Plan Overview

A Data Management Plan created using DMPTool

Title: Structural basis of eukaryotic clamp loading

Creator: Dirk Remus - **ORCID:** 0000-0002-5155-181X

Affiliation: Memorial Sloan Kettering Cancer Center (mskcc.org)

Funder: National Institute of Environmental Health Sciences (niehs.nih.gov)

Funding opportunity number: PA-20-185

Template: NIH-Default DMSP

Project abstract:

The maintenance of genome stability across generations is critical for human health and relies on the efficient repair of spontaneous DNA lesions and the faithful duplication of all chromosomal DNA prior to cell division. The highly conserved DNA sliding clamps PCNA and 9-1-1 are critical for these genome maintenance mechanisms in eukaryotic cells by acting as mobile hubs for the assembly of the protein complexes mediating the signaling and repair of DNA damage and conducting the faithful replication of the nuclear chromosomes. The dynamic association of PCNA and 9-1-1 with chromosomal DNA is controlled by a set of four conserved and related ATP-dependent clamp loader complexes that each perform non-redundant genome maintenance functions in the cell. How eukaryotic clamp loaders target their client clamps to sites of DNA replication and repair and load the clamps around DNA has been the subject of intense investigation for several decades. Using advanced biochemical reconstitution approaches and cryogenic electron-microscopy (cryo-EM), we have recently determined the first structures of active eukaryotic clamp loader:clamp:DNA complexes, which revealed the molecular basis for the substrate specificities of the yeast RFC and Rad24-RFC complexes and resulted in a significant revision of current clamp loading models. Building on this work, here we propose to extend and advance the approaches established by us for the yeast RFC:PCNA and Rad24-RFC:9-1-1 clamp loader systems to characterize the molecular mechanisms of the yeast and human orthologues of Ctf18-RFC:PCNA, Elg1ATAD5-RFC:PCNA, and RAD17-RFC:9-1-1. The innovative approach leverages the expertise of Dr. Remus' laboratory in the biochemical reconstitution of eukaryotic DNA replication and the expertise of Dr. Hite's laboratory in the characterization of the conformational landscapes of protein complexes by cryo-EM. The proposed studies will determine the mechanistic basis for the functional specialization of CTF18-RFC (Aim 1), uncover the currently unknown mechanism of PCNA unloading by Elg1ATAD5-RFC:PCNA (Aim 2) and visualize the spectrum of conformational states of the human RAD17-RFC:9-1-1 complex to reveal the evolutionary conservation or divergence of the 9-1-1 checkpoint clamp loading mechanism (Aim 3). Collectively, this work will provide novel mechanistic insight into fundamental genome maintenance pathways.

Start date: 01-01-2024

End date: 12-30-2028

Last modified: 06-05-2023

Copyright information:

Structural basis of eukaryotic clamp loading

This project will produce molecular, structural and cell biological data. The data types being generated will include fluorescence, cryo-EM, and phosphorimaging/autoradiography data. The data will be collected on an in-house Titan Krios, PhosphorImager (Typhoon) and epifluorescence microscope (Deltavision Elite, Cytiva). The cryo-EM data will be collected from at least 2 grids per specimen, generating ~500 tB of total data. Raw images will be processed using state-of-the-art software packages that currently include relion and cryosparc. PhosphorImager scans and fluorescence microscopy will produce < 1TB of image files. Quantitative analysis of autoradiography scans will be performed with ImageJ (https://imagej.nih.gov/ij/index.html).

In this proposed project, the raw cryo-EM data, along with processed maps and models will be shared openly by public databases for preservation and reuse. Final files used to generate specific analyses to answer the Specific Aims and related results will also be shared. The rationale for sharing only cleaned data is to foster ease of data reuse.

To facilitate the interpretation and reuse of the data, metadata will be generated and deposited into a repository along with all shared datasets. Protocol DOIs issued from protocols.io.

Cryo-EM data will be collected using SerialEM (PMID: 31086343). Cryo-EM images will be processed with Cryosparc (PMID: 28165473) and Relion (PMID: 34783343). Maps derived from cryo-EM analysis will be processed with Phenix (PMID: 31588918). All cryo-EM software is available via an academic license.

In accordance with FAIR Principles for data, we will use open file formats (e.g. JPEG, MP4, CSV, TXT, PDF, HTML, etc.) and persistent unique identifiers (PIDs) such as RRIDs for resources (e.g., organisms, plasmids, antibodies, cell lines, software tools, and databases) and DOIs for protocols using protocols.io. The bioimaging community has not yet agreed on a single standard data format that is generated by all acquisition systems, but we will use OME-Files for data that will be preserved and shared.

Raw cryo-EM data will be archived at the Electron Microscopy Public Image Archive. Processed cryo-EM maps will be deposited at the Electron Microscopy Data Bank. Atomic models will be deposited at the Protein Data Bank.

The Electron Microscopy Public Image Archive, Electron Microscopy Data Bank and Protein Data Bank all provide metadata, DOI and long-term access. Data will be discoverable online through standard web search of the study-level metadata as well as the persistent pointer from the DOI to the dataset. Persistent Unique Identifiers (PIDs) to improve data findability across all dissemination outputs will include ORCID IDs for people and DOIs for outputs (e.g., datasets, protocols). We will keep our ORCID Records up to date with DOIs for our datasets and publications to increase findability.

All scientific data generated from this project will be made available as soon as possible, and no later than the time of publication or the end of the funding period, whichever comes first. The duration of preservation and sharing of the data will be a minimum of 5 years after the funding period.

There are no anticipated factors or limitations that will affect the access, distribution or reuse of the scientific data generated by the proposal.

Controlled access will not be used. The data that is shared will be shared by unrestricted download.

No human subjects.

Co-PIs Dirk Remus and Richard Hite will be responsible for the day-to-day oversight of lab/team data management activities and data sharing.

Planned Research Outputs

Dataset - "Cryo-EM analysis of CTF18-RFC:PCNA"

Raw images and processed map and models of CTF18-RFC:PCNA complexes in functionally relevant conformations

Dataset - "Cryo-EM analysis of Elg1-RFC:PCNA"

Raw images and processed map and models of Elg1-RFC:PCNA complexes in functionally relevant conformations

Dataset - "Cryo-EM analysis of RAD17-RFC:PCNA"

Raw images and processed map and models of RAD17-RFC:PCNA complexes in functionally relevant conformations

---

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Cryo-EM analysis of CTF18-RFC:PCNA	Dataset	2026-05-30	Open	Worldwide Protein Data Bank Electron Microscopy Public Image Archive The Electron Microscopy Data Bank	100 TB	Creative Commons Attribution 4.0 International	None specified	No	No
Cryo-EM analysis of Elg1-RFC:PCNA	Dataset	2026-05-30	Open	Worldwide Protein Data Bank Electron Microscopy Public Image Archive The Electron Microscopy Data Bank	100 TB	Creative Commons Attribution 4.0 International	None specified	No	No
Cryo-EM analysis of RAD17-RFC:PCNA	Dataset	Unspecified	Open	Worldwide Protein Data Bank Electron Microscopy Public Image Archive The Electron Microscopy Data Bank	100 TB	Creative Commons Attribution 4.0 International	None specified	No	No