

PI: Liu, Shixin	Title: Probing Symmetry Breaking in Epigenetic Inheritance: From Single Molecules to Systems Biology	
Received: 09/07/2017	FOA: RM17-006	Council: 05/2018
Competition ID: FORMS-D	FOA Title: NIH Director's New Innovator Award Program (DP2)	
1 DP2 HG010510-01	Dual: OD, RM	Accession Number: 4087047
IPF: 7056601	Organization: ROCKEFELLER UNIVERSITY	
Former Number: 1DP2OD025835-01	Department: Nanoscale Biophysics and Bioch	
IRG/SRG: ZRG1 MOSS-R (70)R	AIDS: N	Expedited: N
Subtotal Direct Costs (excludes consortium F&A) Year 1: 1,500,000	Animals: N Humans: N Clinical Trial: N Current HS Code: 10 HESC: N HFT: N	New Investigator: Y Early Stage Investigator: Y
<i>Senior/Key Personnel:</i>		
	<i>Organization:</i>	<i>Role Category:</i>
Shixin Liu	The Rockefeller University	PD/PI

APPLICATION FOR FEDERAL ASSISTANCE
SF 424 (R&R)

3. DATE RECEIVED BY STATE		State Application Identifier
1. TYPE OF SUBMISSION*		4.a. Federal Identifier
<input type="radio"/> Pre-application <input checked="" type="radio"/> Application <input type="radio"/> Changed/Corrected Application		b. Agency Routing Number 6 MCB; 8 HIB
2. DATE SUBMITTED 2017-09-07	Application Identifier 0000002316	c. Previous Grants.gov Tracking Number
5. APPLICANT INFORMATION		Organizational DUNS*: [REDACTED]
Legal Name*: The Rockefeller University Department: Nanoscale Biophysics and Bioch Division: Street1*: [REDACTED] Street2: City*: [REDACTED] County: State*: [REDACTED] Province: Country*: [REDACTED] ZIP / Postal Code*: [REDACTED]		
Person to be contacted on matters involving this application Prefix: First Name*: [REDACTED] Middle Name: Last Name*: [REDACTED] Suffix: Position/Title: [REDACTED] Street1*: [REDACTED] Street2: [REDACTED] City*: [REDACTED] County: State*: [REDACTED] Province: Country*: [REDACTED] ZIP / Postal Code*: [REDACTED] Phone Number*: [REDACTED] Fax Number: N/A Email: [REDACTED]		
6. EMPLOYER IDENTIFICATION NUMBER (EIN) or (TIN)* [REDACTED]		
7. TYPE OF APPLICANT*		<input type="radio"/> Private Institution of Higher Education
Other (Specify): <input checked="" type="radio"/> Small Business Organization Type <input type="radio"/> Women Owned <input type="radio"/> Socially and Economically Disadvantaged		
8. TYPE OF APPLICATION*		If Revision, mark appropriate box(es).
<input checked="" type="radio"/> New <input type="radio"/> Resubmission <input type="radio"/> Renewal <input type="radio"/> Continuation <input type="radio"/> Revision		<input type="radio"/> A. Increase Award <input type="radio"/> B. Decrease Award <input type="radio"/> C. Increase Duration <input type="radio"/> D. Decrease Duration <input type="radio"/> E. Other (specify) :
Is this application being submitted to other agencies?* <input type="radio"/> Yes <input checked="" type="radio"/> No What other Agencies?		
9. NAME OF FEDERAL AGENCY* National Institutes of Health		10. CATALOG OF FEDERAL DOMESTIC ASSISTANCE NUMBER TITLE:
11. DESCRIPTIVE TITLE OF APPLICANT'S PROJECT* Probing Symmetry Breaking in Epigenetic Inheritance: From Single Molecules to Systems Biology		
12. PROPOSED PROJECT		13. CONGRESSIONAL DISTRICTS OF APPLICANT
Start Date* 09/30/2018	Ending Date* 06/30/2023	NY-012

14. PROJECT DIRECTOR/PRINCIPAL INVESTIGATOR CONTACT INFORMATION

Prefix: First Name*: Shixin Middle Name: Last Name*: Liu Suffix:

Position/Title:

Organization Name*: The Rockefeller University

Department: Nanoscale Biophysics and Bioch

Division:

Street1*: [REDACTED]

Street2: [REDACTED]

City*: [REDACTED]

County:

State*: [REDACTED]

Province:

Country*: [REDACTED]

ZIP / Postal Code*: [REDACTED]

Phone Number*: [REDACTED] Fax Number: [REDACTED] Email*: [REDACTED]

15. ESTIMATED PROJECT FUNDING

a. Total Federal Funds Requested* \$1,500,000.00

b. Total Non-Federal Funds* \$0.00

c. Total Federal & Non-Federal Funds* \$1,500,000.00

d. Estimated Program Income* \$0.00

16. IS APPLICATION SUBJECT TO REVIEW BY STATE EXECUTIVE ORDER 12372 PROCESS?*

a. YES THIS PREAPPLICATION/APPLICATION WAS MADE AVAILABLE TO THE STATE EXECUTIVE ORDER 12372 PROCESS FOR REVIEW ON:

DATE:

b. NO PROGRAM IS NOT COVERED BY E.O. 12372; OR

PROGRAM HAS NOT BEEN SELECTED BY STATE FOR REVIEW

17. By signing this application, I certify (1) to the statements contained in the list of certifications* and (2) that the statements herein are true, complete and accurate to the best of my knowledge. I also provide the required assurances * and agree to comply with any resulting terms if I accept an award. I am aware that any false, fictitious, or fraudulent statements or claims may subject me to criminal, civil, or administrative penalties. (U.S. Code, Title 18, Section 1001)

I agree*

* The list of certifications and assurances, or an Internet site where you may obtain this list, is contained in the announcement or agency specific instructions.

18. SFLL or OTHER EXPLANATORY DOCUMENTATION File Name:

19. AUTHORIZED REPRESENTATIVE

Prefix: First Name*: [REDACTED] Middle Name: Last Name*: [REDACTED] Suffix:

Position/Title*: [REDACTED]

Organization Name*: [REDACTED]

Department: [REDACTED]

Division:

Street1*: [REDACTED]

Street2: [REDACTED]

City*: [REDACTED]

County:

State*: [REDACTED]

Province:

Country*: [REDACTED]

ZIP / Postal Code*: [REDACTED]

Phone Number*: [REDACTED] Fax Number: N/A Email*: [REDACTED]

Signature of Authorized Representative* [REDACTED] **Date Signed*** [REDACTED]

20. PRE-APPLICATION File Name:

21. COVER LETTER ATTACHMENT File Name: Cover Letter.pdf

**424 R&R and PHS-398 Specific
Table Of Contents**

SF 424 R&R Cover Page.....	1
Table of Contents.....	3
Performance Sites.....	4
Research & Related Other Project Information.....	5
Project Summary/Abstract(Description).....	6
Project Narrative.....	7
Facilities & Other Resources.....	8
Research & Related Senior/Key Person.....	9
PHS398 Cover Page Supplement.....	17
PHS 398 Research Plan.....	19
Research Strategy.....	20
Authentication of Key Biological and/or Chemical Resources.....	30

Project/Performance Site Location(s)

Project/Performance Site Primary Location

I am submitting an application as an individual, and not on behalf of a company, state, local or tribal government, academia, or other type of organization.

Organization Name: The Rockefeller University
Duns Number: [REDACTED]
Street1*: [REDACTED]
Street2:
City*: [REDACTED]
County:
State*: [REDACTED]
Province:
Country*: [REDACTED]
Zip / Postal Code*: [REDACTED]
Project/Performance Site Congressional District*: NY-012

Additional Location(s)

File Name:

RESEARCH & RELATED Other Project Information

1. Are Human Subjects Involved?* <input type="radio"/> Yes <input checked="" type="radio"/> No	
1.a. If YES to Human Subjects Is the Project Exempt from Federal regulations? <input type="radio"/> Yes <input type="radio"/> No If YES, check appropriate exemption number: — 1 — 2 — 3 — 4 — 5 — 6 If NO, is the IRB review Pending? <input type="radio"/> Yes <input type="radio"/> No IRB Approval Date: Human Subject Assurance Number	
2. Are Vertebrate Animals Used?* <input type="radio"/> Yes <input checked="" type="radio"/> No	
2.a. If YES to Vertebrate Animals Is the IACUC review Pending? <input type="radio"/> Yes <input type="radio"/> No IACUC Approval Date: Animal Welfare Assurance Number	
3. Is proprietary/privileged information included in the application?* <input type="radio"/> Yes <input checked="" type="radio"/> No	
4.a. Does this project have an actual or potential impact - positive or negative - on the environment?* <input type="radio"/> Yes <input checked="" type="radio"/> No	
4.b. If yes, please explain: 4.c. If this project has an actual or potential impact on the environment, has an exemption been authorized or an environmental assessment (EA) or environmental impact statement (EIS) been performed? <input type="radio"/> Yes <input type="radio"/> No 4.d. If yes, please explain:	
5. Is the research performance site designated, or eligible to be designated, as a historic place?* <input type="radio"/> Yes <input checked="" type="radio"/> No	
5.a. If yes, please explain:	
6. Does this project involve activities outside the United States or partnership with international collaborators?* <input type="radio"/> Yes <input checked="" type="radio"/> No	
6.a. If yes, identify countries: 6.b. Optional Explanation:	
7. Project Summary/Abstract*	Filename Project Summary.pdf
8. Project Narrative*	Project Narrative.pdf
9. Bibliography & References Cited	
10. Facilities & Other Resources	Facilities and Resources_1.pdf
11. Equipment	

PROJECT SUMMARY

Nucleosomes are not only the building blocks of eukaryotic chromatin, but also major carriers of epigenetic information. Shortly following DNA replication, new nucleosomes are assembled on the nascent strands from a mixture of recycled parental histones and newly synthesized ones. This process allows the chromatin state to be faithfully transmitted for many generations, while also enabling the offspring to reset its epigenome if necessary. Another principal, somewhat less appreciated, characteristic of chromatin replication is that it is inherently asymmetric, with the leading and lagging strands synthesized by drastically different machineries and pathways. Such asymmetry entails that each of the two daughter strands incorporates a distinct set of histones, thereby capable of adopting different chromatin states. Asymmetric histone segregation plays a pivotal role in cell fate decisions during development and stem cell maintenance. However, whether histones are deterministically or randomly distributed on nascent DNA, and how the distribution pattern depends on chromatin contexts, remain poorly understood. This is largely due to a lack of experimental tools with which to distinguish between leading- and lagging-strand deposition of histones. Using budding yeast as a model organism, this project aims to develop novel methodologies to probe histone inheritance in a strand-specific manner. At the molecular level, the chain of events during replication-coupled chromatin assembly will be directly visualized by correlative single-molecule fluorescence-optical tweezers microscopy. The leading and lagging strands will be spatially separated via microfluidic control so that the destination of deposited histones can be unambiguously assigned. This assay will then be used to dissect the functions of histone chaperones in mediating nucleosome disruption and reassembly. At the cellular level, a cell-cycle-dependent pulse-labeling strategy will be adopted to distinguish newly synthesized histones from preexisting ones. Histone distribution at the replication fork will be imaged and cell-to-cell variability assessed by super-resolution microscopy. At the systems level, nucleosomes on the leading or lagging strands will be selectively isolated through the usage of yeast strains carrying corresponding error-prone replicative polymerase variants. The genomic positions of these strand-selective nucleosomes will be mapped by next-generation sequencing and their chemical compositions analyzed by quantitative mass spectrometry. Overall, this multi-pronged approach will establish a framework for characterizing the differential partitioning of histones, their variants and post-translational modifications behind the replication fork, and for understanding the role of asymmetry in driving epigenome evolution and creating heterogeneity of cell identity. This project also paves the way for investigating how the dynamics of epigenetic inheritance alter in development and disease, thereby providing new therapeutic targets and strategies for developmental disorders, ageing, and cancer.

PROJECT NARRATIVE

Chromatin replication is an inherently asymmetric process, with two daughter strands produced by drastically different mechanisms. Such asymmetry allows sister chromatids to adopt distinct chromatin states and plays essential roles in development and stem cell maintenance. This project aims at elucidating how epigenetic information is differentially inherited on leading and lagging strands during DNA replication, thereby contributing to knowledge of basic biology and informing therapeutic strategies for developmental disorders, premature ageing, and cancer.

FACILITIES AND RESOURCES

Laboratory Space

I established the Laboratory of Nanoscale Biophysics and Biochemistry, located on the B and C levels of Theobald Smith Hall at the Rockefeller University, New York City. The newly renovated lab space (~ 3,100 square feet) includes a wet lab with eight benches, a 4°C cold room, a 37°C warm room, a dedicated radioactivity biochemistry room, equipment rooms, four optics suites, and office space. The optics suites are equipped with environmental control ($\pm 0.3^\circ\text{C}$ temperature stability, $\pm 5\%$ humidity stability) and are acoustically and vibrationally shielded to accommodate ultrahigh-resolution measurements.

Major Equipment

My lab owns several single-molecule instruments required for the proposed project, including a multicolor total-internal-reflection fluorescence microscope, an analytical optical tweezers instrument, and a single-molecule fluorescence-force combo instrument featuring dual-trap optical tweezers combined with confocal fluorescence microscopy and automated microfluidics (the first of its kind in the United States). In addition, the lab houses a variety of equipment items for molecular biology and biochemical assays, including several centrifuges, an ultracentrifuge, a Typhoon gel imager, an AKTA pure chromatography system, a real-time PCR system, a Nanodrop spectrophotometer, two PCR thermocyclers, multiple freezers and refrigerators.

University Resource Centers

The Rockefeller University houses a number of state-of-the-art core facilities. Facilities relevant to the proposed project are: Bio-Imaging Resource Center (instrumentation and analysis software for cell imaging with wide-field, confocal, multi-photon, and super-resolution fluorescence microscopy); Genomics Resource Center (services and instruments for high-throughput sequencing, including Illumina HiSeq2500, NextSeq500, and MiSeq platforms); Proteomics Resource Center (mass spectrometry-based proteomics and production of custom peptides); High-Throughput and Spectroscopy Resource Center (spectroscopic equipment for measuring biomolecular interactions, such as Isothermal Titration Calorimetry and Microscale Thermophoresis); Precision Fabrication Facility (nanofabrication, microfluidics, design and construction of prototype equipment); Antibody Core Facility (custom monoclonal antibodies).

Institutional Environment

The open, department-less structure of the Rockefeller University encourages collaboration across disciplines and empowers investigators to take on high-risk, high-reward projects. The university's 81 laboratories work in a broad range of fields including both basic and clinical research. For the proposed project, we will consult

[REDACTED]

[REDACTED]

[REDACTED] In addition to the resources within my laboratory and university, this project will also benefit from various resources and facilities at the nearby Weill Cornell Medicine and Memorial Sloan Kettering Cancer Center. This Tri-Institutional community is highly collaborative and attracts leading scientists from around the world to visit and communicate ideas and results.

Support for Early Stage Investigators

The Rockefeller University fosters an exceptionally supportive environment for its junior faculty. [REDACTED]

[REDACTED]

[REDACTED]

RESEARCH & RELATED Senior/Key Person Profile (Expanded)

PROFILE - Project Director/Principal Investigator				
Prefix:	First Name*: Shixin	Middle Name	Last Name*: Liu	Suffix:
Position/Title*:				
Organization Name*:	The Rockefeller University			
Department:	Nanoscale Biophysics and Bioch			
Division:				
Street1*:	[REDACTED]			
Street2:	[REDACTED]			
City*:	[REDACTED]			
County:				
State*:	[REDACTED]			
Province:				
Country*:	[REDACTED]			
Zip / Postal Code*:	[REDACTED]			
Phone Number*:	[REDACTED]	Fax Number:	[REDACTED]	
E-Mail*:	[REDACTED]			
Credential, e.g., agency login:	[REDACTED]			
Project Role*:	PD/PI	Other Project Role Category:		
Degree Type:	Degree Year:			
Attach Biographical Sketch*:	File Name:	Biographical Sketch.pdf		
Attach Current & Pending Support:	File Name:	Other Support.pdf		

BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.

Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Liu, Shixin

eRA COMMONS USER NAME (credential, e.g., agency login): XXXXXXXXXX

POSITION TITLE: Assistant Professor, Head of Laboratory of Nanoscale Biophysics and Biochemistry

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)

INSTITUTION AND LOCATION	DEGREE (if applicable)	Completion Date MM/YYYY	FIELD OF STUDY
University of Science & Technology of China, Hefei, China	B.S.	06/2003	Biological Sciences
Harvard University, Cambridge, MA	Ph.D.	06/2009	Chemistry
University of California, Berkeley, CA	Postdoctoral	12/2015	Biophysics

A. Personal Statement

I have extensive research experiences in the fields of single-molecule biophysics and molecular motors. Trained as a molecular biologist in college, I did my Ph.D. research at the Department of Chemistry and Chemical Biology of Harvard University, where I used single-molecule fluorescence methods to study enzymatic reactions and protein-nucleic acid interactions. I then performed postdoctoral research at the University of California, Berkeley, using high-resolution optical tweezers to study the mechanical processes catalyzed by biomolecular motors. My work as a graduate student and as a postdoctoral fellow produced a number of high-impact publications. In my independent laboratory at the Rockefeller University, I am developing next-generation single-molecule assays to study the interaction, cooperation and competition between macromolecular complexes involved in prokaryotic and eukaryotic gene replication and expression. This New Innovator project concerns the fundamental questions of how epigenetic information is transmitted to newly synthesized daughter strands and how asymmetric inheritance between leading and lagging strands leads to distinct cell fates. This project builds upon my expertise in single-molecule biophysics. It also breaks new ground by proposing complementary cell imaging, genomics and proteomics experiments, thereby broadening the scope of my research program. The Rockefeller University provides an exceptionally nurturing environment for early-career investigators, a strong and collaborative group of faculty, and state-of-the-art facilities and resources, which empower me to take on this high-risk, high-reward project. Since becoming a PI, I have successfully obtained funding from federal and private sources, and established collaborations with scientists both within and outside my institution. In summary, I have a strong record of productive research in areas of high relevance to the basic mechanism of cell operation and the molecular origin of human disease; I am in an environment that offers optimal intellectual and technical support for the success of this project.

1. S. Liu, E. A. Abbondanzieri, J. W. Rausch, S. F. Le Grice, X. Zhuang. (2008) Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science* 322, 1092-1097. PMID: PMC2717043

2. S. Liu*, B. T. Harada*, J. T. Miller, S. F. Le Grice, X. Zhuang. (2010) Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription. *Nature Structural & Molecular Biology* 17, 1453-1460. PMID: PMC3058889
3. G. Chistol*, S. Liu*, C. L. Hetherington, J. R. Moffitt, S. Grimes, P. J. Jardine, C. Bustamante. (2012) High degree of coordination and division of labor among subunits in a homomeric ring ATPase. *Cell* 151, 1017-1028. PMID: PMC3652982
4. S. Liu*, G. Chistol*, C. L. Hetherington*, S. Tafoya, K. Aathavan, J. Schnitzbauer, S. Grimes, P. J. Jardine, C. Bustamante. (2014) A viral packaging motor varies its DNA rotation and step size to preserve subunit coordination as the capsid fills. *Cell* 157, 702-713. PMID: PMC4003460

(* denotes co-first author)

B. Positions and Honors

Positions and Employment

2010-2015 Postdoctoral Fellow, University of California, Berkeley, CA
 2016- Assistant Professor, Head of Laboratory, The Rockefeller University, New York, NY
 2016- Faculty Member, Tri-Institutional PhD Program in Chemical Biology, New York, NY

Professional Memberships

2005- Member, Biophysical Society
 2017- Member, American Society for Biochemistry and Molecular Biology

Honors

1998 Ranked 1st Place in the Special Class for the Gifted Young, Top Freshman Scholarship, University of Science & Technology of China, Hefei, China
 2003 Fieser Graduate Research Grant, Harvard University, Cambridge, MA
 2008 Chinese Government Award for Outstanding Student Abroad
 2012 Education Travel Award, Biophysical Society
 2013 NIH Pathway to Independence Award (K99/R00), National Institute of General Medical Sciences
 2016 Monique Weill-Caulier Career Scientist Award, Irma T. Hirschl/Monique Weill-Caulier Trusts
 2017 Quadrivium Award for Innovative Research in Epigenetics, The Quadrivium Foundation
 2017 Basil O'Connor Starter Scholar Research Award, March of Dimes Foundation
 2017 Kimmel Scholar Award, Sidney Kimmel Foundation

C. Contribution to Science

1. Many fundamental cellular processes are inherently directional and are carried out by nanoscale machine-like devices known as molecular motors. A large number of molecular motors are oligomers organized in ring-shaped structures. I used single-molecule manipulation methods to study the mechanism by which ring motors couple chemical energy into mechanical work and coordinate the activity of their constituent subunits. Our model enzyme, the bacteriophage $\phi 29$ DNA packaging motor, is a ring ATPase that packages the viral genome into a protein capsid, overcoming a large energy barrier. We found that the five ring subunits coordinate their chemical and mechanical transitions in a precisely timed manner. Moreover, we discovered an unexpected division of labor in this homomeric ring motor: one subunit plays a regulatory role, while the other four participate in substrate translocation. We then tracked DNA packaging in three dimensions and found that the motor rotates DNA during translocation. As the capsid fills and the internal pressure increases, the packaging motor modulates multiple aspects of its mechanochemical cycle. Remarkably, the motor simultaneously changes its DNA rotation and step size in

order to preserve intersubunit coordination at high filling. These results illustrate how a biological motor adjusts its operation in response to environmental changes. I was the first or co-first author in all of the following publications. One other manuscript is currently under review.

- a. G. Chistol*, S. Liu*, C. L. Hetherington, J. R. Moffitt, S. Grimes, P. J. Jardine, C. Bustamante. (2012) High degree of coordination and division of labor among subunits in a homomeric ring ATPase. *Cell* 151, 1017-1028. PMID: PMC3652982
- b. S. Liu*, G. Chistol*, C. L. Hetherington*, S. Tafoya, K. Aathavan, J. Schnitzbauer, S. Grimes, P. J. Jardine, C. Bustamante. (2014) A viral packaging motor varies its DNA rotation and step size to preserve subunit coordination as the capsid fills. *Cell* 157, 702-713. PMID: PMC4003460
- c. S. Liu*, G. Chistol*, C. Bustamante. (2014) Mechanical operation and intersubunit coordination of ring-shaped molecular motors: insights from single-molecule studies. *Biophysical Journal* 106, 1844-1858. PMID: PMC4017299
- d. S. Liu#, S. Tafoya, C. Bustamante#. (2017) Deciphering the molecular mechanism of the bacteriophage ϕ 29 DNA packaging motor. *Methods in Molecular Biology* 1486, 343-355.
(# denotes co-corresponding author)

2. DNA-protein or RNA-protein complexes play crucial roles in a multitude of biological processes. Single-molecule fluorescence spectroscopy is well suited for monitoring the assembly and dynamics of these complexes. My research primarily focused on HIV reverse transcriptase (RT), a key enzyme in the HIV life cycle and a major target for AIDS therapy. RT encounters various nucleic acid substrates during reverse transcription, converting single-stranded viral RNA into double-stranded DNA. We discovered that RT is a remarkably dynamic enzyme that can spontaneously flip between two opposite binding orientations and slide over long distances on nucleic acid duplexes. Flipping and sliding allow RT to efficiently switch between multiple functional modes, including DNA polymerization, RNase H cleavage, and strand displacement synthesis. Such dynamic flexibility may be a general design principle for multifunctional enzymes to rapidly access different binding configurations and perform specific activities. I also used the single-molecule assay to explore the effects of various RT inhibitors on the enzyme's motions. These experiments revealed new inhibition mechanisms that can be exploited to develop novel antiretroviral drugs. I was the leading author in two of these studies and a co-author in two others.

- a. S. Liu, E. A. Abbondanzieri, J. W. Rausch, S. F. Le Grice, X. Zhuang. (2008) Slide into action: dynamic shuttling of HIV reverse transcriptase on nucleic acid substrates. *Science* 322, 1092-1097. PMID: PMC2717043
- b. S. Liu*, B. T. Harada*, J. T. Miller, S. F. Le Grice, X. Zhuang. (2010) Initiation complex dynamics direct the transitions between distinct phases of early HIV reverse transcription. *Nature Structural & Molecular Biology* 17, 1453-1460. PMID: PMC3058889
- c. S. Chung, M. Wendeler, J. W. Rausch, G. Beilhartz, M. Gotte, B. R. O'Keefe, A. Bermingham, J. A. Beutler, S. Liu, X. Zhuang, S. F. Le Grice. (2010) Structure-activity analysis of vinylogous urea inhibitors of human immunodeficiency virus-encoded ribonuclease H. *Antimicrobial Agents and Chemotherapy* 54, 3913-3921. PMID: PMC2935023

3. Single-molecule methods are powerful tools for probing the kinetic mechanism of complex enzymatic reactions. However, some intermediate steps, especially chemical ones, do not yield a significant structural change required for single-molecule detection. I developed a method that combines single-molecule FRET with pulse-chase experimental schemes to differentiate each reaction intermediate by a distinct time sequence of FRET signal—a kinetic “fingerprint”. This method was applied to fully characterize the multistep RNA-cleavage reaction by the hairpin ribozyme. I also studied the kinetics of transcription

elongation by eukaryotic RNA polymerase II. By measuring the elongation and pausing dynamics of the polymerase on its natural track—nucleosomal DNA—complemented with theoretical modeling, we determined all the major on- and off-pathway kinetic parameters in the polymerase’s elongation cycle. The resulting model reconciles existing structural, biochemical, and single-molecule data, and serves as a kinetic framework for studying transcriptional regulation. I was the first or co-first author in these two studies.

- a. S. Liu, G. Bokinsky, N. G. Walter, X. Zhuang. (2007) Dissecting the multistep reaction pathway of an RNA enzyme by single-molecule kinetic “fingerprinting”. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12634-12639. PMID: PMC1937518
 - b. M. Dangkulwanich*, T. Ishibashi*, S. Liu*, M. L. Kireeva, L. Lubkowska, M. Kashlev, C. Bustamante. (2013) Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. *eLife* 2, e00971. PMID: PMC3778554
4. Life processes often occur far from equilibrium, making existing thermodynamic models insufficient for interpreting their dynamic behavior and energy consumption. I initiated a project that aimed to delineate minimum-dissipation protocols for out-of-equilibrium biological systems. Using DNA hairpin unfolding as a model system, we found that theoretically predicted optimal protocols indeed require significantly less work than naive ones across a wide range of driving velocities. This validates a simple method to predict energetically efficient nonequilibrium operational protocols from equilibrium behavior, providing insights into design principles of biological functioning. A manuscript is in preparation. We are now building a composite theoretical and experimental framework to study gene regulation by pulsatile transcription factor (TF) binding and active nucleosome remodeling. We will make in vitro and in vivo measurements of TF-nucleosome interaction and transcription to test the hypothesis that far-from-equilibrium energy consumption allows for more versatile and precise control of gene expression.
- a. B. Cheng, S. Wu, S. Liu, P. Rodriguez-Aliaga, J. Yu, S. Cui. (2015) Protein denaturation at a single-molecule level: the effect of nonpolar environment and its implications on the unfolding mechanism by proteases. *Nanoscale* 7, 2970-2977.

Complete List of Published Work:

<http://www.ncbi.nlm.nih.gov/sites/myncbi/1heKcstQAX-k6/collections/48744514/public/>

D. Research Support

Ongoing Research Support

R00 GM107365, NIH/NIGMS

Liu (PI)

01/26/16 – 12/31/18

Single-Molecule Study of Cooperation Between Bacterial Gene Expression Machines

The goal of this project is to understand how essential bacterial gene expression machines interact with one another and how they act in concert in response to environmental changes.

Role: PI

[REDACTED]

[Redacted text block]

[Redacted text block]

[Redacted text block]

[Redacted text block]

Role: PI

Completed Research Support

K99 GM107365, NIH/NIGMS

Liu (PI)

09/15/13 – 12/31/15

Single-Molecule Study of Cooperation Between Bacterial Gene Expression Machines

The goal of this project is to study the interplay between three key molecular machines involved in bacterial gene expression, namely the RNA polymerase, the ribosome, and the transcription termination factor Rho.

Role: PI

PENDING

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

1. Human Subjects Section

Clinical Trial? Yes No

*Agency-Defined Phase III Clinical Trial? Yes No

2. Vertebrate Animals Section

Are vertebrate animals euthanized? Yes No

If "Yes" to euthanasia

Is the method consistent with American Veterinary Medical Association (AVMA) guidelines?

Yes No

If "No" to AVMA guidelines, describe method and provide scientific justification

.....

3. *Program Income Section

*Is program income anticipated during the periods for which the grant support is requested?

Yes No

If you checked "yes" above (indicating that program income is anticipated), then use the format below to reflect the amount and source(s). Otherwise, leave this section blank.

*Budget Period *Anticipated Amount (\$) *Source(s)

PHS 398 Cover Page Supplement

4. Human Embryonic Stem Cells Section

*Does the proposed project involve human embryonic stem cells? Yes No

If the proposed project involves human embryonic stem cells, list below the registration number of the specific cell line(s) from the following list: http://grants.nih.gov/stem_cells/registry/current.htm. Or, if a specific stem cell line cannot be referenced at this time, please check the box indicating that one from the registry will be used:

Specific stem cell line cannot be referenced at this time. One from the registry will be used.

Cell Line(s) (Example: 0004):

5. Inventions and Patents Section (RENEWAL)

*Inventions and Patents: Yes No

If the answer is "Yes" then please answer the following:

*Previously Reported: Yes No

6. Change of Investigator / Change of Institution Section

Change of Project Director / Principal Investigator

Name of former Project Director / Principal Investigator

Prefix:

*First Name:

Middle Name:

*Last Name:

Suffix:

Change of Grantee Institution

*Name of former institution:

RESEARCH STRATEGY (ESSAY)

Probing Symmetry Breaking in Epigenetic Inheritance: From Single Molecules to Systems Biology

PROJECT SCIENCE AREAS

6 MCB; 8 HIB

PROJECT DESCRIPTION

I. INTRODUCTION

Genome replication is fundamental to biological inheritance in all kingdoms of life [1]. Through a dramatic set of events, tightly packaged chromosomes are unraveled, duplicated, and reassembled into two new daughter strands. This remarkable task is accomplished by a multi-component cellular machinery known as the replisome [2]. In eukaryotes, the chromatin is organized into nucleosomes, in which 147 base pairs (bp) of DNA are wrapped around a histone octamer consisting of two copies each of the core histones H2A, H2B, H3, and H4. In addition to the genetic code stored in the DNA sequence, there exists in the genome another type of code in the form of chemical modifications of DNA or histones, such as methylation and acetylation. These *epigenetic* marks play a central role in governing chromatin architecture and gene expression [3]. Replication-coupled chromatin assembly allows the epigenetic state to be faithfully inherited across generations, while also enabling the offspring to remove undesirable marks to enhance fitness (**Fig. 1**) [4,5]. Disruption of the balance between fidelity and plasticity can compromise genome and epigenome integrity, and is associated with ageing and many pathological conditions including cancer [6-8].

Asymmetry in DNA replication:

It is well known that the genetic information is transmitted from one generation to the next via a semi-conservative mechanism. Whereas the end product is two identical copies of the genome, the underlying process is inherently asymmetric [9]. The synthesis of one strand, known as the leading strand, is carried out by DNA polymerase ϵ (Pol ϵ) in a largely continuous manner in the same direction as the replicative helicase CMG (Cdc45/MCM2-7/GINS) that unwinds the parental duplex. In contrast, the other strand known as the lagging strand is made discontinuously in the opposite direction, pieced together from a series of discrete DNA segments called Okazaki fragments that are 150 to 200 nucleotides (nt) in length. Each Okazaki fragment is initiated by Pol α -primase, which synthesizes an RNA primer and a short DNA, and further extended by Pol δ . Two adjacent Okazaki fragments are joined together after the coordinated events of strand displacement synthesis by Pol δ , cleavage by FEN1 and other nucleases, and ligation by DNA Ligase I [10].

Such asymmetry entails that the two daughter strands are differentially regulated. In fact, the mutational landscape has been found to be asymmetrical in cancer genomes [11] (we will revisit this

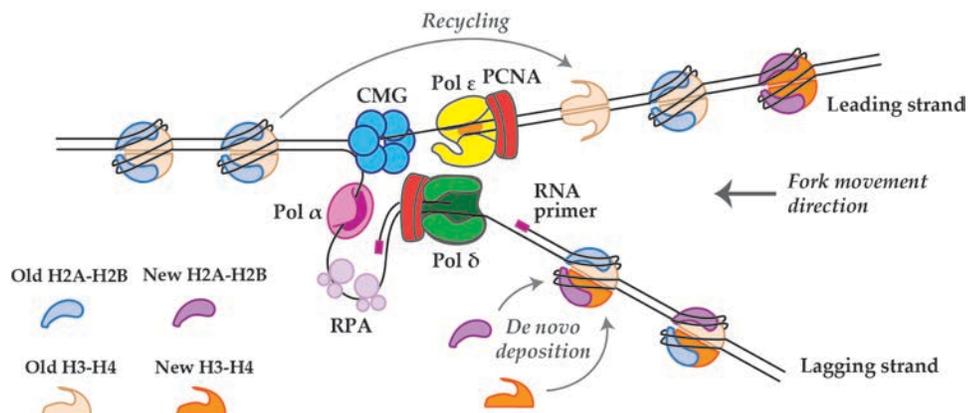


Figure 1 | Schematic of eukaryotic chromatin replication. Nucleosomes are reassembled on nascent DNA strands from a mixed pool of recycled parental histones and newly synthesized histones. Old nucleosome disassembly and new nucleosome assembly are mediated by histone chaperones (not shown). Maturation of nascent chromatin is facilitated by various chromatin remodeling and modifying enzymes (not shown).

issue later in the essay). In eukaryotic cells, the potential significance of symmetry breaking for biological inheritance is further underscored by the existence of histones—major carriers of epigenetic information.

Asymmetry in histone distribution?

How is the epigenetic information inherited? The histone content needs to double after replication in order to fully package the two nascent daughter strands. New nucleosomes are assembled from a mixture of old histones recycled from the parental nucleosomes and newly synthesized histones that are imported into the nucleus (Fig. 1). Note that new and old histones have distinct post-translational modification (PTM) characteristics. Thus, the nascent nucleosome composition, unlike the nascent DNA sequence, does not resemble its parental counterpart. Elucidating how daughter strands acquire a correct dosage of old and new histones is essential for understanding the maintenance of epigenetic memory.

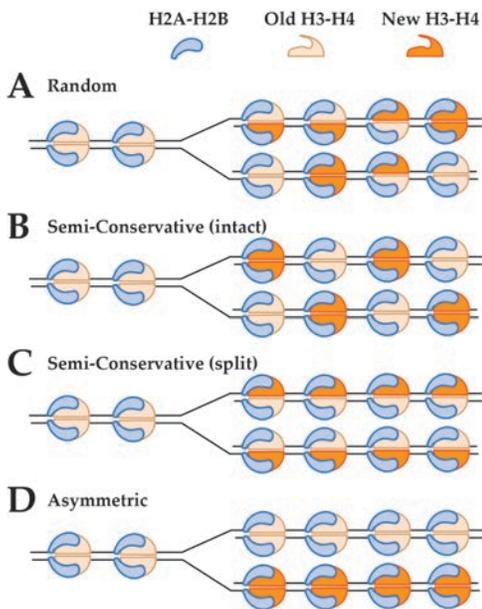


Figure 2 | Different models of histone distribution at the replication fork. Only the deposition scenarios for H3-H4 histones are illustrated because they contain the majority of epigenetic marks and have been more extensively studied. Another set of scenarios likely exists for H2A-H2B.

Several models have been proposed to describe histone distribution at the replication fork. In the first model, new and old histones are randomly deposited onto the leading and lagging strands (Fig. 2A). Because of the lack of organized inheritance pattern, the parental chromatin state is essentially overhauled in the wake of replication. The second model, known as the semi-conservative model, posits that parental H3-H4 histones are equally distributed between the two strands, either as intact tetramers (Fig. 2B) or as split dimers (Fig. 2C). The histone PTM profile is thus semi-inherited, albeit diluted. It requires additional machinery to fully restore the pre-replication chromatin state. Finally, in the asymmetric model (Fig. 2D), old histones are exclusively installed on one of the sister chromatids such that it faithfully copies the parental chromatin state, whereas the other sister incorporates an entirely new set of histones.

The asymmetric model, first put forth by Harold Weintraub in the 1970s [12], received early experimental support. The random and semi-conservative models have since gained favor, as investigators found no clear strand bias in histone deposition. The semi-conservative ones are particularly attractive because they can easily explain how the parental chromatin state is reinstated on daughter strands. Nonetheless, recent studies suggest that asymmetric histone segregation may indeed be at work in certain developmental and disease contexts, such as during the asymmetric division of germline and cancer stem cells [13-15]. Whether

these models are mutually exclusive and how they are dictated by the chromatin state is still much debated.

Our Goal: Detecting differential histone deposition on leading versus lagging strands.

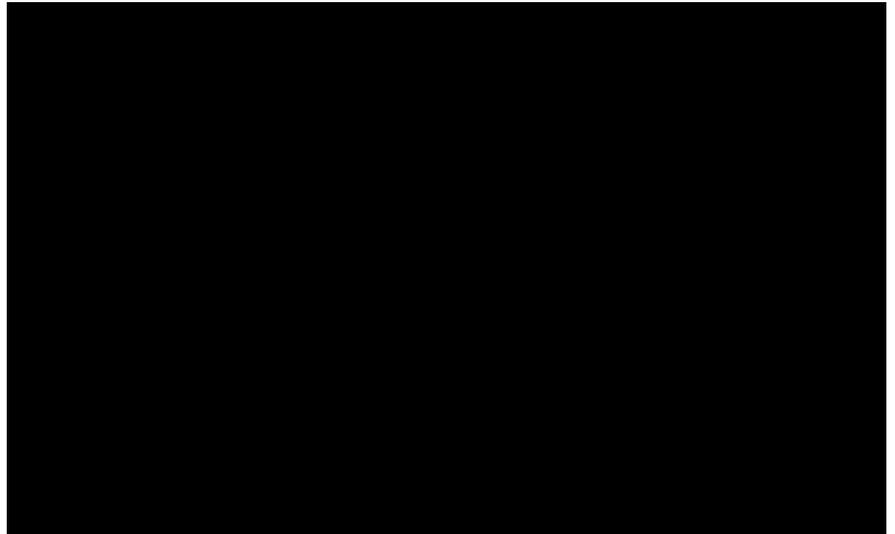
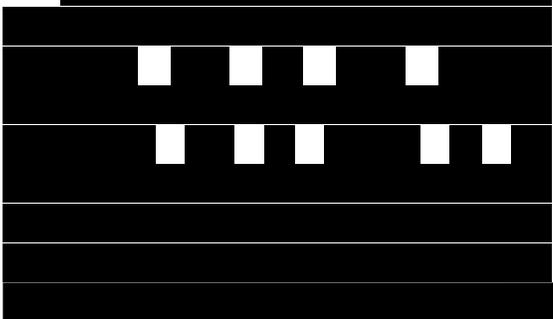
The crux of such contention is that direct observation of replication-coupled nucleosome disruption and reassembly is extremely challenging, due to the multitude of molecular players and dynamic events involved. In particular, there has been no good way to unambiguously determine whether a histone is deposited on the leading strand or the lagging strand. As a result, the exact mechanism of histone distribution and epigenetic inheritance remains elusive. In this New Innovator proposal, we will develop novel experimental tools to probe the pattern of histone deposition in a strand-specific manner at molecular, cellular, and systems levels. These integrated approaches promise to illuminate the interplay between the replisome and nucleosomes with unprecedented spatial and temporal resolution. The resultant experimental framework can be employed to dissect the specific and distinct roles of diverse *cis*- and *trans*-regulatory factors in epigenetic inheritance during development, physiology, and disease.

II. SINGLE MOLECULES

Single-molecule methods circumvent the need for synchronization, thereby providing a powerful means to elucidate the kinetic and structural features of stochastic and transient biomolecular events. These methods have been applied to studies of phage and bacterial DNA replication, revealing the intricate coordination between different components of the replisome [16]. However, the same level of understanding has not yet been achieved for eukaryotic replisomes due to their more complex architectures and the nucleosomal barriers that they have to deal with. Recently, several groups reported the reconstitution of the budding yeast *Saccharomyces cerevisiae* replisome with purified proteins [17-19], paving the way for a detailed biophysical characterization of eukaryotic replication.

Experimental Design:

Conventional single-molecule approaches either follow the processive movement of molecular motors by force manipulation or the compositional dynamics of molecular assemblies by fluorescence detection, but not both. To correlate the activity of DNA replication to the transfer of histones, we will utilize a next-generation single-molecule instrument that combines dual-trap optical tweezers, multi-color confocal microscopy, and automated microfluidics [20]. The experimental geometry is depicted in Fig. 3A.



This design enables us to, for the first time, visualize histone transfer from parental to daughter strand and unambiguously assign its destination. Possible scenarios include: exclusively leading strand, exclusively lagging strand, equal distribution between both strands, random deposition, and dissociation into solution. Not only can our assay distinguish among these scenarios, but it can also capture intermediate states and heterogeneous pathways in the transfer process. For instance, H2A-H2B dimers are generally thought to be incorporated into the nucleosome following the deposition of the (H3-H4)₂ tetramer, but this model has never been proved. We will follow the exact order of events during core histone incorporation. We will also examine whether the (H3-H4)₂ tetramer is split or remains intact during the transfer process, a topic that remains controversial [21].

Given that new histones must also be recruited to complement the old histone pool, it is conceivable that histone transfer is regulated by the concentration of free histones in solution, which we can easily control in this *in vitro* assay. Moreover, we will delineate the precise functions of specific replisome components that have been suggested to play a role in mediating replication-coupled nucleosome assembly, such as the lagging-strand synthesis machinery [22] and the single-stranded DNA binding protein RPA [23].

The disruption and reassembly of nucleosomes at the replication fork is aided by histone chaperones (*e.g.*, H3-H4 chaperones CAF-1 and ASF1, H2A-H2B chaperones FACT and NAP1), many of which are implicated in human disease [7,24]. We will interrogate how these factors regulate histone transfer patterns and fork stability. For example, by fluorescently label CAF-1, we will monitor the formation of transient CAF-1/H3-H4/DNA intermediates and assess their stoichiometry and dynamics during replication [25]. Furthermore, by adding factors individually or in combination, we can systematically dissect their functional synergy, antagonism, and redundancy.

Preliminary Results:

A single-molecule fluorescence/optical-tweezers combo instrument has been installed in my laboratory (the only one of its kind currently available in the United States). We have site-specifically labeled all four core histones and assembled nucleosomes (**Fig. 4**). In collaboration with the O'Donnell laboratory at Rockefeller, we have tailored the *in vitro* reconstituted yeast replisome for single-molecule experiments.

We have started to test replisome function at the single-molecule level (**Fig. 5A**), and have already observed the DNA unwinding activity of the replicative helicase CMG—a hallmark of replication fork movement.

These early results warrant further exploration in their own right. In fact, the basic properties of CMG as a molecular motor (step size, processivity, directionality, force-velocity relationship, *etc.*) are not at all well understood. For the purpose of this proposal, however, we will focus on investigating how CMG and the rest of the replisome negotiate with downstream nucleosomes and carry out the histone recycling process, in which CMG likely plays an integral part [26].

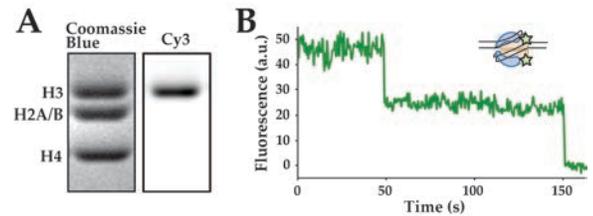


Figure 4 | Preliminary data: Site-specific histone labeling. (A) H3 histones are labeled with Cy3 via an engineered cysteine at position 33. Shown are the Coomassie Blue staining and Cy3 fluorescence signals of histone octamers in an SDS-PAGE gel. (B) Cy3-labeled octamers are reconstituted into nucleosomes on a biotinylated DNA template. The nucleosomes are then surface-immobilized and evaluated by total-internal-reflection fluorescence microscopy. The representative single-molecule fluorescence trajectory displays two photobleaching steps, indicating the presence of two Cy3-H3 molecules within the nucleosome. Site-specifically labeled H2A, H2B, and H4 have also been obtained (not shown).

Potential Pitfalls and Alternative Approaches:

The immediate availability of the instrument and the preliminary single-molecule results already obtained boost this assay's likelihood of success. A number of early hurdles, such as site-specific labeling of proteins, reconstitution of a fully functional replisome, and construction of the replication fork substrate, have been cleared. A few potential issues remain, which we address below.

The reconstituted yeast replisome has been shown to be capable of performing leading and lagging strand synthesis in bulk biochemical assays [17]. However, these experiments were not conducted in the presence of external force.

The core replisome may need additional factors, such as FACT, Mcm10, Mrc1, Csm3 and Tof1, to achieve *in vivo* replication rate and to efficiently read through nucleosomes [18]. These factors are either readily available from our collaborators or actively being purified in my lab. They will be included in the single-molecule assay if necessary. In addition, it may prove difficult to displace histones from the high-affinity "601" nucleosome positioning sequence, which is what we plan to use initially for nucleosome reconstitution. This potential problem can be alleviated by using a weaker nucleosome positioning sequence or by adjusting the salt concentration to weaken the histone-DNA interaction.

The speed of a eukaryotic replisome (1000-2000 bp/min) suggests that it may take up to 30 seconds for a fork to travel through a nucleosome. The lifetime of commonly used dyes can routinely reach several minutes in our experimental setting (Fig. 5D), which is enough time for us to monitor the entire histone transfer process. If photobleaching still poses a problem, fluorophores with improved photostabilities are available [27].

III. SINGLE CELLS

The *in vitro* biophysical approach described above allows us to examine the behaviors of individual molecules during chromatin replication with exquisite control over the reaction conditions. To test whether the observed histone distribution patterns can be recapitulated inside the cell, next we seek to visualize histone deposition at individual replication forks in single yeast cells. Single-molecule fluorescence microscopy has been employed to observe replisome dynamics in bacteria, which typically contain one origin of replication and at most two active replication forks [28,29]. By contrast, yeast cells contain multiple origins per chromosome. At the peak of DNA replication during S phase, there are an estimated 200-300 replication forks all confined in a nucleus of 1-2 μm in diameter, preventing them from being individually resolved by conventional light microscopy. The advent of super-resolution nanoscopy broke the diffraction limit. Spatial resolutions of ~ 20 nm can now be achieved [30]. Super-resolution techniques based on fluorophore photoswitching, such as STORM and PALM, have been utilized to image nucleosomes in

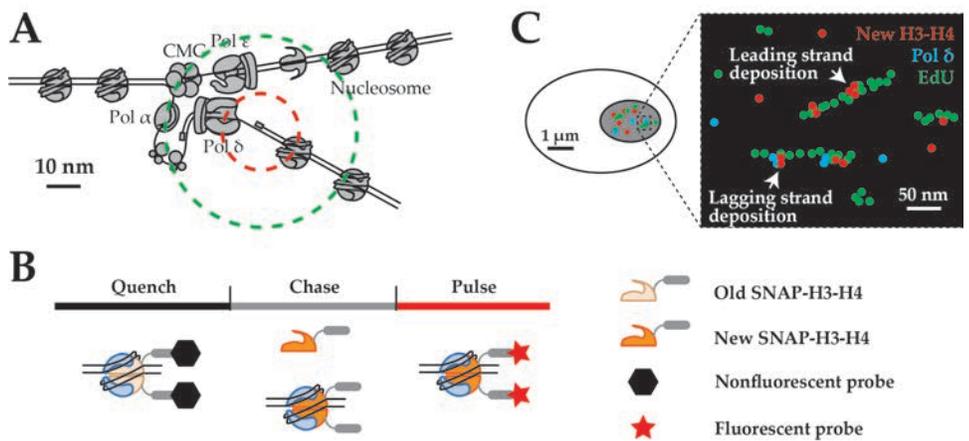


Figure 6 | Visualizing strand-specific new histone deposition in single cells using super-resolution microscopy. (A) Estimated dimension of the budding yeast replication fork. Circles of a 20-nm and 50-nm diameter are shown in red and green, respectively. (B) A quench-chase-pulse scheme to specifically label newly synthesized histones. See main text for details. (C) Putative super-resolution images illustrating two scenarios in which new H3-H4 is deposited on the leading or lagging strand. Nascent DNA is labeled by fluorescently tagged EdU.

eukaryotic nuclei [31,32]. However, due to the high copy number of histones, imaging their localizations and dynamics at replication foci remains a major challenge. Here we will make the first such attempt by combining super-resolution technology with a fluorescent pulse-labeling strategy [33].

Experimental Design:

The dimension of the yeast replication fork suggests that nucleosomes assembled on the leading strand can be potentially distinguished from those on the lagging strand by super-resolution imaging (**Fig. 6A**). To differentiate between old and new histones,

Meanwhile, we will concomitantly pulse-label nascent DNA with the thymidine analog EdU (5-ethynyl-2'-deoxyuridine) that can be fluorescently tagged through “click” reaction. A 1-minute EdU incubation can label 1000-2000 bp of nascent DNA (~125-250 nm in chromatin length), which can load 5-10 nucleosomes. Moreover, we will mark the leading or lagging strands with their characteristic bound proteins (*e.g.*, Pol ϵ and CMG for the leading strand, Pol δ and Pol α for the lagging strand), either by immunostaining or by fusion with photoactivatable fluorescent proteins (*e.g.*, PA-mCherry and mEos2). By obtaining dual-color STORM images of new histones and strand-specific proteins and then analyzing their co-localization patterns, we can evaluate whether and to what extent newly synthesized histones are preferentially or randomly deposited (**Fig. 6C**).

Since cell-permeable dyes are used to label SNAP-tagged histones, this assay is compatible with live-cell imaging. The temporal resolution for live-cell STORM (with a 20-nm spatial resolution) is on the order of 0.5 seconds [35], in principle sufficient for resolving histone dynamics at the fork where a new nucleosome is loaded every few seconds.

The strength of this assay lies in its ability to detect differentially distributed histones in single cells. It will be of great interest to examine whether there exists cell-to-cell variation in the distribution patterns. When combined with DNA barcoding technology, this assay can report how the histone inheritance mechanism evolves along a cell lineage [36]. Further, by using chromosomal insertion [37] or CRISPR/dCas9 tools [38] to mark specific genomic regions, we can distinguish histone dynamics at different replication loci.

Potential Pitfalls and Alternative Approaches:

It was reported that SNAP-tagged histones are correctly incorporated into nucleosomes in mammalian cells [34]. Given the highly conserved nature of core histones, we expect that the same holds true for yeast and will verify this with biochemical assays. We will also check the expression level of SNAP-tagged histones by western blot and make sure that they are not grossly overexpressed, which may affect the deposition pattern.

The quality of STORM imaging depends critically on the choice of fluorophores. We will test multiple labeling strategies and cross-examine them in order to eliminate potential labeling artifacts. Generally speaking organic dyes have superior photoswitching properties compared to fluorescent proteins and hence are our preferred choice. In the case that further improvement in signal-to-noise is required, we will add a triton extraction step after pulse-labeling to remove soluble histones, albeit sacrificing the live-cell imaging capacity.

It was proposed that replisomes are organized in “replication factories” in eukaryotic nuclei [37]. In budding yeast there may exist 8-20 factories (each ~200 nm in size) per cell; each factory may contain 2-10 replisomes. These numbers remain to be validated and our super-resolution assay should provide a more definitive answer.

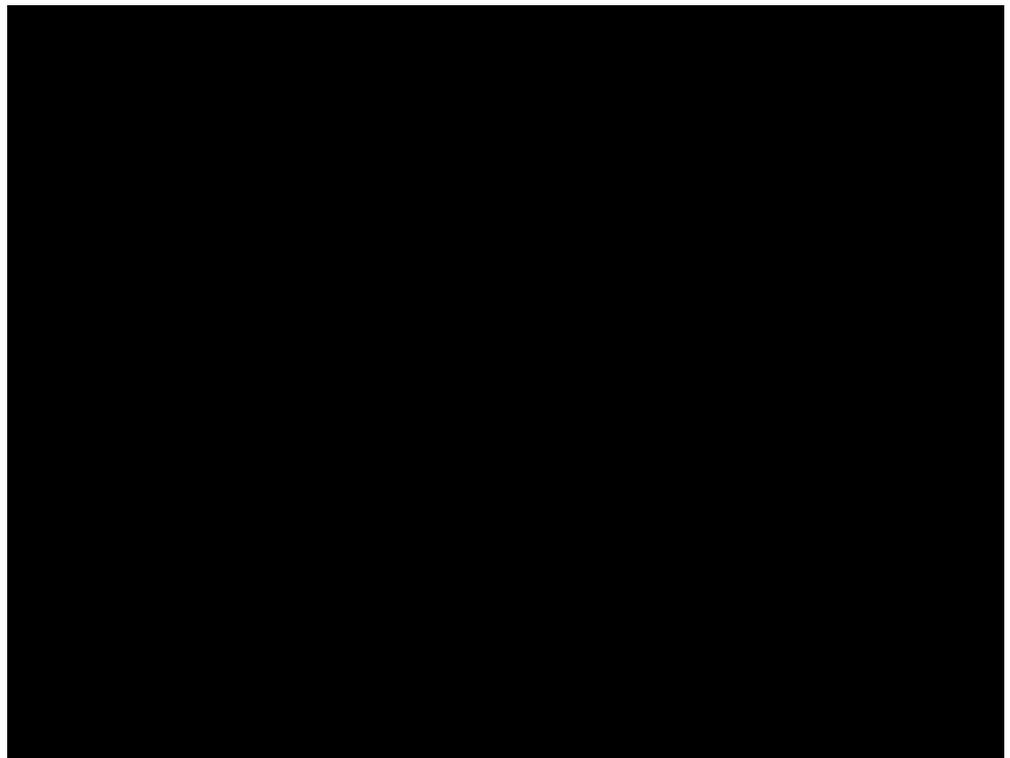
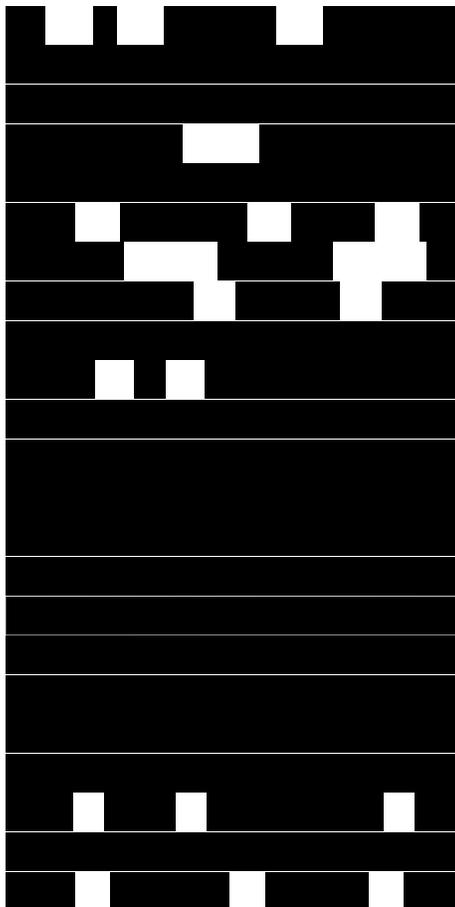
Regarding the choice of leading/lagging-strand-specific proteins, an important consideration is their occupancy density and labeling efficiency. Besides the polymerases, a few more options are available for marking the lagging strand, such as PCNA, FEN1, and DNA Ligase I. There are fewer options for marking the leading strand. If necessary we will adopt alternative methods such as the proximity ligation assay (PLA) [39], which requires cell fixation but affords single-molecule sensitivity.

We have thus far focused on imaging *de novo* deposited histones. While technically more difficult, it is also possible to track both pre-existing histones and newly synthesized ones through the usage of chemically orthogonal tags.

IV. GENOMICS AND PROTEOMICS

The assays proposed above promise to yield high-resolution pictures and movies of histone transfer at the replication fork *in vitro* and *in vivo*. Here we plan to interrogate strand-specific histone deposition at a systems level. Replication-independent histone turnover, which represents the other major mechanism for epigenome maintenance, has been investigated genome-wide [40]. Replication-coupled chromatin assembly has also begun to be understood thanks to the recently developed genomic methods that monitor how the nucleosome landscape emerges and matures on nascent DNA [41,42]. However, these methods either cannot differentiate between leading and lagging strands, or can only do so in regions immediately adjacent to high-usage replication origins. Taking advantage of the fact that eukaryotes use distinctive polymerases to perform leading versus lagging strand synthesis, we will establish a new method that enables us to confidently assign the strand identity of newly assembled nucleosomes behind a replication fork throughout the entire genome.

Experimental Design:



How can we map strand-specific nucleosome positions? EdU is commonly used to label nascent chromatin and isolate newly formed nucleosomes on both strands [41,42].

The products are then prepared for high-throughput sequencing, which yields the nucleosome profile on the newly replicated chromatin. By comparing the profiles for different strains, we will gain insight into how nucleosomes are differentially assembled on leading versus lagging strands at various genomic loci.

Methods such as iPOND and nascent chromatin capture (NCC) utilize EdU or biotin-dUTP to isolate nascent chromatin followed by immunoblotting or mass spectrometry [47,48]. By contrast,

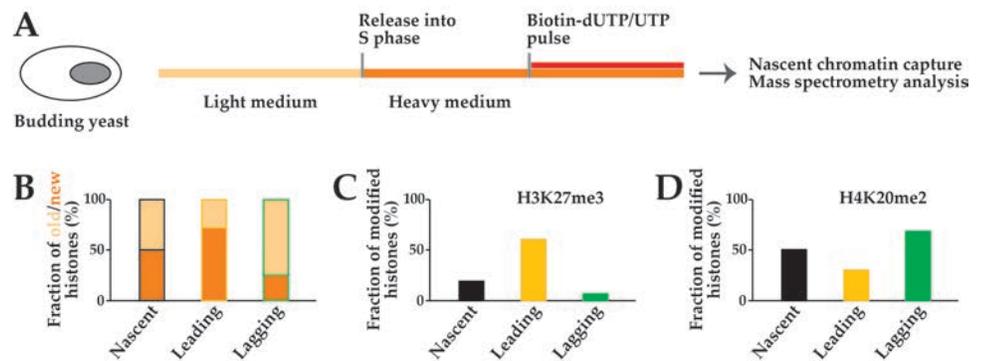


Figure 8 | Proteomic analysis of histones on newly replicated chromatin. (A) Protocol for SILAC-based nascent chromatin capture. (B) Putative distributions of old vs. new histones on all nascent DNA, leading DNA, and lagging DNA. (C-D) Putative fractions of histones on indicated strands that harbor specific PTMs, such as H3K27me3 (C) and H4K20me2 (D), revealed by quantitative mass spectrometry.

The nascent chromatin as a whole is expected to contain roughly equal amounts of old and new histones. But they may be asymmetrically distributed between leading and lagging strands (Fig. 8B). This assay can be further applied to depict the differential histone PTM profiles on specific nascent strands and, when combined with

Potential Pitfalls and Alternative Approaches:

There are a few ways to potentially reduce the required cell number. First, we will try to synchronize cells to near homogeneity such that they can be simultaneously released into the S phase.

These ensemble methods report the average behavior of a large population of cells, thus blind to any inherent heterogeneity in histone deposition. Nonetheless, their ability to unveil the genomic positions and chemical compositions of new nucleosomes makes them complementary to the single-molecule and single-cell assays.

V. OUTLOOK

The experimental platforms developed here will pave the way for studying the mechanisms by which proper histone variants and PTMs are differentially incorporated into two sister chromatids [49,50]. We are also interested in examining the redistribution patterns of transcription factors, whose occupancy profiles, together with the nucleosome landscape, provide instructions for gene expression programs. Transcription factors generally exist in much lower copy numbers in the cell than histones, hence more likely to bind asymmetrically and cause two gene copies to be differentially transcribed. Finally, due to the highly conserved nature of the replisome, nucleosome, and chromatin assembly factors, lessons learned from yeast will be of relevance to all eukaryotes. The roles that asymmetric inheritance plays in cell fate decision during development, in replicative exhaustion during ageing, and in epigenetic alteration during tumorigenesis are fascinating subjects to pursue in future studies.

INNOVATIVENESS

Conceptually, this proposal concerns a foundational question in biology: how is epigenetic information inherited? Replication-coupled nucleosome assembly must occur on two separate daughter strands, raising the possibility that the two offspring may adopt distinct transcriptional programmes and cell fates. Therefore, symmetry breaking may be essential for driving epigenome evolution and creating diverse cell identities in multicellular organisms. No existing method focuses on probing whether histones—together with the epigenetic marks that they carry—are deterministically or stochastically distributed between the leading and lagging strands. Finding the answer to this question, which has eluded scientists for decades, is precisely the motivation and pivot of this proposal. If successful, this project will present a new paradigm for understanding the establishment, maintenance, and modulation of chromatin states and epigenetic memory.

Technically, this proposal features a series of first-of-its-kind experimental designs. We use cutting-edge single-molecule tools to make correlative fluorescence-force measurements while exerting controlled

The confluence of diverse assays boosts the chance of making transformative discoveries and mitigates the risk of each individual assay. These assays can also be adapted to study a broad range of chromatin-based processes such as transcription, DNA repair, and centromere/telomere metabolism.

INVESTIGATOR QUALIFICATIONS

I have the right scientific training for this project. I have more than a decade of experience in single-molecule methodology [51,52]. Having done my graduate work in Xiaowei Zhuang's lab, I am also familiar with super-resolution microscopy. I am fully aware of the strengths and limitations of these tools, and know how to tailor them to specific questions and further develop them for novel applications. **I have a track record of conducting innovative research.** This is demonstrated by the publications and awards listed in my Biosketch. I do not hesitate to leave my comfort zone and break new ground. For instance, in my K99/R00 grant I proposed to use optical tweezers to study transcription-translation coupling in bacteria. During the course of my research I realized that a higher-throughput approach would be extremely profitable. So I learned from scratch and developed a novel RNA-sequencing-based method, which is now producing exciting results on the life

cycle of RNA. **I am at an ideal place to execute this project.** The Rockefeller University gives junior faculty maximum freedom to pursue high-risk, high-reward research. I have the privilege to be colleague with world leaders in DNA replication (Michael O'Donnell) and epigenetics (David Allis). They have already provided us with tremendous intellectual and technical help, and will continue to be invaluable sources of expertise.

SUITABILITY FOR THE NEW INNOVATOR AWARD PROGRAM

Work in my lab thus far mainly involves *in vitro* single-molecule experiments. This project represents our first attempt to directly link biophysical principles with cellular events and systems biology. While this is an ambitious undertaking, the potential implications of this multi-pronged approach for one of the most fundamental biological problems make it suitable for the New Innovator Program. The first steps are being made and the preliminary data are promising. If any of the proposed assays is successful, this project will open up a whole new field of inquiry that can be morphed into an R01-type project in the future.

STATEMENT OF RESEARCH EFFORT COMMITMENT

My chief responsibility at Rockefeller is to carry out a research program. **If chosen to receive a New Innovator Award, I will commit at least 25% of my research effort to this project.**

BIBLIOGRAPHY

1. Kornberg A, Baker TA (2005) *DNA Replication*, 2nd Edition. University Science Books, California.
2. O'Donnell M, Langston L, Stillman B (2013) *Cold Spring Harb. Perspect. Biol.* **5**, a010108.
3. Allis CD, Jenuwein T (2016) *Nat. Rev. Genet.* **17**, 487.
4. MacAlpine DM, Almouzni G (2013) *Cold Spring Harb. Perspect. Biol.* **5**, a010207.
5. Margueron R, Reinberg D (2016) *Nat. Rev. Genet.* **11**, 285.
6. Alabert C, Groth A (2012) *Nat. Rev. Mol. Cell Biol.* **13**, 153.
7. Burgess RJ, Zhang Z (2013) *Nat. Struct. Mol. Biol.* **20**, 14.
8. Gaillard H, Garcia-Muse T, Aguilera A (2015) *Nat. Rev. Cancer* **15**, 276.
9. Snedeker J, Wooten M, Chen X (2017) *Annu. Rev. Cell Dev. Biol.* [Epub ahead of print] doi:10.1146/annurev-cellbio-100616-060447.
10. Balakrishnan L, Bambara RA (2013) *Cold Spring Harb. Perspect. Biol.* **5**, a010173.
11. Haradhvala NJ *et al.* (2016) *Cell* **164**, 538.
12. Weintraub H (1973) *Cold Spring Harb. Symp. Quant. Biol.* **38**, 247.
13. Nakano S, Stillman B, Horvitz HR (2011) *Cell* **147**, 1525.
14. Tran V *et al.* (2012) *Science* **338**, 679.
15. Lan X *et al.* (2017) *Nature* [Epub ahead of print] doi:10.1038/nature23666.
16. Stratmann SA, van Oijen AM (2014) *Chem. Soc. Rev.* **43**, 1201.
17. Georgescu RE *et al.* (2015) *eLife* **4**, e04988.
18. Yeeles JT *et al.* (2017) *Mol. Cell* **65**, 105.
19. Devbhandari S *et al.* (2017) *Mol. Cell* **65**, 131.
20. Hashemi Shabestari M *et al.* (2017) *Methods Enzymol.* **582**, 85.
21. Xu M *et al.* (2010) *Science* **328**, 94.
22. Smith DJ, Whitehouse I (2012) *Nature* **483**, 434.
23. Liu S *et al.* (2017) *Science* **355**, 415.
24. Ransom M, Dennehey BK, Tyler JK (2010) *Cell* **140**, 183.
25. Mattioli F *et al.* (2017) *eLife* **6**, e22799.
26. Huang H *et al.* (2015) *Nat. Struct. Mol. Biol.* **22**, 618.
27. Zheng Q *et al.* (2014) *Chem. Soc. Rev.* **43**, 1044.
28. Reyes-Lamothe R, Sherratt DJ, Leake MC (2010) *Science* **328**, 498.
29. Mangiameli SM *et al.* (2017) *eLife* **6**, e19848.
30. Huang B, Babcock H, Zhuang X (2010) *Cell* **143**, 1047.
31. Ricci MA *et al.* (2015) *Cell* **160**, 1145.
32. Nozaki T *et al.* (2017) *Mol. Cell* **67**, 10.
33. Jansen LE *et al.* (2007) *J. Cell Biol.* **176**, 795.
34. Clément C *et al.* (2016) *Methods Enzymol.* **573**, 97.
35. Jones SA *et al.* (2011) *Nat. Methods* **8**, 499.
36. Frieda KL *et al.* (2017) *Nature* **541**, 107.
37. Kitamura E, Blow JJ, Tanaka TU (2006) *Cell* **125**, 1297.
38. Chen B *et al.* (2013) *Cell* **155**, 1479.
39. Petruk S *et al.* (2012) *Cell* **150**, 922.
40. Dion MF *et al.* (2007) *Science* **315**, 1405.
41. Ramachandran S, Henikoff S (2016) *Cell* **165**, 580.
42. Vasseur P *et al.* (2016) *Cell Rep.* **16**, 2651.
43. [REDACTED]
44. [REDACTED]
45. [REDACTED]
46. [REDACTED]
47. Sirbu BM, Couch FB, Cortez D (2012) *Nat. Protoc.* **7**, 594.
48. Alabert C *et al.* (2014) *Nat. Cell Biol.* **16**, 281.
49. Talbert PB, Henikoff S (2017) *Nat. Rev. Mol. Cell Biol.* **18**, 115.
50. David Y, Muir TW (2017) *J. Am. Chem. Soc.* **139**, 9090.
51. Liu S *et al.* (2008) *Science* **322**, 1092.
52. Liu S *et al.* (2014) *Cell* **157**, 702.

AUTHENTICATION OF KEY BIOLOGICAL AND/OR CHEMICAL RESOURCES

Beads

Polystyrene beads for optical tweezers experiments are commercially available (Spherotech). The stock solution is diluted prior to use according to the manufacturer's manual. To control for size variation, each optically trapped bead is subject to visual inspection and its Brownian motion profile is recorded. Beads with profiles that deviate from the theoretical prediction based on the size distribution data from the vendor are discarded without further analysis.

DNA Constructs

The size and purity of ligation products is examined by gel electrophoresis. DNA tethered between two optically trapped beads is further evaluated by its force-extension curve. Only those with the correct length and force response as predicted by the worm-like-chain model are used for single-molecule imaging and manipulation.

Protein Constructs

Proteins conjugated with short peptide tags or fluorescent proteins [REDACTED] are verified by changes in gel migration and/or western blot.

Fluorophores

Commercial fluorophores for single-molecule and single-cell imaging are shielded from light exposure during storage. Fluorescently labeled samples are periodically evaluated (every three months) by their absorption spectra with a spectrophotometer. To determine the expected localization precision for super-resolution imaging, we will characterize the photoswitchable fluorophores in terms of their on/off duty cycles, number of photoswitching cycles and photon counts, and then compare these parameters to published standards [Dempsey *et al.*, *Nat. Methods* 8, 1027 (2011)].

Yeast Strains

[REDACTED]

Antibodies

All antibodies used in this study, such as those against histones, histone variants, and post-translational modifications, are purchased from well-recognized vendors (Cell Signaling Technology, Life Technologies, Abcam, *etc.*). They are diluted according to the manufacturer's manual. Different antibody lots are independently evaluated by western blot or immunoprecipitation, and a new dilution is selected if necessary. The specificity of antibodies is validated by lack of western blot bands or immunostaining in negative controls. Fluorescently tagged antibodies are also assessed for their labeling ratio at regular intervals (every six months).

Chemicals

Thymidine and uridine analogs used in this study, [REDACTED] are commercially available (Invitrogen, Thermo Scientific, Sigma, *etc.*). Their quality is evaluated by relevant functional assays [REDACTED]. Their purity and stability is assessed over time by high-performance liquid chromatography and, if necessary, liquid chromatography-mass spectrometry.