

CHANGING SHAPE: AN INVESTIGATION INTO ALLOSTERY AND PROTEIN
CONFORMATIONAL ENSEMBLES

by

Jenna Rose Mattice

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Biochemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

June 2023

©COPYRIGHT

by

Jenna Rose Mattice

2023

All Rights Reserved

ACKNOWLEDGEMENTS

First, I'd like to thank my mentor and advisor Dr. Bothner. For your support and guidance, and belief that I can do hard things. You have helped me grow as a scientist and as a person, and I thank you. I'd also like to thank the Bothner lab past and present: undergrads, grads, post-bacs, and research scientists. What a special and talented group of people. Thank you for your support, help, collaboration, and general shenanigans. Thank you to my committee for the opportunity and guidance. I would also like to thank the wonderful collaborators that I've had the chance to work with and learn from. Science is better when we work together and thank you for working with me.

TABLE OF CONTENTS

1. INTRODUCTION	1
Protein Folding.....	2
Protein Dynamics.....	6
Allostery.....	7
Ligand Binding Alters Protein Energy Landscape	11
Hydrogen Deuterium Exchange Mass Spectrometry to Study Proteins	14
Ion Mobility-Native Mass Spectrometry to Study Proteins.....	17
Summary	21
References.....	22
2. A CATALYTIC DYAD MODULATES CONFORMATIONAL CHANGE IN THE CO ₂ FLAVOBENZYLME 2-KETOPROPYL COENZYME M OXIDOREDUCTASE/CARBOXYLASE	26
Contributions of Authors and Co-Authors.....	26
Manuscript Information	27
Abstract	28
Introduction.....	28
Results and Discussion	30
Addition of 2-KPC Induced a More Compact 2-KPCC Structure.....	30
Limited Proteolysis Revealed Ligand-Dependent Protection of Two Protease Cleavage Sites	30
Global Thermal Stability is Increased by Substrate Binding in 2-KPCC.....	31
Influence of 2-KPCC Catalytic Dyad on Conformation.....	32
Limited Proteolysis Indicates Dyad-Dependent Protection of Proteolytic Sites	32
Active Site Dyad Contributes to Global Thermal Stability	33
Computational Analysis Implicates Catalytic Dyad in Conformational Change	33
Conclusions.....	34
Experimental Procedures	35
Expression and Purification of Native 2-KPCC and Amino Acid-Substituted Variants.....	35
NMS and IM of 2-KPCC and GR.....	36
Limited Proteolysis and Peptide Mapping of 2-KPCC and GR	36
DSF of 2-KPCC and GR.....	36
AlloSigMA analysis of 2-KPCC and GR	36
References.....	37

TABLE OF CONTENTS CONTINUED

3. MESSAGE RECEIVED: LONG RANGE ALLOSTERIC COMMUNICATION IS ESSENTIAL IN A MICROBIAL CARBOXYLASE	39
Contributions of Authors and Co-Authors.....	39
Manuscript Information	41
Abstract	42
Introduction	43
Results	45
E89 is Critical for Mn(II) Coordination.....	45
E89A Exhibited No Acetoacetate Production.....	47
Ligand Binding Alters the Hydrogen Bonding Network of AC	48
Communication Pathway from Substrate Binding to Active Site in WT AC.....	50
E89A Disrupts the Hydrogen-bonding Network of AC	52
Discussion	55
Materials and Methods.....	58
Expression and Purification of Recombinant Acetone Carboxylase	58
EPR Spectroscopy	58
Enzymatic Assays	59
Intact HDX.....	60
Peptide-level HDX.....	61
References.....	63
4. C-TERMINUS INDUCED ASYMMETRY WITHIN A RAD52 HOMODECAMER DICTATES SINGLE-POSITION RAD51 NUCLEATION IN HOMOLOGOUS RECOMBINATION	65
Contributions of Authors and Co-Authors.....	65
Manuscript Information	67
Abstract	68
Introduction	68
Results.....	71
<i>Saccharomyces cerevisiae</i> Rad52 Functions as a Homodecamer.....	71
Cryo-EM Structure of Full-length <i>S. cerevisiae</i> Rad52.....	72
Rad52 Displays Biphasic ssDNA Binding Activity	73
Hydrogen-Deuterium Exchange Mass-Spectrometry (HDX-MS) Reveals Conformational Changes in the C-terminus of Rad52 upon ssDNA Binding	75
Cross-linking Mass Spectrometry (XL-MS) Analysis Reveals Interactions between the Ordered N-terminal and Disordered C-terminal Halves of Rad52.....	78
The C-tail of Rad52 Influences the Dynamics of ssDNA Wrapping.....	80
The C-tail Regulates the Dynamic Movement/Diffusion of ssDNA on Rad52.....	83

TABLE OF CONTENTS CONTINUED

The C-tail Induced Asymmetry Dictates Rad51 Binding to a Single Rad52 Subunit.....	84
Formation of proper Rad51 Filaments Requires the C-tail of Rad52.....	86
Discussion.....	88
Materials and Methods.....	92
Reagents and Chemicals	92
Plasmids for Protein Overproduction.....	92
Purification of Rad52 and Rad52 mutants	92
Determination of Stoichiometry and Molecular Weight using Mass Photometry	94
Analytical Ultracentrifuge (AUC) Analysis	95
Cryo-EM Data Collection	96
Cryo-EM Imaging Processing.....	97
Model Building	98
Measurement of ssDNA binding to Rad52 Using Fluorescence Anisotropy	98
Tryptophan Quenching Assays	99
Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS) Analysis.....	101
Crosslinking Mass Spectrometry (XL-MS) Analysis	102
Cy3/Cy5 FRET Based Steady State ssDNA Wrapping Experiments.....	102
Single-molecule Analysis of ssDNA-Rad52 Wrapping	104
Steady-State Analysis of ssDNA handoff to Rad51	107
Stopped Flow Analysis of DNA Binding and Rad51 Filament Formation	107
References.....	108
 5. RTT105 REGULATES RPA FUNCTION BY CONFIGURATIONALLY STAPLING THE FLEXIBLE DOMAINS.....	115
Contributions of Authors and Co-Authors.....	115
Manuscript Information	117
Abstract	118
Introduction.....	118
Results.....	120
Rtt105 Forms a Stoichiometric Complex with RPA.....	120
Rtt105 Alters the ssDNA Bound Configurations of the DNA Binding Domains (DBDs) of RPA.....	121
Rtt105 Influences the ssDNA Binding Kinetics and Configurations of Individual DBDs of RPA	122
Rtt105-RPA-ssDNA Complexes are Detected on Short DNA Substrates.....	122
Rtt105 Configurationally Staples RPA through Contacts with Multiple Domains in RPA	123
Length of ssDNA and Assembly of Multiple RPA Molecules	

TABLE OF CONTENTS CONTINUED

Promote Release of Rtt105	125
Rtt105 Drives Formation of High-Density RPA-ssDNA Filaments.....	125
Rtt105 Impedes Facilitated Exchange of RPA	125
Complex Formation with Rtt105 Negatively Regulates Interactions With RPA-interacting proteins like Rad52	126
Discussion	127
Methods.....	128
Reagents and Chemicals	128
Plasmids for Protein Overproduction.....	128
Purification of RPA, Rtt105, and RPAIN	128
Generation of Fluorescently-labeled Rtt105	128
Yeast Strains and Genetic Techniques.....	129
Co-Immunoprecipitation.....	129
Secondary Structure Analysis using Circular Dichroism	129
Analysis of Complex Formation using Size Exclusion Chromatography (SEC)	129
Measurement of RPA-ssDNA Interactions using Electrophoretic Mobility Band Shift Analysis (EMSA).....	129
Measurement of RPA and ssDNA Binding in Presence or Absence of Rtt105 using Fluorescence Anisotropy	129
Crosslinking Mass Spectrometry (XL-MS) Analysis	130
Hydrogen-Deuterium Exchange Mass Spectrometry (HDX-MS) Analysis.....	130
MB543 Fluorescence Quenching Assay to Estimate Binding Affinity between RPA and Rtt105	130
Measurements of DNA Binding Kinetics using Stopped Flow Fluorescence	131
Steady State Forster Resonance Energy Transfer (FRET) Measurement of Protein-DNA and Protein-Protein Complexes.....	131
C-Trap Optical Tweezer Analysis of RPA-Rtt105 Interactions	131
Analytic Ultracentrifugation (AUC) Analysis	131
References	132
6. CONCLUDING REMARKS.....	134
Introduction.....	134
Conformational Change and Allostery across Systems	135
Final Remarks on Conformational Ensembles and Allostery	138
References.....	140
APPENDICES.....	142
APPENDIX A: 2-KPCC Supplemental	143
APPENDIX B: Acetone Carboxylase Supplemental.....	147

TABLE OF CONTENTS CONTINUED

APPENDIX C: Rad52 Supplemental.....	151
APPENDIX D: RPA and Rtt105 Supplemental	174
REFERENCES CITED.....	206

LIST OF TABLES

Table	Page
2.1 Rate of ATP Consumption and Rates of Acetoacetate, ADP, and AMP Production are Shown for the WT, E89A, E89D Variant With and Without Acetone	48
5.1 <i>S. cerevisiae</i> strains used in this study	126

LIST OF FIGURES

Figure	Page
1.1 Protein Folding Funnel	4
1.2 Rugged Bottom of the Protein Folding Funnel	5
1.3 Timescale of Protein Motion	6
1.4 Three Schematic Models of Protein Allostery	8
1.5 External Effectors Alterations to Conformational Ensemble	11
1.6 Substrate Binding Models	12
1.7 Overview of HDX Reaction	14
1.8 Overview of IMS-NMS	17
1.9 Schematic of Waters Synapt G2-Si	18
2.1 Reactions of Glutathione Reductase and 2-KPCC Showing the Active Site Residues and Intermediates	28
2.2 Structural Comparison of 2-KPCC and Glutathione Reductase	28
2.3 Mobiligrams and Charge State Distributions of WT, Mutant 2-KPCC, and Glutathione Reductase	29
2.4 Limited Proteolysis of WT 2-KPCC	30
2.5 Differential Scanning Fluorimetry Curves of Glutathione Reductase and WT 2-KPCC	31
2.6 Limited Proteolysis of 2-KPCC Variants	32
2.7 Differential Scanning Fluorimetry Curves of 2-KPCC Variants	33
2.8 Change in Free Energy Calculated by AlloSigMA as a Result of Mutations in the Catalytic Dyad Plotted on Structure of 2-KPCC	33
2.9 Conformational Ensembles of GR and 2-KPCC	34

LIST OF FIGURES CONTINUED

Figure	Page
3.1 Overview of Acetone Carboxylase Mechanism and Structure	43
3.2 EPR Reveals No Change in Mn(II) Signal in E89A when AMP is Bound	45
3.3 Intact Hydrogen-Deuterium Exchange Highlights Influence Ligand Binding has on Hydrogen Bonding Network in WT	48
3.4 Peptide-level HDX Exchange shows Specific Changes upon Ligand Binding	50
3.5 Intact Hydrogen-Deuterium Exchange Highlights Influence on Hydrogen Bonding Network Ligand Binding and Mutation has on E89A	52
3.6 Peptide-level HDX Exchange Highlight Exact Areas of Change in Deuterium Uptake due to Ligand Binding in WT	53
3.7 Overview of AC Structural Changed during Reaction for WT and E89A	55
4.1 <i>S. cerevisiae</i> Rad52 functions as a homodecamer	71
4.2 Rad52 binds to ssDNA using multiple binding sites	73
4.3 Hydrogen-Deuterium Exchange Mass-Spectrometry Reveals ssDNA Driven Changes in both the N- and C-terminal Halves of Rad52	76
4.4 Cross Linking Mass-Spectrometry (XL-MS) Reveals Interactions between the N- and C-terminal Halves of Rad52	78
4.5 Interactions between N- and C-terminal halves of Rad52 Influences ssDNA Binding Interactions	80
4.6 The C-terminus Controls the Intrinsic Diffusion of Rad52 on ssDNA	82
4.7 The C-tail of Rad52 Regulates Allosteric Interactions with Rad51	84
4.8 Model Describing the Mechanism of Action of Rad52 and Regulation by the C-tail	85
5.1 Rtt105 Forms a Stoichiometric Complex with RPA.....	115

LIST OF FIGURES CONTINUED

Figure	Page
5.2 Rtt105 Alters the Configuration of DBD-A on ssDNA and Kinetics of DBD-D Binding	116
5.3 Rtt105 Forms a Higher Order Complex with RPA Bound to Short ssDNA Substrates.....	117
5.4 Rtt105 Configurationally Staples RPA through Multiple Contacts with DBDs and PIDs.....	119
5.5 Binding of Multiple RPA Molecules on Longer ssDNA Substrates Triggers the Release of Rtt105 and Formation of a High-Density RPA Nucleoprotein Filament.....	120
5.6 Rtt105 Inhibits Facilitated Exchange Activity of RPA.....	122
5.7 Rtt105 Inhibits Spurious RPA-Rad52 Interactions in the Absence of ssDNA.....	123
6.1 Summary of the Findings Contained in this Thesis	132

ABSTRACT

Allostery is the mechanism by which action at one site on a protein causes a functional change at a distant site. An allosteric change can manifest as conformational change or a change in protein dynamics. In this way, the study of allostery, protein dynamics, and structural biology are individual, yet related fields. Progress and technical advancements in one field inform and drive the others. In this thesis, four protein complexes 2-ketopropyl-coenzyme M oxidoreductase/carboxylase (2-KPCC), acetone carboxylase (AC), Replication protein A (RPA), and Radiation sensitive 52 (Rad52) were studied to elucidate conformational change and allostery during catalysis. A variety of orthogonal biophysical approaches were used to study these systems. To infer changes in protein dynamics and conformation, hydrogen-deuterium exchange coupled to mass spectrometry was used in three of these studies. This technique allows for the probing of the hydrogen bonding network based on ligand binding or mutation. Probing of AC, RPA and Rad52 has led to the description of conformational changes essential for function. Ion mobility coupled to native mass spectrometry was used to investigate the available conformations of 2-KPCC during catalysis and led to the discovery of residues essential for modulating those conformations. The concepts of allostery, conformational ensemble, and protein dynamics have evolved since they were first described. Utilizing mass spectrometry-based techniques, my work helped expand the knowledge of several protein systems that contain allosteric networks which are necessary for function. The studies presented in this thesis increase the understanding not only for these protein systems, but also of protein function on a deeper level.

CHAPTER ONE

INTRODUCTION

The study of allostery, protein dynamics, and structural biology developed hand in hand, with advancements in one field informing and driving the others. The first structural model of a protein was solved for myoglobin in 1958, followed shortly thereafter by hemoglobin in 1960.^{1,2} While these first crystal structures were milestone accomplishments, pushing the limits of biological insights, they also revealed new and exciting questions. In a landmark 1965 paper Monod, Changeux, and Jacob describe allostery in their introduction:³

Ever since the haem-haem interactions of haemoglobin were first observed, this remarkable phenomenon has excited much interest, both because of its physiological significance and because of the challenge which its physical interpretation offered. The elucidation of the structure of haemoglobin has, if anything, made this problem more challenging, since it has revealed that the haems lie far apart from one another in the molecule.

Since this statement was made more than 60 years ago, the study of protein structure, dynamics, and allostery have each grown into their own fields, yet continue to challenge scientists.

Structural techniques (X-ray crystallography, cryo-EM), biophysical techniques (mass spectrometry, nuclear magnetic resonance, fluorescence) and the advancement of computational techniques (molecular dynamics simulations, structure prediction) has led to a much deeper understanding of protein systems. However, numerous nuanced and fundamental questions remain. For example: Is every protein allosteric?⁴ Would proteins at large benefit from being allosteric? Can we predict amino acid residues which are essential for allosteric communication? This thesis aims to expand the study of protein conformational ensembles and allosteric networks

by focusing on four unique systems and how structure, conformational ensembles, and allostery facilitate catalysis.

Protein Folding

The primary functional roles of proteins are structural, signaling, regulation, energy production, and the synthesis and degradation of biomolecules. The functions of proteins are dependent on their structures which are encoded in their amino acid sequence. The amino acid sequence of proteins is composed of twenty common amino acids which differ in physical and chemical properties. In most cases, the amino acid sequence contains all the information needed for the protein to fold into its native structure. Multiple amino acid sequences can adopt a given fold, which is known as sequence capacity.⁵ Uniprot, a database of protein sequences and functions, contains over 70 million protein sequences with more being added as sequencing of microbial genomes continues.⁵ How do proteins fold into their native structure based solely on the amino acid sequence? What forces define their native fold? How do they fold so fast? These are a few of the questions scientists have asked and that have been explored at length.

Before discussing how proteins fold, it is necessary to define the forces on the amino acid sequence which will define the native structure. These forces are primarily van der Waals forces, electrostatic interactions, hydrophobic interactions, and hydrogen bonding.^{6,7} In addition, the backbone of a protein can only adopt certain angles because of steric hindrances resulting from sidechain clashes with the carbonyl group. Van der Waals forces are very short-range attractive or repulsive forces observed in neutral atoms. Electrostatic interactions are the attraction or repulsion based on positive or negative charge. Hydrophobic interactions arise from the tendency of water to exclude nonpolar molecules. For this reason, protein cores contain predominantly

amino acids with nonpolar sidechains (hydrophobic) and the exterior or the surface of the protein is made up of amino acids with predominantly polar sidechains (hydrophilic). Hydrogen bonds between the sidechains and protein backbone create a network that helps stabilize the three-dimensional structure. Together these forces lead to the native structure of a protein, but the details of this process have remained an open question.

Levinthal proposed protein folding to be a random search problem, where all conformations of the peptide chain, except for the native state, are equally probable.⁸ It is estimated that before folding, each residue in a protein can potentially take on 10 different conformations.⁹ If a peptide chain of 60 amino acids were to sample all possible conformations, it would need to sample 10^{60} conformations, and it would take years to adopt the native fold of the protein. Proteins can fold on a microsecond to seconds timescale, so logic dictates that protein folding is not a random search problem.^{10,11} The energy landscape theory accounts for the discrepancy in timescale. In the energy landscape theory, proteins fold as they travel down a funnel-shaped energy landscape (Figure 1).¹² The funnel is wider at the top where more conformational entropy and higher energy conformations are present. As the protein folds, it loses entropy, however, this is compensated for by a greater increase in the entropy of the solvent (water). The hydrophobic effect is based on the tendency of water to exclude nonpolar molecules. In most cases, this is the major force driving protein folding. As the protein travels down the funnel, the conformational landscape narrows and the protein eventually adopts its native fold.⁹⁻¹⁴ The energy landscapes can be rugged and have many energy maxima and minima. Meaning there are many potential pathways for a protein to fold into its native state. Further, the bottom of the energy landscape can have multiple energy minima (valleys) separated

by low barriers this leads to native conformational microstates where sample frequency is based on the energy barriers between them.⁶

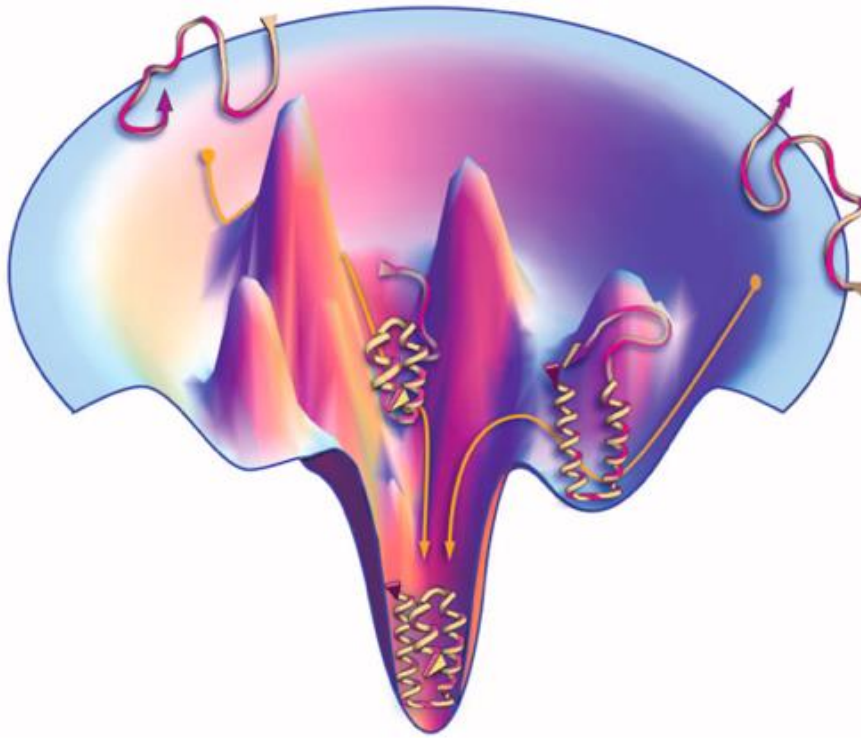


Figure 1. Protein folding funnel depicting unfolded proteins at the top of the funnel. As the proteins travel down the energy landscape, they fold while exhibiting intermediate states during energy maxima on their way to the native fold. Image taken from Dill *et al* ,2012 *Science*.

The amplitude of conformational change and timescale between the conformational microstates can be defined by the bottom of the energy landscape. Conformational microstates can be divided into three tiers: tier-0, tier-1, and tier-2 (Figure 2). Where tier-0 is defined by conformational microstates that are separated by energy barriers that are more than one kT , (Boltzmann's constant and absolute temperature) or greater than 4.11×10^{-24} kJ.^{6,15} Tier-0 occurs on a relatively slow timescale (μ s-ms) and conformational transitions in this tier are not frequent

because of their high energy barrier. Transitions that do occur are generally large, resulting in changes throughout the whole protein complex. The tier-1 and tier-2 microstates occur on fast timescales (ps-ns) and inside the wells of the tier-0 microstates. The energy barriers separating these conformations are generally less than one kT (less than 4.11×10^{-24} kJ) and result in conformational transitions that occur with greater frequency and smaller changes to the protein structure.^{6,15} The conversion between conformational microstates within an energy landscape manifest as protein dynamics.

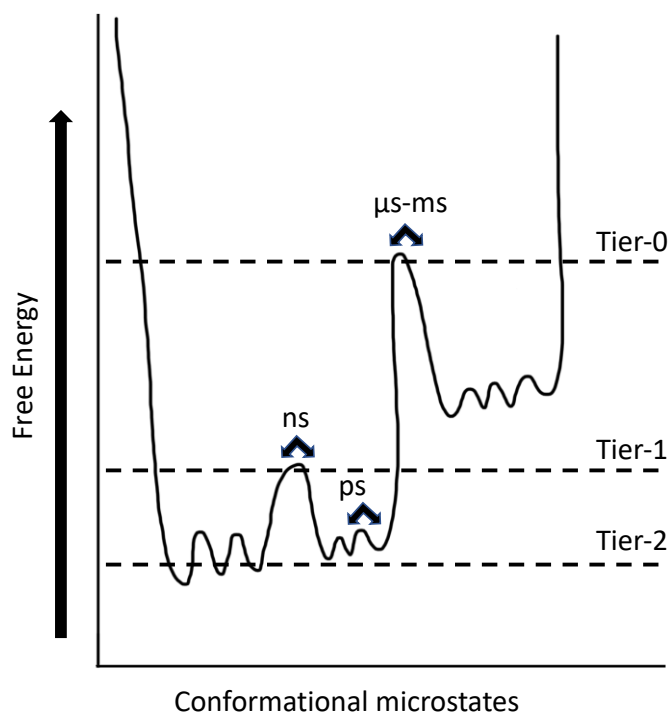


Figure 2. Rugged bottom of the protein folding funnel depicting tier-0, tier-1, and tier-2 conformational microstates. tier-0 occurs on slow timescales, while tier-1 and tier-2 occur on fast timescales. Tier-1 and tier-2 microstates occur inside of tier-0 microstates. Figure adapted from Henzler-Wildman *et al*, 2007 *Nature* and Liu *et al*, 2012 *Protein Eng*.

Protein Dynamics

Interconversion between conformational microstates leads to protein motion. Protein motion occurs over a wide range of timescales (fs-minutes)(Figure 3).^{15,16} Also, the scale of these motions occurs over a broad range, from bond vibration (fs) to local unfolding (ms). Smaller motions such as bond vibrations happen at fast timescales, while larger (global) motions may occur on longer timescales. My work has focused on conformational changes that are associated with the millisecond to minutes timescale, which is well-suited to analysis by mass spectrometry-based approaches.¹⁶

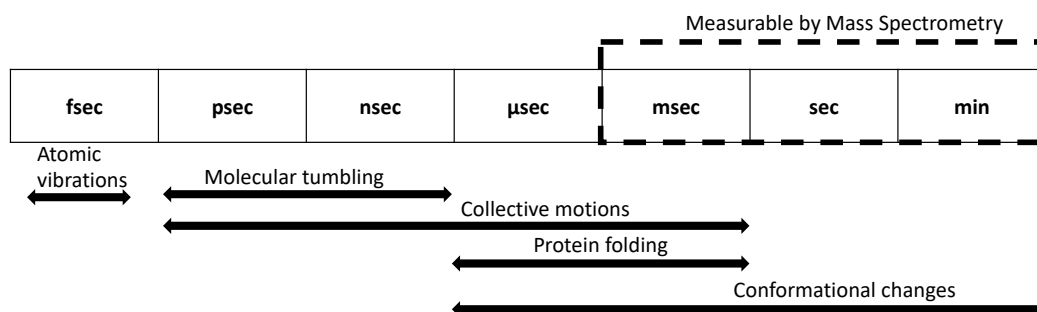


Figure 3. Timescale of protein motion. Conformational change occurs on the μ sec to minutes timescale. Mass spectrometry is ideally suited to measure msec to minutes of protein motion. Figure adapted from Ben-Nissan *et al*, 2011 *Chem Soc Rev*.

Protein dynamics describes a wide variety of motion.¹³ At one end of the continuum is the rigid body motion of the subunits in hemoglobin, where the subunits reorient with a pivot motion, while the fold is relatively static.¹³ Slightly more dynamic proteins can have side chains that participate in local protein motions. When backbone dynamics become involved, larger scale conformational changes can occur. In the case of a local unfolding event in a protein, where

stretches of backbone unfold, resulting in a flexible disordered region. At the most dynamic end of the continuum are intrinsically disordered proteins (IDPs) which contain one or more stretches of amino acids which lack a defined structure and have extreme conformational flexibility.^{13,17,18} IDPs have the freedom to sample a wide variety of conformations, making them highly dynamic. Conformational freedom is advantageous in some protein systems because it will allow the same sequence of amino acids to interact with high specificity with different folded domains. Proteins containing IDPs are often involved in signaling and the assembly of protein complexes. When interacting with a ligand or target domain, IDPs will often undergo a disordered to ordered transition.¹⁸

Allostery

Allostery as a means to describe protein function has gone so far as to be labeled the ‘second secret of life’.¹⁹ Allostery’s importance in biology has been known for over 50 years.²⁰ Allosteric regulation is essential for understanding a multitude of important processes, including, signal transduction, molecular machine function, transcriptional regulation, and metabolism. However, a clear molecular understanding of allostery is lacking. The definition of allostery that I find the most informative is: allostery is ligand binding, or other energetic perturbation, that causes a change in protein structure or dynamics at a distal site.¹⁹ Many models and studies have been put forth to study allostery as we continue to elucidate the molecular understanding of allostery.

The first observation of allostery was made by Bohr in 1903. The Bohr effect is the decreased affinity of oxygen binding to hemoglobin in response to higher concentrations of carbon dioxide and lowered blood pH.^{21,22} Bohr and contemporaries also reported the sigmoidal

nature of the oxygen binding curves, which are hallmarks of cooperativity. Ligand or effector binding increasing or decreasing the probability of binding in other subunits is called cooperativity. Half a century after these observations, the concept of allostery would be used to describe hemoglobin. The first model put forth to describe allostery was MWC (Monod-Wyman-Changeux), also known as the concerted or symmetry model (Figure 4).^{3,23} The MWC model describes positive cooperativity in allostery. In this model the protein complex is composed of multiple binding sites that can only be in two states, the tense (T) low binding affinity state, or the relaxed (R) high affinity binding state. These two states exist in an equilibrium and all protein subunits or binding sites remain in the same conformation. When an allosteric effector binds to either state the equilibrium is then shifted to favor the bound state. The MWC model fails to fully describe allostery because it does not allow for the multiple binding sites in a protein complex to be in separate conformations. Additionally, this model does not allow for negative cooperativity, where ligand binding decreases likelihood of subsequent binding.

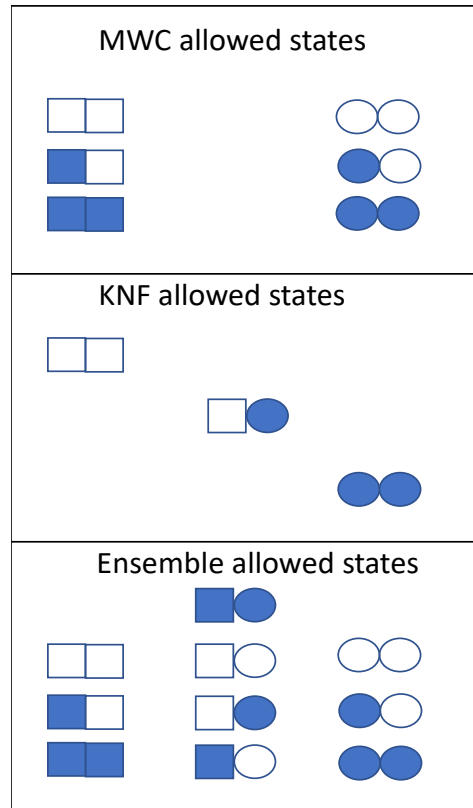


Figure 4. Allowed conformational states for each allosteric model in the case of a homodimeric protein. Squares represent low affinity states and circles represent high affinity. Blue shapes represent ligands bound while white shapes represent unbound forms of the protein. Figure adapted from Hisler *et al*, 1996 *J Mol Bio*.

The KNF (Koshland-Nemethy-Filmer) or sequential model, is another model of allostery (Figure 4).^{21,22,24} Like the MWC model, the KNF model assumes two states, high and low affinity. However, it allows each subunit in the protein complex to undergo conformational change independently. Each binding event will increase or decrease the likelihood of binding in the subsequent subunit, but the unbound subunit will not undergo conformational change until bound. The KNF model better describes cooperativity in allostery but still falls short of fully explaining the complexity of these processes. Although neither model accurately explains all examples of allostery, both the KNF and MWC model have been useful in providing insight into

the relationship between protein structure and function. One insight gained from these models is that the structural conformations available to a protein cannot be categorized into a two-state model.

The concept of allostery raises an important point; proteins do not exist as static molecules, but instead adopt an ensemble of conformations that are in equilibrium. The ensemble allosteric model (EAM) was developed to account for conformational ensemble flexibility (Figure 4).²⁵⁻²⁷ In the EAM model there is less importance placed on the adopted structural changes during allostery but instead on energetic changes occurring to energy landscape of the protein when an allosteric effector is present. Allostery was initially thought to induce conformational change that resulted in a distinctive shape, but that is not always the case, and some proteins undergo subtler conformational changes. Allostery can also be present in a protein even in the apparent absence of conformational change.^{4,13,14,22,28,29} Changes can be characterized by alterations in entropy when an allosteric effector is bound. In this case the binding of the effector does not change the average conformation but instead the conformational ensemble is altered.^{28,29} One common hypothesis is that proteins contain a conserved allosteric pathway that is responsible for transmitting the allosteric signal.²⁵ In other words, the conserved allosteric pathway is how communication occurs between distal sites. For a signal to be sent between sites, effector binding must alter the entropy or enthalpy, and this must then be converted into a mechanical signal. The mechanical signal can then transmit changes in protein dynamics as well as conformation. By tracking the changes in dynamics and conformation, we can probe the allosteric network within the protein. A variety of techniques can be utilized to provide information on the allosteric network of a protein, in this thesis I focus primarily on hydrogen-

deuterium exchange coupled to mass spectrometry and ion mobility coupled to mass spectrometry, discussed more below.

Ligand binding alters protein energy landscape

The free energy and entropy a protein exhibits are directly related to conditions such as pH, salt concentration, and temperature. Changes to these conditions will lead to alterations in the energy landscape of a protein (Figure 5).^{24,26,30} Factors such as point mutations and ligand binding can also alter the energy landscape and the conformational ensemble of the protein. Three scenarios have been proposed for how energy landscapes would be altered by an external effector.^{13,22} The first occurs when one conformation of the protein is stabilized. In this scenario, the energy landscape is redistributed to create a single “deeper” energy leading to a favored conformation. The second case is when the energy landscape is altered such that a new conformation becomes accessible. The last scenario occurs when the external effector leads to more available conformations by remodeling the energy barriers between conformations, essentially changing the roughness of the landscape. If this leads to lower energy barriers, the protein will sample more conformations in a given time period. My research has focused on, ligand binding and point mutation, as effectors of protein conformational ensembles.

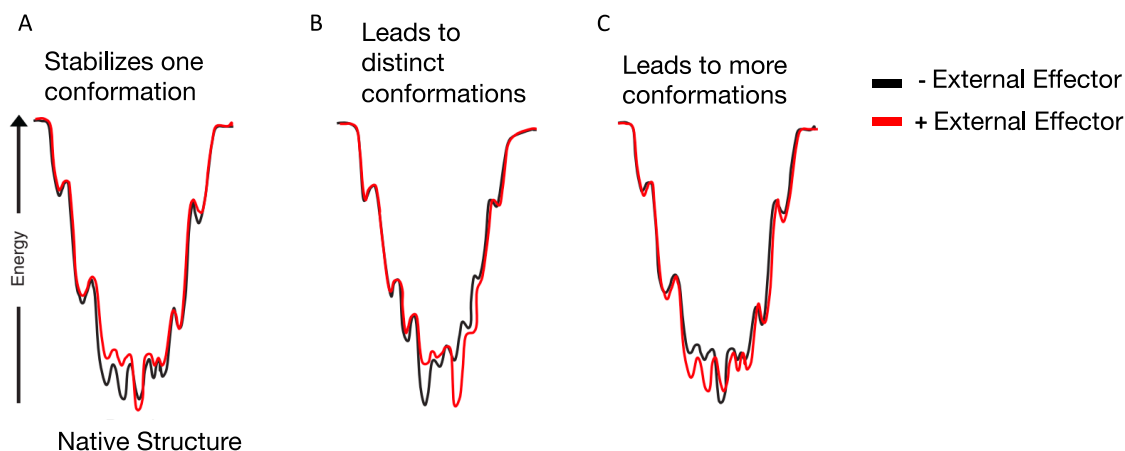
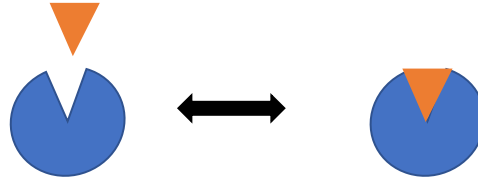


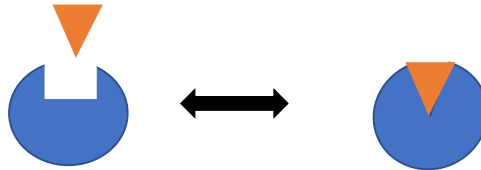
Figure 5. External effectors alterations to the protein conformational ensemble. A) The binding of an effector stabilizes one conformation. B) The binding of an effector leads to distinct conformations compared to the native structure. C) The binding of an effector leads to more available conformations. Figure adapted from Goodey *et al*, 2008 *Nat Chem Biol*.

The description of protein-ligand interactions has changed since the concept was first proposed by Emil Fischer in 1894 (Figure 6).^{30,31} One of the currently favored models, called the conformational selection model, considers the ensemble of conformations proteins can adopt. In the conformational selection model, ligand binding shifts the conformational population and redistributes the conformational states and energy landscape in the ways described in the previous paragraph.^{13,22,25,26,28} Historically the previous models of ligand binding fell short of the complexity of the systems. The first model of ligand binding was lock and key, where enzyme and substrate exist in complimentary shapes to one another.^{24,26,32} However, this could not explain enzyme promiscuity or affinities for different shaped ligands. To address this the induced fit model was conceived. In this model, ligand binding will induce a conformational change in the protein. But this model does not allow for multiple conformations to exist.^{24,26,32} The lock and key and induced fit models also informed how scientists first described allostery.

A) Lock and key model



B) Induced fit model



C) Conformational selection model

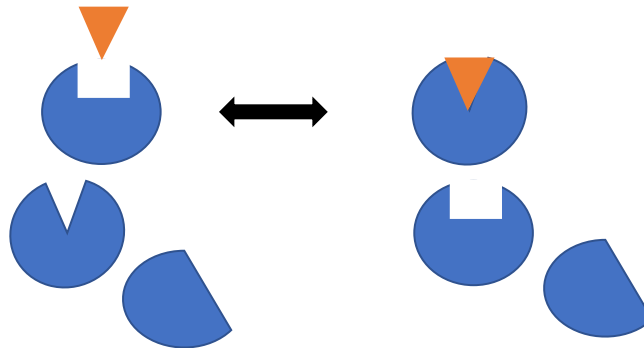


Figure 6. Substrate binding models A) Lock and key model, where substrate and protein active site are complimentary shapes. B) Induced fit model, substrate induces a change in protein active site shape to allow binding. C) Conformational selection model, conformations of the protein exist in equilibrium. Presence of substrate shifts equilibrium to populations which bind substrate.

Hydrogen deuterium exchange mass spectrometry to study proteins

As described above, there are several noncovalent forces that work in concert to stabilize the three-dimensional shape of a protein. One of the most important with respect to conformational plasticity is hydrogen bonding. Hydrogen bonds are formed when a hydrogen covalently bound to an electronegative donor is shared with an electronegative acceptor. Proteins contain many hydrogen bond donors and acceptors, such as the amide and carbonyl groups of the peptide backbone and polar functional groups on the sidechains of amino acids.³³ Protein hydrogen bonds are generally made whenever two criteria are satisfied, the donor and acceptor atoms are within 3.5 Å of each other, and the angle between them is greater than 90°. ³⁴ The energy of the hydrogen bond is dependent on the length of the bond and is normally 20-25 kJ mol⁻¹.³⁴

Hydrogen bonding in a protein can be thought of as a cooperative network that plays an important role in the overall stability and structure of protein. Removing even one hydrogen bond from the protein will weaken the hydrogen bonding network. Probing hydrogen bonding networks can therefore reveal changes in overall stability, which in turn can be used to infer changes in dynamics or conformation of a protein. One technique used to investigate hydrogen bonding networks is hydrogen-deuterium exchange coupled to mass spectrometry (HDX-MS). Hydrogen atoms in O-H, N-H, C-H and S-H found in amino acids are regularly exchanged with surrounding water molecules.^{33,35,36} Exposing a protein to a D₂O-containing buffer, the hydrogens will be replaced with deuterium. Sidechain O-H, C-H, and S-H will exchange immediately, while the N-H in the peptide backbone will exchange on the milliseconds to hours timescale. The incorporation of deuterium into the peptide backbone can be tracked using mass spectrometry, which will measure the net increase of one atomic mass unit (amu) per deuterium

exchanged (Figure 7). Only exchange on the amide backbone hydrogens will be measured by mass spectrometry because of the slower timescale of exchange.

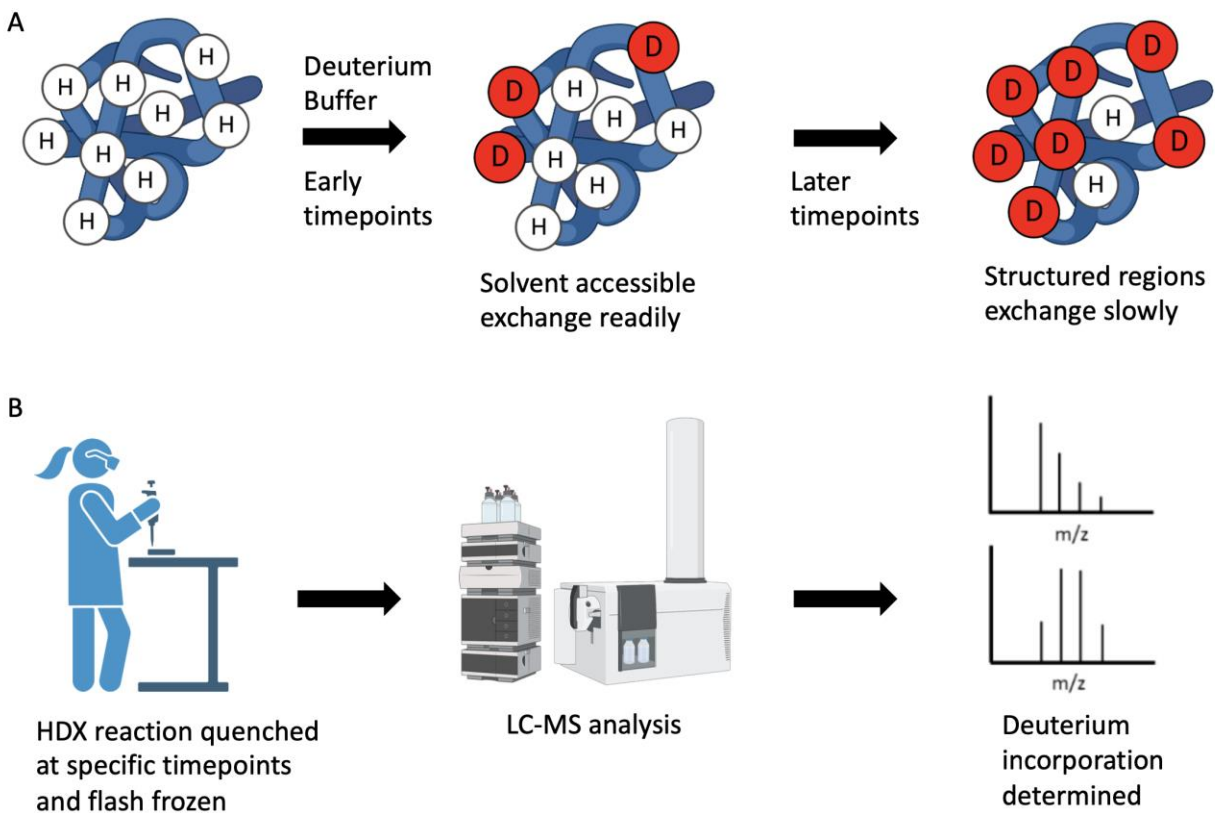


Figure 7. Overview of HDX reaction. A) Principles of HDX reaction. Protein is exposed to deuterium buffer and solvent accessible regions of the protein exchange quickly while structured regions exchange slowly. B) Benchtop HDX reaction is initiated by diluting protein 1:10 into deuterium buffer. The reaction is quenched by lowering the pH and temperature which slows the exchange of hydrogen/deuterium and a pepsin digestion is performed before flash freezing with liquid nitrogen. The peptide mixture is then stored in liquid nitrogen until LC-MS analysis. LC-MS allows for the determination of how much deuterium is incorporated per peptide through the analysis of the isotopic distributions for each peptide. Figure adapted from Tokimina-Lukaszewska *et al*, 2018 *Front Microbiol*.

The wide range in exchange time for amide backbone hydrogens occurs because of the structured nature of proteins. Amide hydrogens that are solvent accessible and not participating

in hydrogen bonding will exchange rapidly. Alternatively, amide hydrogens that are participating in hydrogen bonding will exchange more slowly. The ability of the amide hydrogens participating in hydrogen bonding to exchange is related to the local fluctuations in the hydrogen bonding network. Local hydrogen bonding fluctuations have been termed “closed” to signify the amide hydrogen is participating in a hydrogen bond, and “open” to signify the amide hydrogen is no longer participating in a hydrogen bond.³³ The rate of conversion between the “open” and “closed” states is $k(\text{open})$ and $k(\text{closed})$, respectively. If the rate of deuterium exchange $k(\text{exchange})$ is equal to $k(\text{open})$, this is referred to as EX1 kinetics and only one opening event occurred before exchange.³³ If multiple opening and closing events occurred before exchange this is referred to as EX2 kinetics. EX2 kinetics are commonly used in probing structured proteins hydrogen bonding networks.

Measuring deuterium incorporation over time enables the assessment of the stability of the hydrogen bonding network. The more often local fluctuations occur, the faster the hydrogens will be exchanged for deuterium. If the hydrogen bonding network is more stable less deuterium incorporation will occur. Comparing different conditions, like, unbound and ligand bound proteins, and point mutations, alterations in the hydrogen bonding network can be assessed. Measuring deuterium exchange in an intact protein provides information on the global stability of the hydrogen bonding network of the protein. By comparing the deuterium uptake curves under different conditions, we can ascertain the global disruptions in the hydrogen bonding network and infer disruptions in the stability, solvent accessibility, and dynamics or conformational change of a protein. This is explored in the third chapter in this thesis.

By adding a proteolytic digestion step to the hydrogen/deuterium exchange experiment, the number of deuterium incorporated into a specific peptide can be measured. This locates the exchange to the specific region of the peptide. While intact HDX-MS elucidates global, overall changes in deuterium exchange, peptide-level HDX-MS illuminate local disruptions in the hydrogen-bonding network and provides a more detailed picture of regions that experience exposure or protection from exchange.³⁷ Additionally, by examining enzymes poised with substrate or variants with amino acid substitutions, which can cause local disruptions in stability, we can infer changes in solvent accessibility and/or dynamics/conformational change of key regions of protein that are important for catalysis. Peptide-level HDX-MS is used to examine the catalytic cycle in the third, fourth, and fifth chapters.

Ion-mobility- native mass spectrometry to study proteins

In native mass spectrometry (NMS), when measurements are conducted under non-denaturing conditions, structure is maintained during the desolvation and ionization process of a protein or protein complex.^{38,39} Observing proteins under native-like conditions using mass spectrometry has been well documented and protocols for keeping proteins in their folded structural state include appropriate salt, pH, and soft ionization techniques when transitioning to the gas phase. From employing these protocols information about protein conformation can be obtained because most non-covalent interactions are maintained, permitting secondary and tertiary structure to be preserved. Mass spectrometry can distinguish between native and denatured conformers of a protein because denaturing leads to enhanced ionization and increases the number of sites for available protonation.

Therefore, charge state distributions can be used to interrogate the conformations of a protein. A more extended protein conformation will result in a higher average charge on the protein and a broader charge state distribution; therefore, comparing the charge state distribution of a protein under varying conditions can aid in determining a protein's fold (Figure 8A & 8B).^{38,40,41} Ligand binding or mutation which has an effect on the fold of the protein will produce a change in the charge state distribution. By comparing the unbound/WT protein with the ligand bound/ mutant protein it is possible to determine if conformation has changed.

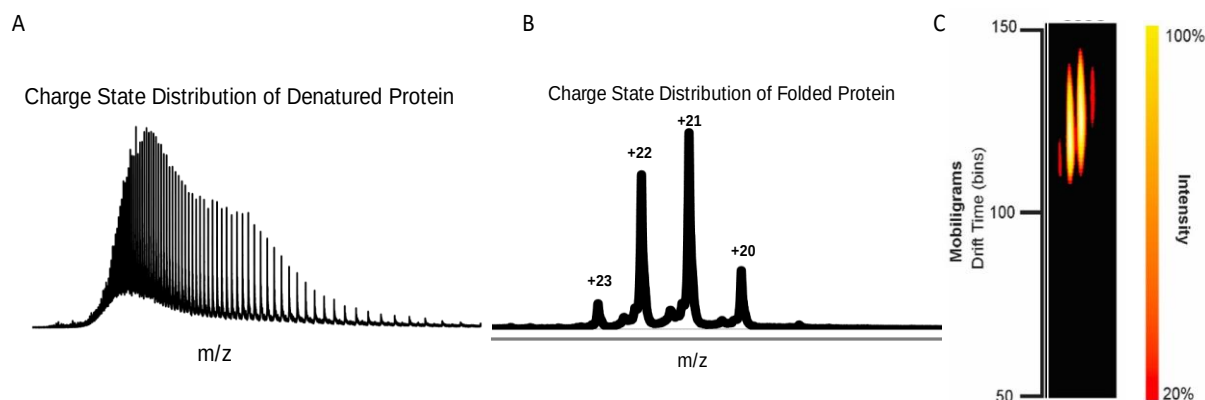


Figure 8. Representative charge state distributions and mobiligram. A) Charge state distribution of a denatured protein highlighting the plethora of charges available. B) Charge state distribution of a folded protein highlighting the limited number of charges available. C) Mobiligram of a folded protein.

NMS can be coupled to ion mobility spectroscopy (IM). IM measures the time it takes an ion to travel a drift cell filled with an inert gas, where larger ions will collide more often with the collision gas and take more time to traverse the cell. Here at MSU the mass spectrometry facility has a traveling wave ion mobility spectrometry (TWIMS) instrument, a variation of IMS. The drift cell of the TWIMS instrument contains a T-Wave mobility separator which produces an oscillating electric field through stacked ion ring electrodes (Figure 9). In the TWIMS instrument, the maximum RF (radio frequency) applied is 250V peak-to-peak with a typical

frequency of 2.8 MHz which will produce a maximum wave amplitude of 40V.⁴²⁻⁴⁴ These can be adjusted to ensure proteins are remaining intact throughout the measurement. Further, the drift gas pressure (mbar) is kept at an optimum to ensure proteins are being driven through the cell but low enough to not cause fragmentation or denaturing of the protein. The time it takes for an ion to traverse the T-Wave drift cell is related to its mobility.⁴³ Ions with lower mobility will roll over the voltage waves multiple times and take longer to traverse the cell, while ions with higher mobility will ride the wave and have faster traverse times. Mobility is related to the size, shape, and charge of the molecule (Figure 8C). IM-NMS measurements can also provide the collisional cross section of a protein (CCS or Ω). The Mason-Schamp equation is used to calculate CCS (Equation 1).³⁸ In this equation the charge of electrons (e) and ions (z), buffer gas density (N_0), mass of collision partners (μ), Boltzman's constant (k_b) and, drift region temperature (T) are all variables related to CCS. Also, the analyte mobility (K_0), is considered, which is measured as a function of temperature and pressure. Collisional cross-section relates the 3D structure and mobility of the protein under specific instrument parameters so they can be better identified and compared.⁴⁵ TWIMS wave instruments however cannot measure CCS directly and standards of known CCS must be used for calibration.

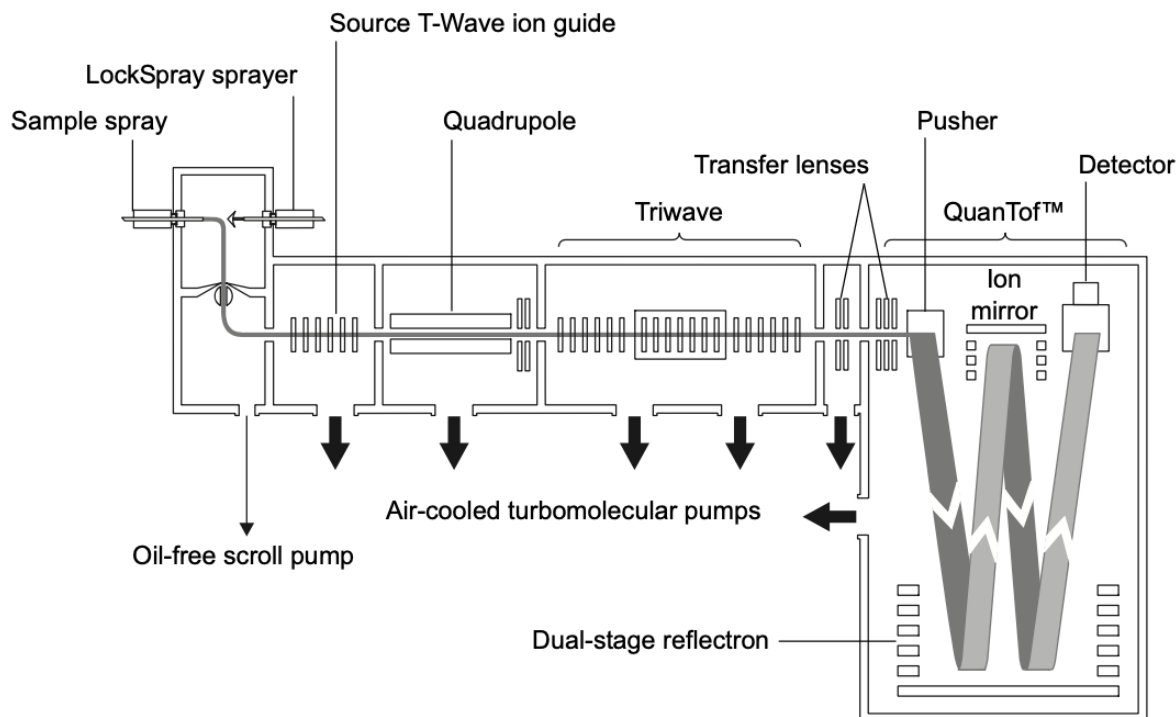


Figure 9. Schematic of Waters Synapt G2-Si. The sample will travel from the sample inlet through the ion guide, quadrupole mass analyzer, Tri-wave cell, QuanToF, and finally to the detector. The Triwave cell contains stacked ion ring electrodes which will produce oscillating electric field waves. Figure from Synapt G2-Si Manual, Waters Corp 2009.⁴²

$$\Omega = \frac{\frac{3}{16} \left(\frac{2\pi}{\mu k_b T} \right)^{\frac{1}{2}} ze}{N_0 K_0}$$

Equation 1. Mason-Schamp equation for calculating CCS. The variables of this equation are: e , charge of an electron; z , ion charge; N_0 , buffer gas density, K_0 , analyte mobility; μ , reduced mass of the collision partners; k_b , Boltzmann's constant; and T , the drift region temperature. Equation take from Dodds *et al* 2019 *JASMS*.

Comparisons of collisional cross-sections between proteins poised in different states (ligands bound or mutation) allow for the probing of protein conformations and their relative populations. IM-NMS technique is highly sensitive and can detect changes in collisional cross-sections small

enough to distinguish between isomeric peptides. IM-NMS was utilized to probe conformations of the protein 2-KPCC in chapter two of this thesis.

Summary

In this thesis I will explore the effects of ligand binding and mutation on conformational ensemble and allostery in four different systems. In the first two manuscript chapters I explore the microbial carboxylases 2-ketopropyl-coenzyme M oxidoreductase/carboxylase (2-KPCC) and acetone carboxylase (AC). Both enzymes undergo conformational change to produce acetoacetate. In both studies the effect of ligand binding and mutations were explored to understand the conformational ensembles essential for catalysis. The last two manuscript chapters of this thesis focus on key proteins in DNA metabolism and their conformational fluctuations. Replication protein A (RPA) is a heterotrimeric protein which recruits and binds over three dozen proteins during DNA metabolism. The key proteins in the DNA metabolism chapters of this thesis are RPA and its interactions with single-stranded DNA (ssDNA) and regulator of Ty1 transposition 105 (Rtt105), and Radiation sensitive 52 (Rad52) which are essential for a variety of aspects in DNA metabolism. Utilizing mass spectrometry-based techniques, my work helped expand the knowledge of several protein systems which undergo a change in conformation or dynamics and have implications in carbon and DNA metabolism.

References

1. Kendrew J, Bodo G, Dintzis H, Parrish R, Wyckoff H. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature*. 1958;181.
2. Perutz M, Rossmann M, Cullis AF, Muirhead H, Will G, North ACT. Structure of hoemoglobin. *Nature*. 1960;185:416-422.
3. Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: A plausible model. *J Mol Biol*. 1965;12(1):88-118. doi:10.1016/S0022-2836(65)80285-6
4. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? *Proteins Struct Funct Genet*. 2004;57(3):433-443. doi:10.1002/prot.20232
5. Tian P, Best RB. How Many Protein Sequences Fold to a Given Structure? A Coevolutionary Analysis. *Biophys J*. 2017;113(8):1719-1730. doi:10.1016/j.bpj.2017.08.039
6. Liu S-Q, Ji X-L, Tao Y, Tan D-Y, Zhang K-Q, Fu Y-X. Protein Folding, Binding and Energy Landscape: A Synthesis. *Protein Eng*. 2012. doi:10.5772/30440
7. Anfinsen. Folding of Protein Chains. *Science (80-)*. 1973;181(4096):223-230.
8. Karplus M. The Levinthal paradox: Yesterday and today. *Fold Des*. 1997;2(4):69-75. doi:10.1016/S1359-0278(97)00067-9
9. Wolynes PG, Onuchic JN, Thirumalai D. Navigating the folding routes. *Science (80-)*. 1995;267(5204):1619-1620. doi:10.1126/science.7886447
10. Dill KA, Maccallum JL, Folding P. The Protein-Folding Problem , 50 Years On. 2012;(NOVEMBER):1042-1046.
11. Hilser VJ, Freire E. Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. *J Mol Biol*. 1996;262(5):756-772. doi:10.1006/jmbi.1996.0550
12. Tsai C-J, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels, and protein function. *Protein Sci*. 1999;8(6):1181-1190. doi:10.1110/ps.8.6.1181
13. Motlagh HN, Wrabl JO, Li J, Hilser VJ. The ensemble nature of allostery. *Nature*. 2014;508(7496):331-339. doi:10.1038/nature13001
14. Bryngelson JD, Wolynes PG. Spin glasses and the statistical mechanics of protein folding. *Proc Natl Acad Sci U S A*. 1987;84(21):7524-7528. doi:10.1073/pnas.84.21.7524

15. Henzler-Wildman K, Kern D. Dynamic personalities of proteins. *Nature*. 2007;450(7172):964-972. doi:10.1038/nature06522
16. Ben-Nissan G, Sharon M. Capturing protein structural kinetics by mass spectrometry. *Chem Soc Rev*. 2011;40(7):3627-3637. doi:10.1039/c1cs15052a
17. Wright PE, Dyson HJ. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J Mol Biol*. 1999;293(2):321-331. doi:10.1006/jmbi.1999.3110
18. Uversky VN. Intrinsically disordered proteins from A to Z. *Int J Biochem Cell Biol*. 2011;43(8):1090-1103. doi:10.1016/j.biocel.2011.04.001
19. Fenton. Aron W. Allostery: an illustrated definition for the ‘second secret of life.’ *Trends Biochem Sci*. 2008;33(9):420-425. doi:10.1016/j.tibs.2008.05.009.Allostery
20. Changeux JP. 50Th Anniversary of the Word “Allosteric.” *Protein Sci*. 2011;20(7):1119-1124. doi:10.1002/pro.658
21. Cui Q, Karplus M. Allostery and cooperativity revisited. *Protein Sci*. 2008;17(8):1295-1307. doi:10.1110/ps.03259908
22. Goodey NM, Benkovic SJ. Allosteric regulation and catalysis emerge via a common route. *Nat Chem Biol*. 2008;4(8):474-482. doi:10.1038/nchembio.98
23. Monod J, Changeux JP, Jacob F. Allosteric proteins and cellular control systems. *J Mol Biol*. 1963;6(4):306-329. doi:10.1016/S0022-2836(63)80091-1
24. Koshland DE, Némethy G, Filmer D. Comparison of Experimental Binding Data. *Theor Model proteins Contain subunits*. 1966;5(1):365-385.
25. Hilser VJ, Wrabl JO, Motlagh HN. Structural and energetic basis of allostery. *Annu Rev Biophys*. 2012;41(1):585-609. doi:10.1146/annurev-biophys-050511-102319
26. Boehr DD, Nussinov R, Wright PE. The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol*. 2009;5(11):789-796. doi:10.1038/nchembio.232
27. Byun JA, VanSchouwen B, Akimoto M, Melacini G. Allosteric inhibition explained through conformational ensembles sampling distinct “mixed” states. *Comput Struct Biotechnol J*. 2020;18:3803-3818. doi:10.1016/j.csbj.2020.10.026
28. Swain JF, Gierasch LM. The changing landscape of protein allostery. *Curr Opin Struct*

- Biol.* 2006;16(1):102-108. doi:10.1016/j.sbi.2006.01.003
29. Cooper A, Dryden DTF. Allostery without conformational change - A plausible model. *Eur Biophys J.* 1984;11(2):103-109. doi:10.1007/BF00276625
 30. Koshland DE. The Key–Lock Theory and the Induced Fit Theory. *Angew Chemie Int Ed English.* 1995;33(23-24):2375-2378. doi:10.1002/anie.199423751
 31. Lichtenthaler FW. 100 Years “Schlüssel-Schloss-Prinzip”: What Made Emil Fischer Use this Analogy? *Angew Chemie Int Ed English.* 1995;33(23-24):2364-2374. doi:10.1002/anie.199423641
 32. Koshland DE. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci.* 1958;44(2):98-104. doi:10.1073/pnas.44.2.98
 33. Konermann L, Pan J, Liu YH. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem Soc Rev.* 2011;40(3):1224-1234. doi:10.1039/c0cs00113a
 34. Hubbard RE, Kamran Haider M. Hydrogen Bonds in Proteins: Role and Strength. *eLS.* 2010;(February). doi:10.1002/9780470015902.a0003011.pub2
 35. Sowole MA, Konermann L. Effects of protein-ligand interactions on hydrogen/deuterium exchange kinetics: Canonical and noncanonical scenarios. *Anal Chem.* 2014;86(13):6715-6722. doi:10.1021/ac501849n
 36. Tokmina-Lukaszewska M, Patterson A, Berry L, Scott L, Balasubramanian N, Bothner B. The role of mass spectrometry in structural studies of flavin-based electron bifurcating enzymes. *Front Microbiol.* 2018;9(JUL):1-18. doi:10.3389/fmicb.2018.01397
 37. Masson GR, Burke JE, Ahn NG, et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods.* 2019;16(7):595-602. doi:10.1038/s41592-019-0459-y
 38. Dodds JN, Baker ES. Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *J Am Soc Mass Spectrom.* 2019. doi:10.1007/s13361-019-02288-2
 39. Uetrecht C, Rose RJ, Van Duijn E, Lorenzen K, Heck AJR. Ion mobility mass spectrometry of proteins and protein assemblies. *Chem Soc Rev.* 2010;39(5):1633-1655. doi:10.1039/b914002f
 40. Smith DP, Knapman TW, Campuzano L, et al. Deciphering drift time measurements from travelling wave ion mobility spectrometry- Mass spectrometry studies. *Eur J Mass Spectrom.* 2009;15(2):113-130. doi:10.1255/ejms.947

41. Shvartsburg AA, Smith RD. Fundamentals of traveling wave ion mobility spectrometry. *Anal Chem.* 2008;80(24):9689-9699. doi:10.1021/ac8016295
42. Waters Corporation. Waters SYNAPT G2 Mass Spectrometry System Operator's Overview and Maintenance Guide Revision A. 2009.
43. Giles K, Williams JP, Campuzano I. Enhancements in travelling wave ion mobility resolution. *Rapid Commun Mass Spectrom.* 2011;25(11):1559-1566. doi:10.1002/rcm.5013
44. Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun Mass Spectrom.* 2004;18(20):2401-2414. doi:10.1002/rcm.1641
45. Bush MF, Hall Z, Giles K, Hoyes J, Robinson C V., Ruotolo BT. Collision cross sections of proteins and their complexes: A calibration framework and database for gas-phase structural biology. *Anal Chem.* 2010;82(22):9557-9565. doi:10.1021/ac1022953

CHAPTER TWO

A CATALYTIC DYAD MODULATES CONFORMATIONAL CHANGE IN THE CO₂-
FIXING FLAVOENZYME 2-KETOPROPYL COENZYME M
OXIDOREDUCTASE/CARBOXYLASE

Contribution of Authors and Co-Authors

Manuscript in Chapter 2

Author: Jenna R. Mattice

Contributions: Developed and performed experiments. Wrote and edited the manuscript.

Co-Author: Krista A. Shisler

Contributions: Contributed to the understanding and application of the experiments. Reviewed and edited the manuscript.

Co-Author: Jennifer L. DuBois

Contributions: Supervised experiments. Reviewed and edited the manuscript.

Co-Author: John W. Peters

Contributions: Supervised experiments. Reviewed and edited the manuscript.

Co-Author: Brian Bothner

Contributions: Supervised experiments. Reviewed and edited the manuscript.

Manuscript Information

Jenna R. Mattice, Krista A. Shisler, Jennifer L. DuBois, John W. Peters, and Brian Bothner

Journal of Biological Chemistry

Status of Manuscript: [Put an x in one of the options below, delete this]

☐ Prepared for submission to a peer-reviewed journal

☐ Officially submitted to a peer-reviewed journal

☐ Accepted by a peer-reviewed journal

☒ Published in a peer-reviewed journal

Journal of Biological Chemistry

Volume 298, Issue 5

DOI: <https://doi.org/10.1016/j.jbc.2022.101884>



A catalytic dyad modulates conformational change in the CO₂-fixing flavoenzyme 2-ketopropyl coenzyme M oxidoreductase/carboxylase

Received for publication, November 12, 2021, and in revised form, March 21, 2022. Published, Papers in Press, March 31, 2022.
<https://doi.org/10.1016/j.jbc.2022.101884>

Jenna R. Mattice¹, Krista A. Shisler², Jennifer L. DuBois¹, John W. Peters², and Brian Bothner^{1,*}

From the ¹Department of Chemistry and Biochemistry, Montana State University, Bozeman, Montana, USA; ²Institute of Biological Chemistry, Washington State University, Pullman, Washington, USA

Edited by Ruma Banerjee

2-Ketopropyl-coenzyme M oxidoreductase/carboxylase (2-KPCC) is a member of the flavin and cysteine disulfide containing oxidoreductase family (DSOR) that catalyzes the unique reaction between atmospheric CO₂ and a ketone/enolate nucleophile to generate acetoacetate. However, the mechanism of this reaction is not well understood. Here, we present evidence that 2-KPCC, in contrast to the well-characterized DSOR enzyme glutathione reductase, undergoes conformational changes during catalysis. Using a suite of biophysical techniques including limited proteolysis, differential scanning fluorimetry, and native mass spectrometry in the presence of substrates and inhibitors, we observed conformational differences between different ligand-bound 2-KPCC species within the catalytic cycle. Analysis of site-specific amino acid variants indicated that 2-KPCC-defining residues, Phe501-His506, within the active site are important for transducing these ligand induced conformational changes. We propose that these conformational changes promote substrate discrimination between H⁺ and CO₂ to favor the metabolically preferred carboxylation product, acetoacetate.

Atmospheric carbon dioxide is a carbon source for building biomass in plants and microbes. In plants, RuBisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase), the most abundant enzyme on earth, catalyzes the first step in CO₂ fixation (1, 2). A variety of lesser known, nonphotosynthetic microbial processes that fix CO₂ or bicarbonate have also been described. However, how these processes contribute to the global carbon cycle and their catalytic mechanisms is not well understood (1).

2-ketopropyl-coenzyme M oxidoreductase/carboxylase (2-KPCC) is a unique CO₂ fixing carboxylase which is part of the flavin/cysteine disulfide oxidoreductase (DSOR) family, members of which are best known for reducing metal ion or disulfide substrates (3). Glutathione reductase (GR) is a well characterized, paradigmatic DSOR that is responsible for cellular redox homeostasis (4) (Fig. 1A). Like GR, 2-KPCC is a homodimer with a redox active disulfide/FAD cofactor in each

subunit. However, 2-KPCC has distinguishing features that appear to permit it to perform a carboxylation (3, 5, 6). Specifically, typical DSORs have a conserved His-Glu catalytic dyad (4, 7). The conserved His serves as a source of protons during the catalytic cycle and stabilizes an important reaction intermediate. The conserved Glu moderates the sidechain orientation and pKa of the conserved His (4, 7). In 2-KPCC, this catalytic dyad is substituted by a Phe-His pair. Previous studies have shown that the Phe-His in 2-KPCC produces a unique reduced active form and facilitates carboxylation (8). The reduced, reactive form of 2-KPCC reductively cleaves a thioether substrate 2-KPC (2-ketopropyl-coenzyme M). The product of this cleavage is an unstable enol-acetone anion, which as a nucleophile attacks enzyme-bound CO₂ to form acetoacetate (Fig. 1B). In the absence of CO₂ and in the presence of available protons, acetone forms as an undesirable side product. Therefore, 2-KPCC has a unique requirement, compared to GR, in that it must protect the enol-acetone intermediate from bulk solvent to produce the carboxylation product, acetoacetate (Fig. 1B) (1). Prior work showed that the Phe of the Phe-His dyad suppresses protonation of the enol-acetone intermediate (9), while the His helps stabilize acetoacetate (10), though whether their influence on the reaction is wholly electrostatic or something else is not known.

The overall structures of 2-KPCC and GR are very similar. Both are homodimers of subunits that each contain FAD binding, NADPH binding, and interfacial domains (5, 6, 11). Unlike GR, 2-KPCC has extended N and C termini and a 13 amino acid insertion (Fig. 2) (5). The structure of 2-KPCC has been determined in the absence of added ligands and in several ligand-bound states (5, 6, 10–12). The structure of 2-KPCC in the presence of substrates and intermediates reveals conformational shifts in which the substrate access channel has collapsed. This is mainly due to differences in the conformation of the aforementioned extended N and C termini and 13 amino acid insertion not present in GR. In the presence of substrates and intermediates, structures of 2-KPCC exist with a smaller hydrophobic channel that we have proposed facilitates the access of CO₂ but excluding other molecules including aqueous solvent. We have hypothesized that this arrangement and implied conformation change is of key

* For correspondence: Brian Bothner, bbothner@montana.edu.



Catalytic residues and conformational change in CO₂ fixation

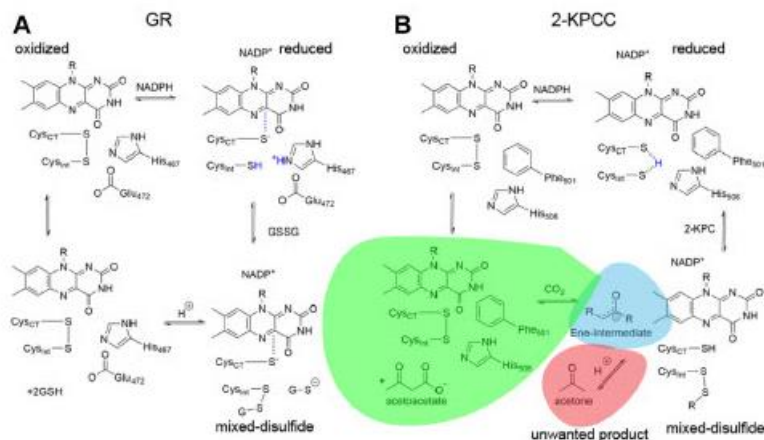


Figure 1. Reactions of glutathione reductase and 2-KPCC showing the active site residues and intermediates. A, oxidized active site of GR is reduced by NADPH, the reduced reactive form then reacts with GSSG to make a mixed disulfide and one glutathione (GSH). Lastly, the active site returns to its oxidized form and releases two GSH. B, oxidized active site of 2-KPCC is reduced by NADPH, the reduced reactive form then reacts with 2-KPC to produce a mixed disulfide and the Ene-intermediate (blue). The Ene-intermediate needs to be protected from protons or it will form acetone (red) instead of acetoacetate (green). Lastly, after carboxylation is complete, the active site returns to its oxidized form and acetoacetate and CoM are released. 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; CoM, coenzyme M; GR, glutathione reductase; GSSG, glutathione disulfide.

importance in protecting reaction intermediates promoting carboxylation and the formation of acetoacetate relative to protonation and the formation of the unproductive product acetone (5, 6)

To test this hypothesis and gain insight into solution behavior, we have compared 2-KPCC and GR using a suite of

biophysical techniques sensitive to conformational differences including ion mobility (IM)–native mass spectrometry (NMS), limited proteolysis, differential scanning fluorimetry (DSF), and the computational tool, Allosteric Signaling and Mutation Analysis (AlloSigMA). Experimental conditions include different ligand-bound and oxidation states of the enzyme.

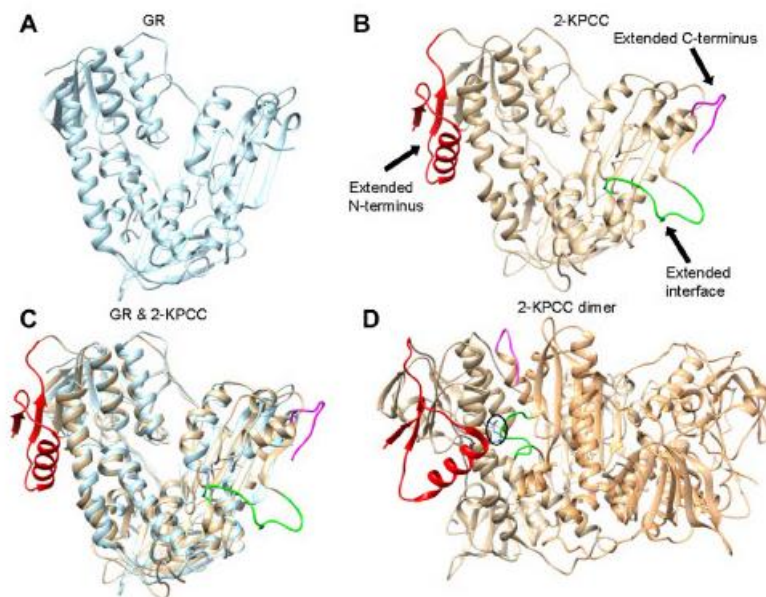


Figure 2. Structural comparison of 2-KPCC and glutathione reductase. A, structure of a glutathione reductase monomer (PDB: 1GRG). B, 2-KPCC (PDB: 1MO9) monomer highlighting extended regions, N terminus (red), C terminus (magenta), and extended interface domain (green). C, superimposed structures of 2-KPCC and glutathione reductase. D, dimeric structure of 2-KPCC. The regions unique to 2-KPCC surround the opening to the active site (black oval). 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; GR, glutathione reductase.

Catalytic residues and conformational change in CO₂ fixation

One condition of particular note is the oxidized enzyme and oxidized ligand complex, which was used to prevent product turnover and effectively explore the conformational changes reported in prior crystallographic work. Further, wildtype (WT) and amino acid-substituted variants of 2-KPCC in which one or both catalytic dyad residues (Phe501 and His506) are exchanged for the characteristic His-Glu pair found in GR were examined to determine whether the catalytic dyad, which promotes formation of the carboxylation product, might likewise play a role in producing conformational change. The evidence indeed indicates that 2-KPCC's noncanonical Phe501-His506 active site dyad contributes to regulating dynamics associated with carboxylation.

Results and discussion

Addition of 2-KPC induced a more compact 2-KPCC structure

To begin a solution phase structural investigation of GR and 2-KPCC, NMS was conducted. In NMS, noncovalent interactions are maintained, facilitating an analysis of protein conformational changes upon ligand binding (13, 14). Solutions containing 5 μ M GR or 2-KPCC with and without substrate (2 mM 2-KPC or glutathione disulfide (GSSG) intended to maximize the enzyme-substrate bound form) (15) were prepared in 300 mM ammonium acetate and analyzed by direct infusion NMS. The charge state distributions of 2-KPCC in the presence and absence of 2-KPC indicated that 2-KPC narrowed the distribution of charge states and shifted it to a higher mass to charge (m/z) ratio (Fig. 3A). This suggests a more compact protein fold when 2-KPC is present. In contrast,

GR exhibited no change in charge state distribution upon GSSG binding (Fig. 3A).

IM coupled to NMS was then employed using the same experimental conditions. IM combines transfer time through an ion mobility cell and m/z to separate ions based on collisional cross section (CCS). Comparison of the observed CCS with values calculated from structural models provides information on protein conformation. Differences in protein conformation are detected as a change in CCS (13, 14). 2-KPCC in the absence of 2-KPC had a longer drift time than in its presence, corresponding to a larger CCS value for 2-KPCC in the absence of 2-KPC (Fig. 3A). Also, in the absence of substrate, 2-KPCC had a wider range of drift times, indicating a spread in the CCS which translates to a more diverse ensemble of conformations. This can be interpreted as greater conformational freedom in the absence of 2-KPC. In contrast, GR did not exhibit a shift in drift time with GSSG binding (Fig. 3A), giving strong evidence of no apparent difference in conformation.

Limited proteolysis revealed ligand-dependent protection of two protease cleavage sites

The IM-NMS data provided evidence for a substrate-dependent conformational change, but no information on its nature or location. To further our investigation, we used limited proteolysis to track the rate and location of kinetically favored sites of cleavage. The rate of proteolysis is expected to differ between dynamic and stable regions, with unstructured/dynamic regions of a protein exhibiting faster peptide bond hydrolysis (16, 17). Thermolysin was selected for this

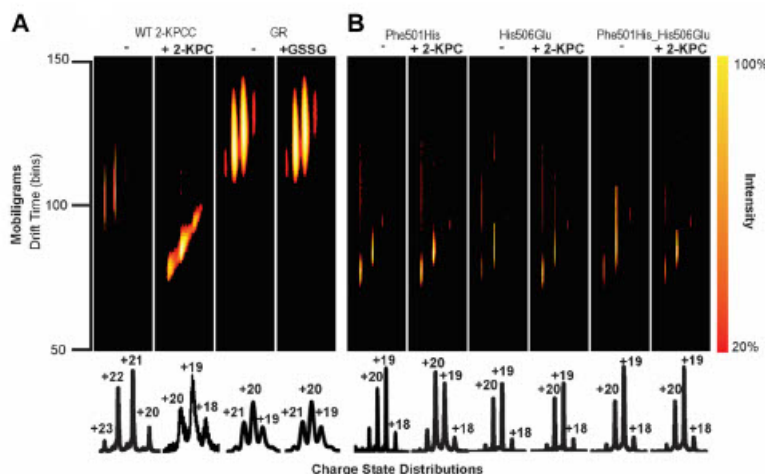


Figure 3. Mobligrams and charge state distributions of WT, mutant 2-KPCC, and glutathione reductase. Charge state distribution of proteins in the native state is on the bottom of each panel. Ion mobility data for the specific charge species are directly above. The vertical axis is time for ion mobility with lower times indicating a smaller CCS. The color scale on the right shows relative intensity starting at 20%. A, from left to right WT 2-KPCC $\pm$ 2 mM 2-KPC and GR $\pm$ 2 mM GSSG. The comparison of WT 2-KPCC $\pm$ 2 mM 2-KPC depicts a change in ion mobility and charge state distribution based on 2-KPC binding. The comparison of GR $\pm$ 2 mM GSSG depicts no change in ion mobility or charge state distribution based on 2-KPC binding. B, from left to right Phe501His, His506Glu, Phe501His_His506Glu 2-KPCC $\pm$ 2 mM 2-KPC. The variants showed no change in ion mobility or charge state distribution based on 2-KPC binding. 2-KPC, 2-ketopropyl coenzyme M; 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; CCS, collisional cross section; GR, glutathione reductase; GSSG, glutathione disulfide.

Catalytic residues and conformational change in CO₂ fixation

experiment because of the high frequency of potential cleavage sites in 2-KPCC (168) and GR (149) based on predicted sequence specificity (Fig. S1). Substrate binding can affect the rate and pattern of hydrolysis by changing accessibility to cleavable sites. Solutions of 2-KPCC and GR at 0.5 mg/ml were incubated with thermolysin at a 1:100 (wt:wt) ratio of protease to protein. To probe for substrate-induced conformational change, each enzyme was poised with different ligands (NADPH, 2-KPC, GSSG, and acetoacetate and NADP⁺). Aliquots were quenched for analysis at 0 min, 30 min, 2 h, and 5 h. The reaction was then visualized by SDS-PAGE.

For WT 2-KPCC, there were significant differences in the cleavage kinetics and SDS-PAGE gel banding pattern depending on the ligand present (Fig. 4A). 2-KPCC in the absence of ligands was completely digested within 2 h: more rapidly than any of the conditions where a ligand was present. This suggests that the ligand-free state of the enzyme has the greatest conformational freedom. Further, the banding pattern on the gel suggests there are two primary sites of cleavage in the 2-KPCC 57 kDa monomer, termed sites 1 and 2. Cleavage at site 1 generates a band with an apparent molecular weight just slightly smaller than the parent protein (51 kDa), while cleavage at site 2 gives rise to a 42 kDa band (Fig. 4A). To identify the main proteolysis cleavage sites on 2-KPCC, solution phase LC-MS and in-gel digestion were employed. Both sites are found on the N terminus, site 1 was identified as amino acid A26 and site 2, which gives rise to the formation of the 42 kDa band, is F142 (Fig. 4B). The expected 9 kDa peptide accompanying the 42 kDa fragment was too small to be resolved by the SDS-PAGE experiment.

In sharp contrast to the ligand-free state, the presence of 2-KPC dramatically slowed proteolysis. The intact protein was present throughout the 5 h incubation, with the 51 kDa band present after 2-h and the 42 kDa band not appearing at all. These data suggest that 2-KPC supplies substantial protection from proteolysis at both cleavage positions. Binding the products NADP⁺ and acetoacetate had very similar effects to 2-KPC, while in the NADPH present condition, the 42 kDa band appeared after 30 min and persisted to the 5-h time

point. For GR, there was no significant change in cleavage kinetics or SDS-PAGE gel banding pattern upon substrate binding (Fig. 4C). In ligand-free, NADPH present, and GSSG present GR, the intact band was present at 5 h. GR did not exhibit change in proteolytic rate in the absence or presence of ligands, and this suggests little conformational flexibility regardless of ligand binding.

Collectively, the limited proteolysis data indicate that 2-KPCC underwent ligand-induced conformational changes that restricted protease cleavage in the presence of 2-KPC, NADPH, or the acetoacetate/NADP⁺ products, while GR did not. The ligand-induced conformational changes seen in 2-KPCC also depend on the ligand present, suggesting that the conformational change undergone during reaction may follow an ordered change which will help protect the reaction intermediate.

Global thermal stability is increased by substrate binding in 2-KPCC

The limited proteolysis data revealed that substrate binding altered susceptibility to proteolysis. This inspired us to probe whether the same factors also influence global protein stability. To investigate this, we employed DSF to provide a read out of thermal stability of the proteins based on their melting profiles. A hydrophobic, fluorescent dye, which is quenched in a polar environment, is added to the protein solution. As the temperature is raised, the dye binds to hydrophobic pockets that become accessible as the protein structure denatures, producing a melting curve (18–20). DSF reports on local and global thermal stability. It has been shown that this technique can be utilized to investigate ligands that alter thermal stability (20, 21) and is particularly well-suited for protein complexes (22, 23). Therefore, we employed it to investigate the substrate binding effects on global thermal stability of 2-KPCC.

Reaction solutions of 0.25 mg/ml 2-KPCC or GR were incubated with 50× concentration of Sypro Orange dye with or without substrate (0–5 mM). 2-KPCC has a thermal melting point (T_m) of 65.5 °C (Fig. 5A). Binding 2-KPC and NADPH

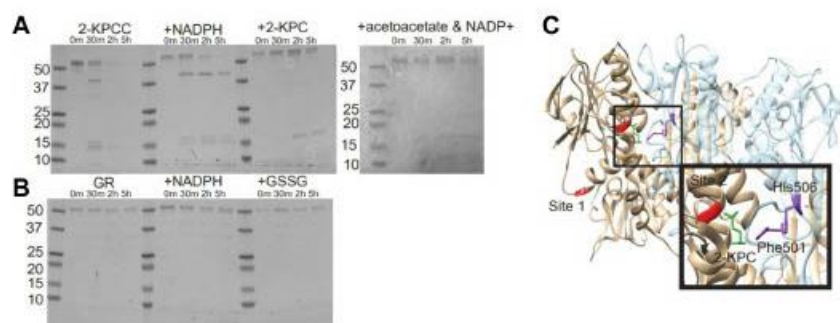


Figure 4. Limited proteolysis of WT 2-KPCC. The presence of substrate alters the kinetics and pattern of proteolysis. A, SDS-PAGE gel of limited proteolysis showing banding pattern of WT 2-KPCC under ligand-free, NADPH-bound, 2-KPC-bound, and acetoacetate & NADP⁺ conditions. B, SDS-PAGE gel of limited proteolysis showing banding pattern of GR under ligand-free, NADPH-bound, and GSSG-bound conditions. C, structure of 2-KPCC with two thermolysin cleavage sites shown in red. Inset, active site residues Phe501 and His506 (purple) and 2-KPC-bound (green) and proximity to the cleavage site (red) which produces the 42 kDa form. 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; GR, glutathione reductase; GSSG, glutathione disulfide.

Catalytic residues and conformational change in CO₂ fixation

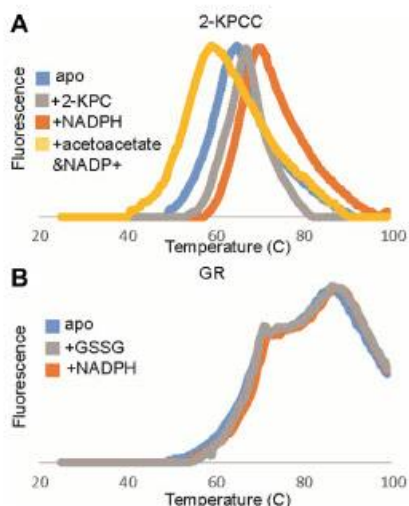


Figure 5. Differential scanning fluorimetry curves of glutathione reductase and WT 2-KPCC. A, DSF curves showing substrate binding effects on melting curves of WT 2-KPCC. Apo (blue) $T_m = 66.5^\circ\text{C}$ SD = 0.05, +2-KPC (gray) $T_m = 67^\circ\text{C}$ SD = 0.34, +NADPH (orange) $T_m = 70.5^\circ\text{C}$ SD = 0.74, +acetoacetate and NADP+ $T_m = 60^\circ\text{C}$ SD = 0.071. B, DSF curves showing substrate binding effects on melting curves for GR. Apo (blue) SD = 0.013, +GSSG (gray) SD = 0.017, +NADPH (orange) SD = 0.017 $T_m = 87.5^\circ\text{C}$. Overall 2-KPCC exhibits lower melting temperatures than glutathione reductase. 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; DSF, differential scanning fluorimetry; GR, glutathione reductase; GSSG, glutathione disulfide.

shifted the T_m to 67°C and 70.5°C , respectively. This indicates that the ligands have a stabilizing effect on 2-KPCC. Increased thermal stability is consistent with a model in which the substrate-bound conformation can exclude solvent and a more compact structure is present. The stability and compact form arise from enhanced noncovalent interactions in the protein–ligand complex (20, 21, 24). We would expect a more compact structure to have a higher T_m , like what is exhibited by binding of 2-KPC and NADPH. In contrast, binding of the reaction products acetoacetate and NADP⁺ lowered the T_m . This destabilization could characterize a product release form of 2-KPCC which was too small to be seen by limited proteolysis. Interestingly, GR has the highest T_m at 87°C and exhibits no shift in T_m upon ligand binding but does exhibit a local destabilization at 72°C (Fig. 5B). Together the IM-NMS, proteolysis, and thermal stability data reflect the crystallographic and catalytic models in which structural change is part of the catalytic cycle. The binding of different ligands produces different structural changes within 2-KPCC. These changes are not shared by GR, which has a high global thermal stability with none of the structural parameters interrogated here altered by substrate binding.

Influence of 2-KPCC catalytic dyad on conformation

To better elucidate the cause for 2-KPCC's conformational changes compared to GR's static behavior, we wanted to

explore amino acid–substituted variants of 2-KPCC in which one or both catalytic dyad residues (Phe501 and His506) are substituted by the characteristic His–Glu pair found in GR (Phe501His and/or His506Glu). Prior work showed that the Phe of the Phe–His dyad suppresses protonation of the enol-acetone intermediate (9), while the His helps stabilize the formation of acetoacetate (10). We now sought to determine if the catalytic dyad, which is essential for promoting formation of the carboxylation product, might likewise play a role in the conformational changes that are proposed to exclude solvent from the enol-acetone intermediate.

To probe the structures of these variants, NMS and IM were employed in the presence and absence of 2-KPC. The charge state distributions of Phe501His, His506Glu, and Phe501His_His506Glu were lower than WT in the absence of 2-KPC, but not as low as WT in the presence of 2-KPC (Fig. 3B). In contrast to the WT enzyme, the variants did not have a shift in charge state upon addition of 2-KPC. The catalytic dyad of 2-KPCC promotes carboxylation and, when mutated, alters the product outcome (9, 10). Our results suggest a role for the unique catalytic dyad in carboxylation is manifested at least in part in transducing an essential conformational change. In the variants, the catalytic dyad stabilizes a more compact structure even without substrate. The 2-KPCC variants were also investigated in the presence and absence of 2-KPC (Fig. 3B). Importantly, they did not show a decrease in drift time with 2-KPC. However, they did exhibit a wider distribution of drift times. Taken with the NMS data, the variants appear to be in a conformation that closely resembles WT 2-KPCC in the presence of 2-KPC. This further implies that the Phe501His and His506Glu substitutions alter the conformational freedom of 2-KPCC.

Limited proteolysis indicates dyad-dependent protection of proteolytic sites

The 2-KPCC variants were also investigated by proteolysis. In the Phe501His variant, the ligand-free and NADPH present states exhibit lower proteolysis than WT (Fig. 6A), with the intact protein band persisting up to the 2 h time point. Addition of 2-KPC or NADPH extends the parent band lifetime to the 5 h mark and also stabilized the 42 kDa band so that it was likewise observable on this timescale. The His506Glu and the Phe501His_His506Glu reactions had intact protein remaining at the 2-h time point in all conditions and no observable 42 kDa species (Fig. 6, B and C). These results suggest that the His506Glu substitution stabilized the intact protein by limiting protease access to the helix that is cleaved to generate the 42 kDa form, but inhibited substrate-based protection.

The active site residues Phe501 and His506 modulated protease access in 2-KPCC. When the active site was mutated to the canonical DSOR catalytic dyad residues, substrate-based protease protection was not observed in the variants with His506Glu substitution. This suggests that the protease-protective conformational changes could be the same as those needed for carboxylation. The His506Glu substitution

Catalytic residues and conformational change in CO₂ fixation

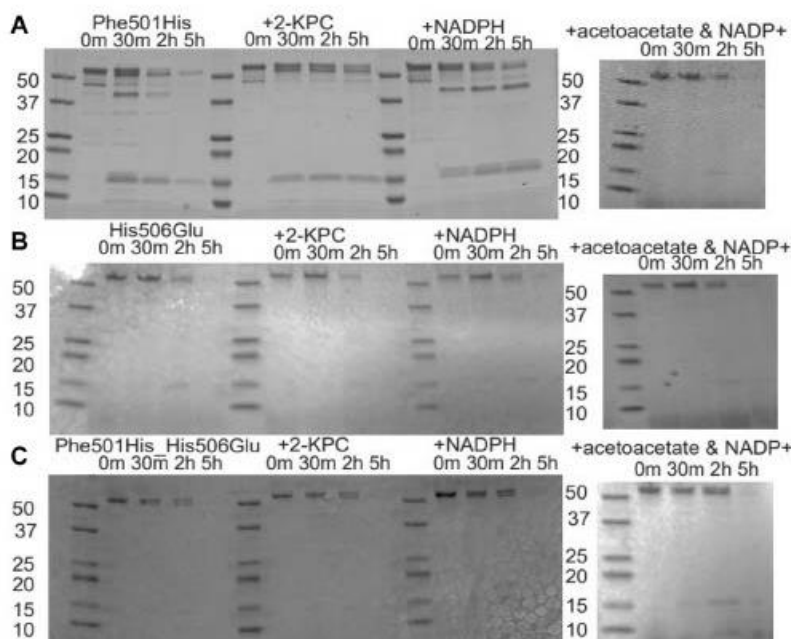


Figure 6. Limited proteolysis of 2-KPCC variants. The presence of substrate alters the kinetics and pattern of proteolysis. *A*, SDS-PAGE gel of limited proteolysis showing banding pattern of Phe501His 2-KPCC under ligand-free, 2-KPC-bound, NADPH-bound, and acetoacetate & NADP⁺-bound conditions. *B*, SDS-PAGE gel of limited proteolysis showing banding pattern of His506Glu 2-KPCC under ligand-free, 2-KPC-bound, NADPH-bound, and acetoacetate & NADP⁺-bound conditions. *C*, SDS-PAGE gel of limited proteolysis showing banding pattern of Phe501His_His506Glu 2-KPCC under ligand-free, 2-KPC-bound, NADPH-bound, and acetoacetate & NADP⁺-bound conditions. 2-KPC, 2-ketopropyl coenzyme M; 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase.

appeared to alter conformational freedom more significantly than the Phe501His, as the substrate- and product-dependent modulation of protease access was not observed for this variant.

Active site dyad contributes to global thermal stability

The IM-NMS and limited proteolysis data demonstrated that the active site variants alter conformational freedom. We then employed DSF to investigate if the variants alter global protein stability. The ligand-free His506Glu variant has a decreased T_m relative to the WT and other variants, with a T_m of 57 °C (Fig. 7). The His506Glu variant exhibited small shifts to higher T_m upon binding of 2-KPC and NADPH and a slight shift to a lower T_m with acetoacetate and NADP⁺. This supports that the His506Glu substitution structure is possibly locked in a conformation not allowing for conformational freedom. Also, the melting temperatures of all conditions tested were similar to the acetoacetate and NADP⁺ condition of WT, which we interpret as the H506E substitution having a destabilizing effect on protein structure. The ligand free Phe501His and Phe501His_His506Glu variants have a similar T_m to WT with a T_m of 67 °C and 66 °C, respectively. However, the Phe501His variant did not have a large shift in T_m upon ligand binding except for with acetoacetate and

NADP⁺ present, with a T_m of 71 °C. The Phe501His_His506Glu variant exhibited some hallmarks of both the single variants. A shoulder was present in some of the curves of Phe501His and Phe501His_His506Glu 2-KPCC. This shoulder suggests a local destabilization in the protein structure as the temperature is raised. Interestingly, GR shares this local destabilization.

Computational analysis implicates catalytic dyad in conformational change

To further elucidate how these active site residues influence conformational change, we turned to a computational analysis using AlloSigMA (25, 26). AlloSigMA uses a structure-based statistical mechanical model of allostery to calculate the per-residue change in free energy for all possible protein configurations due to interactions with ligands or mutations (27, 28). AlloSigMA defines the protein contact network of a protein by solving the structure-based statistical mechanical model of allostery for all possible local configurations in all available states. AlloSigMA defines mutations as either: UP-mutations, which cause over-stabilizing effects on the protein contact network, and DOWN-mutations, which cause destabilizing effects on a protein contact network (29). To test the impact of an amino acid substitution at a specified position, the user

Catalytic residues and conformational change in CO₂ fixation

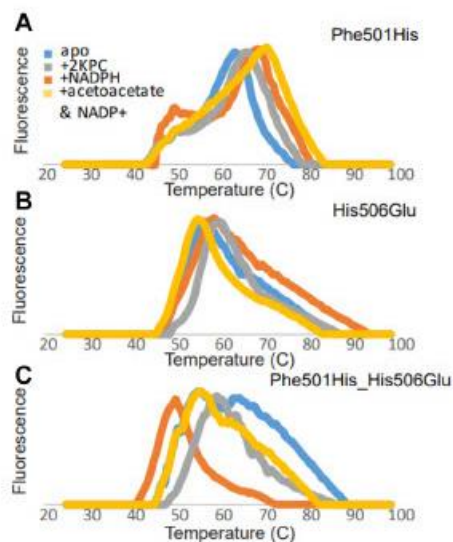


Figure 7. Differential scanning fluorimetry curves of 2-KPCC variants. A, DSF curves showing substrate binding effects on melting curves of Phe501His 2-KPCC. Apo (blue) $T_m = 67^\circ\text{C}$ SD = 0.06, +2-KPC (gray) $T_m = 67.5^\circ\text{C}$ SD = 0.02, +NADPH (orange) $T_m = 68^\circ\text{C}$ SD = 0.09, +acetoacetate and NADP+ $T_m = 71^\circ\text{C}$ SD = 0.05 $n = 3$. B, DSF curves showing substrate binding effects on melting curves for His506Glu 2-KPCC. Apo (blue) $T_m = 57^\circ\text{C}$ SD = 0.04, +2-KPC (gray) $T_m = 60^\circ\text{C}$ SD = 0.26, +NADPH (orange) $T_m = 58^\circ\text{C}$ SD = 0.04, +acetoacetate and NADP+ $T_m = 55.5^\circ\text{C}$ SD = 0.06 $n = 3$. C, DSF curves showing substrate binding effects on melting curves for Phe501His_His506Glu 2-KPCC. Apo (blue) $T_m = 55$ & 66°C SD = 0.06, +2-KPC (gray) $T_m = 57.5^\circ\text{C}$ SD = 0.02, +NADPH (orange) $T_m = 50^\circ\text{C}$ SD = 0.06, +acetoacetate and NADP+ $T_m = 56^\circ\text{C}$ SD = 0.017 $n = 3$. 2-KPC, 2-ketopropyl coenzyme M; 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; DSF, differential scanning fluorimetry.

selects the residue to be probed and the type of mutation to be modeled, UP or DOWN. This information is used to predict changes in allosteric communication across a protein model. Allostery, in its broadest sense, is the functional change at one site on a protein caused by a change at a distant site (28–30). AlloSigMA interprets allostery as a change at any site in the structure that can lead to a shift in occupancy of the conformational states in which the protein can exist.

The crystal structure of 2-KPCC (PDB: 1MOK) was explored using AlloSigMA, where the Phe-His dyad was singly or doubly mutated to a DOWN-mutation (destabilizing residue). AlloSigMA then calculated the per-residue change in allosteric free energy (kcal/mol) caused by the mutation. A positive change in allosteric free energy indicates a change in force acting on the residue because of the mutation, and it can be interpreted as a possible increase in dynamics that could lead to a conformational change (31). A negative change in allosteric free energy indicates a stabilization of the residue caused by the mutation and possibly preventing conformational change or a decrease in dynamics. DOWN-mutations at Phe501, His506, and Phe501_His506 revealed an increase in free energy in the active site while the rest of the protein structure had a decrease in free energy. These values were color coded positive (red) or negative (blue) and displayed on

the protein sequence or the structure (Fig. 8). These results are interpreted as an increase in conformational flexibility, where the number of possible conformational states is increased in the active site and a decrease in the rest of the protein. This was exhibited by all three variants, but the most prominent changes were seen in His506. A control experiment was conducted by mutating noncatalytically active residues in the active site and surface of the protein. These mutations produced very little change in free energy (Fig. S2).

GR was also explored by AlloSigMA with mutations to the His-Glu catalytic dyad (Fig. S3). Interestingly, single substitutions at either residue produced similar behavior to 2-KPCC, where the active site showed an increase in free energy while the rest of the protein had a decrease. Tandem substitutions exhibited an increase in free energy near the active site. This further supports that the active site catalytic dyad has an important role in modulating conformational change in 2-KPCC and differences between the 2-KPCC and canonical DSOR enzymes such as GR.

Conclusion

Taken together, the DSF, limited proteolysis, IM-NMS data, and AlloSigMA analyses support a functional model of catalysis in which 2-KPCC transitions between multiple conformations, the equilibrium of which is controlled by substrate and product. The IM-NMS data showed a change in conformation of 2-KPCC when 2-KPC was present, and that active site substitutions altered the conformational ensemble of 2-KPCC. Limited proteolysis allowed us to localize conformational change upon binding of substrates and products to the N terminus. A change we believe serves to exclude solvent from the active site (see Fig. 2D). A key finding of this work is the role of the catalytic dyad (Phe-His) in the modulation of 2-KPCC's conformational ensemble. DSF and AlloSigMA contributed significantly to this conclusion.

In recent years, our understanding of conformational change and allostery has shifted to a more nuanced model in which activity is best described in the context of conformational ensembles where population re-distributions influence catalysis (29, 32). Here we provide evidence that 2-KPCC is such a system, with a conformational ensemble that is influenced by substrate binding and modulated by the catalytic

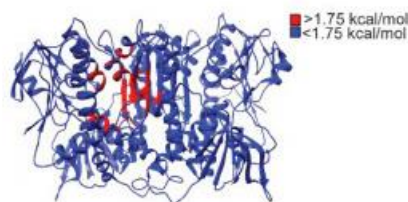


Figure 8. Change in free energy calculated by AlloSigMA as a result of mutations in the catalytic dyad plotted on structure of 2-KPCC (PDB: 1MOK). Increase in free energy (red) indicates more perturbation and higher dynamics. Decrease in free energy (blue) indicates a stabilization and decrease in dynamics. 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; AlloSigMA, Allosteric Signaling and Mutation Analysis.

Catalytic residues and conformational change in CO₂ fixation

dyad. We propose the ligand free, oxidized state of 2-KPCC has full conformational freedom. During the catalytic cycle, binding of NADPH and reduction of the disulfide restricts the conformational ensemble to forms that support the binding of 2-KPC and formation of the mixed disulfide (Fig. 9). The binding of 2-KPC results in a more compact structure that presumably protects the active site, only allowing access of the second substrate CO₂. Carboxylation then results in a form of 2-KPCC which supports the release of acetoacetate. In this scheme, changes to the conformational ensemble brought upon by substrate binding, help protect the enolacetone intermediate and promote carboxylation. Mutations of the catalytic Phe-His resulted in alterations to the conformational ensemble of 2-KPCC, and show the important role for the catalytic dyad in the conformational state of 2-KPCC.

2-KPCC's mechanism is in contrast to catalysis by other DSOR proteins such as GR, which undergoes reduction, mixed disulfide formation, and product release with no detectable change in conformation (33). This suggests that the carboxylation is a derived function arising from adaptations by the active site chemistry and protein conformational stability.

Experimental procedures

Expression and purification of native 2-KPCC and amino acid-substituted variants

E. coli BL21 (DE3)pLysS cells were transformed with the pBAD plasmid harboring the WT or corresponding 2-KPCC

variant gene from *Xanthobacter autotrophicus* Py2. The construct contained a His affinity tag, a thioredoxin tag, and a TEV-cleavage site. Cells were plated on LB agar + kanamycin (25 µg/ml) and grown overnight. A single colony from the plate was used to inoculate a 5 ml overnight culture in LB. The 5 ml overnight culture was used as the inoculum for a 500 ml baffled flask containing 500 ml of ZYP-rich medium + kanamycin (25 µg/ml). Cells were grown at 37 °C with agitation at 225 rpm until the A₆₀₀ reached 0.6 to 1.0. The temperature was reduced to 25 °C, L-arabinose was added to 0.02%, and the cells were grown for an additional 16 to 18 h. The cells were pelleted by centrifugation, frozen, and stored at -80 °C. The cell pellet was re-suspended in three volumes of lysis buffer [20 mM Tris, pH 8.0, 300 mM NaCl, 5 mM imidazole, and protease inhibitor tablets (Pierce)] and thawed at 30 °C. All subsequent treatments were performed on ice or at 4 °C. The resuspended cells were lysed *via* multiple rounds of sonication (Branson Ultrasonifier). Cell lysates were clarified *via* centrifugation at 75,000g for 45 min. Clarified lysates were loaded onto a nickel-nitrilotriacetic-resin column (Bio-Rad) *via* a Bio-Rad FPLC, washed with lysis buffer, and eluted using a 0.005 to 0.4 M imidazole gradient in lysis buffer at 3 ml/min. The eluted 2-KPCC was diluted 6-fold into buffer A (20 mM Tris-Cl, pH 6.5, 5% w/v glycerol), applied to a Q-Sepharose ion-exchange column (Bio-Rad), and eluted using a 0 to 1 M NaCl gradient in buffer A. Fractions were screened using SDS-PAGE. Pure 2-KPCC protein was pooled and run on a desalting column (Bio-Rad) in 20 mM Tris, pH 7.4, 10%

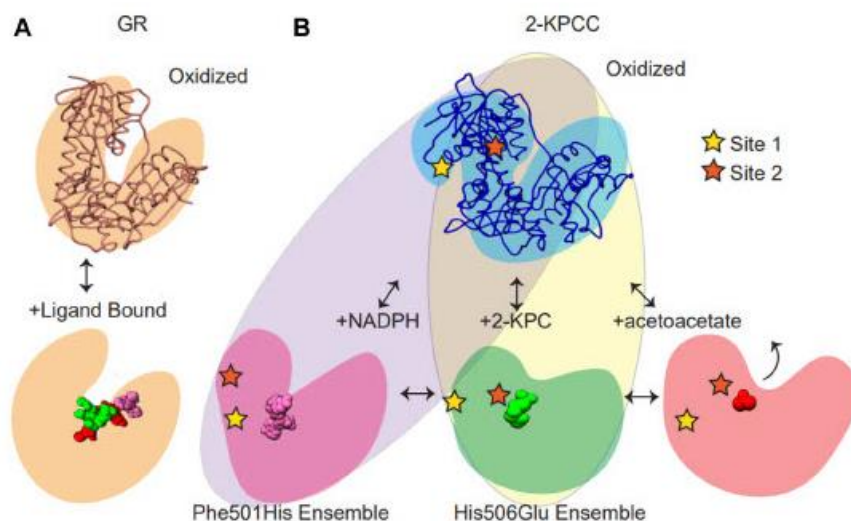


Figure 9. Conformational ensembles of GR and 2-KPCC. Schematic showing conformational changes associated with catalysis. One subunit of the dimer is shown. A, GR occupies a limited conformational space, remaining in the "open" conformation throughout the catalytic cycle (shown in tan). B, representative conformations of 2-KPCC highlighting the available conformational space sampled in solution. The oxidized, "open" conformations of 2-KPCC (blue) are shown with protease site 1 (yellow star) and site 2 (orange star). Binding of NADPH (pink) leads to altered access to site 2 and changes in conformational ensemble. 2-KPC binding (green) alters access to site 1 and the observed conformations of 2-KPCC. Binding of acetoacetate (red) represents yet another conformational ensemble favoring product release. The Phe501His variant restricts the available conformations which are populated in the ensemble (purple oval). Phe501His favors the oxidized and NADPH-bound conformations. The His506Glu variant restricts the accessible conformations (yellow oval) favoring oxidized and 2-KPC-bound conformations. 2-KPC, 2-ketopropyl coenzyme M; 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; GR, glutathione reductase.

Catalytic residues and conformational change in CO₂ fixation

glycerol, and 200 mM NaCl. Protein was concentrated using a 30 kDa MWCO spin filter concentrator. To remove the 6x-His tag, 2-KPCC protein was incubated with TEV protease at a ratio of 1:100 (w/w) (Millipore) overnight at 4 °C. TEV protease was removed using a nickel–nitrilotriacetic–resin column *via* gravity. Fractions were screened using SDS-PAGE. Cleaved 2-KPCC protein was pooled and concentrated using 30kDaMWCO spin filter concentrator (Millipore). Total protein concentration was determined using a BCA protein assay (Thermo Scientific), and flavin concentration was determined from its UV-visible absorbance at 450 nm using ϵ_{450} of 11,828 M⁻¹ cm⁻¹. All concentrations of protein cited in the text refer to flavin-containing protein.

NMS and IM of 2-KPCC and GR

Protein stock solutions were made anaerobic using a double-manifold Schlenk line with alternating cycles of argon gas purging and evacuation. Protein stocks were buffer exchanged into deoxygenated 300 mM ammonium acetate at pH 7.5 using 7.5 kDa mini-dialysis filters (Thermo-Fisher) overnight at 4 °C. Buffer exchanged proteins for bound conditions were then incubated for 15 min with either 2 mM 2-KPC, 2 mM GSSG, or 5 mM NADPH. Proteins were then loaded into in-house-prepared gold-coated borosilicate capillaries prepared as outlined in Luo *et al.* and analyzed using a Waters Synapt G2-Si with the following settings: capillary kV was set to 1.75; sampling cone, 95; source offset, 70; trap collision energy, 110 V; transfer collision energy, 105 V, source temperature, 30 °C; and trap gas, 7.0 ml/min. Data were analyzed using Waters MassLynx version 4.1. IM analysis was performed under the same conditions outlined above. The Synapt G2-Si instrument was set in IM mode, and data were analyzed by MassLynx 4.1 and DriftScope.

Limited proteolysis and peptide mapping of 2-KPCC and GR

2-KPCC (WT and variants) and GR samples at 0.5 mg/ml were prepared anaerobically (Mbraun anaerobic chamber Unilabpro) with deoxygenated buffer (200 mM Tris-HCl, 0.5 M CaCl₂ pH 8) and incubated with either 2 mM 2-KPC, 2 mM GSSG, 5 mM NADPH, or 2 mM acetoacetate and 5 mM NADP⁺. Prepared reactions were then sealed using crimp seal vials and removed from the Mbraun. Using a gas-tight syringe, thermolysin (Promega) was added to reactions at a ratio of 1:100 (wt:wt) and incubated at 37 °C. At timepoints (0 min, 30 min, 2 h, 5 h) aliquots of 9 µl were removed from the reaction and immediately diluted into 36 µl of gel loading buffer, vortexed, and heated to 100 °C for 5 min. Following boiling, samples were frozen. Prior to gel loading, the frozen samples were simultaneously boiled at 100 °C for 5 min, centrifuged, vortexed, and maintained at 25 °C during gel loading. Samples were run on 4% to 20% (w/v) gradient SDS-PAGE gels (Bio-Rad) with tris-glycine buffer. Gels were stained with GelCode blue safe protein stain (Thermo).

Proteolysis reactions for mass spectrometry analysis were conducted identical to those for gel-based analysis except for quenching. Quenching of proteolysis was performed by

multiple methods, these included dropping the pH, boiling, spotting directly onto a MALDI plate, and direct LCMS injection. LC-MS analysis of proteolysis reactions was completed on a 1200 infinity HPLC (Agilent) coupled to a MicroTOF LC mass spectrometer (Bruker). Peptides were separated on a size exclusion column (Phenomenex 150 × 2 mm) using a flow rate of 500 µl/min with solvent conditions 90% solvent A [0.1% FA (Sigma) in water (Thermo) and 10% solvent B (0.1% FA in acetonitrile) (Thermo)]. Data were acquired at 2 Hz over the scan range 50 to 1700 m/z in positive mode. Electrospray settings were as follows: nebulizer set to 3.7 bar, drying gas at 8.0 l/min, drying temperature at 350 °C, and capillary voltage at 3.5 kV. Data processing was carried out with Bruker DataAnalysis software. Unique peptides were identified by using Peptide Analysis Worksheet version 2000.6.8.0 with an accuracy of 100 ppm. Reactions were also analyzed by in-gel digestion and mass spectrometry. Protein bands from SDS-PAGE were cut and digested with trypsin (Promega) according to a standard protocol recommended by the manufacturer (1:50 wt:wt, overnight). Proteins were identified as described in (34) using a maXis Impact UHR-QTOF instrument (Bruker) interfaced with Dionex 3000 nano-uHPLC (Thermo). Data were analyzed and mapped by using Peptide Shaker, version 0.41.1, paired with Search GUI, version 1.30.1 (Compomics).

DSF of 2-KPCC and GR

Differential reactions containing 2 µg of protein in 200 mM Tris-HCl pH 8, a variety of ligands (2 mM 2-KPC, 2 mM GSSG, 5 mM NADPH, or 2 mM acetoacetate & 5 mM NADP⁺), and 1.2 µl of 50× stock SYPRO orange dye (Invitrogen) for a total reaction volume of 30 µl. The samples were loaded into a Roto-Gene Q instrument (Qiagen) with SYPRO orange absorbance excited at 470 nm and the fluorescence monitored at 570 nm as the temperature was ramped from 25 to 95 °C at 1 °C/min. All samples were run in triplicate.

AlloSigMA analysis of 2-KPCC and GR

The online computational tool AlloSigMA was used to explore the PDB files of both 2-KPCC and GR (PDB: 1MOK, and PDB: 1GRG), respectively. Published protocols were followed for this analysis (31). Upon uploading the files, an allosteric signaling map was calculated to exhaustively compute the allosteric modulation of each residue due to mutation. For 2-KPCC, calculations were computed for DOWN-mutations at residues, Phe501, His506, Phe501-His506, Leu502, and Gly377. For GR, calculations were computed for DOWN-mutations of residues His467, Glu472, and His467_Glu472. The change in free energy calculated for each mutational study was then plotted on to the PDB structures 1MOK and 1GRG.

Source of glutathione reductase

Recombinant expressed in *Escherichia coli*, human GR was purchased from Sigma-Aldrich. Cas Number: 9001-48-3.

Catalytic residues and conformational change in CO₂ fixation

Data availability

Data will be shared upon request.

Supporting information—This article contains supporting information.

Author contributions—J. R. M. investigation; J. R. M. formal analysis; J. R. M. validation; J. R. M. visualization; J. R. M. data curation; J. R. M. writing-original draft; J. R. M., K. A. S., J. L. D., J. W. P., and B. B. writing-review & editing; K. A. S. resources; J. L. D., J. W. P., and B. B. conceptualization; J. L. D., J. W. P., and B. B. supervision; J. W. P. and B. B. methodology; J. W. P. project administration; J. W. P. funding acquisition.

Funding and additional information—This work was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award DE-FG02-04ER15563 to J. L. D., J. W. P., and B. B. Partial salary support for J. W. P. was supported by the USDA National Institute of Food and Agriculture, Hatch umbrella project #1015621. Funding for the Mass Spectrometry Facility used in this publication was made possible in part by the MJ Murdock Charitable Trust, the National Institute of General Medical Sciences of the National Institutes of Health under Award Numbers P20GM103474 and S10OD28650, and the MSU Office of Research, Economic Development and Graduate Education. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The authors declare that they have no conflict of interest with the contents of this article.

Abbreviations—The abbreviations used are: 2-KPC, 2-ketopropyl coenzyme M; 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; AlloSigMA, Allosteric Signaling and Mutation Analysis; CCS, collisional cross section; CoM, coenzyme M; DSF, differential scanning fluorimetry; DSOR, disulfide oxidoreductase; FAD, flavin adenine dinucleotide; GR, glutathione reductase; GSSG, glutathione disulfide; IM, ion mobility; m/z, mass to charge; NADPH, β -nicotinamide adenine dinucleotide phosphate; NMS, native mass spectrometry; WT, wildtype.

References

- Erb, T. J. (2011) Carboxylases in natural and synthetic microbial pathways. *Appl. Environ. Microbiol.* 77, 8466–8477.
- Frey, P., and Hegeman, A. (2007) *Enzymatic Reaction Mechanisms*. Oxford University Press, Oxford, UK: 418–430.
- Clark, D. D., Allen, J. R., and Ensign, S. A. (2000) Characterization of five catalytic activities associated with the NADPH2-Ketopropyl-coenzyme M [2-(2-ketopropylthio)ethanesulfonate] oxidoreductase/carboxylase of the *Xanthobacter* strain Py2 epoxide carboxylase system. *Biochemistry* 39, 1294–1304.
- Rietveld, P., Arscott, L. D., Berry, A., Scrutton, N. S., Deonarain, M. P., Perham, R. N., and Williams, C. H., Jr. (1994) Reductive and oxidative half-reactions of glutathione reductase from *Escherichia coli*. *Biochemistry* 33, 13888–13895.
- Nocek, B., Jang, S. B., Jeong, M. S., Clark, D. D., Ensign, S. A., and Peters, J. W. (2002) Structural basis for CO₂ fixation by a novel member of the disulfide oxidoreductase family of enzymes, 2-ketopropyl-coenzyme M oxidoreductase/carboxylase. *Biochemistry* 41, 12907–12913.
- Pandey, A. S., Nocek, B., Clark, D. D., Ensign, S. A., and Peters, J. W. (2006) Mechanistic implications of the structure of the mixed-disulfide intermediate of the disulfide oxidoreductase, 2-ketopropyl-coenzyme M oxidoreductase/carboxylase. *Biochemistry* 45, 113–120.
- Deonarain, R., Mahendra, P., Berry, A., Scrutton, N. S., and Perham, R. N. (1989) Alternative proton donors/acceptors in the catalytic mechanism of the glutathione reductase of *Escherichia coli*: The role of histidine-439 and tyrosine-99. *Biochemistry* 28, 9602–9607.
- Streit, B. R., Matice, J. R., Prussia, G. A., Peters, J. W., and DuBois, J. L. (2019) The reactive form of a C-S bond-cleaving, CO₂-fixing flavoenzyme. *J. Biol. Chem.* 294, 5137–5145.
- Prussia, G. A., Gauss, G. H., Mus, F., Conner, L., DuBois, J. L., and Peters, J. W. (2016) Substitution of a conserved catalytic dyad into 2-KPCC causes loss of carboxylation activity. *FEBS Lett.* 590, 2991–2996.
- Prussia, G. A., Shisler, K. A., Zadovnyy, O. A., Streit, B. R., DuBois, J. L., and Peters, J. W. (2021) The unique Phe-His dyad of 2-ketopropyl coenzyme M oxidoreductase/carboxylase selectively promotes carboxylation and S–C bond cleavage. *J. Biol. Chem.* 297, 100961.
- Pandey, A. S., Mulder, D. W., Ensign, S. A., and Peters, J. W. (2011) Structural basis for carbon dioxide binding by 2-ketopropyl coenzyme M oxidoreductase/carboxylase. *FEBS Lett.* 585, 459–464.
- Kofoed, M. A., Wampler, D. A., Pandey, A. S., Peters, J. W., and Ensign, S. A. (2011) Roles of the redox-active disulfide and histidine residues forming a catalytic dyad in reactions catalyzed by 2-ketopropyl coenzyme M oxidoreductase/carboxylase. *J. Bacteriol.* 193, 4904–4913.
- Tokmina-Lukaszewska, M., Patterson, A., Berry, L., Scott, L., Balasubramanian, N., and Bothner, B. (2018) The role of mass spectrometry in structural studies of flavin-based electron bifurcating enzymes. *Front. Microbiol.* 9, 1–18.
- Uetrecht, C., Rose, R. J., Van Duijn, E., Lorenzen, K., and Heck, A. J. R. (2010) Ion mobility mass spectrometry of proteins and protein assemblies. *Chem. Soc. Rev.* 39, 1633–1655.
- Janes, W., and Schulz, G. E. (1990) The binding of the retro-analogue of glutathione disulfide to glutathione reductase. *J. Biol. Chem.* 265, 10443–10445.
- Fontana, A., Zamboni, M., Polverino De Laureto, P., De Filippis, V., Clementi, A., and Scaramella, E. (1997) Probing the conformational state of apomyoglobin by limited proteolysis. *J. Mol. Biol.* 266, 223–230.
- Bothner, B., Taylor, D., Jun, B., Lee, K. K., Siuzdak, G., Schlutz, C. P., and Johnson, J. E. (2005) Maturation of a tetra virus capsid alters the dynamic properties and creates a metastable complex. *Virology* 334, 17–27.
- Rayaproli, V., Kruse, S., Kant, R., Venkatakrishnan, B., Movahed, N., Brooke, D., Lins, B., Bennett, A., Potter, T., McKenna, R. M., Agbandje-McKenna, M., and Bothner, B. (2013) Comparative analysis of adeno-associated virus capsid stability and dynamics. *J. Virol.* 87, 13150–13160.
- Rayaproli, V., Kruse, S., Kant, R., Movahed, N., Brooke, D., and Bothner, B. (2014) Fluorometric estimation of viral thermal stability. *Bio Protoc.* 4, e1199.
- Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* 2, 2212–2221.
- Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J., Jr., Wasney, G. A., Yeung, R., Arrowsmith, C., Ball, L. J., Berglund, H., Hui, R., Marsden, B. D., Nordlund, P., Sundstrom, M., Weigelt, J., et al. (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. U. S. A.* 103, 15835–15840.
- Wright, T. A., Stewart, J. M., Page, R. C., and Konkolewicz, D. (2017) Extraction of thermodynamic parameters of protein unfolding using parallelized differential scanning fluorimetry. *J. Phys. Chem. Lett.* 8, 553–558.
- Kopeck, J., and Schneider, G. (2011) Comparison of fluorescence and light scattering based methods to assess formation and stability of protein-protein complexes. *J. Struct. Biol.* 175, 216–223.
- Matulis, D., Kranz, J. K., Salemme, F. R., and Todd, M. J. (2005) Thermodynamic stability of carbonic anhydrase: Measurements of

Catalytic residues and conformational change in CO₂ fixation

- binding affinity and stoichiometry using thermofluor. *Biochemistry* 44, 5258–5266
25. Guamera, E., Tan, Z. W., Zheng, Z., and Berezovsky, I. N. (2017) Allo-SigMA: Allosteric signaling and mutation analysis server. *Bioinformatics* 33, 3996–3998
26. Guamera, E., and Berezovsky, I. N. (2016) Structure-based statistical mechanical model accounts for the causality and energetics of allosteric communication. *PLoS Comput. Biol.* 12, 1–27
27. Kurochkin, I. V., Guarnera, E., Wong, J. H., Eisenhaber, F., and Berezovsky, I. N. (2017) Toward allosterically increased catalytic activity of insulin-degrading enzyme against amyloid peptides. *Biochemistry* 56, 228–239
28. Swain, J. F., and Gierasch, L. M. (2006) The changing landscape of protein allostery. *Curr. Opin. Struct. Biol.* 16, 102–108
29. Gunasekaran, K., Ma, B., and Nussinov, R. (2004) Is allostery an intrinsic property of all dynamic proteins? *Proteins* 57, 433–443
30. Tee, W. V., Guarnera, E., and Berezovsky, I. N. (2018) Reversing allosteric communication: From detecting allosteric sites to inducing and tuning targeted allosteric response. *PLoS Comput. Biol.* 14, 1–26
31. Tan, Z. W., Guamera, E., Tee, W. V., and Berezovsky, I. N. (2020) Allo-SigMA 2: Paving the way to designing allosteric effectors and to exploring allosteric effects of mutations. *Nucleic Acids Res.* 48, W116–W124
32. Motlagh, H. N., Wrabl, J. O., Li, J., and Hilser, V. J. (2014) The ensemble nature of allostery. *Nature* 508, 331–339
33. Pai, E. F., and Schulz, G. E. (1983) The catalytic mechanism of glutathione reductase as derived from x-ray diffraction analyses of reaction intermediates. *J. Biol. Chem.* 258, 1752–1757
34. Yang, Z. Y., Ledbetter, R., Shaw, S., Pence, N., Tokmina-Lukaszewska, M., Eilers, B., Guo, Q., Pokhrel, N., Cash, V. L., Dean, D. R., Antony, E., Peters, J. W., and Seefeldt, L. C. (2016) Evidence that the Pi release event is the rate-limiting step in the nitrogenase catalytic cycle. *Biochemistry* 55, 3625–3635

CHAPTER THREE

MESSAGE RECEIVED: LONG RANGE ALLOSTERIC COMMUNICATION IS ESSENTIAL
IN A MICROBIAL CARBOXYLASEContribution of Authors and Co-Authors

Manuscript in Chapter 3

Author: Jenna R. Mattice

Contributions: Developed and performed intact and peptide-level HDX experiments. Wrote and edited the manuscript.

Co-Author: Krista A. Shisler

Contributions: Developed and performed the EPR and protein assay experiments. Reviewed and edited the manuscript.

Co-Author: J. R. Malone

Contributions: Assisted in performing the intact and peptide-level HDX experiments.

Co-Author: N.A Murray

Contributions: Assisted in performing the intact and peptide-level HDX experiments.

Co-Author: Monika Tokimina-Lukaszewska

Contributions: Contributed to the understanding and application of the experiments. Reviewed and edited the manuscript.

Co-Author: Jennifer L. DuBois

Contributions: Supervised experiments. Reviewed and edited the manuscript.

Co-Author: John W. Peters

Contributions: Supervised experiments. Reviewed and edited the manuscript.

Co-Author: Brian Bothner

Contributions: Supervised experiments. Reviewed and edited the manuscript.

Manuscript Information

Jenna R. Mattice, Krista A. Shisler, J.R. Malone, N.A. Murray, Monika Tokmina-Lukaszewska,

Jennifer L DuBois, John W. Peters

Journal of Biological Chemistry

Status of Manuscript: [Put an x in one of the options below, delete this]

☒ Prepared for submission to a peer-reviewed journal

☐ Officially submitted to a peer-reviewed journal

☐ Accepted by a peer-reviewed journal

☐ Published in a peer-reviewed journal

Abstract

Acetone carboxylase (AC) is a non-canonical microbial carboxylase. In contrast to other carboxylases, it uses one ATP to phosphorylate acetone and bicarbonate to generate acetoacetate. Phosphorylation of two intermediates from a single molecule of ATP and the absence of other cofactors make this a unique reaction. Acetone carboxylase is a large (360 KDa) dimer of trimers with two active sites. Structural models of AC with and without ligands show that substrate binding and phosphorylation take place ~ 40 Å from the Mn(II) active site where acetoacetate is formed. Based on the crystal structures, it was suggested that a significant conformational change opens and closes a tunnel that facilitates passage of reaction intermediates between the binding and active sites. The goal of this study is to elucidate the conformational change and allosteric communication that such a mechanism would require. In this effort we have employed a suite of biophysical techniques, including electro paramagnetic resonance (EPR), kinetic assays, and hydrogen/deuterium exchange mass spectrometry (HDX-MS). Data from poised ligand bound states and site-specific amino acid variants indicate that the active site E89 residue acts as a molecular relay, having a direct role in catalysis and allosteric communication that is essential for catalysis.

Introduction

Atmospheric carbon dioxide is a source for building biomass in plants and microbes. Carboxylases are enzymes that catalyze fixing CO₂ into organic substrates.¹ The canonical carboxylases use biotin, or another cofactor, to derivatize their substrates and are well characterized. Acetone carboxylase is a novel carboxylase that does not share sequence identity with other carboxylase. *X. autotrophicus* acetone carboxylase (AC) uses a single ATP to phosphorylate acetone and bicarbonate to form acetoacetate.² The reaction is unique in that biotin or other cofactors are not employed. The proposed reaction mechanism involves the binding of acetone and bicarbonate which are then phosphorylated and travel ~40 Å from the substrate binding site to an active site where they coordinate Mn(II) and acetoacetate is formed (Figure 1).

AC is a 360 kDa heterohexameric complex made up of alpha (75 kDa), beta (85 kDa), and gamma (20 kDa) subunits, in a ($\alpha\beta\gamma$)₂ fashion (Figure 1). Structural models are available for AC trapped in multiple steps in the catalytic cycle: ligand free (PDB: 5SVC), AMP-bound (PDB: 5SVB) and AMP-acetate bound (PDB: 5M45).² These crystal structures suggest conformational change occurs when ligands are bound. The binding of ATP, acetone and bicarbonate are proposed to open up an interior channel where phosphorylated intermediates would be protected from solvent as they travel to the Mn(II) active site and form acetoacetate. The presence of AMP and acetoacetate has been proposed to induce a product release conformation of the protein. Previous examination of the structural models suggests that E89 in the alpha subunit has an allosteric communication pathway between the substrate binding site. E89 in the alpha subunit changes coordination to the Mn metal when different ligands are bound in the beta subunit about

~40 Å away in the crystal structures (Figure 1). The mechanism of communication linking these two sites is unknown.

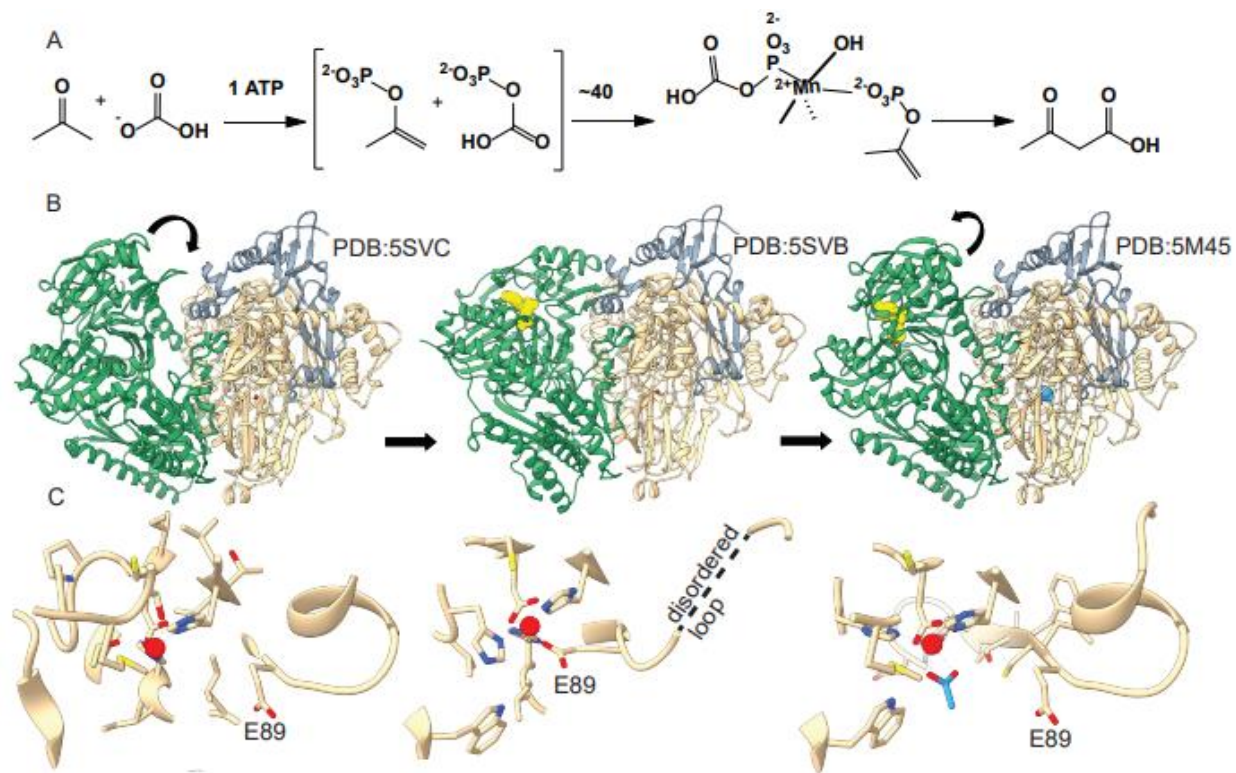


Figure 1. Overview of acetone carboxylase mechanism and structure. A) Truncated mechanism of acetone carboxylase showing acetone and bicarbonate becoming acetoacetate. B) Half of the dimer of acetone carboxylase, beta subunit (green), alpha subunit (tan) and gamma (grey). Undergoing the proposed conformational change during reaction. C) Detailed view of E89 changing coordination during reaction. The alpha helix associated with E89 becoming a disordered loop when coordinated to Mn(II) and reforming upon acetate coordination.

This work aims to elucidate key questions raised from the crystal structures. We hypothesize that ligand binding is essential for catalysis and that there is communication between the ligand binding site and the Mn(II) active site. To explore this, we employ a host of techniques such as electron paramagnetic resonance, protein assays, and hydrogen/deuterium

exchange coupled to mass spectrometry. Experimental conditions include different ligand-bound states. Mutagenesis of the E89 residue was also explored using variants E89D and E89A. E89D has one carbon shorter in the side chain and may be able to coordinate Mn(II), while E89A should not be able to coordinate Mn(II). These variants will help us elucidate the importance of E89 during catalysis.

Results

E89 is critical for Mn(II) coordination.

It was proposed that ATP/AMP binding is communicated to the Mn(II) site via a series of structural changes that result in rotation of the E89 side chain into bidentate coordination and concomitant loss of alpha helical structure in nearby residues.² To investigate this, we began using Electron Paramagnetic Resonance (EPR). EPR is a technique for studying unpaired electron spins and is used to provide information on the structure and bonding of a molecule. Unpaired electrons, like those found in transition metals, have both orbital and angular spin momentum which will generate a magnetic dipole moment.³ In a basic EPR experiment this is achieved by irradiating the sample with a set frequency of microwaves and sweeping a constant magnetic field. The electrons will align (parallel or antiparallel) to the magnetic field and have a specific energy which is proportional to the magnetic field's strength and the first derivative of the absorption of electromagnetic radiation will be reported. When transition metals with unpaired electrons are exposed to an external magnetic field, this can be exploited to determine the metal's identity, its oxidation state, and local symmetry. Furthermore, the types of nuclei interacting with the metal can be determined. In studying metalloproteins, EPR is advantageous because there is no limit to the protein size that can be probed, and many different paramagnetic

metal centers can be poised in varying oxidation states to provide information on substrate binding and catalysis. AC is large protein with an Mn(II) active site that contains five unpaired electrons, which can be investigated by EPR to provide information on the Mn(II) coordination inside the protein environment. Analysis of AC in the absence of AMP, revealed a 6-line Mn(II) EPR spectra centered at $g=2$ (Figure 2, Supplemental Figure 1).⁴ A 6-line EPR spectra coincides with Mn(II) based on the equation $2NI+1 = \text{number of lines in EPR spectra}$ where N is number of nuclei ($N=1$) and I ($I=5/2$) is electron spin. With the addition of Mg-AMP, more defined hyperfine coupling of a $5/2$ Mn(II) center is observed. This has been previously observed and is consistent with coordination of Mn by E89 as observed in the structural model.^{4,5}

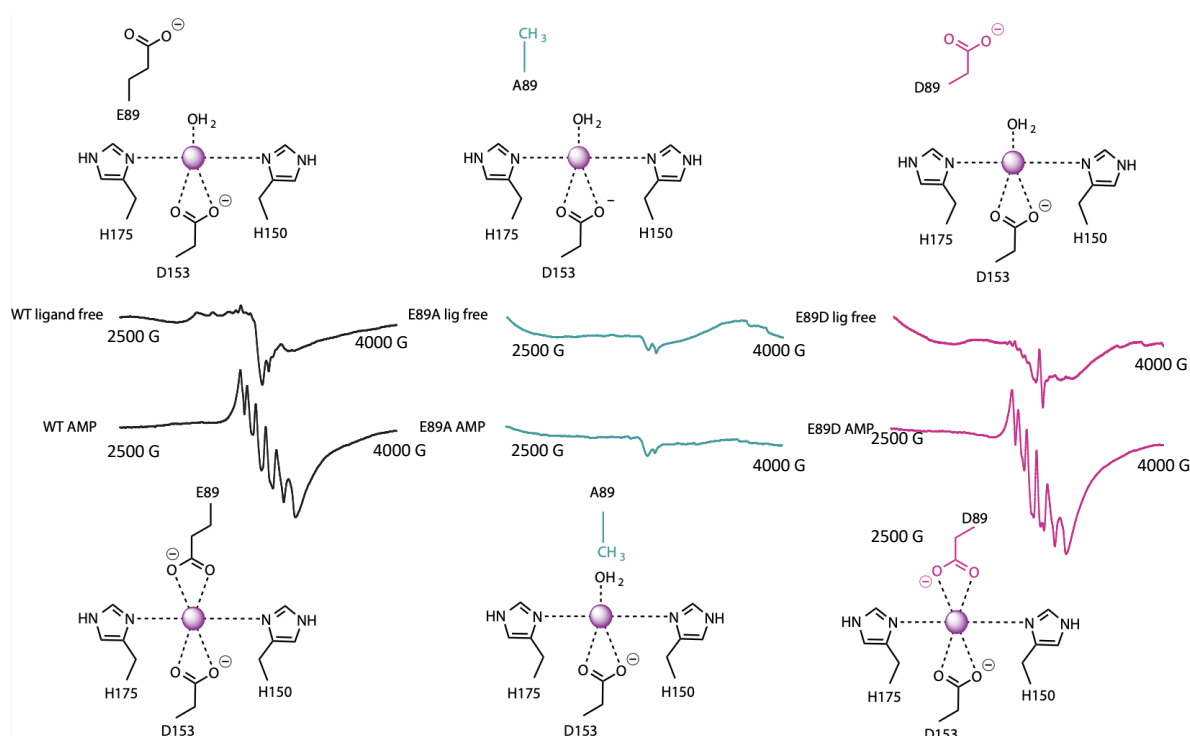


Figure 2. EPR reveals no change in Mn(II) signal in E89A when AMP is bound. AC contains a Mn(II) signal at $g=2$. With the addition of Mg-AMP, more defined hyperfine coupling of the $5/2$ Mn(II) center is observed, indicating coordination of the E89 residue as observed in the crystal structure. In the E89A variant, no spectral changes are observed with the addition of Mg-AMP due to the lack of a sixth coordinating ligand. In E89D the more defined hyperfine coupling occurs indicating coordination to the Mn(II).

In this study and previous studies AMP is used to explore the conformational change in AC caused by nucleotide binding. A change in the signal is observed in the E89D variant upon addition of Mg-AMP similar to WT. This suggests that the one carbon shorter side chain of aspartate can still coordinate with the Mn site. We next tested the E89A variant, where no change in the EPR signal was observed upon addition of Mg-AMP, indicating a sixth coordinate of Mn was not available. The EPR data confirm that in WT AC the binding of AMP causes Mn(II) to change coordination. Change in coordination is also observed in the E89D variant, but not in E89A variant. This suggests that AMP binding is communicated to the Mn(II) site via the residue E89.

E89A exhibited no acetoacetate production.

To probe the role of E89 in the catalytic cycle, ATP hydrolysis and the production of ADP, AMP, and acetoacetate were monitored in WT, E89D, and E89A. Acetoacetate production was measured using a coupled assay with β -hydroxybutyrate dehydrogenase (β -HBDH) which converts acetoacetate to 3-hydroxybutanoate in the reverse reaction upon the oxidation of NADH.⁵ Parallel assays were conducted to analyze acetoacetate production alongside ATP hydrolysis and ADP and AMP production. In WT, the rate of ATP consumption is 76.3 $\mu\text{M}/\text{min}$ and the acetoacetate formation is 33.5 $\mu\text{M}/\text{min}$. The E89D was slightly more active, with a rate of ATP consumption of 89.2 $\mu\text{M}/\text{min}$ and acetoacetate formation of 41.6 $\mu\text{M}/\text{min}$. The E89A variant does not produce acetoacetate. WT and both variants have non-productive ATPase activity, hydrolyzing ATP even in the absence of acetone (Table 1). The inactivity in the E89A variant suggests the inability of the residue to coordinate the Mn(II) during reaction is important

for the carboxylation activity of the enzyme. From these data we then sought to elucidate the mode of communication between the ligand binding's site and the E89.

	ATP ($\mu\text{M}/\text{min}$)	ADP ($\mu\text{M}/\text{min}$)	AMP ($\mu\text{M}/\text{min}$)	Acetoacetate ($\mu\text{M}/\text{min}$)	ATP/Acetoacetate ratio
WT	76.3 ± 9.9	23.7 ± 4.5	67.6 ± 9	33.5 ± 1.5	0.8
WT no acetone	49.3 ± 2.2	15.4 ± 3.7	21.5 ± 8.8	1.3 ± 0.2	
E89A	49.6 ± 5.1	12.1 ± 1.1	51.9 ± 2.6	0.5 ± 0.1	1.4
E89A no acetone	50.3 ± 15.6	11.5 ± 1.3	43.5 ± 1.8	0.2 ± 0.2	
E89D	89.2 ± 2.3	25 ± 6.7	86.5 ± 4.1	41.6 ± 1.0	1.1
E89D no acetone	42.3 ± 25.3	14 ± 0.4	41.7 ± 0.7	0.3 ± 0.1	

Table 1. Rate of ATP consumption and rates of acetoacetate, ADP, and AMP production are shown for the WT, E89A, E89D variant with and without acetone.

Ligand binding alters the hydrogen bonding network of AC.

Hydrogen-deuterium exchange (HDX) can provide insights into the effects of ligand binding and can be used investigate protein conformational change.⁶ Amide backbone hydrogens exchange readily with deuterium. By exposing WT AC to a D₂O-containing buffer overtime (1, 8, 20, 60, 180, 1440 minutes) the exchange can be exploited to determine changes in the hydrogen bonding network due to ligand binding (AMP or AMP-acetate). By comparing the deuterium uptake under these different binding conditions, we can ascertain the global disruptions in the hydrogen bonding network and therefore infer changes in the stability, solvent

accessibility, and dynamics or conformational change of a protein.^{6,7} We have employed this technique to probe the changes in WT upon ligand binding.

Deuterium uptake was measured at multiple timepoints in WT AC in unbound, AMP-bound, and AMP-acetate bound conditions to examine the hydrogen bonding network in the three subunits (alpha, beta, gamma). With AMP simulating the ATP/AMP structure where E89 should coordinate the Mn(II) and AMP-acetate simulating a product release structure where E89 is no longer coordinated to the Mn(II). There were large differences in deuterium uptake for the beta-subunit while the alpha and gamma subunits did not show distinct changes upon ligand binding (Figure 3).

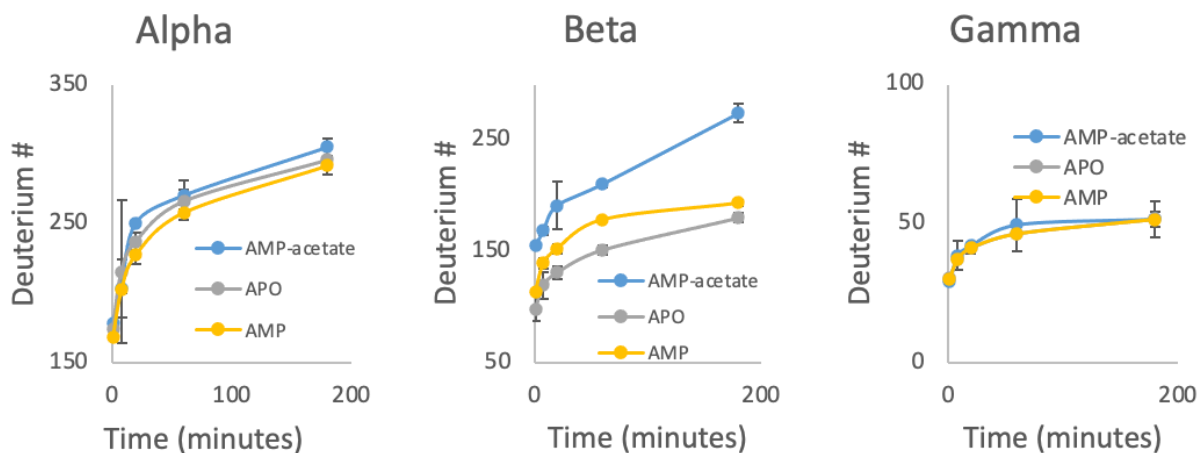


Figure 3. Intact hydrogen-deuterium exchange highlights influence ligand binding has on hydrogen bonding network in WT. Deuterium uptake curves over time shown for all subunits. In the Beta subunit more uptake is seen in the AMP and AMP-acetate conditions compared to APO. While the alpha and gamma do not see changed in deuterium uptake based upon binding.

In the beta-subunit, the addition of AMP lead to an increase in deuterium uptake across all timepoints compared to the unbound (APO). While the AMP-acetate condition lead to a greater increase in deuterium uptake than AMP alone. The difference in deuterium uptake between AMP

and AMP-acetate support that the addition of acetate leads to a unique conformation which may be the product release structure. Furthermore, the changes in deuterium uptake being predominately in the beta subunit support the proposed rearrangement upon AMP binding to be predominantly in the beta subunit, where binding occurs. However, the acetate binds in the alpha subunit and no change in deuterium is observed. Perhaps the rearrangement in binding at the alpha subunit produces a net zero deuterium incorporation (ie the areas that undergo exposure are cancelled out by areas that are now undergoing protection). While the beta subunit appears to have more areas that undergo exposure under ligand binding.

Communication pathway from substrate binding to active site in WT AC

To provide information on specific regions of AC that have altered hydrogen bonding stability, pepsin digestion was added to the experiment. We can then infer changes in solvent accessibility and/or conformational change of specific peptides that are correlated with catalysis. Using the same conditions (APO, AMP, AMP-Acetate) with a slightly altered timescale (0.5, 3, 30, 180, 1440 minutes) from the intact HDX experiment we recorded peptide-level HDX of the WT enzyme.

Pepsin-based peptide mapping of WT AC yielded sequence coverage of the 95 % for the alpha subunit, 96 % for beta, and 85% for gamma (Supplemental 1). In the HDX experiments, we focused on 33 peptides from critical locations such as the interior channel from the substrate binding site to the Mn(II) active site (Supplemental 2). Interestingly, in the alpha subunit more exchange was observed in peptides that line the interior channel when AMP bound than AMP-acetate (Figure 4).

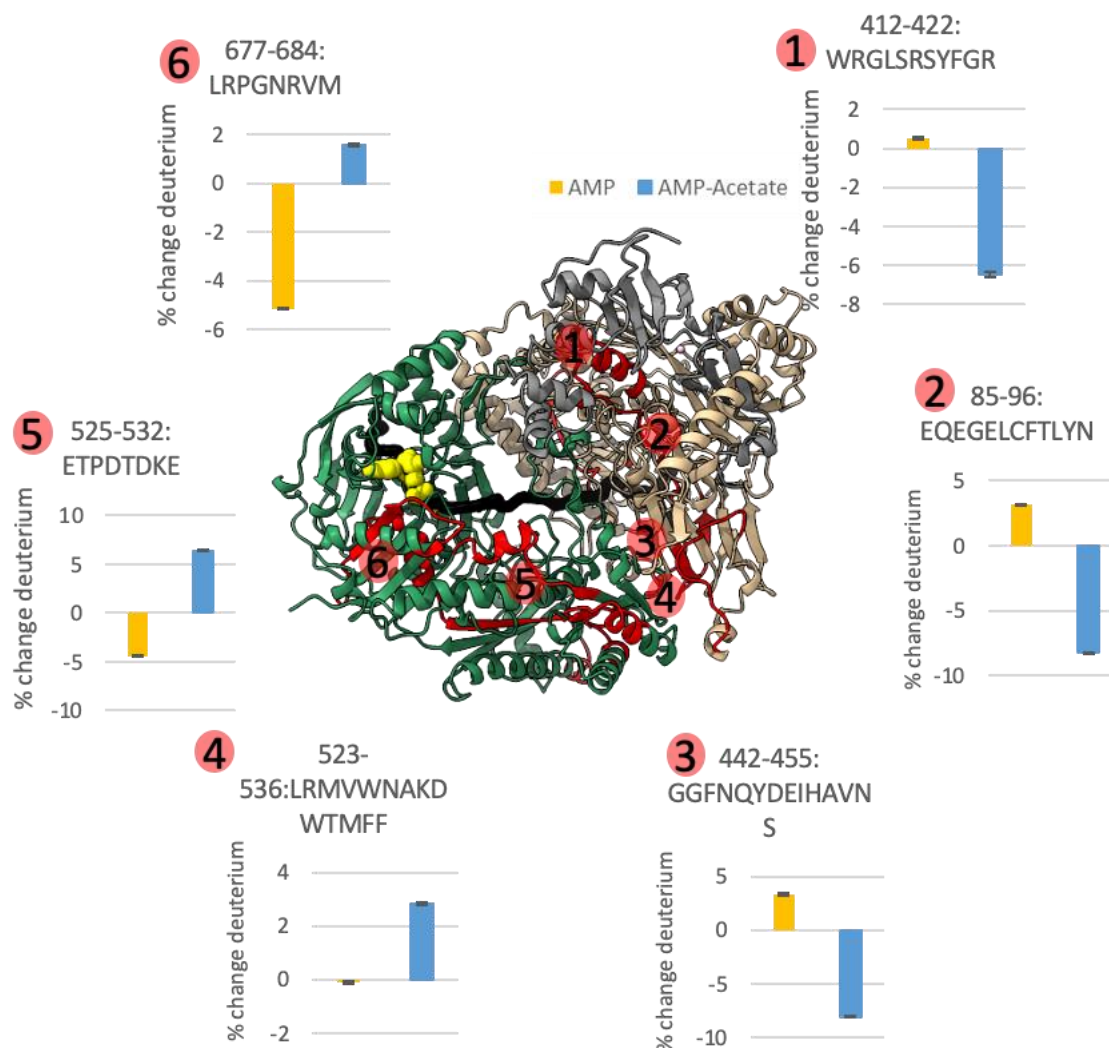


Figure 4. Peptide-level HDX exchange shows specific changes upon ligand binding. One half of the AC dimer is shown with the beta subunit (green), alpha subunit (tan) and gamma (gray). A proposed interior channel (black line) in the enzyme which leads from AMP(yellow) binding site to the Mn(II) active site. Histograms of % change deuterium uptake of AMP and AMP-acetate compared to APO are shown of key peptides which line the channel. Peptides shown on the structure in red.

We were particularly interested in peptide 85-96 which contains E89 and undergoes a helix to coil transition upon AMP binding that allows E89 to coordinate of Mn(II). This peptide shows significant differences in uptake upon ligand binding with AMP-bound having an increase

in deuterium uptake compared to APO, while AMP-acetate bound showed a decrease in uptake compared to APO. The uptake patterns coincide with the proposed changes in the AMP bound structure compared to the unbound. With the alpha helix of the E89 becoming disordered upon coordinating the Mn(II), resulting in higher exchange. While in the AMP-acetate bound enzyme the E89 is not coordinated to the Mn(II) and the alpha-helix is intact, resulting in lower exchange. Other peptides analyzed support AMP and AMP-acetate inducing unique conformational changes. With the peptides exhibiting opposite deuterium uptake trends when bound. Comparing the unbound (APO) to the AMP and AMP-acetate bound conditions we have successfully identified areas where there are changes in the deuterium exchange and infer an interior channel in the beta- and alpha- subunit upon AMP binding. By mapping these areas on the structure, we can propose where an interior channel opens and allows for intermediates to pass through. We propose that these residues are not only involved in conformational change but also for the communication between binding sites.

E89A disrupts the hydrogen-bonding network of AC

The E89 residue in the alpha subunit had been shown to be key for catalysis, perhaps through Mn(II) coordination. The variant E89A being unable to undergo Mn(II) coordination is the probe necessary to further elucidate the proposed conformational change and communication essential for catalysis. In order to do this, we repeated the intact protein and peptide-level HDX in the same manner as WT AC. Data from intact HDX showed different patterns E89A and WT enzyme (Figure 5).

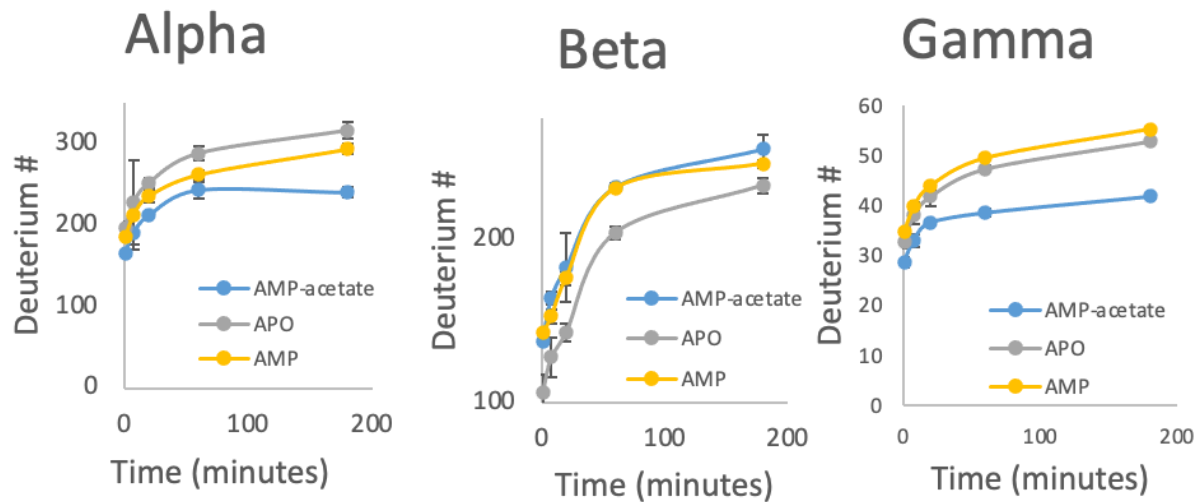


Figure 5. Intact hydrogen-deuterium exchange highlights influence on hydrogen bonding network ligand binding and mutation has on E89A. In the Beta subunit more uptake is seen in the amp and amp-acetate conditions compared to apo, as seen in WT. However, the alpha and gamma subunits are now also exhibiting changes due to ligand binding, which was not seen in WT.

In all subunits there are measurable changes in deuterium incorporation based on ligand binding. In the beta subunit the AMP and AMP-acetate conditions exhibited an increase in uptake similar to each other, while in WT they showed different deuterium uptake patterns. In the gamma subunit the AMP-acetate condition exhibited less deuterium uptake than the apo and AMP conditions, while no change was seen in WT. In the alpha subunit, the bound conditions exhibited the more uptake compared to unbound, this is in contrast to WT where no change was seen. The changes seen in the beta subunit are similar to WT, except the AMP-acetate not incorporating more than AMP alone. The alpha and gamma disruptions are distinctive to the E89A variation and suggest the E89A is disrupting the hydrogen bonding network of the protein complex on its own.

To elucidate the specific protein regions responsible for the distinctive changes in deuterium uptake caused by ligand binding and the E89A substitution, peptide-level HDX was used. We focused our analysis on peptides that are part of the interior channel (Figure 6).

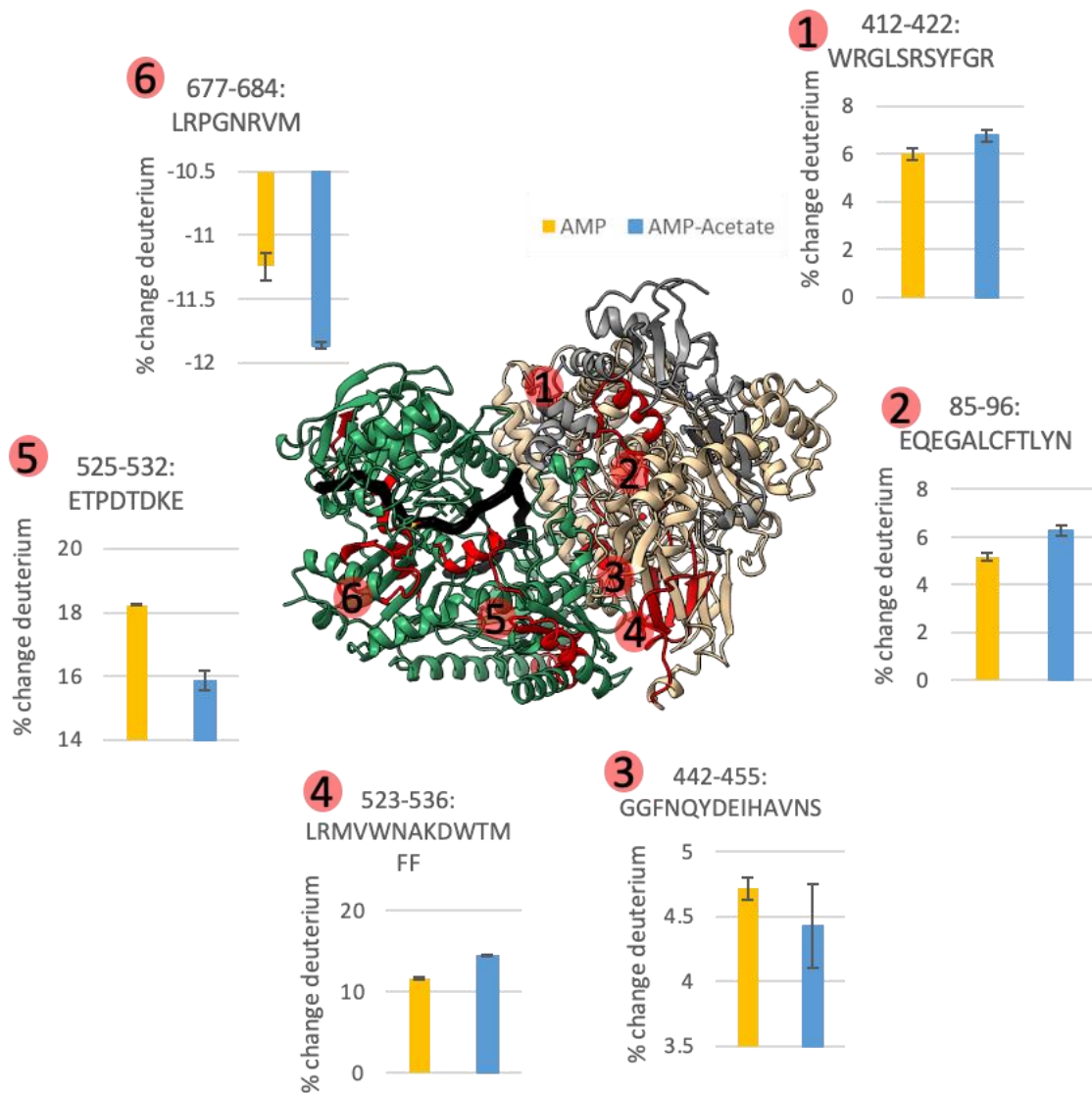


Figure 6. Peptide-level HDX exchange highlight exact areas of change in deuterium uptake due to ligand binding in WT. One half of the dimer is shown with the beta subunit (green), alpha subunit (tan) and gamma (gray). A proposed disrupted interior channel (black line) in the enzyme. Histograms of % change deuterium uptake of AMP and AMP-acetate compared to APO are shown of key peptides which line the channel. Peptides shown on the structure in red.

In the beta subunit of E89A the deuterium uptake of AMP and AMP-acetate bound forms exhibited similar patterns of exchange than the WT bound conditions. In the alpha subunit of E89A this also is the case with both conditions exhibiting the same trends and overall uptake. In the E89A containing peptide 85-96, both AMP and AMP-acetate exhibited more uptake than APO. This is in direct contrast to the WT behavior which had AMP-acetate exhibited significantly decreased uptake than APO while AMP alone showed more uptake than APO. These data is consistent with the E89A not being able to coordinate Mn(II), with the helix if E89A unable to undergo conformational change. The E89A substitution alters the active site environment as well the hydrogen bonding network around the interior substrate channel.

Discussion

AC is a large non-canonical carboxylase that manages to phosphorylate acetone and bicarbonate in the formation of acetoacetate from a single molecule of ATP. This is done without the use of biotin or other cofactors that canonical carboxylases employ. Further, previous studies have shown AC to be a large complex protein with the ligand binding and phosphorylation take place ~ 40 Å from the Mn(II) active site. Conformational change was hypothesized to be important for catalysis. We also hypothesized that communication between these two sites must be essential for modulating the conformational change. In this work we employed a variety of techniques to elucidate the necessary conformations for catalysis.

All the data taken together elucidate the interior channel in the protein which is dependent on ATP/AMP binding and the E89 residue coordinating the Mn(II) (Figure 7).

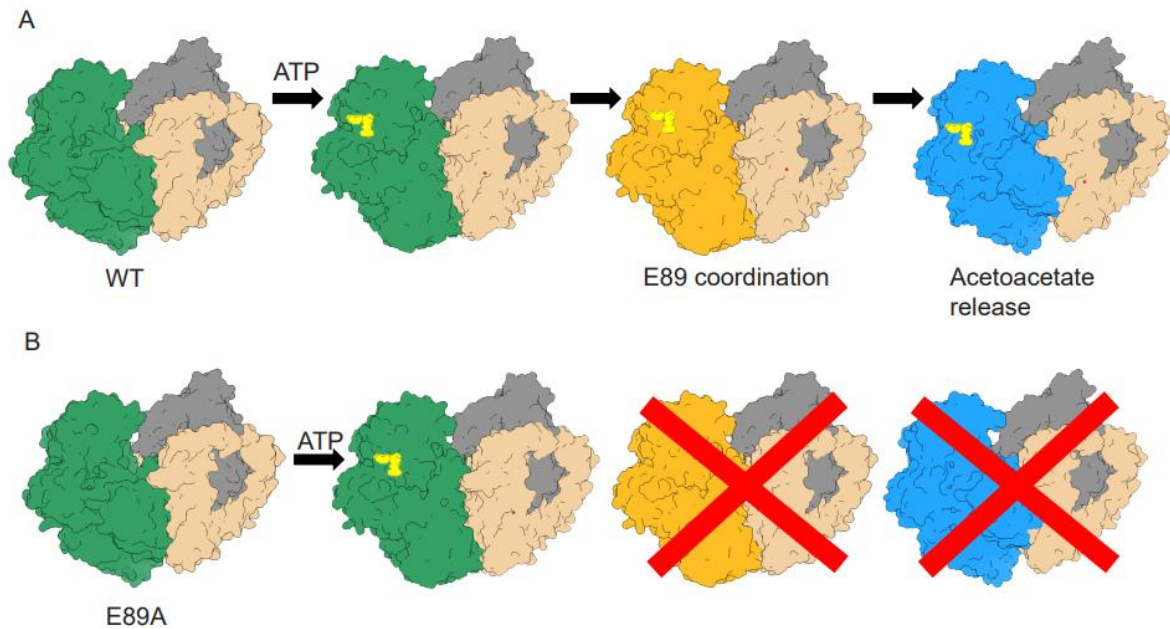


Figure 7. Overview of AC structural changed during reaction for WT and E89A. A) Half the dimer of WT showing illustrating the structural changed during reaction. Where nucleotide binding (yellow), where then allosteric communication occurs from that binding site to the E89 residue where then acetoacetate is formed. B) The E89A half of the dimer is shown where nucleotide binding (yellow) invokes a small change but allosteric communication is disrupted to the E89 residue and therefore the reaction is disrupted.

EPR data highlighting changes in the hyperfine spectra of the Mn(II) caused by AMP binding. This is evidence of the communication between the ligand binding site and the Mn(II) active site where the E89 coordinates the Mn(II) upon ligand binding. The E89D variant exhibited changes in the EPR spectra upon AMP binding similar to WT while the E89A did not. This suggests the E89 residue is part of the communication between the two sites. The rates of catalysis were measured for the WT and variants revealing the E89A variant is inactive and does not produce acetoacetate. This implies the ability to coordinate Mn(II) is essential for catalysis, perhaps by inducing conformational change.

To investigate ligand induced conformational change in WT AC we employed intact protein HDX. These data exhibited large changes in deuterium uptake in the beta subunit upon AMP-binding. This was congruent with a large rearrangement in the beta subunit when ligands bind, which opens an interior channel to the Mn(II) active site. Furthermore, the intact HDX of the AMP-acetate bound AC produced different uptake patterns suggesting a distinctive product release conformation which collapses the interior channel. We next performed peptide-level HDX to provide amino acid level information about the interior channel to the Mn(II) active site. The AMP and AMP-acetate conditions exhibited opposite deuterium uptake trends along the interior channel which runs from the substrate binding site to the Mn(II) active site. This highlights the opening and closing of the channel which is essential for catalysis. The peptide 85-96, which contains the E89 residue also exhibited deuterium uptake trends consistent with Mn(II) coordination inducing unfolding of its alpha helix upon AMP binding, and reforming upon AMP-acetate binding. From this we propose AMP and AMP-acetate are exhibiting different conformations which are both essential for catalysis.

Taken a step further, the E89A variant produced distinctive conformations of the protein. With the EPR showing the E89A cannot coordinate the Mn(II) and so therefore unable to generate the interior channel. We investigated the E89A variant using intact HDX. This exposed that the E89A variant disrupted the hydrogen bonding network of AC and exhibited distinct uptake patterns when ligands were bound. We did not anticipate this large of disruption of the hydrogen bonding network based on a discrete mutation. This suggests E89A may exhibit a change to its conformational ensembles from mutation alone. To probe the available conformations, we then performed peptide-level HDX. This provided evidence of E89A

undergoing conformational change that can bind AMP but is not conducive for opening an interior channel and therefore acetoacetate production. E89 is therefore essential for allosteric communication in the protein complex and carboxylation.

Materials and Methods

Expression and purification of recombinant acetone carboxylase

WT, E89A and E89D AC from *X. autotrophicus* were expressed as previously described (Mus Sci Rep 2017) and purified as previously described (Mus Sci Rep 2017) with slight modification. The *E. coli* cells were lysed by sonification in a 25 mM Tris, pH 7.6, 0.1 mM EDTA, and 20% glycerol buffer. The lysis was applied to a His-NTA column (Bio-Rad) and eluted with a linear gradient with an imidazole buffer (25 mM Tris, pH 7.6, 0.1 mM EDTA, 20% glycerol, and 400 mM imidazole). Fractions containing AC were pooled and buffer exchanged into a 20 mM Tris, pH 7.6 buffer and applied to a Q-sepharose (Bio-Rad) column. The protein was eluted with a linear gradient with an NaCl buffer (20 mM Tris, pH 7.6, 800 mM NaCl). Fractions were pooled and buffer exchanged into a 50 mM Tris, pH 7.6, 150 mM NaCl, 10% glycerol storage buffer and stored at -80°C. Concentrations were determined by Bradford assays.

EPR spectroscopy

EPR samples of as-purified WT, E89A, and E89D *X. autotrophicus* ACs were diluted to 36 mg/mL in 50 mM Tris, pH 7.6, 150 mM NaCl buffer. For AMP containing samples, 10 mM AMP and 1 mM MgCl₂ were added. For metal-depleted samples, 900 µL aliquots of EDTA/TEA treated AC were removed (after buffer exchanging into AMP-free buffer to remove excess EDTA/TEA) and concentrated to 250 µL. Final concentrations were 27 mg/mL (X.

autotrophicus) 73 mM protein. For MnCl_2 reconstituted samples, 900 μL aliquots of MnCl_2 treated AC were removed and buffer exchanged against metal-free 50 mM Hepes, pH 7.6, 150 NaCl with or without 1 mM AMP and concentrated to 250 μL . Final concentrations for *X. autotrophicus* AC were 20 mg/mL (AMP present) and 21 mg/mL (AMP-free), or 55 and 57 mM protein, respectively.

Samples were loaded into quartz EPR tubes and immediately flash frozen in liquid nitrogen. X-band EPR spectra were measured using a Bruker EMX continuous-wave (CW) spectrometer equipped with a continuous helium flow Oxford cryostat set to 12K. Normal EPR parameters were: modulation amplitude 10 G, modulation frequency 100 kHz, 2 mW power, and microwave frequency 9.38 GHz. Spectra for the Mn(II)-bound AC samples were re-measured under identical conditions but 10 μW power, which we had previously determined was well below the power at which spectral intensity saturated.

Enzymatic assays

As-purified WT, E89A, and E89D AC was assayed for acetoacetate production via a couple assay with β -hydroxybutyrate dehydrogenase (β -HBDH) which converts acetoacetate to 3-hydroxybutanoate in the reverse reaction upon the oxidation of NADH, as previously described (Sluis PNAS, 1997, 94, 8456-8461). The rate of acetoacetate is proportional to the rate of NADH oxidization which is observed at 340 nm. Parallel assays were conducted to analyze acetoacetate production alongside ATP hydrolysis and ADP and AMP production. In a rubber-stoppered UV-vis quartz cuvette, the reaction components were added to a final concentration of 100 mM Tris, pH 7.6, 80 mM KCl, 54 mM KHCO_3 , 1 mM MgCl_2 , 343 μM NADH, 2 mM acetone, 0.25 mg/mL AC and 0.05 mg/mL β -HBDH and incubated at 37°C for 4 minutes prior to initiation by

the addition of 2 mM ATP. Once ATP was added, the cuvette was inverted to mix and a 200 μ L aliquot was removed and quenched with 10 mM EDTA (final concentration), pH 8.0 while the cuvette was inserted into a Cary 6000i for UV-vis analysis at 340 nm. The reaction was allowed to proceed for 1.5 minutes, upon which another 200 μ L aliquot was removed and quench with 10 mM EDTA. The 200 μ L aliquots were placed in a 100 kDa MWCO spin concentrator (pre-chilled on ice) and spun at 14,000xg to separate protein from reaction products. The flow-through was then injected onto an HPLC (Agilent) and run with a 3% methanol (buffer B) and 97% 20 mM ammonium acetate, pH 4.5 buffer (buffer A) method for 15 minutes to separate the peaks. ATP, ADP, and AMP eluted at 2.2, 2.9, and 5.8 min, respectively. ATP, ADP, and AMP were quantified using a standard curve and integration of the peaks. Trials were done in triplicate.

Intact HDX

Stock solutions of WT AC (44.8 mg/mL) and E89A AC (41.7 mg/mL) were mixed with MgCl_2 (1 mM) in the presence and absence of AMP (10 mM) or AMP-acetate (10 mM). Reactions were then diluted 1:10 into deuterated reaction buffer (100mM Tris, 100mM KCl, pH 7.6) Control samples were diluted into a non-deuterated reaction buffer. At each time point (1, 8, 20, 60, 180, 1440 minutes) 10 μ L of the reaction was removed and direct injected into LCMS. LC-MS analysis of proteolysis reactions was completed on a 1200 infinity HPLC (Agilent) coupled to a MicroTOF LC mass spectrometer (Bruker). Peptides were separated on a size exclusion column (Phenomenex 150 $\times$ 2mm) using a flow rate of 500 μ L/min with solvent conditions 90% solvent A [0.1% FA (Sigma) in water (Thermo) and 10% solvent B (0.1% FA in acetonitrile) (Thermo)]. Data were acquired at 2 Hz over the scan range 50 to 1700 m/z in

positive mode. Electrospray settings were as follows: nebulizer set to 3.7 bar, drying gas at 8.0 l/min, drying temperature at 350 C, and capillary voltage at 3.5 kV. Data processing was carried out with Bruker DataAnalysis software.

Peptide-level HDX

Stock solutions of WT AC (44.8 mg/mL) and E89A AC (41.7 mg/mL) were mixed with MgCl_2 (1 mM) in the presence and absence of AMP (10 mM) or AMP-acetate (10 mM). Reactions were then diluted 1:10 into deuterated reaction buffer (100mM Tris, 100mM KCl, pH 7.6). Control samples were diluted into a non-deuterated reaction buffer. At each time point (0.5, 3, 30, 180, 1440 minutes) 10 μL of the reaction was removed and quenched by adding it to 60 μL of 0.75 % formic acid (FA, Sigma) and 0.25 mg/mL porcine pepsin (Sigma) at pH 2.5 on ice. Each sample was digested for 2 min with vortexing every 30 s and then flash-frozen in liquid nitrogen. Samples were stored in liquid nitrogen until the LC-MS analysis. The LC-MS analysis of RPA was completed as described (Patterson et al 2020). Briefly, the LC-MS analysis of RPA was completed on a 1290 UPLC series chromatography stack (Agilent Technologies) coupled with a 6538 UHD Accurate-Mass QTOF LC/MS mass spectrometer (Agilent Technologies). Peptides were separated on a reverse phase column (Phenomenex Onyx Monolithic C18 column, 100 x 2 mm) at 1 C using a flow rate of 500 $\mu\text{L}/\text{min}$ under the following conditions: 1.0 min, 5% B; 1.0 to 9.0 min, 5 to 45% B; 9.0 to 11.8 min, 45 to 95% B; 11.8 to 12.0 min, 5% B; solvent A = 0.1 % FA (Sigma) in water (Thermo Fisher) and solvent B = 0.1% FA in acetonitrile (Thermo Fisher). Data were acquired at 2 Hz s^{-1} over the scan range 50 to 1700 m/z in the positive mode. Electrospray settings were as follows: the nebulizer set to 3.7 bar, drying gas at 8.0 L/min, drying temperature at 350 C, and capillary voltage at 3.5 kV. Peptides were identified as previously

described (Berry et al 2018) using MassHunter Qualitative Analysis, version 6.0 (Agilent Technologies), Peptide Analysis Worksheet (ProteoMetrics LLC), and PeptideShaker, version 1.16.42, paired with SearchGUI, version 3.3.16 (CompOmics). Deuterium uptake was determined and manually confirmed using HDExaminer, version 2.5.1 (Sierra Analytics). Heat maps were created using MSTools (Kavan et al 2011).

Acknowledgements

We would like to thank Dr. Eric Shepard for collecting and contributing to analyzing the EPR data. As well as teaching graduate students and post-doctoral researchers the fundamentals of interpreting EPR spectra, collecting the data, and analyzing and plotting the data.

References

1. Erb TJ. Carboxylases in natural and synthetic microbial pathways. *Appl Environ Microbiol.* 2011. doi:10.1128/AEM.05702-11
2. Erb TJ, Zarzycki J. A short history of RubisCO: the rise and fall (?) of Nature's predominant CO₂ fixing enzyme. *Curr Opin Biotechnol.* 2018;49:100-107. doi:10.1016/j.copbio.2017.07.017
3. Durall C, Lindblad P. Mechanisms of carbon fixation and engineering for increased carbon fixation in cyanobacteria. *Algal Res.* 2015. doi:10.1016/j.algal.2015.07.002
4. Mus F, Eilers BJ, Alleman AB, et al. Structural Basis for the Mechanism of ATP-Dependent Acetone Carboxylation. *Sci Rep.* 2017. doi:10.1038/s41598-017-06973-8
5. Mus F, Wu HH, Alleman AB, et al. Insights into the unique carboxylation reactions in the metabolism of propylene and acetone. *Biochem J.* 2020;477(11):2027-2038. doi:10.1042/BCJ20200174
6. Goodey NM, Benkovic SJ. Allosteric regulation and catalysis emerge via a common route. *Nat Chem Biol.* 2008;4(8):474-482. doi:10.1038/nchembio.98
7. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? *Proteins Struct Funct Genet.* 2004;57(3):433-443. doi:10.1002/prot.20232
8. Motlagh HN, Wrabl JO, Li J, Hilser VJ. The ensemble nature of allostery. *Nature.* 2014;508(7496):331-339. doi:10.1038/nature13001
9. Boyd JM, Ellsworth H, Ensign SA. Bacterial acetone carboxylase is a manganese-dependent metalloenzyme. *J Biol Chem.* 2004. doi:10.1074/jbc.M407177200
10. Boyd JM, Ensign SA. ATP-dependent enolization of acetone by acetone carboxylase from *Rhodobacter capsulatus*. *Biochemistry.* 2005. doi:10.1021/bi050393k
11. Sluis MK, Ensign SA. Purification and characterization of acetone carboxylase from *Xanthobacter* strain Py2. *Proc Natl Acad Sci U S A.* 1997;94(16):8456-8461. doi:10.1073/pnas.94.16.8456
12. Tokmina-Lukaszewska M, Patterson A, Berry L, Scott L, Balasubramanian N, Bothner B. The role of mass spectrometry in structural studies of flavin-based electron bifurcating enzymes. *Front Microbiol.* 2018;9(JUL):1-18. doi:10.3389/fmicb.2018.01397
13. Konermann L, Pan J, Liu YH. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem Soc Rev.* 2011;40(3):1224-1234. doi:10.1039/c0cs00113a

14. McAllister RG, Konermann L. Challenges in the interpretation of protein H/D exchange data: A molecular dynamics simulation perspective. *Biochemistry*. 2015;54(16):2683-2692. doi:10.1021/acs.biochem.5b00215
15. Masson GR, Burke JE, Ahn NG, et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods*. 2019;16(7):595-602. doi:10.1038/s41592-019-0459-y
16. Gao S, Klinman JP. Functional roles of enzyme dynamics in accelerating active site chemistry: Emerging techniques and changing concepts. *Curr Opin Struct Biol*. 2022;75:102434. doi:10.1016/j.sbi.2022.102434
17. Sowole MA, Konermann L. Effects of protein-ligand interactions on hydrogen/deuterium exchange kinetics: Canonical and noncanonical scenarios. *Anal Chem*. 2014;86(13):6715-6722. doi:10.1021/ac501849n

CHAPTER FOUR

C-TERMINUS INDUCED ASYMMETRY WITHIN A RAD52 HOMODECAMER
DICTATES SINGLE-POSITION RAD51 NUCLEATION IN HOMOLOGOUS
RECOMBINATION

Contribution of Authors and Co-Authors

Manuscript in Chapter 4

Author: Jaigeeth Deveryshetty

Contributions: Collected single-particle Cryo-EM data and solved the structure. R.C. performed fluorescence and single-molecule experiments. Reviewed and edited the manuscript.

Author: Rahul Chadda

Contributions: Performed fluorescence and single-molecule experiments. Reviewed and edited the manuscript.

Co-Author: Jenna Mattice

Contributions: Performed and analyzed the HDX-MS and XL-MS experiments. Reviewed and edited the manuscript.

Co-Author: Simrithaa Karunakaran

Contributions: Assisted with cloning, protein purification, and experiments. Reviewed and edited the manuscript.

Co-Author: Michael J. Rau

Contributions: Collected single-particle Cryo-EM data and solved the structure. R.C. performed fluorescence and single-molecule experiments. Reviewed and edited the manuscript.

Co-Author: Katherine Basore

Contributions: collected single-particle Cryo-EM data and solved the structure. R.C. performed fluorescence and single-molecule experiments. Reviewed and edited the manuscript.

Co-Author: Nilisha Pokhrel

Contributions: Assisted with cloning, protein purification, and experiments. Reviewed and edited the manuscript.

Co-Author: Noah Englander

Contributions: Assisted with cloning, protein purification, and experiments. Reviewed and edited the manuscript.

Co-Author: James A.J. Fitzpatrick

Contributions: Collected single-particle Cryo-EM data and solved the structure. R.C. performed fluorescence and single-molecule experiments. Reviewed and edited the manuscript.

Co-Author: Brian Bothner

Contributions: Supervised Experiments. Reviewed and edited the manuscript.

Co-Author: Edwin Antony

Contributions: Supervised Experiments. Wrote and edited the manuscript.

Manuscript Information

Jaigeeth Deveryshetty, Rahul Chadda, Jenna Mattice, Simrithaa Karunakaran, Michael J. Rau,
Katherine Basore, Nilisha Pokhrel, Noah Englander, James A.J. Fitzpatrick, Brian Bothner and
Edwin Antony

Nature Structural and Molecular Biology

Status of Manuscript: [Put an x in one of the options below, delete this]

☐ Prepared for submission to a peer-reviewed journal

☒ Officially submitted to a peer-reviewed journal

☐ Accepted by a peer-reviewed journal

☐ Published in a peer-reviewed journal

March 2023

Abstract

Homologous recombination (HR) is a pathway for the accurate repair of double-stranded DNA breaks, which are resected to yield single-strand DNA (ssDNA). Rad52 promotes HR by facilitating formation of Rad51 nucleoprotein filaments on ssDNA. Using single-particle cryo-electron microscopy, we show that Rad52 functions as a homodecamer. The N-terminal half of Rad52 is a well-ordered ring, while the C-terminal half is disordered. An intrinsic asymmetry within Rad52 is observed, where one or a few of the C-terminal halves interacts with the ordered N-terminal ring. Within the C-terminal half, we identify two conserved charged patches that harbor the Rad51 and RPA interacting motifs. Interactions between these two charged patches promote asymmetry and regulate a two-stage ssDNA binding mechanism. Interestingly, the intrinsic asymmetry allosterically drives Rad51 binding to a single position on the Rad52 ring. We propose a Rad52 catalyzed single-position nucleation model for the formation of pre-synaptic Rad51 filaments in HR.

Introduction

Homologous recombination (HR) protects genomic integrity by repairing double-stranded breaks (DSBs).¹⁻³ Mutations to genes that encode for HR-related proteins are implicated in a plethora of human cancers and genetic disorders.^{4,5} For accurate repair through HR, information from the undamaged sister allele is used as a template. In *Saccharomyces cerevisiae*, HR progresses through a series of defined steps beginning with resection, where the DSB is processed by the Mre11-Rad50-Nbs1 (MRN) complex yielding 3' ssDNA overhangs.¹ The transiently exposed ssDNA is coated by the high-affinity ssDNA binding protein Replication

Protein A (RPA)^{6,7} which then recruits the Mec1-Ddc2 DNA-damage-sensing kinase to regulate downstream events in HR.^{8,9}

Rad51, the recombinase, catalyzes identification and pairing of the homologous sequence in the undamaged allele. Rad51 forms a helical filament on ssDNA with defined ATP-driven nucleation and filament growth processes.^{10,11} Formation of the Rad51 filament on RPA-coated ssDNA occurs during the pre-synaptic phase of HR. Since Rad51 binds to ssDNA with lower affinity ($K_D \sim 10^{-6}$ M)¹² compared to RPA ($K_D < 10^{-10}$ M)^{13,14}, mediator proteins are required to overcome the thermodynamic barrier to facilitate ssDNA handoff. Pro-HR mediators function to promote formation of Rad51 filaments whereas anti-HR mediators facilitate Rad51 disassembly.¹⁰ In yeast, Rad52 is the sole pro-HR mediator. In mammals, Rad52 and Brca2 play overlapping roles in promoting Rad51 filament formation.^{15,16} Deletion of *Rad52* in *S. cerevisiae* has severe defects in various recombination processes including Rad51-dependent allelic recombination and Rad51-independent single-strand annealing.^{3,17-21}

In terms of the mechanism of action, yeast and human Rad52 are considered to function as heptameric rings.²²⁻²⁵ However, structural studies of the C-terminal truncated human Rad52 show that the complex is an undecamer.^{26,27} Recent Cryo-EM analysis of full-length human Rad52 also behaves as an undecamer.²⁸ Thus, the functional stoichiometry of yeast Rad52 in solution is uncertain. Both DNA binding and protein-protein interactions are integral to the recombinogenic functions of Rad52. When promoting Rad51 filament formation, Rad52 requires binding to ssDNA and substrates with a 3' overhang. During strand annealing, Rad52 must simultaneously bind to two strands of ssDNA or one ssDNA and another dsDNA fragment. A new function for Rad52 in RNA-templated DNA repair was recently uncovered^{29,30} that would

require simultaneous binding to both DNA and RNA fragments.³¹ Each Rad52 subunit is composed of an ordered N-terminal and disordered C-terminal half (Fig. 1a). The primary DNA binding activity and the oligomerization regions are harbored in the N-terminal half³² whereas Rad51 and RPA binding motifs reside in the disordered C-terminal half of Rad52.³³⁻³⁸ A weaker DNA binding site has also been identified in the C-terminal region.^{39,40} Two crystal structures of the N-terminal half of human Rad52 (residues 1-212) bound to ssDNA were recently solved.⁴¹ An outer and inner DNA binding site in the ordered N-terminal half coordinates DNA binding in Rad52.^{23,42,43}

Physical interactions of Rad52 with Rad51 and RPA are essential for the mediator functions in HR, and both binding sites reside in the disordered C-terminal half of Rad52.³³⁻³⁸ The C-terminus also promotes phase separation of Rad52 into liquid droplets and cluster DNA damage sites.⁴⁴ In addition to the above mentioned roles, Rad52 regulates the extent of DNA resection during HR by regulating the Rqh1 helicase in fission yeast.⁴⁵ Furthermore, Rad52 functions in other processes such as regulating retrotransposition,⁴⁶ CRISPR-Cas9 genome editing,⁴⁷ and IgD class-switch recombination.⁴⁸ Given the emerging multifaceted functions of Rad52, knowledge of its functions and mechanism of action is critical.

We show that *S. cerevisiae* Rad52 functions as a homodecamer with each subunit possessing an ordered N-terminal and disordered C-terminal half. Upon ssDNA binding, extensive conformational changes are observed in both the ordered N-terminal and disordered C-terminal halves of Rad52. More interestingly, we find an intrinsic asymmetry in Rad52 imparted by the C-terminal half. One, or a small subset of the C-termini interacts with the N-terminal ring and dictates ssDNA wrapping and diffusion properties. The allostery also dictates Rad51 binding

to a single position in the Rad52 ring and guides Rad51 filament formation. Furthermore, interactions between Rad51 and N-terminal half are observed. Our findings reveal how the ordered decameric N-terminal ring and the disordered C-terminal region cooperate to perform the mediator role of Rad52 to facilitate Rad51 nucleation during HR.

Results

Saccharomyces cerevisiae Rad52 functions as a homodecamer

Yeast and human Rad52 are canonically considered to function as heptamers based on negative-stain transmission electron microscopy and analytical ultracentrifugation (AUC) analysis.²²⁻²⁵ However, crystal structures of the N-terminal half of human Rad52 revealed an undecamer both in the absence or presence of ssDNA^{22,23}, with ssDNA encircling the ring in a uniform manner.^{26,27,41} Thus, to establish the functional stoichiometry, we performed single-particle Cryo-EM analysis of full-length *S. cerevisiae* Rad52. Recombinantly overproduced and purified Rad52 is predominantly a single species in both mass photometry (MP) and analytical ultracentrifugation sedimentation velocity (AUC^{SV}) analyses over a range of protein concentrations tested (Fig. 1B and Supplementary Fig. 1). A small fraction (~5-10%) of a higher order species is observed and corresponds to double the mass of the major peak. The predicted mass of one Rad52 subunit is 52.4 kDa and measurements from MP and AUC yield molecular weights of ~500 kDa (~10 monomers) and ~1000 kDa (~19-20 monomers) for the two species, respectively (Supplementary Tables 1 & 2). These data suggest that yeast Rad52 functions as a homodecamer in solution.

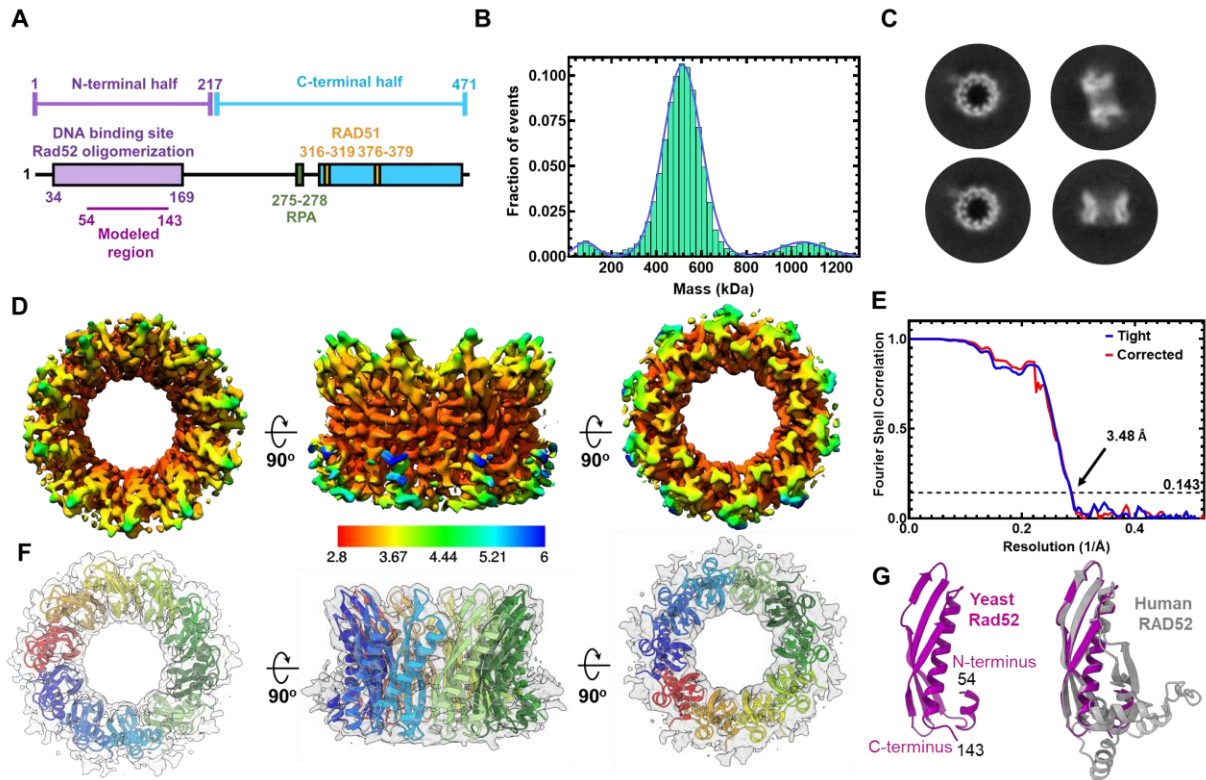


Figure 1. *S. cerevisiae* Rad52 functions as a homodecamer. A) Schematic representation of a yeast Rad52 subunit depicts N- and C-terminal regions that promote DNA binding, oligomerization, & protein-protein interactions. B) Mass photometry analysis predominantly shows a single Rad52 species. The major species (>85%) shows a mass of 498.4 ± 23.8 kDa. A minor 991.8 ± 75.34 kDa species (~8-12%) is observed along with a small fraction of free monomers (72.9 ± 13.9 kDa). C) Representative 2D classes of full-length yeast Rad52 observed in Cryo-EM analysis and used for determination of the structure. D) Color coded depiction of the local resolution mapped on to the Rad52 Cryo-EM map. E) Gold Standard Fourier shell correlation (FSC) curves estimated using Cryosparc. F) De novo built atomic model of Rad52 overlaid on the Cryo-EM map. G) Atomic model of a Rad52 subunit (magenta) aligned with the crystal structure of human RAD52 (grey; PDB 1KN0) using Chimera with Root Mean Square Deviation (RMSD) of 0.849 Å for Ca atoms of 87 amino acid pairs.

Cryo-EM structure of full-length *S. cerevisiae* Rad52

2D classification of Rad52 single particles in Cryo-EM analysis confirmed that yeast Rad52 is predominantly a homodecameric ring (Fig. 1C & Supplementary Fig. 2). Rings with 9 or 11 subunits were also rarely observed (<0.2% of all particles). 3D reconstruction and

subsequent refinement revealed strong density for the N-terminal half (Supplementary Fig. 3). However, only partial density was observed for the C-terminal half, which is consistent with the disordered nature. The final cryo-EM map reconstruction reached an average resolution of 3.5 Å with local resolution ranging from 2.9 – 6 Å (Figs. 1D-F & Supplementary Table 3). Using the refined map, a *de novo* model was built for the N-terminal region encompassing amino acids 54-143. This region of yeast Rad52 is structurally similar to human Rad52 (41% sequence identity)²⁷ with a RMSD of 0.849 Å for C α atoms of 87 amino acid pairs (Fig. 1G and Supplementary Fig. 4). However, a key difference is the absence of resolvable density for the petal-like extensions observed around the periphery of the ring structure of human Rad52 suggesting their dynamic nature in yeast Rad52 (Fig. 1G and Supplementary Fig. 4).^{26,27,41} The overall charge distribution and the interfaces that coordinate oligomerization are conserved between the two proteins (Supplementary Fig. 4).

Rad52 displays biphasic ssDNA binding activity

Next, to investigate how the Rad52 decamer binds to ssDNA, we followed the change in intrinsic tryptophan (Trp) fluorescence (Figs. 2A & B). We used poly(dT) or (dT)₈₄ to determine the occluded site-size of Rad52 which is defined as the number of ssDNA nucleotides required to saturate all the potential binding sites on the protein. Two binding phases are observed with a high affinity ($K_D = 1.3 \times 10^{-12}$ M) and relatively lower affinity ($K_D = 28 \times 10^{-9}$ M) site saturating at 55 $\pm$ 3 nt and 158 $\pm$ 3 nt per Rad52 decamer, respectively (Fig. 2A & B, Supplementary Fig. 5A & B). In the DNA-bound crystal structures of human Rad52⁴¹, depending on the mutations introduced to promote crystallization, an inner binding site was observed where a 40 nt ssDNA oligonucleotide wraps around the undecameric ring (~ 3.6 nt/subunit; Supplementary Fig. 6A)⁴¹.

An outer binding site was also captured with 6 nt of ssDNA sandwiched between two Rad52 undecameric rings (Supplementary Fig. 6B)⁴¹. In comparison, our data reveals that for yeast Rad52 ~5.5 nt/subunit and ~15 nt/subunit are required to saturate the high and low affinity binding sites, respectively. These data suggest that even though the Rad52 decamer contains one subunit less as compared to human Rad52 the high affinity binding site is larger: 5.5 nucleotides/subunit as compared to 3-4 nucleotides/subunit indicating that the ssDNA wrapping by yeast Rad52 may not be uniform as observed for human Rad52.

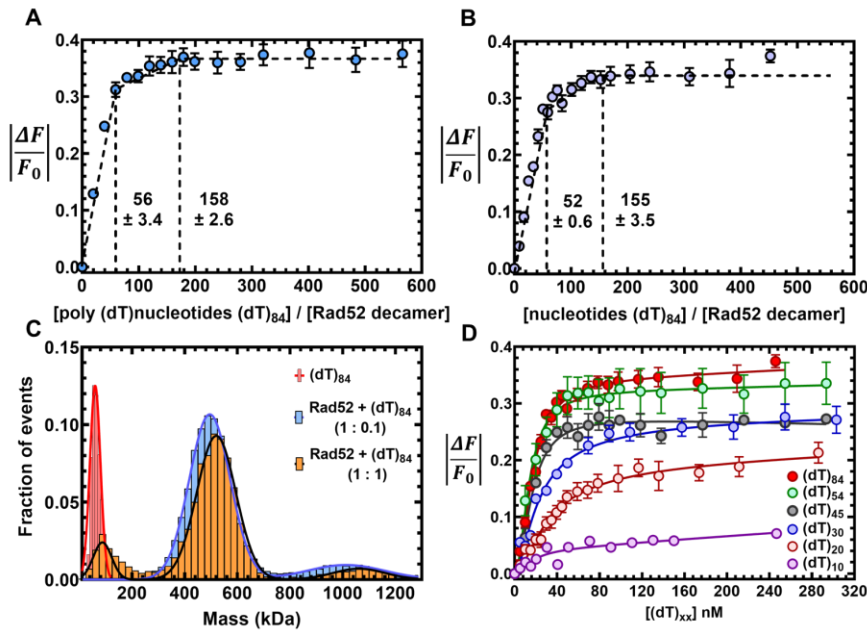


Figure 2. Rad52 binds to ssDNA using multiple binding sites. Tryptophan quenching experiments with A) poly(dT) or B) (dT)84 reveal two ssDNA binding phases for Rad52 with sites sizes of 56 ± 3 and 158 ± 3 , respectively. C) Mass photometry analysis of (dT)84 or Rad52+(dT)84 (1:0.1 or 1:1, Rad52decamer:ssDNA) show that a single Rad52 homodecimer binds to (dT)84. The small fraction of the higher order Rad52 molecules also bind to ssDNA and behave as a single DNA-bound complex. D) Tryptophan quenching of Rad52 as a function of ssDNA length shows that DNA lengths of $>(\text{dT})_{45}$ are required to observe maximal signal quenching.

Two binding phases could also be explained by the accumulation of ssDNA bound Rad52 aggregates as has been proposed for human Rad52.⁴¹ To investigate this possibility, we performed MP and AUC^{SV} analysis to ascertain the stoichiometry of yeast Rad52 decamers bound to ssDNA. When binding of Rad52 to short (dT)₄₅ or long ssDNA (dT)₉₇ is assessed, stoichiometric Rad52 decamer:DNA complexes are observed (Supplementary Fig. 5C). The minor fraction of two-Rad52 ring species also binds to equimolar amounts of ssDNA. On longer (dT)₉₇ or (dT)₈₄ ssDNA substrates, predominantly 1:1 complexes are observed even under conditions where a 10-fold excess of the Rad52 ring is present (Fig. 2C). Fluorescence anisotropy analysis reveals that Rad52 binding to ssDNA is highly specific (Supplementary Fig. 5B), and shows hallmarks of an equilibrium process including saturation, high affinity (Supplementary Fig. 5A, B & C), and reversibility (Supplementary Fig. 5E). Furthermore, the affinity for ssDNA interactions is dictated by the length of the ssDNA and (dT)₅₄ binding gives a similar readout as (dT)₈₄ or poly(dT) providing an independent confirmation of the binding site size analysis (Fig. 2D). These data suggest that there are intrinsic differences in the ssDNA binding properties between the yeast and human Rad52 proteins. Next, we investigated whether the disordered C-terminal half of Rad52 contributes to non-uniform ssDNA binding.

Hydrogen-deuterium exchange mass-spectrometry (HDX-MS) reveals conformational changes in the C-terminus of Rad52 upon ssDNA binding.

While the C-terminal half of yeast Rad52 is disordered in our structure, interactions between this region and DNA have been captured.⁴⁰ Thus, to map the overall DNA binding driven conformational changes and to identify other potential DNA binding regions, we performed HDX-MS analysis of Rad52 in the absence or presence of ssDNA. Deuterium incorporation was quantified as a function of time for Rad52 and the Rad52-ssDNA [(dT)₉₇] complex. Peptide level

deuterium uptake was compared between conditions (Supplementary Figs. 7-10). The difference in HDX between Rad52 and the Rad52-(dT)₉₇ complex (differential-HDX) were mapped onto an Alphafold model of the Rad52 monomer and shown as a heatmap (Fig. 3A). Regions colored blue and red depict increased and decreased differential-HDX, respectively. The Alphafold model aligns well with the cryoEM structure of the N-terminal half of Rad52 (Supplementary Fig. 10). Differential-HDX changes are observed in the N-terminal half upon DNA binding as well as inner, outer, and transition regions as expected based the homologous DNA binding regions in human Rad52 (Fig. 3B-D and Supplementary Figs. 7 & 8). Interestingly, a surprising number of peptides in the disordered C-terminal half also show changes in differential-HDX upon ssDNA binding (Fig. 3A, E-F, and Supplementary Figs. 9). One peptide in particular (residues 383-413) displays the largest HDX changes (Fig. 3E). A more detailed investigation of the MS data for this peptide shows an intriguing bimodal deuterium uptake pattern (Fig. 3F and Supplementary Fig. 11). In the absence of DNA, two populations of this peptide are observed that show low or high deuterium uptake. Calculation of the distribution based on ion intensity shows that ~70% of this peptide displays higher uptake (Fig. 3G). In the presence of ssDNA, a loss of the higher uptake population occurs. The data suggests an intrinsic asymmetry in the structural organization/accessibility of the C-terminal half (C-tails) of Rad52 where a small subset of the C-tails are protected from deuterium uptake.

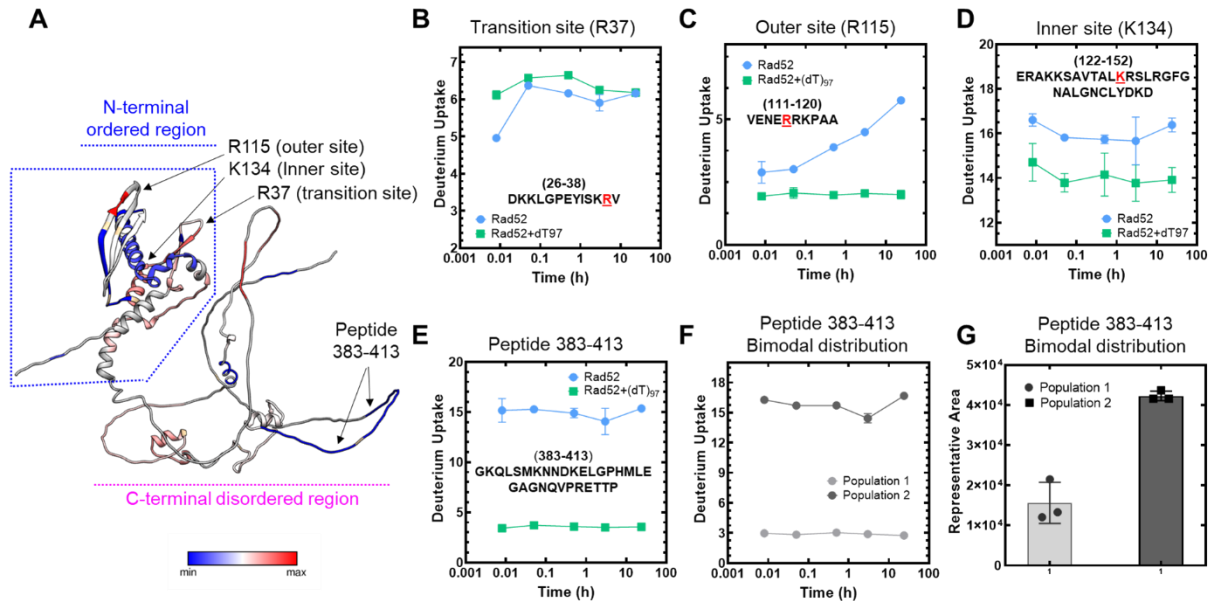


Figure 3. Hydrogen-deuterium exchange mass-spectrometry reveals ssDNA driven changes in both the N- and C-terminal halves of Rad52. A) Differential-HDX changes between Rad52 and the Rad52-(dT)97 complex are projected on to the AlphaFold predicted structure (AF-P06778-F1). DNA-induced HDX changes in the known DNA binding regions (inner, outer, and transition binding sites) are noted. Raw HDX data showing differences within peptides harboring B) Transition (R37), C) Outer (R115), and D) Inner (K134) ssDNA binding sites are shown with the key residues highlighted in red. E) Deuterium uptake profile for peptide 383-413 in the disordered C-terminal half shows a significant reduction in deuterium uptake upon ssDNA binding. F) & G) A bimodal distribution of deuterium uptake was observed in the Rad52 apo sample with least uptake by population 1 and high uptake by population 2. Population 1 represents only 30% of the total signal suggesting that not all subunits in Rad52 are uniform in structure or conformation.

Since there is a dichotomy in the behavior of C-tails, we reanalyzed our 2D classes from the Cryo-EM dataset to gain further insights into this asymmetry. Several 2D classes show the presence of extra bright density adjacent to one part of the decameric ring (Supplementary Fig. 12A). Unfortunately, detailed structural refinement was not possible since this is a disordered region. Low resolution 3D-volume calculations shows the extra density situated on the outer side of the ring closer to the N-terminal half of one side of the Rad52 decamer (Supplementary Fig.

12B). The data reveal an interesting structural feature where only a small subset of the C-tails interact with the decameric ring of Rad52.

Cross-linking mass spectrometry (XL-MS) analysis reveals interactions between the ordered N-terminal and disordered C-terminal halves of Rad52

To identify the regions of the disordered C-terminal half that interact with the ordered N-terminal region we performed XL-MS analysis of Rad52 with bis(sulfosuccinimidyl)suberate (BS3).⁴⁹ The *N*-hydroxysulfosuccinimide (NHS) esters positioned at either end of the BS3 crosslinker are spaced 8-atoms apart and likely report on crosslinks with primary amines in Rad52 that are up to 12Å apart⁵⁰. XL-MS analysis on Rad52 alone or in complex with ssDNA [(dT)₉₇] showed an extensive array of crosslinks (Fig. 4). ~38% and ~19% of the crosslinks are unique to the unbound and DNA bound Rad52 conditions, respectively. For clarity in discussion, we present the data from the perspective of the two halves of Rad52 and focus on the crosslinks between the two halves (Figs. 4A & B). Specifically, Lys-84, Lys-117, and Lys-134 in the N-terminal half make contacts with a defined region in the C-terminal half (residues 340-380; Fig. 4C). Lys-84 and Lys-117 are close to the outer ssDNA binding site whereas Lys-134 is the inner binding site residue (Fig. 4C). Closer examination of the interaction region in the C-terminal half reveals that it harbors the Rad51 binding site. Mutations in this region perturb Rad51 interactions and give rise to HR defects in yeast.^{34,51} Given the abundance of conserved positively charged amino acid residues in this region, we term this the ‘*positive patch*’ (Supplementary Fig. 13).

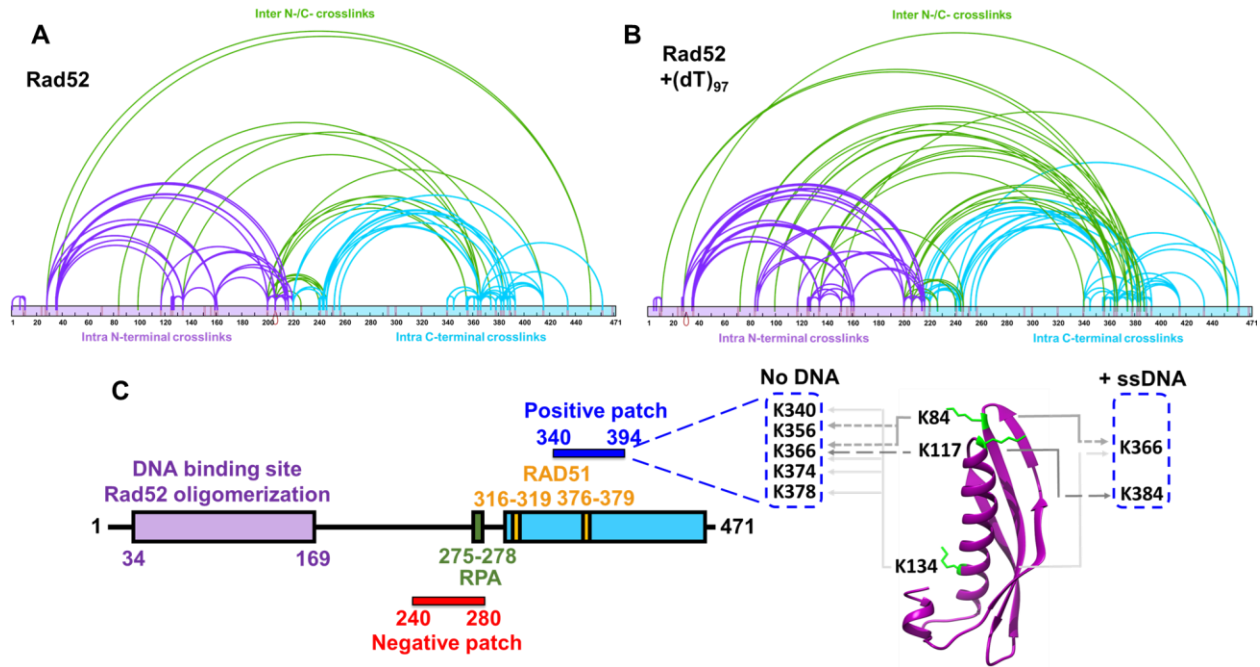


Figure 4. Cross linking mass-spectrometry (XL-MS) reveals interactions between the N- and C-terminal halves of Rad52. A) Rad52 or B) the Rad52-(dT)₉₇ complex was subject to crosslinking with BS3, and the resulting crosslinks were identified using MS. For clarity, data are mapped onto the N-terminal (purple) and C-terminal (cyan) halves of Rad52 as two distinct regions. Intra-N and intra-C half crosslinks are colored purple and cyan, respectively. Inter-crosslinks between the N- and C-terminal halves are colored green. C) Schematic representation of Rad52 shows positions of the RPA and Rad51 binding regions. The proposed positive and negative patches are denoted and the key crosslinks between the DNA binding sites, and the positive patch identified in the XL-MS analysis are highlighted.

Comparison of the crosslinks observed in the presence of ssDNA reveals both loss and gain of interactions (Fig 4). In particular, several of the crosslinks observed in the positive patch are lost (Fig 4B). These data suggest that ssDNA binding modulates interactions between the N- and C-terminal half of Rad52. In addition, substantial intra-C-terminal crosslinks are captured, suggesting that interactions within the disordered region might be important for function. We note that a conserved ‘*negative patch*’ resides in the C-terminal half of Rad52 (residues 240-280) that harbors the site for RPA binding (Supplementary Fig. 13).³⁷ Significant crosslinks are observed between the positive and negative patches (Fig. 4). We propose that the network of

interactions between the patches and the N-terminal half are likely important for function and may dictate the DNA binding and other functional properties of Rad52 (discussed below). Correspondingly, mutations in these patches result in loss of function phenotypes.^{34,37,51}

The C-tail of Rad52 influences the dynamics of ssDNA wrapping

To decipher the contributions of the C-tail to Rad52 function, we first assessed the ssDNA binding activity of Rad52 mutants carrying deletions in either the C-terminal (Rad52^{ΔC}; residues 1-217) or N-terminal (Rad52^{ΔN}; residues 218-417) halves (Fig. 5A). In MP and AUC^{SV} analysis, majority of Rad52^{ΔC} behaves as a homodecamer and a smaller fraction (<5%) of higher order species of twice the mass (Fig. 5B and Supplementary Fig. 14A). In the cryo-EM analysis, 2D classes reveal that Rad52^{ΔC} also forms homodecameric rings in agreement with the N-terminal half harboring the oligomerization region (Supplementary Fig. 14B). ssDNA binding analysis reveals that Rad52^{ΔC} binds ssDNA with similar or higher affinity (Fig. 5C & D), but only the first phase is apparent which saturates with a site size of 55±3 nt/decamer (Fig. 5C). This finding strongly suggests that the relatively weaker second DNA binding site either resides in the deleted C-terminal half, or emerges from an inhibited state through protein-protein interactions between the N- and C-terminal halves. In addition, the Rad52^{ΔC} decamer binds to one molecule of ssDNA (Fig. 5B and Supplementary Fig. 14 B & C).

Although the C-tail appears to influence ssDNA binding in the context of full-length Rad52, on its own lacks any DNA binding activity (Fig. 5E). Interestingly, when Rad52-ssDNA binding is measured in the presence of Rad52^{ΔN} (the C-tail by itself, residues 218-47), a reduction in the second DNA binding phase is observed (Fig. 5F). Furthermore, when Rad52^{ΔC}-ssDNA binding is measured in the presence of Rad52^{ΔN}, an inversion of the Trp quenching

profile is observed (Fig. 5G). These data suggest that the second (weaker) binding site likely emerges from the protein-protein interactions between the N- and C-terminal halves of. A shorter version of the C-tail (Rad52^{ΔN*} residues 294-471) was shown to interact with ssDNA in band-shift analysis.⁴⁰ We observe a weak ssDNA binding activity for Rad52^{ΔN*} (Fig. 5E). Challenging the ssDNA binding activity of Rad52^{ΔC} by addition of increasing concentrations of Rad52^{ΔN*} results in a robust recovery of Trp fluorescence or dequenching (Fig. 5H). We reemphasize that Rad52^{ΔN} possesses both the positive and negative patches whereas the shorter Rad52^{ΔN*} protein carries only the positive patch (Fig. 5A). Since we detected crosslinks between these two patches in the C-terminus (Fig. 4), we propose that these two regions play opposing roles in regulating ssDNA binding. The positive patch may promote interactions with the phosphate backbone of ssDNA. The negative patch likely suppresses DNA binding/remodeling by sequestering away the positive patch.

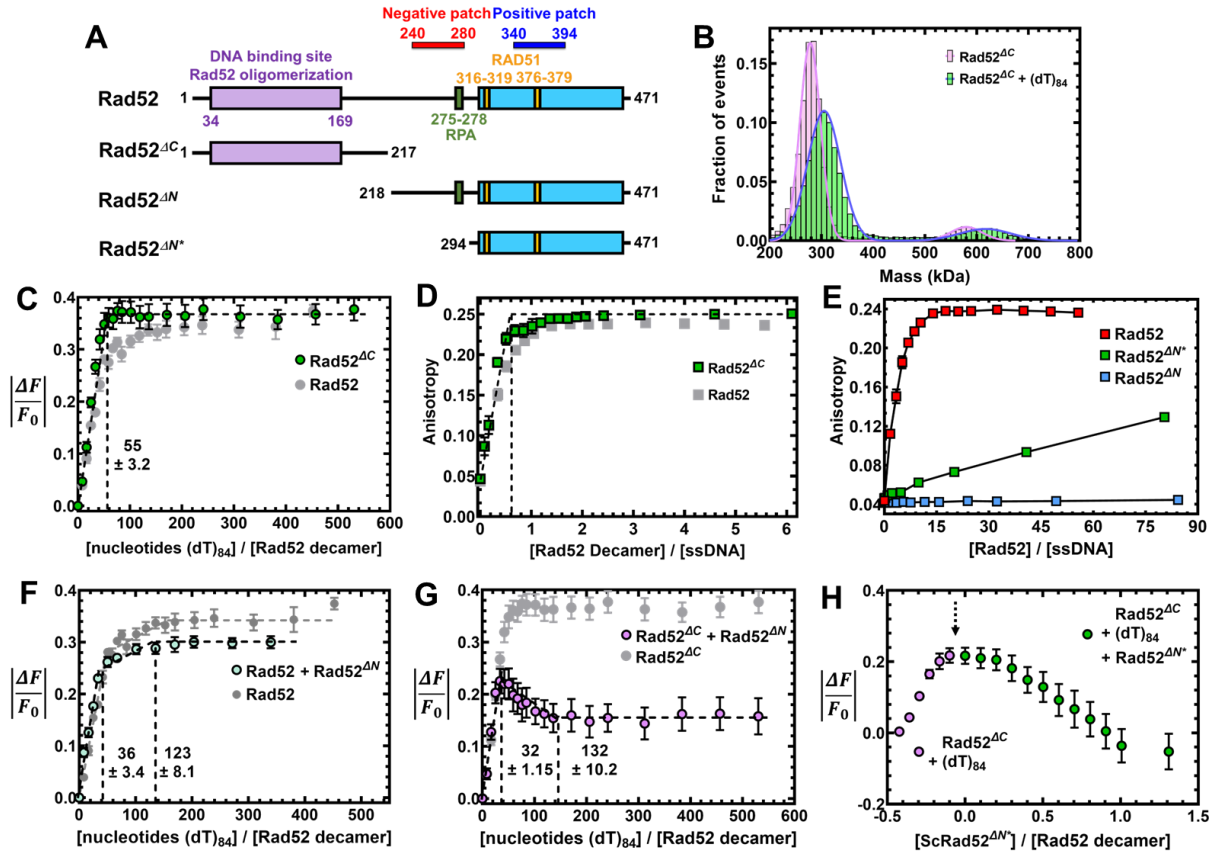


Figure 5. Interactions between the N- and C-terminal halves of Rad52 influence ssDNA binding interactions. A) Schematic of the N- and C-terminal truncations of Rad52. B) Mass photometry analysis shows Rad52DC as a predominant single species in solution with mass corresponding to a homodecamer (254.7 ± 11.9 kDa). The complex remains a homodecamer when bound to ssDNA. C) Trp. quenching analysis of Rad52DC binding to poly(dT) reveals high-affinity monophasic ssDNA binding with site-size of 55 ± 3.2 nt/decamer. D) Fluorescence anisotropy analysis of Rad52DC binding to a FAM-(dT)₃₅ oligonucleotide shows stoichiometric high affinity binding to ssDNA. E) Fluorescence anisotropy ssDNA binding analysis of N-terminal truncated versions of Rad52 show no ssDNA binding activity for Rad52DN, but weak binding for Rad52DN*. F) When Rad52DN is premixed with Rad52, a reduction in the Trp quenching signal upon ssDNA (dT)₈₄ binding is observed. The signal corresponds to the loss of ssDNA to the low affinity binding site in Rad52. G) When Rad52DC and Rad52DN are premixed and binding to ssDNA (dT)₈₄ is assessed through Trp. quenching, an initial increase in binding is observed as the high-affinity site is occupied. This was followed by a sharp loss in binding. H) When similar experiments are performed by sequential addition, titration of ssDNA (dT)₈₄ to Rad52DC produces an increase in Trp quenching as expected (purple data points). After addition of ssDNA, increasing concentrations of Rad52DN* was added, leading to loss in Trp quenching. This data shows defined regions in the C-terminal half of Rad52 modulate ssDNA binding interactions.

The C-tail regulates the dynamic movement/diffusion of ssDNA on Rad52

The data shown thus far argues that the two halves of Rad52 harbor distinct ssDNA binding activities: N-terminal half binds a ~ 55 base long ssDNA tightly followed by a weaker binding to an additional 100 bases either directly to the C-terminal half or to a compound site created by both halves (Fig. 5). Furthermore, a small subset of the C-tails interacts with the N-terminal region. For Rad52 to act as an ssDNA annealing machine, the binding of ssDNA must be dynamic. Such ‘breathing’ of ssDNA in nucleic acid-protein complexes may facilitate base pairing and annealing. Thus, we hypothesized that allostery and asymmetry might regulate the dynamics and diffusion of Rad52 on ssDNA. Functionally, such regulation could help position Rad52 at a defined position on the DNA (either at the ss-dsDNA junction or at the 3’ end). Using single-molecule total internal reflectance fluorescence (smTIRF) microscopy, we investigated whether the binding of ssDNA to Rad52 is dynamic. Cy3 and Cy5 fluorophores were positioned on the DNA such that even partial wrapping or unwrapping of ssDNA bound to Rad52 could be captured as distinct FRET states. A ssDNA length of 45nt is sufficient to capture FRET (Supplementary Fig. 15). The addition of Rad52 to the smFRET DNA substrate led to the appearance of transients *i.e.* bursts of short-lived, high FRET states (Fig. 6B) which are completely absent in the DNA substrate alone sample (Fig. 6A). This indicates that Rad52 very rarely allows complete wrapping (high FRET state) of the overhang substrate. This can also be observed in the smFRET histogram (Fig. 6D) with the peak centered at FRET efficiency of 0.2. On the other hand, Rad52^{ΔC} promotes long lived high FRET states (Fig. 6C-E and Supplementary Figs. 16 & 17). Taken together, these data strongly indicate that the ssDNA binding/wrapping by Rad52 is highly dynamic with the DNA sampling multiple fully/partially

wrapped conformations. Deletion of the C-tail allows the ssDNA to sample higher FRET states more frequently. This indicates that in addition to creating a second (weaker) binding site, the C-tail imposes restrictions on the diffusion of bound ssDNA in the primary (or the high affinity) binding site (Fig. 6B-E).

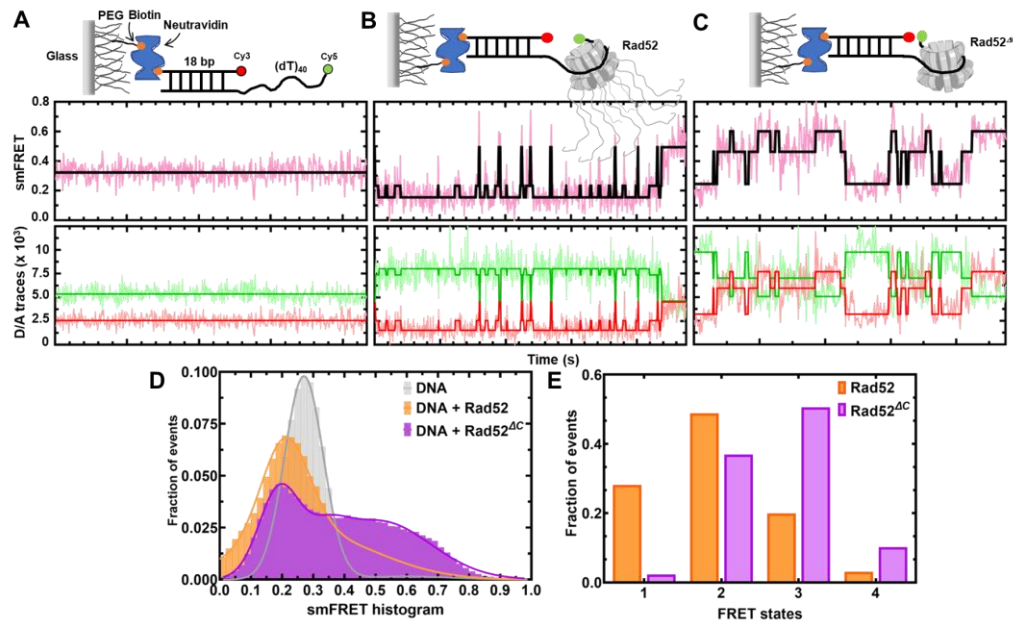


Figure 6. The C-terminus controls the intrinsic diffusion of Rad52 on ssDNA. Single-molecule FRET distributions of A) (dT)40 overhang-ssDNA alone or (dT)40 overhang bound to B) Rad52, or C) Rad52DC. D) & E) HMM analyzed FRET state distributions of smFRET traces show Rad52DC accessing higher FRET states better than Rad52.

The C-tail induced asymmetry dictates Rad51 binding to a single Rad52 subunit.

In both the HDX-MS and cryo-EM analyses, only a select number of C-tails appears to asymmetrically interact with the N-terminus (Fig. 3G and Supplementary Fig. 12). Since the C-tails harbor Rad51 and RPA interaction sites, allosteric regulation of asymmetric binding of these proteins might be functionally important for Rad52 roles in HR. Thus, we investigated the

interactions between Rad52 and Rad51 and the role of the C-terminus. Since each Rad52 subunit in the decamer possesses a Rad51 binding region, multiple Rad51 molecules can be expected to be coated around the ring. Furthermore, Rad51 forms multiple oligomeric complexes in solution as observed by AUC^{SV} (Fig. 7A) and MP (Supplementary Fig. 18). When Rad51 and Rad52 were mixed in equimolar amounts (1 Rad51 per Rad52 monomer) all the Rad51 molecules form a complex with the Rad52 decamer (Fig. 7A). Increasing the ratio of Rad51:Rad52 results in formation of complexes with higher molecular masses (Fig. 7A). Thus, we assumed that all subunits in the Rad52 ring were bound by Rad51. To better understand the structural organization of this complex we used single-particle Cryo-EM (Supplementary Fig. 19). Surprisingly, 2D classes show strong density for Rad51 bound alongside only one subunit in the Rad52 decamer (Supplementary Fig. 19A). Even when the concentration of Rad51 is doubled in the complex, selective binding to one part of the Rad52 ring is observed in the 2D classes (Supplementary Fig. 19B). Unfortunately, the dynamic nature of the complex and poor sampling of orientations hampered our ability to obtain high-resolution structures or reliable 3D volumes (Supplementary Fig. 19C). These studies imply that an intrinsic asymmetry within Rad52 exists and controls Rad51 interactions. This asymmetry and allostery was also observed in the interactions between the C-tails and the N-terminal half of Rad52 (Fig. 3). Deletion of the C-terminal tail (Rad52^{ΔC}) reduced complex formation with Rad51 and more polydisperse complexes were observed (Supplemental Fig. 20). Thus, it is likely that the C-tail allosterically contributes to modulation of Rad52-Rad51 binding asymmetry. In addition, since Rad51 binds to Rad52^{ΔC}, a binding site for Rad51 exists within the N-terminal half of Rad52. However, the relative affinities for the different sites are to be established.

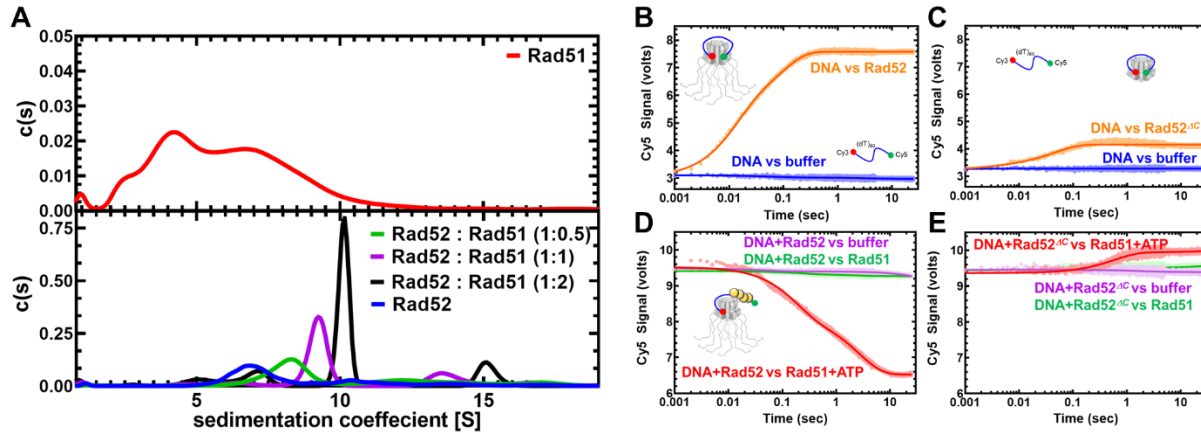


Figure 7. The C-tail of Rad52 regulates allosteric interactions with Rad51. A) AUCSV analysis (top panel) shows polydispersity of Rad51 with multiple oligomeric states. When increasing amounts of Rad51 (1, 10, or 10 μ M) are mixed with a fixed concentration of Rad52 (10 μ M monomer), formation of Rad52-Rad51 complexes is observed. The size of the complexes increases as a function of Rad51 concentration in the solution. B) Stopped flow FRET experiments were performed with a (dT)₉₇ ssDNA labeled at the ends with Cy3 (3'; donor) and Cy5 (5'; acceptor). In the absence of protein, no changes in FRET are observed (blue) and a robust increase is captured when Rad52 is introduced (orange). Similar experiments done with C) Rad52DC show wrapping of ssDNA. D) FRET experiments were performed using prewrapped Rad52-ssDNA complexes and change in FRET was measured upon addition of Rad51 in the absence (green) or presence (red) of ATP. A loss in FRET is observed when Rad51 filaments are formed on the ssDNA in the presence of ATP. E) Similar experiments performed with Rad52DC do not show a decrease in FRET suggesting an impediment to Rad51 filament formation.

Formation of proper Rad51 filaments requires the C-tail of Rad52

To investigate the role of Rad52 in promoting Rad51 binding to ssDNA and formation of a filament, we utilized a FRET assay using end-labeled ssDNA oligonucleotides. Cy3 (donor) and Cy5 (acceptor) were positioned at the 3' and 5' end of a (dT)₉₇ ssDNA oligonucleotide and changes in Cy5 fluorescence were measured by exciting Cy3. Upon binding to Rad52 or Rad52^{4C}, the ssDNA is wrapped around the decameric ring, and the fluorophores are brought in proximity resulting in sensitized Cy5 emission (Supplementary Fig. 15 & Figs. 7A & B). To measure the pre-steady state kinetics of this process, experiments were conducted in a stopped

flow instrument. Next, preformed Rad52-DNA or Rad52^{ΔC}-DNA complexes were challenged with Rad51 in the absence or presence of ATP. ATP-binding promotes Rad51 binding and filament formation¹², and thus, in the presence of ATP a rapid loss in the FRET signal is observed in experiments with Rad52 (Fig. 7D). In contrast, no reduction in FRET was observed for the Rad52^{ΔC}-DNA complex suggesting that Rad51 does not immediately form a filament when the C-terminal half of Rad52 is absent (Fig. 7E). Interestingly, a minor increase in the FRET signal is observed which reflects Rad51 binding directly to ssDNA as free DNA would be available when Rad52^{ΔC} is wrapped by a (dT)₉₇ substrate (Supplementary Fig. 15). Please note that these changes are only captured in shorter time regimes in the stopped flow experiments. In steady state FRET measurements, Rad51 eventually outcompetes both Rad52 and Rad52^{ΔC} leading to similar signal changes (Supplemental Fig. 21 B & C). Thus, the C-terminus of Rad52 regulates formation of Rad51 filaments on ssDNA.

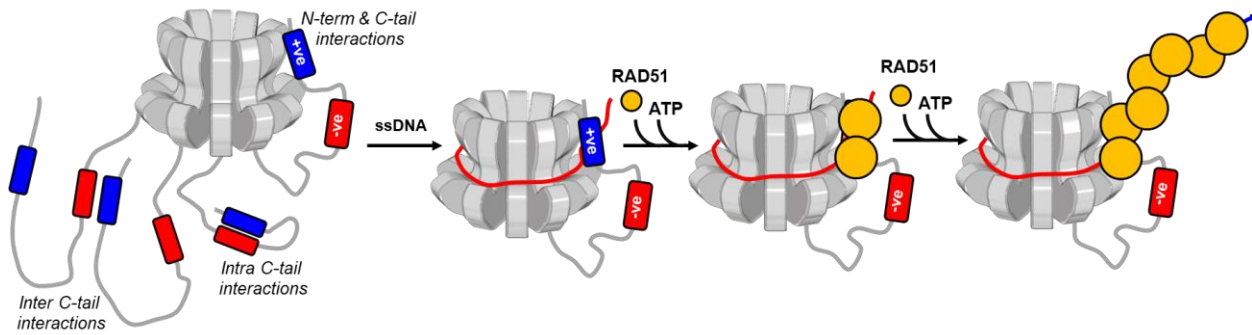


Figure 8. Model describing the mechanism of action of Rad52 and regulation by the C-tail. The homodecameric ring structure of the N-terminal half of *S. cerevisiae* was depicted and three of the ten disordered C-terminal tails were shown (for clarity). The negative and positive patches were shown and the asymmetric interaction between one tail and the N-terminal half was depicted. We propose that such interactions in remaining tails were allosterically suppressed through either intra or inter C-tail interactions between the negative and positive patches. The path of ssDNA binding is modulated through interactions between the DNA backbone and the positive patch bound to the N-terminal half. This unique asymmetric region of interaction between the two halves of Rad52 serves as the point of Rad51 binding and nucleation on to ssDNA. In the presence of ATP, this allosterically controlled process promotes single-strand nucleation and growth of presynaptic Rad51 filaments in HR.

Discussion

A barebones view of events during formation of the Rad51 filament during HR pre-synapsis encompasses a resected 3' ssDNA overhang coated by RPA, Rad51, and Rad52. Previous models suggested that the mediator functions of Rad52 promote removal of the high-affinity DNA binding RPA ($K_D < 10^{-10}$ M) to enable binding/nucleation of the lower affinity Rad51 ($K_D \sim 10^{-6}$ M)⁵². Physical interactions between Rad52 and Rad51 and RPA are required for this exchange^{34,38,53}. Recent single-molecule experiments show that RPA-Rad52 function as a complex on the DNA to promote Rad51 binding, with additional roles in downstream second-strand capture⁵⁴. In terms of the structure-function assessments, the N-terminal half of each Rad52 subunit is thought to promote the oligomerization and DNA binding activities^{26,27,41,55}.

Protein-interaction domains have been mapped to the C-terminal half^{25,34,56-58}. How the various activities of Rad52 are coordinated to enable the mediator functions in HR are poorly resolved.

Using single-particle Cryo-EM analysis, we illustrate that *S. cerevisiae* Rad52 functions as a homodecamer. As shown for human Rad52, the N-terminal half is ordered in the homodecameric ring architecture and the C-terminal half is disordered. Remarkably, an intrinsic asymmetry in Rad52 is imparted by the C-terminal region from one (or a select few) subunit that interacts with the ordered ring. This asymmetry in turn regulates a two-state ssDNA binding activity driven through high- and low-affinity ssDNA interactions. When the C-tail is removed, the low-affinity binding is lost. We propose the presence of two charged regions (termed as positive and negative patches) in the C-tail that contribute to the asymmetry within Rad52. The positive patch interacts weakly with ssDNA, and interactions between the positive and negative patches inhibit this activity. In the subunits where interactions between the N- and C-terminal halves are observed, the positive charge is accessible to ssDNA interactions. We speculate that such asymmetric interactions could guide the path of ssDNA around the Rad52 ring and likely setups either polarity or define the position of Rad51 or RPA binding (Fig. 8).

In support of this model, the asymmetry identified in the Rad52 structure also extends to interactions with Rad51. In cryo-EM 2D class averages, Rad51 is bound to only one region in the Rad52 ring (Supplementary Fig. 19). This observation persists even when a vast excess of Rad51 is present in the reaction. These data suggest that the asymmetry in ssDNA and Rad51 binding within Rad52 might serve as a mechanism for ‘single-position nucleation’ of the Rad51 nucleoprotein filament (Fig. 8). This phenomenon might be essential for accurate DNA repair in HR. Identification of perfect homology within the undamaged strand (sister allele) is of

paramount importance⁵⁹. For Rad51 (and the prokaryotic homolog RecA) a minimum of 8 nt homology is required to promote strand invasion⁶⁰⁻⁶³. Furthermore, a ‘search entity’ has been proposed where Rad51 is in complex with additional proteins (including Rad52 and RPA) to catalyze homology search and pairing⁵⁹. Thus, if Rad51 binds to multiple subunits on Rad52, formation of a uniform search entity that engages and handles ssDNA at a single position might not be possible. We envision a minimal search entity to contain a Rad52 decamer bound to Rad51 at a single defined position. This binding position is dictated by one (or a select few) of the C-tails that interacts with both the ordered N-terminal half and ssDNA. The mechanisms of negative cooperativity and allostery enacted to suppress such binding events in the other C-tails in the Rad52 ring needs to be deciphered. The positive patch we define also harbors the Rad51 binding motif. Thus, the C-terminal ssDNA and Rad51 binding sites are likely coupled.

Another layer of complexity is the position of RPA within the RPA-Rad52-Rad51 complex. Does the minimal search entity comprise RPA-Rad52-Rad51? Does the RPA-Rad52 complex form first and then recruit Rad51, or is it the inverse? Can such complexes be formed in the absence of ssDNA? Many such questions need to be answered. In the absence of ssDNA we can reliably capture the RPA-Rad52-Rad51 complex in analytical ultracentrifugation analysis (Supplementary Fig. 20C). Using single-molecule tools, we previously showed that Rad52 can selectively remodel DNA-binding domain D of RPA of an RPA-ssDNA complex¹³. In the absence of ssDNA, the Rtt105 chaperone inhibits interactions between Rad52 and RPA by configurationally stapling the flexible architecture of RPA⁶⁴. Thus, many other regulatory mechanisms need to be accounted for to comprehensively understand the nature of the ‘search entity’ and its regulation.

Acknowledgements

The authors thank members of our respective research laboratories for critical reading of the manuscript. This work was supported by grants from the National Institutes of Health (R01 GM130746 and R01 GM133967) to E.A. Funding for Proteomics, Metabolomics and Mass Spectrometry Facility at MSU was made possible in part by the MJ Murdock Charitable Trust and NIGMS of the National Institutes of Health under Award Number P20 GM103474. The analytical ultracentrifugation experiments were supported by an instrumentation grant from the Office of the Director, National Institutes of Health (S10 OD030343 to E.A.). M.J.R. and K.B. gratefully acknowledge support from the Washington University Center for Cellular Imaging (WUCCI) which is funded by Washington University School of Medicine, The Children's Discovery Institute of Washington University and St. Louis Children's Hospital (CDI-CORE-2015-505 and CDI-CORE-2019-813), the Foundation for Barnes-Jewish Hospital (3770) and the Alvin J. Siteman Cancer Center at Washington University School of Medicine and Barnes-Jewish Hospital under NCI Cancer Center Support Grant P30CA091842. J.A.J.F gratefully acknowledges the Chan Zuckerberg Initiative for their support as a CZI Imaging Scientist under award 2020-225726.

Author Contributions

J.D., M.J.R., K.B. and J.A.J.F. collected single-particle Cryo-EM data and solved the structure. R.C. performed fluorescence and single-molecule experiments. J.M. and B.B. performed and analyzed the HDX-MS and XL-MS experiments. S.K., N.P., and N.E. assisted

with cloning, protein purification, and experiments. E.A. conceived the project and wrote the manuscript with assistance and contributions from all authors.

Data availability

The coordinates for the Rad52 structure are available in the PDB with code 8G3G and the Cryo-EM maps are available in EMBD with code EMD-29695. The source data for all experiments shown are provided as a Source Data file. Constructs for protein expression are available from the corresponding author upon request.

Competing interests

The authors declare no competing interests.

Materials and Methods

Reagents and Chemicals

Fluorescently labeled and non-labeled oligonucleotides were synthesized by Integrated DNA technologies Inc. Poly(dT) was purchased from Midland Certified Reagent Company. Chemicals were purchased from Millipore-Sigma, Research Products International, Gold Biotechnology, and Fisher Scientific. Resins for protein purification were from Cytiva Life Sciences, Bio-Rad laboratories, and Gold Biotechnology.

Plasmids for protein overproduction

Plasmids carrying the coding sequence for yeast Rad52 with a cleavable C-terminal chitin binding domain (pTXB1-CBD) was a kind gift from Dr. Eric Greene (Columbia University).

Rad52^{ΔC} was generated by introducing a stop codon at the nucleotide position corresponding to amino acid position 218. *Rad52^{ΔN}* and *Rad52^{ΔN*}* were generated by PCR amplifying the respective fragments and subcloning into the original pTXB1 plasmid encoding a C-terminal CBD tag using *NdeI* and *XhoI* restriction sites. Plasmids for overproduction of RPA and Rad51 were as described.^{12,14}

Purification of Rad52 and Rad52 mutants

Rad52 plasmids were transformed into Rosetta-2 PlysS cells, plated on LB agar plates (with 100 µg/mL ampicillin), and incubated overnight at 37 °C. A single colony was inoculated in a 50 mL LB starter culture supplemented with ampicillin and incubated overnight at 250 rpm at 37 °C. 10 mL of the overnight culture was inoculated into a 2.8 L baffles flask with 1.5 L LB media supplemented with ampicillin and grown at 250 rpm at 37 °C until the growth reached an OD₆₀₀ of 0.6. The culture was induced by the addition of 0.5 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG). Following induction, the culture was incubated overnight at 150 rpm, 18 °C. All the steps from here on were performed at 4 °C or on ice. Cells were pelleted by centrifugation at 5000 rpm for 30 minutes and resuspended in 200 mL cell lysis buffer (50 mM Tris-HCl, pH 8.0, 600 mM NaCl, 1 mM EDTA, and 2X protease inhibitor cocktail (PIC)). Cells were lysed by incubation for 30 minutes with 0.4 mg/ml lysozyme at 4 °C, followed by 2 cycles of sonication (50 % amplitude, 2 seconds pulses for 60 seconds) with a minute pause in between. Cell lysate was clarified by centrifugation at 17,000 rpm for 1 hour. The clarified supernatant was then batch-bound for 3 hours on to 20 mL chitin resin (New England Biolabs) equilibrated with 200 mL of cell lysis buffer. The protein-bound beads were sequentially washed with 100 mL each of cell lysis buffer, cell lysis buffer + 1.5 M NaCl and finally with the cell lysis buffer.

Rad52 protein was batch eluted with CBD elution buffer (cell lysis buffer + 50 mM DTT). Chitin column eluates were analyzed by 10 % SDS-PAGE and eluates containing Rad52 were pooled, diluted with R₀ (50 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 2 mM β ME, 10 % glycerol, and 2X PIC) to a final conductivity of R₂₀₀ and applied on a 20 mL heparin column (Cytiva Inc.). Bound Rad52 was fractioned using a linear gradient of R₂₀₀-R₁₀₀₀. Fractions containing Rad52 were pooled and loaded on a 20 mL Q-Sepharose column. Fractions containing Rad52 were pooled, concentrated with a 30 kDa cut-off spin concentrator to a volume ~4.5 mL and applied on to HiLoad 16/600 200 Superdex-200 pg size exclusion column. Fractions containing Rad52 were pooled, concentrated with a 30 kDa cut-off spin concentrator, and stored at -80 °C as small aliquots after flash freezing with liquid nitrogen. Before flash freezing, concentration of Rad52 was determined spectroscopically using molar extinction coefficient (ϵ_M) 25,900 M⁻¹cm⁻¹. Rad52 ^{ΔN} , and Rad52 ^{ΔN^*} lack tryptophan residues so their concentration was first estimated using the predicted molar extinction coefficient for the peptide bond measured at 214 nm. In addition, the concentration of Rad52, Rad52 ^{ΔC} , Rad52 ^{ΔN} , and Rad52 ^{ΔN^*} were also verified using the Bradford method before performing Tryptophan quenching experiments. Rad51 and RPA were purified as described.^{12,14}

Determination of stoichiometry and molecular weight using Mass Photometry

Mass photometry measurements were carried out on a TwoMP instrument (Refeyn Inc.). Glass coverslips (No. 1.5H thickness, 24x50 mm, VWR) were cleaned by sonication, first in isopropanol, and then in deionized water for 15 min. Cleaned coverslips were dried under a stream of nitrogen. For some measurements (e.g., DNA-bound complexes) coverslips were

coated with a 0.1% w/v poly-lysine solution. For each measurement, a clean coverslip with a 6-well silicone gasket was placed onto the water-immersion objective and samples were added onto each well as described below. Molecular weight standards were resuspended in their matching buffers and used to calibrate and convert the particle-image contrast due to scattered light into MW units. Rad52 was diluted to a final concentration of 130 nM decamers (or 1.3 μ M monomers) in buffer (50 mM Tris-acetate pH 7.5, 50 mM KCl, and 5 mM MgCl₂). After focus was obtained and the image stabilized, 1 μ L of the 130 nM protein solution was mixed in with 15 μ L of 1X buffer taken in a silicon gasket well resulting in an 8.1 nM solution of decamers (final concentration in the well). Standards included β -amylase (3 species of 56, 112, & 224 kDa, respectively) and thyroglobulin (single species of 667 kDa). 8-10 nM of Rad52 decamers produced statistically significant number of particles ($\sim$ 5,000) over a 60 s recording. For data analysis, single particle landings (events) were identified and converted to mass units using the standard calibration, extracted from videos, and non-linear least squares fitted with Gaussian mixture model (Eq. 1) to quantify the underlying populations (Supplementary Table). For Rad52, the molecular weight upon addition of (dT)₈₄ (25.4 kDa) did not appear to shift noticeably. This is likely since the limit of sensitivity of the instrument is $\sim$ 50 kDa and any increments in molecular weight lower than that cannot be distinguished from noise. However, a change was readily visible for Rad52^{AC}.

Eq. 1.

$$f(x) = \sum_{i=1}^n a_i e^{\left[-\left(\frac{x-b_i}{c_i}\right)^2\right]}$$

$$f(x) = a_1 * \exp^{-\left(\frac{x-b_1}{c_1}\right)^2} + a_2 * \exp^{-\left(\frac{x-b_2}{c_2}\right)^2} + \dots a_n * \exp^{-\left(\frac{x-b_n}{c_n}\right)^2}$$

Where a_n , b_n , and c_n are the amplitude, mean, and the standard deviation of the n^{th} Gaussian component. For Rad52 and Rad52^{4C}, the n equaled 3, but Rad51 showed significant heterogeneity and the Rad51 nano-filament mass size ranged from a dimer to a pentadecamer ($n = 15$; Supplementary Figure 18).

Analytical Ultracentrifuge (AUC) Analysis

Analytical ultracentrifugation sedimentation velocity experiments were performed with an Optima Analytical Ultracentrifuge (Beckman-Coulter Inc.) using a An-50 Ti rotor. Samples were spun at 40,000 rpm at 20 °C. Samples were prepared by thoroughly dialyzing against 30 mM HEPES, pH 7.8, 100 mM KCl, 10% glycerol, and 1 mM TCEP. Sample (380 µL) and buffer (400 µL) were filled in each sector of 2-sector charcoal quartz cells. Absorbance was monitored at 280 nm for samples using absorbance optics and scans were obtained at 3 min intervals. The density and viscosity of the buffer at 20 °C was calculated using SEDNTERP. The continuous distribution sedimentation coefficient (c(s)) model was used to fit the AUC data using SEDFIT.⁶⁵

Cryo-EM Data collection

3 µL of 0.31 mg/mL Rad52 or 0.44 mg/mL of the Rad51-Rad52 complex were applied on to Quantifoil R 2/2 300 copper mesh grids, then plunge-frozen into liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific, Brno, Czech Republic) set to 4 °C and 100% humidity. Prior to vitrification, the grids were plasma cleaned in an H₂O₂ plasma for 1 minute using a Solarus 950 (Gatan, Warrendale, PA). The sample was allowed to incubate for 20 s on

the grids prior to blotting. Single-particle cryo-EM data was acquired on a 200 kV Glacios cryo-TEM equipped with a Falcon IV direct electron detector (Thermo Fisher Scientific, Eindhoven, Netherlands). The nominal magnification was 150,000x, resulting in a pixel size of 0.94 Å. The total electron dose was $50.7 \text{ e}^-/\text{Å}^2$ per movie with 48 frames, and the defocus value varied between -0.8 and -2.4 μm.

Cryo-EM Image processing

Single-particle cryo-EM data was processed using cryoSPARC v3.2.0.⁶⁶ For the Rad52 dataset, 2,447 raw movies were motion corrected using patch motion correction. The CTF estimation of the subsequent micrographs was performed using patch CTF estimation. Initial particle picking was accomplished using blob picker, the results of which were subjected to 2D classification. The best templates were chosen for a subsequent template picker job. After inspection of the template picks, 701,962 particles were extracted using a box size of 300 pixels and subjected to three rounds of 2D classification, resulting in 180,545 particles. Three initial models were generated using an *ab-initio* reconstruction job, which were then refined by a heterogeneous refinement job with C1 symmetry. After visualizing the volumes in UCSF ChimeraX⁶⁷ the best class (consisting of 46,845 particles) was chosen for further processing. The class was subjected to non-uniform refinement with C10 symmetry, then the particle stack was symmetry expanded and underwent local refinement. The resulting cryo-EM density map had an estimated gold standard Fourier shell correlation (FSC) resolution of 3.48 Å at a threshold of 0.143.

For the Rad52-Rad51 dataset, 5,054 raw movies were motion corrected using patch motion correction, then the CTF estimation of the subsequent micrographs was performed using

patch CTF estimation. Initial particle picking was accomplished using blob picker, the results of which were subjected to 2D classification. The best templates were chosen for a subsequent topaz picker job. After removing duplicates, 1,111,384 particles were extracted using a box size of 360 pixels and subjected to three rounds of 2D classification, resulting in 332,737 particles. Five initial models were generated using an ab-initio reconstruction job, which were then refined by a heterogeneous refinement job with C1 symmetry. The disordered nature of the Rad51 density in these volumes made it difficult to obtain a higher resolution structure of the Rad52-Rad51 complex.

Model building:

The Atomic model for yeast rad52 was built *de novo* using Model Angelo tool.⁶⁸ The Model Angelo generated model was curated manually in Chimera⁶⁹ and Coot v0.9.3⁷⁰. The manually curated atomic model was real space refined followed by manual curation in Coot. The final model was validated by Ramachandran plot in Phenix.

Measurement of ssDNA binding to Rad52 using fluorescence anisotropy

5'-FAM labeled (dT)₃₅ or (dT)₇₀ ssDNA was diluted to 30 nM in buffer (50 mM Tris-acetate pH 7.5, 50 mM KCl, 5 mM MgCl₂, 10% v/v glycerol, 1 mM DTT and 0.2 mg/mL BSA) and taken in a 10 mm pathlength cuvette (Starna Cells Inc). The temperature was maintained at 23 °C as the fluorescein labelled-ssDNA molecules were excited with vertically polarized 488 nm light and emission was collected using a 520 nm bandpass filter in parallel and perpendicular orientation in a PC1 spectrofluorometer (ISS Inc.). The samples were taken in triplicate and five consecutive measurements were averaged for each data point as G-factor corrected anisotropy values were measured after addition of Rad52 (full-length or Rad52^{ΔC}). The titration was

continued until the anisotropy value plateaued indicating a steady state. The raw anisotropy values were corrected for the reduction in quantum yield of the fluorescein moiety upon Rad52 addition, brought about the proximity based quenching effects of the bound protein molecule as follows,

Eq. 2.

$$A_c = \frac{\left[\left(\frac{A - A_f}{A_b - A} \right) \times \left(\frac{Q_f}{Q_b} \right) \times (A_b) \right] + A_f}{1 + \left[\left(\frac{A - A_f}{A_b - A} \right) \times \left(\frac{Q_f}{Q_b} \right) \right]}$$

Where, 1) F_b , and F_f are the bound, and free concentrations of the FAM-labeled fluorescent ssDNA in nM, 2) Q_b , and Q_f are the fluorescence quantum yields of the bound and free form of the FAM-labeled fluorescent ssDNA (arbitrary units), 3) A_b , and A_f are the anisotropy values of the bound and free forms of the FAM-labeled fluorescent ssDNA, 4) A , is the measured anisotropy, and 5) A_c is the corrected anisotropy value. Care was taken to limit the maximal dilution to <5% of initial volume and the protein, and ssDNA concentration was corrected for this effect. Due to extremely tight binding between Rad52 (full length, and Rad52^{4C}), and (dT)₃₅, no attempt was made to fit a binding curve to the data points.

Tryptophan quenching assays

Titration monitoring the tryptophan fluorescence output of Rad52 (full length or Rad52^{4C}) were performed with a PC1 spectrofluorometer. The tryptophan residues in the protein samples were selectively excited with 296 nm excitation light and the emitted fluorescence (320 – 360 nm) was collected as an emission scan in triplicate using a 10 mm path length cuvette; the peak being located at 345 nm. The temperature of the sample was maintained at 23 °C. ssDNA

(poly(dT) or (dT)₈₄) was added to a 1.2 mL, 50 nM (Rad52 or Rad52^{ΔC} decamer concentration) protein taken in buffer (50 mM Tris-acetate pH 7.5, 50 mM KCl, 5 mM MgCl₂, 10% v/v glycerol, and 1 mM DTT) with constant stirring. After a 2 min incubation, the fluorescence emission scan was reacquired. To minimize photobleaching the excitation shutter was only open during data acquisition. The fluorescence values were corrected for the effects of dilution, photobleaching, and inner filter effects using Eq. 3.

Eq. 3

$$F_{i,corr} = F_{i,raw} \times \left(\frac{V_{i,raw}}{V_0} \right) \times \left(\frac{1}{C} \right) \times \left(\frac{f_0}{f_i} \right)$$

Where, $F_{i,raw}$, and $F_{i,corr}$ are measured, and corrected Rad52 fluorescence values after each (i^{th}) addition of nucleic acid stock. V_0 is the initial volume of the solution before titration, V_i is the volume after addition of the i^{th} aliquot. f_0 , and f_i are the initial fluorescence and fluorescence of Rad52 due to photobleaching alone under identical solution conditions. C or the correction factor for the inner filter effect was assumed to be 1 because of following reasons: 1) The optical density (OD₂₈₀) of Rad52 at 50 nM decamer concentration (or 0.5 μM monomer concentration) is only 0.013 which is at least 7-fold lower than the O.D. = 0.1 cutoff where inner filter effects become significant. 2) In addition, the normalized fluorescence spectra of Rad52 at 10, 25, 50, and 70 nM decamer concentration appear comparable when normalized for concentration (data not shown). 3) The maximum concentration of ssDNA in the cuvette, at the end of titration, rarely exceeded 250 nM. The corrected data was plotted as the absolute change in fluorescence normalized to the initial fluorescence ($\left| \frac{\Delta F}{F_0} \right|$) on Y-axis, with [nucleotides] / [Protein] ratio on the X-axis. The biphasic curve resulting from the TRP quenching of Rad52

with poly(dT) or (dT)₈₄ could be fitted with a model for two independent, sequential DNA binding sites as follows:

Eq. 4

$$TQ = \frac{TQ_1 K_{D1} D_f + TQ_2 K_{D1} K_{D2} D_f^2}{1 + K_{D1} D_f + K_{D1} K_{D2} D_f^2}$$

Eq. 5

$$D = D_f + D_b = D_f + \frac{K_{D1} D_f + 2K_{D1} K_{D2} D_f^2}{1 + K_{D1} D_f + K_{D1} K_{D2} D_f^2} [Rad52]$$

Where, TQ is the observed Tryptophan quenching, which is sum of two components TQ_1 , and TQ_2 due to the strong and the weak binding components. K_{D1} and K_{D2} are macroscopic dissociation constants for the first and the second phase, respectively. D is the total DNA concentration which is sum of free and bound components D_f , and D_b . Note that D_b is a function of total protein concentration or $[Rad52]$. $[Rad52]$ is assumed to be constant at 50 nM (decamer) during fitting as there is less than 5% dilution by the end of titration. K_{D1} , and K_{D2} are determined from the non-linear least squares fitting of equations 2, and 3.

Hydrogen-Deuterium Exchange mass spectrometry (HDX-MS) analysis

Stock solutions of Rad52 (15.3 mg/mL) were mixed in the presence or absence of (dT)₉₇ ssDNA in a 1:1.2 ratio. Reactions were diluted 1:10 into deuterated reaction buffer (30 mM HEPES, 200 mM KCl, pH 7.8). Control samples were diluted into a non-deuterated reaction buffer. At each time point (0, 0.008, 0.05, 0.5, 3, 30 h), 10 μ L of the reaction was removed and quenched by adding 60 μ L of 0.75% formic acid (FA, Sigma) and 0.25 mg/mL porcine pepsin (Sigma) at pH 2.5 on ice. Each sample was digested for 2 min with vortexing every 30 s and

flash-frozen in liquid nitrogen. Samples were stored in liquid nitrogen until the LC-MS analysis. LC-MS analysis of Rad52 was completed as described⁷¹ Briefly, the LC-MS analysis of Rad52 was performed on a 1290 UPLC series chromatography stack (Agilent Technologies) coupled with a 6538 UHD Accurate-Mass QTOF LC/MS mass spectrometer (Agilent Technologies). Peptides were separated on a reverse phase column (Phenomenex Onyx Monolithic C18 column, 100x2mm) at 1°C using a flow rate of 500 µL/min under the following conditions: 1.0 min, 5% B; 1.0 to 9.0 min, 5 to 45% B; 9.0 to 11.8 min, 45 to 95% B; 11.8 to 12.0 min, 5% B; solvent A = 0.1% FA (Sigma) in water (Thermo Fisher) and solvent B = 0.1% FA in acetonitrile (Thermo Fisher). Data were acquired at 2 Hz s⁻¹ over the scan range 50 to 1700 m/z in the positive mode. Electrospray settings were as follows: the nebulizer set to 3.7 bar, drying gas at 8.0 L/min, drying temperature at 350 °C, and capillary voltage at 3.5 kV. Peptides were identified as previously described⁷² using Mass Hunter Qualitative Analysis, version 6.0 (Agilent Technologies), Peptide Analysis Worksheet (ProteoMetrics LLC), and PeptideShaker, version 1.16.42, paired with SearchGUI, version 3.3.16 (CompOmics). Deuterium uptake was determined and manually confirmed using HDExaminer, version 2.5.1 (Sierra Analytics). Heat maps were created using MSTools⁷³.

Crosslinking Mass Spectrometry (XL-MS) analysis

Stock solutions of Rad52 and ssDNA (dT)₉₇ were diluted to 1.8 mg/mL and 2.2 mg/mL, respectively in buffer (30 mM HEPES, 200 mM KCl, pH 7.8) and incubated together for 30 min. The diluted proteins were reacted with primary amine reactive 5 mM bis(sulphosuccinimidyl)suberate (BS3) and 20 µL of the sample was taken at various time points (0, 15, and 30 minutes) and immediately quenched with 2 µL of 1 M ammonium acetate.

Quenched samples were diluted with 1.5X Laemmli gel loading buffer to a final volume of 40 μ L, vortexed, and heated to 100 °C for 5 min and resolved on 4-20% (w/v) gradient SDS-PAGE gels (Bio-Rad) with Tris-glycine buffer. Gels were stained with Gelcode blue safe protein stain (Thermo Scientific). Gel bands were excised for protein identification and analysis. Excised bands were destained with a 50 mM ammonium bicarbonate and 50% acetonitrile mixture and reduced with a mixture of 100 mM DTT and 25 mM ammonium bicarbonate for 30 min at 56 °C. The reaction was subsequently exchanged for the alkylation step with 55 mM iodoacetamide and 25 mM ammonium bicarbonate and incubated in the dark at room temperature for 25 min. The solution was then washed with the 50 mM ammonium bicarbonate and 50% acetonitrile mixture. The gel pieces were then first dehydrated with 100% acetonitrile and subsequently rehydrated with sequence grade trypsin solution (0.6 μ g, Promega) and incubated overnight at 37 °C. The reaction was quenched with 10 μ L of 50% acetonitrile and 0.1% formic acid (FA, Sigma) and transferred to new microfuge tubes, vortexed for 5 min, and centrifuged at 15,000 rpm for 30 min. Samples were transferred to mass spectrometry vials and quantitated by LC-MS as described for peptide identification^{74,75}. Peptides were identified as previously described⁷² using MassHunter Qualitative Analysis, version 6.0 (Agilent Technologies), Peptide Analysis Worksheet (ProteoMetrics LLC), and PeptideShaker, version 1.16.42, paired with SearchGUI, version 3.3.16 (CompOmics). Crosslinks were then determined using Spectrum Identification Machine (SIMXL 1.5.5.2).

Cy3/Cy5 FRET based steady state ssDNA wrapping experiments

FRET-based analyses of ssDNA binding, and wrapping by full-length Rad52 and Rad52^{ΔC} proteins was carried out by titrating the protein into ssDNA with Cy5 and Cy3 labels at either ends. A 2 mL solution containing 30 nM ssDNA (5'-Cy5-(dT)₈₀-Cy3-3') was taken in 50 mM Tris-acetate pH 7.5, 50 mM KCl, 5 mM MgCl₂, 10% v/v glycerol, 1 mM DTT, and 0.2 mg/mL BSA in a 10 mm pathlength quartz cuvette in spectrofluorometer. The temperature was maintained at 23 °C with constant stirring. The samples were taken in triplicate, excited with 535 nm light, and the resulting emission from the donor (Cy3), and the sensitized emission from the acceptor (Cy5) was collected as a FRET scan from 550 – 700 nm. Ratiometric FRET efficiency was calculated from the background subtracted FRET spectra as follows:

Eq. 6

$$FRET_e = \frac{I_A}{I_A + I_D}$$

Where $FRET_e$ is the FRET efficiency calculated by normalizing the sensitized emission counts (I_A ; fluorescence intensity collected at 667 nm) with total number of photons collected in the donor and the acceptor channel ($I_A + I_D$; fluorescence intensity counts 667 nm + that collected at 568 nm). We did not estimate or apply correction factors for spectral cross-stalk (i.e. Cy3 emission spilling over into Cy5 channel or *vice versa*, or cross-excitation (i.e. 535 nm excitation light directly exciting Cy5 fluorophores) because following two reasons: 1) Equation (6) is a simple means to estimate ratiometric FRET. We are merely interested in monitoring the relative changes in the distance between the ends of the ssDNA upon binding to Rad52 rather than any precise value in nm. 2) A complete battery of correction factors will also include the careful estimation of proximity-based effects of bound protein molecules on the quantum efficiency of Cy3 and Cy5 fluorescence. This will involve additional experiments with Cy3- or Cy5- alone

labelled ssDNA of comparable lengths something which is beyond the scope of this present study. After an initial FRET scan of ssDNA alone, full length Rad52 or Rad52^{ΔC} were titrated in multiple steps gradually increasing the protein concentration. After each addition, FRET spectra were reacquired after an incubation period of 2 min allowing the binding reaction to reach equilibrium. The estimated $FRET_e$ was plotted against the concentration of Rad52 in decamer units (see supplementary figure 15) for ssDNA of different lengths. (dT)_{xx}, where xx refers to 45, 60, 80 or 97 nucleotides in length.

Single-molecule analysis of ssDNA-Rad52 wrapping

Single-molecule FRET measurements were performed on an inverted, objective based total-internal-reflection fluorescence microscope (TIR-FM; IX71 Olympus). TIR excitation was achieved through an oil-immersion objective (100X Olympus UplanApo Numerical Aperture 1.5). The sample was excited with a 532nm laser and the emitted fluorescence was split into two channels using Optosplit II (Cairn-Research, UK); Cy3 and Cy5 emissions was collected on the half chip of the same electron-multiplying charge-coupled device (EM-CCD) camera (iXon Ultra DU-897U-CS0). Images were acquired with 150 ms frame time. Donor and Acceptor images were subpixel registered in Fiji using an image of multicolor beads as fiducial markers. Next, background corrected intensity traces were extracted from immobilized single-molecule spots using Matlab scripts. FRET values were calculated for each fluorescent spot as a ratio between the acceptor intensity and the sum of the intensities of the donor and acceptor, corrected for crosstalk and cross excitation. The single-molecule FRET vs time traces were further analyzed in vbFRET⁷⁶ to ascribe FRET states and dwell times.

Glass coverslips for single-molecule imaging were cleaned and coated as follows: Briefly No. 1.5 thickness Gold Seal coverslips (22 x 60 x 0.17) were sonicated in deionized water, 2% Micro-90, 200-proof ethanol, and finally in 1N KOH all at 60 °C for 15 minutes each interspersed with exhaustive washes with deionized water. After drying under a stream of filtered N₂, the coverslips were first coated with Vectabond (1% v/v Vectabond in 1:18 methanol/acetic acid mixture) followed by washes with deionized water. Next, dried coverslips were incubated overnight with a freshly prepared 18:1 mixture of mPEG-SVA and Biotin-PEG-SVA in 0.1 M NaHCO₃. Nonspecific binding, which was measured by flowing in non-biotinylated Cy3- or Cy5- labelled DNA or by flowing in biotinylated fluorescent DNA substrate in absence of Neutravidin, was <1 %. All measurements were performed at room temperature in buffer (50 mM Tris-acetate pH 7.5, 50 mM KCl, 5 mM MgCl₂, 10% v/v glycerol, 1 mM DTT, and 0.2-0.5 mg/mL BSA) supplemented with oxygen scavenging solution i.e. 0.1 mg/mL glucose oxidase, 0.2 mg/mL catalase, and 0.4% (wt/wt) β -D-glucose (or 0.1 μ M PCD / 10 mM PCA), and 2.5 mM Trolox to reduce blinking. To a neutravidin (0.2 mg/mL) coated slide, biotin tagged partial duplexes (18 bp duplex region) with 3' dT overhangs and Cy3/Cy5 pairs (Supplementary Table 4) were added at 50-100 pM concentration to arrive at optimal single-molecule number- density of ~ 1.2 spot per 3.5 μ m² area on the EM-CCD chip. Basal FRET (w/o any protein) trajectories were recorded for 5 minutes. Following which 0.9-1.3 nM (decamer) Rad52 or Rad52^{AC} was introduced. After a 1-2 minute incubation period to re-achieve anoxic conditions and the Rad52-ssDNA binding to arrive at an equilibrium, 5 min long videos were again recorded from the same field of view.

Steady-State analysis of ssDNA handoff to Rad51

To 30 nM ssDNA (5'-Cy5-(dT)₈₀-Cy3-3') taken in buffer (50 mM Tris-acetate pH 7.5, 50 mM KCl, 5 mM MgCl₂, 10% v/v glycerol, 1 mM DTT, and 0.2 mg/mL BSA) in a 10 mm path-length cuvette full length Rad52 or Rad52^{ΔC} was added until the FRET peaked indicating the end of binding reaction. Next, increasing concentration of Rad51 ($\pm$ 2.5 mM ATP) was added to the reaction in a stepwise manner. As Rad51 concentration was increased, average FRET begins to decrease especially in the case where ATP was present in the reaction. These results are consistent with the model that ATP-dependent polymerization of Rad51 extracts ssDNA away from Rad52-wrapped configuration and thereby linearizing it and resulting in a loss in FRET.

Stopped Flow analysis of DNA binding and Rad51 filament formation

Stopped-flow experiments were performed using an Applied Photophysics SX20 instrument (Applied Photophysics Inc.) at 25 °C in reaction buffer (50 mM Tris-acetate pH 7.5, 5 mM MgCl₂, 10% glycerol, and 1mM DTT). Reactions from individual syringes were rapidly mixed and fluorescence data were collected. The respective mixing schemes are denoted by cartoon schematics within the figure panel. Five individual shots were averaged for each experiment. All experiments were repeated a minimum of 3 times. In FRET experiments, samples were excited at 535 nm and Cy5 emission was captured using a 645 nm band-pass filter. For the Cy3-dT₍₈₀₎-Cy5 Rad52-Rad51 interactions, Rad52-DNA and Rad52-DNA-Rad51 interactions, experiments were performed with 100 nM each of Rad52 (decamer) and Cy3-dT₍₈₀₎-Cy5 ssDNA substrates, along with 3 μ M Rad51. All concentrations denoted are post-mixing.

References

- 1 Symington, L. S., Rothstein, R. & Lisby, M. Mechanisms and regulation of mitotic recombination in *Saccharomyces cerevisiae*. *Genetics* **198**, 795-835, doi:10.1534/genetics.114.166140 (2014).
- 2 Sung, P. & Klein, H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* **7**, 739-750, doi:10.1038/nrm2008 (2006).
- 3 Symington, L. S. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol Mol Biol Rev* **66**, 630-670, table of contents, doi:10.1128/MMBR.66.4.630-670.2002 (2002).
- 4 van Gent, D. C., Hoeijmakers, J. H. & Kanaar, R. Chromosomal stability and the DNA double-stranded break connection. *Nat Rev Genet* **2**, 196-206, doi:10.1038/35056049 (2001).
- 5 Stewart, M. D. *et al.* Homologous Recombination Deficiency: Concepts, Definitions, and Assays. *Oncologist* **27**, 167-174, doi:10.1093/oncolo/oyab053 (2022).
- 6 Marechal, A. & Zou, L. RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. *Cell Res* **25**, 9-23, doi:10.1038/cr.2014.147 (2015).
- 7 Wold, M. S. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66**, 61-92, doi:10.1146/annurev.biochem.66.1.61 (1997).
- 8 Deshpande, I. *et al.* Structural Basis of Mec1-Ddc2-RPA Assembly and Activation on Single-Stranded DNA at Sites of Damage. *Mol Cell* **68**, 431-445 e435, doi:10.1016/j.molcel.2017.09.019 (2017).
- 9 Navadgi-Patil, V. M. & Burgers, P. M. The unstructured C-terminal tail of the 9-1-1 clamp subunit Ddc1 activates Mec1/ATR via two distinct mechanisms. *Mol Cell* **36**, 743-753, doi:10.1016/j.molcel.2009.10.014 (2009).
- 10 Andriuskevicius, T., Kotenko, O. & Makovets, S. Putting together and taking apart: assembly and disassembly of the Rad51 nucleoprotein filament in DNA repair and genome stability. *Cell Stress* **2**, 96-112, doi:10.15698/cst2018.05.134 (2018).
- 11 San Filippo, J., Sung, P. & Klein, H. Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem* **77**, 229-257, doi:10.1146/annurev.biochem.77.061306.125255 (2008).

- 12 Antony, E. *et al.* Srs2 disassembles Rad51 filaments by a protein-protein interaction triggering ATP turnover and dissociation of Rad51 from DNA. *Mol Cell* **35**, 105-115, doi:10.1016/j.molcel.2009.05.026 (2009).
- 13 Pokhrel, N. *et al.* Dynamics and selective remodeling of the DNA-binding domains of RPA. *Nat Struct Mol Biol* **26**, 129-136, doi:10.1038/s41594-018-0181-y (2019).
- 14 Pokhrel, N. *et al.* Monitoring Replication Protein A (RPA) dynamics in homologous recombination through site-specific incorporation of non-canonical amino acids. *Nucleic Acids Res* **45**, 9413-9426, doi:10.1093/nar/gkx598 (2017).
- 15 Heyer, W. D., Ehmsen, K. T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* **44**, 113-139, doi:10.1146/annurev-genet-051710-150955 (2010).
- 16 Davies, A. A. *et al.* Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol Cell* **7**, 273-282, doi:10.1016/s1097-2765(01)00175-7 (2001).
- 17 Paques, F. & Haber, J. E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **63**, 349-404, doi:10.1128/MMBR.63.2.349-404.1999 (1999).
- 18 Mortensen, U. H., Bendixen, C., Sunjevaric, I. & Rothstein, R. DNA strand annealing is promoted by the yeast Rad52 protein. *Proc Natl Acad Sci U S A* **93**, 10729-10734, doi:10.1073/pnas.93.20.10729 (1996).
- 19 Prado, F., Cortes-Ledesma, F., Huertas, P. & Aguilera, A. Mitotic recombination in *Saccharomyces cerevisiae*. *Curr Genet* **42**, 185-198, doi:10.1007/s00294-002-0346-3 (2003).
- 20 Davis, A. P. & Symington, L. S. RAD51-dependent break-induced replication in yeast. *Mol Cell Biol* **24**, 2344-2351, doi:10.1128/MCB.24.6.2344-2351.2004 (2004).
- 21 Bartsch, S., Kang, L. E. & Symington, L. S. RAD51 is required for the repair of plasmid double-stranded DNA gaps from either plasmid or chromosomal templates. *Mol Cell Biol* **20**, 1194-1205, doi:10.1128/MCB.20.4.1194-1205.2000 (2000).
- 22 Stasiak, A. Z. *et al.* The human Rad52 protein exists as a heptameric ring. *Curr Biol* **10**, 337-340, doi:10.1016/s0960-9822(00)00385-7 (2000).
- 23 Grimme, J. M. *et al.* Human Rad52 binds and wraps single-stranded DNA and mediates annealing via two hRad52-ssDNA complexes. *Nucleic Acids Res* **38**, 2917-2930, doi:10.1093/nar/gkp1249 (2010).

- 24 Shinohara, A., Shinohara, M., Ohta, T., Matsuda, S. & Ogawa, T. Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing. *Genes Cells* **3**, 145-156, doi:10.1046/j.1365-2443.1998.00176.x (1998).
- 25 Kagawa, W. *et al.* Functional analyses of the C-terminal half of the *Saccharomyces cerevisiae* Rad52 protein. *Nucleic Acids Res* **42**, 941-951, doi:10.1093/nar/gkt986 (2014).
- 26 Singleton, M. R., Wentzell, L. M., Liu, Y., West, S. C. & Wigley, D. B. Structure of the single-strand annealing domain of human RAD52 protein. *Proc Natl Acad Sci U S A* **99**, 13492-13497, doi:10.1073/pnas.212449899 (2002).
- 27 Kagawa, W. *et al.* Crystal structure of the homologous-pairing domain from the human Rad52 recombinase in the undecameric form. *Mol Cell* **10**, 359-371, doi:10.1016/s1097-2765(02)00587-7 (2002).
- 28 Kinoshita, C. *et al.* The cryo-EM structure of full-length RAD52 protein contains an undecameric ring. *FEBS Open Bio*, doi:10.1002/2211-5463.13565 (2023).
- 29 Welty, S. *et al.* RAD52 is required for RNA-templated recombination repair in post-mitotic neurons. *J Biol Chem* **293**, 1353-1362, doi:10.1074/jbc.M117.808402 (2018).
- 30 McDevitt, S., Rusanov, T., Kent, T., Chandramouly, G. & Pomerantz, R. T. How RNA transcripts coordinate DNA recombination and repair. *Nat Commun* **9**, 1091, doi:10.1038/s41467-018-03483-7 (2018).
- 31 Mazina, O. M., Keskin, H., Hanamshet, K., Storici, F. & Mazin, A. V. Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair. *Mol Cell* **67**, 19-29 e13, doi:10.1016/j.molcel.2017.05.019 (2017).
- 32 Shi, I. *et al.* Role of the Rad52 amino-terminal DNA binding activity in DNA strand capture in homologous recombination. *J Biol Chem* **284**, 33275-33284, doi:10.1074/jbc.M109.057752 (2009).
- 33 Sugiyama, T., New, J. H. & Kowalczykowski, S. C. DNA annealing by RAD52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proc Natl Acad Sci U S A* **95**, 6049-6054, doi:10.1073/pnas.95.11.6049 (1998).
- 34 Krejci, L. *et al.* Interaction with Rad51 is indispensable for recombination mediator function of Rad52. *J Biol Chem* **277**, 40132-40141, doi:10.1074/jbc.M206511200 (2002).
- 35 Song, B. & Sung, P. Functional interactions among yeast Rad51 recombinase, Rad52 mediator, and replication protein A in DNA strand exchange. *J Biol Chem* **275**, 15895-15904, doi:10.1074/jbc.M910244199 (2000).

- 36 Deng, X. *et al.* Human replication protein A-Rad52-single-stranded DNA complex: stoichiometry and evidence for strand transfer regulation by phosphorylation. *Biochemistry* **48**, 6633-6643, doi:10.1021/bi900564k (2009).
- 37 Plate, I. *et al.* Interaction with RPA is necessary for Rad52 repair center formation and for its mediator activity. *J Biol Chem* **283**, 29077-29085, doi:10.1074/jbc.M804881200 (2008).
- 38 Park, M. S., Ludwig, D. L., Stigger, E. & Lee, S. H. Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *J Biol Chem* **271**, 18996-19000, doi:10.1074/jbc.271.31.18996 (1996).
- 39 Khade, N. V. & Sugiyama, T. Roles of C-Terminal Region of Yeast and Human Rad52 in Rad51-Nucleoprotein Filament Formation and ssDNA Annealing. *PLoS One* **11**, e0158436, doi:10.1371/journal.pone.0158436 (2016).
- 40 Seong, C. *et al.* Molecular anatomy of the recombination mediator function of *Saccharomyces cerevisiae* Rad52. *J Biol Chem* **283**, 12166-12174, doi:10.1074/jbc.M800763200 (2008).
- 41 Saotome, M. *et al.* Structural Basis of Homology-Directed DNA Repair Mediated by RAD52. *iScience* **3**, 50-62, doi:10.1016/j.isci.2018.04.005 (2018).
- 42 Arai, N. *et al.* Vital roles of the second DNA-binding site of Rad52 protein in yeast homologous recombination. *J Biol Chem* **286**, 17607-17617, doi:10.1074/jbc.M110.216739 (2011).
- 43 Kagawa, W. *et al.* Identification of a second DNA binding site in the human Rad52 protein. *J Biol Chem* **283**, 24264-24273, doi:10.1074/jbc.M802204200 (2008).
- 44 Oshidari, R. *et al.* DNA repair by Rad52 liquid droplets. *Nat Commun* **11**, 695, doi:10.1038/s41467-020-14546-z (2020).
- 45 Yan, Z. *et al.* Rad52 Restrains Resection at DNA Double-Strand Break Ends in Yeast. *Mol Cell* **76**, 699-711 e696, doi:10.1016/j.molcel.2019.08.017 (2019).
- 46 Li, F. *et al.* Identification of an integrase-independent pathway of retrotransposition. *Sci Adv* **8**, eabm9390, doi:10.1126/sciadv.abm9390 (2022).
- 47 Paulsen, B. S. *et al.* Ectopic expression of RAD52 and dn53BP1 improves homology-directed repair during CRISPR-Cas9 genome editing. *Nat Biomed Eng* **1**, 878-888, doi:10.1038/s41551-017-0145-2 (2017).

- 48 Xu, Y., Zhou, H., Post, G., Zan, H. & Casali, P. Rad52 mediates class-switch DNA recombination to IgD. *Nat Commun* **13**, 980, doi:10.1038/s41467-022-28576-2 (2022).
- 49 Tokmina-Lukaszewska, M. *et al.* The Role of Mass Spectrometry in Structural Studies of Flavin-Based Electron Bifurcating Enzymes. *Front Microbiol* **9**, 1397, doi:10.3389/fmicb.2018.01397 (2018).
- 50 Gong, Z., Ye, S. X. & Tang, C. Tightening the Crosslinking Distance Restraints for Better Resolution of Protein Structure and Dynamics. *Structure* **28**, 1160-1167 e1163, doi:10.1016/j.str.2020.07.010 (2020).
- 51 Ma, E. *et al.* Rad52-Rad51 association is essential to protect Rad51 filaments against Srs2, but facultative for filament formation. *Elife* **7**, doi:10.7554/eLife.32744 (2018).
- 52 New, J. H., Sugiyama, T., Zaitseva, E. & Kowalczykowski, S. C. Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature* **391**, 407-410, doi:10.1038/34950 (1998).
- 53 Sugiyama, T. & Kowalczykowski, S. C. Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation. *J Biol Chem* **277**, 31663-31672, doi:10.1074/jbc.M203494200 (2002).
- 54 Gibb, B. *et al.* Protein dynamics during presynaptic-complex assembly on individual single-stranded DNA molecules. *Nat Struct Mol Biol* **21**, 893-900, doi:10.1038/nsmb.2886 (2014).
- 55 Saotome, M., Saito, K., Onodera, K., Kurumizaka, H. & Kagawa, W. Structure of the human DNA-repair protein RAD52 containing surface mutations. *Acta Crystallogr F Struct Biol Commun* **72**, 598-603, doi:10.1107/S2053230X1601027X (2016).
- 56 Sugawara, N., Wang, X. & Haber, J. E. In vivo roles of Rad52, Rad54, and Rad55 proteins in Rad51-mediated recombination. *Mol Cell* **12**, 209-219, doi:10.1016/s1097-2765(03)00269-7 (2003).
- 57 Sung, P. Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. *J Biol Chem* **272**, 28194-28197, doi:10.1074/jbc.272.45.28194 (1997).
- 58 Jaffe, E., Bornstein, P. & Disteché, C. M. Mapping of the thrombospondin gene to human chromosome 15 and mouse chromosome 2 by in situ hybridization. *Genomics* **7**, 123-126, doi:10.1016/0888-7543(90)90528-3 (1990).

- 59 Greene, E. C. DNA Sequence Alignment during Homologous Recombination. *J Biol Chem* **291**, 11572-11580, doi:10.1074/jbc.R116.724807 (2016).
- 60 Qi, Z. *et al.* DNA sequence alignment by microhomology sampling during homologous recombination. *Cell* **160**, 856-869, doi:10.1016/j.cell.2015.01.029 (2015).
- 61 Lee, J. Y. *et al.* DNA RECOMBINATION. Base triplet stepping by the Rad51/RecA family of recombinases. *Science* **349**, 977-981, doi:10.1126/science.aab2666 (2015).
- 62 Rangunathan, K., Liu, C. & Ha, T. RecA filament sliding on DNA facilitates homology search. *Elife* **1**, e00067, doi:10.7554/eLife.00067 (2012).
- 63 Hsieh, P., Camerini-Otero, C. S. & Camerini-Otero, R. D. The synapsis event in the homologous pairing of DNAs: RecA recognizes and pairs less than one helical repeat of DNA. *Proc Natl Acad Sci U S A* **89**, 6492-6496, doi:10.1073/pnas.89.14.6492 (1992).
- 64 Kuppa, S. *et al.* Rtt105 regulates RPA function by configurationally stapling the flexible domains. *Nat Commun* **13**, 5152, doi:10.1038/s41467-022-32860-6 (2022).
- 65 Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606-1619, doi:10.1016/S0006-3495(00)76713-0 (2000).
- 66 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296, doi:10.1038/nmeth.4169 (2017).
- 67 Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70-82, doi:10.1002/pro.3943 (2021).
- 68 Jamali, K., Kimanius, D., Scheres, S.H.W. A graph neural network approach to automated model building in cryo-EM maps. *arXiv*, doi:https://doi.org/10.48550/arXiv.2210.00006 (2022).
- 69 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).
- 70 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
- 71 Patterson, A., Zhao, Z., Waymire, E., Zlotnick, A. & Bothner, B. Dynamics of Hepatitis B Virus Capsid Protein Dimer Regulate Assembly through an Allosteric Network. *ACS Chem Biol* **15**, 2273-2280, doi:10.1021/acscchembio.0c00481 (2020).

- 72 Berry, L., Patterson, A., Pence, N., Peters, J. W. & Bothner, B. Hydrogen Deuterium Exchange Mass Spectrometry of Oxygen Sensitive Proteins. *Bio Protoc* **8**, doi:10.21769/BioProtoc.2769 (2018).
- 73 Kavan, D. a. M., P. MSTools—Web based application for visualization and presentation of HXMS data. *International Journal of Mass Spectrometry* **302**, 53-58 (2010).
- 74 Burns, E. E., Keith, B. K., Refai, M. Y., Bothner, B. & Dyer, W. E. Proteomic and biochemical assays of glutathione-related proteins in susceptible and multiple herbicide resistant *Avena fatua* L. *Pestic Biochem Physiol* **140**, 69-78, doi:10.1016/j.pestbp.2017.06.007 (2017).
- 75 Kim, C. *et al.* Core Protein-Directed Antivirals and Importin beta Can Synergistically Disrupt Hepatitis B Virus Capsids. *J Virol* **96**, e0139521, doi:10.1128/JVI.01395-21 (2022).
- 76 Bronson, J. E., Fei, J., Hofman, J. M., Gonzalez, R. L., Jr. & Wiggins, C. H. Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data. *Biophys J* **97**, 3196-3205, doi:10.1016/j.bpj.2009.09.031 (2009).

CHAPTER FIVE

RTT105 REGULATES RPA FUNCTION BY CONFIGURATIONALLY STAPLING THE
FLEXIBLE DOMAINS

Contribution of Authors and Co-Authors

Manuscript in Chapter 5

Author: Sahiti Kuppa

Contributions: Purified proteins, generated fluorescent protein, designed, and performed the steady-state and pre-steady state experiments. Reviewed and edited the manuscript.

Author: Jaigeeth Deveryshetty

Contributions: Purified proteins, generated fluorescent protein, designed, and performed the steady-state and pre-steady state experiments. Reviewed and edited the manuscript.

Co-Author: Rahul Chadda

Contributions: Reviewed and edited the manuscript.

Co-Author: Jenna R. Mattice

Contributions: Performed and analyzed the HDX-MS and XLMS experiments. Reviewed and edited the manuscript.

Co-Author: Nilisha Pokhrel

Contributions: Purified proteins, generated fluorescent protein, designed, and performed the steady-state and pre-steady state experiments. Reviewed and edited the manuscript.

Co-Author: Vikas Kaushik

Contributions: Purified proteins, generated fluorescent protein, designed, and performed the steady-state and pre-steady state experiments. Reviewed and edited the manuscript.

Co-Author: Angela Patterson

Contributions: Performed and analyzed the HDX-MS and XLMS experiments.

Reviewed and edited the manuscript.

Co-Author: Nalini Dhingra

Contributions: performed and analyzed the in vivo experiments.
Reviewed and edited the manuscript.

Co-Author: Sushil Pangen

Contributions: Performed and analyzed the C-trap experiments.
Reviewed and edited the manuscript.

Co-Author: Marisa K. Sadauskas

Contributions: Purified proteins, generated fluorescent protein, designed, and performed the steady-state and pre-steady state experiments. Reviewed and edited the manuscript.

Co-Author: Sajad Shiekh

Contributions: Performed and analyzed single molecule FRET measurements of GQ-unwinding.
Reviewed and edited the manuscript.

Co-Author: Hamza Balci

Contributions: performed and analyzed single molecule FRET measurements of GQ-unwinding.
Reviewed and edited the manuscript.

Co-Author: Taekjip Ha

Contributions: Performed and analyzed the C-trap experiments.
Reviewed and edited the manuscript.

Co-Author: Xiaolan Zhao

Contributions: Performed and analyzed the in vivo experiments. Reviewed and edited the manuscript.

Co-Author: Brian Bothner

Contributions: Supervised Experimentes. Reviewed and edited the manuscript.

Co-Author: Edwin Antony

Contributions: Directed the project, designed experiments. Wrote and edited the manuscript.

Manuscript Information

Sahiti Kuppa, Jaigeeth Deveryshetty, Rahul Chadda, Jenna R. Mattice,

Nilisha Pokhrel, Vikas Kaushik, Angela Patterson, Nalini Dhingra Sushil Pangen Marisa K.

Sadauskas, Sajad Shiekh , Hamza Balci Taekjip Ha, Xiaolan Zhao , Brian Bothner²& Edwin

Antony

Nature Communications

Status of Manuscript: [Put an x in one of the options below, delete this]

☐ Prepared for submission to a peer-reviewed journal

☐ Officially submitted to a peer-reviewed journal

☐ Accepted by a peer-reviewed journal

☒ Published in a peer-reviewed journal

Nature Communications

Volume 13, Issue 1

DOI: 10.1038/s41467-022-32860-6



Rtt105 regulates RPA function by configurationally stapling the flexible domains

Received: 7 March 2022

Accepted: 18 August 2022

Published online: 02 September 2022



Sahiti Kuppa¹, Jaigeeth Deveryshetty^{1,10}, Rahul Chadda^{1,10}, Jenna R. Mattice^{2,10}, Nilisha Pokhrel^{3,9}, Vikas Kaushik¹, Angela Patterson², Nalini Dhingra⁴, Sushil Pangen⁵, Marisa K. Sadauskas¹, Sajad Shiekh⁶, Hamza Balci⁶, Taekjip Ha^{5,7,8}, Xiaolan Zhao⁴, Brian Bothner² & Edwin Antony^{1,3} ✉

Replication Protein A (RPA) is a heterotrimeric complex that binds to single-stranded DNA (ssDNA) and recruits over three dozen RPA-interacting proteins to coordinate multiple aspects of DNA metabolism including DNA replication, repair, and recombination. Rtt105 is a molecular chaperone that regulates nuclear localization of RPA. Here, we show that Rtt105 binds to multiple DNA binding and protein-interaction domains of RPA and configurationally staples the complex. In the absence of ssDNA, Rtt105 inhibits RPA binding to Rad52, thus preventing spurious binding to RPA-interacting proteins. When ssDNA is available, Rtt105 promotes formation of high-density RPA nucleoprotein filaments and dissociates during this process. Free Rtt105 further stabilizes the RPA-ssDNA filaments by inhibiting the facilitated exchange activity of RPA. Collectively, our data suggest that Rtt105 sequesters free RPA in the nucleus to prevent untimely binding to RPA-interacting proteins, while stabilizing RPA-ssDNA filaments at DNA lesion sites.

Replication protein A (RPA) is an essential single-stranded DNA (ssDNA) binding protein that coordinates almost all aspects of DNA metabolism including DNA replication, repair, and recombination¹. The myriad functions of RPA are facilitated through high affinity interactions with ssDNA ($K_D < 10^{-10}$ M) and physical interactions with over three dozen proteins². With respect to the functions of RPA in double strand DNA break repair, current models posit that recognition of ssDNA by RPA triggers the DNA damage response (DDR). Coating of RPA on ssDNA serves as a nucleoprotein hub for the recruitment of several proteins including kinases that trigger downstream DNA repair pathways³. For example, yeast Mec1 or human ATM/ATR kinases are recruited on to the RPA-ssDNA substrate⁴. How such protein-protein

interactions with RPA are prevented in the absence of ssDNA remains a mystery.

Functional specificity for the myriad DNA metabolic functions of RPA is facilitated by a complex series of oligosaccharide/oligonucleotide binding (OB) folds/domains that are distributed across a heterotrimeric structural complex made of RPA70, RPA32 and RPA14 subunits (Fig. 1a)^{5,6}. OB-domains A, B, C, and F are in RPA70⁷. OB-domain D and a winged-helix (wh) domain are situated in RPA32, and RPA14 houses OB-domain E. OB-F and the wh-domain are primarily responsible for mediating protein-protein interactions and are thus termed Protein-Interaction-Domains (PID^{70A} and PID^{32C}, respectively)^{8,9}. OB-A, B, C and D contribute most to ssDNA binding and are termed DNA-Binding-

¹Department of Biochemistry and Molecular Biology, Saint Louis University School of Medicine, St. Louis, MO 63104, USA. ²Department of Chemistry and Biochemistry, Montana State University, Bozeman, MT 59717, USA. ³Department of Biological Sciences, Marquette University, Milwaukee, WI 53201, USA.

⁴Molecular Biology Department, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. ⁵Department of Biophysics, Johns Hopkins University, Baltimore, MD 21218, USA. ⁶Department of Physics, Kent State University, Kent, OH 44242, USA. ⁷Department of Biophysics and Biophysical Chemistry, Johns Hopkins University, Baltimore, MD 21205, USA. ⁸Howard Hughes Medical Institute, Baltimore, MD 21205, USA. ⁹Present address: Laronde Bio, Cambridge, MA, USA.

¹⁰These authors contributed equally: Jaigeeth Deveryshetty, Rahul Chadda, Jenna Mattice. ✉ e-mail: edwin.antony@health.slu.edu

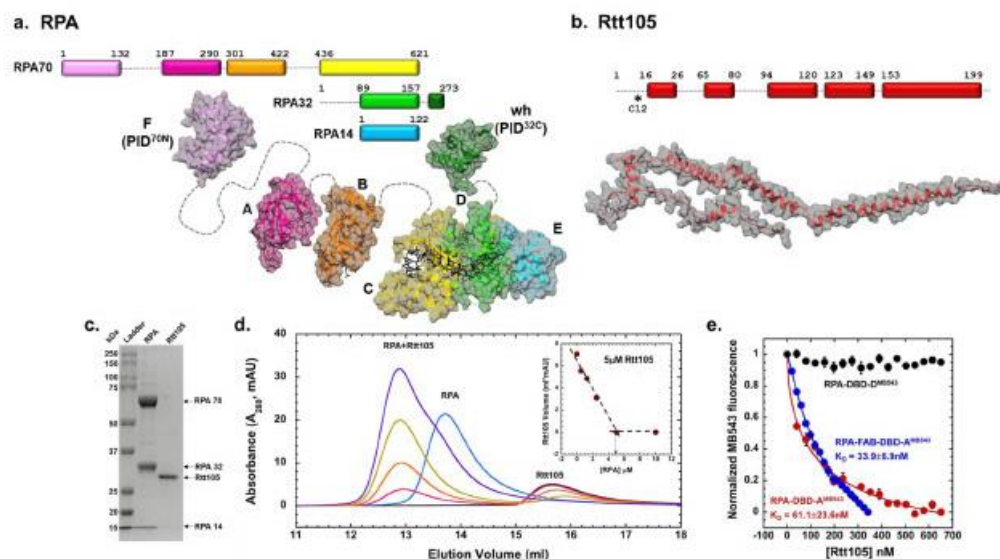


Fig. 1 | Rtt105 forms a stoichiometric complex with RPA. **a** Schematic of the OB-domains of the three *S. cerevisiae* RPA subunits is depicted along with the structures of the individual domains. The domains are spread apart for clarity and the dotted lines denote the flexible linkers connecting the domains. ssDNA bound to the individual DBDs (A, B, C and D) is shown as sticks. DBDs C, D and E are from the cryoEM structure of *S. cerevisiae* RPA (PDB: 6LS2). Homology models for DBD-F (PDB: 5N8A), DBD-A & DBD-B (PDB: 1JMC), and winged helix domains (PDB: 4OU0) were built with Swiss-Model using the listed PDB files as template. DNA from 1JMC is docked onto DBD-A and B homology models. **b** AlphaFold (AF-P40063-F1) model of Rtt105 is shown with predicted α -helices. *Denotes Cys12 used to generate fluorescently labeled Rtt105. **c** SDS-PAGE analysis of recombinantly purified RPA and

Rtt105. **d** Size exclusion chromatographic (SEC) analysis of Rtt105 (5 μ M) and RPA (0–5 μ M) show RPA, Rtt105, and the Rtt105-RPA complex migrating as defined species. Inset shows the change in the Rtt105 peak volume as a function of RPA concentration and stoichiometric complex formation between Rtt105 and RPA. **e** Changes in RPA-DBD-A^{MB543} or RPA-DBD-B^{MB543} fluorescence was measured after addition of increasing concentrations of Rtt105. Rtt105 binding does not influence the fluorescence of RPA-DBD-A^{MB543} but quenches RPA-DBD-B^{MB543} fluorescence. F-A-B labeled at DBD-A^{MB543} binding to Rtt105 reveals a higher binding affinity compared to full-length RPA. Data were fit as described in the Methods to obtain K_D values. Mean values $\pm$ SE from $n = 3$ biologically independent experiments are shown.

Domains (DBDs). The roles of OB-F and OB-E in binding ssDNA are uncertain. The DBDs and PIDs are connected by flexible linkers of varying lengths and thus can adopt a wide variety of structural assemblies on and off the DNA³, collectively termed ‘configurational arrangements’ in RPA. Conformational changes upon ssDNA binding and protein interactions are also quite extensive within the individual domains¹⁰.

In response to cellular DNA metabolic needs, RPA is shuttled from the cytoplasm into the nucleus. In *Saccharomyces cerevisiae*, the nuclear localization of RPA is facilitated by a chaperone-like protein called Regulator of Ty1 Transposition 105 (Rtt105)¹¹. In higher eukaryotes, RPA-interacting protein (RPAIN or RIPA) appears to be the functional homolog. *RTT105* was originally discovered in a genome-wide screen as interacting with genes associated with genome maintenance pathways¹². Rtt105 physically interacts with RPA in the cytoplasm and mediates nuclear localization in complex with Kap95, a key karyopherin-beta protein that belongs to the importin-exportin family of nuclear transport protein complexes¹¹. Consequently, deletion of Rtt105 or mutations that perturb the RPA-Rtt105 interaction give rise to defects in cell growth, DNA repair, and recombination^{13,14}.

Rtt105 is a 24 kDa protein (Fig. 1b) and proposed to stabilize an extended conformation of RPA where all the OB-domains are stretched out^{11,14,15}. In addition to the nuclear transport functions, recent studies propose Rtt105 to enhance the ssDNA binding properties of RPA^{11,14}. Depending on the number of DBDs bound to ssDNA, RPA can adopt distinct DNA binding modes where varying number of nucleotides are occluded by the RPA complex¹⁶. The DBDs of RPA are also dynamic in

nature and thus, while the protein is macroscopically bound to the ssDNA, each domain can exist in microscopically unbound states^{17–19}. We recently showed that RPA can be envisioned to interact with ssDNA as two functional halves: a dynamic half consisting of the OB-F, DBD-A and DBD-B domains, and a less-dynamic half with DBD-C, DBD-D and OB-E (trimerization core or Tri-C) contributing more to the stability of RPA-ssDNA interactions²⁰.

RPA binds to ssDNA with very high affinity ($K_D < 10^{-10}$ M)^{17,21}. Yet, a substantial increase in RPA binding affinity to ssDNA was shown in the presence of Rtt105 which led to a model where Rtt105 binds and promotes a configuration of RPA that extends out its DBDs such that maximal DNA binding contacts are promoted^{11,14,15}. Rtt105 inherently does not possess DNA binding activity and RPA intrinsically binds ssDNA with very high affinity. More strikingly, Rtt105 is not bound to the RPA-ssDNA complex. Thus, the finding that Rtt105 enhances the DNA binding activity of RPA, while not part of the DNA-RPA complex is perplexing, given that experimental tools such as electrophoretic mobility shift analysis (EMSA) do not have the required resolution to detect changes in sub-nanomolar biomolecular interactions. Furthermore, the rationale for Rtt105 stabilizing an extended conformation of RPA whilst having to transport it across the nuclear pore is also puzzling. To address these disparities, we undertook a detailed structural and mechanistic investigation of the Rtt105-RPA complex.

Here, we show that Rtt105 contacts multiple regions in RPA and conformationally compacts the multiple DBDs and PIDs through a configurational stapling mechanism. Shorter ssDNA substrates (~15–35 nt) engage the Rtt105-RPA complex and promote remodeling of Rtt105, and a transition-state Rtt105-RPA-ssDNA complex is

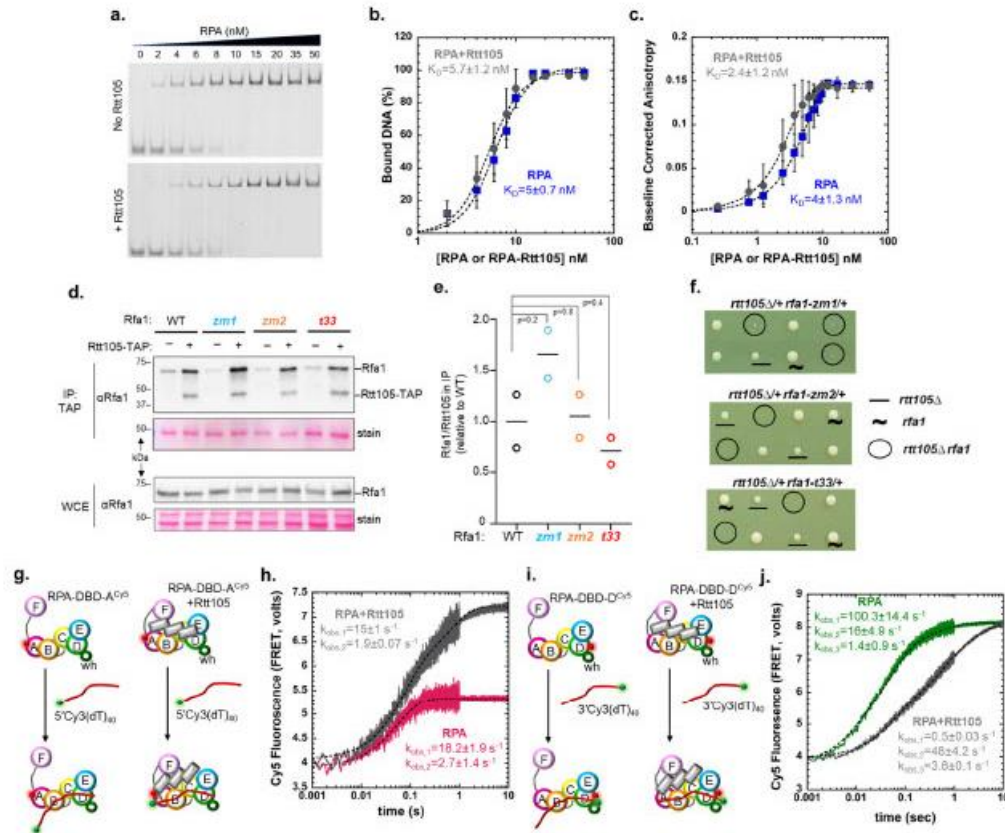


Fig. 2 | Rtt105 alters the configuration of DBD-A on ssDNA and kinetics of DBD-B binding. **a** EMSA of RPA binding to 5'-Cy3-(dT)₃₀ ssDNA and **b** quantitation shows no appreciable difference in DNA binding affinity in the absence ($K_D = 5 \pm 0.7$ nM) or presence of Rtt105 ($K_D = 5.7 \pm 1.2$ nM). **c** Fluorescence anisotropy measurements of RPA binding to 3'-FAM-(dT)₃₀ ssDNA shows no appreciable difference in DNA binding affinity in the absence ($K_D = 4 \pm 1.3$ nM) or presence of Rtt105 ($K_D = 2.4 \pm 1.2$ nM). **d** Rtt105 association with RPA is unaffected by its ssDNA binding ability. TAP-tagged Rtt105 was immunoprecipitated and RPA was detected with an anti-Rfa1 antibody. Representative immunoblots examining elute from the eluate (IP) and the whole cell extract (WCE) are shown. **e** Enrichment of Rfa1 present in the IP eluate relative to Rtt105 and normalized to wild type (WT). Mean and SEMs from two biological replicates are plotted. *P* value obtained from Student's *t*-test show that the differences in relative Rfa1 enrichments are not significant. **f** *rtt105Δ* sensitizes *rfa1* mutants with impaired ssDNA association. Representative tetrads of diploids heterozygous for indicated mutations are shown. **g, h** Stopped flow

fluorescence measurements of DNA binding were monitored by FRET-induced enhancement of Cy5 fluorescence upon Cy3 excitation. FRET pairs positioned on the 5'-DNA end (5'-Cy3-(dT)₄₀) and DBD-A of RPA (RPA-DBD-A^{Cy3}) show similar rates of DNA-induced changes $\pm$ Rtt105 ($k_{\text{obs},1} = 18.2 \pm 1.9$ s⁻¹, 15 ± 1 s⁻¹; $k_{\text{obs},2} = 2.7 \pm 1.4$ s⁻¹, 1.9 ± 0.07 s⁻¹; in the absence and presence of Rtt105, respectively), but Rtt105 markedly enhances the amplitude of Cy5 fluorescence. **i, j** Similar experiments performed with FRET pairs positioned on the 3'-DNA end (3'-Cy3-(dT)₄₀) and DBD-B of RPA (RPA-DBD-B^{Cy3}) show an Rtt105 induced reduction in the rate of DNA binding ($k_{\text{obs},1} = 100.3 \pm 14.4$ s⁻¹, 0.5 ± 0.03 s⁻¹; $k_{\text{obs},2} = 16 \pm 4.9$, 48 ± 4.2 s⁻¹; $k_{\text{obs},3} = 1.4 \pm 0.9$, 3.6 ± 0.1 s⁻¹ in the absence and presence of Rtt105, respectively). But, unlike in **d**, both traces show similar changes in fluorescence amplitude. Representative stopped flow data averaged from seven to eight shots from one experiment are shown. For **b, c, h**, and **j**, mean values $\pm$ SE from *n* = 3 biologically independent experiments are denoted.

observed. Longer ssDNA substrates (>35 nt) promote the binding of multiple RPA molecules and lead to dissociation of Rtt105. Interestingly, Rtt105 promotes formation of high-density RPA nucleoprotein filaments on ssDNA. In the absence of ssDNA, we find that Rtt105 blocks interactions between RPA and RPA-interacting proteins (RIPs) such as the homologous recombination mediator Rad52. ssDNA binding to Rtt105-RPA remodels the Rtt105-RPA-ssDNA complex and promotes RIP engagement. Posttranslational modifications of RPA further contribute to the remodeling of the RPA-Rtt105 complex. Finally, Rtt105 blocks the facilitated exchange activity of RPA thereby contributing to the stability of RPA nucleoprotein filaments. Thus, we

here present new functional roles for Rtt105 where it sequesters RPA in the nucleus, drives formation of densely packed RPA filaments, and serves as a negative regulator by blocking spurious interactions with RPA-interacting proteins in the absence of ssDNA.

Results

Rtt105 forms a stoichiometric complex with RPA

Rtt105 physically interacts with RPA and was shown to copurify as a complex⁴¹. To obtain the stoichiometry of the complex, using recombinantly purified proteins (Fig. 1c) we analyzed formation of the Rtt105-RPA complex as a function of RPA concentration using size exclusion

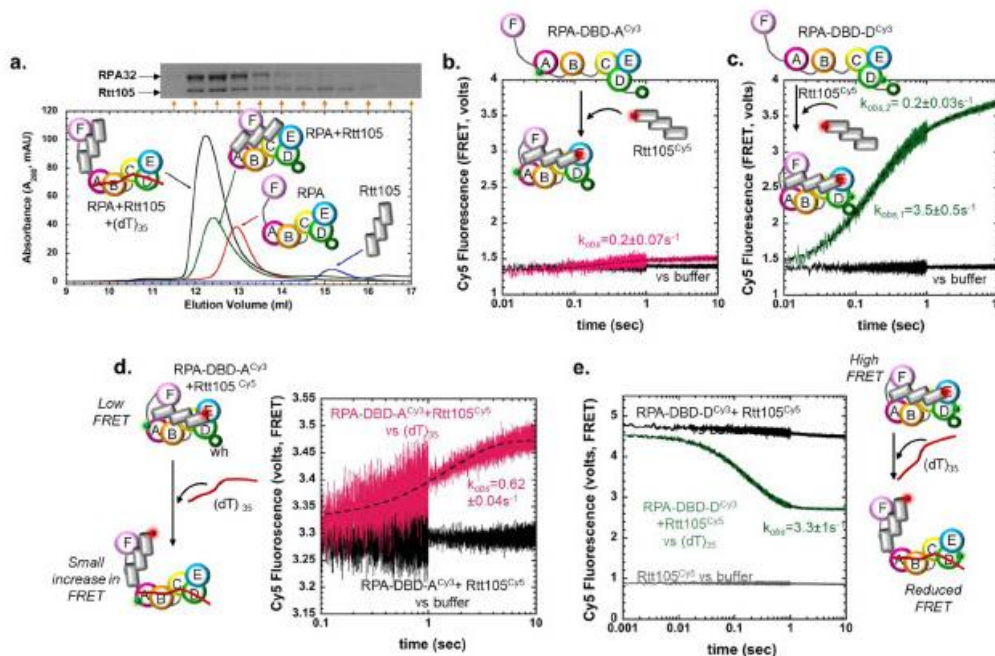


Fig. 3 | Rtt105 forms a higher order complex with RPA bound to short ssDNA substrates. a Size exclusion chromatography (SEC) analysis of Rtt105, RPA, and Rtt105-RPA show complex formation between the proteins. Addition of equimolar amounts of (dT)₃₅ drive formation of a higher order Rtt105-RPA-(dT)₃₅ complex. SDS page analysis of the SEC fractions (insert) shows RPA and Rtt105 comigrating as a complex in the presence of DNA. A small fraction of free Rtt105 is also observed. b The kinetics of RPA binding to Rtt105 was measured using FRET with Cy5-labeled Rtt105 and either RPA-DBD-A^{Cy5} or c RPA-DBD-D^{Cy5}. Cy3 positioned on DBD-D produces a larger FRET-induced Cy5 enhancement compared to DBD-A, suggesting that the N terminal region of Rtt105 resides closer to DBD-D of RPA. The data for RPA-DBD-A in b fits well to a single exponential model ($k_{obs} = 0.2 \pm 0.07 \text{ s}^{-1}$) whereas the RPA-DBD-D data is better described by a double exponential ($k_{obs,1} =$

$3.5 \pm 0.5 \text{ s}^{-1}$, $k_{obs,2} = 0.2 \pm 0.03 \text{ s}^{-1}$). These data suggest faster, initial engagement of Rtt105 towards the trimerization core (Tri-C) of RPA. Stopped flow experiments show transition from either a preformed d RPA-DBD-A^{Cy5}-Rtt105^{Cy5} or e RPA-DBD-D^{Cy5}-Rtt105^{Cy5} complex to remodeled Rtt105^{Cy5}-RPA^{Cy5}-(dT)₃₅ complexes upon addition of ssDNA. A small increase in Cy5 fluorescence is observed for the RPA-DBD-A^{Cy5}-Rtt105^{Cy5} complex. In contrast, a sharp decrease in Cy5 signal is observed for the RPA-DBD-D^{Cy5}-Rtt105^{Cy5} complex. The Cy5 fluorescence does not reach the baseline signal for Rtt105^{Cy5}. The remodeling data fits well to a single step model ($k_{obs} = 3.3 \pm 0.96 \text{ s}^{-1}$). Representative stopped flow data averaged from seven to eight shots from one experiment are shown. Mean values $\pm$ SE of the observed rates from $n = 3$ biologically independent experiments are denoted.

chromatography (SEC). We observe concentration-dependent stoichiometric binding between RPA and Rtt105 (Fig. 1d). To determine a K_D value for the interaction, we followed the change in fluorescence in RPA upon complex formation with Rtt105. We used fluorescent versions of RPA where either DBD-A or DBD-D carry a site-specific environment-sensitive MBS43 fluorophore^{17,22,23}. Upon binding to Rtt105 we observe a change in fluorescence for RPA-DBD-A^{MBS43}, but not for RPA-DBD-D^{MBS43} (Fig. 1e). Rtt105 binds to RPA with high affinity ($K_D = 61.1 \pm 23.6 \text{ nM}$; Fig. 1e) and the selective change in RPA-DBD-A^{MBS43} fluorescence suggests that at least a part of Rtt105 is situated/bound in proximity to DBD-A. Furthermore, an F-A-B version of RPA containing just OB-F, DBD-A and DBD-B interacts with Rtt105 (Supplementary Fig. 1a). Thus, the region around OB-F and DBD-A likely makes physical contacts with Rtt105. The F-A-B domain binds to Rtt105 with higher affinity ($K_D = 33.9 \pm 6.9 \text{ nM}$; Fig. 1e). These data suggest that one or more of the Rtt105 binding sites in the F-A-B region is occluded in the context of the full-length RPA complex. A recent study suggested that a Val106 to Ala substitution in OB-F (RPA70 subunit) was sufficient to abolish interaction with Rtt105⁵⁴. However, in our analysis, this single point substitution in RPA (RPA^{V106A}) is not sufficient to perturb the interaction (Supplementary Fig. 1b) suggesting that additional contacts must exist between Rtt105 and RPA.

Rtt105 alters the ssDNA bound configurations of the DNA binding domains (DBDs) of RPA

Recent studies have shown that Rtt105 enhances the ssDNA binding activity of RPA¹¹. More perplexingly, Rtt105 is not bound to the RPA-ssDNA complex in pull-down experiments. Thus, how Rtt105 modulates the DNA binding activity of RPA, while not part of the RPA-ssDNA complex, is unresolved. In both these earlier studies, control experiments performed in the absence of Rtt105 showed only ~50% of RPA bound to ssDNA oligonucleotides at equimolar concentrations^{8,14}. These data are inconsistent with the high-affinity and stoichiometric DNA binding behavior of RPA^{17,22}. Thus, we first repeated these experiments, under similar experimental conditions, and do not observe this behavior. RPA binds stoichiometrically to a 5'-Cy5-(dT)₃₀ oligonucleotide in electrophoretic mobility band-shift analysis (EMSA) and a preformed Rtt105-RPA complex does not influence the ssDNA binding activity of RPA ($K_D = 5 \pm 0.7$ and $5.7 \pm 1.2 \text{ nM}$ for RPA and RPA-Rtt105, respectively; Fig. 2a, b). Since EMSA experiments may not have the resolution to tease apart subtle differences in high-affinity biomolecular interactions, we used fluorescence anisotropy to further quantitate binding. RPA and the RPA-Rtt105 complex bound stoichiometrically to a 5'-FAM-(dT)₃₅ ssDNA substrate and we do not see a measurable Rtt105-induced

enhancement of RPA-ssDNA binding activity ($K_D = 4 \pm 1.3$ and 2.4 ± 1.2 nM for RPA and RPA-Rtt105, respectively; Fig. 2c). Nevertheless, in both these experiments, the high affinity stoichiometric interactions between RPA and ssDNA make it experimentally difficult to decipher a role for Rtt105 in modulating the overall binding affinity of RPA to ssDNA.

These results suggest that to a large degree, RPA association with Rtt105 is a biochemical feature that can be functionally distinct from its ssDNA binding activity. This notion of two separable features of RPA is supported by our *in vivo* data, where we tested the cellular effects of deletion of Rtt105 along with three distinct DNA binding mutants of RPA: *zml* (K494A) and *zm2* (N492D, K494R, K494R) mutants of RPA carry amino acid substitutions close to the Zn²⁺-finger binding domains in DBD-C and have reduced ssDNA binding activity²⁴. *t33* is a well characterized RPA mutant with a S373P substitution in DBD-B^{32,26}. Immunoprecipitation of Rtt105 from cells expressing these mutant variants of RPA show that the physical interaction between the two proteins is not perturbed by these mutations in RPA (Fig. 2d, e). From a genetic point view, the lack of measurable influence on global RPA-DNA binding by Rtt105 (Fig. 2b, c) predicts that mutants impairing RPA binding to ssDNA in cells lacking Rtt105 should have additive genetic interactions. Indeed, we find that RPA mutations that displayed reduced ssDNA binding activities were additive when combined with *rtt105* null cells when assayed for growth (Fig. 2f). These *in vivo* data support the idea that RPA-Rtt105 interactions and RPA-ssDNA binding are two separable functions.

Rtt105 influences the ssDNA binding kinetics and configurations of individual DBDs of RPA

Since RPA has four DBDs, its interaction with ssDNA can be defined in terms of macroscopic binding and microscopic dynamics of DBDs^{27,28}. Each DBD has intrinsic on/off ssDNA binding properties that describe the kinetics of their interactions. Thus, at any given point, RPA as a complex can remain stably bound to ssDNA while one or more DBDs can be remodeled or displaced². Such dynamic interactions allow RPA to interact with other proteins while bound to ssDNA¹⁰. Since the influence of Rtt105 on the equilibrium ssDNA binding properties of RPA cannot be reliably measured because of the macroscopic binding interactions, we next tested whether the kinetics of RPA-ssDNA interactions (especially the DBDs) were influenced by Rtt105. We used fluorescent RPA carrying Cy5 on either DBD-A or DBD-D and captured the kinetics of binding to (dT)₄₀ oligonucleotides labeled with Cy3 at either the 5' or 3' end. Changes in Förster resonance energy transfer (FRET) induced Cy5 fluorescence were monitored by exciting Cy3 in a stopped flow fluorometer (Fig. 2g–j). Since the DBDs of RPA bind ssDNA with defined polarity^{27,28}, with DBD-A and DBD-D residing close to the 5' and 3' ends, respectively, the change in fluorescence in each experiment reflects the configurational changes of the respective DBDs with respect to their cognate DNA termini positions². These data show that DBD-A binds with similar rates to the 5' end-labeled DNA in the absence or presence of Rtt105 (Fig. 2g, h). However, the fluorescence signal amplitude in the presence of Rtt105 is twice that observed for RPA alone (Fig. 2h). In contrast, Rtt105 reduces the rate of DBD-D binding to the 3' end-labeled DNA (Fig. 2i, j), while the signal amplitudes remain similar (Fig. 2j). These data suggest that Rtt105 differentially influences the two DBDs of RPA and support a model where Rtt105 could drive formation of specific configurations RPA (described below).

Rtt105-RPA-ssDNA complexes are detected on short DNA substrates

The experiments above show that Rtt105 influences the configurations of the DBDs of RPA but does not reveal whether Rtt105 remains in complex with RPA in the presence of ssDNA. To directly test formation of the Rtt105-RPA-ssDNA complex we performed SEC analysis of the

individual proteins and their complexes. Rtt105 (24 kDa) and RPA (114 kDa) migrate as distinct species due to large differences in their molecular weight and migrate together as a complex when pre-mixed (Fig. 3a). However, when equimolar amounts of (dT)₃₅ are added to the complex, Rtt105 primarily remains bound as a higher order Rtt105-RPA-(dT)₃₅ complex. A small fraction of free Rtt105 and RPA-(dT)₃₅ complexes are also observed (Fig. 3a and Supplementary Fig. 1c, d). These data show that the configurational changes in RPA observed in the stopped flow experiments (Fig. 2) are induced while both Rtt105 and DNA are simultaneously bound to RPA. We propose that the RPA-Rtt105 complex binds ssDNA and both proteins are likely reconfigured as the DBDs engage onto ssDNA. Thus, the interactions between RPA and Rtt105 are different in the DNA bound versus unbound states. Based on the changes in observed fluorescence between the ends of the DNA and the DBDs, and the higher Rtt105 binding affinity for the F-A-B half of RPA, we propose that upon ssDNA binding Rtt105 interactions with RPA are likely shifted towards the F-A-B part of RPA. This model suggests an active repositioning of Rtt105 on RPA upon ssDNA binding.

To experimentally capture this process, we generated fluorescently labeled Rtt105 to perform FRET analysis of the RPA-Rtt105 complex. Rtt105 has two non-conserved native Cys residues situated adjacently (Cys12 and Cys13, Supplementary Fig. 2a). Substitution of either or both Cys residues to Ser does not change the secondary structure, RPA binding, or RPA remodeling activity of Rtt105 (Supplementary Fig. 2). Thus, we generated a single Cys version of Rtt105 by converting Cys13 to Ser and fluorescently labeled Rtt105^{Cys} with Cy5 using maleimide chemistry (Supplementary Fig. 2e, f). Generation of a functional fluorescent Rtt105 enabled us to investigate the assembly, remodeling, and disassembly of the Rtt105-RPA complex using stopped flow FRET. First, we measured the kinetics of Rtt105^{Cys} binding to RPA-DBD-A^{Cy3} or RPA-DBD-D^{Cy3} by exciting Cy3 and monitoring the changes in Cy5 fluorescence. A small increase in FRET-induced Cy5 fluorescence is observed when Rtt105^{Cys} binds to RPA-DBD-A^{Cy3} (Fig. 3b), but a robust signal change is observed for RPA-DBD-D^{Cy3} (Fig. 3c). These data suggest that in the absence of DNA, the N-terminal region of Rtt105 (where the Cy5 is positioned) is situated closer to DBD-D. The data for DBD-A^{Cy3} and Rtt105^{Cys} binding fits to a single-step model ($k_{obs} = 0.2 \pm 0.07$ s⁻¹; Fig. 3b) whereas the DBD-D^{Cy3} and Rtt105^{Cys} data fits better to a two-step model ($k_{obs,1} = 3.5 \pm 0.5$ s⁻¹, $k_{obs,2} = 0.2 \pm 0.03$ s⁻¹; Fig. 3c). Conservative interpretation of the kinetics suggests faster initial binding of the N-terminal region of Rtt105 closer to DBD-D followed by slower binding/rearrangements to other regions of RPA.

Next, to obtain the kinetics of ssDNA-induced remodeling of the Rtt105-RPA complex, we preformed either the low-FRET RPA-DBD-A^{Cy3}:Rtt105^{Cys} or high-FRET RPA-DBD-D^{Cy3}:Rtt105^{Cys} or complex and monitored the change in Cy5 fluorescence upon addition of (dT)₃₅ ssDNA. When the preformed RPA-DBD-A^{Cy3}:Rtt105^{Cys} complex is mixed with (dT)₃₅, a small enhancement in Cy5 fluorescence is observed ($k_{obs} = 0.62 \pm 0.04$ s⁻¹, Fig. 3d). In the corollary experiment, a rapid, single-step transition from the high-FRET to low-FRET state is observed when the RPA-DBD-D^{Cy3}:Rtt105^{Cys} complex is mixed with (dT)₃₅ ($k_{obs} = 3.3 \pm 1.5$ s⁻¹; Fig. 3e). However, the fluorescence signal does not decrease to the baseline Rtt105^{Cys} level. These data further support a model where Rtt105 does not completely dissociate from the RPA-(dT)₃₅ complex but is remodeled such that the N-terminal region of Rtt105 is moved away from DBD-D upon DNA binding. This movement likely encompasses movement of Rtt105 away from the trimerization core (DBD-C, RPA32, and RPA14; Tri-C) and towards the F-A-B region (OB-F, DBD-A and DBD-B) of RPA. Since ssDNA binding to RPA causes significant configurational changes, we next tested if free Rtt105 in solution can bind to preformed RPA-ssDNA complexes. FRET pairs were positioned on RPA (either RPA-DBD-A^{Cy3} or RPA-DBD-D^{Cy3}) and Rtt105 (Rtt105^{Cys}). Stoichiometrically bound RPA^{Cy3}-(dT)₃₅ complexes were

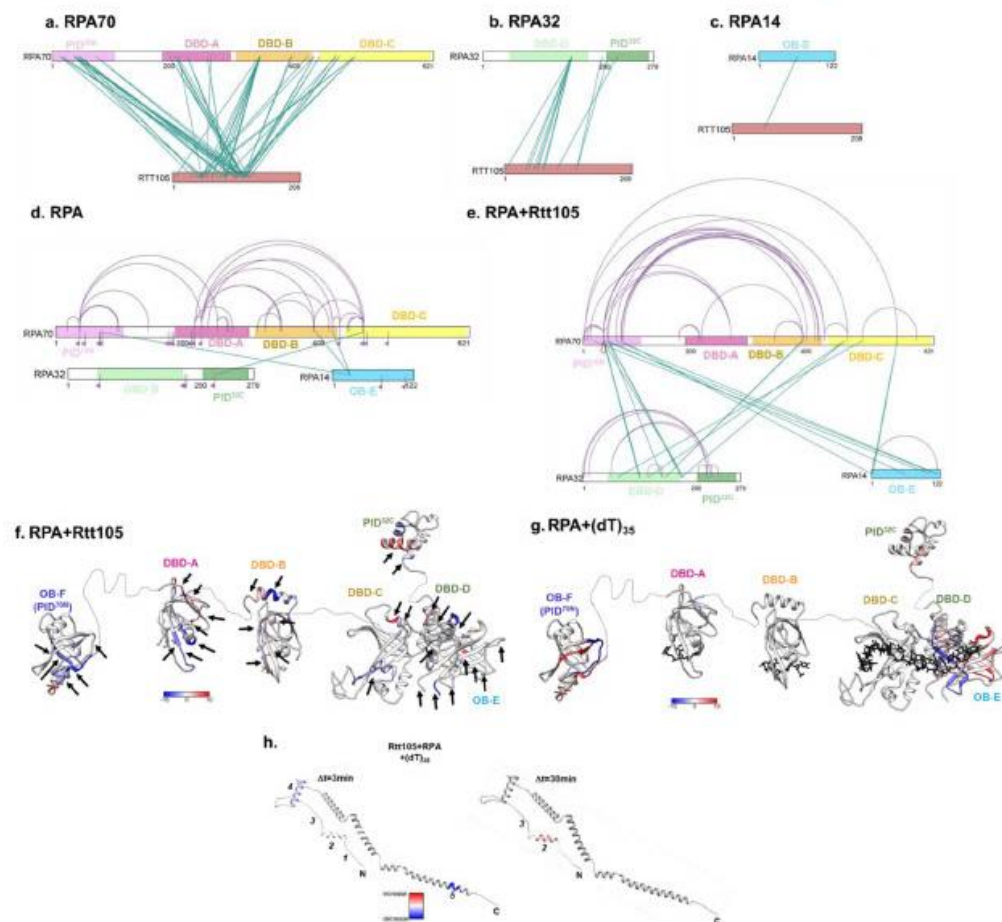


Fig. 4 | Rtt105 configurationally staples RPA through multiple contacts with DBDs and PIDs. Crosslinking mass spectrometry (XL-MS) analysis of the RPA-Rtt105 complex reveals crosslinks between Rtt105 and a RPA70, b RPA32, and c RPA14 subunits. Crosslinks are observed in all OB-folds and the BC-linker. Inter-subunit (green lines) and intra-subunit (purple traces) crosslinks in RPA are shown in the d absence or e presence of Rtt105, respectively. An increase in both inter- and intra-subunit crosslinks are observed upon Rtt105 binding to RPA. The inverted flag marks in d denote the sites of mono-crosslinks and show that they have no partner sites available in proximity for the BS3 crosslinker due to the extended

configuration of RPA in the absence of Rtt105. Hydrogen-deuterium mass spectrometry (HDX-MS) analysis of f RPA-Rtt105 and g RPA-(dT)₂₅ complexes show net changes in deuterium uptake or loss (ΔHDX) in almost all domains of RPA. The arrows point to ΔHDX that are unique to the Rtt105-RPA complex and not observed in the RPA-(dT)₂₅ complex. h ΔHDX changes in Rtt105 bound to RPA in the absence and presence of (dT)₂₅ are denoted in the AlphaFold derived model of Rtt105. The regions of change are denoted 1–5. The ΔHDX changes upon DNA binding suggests remodeling of Rtt105. The scales (blue to red) denote the net ΔHDX.

preformed and mixed with equimolar amounts of Rtt105^{Cys}. In both cases, only a very small change in Cys fluorescence is observed (Supplementary Fig. 3a, b) suggesting that Rtt105 poorly binds to preformed RPA-DNA complexes. We do note that when RPA is in complex with ssDNA, its F-A-B is likely more exposed (described below) and thus, within the small fraction where Rtt105 binding is observed, we see a faster rate of Rtt105 binding to the F-A-B region (Supplementary Fig. 3c).

Rtt105 configurationally staples RPA through contacts with multiple domains in RPA

A structure of the Rtt105-RPA complex is not available and our efforts to obtain one using cryo-electron microscopy have not been

successful because of the dynamic nature of the interactions. Thus, to understand how Rtt105 interacts with RPA and decipher how the complex is remodeled by ssDNA, we performed cross-linking mass spectrometry (XL-MS) and hydrogen-deuterium exchange mass spectrometry (HDX-MS) analysis of the Rtt105-RPA complex in the absence or presence of ssDNA. We observe good peptide coverage for both RPA and Rtt105 in MS analysis (92% RPA70, 67% RPA32, 74% RPA14, and 67% Rtt105; Supplementary Fig. 4) and thus can comprehensively assess the global conformational changes.

For XL-MS, RPA or RPA-Rtt105 were cross-linked with bis(sulphosuccinimidyl)suberate (BS3), which reacts with primary amines in lysine side chains and the N-termini²⁹. RPA and Rtt105 cross-linked readily, as observed by the shift in the protein bands in SDS-PAGE

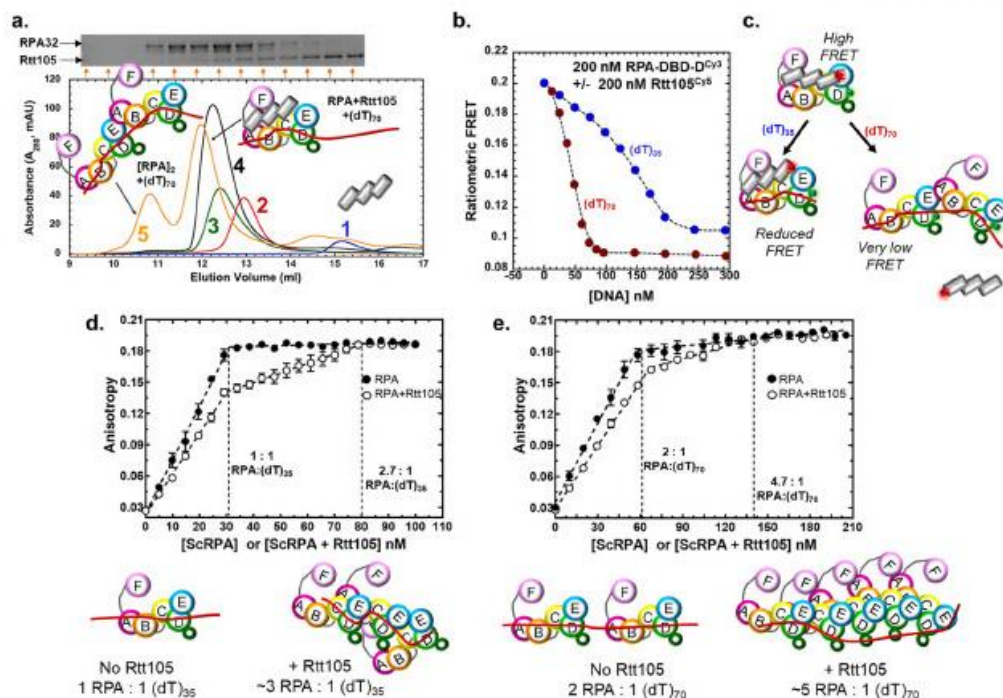


Fig. 5 | Binding of multiple RPA molecules on longer ssDNA substrates triggers the release of Rtt105 and formation of a high-density RPA nucleoprotein filament. **a** Size exclusion chromatography (SEC) analysis of the Rtt105-RPA complex in the presence of equimolar amounts of (dT)₃₅ and (dT)₇₀ show the presence of three distinct species. A larger [RPA]₂-(dT)₃₅ and smaller remodeled [Rtt105-RPA-(dT)₃₅] complex are observed. Rtt105 is displaced from the larger complex and seen as the third peak. The distribution of Rtt105 in the three peaks can be noted in the SDS-PAGE gel shown as inset. Traces 1, 2, 3 and 4 are shown as reference and denote Rtt105, RPA, Rtt105-RPA, and the Rtt105-RPA-(dT)₃₅ complex, respectively. **b** Preformed RPA-DBD-D^{Cy3}-Rtt105⁵⁹⁵ complexes were mixed with increasing concentrations of either (dT)₃₅ or (dT)₇₀ and the change in FRET was measured. At ~1:1 ratios of RPA:(dT)₃₅

the FRET signal is significantly reduced. At ~1:1 ratio and lower, an alternate reduced FRET state is observed for (dT)₃₅. **c** The reduced FRET state likely corresponds to the remodeled [Rtt105-RPA-(dT)₃₅] complex and the very low FRET state corresponds to the [RPA]₂-(dT)₃₅ state lacking Rtt105. **d** Fluorescence anisotropy experiments were performed with 30 nM (dT)₃₅ ssDNA by adding increasing amounts of RPA or the RPA-Rtt105 complex. On (dT)₃₅, RPA binds with 1:1 stoichiometry, but the Rtt105 increases the binding density of RPA to ~3 RPA:(dT)₃₅ molecule. **e** A similar phenomenon is observed on the longer (dT)₇₀ substrate where a higher density of RPA binding (~5 RPA:(dT)₇₀) is observed. Mean values ± SE from *n* = 3 biologically independent experiments are denoted.

analysis after BS3 addition (Supplementary Fig. 5). MS analysis of the crosslinked peptides yielded several linkage pairs for the RPA and RPA-Rtt105 complexes, respectively (Fig. 4a–e). These data mapped onto the individual subunits of RPA reveal extensive crosslinks between the N-terminal two-thirds of Rtt105 and all DBDs and PIDs in RPA (Fig. 4a–c). Multiple contacts to OB-F, DBD-A, DBD-B, and DBD-C (all in RPA70) are observed. The F-A and A-B flexible linkers show no crosslinks, but a few are observed in the B-C linker (Fig. 4a). A limited number of crosslinks are observed in RPA32 with one peptide in DBD-D making 9 contacts in Rtt105 (Fig. 4b). Finally, a single crosslink is observed in OB-E (RPA14; Fig. 4c). These data agree with our observation that Rtt105 interacts more closely with the F-A-B half of RPA (Fig. 1e).

A closer look at the crosslinked peptides in the individual DBDs reveal that the crosslinks in DBD-A and DBD-B are not in the DNA binding pockets (Supplementary Fig. 6a, b). In contrast, some of the crosslinked peptides in DBD-C and DBD-D overlap with the DNA binding region (Supplementary Fig. 6c, d). Based on the predicted structure and charge distribution of Rtt105 (Supplementary Fig. 6e), the elongated helical regions likely extend and bind across the many domains of RPA. Circular dichroism (CD) analysis of Rtt105 confirm the

helical nature of Rtt105 (Supplementary Fig. 2b). Due to the extensive crosslinks observed we propose a configurational stapling model for the RPA-Rtt105 interaction. In this model, Rtt105 can be envisioned sitting on the surface of RPA, interacting with multiple domains, and configurationally constraining (or stapling) the DBDs and PIDs that are connected by flexible linkers.

Comparison of XL-MS data for RPA in the absence or presence of Rtt105 lends further support to the configurational stapling model. In the absence of Rtt105, a total of four inter-subunits crosslinks are captured between the three RPA subunits (Fig. 4d). In contrast, Rtt105 binding induces 16 distinct inter-subunits crosslinks (Fig. 4e). These data show that the three subunits are brought in proximity upon Rtt105 binding. Analysis of the intra-subunit crosslinks further highlight the configurational stapling mechanism as the crosslinks between the individual DBDs and PIDs (within each subunit) are extensively increased in the presence of Rtt105 (Fig. 4d, e). Thus, in contrast to existing models that propose a stretching of the RPA domains upon Rtt105 binding^{43,45}, we show that the many DBDs and PIDs in RPA are compacted by Rtt105 through a configurational stapling mechanism.

To further assess the degree of configurational and conformational changes in RPA induced by Rtt105, we performed HDX-MS

analysis of RPA in the absence and presence of Rtt105 (Fig. 4f, g and Supplementary Figs. 7–11). In agreement with the XL-MS observations, changes in deuterium uptake/loss are observed in all the DBDs and PIDs of RPA. Most of the changes favor an increase in deuterium uptake (denoted as red) suggesting enhanced solvent exposure of RPA domains upon Rtt105 binding. To delineate whether Rtt105 binding overlaps with ssDNA binding to RPA, we compared the Δ HDX between the Rtt105-RPA and RPA-ssDNA complexes (Fig. 4f, g; regions marked with arrows are unique to Rtt105 binding). Thus, ssDNA ((dT)₃₅) binding remodels the Rtt105-RPA complex in agreement with the kinetic data.

Additional support for the remodeling of the RPA-Rtt105 complex by ssDNA is evident in the Δ HDX changes observed within Rtt105. Upon addition of DNA to the preformed Rtt105-RPA complex we observe deuterium uptake/release in five distinct regions, primarily around the N-terminal half (Fig. 4h and Supplementary Fig. 12). When assessed as a function of time, the remodeling of the N-terminal region of Rtt105 is observed at shorter time points. At later time points, the exchange patterns are preserved in only two of the original five regions. These data further support our model where ssDNA binding remodels the N-terminal region of Rtt105 away from Tri-C region of RPA and Rtt105 interactions with the RPA-ssDNA complex are maintained through interactions between Rtt105 and F-A-B region of RPA. The precise nature of the interactions will have to be established through future structural studies.

Length of ssDNA and assembly of multiple RPA molecules promote release of Rtt105

In pull-down and single-molecule DNA curtain analysis Rtt105 has been proposed to not remain in complex with RPA on ssDNA^{13,45}. Since we clearly observe a Rtt105-RPA-ssDNA complex on short ssDNA oligonucleotides (e.g. (dT)₃₅, Fig. 3a) we next tested whether ssDNA length influenced Rtt105 dissociation. Longer ssDNA substrates offer binding sites for more than one RPA molecule. In this scenario, we previously showed that DBD-A from one RPA can interact with DBD-E of the neighboring RPA while being stably bound to DNA through Tri-C-DNA interactions³⁰. We hypothesized that since the F-A-B region is in an alternate configuration when multiple RPA are bound on DNA, Rtt105 might fully dissociate from the RPA-ssDNA complex. Using SEC, we tested whether Rtt105 remained bound to RPA when (dT)₇₀ ssDNA was added to the reaction (Fig. 5a). The elution profile shows three distinct species: [RPA]₂-(dT)₇₀, Rtt105-RPA-(dT)₇₀, and free Rtt105. The larger, earlier eluting, species does not contain Rtt105 (Fig. 5a and Supplementary Fig. 13). Thus, assembly of multiple RPA molecules serves as a trigger for Rtt105 dissociation.

To further probe the relationship between DNA length, multiple RPA binding, and Rtt105 dissociation, we used the FRET signal between RPA-DBD-D⁹³ and Rtt105⁹⁵ and followed complex dynamics as a function of either (dT)₃₅ or (dT)₇₀ concentration (Fig. 5b). Mixing 200 nM each of RPA-DBD-D⁹³ and Rtt105⁹⁵ yields a high FRET state, and the signal decreases as ssDNA is added to the reaction. (dT)₇₀ drives dissociation of Rtt105 at lower DNA concentrations (~100 nM) and produces a very low FRET state. At these ratios (2:1, RPA:(dT)₇₀) multiple RPA are expected to be bound to a single DNA molecule. In contrast, similar experiments done with (dT)₃₅ results in formation of an alternate low-FRET [Rtt105:RPA:(dT)₃₅] state that corresponds to the remodeled complex where most Rtt105 remains bound, but in a remodeled state. Furthermore, the longer (dT)₇₀ substrate is 3-fold better at arriving at the low FRET state compared to (dT)₃₅ ($K_{1/2}$ = 130.4 ± 10.5 and 42.8 ± 0.37 nM for (dT)₃₅ and (dT)₇₀, respectively; Fig. 5b). Performing this experiment as a function of ssDNA length reveals that ssDNA >35 nt and longer are effective in promoting Rtt105 dissociation from RPA (Supplementary Fig. 14). Thus, when adequate ssDNA is present to saturate all the individual DBDs (≤ (dT)₃₅) within one RPA molecule, Rtt105 is

remodeled. Longer DNA (≥(dT)₃₅), and higher RPA:DNA ratios promote binding of multiple RPA molecules and drive Rtt105 dissociation (Fig. 5c).

Rtt105 drives formation of high-density RPA-ssDNA filaments

In single molecule DNA curtain experiments, Rtt105 was shown to drive formation of stretched RPA-ssDNA filaments¹¹. Since Rtt105 was proposed to stretch out the domains of RPA, the model invoked elongated RPA bound to longer ssDNA with Rtt105 not bound¹⁵. Our data show a compaction of RPA by Rtt105. To reconcile these differences, we wondered if Rtt105 promoted alternate DNA-bound RPA configurations. For example, if Rtt105 were bound to the F-A-B part of RPA, one could envision an RPA-ssDNA filament where multiple RPA are bound using only the Tri-C region. Since the Tri-C region is more stably bound to ssDNA²⁰, the resulting filament could be more rigid and elongated as observed in the DNA curtain experiments. To test this model, we used fluorescence anisotropy to monitor RPA binding to (dT)₃₅ or (dT)₇₀ ssDNA in the absence or presence of Rtt105. On the shorter (dT)₃₅ substrate, the signal saturates at 1:1 (RPA:DNA) in the absence of Rtt105 (Fig. 5d). In contrast, the ratio shifts to 2.7:1 (RPA:DNA) when the RPA-Rtt105 complex is used (Fig. 5d). On the longer (dT)₇₀ substrate, this effect is further exaggerated as ~5 molecules of RPA are loaded onto the ssDNA in the presence of Rtt105 compared to ~2 when no Rtt105 is present in the reaction (Fig. 5e). Thus, remarkably, the configuration of RPA on ssDNA formed in the presence of Rtt105 is stable and distinctly different with a higher density of RPA molecules bound to the ssDNA. Interestingly, the difference in the amplitude of the anisotropy signals at the 1:1 saturation point for both ssDNA substrates suggest that the hydrodynamic radius of the RPA-DNA complex in the absence of Rtt105 is larger. We propose that the RPA molecules formed in the presence of Rtt105 likely have the F-A-B region (the dynamic half of RPA) not bound to DNA.

In all the above experiments, we investigated modulation of the RPA-Rtt105 complex using short ssDNA oligonucleotides. To test whether our observations of RPA-Rtt105 remodeling were recapitulated on longer kilobase-long ssDNA we utilized C-trap analysis. Here lambda DNA (~48.5 kbp) is tethered between two beads and stretched to produce ssDNA (Supplementary Fig. 15a–c). Fluorescent proteins binding to DNA, remodeling, and/or dissociating are visualized as kymographs as a function of time. Rtt105⁹⁵ alone does not interact with DNA at either low or high concentrations and thus no fluorescent spots are observed (Supplementary Fig. 15d). When the Rtt105⁹⁵-RPA complex is premixed at 1:1 ratio and first incubated with the DNA in the first chamber and then moved to chamber containing buffer, a few spots of Rtt105⁹⁵ are occasionally encountered (Supplementary Fig. 15e, f). Even in these rare instances, Rtt105⁹⁵ dissociates from these spots with a $t_{1/2}$ ~5 s. Quantitation of these data over multiple experiments reveal a distribution of 1–4 binding/dissociation events per kymograph recorded (Supplementary Fig. 15f). These experiments provide direct visual evidence for engagement of Rtt105-RPA onto DNA and Rtt105 dissociation under conditions where multiple RPA can bind onto a single DNA lattice and form a nucleoprotein filament.

Rtt105 impedes facilitated exchange of RPA

Free RPA has been shown to exchange with RPA bound to ssDNA through a process called facilitated exchange (FE)³¹. This activity arises from the dynamic binding, rearrangement, and dissociation of one or more DBDs on the DNA and thus transiently exposed short segments of ssDNA allow free RPA to gain access. In this scenario, while multiple RPA molecules are bound on ssDNA, we expect only short segments DNA exposed at any given point during FE. The biological significance of FE is poorly understood but thought to contribute to RPA dynamics and RPA-promoted enhancement or impediment of DNA metabolic processes in the cell. FE could also be

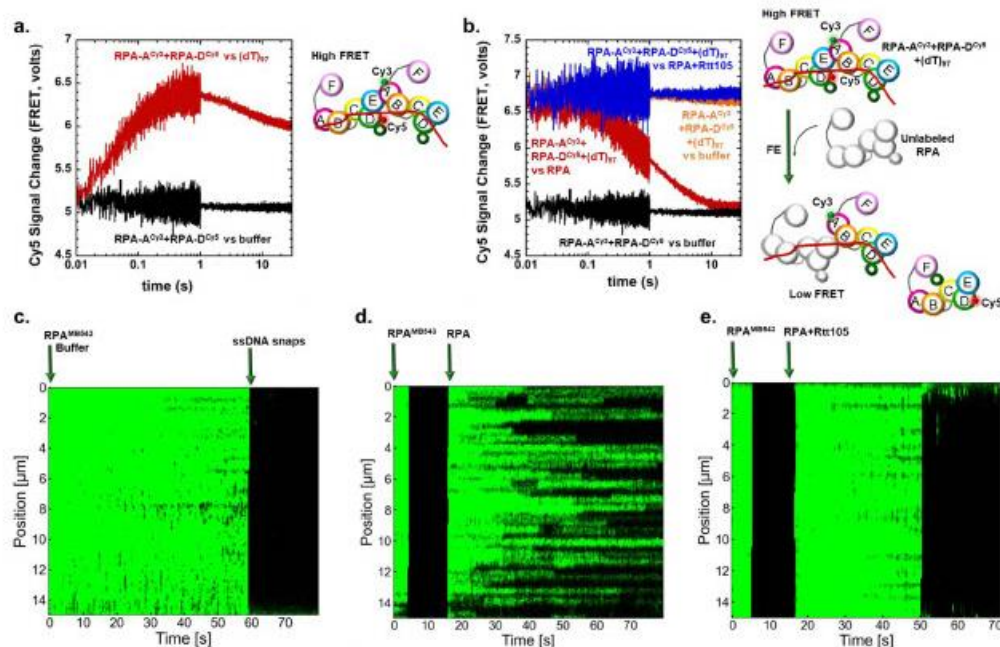


Fig. 6 | Rtt105 inhibits facilitated exchange activity of RPA. **a** A high FRET complex is formed when multiple molecules of RPA-DBD-A⁹³ and RPA-DBD-D⁹³ bind to a (dT)₉₇ substrate (red). FRET-induced Cy5 fluorescence is not observed in the absence of DNA (black). **b** Facilitated exchange (FE) activity was measured by mixing preformed [RPA-DBD-A⁹³:RPA-DBD-D⁹³:(dT)₉₇] complexes with unlabeled RPA (red) resulting in a loss of Cy5 fluorescence. Rtt105 inhibits the FE activity as preformed Rtt105/RPA complexes do not perturb the Cy5 signal (blue). Control experiments in the absence of unlabeled RPA (orange) or DNA (black) are also shown as reference for the high and low FRET-induced Cy5 fluorescence states. Models for FE are denoted on the right. Representative stopped flow data averaged from seven to eight shots from one experiment are shown. **c** C-trap experiments to

investigate FE of RPA were performed on ssDNA that was generated by mechanically unfolding lambda dsDNA (~48.5 kbp) in the optical trap and subsequently incubated with RPA-DBD-M8543 (10 nM) for 15–20 s and stably formed fluorescent RPA coated ssDNA nucleoprotein filaments are observed. **d** The fluorescent RPA nucleoprotein filament was subsequently moved to a channel containing unlabeled RPA (10 nM) and after a 10 s incubation in the dark FE can be observed. Spots where loss of fluorescence is observed are events where the unlabeled RPA have replaced fluorescent RPA molecules. **e** When challenged with unlabeled RPA+Rtt105, no loss in fluorescence is observed. Thus, Rtt105 inhibits the FE activity of RPA. C-trap data traces from one DNA molecule per condition are shown, but more than 10 DNA molecules were visualized and recorded for analysis.

envisioned as a counterproductive process in the cell where excess RPA perturbs the stability of an RPA nucleoprotein filament. We hypothesized that Rtt105 might influence FE by serving as a sink to sequester free RPA. To test this idea, we performed FE experiments using RPA labeled with FRET pairs. Binding of equimolar amounts of RPA-DBD-A⁹³ and RPA-DBD-D⁹³ (100 nM each) to (dT)₉₇ ssDNA (30 nM) results in a high FRET signal as multiple RPA molecules assemble (Fig. 6a). When excess unlabeled RPA is added to a preformed RPA-DBD-A⁹³:RPA-DBD-D⁹³:(dT)₉₇ complex, FE occurs leading to a loss in the FRET signal (Fig. 6b, red trace). In contrast, when challenged with RPA prebound to Rtt105, no FE is observed, and the FRET signal is not perturbed (Fig. 6b, blue trace). Thus, the short segments of free ssDNA available during FE does not allow binding of the Rtt105-RPA complex. To further investigate whether Rtt105 blocks FE on long RPA nucleoprotein filaments, we used C-trap analysis where lambda DNA (~48.5 kbp) was first coated with fluorescent RPA (RPA-DBD-M8543) and then moved over to a channel containing either buffer only, buffer with unlabeled RPA, or unlabeled RPA plus Rtt105 (Fig. 6c–e). Extensive FE is observed when unlabeled RPA is introduced as evidenced by the loss in fluorescence. When challenged with the RPA + Rtt105 complex, no FE is observed. Thus, Rtt105 blocks FE activity of RPA and thus likely contributes to the formation of stable RPA nucleoprotein filaments.

Complex formation with Rtt105 negatively regulates interactions with RPA-interacting proteins like Rad52
Since Rtt105 configurationally staples RPA and restricts the DBDs and PIDs, we wondered if the protein-protein interaction activities of RPA were regulated by Rtt105. This could prevent spurious RPA interactions with its myriad binding partners upon nuclear import from the cytoplasm and add to the functions of Rtt105 in genomic integrity. To test this hypothesis, we tested whether Rtt105 influenced RPA binding to Rad52, a mediator protein that physically interacts with RPA and functions to promote homologous recombination (HR)³². SEC has limitations in resolving larger molecular weight complexes, thus we used analytical ultracentrifugation (AUC) to monitor complex formation. Rtt105, RPA, Rad52 all sediment as distinct complexes (Fig. 7a). RPA and Rad52 form a complex in the absence or presence of ssDNA (dT)₃₅ (Fig. 7a). We and others have previously shown that Rad52 interacts with RPA on and off the DNA and these data agree with our knowledge of these interactions^{32,33}. The only key difference is the oligomeric state of yeast Rad52 which is widely considered to be a heptamer³⁴. In our ongoing cryoEM studies, *S. cerevisiae* Rad52 is a homodecamer (Deveryshetty and Antony, unpublished data) and the stoichiometries denoted here reflect this finding. In the absence of ssDNA, Rtt105 remains as a complex with RPA and prevents interaction with Rad52 (Fig. 7b). Thus, Rtt105 functions as a negative regulator of RPA-Rad52 interactions in the

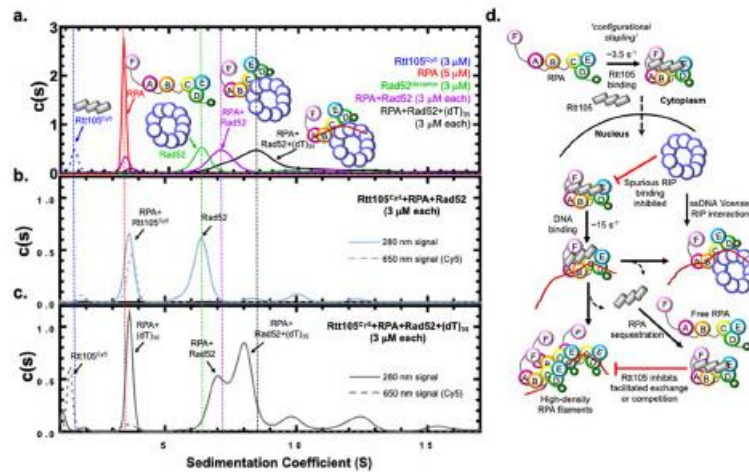


Fig. 7 | Rtt105 inhibits spurious RPA-Rad52 interactions in the absence of ssDNA. **a** Analytical ultracentrifugation (AUC) analysis of RPA, Rtt105^{WT}, Rad52, and the RPA-Rad52 and RPA-Rad52-(dT)₃₅ complexes show them sedimenting as distinct complexes. The 280 nm signals were monitored for all the proteins and complexes except Rtt105^{WT} which was monitored at 650 nm (dashed line). The horizontal dotted lines serve as a visual guide for the sedimentation positions of the denoted proteins across all the AUC experiment panels. **b** Rtt105^{WT}, RPA, and Rad52 were mixed in equimolar concentrations and analyzed. In the presence of Rtt105, Rad52 does not interact with RPA and sediments independently while RPA and Rtt105 remain as a complex. **c** When an equimolar amount of ssDNA [(dT)₃₅] is present, Rtt105 is displaced and sediments similar to free Rtt105. RPA and Rad52 partition into multiple RPA+DNA, RPA+Rad52, and RPA+Rad52+DNA complexes.

Importantly, Rtt105 is not part of most of these complexes. A small fraction of the Rtt105 is present along with the RPA-DNA complex. Representative data from a single run are shown, but repeated $n = 3$ with biological replicates. **d** Rtt105 configurationally staples RPA through interactions with multiple domains including the DNA binding and protein interaction domains. The complex is shuttled into the nucleus where promiscuous interactions with RPA binding proteins such as Rad52 are inhibited by Rtt105 in the absence of ssDNA. Availability of ssDNA triggers configurational and conformational changes within the RPA-Rtt105 complex and this higher order complex can interact with RPA binding proteins such as Rad52 or drive assembly of multiple RPA-bound DNA structures. Both these scenarios drive dissociation of Rtt105. Rtt105 can also sequester free RPA to prevent untimely facilitated exchange (FE) events as the RPA-Rtt105 complex cannot perform FE.

absence of DNA. In the presence of ssDNA Rtt105 is released and formation of RPA-DNA, RPA-Rad52, and RPA-Rad52-DNA complexes are observed (Fig. 7c). These data suggest that ssDNA acts as a licensing agent to release Rtt105 from RPA and promote RPA interactions with Rad52, and likely with other RPA-interacting proteins (Fig. 7d).

Discussion

Rtt105 functions akin to a chaperone-like protein to transport RPA into the nucleus, and deletion of Rtt105 reduces RPA availability for DNA metabolic roles including DNA repair and recombination^{11b-15}. Previous studies have proposed a model where Rtt105 stretches the many domains of RPA to enhance DNA binding¹⁵. Here, we present a modified model for Rtt105 mediated regulation of RPA. First, our findings show extensive contacts between Rtt105 and RPA that promotes a compaction of the four DNA binding and two protein-interaction domains of RPA. The CL-MS and HDX MS analysis show configurational and conformational changes in all domains and a few of the intervening flexible linkers. While a structure of Rtt105 or the RPA-Rtt105 complex is not yet available, the AlphaFold prediction renders a helical protein that could be stretched over a large radius (Fig. 1b). Thus, we propose that configurational stapling compacts RPA through extensive contacts with Rtt105 and promotes nuclear import (Fig. 7d). Rtt105 is an extensively charged protein and stretches of negatively charged patches are observed that likely bind close to the positively charged DNA binding pockets in RPA. In support of this model, we see differences in the DNA binding properties of the Rtt105-RPA complex.

While Rtt105 does not bind directly to DNA, Rtt105 complexed with RPA influences the kinetics and conformations/configurations of the DBDs on ssDNA. Positioning of site-specific fluorophores on either DBD-A or DBD-D enabled us to quantify how each of these domains

bind to DNA and are remodeled. Rtt105 reduces the rate of ssDNA binding by DBD-D by ~2-fold but does not change the final configuration of this domain on DNA. In contrast, Rtt105 does not change the rate of ssDNA binding to DBD-A, but significantly alters the final configuration. Thus, we envision Rtt105 situated on both the Tri-C and FAB halves of RPA in the absence of ssDNA as supported by the XL-MS analysis (Fig. 4). The region close to Tri-C occludes the DNA binding site in DBD-D (possibly in DBD-C as well) and is remodeled in the presence of ssDNA. The Rtt105 region close to the FAB region likely does not block DNA binding but restrains the configurational freedom of those domains. In the [Rtt105-RPA-(dT)₃₅] complex, ssDNA must be bound close to the Tri-C region. Further evidence for this model arises from the dissociation of Rtt105 on ssDNA longer than (dT)₃₅ where multiple RPA can bind. Here, DBD-A (RPA70) from one RPA contacts the OB-E (RPA14) of the neighboring RPA³⁰. This interaction likely triggers the dissociation of Rtt105. Moreover, Rtt105 also promotes formation of RPA nucleoprotein filaments that have a higher density of RPA molecules bound. This property renders the RPA nucleoprotein filament more stable and explains the additional stretching of ssDNA observed in DNA curtain experiments¹¹. In the cell, formation of such RPA nucleoprotein platforms triggers the DNA damage response³ and Rtt105 might contribute to such functions by not releasing RPA until longer stretches of ssDNA become exposed.

Contradictory to earlier studies^{11a}, we do not see observable differences in the K_D for RPA ssDNA interactions in the presence of Rtt105. This is not surprising as RPA will be macroscopically bound with high affinity to ssDNA while select DBDs can undergo dynamic rearrangements¹⁸. Moreover, because of the high ssDNA binding affinity ($K_D < 10^{-10}$ M), enhancements of DNA binding affinity cannot be measured using traditional bulk ensemble measurements. Our in vivo

studies show a distinction between functions associated with Rtt105 and DNA interactions of RPA. Loss of Rtt105 (*Artt105*) or reduction in RPA-ssDNA interactions (*rfa1 zm1, zm2*, or *t33* mutations) shows defects in cell growth and a combination of *Artt105* and *rfa1* mutations result in severe growth defects. The additive nature of these phenotypes again suggests that Rtt105 does not directly influence the macroscopic ssDNA binding properties of RPA, but regulates it through either affecting nuclear cytoplasmic shuttling, RIP interactions, or processes such as facilitated exchange and the differences in the density of RPA bound to ssDNA. Rtt105 contributes to the levels of RPA in the nucleus, however, no severe defects in global DNA synthesis are observed under non-stressed conditions¹³. In contrast, at perturbed replication forks, the lack of Rtt105 becomes important for RPA loading²¹. In this scenario, longer ssDNA intermediates are generated and thus, the inhibition of FFE by Rtt105 and formation of higher density RPA binding could explain the formation of more stable RPA filaments.

Since RPA interactions with Rtt105 occur in the cell well before the availability of ssDNA, we hypothesize that posttranslational modifications of RPA (and/or Rtt105), such as phosphorylation by kinases, might influence the interaction. In support of this idea, a phosphomimetic of RPA carrying a Ser to Asp substitution at position 178 in RPA70 (RPA^{S178D}) shows reduced binding to Rtt105 (Supplementary Fig. 16). Rtt105 has also been implicated in the resolution of G-quadruplex structures by RPA³⁵. In vitro, we do not see any direct Rtt105 enhancement of RPA promoted G-quadruplex unwinding (Supplementary Fig. 17). Thus, in this scenario, we propose that Rtt105 regulates the activity of RPA by controlling its nuclear transport, inhibiting RIP interactions in the absence of ssDNA, and regulating the availability of free RPA.

Rtt105 appears to block spurious interactions with RIPs in the absence of ssDNA (Fig. 7a). Such protein-protein interactions are primarily mediated by either OB-F in RPA70 (PID⁷⁰⁰) or the winged helix domain in RPA32 (PID³²⁵)¹⁸. A few interacting proteins such as Rad52 are composite binders that interact with a second region in RPA^{8,32,36,37}. In the case of Rad52, interactions between PID³²⁵ and a buried site in DBDs-A & B have been identified^{32,36–38}. Rad52 is a mediator protein that facilitates the nucleation of the Rad51 recombinase on RPA-coated ssDNA³⁹. Rad52 selectively remodels DBD-D in RPA32 and gains access towards the 3'-end of the resected ssDNA during homologous recombination^{17,28}. Upon Rtt105 binding, PID⁷⁰⁰ is positioned closer to DBD-B and DBD-C (Fig. 4e) and thus is likely occluded from interacting proteins that use this region for binding. Similarly, new crosslinks are observed between PID³²⁵ and the N-terminal region of RPA32 in the presence of Rtt105 (Fig. 4e). Thus, we propose that the configurational stapling of RPA by Rtt105 extends to the protein-protein interaction domains. Functionally, Rtt105 binding protects RPA from binding to potential RIPs upon nuclear localization. When ssDNA is exposed, removal/remodeling of Rtt105 occurs which in turn likely releases the protein-interaction domains and leads to recruitment of RIPs onto the RPA-coated ssDNA. The extent of release could also be facilitated by posttranslational modifications of RPA such as phosphorylation (Supplementary Fig. 16).

Finally, in higher eukaryotes, RPAIN (RPA interacting protein) or RIP α (RPA interacting protein α) is proposed to be the functional ortholog of Rtt105 as they share poor sequence similarity⁴⁴. The AlphaFold predicted structure of RPAIN shows a reasonable degree of resemblance to Rtt105 (Supplementary Fig. 18). RPAIN has also been proposed to function by enhancing the DNA binding activity of human RPA (hRPA)⁴⁴. Similar to our observations for Rtt105 and yeast RPA, we see complex formation between hRPA and RPAIN, but do not observe a stimulation of the macroscopic ssDNA binding properties (Supplementary Fig. 18). Interestingly, shorter DNA (dT)₃₅ can promote dissociation of the RPAIN-hRPA complex, and future work will focus on deciphering the differences in the mechanisms of action of RPAIN.

Methods

Reagents and buffers

Chemicals were purchased from Sigma-Millipore Inc., Research Products International Inc. and Gold Biotechnology Inc. Fluorescent and unlabeled oligonucleotides were synthesized by Integrated DNA Technologies. Enzymes for molecular biology were purchased from New England Biolabs. Resins for protein purification were sourced from GE-Cytiva Life Sciences Inc. Fluorophores for protein labeling were from Click Chemistry Tools Inc.

Plasmids for protein overproduction

An RSF-Duet1 plasmid coding for Rtt105 with a N-terminal 6x-polyhistidine tag was used. Mutations in *Rtt105* were introduced using site-directed mutagenesis. Plasmids for RPA and 4-azidophenylalanine (4AZP) incorporation were as described^{17,22,23}. A codon-optimized open reading frame for human RPAIN was synthesized (Genescript Inc.) and carries a SUMO protease cleavable N-terminal Strep-6xHis-SUMO tag.

Purification of RPA, Rtt105, and RPAIN

Rtt105 was overproduced in Rosetta-2 PlysS *E. coli* cells and purified as described⁴¹. *Saccharomyces cerevisiae* and human RPA were purified as described^{22,40}. Non-canonical amino acid (4AZP) incorporation based fluorescently-labeled RPA were generated as described^{17,22,23}. Concentration of Rtt105 was determined spectroscopically using $\epsilon_{280} = 18,450 \text{ M}^{-1}\text{cm}^{-1}$. Rtt105 was flash frozen and stored at -70°C . Concentration of unlabeled and labeled RPA was measured spectroscopically using $\epsilon_{280} = 98,500 \text{ M}^{-1}\text{cm}^{-1}$. Labeling efficiency for fluorescent RPA was calculated as described using absorption values measured at 280 nm and $\epsilon_{280} = 98500 \text{ M}^{-1}\text{cm}^{-1}$ for RPA, at 550 nm with $\epsilon_{550} = 105,000 \text{ M}^{-1}\text{cm}^{-1}$ for RPA-MB543, at 555 nm with $\epsilon_{555} = 150,000 \text{ M}^{-1}\text{cm}^{-1}$ for RPA-Cy3, and at 650 nm with $\epsilon_{650} = 250,000 \text{ M}^{-1}\text{cm}^{-1}$ for RPA-Cy5 versions^{17,23}. Human RPAIN was overproduced in Rosetta-2 PlysS *E. coli* cells by inducing the cells at OD₆₀₀ = 0.6 with 1 mM IPTG and growing them overnight at 18 °C. The cell pellets were lysed in buffer containing 30 mM HEPES pH 7.8, 300 mM KCl, 1 mM TCEP-HCl, 10 % v/v glycerol, 0.04 mg/ml of lysozyme and protease inhibitor cocktail and then sonicated on ice for a total of 2 min. After centrifugation at 4 °C for 60 mins at 41,107 g, the clarified lysate was batch bound to Ni²⁺-NTA beads for 18 h at 4 °C. Beads were washed sequentially with wash buffer (30 mM HEPES pH 7.8, 0.02% Tween-20, 10 mM imidazole, 1 mM TCEP-HCl, 10% v/v glycerol, and protease inhibitor cocktail) containing varying amounts of KCl (0.3 M, 2 M and 0.05 M). RPAIN was eluted with elution buffer (30 mM HEPES pH 7.8, 50 mM KCl, 0.02% Tween-20, 400 mM imidazole, 1 mM TCEP-HCl, 10% v/v glycerol, and protease inhibitor cocktail). RPAIN containing fractions were pooled and further fractionated using a 10 mL Q-sepharose Fast Flow column (Cytiva Life Sciences) equilibrated in buffer (30 mM HEPES pH 7.8, 0.02% Tween-20, 0.1 M KCl, 0.25 mM EDTA pH 8.0, 1 mM TCEP-HCl, 10% glycerol, and protease inhibitor cocktail). After loading and washing, RPAIN was eluted in the same buffer with a 0.1–1.5 M KCl gradient. RPAIN containing fractions were pooled and digested for 18 h at 4 °C with SUMO protease (1:10 ratio) to remove the N-terminal tag. Cleaved RPAIN was separated using a Hi-Load 16/600 Superdex 200 pg size exclusion column (Cytiva Life Sciences) using buffer (30 mM HEPES pH 7.8, 300 mM KCl, 0.02 % Tween-20, 0.25 mM EDTA pH 8.0, 1 mM TCEP-HCl, and 10 % v/v glycerol). RPAIN containing fractions were pooled, concentrated using a spin-concentrator, flash frozen, and stored at -80°C . RPAIN concentration was determined spectroscopically using extinction coefficient $\epsilon_{280} = 28,585 \text{ M}^{-1}\text{cm}^{-1}$.

Generation of fluorescently-labeled Rtt105

Rtt105 has two consecutive Cys residues at positions 12 and 13, respectively. Either or both Cys residues can be substituted with Ser without loss of binding to RPA (Supplementary Fig. 2). We used Cys-

Table 1 | *S. cerevisiae* strains used in this study

Strain	Genotype	Source
X8584-2A	<i>Rtt105-TAP::TRP</i>	This study
X8049-7B	<i>MATa rfa1-zm1</i>	Ref. 24
X8047-1B	<i>MATa rfa1-zm2</i>	Ref. 24
X8584-2C	<i>Rtt105-TAP::TRP rfa1-zm1</i>	This study
X8585-11C	<i>Rtt105-TAP::TRP rfa1-zm2</i>	This study
G996	<i>MATa rfa1-t33</i>	L. Symington
X8587-1A	<i>Rtt105-TAP::TRP rfa1-t33</i>	This study
T2159-16	<i>MATa rtt105Δ::hphMX6</i>	This study

12 for attachment of fluorophores using maleimide chemistry and converted Cys-13 to Ser. This version of Rtt105 (Rtt105^{CS}) was purified similar to the wild-type protein and ~5 ml of 100 μM Rtt105^{CS} was dialyzed extensively in labeling buffer (30 mM HEPES, pH 7.8, 200 mM KCl, 0.25 mM EDTA, pH 8.0, 0.01 % Tween-20, and 10 % v/v glycerol) to remove from the storage buffer. After 4 buffer exchanges, a 1.5-fold molar excess of Cy5-maleimide dye was added to the dialyzed protein and incubated at 4 °C for 3 h. The reaction was then quenched with 0.5 % β-mercaptoethanol (βME). Excess dye was separated from the protein using a Biogel-P4 column (Bio-rad laboratories) and resolved with labeling buffer with 2 mM TCEP. The concentration of Rtt105 calculated spectroscopically using extinction coefficient $\epsilon_{280\text{nm}}$ 18,450 M⁻¹cm⁻¹ and labeling efficiency was calculated using the Cy5 absorbance signal and ϵ_{650} = 250,000 M⁻¹cm⁻¹. Rtt105 absorbance values at 280 nm were also corrected for minor signal interference from Cy5 by measuring the percent contribution of free Cy5.

Yeast strains and genetic techniques

Standard procedures were used for cell growth and media preparation. Strains used are provided in Table 1 and are isogenic to W1588-4C, a *RADS* derivative of W303 (*MATa ade2-1 can1-100 ura3-1 his3-11,15, leu2-3, 112 trp1-1 rad5-S35*⁴¹). Gene deletion to generate *rtt105Δ* strain was performed following standard PCR based method. Standard yeast genetic procedures were used for tetrad analyses and at least two biological duplicates were used for each genotype.

Co-Immunoprecipitation

Yeast cells from log phase cultures growing in YPD were harvested and lysed by bead beating in TMG-140 buffer (10 mM Tris-HCl, pH 8.0, 4 mM MgCl₂, 10 % v/v glycerol, 140 mM NaCl, 0.1 mM EDTA, 0.5 % Tween20, 1 mM DTT and Roche cOmplete-Ultra EDTA free protease inhibitor). DNA was digested by incubation with benzonase for 30 min at 4 °C. Lysates were cleared by centrifugation and incubated in TMG-140 buffer with IgG sepharose beads for 2 h at 4 °C. After incubation, beads were washed with TMG-140 and proteins were eluted with Laemmli buffer. Proteins were separated on gradient gels followed by western blotting with antibodies against Rfa1 (a kind gift from Dr. Steven Brill at Rutgers University). Rfa1 primary antibody was used at 1:6000 dilution. Anti-rabbit HRP secondary body was used at 1:8000 dilution (VWR Scientific).

Secondary structure analysis using circular dichroism (CD)

CD measurements were used to compare the secondary structures of Rtt105 and the Cys variants of Rtt105. A nitrogen-fused observation chamber with a cell pathlength of 10 mm was used. All CD traces were obtained between 200–260 nm at 20 °C on a Chirascan CD spectrometer (Applied Photophysics Inc. using Pro-data Chirascan software). 600 nM of Rtt105^{WT}, Rtt105^{CS}, Rtt105^{CS}, Rtt105^{CS} in CD reaction buffer (5 mM Tris-Cl pH 7.8, 100 mM KCl, 5 mM MgCl₂, and 6 % v/v

glycerol) was used to obtain CD spectra. The results were collected using 1 nm step size, 1 nm bandwidth and 5 traces were averaged.

Analysis of complex formation using size exclusion chromatography (SEC)

600 μl of the noted concentrations of RPA, Rtt105, or the complex in the absence or presence of equimolar amounts of DNA were resolved on a 10/300 Superose 6 Increase column using an AKTA-pure FPLC system. Protein and protein-DNA complexes were incubated at 4 °C for 10 min before analysis. Resolution was performed using Rtt105-SEC buffer (30 mM HEPES, pH 7.8, 100 mM KCl, 1 mM TCEP-HCl, and 10 % v/v glycerol). A total of 30 ml elution volume was collected as 0.5 ml fractions and further analyzed on 10 % SDS-PAGE.

Measurement of RPA-ssDNA interactions using electrophoretic mobility band shift analysis (EMSA)

10 nM 5'-Cy5-(dT)₃₀ ssDNA was incubated with indicated amounts of RPA at 25 °C in EMSA buffer (25 mM Tris-Cl, pH 7.5, 200 mM NaCl, 5 mM MgCl₂, 5 % v/v glycerol, and 0.05 % Tween-20) for 10 min. For experiments performed in the presence of Rtt105, RPA and Rtt105 were premixed in equimolar ratios at 25 °C for 5 min before DNA was introduced. The reaction mixture (20 μl) was mixed with 10 μl of 70 % v/v glycerol, mixed, and loaded onto a 6–15 % bis-acrylamide gradient gel and resolved using 1x TBE buffer. The gels were scanned using an iBright 1500 imager (Thermo Fischer Inc.) and the Cy5 fluorescence associated with ssDNA (unbound fraction) or ssDNA+protein (bound fraction) bands was background subtracted and quantified with the associated iBright software. The mean values and standard deviation from three independent experiments were plotted for analysis. K_D was estimated by nonlinear least squares fitting to Eq. (1):

$$F_b = \frac{F_{\max} [RPA]^h}{K_D^h + [RPA]^h} \quad (1)$$

F_b is the fraction of bound ssDNA determined from fluorescence intensities of two bands, namely:

$$F_b = \frac{\text{bound}}{\text{bound} + \text{unbound}}$$

$F_{\max}$ is the fraction of bound ssDNA at saturating protein concentration, K_D is the apparent dissociation constant, $[RPA]$ is concentration of RPA or RPA+Rtt105 (1:1) in each well, and h is the Hill coefficient.

Measurement of RPA and ssDNA binding in presence or absence of Rtt105 using fluorescence anisotropy

5'-FAM-(dT)₃₅ ssDNA was diluted to 10 nM in 1x RPA reaction buffer (30 mM HEPES pH 7.8, 100 mM KCl, 6 % v/v glycerol, 5 mM MgCl₂, and 1 mM βME). 180 μl of this working stock was added to a 3 mm path-length quartz cuvette (Starna Cells Inc.) and the temperature was maintained at 23 °C. Fluorescence anisotropy of the FAM-labeled ssDNA was measured using PC1 spectrofluorometer (ISS Inc.) and data collected using the associated Vinci 3 software. Samples were excited at 488 nm and the resulting emission was collected using a 520 nm band pass emission filter. Five consecutive anisotropy readings were acquired from ssDNA alone or after stepwise addition of RPA alone, or 1:1 stock of RPA and Rtt105 (each protein at 5 μM). The concentrations of ssDNA, the added protein, and the reduction in intensity were corrected for effects due to dilution alone. Measured anisotropy values were corrected for the G-factor, and any changes in fluorescence intensity using Eq. (2). Finally mean ± SEM were estimated from four experiments and plotted.

Binding of proteins in the vicinity of the fluorescein moiety often leads to quenching because the fluorescence quantum yield of the

bound species is typically lower than that of free ssDNA. Unless corrected for, this artifact can result in significant errors in the estimation of binding affinity⁴².

$$\frac{F_b}{F_f} = \left[\frac{A - A_f}{A_b - A} \right] \times \frac{Q_f}{Q_b} = \frac{A_c - A_f}{A_b - A_c}$$

Rearranging, we get

$$A_c = \frac{\left[\frac{(A - A_f)}{(A_b - A)} \cdot \left(\frac{Q_f}{Q_b} \right) \cdot (A_b) \right] + A_f}{1 + \left[\frac{(A - A_f)}{(A_b - A)} \cdot \left(\frac{Q_f}{Q_b} \right) \right]} \quad (2)$$

Where, (1) F_b and F_f are the bound, and free concentrations of the FAM-labeled fluorescent ssDNA in μM , (2) Q_b and Q_f are the fluorescence quantum yields of the bound and free form of the FAM-labeled fluorescent ssDNA (arbitrary units), (3) A_b and A_f are the anisotropy values of the bound, and free forms of the FAM-labeled fluorescent ssDNA, (4) A is the measured anisotropy, and (5) A_c is the corrected anisotropy value

K_D was determined after fitting corrected anisotropy values (A_c using Eq. (2)) with a model for one site specific binding with Hill slope as defined by Eq. (3) using GraphPad Prism 9.

$$A_c = \frac{A_{c,\text{max}}[RPA]^h}{K_D^h + [RPA]^h} \quad (3)$$

Where, A_c is the corrected anisotropy value from Eq. (2), $A_{c,\text{max}}$ is the maximum anisotropy when 100% of ssDNA is complexed with RPA, K_D is the apparent dissociation constant, $[RPA]$ is the concentration RPA in the cuvette after each successive addition, and h is the Hill coefficient.

For the anisotropy experiments where the binding density of RPA was measured, 30 nM 5'-FAM-(dT)₃₅ or 5'-FAM-(dT)₇₀ were taken in 1x RPA reaction buffer in a 10 mm pathlength quartz cuvette (Firefly Sci) with stirring. RPA alone or RPA + Rtt105 (1:1) were titrated and after an incubation period of 3 min fluorescence anisotropy was measured and plotted against the concentration of proteins. RPA and Rtt105 were mixed at 1:1 molar ratio (1.1 μM each) in 1x RPA reaction buffer and incubated on ice for 30 min prior to each experiment. Measured anisotropy values were corrected for the G-factor and any changes in fluorescence intensity using Eq. (2); mean $\pm$ sem were estimated from three experiments and plotted.

The saturation points were taken as the intersection of biphasic or triphasic curves from the linear fits of the initial data points reflecting the change in anisotropy upon binding of sub-saturating amounts of proteins. For (dT)₃₅, the first dotted line represents a stoichiometry of 1:1, and the other represents 2.7:1. For (dT)₇₀, first dotted line represents a stoichiometry of 2:1, and the other represents 4.7:1. However, note that stoichiometry estimates with (dT)₇₀ is probably an underestimate because an anisotropy value of 0.19 is close to the limiting value of anisotropy measurements for fluorescein.

Crosslinking mass spectrometry (XL-MS) analysis

Stock solutions of Rtt105 (13.4 mg/mL) and RPA (1.77 mg/mL) were diluted to 0.89 mg/mL and 0.3752 mg/mL, respectively in buffer (30 mM HEPES, 200 mM KCl, pH 7.8) and incubated together for 30 min. The diluted proteins were reacted with 5 mM bis(sulphosuccinimidyl)suberate (BS3) and 20 μL of the sample was taken at various time points (0, 15 and 30 min) and immediately quenched with 2 μL of 1 M ammonium acetate. Quenched samples were diluted with 1.5X Laemmli gel loading buffer to a final volume of 40 μL , vortexed, and heated to 100 $^{\circ}\text{C}$ for 5 min and resolved on 4–20% (w/v) gradient SDS-PAGE gels (Bio-Rad) with Tris-glycine buffer. Gels were stained with Gelcode blue safe protein stain (Thermo Scientific). Gel bands were excised for protein identification and analysis.

Excised bands were destained with a 50 mM ammonium bicarbonate and 50% acetonitrile mixture and reduced with a mixture of 100 mM DTT and 25 mM ammonium bicarbonate for 30 min at 56 $^{\circ}\text{C}$. The reaction was subsequently exchanged for the alkylation step with 55 mM iodoacetamide and 25 mM ammonium bicarbonate and incubated in the dark at room temperature for 25 min. The solution was then washed with the 50 mM ammonium bicarbonate and 50% acetonitrile mixture. The gel pieces were then first dehydrated with 100% acetonitrile and then rehydrated with sequence grade trypsin solution (0.6 μg , Promega) and incubated overnight at 37 $^{\circ}\text{C}$. The reaction was quenched with 10 μL of 50% acetonitrile and 0.1% formic acid (FA, Sigma) and transferred to new microfuge tubes, vortexed for 5 min, and centrifuged at 16,000 g for 30 min. Samples were transferred to mass spectrometry vials and quantitated by LC-MS as described for peptide identification^{43,44}. Peptides were identified as previously described⁴⁵ using MassHunter Qualitative Analysis, version 6.0 (Agilent Technologies), Peptide Analysis Worksheet (ProteoMetrics LLC), and PeptideShaker, version 1.16.42, paired with SearchGUI, version 3.3.16 (CompOmics). Crosslinks were then determined using Spectrum Identification Machine (SIMXL1.5.5.2).

Hydrogen-deuterium exchange mass spectrometry (HDX-MS) analysis

Stock solutions of RPA (13.4 mg/mL) and Rtt105 (1.77 mg/mL) were mixed in the presence or absence of (dT)₃₅ ssDNA in a 1:1.2 ratio. Reactions were diluted 1:10 into deuterated reaction buffer (30 mM HEPES, 200 mM KCl, pH 7.8). Control samples were diluted into a non-deuterated reaction buffer. At each time point (0, 0.008, 0.05, 0.5, 3, 30 h), 10 μL of the reaction was removed and quenched by adding 60 μL of 0.75% formic acid (FA, Sigma) and 0.25 mg/mL porcine pepsin (Sigma) at pH 2.5 on ice. Each sample was digested for 2 min with vortexing every 30 s and flash-frozen in liquid nitrogen. Samples were stored in liquid nitrogen until the LC-MS analysis. LC-MS analysis of RPA was completed as described⁴⁶. Briefly, the LC-MS analysis of RPA was completed on a 1290 UPLC series chromatography stack (Agilent Technologies) coupled with a 6538 UHD Accurate-Mass QTOF LC/MS mass spectrometer (Agilent Technologies). Peptides were separated on a reverse phase column (Phenomenex Onyx Monolithic C18 column, 100 $\times$ 2 mm) at 1 $^{\circ}\text{C}$ using a flow rate of 500 $\mu\text{L}/\text{min}$ under the following conditions: 1.0 min, 5% B; 1.0 to 9.0 min, 5 to 45% B; 9.0 to 11.8 min, 45 to 95% B; 11.8 to 12.0 min, 5% B; solvent A = 0.1% FA (Sigma) in water (Thermo Fisher) and solvent B = 0.1% FA in acetonitrile (Thermo Fisher). Data were acquired at 2 Hz s^{-1} over the scan range 50 to 1700 m/z in the positive mode. Electrospray settings were as follows: the nebulizer set to 3.7 bar, drying gas at 8.0 L/min, drying temperature at 350 $^{\circ}\text{C}$, and capillary voltage at 3.5 kV. Peptides were identified as previously described⁴⁵ using MassHunter Qualitative Analysis, version 6.0 (Agilent Technologies), Peptide Analysis Worksheet (ProteoMetrics LLC), and PeptideShaker, version 1.16.42, paired with SearchGUI, version 3.3.16 (CompOmics). Deuterium uptake was determined and manually confirmed using HDExaminer, version 2.5.1 (Sierra Analytics). Heat maps were created using MStools⁴⁷.

MB543 fluorescence quenching assay to estimate binding affinity between RPA and Rtt105

RPA-DBD-A^{MB543}, RPA-DBD-D^{MB543}, or F-A-B-DBD-A^{MB543} were diluted to 200 nM in 1x RPA reaction buffer. 180 μL of either fluorescent protein was added to a 3 mm path length quartz cuvette (Starna Cells Inc.) and maintained at 23 $^{\circ}\text{C}$ in a PC1 spectrofluorometer (ISS Inc.). The MB543 dye, conjugated to either DBD-A or DBD-D domain of RPA, was excited at 535 nm and the resulting fluorescence emission spectra were collected between 558 nm to 578 nm with the λ_{max} located at 568 nm. Unlabeled Rtt105, diluted to a 4 μM stock in RPA

reaction buffer, was added to the cuvette in a stepwise manner, mixed, and incubated for 3 min to achieve equilibrium before emission spectra were measured. Emission scans were recorded twice from at least two cuvettes, and the experiment was repeated three to four times. The fluorescence at $\lambda_{\max}$ from 3–4 trials was corrected for stepwise dilution of sample (<8%), normalized to the initial fluorescence (fluorescence intensity in the absence of Rtt105), and plotted as mean and SEM. The fluorescence intensity values were transformed to fraction quenched versus Rtt105 concentration and fitted to a quadratic Eq. (4) using non-linear least squares regression, accounting for ligand depletion, to yield an apparent equilibrium constant (K_D). The dilution factor corrected RPA-DBD- D^{MBS43} fluorescence remained nearly constant (within error) and served as a control for change in fluorescence due to photobleaching alone.

$$i = i_{\min} + \frac{(i_{\max} - i_{\min}) \times \left(\left(K_D + [RPA] + [Rtt105] \right) - \sqrt{\left(K_D + [RPA] + [Rtt105] \right)^2 - 4[RPA][Rtt105]} \right)}{2[RPA]} \quad (4)$$

Where, i , is the measured fluorescence intensity, $i_{\min}$, & $i_{\max}$ are minimum and maximum values of the fluorescence intensity of 100% free, and 100% bound RPA determined from the fit, respectively. $[RPA]$, is the concentration of fluorescent RPA taken in the cuvette and the value is constrained at 200 nM. However, note that due to the stepwise addition of Rtt105 there is a 5% dilution by the end of the measurement. $[Rtt105]$ is the dilution factor corrected concentration of Rtt105 in the cuvette in nM, and K_D is the dissociation constant determined from the fit.

Measurement of DNA binding kinetics using stopped flow fluorescence

All stopped flow experiments were performed on a SX20 instrument (Applied Photophysics Inc.) at 25 °C in 1x RPA reaction buffer. Protein and/or DNA reactions from individual syringes were rapidly mixed and fluorescence data were collected. The respective mixing schemes are denoted by cartoon schematics within the figure panels. Seven to eight individual shots were averaged for each experiment. All experiments were repeated a minimum of 3 times and SEM from the individual fits are noted in the figure legends. For the FRET experiments, samples were excited at 535 nm (Cy3 wavelength) and Cy5 emission was captured using a 645 nm long-pass filter. For the RPA-Rtt105 interactions, RPA-DNA and Rtt105-RPA-DNA interactions, experiments were performed with 100 nM each of RPA, Rtt105, and (dT)₃₅ or (dT)₄₀ ssDNA substrates (1:1:1 ratio). For facilitated exchange stopped flow experiments, 200 nM RPA-DBD- A^{Cy5} and 200 nM RPA-DBD- D^{Cy5} were premixed with 120 nM (dT)₉₇ and shot against unlabeled RPA (500 nM) or the RPA-Rtt105 complex (500 nM each).

Steady State Förster resonance energy transfer (FRET) measurement of protein-DNA and protein-protein complexes

RPA-DBD- D^{Cy5} and Rtt105- C^{Cy5} were mixed in 1x RPA reaction buffer at 1:1 ratio such that the final concentration of each protein was 200 nM. The complex was incubated on ice for 30 min and then transferred to a 3 mm pathlength cuvette maintained at 23 °C in the PCI spectrofluorometer. The sample was excited at 535 nm and the resulting fluorescence between 550 nm to 700 nm was collected as a FRET spectrum. Next, ssDNA of different lengths (dT)_x; where $x = (\text{dT})_8, (\text{dT})_{15}, (\text{dT})_{25}, (\text{dT})_{35}, (\text{dT})_{45}, (\text{dT})_{54}, (\text{dT})_{64}, (\text{dT})_{70},$ or $(\text{dT})_{84}$ were added in a stepwise manner to the cuvette, mixed, and incubated for 3 min at 23 °C before FRET spectra was again recorded after each addition. Raw spectra were corrected by incorporating the estimated dilution factor and then area normalized to account

for any fluctuations in lamp intensity. Finally, ratiometric FRET was calculated as defined by Eq. (5):

$$FRET = \frac{I_A}{I_A + I_D} \quad (5)$$

Where, I_A and I_D are acceptor and donor fluorescence emission intensities at respective $\lambda_{\max} = 673$, and 573 nm, respectively.

Quantitative FRET between RPA-DBD- D^{Cy5} or - A^{Cy5} and 3' or 5' labelled Cy3-(dT)₄₀, respectively was estimated by correcting for the increase in acceptor fluorescence during titration. This monotonic increase was estimated and subtracted from FRET spectra. In addition, protein induced fluorescence enhancement or PIFE, was estimated by titrating in - D^{Cy5} or - A^{Cy5} with 3' or 5' labelled Cy3-(dT)₄₀. This was also subtracted to estimate true energy transfer values.

C-Trap optical tweezer analysis of RPA-Rtt105 interactions

48.5 kbp Lambda DNA construct was prepared with three biotins on either end of DNA. Lambda DNA was purchased from Roche Inc. Short oligos to anneal to the sticky ends were purchased from Integrated DNA Technologies Inc. DNA were stored in TE buffer (10 mM Tris-HCl, pH 8.0 and 0.1 mM EDTA). Optical trap experiments were performed using a commercial dual optical trap combined with confocal microscopy and microfluidics [C-trap] from Lumicks BV Inc. Streptavidin coated polystyrene particle beads of average size 4.8 μm [0.5% w/v] (Spherotech Inc.) were diluted 1:250 in 1x PBS and 1–2 nM of DNA were made in 1x PBS. DNA was captured between two streptavidin beads and mechanically denatured by moving one bead to create ssDNA. ssDNA was confirmed by fitting force-distance (FD) curve to Freely Jointed Chain model [FJC] (contour length 48.5 kbp/16.49 μm ; persistence length 46 nm; stretch modulus 1000 pN) in real time. DNA was held for 5 s in the fully ssDNA state and then returned to 5 pN tension for the fluorescence experiments. RPA-DBD- D^{MBS43} in storage buffer (30 mM HEPES pH 7.8, 200 mM KCl, 0.02% Tween-20, 10% glycerol, and 0.2 mM EDTA) was added to the DNA in the absence or presence of Rtt105- C^{Cy5} . Rtt105- C^{Cy5} was kept in storage buffer with 1 mM TCEP-HCl. Both proteins were diluted to 1 nM with experimental buffer (30 mM HEPES pH 7.8, 100 mM KCl, 6% Glycerol, 5 mM MgCl₂ and incubated together (for the experiments where Rtt105-RPA complexes were tested) at 1:1 molar ratio (10 pM final concentration each). Imaging buffer 0.8% (w/v) dextrose, 165 U/mL glucose oxidase, 2170 U/mL catalase, and 2–3 mM Trolox was used to increase the fluorescence lifetime of the fluorophores. Imaging settings were 2–3 ms exposure time (per pixel), red excitation 638 nm, and green excitation 561 nm. Data was analyzed using custom python script [Pylake API from Lumicks].

Analytic ultracentrifugation (AUC) analysis

AUC sedimentation velocity experiments were performed on an Optima analytical ultracentrifuge (Beckman-Coulter Inc.) using an An-50Ti rotor at 40,000 rpm at 20 °C. Proteins and DNA either alone or in complex were dialyzed against 30 mM HEPES, pH 7.8, 100 mM KCl, 10% glycerol, and 1 mM TCEP-HCl before each experiment. Concentrations used for the experiments are mentioned in the appropriate figures. Sample (380 μL) and buffer (400 μL) were filled in each chamber of a 2-sector charcoal quartz cell. Absorbance was monitored at 280 nm and/or 650 nm. Since the absorbance signal from Rtt105 was low, Rtt105- C^{Cy5} was used and tracked at 650 nm. Scans were recorded at 3 min intervals. The density and viscosity of the buffer at 20 °C were calculated using SEDNTERP. Continuous distribution (c(s)) model was used to fit the data in SEDFIT⁴⁸.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data are available within this manuscript and its supplementary information files. Plasmids used for protein overexpression and the data supporting the findings of this study are available from the corresponding author upon request. Source data are provided with this paper.

Code availability

Code for C-trap data analysis is available at <https://github.com/spangeni/Rtt105-Analysis-Pipeline/tree/main>.

References

- Wold, M. S. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev. Biochem.* **66**, 61–92 (1997).
- Fanning, E., Klimovich, V. & Nager, A. R. A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res.* **34**, 4126–4137 (2006).
- Marechal, A. & Zou, L. RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. *Cell Res.* **25**, 9–23 (2015).
- Deshpande, I. et al. Structural Basis of Mec1-Ddc2-RPA Assembly and Activation on Single-Stranded DNA at Sites of Damage. *Mol. Cell* **68**, 431–445 e435 (2017).
- Bochkareva, E., Korolev, S., Lees-Miller, S. P. & Bochkarev, A. Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J.* **21**, 1855–1863 (2002).
- Fan, J. & Pavletich, N. P. Structure and conformational change of a replication protein A heterotrimer bound to ssDNA. *Genes Dev.* **26**, 2337–2347 (2012).
- Bochkarev, A., Pfuetzner, R. A., Edwards, A. M. & Frappier, L. Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385**, 176–181 (1997).
- Feldkamp, M. D., Mason, A. C., Eichman, B. F. & Chazin, W. J. Structural analysis of replication protein A recruitment of the DNA damage response protein SMARCA1. *Biochemistry* **53**, 3052–3061 (2014).
- Guilliam, T. A. et al. Molecular basis for PrimPol recruitment to replication forks by RPA. *Nat. Commun.* **8**, 15222 (2017).
- Brosey, C. A. et al. A new structural framework for integrating replication protein A into DNA processing machinery. *Nucleic Acids Res.* **41**, 2313–2327 (2013).
- Li, S. et al. Rtt105 functions as a chaperone for replication protein A to preserve genome stability. *EMBO J.* **37**, <https://doi.org/10.15252/embj.201899154> (2018).
- Collins, S. R. et al. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* **446**, 806–810 (2007).
- Corda, Y., Maestroni, L., Luciano, P., Najem, M. Y. & Geli, V. Genome stability is guarded by yeast Rtt105 through multiple mechanisms. *Genetics* **217**, <https://doi.org/10.1093/genetics/yaa035> (2021).
- Wang, X. et al. Rtt105 promotes high-fidelity DNA replication and repair by regulating the single-stranded DNA-binding factor RPA. *Proc. Natl Acad. Sci. USA* **118**, <https://doi.org/10.1073/pnas.2106393118> (2021).
- Li, S., Dong, Z., Yang, S., Feng, J. & Li, Q. Chaperoning RPA during DNA metabolism. *Curr. Genet.* **65**, 857–864 (2019).
- Kumaran, S., Kozlov, A. G. & Lohman, T. M. Saccharomyces cerevisiae replication protein A binds to single-stranded DNA in multiple salt-dependent modes. *Biochemistry* **45**, 11958–11973 (2006).
- Pokhrel, N. et al. Dynamics and selective remodeling of the DNA-binding domains of RPA. *Nat. Struct. Mol. Biol.* **26**, 129–136 (2019).
- Caldwell, C. C. & Spies, M. Dynamic elements of replication protein A at the crossroads of DNA replication, recombination, and repair. *Crit. Rev. Biochem. Mol. Biol.* **55**, 482–507 (2020).
- Chen, R., Subramanyam, S., Elcock, A. H., Spies, M. & Wold, M. S. Dynamic binding of replication protein A is required for DNA repair. *Nucleic Acids Res.* **44**, 5758–5772 (2016).
- Ahmad, F. et al. Hydrogen-deuterium exchange reveals a dynamic DNA-binding map of replication protein A. *Nucleic Acids Res.* **49**, 1455–1469 (2021).
- Kim, C., Paulus, B. F. & Wold, M. S. Interactions of human replication protein A with oligonucleotides. *Biochemistry* **33**, 14197–14206 (1994).
- Pokhrel, N. et al. Monitoring Replication Protein A (RPA) dynamics in homologous recombination through site-specific incorporation of non-canonical amino acids. *Nucleic Acids Res.* **45**, 9413–9426 (2017).
- Kuppa, S., Pokhrel, N., Corless, E., Origanti, S. & Antony, E. Generation of Fluorescent Versions of Saccharomyces cerevisiae RPA to Study the Conformational Dynamics of Its ssDNA-Binding Domains. *Methods Mol. Biol.* **2281**, 151–168 (2021).
- Dhingra, N. et al. The Srs2 helicase dampens DNA damage checkpoint by recycling RPA from chromatin. *Proc. Natl Acad. Sci. USA* **118**, <https://doi.org/10.1073/pnas.2020185118> (2021).
- Chen, C., Umez, K. & Kolodner, R. D. Chromosomal rearrangements occur in S. cerevisiae rfa1 mutator mutants due to mutagenic lesions processed by double-strand-break repair. *Mol. Cell* **2**, 9–22 (1998).
- Deng, S. K., Gibb, B., de Almeida, M. J., Greene, E. C. & Symington, L. S. RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. *Nat. Struct. Mol. Biol.* **21**, 405–412 (2014).
- Kolpashchikov, D. M. et al. Polarity of human replication protein A binding to DNA. *Nucleic Acids Res.* **29**, 373–379 (2001).
- Ifode, C. & Borowiec, J. A. 5'→3' molecular polarity of human replication protein A (hRPA) binding to pseudo-origin DNA substrates. *Biochemistry* **39**, 11970–11981 (2000).
- Rappsilber, J. The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173**, 530–540 (2011).
- Yates, L. A. et al. A structural and dynamic model for the assembly of Replication Protein A on single-stranded DNA. *Nat. Commun.* **9**, 5447 (2018).
- Ma, C. J., Gibb, B., Kwon, Y., Sung, P. & Greene, E. C. Protein dynamics of human RPA and RAD51 on ssDNA during assembly and disassembly of the RAD51 filament. *Nucleic Acids Res.* **45**, 749–761 (2017).
- Park, M. S., Ludwig, D. L., Stigger, E. & Lee, S. H. Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *J. Biol. Chem.* **271**, 18996–19000 (1996).
- Plate, I. et al. Interaction with RPA is necessary for Rad52 repair center formation and for its mediator activity. *J. Biol. Chem.* **283**, 29077–29085 (2008).
- Shinohara, A., Shinohara, M., Ohta, T., Matsuda, S. & Ogawa, T. Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing. *Genes Cells* **3**, 145–156 (1998).
- Maestroni, L. et al. RPA and Pif1 cooperate to remove G-rich structures at both leading and lagging strand. *Cell Stress* **4**, 48–63 (2020).
- Hays, S. L., Firnenich, A. A., Massey, P., Banerjee, R. & Berg, P. Studies of the interaction between Rad52 protein and the yeast single-stranded DNA binding protein RPA. *Mol. Cell Biol.* **18**, 4400–4406 (1998).
- Deng, X. et al. Human replication protein A-Rad52-single-stranded DNA complex: stoichiometry and evidence for strand transfer regulation by phosphorylation. *Biochemistry* **48**, 6633–6643 (2009).

38. Mer, G. et al. Structural basis for the recognition of DNA repair proteins UNG2, XPA, and RAD52 by replication factor RPA. *Cell* **103**, 449–456 (2000).
39. New, J. H., Sugiyama, T., Zaitseva, E. & Kowalczykowski, S. C. Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature* **391**, 407–410 (1998).
40. Hormeno, S. et al. Human HELB is a processive motor protein that catalyzes RPA clearance from single-stranded DNA. *Proc. Natl Acad. Sci. USA* **119**, e2112376119 (2022).
41. Zhao, X. & Blobel, G. A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization. *Proc. Natl Acad. Sci. USA* **102**, 4777–4782 (2005).
42. Walter, B. D., Mao-Lin, H. & Jacques, L. B. 1 Equilibrium and kinetic inhibition assays based upon fluorescence polarization. *Methods Enzymol.* **74**, 3–28 (1981).
43. Burns, E. E., Keith, B. K., Refai, M. Y., Bothner, B. & Dyer, W. E. Proteomic and biochemical assays of glutathione-related proteins in susceptible and multiple herbicide resistant *Avena fatua* L. *Pestic. Biochem. Physiol.* **140**, 69–78 (2017).
44. Kim, C. et al. Core Protein-Directed Antivirals and Importin beta Can Synergistically Disrupt Hepatitis B Virus Capsids. *J. Virol.* **96**, e0139521 (2022).
45. Berry, L., Patterson, A., Pence, N., Peters, J. W. & Bothner, B. Hydrogen Deuterium Exchange Mass Spectrometry of Oxygen Sensitive Proteins. *Bio. Protoc.* **8**, <https://doi.org/10.21769/BioProtoc.2769> (2018).
46. Patterson, A., Zhao, Z., Waymire, E., Zlotnick, A. & Bothner, B. Dynamics of Hepatitis B Virus Capsid Protein Dimer Regulate Assembly through an Allosteric Network. *ACS Chem. Biol.* **15**, 2273–2280 (2020).
47. Kavan, D. A. M. P. MStools—Web based application for visualization and presentation of HXMS data. *Int. J. Mass Spectrom.* **302**, 53–58 (2010).
48. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).

Acknowledgements

The authors thank members of our respective research laboratories for critical reading of the manuscript. This work was supported by grants from the National Institutes of Health (R01 GM130746 and R01 GM133967) to E.A., (R15 GM123443) to H.B., (R35 GM122569) to T.J. and (R01 GM131058 and R35 GM145260) to X.Z. Funding for Proteomics, Metabolomics and Mass Spectrometry Facility at MSU was made possible in part by the MJ Murdock Charitable Trust and NIGMS of the National Institutes of Health under Award Number P20 GM103474. The analytical ultracentrifuge experiments and C-trap experiments were supported by instrumentation grants from the Office of the Director, National Institutes of Health (S10 OD030343 to E.A. and S10 OD025221 to Johns Hopkins University, respectively). M.K.S. was supported by a

summer undergraduate research supplement from the National Institutes of Health (R01 GM130746-04S1 to E.A.).

Author contributions

S.K., J.D., R.C., N.P., V.K. and M.K.S. purified proteins, generated fluorescent protein, designed, and performed the steady-state and pre-steady state experiments. J.M., A.P., S.K., and B.B. performed and analyzed the HDX-MS and XLMS experiments. S.P. and T.H. performed and analyzed the C-trap experiments. N.D. and X.Z. performed and analyzed the in vivo experiments. S.S. and H.B. performed and analyzed single molecule FRET measurements of GQ-unwinding. E.A. directed the project, designed experiments, and composed the manuscript. All authors contributed to manuscript preparation.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-32860-6>.

Correspondence and requests for materials should be addressed to Edwin Antony.

Peer review information *Nature Communications* thanks Karl Duderstadt and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022

CHAPTER SIX

CONCLUDING REMARKS

Each of the proteins investigated in this thesis has evolved to take advantage of conformational change and allostery during catalysis. Though they have different folds, are involved in different pathways, and different biochemical processes, they all exhibit and need conformational change/allostery to function properly. Using these proteins as examples, I suggest that to fully explore the structure and function of proteins we must expand our scope to encompass the conformational ensemble of the protein of interest. An interesting quote that encompasses this work was posited by Hilser and coworkers:¹

Finally, taken as a whole, the results presented here have important evolutionary implications as well. If allosteric signaling is not only quantitatively, but qualitatively, dependent on the conformational ensemble of the protein, it must be the case that protein sequences code not just for the high-resolution structure of the protein, but also for its entire energy landscape. Because the sign of the signal (i.e., activation or repression) can be changed by modulating the stability of just one domain, even without a structural change, the notion that structure codes for function, or more precisely that function can be inferred from structure, must be critically revisited.

This statement is perhaps even more relevant today than when it was made nearly ten years ago. The functions of proteins are dependent on their structure which are encoded in their amino acid sequence. The amino acid sequence must also contain the “information” for the conformational ensemble of a protein in solution. We now have powerful computational and experimental tools for sampling significantly populated states of complex protein systems. However, even with these advances, current computational methods are still unable to chart the free-energy landscape of allosteric systems without prior knowledge of significantly populated states of the system.²⁻⁴ Even when such knowledge is available, computational procedures employ various levels of

approximations to predict the pathways used for conformational transitions between states.^{5,6} Experimental approaches to elucidate conformational ensembles can be hindered by the size and complexity of the system under study, making it challenging to evaluate the information they provide experimentally. Contained in this thesis are four large and complex systems for which we have successfully probed the conformational ensemble and allosteric network, leading to greater understanding of their catalysis.

Conformational change and allostery across systems

In the first half of this thesis, I described two microbial carboxylases. Carboxylases must capture carbon dioxide or its hydrated form bicarbonate, derivatize substrates for reaction, all while protecting reaction intermediates from unproductive side reactions.^{7,8} The two microbial carboxylases in this study, 2-KPCC and AC undergo conformational change induced by ligand binding to produce acetoacetate. My work also shows that both enzymes rely on residues in their active site to modulate stability and the conformational ensemble during catalysis. I showed that mutation of active site residues could alter the global conformational stability at different steps in their reaction. The studies presented in this thesis therefore expand the knowledge not only of carboxylases but allosteric enzymes.

In 2-KPCC, I was able to highlight the precise regions of conformational change, as well as probe the conformations that are accessible during catalysis. By poisoning the enzyme at different states in the catalytic cycle using a series of ligands, I was able to measure their mobility and analyze the conformational distribution using ion mobility mass spectrometry. In this way it was possible to compare conformational states and the overall ensemble at specific steps in the catalytic cycle. By using site-directed amino acid substitution, I was able to elucidate

that active site residues facilitate conformational changes critical to the overall reaction. From this study ligand binding in 2-KPCC was shown to alter the conformational ensemble by redistributing the energy landscape to favor one conformation which is essential for carboxylation (Figure 1A).

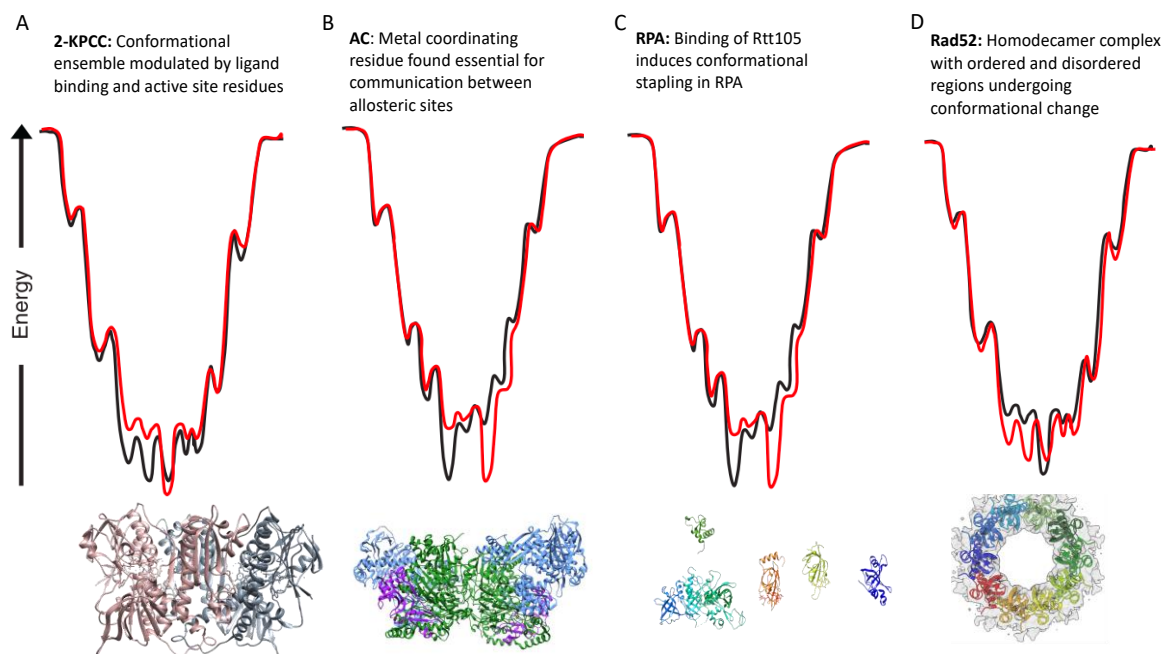


Figure 1. Summary of the findings contained in this thesis. Four protein complexes were studied and in all four conformational change, allostery, and conformational ensemble were found to be essential for function.

In AC, I was able to provide evidence for complex conformational change and propose an essential residue for allosteric communication. To achieve this HDX-MS was employed to compare hydrogen bonding networks of AC with and without ligands bound. Probing the hydrogen bonding network of AC resulted in the elucidating large conformational change in the alpha and beta subunits of the protein complex and the identification of a hydrophobic tunnel. The hydrophobic tunnel forms to transport reaction intermediates during catalysis and collapses

during product release. Further, by site-directed mutagenesis and HDX-MS the residues of allosteric communication necessary for the formation and collapse of the hydrophobic tunnel were illuminated. From this study ligand binding in AC was shown to redistribute the energy landscape to allow for a new conformation to be accessed during catalysis (Figure 1B).

In the second half of this thesis, I turned my attention to enzymes involved in DNA metabolism.⁹⁻¹¹ RPA is a heterotrimeric complex that is essential in DNA replication, repair, and recombination. RPA binds ssDNA and recruits over three dozen interacting proteins during DNA metabolism. This thesis contains investigations into two RPA binding partners Rtt105 and Rad52 and expand the knowledge not only of DNA metabolism but allosteric enzymes.

In the study elucidating the interactions between RPA and Rtt105 it was shown the binding of Rtt105 induces conformational change in RPA. RPA is a heterotrimer composed of RPA70, RPA32, and RPA14. The complex is held together by the trimerization core (Tri-C) which consists of three domains (C, D, and E domains). The other three domains (A, B, and F domains) in the protein complex are connected by flexible linker and extended away from the Tri-C. Through HDX-MS and crosslinking, interactions between RPA and Rtt105 were elucidated. The binding of Rtt105 to RPA induces a conformational change in RPA where the domains connected by flexible linker will directly interact with Rtt105 and configurationally staple those domains. The stapling of RPA leads to a more compact structure of the multidomain protein which is more stable and is protected from interacting with other binding partners in the absence of ssDNA. The binding of Rtt105 to RPA redistributes the energy landscape of RPA to stabilize a new compact conformation which is not accessible otherwise (Figure 1C).

In studying Rad52, it was shown that *S. cerevisiae* Rad52 functions as a homodecamer with each subunit possessing an ordered N-terminal and disordered C-terminal half. Binding of ssDNA induced large conformational changes in the N- and C- termini. An exciting finding was that Rad52 has an inherent asymmetry in the complex. This was demonstrated by investigating the bimodal behavior of a key C- terminal peptide in our HDX experiments. The asymmetry of the complex regulates ssDNA binding and interactions of multiple binding partners. Upon binding of ssDNA, the energy landscape of Rad52 redistributes and becomes more rugged with lower energy barriers between conformations, allowing for more conformations to be sampled (Figure 1D).

Final remarks on conformational ensembles and allostery

The concepts of allostery, conformational ensemble, and protein dynamics have evolved significantly since they were first proposed.^{12–15} It is now understood that many proteins are not rigid, static structures, but exist as conformational ensembles which undergo population shifts when exposed to external effectors. The re-distribution of conformational ensembles is an inherent physical attribute of all dynamic proteins.¹⁶ When this is applied to the concept of allostery, it can be understood that allostery is derived from a population shift in conformational ensemble. Thus, allosteric properties can vary widely in how they manifest in the protein. Some allosteric changes result in large conformational changes while others are more subtle and challenging to detect. Advancements in tools and techniques for detecting conformational change and measuring protein stability has led to the identification of more proteins and complexes that have allosteric properties. Allostery remains the coupling of two distal sites, however, how we investigate and think about allostery continues to expand. The classical models of the past have

given way to more diverse and robust models, which are themselves being challenged as more information becomes available on the properties of allostery and allosteric systems.

At this point, I think it is fair to ask the question; are all proteins under allosteric control? While this may not be true, I hypothesize that, except for fibrous proteins, some combination of allostery and dynamics have a functional role in many, if not most, proteins. Altogether, the research presented in this thesis provides measurable examples of allostery in the forms of substrate and binding partner induced re-distributions to the conformational ensemble. Though these studies employ allostery for function (catalysis or regulation) the manifestation of allostery is distinctive and dependent on the structure, conformational ensemble, and function of the protein. Studying large and complex protein systems such as those discussed here, is nearly always a challenge. By accepting the challenge, I have demonstrated that it is possible to elucidate functional elements of large protein and protein- nucleic acid complexes. This work also highlights how and where we can push the limits of current methodologies. While it is unlikely that we will ever have a definitive answer to the question – are all proteins allosterically regulated? – it is certain that we will break new ground and understand protein function on a deeper level.

References

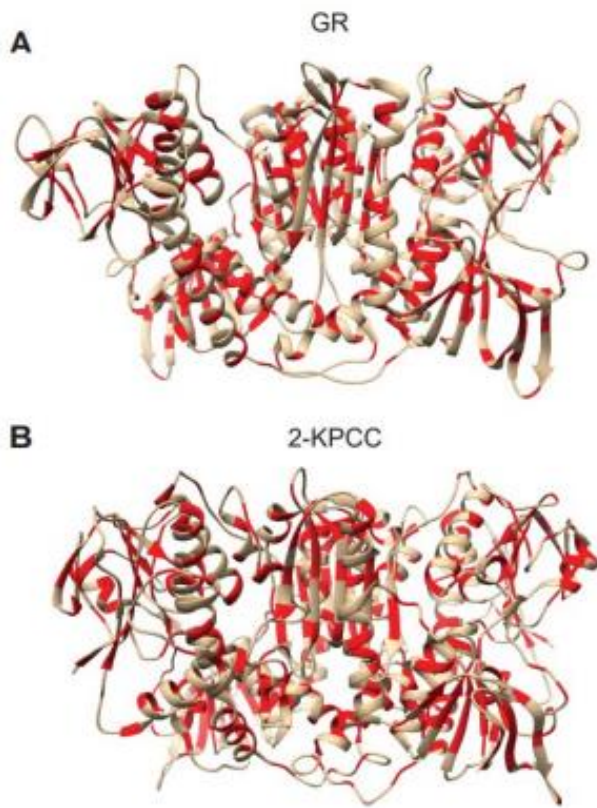
1. Hilser VJ, Wrabl JO, Motlagh HN. Structural and energetic basis of allostery. *Annu Rev Biophys.* 2012;41(1):585-609. doi:10.1146/annurev-biophys-050511-102319
2. Changeux JP. 50Th Anniversary of the Word “Allosteric.” *Protein Sci.* 2011;20(7):1119-1124. doi:10.1002/pro.658
3. Guarnera E, Tan ZW, Zheng Z, Berezovsky IN. AlloSigMA: Allosteric signaling and mutation analysis server. *Bioinformatics.* 2017;33(24):3996-3998. doi:10.1093/bioinformatics/btx430
4. Tan ZW, Guarnera E, Tee WV, Berezovsky IN. AlloSigMA 2: Paving the way to designing allosteric effectors and to exploring allosteric effects of mutations. *Nucleic Acids Res.* 2020;48(W1):W116-W124. doi:10.1093/NAR/GKAA338
5. Tee WV, Guarnera E, Berezovsky IN. Reversing allosteric communication: From detecting allosteric sites to inducing and tuning targeted allosteric response. *PLoS Comput Biol.* 2018;14(6):1-26. doi:10.1371/journal.pcbi.1006228
6. Greener JG, Sternberg MJ. Structure-based prediction of protein allostery. *Curr Opin Struct Biol.* 2018;50:1-8. doi:10.1016/j.sbi.2017.10.002
7. Erb TJ. Carboxylases in natural and synthetic microbial pathways. *Appl Environ Microbiol.* 2011. doi:10.1128/AEM.05702-11
8. Schada von Borzyskowski L, Rosenthal RG, Erb TJ. Evolutionary history and biotechnological future of carboxylases. *J Biotechnol.* 2003. doi:10.1016/j.jbiotec.2013.05.007
9. Ahmad F, Patterson A, Deveryshetty J, et al. Hydrogen-deuterium exchange reveals a dynamic DNA-binding map of replication protein A. *Nucleic Acids Res.* 2021;49(3):1455-1469. doi:10.1093/nar/gkaa1288
10. Li S, Xu Z, Xu J, et al. Rtt105 functions as a chaperone for replication protein A to preserve genome stability. *EMBO J.* 2018;37(17):1-16. doi:10.15252/embj.201899154
11. Kuppa S, Deveryshetty J, Chadda R, et al. Rtt105 regulates RPA function by configurationally stapling the flexible domains. *Nat Commun.* 2022;13(1):1-16. doi:10.1038/s41467-022-32860-6
12. Goodey NM, Benkovic SJ. Allosteric regulation and catalysis emerge via a common route. *Nat Chem Biol.* 2008;4(8):474-482. doi:10.1038/nchembio.98

13. Swain JF, Gierasch LM. The changing landscape of protein allostery. *Curr Opin Struct Biol.* 2006;16(1):102-108. doi:10.1016/j.sbi.2006.01.003
14. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? *Proteins Struct Funct Genet.* 2004;57(3):433-443. doi:10.1002/prot.20232
15. Wodak SJ, Paci E, Dokholyan N V., et al. Allostery in Its Many Disguises: From Theory to Applications. *Structure.* 2019;27(4):566-578. doi:10.1016/j.str.2019.01.003
16. Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? *Proteins Struct Funct Genet.* 2004;57(3):433-443. doi:10.1002/prot.20232

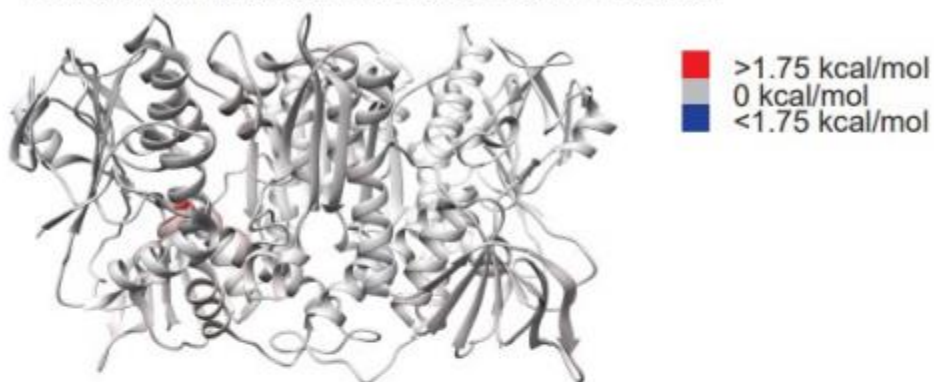
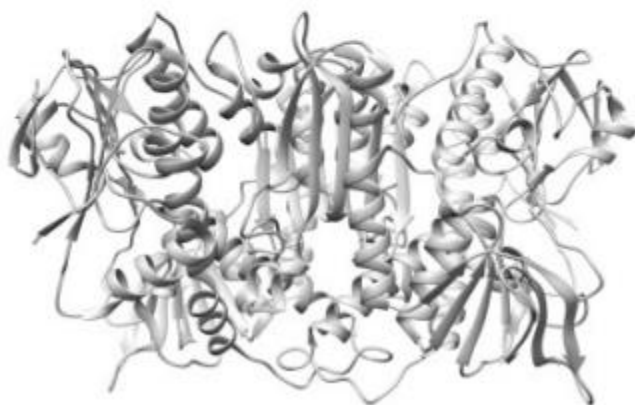
APPENDICES

APPENDIX A

2-KPCC SUPPLEMENTAL MATERIAL

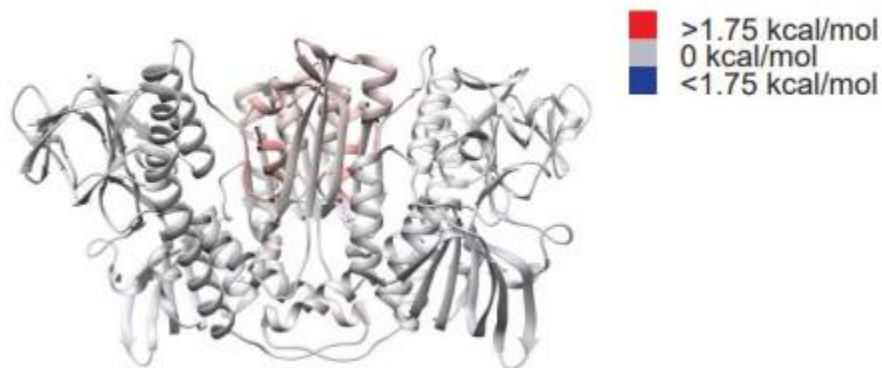


Supplemental Figure. Potential protease cleavage sites in GR and 2-KPCC. *A*, structure of GR (PDB: 1GRG) with potential thermolysin cleavage sites in *red*. *B*, structure of 2-KPCC (PDB: 1MO9) with potential thermolysin cleavage sites in *red*. 2-KPCC, 2-ketopropyl coenzyme M oxidoreductase/carboxylase; GR, glutathione reductase.

A 2-KPCC active site residue Leu502 DOWN mutated**B** 2-KPCC surface residue Gly377 DOWN mutated

Supplemental Figure. AlloSigMA control experiment mutating noncatalytically active residues. *A*, noncatalytically active site residue Leu502 was down mutated in AlloSigMA. Increase in free energy (*red*) indicates more perturbation and higher dynamics. Decrease in free energy (*blue*) indicates a stabilization and decrease in dynamics. No change in free energy (*gray*). *B*, noncatalytically active site residue Gly377 was down mutated in AlloSigMA. AlloSigMA, Allosteric Signaling and Mutation Analysis.

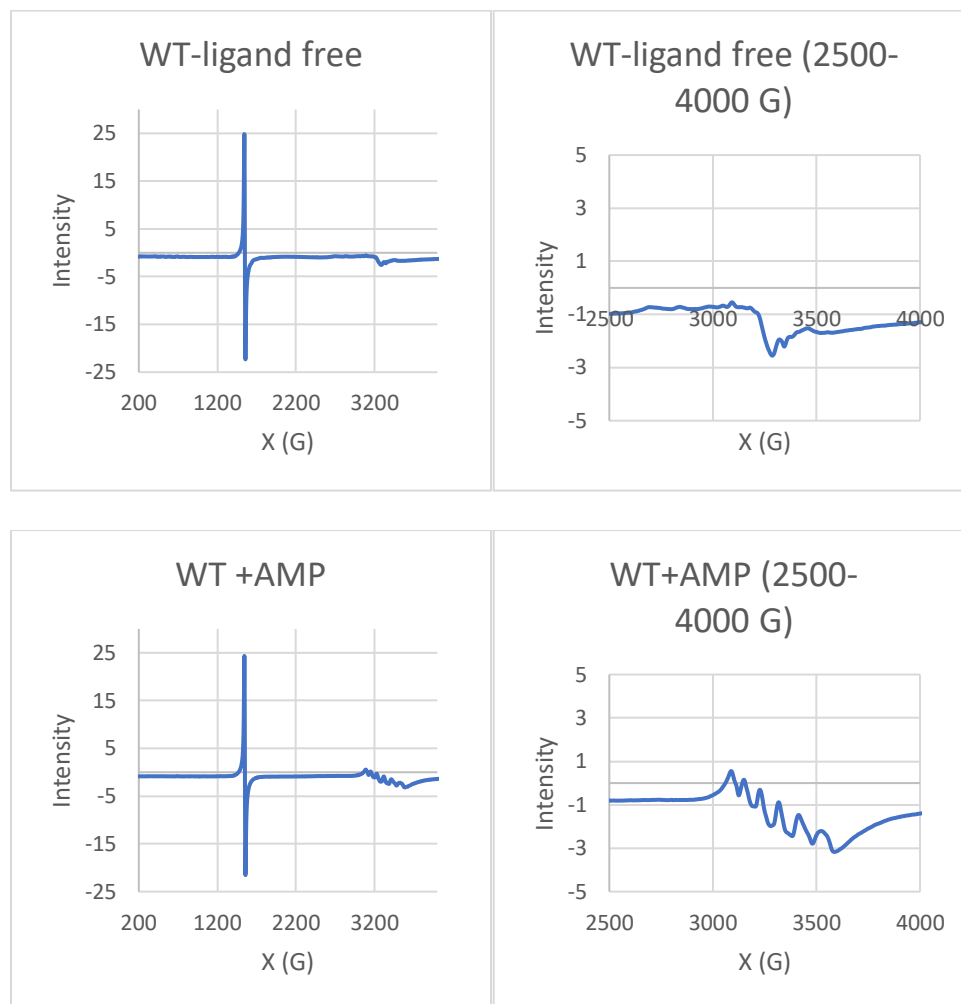
GR with catalytic dyad His467-Glu427 down mutated



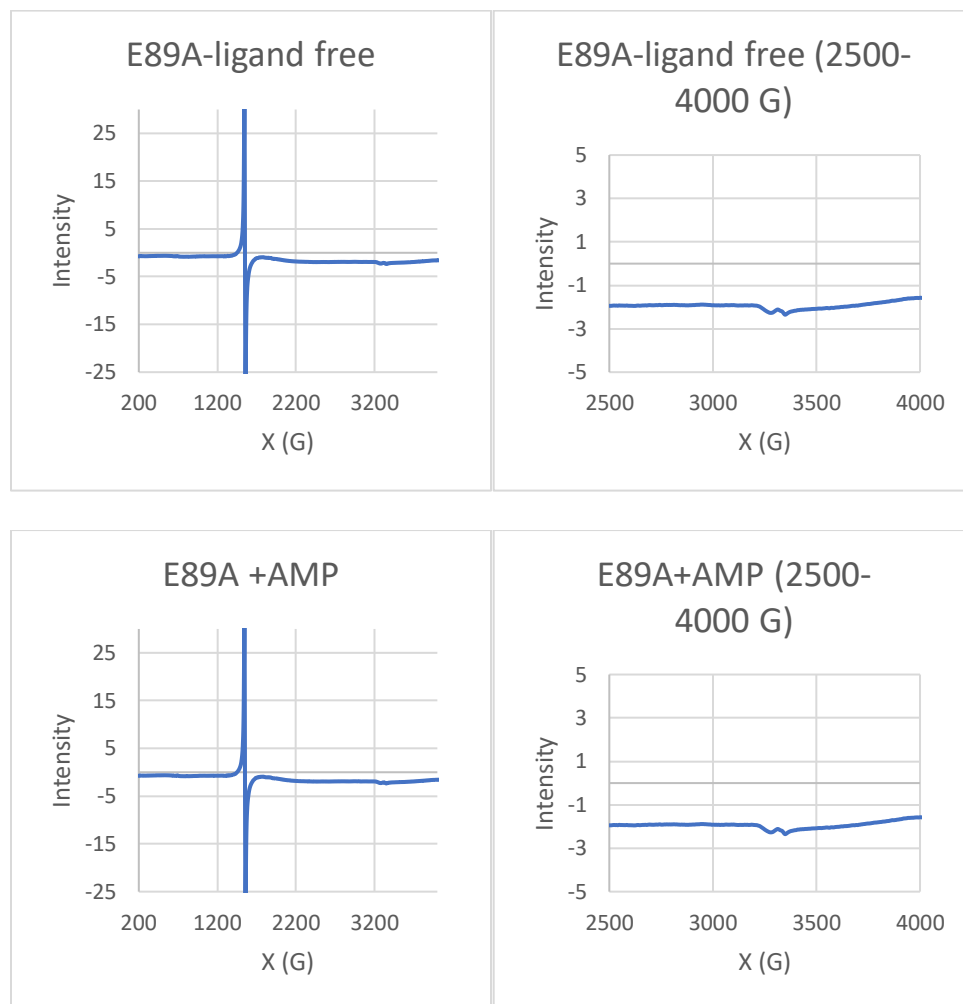
Supplemental Figure. Change in free energy calculated by AlloSigMA plotted on structure of GR (PDB:). Increase in free energy (*red*) indicates more perturbation and higher dynamics. Decrease in free energy (*blue*) indicates a stabilization and decrease in dynamics. No change in free energy (*gray*). AlloSigMA, Allosteric Signaling and Mutation Analysis; GR, glutathione reductase.

APPENDIX B

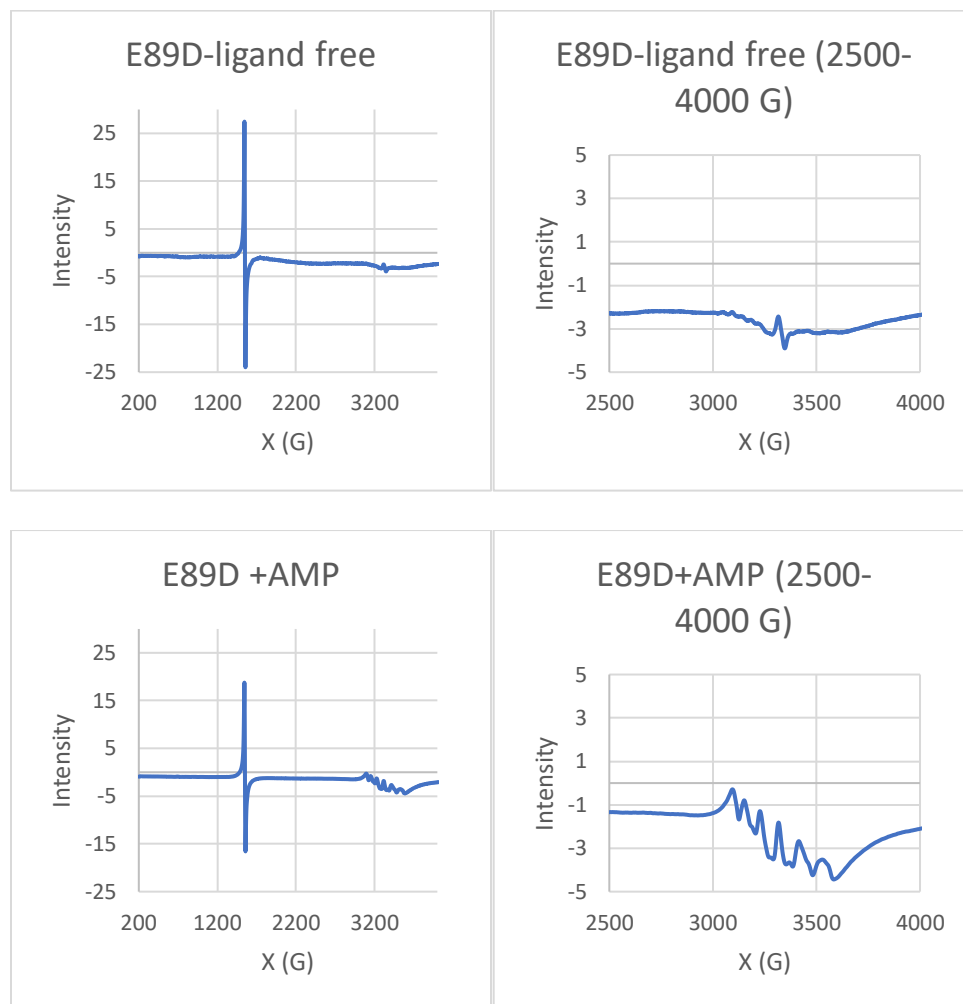
ACETONE CARBOXYLASE SUPPLEMENTAL



Supplemental WT EPR Spectra. WT acetone carboxylase with and without AMP bound. Spectra on the left column is the full spectra collected plotted. Spectra on the right are a zoomed in view of the 2500-4000G plotted.



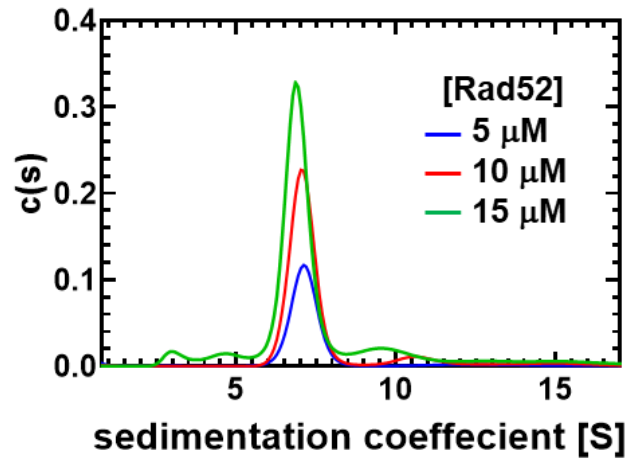
Supplemental E89A EPR Spectra. E89A variant acetone carboxylase with and without AMP bound. Spectra on the left column is the full spectra collected plotted. Spectra on the right are a zoomed in view of the 2500-4000G plotted.



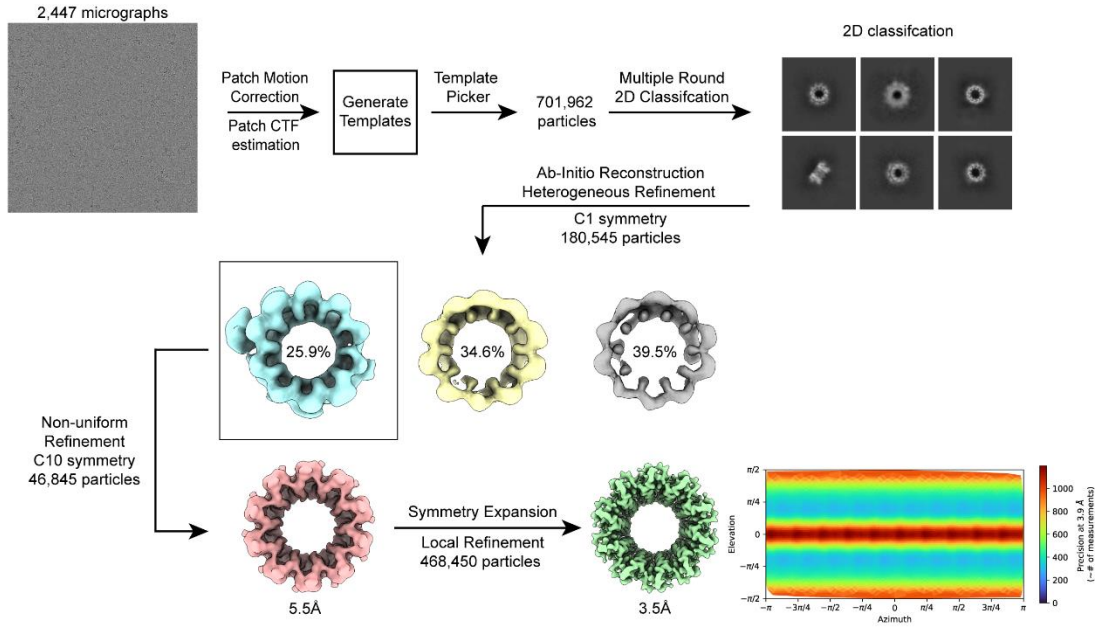
Supplemental E89D EPR Spectra. E89D variant acetone carboxylase with and without AMP bound. Spectra on the left column is the full spectra collected plotted. Spectra on the right are a zoomed in view of the 2500-4000G plotted.

APPENDIX C

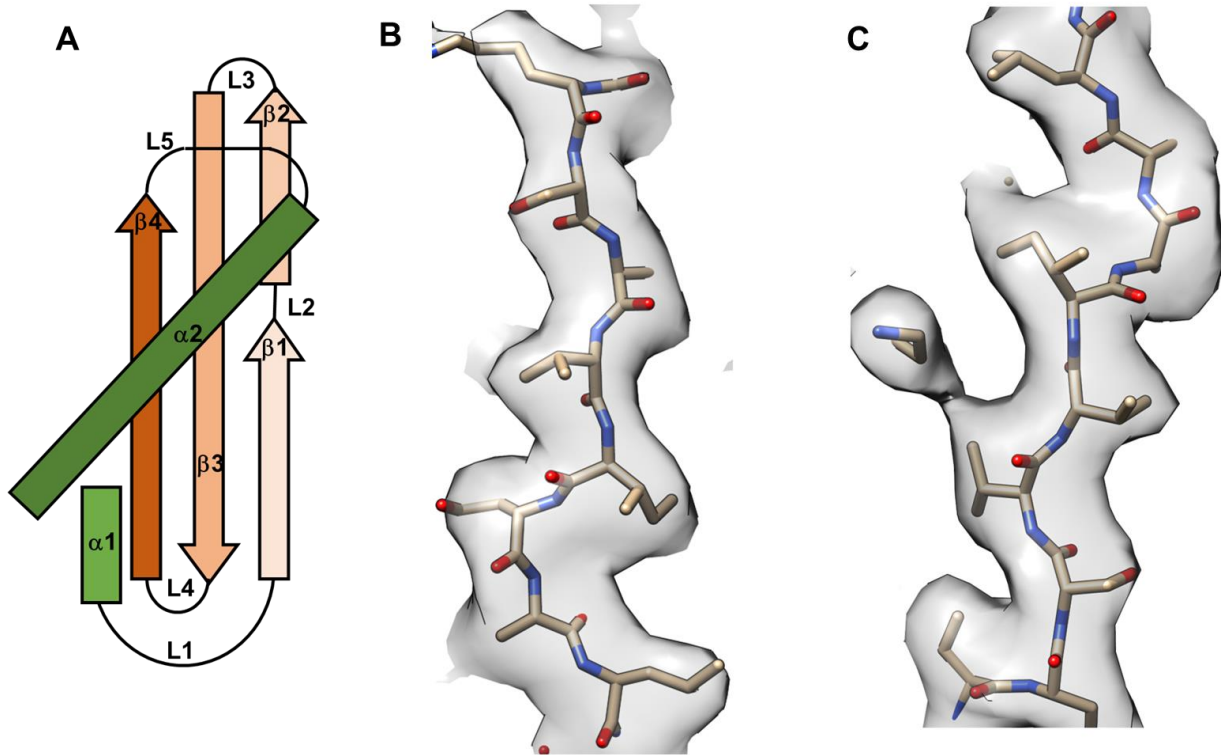
RAD52 SUPPLEMENTAL MATERIAL



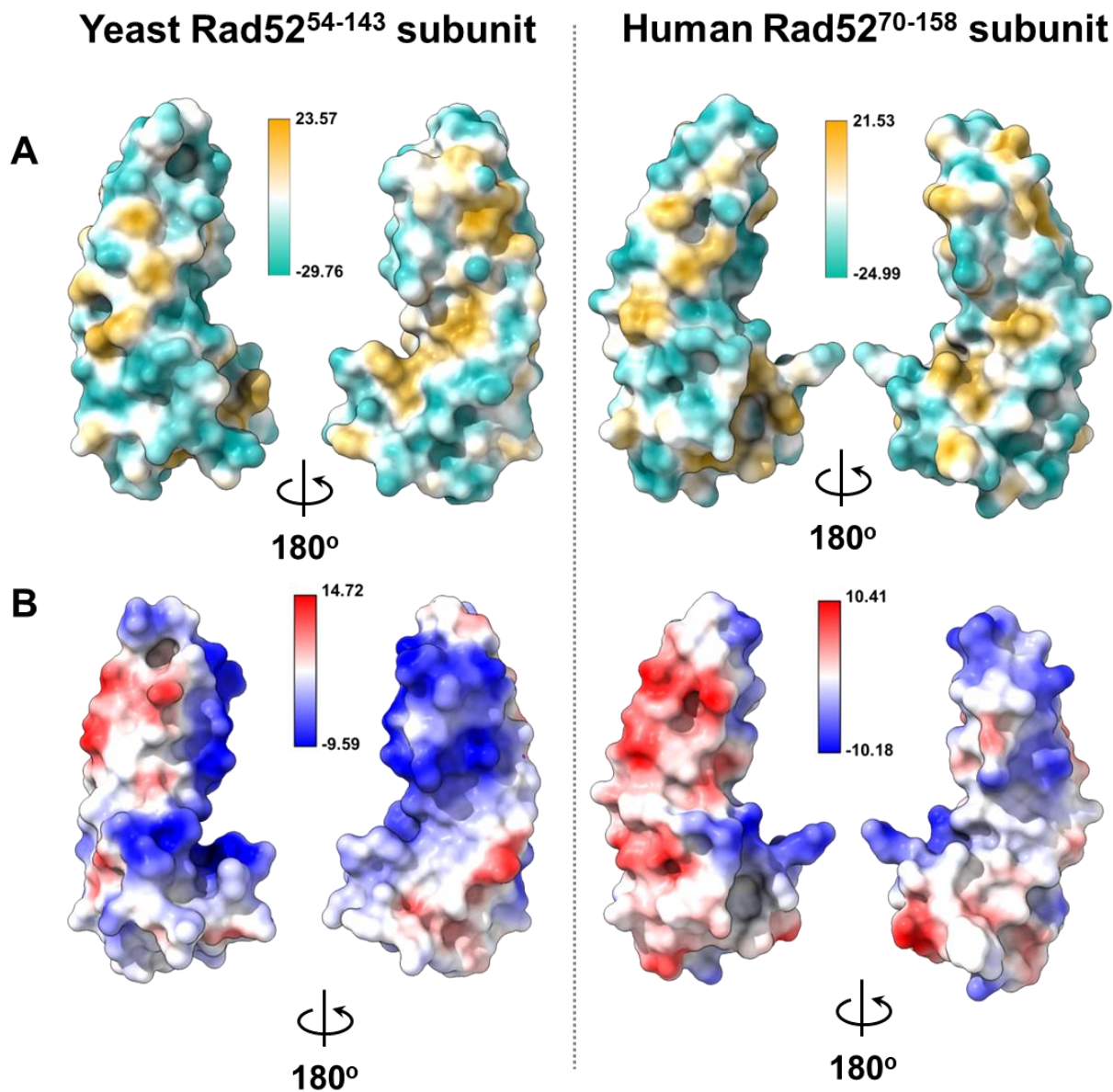
Supplementary Figure 1. *Saccharomyces cerevisiae* Rad52 is predominantly a single species solution. Sedimentation velocity analytical ultracentrifugation (AUCSV) analysis of full-length *S. cerevisiae* Rad52 was performed at various concentrations of protein (monomer concentrations noted). Rad52 is predominantly a single species in solution with an estimated molecular mass corresponding to a homodecamer (please refer to Supplementary Table 2 for measured parameters).



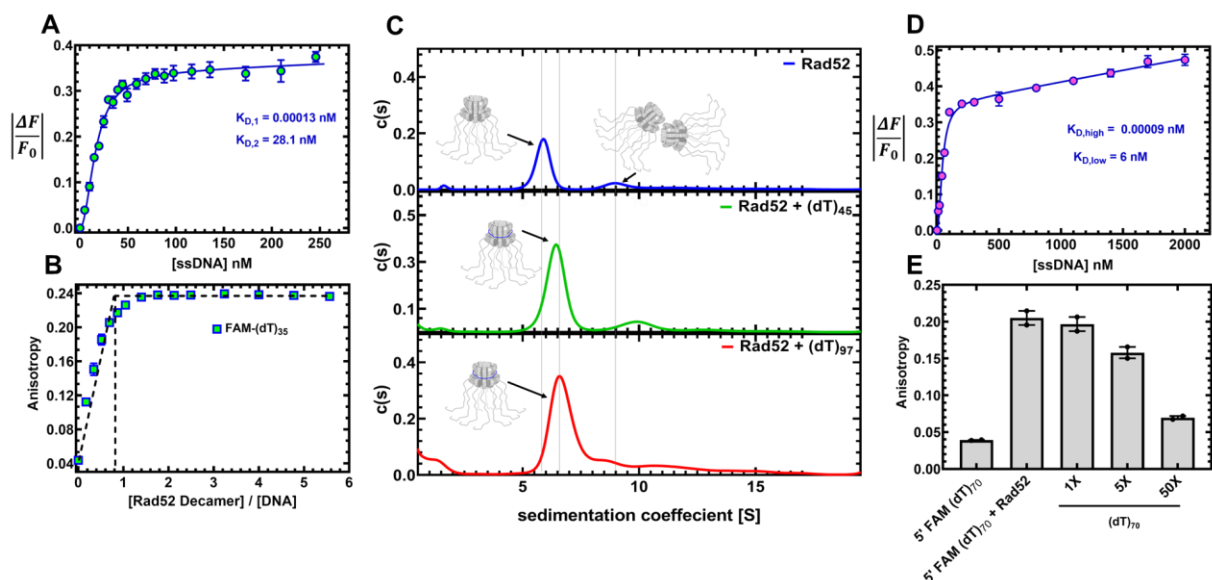
Supplementary Figure 2. Workflow of single particle analysis of the yeast Rad52 Cryo-EM data. Micrographs collected during imaging of Rad52 was analyzed as shown.



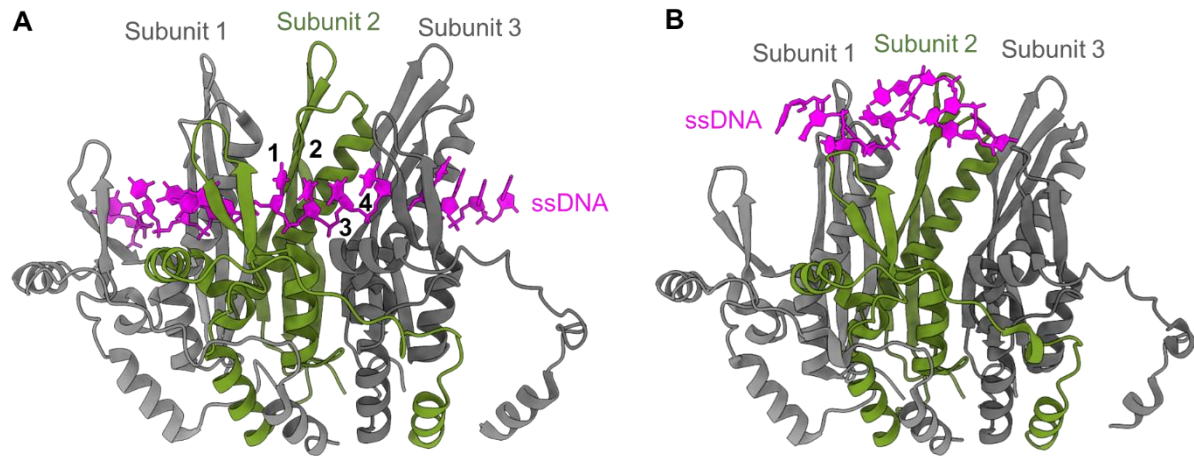
Supplementary Figure 3. Structural details of *S. cerevisiae* Rad52. A) Cartogram of the Rad52 monomer observed in the Cryo-EM structure. B) & C) Examples of sample maps and models of yeast Rad52.



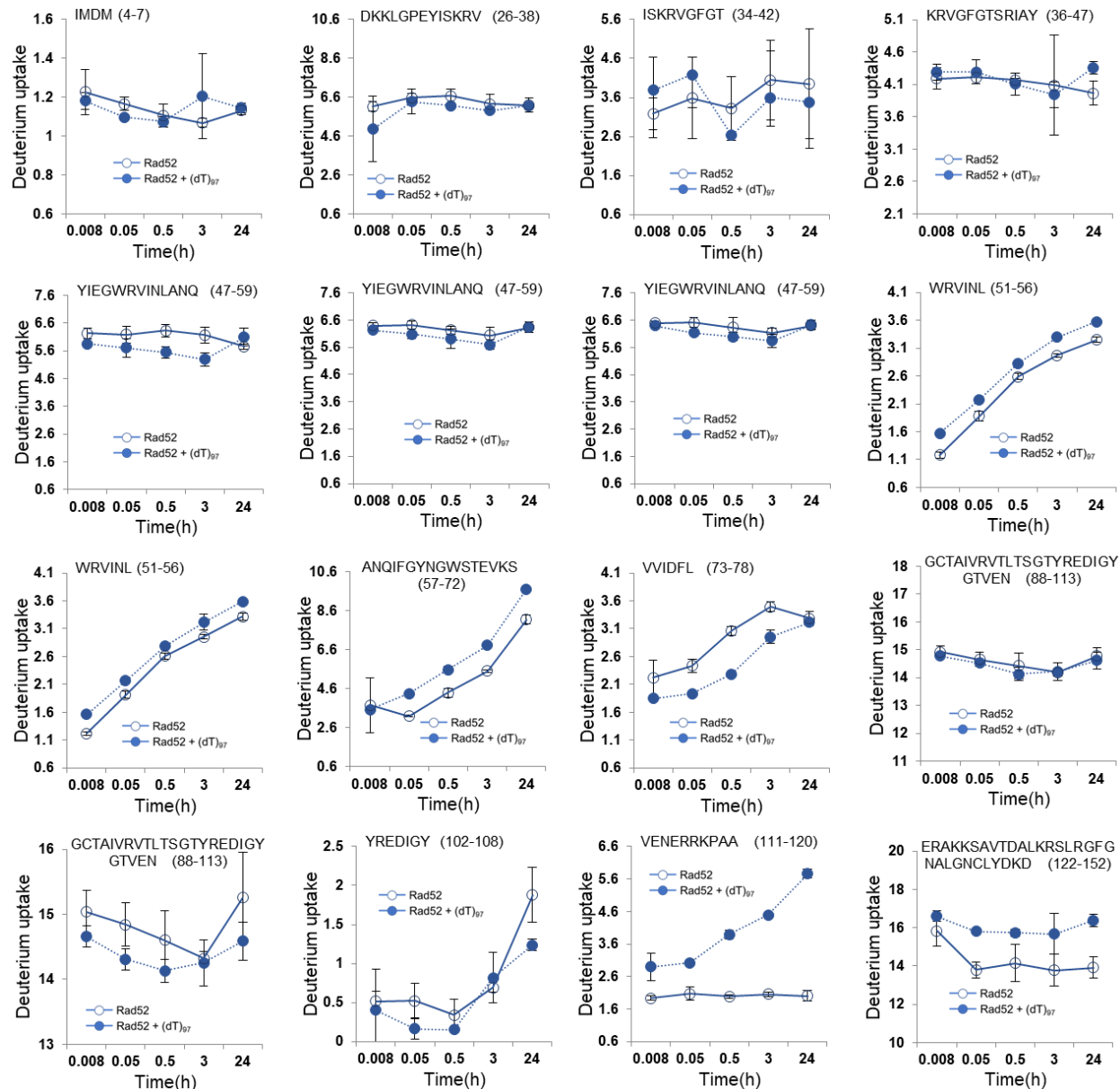
Supplementary Figure 4. Comparison of yeast and human Rad52 structures. A comparison of the A) hydrophobicity and B) electrostatic maps between yeast Rad52 and the corresponding residues in the human Rad52 crystal structure (PDB: 1KN0). The comparisons reveal the minor differences between the ordered part of the two proteins.



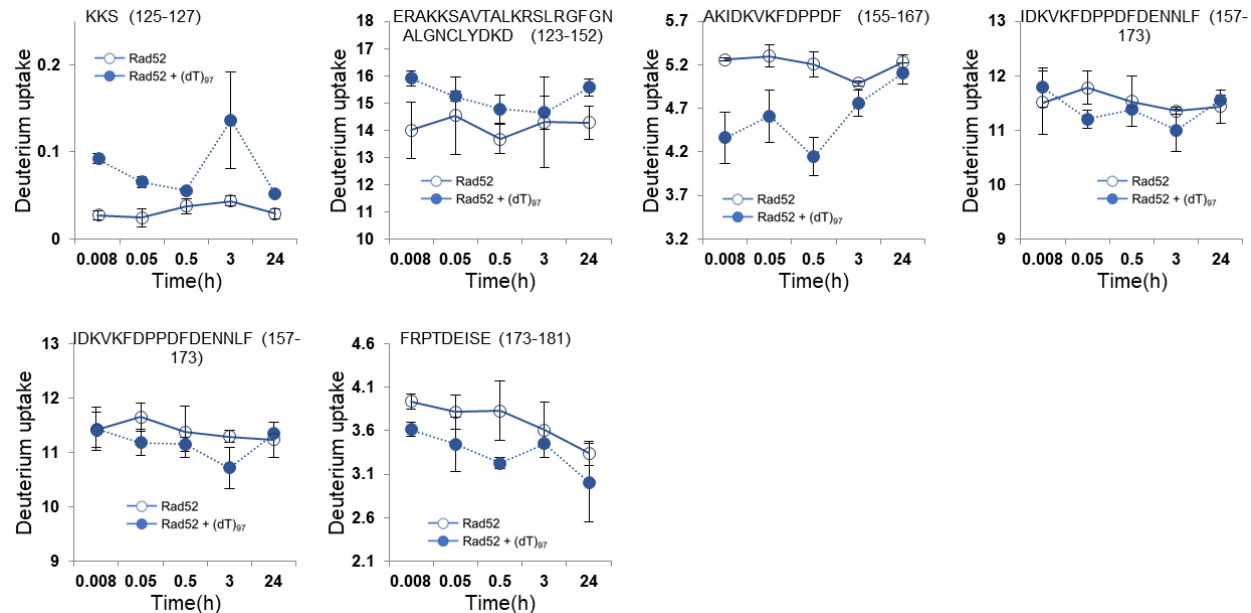
Supplementary Figure 5. Rad52 binds to ssDNA as a homodecamer. A) Intrinsic tryptophan quenching titration-based analysis of ssDNA (dT)84 binding to Rad52 is consistent with two binding sites. The data are well described by a model describing two sequential binding sites: high affinity ($K_D = 1.3 \times 10^{-12}$ M) and lower affinity binding site ($K_D = 28 \times 10^{-9}$ M). B) Fluorescence anisotropy measurement of Rad52 binding to a 5'-FAM-(dT)35 oligonucleotide shows formation of a stoichiometric complex. As expected, (dT)35 is long enough to satisfy only the high affinity site. C) Analytical ultracentrifugation analysis of Rad52 in the absence or presence of short [(dT)45] or longer [(dT)97] ssDNA substrates predominantly show a decamer forming 1:1 complex with DNA. The double decamer species is also capable of binding to ssDNA. D) Investigation of the Trp quenching of Rad52 at significantly higher ssDNA concentrations shows additional signal changes. These changes could be attributed to exposure and binding of multiple ssDNA to lower affinity cryptic sites, formation of protein-protein, and protein-ssDNA dependent aggregates, or that ssDNA is not uniformly bound around a single Rad52 homodecameric ring and can likely be looped out between the subunits. E) Competition experiments were performed by titrating in increasing concentrations of unlabeled (dT)70 into a preformed Rad52:5'-FAM-(dT)70 complex. Data shows that the Rad52 possesses a limited number of ssDNA binding sites, and the bound ssDNA is in dynamic equilibrium with the pool of free ssDNA in the solution.



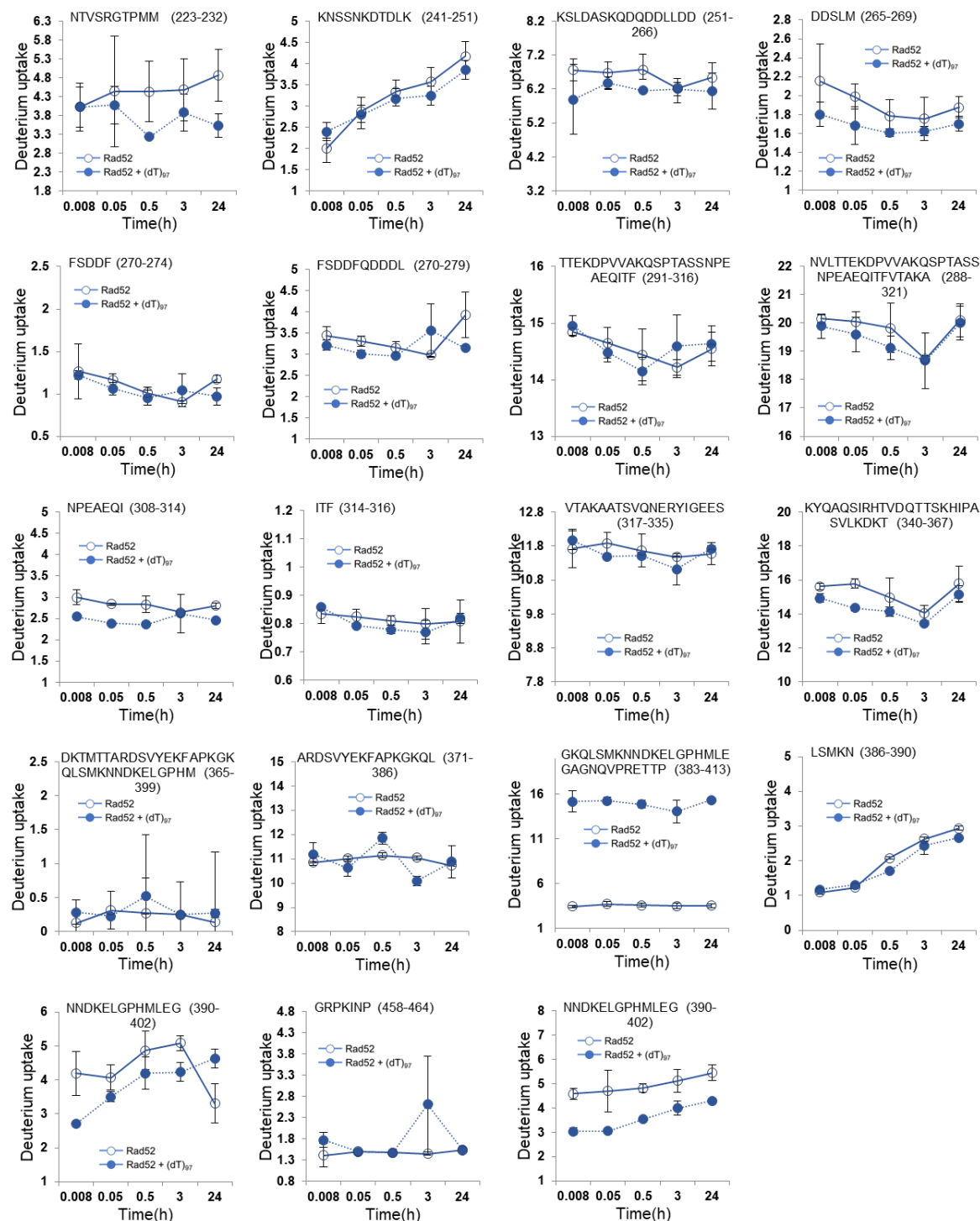
Supplementary Figure 6. Crystal structures of ssDNA bound human Rad52. Three subunits (colored individually) from the undecameric human Rad52 crystal structures are shown. A) ssDNA (dT)40 bound to the inner binding site was captured when the outer binding site was perturbed by mutating Lys-102 and Lys-133 (PDB ID 5XRZ). B) 6 nt of a (dT)40 ssDNA substrate is observed bound to outer site was captured sandwiched between two Rad52 undecameric rings (PDB ID 5XS0). Shown here are contacts between a single ring and ssDNA.



Supplementary Figure 7. Deuterium uptake changes captured in HDX-MS analysis of yeast Rad52. The plotted data for deuterium incorporation as a function of time are shown for individual peptides identified in the HDX-MS analysis. This panel depicts peptides corresponding to amino acid positions 1-152 in the N-terminal half. Data from both Rad52 and the Rad52-(dT)₉₇ complex are compared for each peptide.



Supplementary Figure 8. Deuterium uptake changes captured in HDX-MS analysis of yeast Rad52. The plotted data for deuterium incorporation as a function of time are shown for individual peptides identified in the HDX-MS analysis. This panel depicts peptides corresponding to amino acid positions 125-181 in the N-terminal half. Data from both Rad52 and the Rad52-(dT)₉₇ complex are compared for each peptide.



Supplementary Figure 9. Deuterium uptake changes captured in HDX-MS analysis of yeast Rad52. The plotted data for deuterium incorporation as a function of time are shown for individual peptides identified in the HDX-MS analysis. This panel depicts peptides corresponding to amino acid positions 223-464 in the C-terminal half. Data from both Rad52 and the Rad52-(dT)97 complex are compared for each peptide.

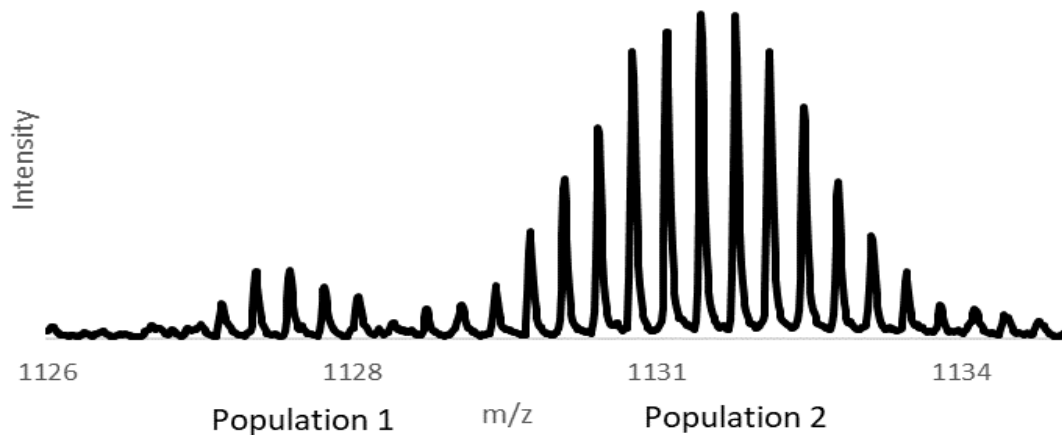


Yeast Rad52 –Alphafold

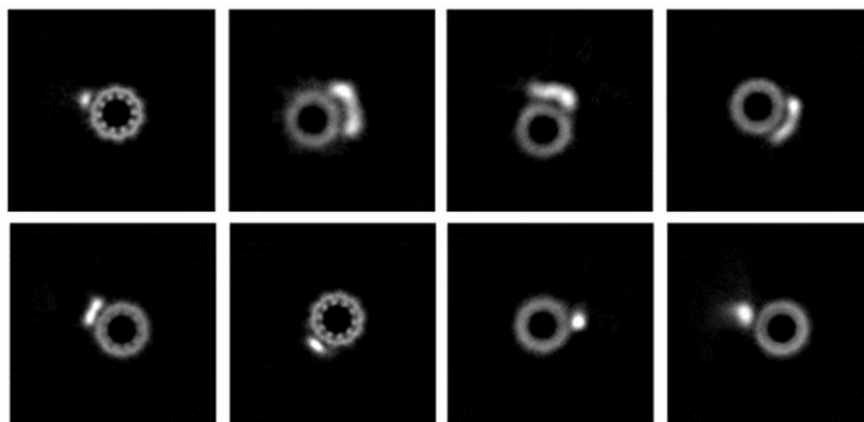
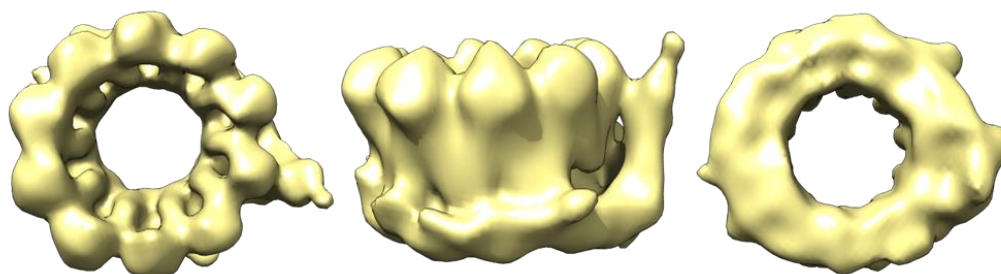
Yeast Rad52 -CryoEM

Supplementary Figure 10. Comparison of the AlphaFold model and cryo-EM structure of yeast Rad52. The AlphaFold model of *S. cerevisiae* Rad52 (AF-P06778-F1; green) overlaid with the ordered part of the N-terminal half observed in the cryo-EM structure (purple). Excellent agreement in the structures is observed in the N-terminal half between the two.

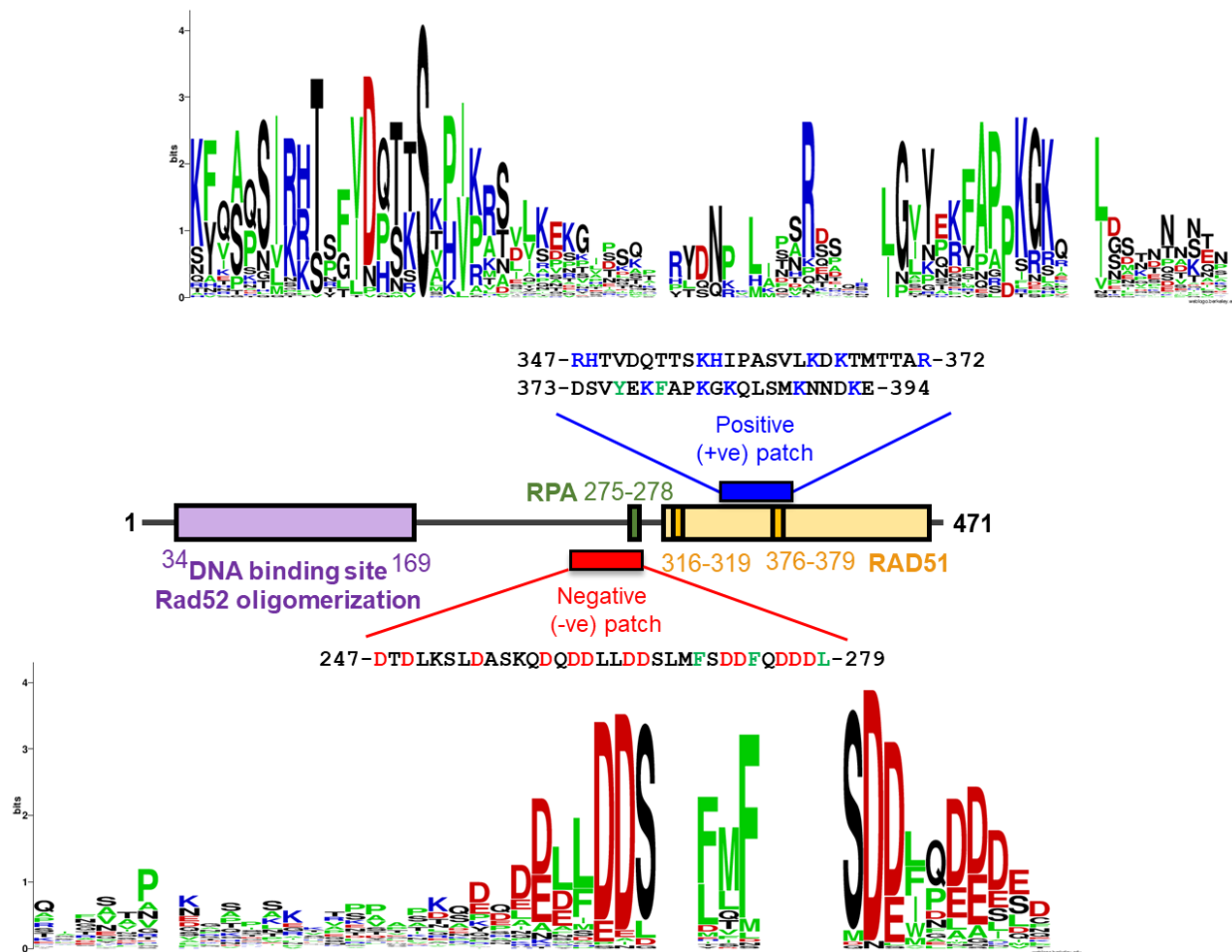
384-413: +3 GKQLSMKNNDKELGPHMLEGAGNQVPRETTP



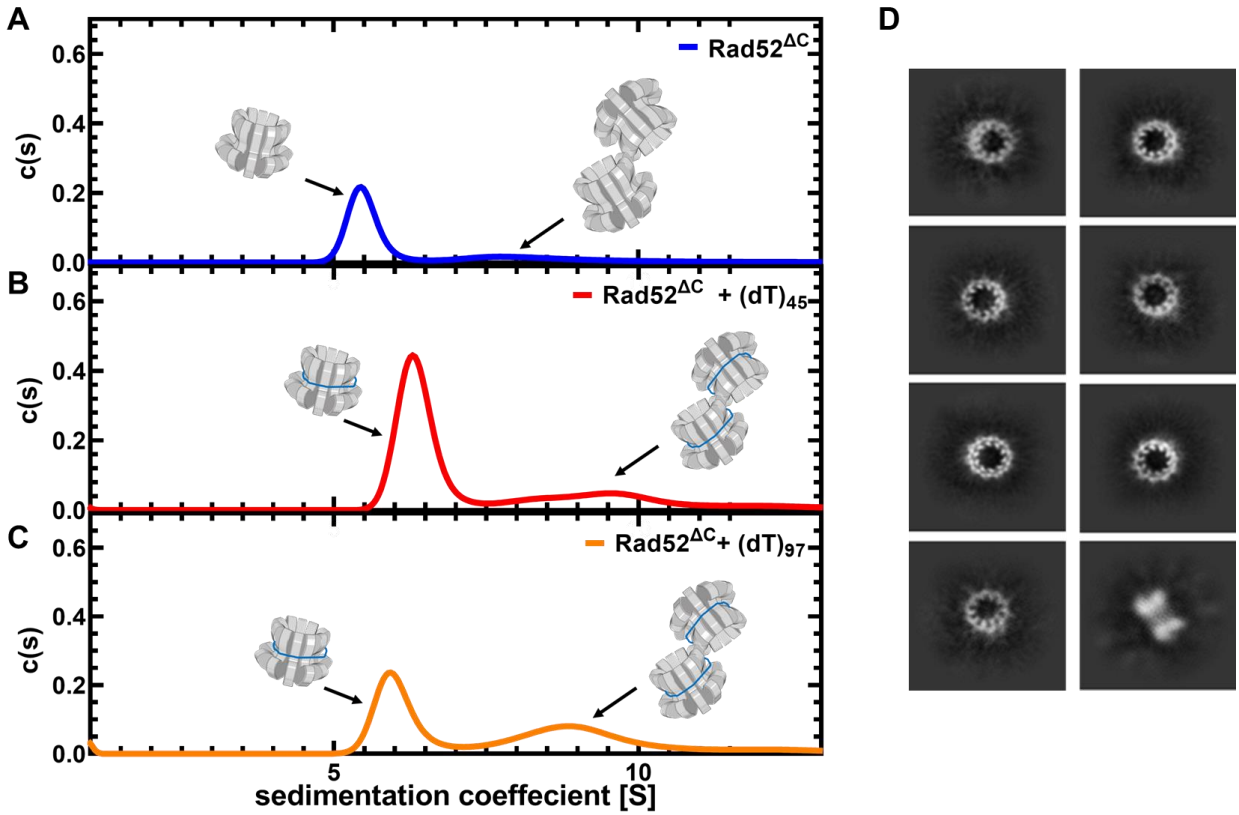
Supplementary Figure 11. Hydrogen/deuterium exchange data of Rad52 peptide 384-413 has a bimodal distribution. Mass spectra of the (M+3H)³⁺ charge state of peptide 384-413 following incubation in 90% deuterium buffer for 30 seconds. Two distinct populations (envelopes) of deuterium uptake were observed. Relative abundance of the different peptide forms is shown in Fig 3E and 3F.

A 2D classes with C-terminal density**B** 3D volume with C-terminal density

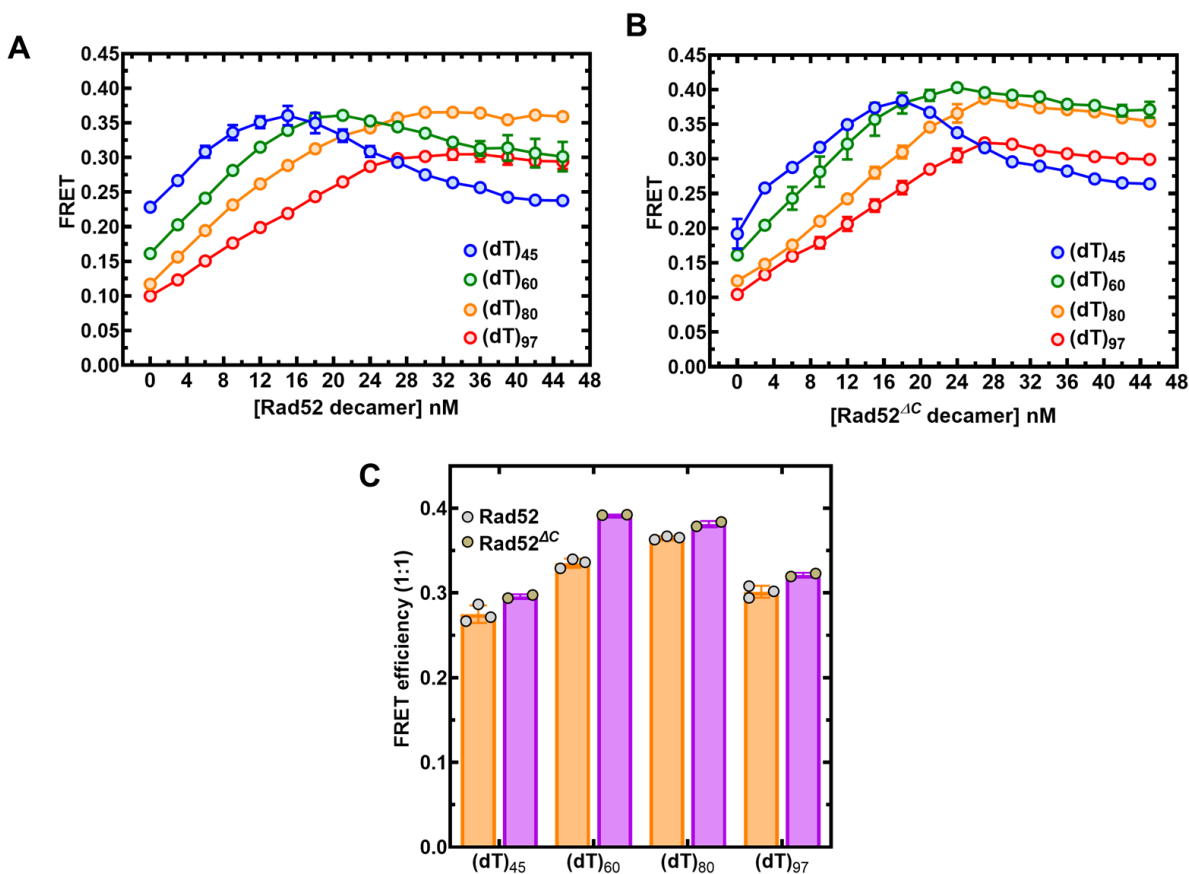
Supplementary Figure 12. Additional density for C-terminal region positioned adjacent to the ordered N-terminal homodecameric ring. A) Examples of 2D classes that are captured in the cryo-EM analysis of Rad52. These classes were discarded in the structure determination, but clearly show presence of additional density adjacent to the ordered N-terminal ring. B) Subsequent 3D volumes show that the C-terminal region sits adjacent to the ring but appears to be asymmetrically positioned to one part. The underlying allosteric mechanisms that prevent other subunits in the ring from interacting with their respective C-terminal regions remains to be determined.



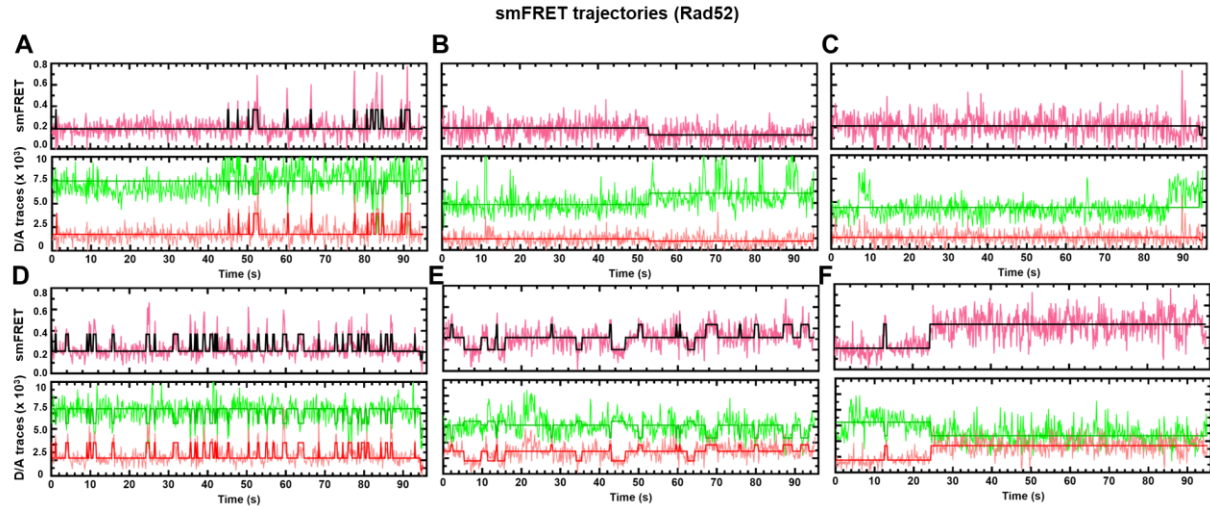
Supplementary Figure 13. Sequence conservation analysis of the C-tail. WebLogo represents sequence similarity analysis of regions in the disordered C-terminal half of Rad52 carrying the Rad51 binding site (positive patch) and the RPA binding site (negative patch). The analysis encompasses Rad52 proteins from 66 different fugal species. Several residues are highly conserved.



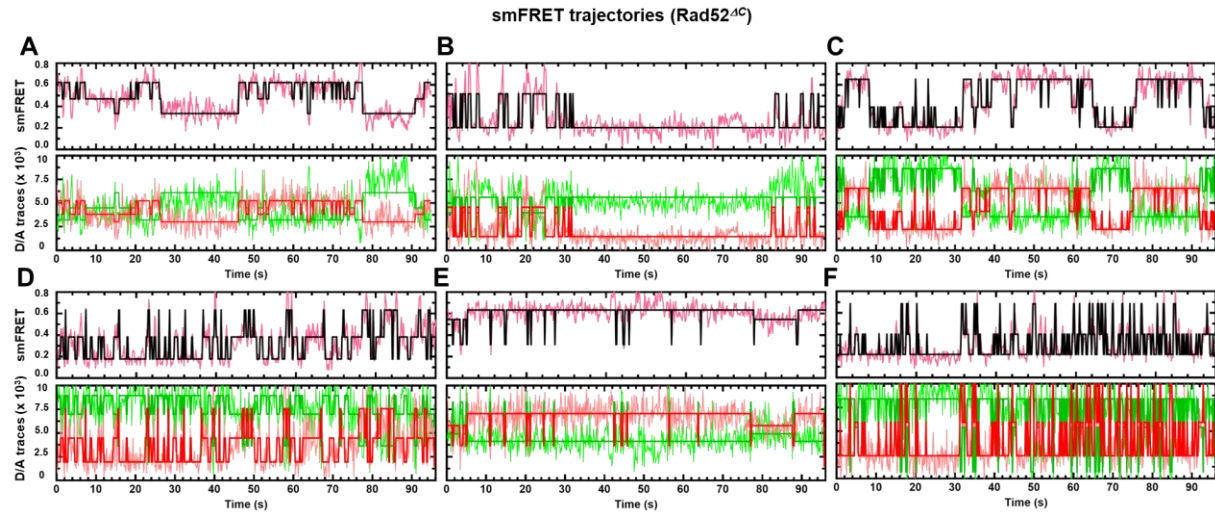
Supplementary Figure 14. $\text{Rad52}^{\Delta C}$ is also a homodecamer in solution. Analytical ultracentrifugation sedimentation velocity analysis of A) $\text{Rad52}^{\Delta C}$ in the absence or presence of B) $(\text{dT})_{45}$ or C) $(\text{dT})_{97}$ show that the N-terminal half of the protein primarily behaves as a single higher order species corresponding to a homodecamer. A small fraction of the double-ring species is also observed. Both species bind to short and longer ssDNA, however, the species distribution does not change. D) 2D classes of $\text{Rad52}^{\Delta C}$ in cryo-EM analysis show homodecameric rings.



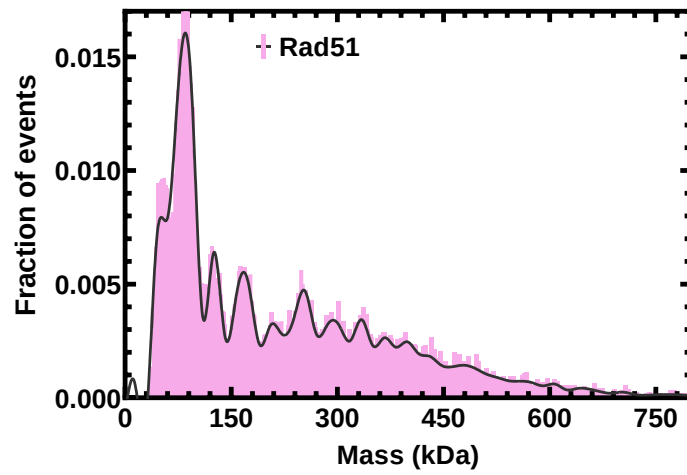
Supplementary Figure 15. ssDNA length dependent change in Rad52 binding induced FRET. ssDNA of varying lengths carrying 5'-Cy5 and 3'-Cy3 fluorophores were subject to A) Rad52 or B) Rad52 Δ C binding. Changes in FRET were captured by exciting Cy3 and monitoring changes in Cy5 emission. C) Optimal FRET for Rad52 is observed on the (dT)₈₀ substrate. For Rad52 Δ C, highest FRET efficiency is observed on the (dT)₆₀ substrate. These data also reveal that the C-tail influences wrapping of the ssDNA around Rad52.



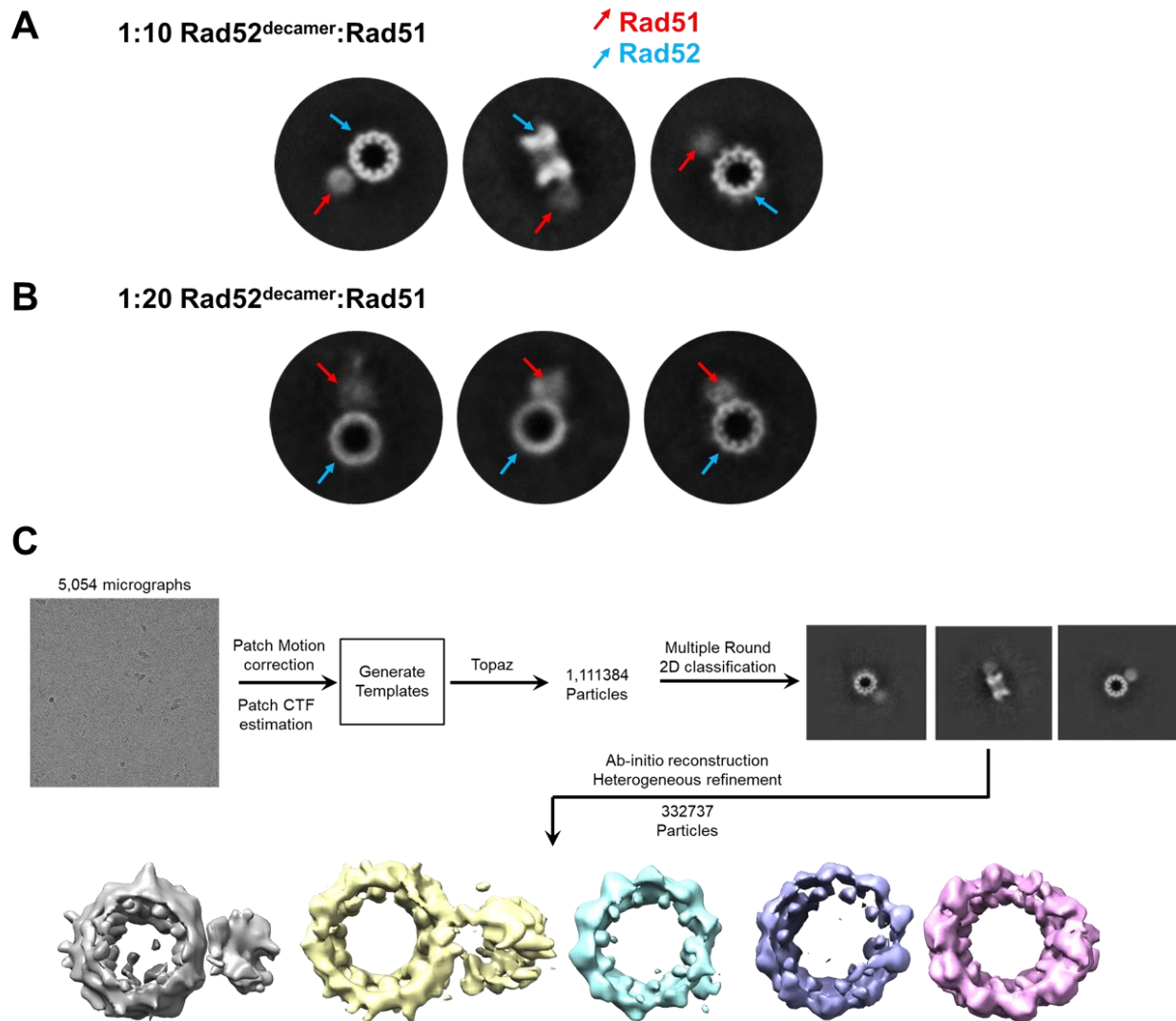
Supplementary Figure 16. Additional single-molecule FRET trajectories for Rad52. Trajectories from six different single DNA molecules are shown from the analysis of full length Rad52.



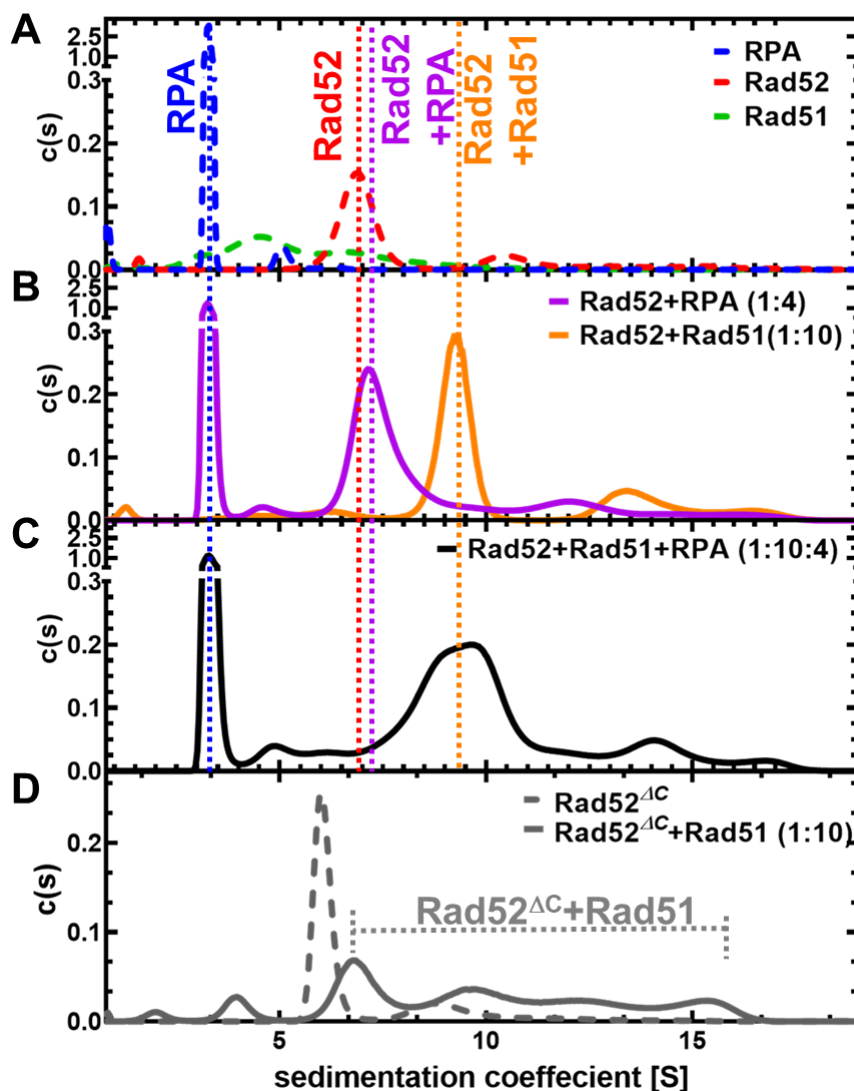
Supplementary Figure 17. Additional single-molecule FRET trajectories for Rad52^{ΔC}. Trajectories from six different single DNA molecules are shown from the analysis of Rad52^{ΔC}.



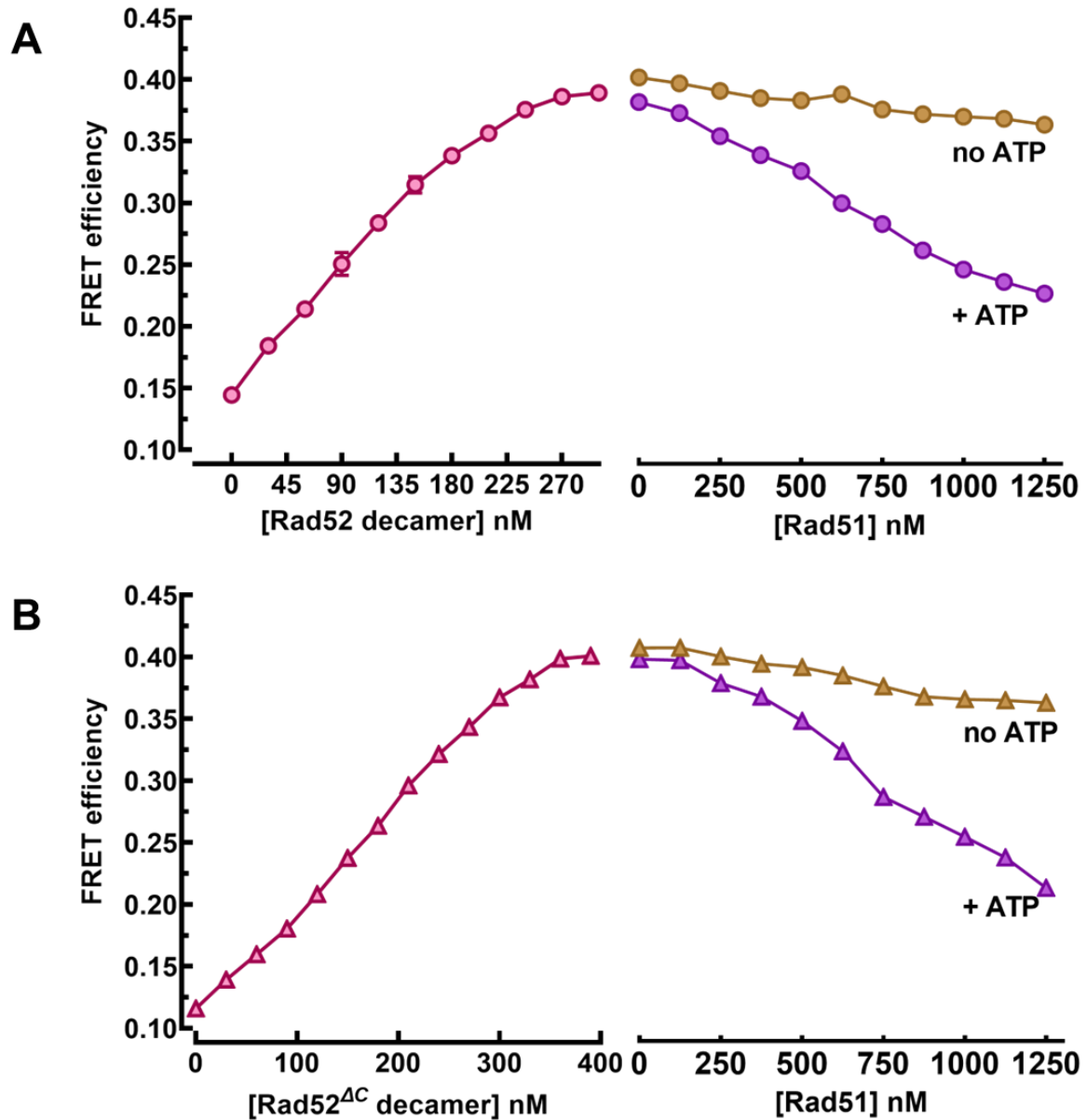
Supplementary Figure 18. Mass photometry analysis of yeast Rad51. Mass photometry analysis of *S. cerevisiae* Rad51 shows the presence of multiple oligomeric species.



Supplementary Figure 19. Interaction of Rad52-Rad51 complex. 2D class averages from the cryo-EM analysis of the Rad51-Rad52 complex when formed using A) 1:10 and B) 1:20 molar ratios of Rad52 (decamer) to Rad51 are shown. Rad51 molecules are bound to the Rad52 ring, but only at a single defined position. C) Workflow of Single particle analysis of Rad52-Rad51 Cryo-EM data



Supplementary Figure 20. RPA, Rad52, and Rad51 form a stable complex in the absence of ssDNA. Analytical ultracentrifugation sedimentation velocity analysis of RPA, Rad51, and Rad52 alone and in complex. **A)** Sedimentation profiles of RPA (4 μ M), Rad52 (10 μ M monomer), and Rad51 (10 μ M). RPA is a single heterotrimer. Rad52 is predominantly a homodimer. Rad51 migrates as several polydisperse complexes. **B)** A complex between Rad52 (10 μ M) and RPA (4 μ M) or Rad51 (10 μ M) shows formation of complexes. Even though 10 RPA binding sites are present in the Rad52 decamer, only a few RPA molecules are bound to the Rad52 ring. **C)** When all three proteins are mixed, a large species corresponding to RPA-Rad52-Rad51 are observed. **D)** Rad52 $^{\Delta C}$ displays reduced complex formation with Rad51. The sedimentation profiles of Rad52 $^{\Delta C}$ or Rad52 $^{\Delta C}$ +Rad51 (10 μ M monomer of Rad52 $^{\Delta C}$ with or without 10 μ M Rad51) show a defined Rad52 $^{\Delta C}$ species. In contrast, the Rad52 $^{\Delta C}$ +Rad51 complex displays more polydisperse behavior suggesting non-uniform complex formation



Supplementary Figure 21. Ensemble FRET measurements of Rad51-binding to Cy3-Cy5 end labeled (dT)80 ssDNA. A) Rad52 or B) Rad52^{ΔC} were prebound to a Cy3-(dT)80-Cy5 substrate and Rad51 was introduced in the absence or presence of ATP. Data shows that both Rad52 and Rad52^{ΔC}-DNA complexes can be remodeled by Rad51 when sufficient time is provided to reach equilibrium.

(dT)₈₄	Rad52(n=	Rad52 +	Rad52^{4C}	Rad52^{4C} +
kDa	8)	(dT)₈₄	(n=3)	(dT)₈₄
	kDa	(n = 11)kDa	kDa	(n=3)kDa
63±19	72.9±14	87±26.	66±33.5	97
-	498.4	522.5±25	255±12	293±10
	±24			
-	991.8±7	1078±68	517	606±17
	5		±29	
-	1551(±9	1562±37	782	931.5±29
	8.99)		±46	

Supplementary Table I: Mass Photometer mass distribution. 1) Major peak (>85% of the population) is highlighted in bold 2) All the different Rad52 and Rad52^{4C} peaks show ssDNA binding as can be seen in the systematic shift in the entire distribution 3) The (dT)₈₄ mass estimation is prone to lot of error because the actual mass of the ssDNA is 25,490.5 gm/mole which is at/below the detection limit of the of the scope

Rad52	sw	Sw	Mw	Stokes	a/b	a/b	%
		(20,	(kDa)	Radius	(oblate)	(prolate)	
		w)		20C			
				(nm)			
5μM	7.25	11.3	532	11.2	19.0	16.7	73.7
	11.0	17.1	991	13.8	19.0	16.7	04.4
	14.4	22.4	1484	15.8	19.0	16.7	02.2
10μM	7.14	11.1	521	11.2	19.0	16.8	81.8
	10.6	16.5	944	13.6	19.0	16.8	05.9
	13.9	21.7	1421	15.6	19.0	16.8	04.2
15μM	7.46	11.6	537	11.0	17.9	15.7	70.8
	13.11	20.4	1250	14.6	17.9	15.7	3.0

Supplementary Table 2: Velocity AUC c(s) numbers

	Rad52 Apo
EMD	29695
PDB	8G3G
Data Collection and processing	
Magnification	150,000x
Voltage (kV)	200
Electron exposure (e ⁻ / Å ²)	50.7
Defocus range (μm)	-0.8 to -2.4
Nominal Pixel Size (Å)	0.94
Symmetry imposed	C10
Initial Particle images	701,962
Final Particle images	180,545
Map resolution (Å)	3.3/3.4/3.8
FSC threshold	0/0.143/0.5
Map resolution range (Å)	2.8 to 6.0
Refinement	
Model resolution (Å)	3.5
FSC threshold	0.143
Map sharpening B-factor (Å ²)	
Model composition	

Chains	10
Non-hydrogen atoms	6930
Protein residues	890
Ligands	0
Water	0
Average B-factors ($\text{\AA}^2$)	
Protein	71.88
Bond RMSD	
Bond lengths ($\text{\AA}$)	0.003
Bond angles ($^\circ$)	0.719

Validation

MolProbity score	1.49
All-atom clash score	8.27
Ramachandran plot	
Favored (%)	97.87
Allowed (%)	2.18
Outliers (%)	0

Supplementary Table 4: DNA oligonucleotides used in various assays.

Name	Sequence (5' to 3')	Comments
5'-FAM-(dT) ₃₅	5'- /56-FAM/TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TT -3'	Fluorescence Anisotropy
5'-FAM-(dT) ₇₀	5'- /56-FAM/TTT T -3'	Fluorescence Anisotropy
5'-FAM-32bp-duplex	5'-/56-FAM/CGA ATG TAA CCA TCG TTG GTC GGC AGC AGG GC -3'	5'-FAM labelled oligo (sequence shown on the left) was hybridized with an unlabeled complementary strand to form a 32 base paired duplex for Rad52 binding experiments
(dT) ₂₀	5'- TTT TTT TTT TTT TTT TTT TT -3'	Trp quenching
(dT) ₃₀	5'- TTT TTT TTT TTT TTT TTT TTT TTT TTT -3'	Trp quenching
(dT) ₄₅	5'- TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT -3'	Trp quenching
(dT) ₆₀	5'- TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT -3'	Trp quenching
(dT) ₈₄	5'- TTT -3'	Trp quenching, Mass Photometry
(dT) ₉₇	5'- TTT	Cryo-EM substrate, AUC

	TTT TTT TTT TTT TTT T -3'	
poly(dT)	-	Mean oligo length is 200 bases
5'-Cy5-(dT) ₄₅ -Cy3-3'	5'- /5Cy5/TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT /3Cy3Sp/ -3'	Fluorometer based ssDNA wrapping FRET assay
5'-Cy5-(dT) ₆₀ -Cy3-3'	5'- /5Cy5/TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT /3Cy3Sp/ -3'	Fluorometer based ssDNA wrapping FRET assay
5'-Cy5-(dT) ₈₀ -Cy3-3'	5'- /5Cy5/TTT TT/3Cy3Sp/ -3'	Fluorometer based ssDNA wrapping FRET assay
5'-Cy5-(dT) ₉₇ -Cy3-3'	5'- /5Cy5/TTT T/3Cy3Sp/ - 3'	Fluorometer based ssDNA wrapping FRET assay
5' -Cy5-18mer-Bio-3'	5' /5Cy5/GCC TCG CTG CCG TCG CCA/3Bio/ 3'	Biotin-ssDNA handle with 5'- Cy5
5' - 18mer-(dT) ₄₀ -Cy3-3'	5'- TGG CGA CGG CAG CGA GGC (dT) ₄₀ /3Cy3Sp/ -3'	Partial duplex with (dT) ₄₀ overhang
5' - 18mer-(dT) ₆₀ -Cy3-3'	5'- TGG CGA CGG CAG CGA GGC (dT) ₆₀ /3Cy3Sp/ -3'	Partial duplex with (dT) ₆₀ overhang
5' - 18mer-(dT) ₈₀ -Cy3-3'	5'- TGG CGA CGG CAG CGA GGC (dT) ₈₀ /3Cy3Sp/ -3'	Partial duplex with (dT) ₈₀ overhang

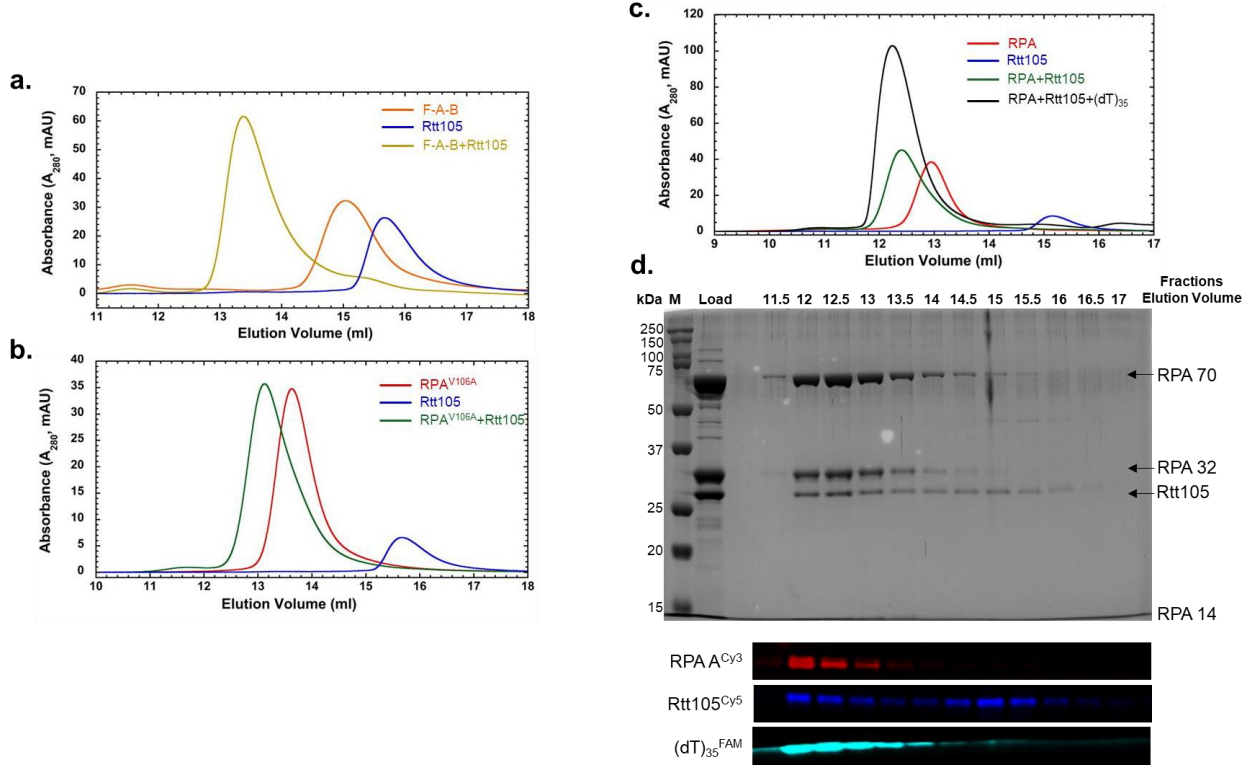
Supplementary Table 4: DNA oligonucleotides used in various assays.

Name	Sequence (5' to 3')	Comments
Rad52 ¹⁻²¹² F	5'-ATATACATATGAATGAAATTATGGATATGGATGAG-3'	Rad52 ^{ΔC}
Rad52 ¹⁻²¹² R	5'-GCAGTACTCGAGCTTCGTCGAGTCGGGATTG-3'	Rad52 ^{ΔC}
Rad52 ²¹³⁻⁴⁷¹ F	5'- ATATACATATGAACCTGGTGAAAATAGAAAATACAGTA AGTCG-3'	Rad52 ^{ΔN}
Rad52 ²¹³⁻⁴⁷¹ R	5'-GCAGTACTCGAGTCAAGTAGGCTTGCGTGCA-3'	Rad52 ^{ΔN}
Rad52 ²⁹³⁻⁴⁷¹ F	5'-ATATACATATGGATCCCGTTGTAGCTAAGCAAAGCCC- 3'	Rad52 ^{ΔN*}
Rad52 ²⁹³⁻⁴⁷¹ R	5'-GCAGTACTCGAGTCAAGTAGGCTTGCGTGCATGCA-3'	Rad52 ^{ΔN*}

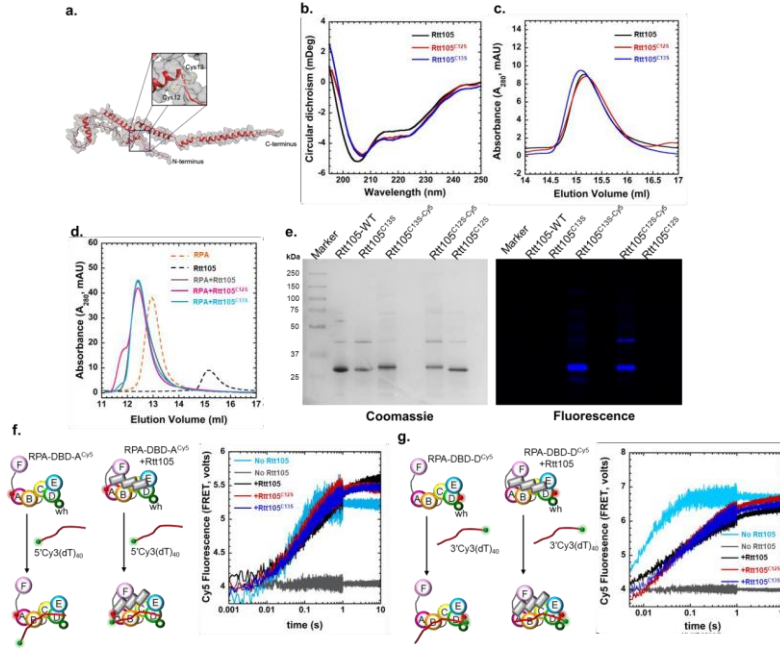
Supplementary Table 5: DNA Primers used for Rad52 mutagenesis.

APPENDIX D

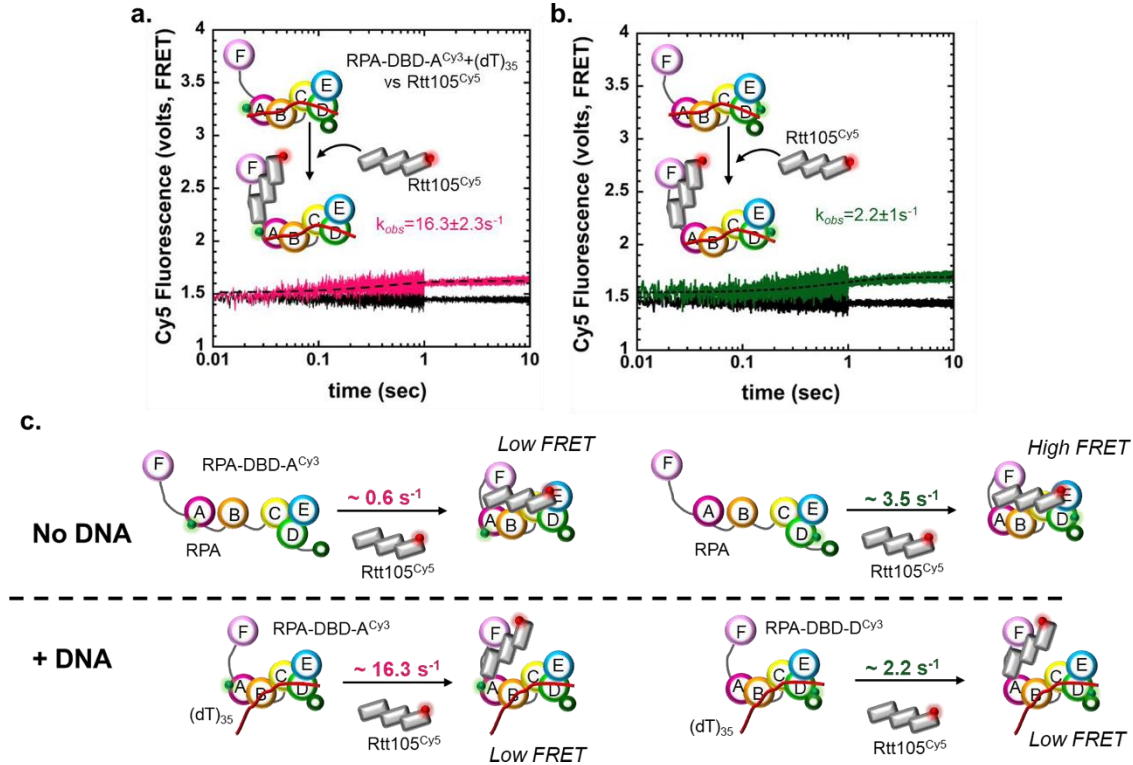
RPA AND RTT105 SUPPLEMENTAL MATERIAL



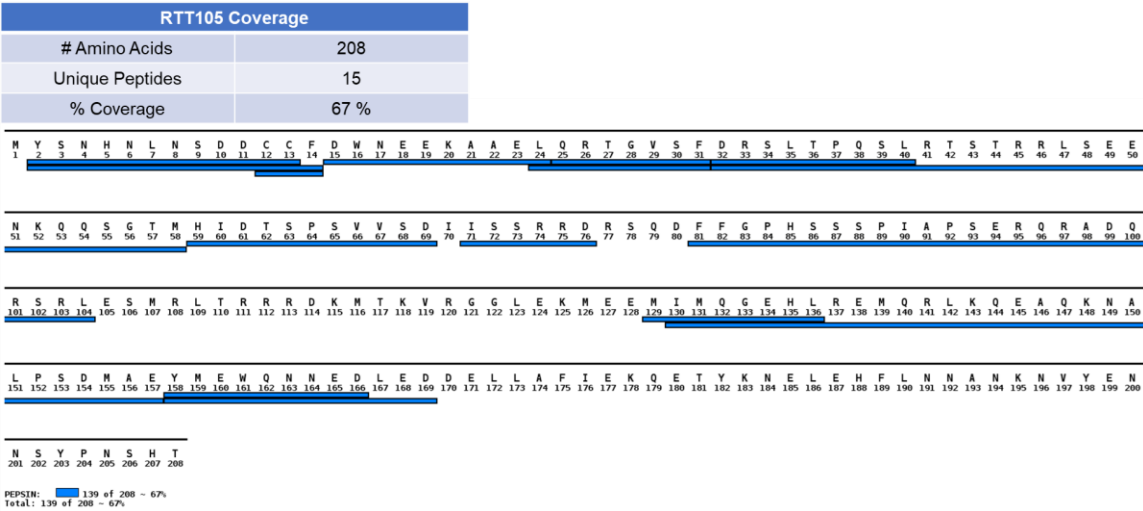
Supplementary Figure 1: Size exclusion chromatography (SEC) analysis of complex formation between RPA and Rtt105. a) SEC analysis of the F-A-B version of RPA (containing the OB-F, DBD-A and DBD-B domains) shows interaction with Rtt105 as observed by formation of the higher molecular weight species. b) A Val106 to Ala substitution in the OB-F domain of RPA70 (RPAV106A) was predicted to not interact with Rtt105, but we observe complex formation in SEC suggesting interaction between RPAV106A and Rtt105. c) SEC analysis of 5 μ M each of RPA-DBD-ACy3, Rtt105Cy5, or a complex of RPACy3 (5 μ M) and Rtt105Cy5 (5 μ M) in presence of 5 μ M 5'-FAM-(dT)₃₅ ssDNA. Rtt105, RPA, and the Rtt105-RPA complex migrate as single species in SEC. When ssDNA is added to this complex, a larger Rtt105-RPA-(dT)₃₅ complex is observed suggesting that Rtt105 remains in complex with RPA in the presence of short ssDNA oligonucleotides. d) Fractions from the SEC analysis in panel c were analyzed by SDS-PAGE and imaged by Coomassie staining (top) or fluorescence (bottom) using the respective excitation channels. Rtt105 migrates as a co-complex with RPA and the (dT)₃₅ ssDNA substrate. A small fraction of free Rtt105 (vol. 15.5 ml onwards) and the RPA-(dT)₃₅ complex (vol. 11.5 ml) are also observed. Representative images from n=3 independent experiments are shown.



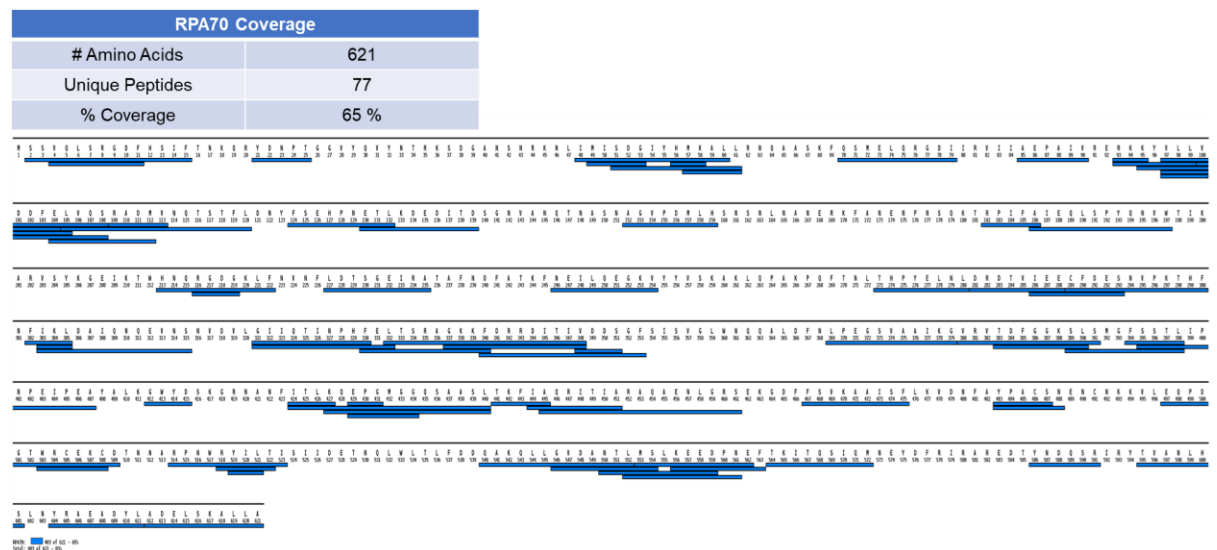
Supplementary Figure 2. Fluorescent Rtt105 interacts with RPA similar to unlabeled Rtt105.a) Positions of Cys12 and Cys13 residues are depicted in the Rtt105 structure generated using AlphaFold. Cysteine residues are represented as ball and stick on the ribbon structure (red) with a transparent surface (grey). b) Circular dichroism analysis of the secondary structure of Rtt105 (black) and the two single Cys variants carrying either C12S (red) or C13S (blue) substitutions show no change in the alpha helical secondary structure of Rtt105 or the mutant variants of Rtt105.c) Analytical size-exclusion chromatography analysis shows similar elution profiles for Rtt105 and the single Cys Rtt105 variants. d) SEC analysis showing complex formation between RPA and Rtt105 or the single-Cys variants of Rtt105. Data show that all three proteins form stable complexes with RPA. e) SDS-PAGE analysis of the Rtt105, Rtt105C13S, Rtt105C13S-Cy5, Rtt105C12S-Cy5, and Rtt105C12S protein samples by Coomassie (left) and fluorescence (on right) imaging. f) Rapid kinetic analysis of the effect of Rtt105 on RPA-DBD-A interactions with ssDNA show both Rtt105 and the single Cys variants similarly influencing the binding properties. For all Rtt105 proteins tested, the rate of binding is not affected, but the amplitudes are reduced. Cy3 placed on the (dT)40 is excited (5'-Cy3-(dT)40) and emission of Cy5 positioned on DBD-A (RPA-DBD-ACy5) via FRET were measured. Since DBD-A resides close to the 5' end of the DNA, a high FRET signal is observed. g) Similar rapid kinetic experiments were used to quantitate the influence of the single Cys variants on FRET induced kinetics of Cy5 change between a 3'-Cy3-(dT)40 substrate and RPA-DBD-D^{Cy5}. In this experiment, DBD-D is situated close to the 3' end of the DNA and generates a high FRET signal. Again, no differences are observed between Rtt105 and the single- Cys variants. The rates of DNA binding are reduced, but the amplitudes are similar. Thus, substitution of either Cys residue in Rtt105 does not affect RPA binding or RPA-DNA interactions. Cys at position 12 was labeled with Cy5 to generated fluorescently labeled Rtt105. Representative images from n=3 independent CD, SEC, and stopped flow experiments are shown.



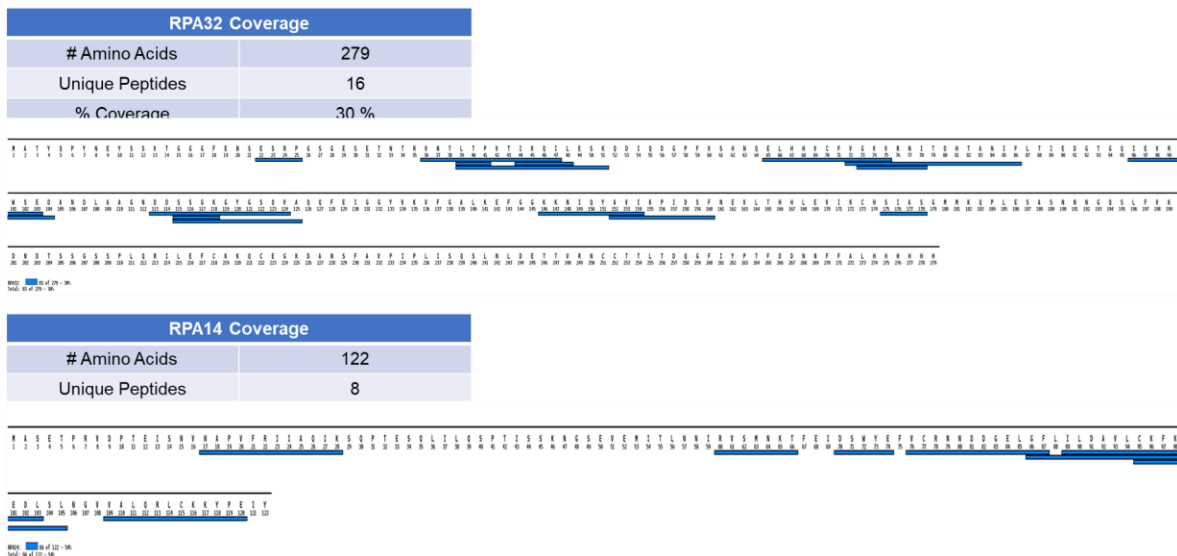
Supplementary Figure 3: Rtt105 binds poorly to a preformed RPA-ssDNA complex. Binding of fluorescent Rtt105 (100 nM Rtt105Cy5) to ssDNA-bound fluorescent RPA (100 nM (dT)₃₅ with either a) 100 nM RPA-DBD-ACy3 or b) 100 nM RPA-DBD-DCy3 was measured using stopped flow analysis. FRET between Rtt105 and respectively labeled DBDs were monitored by exciting Cy5 and monitoring the change in Cy5 fluorescence. In both cases, only a very small change in fluorescence is observed suggesting that Rtt105 binding to a preformed RPA-ssDNA complex is poor. c) Fitting the signals to a single step binding model reveals that Rtt105 binding to DBD-D in the absence or presence of ssDNA is similar (3.5 s⁻¹ versus 2.5 s⁻¹). In contrast, the binding induced FRET signal change in DBD-A is ~25-fold faster. This suggests that Rtt105 engages closer to DBD-A faster when RPA is prebound to ssDNA. Since the F-A-B region of RPA is more dynamic and predominantly more accessible when bound to DNA it likely enables better access to OB-F and thus the increase in binding kinetics. In both panels a and b, no changes in the FRET induced Cy5 fluorescence are observed upon addition of buffer (black). A representative image from three independent stopped flow experiments is shown and the data were fit to a single exponential equation. Mean values +/- SE from fits of n=3 independent experiments are noted.



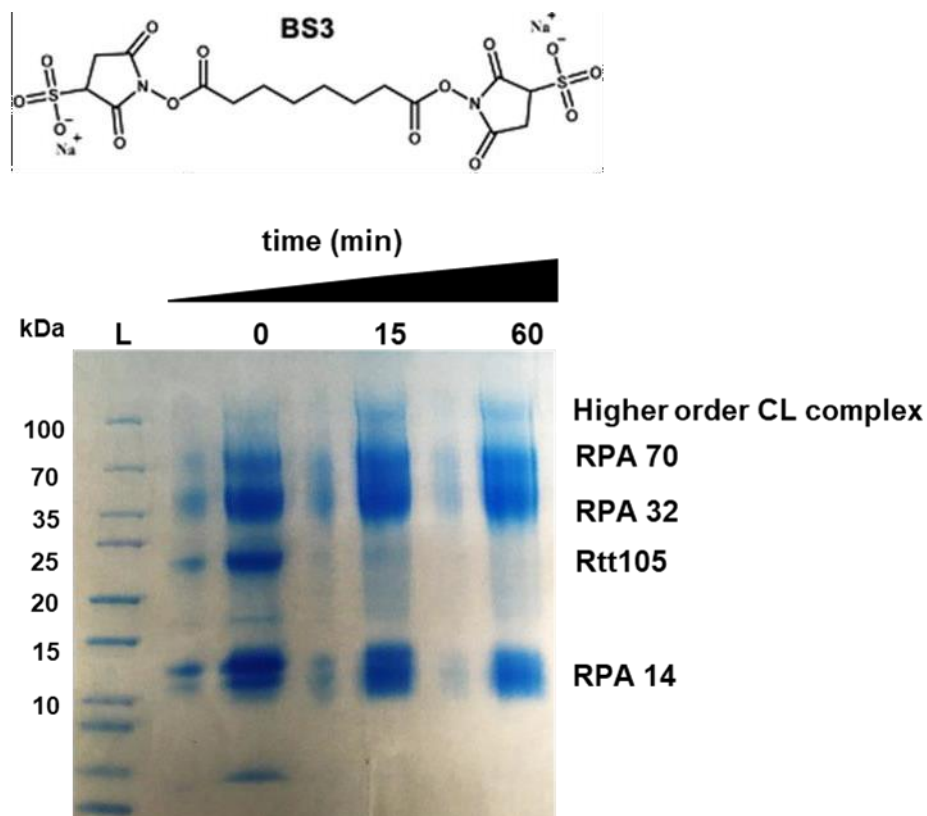
Supplementary Figure 4a. Peptide coverage in mass spectrometry analysis. Good sequence coverage is observed for Rtt105 upon peptide digestion during mass spectrometry analysis. The total number of amino acids, unique peptides, and percentage coverage for the individual polypeptides are denoted.



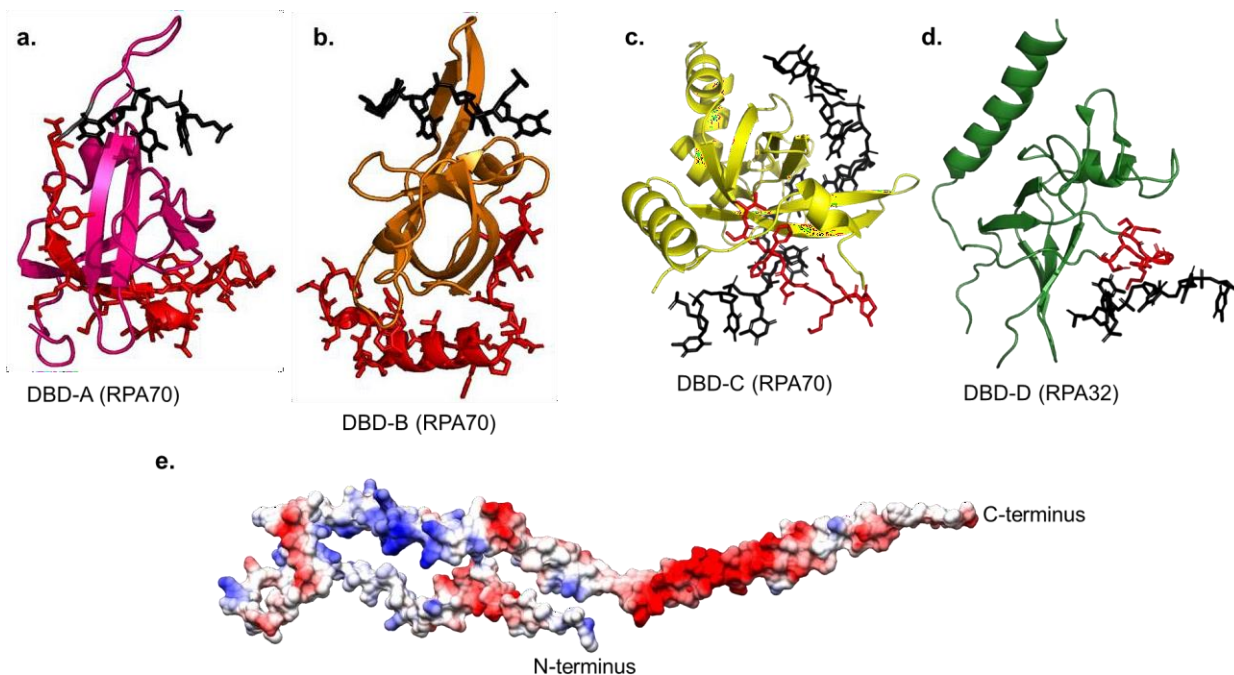
Supplementary Figure 4b. Peptide coverage in mass spectrometry analysis. Good sequence coverage is observed for RPA70 upon peptide digestion during mass spectrometry analysis. The total number of amino acids, unique peptides, and percentage coverage for the individual polypeptides are denoted.



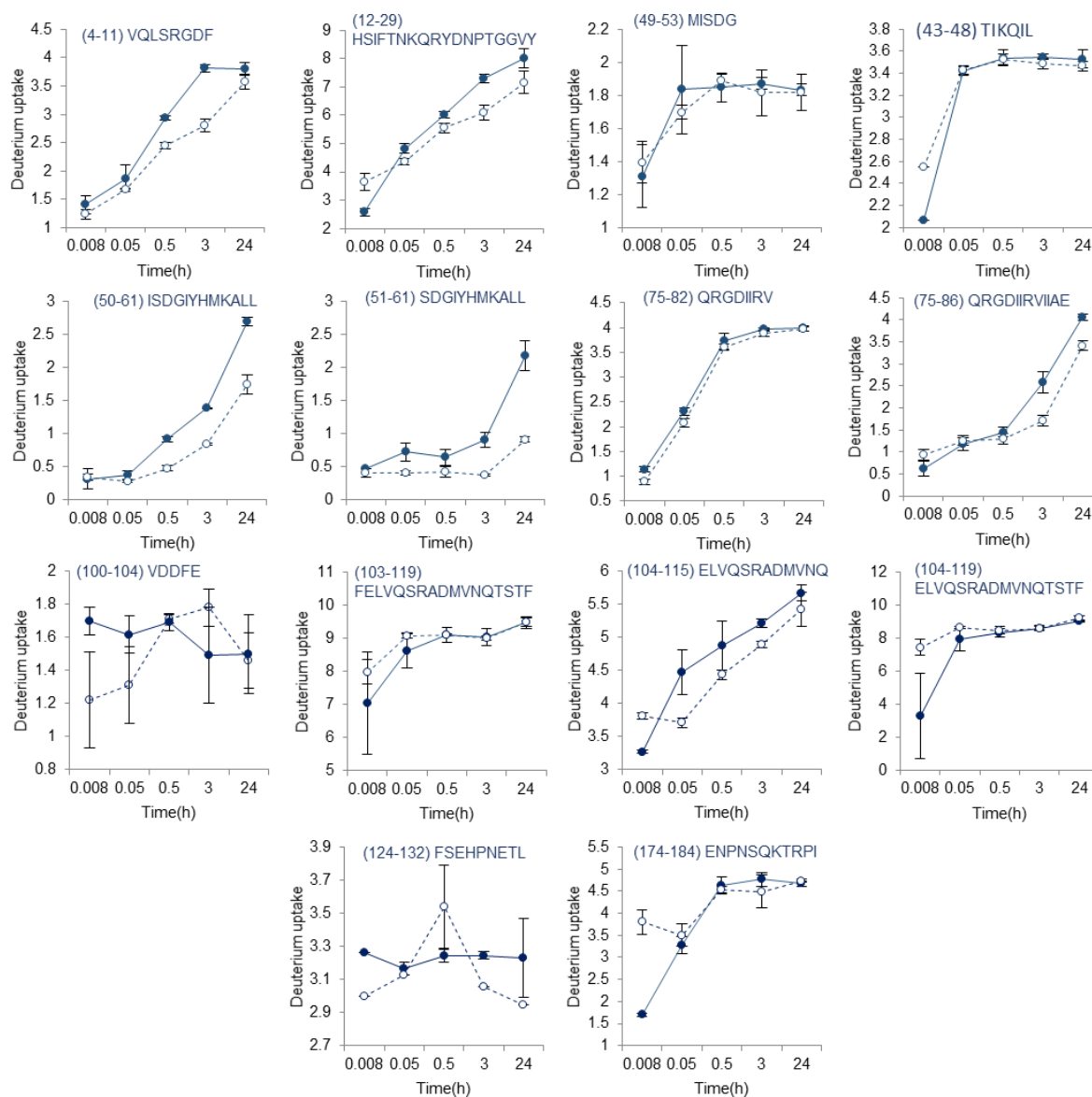
Supplementary Figure 4c. Peptide coverage in mass spectrometry analysis. Good sequence coverage is observed for RPA32 and RPA14 subunits upon peptide digestion during mass spectrometry analysis. The total number of amino acids, unique peptides, and percentage coverage for the individual polypeptides are denoted.



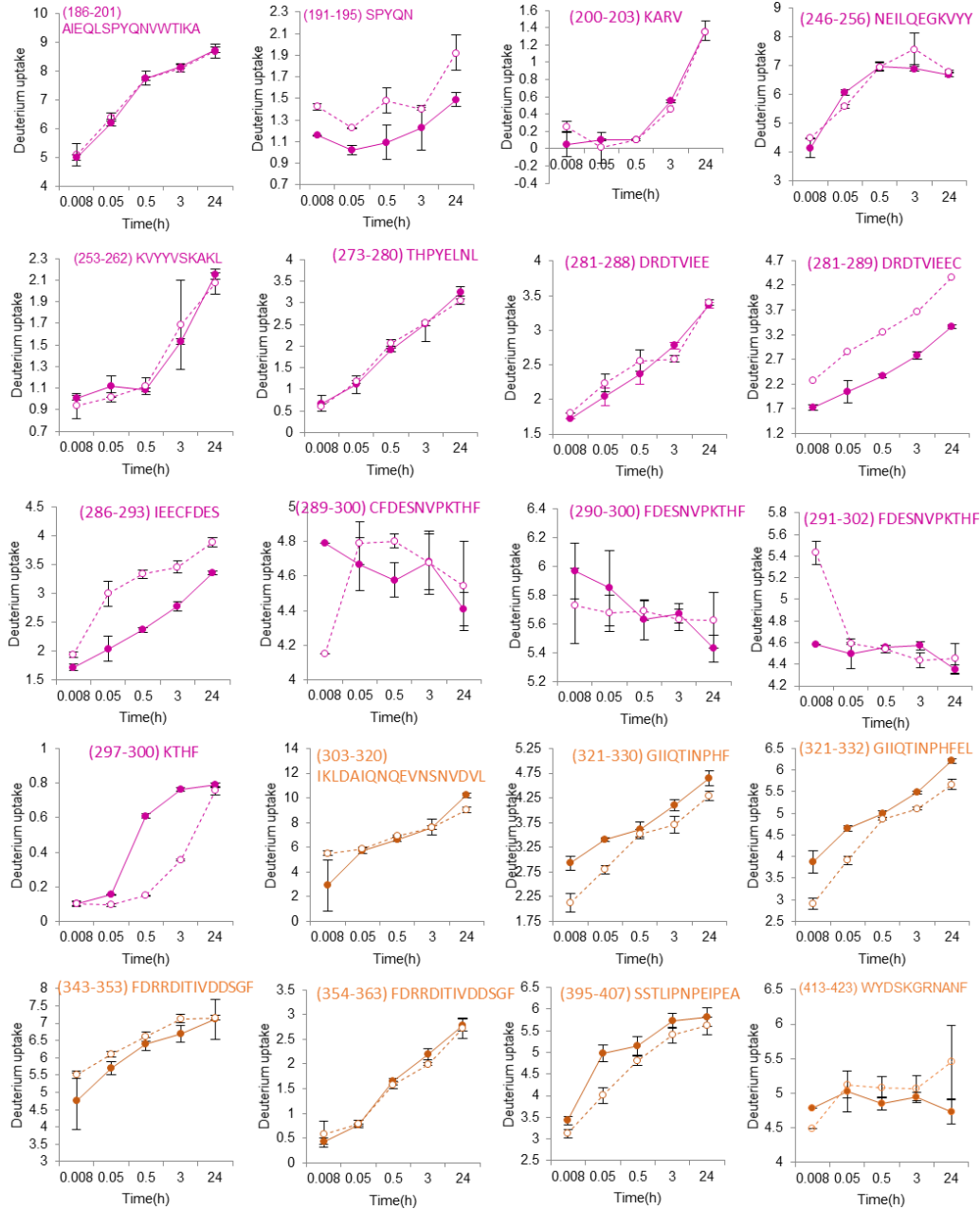
Supplementary Figure 5. Crosslinking analysis of the RPA-Rtt105 complex. Structure of the bis(sulphosuccinimidyl)suberate (BS3) crosslinker is shown (top). 7.75 μ M RPA and 15.5 μ M Rtt105 were mixed and treated with 5 mM BS3 for increasing times (0, 15 and 60 min). Samples at each time point were analyzed by SDS-PAGE and show a time-dependent increase in the crosslinking. Crosslinking of Rtt105 to RPA70 and RPA32 is readily observed with a complete shift in the Rtt105 band. The RPA14 subunit is not readily crosslinked and correspondingly, only a single crosslink between Rtt105 and RPA14 is detected in the XL-MS analysis.



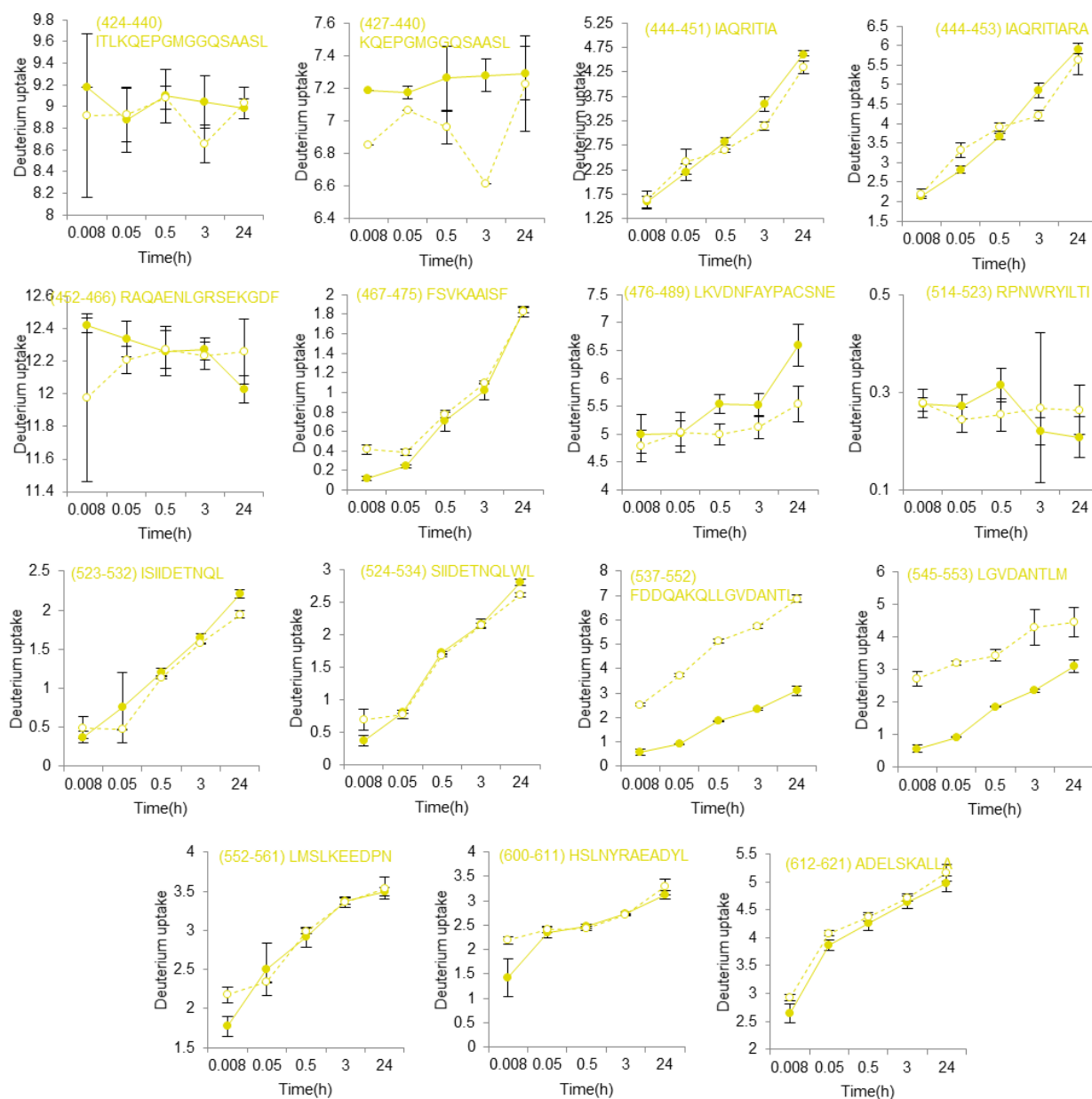
Supplementary Figure 6. Overlap between DNA binding and Rtt105 binding regions in DBD-C and DBD-D revealed by crosslinking analysis. The peptides identified in XL-MS are denoted as red sticks for a) DBD-A, b) DBD-B, c) DBD-C and d) DBD-D along with the bound DNA from the respective structures (black sticks). The crosslinked regions do not (or minimally) overlap with the DNA binding pockets in DBD-A or DBD-B, but significantly overlap in DBD-C and DBD-D. DBD-C and DBD-D are from the cryoEM structure of *S. cerevisiae* RPA (PDB ID: 6I52). The ssDNA-bound DBD-A and DBD-B structures were generated using SWISS-MODEL based on the DNA bound structure of human RPA (PDB ID: 1JMC). e) The surface charges of Rtt105 are displayed in the AlphaFold generated model (AF-P40063-F1). Multiple pockets of negative charges are noted and can potentially bind close to the DNA binding regions in RPA.



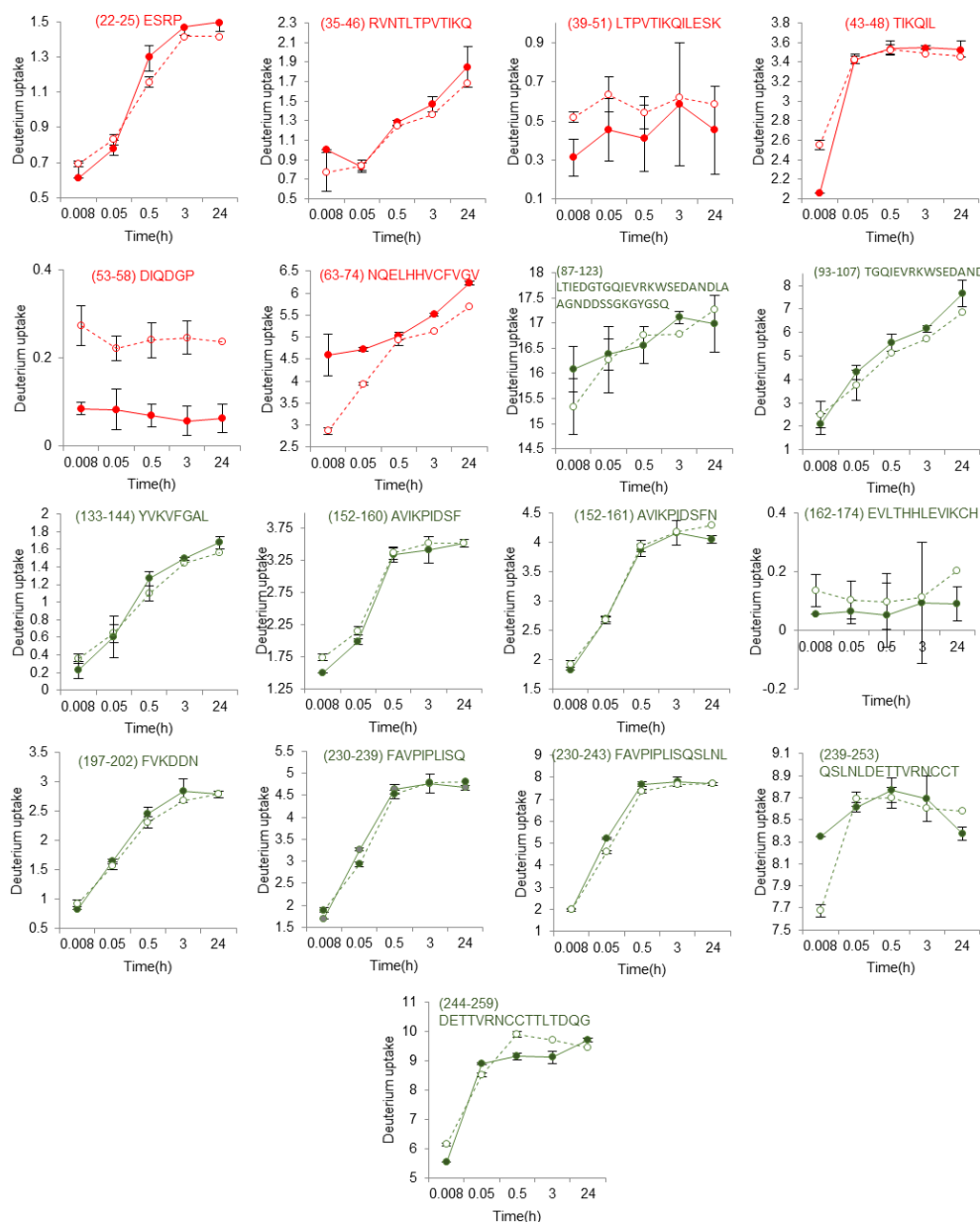
Supplementary Figure 7. HDX-MS analysis of RPA-Rtt105 interactions and OB-F (RPA70) peptides. HDX-MS data corresponding to specific peptides from PID70N (OB-F domain) are shown. Deuterium uptake for each peptide collected in the absence (●) and presence (○) of Rtt105 are plotted as a function of time. The amino acid residue number and sequence of each corresponding peptides are noted. Mean values $\pm$ SD from n=3 biologically independent experiments are shown.



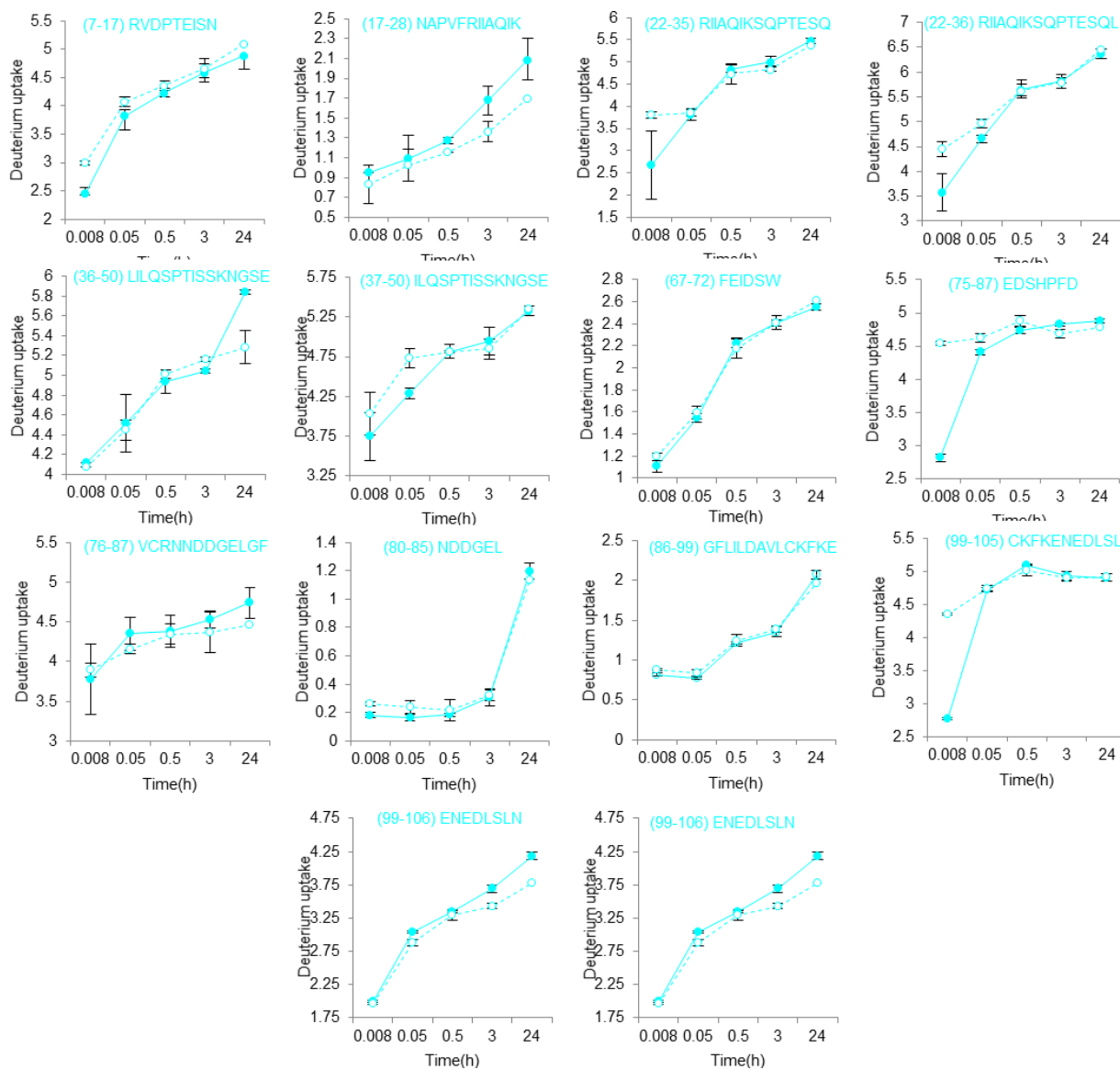
Supplementary Figure 8. HDX-MS analysis of RPA-Rtt105 interactions and DBD-A & DBD-B (RPA70) peptides. HDX-MS data corresponding to specific peptides from DBD-A (purple) and DBD-B (orange) are shown. Deuterium uptake for each peptide collected in the absence (●) and presence (○) of Rtt105 are plotted as a function of time. The amino acid residue number and sequence of each corresponding peptides are noted. Mean values $\pm$ SD from $n=3$ biologically independent experiments are shown.



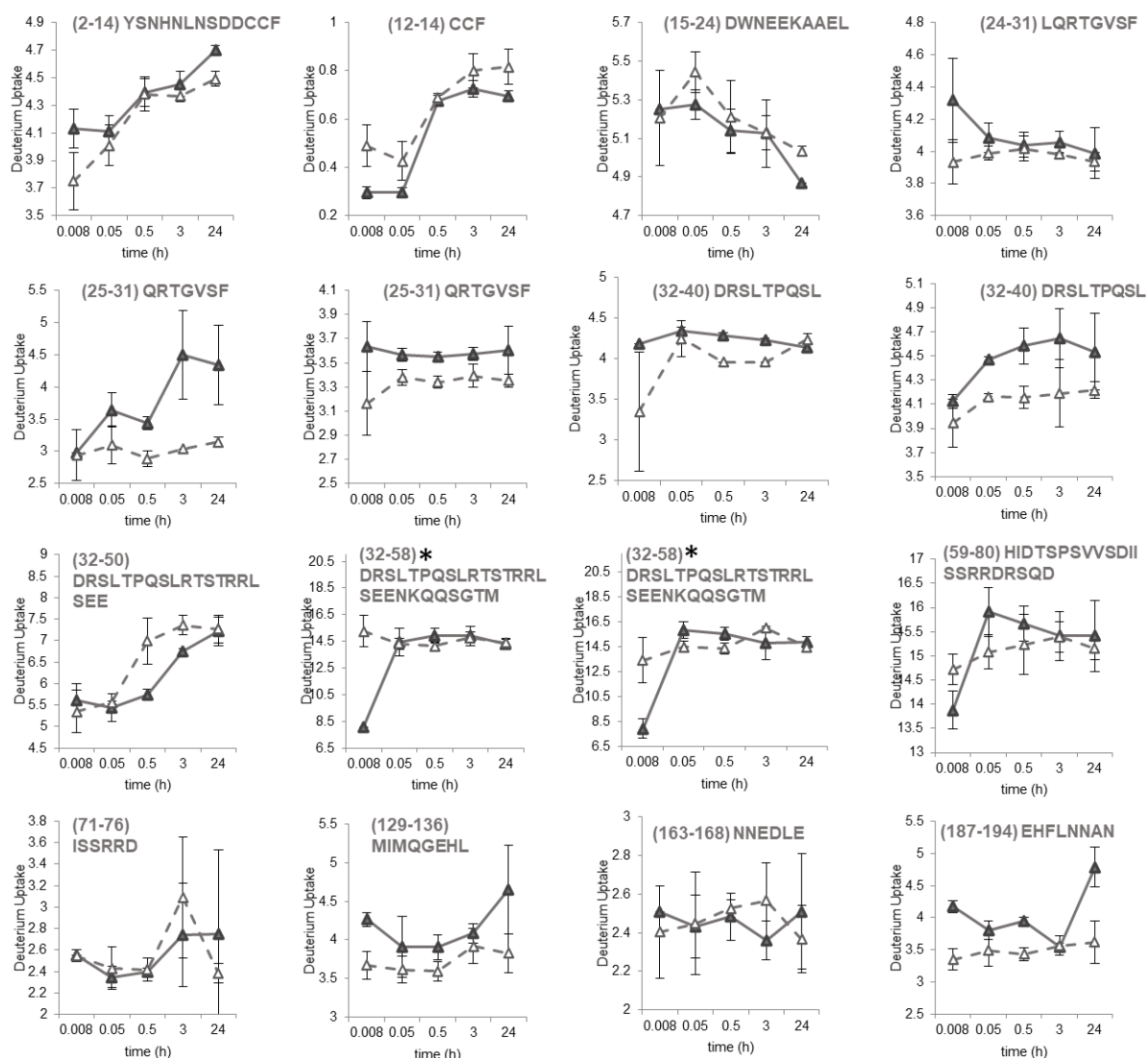
Supplementary Figure 9. HDX-MS analysis of RPA-Rtt105 interactions and DBD-C (RPA70) peptides. HDX-MS data corresponding to specific peptides from DBD-C are shown. Deuterium uptake for each peptide collected in the absence (●) and presence (○) of Rtt105 are plotted as a function of time. The amino acid residue number and sequence of each corresponding peptides are noted. Mean values $\pm$ SD from $n=3$ biologically independent experiments are shown.



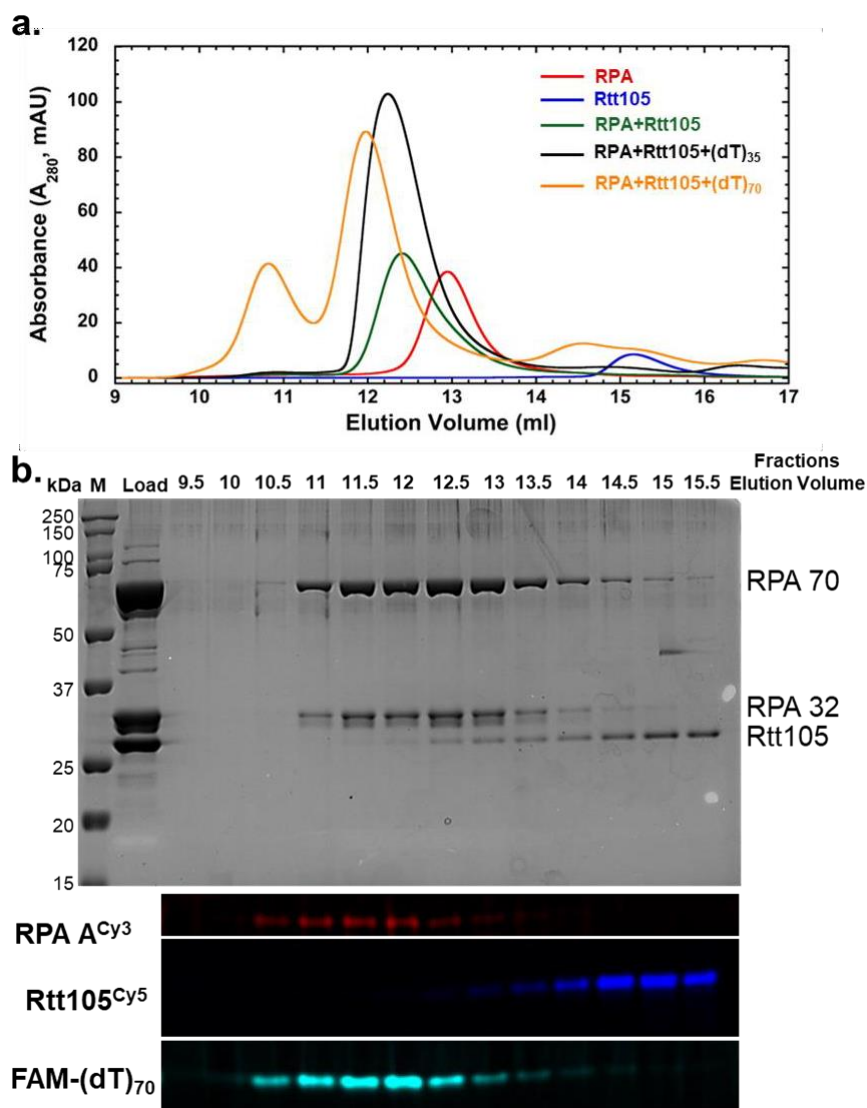
Supplementary Figure 10. HDX-MS analysis of RPA-Rtt105 interactions and RPA32 (DBD-D and winged helix or PID32C) peptides. HDX-MS data corresponding to specific peptides from RPA32 are shown. The peptides in the N-terminal region of hyperphosphorylation are shown in red and DBD-D peptides are in green. Deuterium uptake for each peptide collected in the absence (●) and presence (○) of Rtt105 are plotted as a function of time. The amino acid residue number and sequence of each corresponding peptides are noted. Mean values $\pm$ SD from $n=3$ biologically independent experiments are shown.



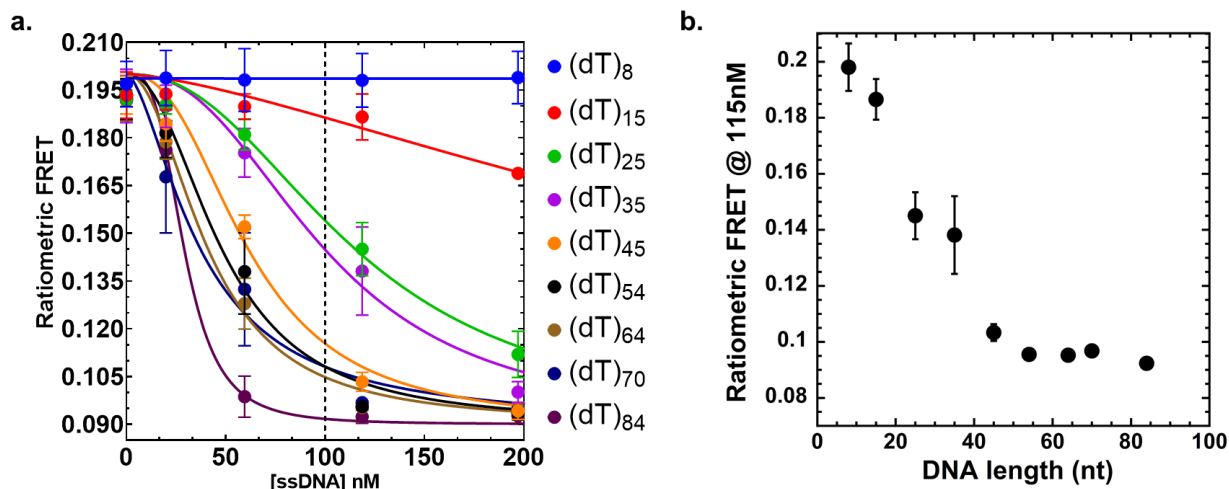
Supplementary Figure 11. HDX-MS analysis of RPA-Rtt105 interactions and RPA14 (OB-E) peptides. HDX-MS data corresponding to specific peptides from RPA14 are shown. Deuterium uptake for each peptide collected in the absence (●) and presence (○) of Rtt105 are plotted as a function of time. The amino acid residue number and sequence of each corresponding peptides are noted. Mean and SD values from n=3 biologically independent experiments are shown.



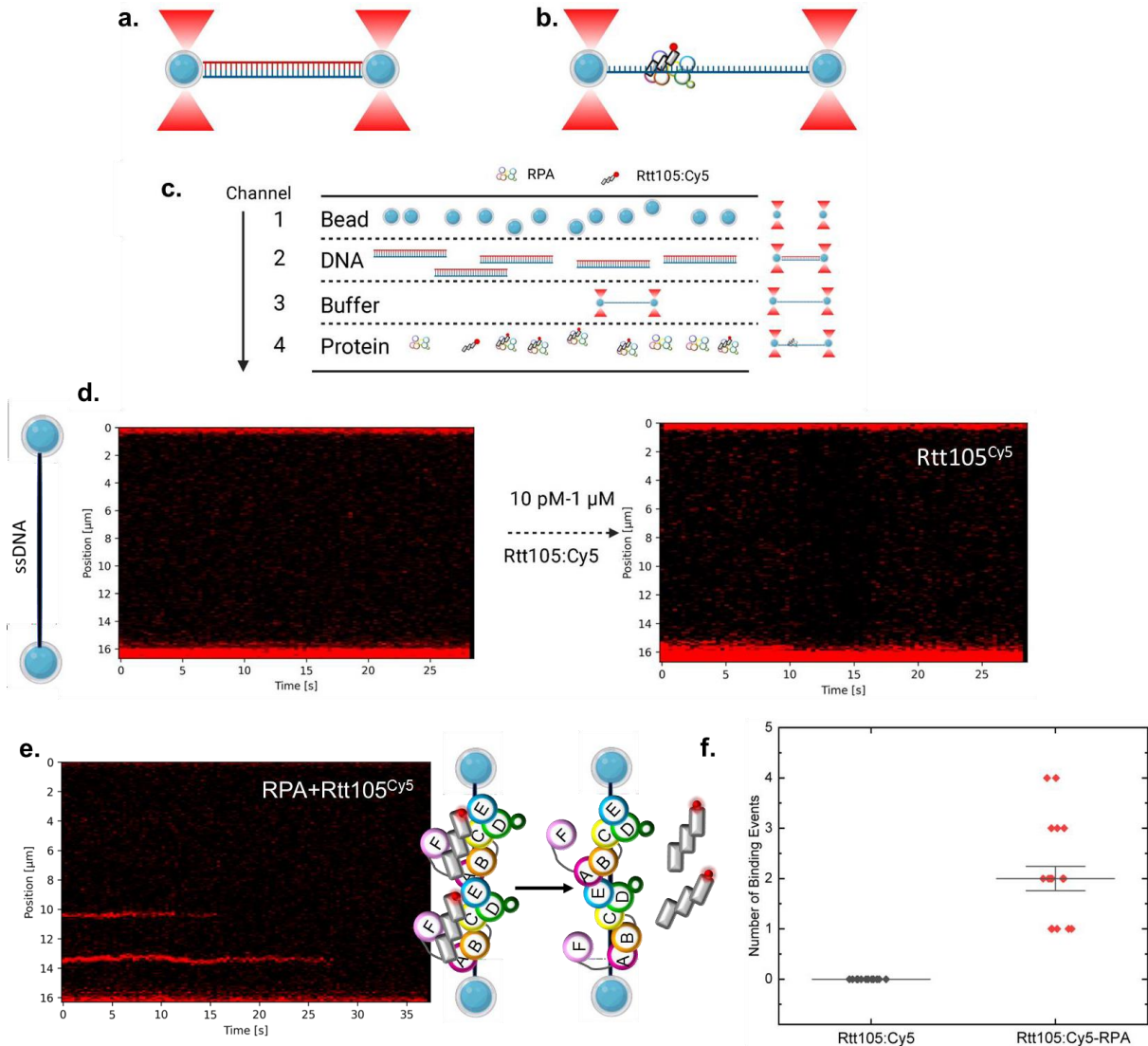
Supplementary Figure 12. HDX-MS analysis of Rtt105 peptides from the Rtt105-RPA- ssDNA complex reveals DNA-induced remodeling of Rtt105. HDX-MS data corresponding to specific peptides from Rtt105 are shown for samples measured in the absence ((▲) Rtt105-RPA) or presence of DNA ((△) Rtt105-RPA-(dT)₃₅). The amino acid residue number and sequence of each corresponding peptides are noted. Mean values +/- SD from n=3 biologically independent experiments are shown. *Denotes similar peptides carrying different charge.



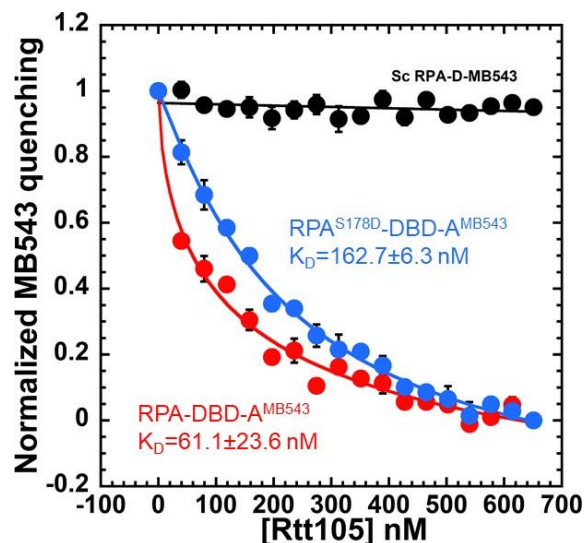
Supplementary Figure 13. SEC analysis of Rtt105-RPA-(dT)70 complex shows dissociation of Rtt105. a) SEC analysis of 5 μ M each of RPA-DBD-ACy3, Rtt105Cy5, or preformed complex RPACy3 (5 μ M) +Rtt105Cy5 (5 μ M) in presence of 5 μ M 5'-FAM-(dT)35 or 5 μ M 5'-FAM-(dT)70 ssDNA oligonucleotides. b) The fractions from the Rtt105-RPA-(dT)70 run in panel a were analyzed by SDS-PAGE and imaged by Coomassie staining (top) or fluorescence (bottom) using the respective excitation channels. Rtt105 dissociates from RPA in the presence of (dT)70 ssDNA substrate. Representative images from n=3 independent experiments are shown.



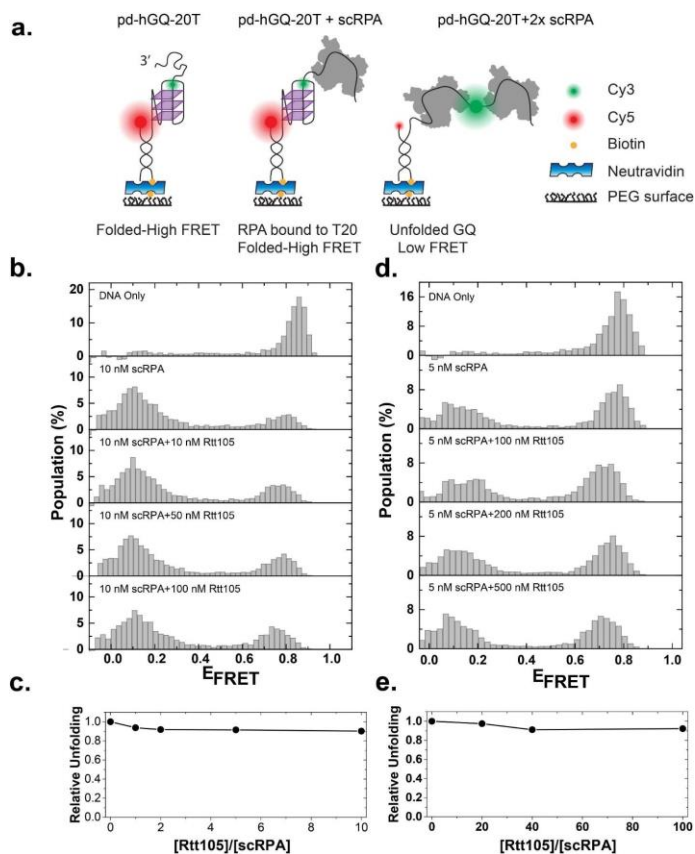
Supplementary Figure 14: ssDNA substrates ≥ 35 nt in length trigger the release of Rtt105 from RPA. a) Preformed RPA-DBD-DCy3-Rtt105Cy5 complexes were mixed with increasing concentrations of (dT)_n (where n = length of ssDNA) and the change in FRET was measured. b) Respective FRET values observed at 115 nM ssDNA concentration ($\sim 2:1$ ratio of RPA:(dT)_n) were plotted as a function of (dT)_n. Increasing the length beyond (dT)₃₅ leads to a very low-FRET state which corresponds to Rtt105 remodeling or dissociation. The moderate FRET state (~ 0.14 FRET) observed around $n=35$ corresponds to a complex where Rtt105 is remodeled, but still bound to the RPA-DNA complex. Mean values $\pm$ SE from $n=3$ biologically independent experiments are shown.



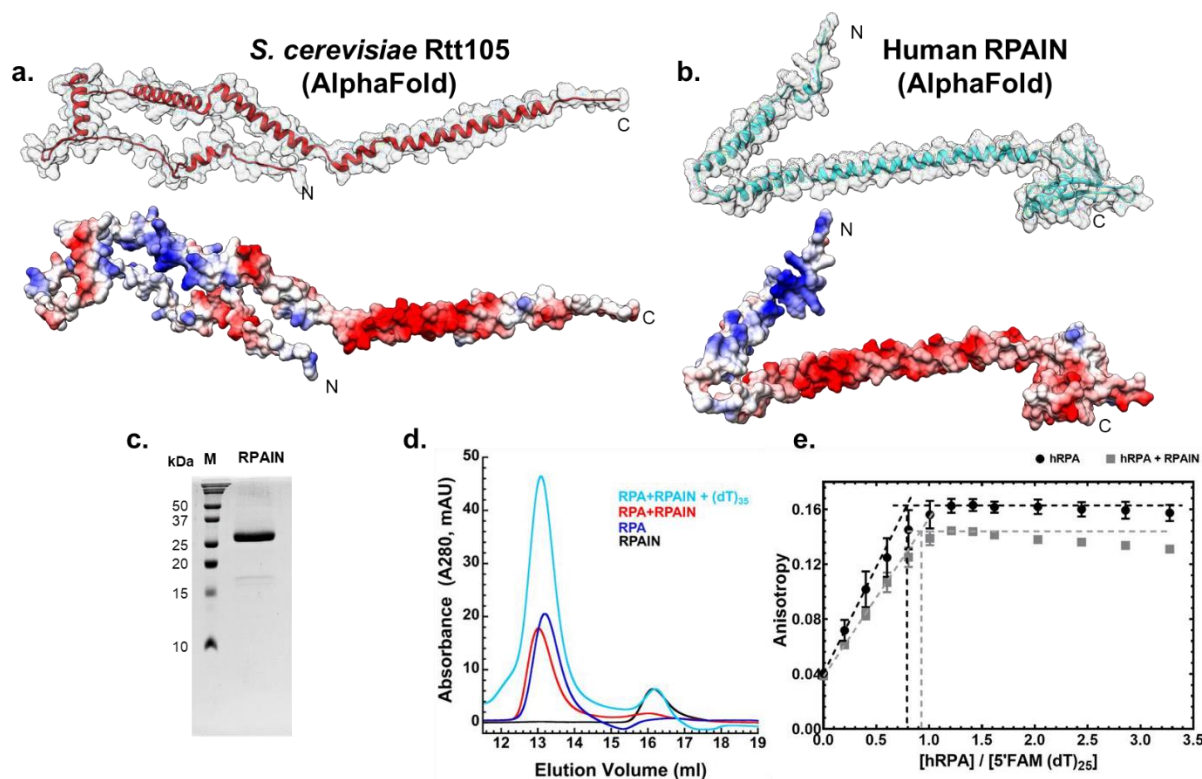
Supplementary Figure 15. C-trap analysis shows Rtt105 poorly bound to RPA on long ssDNA. a) Schematic of the C-trap setup to investigate RPA-Rtt105 interactions. Lambda dsDNA is captured between the beads and b) stretched to generate ssDNA. c) The ssDNA is introduced into channels containing the appropriate proteins. d) Analysis of Rtt105Cy5-RPA interactions on ~48.5 knt long DNA substrates reveal no Rtt105 binding to ssDNA in the absence of RPA. e) When the fluorescence is monitored upon exposure of the ssDNA to the Rtt105Cy5-RPA, very few Rtt105Cy5 binding events are observed, and these molecules also dissociate. f) Quantitation of the number of DNA binding events in C-trap analysis shows no Rtt105 binding to ssDNA in the absence of RPA. Roughly 2 transient binding events per RPA-coated 48.5 knt long ssDNA are observed under these experimental conditions. Data show a $t_{1/2}$ of ~5 s for the dissociation of Rtt105Cy5 from RPA on these DNA molecules. Mean values $\pm$ SE from $n > 10$ biologically independent experiments (ssDNA molecules captured and analyzed) are shown.



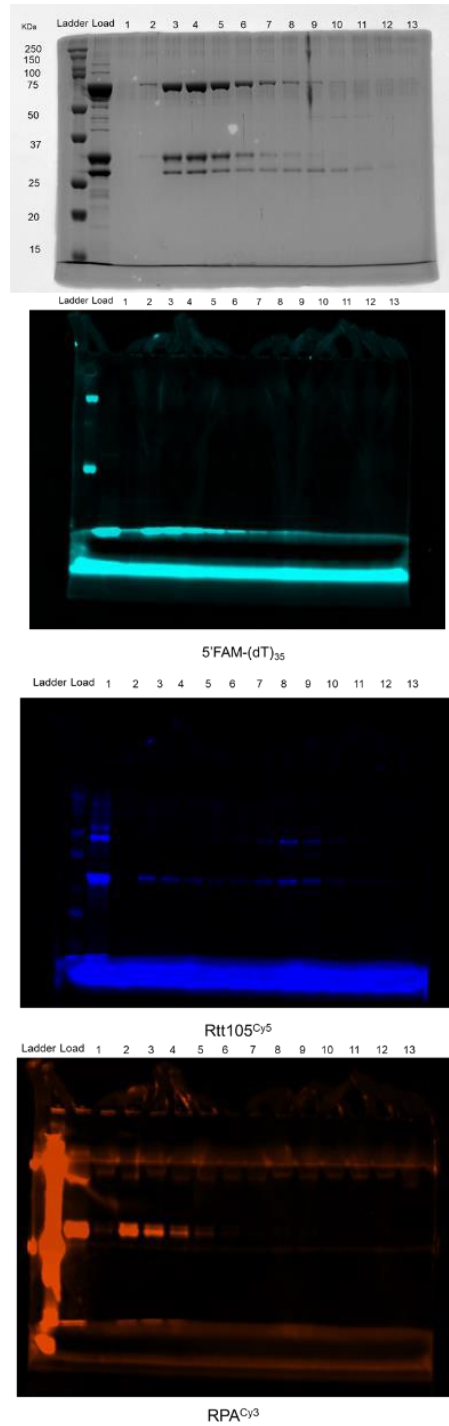
Supplementary Figure 16. Phosphorylation of RPA at S178 reduces Rtt105 binding affinity. Rtt105 binding to RPA or a phosphomimetic RPAS178D (in RPA70) was measured by following the change in fluorescence of a MB543 fluorophore positioned in DBD-A. The fluorophore, when positioned in DBD-D, does not produce a change in fluorescence upon Rtt105 binding (black). RPAS178D binds with reduced affinity to Rtt105 ($K_D = 61.1 \pm 23.6$ nM versus 162.7 ± 6.3 nM for RPA and RPAS178D, respectively). Mean values $\pm$ SE from $n=3$ biologically independent experiments are shown.



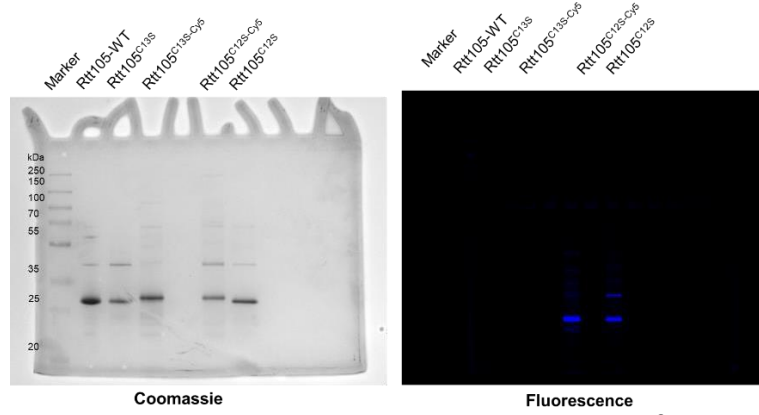
Supplementary Figure 17. Rtt105 does not influence the G-quadruplex unwinding activity of RPA. **a)** Schematics of the smFRET assay. A partial duplex DNA construct with a hGQ forming sequence and 20T overhang is immobilized on the surface of a microfluidic channel. The donor-acceptor fluorophores are placed across the hGQ. scRPA binding to the 20T overhang does not result in a FRET change until the hGQ is unfolded. Unfolding of the hGQ exposes an additional 21 nt ssDNA, which is long enough to accommodate binding of a second scRPA. Binding of scRPA to this newly exposed ssDNA results in emergence of a new low FRET peak. **b)** Top panel shows the smFRET histogram in the absence of any proteins (DNA Only) where a single high FRET peak is observed. Adding 10 nM scRPA results in ~70% of hGQ molecules being unfolded. Adding mixtures of scRPA and increasing concentrations of Rtt105 does not result in a significant change in the unfolded population. **c)** Quantification of the unfolded population for different $[Rtt105]/[scRPA]$ ratios. The unfolded population was calculated from the cumulative population of $E_{FRET} < 0.5$ states and normalized such that the 10 nM scRPA (without Rtt105) has a relative unfolding of 1.0. Before adding a new scRPA and Rtt105 mixture, the channel was washed with a high salt buffer (1 M KCl) to remove all bound proteins. The reset of the system was confirmed by restoring the histogram with a single high FRET peak after adding imaging buffer at original salt concentration (50 mM KCl+2 mM MgCl₂). **d)** and **e)** Similar measurements and analysis as those performed for **(b)** and **(c)** when scRPA concentration was kept at 5 nM and higher $[Rtt105]/[scRPA]$ ratios were reached. Even at $[Rtt105]/[scRPA]=100$, the unfolded population was within 10% of its original value.



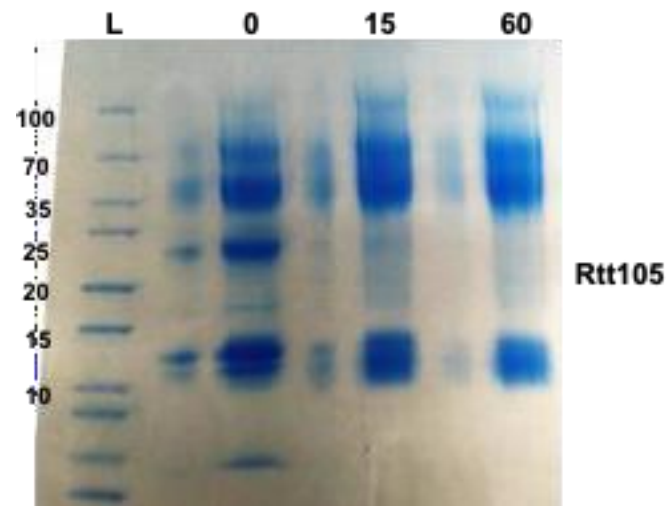
Supplementary Figure 18. Rtt105 does not enhance the ssDNA binding activity of RPAIN, the human ortholog of Rtt105. AlphaFold generated models of a) ScRtt105 (AF-P40063-F1) and b) human RPAIN (AF-Q86UA6-F1). The electrostatic maps of both structures are displayed as well and show regions of strong negative charge. c) SDS-PAGE analysis of recombinantly purified RPAIN. d) SEC analysis of 3 μ M hRPA (blue) and 3 μ M RPAIN (black) show singly migrating species of each protein. The RPA-RPAIN complex generated by mixing 3 μ M of each protein (red) migrates as a larger complex and shows interaction between the two proteins. Addition of 3 μ M (dT)₃₅ to the hRPA-RPAIN complex results in the release of RPAIN from RPA (cyan). e) The ssDNA binding activity of hRPA was measured by fluorescence anisotropy in the absence or presence of 2-fold molar excess of RPAIN. Experiments were performed using 20 nM 5'-FAM-(dT)₂₅ ssDNA and titrating increasing concentrations of hRPA or the hRPA-RPAIN (1:2 molar ratio) complex. Data shows an increase in anisotropy upon ssDNA binding by hRPA. Due to the high affinity interactions, stoichiometric binding is observed in the absence (●) or presence (■) of RPAIN. Interestingly, the maximum anisotropy values for the RPA-RPAIN complex are lower suggesting that the configuration of the hRPA-RPAIN-ssDNA complex is likely different than the hRPA-ssDNA complex. Mean values $\pm$ SE are plotted from n=4 biologically independent experiments.



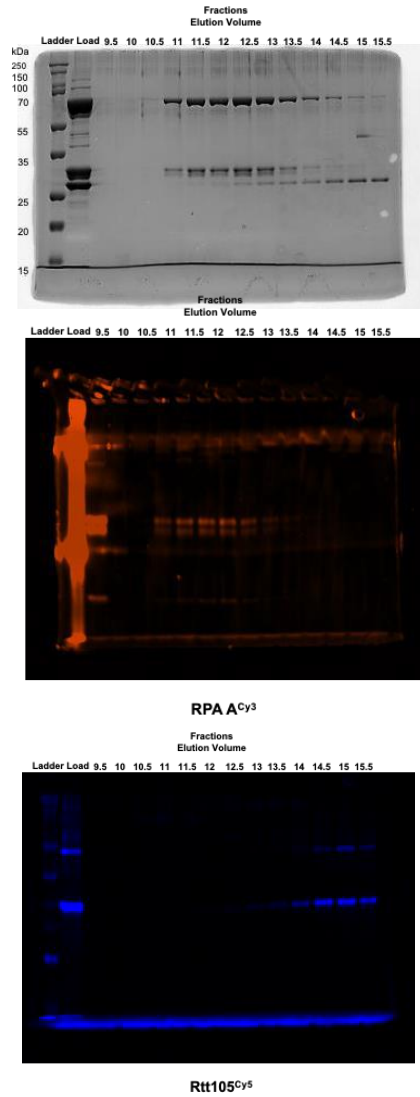
Supplemental size exclusion chromatography (SEC) analysis of Rtt105, RPA, and Rtt105-RPA and (dT)35



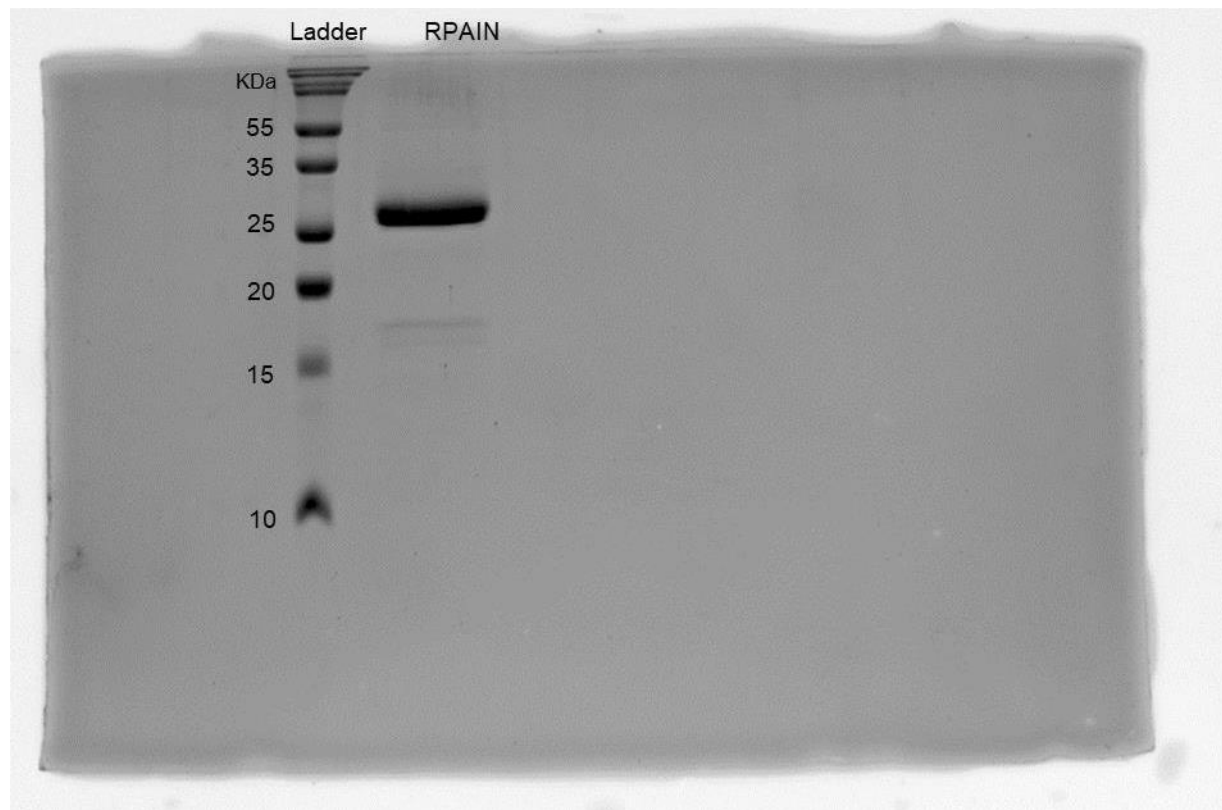
Supplemental size exclusion chromatography (SEC) analysis of Rtt105, RPA, and Rtt105-RPA and (dT)35Uncropped images for Supplementary Figure 2e SDS-PAGE analysis of Rtt105-Cys variants.



Supplemental uncropped image for Supplementary Figure 5 SDS-PAGE analysis of Crosslinking with BS3.



Supplemental. Uncropped images for Supplementary Figure 13b Size exclusion chromatography (SEC) analysis of Rtt105, RPA, and Rtt105-RPA and (dT)70



Supplemental. Uncropped image for Supplementary Figure 18c

REFERENCES CITED

- Andriuskevicius, T., Kotenko, O. & Makovets, S. Putting together and taking apart: assembly and disassembly of the Rad51 nucleoprotein filament in DNA repair and genome stability. *Cell Stress* 2, 96-112, doi:10.15698/cst2018.05.134 (2018).
- Anfinsen. Folding of Protein Chains. *Science* (80-). 1973;181(4096):223-230.
- Antony, E. et al. Srs2 disassembles Rad51 filaments by a protein-protein interaction triggering ATP turnover and dissociation of Rad51 from DNA. *Mol Cell* 35, 105-115, doi:10.1016/j.molcel.2009.05.026 (2009).
- Arai, N. et al. Vital roles of the second DNA-binding site of Rad52 protein in yeast homologous recombination. *J Biol Chem* 286, 17607-17617, doi:10.1074/jbc.M110.216739 (2011).
- Bartsch, S., Kang, L. E. & Symington, L. S. RAD51 is required for the repair of plasmid double-stranded DNA gaps from either plasmid or chromosomal templates. *Mol Cell Biol* 20, 1194-1205, doi:10.1128/MCB.20.4.1194-1205.2000 (2000).
- Ben-Nissan G, Sharon M. Capturing protein structural kinetics by mass spectrometry. *Chem Soc Rev.* 2011;40(7):3627-3637. doi:10.1039/c1cs15052a
- Berry, L., Patterson, A., Pence, N., Peters, J. W. & Bothner, B. Hydrogen Deuterium Exchange Mass Spectrometry of Oxygen Sensitive Proteins. *Bio Protoc* 8, doi:10.21769/BioProtoc.2769 (2018).
- Boehr DD, Nussinov R, Wright PE. The role of dynamic conformational ensembles in biomolecular recognition. *Nat Chem Biol.* 2009;5(11):789-796. doi:10.1038/nchembio.232
- Bothner B, Johnson JE, Schlutz CP, et al. Maturation of a tetravirus capsid alters the dynamic properties and creates a metastable complex. *Virology.* 2005;334(1):17-27. doi:10.1016/j.virol.2005.01.017
- Boyd JM, Ellsworth H, Ensign SA. Bacterial acetone carboxylase is a manganese-dependent metalloenzyme. *J Biol Chem.* 2004. doi:10.1074/jbc.M407177200
- Boyd JM, Ensign SA. ATP-dependent enolization of acetone by acetone carboxylase from *Rhodobacter capsulatus*. *Biochemistry.* 2005. doi:10.1021/bi050393k
- Bronson, J. E., Fei, J., Hofman, J. M., Gonzalez, R. L., Jr. & Wiggins, C. H. Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data. *Biophys J* 97, 3196-3205, doi:10.1016/j.bpj.2009.09.031 (2009).
- Bryngelson JD, Wolynes PG. Spin glasses and the statistical mechanics of protein folding. *Proc Natl Acad Sci U S A.* 1987;84(21):7524-7528. doi:10.1073/pnas.84.21.7524

Burns, E. E., Keith, B. K., Refai, M. Y., Bothner, B. & Dyer, W. E. Proteomic and biochemical assays of glutathione-related proteins in susceptible and multiple herbicide resistant *Avena fatua* L. *Pestic Biochem Physiol* 140, 69-78, doi:10.1016/j.pestbp.2017.06.007 (2017).

Bush MF, Hall Z, Giles K, Hoyes J, Robinson C V., Ruotolo BT. Collision cross sections of proteins and their complexes: A calibration framework and database for gas-phase structural biology. *Anal Chem*. 2010;82(22):9557-9565. doi:10.1021/ac1022953

Byun JA, VanSchouwen B, Akimoto M, Melacini G. Allosteric inhibition explained through conformational ensembles sampling distinct “mixed” states. *Comput Struct Biotechnol J*. 2020;18:3803-3818. doi:10.1016/j.csbj.2020.10.026

Changeux JP. 50Th Anniversary of the Word “Allosteric.” *Protein Sci*. 2011;20(7):1119-1124. doi:10.1002/pro.658

Clark DD, Allen JR, Ensign SA. Characterization of five catalytic activities associated with the NADPH:2-Ketopropyl-coenzyme M [2-(2-ketopropylthio)ethanesulfonate] oxidoreductase/carboxylase of the *Xanthobacter* strain Py2 epoxide carboxylase system. *Biochemistry*. 2000;39(6):1294-1304. doi:10.1021/bi992282p

Cooper A, Dryden DTF. Allostery without conformational change - A plausible model. *Eur Biophys J*. 1984;11(2):103-109. doi:10.1007/BF00276625

Cui Q, Karplus M. Allostery and cooperativity revisited. *Protein Sci*. 2008;17(8):1295-1307. doi:10.1110/ps.03259908

Davies, A. A. et al. Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol Cell* 7, 273-282, doi:10.1016/s1097-2765(01)00175-7 (2001).

Davis, A. P. & Symington, L. S. RAD51-dependent break-induced replication in yeast. *Mol Cell Biol* 24, 2344-2351, doi:10.1128/MCB.24.6.2344-2351.2004 (2004).

Deng, X. et al. Human replication protein A-Rad52-single-stranded DNA complex: stoichiometry and evidence for strand transfer regulation by phosphorylation. *Biochemistry* 48, 6633-6643, doi:10.1021/bi900564k (2009).

Deonarain R, Mahendra P, Berry A, Scrutton NS, Perham. RN. Alternative Proton Donors/ Acceptors in the Catalytic Mechanism of the Glutathione Reductase of *Escherichia coli*: The Role of Histidine-439 and Tyrosine-99. *Biochemistry*. 1989;28:9602-9607.

Deshpande, I. et al. Structural Basis of Mec1-Ddc2-RPA Assembly and Activation on Single-Stranded DNA at Sites of Damage. *Mol Cell* 68, 431-445 e435, doi:10.1016/j.molcel.2017.09.019 (2017).

Dill KA, Maccallum JL, Folding P. The Protein-Folding Problem , 50 Years On. 2012;(NOVEMBER):1042-1046.

Dodds JN, Baker ES. Ion Mobility Spectrometry: Fundamental Concepts, Instrumentation, Applications, and the Road Ahead. *J Am Soc Mass Spectrom*. 2019. doi:10.1007/s13361-019-02288-2

Durall C, Lindblad P. Mechanisms of carbon fixation and engineering for increased carbon fixation in cyanobacteria. *Algal Res*. 2015. doi:10.1016/j.algal.2015.07.002

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66, 486-501, doi:10.1107/S0907444910007493 (2010).

Erb TJ, Zarzycki J. A short history of RubisCO: the rise and fall (?) of Nature's predominant CO₂ fixing enzyme. *Curr Opin Biotechnol*. 2018;49:100-107. doi:10.1016/j.copbio.2017.07.017
Erb TJ. Carboxylases in natural and synthetic microbial pathways. *Appl Environ Microbiol*. 2011. doi:10.1128/AEM.05702-11

Fenton. Aron W. Allostery: an illustrated definition for the 'second secret of life.' *Trends Biochem Sci*. 2008;33(9):420-425. doi:10.1016/j.tibs.2008.05.009.Allostery

Fontana A, Zambonin M, Polverino De Laureto P, De Filippis V, Clementi A, Scaramella E. Probing the conformational state of apomyoglobin by limited proteolysis. *J Mol Biol*. 1997;266(2):223-230. doi:10.1006/jmbi.1996.0787

Frey, Perry. Hegeman A. Enzymatic Reaction Mechanisms. In: Oxford University Press; 2007:418-430.

Gibb, B. et al. Protein dynamics during presynaptic-complex assembly on individual single-stranded DNA molecules. *Nat Struct Mol Biol* 21, 893-900, doi:10.1038/nsmb.2886 (2014).

Giles K, Pringle SD, Worthington KR, Little D, Wildgoose JL, Bateman RH. Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun Mass Spectrom*. 2004;18(20):2401-2414. doi:10.1002/rcm.1641

Giles K, Williams JP, Campuzano I. Enhancements in travelling wave ion mobility resolution. *Rapid Commun Mass Spectrom*. 2011;25(11):1559-1566. doi:10.1002/rcm.5013

Gong, Z., Ye, S. X. & Tang, C. Tightening the Crosslinking Distance Restraints for Better Resolution of Protein Structure and Dynamics. *Structure* 28, 1160-1167 e1163, doi:10.1016/j.str.2020.07.010 (2020).

Goodey NM, Benkovic SJ. Allosteric regulation and catalysis emerge via a common route. *Nat Chem Biol*. 2008;4(8):474-482. doi:10.1038/nchembio.98

Greene, E. C. DNA Sequence Alignment during Homologous Recombination. *J Biol Chem* 291, 11572-11580, doi:10.1074/jbc.R116.724807 (2016).

Greener JG, Sternberg MJ. Structure-based prediction of protein allostery. *Curr Opin Struct Biol.* 2018;50:1-8. doi:10.1016/j.sbi.2017.10.002

Grimme, J. M. et al. Human Rad52 binds and wraps single-stranded DNA and mediates annealing via two hRad52-ssDNA complexes. *Nucleic Acids Res* 38, 2917-2930, doi:10.1093/nar/gkp1249 (2010).

Guarnera E, Berezovsky IN. Structure-Based Statistical Mechanical Model Accounts for the Causality and Energetics of Allosteric Communication. *PLoS Comput Biol.* 2016;12(3):1-27. doi:10.1371/journal.pcbi.1004678

Guarnera E, Tan ZW, Zheng Z, Berezovsky IN. AlloSigMA: Allosteric signaling and mutation analysis server. *Bioinformatics.* 2017;33(24):3996-3998. doi:10.1093/bioinformatics/btx430

Gunasekaran K, Ma B, Nussinov R. Is allostery an intrinsic property of all dynamic proteins? *Proteins Struct Funct Genet.* 2004;57(3):433-443. doi:10.1002/prot.20232

Henzler-Wildman K, Kern D. Dynamic personalities of proteins. *Nature.* 2007;450(7172):964-972. doi:10.1038/nature06522

Heyer, W. D., Ehmsen, K. T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* 44, 113-139, doi:10.1146/annurev-genet-051710-150955 (2010).

Hilser VJ, Freire E. Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. *J Mol Biol.* 1996;262(5):756-772. doi:10.1006/jmbi.1996.0550

Hilser VJ, Wrabl JO, Motlagh HN. Structural and energetic basis of allostery. *Annu Rev Biophys.* 2012;41(1):585-609. doi:10.1146/annurev-biophys-050511-102319

Hsieh, P., Camerini-Otero, C. S. & Camerini-Otero, R. D. The synapsis event in the homologous pairing of DNAs: RecA recognizes and pairs less than one helical repeat of DNA. *Proc Natl Acad Sci U S A* 89, 6492-6496, doi:10.1073/pnas.89.14.6492 (1992).

Hubbard RE, Kamran Haider M. Hydrogen Bonds in Proteins: Role and Strength. *eLS.* 2010;(February). doi:10.1002/9780470015902.a0003011.pub2

Jaffe, E., Bornstein, P. & Disteché, C. M. Mapping of the thrombospondin gene to human chromosome 15 and mouse chromosome 2 by in situ hybridization. *Genomics* 7, 123-126, doi:10.1016/0888-7543(90)90528-3 (1990).

Jamali, K., Kimanius, D., Scheres, S.H.W. A graph neural network approach to automated model building in cryo-EM maps. arXiv, doi:<https://doi.org/10.48550/arXiv.2210.00006> (2022).

Janes W, Schulz GE. The binding of the retro-analogue of glutathione disulfide to glutathione reductase. *J Biol Chem*. 1990;265(18):10443-10445. doi:10.2210/pdb4gr1/pdb

Kagawa, W. et al. Crystal structure of the homologous-pairing domain from the human Rad52 recombinase in the undecameric form. *Mol Cell* 10, 359-371, doi:10.1016/s1097-2765(02)00587-7 (2002).

Kagawa, W. et al. Functional analyses of the C-terminal half of the *Saccharomyces cerevisiae* Rad52 protein. *Nucleic Acids Res* 42, 941-951, doi:10.1093/nar/gkt986 (2014).

Kagawa, W. et al. Identification of a second DNA binding site in the human Rad52 protein. *J Biol Chem* 283, 24264-24273, doi:10.1074/jbc.M802204200 (2008).

Karplus M. The Levinthal paradox: Yesterday and today. *Fold Des*. 1997;2(4):69-75. doi:10.1016/S1359-0278(97)00067-9

Kavan, D. a. M., P. MStools—Web based application for visualization and presentation of HXMS data. *International Journal of Mass Spectrometry* 302, 53-58 (2010).

Kendrew J, Bodo G, Dintzis H, Parrish R, Wyckoff H. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature*. 1958;181.

Khade, N. V. & Sugiyama, T. Roles of C-Terminal Region of Yeast and Human Rad52 in Rad51-Nucleoprotein Filament Formation and ssDNA Annealing. *PLoS One* 11, e0158436, doi:10.1371/journal.pone.0158436 (2016).

Kim, C. et al. Core Protein-Directed Antivirals and Importin beta Can Synergistically Disrupt Hepatitis B Virus Capsids. *J Virol* 96, e0139521, doi:10.1128/JVI.01395-21 (2022).

Kinoshita, C. et al. The cryo-EM structure of full-length RAD52 protein contains an undecameric ring. *FEBS Open Bio*, doi:10.1002/2211-5463.13565 (2023).

Kofoed MA, Wampler DA, Pandey AS, Peters JW, Ensign SA. Roles of the redox-active disulfide and histidine residues forming a catalytic dyad in reactions catalyzed by 2-ketopropyl coenzyme M oxidoreductase/carboxylase. *J Bacteriol*. 2011. doi:10.1128/JB.05231-11

Konermann L, Pan J, Liu YH. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem Soc Rev*. 2011;40(3):1224-1234. doi:10.1039/c0cs00113a

- Kopec J, Schneider G. Comparison of fluorescence and light scattering based methods to assess formation and stability of protein-protein complexes. *J Struct Biol*. 2011;175(2):216-223. doi:10.1016/j.jsb.2011.04.006
- Koshland DE, Némethy G, Filmer D. Comparison of Experimental Binding Data. *Theor Model proteins Contain subunits*. 1966;5(1):365-385.
- Koshland DE. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci*. 1958;44(2):98-104. doi:10.1073/pnas.44.2.98
- Koshland DE. The Key–Lock Theory and the Induced Fit Theory. *Angew Chemie Int Ed English*. 1995;33(23-24):2375-2378. doi:10.1002/anie.199423751
- Krejci, L. et al. Interaction with Rad51 is indispensable for recombination mediator function of Rad52. *J Biol Chem* 277, 40132-40141, doi:10.1074/jbc.M206511200 (2002).
- Kuppa, S. et al. Rtt105 regulates RPA function by configurationally stapling the flexible domains. *Nat Commun* 13, 5152, doi:10.1038/s41467-022-32860-6 (2022).
- Kurochkin I V., Guarnera E, Wong JH, Eisenhaber F, Berezovsky IN. Toward allosterically increased catalytic activity of insulin-degrading enzyme against amyloid peptides. *Biochemistry*. 2017;56(1):228-239. doi:10.1021/acs.biochem.6b00783
- Lee, J. Y. et al. DNA RECOMBINATION. Base triplet stepping by the Rad51/RecA family of recombinases. *Science* 349, 977-981, doi:10.1126/science.aab2666 (2015).
- Li, F. et al. Identification of an integrase-independent pathway of retrotransposition. *Sci Adv* 8, eabm9390, doi:10.1126/sciadv.abm9390 (2022).
- Lichtenthaler FW. 100 Years “Schlüssel-Schloss-Prinzip”: What Made Emil Fischer Use this Analogy? *Angew Chemie Int Ed English*. 1995;33(23-24):2364-2374. doi:10.1002/anie.199423641
- Liu S-Q, Ji X-L, Tao Y, Tan D-Y, Zhang K-Q, Fu Y-X. Protein Folding, Binding and Energy Landscape: A Synthesis. *Protein Eng*. 2012. doi:10.5772/30440
- Ma, E. et al. Rad52-Rad51 association is essential to protect Rad51 filaments against Srs2, but facultative for filament formation. *Elife* 7, doi:10.7554/eLife.32744 (2018).
- Marechal, A. & Zou, L. RPA-coated single-stranded DNA as a platform for post-translational modifications in the DNA damage response. *Cell Res* 25, 9-23, doi:10.1038/cr.2014.147 (2015).
- Masson GR, Burke JE, Ahn NG, et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods*. 2019;16(7):595-602. doi:10.1038/s41592-019-0459-y

Matulis D, Kranz JK, Salemme FR, Todd MJ. Thermodynamic stability of carbonic anhydrase: Measurements of binding affinity and stoichiometry using thermofluor. *Biochemistry*. 2005;44(13):5258-5266. doi:10.1021/bi048135v

Mazina, O. M., Keskin, H., Hanamshet, K., Storici, F. & Mazin, A. V. Rad52 Inverse Strand Exchange Drives RNA-Templated DNA Double-Strand Break Repair. *Mol Cell* 67, 19-29 e13, doi:10.1016/j.molcel.2017.05.019 (2017).

McDevitt, S., Rusanov, T., Kent, T., Chandramouly, G. & Pomerantz, R. T. How RNA transcripts coordinate DNA recombination and repair. *Nat Commun* 9, 1091, doi:10.1038/s41467-018-03483-7 (2018).

Monod J, Changeux JP, Jacob F. Allosteric proteins and cellular control systems. *J Mol Biol*. 1963;6(4):306-329. doi:10.1016/S0022-2836(63)80091-1

Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: A plausible model. *J Mol Biol*. 1965;12(1):88-118. doi:10.1016/S0022-2836(65)80285-6

Mortensen, U. H., Bendixen, C., Sunjevaric, I. & Rothstein, R. DNA strand annealing is promoted by the yeast Rad52 protein. *Proc Natl Acad Sci U S A* 93, 10729-10734, doi:10.1073/pnas.93.20.10729 (1996).

Motlagh HN, Wrabl JO, Li J, Hilser VJ. The ensemble nature of allostery. *Nature*. 2014;508(7496):331-339. doi:10.1038/nature13001

Mus F, Eilers BJ, Alleman AB, et al. Structural Basis for the Mechanism of ATP-Dependent Acetone Carboxylation. *Sci Rep*. 2017. doi:10.1038/s41598-017-06973-8

Mus F, Wu HH, Alleman AB, et al. Insights into the unique carboxylation reactions in the metabolism of propylene and acetone. *Biochem J*. 2020;477(11):2027-2038. doi:10.1042/BCJ20200174

Navadgi-Patil, V. M. & Burgers, P. M. The unstructured C-terminal tail of the 9-1-1 clamp subunit Ddc1 activates Mec1/ATR via two distinct mechanisms. *Mol Cell* 36, 743-753, doi:10.1016/j.molcel.2009.10.014 (2009).

New, J. H., Sugiyama, T., Zaitseva, E. & Kowalczykowski, S. C. Rad52 protein stimulates DNA strand exchange by Rad51 and replication protein A. *Nature* 391, 407-410, doi:10.1038/34950 (1998).

Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. 2007;2(9):2212-2221. doi:10.1038/nprot.2007.321

Niesen FH, Berglund H, Vedadi M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc.* 2007;2(9):2212-2221.

doi:10.1038/nprot.2007.321

Nocek B, Jang SB, Jeong MS, Clark DD, Ensign SA, Peters JW. Structural basis for CO₂ fixation by a novel member of the disulfide oxidoreductase family of enzymes, 2-ketopropyl-coenzyme M oxidoreductase/carboxylase. *Biochemistry.* 2002;41(43):12907-12913.

doi:10.1021/bi026580p

Oshidari, R. et al. DNA repair by Rad52 liquid droplets. *Nat Commun* 11, 695,

doi:10.1038/s41467-020-14546-z (2020).

Pai EF, Schulz GE. The catalytic mechanism of glutathione reductase as derived from x-ray diffraction analyses of reaction intermediates. *J Biol Chem.* 1983;258(3):1752-1757.

Pandey AS, Mulder DW, Ensign SA, Peters JW. Structural basis for carbon dioxide binding by 2-ketopropyl coenzyme M oxidoreductase/carboxylase. *FEBS Lett.* 2011.

doi:10.1016/j.febslet.2010.12.035

Pandey AS, Mulder DW, Ensign SA, Peters JW. Structural basis for carbon dioxide binding by 2-ketopropyl coenzyme M oxidoreductase/carboxylase. *FEBS Lett.* 2011;585(3):459-464.

doi:10.1016/j.febslet.2010.12.035

Pandey AS, Nocek B, Clark DD, Ensign SA, Peters JW. Mechanistic implications of the structure of the mixed-disulfide intermediate of the disulfide oxidoreductase, 2-ketopropyl-coenzyme M oxidoreductase/carboxylase. *Biochemistry.* 2006;45(1):113-120.

doi:10.1021/bi051518o

Paques, F. & Haber, J. E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* 63, 349-404,

doi:10.1128/MMBR.63.2.349-404.1999 (1999).

Park, M. S., Ludwig, D. L., Stigger, E. & Lee, S. H. Physical interaction between human RAD52 and RPA is required for homologous recombination in mammalian cells. *J Biol Chem* 271, 18996-19000, doi:10.1074/jbc.271.31.18996 (1996).

Patterson, A., Zhao, Z., Waymire, E., Zlotnick, A. & Bothner, B. Dynamics of Hepatitis B Virus Capsid Protein Dimer Regulate Assembly through an Allosteric Network. *ACS Chem Biol* 15, 2273-2280, doi:10.1021/acscchembio.0c00481 (2020).

Paulsen, B. S. et al. Ectopic expression of RAD52 and dn53BP1 improves homology-directed repair during CRISPR-Cas9 genome editing. *Nat Biomed Eng* 1, 878-888, doi:10.1038/s41551-017-0145-2 (2017).

Perutz M, Rossmann M, Cullis AF, Muirhead H, Will G, North ACT. Structure of hoemoglobin. *Nature*. 1960;185:416-422.

Pettersen, E. F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605-1612, doi:10.1002/jcc.20084 (2004).

Pettersen, E. F. et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* 30, 70-82, doi:10.1002/pro.3943 (2021).

Plate, I. et al. Interaction with RPA is necessary for Rad52 repair center formation and for its mediator activity. *J Biol Chem* 283, 29077-29085, doi:10.1074/jbc.M804881200 (2008).

Pokhrel, N. et al. Dynamics and selective remodeling of the DNA-binding domains of RPA. *Nat Struct Mol Biol* 26, 129-136, doi:10.1038/s41594-018-0181-y (2019).

Pokhrel, N. et al. Monitoring Replication Protein A (RPA) dynamics in homologous recombination through site-specific incorporation of non-canonical amino acids. *Nucleic Acids Res* 45, 9413-9426, doi:10.1093/nar/gkx598 (2017).

Prado, F., Cortes-Ledesma, F., Huertas, P. & Aguilera, A. Mitotic recombination in *Saccharomyces cerevisiae*. *Curr Genet* 42, 185-198, doi:10.1007/s00294-002-0346-3 (2003).

Prussia GA, Gauss GH, Mus F, et al. Substitution of a conserved catalytic dyad into 2-KPCC causes loss of carboxylation activity. *FEBS Lett*. 2016;590:2991-2996. doi:10.1002/1873-3468.12325

Prussia GA, Shisler KA, Zadvornyy OA, Streit BR, DuBois JL, Peters JW. The unique Phe–His dyad of 2-ketopropyl coenzyme M oxidoreductase/carboxylase selectively promotes carboxylation and S–C bond cleavage. *J Biol Chem*. 2021;297(2):100961. doi:10.1016/j.jbc.2021.100961

Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14, 290-296, doi:10.1038/nmeth.4169 (2017).

Qi, Z. et al. DNA sequence alignment by microhomology sampling during homologous recombination. *Cell* 160, 856-869, doi:10.1016/j.cell.2015.01.029 (2015).

Ragunathan, K., Liu, C. & Ha, T. RecA filament sliding on DNA facilitates homology search. *Elife* 1, e00067, doi:10.7554/eLife.00067 (2012).

Rayaprolu V, Kruse S, Kant R, et al. Comparative Analysis of Adeno-Associated Virus Capsid Stability and Dynamics. *J Virol*. 2013;87(24):13150-13160. doi:10.1128/jvi.01415-13

- Rayaprolu V, Kruse S, Kant R, Movahed N, Brooke D, Bothner B. Fluorometric Estimation of Viral Thermal Stability. *Bio-protocol*. 2014;4(15).
<http://www.ncbi.nlm.nih.gov/pubmed/27148558><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4852874>.
- Rietveld P, Arscott LD, Berry A, et al. Reductive and Oxidative Half-Reactions of Glutathione Reductase from *Escherichia coli*. *Biochemistry*. 1994. doi:10.1021/bi00250a043
- San Filippo, J., Sung, P. & Klein, H. Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem* 77, 229-257, doi:10.1146/annurev.biochem.77.061306.125255 (2008).
- Saotome, M. et al. Structural Basis of Homology-Directed DNA Repair Mediated by RAD52. *iScience* 3, 50-62, doi:10.1016/j.isci.2018.04.005 (2018).
- Saotome, M., Saito, K., Onodera, K., Kurumizaka, H. & Kagawa, W. Structure of the human DNA-repair protein RAD52 containing surface mutations. *Acta Crystallogr F Struct Biol Commun* 72, 598-603, doi:10.1107/S2053230X1601027X (2016).
- Schada von Borzyskowski L, Rosenthal RG, Erb TJ. Evolutionary history and
- Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* 78, 1606-1619, doi:10.1016/S0006-3495(00)76713-0 (2000).
- Seong, C. et al. Molecular anatomy of the recombination mediator function of *Saccharomyces cerevisiae* Rad52. *J Biol Chem* 283, 12166-12174, doi:10.1074/jbc.M800763200 (2008).
- Shi, I. et al. Role of the Rad52 amino-terminal DNA binding activity in DNA strand capture in homologous recombination. *J Biol Chem* 284, 33275-33284, doi:10.1074/jbc.M109.057752 (2009).
- Shinohara, A., Shinohara, M., Ohta, T., Matsuda, S. & Ogawa, T. Rad52 forms ring structures and co-operates with RPA in single-strand DNA annealing. *Genes Cells* 3, 145-156, doi:10.1046/j.1365-2443.1998.00176.x (1998).
- Shvartsburg AA, Smith RD. Fundamentals of traveling wave ion mobility spectrometry. *Anal Chem*. 2008;80(24):9689-9699. doi:10.1021/ac8016295
- Singleton, M. R., Wentzell, L. M., Liu, Y., West, S. C. & Wigley, D. B. Structure of the single-strand annealing domain of human RAD52 protein. *Proc Natl Acad Sci U S A* 99, 13492-13497, doi:10.1073/pnas.212449899 (2002).
- Sluis MK, Ensign SA. Purification and characterization of acetone carboxylase from *Xanthobacter* strain Py2. *Proc Natl Acad Sci U S A*. 1997;94(16):8456-8461. doi:10.1073/pnas.94.16.8456

- Smith DP, Knapman TW, Campuzano L, et al. Deciphering drift time measurements from travelling wave ion mobility spectrometry- Mass spectrometry studies. *Eur J Mass Spectrom.* 2009;15(2):113-130. doi:10.1255/ejms.947
- Song, B. & Sung, P. Functional interactions among yeast Rad51 recombinase, Rad52 mediator, and replication protein A in DNA strand exchange. *J Biol Chem* 275, 15895-15904, doi:10.1074/jbc.M910244199 (2000).
- Sowole MA, Konermann L. Effects of protein-ligand interactions on hydrogen/deuterium exchange kinetics: Canonical and noncanonical scenarios. *Anal Chem.* 2014;86(13):6715-6722. doi:10.1021/ac501849n
- Stasiak, A. Z. et al. The human Rad52 protein exists as a heptameric ring. *Curr Biol* 10, 337-340, doi:10.1016/s0960-9822(00)00385-7 (2000).
- Stewart, M. D. et al. Homologous Recombination Deficiency: Concepts, Definitions, and Assays. *Oncologist* 27, 167-174, doi:10.1093/oncolo/oyab053 (2022).
- Streit BR, Mattice JR, Prussia GA, Peters JW, DuBois JL. The reactive form of a C-S bond-cleaving, CO₂-fixing flavoenzyme. *J Biol Chem.* 2019;294:jbc.RA118.005554. doi:10.1074/jbc.ra118.005554
- Sugawara, N., Wang, X. & Haber, J. E. In vivo roles of Rad52, Rad54, and Rad55 proteins in Rad51-mediated recombination. *Mol Cell* 12, 209-219, doi:10.1016/s1097-2765(03)00269-7 (2003).
- Sugiyama, T. & Kowalczykowski, S. C. Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation. *J Biol Chem* 277, 31663-31672, doi:10.1074/jbc.M203494200 (2002).
- Sugiyama, T., New, J. H. & Kowalczykowski, S. C. DNA annealing by RAD52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proc Natl Acad Sci U S A* 95, 6049-6054, doi:10.1073/pnas.95.11.6049 (1998).
- Sung, P. & Klein, H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* 7, 739-750, doi:10.1038/nrm2008 (2006).
- Sung, P. Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. *J Biol Chem* 272, 28194-28197, doi:10.1074/jbc.272.45.28194 (1997).
- Swain JF, Gierasch LM. The changing landscape of protein allostery. *Curr Opin Struct Biol.* 2006;16(1):102-108. doi:10.1016/j.sbi.2006.01.003

Symington, L. S. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol Mol Biol Rev* 66, 630-670, table of contents, doi:10.1128/MMBR.66.4.630-670.2002 (2002).

Symington, L. S., Rothstein, R. & Lisby, M. Mechanisms and regulation of mitotic recombination in *Saccharomyces cerevisiae*. *Genetics* 198, 795-835, doi:10.1534/genetics.114.166140 (2014).

Tan ZW, Guarnera E, Tee WV, Berezovsky IN. AlloSigMA 2: Paving the way to designing allosteric effectors and to exploring allosteric effects of mutations. *Nucleic Acids Res.* 2020;48(W1):W116-W124. doi:10.1093/NAR/GKAA338

Tee WV, Guarnera E, Berezovsky IN. Reversing allosteric communication: From detecting allosteric sites to inducing and tuning targeted allosteric response. *PLoS Comput Biol.* 2018;14(6):1-26. doi:10.1371/journal.pcbi.1006228

Tian P, Best RB. How Many Protein Sequences Fold to a Given Structure? A Coevolutionary Analysis. *Biophys J.* 2017;113(8):1719-1730. doi:10.1016/j.bpj.2017.08.039

Tokmina-Lukaszewska M, Patterson A, Berry L, Scott L, Balasubramanian N, Bothner B. The role of mass spectrometry in structural studies of flavin-based electron bifurcating enzymes. *Front Microbiol.* 2018;9(JUL):1-18. doi:10.3389/fmicb.2018.01397

Tokmina-Lukaszewska, M. et al. The Role of Mass Spectrometry in Structural Studies of Flavin-Based Electron Bifurcating Enzymes. *Front Microbiol* 9, 1397, doi:10.3389/fmicb.2018.01397 (2018).

Tsai C-J, Kumar S, Ma B, Nussinov R. Folding funnels, binding funnels, and protein function. *Protein Sci.* 1999;8(6):1181-1190. doi:10.1110/ps.8.6.1181

Uetrecht C, Rose RJ, Van Duijn E, Lorenzen K, Heck AJR. Ion mobility mass spectrometry of proteins and protein assemblies. *Chem Soc Rev.* 2010;39(5):1633-1655. doi:10.1039/b914002f

Uversky VN. Intrinsically disordered proteins from A to Z. *Int J Biochem Cell Biol.* 2011;43(8):1090-1103. doi:10.1016/j.biocel.2011.04.001

Van Gent, D. C., Hoeijmakers, J. H. & Kanaar, R. Chromosomal stability and the DNA double-stranded break connection. *Nat Rev Genet* 2, 196-206, doi:10.1038/35056049 (2001).

Vedadi M, Niesen FH, Allali-Hassani A, et al. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A.* 2006;103(43):15835-15840. doi:10.1073/pnas.0605224103

Waters Corporation. Waters SYNAPT G2 Mass Spectrometry System Operator's Overview and Maintenance Guide Revision A. 2009.

Welty, S. et al. RAD52 is required for RNA-templated recombination repair in post-mitotic neurons. *J Biol Chem* 293, 1353-1362, doi:10.1074/jbc.M117.808402 (2018).

Wold, M. S. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* 66, 61-92, doi:10.1146/annurev.biochem.66.1.61 (1997).

Wolynes PG, Onuchic JN, Thirumalai D. Navigating the folding routes. *Science* (80-). 1995;267(5204):1619-1620. doi:10.1126/science.7886447

Wright PE, Dyson HJ. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J Mol Biol.* 1999;293(2):321-331. doi:10.1006/jmbi.1999.3110

Wright TA, Stewart JM, Page RC, Konkolewicz D. Extraction of Thermodynamic Parameters of Protein Unfolding Using. 2017:6-11. doi:10.1021/acs.jpcllett.6b02894

Xu, Y., Zhou, H., Post, G., Zan, H. & Casali, P. Rad52 mediates class-switch DNA recombination to IgD. *Nat Commun* 13, 980, doi:10.1038/s41467-022-28576-2 (2022).

Yan, Z. et al. Rad52 Restrains Resection at DNA Double-Strand Break Ends in Yeast. *Mol Cell* 76, 699-711 e696, doi:10.1016/j.molcel.2019.08.017 (2019).