-Au distance distributions are shifted to lower mean values compared to the substrate alone for samples where nuclease activity was not observed indicating structural changes in the DNA associated with MRE11 binding. As expected, MRE11 nuclease activity decreases the population of doubly Au Nps-labelled dsDNA as observed by the decrease in the amplitude of $P(r)$ corresponding only to inter-Au distances. The intra-Au regions remain unperturbed as the amount of Au NPs is not changing. The $P(r)$ functions are normalized to the intra-Au peak during analysis to account for any minor fluctuations in Au NP concentration. From these XSI assays, we observe that MRE11 is not active when the active site metal (Mn^{2+}) is not present in the reaction buffer or a nuclease-dead mutant (H129N) is used in the reaction instead of the wild-type (WT) enzyme.

Prior to analysis, we validated that the substrate is cleaved by MRE11 via gel-based assays using the same substrates as in the XSI experiments except with Fluorescein (6-FAM) substituted for the TrT linkers and Au NPs on the 5′-ends. In general, it is useful (but not essential) to have an independent assay for protein-DNA interactions. For MRE11, the gel-based activity assay data agreed with our XSI assays showing that the nuclease activity is only observed in the reaction with the WT enzyme in the presence of $MnCl_2$. Additionally, these gel-based assays indicate that MRE11 can cut on both strands, and on the longer strand, it can chew all the way to the 5′-end of the DNA.

Since SAXS probes all molecules in a solution, homogeneous samples are often used [9]. In DNA repair and damage responses, there is a need to examine enzyme activities where their active states may be in complexes that are transient and dynamic. To address this challenge, we combined the efficiency of SAXS with the high contrast of Au and present here a SAXS method with Au-labeled DNA as a robust prototypic assay on DNA processing. More specifically, these experiments can be carried out in a variety of solution conditions, in high throughput, provide sub-nm resolution at low concentrations, and have the inherent potential to categorize sub-millisecond reaction steps. Furthermore, many DNA repair processes have longer DNA footprints than are comfortably assayed using FRET. This method and the approach defined here for MRE11 can complement and extend more traditional, fluorescent-based assays. High-throughput XSI has a robust ability to test combinations and additives including other macromolecules without loss of signal. These protocols offer strategic and tactical advantages for studies to identify novel inhibitors from screening chemical libraries with the expectation that 1000 experiments can be done weekly and with batched compounds 10,000 compounds can be screened in 1 week.

In the following sections, we describe our XSI method to conjugate DNA substrate to the Au NPs for XSI experiments and data analysis protocol. We employ MRE11 as an example, but by changing the design of the DNA substrate, this method can be applied for many other enzymes that are known to cause structural changes in DNA including major types of DNA damage responses. For example, some DNA repair proteins of biological interest bind DNA without making any chemical alterations to control pathway selection [28]; however, if these bend DNA or otherwise alter the distance between DNA ends as they typically do, then XSI will provide a sensitive high-throughput measure of their interactions. We therefore expect XSI will be able to interrogate the impacts of proteins and RNA binding to DNA repair and replication complexes, ranging from scaffold proteins such as XRCC1 that is essential to micro-homology-mediated end joining [29] to RNA that can act in efficient DSB repair machines [30], to PAR clouds at

DNA damage controlled by poly(ADP-ribose) polymerase (PARP1) and poly(ADP-ribose) glycohydrolase (PARG) whose inhibitors are actively being pursued for cancer therapy [31, 32], and even to G-quadruplex, repetitive sequence elements, and other non-B DNA sequences associated with DNA instability and mutation sites [33–35]. We demonstrate how to leverage this technique for use at a researcher's home institution as well as how to take advantage of the mail-in user program of the SIBYLS beamline at the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory (LBNL) helping to design and carry out experiments like those mentioned herein.

2 Materials

2.1 Preparation of BSPP Protected Au Nanoparticles Via BSPP-Citrate Exchange (Au-BSPP)

1. 15 mg/mL BSPP solution: Dissolve 375 mg bis(*p*-sulfonatophenyl)phenylphosphine (BSPP) in 25 mL ddH_2O.
2. 5 M NaCl solution: Dissolve 146 g sodium chloride (NaCl) in 400 mL ddH_2O, and then add ddH_2O until total volume equals 500 mL.
3. 100 mM Phosphate buffer (PBS) pH 7: Dissolve 7.744 g of sodium phosphate dibasic heptahydrate ($Na_2HPO_4{\cdot}7H_2O$) and 2.913 g of sodium phosphate monobasic monohydrate ($NaH_2PO_4{\cdot}H_2O$) in 400 mL ddH_2O. Adjust pH to 7 using HCl or NaOH, and then add ddH_2O until total volume equals 500 mL.
4. Au-BSPP storage buffer (15 mM PBS, 1 mg/mL BSPP, 1 mM TCEP, pH 6.4): Dissolve 50 mg BSPP and 14 mg Tris(2-carboxyethyl)phosphine hydrochloride (TCEP) in 30 mL ddH_2O. Add 15 mL 100 mM phosphate buffer (PBS) pH 7 to solution. Adjust pH to 6.4 using HCl or NaOH, and then add ddH_2O until total volume equals 50 mL.
5. 10 nm Au NP: Purchased from Ted Pella.

2.2 Au-ssDNA Conjugation, Anion Exchange Chromatography, and Au-dsDNA Annealing

1. 5′ Tri-thiolated ssDNA in solution: Purified/lyophilized ssDNA sequences with a Trithiol (Letsinger's type) modification to the 5′-end are purchased from Fidelity Oligos at ~100 nmole scale (Table 1) and are re-hydrated in 0.5 mL ddH_2O.
2. SH-PEG solution: Thiolated poly(ethylene glycol) (SH-PEG), MW = 356.5 was purchased from PolyPure (Oslo, Norway), and 20 μL SH-PEG is added to 480 μL ddH_2O.
3. High salt FPLC buffer (15 mM Tris, 1 M NaCl, pH 8): Dissolve 58.44 g NaCl and 1.82 g tris(hydroxymethyl)-aminomethane (Tris) in 900 mL ddH_2O. Adjust pH to 8 using HCl or NaOH, and then add ddH_2O until total volume equals 1 L.

Table 1
Table showing DNA substrate sequences

[a]37-bp-Au-F	5′-FAM-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGGGCGC-3′ 3′-CGGCCCGCGT-5′
[a]37-bp-Au-R	5′-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGGGCGC-3′ 3′-CGGCCCGCGT-5′-FAM
[a]57-bp-Au-F	FAM- 5′-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGGGCGC-3′ 3′-CGGCCCGCGT-5′
[a]57-bp-Au-R	5′-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGGGCGC-5′ 3′-CGGCCCGCGT-5′-FAM
[b]37-bp-XSI substrate	5′-(Au-NP)-TrT-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGGGCGC-3′ 3′-CGGCCCGCG-TrT-(Au-NP)-5′
[b]57-bp-XSI substrate	5′-(Au-NP)-TrT-TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCCGGGCGC-3′ 3′-CGGCCCGCG-TrT-(Au-NP)-5′

[a]FAM = fluorescein (6-FAM)
[b]TrT = trithiol linker (Letsinger's type)

4. No salt FPLC buffer (15 mM Tris, pH 8): Dissolve 1.82 g Tris in 990 mL ddH_2O. Adjust pH to 8 using HCl or NaOH, and then add ddH_2O until total volume equals 1 L.
5. FPLC AKTA purifier for an automated anion-exchange chromatography.

2.3 Protein Expression and Purification

1. The catalytic domain of human MRE11 nuclease (1–411) selected is based on the previous report [6] and is cloned into pET series expression vector with an N-terminus His-tag (Addgene#29653). Surface-exposed methionines (M26, M84, M157, M309, M343) are modified to leucines for improving the protein stability; the modified MRE11 construct maintains the nuclease activity as the parental construct.
2. DH5α chemical competent cells (Thermo Fisher).
3. Rosetta™ chemical competent cells (Novagen).
4. BD Difco™ LB Broth, Miller (Luria-Bertani) media.
5. BD Difco™ LB Agar, Miller (Luria-Bertani) media.
6. Kanamycin sulfate UPS grade (Teknova).
7. 37 °C incubator and refrigerated shaker.
8. Lysis buffer (50 mM Tris (pH = 7.5), 500 mM KCl, 5% glycerol, 0.5% T-20, 1 mM TCEP, protease inhibitors).
9. Buffer A (25 mM Tris (pH = 7.5), 300 mM NaCl, 2.5% glycerol, 1 mM TCEP, 20 mM imidazole).

10. Buffer B (25 mM Tris (pH = 7.5), 300 mM NaCl, 2.5% glycerol, 1 mM TCEP, 500 mM imidazole).
11. SEC buffer (20 mM Tris (pH = 8.0), 200 mM NaCl, 0.1 mM EDTA, 5 mM DTT).
12. HisTrap FF Crude pre-packed 5 mL column (GE/Cytiva).
13. Hi Load™ 16/600 Superdex200 pg (GE/Cytiva).
14. FPLC AKTA Pure system (GE/Cytiva) for automated affinity and size-exclusion chromatography.
15. Thermo Scientific NanoDrop 2000 Spectrophotometer.

2.4 DNA Substrate Preparation for the Fluorescence-Based Nuclease Reaction

DNA sequences (from IDT) used in the gel-based assay are given in Table 1 and annealed using a PCR machine.

2.5 Fluorescence-Based Nuclease Reaction to Validate Substrates and the Activity

1. Nuclease reaction buffer for FAM-based detection (25 mM HEPES (pH = 7), 50 mM KCl, ±1 mM $MnCl_2$).
2. PCR machine.
3. To make 3× Stop Buffer, mix 0.5 mL formamide, 0.12 mL of 0.5 M EDTA (pH = 8), 0.25 mL of 100% glycerol, and 0.15 mL of 10% SDS.
4. 4–20% Mini-PROTEAN® TGX Stain-Free protein gels (Bio-Rad).
5. 15% Mini-PROTEAN® TBE-Urea gel (Bio-Rad).
6. Mini-PROTEAN® Tetra Vertical electrophoresis cell and PowerPac™ power supply.

2.6 Sample Preparation of XSI Experiments

1. Nuclease reaction buffer for SAXS-based detection (25 mM MOPS, 60 mM KCl, 0.2% T-20, pH 7): Dissolve 2.89 g (3-(*N*-morpholino)propanesulfonic acid) (MOPS), 2.24 g potassium chloride (KCl), and 1.095 g Tween-20 in 980 mL ddH_2O. Adjust pH to 7 using HCl or KOH, and then add ddH_2O until total volume equals 1 L.
2. 20 mM $MnCl_2$ solution: Dissolve 125.8 mg in 50 mL ddH_2O.
3. PCR machine.

3 Methods

3.1 Preparation of BSPP Protected Au Nanoparticles Via BSPP-Citrate Exchange (Au-BSPP)

1. Add 25 mL of 15 mg/mL BSPP to 400 mL of either purchased 10 nm colloidal Au NPs, and filter solution through 0.22 μm filter.
2. Stir 400 mL citrate-stabilized colloidal Au NPs with BSPP overnight.

3. Add 5 M NaCl until the solution turns from red to dark red/-purple (~75 mL).
4. Pour into Beckman 100 mL polypropylene bottles w/cap assembly.
5. Spin in Beckman centrifuge in JA-18 rotor @ 12,000 G for 10 min.
6. Decant slowly, or pipette off supernatant (*see* **Note 1**).
7. Use 0.5 M NaCl solution to wash NPs, sonicate, and repeat **step 6**.
8. Repeat **step 7** twice.
9. Resuspend in 25 mL Au-BSPP storage buffer.

3.2 Au-ssDNA Conjugation, Anion Exchange Chromatography, and Au-dsDNA Annealing

1. One limitation of Au NP conjugation to DNA is feasible only in the 5′ end of the DNA; thus, we designed a substrate (Table 1) that would be cut by MRE11 as well as leads to separation of paired Au NPs upon the nuclease reaction.
2. For the conjugation, measure the concentrations of Au-BSPP and ssDNA solutions (diluted appropriately; *see* **Note 2**) using Thermo Scientific NanoDrop 2000 Spectrophotometer at 520 and 260 nm, respectively.
3. Calculate concentration using Beer's law and the appropriate extinction coefficients for Au NPs and ssDNA (*see* calculation **Note 3**).
4. Colloidal Au-BSPP and selected ssDNA solutions are mixing at a mole ratio of 3:1 and shaken gently at room temperature (RT) overnight.
5. SH-PEG solution is added to final to mixture at v/v% ratio of 10% (i.e., 100 μL added to 1000 μL solution), and mixture is shaken gently at RT for 2 h.
6. Separate and collected mono-conjugated Au-ssDNA from multi-conjugated using a Dionex DNA-Pac PA100 anion exchange column on an AKTA series fast protein liquid chromatography (FPLC) (*see* FPLC Method **Note 4,** Fig. 1a, b).
7. Complementary Au-ssDNA conjugates are annealed by heating at 94 °C for 3 min and allowing to cool to RT slowly to form final Au-dsDNA substrates (Fig. 1c).
8. Final Au-dsDNA substrates are observed via XSI to ensure inter-particle signal only seen from the properly annealed substrate (Fig. 1c).

3.3 Protein Expression and Purification

1. After expression plasmids are verified through DNA sequencing, plasmids are amplified by transforming into DH5α cells, and cells are grown on LB-agar plates with kanamycin selection (50 μg/mL) overnight at 37 °C and are extracted using Qiagen® miniprep kit as per the manufacturer protocol.

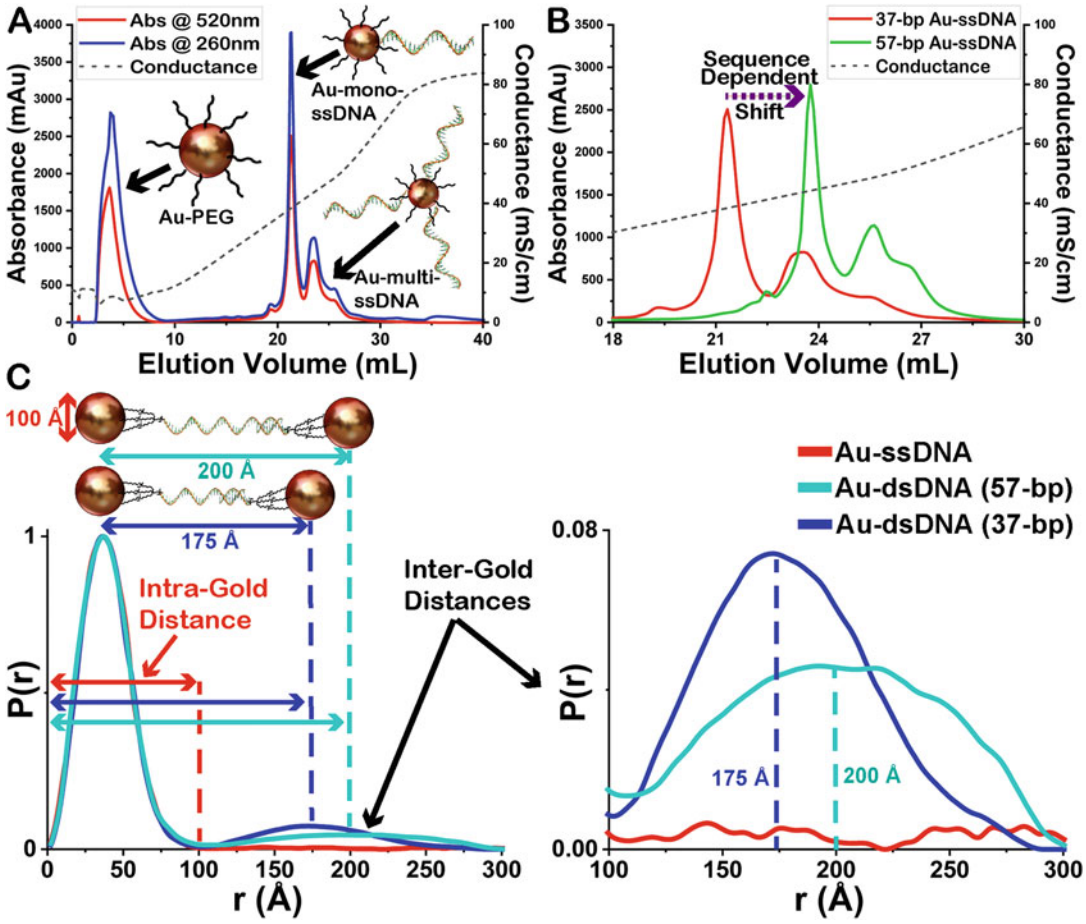


Fig. 1 Anion exchange chromatograms showing (**a**) the separation of mono-conjugated Au-ssDNA from un-conjugated Au-PEG and multi-conjugated Au-ssDNA using fast protein liquid chromatography (FPLC) and (**b**) the sequence-dependent shift in the elution volume. Both as measured by diode array detector (DAD) at 520 and 280 nm for Au NPs and ssDNA, respectively, as well as measurement of conductance showing salt gradient conditions. (**c**) Demonstration of the two Au-dsDNA substrates used and the normalized electron-pair distance distribution $P(r)$ functions from these experiments showing the peak regions corresponding to the intra-Au and inter-Au distances as well as the disappearance of the inter-Au distances in the Au-ssDNA sample. The $P(r)$ functions are normalized to the intra-Au peak to compensate for fluctuations in concentration (Au NP conc 200 nM $\pm$ 10)

2. For protein expression, extracted plasmids are transformed into Rosetta™ competent cells in a similar fashion as above (*see* **Note 5**).
3. Expression-plasmid transformed Rosetta™ cells are inoculated into a small LB culture medium (200 mL) supplemented with kanamycin (50 μg/mL) and grown overnight at 37 °C in a shaker.

4. Overnight culture is further utilized to inoculate large-scale (6 L) LB media (1.5 L/flask) supplemented with kanamycin (50 μg/mL), and protein expression is induced with 0.75 mM IPTG at 16 °C overnight.
5. Cells are harvested and stored in −80 °C deep freezer until further use.
6. Cell pellets are thawed and resuspended in the lysis buffer and homogenized using a Dounce homogenizer.
7. Homogenized cells are lysed by sonication.
8. Lysed cells are clarified by centrifugation at 39,191 × *g* for 45 min.
9. Automated affinity purification is performed on an FPLC system (e.g., AKTA Pure). The clarified lysate is loaded onto a prepacked 5 mL HisTrap column. Prior to loading the lysate, the column is pre-equilibrated with Buffer A. The following steps are used for the automated affinity purification: column wash, 100 mL of Buffer A; second wash, 25 mL of 10% Buffer B; and elution, 50 mL of 60% Buffer B, second elution, 50 mL of 100% Buffer B. Protein eluted with 60% Buffer B is used for downstream activity assays.
10. Eluted protein fractions are verified by protein gel electrophoresis, and protein-containing fractions are pooled and concentrated and loaded onto pre-equilibrated (with SEC buffer) Superdex 200 16/600 column mounted on an AKTA pure machine for further purification by size-exclusion chromatography.
11. Protein fractions are verified by gel electrophoresis, and protein-containing fractions are pooled and concentrated and quantified by NanoDrop.
12. Protein is distributed into 20–30 μL fractions and flash frozen in the liquid nitrogen and stored in −80 °C deep freezer until further use.
13. Plasmid for the nuclease-dead version of the enzyme (H129N) was generated through mutagenesis and purified exactly as the wild-type (WT) enzyme.
14. Given the composition of SEC buffer contains 0.1 mM EDTA, the purified proteins at the end are in a metal-free state.

3.4 DNA Substrate Preparation for the Fluorescence-Based Nuclease Reaction

1. Identical DNA substrates are used in Au-SAXS and gel-based nuclease reactions (*see* **Note 6** and Table 1).
2. All 5′-Fluorescein (FAM) labelled DNA oligos are purchased from IDT with HPLC purification.
3. We verified that MRE11 cuts our substrate through monitoring the cleavage in a fluorescence-based nuclease assay (Fig. 2).

4. To monitor how MRE11 cuts the DNA substrates on both strands, both 37-bp and 57-bp substrates are labelled with FAM at 5′ individually resulting in four different substrates: (1) duplex with a 5′-FAM on longer strand of 37-bp, (2) duplex with 5′-FAM on shorter strand of 37-bp, (3) duplex with a 5′-FAM on longer strand of 57-bp, and (4) duplex with a 5′-FAM on shorter strand of 57-bp.
5. DNA substrates (in Table 1) used in gel-based nuclease reaction are prepared by annealing complementary non-labelled strand with fluorescently labelled oligo (in 1.3:1 ratio) and by heating at 95 °C for 5 min followed by gradual cooling to room temperature for the duplex formation.
6. Substrates are stored at −20 °C until further use (@ 1 μM stock concentration).

3.5 Fluorescence-Based Nuclease Reaction to Validate Substrates and the Activity (Fig. 2)

1. Proteins (WT or H129N) are diluted to the desired concentration in the nuclease reaction buffer with or without $MnCl_2$.
2. Nuclease reaction is initiated by adding the substrate to the reaction mixture and incubating at 37 °C for 1 h.
3. Nuclease reaction is stopped by adding a stop buffer and incubated further at 37 °C for 15 min (*see* **Note 7**).
4. For each substrate, a non-labelled version of the cleaved FAM-labelled strand is added (100–200-fold excess) to the reaction mixture to visualize only FAM-labelled ssDNA product.
5. Reaction mixture is run on a denaturing TBE-UREA gel for 50 min at 185 V (*see* **Note 8**).
6. Gel can be imaged with FAM excitation/emission filter on any gel imager (Fig. 2).

3.6 Sample Preparation of XSI Experiments

1. Dialyze Au-dsDNA substrates overnight at 4 °C in 1 L reaction buffer using 4 kDa dialysis membranes. Be cautious of strong reducing agents in the buffer (*see* **Note 9**).
2. Measure the concentrations of Au-dsDNA (diluted appropriately; *see* **Note 2**) using Thermo Scientific NanoDrop 2000 Spectrophotometer at 520 nm. Adjust concentration if needed (*see* **Note 10**).
3. Combine enzymes with Au-dsDNA in an Axygen 96-well Polypropylene PCR Microplate at a final molar ratio of 10:1 (MRE11 2 μM and Au NPs 200 nM) in nuclease reaction buffer (with or without 2 mM $MnCl_2$), and then bracket the samples with a blank buffer sample on either end for buffer subtraction (*see* **Note 11**, Fig. 3).
4. Incubate plate containing samples at 37 °C for 1 h for the reaction to take place.

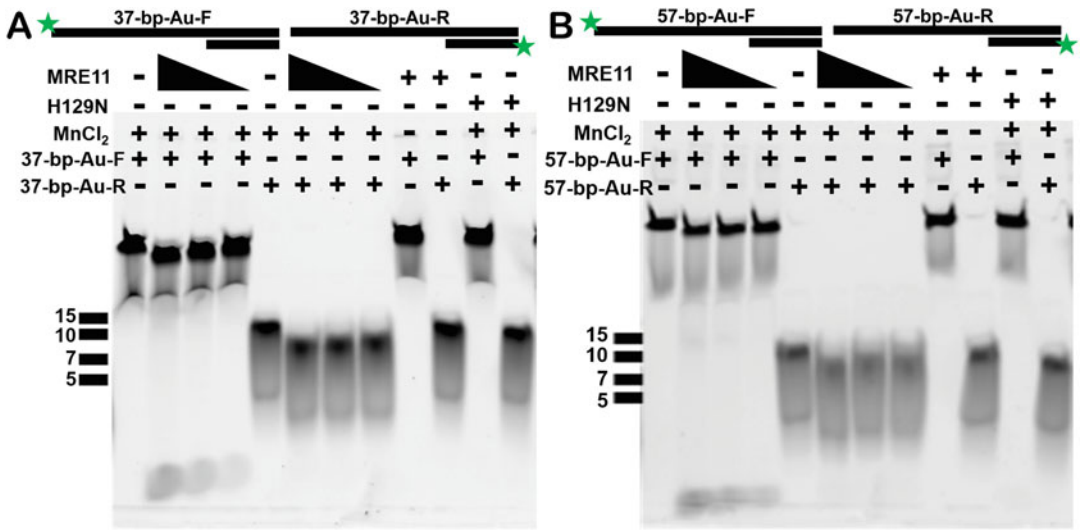


Fig. 2 MRE11 nuclease activity as monitored in gel-based assays. Both substrates (37-bp and 57-bp) used in the XSI experiments were used in fluorescence-based nuclease assay. To monitor the nuclease activity of MRE11 on both strands of the duplex substrate, a 5′-FAM label is added on either end resulting in four substrates as shown above (37-bp-Au-F, 37-bp-Au-R, 57-bp-Au-F, and 57-bp-Au-R). (**a**) MRE11 shows nuclease activity (at 2, 1, and 0.5 μM concentration) on both strands of 37-bp substrate, and the activity is dependent on the presence of $MnCl_2$ in the reaction buffer. As expected, the nuclease-dead mutant H129N is not active even in the presence of $MnCl_2$ at 2 μM enzyme concentration. (**b**) MRE11 shows nuclease activity (at 2, 1, and 0.5 μM concentration) on both strands of 57-bp substrate, and the activity is dependent on the presence of $MnCl_2$ in the reaction buffer. As expected, the nuclease-dead mutant H129N is not active even in the presence of $MnCl_2$ at 2 μM enzyme concentration. ssDNA markers are indicated for each gel

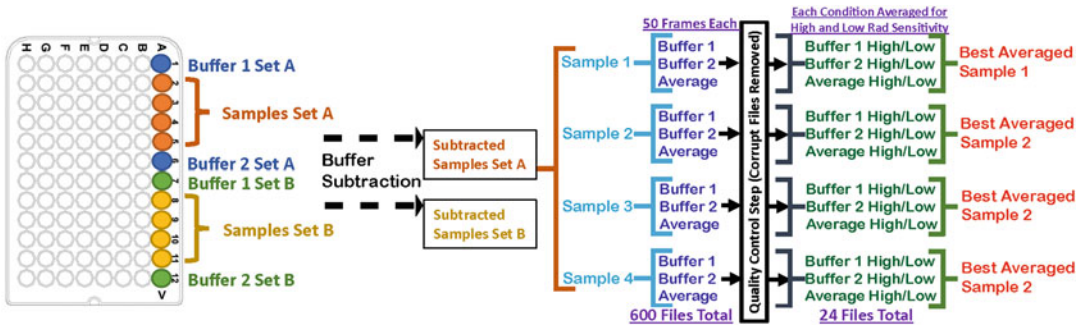


Fig. 3 Exemplary demonstration of how to set up a 96-well plate for high-throughput XSI assay and the subsequent data processing pipeline

3.7 XSI Data Collection at the SIBYLS Beamline (See Note 12)

1. XSI data is collected at the SIBYLS beamline (BL12.3.1), at the Advanced Light Source at Lawrence Berkeley National Laboratory, Berkeley, California [36]. To send samples for collection, *see* **Note 12**.

2. Load 96-well sample plate onto cooled 10 °C sampling position.
3. Samples are transferred from a 96-well plate via a Tecan Evo liquid handling robot with modified pipetting needles acting as sample cells to the X-ray beam as described previously [9].
4. X-ray wavelength is set at $\lambda = 1.24$ Å, and the sample-to-detector distance is 2.1 m, resulting in scattering vector q, ranging from 0.01 to 0.45 $Å^{-1}$. The scattering vector is defined as $q = 4\pi \sin \theta/\lambda$, where 2θ is the scattering angle. Data is collected using a Dectris PILATUS3X 2M detector at 20 °C and processed as previously described [37]. Samples are exposed to X-ray synchrotron radiation for a total of 10 s at a frame rate of 0.2 s for a total of 50 images.
5. For each sample collected, two sample-free buffer samples are also collected to reduce error in subtraction (Fig. 3).
6. Each collected image is circularly integrated and normalized for beam intensity to generate a one-dimensional scattering profile by beamline specific software (Fig. 3).
7. Buffer subtraction is performed for the one-dimensional scattering profile of each sample by using each of the two corresponding buffers, producing two sets of buffer-subtracted sample profiles to ensure the subtraction process was not subject to instrument variations (Fig. 3).

3.8 Setting up XSI Data Processing Pipeline

1. Once data collection has been completed, you will receive your data back with the following file hierarchy:

```
Username_Date:
   Results
     Subtracted
     A2_results
      Average
     50 Integrated Curves (.dat files)
      Buffer1
     50 Integrated Curves (.dat files)
      Buffer2
     50 Integrated Curves (.dat files)
     Unsubtracted
        A1b_results
        A2_results
        A3b_results
```

2. Scattering profiles over the 10 s exposure (50 frames total) should be sequentially averaged to eliminate any radiation damage affects. This can be done either manually, for each

sample using our web-based beamline software FrameSlice (sibyls.als.lbl.gov/ran), or by batch processing using our XSI data processing pipeline (Fig. 3) which is recommended for large data sets.

3. To set up your system for running the frame averaging pipeline, we have recommended Bash terminal environment (*see* **Note 13**), but it should be able to run on any platform with Python 3 and pip (both required).
4. Check your versions of python and pip.

   To check your python version from terminal:

```
$ python3 --version
```

   If no version of python 3 (*see* **Note 13**).

   To check your pip version from terminal:

```
$ python3 -m pip --version
```

   If no version of pip type:

```
$ python3 get-pip.py
```

5. In a new bash terminal clone our gitlab repository (*see* **Note 14**):

```
$ git clone https://git.bl1231.als.lbl.gov/djrosenberg/frame_averaging_pipeline.git
```

6. Go to the folder called frame_averaging_pipeline, and install:

```
$ cd frame_averaging_pipeline
$ pip install .
```

   To make sure pip has installed frame_averaging_pipeline:

```
$ pip list
```

   Note the location of the repository folder, frame_averaging_pipeline:

```
$ pwd
```

Example Output: *folder_path/*frame_averaging_pipeline.

This output we will call *folder_path* (needed in **step 7** to run the main script).

7. Start Xserver if on Windows or Mac, and leave it running in the background (*see* **Note 15**).
8. If your data is **local,** *cd* to the folder containing the *Results* folder you would like to process (called *Username_Date* in the file hierarchy example above), and run the *xsi_batch_processing* shell script.

   Here we use the example of the *Test_Data* included in the *frame_averaging_pipeline* folder:

```
$ cd tests/Test_Data
```

Run *xsi_batch_processing.sh* using the path from **step 4**:

```
$ folder_path/frame_averaging_pipeline /xsi_batch_processing.sh
```

9. When asked "Is your data on your local machine and are you in the folder containing your Results folder," answer "y" or "yes."
10. The data processing pipeline should start. Once complete you are asked to "Please Review Output In" the folder. Scroll through the .png images in the Xserver window or preferred image viewer if prompted (*see* **Note 15**), and decide whether buffer subtraction one, two, or the average should be used, and enter 1, 2, or A, respectively (if buffers match closely, use average). Then select whether the samples show high or low sensitivity to radiation, and enter either H or L, respectively (*see* Fig. 4).
11. Make sure the desired data has been selected and the output directory is correct, and enter "Y" to continue or "N" to repeat selection **step 9**.
12. To pull your data directly from the SIBYLS beamline database (must have an account; *see* **Note 12**), answer "n" or "no" when asked "Is your data on your local machine and are you in the folder containing your Results folder."
13. When asked, "Please Enter Your SIBYLS User Name: (this is caps sensitive)."
14. When asked, "Is *Current_Year* the correct year of your data collection?," answer "Y" to continue or "N" to enter the year of your data collection as YYYY.

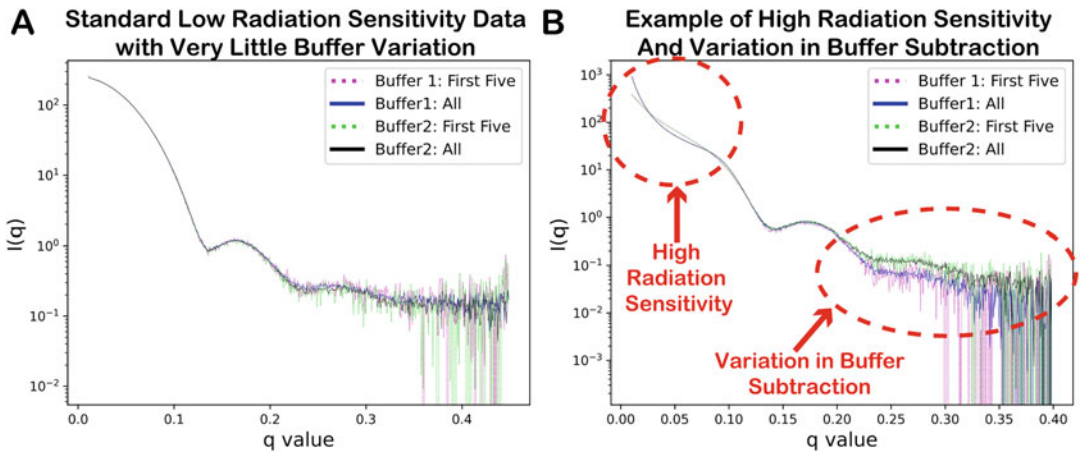


Fig. 4 Output from the xsi_batch_processing.sh script with examples of (**a**) standard low radiation sensitivity data and (**b**) high radiation sensitivity with variations in buffer subtraction

15. Enter user *password*.
16. When asked, "Please Select Data Folder Name You'd Like to Work on (this is caps sensitive) (For example: 2020_02_25_*username*_results)."
17. Make sure the folder and path are correct, and answer "Y" to continue or "N" to repeat data selection **step 14.**
18. Enter user *password*.
19. Batch data processing will start automatically. Follow **steps 9** and **10**.
20. Once XSI batch processing is completed, the following file hierarchy is output:

```
Username_Date:
    Best_Curves (best selected averaged curves)
    Results (original subtracted/unsubtracted data)
    Username_Date _Ave_Curves
     Full_Ave (full frame averaged curves (low radiation))
    Region_Ave (Regionally averaged curves (high radiation))
    Username_Date _Ave_Graphs (.png images of averaged curves)
    logs (logs for debugging code)
```

3.9 XSI Data Analysis and Interpretation

1. Once you have your best frame averaged XSI curves (*best_curves* in the above file hierarchy), you can use those for analysis.
2. The simplest analysis is to generate pair-distribution functions, $P(r)$, from the inverse Fourier transformation of the best

averaged XSI profiles [14, 38]. For this we recommend using SCÅTTER [39] (*see* **Note 17**), but there are many other options [40, 41].

3. $P(r)$ functions were normalized to the intra-Au peak to account for variations in concentration.
4. A shift in the peak maximum in the inter-Au region indicates a highly accurate change in inter-particle distance suggesting a change in the substrate. For the example of MRE11, the inter-Au distance distributions are shifted to lower mean values compared to the substrate alone for samples where nuclease activity was not observed indicating structural changes in the DNA associated with MRE11 binding (Fig. 5). These finding are consistent with both Au-dsDNA substrate lengths.
5. The integration of the inter-Au peaks is used to estimate the relative changes in concentration for intact Au-dsDNA

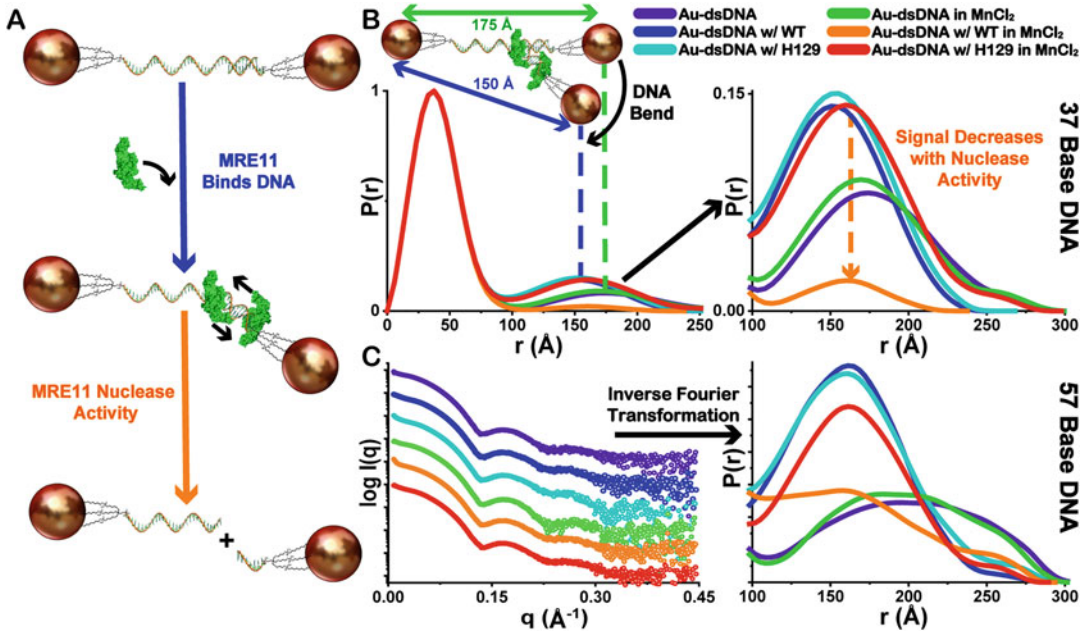


Fig. 5 Demonstration of overall XSI assay scheme. (**a**) The proposed mechanism of MRE11 interaction with intact Au-dsDNA substrates and the subsequent nuclease activity leading to separation of the fixed inter-particle distances as Au-ssDNA. (**b**) Demonstration of the shifts in the distribution of inter-Au electron-pair distances, seen in the normalized $P(r)$ functions (37-bp DNA), to lower mean values compared to the substrate alone representing the structural changes in the Au-dsDNA substrates associated with MRE11 binding. Additionally, a decrease in the amplitude of $P(r)$ corresponding only in the inter-Au regions is observed only with WT-MRE11 in the presence of $MnCl_2$ (orange) suggesting increased MRE11 nuclease activity (Au-dsDNA to Au-ssDNA). Legend for sample identity shown. (**c**) Exemplary experimental XSI curves and derived $P(r)$ functions for 57-bp DNA colored as in the **Panel** (**b**). Curves have scaled $I(q)$ for visualization purposes. $P(r)$ function plot is scaled to show inter-Au distance region. All $P(r)$ functions are normalized to the intra-Au peak to compensate for fluctuations in concentration as depicted in Fig. 1c

substrates after the enzymatic reaction takes place. As expected, MRE11 nuclease activity decreases the population of doubly Au-NP-labelled dsDNA as observed by the decrease in the amplitude of $P(r)$ corresponding only to inter-Au distances. From these XSI assays, we observe that MRE11 is not active when the active site metal (Mn^{2+}) is not present in the reaction buffer or a nuclease-dead mutant (H129N) is used in the reaction instead of the wild-type (WT) enzyme (Fig. 5). These findings are consistent with both Au-dsDNA substrate lengths.

4 Notes

1. Au NP pellets can be disturbed easily if not careful. It is recommended to decant with particles held at side closest to the floor if decanting or to use a pipette.
2. To measure concentrations accurately within the confines of Beer's law, analytes should be diluted so that the absorbance range is with 0.1 and 1 mAu.
3. Extinction Coefficient Au @ 520 nm = 9.69×10^6 L/mol cm, Path Length NanoDrop = 0.1 cm, Extinction Coefficient ssDNA @ 260 nm sequence dependent.

$$\frac{\text{Absorbance} * \text{Dilution Factor}}{\text{Path Length} * \text{Extinction Coefficient}} = \text{Concentration}$$

4. Using no salt and high salt FPLC buffers, a salt gradient from 10 to 1000 mM is created over a period of 50 min (Fig. 1a, b). Sample elution monitored UV-Vis absorption at the Au plasmon maximum of 520 nm. Typical final concentrations for collected conjugates were 0.1–0.2 μM.
5. The catalytic domain of MRE11 used (with Met-to-Leu mutations) in this chapter is not super-soluble when expressed in *E. coli*; however, soluble fraction of the purified enzyme is active in the nuclease reaction.
6. Prior knowledge of enzyme-DNA substrate reaction can be quite useful in designing the substrates. By carefully modifying the substrate, this method can be adapted for other DNA nucleases.
7. Enzymes can be removed from the substrate if desired so by adding proteinase K at the end of the reaction.
8. It is recommended to pre-run the TBE-Urea gel (@200 V for 60 min) prior to running the nuclease reaction products on the gel.
9. DNA is conjugated to Au NPs through Au-S interaction. Strong reducing agents in the reaction buffer can disrupt this

interact and cause aggregation of the Au NPs, especially when exposed to strong synchrotron X-ray radiation. Thus, the reaction conditions need to be optimized accordingly.

10. Samples can be diluted with reaction buffer or concentrated by using 4 kDa centrifuge concentrator tubes and spinning at 10,000 × *g*. Generally, concentrations >100 nM give great scattering signal.
11. SAXS is a contrast measurement, as such the buffer used for subtraction must be as close to the buffer containing the sample as possible. Dialyze samples and use the dialysis buffer for best subtraction.
12. While it should be possible to leverage this technique for use at a researcher's home institution, we also offer a mail-in user program at the SIBYLS beamline where we can help to design and carry out experiments like those mentioned herein. To obtain XSI data collection time at the SIBYLS beamline, please follow the directions on our website https://bl1231.als.lbl.gov/htsaxs/instructions/htsaxs and/or contact us.
13. The use of Python 3 is required. The desired version of Python 3 can be installed by following the instruction on https://www.python.org/downloads/, and any version should work. We also recommend running bash terminal in a Conda, Python 3.7+ environment as it may streamline the setup of the code, but Conda is not required. To set up your own conda environment, follow the instructions on https://docs.conda.io/projects/conda/en/latest/user-guide/install/index.html for setting up miniconda on your system if desired.
14. This code under active development and the newest setup and usage information can be found in the README file at our gitlab for frame_averaging_pipeline at: https://git.bl1231.als.lbl.gov/djrosenberg/frame_averaging_pipeline.git.
15. If running *xsi_bacth_processing.sh* on Windows or Mac, you will likely need an Xserver (https://kb.thayer.dartmouth.edu/article/336-x11-for-windows-and-mac) to run graphical interfaces, and you will also need Eye of Gnome (eog) installed in your terminal if you don't already have it (*see* **Note 16 below**). Alternatively, just open the .png images in your preferred image viewer, and *ignore* **Notes 15 and 16**.
16. The *xsi_bacth_processing.sh* will prompt you if "Eye of Gnome (eog) could not be found. Please use your prefer image viewer to view .png files in: *username_date*_Ave_Graphs folder." To install eog:

```
$ eog --version # first check if you have eog installed
$ sudo apt-get install eog #if you don't have it (ubuntu)
$ sudo yum install eog #if you don't have it (centos/redhat)
```

17. The SAXS analysis software SCÅTTER can be downloaded from http://www.bioisis.net/tutorials, and the website includes tutorials for its use.

Acknowledgments

The work is supported by R35 CA220430, by the Cancer Prevention Research Institute of Texas (CPRIT-grant#RP180813), and by a Robert A. Welch Chemistry Chair. Efforts to apply SAXS to characterize eukaryotic pathways relevant to human cancers and merge nano- to mesoscale structures are supported in part by National Cancer Institute grants Structural Biology of DNA Repair (SBDR) CA092584. SAXS data was collected at the Advanced Light Source (ALS) beamline SIBYLS which is supported by the DOE-BER IDAT DE-AC02-05CH11231 and NIGMS ALS-ENABLE (P30 GM124169 and S10OD018483).

References

1. Syed A, Tainer JA (2018) The MRE11–RAD50–NBS1 complex conducts the orchestration of damage signaling and outcomes to stress in DNA replication and repair. Annu Rev Biochem 87:263–294. https://doi.org/10.1146/annurev-biochem-062917-012415
2. Paull TT (2018) 20 years of Mre11 biology: no end in sight. Mol Cell 71:419–427. https://doi.org/10.1016/j.molcel.2018.06.033
3. Stracker TH, Petrini JHJ (2011) The MRE11 complex: starting from the ends. Nat Rev Mol Cell Biol 12:90–103. https://doi.org/10.1038/nrm3047
4. Hopfner K-P, Karcher A, Craig L et al (2001) Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. Cell 105:473–485. https://doi.org/10.1016/S0092-8674(01)00335-X
5. Williams RS, Moncalian G, Williams JS et al (2008) Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. Cell 135:97–109. https://doi.org/10.1016/j.cell.2008.08.017
6. Park YB, Chae J, Kim YC, Cho Y (2011) Crystal structure of human Mre11: understanding tumorigenic mutations. Structure 19:1591–1602. https://doi.org/10.1016/j.str.2011.09.010
7. Shibata A, Moiani D, Arvai AS et al (2014) DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. Mol Cell 53:7–18. https://doi.org/10.1016/j.molcel.2013.11.003
8. Chen C, Hildebrandt N (2020) Resonance energy transfer to gold nanoparticles: NSET defeats FRET. TrAC Trends Anal Chem 123:115748. https://doi.org/10.1016/j.trac.2019.115748
9. Hura GL, Menon AL, Hammel M et al (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). Nat Methods 6:606–612. https://doi.org/10.1038/nmeth.1353
10. Brosey CA, Tainer JA (2019) Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology. Curr Opin Struct Biol 58:197–213. https://doi.org/10.1016/j.sbi.2019.04.004
11. Tang HYH, Tainer JA, Hura GL (2017) High resolution distance distributions determined by X-ray and neutron scattering. In: Chaudhuri B, Muñoz IG, Qian S, Urban VS (eds) Biological small angle scattering: techniques, strategies and tips. Springer, Singapore, pp 167–181
12. Rambo RP, Tainer JA (2013) Super-resolution in solution X-ray scattering and its applications to structural systems biology. Annu Rev Biophys 42:415–441. https://doi.org/10.1146/annurev-biophys-083012-130301
13. Rambo RP, Tainer JA (2013) Accurate assessment of mass, models and resolution by small-angle scattering. Nature 496:477–481. https://doi.org/10.1038/nature12070

14. Putnam CD, Hammel M, Hura GL, Tainer JA (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. Q Rev Biophys 40:191–285. https://doi.org/10.1017/S0033583507004635
15. Rambo RP, Tainer JA (2011) Characterizing flexible and instrinsically unstructured biological macromolecules by SAS using the Porod-Debye Law. Biopolymers 95:559–571. https://doi.org/10.1002/bip.21638
16. Young T (1804) I. The Bakerian Lecture. Experiments and calculations relative to physical optics. Philos Trans R Soc Lond 94:1–16. https://doi.org/10.1098/rstl.1804.0001
17. Hura GL, Tsai C-L, Claridge SA et al (2013) DNA conformations in mismatch repair probed in solution by X-ray scattering from gold nanocrystals. Proc Natl Acad Sci 110:17308–17313. https://doi.org/10.1073/pnas.1308595110
18. Vainshtein BK, Feigin LA, Lvov YM et al (1980) Determination of the distance between heavy-atom markers in haemoglobin and histidine decarboxylase in solution by small-angle X-ray scattering. FEBS Lett 116:107–110. https://doi.org/10.1016/0014-5793(80)80539-4
19. Mathew-Fenn RS, Das R, Silverman JA et al (2008) A molecular ruler for measuring quantitative distance distributions. PLoS One 3:e3229. https://doi.org/10.1371/journal.pone.0003229
20. Mathew-Fenn RS, Das R, Harbury PAB (2008) Remeasuring the double helix. Science 322:446–449. https://doi.org/10.1126/science.1158881
21. Mastroianni AJ, Sivak DA, Geissler PL, Alivisatos AP (2009) Probing the conformational distributions of subpersistence length DNA. Biophys J 97:1408–1417. https://doi.org/10.1016/j.bpj.2009.06.031
22. Shi X, Herschlag D, Harbury PAB (2013) Structural ensemble and microscopic elasticity of freely diffusing DNA by direct measurement of fluctuations. Proc Natl Acad Sci 110:E1444–E1451. https://doi.org/10.1073/pnas.1218830110
23. Shi X, Beauchamp KA, Harbury PB, Herschlag D (2014) From a structural average to the conformational ensemble of a DNA bulge. Proc Natl Acad Sci 111:E1473–E1480. https://doi.org/10.1073/pnas.1317032111
24. Shi X, Huang L, Lilley DMJ et al (2016) The solution structural ensembles of RNA kink-turn motifs and their protein complexes. Nat Chem Biol 12:146–152. https://doi.org/10.1038/nchembio.1997
25. Zettl T, Mathew RS, Shi X et al (2018) Gold nanocrystal labels provide a sequence–to–3D structure map in SAXS reconstructions. Sci Adv 4:eaar4418. https://doi.org/10.1126/sciadv.aar4418
26. Zettl T, Mathew RS, Seifert S et al (2016) Absolute intramolecular distance measurements with angstrom-resolution using anomalous small-angle X-ray scattering. Nano Lett 16:5353–5357. https://doi.org/10.1021/acs.nanolett.6b01160
27. Claridge SA, Mastroianni AJ, Au YB et al (2008) Enzymatic ligation creates discrete multinanoparticle building blocks for self-assembly. J Am Chem Soc 130:9598–9605. https://doi.org/10.1021/ja8026746
28. Tubbs JL, Latypov V, Kanugula S et al (2009) Flipping of alkylated DNA damage bridges base and nucleotide excision repair. Nature 459:808–813. https://doi.org/10.1038/nature08076
29. Eckelmann BJ, Bacolla A, Wang H et al (2020) XRCC1 promotes replication restart, nascent fork degradation and mutagenic DNA repair in BRCA2-deficient cells. NAR Cancer 2:zcaa013. https://doi.org/10.1093/narcan/zcaa013
30. Thapar R, Wang JL, Hammel M et al (2020) Mechanism of efficient double-strand break repair by a long non-coding RNA. Nucleic Acids Res 48:10953–10972. https://doi.org/10.1093/nar/gkaa784
31. Zandarashvili L, Langelier M-F, Velagapudi UK et al (2020) Structural basis for allosteric PARP-1 retention on DNA breaks. Science 368:eaax6367. https://doi.org/10.1126/science.aax6367
32. Houl JH, Ye Z, Brosey CA et al (2019) Selective small molecule PARG inhibitor causes replication fork stalling and cancer cell death. Nat Commun 10:5654. https://doi.org/10.1038/s41467-019-13508-4
33. Bacolla A, Tainer JA, Vasquez KM, Cooper DN (2016) Translocation and deletion breakpoints in cancer genomes are associated with potential non-B DNA-forming sequences. Nucleic Acids Res 44:5673–5688. https://doi.org/10.1093/nar/gkw261
34. Bacolla A, Ye Z, Ahmed Z, Tainer JA (2019) Cancer mutational burden is shaped by G4 DNA, replication stress and mitochondrial dysfunction. Prog Biophys Mol Biol 147:47–61. https://doi.org/10.1016/j.pbiomolbio.2019.03.004
35. Seo SH, Bacolla A, Yoo D et al (2020) Replication-based rearrangements are a common mechanism for *SNCA* duplication in Parkinson's disease. Mov Disord 35:868–876. https://doi.org/10.1002/mds.27998

36. Classen S, Hura GL, Holton JM et al (2013) Implementation and performance of SIBYLS: a dual endstation small-angle X-ray scattering and macromolecular crystallography beamline at the advanced light source. J Appl Crystallogr 46:1–13. https://doi.org/10.1107/S0021889812048698
37. Dyer KN, Hammel M, Rambo RP et al (2014) High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. Methods Mol Biol 1091:245–258
38. Glatter O (1977) A new method for the evaluation of small-angle scattering data. J Appl Crystallogr 10:415–421. https://doi.org/10.1107/S0021889877013879
39. Tully MD, Tarbouriech N, Rambo RP, Hutin S (2021) Analysis of SEC-SAXS data via EFA deconvolution and scatter. J Vis Exp (167): e61578. https://doi.org/10.3791/61578
40. Hopkins JB, Gillilan RE, Skou S (2017) *BioXTAS RAW*: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. J Appl Crystallogr 50:1545–1553. https://doi.org/10.1107/S1600576717011438
41. Manalastas-Cantos K, Konarev PV, Hajizadeh NR, et al (2021) *ATSAS 3.0*: expanded functionality and new tools for small-angle scattering data analysis. J Appl Crystallogr 54: 343–355. https://doi.org/10.1107/S1600576720013412

Chapter 13

In Vitro Reconstitution of BRCA1-BARD1/RAD51-Mediated Homologous DNA Pairing

Meiling Wang, Cody M. Rogers, Dauren Alimbetov, and Weixing Zhao

Abstract

RAD51-mediated homologous recombination (HR) is a conserved mechanism for the repair of DNA double-strand breaks and the maintenance of DNA replication forks. Several breast and ovarian tumor suppressors, including BRCA1 and BARD1, have been implicated in HR since their discovery in the 1990s. However, a holistic understanding of how they participate in HR has been hampered by the immense challenge of expressing and purifying these large and unstable protein complexes for mechanistic analysis. Recently, we have overcome such a challenge for the BRCA1-BARD1 complex, allowing us to demonstrate its pivotal role in HR via the promotion of RAD51-mediated DNA strand invasion. In this chapter, we describe detailed procedures for the expression and purification of the BRCA1-BARD1 complex and in vitro assays using this tumor suppressor complex to examine its ability to promote RAD51-mediated homologous DNA pairing. This includes two distinct biochemical assays, namely, D-loop formation and synaptic complex assembly. These methods are invaluable for studying the BRCA1-BARD1 complex and its functional interplay with other factors in the HR process.

Key words Homologous recombination, BRCA1-BARD1, RAD51, D-loop formation, Synaptic complex assembly

1 Introduction

During homologous recombination (HR) of DNA double-strand breaks (DSBs), the 5′ strands of both DNA ends are resected to yield 3′ single-stranded DNA (ssDNA) tails [1]. These ssDNA tails are then bound by replication protein A (RPA), which is subsequently replaced by the recombinase enzyme RAD51 to form a highly ordered, helical nucleoprotein complex referred to as the presynaptic filament [2, 3]. The presynaptic filament performs DNA homology search to engage a duplex target (e.g., sister chromatid during mitosis or homologous chromosome during meiosis) and then invades the duplex target to form a nascent heteroduplex DNA joint called the displacement loop (D-loop) [2, 3]. This is followed by DNA synthesis and resolution of DNA intermediates

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_13,

to complete repair [2–4] (Fig. 1). DNA end resection, presynaptic filament assembly, and D-loop formation thus represent three distinct, obligatory steps in HR. Importantly, these key steps of HR involve a suite of well-known tumor suppressors, including BRCA1 and BARD1 [2–5].

BRCA1 is the first familial breast cancer susceptibility locus identified based on linkage analysis in the 1990s [6–9], and *BRCA1* mutations also cause familial ovarian cancer and sporadic cancer in other organs [10]. The BRCA1 protein comprises 1863 amino acids and harbors a RING (Really Interesting New Gene) domain at its N-terminus, 2 copies of the BRCT (BRCA1 Carboxyl-Terminal) repeat, and a coiled-coil domain immediately preceding the BRCT repeats [3, 11]. BRCA1 heterodimerizes with another tumor suppressor BARD1 (BRCA1-Associated RING Domain protein 1) [12], which is 777 residues long and contains a RING domain, 4 ankyrin repeats, and 2 tandem BRCT domains (Fig. 2). The BRCA1-BARD1 complex possesses DNA binding and ubiquitin E3 ligase activities [3, 13, 14] and interacts with over 100 proteins that function in diverse biological processes [15–28], especially RAD51-mediated HR of DSB repair [3, 5, 11, 18, 19]. BRCA1 was discovered to function in HR via its colocalization and coimmunoprecipitation with RAD51 in 1997 [18], and subsequent investigations have suggested a multifaceted role of BRCA1-BARD1 in HR [3, 5, 11]. In part, BRCA1 facilitates DNA end resection by acting as an antagonist of 53BP1 and regulating the MRE11-RAD50-NBS1-CtIP resection nuclease complex and also promotes RAD51-mediated presynaptic filament formation together with another two tumor suppressors BRCA2 and PALB2 [3, 5] (Fig. 1). Our recent biochemical and cellular data have clearly demonstrated a previously unrecognized role of BRCA1-BARD1 in a late step of HR by enhancing RAD51-mediated D-loop formation via the assembly of a three-stranded synaptic complex (Fig. 1). We revealed for the first time that BARD1 (in addition to BRCA1 [29, 30]) provides a major interface for structured DNA species (e.g., D-loop) resembling HR intermediates and both BRCA1 and BARD1 physically interact with RAD51 [14] (Fig. 2**).** Importantly, we and others have shown that the BRCA1-BARD1/RAD51 complex formation is critical for RAD51-mediated HR of DSBs [14] and for DNA replication fork protection in cells [31].

Here, we describe detailed protocols for how to obtain biochemically amenable amounts of high-quality BRCA1-BARD1 preparations and how to set up the in vitro assays for its key properties in RAD51-mediated HR (i.e., D-loop formation and synaptic complex assembly; Fig. 3).

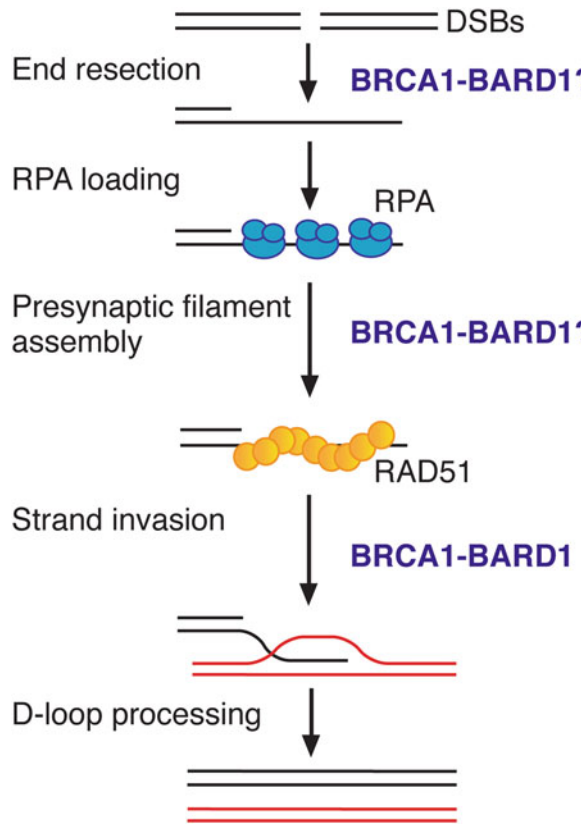


Fig. 1 Schematic of homologous recombination and the potential roles of BRCA1-BARD1 therein

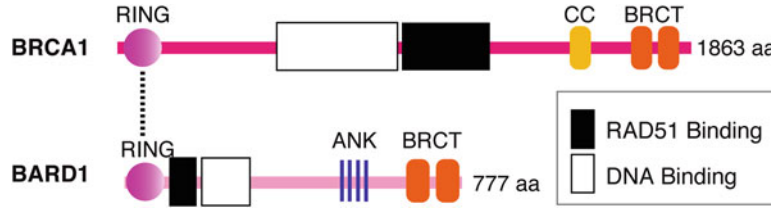


Fig. 2 Functional motifs and domains in BRCA1 and BARD1**.** RAD51 binding domains in BRCA1 and BARD1 are indicated with black boxes, while DNA binding domains in BRCA1 and BARD1 are indicated with white boxes. *RING* really interesting new gene, *ANK* ankyrin repeats, *BRCT* BRCA1 carboxy-terminal, *CC* coiled-coil

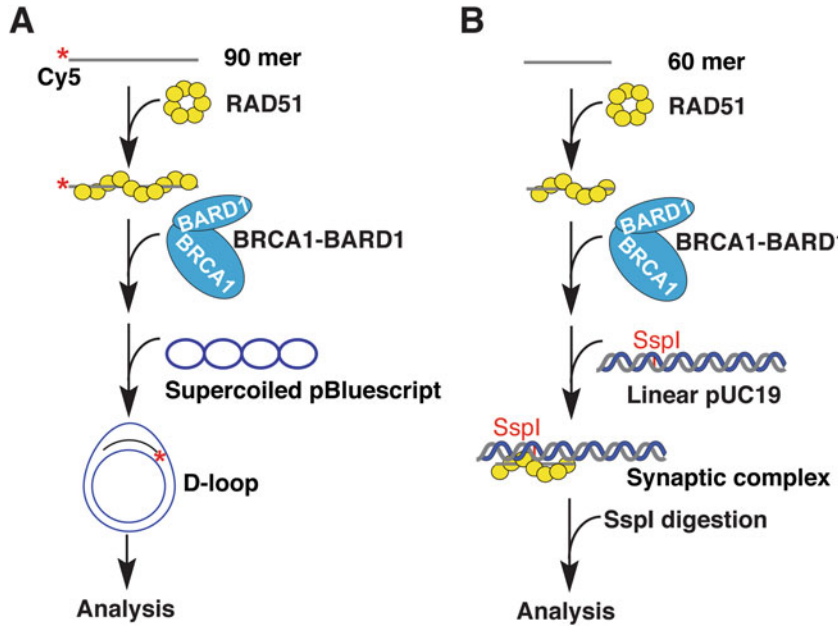


Fig. 3 Schematic of BRCA1-BARD1/RAD51-mediated homologous DNA pairing assays: (**a**) D-loop formation with Cy5-90mer ssDNA; (**b**) synaptic complex formation

2 Materials

2.1 Purification of the BRCA1-BARD1 Complex

2.1.1 Generating Recombinant Bacmid of BRCA1 and BARD1

1. MAX efficiency DH10Bac competent cells (Invitrogen, 10361012).
2. SOC medium: 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose.
3. LB medium: 1% tryptone, 0.5% yeast extract, 1% NaCl.
4. LB plate: LB medium with 15% Agar in 10-cm petri dish.
5. Gen-Kan-Tet-Bluo-gal-IPTG plate: LB agar plate supplemented with 7 μg/mL gentamicin, 50 μg/mL kanamycin, 10 μg/mL tetracycline, 100 μg/mL Bluo-gal (5-bromo-3-indolyl-B-D-galactopyranoside), and 40 μg/mL IPTG (isopropyl β-D-1-thiogalactopyranoside) (*see* **Note 1**).
6. 1× TE buffer: 10 mM Tris–HCl pH 8.0, and 1 mM EDTA.
7. Plasmid Midi prep kit.
8. Digital water bath.
9. Temperature-controlled shaker.
10. Benchtop centrifuges.
11. Ultraviolet (UV) spectrophotometer.

2.1.2 Producing the Recombinant Baculovirus of BRCA1 and BARD1

1. Sf9 cells (Gibco, 11496015).
2. Sf-900™ III SFM-insect cell culture media (Gibco, 12658027).
3. Cellfectin™ II-insect transfection reagent (Invitrogen, 10362100).
4. Primary antibody: Baculovirus envelope gp64 antibody (Thermo Fisher, 14699182).
5. Secondary antibody: Mouse IgG, HRP-linked whole antibody (Invitrogen, 31450).
6. 1× PBS buffer: 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4.
7. PBS-T: 1× PBS buffer with 0.1% Tween-20.
8. Clarity max Western ECL substrate (Bio-Rad, 1705062).
9. 6-well TC-treated plate.
10. 175 cm^2 TC treated flask.
11. Refrigerated incubator.
12. Refrigerated shaker with 2 cm rotation radius.
13. 125–2000 mL narrow mouth Erlenmeyer flask with phenolic screw cap.
14. Refrigerated centrifuge with swing rotor.

2.1.3 Expression and Purification of BRCA1-BARD1 in Hi5 Cells

1. High Five™ cells (BTI-TN-5B1-4).
2. HyClone SFX-insect cell culture media (Gibco, SH30278.02).
3. Lysis buffer: 50 mM Tris–HCl pH 7.5, 500 mM KCl, 1 mM 2-mercaptoethanol, 0.5% NP-40, 5 mM $MgCl_2$, 2 mM ATP, and the following protease inhibitors: aprotinin, chymostatin, leupeptin, and pepstatin A at 3 μg/mL each and 1 mM PMSF.
4. FLAG wash buffer: 25 mM Tris–HCl pH 7.5, 300 mM KCl, 10% glycerol, 0.5 mM EDTA, 0.1% Igepal CA-630, 1 mM 2-mercaptoethanol, 5 mM $MgCl_2$, and 2 mM ATP.
5. FLAG elution buffer: FLAG wash buffer containing 200 μg/mL of FLAG peptide.
6. Base buffer: 25 mM Tris–HCl pH 7.5, 10% glycerol, 0.5 mM EDTA, 0.01% Igepal CA-630, 1 mM 2-mercaptoethanol.
7. T75 buffer: Base buffer with 75 mM KCl. T500 buffer: Base buffer with 500 mM KCl.
8. Mini-Protean 3 gel electrophoresis kit and PowerPac™ power supply.
9. SDS-PAGE gel for protein analysis (10 × 7 cm, 0.75 and 1 mm thick, 10% polyacrylamide, 1% SDS, prepared according to instructions of Mini-Protean 3).

10. 4× SDS-PAGE sample loading buffer: 200 mM Tris–HCl pH 6.8, 10% SDS, 40% glycerol, and 0.04% bromophenol blue and 400 mM DTT.
11. Tris-glycine-SDS (TGS) running buffer: 25 mM Tris–HCl, 192 mM glycine, and 0.1% SDS.
12. Erlenmeyer flasks, 50-mL centrifuge tubes; 1× PBS buffer (*see* Subheading 2.1.2).
13. High speed centrifuge with compatible rotors capable of handling different volumes (50–250 mL) at 4 °C.
14. ÄKTA pure chromatography system with compatible columns.
15. 1 mL HiTrap SP Sepharose HP column.
16. Econo-Column glass chromatography columns.

2.2 RAD51 Purification

Human RAD51 was purified as previously described [32].

2.3 DNA Preparation

2.3.1 Oligo DNA Preparation

1. DNA substrate: the sequences of synthesized oligonucleotides are listed in Table 1.
2. 2× Denaturing gel loading buffer: 94% formamide, 20 mM Tris–HCl pH 7.5, 2 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol.
3. 1× TAE buffer: 40 mM Tris-acetate pH 7.4, and 0.5 mM EDTA.
4. 30% acrylamide/Bis-acrylamide (29/1).
5. Ammonium persulfate; prepare 10% solution in water.
6. Tetramethylethylenediamine solution (TEMED).
7. Urea.
8. 1× TE buffer: 10 mM Tris–HCl pH 8.0, and 1 mM EDTA.
9. Protean II vertical gel electrophoresis kit and PowerPac™ power supply.
10. Denaturing gel for purification of oligos (20 × 16 cm, 1.5 mm thick, containing 7 M urea, 10% polyacrylamide, and 1× TAE, prepared according to instructions of Protein II system).
11. Amicon Ultra-4 (10K NMWL, Millipore) centrifugal filter units.
12. Handheld UV lamp.
13. Refrigerated/heated water bath circulator.
14. High-speed centrifuge (*see* Subheading 2.1.3).

1. LB medium with 100 μg/mL ampicillin.
2. Plasmid Maxi prep kit.
3. EcoR I restriction enzyme.

Table 1
Oligonucleotides are used in the assays

	Length	Sequence	Use
Oligo 1	90 nt	5′-/Cy5/AAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGT TACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT-3′	Subheadings 2.4 and 3.3
Oligo 2	60 nt	5′-AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT TATCAGGGTTATT-3′	Subheadings 2.5 and 3.4
Oligo 3	60 nt	5′ -CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCC AGCAAAAGGCCAGGA-3′	Subheadings 2.5 and 3.4

2.3.2 Preparation of pBluescript SK II + Replicative Form I DNA and Linear pUC19 Plasmid

4. 3 M sodium acetate.
5. 100% ethanol; prepare 70% ethanol solution with ddH_2O.
6. Refrigerated shaker; Erlenmeyer flasks; centrifuges (*see* Subheadings 2.1.1–2.1.3).

2.4 D-Loop Formation with RAD51 and BRCA1-BARD1

1. 0.11 μM Cy5-labeled 90-mer ssDNA (Oligo 1, derived from pBluescript SK II + sequence 1932–2021); 50 nM pBluescript SK II + replicative form I (supercoiled form) DNA.
2. 5× D-loop reaction buffer: 125 mM Tris–HCl pH 7.5 and 5 mM DTT.
3. T200 buffer: 25 mM Tris–HCl pH 7.5, 10% glycerol, 0.5 mM EDTA, 0.01% Igepal CA-630, 1 mM DTT, and 200 mM KCl. T300 buffer: T200 buffer with extra 100 mM KCl.
4. 10% SDS.
5. 1× TAE buffer: 40 mM Tris-acetate pH 7.4 and 0.5 mM EDTA.
6. 10 mg/mL proteinase K.
7. 10 mg/mL BSA.
8. 100 mM $MgCl_2$.
9. 100 mM ATP or 100 mM AMP-PNP.
10. 4× Native gel loading buffer: 30 mM Tris–HCl pH 7.5, 50% glycerol, and 0.1% Orange G.
11. Digital dry bath with metal blocks for 1.5 mL microcentrifuge tubes.
12. Horizontal gel electrophoresis system and PowerPac™ power supply.
13. ChemiDoc MP imaging system (Bio-Rad, 12003154).

2.5 Synaptic Complex Formation Assay

1. 2 μM 60-mer ssDNA (Oligo 2 or Oligo 3; derived from pUC19 sequence 1240–1299 or 240–305, respectively). The DNA fragment of 1240–1299 of pUC19 harbors a SspI restriction enzyme site.
2. 200 nM linear pUC19 plasmid DNA.
3. SspI restriction enzyme .
4. 5× synaptic complex formation buffer: 175 mM Tris–HCl pH 7.5 and 5 mM DTT.
5. SDS; proteinase K; 1× TAE buffer; 4× native gel loading buffer; BSA; $MgCl_2$; ATP; T300 buffer (*see* Subheading 2.4).
6. 1 μg/mL ethidium bromide staining solution.
7. Digital dry bath; horizontal gel electrophoresis system; ChemiDoc MP imaging system (*see* Subheading 2.4).

3 Methods

3.1 Purification of the BRCA1-BARD1 Complex

3.1.1 Generation of BRCA1 and BARD1 Bacmid

1. Thaw DH10Bac competent cells on ice for 20 min, and add 1 μL (~5 ng) of pFastBac-Flag-BRCA1 or pFastBac-His-BARD1 plasmid into aliquoted 20 μL of DH10Bac competent cells.
2. Incubate cells on ice for 30 min, and tap the tubes gently every 10 min.
3. Heat-shock the cells for 45 s in a 42 °C water bath, and transfer the tubes to ice immediately for a 2 min incubation.
4. Add pre-warmed 180 μL SOC medium to the DH10Bac competent cells, and incubate with shaking (225 rpm) for 5 h at 37 °C.
5. Dilute 5 and 10 μL of the *E. coli* culture in 90 μL of SOC medium, and plate on two Gen-Kan-Tet-Bluo-gal-IPTG plates (*see* **Note 2**).
6. Incubate the plates for 48 h at 37 °C to see discrete white and blue colonies on the plates clearly.
7. Pick several large white colonies (surrounded with many blue colonies) and restreak them on a new Gen-Kan-Tet-Bluo-gal-IPTG plate for incubation overnight at 37 °C.
8. Inoculate a single true-white colony confirmed from the above restreaking test into the 3 mL of LB media with 50 μg/mL kanamycin, 7 μg/mL gentamicin, 10 μg/mL tetracycline in the shaker at 37 °C for 14–16 h (*see* **Note 3**).
9. Centrifuge 2 mL of the liquid culture in a 2 mL microcentrifuge tube at 15,000 × *g* for 1 min, and discard supernatant.
10. Extract recombinant bacmid DNA from the cell pellet with 300 μL buffer P1, P2, and P3 agents from a commercially available Midi Prep Kit (e.g., Qiagen), according to manufacturer's instructions (*see* **Note 4**).
11. Centrifuge the mixture at ≥15,000 × *g* at room temperature for 10 min, and transfer 900 μL of the supernatant to 800 μL pre-chilled isopropanol, and incubate on ice or at −20 °C for 30 min.
12. Centrifuge the solution at ≥15,000 × *g* at 4 °C for 15 min (*see* **Note 5**).
13. Discard the supernatant, add 1000 μL cold 70% ethanol to wash the DNA pellet, and centrifuge again for 5 min (*see* **Note 6**).
14. Discard the supernatant, and leave the tube on the bench to air-dry the pellet for ~10 min with cap open.

15. Dissolve the recombinant bacmid DNA in 40 μL 1× TE buffer overnight at 4 °C. Mix gently by tapping the tube, and store at 4 °C before transfection (*see* Subheading 3.1.2). The DNA concentration measured by UV spectrophotometer according to manufacturer's instructions is typically 600–1300 ng/μL (*see* **Note 7**).

3.1.2 BRCA1 and BARD1 Bacmid Transfection and Virus Amplification

1. Verify that the Sf9 cells are in the log phase (2–4 × 10^6 cells/mL) with greater than 95% viability, and seed 8 × 10^5 Sf9 cells into the desired number of wells in a 6-well plate (*see* **Note 8**).
2. Transfect Sf9 cells in a six-well plate with the bacmid DNA using Cellfectin™ II-insect transfection reagent according to manufacturer's instructions.
3. Incubate cells at 27 °C for about 7–8 days (cells release from the plate and appear lysed). Collect the cells in a 15 mL conical tube, and spin down for 5 min at 500 × *g*, 4 °C. The supernatant is P1 virus. Store at 4 °C in the dark or proceed directly to the next P2 amplification step.
4. Transfer all of P1 virus (~2 mL) to fresh 30 mL of 1 × 10^6 Sf9 cells in a 175 cm^2 culture flask.
5. Incubate the cells for about 5–6 days in a 27 °C humidified incubator (until 70–80% cells are lysed), remove the media from the 175 cm^2 culture flask to two 15 mL conical tubes, and centrifuge as before. This supernatant is P2 virus, which can be stored at 4 °C in the dark as above or used directly in the next P3 amplification step.
6. Add 0.5 mL P2 virus into fresh 50 mL 2 × 10^6 Sf9 cells in a 250 mL glass flask, incubate cells at 27 °C with shaking (125 rpm) until 70–80% cells are lysed (about 4–5 days), and collect the virus to a 50 mL centrifuge tube named P3 for storage at 4 °C in the dark (*see* **Note 9**).
7. Estimate the relative virus titer with the dot-blot assay as previously published [33]: spot 2 μL of freshly prepared virus on nitrocellulose membrane, wait 5–10 min to dry, and then perform a standard western blot assay with baculovirus envelope gp64 antibody (*see* **Note 10**).

3.1.3 Expression and Purification of BRCA1-BARD1 in Hi5 Cells

1. Verify that the Hi5 cells are in the log phase (2–3 × 10^6 cells/mL) with greater than 95% viability, and seed 600 mL of Hi5 cells (1 × 10^6 cells/mL) in two 1000 mL flasks (300 mL each) (*see* **Note 11**).
2. Infect 300 mL of the Hi5 cells with 5 mL BRCA1 and 5 mL BARD1 P3 viruses, and incubate at 27 °C for shaking (125 rpm) for 40–44 h (*see* **Note 12**).

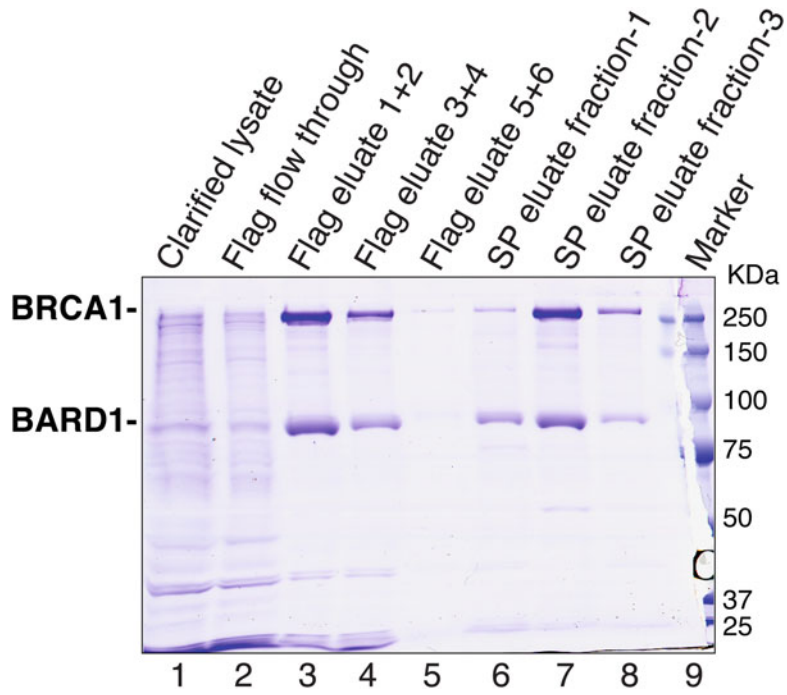


Fig. 4 BRCA1-BARD1 purification. Samples were analyzed by SDS-PAGE and stained with Coomassie Brilliant Blue. Loading order from left to right: clarified lysate after centrifugation, FLAG resin flow through and eluates, SP column peak fractions, Marker

3. Harvest the cells by centrifugation (500 × *g* for 10 min) at 4 °C, discard the supernatant, resuspend the pellet with 50 mL cold 1× PBS buffer, and centrifuge again with 500 × *g* for 10 min at 4 °C. Discard the supernatant, and freeze the pellet in liquid nitrogen for long-term storage at −80 °C.
4. Thaw the frozen cell pellet (~8 g, from 600 mL culture) with 40 mL lysis buffer, and lyse the cells using a Dounce homogenizer type B pestle (30 strokes) during a 20 min incubation on ice.
5. Centrifuge the mixture at 10,000 × *g* for 15 min at 4 °C. Keep a small aliquot of the supernatant (clarified lysate) for gel analysis (Fig. 4). While the lysate is spinning, wash 3 mL anti-FLAG M2 affinity gel with lysis buffer (*see* **Note 13**).
6. Transfer the supernatant onto the pre-equilibrated anti-FLAG M2 affinity resin, and incubate the suspension at 4 °C with gentle agitation for 2 h.
7. Centrifuge the suspension at 1000 × *g* for 2 min, keep a small aliquot of the supernatant (FLAG flow-through) for gel analysis (Fig. 4), and resuspend the resin with 20 mL of FLAG wash buffer after removing the supernatant.

8. Carefully resuspend the resin with FLAG wash buffer several times, and transfer all the FLAG resin into a gravity flow column.
9. Pour 50 mL lysis buffer into the column containing FLAG resin, and let the buffer flowthrough by gravity flow drop by drop. Repeat the process with 50 mL FLAG wash buffer (*see* **Note 14**).
10. Elute the bound protein four times with 2 mL FLAG elution buffer. For every elution step, stop the flow in the column for 15 min, and gently resuspend the resin in the elution buffer. Collect fractions from each elution in separate tubes. If desired, save a small aliquot for gel analysis.
11. Pool the elution fractions (8 mL), and mix them with 32 mL of base buffer (*see* **Note 15**).
12. Load the 40 mL of mixture sample on a 1 mL HiTrap SP Sepharose HP column pre-equilibrated with T75 buffer using a flow rate of 1 mL/min.
13. Develop the SP column with a 12 mL linear gradient of 75–500 mM KCl in base buffer at 0.5 mL/min, and collect 0.5 mL per fraction.
14. Aliquot the peak fractions (should be at 250–350 mM KCl) into 10 μL portions. Additional fractions with lower concentration protein can be kept separately. Snap-freeze all fractions in liquid nitrogen, and store at −80 °C.
15. Mix clarified lysate, FLAG flowthrough, FLAG eluates, and SP peak elutions with 4× SDS-PAGE loading buffer, and analyze them by SDS-PAGE gel electrophoresis in TGS running buffer. The yield of BRCA1-BARD1 (SP fractions) from 600 mL insect cell culture ranges from 400 to 600 μg with a concentration of 50–500 μg/mL (Fig. 4 and *see* **Note 16**).

3.2 DNA Substrate Preparation

3.2.1 Purification of Oligonucleotides

1. Dissolve 200 μg chemically synthesized oligonucleotide (listed in Table 1) in 100 μL 1× TE buffer, and mix with an equal volume of 2× denaturing gel loading buffer.
2. Incubate the sample at 95 °C for 5 min, and then chill on ice for 5 min.
3. Load on the denaturing gel containing 7 M urea, 10% acrylamide, and 1× TAE, and run the gel at 150 V for 4 h at 55 °C using the vertical gel electrophoresis system in conjunction with a circulating water bath.
4. After electrophoresis, disassemble the gel apparatus, and place the gel on plastic Saran wrap over a sheet of white paper.

5. Use a handheld UV lamp to locate the DNA band in the gel, and excise the gel piece containing DNA with a razor blade. Smash the gel piece into small (1 mm) fragments.
6. Transfer the gel fragments to a 15 mL conical tube, and add 4 mL of 1× TAE buffer for soaking overnight with gentle agitation at 4 °C.
7. Transfer the buffer containing DNA (avoid gel fragments) into an Amicon Ultra-4 filter unit, and concentrate by centrifugation at 6000 × *g* until the volume of the solution is reduced to ~100 μL.
8. Add 500 μL of fresh 1× TAE, and repeat centrifugation to remove residual urea and acrylamide. Repeat this step twice.
9. Remove the concentrated DNA solution to a fresh 1.5 mL microcentrifuge tube, and measure the concentration of DNA via absorbance at 260 nm (*see* Subheading 3.1.1). Make ~10 μL aliquots, and store at −20 °C. The yield is typically 40–50 μg.

3.2.2 Preparation of pBluescript SK II + Replicative Form I DNA and Linear pUC19 Plasmid

1. Inoculate a single colony of DH5α cells freshly transformed with either pBluescript SK II + or pUC19 into 100 mL LB medium with 100 μg/mL ampicillin, and culture the cells for 12–13 h to OD_{600} 1.2–1.5.
2. Harvest the cells after centrifugation, and extract plasmid DNA using Qiagen plasmid Maxi prep kit according to manufacturer's instructions (*see* **Note 17**).
3. To obtain linear pUC19, digest 200 μg of pUC19 with 200 units of EcoRI in a 200 μL reaction according to the manufacturer's instructions. Precipitate the DNA with 20 μL 3 M NaOAc (pH 5.2) and 600 μL ice-cold 100% ethanol at −20 °C overnight. The next day centrifuge precipitated DNA (15,000 × *g*) for 30 min at 4 °C, wash the DNA pellet twice with 0.5 mL ice-cold 70% ethanol (with 10 min, 15,000 × *g* spins in between), air dry, and resuspend the DNA in 50 μL 1× TE buffer. The concentration of DNA is measured as above (*see* Subheading 3.1.1).

3.3 D-Loop Formation with RAD51 and BRCA1-BARD1

1. Set up reaction master mixes in a 1.5 mL microcentrifuge tube for the desired number of reactions, and aliquot 7.5 μL into individual tubes (the total final reaction volume will be 12.5 μL):
   (a) 2.5 μL of 5× D-loop reaction buffer.
   (b) 0.125 μL 10 mg/mL BSA.
   (c) 0.25 μL 100 mM ATP or AMP-PNP (*see* **Note 18**).
   (d) 0.125 μL 100 mM $MgCl_2$.

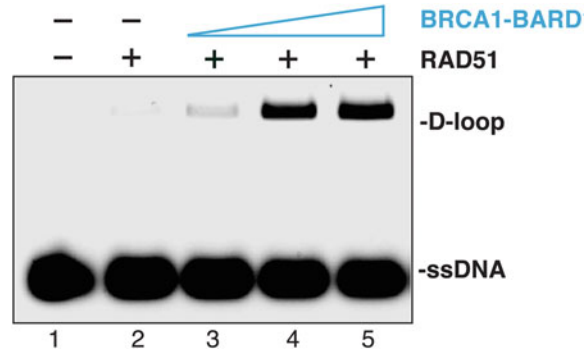


Fig. 5 Promotion of RAD51-mediated D-loop formation by BRCA1-BARD1. D-loop reactions were performed with human RAD51 and different amounts of the BRCA1-BARD1 (30, 61, and 92 nM)

(e) 1 μL 0.11 μM Cy5-labeled 90mer oligonucleotide.

(f) 3.5 μL ddH_2O.

2. Add 1 μL of 3.75 μM human RAD51 (1:3 for RAD51/nucleotide) in T200 buffer to the mixture, mix gently without generating air bubbles, and incubate at 37 °C for 5 min (*see* **Note 19**). For negative control, substitute 1 μL T200 instead of RAD51.
3. Add 1–3 μL BRCA1-BARD1 in T300 to desired concentration, and supplement T300 buffer as necessary such that the total volume of protein and T300 is 3 μL. Thus, the final KCl concentration in the reaction from protein/buffer is ~90 mM. Incubate at 37 °C for 5 min (*see* **Note 20**).
4. Add 1 μL 50 nM pBluescript SK II + replicative form I DNA to initiate the reaction, and incubate at 37 °C for 10 min.
5. Stop the reaction with 0.5 μL 10% SDS and 1 μL of 10 mg/mL proteinase K at 37 °C for 5 min.
6. Add 4 μL 4× native gel loading buffer, and load the sample in a 1% agarose gel for running in 1× TAE buffer.
7. Run the gel at 100 V for 45 min at room temperature, and rinse the gel with new 1× TAE buffer.
8. Image the gel with the Cy5 channel of the ChemiDoc MP Imaging System, and analyze the data with the Image Lab software (Fig. 5).

3.4 Synaptic Complex Formation Assay

1. Set up reaction master mixes in a 1.5 mL microcentrifuge tube for the desired number of reactions, and aliquot 7 μL into individual tubes (the total final reaction volume will be 10 μL):

(a) 2.5 μL of 5× D-loop reaction buffer.

(b) 0.1 μL 10 mg/mL BSA.

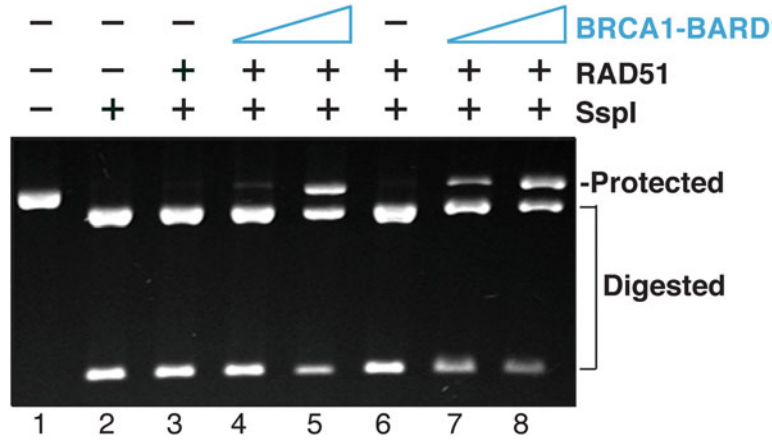


Fig. 6 Promotion of RAD51-mediated synaptic complex assembly by BRCA1-BARD1. Synaptic complex formation by the human RAD51-ssDNA filament and different amounts of BRCA1-BARD1 (75 and 150 nM). Lanes 3–5 were from the reactions with 1 mM $MgCl_2$ and 2 mM ATP, while lanes 6–8 were from the reactions with 2 mM $MgCl_2$ and 2 mM ATP

(c) 0.2 μL 100 mM ATP.

(d) 0.2 μL 100 mM $MgCl_2$.

(e) 1 μL 2 μM 60 mer oligonucleotide.

(f) 3.5 μL ddH_2O.

2. Add 1 μL 40 μM human RAD51 (1:3 for RAD51/nucleotide) in T300 buffer, and mix gently while avoiding the generation of air bubbles. Incubate at 37 °C for 5 min.
3. Add 1 μL BRCA1-BARD1 in T300 to desired concentration or 1 μL T300 buffer for a negative control, making sure that the final KCl concentration in the reaction from protein/buffer is ~60 mM, and incubate at 37 °C for 5 min.
4. Add 1 μL 200 nM linear pUC19 DNA to initiate the reaction, and incubate at 37 °C for 5 min.
5. Add 2.5 units of SspI enzyme, and incubate at 37 °C for 10 min (*see* **Note 21**).
6. Stop the reaction with 0.5 μL 10% SDS and 1 μL of 10 mg/mL proteinase K at 37 °C for 15 min.
7. Add 3 μL 4× native gel loading buffer, load the sample in a 1% agarose gel in 1× TAE buffer, and run the gel at 120 V for 60 min.
8. Stain the gel with ethidium bromide staining solution to visualize DNA species by ChemiDoc MP Imaging System, and analyze the data with the Image Lab software (Fig. 6).

4 Notes

1. The Gen-Kan-Tet-Bluo-gal-IPTG plate should be made fresh on the day of transformation. Typically, spread the mixture of gentamicin, tetracycline, Bluo-gal, and IPTG on the premade Kanamycin LB agar plate with glass beads and incubate the plate at 37 °C for 5–6 h before use.
2. Depending on the competency of the DH10Bac cells used in the transformation, it might be necessary to increase the amount of cells plated (>10 μL of the suspension) to get enough white colonies. Adding too many cells makes harvesting discrete and white colonies difficult.
3. It is recommended to culture DH10Bac cells for 14–16 h and make sure that OD_{600} is about 1.2–1.5 before harvesting, in order to have a high, reproducible yield of bacmid. Make a glycerol stock for the DH10Bac cells harboring the recombinant bacmid, which can be used to reproduce fresh recombinant bacmid for new transfection.
4. Do not vortex, and gently handle the DNA sample after adding the P2 buffer because the recombinant bacmid DNA is greater than 135 kb and prone to breakage.
5. After centrifugation, a small while pellet on bottom of the microcentrifuge tube should be expected. However, the pellet may be small and difficult to visualize. Sometimes, it is necessary to increase the time of the incubation at −20 °C and/or centrifugation.
6. 70% ethanol in this step is used to wash the pellet to remove the residual proteins in the DNA pellet. It is important to avoid disturbing the pellet during the washes.
7. Do not allow the DNA pellet to overdry as it will damage the recombinant bacmid and also make it much harder to dissolve in solution. Store the dissolved recombinant bacmid DNA at 4 °C, and make a fresh one for new transfection once it has been stored at 4 °C for over 2 months.
8. For starting new Sf9 cells from a liquid nitrogen frozen stock, maintain the Sf9 cells on 10-cm dish plate for three generations before transfer to the shaker flask for suspension culture. Normally, the suspension Sf9 cells need to be diluted to 1×10^6 cells/mL every other day. Before transfection with the recombinant bacmid DNA, Sf9 cells should be mixed with trypan blue (Sigma-Aldrich, T6146) to check the viability.
9. The shaker speed and the culture volume for cells in suspension are critical for optimal cell growth and robust protein expression. For Eppendorf INNOVA S44I with shaking at 125 rpm/min, recommended culture volumes are 40–75 mL of cells in 250 mL flasks, 80–150 mL of cells in 500 mL flasks, and 160–300 mL of cells in 1000 mL flasks.

10. The signal intensity of the dots after being developed with ECL reagents can be quantified by Image Lab or similar analytic software for estimating the relative titer of viruses, which is useful for determining the volumes of viruses to be used in the following procedures.
11. For starting new Hi5 cells from a liquid nitrogen frozen stock, maintain the Hi5 cells on 10-cm dish plate for three generations before transfer to the shaker flask for suspension culture. Suspension cultures of Hi5 cells need be passaged to 1×10^6 cells/mL every day.
12. In order to find the best conditions for high expression of BRCA1-BARD1 in large-scale of Hi5 cells, several 50 mL culture volumes of Hi5 cells with different amounts/ratios of both BRCA1 and BARD1 viruses are used for initial small-scale expression tests: Harvest 1.5 mL at the different time points after infection (40–60 h), IP them with anti-FLAG resin, and analyze via SDS-PAGE to compare yields.
13. Sufficient amounts of protease inhibitors should be included in the lysis buffer to reduce the degradation of BRCA1-BARD1 during purification. PMSF is added again into the mixture during the incubation with anti-FLAG resin.
14. Extensive washing with lysis buffer and FLAG wash buffer containing Mg^{2+} and ATP is necessary as it can remove potential non-specific interactors, including molecular chaperones (e.g., Hsp70) and some endo- or exonucleases.
15. Base buffer without any salt is used for diluting the elution fractions to make sure that the conductivity of the final solution is lower than T75 buffer, which is used to equilibrate the HiTrap SP Sepharose HP column. However, do not dilute too much or too fast with base buffer as it can cause protein precipitation.
16. This high yield of BRCA1-BARD is dependent on healthy Hi5 cells, high-quality virus, and optimal infection time. The gel filtration step to further purify BRCA1-BARD1 is optional as the high concentration fractions from the SP column have very similar activity to preps from gel filtration in our biochemical assays. Notably, the purification process should be completed within 16 h, as longer purification process time will increase protein aggregation and degradation.
17. Traditionally, an ideal but complicated protocol without alkaline cell lysate should be applied to separate the intact supercoiled DNA from other forms of plasmid DNA. However, the Qiagen Maxi prep kit using a new colony in a fresh culture is simpler to apply and achieves high quality of plasmid with more than 90% in the supercoiled form.

18. AMP-PNP is used as it is a nonhydrolyzable analogue of ATP that stabilizes the formed RAD51-ssDNA presynaptic filament. This significantly increases the yield of D-loop formation in the assay.
19. Human RAD51 is stable, and its activity does not significantly change if left on ice for a week. The optimal ratio of RAD51 protein to ssDNA to generate maximum D-loop and synaptic complex formation is 1:3 (i.e., one RAD51 to 3 nucleotides). Usually, a titration with various amounts of RAD51 to a fixed amount of ssDNA is required for optimizing the reactions due to small variabilities between RAD51 preps.
20. As BRCA1-BARD1 is unstable, do not freeze and thaw preps repeatedly. High salt (>90 mM) inhibits the activity of RAD51 in D-loop formation assay significantly, so it is critical not to introduce more salt over 90 mM in the reaction. To accommodate the low concentration of purified BRCA1-BARD1, the D-loop assay has been set with a concentration of RAD51 and DNA threefold less than what would normally be used.
21. The amount of SspI and incubation time should be optimized using 200 nM linear pUC19 DNA in 10 μL reactions before the actual synaptic complex formation assay.

Acknowledgments

We thank members of the Zhao laboratory (O'Taveon Fitzgerald, Wenjing Li and Yuxin Huang) for critical reading on the manuscript and Dr. Jeffrey D. Parvin for providing the pFastBac-Flag-BRCA1 and pFastbac-BARD1 plasmids. This work was supported by foundation grants (V Scholar V2019.Q13 from V Foundation for Cancer Research and Young Investigator Award from Max and Minnie Tomerlin Voelcker Fund) and startup funds from University of Texas Health Science Center at San Antonio (all to W.Z.). C.M.R. is supported by a T32 (AG 021890) through the UT Health San Antonio Biology of Aging program.

References

1. Symington LS (2014) DNA repair: making the cut. Nature 514(7520):39–40
2. San Filippo J, Sung P, Klein H (2008) Mechanism of eukaryotic homologous recombination. Annu Rev Biochem 77:229–257
3. Zhao W, Wiese C, Kwon Y et al (2019) The BRCA tumor suppressor network in chromosome damage repair by homologous recombination. Annu Rev Biochem 88:221–245
4. Jasin M, Rothstein R (2013) Repair of strand breaks by homologous recombination. Cold Spring Harb Perspect Biol 5(11):a012740
5. Prakash R, Zhang Y, Feng W et al (2015) Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. Cold Spring Harb Perspect Biol 7(4):a016600

6. Hall JM, Lee MK, Newman B et al (1990) Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250(4988):1684–1689
7. Futreal PA, Liu Q, Shattuck-Eidens D et al (1994) BRCA1 mutations in primary breast and ovarian carcinomas. Science 266(5182):120–122
8. Godwin AK, Vanderveer L, Schultz DC et al (1994) A common region of deletion on chromosome 17q in both sporadic and familial epithelial ovarian tumors distal to BRCA1. Am J Hum Genet 55(4):666–677
9. Miki Y, Swensen J, Shattuck-Eidens D et al (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 266(5182):66–71
10. Petrucelli N, Daly MB, Pal T (2016) BRCA1- and BRCA2-associated hereditary breast and ovarian cancer. In: Adam MP, Ardinger HH, Pagon RA et al (eds) GeneReviews. University of Washington, Seattle
11. Christou CM, Kyriacou K (2013) BRCA1 and its network of interacting partners. Biology (Basel) 2(1):40–63
12. Wu LC, Wang ZW, Tsan JT et al (1996) Identification of a RING protein that can interact in vivo with the BRCA1 gene product. Nat Genet 14(4):430–440
13. Wu W, Koike A, Takeshita T et al (2008) The ubiquitin E3 ligase activity of BRCA1 and its biological functions. Cell Div 3:1
14. Zhao W, Steinfeld JB, Liang F et al (2017) BRCA1-BARD1 promotes RAD51-mediated homologous DNA pairing. Nature 550(7676):360–365
15. Schlacher K, Wu H, Jasin M (2012) A distinct replication fork protection pathway connects Fanconi anemia tumor suppressors to RAD51-BRCA1/2. Cancer Cell 22(1):106–116
16. Hatchi E, Skourti-Stathaki K, Ventz S et al (2015) BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. Mol Cell 57(4):636–647
17. Silver DP, Livingston DM (2012) Mechanisms of BRCA1 tumor suppression. Cancer Discov 2(8):679–684
18. Scully R, Chen J, Plug A et al (1997) Association of BRCA1 with Rad51 in mitotic and meiotic cells. Cell 88(2):265–275
19. Moynahan ME, Chiu JW, Koller BH et al (1999) Brca1 controls homology-directed DNA repair. Mol Cell 4(4):511–518
20. Caestecker KW, Van de Walle GR (2013) The role of BRCA1 in DNA double-strand repair: past and present. Exp Cell Res 319(5):575–587
21. Ray Chaudhuri A, Callen E, Ding X et al (2016) Replication fork stability confers chemoresistance in BRCA-deficient cells. Nature 535(7612):382–387
22. Willis NA, Chandramouly G, Huang B et al (2014) BRCA1 controls homologous recombination at Tus/Ter-stalled mammalian replication forks. Nature 510(7506):556–559
23. Savage KI, Gorski JJ, Barros EM et al (2014) Identification of a BRCA1-mRNA splicing complex required for efficient DNA repair and maintenance of genomic stability. Mol Cell 54(3):445–459
24. Kawai S, Amano A (2012) BRCA1 regulates microRNA biogenesis via the DROSHA microprocessor complex. J Cell Biol 197(2):201–208
25. Kleiman FE, Manley JL (1999) Functional interaction of BRCA1-associated BARD1 with polyadenylation factor CstF-50. Science 285(5433):1576–1579
26. Kleiman FE, Manley JL (2001) The BARD1-CstF-50 interaction links mRNA 3′ end formation to DNA damage and tumor suppression. Cell 104(5):743–753
27. Roy R, Chun J, Powell SN (2012) BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat Rev Cancer 12(1):68–78
28. Deng CX (2006) BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. Nucleic Acids Res 34(5):1416–1426
29. Paull TT, Cortez D, Bowers B et al (2001) Direct DNA binding by Brca1. Proc Natl Acad Sci U S A 98(11):6086–6091
30. Masuda T, Xu X, Dimitriadis EK et al (2016) "DNA binding region" of BRCA1 affects genetic stability through modulating the intra-S-phase checkpoint. Int J Biol Sci 12(2):133–143
31. Daza-Martin M, Starowicz K, Jamshad M et al (2019) Isomerization of BRCA1-BARD1 promotes replication fork protection. Nature 571(7766):521–527
32. Sigurdsson S, Trujillo K, Song B et al (2001) Basis for avid homologous DNA strand exchange by human Rad51 and RPA. J Biol Chem 276(12):8798–8806
33. Anand R, Pinto C, Cejka P (2018) Methods to study DNA end resection I: recombinant protein purification. Methods Enzymol 600: 25–66

Chapter 14

Purification of DNA-Dependent Protein Kinase Catalytic Subunit (DNA-PKcs) from HeLa Cells

Linda Lee, Yaping Yu, and Susan P. Lees-Miller

Abstract

With a predicted molecular mass of 469 kDa, expression of recombinant DNA-dependent protein kinase catalytic subunit (DNA-PKcs) is challenging. However, DNA-PKcs is relatively abundant in human cells, making it possible to purify the endogenous protein. Here we describe a method to purify DNA-PKcs and its binding partner Ku70/80 from HeLa cells and describe conditions for transfer of DNA-PKcs and other large polypeptides for immunoblotting.

Key words Protein purification, DNA-PKcs, Ku70/80, HeLa cells

1 Introduction

The phosphatidylinositol 3 kinase-like (PIKK) family of protein kinases, DNA-dependent protein kinase catalytic subunit (DNA-PKcs), Ataxia-Telangiectasia Mutated (ATM), and ATM-and Rad3-related (ATR), play important roles in the cellular response to DNA damage [1]. At 4128 amino acids and approximately 469 kDa, expression of recombinant DNA-PKcs is challenging. However, DNA-PKcs is relatively abundant in human cells, making it feasible to purify the endogenous protein. Indeed, we have described methods for purification of DNA-PKcs from human placenta [2] and HeLa cells grown in suspension [3, 4]. Another challenge in working with DNA-PKcs, again due to its large size, is difficulty in transferring it from gels to membrane for Western blotting. In this chapter we will describe, in detail, methods to purify DNA-PKcs and its DNA binding subunit KU70/80 from HeLa cells. Furthermore, we will describe a method to improve the transfer of DNA-PKcs (and other large polypeptides) for immunoblot analysis.

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_14,

2 Materials

Prepare all solutions using ultrapure Milli-Q grade water and analytical grade reagents (*see* **Note 1**).

2.1 Purification of DNA-PKcs and Ku70/80 from HeLa Cells

1. *10× Tris Column Buffer (TCB)*: Composition: 500 mM Tris-HCl, pH 8.0, 50% (v/v) glycerol, 2 mM EDTA. For 2 L 10× TCB: Weigh 121 g of Tris-base, and add ~500 mL H_2O, 1000 mL glycerol, 8 mL 0.5 M EDTA pH 8.0, and then water to 1.75 L. Adjust pH to 8 with concentrated HCl and then top off to 2 L with water.
2. *1× TCB*: Dilute 10× TCB (above) 1–10, adding the required amount of 2 M KCl to obtain working solutions with the appropriate salt concentration. For example, for 1 L of 1× TCB containing 50 mM KCl, add 100 mL 10× TCB, 25 mL 2 M KCl, and water to 1 L, and re-adjust pH to 8.0 if needed. Store working solutions of TCB at 4 °C, and add DTT to 1 mM and protease inhibitors immediately before use (*see* Subheading 2.1, **item 10** for details).
3. *2 M KCl*: Weigh out 745.5 g KCl, add ~4 L water, and stir. When dissolved, make up to 5 L.
   Store at 4 °C.
4. *Low Salt Buffer (LSB)*: Composition: 10 mM HEPES-NaOH pH 7.4, 25 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 0.1 mM EDTA. To prepare 1 L 1× LSB, add 10 mL 1 M HEPES-NaOH pH 7.4, 12.5 mL 2 M KCl, 2 mL 5 M NaCl, 1 mL 1 M $MgCl_2$, 0.20 mL 0.5 M EDTA, and water to 800 mL. Adjust pH to 7.4 with 5 M NaOH. Add water to 1 L. Autoclave or sterile filter. Store 4 °C. Add DTT to 1 mM and protease inhibitors immediately before use.
5. *High Salt Buffer (HSB)*: Composition: 50 mM Tris-HCl pH 8.0, 5% (v/v) glycerol, 10 mM $MgCl_2$, 0.2 mM EDTA, 400 mM KCl. To prepare 1 L 1× HSB, add 50 mL 1 M Tris-HCl pH 8.0, 50 mL glycerol, 10 mL 1 M $MgCl_2$, 0.40 mL 0.5 M EDTA, 200 mL 2 M KCl, and water to 800 mL. Adjust pH to 8.0. Add water to 1 L. Store 4 °C. Add DTT to 1 mM and protease inhibitors immediately before use.
6. *1 M DTT*: Dissolve 1.54 g DTT in 10 mL water. Store in 1 mL aliquots at −20 °C. Add to buffers to 1 mM immediately before use.
7. *10× HEPES Column Buffer (HCB)*: Composition: 500 mM HEPES-NaOH pH 7.5, 50% (v/v) glycerol, 2 mM EDTA. To prepare 1 L 10× HCB, dissolve 119 g HEPES and 4 mL 0.5 M EDTA pH 8.0 in about 300 mL water. Adjust pH to 7.5 with

5 M NaOH. Add 500 mL glycerol, re-adjust pH to 7.5, and add water to 1 L. Autoclave or sterile filter. Store at 4 °C.

8. *1× HCB*: Dilute 10× HCB 1–10, adding the required amount of 2 M KCl to obtain working solutions with the appropriate salt concentration. Adjust pH as necessary. Store working solutions of HCB at 4 °C, and add DTT to 1 mM and protease inhibitors immediately before use.
9. *20% (v/v) Tween 20*: To 100 mL of Tween 20, add water to 500 mL. Mix well. Store at 4 °C.
10. *Protease Inhibitors* (*see* **Note 2**).
    (a) *5 M benzamidine*: Dissolve 6.0 g of benzamidine in 10 mL water. Add 0.1 mL per L to LSB for S10 and to HSB for P10 for 0.5 mM final concentration.
    (b) *0.2 M PMSF*: Dissolve 1.74 g of PMSF in 50 mL methanol. Store at 4 °C. Add 1 mL dropwise, with gentle mixing, per 1 L sample for 0.2 mM final concentration.
    (c) *Pepstatin A*: Dissolve 5 mg in 1 mL methanol, and store at −20 °C. Add 10 µL per 100 mL sample for 0.5 µg/mL.
    (d) *Aprotinin and Leupeptin*: Dissolve 5 mg of each protease inhibitor in 1 mL water and store at −20 °C. Add 10 µL per 100 mL of pooled protein fraction for 0.5 µg/mL.
11. *HeLa cells*: If you have the capacity, you can grow your own HeLa-S3 cells in suspension. If not, you can purchase cell pellets resuspended in hypotonic buffer as described below. We purchase cell pellets from 50 to 100 L quantities of HeLa-S3 cells from the Cell Culture Company (https://cellculturecompany.com/national-cell-culture-center/). The cells are grown in suspension to approximately 0.5×10^6 cells/mL in Joklik's modified MEM with 5% Newborn Calf Serum, spun down at 1500 × *g*, washed two times in PBS, and then resuspended in Low Salt Buffer (LSB) containing 1 mM DTT and protease inhibitors. The resuspended cells are then quick frozen in aliquots (each aliquot from 25 L of cell suspension) and shipped by courier on dry ice. The cells can be stored at −80 °C for several months prior to use.

2.2 SDS PAGE and Western Blotting of DNA-PKcs

1. *Buffers and reagents for resolving gel for DNA-PKcs*: Make up or purchase the following reagents: 30% acrylamide in water (*see* **Note 3**); 2% bis-acrylamide in water (*see* **Note 3**); 1 M Tris-HCl, pH 8.8; 20% (w/v) SDS (*see* **Note 4**); 10% (w/v) ammonium persulfate (*see* **Note 5**); *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED).
2. *Buffers and reagents for stacking gel for DNA-PKcs*: Make up or purchase the following reagents: 30% acrylamide in water (*see* **Note 3**); 2% bis-acrylamide in water (*see* **Note 3**); 1 M

Tris-HCl pH 6.8; 20% (w/v) SDS (*see* **Note 4**); 10% (w/v) ammonium persulfate (*see* **Note 5**); TEMED.

3. *5× SDS Running Buffer for DNA-PKcs*: Weigh out 30.0 g Tris Base, 144.2 g glycine, and 5 g SDS (*see* **Note 4**). Make to 1 L with water. Do not adjust pH. Do not autoclave. Store at RT. Dilute 1:5 (e.g., 100–500 mL) with water for use.
4. *10× Electroblot for DNA-PKcs Western Blots*: Weigh out 58 g Tris base and 29.3 g glycine. Add water to 800 mL. Store at RT. For use: pour 100 mL 10× Electroblot-DNA-PKcs into a cylinder, and make to 700 mL with water. Add 200 mL methanol. Add 1.8 mL 20% SDS. The final concentration (at 1×) is 48 mM Tris/39 mM glycine/~0.02% SDS. Store at 4 °C.

3 Methods

3.1 Purification of DNA-PKcs and Ku70/80 from HeLa Cells

1. **Preparation of columns:** Before starting the prep, pour and equilibrate the columns as described below (*see* **Note 6**).
   (a) *DEAE Fast Flow*: Pour a slurry of DEAE Fast Flow resin in water into a 5 cm diameter low-pressure column to approximately 15 cm in height (column volume ~ 300 cm^3). Before use, equilibrate the column in 1× TCB containing 50 mM KCl, 1 mM DTT, and protease inhibitors (referred hereafter as TCB50+) (*see* Subheading 2.1, **item 10**).
   (b) *SP Sepharose Fast Flow*: Pour a slurry of SP Sepharose Fast Flow resin in water into a 5 cm diameter column to approximately 10 cm in height (approx. 200 mL volume). Before use, equilibrate column in TCB50+.
   (c) *DNA cellulose*: Rehydrate the DNA cellulose in TCB containing 100 mM KCl (TCB100), and decant off the fines. Repeat at least one more time until there are no more fines. Pour a slurry of DNA-cellulose in TCB100 into a 5 cm diameter column to approximately 5 cm in height (approximately 100 mL column volume). Before use, equilibrate column in TCB100 containing 1 mM DTT and protease inhibitors (TB100+) (*see* Subheading 2.1, **item 10**).
2. **Preparing HeLa Cell Nuclear Extract**
   (a) Quick thaw the frozen cell extracts (in LSB) under warm water with gently swirling. As soon as the extract starts to melt, decant it into a chilled flask, packed in ice. Add DTT and protease inhibitors to the thawed sample.
   (b) When the extract is thawed, transfer to centrifuge bottles, and spin at 10,000 × *g* at 4 °C for 30 min. Decant the

supernatant, and store on ice. This is the S10 which contains the cytoplasmic contents. Record the volume of the S10. Remove an aliquot for gels, snap freeze the rest in liquid N_2, and store at −80 °C (*see* **Note 7**).

(c) To the residual cell pellet, add a quantity of ice-cold 1× High Salt Buffer containing 1 mM DTT and protease inhibitors equal to 1/5 the volume of the S10. Gently resuspend the pellet using a rubber policeman (*see* **Note 8**). Centrifuge the resuspended pellet at 10,000 × *g* at 4 °C for 30 min. Decant the supernatant, and keep on ice.

(d) Re-extract the pellet with ¼ the original volume of ice-cold 1× High Salt Buffer containing 1 mM DTT and protease inhibitors, centrifuge as above, and combine the two supernatants. This is the high salt wash of the nuclear pellet (HSW). Record the volume of the HSW, and remove 20 μL for gels and protein concentration (*see* **Note 9**).

(e) Dialyze the HSW against 4 L of ice cold TCB50+ for 3–4 h at 4 °C (*see* **Note 10**). Change the dialysis buffer, and repeat for another 3–4 h if necessary. Adjust the conductivity of the dialyzed sample with 1× TCB containing 1 mM DTT and protease inhibitors until it is equivalent to TCB50+. Spin at 10,000 × *g* for 30 min at 4 °C immediately prior to loading onto the DEAE column. A basic outline of the chromatographic steps in the prep is shown in Fig. 1**.**

3. **Chromatography on DEAE Fast Flow**

(a) Apply the dialyzed and/or diluted HSW in TCB50+ to the equilibrated DEAE column, collecting 100 mL fractions. Allow the sample to drain in, wash the sample in with TCB50+, and then wash the column with TCB50+ until the A280 reaches baseline.

(b) Elute DNA-PKcs and Ku70/80 by washing the column with approximately 3 column volumes of 1× TCB containing 175 mM KCl, 1 mM DTT, and protease inhibitors (TCB175+) until the A280 is back to baseline. Collect ~40 mL fractions into 50 mL plastic conical tubes.

(c) Wash the column with approximately 3 column volumes of 1× TCB containing 750 mM KCl, 1 mM DTT, and protease inhibitors (TCB750+) until the A280 is back to baseline. Collect ~40 mL fractions as above.

(d) Remove 20 μL of each column fraction for gels, then snap freeze the remaining fractions in liquid nitrogen, and store at −80 °C.

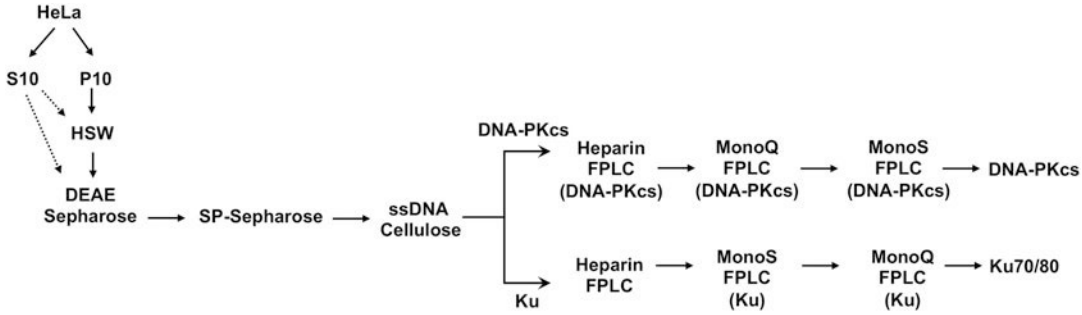


Fig. 1 Flowchart of purification of DNA-PK from HeLa cells. The dashed lines indicate that the S10 can be pooled with the high salt wash (HSW) of the P10 or worked up separately

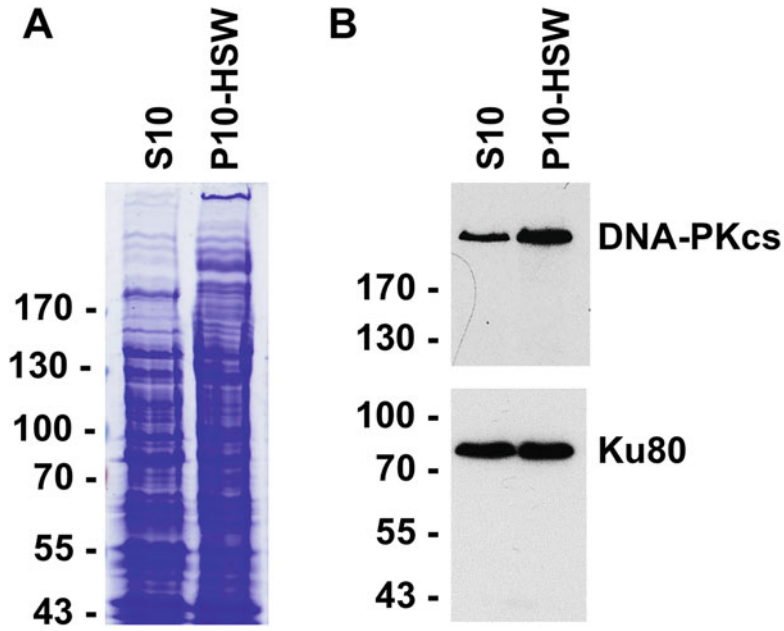


Fig. 2 Coomassie stained gel (**a**) and Western blot (**b**) of 50 μg extract from S10 and high salt wash (HSW) of the P10 probed for DNA-PKcs (top) and KU80 (bottom). The position of molecular weight markers is shown on the left, in kDa

(e) Run 5 μL of S10, HSW, DEAE flow through, and all fractions from the DEAE column on SDS PAGE and Western blot for DNA-PKcs and Ku (*see* Subheadings 2.2 and 3.2 **for details**). Expect DNA-PKcs and Ku70/80 to fractionate approximately equally between the S10 and HSW (Fig. 2) and to elute together in the TCB-175 mM KCl eluate from the DEAE column (Fig. 3). The approximate volumes, protein concentrations, and amounts of total protein are shown in Table 1.

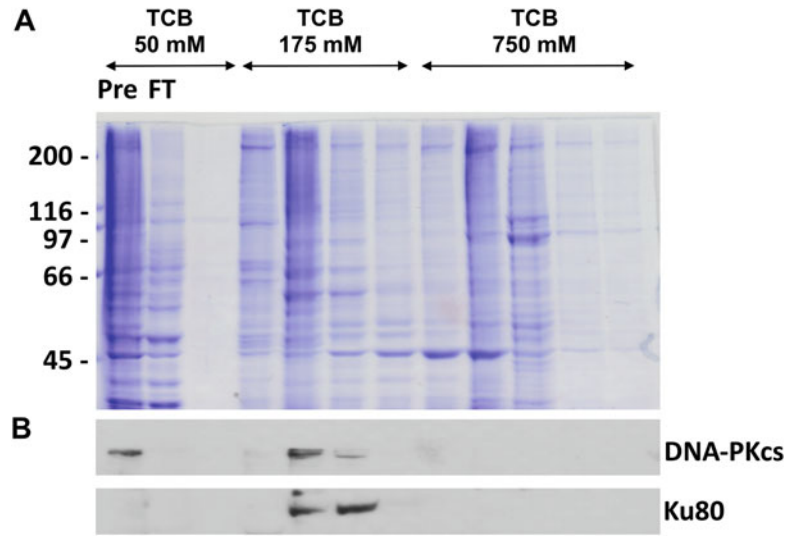


Fig. 3 Coomassie stained gels (**a**) and Western blots of DNA-PKcs and KU (**b**) of Pre-column and fractions from DEAE Fast Flow column of P10 HSW eluted with TCB75+, TCB-175+, and TCB750+. The position of molecular weight markers is shown on the left, in kDa

Table 1
Volumes and protein concentrations from each purification step. Representative values from a 100 L HeLa cell prep are shown

Step	Volume (mL)	Concentration (mg/mL)	Total protein (mg)
S10	330	25.2	8316
P10 (HSW)	220	15.2	3344
Pre-DEAE	310	2.36	732
Pre-SP Sepharose (from pooled DEAE 175 mM fractions)	325	0.58	189
Pre-ssDNA cellulose (from pooled SP Sepharose 175 mM fractions)	120	0.46	55
Pre-heparin (from DNA cellulose fractions pooled for DNA-PKcs)	14	1.36	19
Pre-mono Q	15	1.2	18
DNA-PKcs	2.9	2.2	6.4
KU70/80	1.8	1.7	3.1

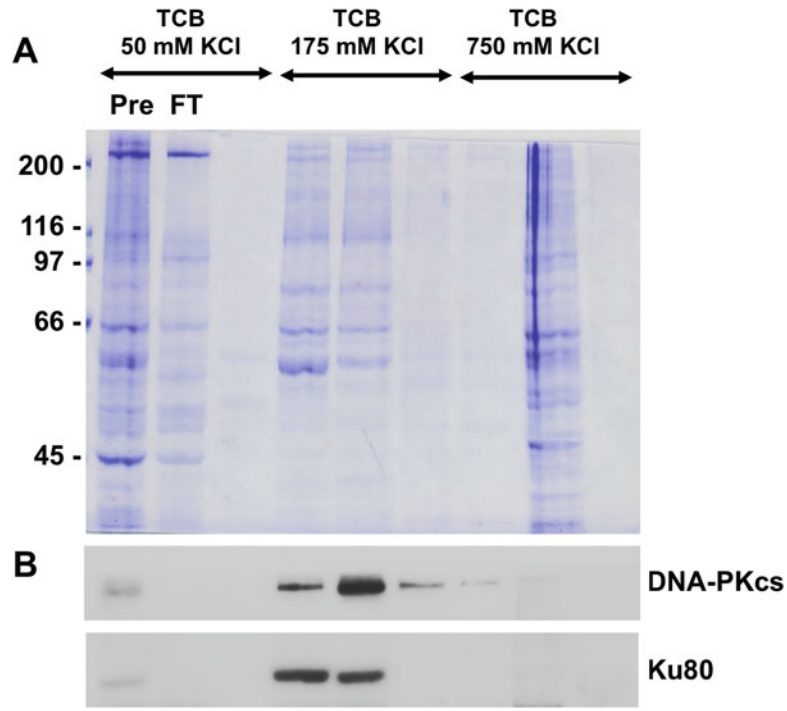


Fig. 4 Coomassie stained gels (**a**) and Western blots of DNA-PKcs and KU (**b**) of Pre-column (Pre) and fractions from SP Sepharose column of TCB175+ fraction from the DEAE column of P10 HSW, eluted with TCB75+ (FT), TCB175+, and TCB750+. The position of molecular weight markers is shown on the left, in kDa

4. **Chromatography on SP Sepharose Fast Flow**

   Dilute the TCB175 fraction from the DEAE column that contains DNA-PKcs and Ku70/80 to 50 mM using 1× TCB containing 1 mM DTT and protease inhibitors (TCB0+) (*see* **Note 10**). Centrifuge the diluted sample at 10,000 × *g* for 30 min before applying to the SP Sepharose column. Apply to column, wash sample with TCB50+, and elute column with TCB175+ and then TCB750+ exactly as in Subheading 3. Analyze samples for DNA-PKcs and Ku as in Subheading 3. Expect DNA-PKcs and Ku70/80 to elute together in the 175 mM eluate (Fig. 4).

5. **Chromatography on DNA Cellulose**

   (a) Dilute or dialyze the 175 mM fraction from the SP Sepharose column containing DNA-PKcs and Ku70/80 to TCB containing 100 mM KCl plus 1 mM DTT and protease inhibitors (TB100+). Centrifuge at 10,000 × *g* for 30 min before applying to the DNA cellulose column.

   (b) Apply the dialyzed/diluted SP Sepharose-TCB-175 fraction to the equilibrated DNA cellulose column, collecting ~10 mL fractions into 15 mL plastic conical tubes. Allow

the sample to drain in, wash the sample in with TCB100+, and wash the column until the A280 reaches baseline.

(c) Elute DNA-PKcs and Ku70/80 using a linear gradient of 150 mL TCB100+ to 150 mL TCB750+, collecting 10 mL fractions (*see* **Note 11**).

(d) Remove 20 μL of each fraction for gels, then snap freeze the remainder in liquid N2, and store at −80 °C until needed.

(e) Run 5 μL of the pre-DNA cellulose, flow through, and all fractions from the DNA cellulose column on SDS PAGE and Western blot for DNA-PKcs and Ku. Also run 5 μL of each fraction on SDS PAGE, and stain with Coomassie blue (*see* **Note 12** and Fig. 5).

(f) Pool the peak DNA-PKcs fractions and the peak Ku70/80 fractions separately so that you end up with two fractions, one enriched for DNA-PKcs and one enriched for Ku70/80. From this step onward, run peak DNA-PKcs-containing fractions and peak Ku-containing fraction separately, pooling peak fractions for each protein after each column, and retaining side fractions of both proteins for later workup.

6. **Chromatography on HiTrap Heparin**

(a) Using a chromatography system such as the Bio-Rad Next Generation Chromatography System or an AKTA FPLC, equilibrate a 5 mL HiTrap Heparin column in TCB100+ plus 0.02% (v/v) Tween 20 (*see* **Note 13**).

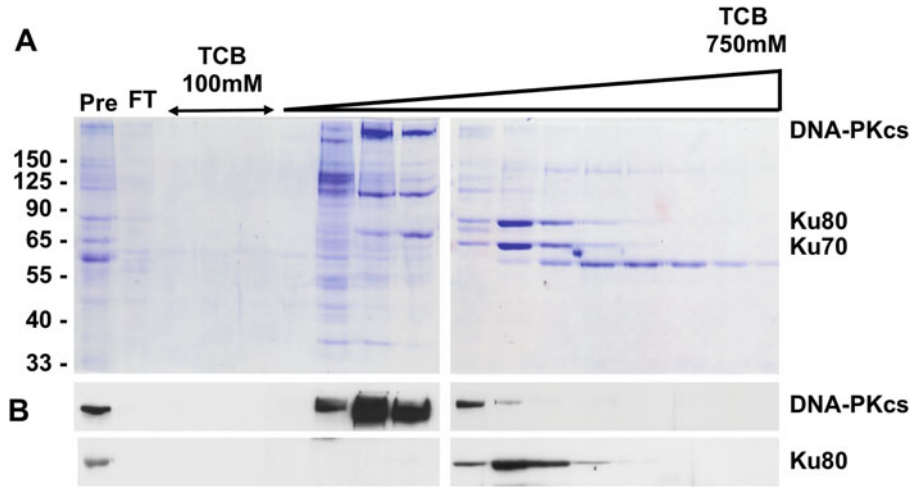


Fig. 5 Coomassie stained gels (**a**) and Western blots of DNA-PKcs and KU (**b**) of Pre-column (Pre) and flow-through (FT) fractions from DNA cellulose column eluted with TCB100+ and a gradient of TCB100+ to TCB750+. The position of molecular weight markers is shown on the left, in kDa

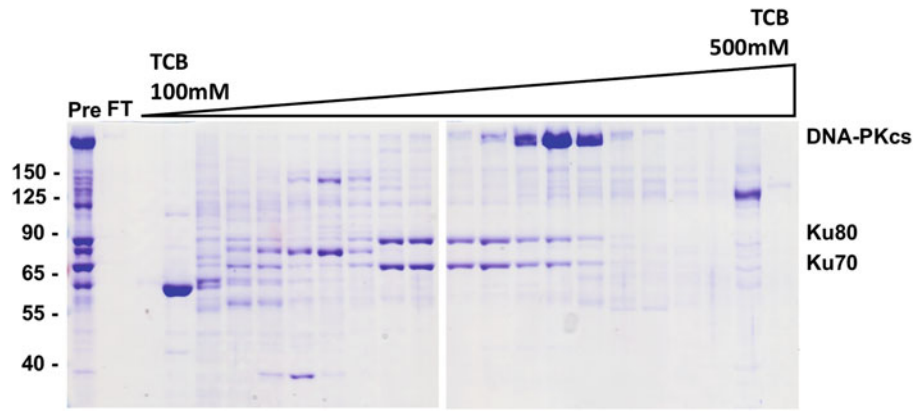


Fig. 6 Coomassie stained gels of Pre-column (Pre), flow through (FT) from Heparin-HiTrap column eluted with a gradient of TCB100+ to TCB500+. The position of molecular weight markers is shown on the left, in kDa

(b) Dilute the DNA-PKcs-containing fractions from the DNA cellulose column to TCB100+ plus 0.02% (v/v) Tween 20. Centrifuge at 10,000 × *g* for 30 min at 4 ° C, and apply the supernatant to the HiTrap Heparin column in TCB100+ plus 0.02% (v/v) Tween 20 at 1 mL/min. Collect 1 mL fractions.

(c) Wash the column with TCB100+ plus 0.02% (v/v) Tween 20 until the A280 is at baseline. Elute with a linear gradient of TCB100+ plus 0.02% (v/v) Tween 20 to TCB500+ plus 0.02% (v/v) Tween 20 over 60 min at 1 mL/min. Collect 1 mL fractions. Remove 20 μL of each fraction for gels, then snap freeze the remainder in liquid N2, and store at −80 °C until needed.

(d) Run 5 μL of the pre-Heparin, flow through, and all fractions from the heparin column on SDS PAGE and Western blot for DNA-PKcs and KU. Also run 5 μL on SDS PAGE, and stain with Coomassie blue (Fig. 6). Pool the peak DNA-PKcs-containing fractions.

(e) Repeat for the KU-containing fractions from DNA cellulose column. For each column, pool the main DNA-PKcs-containing fractions and the main KU70/80-containing fractions, and then pool the DNA-PKcs side fractions and KU side fractions to workup separately.

7. **Chromatography on Mono Q in TCB**

(a) Equilibrate a Mono Q 5/50 GL (GE) column in TCB100+ containing 0.02% (v/v) Tween 20.

(b) Dilute and/or dialyze the DNA-PKcs-containing fractions from the heparin column to TCB100+ plus 0.02% Tween 20.

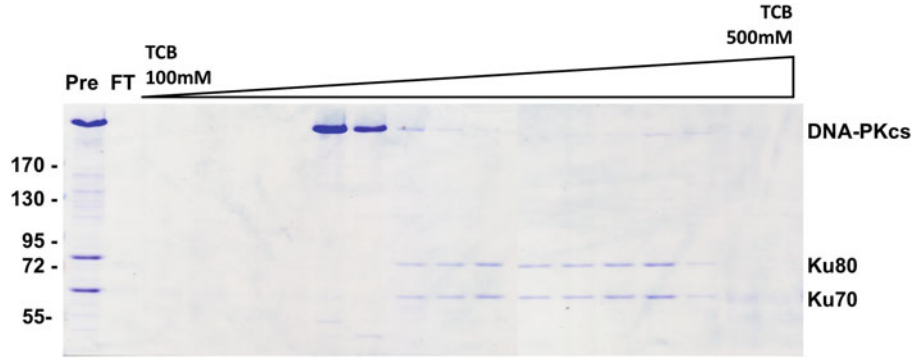


Fig. 7 Coomassie stained gels of Pre-column (Pre), flow through (FT) from the Mono Q column of DNA PKcs-enriched fractions from Heparin HiTrap, eluted with a gradient of TCB100+ to TCB500+. The position of molecular weight markers is shown on the left, in kDa

(c) Centrifuge at 10,000 × *g* for 30 min, and apply to the Mono Q column in 1× TCB100+ plus 0.02% Tween 20 at 1 mL/min, collecting 1 mL fractions.

(d) Wash the column with TCB100+ plus 0.02% Tween 20 until the A 280 is at baseline and then elute DNA-PKcs and KU with a linear gradient of TCB100+ plus 0.02% Tween 20 to TCB500+ plus 0.02% Tween 20 as described above for the HiTrap Heparin column (subheading 3.1, **item 6**) (Fig. 7).

(e) Pool the main DNA-PKcs-containing fractions and DNA-PKcs and KU-containing side fractions separately.

(f) Repeat for the KU-containing fractions.

8. **Chromatography on Mono S in TCB**

(a) Equilibrate a MonoS 5/50 GL column in TCB100+ plus 0.02% (v/v) Tween 20.

(b) Dilute and/or dialyze the peak DNA-PKcs-containing fractions from the Mono Q column to TCB100+ containing 0.02% (v/v) Tween 20.

(c) Centrifuge at 10,000 × *g* for 30 min, apply to the MonoS column and elute and run fractions on SDS PAGE and Western blot as described above. Pool the peak DNA-PKcs-containing fractions (*see* Fig. 8).

(d) Repeat for the KU-containing fractions. Note: KU binds weakly to MonoS and elutes in the flow through.

9. **Chromatography on Mono Q in HEPES Column Buffer**

(a) Equilibrate a Mono Q 5/50 GL column in 1× HEPES Column Buffer (HCB) containing 100 mM KCl, 0.02% (v/v) Tween 20, 1 mM DTT, and protease inhibitors (HCB100+ plus Tween).

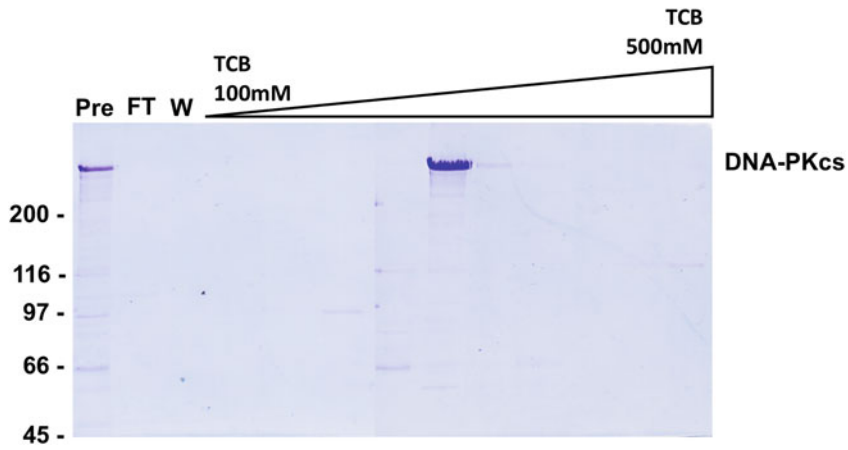


Fig. 8 Coomassie stained gels of Pre-column (Pre), flow through (FT) from the MonoS column of the DNA PKcs-enriched fractions from the first Mono Q column eluted with TCB 100+ (W), and a gradient of TCB100+ to TCB500+. The position of molecular weight markers is shown on the left, in kDa

(b) Dilute and/or dialyze the DNA-PKcs-containing fractions from the MonoS column to HCB100+ plus 0.02% Tween 20.

(c) Centrifuge at 10,000 × *g* for 30 min, and apply to the Mono Q column in HCB100+ plus Tween. Collect 1 mL fractions.

(d) Develop the column, and run fractions on SDS PAGE and Western blot as described above. Pool the peak DNA-PKcs-containing fractions (*see* **Note 14**).

(e) Repeat for the KU-containing fractions (*see* **Note 15**).

10. **Final Yield and Storage of Purified Proteins**

We typically obtain about 3–6 mg of purified DNA PKcs and KU70/80 from 100 L HeLa cells. Final fractions may be concentrated and the buffer changed using a spin concentrator if necessary, but this may reduce your overall yield of protein. Figure 9 shows a Coomassie stained gel of the final purified DNA PKcs and KU. A summary of the protein purification scheme is shown in Fig. 1, and Table 1 shows typical protein yields from a 100 L prep. Assay for DNA-PK activity using previously describe methods [5–7] or using a commercially available assay kit (Promega, catalogue number V4106). We typically assay activity using 30 ng purified KU and DNA-PKcs protein with and without 10 μg/mL sonicated calf thymus DNA (in TE buffer) and then titrate increasing molar equivalents of purified KU70/80 into DNA-PKcs. Maximum kinase activity is seen with an approximately 1:1 molar ratio of KU70/80 to DNA-PKcs. Store the purified protein at 1–2 mg/mL in HCB100+ in small aliquots at −80 °C. The activity is stable for several years, but avoid more than two or three freeze-thaw cycles.

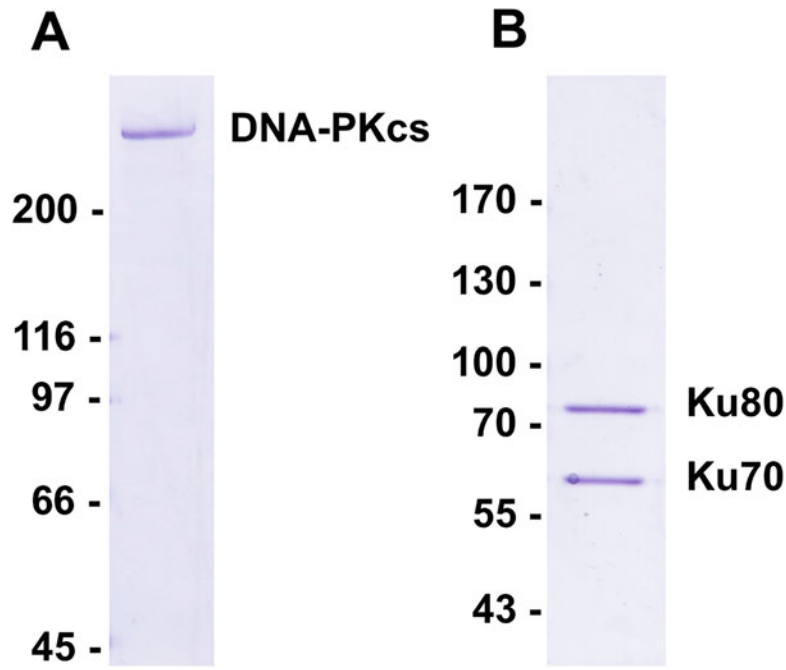


Fig. 9 Coomassie stained gels of purified DNA PKcs (**a**) and KU70/80 (**b**). The position of molecular weight markers is shown on the left, in kDa

3.2 Method for SDS PAGE and Western Blot of DNA-PKcs

1. To make up 8% low-bis DNA-PKcs gels: pipette into a 50 mL conical tube in the order below: 3.5 mL water, 4.2 mL 1 M Tris pH 8.8, 3.0 mL 30% (w/v) acrylamide, 0.4 mL 2% (w/v) bis-acrylamide, and 56 μL 20% (w/v) SDS. Swirl gently to mix, and then add 45 μL 10% (w/v) ammonium persulfate and 10 μL TEMED. Swirl gently to mix and pour into assembled plate cassette immediately. Overlayer the resolving gel with isobutanol-saturated water, and allow to polymerize for at least 1 h. When the gel is polymerized, wash off the isobutanol with a water bottle, wick the top of the resolving gel to remove the water, and add the stacking gel.
2. To make up the stacking gel: gently mix 3.3 mL water, 0.63 mL 1 M Tris-HCl, pH 6.8, 0.83 mL 30% (w/v) acrylamide, 0.175 mL 2% (w/v) bisacrylamide, 25 μL 20% (w/v) SDS, and then add 50 μL 10% ammonium persulfate and 5 μL TEMED and pour onto top of resolving gel containing a comb. Allow the stacking gel to polymerize for about 15 min.
3. Remove the comb, and wash the wells well with water, load sample in SDS sample buffer, and run at 100 V until bromophenol blue dye runs off the end of the resolving gel.
4. Equilibrate the gel and nitrocellulose in DNA-PKcs Electroblot Buffer for 15 min, and then assemble the cassette as per

normal and transfer in DNA-PKcs Electroblot buffer for 1 h at 100 V at RT with an ice pack. Block the membrane, and proceed as normal.

4 Notes

1. Unless otherwise stated, all buffers and reagents should be kept ice cold at all times.
2. Alternatively, you can use commercially available protease inhibitor cocktails.
3. Acrylamide and bis-acrylamide are potential neurotoxins and carcinogens. Use caution and obey all institutional safety protocols when using, especially when weighing the dry powder.
4. SDS is a respiratory irritant. Use caution and obey all institutional safety protocols when using, especially when weighing the dry powder.
5. Make fresh weekly.
6. For DEAE Fast Flow, SP Sepharose Fast Flow, and DNA cellulose columns, we pour resins into Bio-Rad Glass Econo-Column (Bio-Rad catalogue number 7375021) and equilibrate and run using gravity feed at 4 °C. We typically use single-stranded DNA cellulose (Sigma-Aldrich catalogue number D8273). After each prep, transfer DEAE and SP Sepharose resins to a sintered glass funnel, and wash sequentially with 0.1 M NaOH, water, 0.1 M HCl, water, 2 M, KCl, and water. Store in 20% ethanol. DNA cellulose should be washed with 2 M KCl and stored in 10 mM Tris-HCl, 10 mM EDTA, pH 8.0 at 4 °C.
7. DNA-PKcs and KU can also be purified from the S10 fraction using identical methods to those described above. Alternatively, you can combine the S10 and the high salt wash of the P10, dilute to TCB-50 mM KCl, and purify the total DNA-PKcs/KU fractions together using the same methods as described above. If you do this, we suggest starting with 50 L of HeLa cells rather than 100 L so as not to overload the columns.
8. Very important: do not use excessive force to wash the nuclear pellet, or the chromatin will decondense.
9. You can freeze the HSW in liquid N2, store it at −80 °C, and use later or proceed immediately to the DEAE column.
10. Alternatively, you can dialyze the sample into TCB plus 50 mM KCl, DTT, and protease inhibitors or use a combination of dilution and dialysis. Do not re-freeze the sample after dialysis

or reducing the salt concentration. Centrifuge, and load onto the column immediately.

11. Alternately, you can elute DNA-PKcs and KU by washing the column sequentially with 50 mL 1× TCB containing 200 mM KCl, 1 mM DTT, 1 mM PMSF, and protease inhibitor cocktail and then 1× TCB containing 300 mM KCl, 1 mM DTT, 1 mM PMSF, and protease inhibitor cocktail, and then the same buffer but containing 400 mM salt, then 500 mM salt, and then 750 mM salt, collecting ~10 mL fractions of each.
12. This is the first time in the prep that DNA-PKcs and KU70/80 will be visible on Coomassie stained gels. Expect DNA-PKcs to elute between 200 and 400 mM salt and KU70/80 to elute between 400 and 700 mM salt. *See* Fig. 5.
13. We run HiTrap, Mono Q, and MonoS FPLC columns at RT not at 4 °C.
14. If your experiments are not compatible with the presence of Tween 20, repeat the last Mono Q column in HCB in the absence of Tween 20.
15. KU70/80 can also be purified from baculovirus-infected insect cells as described [7].

Acknowledgments

This work was supported by NIH Program Grant CA092584.

References

1. Blackford AN, Jackson SP (2017) ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response. Mol Cell 66:801–817
2. Chan DW, Mody CH, Ting NS, Lees-Miller SP (1996) Purification and characterization of the double-stranded DNA-activated protein kinase, DNA-PK, from human placenta. Biochem Cell Biol 74:67–73
3. Lees-Miller SP, Chen YR, Anderson CW (1990) Human cells contain a DNA-activated protein kinase that phosphorylates simian virus 40 T antigen, mouse p53, and the human Ku autoantigen. Mol Cell Biol 10:6472–6481
4. Goodarzi AA, Lees-Miller SP (2004) Biochemical characterization of the ataxia-telangiectasia mutated (ATM) protein from human cells. DNA Repair 3:753–767
5. Lees-Miller SP, Sakaguchi K, Ullrich SJ, Appella E, Anderson CW (1992) Human DNA-activated protein kinase phosphorylates serines 15 and 37 in the amino-terminal transactivation domain of human p53. Mol Cell Biol 12:5041–5049
6. Yu Y, Mahaney BL, Yano K, Ye R, Fang S, Douglas P, Chen DJ, Lees-Miller SP (2008) DNA-PK and ATM phosphorylation sites in XLF/Cernunnos are not required for repair of DNA double strand breaks. DNA Repair 7:1680–1692
7. Thapar R, Wang JL, Hammel M, Ye R, Liang K, Sun C, Hnizda A, Liang S, Maw SS, Lee L, Villarreal H, Forrester I, Fang S, Tsai MS, Blundell TL, Davis AJ, Lin C, Lees-Miller SP, Strick TR, Tainer JA (2020) Mechanism of efficient double-strand break repair by a long non-coding RNA. Nucleic Acids Res 49:1199–1200

Chapter 15

Purification and Characterization of Human DNA Ligase IIIα Complexes After Expression in Insect Cells

Ishtiaque Rashid, Miaw-Sheue Tsai, Aleksandr Sverzhinsky, Aye Su Hlaing, Brian Shih, Aye C. Thwin, Judy G. Lin, Su S. Maw, John M. Pascal, and Alan E. Tomkinson

Abstract

With improvements in biophysical approaches, there is growing interest in characterizing large, flexible multi-protein complexes. The use of recombinant baculoviruses to express heterologous genes in cultured insect cells has advantages for the expression of human protein complexes because of the ease of co-expressing multiple proteins in insect cells and the presence of a conserved post-translational machinery that introduces many of the same modifications found in human cells. Here we describe the preparation of recombinant baculoviruses expressing DNA ligase IIIα, XRCC1, and TDP1, their subsequent co-expression in cultured insect cells, the purification of complexes containing DNA ligase IIIα from insect cell lysates, and their characterization by multi-angle light scattering linked to size exclusion chromatography and negative stain electron microscopy.

Key words Baculovirus, Bacmid, Insect cells, Affinity chromatography, Size exclusion chromatography, Ion exchange chromatography, Multiple angle light scattering, Negative stain electron microscopy

1 Introduction

It is now generally accepted that many fundamental processes are performed by multi-protein machines. In order to understand how these machines work, it is necessary to understand the architecture and functional flexibility of these machines. While X-ray crystallography provides snapshots of structures at atomic resolution, flexible and unstructured regions of proteins are refractory to crystallization and, even if contained within protein crystals, are poorly resolved. Thus, most crystal structures tend to be of stable, relatively small well-folded regions of a protein. With the development

Ishtiaque Rashid, Miaw-Sheue Tsai and Aleksandr Sverzhinsky contributed equally with all other contributors.

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_15,

of cryo-electron microscopy (EM) combined with improvements in detectors and image analysis, it is now possible to determine high-resolution structures of some large protein complexes with the building of structural models facilitated by docking in X-ray structures corresponding to smaller, well-folded parts of the larger complex [1, 2]. As with X-ray crystallography, flexible regions are not well resolved in cryo-EM. Small angle X-ray scattering (SAXS) provides insights into the shape and conformational flexibility of proteins and protein complexes [3–5]. Similar to cryo-EM, improvement in detectors and data analysis have led to increased use of SAXS to characterize proteins and protein complexes [3–6].

Here we describe the use of Bac-to-Bac Baculovirus Expression systems to efficiently produce separate recombinant baculoviruses encoding DNA ligase IIIα (LigIIIα), XRCC1, or TDP1, each of which contains large unstructured, highly flexible regions. The *LIG3* gene, one of three human genes encoding DNA ligases, encodes mitochondrial and nuclear versions of LigIIIα by alternative translation initiation that are very similar in size after removal of the mitochondrial targeting sequence by proteolysis [7, 8]. In the nucleus, LigIIIα forms a complex with the DNA repair scaffold protein XRCC1 that coordinates the activities of multiple DNA repair enzymes [9, 10]. The nuclear LigIIIα-XRCC1 complex acts as a back-up for DNA ligase I in DNA replication, and there is also functional redundancy with the other nuclear DNA ligases in DNA repair [11–13]. In contrast, LigIIIα is the only DNA ligase in mitochondria where it functions in the replication and repair of the mitochondrial genome in the absence of XRCC1 [11, 12, 14]. The *TDP1* gene also encodes nuclear and mitochondrial versions of this enzyme that interact with LigIIIα and remove 3′ tyrosine residues remaining after degradation of stalled topoisomerase I molecules as well as other 3′ adducts [15, 16]. LigIIIα was co-expressed with XRCC1 or TDP1 as well as with both XRCC1 and TDP1, and, after small-scale experiments to detect complex formation, XRCC1-LigIIIα, LigIIIα-TDP1, and XRCC1-LigIIIα-TDP1 complexes were purified in sufficient amounts for subsequent biochemical and biophysical analysis. Because of the presence of BRCT domains in both LigIIIα and XRCC1 that homo- and heterodimerize and the elongated, asymmetric shape of these proteins [5], it is not possible to accurately estimate the mass and stoichiometry of complexes formed by these proteins by comparing their elution position from a size exclusion column with globular protein standards. Here, we describe the use of multi-angle light scattering coupled to gel filtration (SEC-MALS) in order to estimate the absolute molar mass of LigIIIα complexes independent of their elution volume. Negative stain electron microscopy (NS-EM) is a powerful tool to gain insights into the size and shape of protein complexes with other proteins. Depending on the degree of flexibility and heterogeneity, it is possible to

classify 2D projections of single particles into 2D classes with increased signal-to-noise ratio and/or proceed to create a 3D consensus map that can be subsequently used for 3D classification into different conformers. Here, we describe methods that were used to visualize LigIIIα complexes by NS-EM.

2 Materials

All reagents should be prepared using deionized ultrapure water and using HPLC grade reagents, except that media used for bacterial cultures should be prepared using tap water. Buffers used for cell lysis and protein purification must be vacuum filtered and pre-cooled in 4 °C refrigerator prior to use. Thorough waste disposal and safety regulations must be followed when disposing waste materials.

2.1 Expression of LigIIIα Complexes in Insect Cells

1. pFastBac baculoviral transfer vectors, including pFastBac/LigIIIα, pFastBac/LigIIIα-Strep, pFastBac/6His-XRCC1, pFastBac/Flag-TDP1, and pFastBac/6His-Flag-TDP1 [17, 18], are available from the Expression and Molecular Biology Core directed by Tsai (*see* **Note 1**).
2. DH10Bac chemical competent cells.
3. Buffers P1, P2, and P3 and TE Buffer for bacmid preparation: Buffer P1: 50 mM Tris-HCl pH 8.0, 10 mM EDTA, and 100 μg/mL RNase A. Buffer P2: 0.2N NaOH and 1% SDS. Buffer P3: 3 M potassium acetate, pH 5.5. TE Buffer: 10 mM Tris-HCl, pH 8.0 and 1 mM EDTA.
4. CellFECTIN II transfection reagent.
5. Insect cell lines Sf9 (*Spodoptera frugiperda*) and High Five cells (BTI-TN-5B1-4 *Trichoplusia ni*) cultured in ESF 921 protein-free medium (Expression Systems).
6. Cell culture plasticware (6-well and 100-mm tissue culture-treated plates) and reusable, autoclavable polycarbonate non-baffled shaker flasks (125, 250, 500, and 1000 mL).
7. Environmental chamber incubators with orbital shakers maintained at 27 °C.
8. Cell lysis buffer for Sf9 or HF cells: 50 mM NaH_2PO_4, pH 7.5, 300 mM NaCl, 0.5% Igepal, 1 mM PMSF, 10 mM imidazole, and EDTA-free protease inhibitor tablet.
9. Dounce homogenizer.
10. 1× TAE buffer for DNA gel electrophoresis: 40 mM Tris-acetate and 1 mM ETDA, pH 8.3.

11. 1× Tris-Glycine SDS running buffer for protein gel electrophoresis: 25 mM Tris, 192 mM glycine, and 0.1% SDS, pH 8.3.
12. 1× Tris-glycine buffer for Western blot transfer: 25 mM Tris, 192 mM glycine, and 20% methanol, pH 8.3.
13. Gel electrophoresis apparatus for DNA and proteins and Bio-Rad Mini Trans-Blot transfer apparatus.
14. Ni-NTA Fast Flow resins (Qiagen) and anti-Flag M2 resins (Sigma). Anti-His (Qiagen) and anti-Flag (Sigma)-mouse monoclonal antibodies.

2.2 Purification of LigIIIα Protein Complexes

2.2.1 Buffers

1. Cell lysis buffer: 50 mM sodium phosphate buffer (mono and dibasic) pH 8.0, 300 mM NaCl, 0.5% Igepal, cOmplete mini EDTA free protease inhibitor cocktail.
2. Buffers for ÄKTA FPLC: 40 mM HEPES-NaOH pH 7.5, 10% glycerol, 1 mM benzamidine, 0.2 mM PMSF (phenylmethylsulfonyl fluoride) with NaCl depending on the buffer P0 (0 M NaCl), P200 (200 mM NaCl), and P1000 (1 M NaCl). P200 + I600 buffer corresponds to P200 buffer containing 600 mM imidazole (*see* **Note 15**).

2.2.2 ÄKTA FPLC and Columns

1. ÄKTA FPLC (GE Healthcare) with Unicorn software.
2. ÄKTA FPLC columns: HisTrap HP 1 mL (GE Healthcare), HiTrap Q 1 mL (GE Healthcare), HiTrap SP 1 mL (GE Healthcare), dsDNA cellulose from Calf Thymus DNA (Sigma-Aldrich), and HiLoad Superdex 200 16/600 (GE Healthcare) (*see* **Notes 18** and **19**).
3. Econo column chromatography (Bio-Rad, Cat# 7372511).

2.2.3 Other Equipment

1. Branson Sonifier 150: set to level 6.
2. Beckman Coulter Optima XPN-80 Ultracentrifuge and Ti70 rotor.
3. Bachman Coulter polycarbonate centrifuge bottles, capacity 26.3 mL.
4. VWR Sterile Syringe Filter w/ 0.45 μm cellulose acetate membrane.
5. Millipore Sigma Amicon Ultra 4 (or 15) centrifugal filter unit.

2.3 Characterization of LigIIIα Protein Complexes

2.3.1 SEC-MALS

1. Gel filtration running buffer: 40 mM HEPES pH 8, 200 mM NaCl, 10% glycerol, 0.1 mM Tris (2-carboxyethyl) phosphine (TCEP or a similar reducing agent, such as dithiothreitol). The buffer must be compatible with the gel filtration column and with the sample under study. All solutions should be sterile filtered and degassed.

2. Gel filtration column: 10/300 GL Superdex 200 Increase or similar, depending on expected size of protein complex.
3. Fast Protein Liquid Chromatography (FPLC) system, such as ÄKTA pure (GE Healthcare), ÄKTAmicro (GE Healthcare), DuoFlow (Bio-Rad), Alliance (Waters Corporation). Typically, there is an integrated UV detector able to measure absorbance at 280 nm. UV measurement is optional for SEC-MALS, but it facilitates monitoring elution on the FPLC without using the subsequent differential refractometer. For membrane protein analysis, a UV detector is mandatory.
4. MALS and differential refractometer detectors, such as those offered by Wyatt Technology, Malvern Panalytical, and Waters Corporation.
5. SEC-MALS Analysis Software, such as ASTRA (Wyatt Technology). Here, we use ASTRA version 6.1.6.5.
6. Purified protein standard for calibration should be monodisperse under the combination of the chosen column, flow rate, and buffer (typically BSA for the 10/300 GL Superdex 200 Increase column).

2.3.2 NS-EM

1. Glutaraldehyde (GLT).
2. 1 M Tris or Glycine buffered to the pH of the preferred protein solution.
3. Table-top centrifuge.
4. Spin-column concentrators, such as Amicon Ultra 0.5 mL (10 kDa molecular weight cutoff).
5. 12% SDS PAGE gel and running apparatus.
6. Glow discharge system, such as PELCO easiGlow (Ted Pella Inc.).
7. Uranyl formate or uranyl acetate. Uranyl salts are toxic and mildly radioactive; proper precautions must be taken when handling them.
8. Whatman No. 1 filter paper.
9. Transmission electron microscope capable of operating at room temperature, such as Tecnai 12 (FEI).

3 Methods

3.1 Expression of LigIIIα Complexes in Insect Cells

3.1.1 Transposition

1. Prepare Luria agar plates consisting of 10 g peptone, 5 g yeast extract, 10 g sodium chloride, and 12 g agar in 1 L of tap water. Autoclave and cool down to 55 °C, and add the following to cooled agar solution: 50 μg/mL kanamycin, 7 μg/mL gentamicin, 10 μg/mL tetracycline, 100–200 μg/mL Bluo-gal, and

40 μg/mL IPTG. Mix the agar solution, and pour 20–25 mL per 100-mm Petri dish under sterile conditions (*see* **Notes 2** and **3**).

2. Transform 50–100 ng pFastBac plasmid into 100 μL DH10Bac cells using the general transformation protocol. Add 900 μL SOC media to the heat-shocked cell-plasmid mixture, and incubate at 37 °C at 250 rpm for 4 h.
3. Dilute and plate cells on Luria agar plates to obtain 100–200 colonies per plate. Incubate for 18–20 h at 37 °C. Store the agar plates at 4 °C for 8 h on the following day. Then place the agar plates back to 37 °C, and incubate for another 18–24 h. Blue colonies become evident after 48 h post-transformation.

3.1.2 Isolation of Recombinant Bacmid DNA

1. Pick five white colonies from the plate. Inoculate each single white colony into 2 mL LB with tetracycline, gentamycin, and kanamycin, and incubate at 37 °C at 250 rpm for 18–20 h.
2. Spin down 1.5 mL overnight culture in a microtube at 14,000 × *g* for 3 min. Remove the supernatant, and resuspend bacterial pellets in 0.3 mL Buffer P1. Add 0.3 mL Buffer P2, gently mix, and incubate at room temperature for 5 min. Add 0.3 mL Buffer P3, mix gently, and incubate on ice for 5–10 min. Centrifuge at 14,000 × *g* for 10 min (*see* **Note 4**).
3. Transfer the supernatant to a new tube containing 0.65 mL isopropanol. Mix by inverting the tube a few times, and incubate on ice for 10 min. Centrifuge the sample at 14,000 × *g* for 15 min at room temperature. Pour to discard the supernatant. Add 0.5 mL 70% ethanol to each tube, and invert the tubes a few times to wash the DNA pellet. Centrifuge for 10 min at 14,000 × *g* at room temperature. Carefully remove the supernatant as much as possible using a micropipette. Do not pour to discard ethanol as the DNA pellet may be dislodged. Air dry the bacmid DNA for 10 min until the pellet becomes transparent, and redissolve in 40 μL TE buffer for at least 15 min. Gently tap the tube to mix, and avoid pipetting up and down (*see* **Notes 5** and **6**).
4. Restreak the overnight culture on Luria agar plates to verify the white phenotype (*see* **Note 7**).
5. Analyze 5 μL of bacmid DNA on a 0.5% agarose/TAE gel at 20 V for 18–20 h. Select three bacmid clones that show high intensity of an intact signature bacmid band (Fig. 1) with confirmed white phenotype for transfection (*see* **Note 8**).

3.1.3 Transfection of Sf9 Insect Cells with Recombinant Bacmid DNA

1. Seed 2 mL of Sf9 cells in 1% FBS media at 0.5 million cells/mL in 6-well plates the day before, and incubate at 27 °C for 18–20 h (*see* **Note 9**).

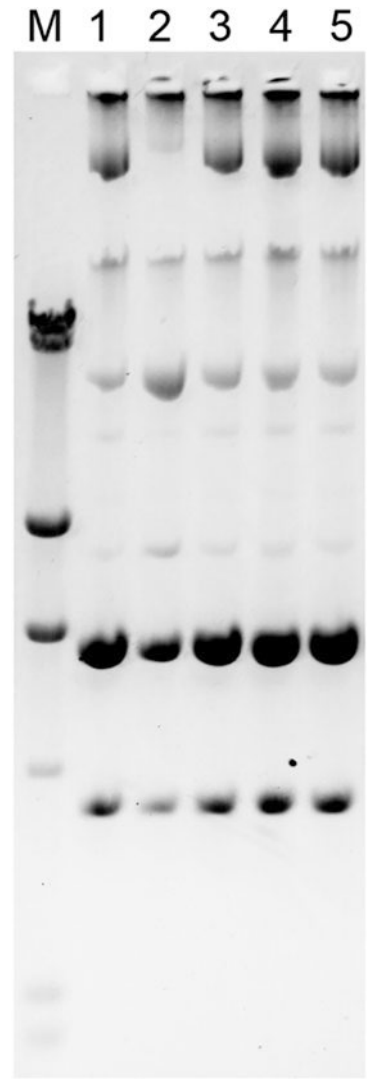


Fig. 1 Analysis of recombinant bacmid DNA by 0.5% agarose gel electrophoresis at 20 V for 18–20 h. Bacmid DNA is isolated from five clones per construct (lanes 1–5). Lambda DNA HindIII digest containing 23.1, 9.4, 6.5, 4.3, 2.3, 2.0, and 0.5 kb fragments is used as a DNA size marker (lane M). The signature intact bacmid DNA band is right above the 23.1 kb fragment as indicated by an arrow

2. For each bacmid clone, prepare a transfection mixture as following:

   Tube A: dilute 5–10 μL of bacmid DNA in 100 μL of ESF 921 SFM.

   Tube B: dilute 5 μL of CellFECTIN in 100 μL of ESF 921 SFM.

   Mix Tubes A and B, and incubate at room temperature for 30–45 min.

3. Include one tube without bacmid DNA as a negative control.
4. Wash Sf9 cells twice with 2 mL of ESF 921 SFM.
5. Add 0.8 mL ESF 291 SFM to the transfection mixture (total 1 mL), and mix gently. Aspirate the media from cells, and carefully add 1 mL diluted transfection mixture to each well without disturbing the attached cells. Also include one well without adding transfection mixture as a control. Then incubate at 27 °C for 5 h.
6. Aspirate transfection mixture, add 2 mL fresh 1% FBS media per well, and incubate at 27 °C for 72–96 h.
7. Harvest virus by transferring the media from each well to a 15-mL conical tube. Clarify by centrifugation at 500 × *g* for 5 min, and transfer the supernatant to a new 15-mL conical tube. This is designated as P0 clones. Store at 4 °C protected from light.

3.1.4 Baculovirus Amplification and Screening

1. Plate 10 mL of Sf9 cells in 1% FBS media at one million cells/mL in 100-mm Petri dishes the day before. Plate four plates for three clones and including an uninfected control. Add 250 μL of each P0 clone to Sf9 cells, gently swirl the plate to mix, and incubate at 27 °C for 72 h.
2. Transfer the media to a 15-mL conical tube, centrifuge at 500 × *g* for 5 min, and transfer the supernatant to a new 15-mL conical tube. This is designated as P1 clones. Store virus stocks at 4 °C and protected from light.
3. Plate 10 mL of Sf9 cells in 1% FBS media at 1 million cells/mL in 100-mm Petri dishes the day before. Add 100 μL of P1 clones to Sf9 cells, gently swirl the plate to mix, and incubate at 27 °C for 72 h.
4. Transfer the media to a 15-mL conical tube, centrifuge at 500 × *g* for 5 min, and transfer the supernatant to a new 15-mL conical tube. This is designated as P2 clones. Store virus stocks at 4 °C and protected from light.
5. Seed 2 mL of High Five cells at 0.5 million cells/mL in 6-well plates, and allow cells to attach for 30 min. Infect High Five cells with 20 μL of P2 clones, gently swirl the plate to mix, and incubate at 27 °C for 45–48 h.
6. Resuspend High Five cells directly in the media, transfer cells to 2 mL microtubes, and centrifuge at 500 × *g* for 5 min to pellet the cells. Aspirate the media, resuspend cell pellets in 150–200 μL 1× SDS loading buffer, and heat denature at 95 °C for 5 min.
7. Total protein expression is analyzed by resolving protein samples (8–10 μL loading per lane) on SDS-PAGE and analyzed with Coomassie staining and Western blotting using anti-

penta-His antibody to detect 6×His-tagged XRCC1 and anti-Flag antibody to detect Flag-tagged TDP1. Select the clone with the best expression for each protein to scale up virus production.

3.1.5 Scaled-Up Baculovirus Production in Suspension Cultures

1. Plate 100 mL of Sf9 cells in 1% FBS media at 1 million cells/mL in a 250-mL shaker flask the day before.
2. To produce P2 stocks, add 1 mL of selected P1 clone to 100 mL cells. Gently swirl the flask to mix and allow cells to sit at 27 °C for 1 h. Then, transfer the flask to an orbital shaker, and incubate at 140 rpm for 72 h at 27 °C.
3. Transfer the media to 50 mL conical tubes, centrifuge at 500 × *g* for 5 min to pellet any floating cells, and transfer the supernatant to new 50-mL conical tubes. This is designated as P2ii stocks. Store virus stocks at 4 °C and protected from light (*see* **Notes 10**–**12**).
4. To test protein expression by the new viral stocks, seed 2 mL of High Five cells as described in **step 5** ***from*** Subheading 3.1.4. Infect High Five cells with 5–40 μL of P2ii stocks per well for 45–48 h at 27 °C. Collect infected cells, and analyze protein samples as described in **steps 6** and **7** ***from*** Subheading 3.1.4.

3.1.6 Co-expression of LigIIIα, XRCC1, and TDP1 in Suspension Insect Cells by Co-infection with Two or Three Baculoviruses

1. To test single protein expression of LigIIIα, XRCC1, and TDP1 in shaker cultures, plate 30 mL of High Five and 50 mL of Sf9 cells at 1 million cells/mL in 125-mL shaker flasks. Add P2ii stocks to insect cells with the amount of 5–40 μL virus per million cells. Gently swirl the flask to mix, allow to sit still for 1 h, and then incubate at 140 rpm for 45–48 h for High-five cells and for 68–72 h for Sf9 cells. Cells (1 mL) were sampled and analyzed for total protein expression by SDS-PAGE and Western blotting as described above. The rest of the cells are pelleted and analyzed for solubility and for affinity purification where applicable (*see* **Notes 13** and **14**).
2. Resuspend High-five (30 mL) or Sf9 (50 mL) cells in 5 mL of lysis buffer, and lyse cells by Dounce homogenizer using a type B, or tight-fitting pestle with repeated 40 strokes. Clarify cell lysates by centrifugation at 15,000 × *g* for 30 min at 4 °C. Sample 50 μL of cell lysates before (total protein fraction) and after (soluble protein fraction) clarification, and mix with 50 μL of 2× SDS loading buffer. Analyze protein solubility by comparing total vs. soluble fractions using SDS-PAGE and Western blotting.
3. Based on the initial shaker expression analysis of each protein in **steps 1** and **2**, plate insect cells as described in **step 1**, co-infect High Five and Sf9 cells with the virus stocks of various ratios,

incubate on a shaker for 48–72 h depending on cell lines, and then collect cells for total protein analysis and solubility analysis as described in **step 2.**

3.1.7 Affinity Co-purification of LigIII Protein Complexes

1. Lyse co-infected insect cells as described in **step 2** ***in*** Subheading 3.1.6.
2. Remove the storage buffer from the selected affinity resins, for example, Ni-NTA beads, and then equilibrate the resins with the same lysis buffer in 1.5 mL microtubes. Use a column volume of 50 μL for 30–50 mL insect cells. Transfer the resins to a 15 mL conical tube.
3. Add soluble extracts to the resins, and allow for batch binding at 4 °C with gentle rocking for 1 h. Centrifuge the tube at 500–700 × *g* for 5 min to collect the resins. Transfer the supernatant (aka flow through fraction) to a new tube. Wash the resin 3 times with 10 mL of Wash Buffer (same as the Lysis Buffer with increased imidazole to 20–25 mM), then transfer resins to a 1.5 mL microtube, and wash 4 times with 1 mL of wash buffer. Centrifuge the tube at 500–700 × *g* for 5 min to collect the resins. Elute proteins from the affinity resins by adding 1 column volume of elution buffer (same as lysis buffer with increased imidazole to 300 mM) and incubating on ice for 5 min, followed by centrifugation. Transfer the supernatant (aka eluted fractions) to new 1.5 mL microtubes. Elute two times. For Flag purification with anti-Flag M2 resins, binding and washing buffers are the same as lysis buffer, and elute proteins from the resins using lysis buffer supplemented with 200 μg/mL 3× Flag peptide and incubate on ice for 15 min.
4. Sample purification fractions (including column load, flow through, eluted fractions, and leftover resins), mix with an equal volume of 2× SDS loading buffer, and heat denature at 95 °C for 5 min. Analyze purification fractions by SDS-PAGE against BSA standards with Coomassie staining and/or Western blotting (Fig. 2).

3.1.8 Scaled-Up Co-expression in Insect Cells

1. After determining the best co-infection ratio and cell line for LigIIIα, XRCC1, and TDP1 from small-scale co-expression and co-purification, plate 400 mL of insect cells at 1 million cells/mL in 1000-mL shaker flasks. Infect cells with the virus amounts, and perform a time course test, that is, sampling cells at 40, 44, and 48 h post-infection for High Five cells and 48, 60, and 72 h for Sf9 cells. Analyze total protein expression by SDS-PAGE and Western blotting, or perform a small-scale affinity co-purification from sampled 25 mL coinfected cells to confirm LigIIIα complexes.

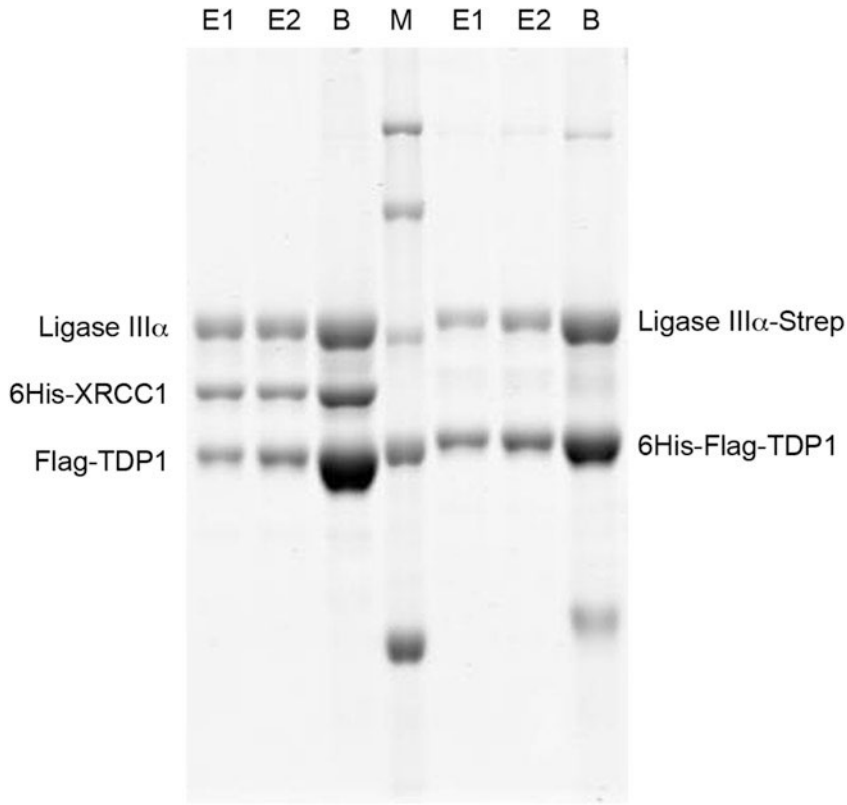


Fig. 2 Small-scale affinity purification of LigIII-XRCC1-TDP1 trimeric complex (left side of gel) and LigIII-TDP1 dimeric complex (right side of gel) from 25 mL of co-infected Sf9 cells using Ni-NTA and anti-Flag M2 columns, respectively. Approximately 10% of first (E1) and second (E2) eluates and 20% of leftover beads (B) were analyzed by 7.5% Tris-Glycine SDS-PAGE with Coomassie staining. Protein molecular weight standards (M) from the top of the gel are 250, 150, 100, 75, and 50 kDa

2. Scale up expression of LigIIIα complexes using the finalized optimal conditions, including selected cell line, viral co-infection ratio, and infection time.

3.2 Purification of LigIIIα Protein Complexes

LigIIIα/6His-XRCC1, LigIIIα/6His-Flag-TDP1, and LigIIIα/6His-XRCC1/Flag-TDP1 protein complexes are purified using the chromatography columns in the order shown in Table 1.

3.2.1 Lysate Preparation

1. Add 20 mL of cell lysis buffer and two cOmplete mini EDTA free protease inhibitor cocktail tablets per 0.4 L cell pellet, and thaw on ice (*see* **Note 15**).
2. Sonicate 25 s, and rest on ice for 30 s. Repeat this four more times (*see* **Note 16**).
3. Transfer the sonicated lysate in the Beckman Coulter polycarbonate centrifuge bottles, and balance the weight for centrifugation.

Table 1
Columns used for LigIIIα complex purification

Protein complex	1st column	2nd column	3rd column	4th column
LigIIIα/6His-XRCC1	HisTrap	HiLoad Superdex 200 16/60	dsDNA Cellulose	
LigIIIa/6His-Flag-TDP1	HisTrap	HiLoad Superdex 200 16/60	HiTrap Q	
LigIIIa/6His-XRCC1/Flag-TDP1	HisTrap	HiLoad Superdex 200 16/60	HiTrap Q	HiTrap SP

4. Centrifuge the lysates at 26,600 × *g*, at 4 °C, for 30 min.
5. Take the clear supernatant, and filter using a sterile syringe filter with 0.45 μm cellulose acetate membrane. Keep the lysate tube on ice throughout the filtering process (*see* **Note 17**).

3.2.2 Nickel Column FPLC Using HisTrap HP Column

1. For this purification step P200 is used as buffer A and P200 + I600 as buffer B. Following the manufacturer's guideline, equilibrate the HisTrap HP (1 mL) disposable column with 5% buffer B (P200 + I600 buffer) (*see* **Notes 18** and **19**).
2. Add imidazole (1 M stock solution) to the filtered lysate right to a final concentration of 30 mM before loading the sample lysate into the sample loading loop for the ÄKTA FPLC system (*see* **Note 20**).
3. Load the sample to the column selecting 5% buffer B and with a flow rate of 0.5 mL/min. Follow the manufacturer's guideline for setting limit for high back pressure alarm.
4. Start collecting flow through as soon as the UV absorbance (280 nm) peak starts going up, and collect till the peak drops near to the basal UV absorbance (280 nm).
5. Stop sample loading, and continue to wash the column with 5% buffer B till the UV reading (280 nm) becomes a flat line.
6. Elute the His-tagged protein bound to the HisTrap HP column in a stepwise manner using 8.3% (50 mM imidazole), 16.7% (100 mM imidazole), 50% (300 mM imidazole), and 100% (600 mM imidazole) buffer B. The elution fraction size should be 1 mL for each step, collecting 10 elution fractions for each step (*see* **Note 21**).
7. Based on the UV absorbance peak (280 nm), select elution fractions, and run an SDS-PAGE gel to verify and identify fraction containing the protein complex of interest. Upon verification, save the fractions containing the protein complex of interest for the next protein purification step (Fig. 3) (*see* **Note 24**).

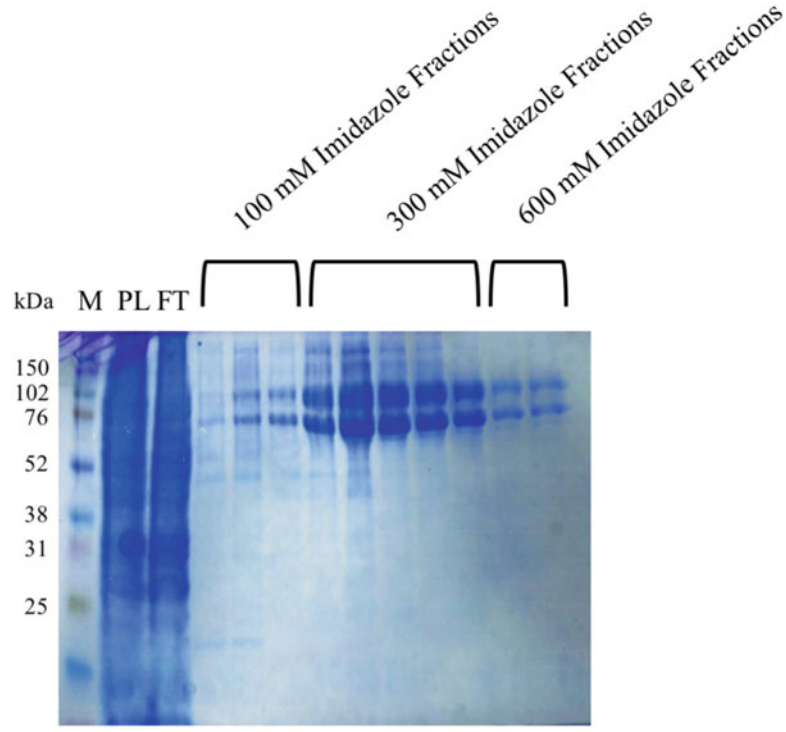


Fig. 3 Purification of LigIIIα/TDP1 complex using a nickel column. PL and FT represent pre-load and flow through, respectively. Proteins in 100, 300, and 600 mM imidazole eluates were detected by Coomassie staining after separation by SDS-PAGE

3.2.3 Size Exclusion Column FPLC Using HiLoad Superdex 200 16/60

1. Prior to running a sample, equilibrate the HiLoad Superdex 200 16/60 column with two column volumes (120 mL × 2) of P200 buffer in ÄKTA FPLC (*see* **Notes 19** and **22**).
2. Load samples from 0.5 to 5 mL onto Superdex 200 16/60 column. If nickel column fraction sample volume is greater than 5 mL, then concentrate the sample in P200 buffer using an Amicon Ultra 4 (or 15) centrifugal filter unit (*see* **Notes 23** and **26**).
3. Load the sample at 0.3–0.5 mL/min, and elute with P200 buffer. It is important to make sure no air bubble gets inside the column while sample is loading. Follow the manufacturer's guideline for setting limit for high back pressure alarm.
4. Start collecting 2 mL elution fractions after the void volume has passed, and continue collection for total 40 tubes.
5. Based on the UV absorbance peak (280 nm), select elution fractions, and run an SDS-PAGE gel to verify and identify fraction containing the protein complex of interest. Upon verification, pool the fractions containing the protein complex for the next protein purification step, or store them in −80 °C freezer (*see* **Note 24**).

3.2.4 dsDNA Cellulose Column FPLC

1. Pack a disposable econo column using 0.7–1 g of dsDNA cellulose from Calf Thymus DNA. The column bed length and diameter should be around 1 and 0.75 in., respectively. To equilibrate the column, use P100 buffer, which can be made by diluting P200 buffer twofold with P0 buffer. Equilibrate the column with 3× 10 mL P100 buffer using gravity (*see* **Note 25**).
2. For this purification step, P0 and P1000 buffers are used as buffers A and B.
3. Before loading the sample, dilute the sample so that the final NaCl concentration is 100 mM. Load the sample to the column selecting 10% buffer B (100 mM NaCl) and with a flow rate of 0.5 mL/min. Set the alarm for high backpressure to 0.5 mPa.
4. Start collecting flow through as soon as the UV absorbance peak (280 nm) starts going up, and collect till the peak drops near to the basal UV reading (280 nm).
5. Stop sample loading, and continue to wash the column with 10% buffer B until the UV absorbance (280 nm) reading becomes a flat line.
6. Elute the protein complex bound to the dsDNA cellulose column using a linear gradient by selecting buffer B range from 10 to 80% and length for 40 min. Collect 0.5 mL fractions and a total of 40 elution fractions.
7. Based on the UV absorbance peak (280 nm), select elution fractions, and run an SDS-PAGE gel to verify and identify fraction containing the protein complex of interest. Upon verification, save the fractions containing the protein complex for the next protein purification step, or store them in −80 °C freezer (Fig. 4) (*see* **Note 27** and **28**).

3.2.5 Ion-Exchange Column FPLC (HiTrap Q or HiTrap SP)

1. Use P0 and P1000 as buffers A and B for both the HiTrap Q and SP columns (*see* **Note 18**).
2. Equilibrate the column with 5 mL of 10% buffer B, and dilute the sample fractions so that the final NaCl concentration is 100 mM.
3. Follow **steps 3**–**7** from the Subheading 3.2.4 *dsDNA cellulose Column FPLC* (Fig. 5) (*see* **Note 27** and **28**).

SEC-MALS analysis can be carried out at room temperature or in a refrigerated chamber. Some proteins are sensitive to room temperature. If using a refrigerated chamber, be sure to cool down the buffer before use. Note that flow rates must be decreased at lower temperatures to respect the pressure limits of the FPLC system and gel filtration column.

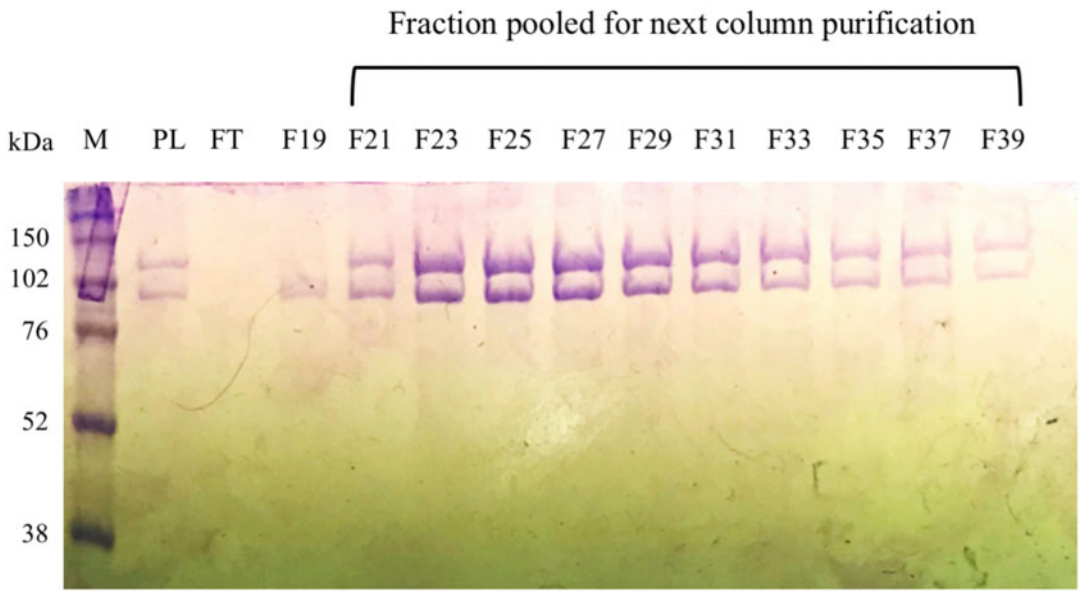


Fig. 4 Purification of LigIIIα/XRCC1 complex using a dsDNA cellulose column. PL and FT represent pre-load and flow through, respectively. Every other fraction from the peak was analyzed by SDS-PAGE to evaluate protein co-elution as an indicator of complex formation

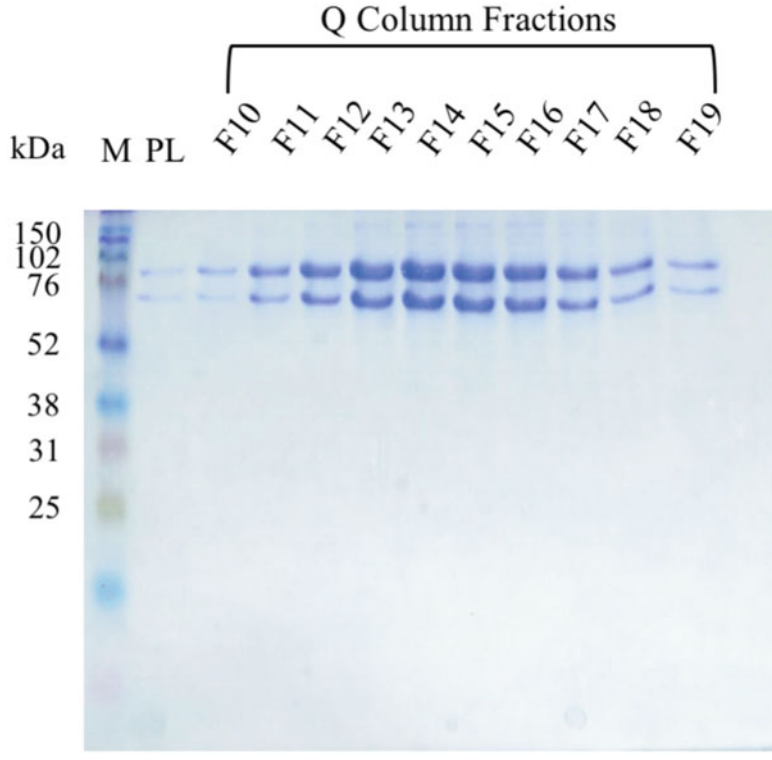


Fig. 5 Purification of LigIIIα/TDP1 complex using a HiTrap Q column. PL represents pre-load. Fractions from the peak were analyzed by SDS-PAGE to evaluate protein co-elution as an indicator of complex formation

3.3 Characterization of LigIIIα Protein Complexes

3.3.1 Characterization by SEC-MALS

Column and System Equilibration in Running Buffer

A FPLC system is used to equilibrate the gel filtration column with gel filtration running buffer. The chosen buffer must be determined somewhat empirically. An appropriate buffer should not induce precipitation of the sample within a short timeframe. Generally, the optimal buffer has already been determined at the protein purification stage. The result from SEC-MALS may indicate that a buffer leads to sample aggregation but without obvious precipitation. In this case the buffer should be altered by modification of the pH, the buffering molecule, addition or removal of salt, glycerol, reducing agents, detergents, or known ligands.

It is important that the only difference between the sample and the buffer is the protein complex to be measured, as the refractive index of the buffer component will be subtracted from the sample to give the concentration of the sample. Detailed articles on theoretical aspects of SEC-MALS are available elsewhere and will not be discussed here [19, 20].

1. Equilibrate the gel filtration column and detectors with the buffer using the FPLC pumps. Set the appropriate high-pressure alarms, and adjust the flow rate accordingly. Maintain the flow rate to no more than ~90% of the high-pressure limit, as there may be short-lived spikes in the pressure that could trigger the alarm and pause the run. *See* Subheading 4, **Note 29**.
2. The flow-path should include the gel filtration column, UV detector (optional, typically first detector), MALS flow cell, differential refractometer (last detector), fraction collector (optional).
3. Set the differential refractometer to purge mode to equilibrate both the reference and sample cells.
4. If present, turn on the sonicator in the MALS flow cell (COMET in Dawn HELEOS II from Wyatt Technology) at least 15 min prior to sample injection.
5. When equilibration is finished, turn off purge on the differential refractometer.

Calibration Using a Known Protein and Obtaining the Molar Mass of a Sample

The SEC-MALS system must be calibrated with a monodisperse protein of known size. BSA is typically used if the gel filtration column can efficiently separate the BSA monomer peak from the dimer peak. For example, the BSA peaks overlap in Fig. 6a when injected onto a 3.2/30 Superdex 200 Increase (2.4 mL) column but show clear separation in Fig. 6b when using a 30/300 Superdex 200 Increase (24 mL) column. Good peak separation is important for calibration since the band broadening (**step 7** below) requires clear peak shapes for correction of peak dispersion (Fig. 6c, d). It is necessary to calibrate the SEC-MALS system

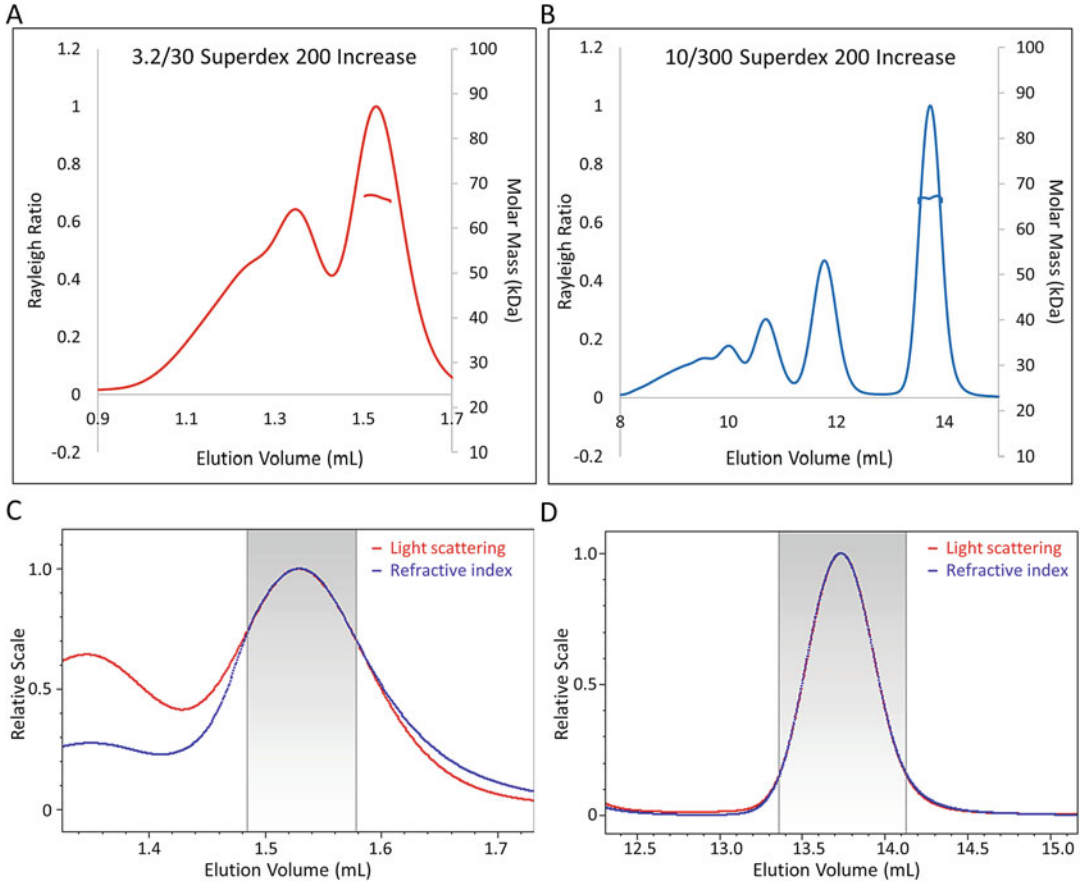


Fig. 6 BSA injected onto 3.2/30 (**a**, **c**) and 10/300 (**b**, **d**) Superdex 200 Increase gel filtration columns showing poor and ideal peak separation, respectively. Adequate peak separation is necessary for proper band broadening correction. The ASTRA software tries to align light scattering (red) and refractive index (blue) detectors in poor (**c**) and ideal (**d**) peak separation

before each set of experiments, as small changes in the buffer can affect results.

Whereas a monodisperse peak is necessary, it does not necessarily have to come from a monomer. The Aldolase tetramer (157 kDa) has been used successfully in calibrating SEC-MALS systems [21]. Regardless of the calibration standard, its Stokes radius must be known (3 nm for BSA monomer, 4.8 nm for aldolase). The calibration standard must be injected in sufficient quantities to provide strong signals in the MALS detectors. We recommend injecting 500 μg of BSA diluted in gel filtration buffer.

It has been observed that changes to the FPLC pump flow rate result in column shedding of fine particles that are visible by MALS detectors and may affect the baseline. It is recommended to inject samples without changing the pumps. Here we provide an overview

of the steps to be taken; exact operation depends on the researcher's specific SEC-MALS system and software.

1. To avoid air entering the FPLC system and the column and to equilibrate the injection loop, change the injection position to *Inject* (or equivalent) to flow running buffer through the injection loop. At this position samples cannot enter the injection loop.
2. Spin down the sample in a table-top centrifuge at top speed for 5 min. Take up the supernatant with an injection syringe, and insert it into the injection valve of the FPLC.
3. Change the injection position to *Load* (or equivalent), and push the injection syringe plunger to expel the sample into the injection loop.
4. On the SEC-MALS software (ASTRA), open a New Experiment from Default. When prompted to inject the sample, press *OK*, and quickly change the injection position on the FPLC to *Inject*. The running buffer is now pushing sample through the injection loop onto the column.
5. Once the run has finished, press *Stop* in the ASTRA software, and save the result. To avoid light scattering noise, the flow rate should not be modified between runs, if possible.
6. In the Procedures sections, select the level of despiking, and define the Baselines of the MALS and refractometer detectors (the UV detector is optional). The baseline limits should extend well beyond the peak(s) of interest. Under the Peaks section, define the monodisperse peak (monomer for BSA) by setting the limits to half the maximum peak height. *See* Subheading 4, (*see* **Note 30**).
7. To calibrate the SEC-MALS, right-click on Configuration, and choose *Alignment*. Highlight the peaks of the displayed detector traces, and click *Align Signals*, and then *Apply*. Next, choose *Brand Broadening* from the Configuration right-click menu, and press *Perform Fit*, using the refractometer as the reference instrument. Finally, right-click on Configuration, and choose *Normalize* and Peak 1 (defined in **step 6**). Enter the radius for the calibration standard (3 nm for BSA monomer), and press *Normalize*.
8. In the Procedures section, choose Molar Mass & Radius from LS. Clicking throughout the BSA monomer peak will show the calculated molar mass at each elution position. *See* Subheading 4, (*see* **Note 30**).
9. Under Results, the Mw should closely match the known molar mass of the calibration protein.
10. Save the Experiment as Method.

It is advisable to run the MALS flow-cell sonicator (COMET) between injections. The sample injection should follow the above steps, with the following modifications to steps.

4. In the ASTRA software, choose New Experiment from Method, and select the calibrated method that was saved using a known protein standard.
6. Under the Peaks section, define each peak of interest by setting the limits to half the maximum peak height. If a peak is asymmetric, one can extend the peak limit to estimate the molar mass of the species causing asymmetry.
7. Calibration is *not* performed on samples of unknown size.
8. The peak limits may be adjusted in the Molar Mass & Radius from LS section.
9. The Mw of each peak and its fraction is presented in the Results Report.

3.3.2 Characterization by Negative Stain Electron Microscopy

Buffer components that stabilize a protein or protein complex may react adversely with the negative stain used in NS-EM. Therefore, it may be necessary to remove these components prior to preparing a sample for observation. One way to retain the sample's integrity and render it compatible with the stain is to mildly crosslink it [22] and then to exchange the buffer to a buffer suitable for negative staining. Another advantage of crosslinking is to preserve the concentration-dependent interactions that may be lost upon sample dilution. Common crosslinking agents include BS3, DSS, and GLT. Here we describe crosslinking with GLT. It is advisable to first perform a titration series of crosslinker concentrations to find an optimal concentration that leads to the generation of a new high-molecular weight band but minimizes aggregated protein in the well of the SDS-PAGE gel. We have found that 0.05% (v/v) GLT works well for LigIIIα complexes. Examples of successful crosslinking of LigIIIα complexes are shown in Fig. 7a.

Mild Chemical Crosslinking of DNA Ligase III Samples

1. Prepare a 10× solution of GLT in the same buffer as the protein complex, and protect from light with aluminum foil.
2. With the protein sample on ice, add 10× GLT to a final concentration of 1×. *See* Subheading 4, (*see* **Note 31**).
3. Leave for 5 min on ice in the dark.
4. Add 1 M buffered Tris or Glycine to final concentration of 50–100 mM final to quench unreacted GLT.
5. Pipette the sample into a spin column, and dilute with NS-EM buffer. *See* Subheading 4, (*see* **Note 32**).
6. Centrifuge for 5 min.

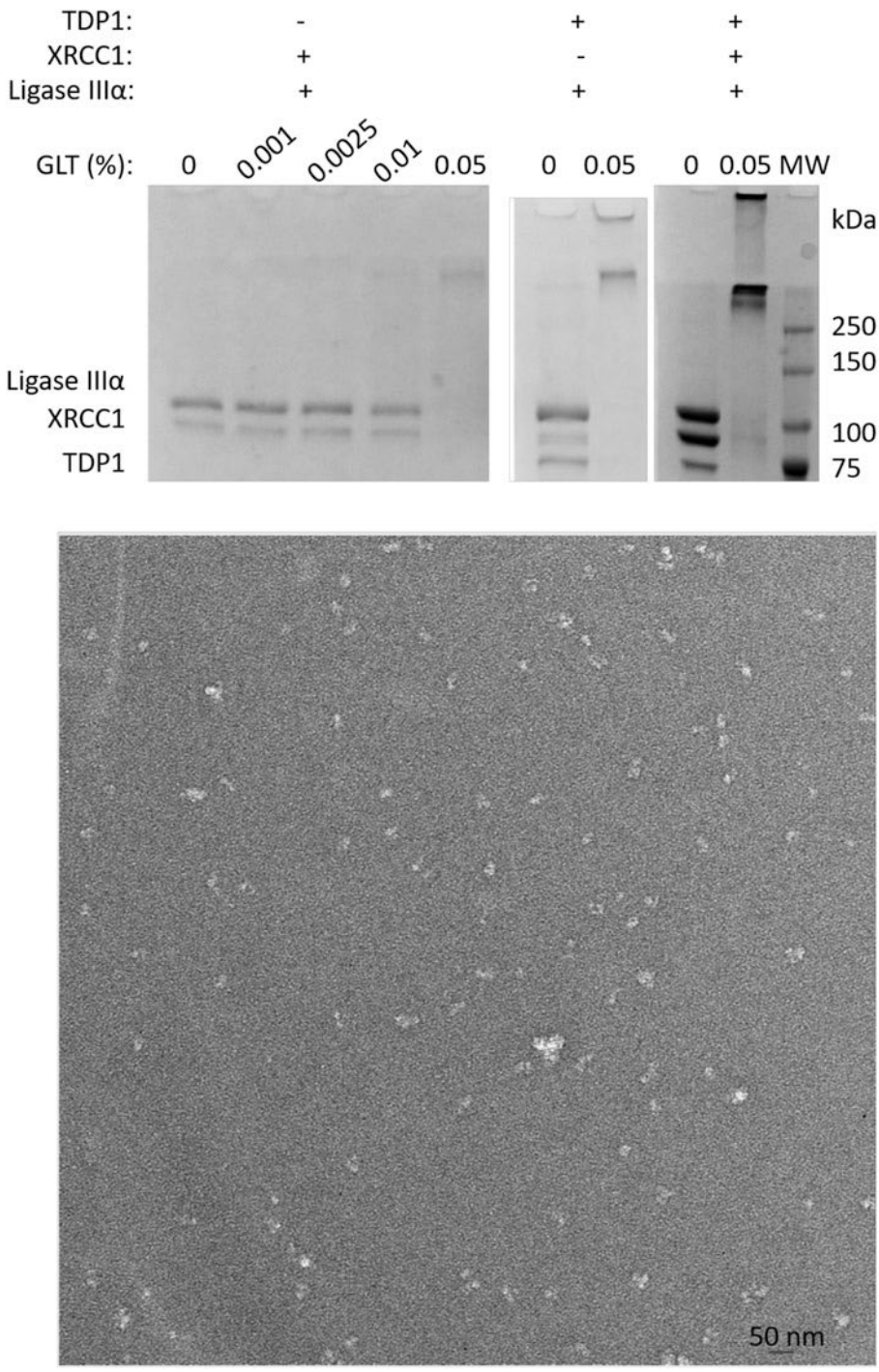


Fig. 7 (**a**) SDS-PAGE gels of LigIIIα complexed with XRCC1, TDP1, or both. In each gel the samples are shown before and after mild chemical crosslinking with GLT. (**b**) Example negative stain micrograph of LigIIIα crosslinked to TDP1 showing a moderate concentration of particles

7. Dilute the spin column retentate further with NS-EM buffer.
8. Repeat **steps 6** and **7** until a sufficient buffer exchange has taken place (typically three 50-fold dilutions).
9. Load samples of non-crosslinked and crosslinked proteins onto an SDS-PAGE gel, stain, and evaluate the crosslinking efficiency (*see* Fig. 7a).

Negative Stain Electron Microscopy

Detailed protocols and variations on the NS-EM technique are well documented elsewhere [23–25]. Here we present a simple overview of NS-EM grid preparation of LigIIIα complexes. A grid with a moderate particle concentration is shown in Fig. 7b. Single particle image analysis is not discussed, as different software have their own workflows.

1. Prepare a solution of uranyl acetate (1–2%) or uranyl formate (1–1.5%). Uranyl acetate may be stored for 1 year at 4 °C, whereas uranyl formate can only be stored for 1–2 days at room temperature. *See* Subheading 4, (*see* **Note 33**).
2. Glow-discharge carbon-coated copper grids for 30 s with 15 mA plasma current, or similar, with the carbon side exposed to the air.
3. With a grid clamped in negative action tweezers, carefully pipette 5 μL of sample onto the grid, and let incubate 1 min. *See* Subheading 4, (*see* **Note 34**).
4. Blot excess sample solution using Whatman paper.
5. Immediately pipette 5 μL of negative stain, and let incubate 1 min.
6. Blot off the negative stain using Whatman paper.
7. Remove the grid by advancing Whatman paper between the tweezer arms.
8. Allow the grid to air dry for 5 min.
9. Insert the grid into a transmission electron microscope, and evaluate the particle concentration. *See* Subheading 4, (*see* **Note 34**) if the concentration is not ideal.
10. Collect micrographs and carry out single particle analysis.

4 Notes

1. To express multi-protein complexes in insect cells, typically, separate viruses for each protein of interest are produced, and then cells are co-infected with multiple viruses at once. This approach allows flexible combinations of individual viruses to create protein complexes of interest, but often requires extensive optimization of co-infection ratios of these viruses and

becomes challenging when more than two viruses are used. Moreover, the probabilities of co-infection of single cells with more than two viruses have been shown to be low [26]. We have developed a multigene baculovirus system, called MacroBac [27], for efficient cloning and assembly of multigenes in the same shuttle vector, thus producing a single baculovirus carrying multiple proteins of interest. Using the MacroBac system significantly increases the reproducibility of protein complex purification in terms of expression levels and subunit stoichiometry. However, the MacroBac single virus does not always work effectively when interactions of proteins in the complex are transient and dynamic, or when there is a vast difference in their sizes. The optimal strategy for efficient production of large, multi-protein complexes should be determined empirically, but can be achieved successfully by co-infection with two MacroBac viruses.

2. Good quality of bacmid DNA is key to producing high-titer baculovirus stocks. Take great care in all steps in producing bacmid DNA.
3. It is important to prepare Luria agar plates as described. Do not use commercial Luria-Bertani agar that contains tryptone, which does not work for the blue/white selection using Bluo-gal.
4. Purify bacmid DNA from 1.5 mL of overnight culture using isopropanol precipitation that generates sufficient amount of DNA for up to 6 transfections. It is not necessary to purify from larger cultures using spin columns as this tends to produce poor quality bacmid DNA.
5. Use clear microcentrifuge tubes when precipitating bacmid DNA by isopropanol, as it should be a small DNA pellet of 1–1.5 mm in diameter. Avoid overprecipitation during bacmid preparation. Discard bacmid samples, and start over if a large, white precipitate (>2 mm in diameter) is obtained.
6. Do not overdry bacmid DNA. Add TE buffer as soon as bacmid DNA becomes transparent, and allow DNA to dissolve for a minimum of 15 min. Gently tap the tubes to dissolve. Avoid pipetting up and down bacmid DNA.
7. Always restreak the white colonies to ensure the white phenotype of the bacmid clones before transfection.
8. Always analyze bacmid DNA by agarose gel electrophoresis to ensure intact, unbroken bacmid is present (Fig. 1) before use for transfection.
9. We grow Sf9 cells in media supplemented with 1% fetal bovine serum (FBS) to produce baculovirus stocks for better viral stability, and use Sf9 cells maintained in serum-free media for

protein expression. Both Sf9 and High Five cells are maintained in antibiotic-free media. Antibiotics are added at half-strength (0.5× concentration) in large-scale cultures for protein expression.

10. Using the described procedures above, high-titered P2ii stocks in the range of 2–10 × 10^8 pfu/mL (i.e., plaque-forming unit per mL of virus) can be reproducibly generated. Traditionally, the viral titers or the infection potency of a viral stock may be determined by plaque formation in immobilized monolayer culture, which is a tedious and long procedure, taking up to 10 days to complete. However, optimal infection conditions vary. The viral titers do not always correlate directly to protein expression levels. The optimal MOI (multiplicity of infection, aka viral particles per cell) has to be determined empirically for each protein of interest. We use a proxy titration method to determine the optimal MOI for protein expression by infecting insect cells with varying volumes of virus, from as low as 1 μL, up to 40 μL, per million cells, and select the best virus/cell ratio based on the total protein expression analysis. This method saves time and streamlines the expression workflow with consistent expression for protein of interest.
11. Store baculovirus at 4 °C in a cardboard box, or use foil to protect the virus from light. The shelf-life of baculovirus is at least 6 months, up to 1 year. However, the stability of every stock is different. Re-test protein expression in 6-well plates every 6 months, and adjust virus amounts accordingly for protein expression.
12. For long-term storage, add 10% FBS to the baculovirus stock, that is, mix 0.5 mL FBS with 4.5 mL virus, and store at −80 °C in 1 mL aliquots. After thawing, one vial of virus can infect up to 100 mL of fresh Sf9 cells, and produce 100 mL new viral stock. Always perform expression tests in six-well plates or small-scale shaker cultures to confirm expression and/or protein complex formation before scaling up expression.
13. When expressing protein complexes using more than two viruses to infect insect cells, test combinations of viral ratios in small-scale shaker cultures, and perform an affinity co-purification to select the best co-infection ratio that yields a protein complex of the best stoichiometry (Fig. 2). To test viral ratios, keep the virus amount constant for the difficult-to-express or larger-sized proteins, and adjust (decrease) the virus amount for the easier-to-express or smaller-sized proteins.
14. Expression conditions determined from 30 to 50 mL cultures are often scalable to 400-mL or larger cultures. However, it is best to run a scale test in the desired culture volume using two of the best possible ratios for a time course and establish optimal expression conditions before scaling up.

15. Benzamidine and PMSF should be added to the buffers right before using them in the ÄKTA FPLC.
16. Keeping lysates on ice in between sonication cycles is crucial in limiting proteolysis and degradation.
17. Filtering the cell lysate after the centrifugation step helps to remove residual debris and particulates, which, if not removed, can cause high back pressure during the FPLC runs.
18. While using a disposable HiTrap FPLC column, make sure the amount of protein loaded onto the column is within the binding capacity of the column beads.
19. Re-useable columns should be stored in 20% ethanol to prevent bacterial growth. Prior to equilibration in running buffer, flush stored columns with water to remove the ethanol.
20. Adding 30 mM imidazole to the lysate before loading on the nickel column helps in reducing non-specific protein binding.
21. For stepwise elution, wait until the UV absorbance becomes flat before moving on to the next step. This usually occurs within ten fractions for the nickel column but if not, wait until it becomes flat and then move on to the next step.
22. Before running the size exclusion column step, it is important to calibrate the column with protein standards. This determines the void volume and expected elution positions of proteins (volume) according to their molecular weight.
23. While loading the sample in a size exclusion column, it is very important not to introduce air bubbles into the column. To avoid bubbles, add buffer to the top of the sample loading loop inlet, and then connect the injection port tube on that end while keeping the bottom outlet of the sample loading loop open and unattached to the other injection port tube. As soon as samples start to come out of the bottom end as droplets, connect the other injection port tube. This will stop air bubbles from getting into the column through the sample loading process.
24. It is important to process the eluted fractions for the next purification step or store (at −80 °C freezer) as soon as they come out of the ÄKTA system. This will limit protein complex degradation.
25. While packing the disposable econo column for dsDNA cellulose from calf thymus, wash the dsDNA cellulose via centrifugation in falcon tubes before packing them into the column using gravity.
26. For concentrating, eluted protein complexes with Amicon Ultra 4 or 15 centrifugal filter units. The cutoff value for the filter molecular weight should be selected to be smaller than

the protein complex molecular weight. These units can also be used for buffer exchange.

27. All the LigIIIα protein complexes were concentrated and stored in P200 buffer (***Buffer,*** Subheading 2.2.1) in small aliquots to avoid repeated freeze/thaw cycles that may cause protein degradation or loss of activity. Aliquots should be flash frozen in liquid nitrogen and then transferred to the −80 °C freezer.
28. Biophysical analysis of protein requires high purity (>95%) and concentration (2–10 mg/mL) of the sample. Concentrations of the purified protein complexes are estimated using the Bradford assay and staining with Coomassie blue after an SDS-PAGE gel. BSA is used as the standard in both assays. The SDS-PAGE gel provides information about the purity of the protein complex with low concentrations revealing the presence of multiple bands with a similar electrophoretic mobility and higher concentrations revealing minor background contaminants.
29. Allow equilibration of at least two column volumes. Ensure that the signals from the various detectors are flat. The pressure over the column may have changed now that the column is equilibrated in the buffer. If this is the case, adjust the flow rate accordingly.
30. Any light scattering detectors that have a low peak signal compared to the baseline noise can be disabled in Molar Mass & Radius from LS.
31. The crosslinking is typically carried out in small volumes (20 μL).
32. In order to recover the original small volume sample without excessive dilution, it is recommended to use a 0.5 mL spin column, as the dead volume is typically 50–60 μL.
33. Uranyl formate requires an adjustment of pH after dissolving in boiling water. Typically 2 μL of 5 M NaOH is added to 500 μL of solution and allowed to equilibrate, and then the solution is centrifuged to remove precipitated uranyl formate.
34. We have found that a sample concentration of ~25–50 nM incubated for ~1 min results in a moderate concentration of particles in the field of view. If the sample concentration is unknown, one may use serial dilutions to find an optimal particle concentration. If the undiluted sample results in too few particles, one may increase on-grid incubation time or apply the sample multiple times after blotting excess.

Acknowledgments

This work was supported in part by National Institutes of Health (NIH) grants R01 ES012512 (A.E.T.) and a Discovery grant RGPIN-2015-05776 from the Natural Science and Engineering Research Council of Canada (J.M.P.). The collaboration between the Tomkinson and Pascal laboratories and the Expression and Molecular Biology Core led by Tsai was supported by the Structural Cell Biology of DNA Repair Program (P01 CA92584). AET acknowledges support from the University of New Mexico Comprehensive Cancer Center.

References

1. Abdella R, Talyzina A, Chen S, Inouye CJ, Tjian R, He Y (2021) Structure of the human mediator-bound transcription preinitiation complex. Science 372(6537):52–56
2. Chen S, Lee L, Naila T, Fishbain S, Wang A, Tomkinson AE, Lees-Miller SP, He Y (2021) Structural basis of long-range to short-range synapsis in NHEJ. Nature 593(7858): 294–298
3. Hammel M, Yu Y, Mahaney BL, Cai B, Ye R, Phipps BM, Rambo RP, Hura GL, Pelikan M, So S, Abolfath RM, Chen DJ, Lees-Miller SP, Tainer JA (2010) Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. J Biol Chem 285:1414–1423
4. Hammel M, Yu Y, Radhakrishnan SK, Chokshi C, Tsai MS, Matsumoto Y, Kuzdovich M, Remesh SG, Fang S, Tomkinson AE, Lees-Miller SP, Tainer JA (2016) An intrinsically disordered APLF links Ku, DNA-PKcs, and XRCC4-DNA ligase IV in an extended flexible non-homologous end joining complex. J Biol Chem 291:26987–27006
5. Hammel M, Rashid I, Sverzhinsky A, Pourfarjam Y, Tsai MS, Ellenberger T, Pascal JM, Kim IK, Tainer JA, Tomkinson AE (2021) An atypical BRCT-BRCT interaction with the XRCC1 scaffold protein compacts human DNA ligase IIIalpha within a flexible DNA repair complex. Nucleic Acids Res 49:306–321
6. Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A (2016) FoXS, FoXSDock and MultiFoXS: single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. Nucleic Acids Res 44: W424–W429
7. Ellenberger T, Tomkinson AE (2008) Eukaryotic DNA ligases: structural and functional insights. Annu Rev Biochem 77:313–338
8. Lakshmipathy U, Campbell C (1999) The human DNA ligase III gene encodes nuclear and mitochondrial proteins. Mol Cell Biol 19: 3869–3876
9. Caldecott KW, McKeown CK, Tucker JD, Ljungquist S, Thompson LH (1994) An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. Mol Cell Biol 14:68–76
10. Caldecott KW (2019) XRCC1 protein; form and function. DNA Repair (Amst) 81:102664
11. Gao Y, Katyal S, Lee Y, Zhao J, Rehg JE, Russell HR, McKinnon PJ (2011) DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair. Nature 471:240–244
12. Simsek D, Furda A, Gao Y, Artus J, Brunet E, Hadjantonakis AK, Van Houten B, Shuman S, McKinnon PJ, Jasin M (2011) Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. Nature 471:245–248
13. Le Chalony C, Hoffschir F, Gauthier LR, Gross J, Biard DS, Boussin FD, Pennaneach V (2012) Partial complementation of a DNA ligase I deficiency by DNA ligase III and its impact on cell survival and telomere stability in mammalian cells. Cell Mol Life Sci 69: 2933–2949
14. Lakshmipathy U, Campbell C (2000) Mitochondrial DNA ligase III function is independent of Xrcc1. Nucleic Acids Res 28:3880–3886
15. Huang SY, Murai J, Dalla Rosa I, Dexheimer TS, Naumova A, Gmeiner WH, Pommier Y (2013) TDP1 repairs nuclear and mitochondrial DNA damage induced by chain-terminating anticancer and antiviral nucleoside analogs. Nucleic Acids Res 41:7793–7803
16. El-Khamisy SF, Saifi GM, Weinfeld M, Johansson F, Helleday T, Lupski JR, Caldecott KW (2005) Defective DNA single-strand break

repair in spinocerebellar ataxia with axonal neuropathy-1. Nature 434:108–113

17. Della-Maria J, Zhou Y, Tsai MS, Kuhnlein J, Carney JP, Paull TT, Tomkinson AE (2011) Human Mre11/human Rad50/Nbs1 and DNA ligase IIIalpha/XRCC1 protein complexes act together in an alternative nonhomologous end joining pathway. J Biol Chem 286: 33845–33853
18. Cannan WJ, Rashid I, Tomkinson AE, Wallace SS, Pederson DS (2017) The human ligase IIIalpha-XRCC1 protein complex performs DNA Nick repair after transient unwrapping of nucleosomal DNA. J Biol Chem 292: 5227–5238
19. Slotboom DJ, Duurkens RH, Olieman K, Erkens GB (2008) Static light scattering to characterize membrane proteins in detergent solution. Methods 46:73–82
20. Wen J, Arakawa T, Philo JS (1996) Size-exclusion chromatography with on-line light-scattering, absorbance, and refractive index detectors for studying proteins and their interactions. Anal Biochem 240:155–166
21. Sverzhinsky A, Fabre L, Cottreau AL, Biot-Pelletier DM, Khalil S, Bostina M, Rouiller I, Coulton JW (2014) Coordinated rearrangements between cytoplasmic and periplasmic domains of the membrane protein complex ExbB-ExbD of Escherichia coli. Structure 22: 791–797
22. Cheng Y, Grigorieff N, Penczek PA, Walz T (2015) A primer to single-particle cryo-electron microscopy. Cell 161:438–449
23. Ohi M, Li Y, Cheng Y, Walz T (2004) Negative staining and image classification - powerful tools in modern electron microscopy. Biol Proced Online 6:23–34
24. Scarff CA, Fuller MJG, Thompson RF, Iadanza, MG (2018) Variations on negative stain electron microscopy methods: tools for tackling challenging systems. J Vis Exp (132): 57199
25. Gallagher JR, Kim AJ, Gulati NM, Harris AK (2019) Negative-stain transmission electron microscopy of molecular complexes for image analysis by 2D class averaging. Curr Protoc Microbiol 54:e90
26. Vijayachandran LS, Viola C, Garzoni F, Trowitzsch S, Bieniossek C, Chaillet M, Schaffitzel C, Busso D, Romier C, Poterszman A, Richmond TJ, Berger I (2011) Robots, pipelines, polyproteins: enabling multiprotein expression in prokaryotic and eukaryotic cells. J Struct Biol 175:198–208
27. Gradia SD, Ishida JP, Tsai MS, Jeans C, Tainer JA, Fuss JO (2017) MacroBac: new technologies for robust and efficient large-scale production of recombinant multiprotein complexes. Methods Enzymol 592:1–26

Chapter 16

Generation of Monoubiquitin and K63-Linked Polyubiquitin Chains for Protein Interaction Studies

Rita Anoh, Kate A. Burke, Dhane P. Schmelyun, and Patrick M. Lombardi

Abstract

Ubiquitylation is a posttranslational modification that utilizes protein-protein binding interactions to regulate cellular processes. In ubiquitin signaling, a vast array of mono- and polyubiquitin modifications to substrate proteins are recognized by a diverse group of ubiquitin-binding proteins. Identifying ubiquitin-binding proteins and characterizing their binding properties is necessary for understanding the structural basis of ubiquitin signaling. This chapter provides a means of studying ubiquitin-binding interactions in vitro by describing how to generate monoubiquitin and K63-linked polyubiquitin chains and perform pull-down assays with ubiquitin-binding proteins, which is of particular relevance for DNA damage and other signaling pathways.

Key words Ubiquitin, Polyubiquitin chains, Cellular signaling, Ubiquitin-binding protein, Pull-down assay, Ubiquitin-activating enzyme, Ubiquitin-conjugating enzyme, Ubiquitin-conjugation reaction

1 Introduction

Ubiquitylation is a posttranslational modification that most commonly involves the enzyme-catalyzed covalent bonding of the C-terminal glycine residue of ubiquitin to the ε-amino group of a lysine residue from a substrate protein. Attachment of the 76-amino acid ubiquitin alters the surface of the substrate protein and creates a binding site for proteins containing one or more of the many ubiquitin-binding domains that have been characterized [1, 2]. Ubiquitin itself has seven lysine residues in its primary structure, allowing for the formation of polyubiquitin chains [3]. The diversity of cellular processes regulated by ubiquitylation is possible because polyubiquitin chains connected at different lysine residues adopt different structures and are recognized by

Rita Anoh, Kate A. Burke and Dhane P. Schmelyun contributed equally with all other contributors.

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_16,

different ubiquitin-binding domains [4]. Identifying proteins with ubiquitin-binding domains and characterizing their ubiquitin-binding properties has provided insight into the mechanisms by which ubiquitin signaling controls various cellular processes. Accordingly, biological roles for many types of mono- and poly-ubiquitin modifications are beginning to be understood in greater detail [5].

In this chapter, we present a collection of protocols for generating monoubiquitin and K63-linked polyubiquitin chains that can be used in pull-down assays to study ubiquitin-binding proteins. The first protocol (Subheading 3.1) describes how to express and purify monoubiquitin from *E. coli* cells using acid precipitation [6]. The next group of protocols explains how to express and purify the E1 ubiquitin-activating enzyme, UBE1 (Subheading 3.2) [7], and the E2 ubiquitin-conjugating enzyme complex, Ubc13/Mms2 (Subheading 3.3) [8], and how to combine these enzymes with monoubiquitin in a conjugation reaction to produce K63-linked polyubiquitin chains (Subheading 3.4) [9]. The last protocol (Subheading 3.5) describes how to perform a pull-down assay with tagged ubiquitin-binding proteins. This chapter focuses on K63-linked polyubiquitin chains because of their well-established role in DNA damage repair, one of the themes of this edition. For example, K63-linked polyubiquitin chains are bound by the tandem ubiquitin-interacting motifs of the protein RAP80 as part of the response to DNA double-strand breaks [10], and K63-linked polyubiquitin chains are recognized by the CUE domain of the protein ASCC2 to target the ALKBH3-ASCC repair complex to sites of DNA alkylation damage [11]. Protocols for assembling other types of polyubiquitin chains have been reported [12, 13], and the approach described in this chapter can be adapted to other types of polyubiquitin chains by substituting Ubc13/Mms2 for other ubiquitin-conjugating enzymes and ubiquitin ligases. In addition to pull-down assays, other techniques such as isothermal titration calorimetry, X-ray crystallography, and nuclear magnetic resonance spectroscopy have been used extensively to determine the affinity, stoichiometry, and structural details of ubiquitin-binding interactions.

2 Materials

The following reagents and instrumentation are necessary for the expression and purification of monoubiquitin, the E1 ubiquitin-activating enzyme UBE1, and the E2 ubiquitin-conjugating enzyme complex Ubc13/Mms2 as described in Subheadings 3.1–3.3:

1. LB-agar plates containing antibiotic.
2. Autoclaved LB media.
3. Refrigerated shaking incubator.
4. UV-Vis spectrophotometer.
5. 1 M isopropyl-β-D-1-thiogalactopuranoside (IPTG).
6. Centrifuge and rotors capable of handling liter-scale volumes.
7. Sonicator for lysis of bacteria.
8. Bicinchoninic acid (BCA) assay kit.
9. Dialysis tubing with a molecular weight cutoff of 3.5 kDa.

2.1 Expressing and Purifying Ubiquitin from E. coli Cells

1. Plasmid containing ubiquitin residues 1–76 suitable for expression in *E. coli* (*see* **Note 1**).
2. Lysis buffer: 50 mM Tris pH 7.6, 1 mM phenylmethylsulfonyl fluoride (PMSF).
3. 70% perchloric acid.

2.2 Expressing and Purifying the E1 Ubiquitin-Activating Enzyme UBE1 from E. coli Cells

1. Plasmid coding for UBE1 suitable for expression in *E. coli* (*see* **Note 2**).
2. Lysis buffer: 25 mM Tris pH 7.5, 150 mM NaCl, 20 mM imidazole, 200 μM TCEP, cOmplete Mini, EDTA-free protease-inhibitor tablet (Roche).
3. Syringe filters with 0.22-μm diameter pores.
4. Fast protein liquid chromatography (FPLC) system, immobilized metal affinity chromatography column, anion-exchange column, size-exclusion column (*see* **Notes 3–5**).
5. Low-imidazole buffer: 20 mM Tris pH 7.5, 150 mM NaCl, 20 mM imidazole, 200 μM TCEP, and high-imidazole buffer: 20 mM Tris pH 7.5, 150 mM NaCl, 400 mM imidazole, 200 μM TCEP.
6. Low ionic strength buffer: 20 mM Tris pH 7.5, 50 mM NaCl, 200 μM TCEP, and high ionic strength buffer: 20 mM Tris pH 7.5, 400 mM NaCl, 200 μM TCEP.
7. Size-exclusion column buffer: 20 mM Tris pH 7.5, 100 mM NaCl, 200 μM TCEP.

2.3 Expressing and Purifying the S. cerevisiae Ubc13/Mms2 E2 Ubiquitin-Conjugating Enzyme Complex from E. coli Cells

1. Plasmid containing *Ubc13 and Mms2* genes suitable for expression in *E. coli* (*see* **Note 6**).
2. Lysis buffer: 25 mM Tris pH 7.5, 150 mM NaCl, 20 mM imidazole, 200 μM TCEP.
3. Syringe filters with 0.22-μm diameter pores.
4. FPLC system, immobilized metal affinity chromatography column, and anion-exchange column (*see* **Notes 3** and **4**).

5. Low-imidazole buffer: 20 mM HEPES pH 7.5, 30 mM imidazole, 200 μM TCEP, and high-imidazole buffer: 20 mM HEPES pH 7.5, 250 mM imidazole, 200 μM TCEP.
6. Low ionic strength buffer: 20 mM HEPES pH 7.2, 25 mM NaCl, 1 mM EDTA, and high ionic strength buffer: 20 mM HEPES pH 7.2, 500 mM NaCl, 1 mM EDTA.

2.4 Assembling and Purifying K63-Linked Polyubiquitin Chains

1. Ubiquitin-conjugation reaction components: HEPES pH 7.5 (50 mM), $MgCl_2$ (10 mM), TCEP pH 7.5 (1 mM), UBE1 ubiquitin-activating enzyme (50 nM), Ubc13/Mms2 ubiquitin-conjugating enzyme complex (2.5 μM), ATP pH 7.2 (10 mM), monoubiquitin (1 mM). The concentrations listed are the final concentrations in the reaction mixture.
2. FPLC system, MonoS 10/100 GL cation-exchange column (Cytiva).
3. Low ionic strength buffer: 50 mM ammonium acetate pH 4.5, 50 mM NaCl, and high ionic strength buffer: 50 mM ammonium acetate pH 4.5, 600 mM NaCl.
4. Dialysis buffer: 20 mM HEPES pH 7.5, 150 mM NaCl, 200 μM TCEP.

2.5 Pull-Down Assay with His-Tagged Ubiquitin-Binding Protein

1. Wash buffer: 50 mM Tris pH 7.5, 150 mM NaCl, 20 mM imidazole, 0.05% Triton X-100, 3 mM β-mercaptoethanol.
2. 2× SDS loading buffer with DTT: 50 μL DTT, 500 μL of 4× SDS loading buffer, 450 μL H_2O.

3 Methods

3.1 Expressing and Purifying Monoubiquitin from E. coli Cells

1. Transform the ubiquitin-containing plasmid into an *E. coli* cell line suitable for protein expression, such as BL21(DE3) cells (*see* **Note 1**).
2. Grow 5-mL cultures by combining a single colony with 5 mL of LB media and antibiotic. The cultures can be grown overnight (~16 h) in an incubator at 37 °C with 225 rpm shaking. Prepare one 5-mL overnight culture for each liter of *E. coli* cells to be used for protein expression.
3. The following morning, inoculate each liter of LB media and antibiotic with one 5-mL overnight culture that has grown to saturation. Grow the 1-L cultures of *E. coli* cells to an OD_{600} between 0.5 and 0.8. Add 0.5 mL of 1 M IPTG, and lower the temperature of the incubator to 16 °C. Allow protein expression to continue at 16 °C overnight with 225 rpm shaking.
4. The following morning, pellet the cells by centrifuging at 7500 × *g* for 15 min at 4 °C. Pour off the supernatant, and

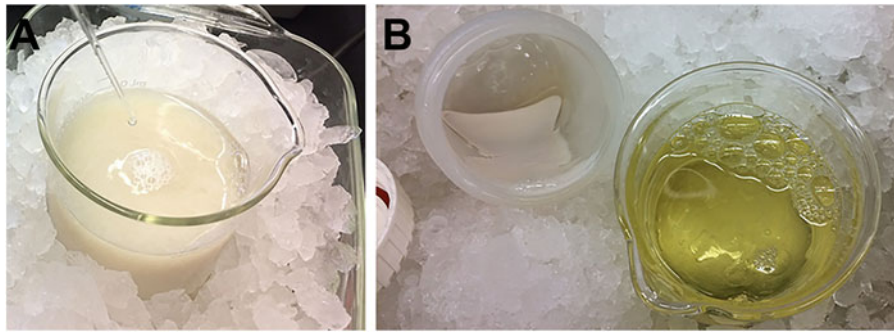


Fig. 1 Purification of monoubiquitin by acid precipitation. (**a**) Adding 70% perchloric acid to cell lysate turns the solution milky white. (**b**) Centrifugation of the acid-treated lysate separates the green-tinted supernatant containing monoubiquitin from the white pellet containing nearly all other *E. coli* proteins

resuspend the cell pellet in lysis buffer. Use approximately 10 mL of lysis buffer per gram of cell pellet (*see* **Note 7**).

5. Lyse the cells by sonicating the resuspended cell pellet on ice (*see* **Note 8**).
6. Centrifuge the cell lysate at 17,500 × *g* for 30 min at 4 °C. Pour the supernatant into a beaker, and place the beaker into a container filled with ice, on a stir plate.
7. While moderately stirring the cell lysate, add 70% perchloric acid dropwise until the volume of added acid is 1% the original volume of the cell lysate. The solution will turn a milky white as the 70% perchloric acid is added (Fig. 1a). Continue to stir the solution for 5 min after the desired volume of acid has been added.
8. Centrifuge the acid-precipitated cell lysate at 17,500 × *g* for 30 min at 4 °C. Pour the supernatant, which contains the purified ubiquitin, into a beaker. It is common for the supernatant to have a green tint (Fig. 1b). The rubbery, white pellet will contain nearly all the other *E. coli* proteins besides ubiquitin.
9. Transfer the supernatant to dialysis tubing, and dialyze overnight in 2 L of 10 mM Tris pH 7.6. Change the dialysis buffer the following day, and continue dialyzing until the ubiquitin-containing solution has reached a neutral pH (*see* **Note 9**).
10. After dialyzing the solution to neutral pH, recover the solution from the dialysis bag, and measure the protein concentration using a BCA assay. Concentrate monoubiquitin to greater than 3 mM if the ubiquitin will be used for subsequent conjugation reactions (*see* **Note 10**).

3.2 Expressing and Purifying the E1 Ubiquitin-Activating Enzyme UBE1 from E. coli Cells

1. Transform the plasmid coding for His-tagged human UBE1 (*see* **Note 2**) into BL21(DE3) *E. coli* cells, or an alternative *E. coli* cell line suitable for protein expression.
2. Prepare 5-mL overnight cultures by adding one colony from the transformation to 5-mL of LB media with antibiotic. Incubate the cultures at 37 °C overnight with 250 rpm shaking. Prepare one 5-mL culture for each liter of cells to be inoculated.
3. Inoculate each 1-L culture of LB media and antibiotic with one 5-mL overnight culture. Grow the cells at 37 °C with 250 rpm shaking until the OD_{600} is between 0.5 and 0.8. Induce protein expression by adding 1 mL of 1 M ITPG and lowering the temperature of the shaking incubator to 16 °C. Allow expression to continue overnight at 16 °C with shaking at 250 rpm.
4. Harvest the cells by centrifuging at 7500 × *g* for 15 min at 4 °C. Pour off the supernatant.
5. Resuspend the cells in lysis buffer. Use approximately 10 mL of lysis buffer per gram of cell pellet (*see* **Note 7**).
6. Lyse the cells by sonication (*see* **Note 8**).
7. Separate the soluble and insoluble fractions of the cell lysate by centrifuging at 17,500 × *g* for 30 min at 4 °C. Discard the pellet, and use a syringe filter with a 0.22-μm pore diameter to filter the supernatant.
8. Load the supernatant onto an immobilized metal affinity chromatography column (*see* **Note 3**) that has been equilibrated in low-imidazole buffer. Once all the unbound protein has flowed through the column, as judged by the absorbance at 280 nm returning to baseline, run a linear gradient from 0% high-imidazole buffer to 50% high-imidazole buffer over 100 mL. The UBE1 protein will elute during the first half of the gradient.
9. Use SDS-PAGE to determine which fractions contain UBE1, and then combine and dialyze these fractions overnight in 2 L of low ionic strength buffer.
10. Recover the protein solution from dialysis and load onto an anion-exchange column (*see* **Note 4**) equilibrated in low ionic strength buffer. Once all the unbound protein has passed through the column, as judged by the absorbance at 280 nm returning to baseline, elute the UBE1 protein by running a linear gradient from 0% high ionic strength buffer to 100% high ionic strength buffer over 100 mL. The UBE1 protein will elute during the middle of the gradient.
11. Use SDS-PAGE to determine which fractions contain UBE1, and concentrate the fractions to a combined volume of less than 5 mL.

12. Further purify the UBE1 protein by passing the concentrated protein solution over a size-exclusion column (*see* **Note 5**) that has been equilibrated in size-exclusion column buffer. The UBE1 protein elutes over a range from 48 to 54 mL after injection.
13. Concentrate the fractions containing the purest UBE1, as judged by SDS-PAGE, to 50 μM, add 10% glycerol to the solution, and store 25-μL aliquots at −80 °C.

3.3 Expressing and Purifying the S. cerevisiae Ubc13/Mms2 E2 Ubiquitin-Conjugating Enzyme Complex from E. coli Cells

1. Transform plasmid(s) containing the *S. cerevisiae* genes *Ubc13* and *Mms2* into BL21(DE3) *E. coli* cells, or into another *E. coli* cell line suitable for protein expression. For this protocol, either Ubc13 or Mms2 must contain a 6× polyhistidine tag.
2. Grow 5-mL overnight cultures by inoculating 5 mL of LB media and antibiotic with a single colony from the transformation. Grow cultures to saturation overnight at 37 °C with 250 rpm shaking. Prepare one overnight culture for each liter of LB media that will be inoculated the following day.
3. The following day, inoculate each 1-L aliquot of LB media and antibiotic with a single overnight culture. Grow the cultures at 37 °C with 250 rpm shaking until the OD_{600} is between 0.5 and 0.8. Induce protein expression by adding 1 mL of 1 M ITPG to the cultures, and lower the temperature of the shaking incubator to 16 °C. Let protein expression continue overnight at 16 °C with 250 rpm shaking.
4. Harvest the cells by centrifuging at 7500 × *g* for 15 min at 4 °C. Discard the supernatant.
5. Resuspend the cells in low imidazole buffer. Use approximately 10 mL of lysis buffer per gram of cell pellet (*see* **Note 7**).
6. Lyse the cells using a sonicator (*see* **Note 8**).
7. Centrifuge the cell lysate at 17,500 × *g* for 30 min at 4 °C. Discard the cell pellet, and filter the supernatant through a syringe filter with 0.22-μm pores.
8. Load the supernatant onto an immobilized metal affinity chromatography column that has been equilibrated in low-imidazole buffer (*see* **Note 3**). Once all the unbound protein has passed through the column, as judged by the absorbance at 280 nm returning to baseline, begin a linear gradient to 100% high-imidazole buffer over 100 mL. Ubc13 and Mms2 will elute near the middle of the gradient.
9. Use SDS-PAGE to determine the fractions from the immobilized-metal affinity chromatography column that contain Ubc13 and Mms2. Combine these fractions in dialysis tubing, along with 0.5 mg of TEV protease (*see* **Note 11**), and dialyze overnight in low ionic strength buffer.

10. Recover the protein solution from dialysis, and load onto an anion-exchange column that has been equilibrated in low ionic strength buffer (*see* **Note 4**). After all the unbound protein has flowed through, as judged by the absorbance at 280 nm returning to baseline, begin a linear gradient to 100% high ionic strength buffer over 100 mL. Ubc13 and Mms2 elute towards the beginning of the gradient.
11. Use SDS-PAGE to determine which fractions contain the purest Ubc13 and Mms2. Concentrate the proteins to approximately 25 μM, and store 100-μL aliquots at −80 °C.

3.4 Assembling and Purifying K63-Linked Polyubiquitin Chains

1. Combine the following reagents in a microcentrifuge tube at the following final concentrations: HEPES pH 7.5 (50 mM), $MgCl_2$ (10 mM), TCEP pH 7.5 (1 mM), UBE1 ubiquitin-activating enzyme (50 nM), Ubc13/Mms2 ubiquitin-conjugating enzyme complex (2.5 μM), ATP pH 7.2 (10 mM), monoubiquitin (1 mM) (*see* **Note 12**). Bring the reaction to the desired final volume with water. Successful conjugation reactions have been run at volumes ranging from hundreds of microliters to several milliliters.
2. Incubate the reaction overnight in a water bath at 37 °C.
3. The following day, dilute the reaction tenfold in low ionic strength buffer. Load the diluted reaction solution onto a MonoS 10/100 GL cation-exchange column (Cytiva) that has been equilibrated in low ionic strength buffer (*see* **Note 13**). Once all the unbound protein has flowed through the column, as judged by the absorbance at 280 nm returning to baseline, run a gradient from 0% high ionic strength buffer to 100% high ionic strength buffer over 300 mL. Monoubiquitin elutes from the column first, and each of the subsequent peaks corresponds to a polyubiquitin chain with one additional ubiquitin protomer (Fig. 2a).
4. Use SDS-PAGE to determine the contents of fractions collected from the peaks in the chromatograph (Fig. 2b). Combine the desired fractions in dialysis tubing, and dialyze in 20 mM HEPES pH 7.5, 150 mM NaCl, and 200 μM TCEP until the protein solutions reach a neutral pH.

3.5 Pull-Down Assay with His-Tagged Ubiquitin-Binding Protein

1. Add 1 mL of cold wash buffer to 400 μL of TALON or nickel NTA agarose beads. Vortex briefly, and then spin down at 850 × *g* for 60 s. Use a pipette to remove the supernatant down to just above bead level. Repeat three times.
2. To block the beads, add an equal volume of wash buffer plus 1% BSA. Rotate the tubes end over end for 1 h at 4 °C.
3. Vortex beads and aliquot 20 μL to each microcentrifuge tube (*see* **Note 14**).

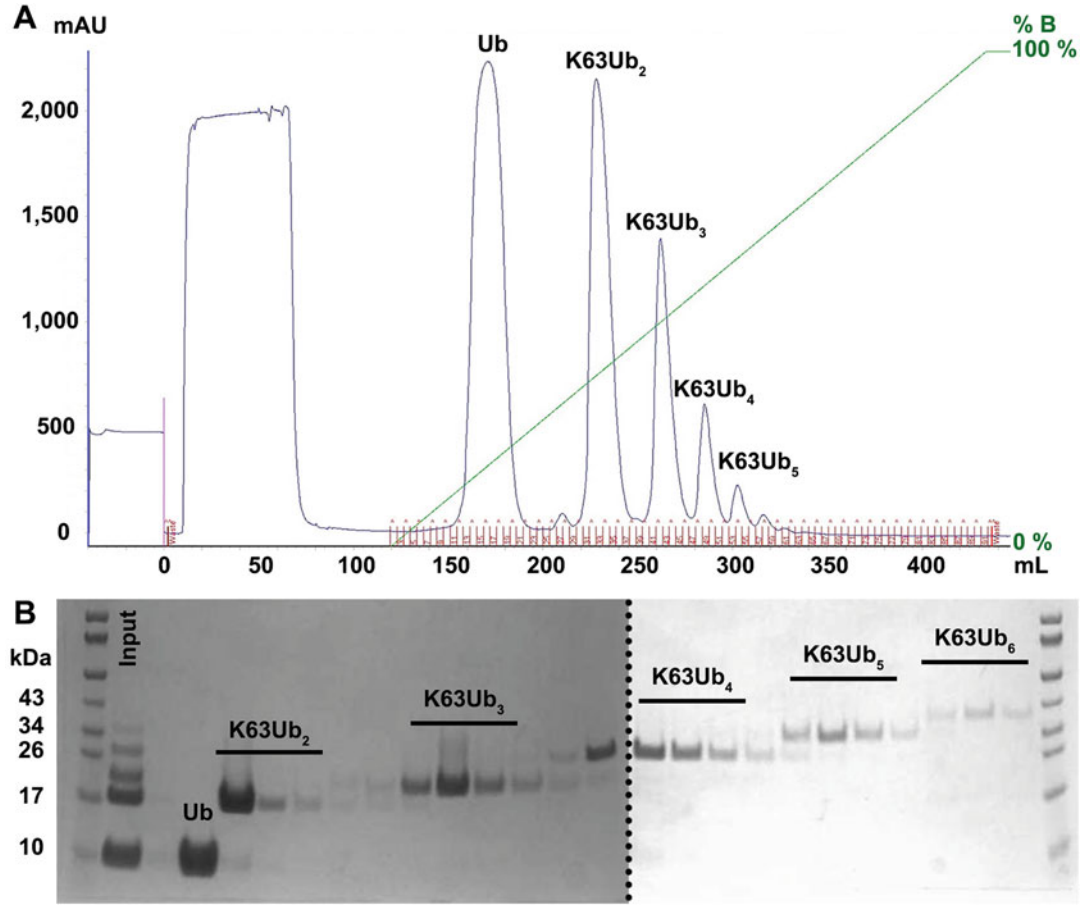


Fig. 2 Purification of K63-linked polyubiquitin chains using cation-exchange chromatography. (**a**) Chromatograph from MonoS 10/100 GL cation-exchange column shows a linear gradient of increasing salt concentration resolves a mixture of K63-linked polyubiquitin chains of different lengths into separate peaks. (**b**) SDS-PAGE reveals that monoubiquitin elutes from the column first and that subsequent peaks contain K63-linked polyubiquitin chains that increase in increments of one ubiquitin

4. Add solution of His-tagged ubiquitin-binding protein to the beads. When deciding on the volume and concentration of His-tagged ubiquitin-binding protein to add, be cognizant of the K_d for ubiquitin binding, if known, and the detection limit for the intended imaging method (e.g., staining with Coomassie blue or Western blotting). Weaker ubiquitin-binding proteins can exhibit K_d values in the hundreds of micromolar range.
5. Add solution of monoubiquitin or polyubiquitin chains to the beads. Using a ubiquitin concentration tenfold greater than that of the His-tagged protein can drive the saturation of the available ubiquitin-binding sites.

6. Bring binding mixture to desired final volume by adding wash buffer.
7. Vortex reactions briefly and rotate the tubes end over end for 1 h at 4 °C to allow for binding.
8. Add 1 mL of cold wash buffer, vortex briefly, and spin down the binding mixtures at 850 × *g* for 30 s.
9. Remove the supernatant down to the level of the beads. Repeat the wash step three times to remove unbound ubiquitin.
10. Add 20 μL of 2× SDS loading buffer with DTT to the beads, and boil at 95 °C for 10 min.
11. Analyze pull-down results by SDS-PAGE followed by the desired imaging technique such as Coomassie staining, silver staining, or Western blotting (*see* **Note 15**).

4 Notes

1. A plasmid containing ubiquitin residues 1–76 in the pET15 vector can be obtained from the plasmid repository, Addgene (Plasmid #12647).
2. A plasmid coding for UBE1 inserted into the pET21d vector is available at the plasmid repository, Addgene (Plasmid #34965).
3. A 5-mL HisTrap HP column from Cytiva can be used for the immobilized metal affinity chromatography step.
4. A 5-mL HiTrap Q FF column from Cytiva can be used for the anion-exchange chromatography step.
5. A HiLoad Superdex 75 16/60 column from Cytiva can be used for the size-exclusion chromatography step.
6. A plasmid coding for TEV-cleavable His-tagged Mms2 and untagged Ubc13 in the pST39 vector is available from the corresponding author.
7. The resuspension process can be sped up by dislodging the cell pellet from the container using a spatula, transferring the cells and lysis buffer to a beaker, and stirring at 4 °C with a magnetic stir bar until the solution is homogeneous.
8. For the YUCHENGTECH Ultrasonic processor, sonicating for alternating intervals of 10 s on and 10 s off at 60% amplification over 30 min is sufficient to lyse the cells. The *E. coli* cells can also be lysed using other instruments, such as a microfluidizer.
9. The ubiquitin solution will be quite acidic following perchloric acid precipitation, and multiple rounds of dialysis may be necessary to bring the solution to neutral pH. Acidic ubiquitin

solutions can cause downstream applications, like the ubiquitin-conjugation reaction, to fail, so care should be taken to ensure that a neutral pH has been achieved.

10. For an estimate of the ubiquitin concentration, measure the absorbance of the solution at 280 nm, and use the approximation that 1 mg/mL of ubiquitin has an absorbance of 0.17.
11. A plasmid coding for His-tagged TEV protease is available at the plasmid repository, Addgene (Plasmid #125194). The His-tagged TEV protease can be expressed in *E. coli* cells and purified by immobilized metal affinity chromatography.
12. To limit the polyubiquitin chain length to diubiquitin, substitute K48R/K63R ubiquitin (1 mM) and D77 ubiquitin (1 mM) for wild-type ubiquitin in the conjugation reaction. The K48R/K63R mutations will prevent chain extension on the distal ubiquitin and the D77 mutation will prevent chain extension on the proximal ubiquitin [9].
13. A 5-mL HiTrap SP HP column (Cytiva) can also be used to purify K63-linked diubiquitin from monoubiquitin if the FPLC system is not compatible with the pressure requirements for running a MonoS 10/100 GL cation-exchange column.
14. Cutting off the end of the pipette tip can increase the diameter of the opening and make it easier to transfer the beads.
15. For anti-ubiquitin Western blots, perform a wet transfer, and then block the membrane for 30 min at room temperature in 5% milk in TBST solution. Wash the membrane twice for 5 min each in TBST solution, and then incubate the membrane overnight in a 1:1000 dilution of mouse anti-ubiquitin (Santa Cruz P4D1) in a solution of TBST with 0.5% BSA and 0.02% sodium azide. The following day, wash the membrane three times for 5 min each with TBST solution. Then, incubate the membrane for 30 min at room temperature in a 1:20,000 dilution of anti-mouse IgG HRP (Sigma-Aldrich NA931) in a TBST solution containing 1% BSA. Finally, wash the blot three times for 5 min each with TBST solution, add 1 mL of enhanced chemiluminescent substrate (Thermo Fisher SuperSignal West Femto Maximum Sensitivity Substrate), and image.

Acknowledgments

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R15GM140410. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Hurley JH, Lee S, Prag G (2006) Ubiquitin-binding domains. Biochem J 399:361–372. https://doi.org/10.1042/BJ20061138
2. Randles L, Walters KJ (2012) Ubiquitin and its binding domains. Front Biosci (Landmark Ed) 17:2140–2157. https://doi.org/10.2741/4042
3. Akutsu M, Dikic I, Bremm A (2016) Ubiquitin chain diversity at a glance. J Cell Sci 129:875–880. https://doi.org/10.1242/jcs.183954
4. Komander D, Rape M (2012) The ubiquitin code. Annu Rev Biochem 81:203–229. https://doi.org/10.1146/annurev-biochem-060310-170328
5. Swatek KN, Komander D (2016) Ubiquitin modifications. Cell Res 26:399–422. https://doi.org/10.1038/cr.2016.39
6. Piotrowski J, Beal R, Hoffman L et al (1997) Inhibition of the 26 S proteasome by polyubiquitin chains synthesized to have defined lengths. J Biol Chem 272:23712–23721. https://doi.org/10.1074/jbc.272.38.23712
7. Berndsen CE, Wolberger C (2011) A spectrophotometric assay for conjugation of ubiquitin and ubiquitin-like proteins. Anal Biochem 418:102–110. https://doi.org/10.1016/j.ab.2011.06.034
8. Berndsen CE, Wiener R, Yu IW et al (2013) A conserved asparagine has a structural role in ubiquitin-conjugating enzymes. Nat Chem Biol 9:154–156. https://doi.org/10.1038/nchembio.1159
9. Pickart CM, Raasi S (2005) Controlled synthesis of polyubiquitin chains. Methods Enzymol 399:21–36. https://doi.org/10.1016/S0076-6879(05)99002-2
10. Sato Y, Yoshikawa A, Mimura H et al (2009) Structural basis for specific recognition of Lys 63-linked polyubiquitin chains by tandem UIMs of RAP80. EMBO J 28:2461–2468. https://doi.org/10.1038/emboj.2009.160
11. Brickner JR, Soll JM, Lombardi PM et al (2017) A ubiquitin-dependent signalling axis specific for ALKBH-mediated DNA dealkylation repair. Nature 551:389–393. https://doi.org/10.1038/nature24484
12. Michel MA, Komander D, Elliott PR (2018) Enzymatic assembly of ubiquitin chains. Methods Mol Biol 1844:73–84. https://doi.org/10.1007/978-1-4939-8706-1_6
13. Dong KC, Helgason E, Yu C et al (2011) Preparation of distinct ubiquitin chain reagents of high purity and yield. Structure 19:1053–1063. https://doi.org/10.1016/j.str.2011.06.010

Correction to: Approaches to Monitor Termination of DNA Replication Using Xenopus Egg Extracts

Tamar Kavlashvili and James M. Dewar

Correction to:
Chapter 7 in: Nima Mosammaparast (ed.), ***DNA Damage Responses: Methods and Protocols,*** **Methods in Molecular Biology, vol. 2444,**
https://doi.org/10.1007/978-1-0716-2063-2

The chapter was inadvertently published with the Acknowledgment section excluded in the chapter.

The correction has been incorporated by including the Acknowledgment as "JMD is supported by NIH grant R35GM128696" in the chapter.

The updated online version of the chapter can be found at
https://doi.org/10.1007/978-1-0716-2063-2_7

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*,
Methods in Molecular Biology, vol. 2444, https://doi.org/10.1007/978-1-0716-2063-2_17,

INDEX

Nima Mosammaparast (ed.), *DNA Damage Responses: Methods and Protocols*, Methods in Molecular Biology, vol. 2444,
https://doi.org/10.1007/978-1-0716-2063-2,

N

O

P

R

S

T

U

V

X