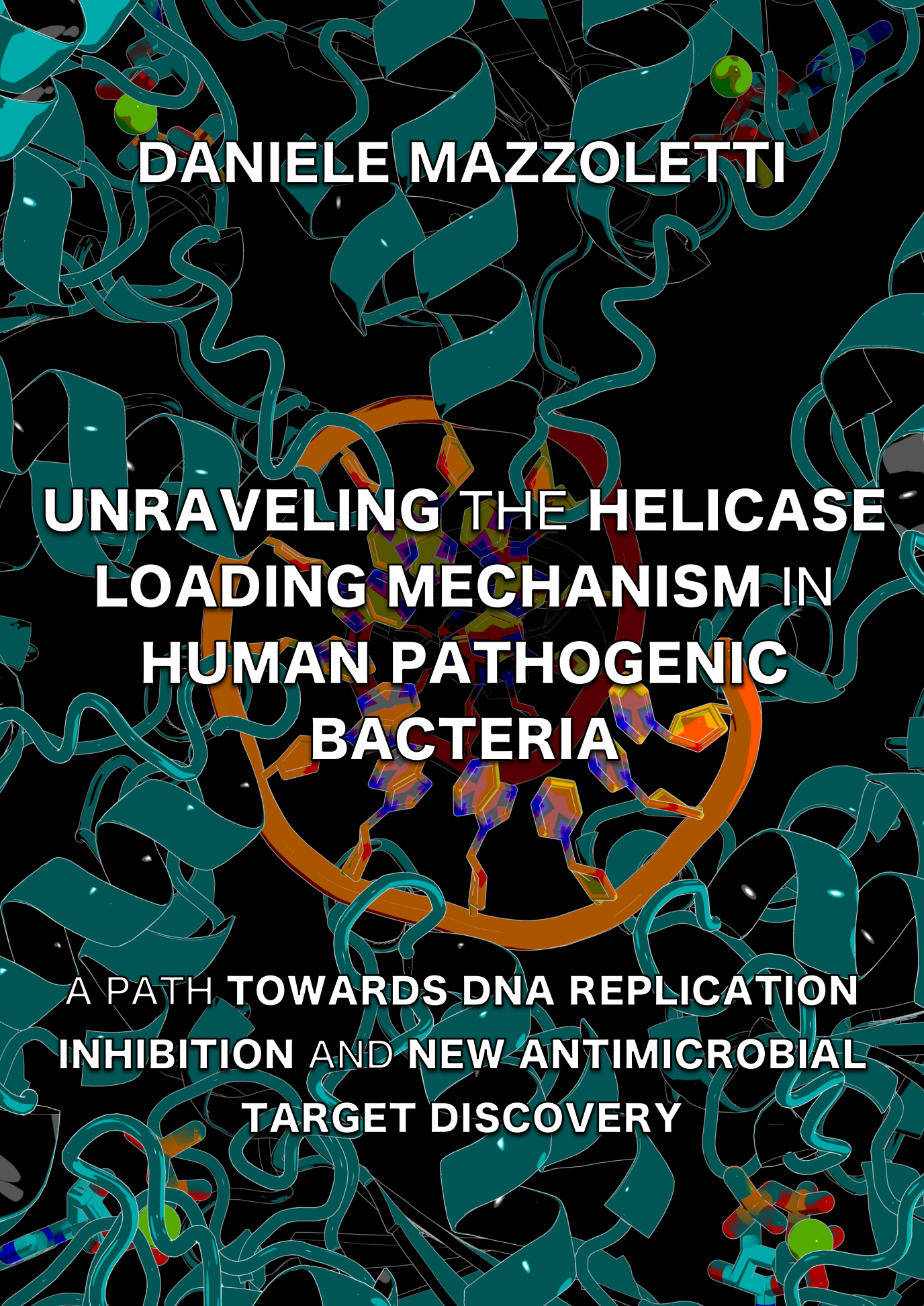


The background of the cover is a complex 3D molecular structure. It features a dense network of teal-colored ribbons, which represent protein chains, intertwined and looping across the entire frame. In the center, there is a prominent orange ribbon that forms a large, circular loop. Attached to this orange loop and scattered throughout the teal network are various small, multi-colored geometric shapes, possibly representing ligands or specific amino acid side chains. The overall composition is highly detailed and scientific in nature.

DANIELE MAZZOLETTI

**UNRAVELING THE HELICASE
LOADING MECHANISM IN
HUMAN PATHOGENIC
BACTERIA**

**A PATH TOWARDS DNA REPLICATION
INHIBITION AND NEW ANTIMICROBIAL
TARGET DISCOVERY**

Università del Piemonte Orientale
"Amedeo Avogadro"

Dipartimento di Scienze del Farmaco

Dottorato di Ricerca in "Drug Innovation"
XXXVII ciclo, a.a. 2021 – 2024

Unraveling the helicase loading mechanism in human
pathogenic bacteria: a path towards DNA replication
inhibition and new antimicrobial target discovery

Daniele Mazzoletti



Supervisor:

Professor Riccardo Miggiano



PhD Program Coordinator:

Professor Gian Cesare Tron

SUMMARY

DECLARATION AND AUTHORIZATION TO ANTIPLAGIARISM DETECTION

Chapter 1

Introduction

Chapter 2

Thesis Outline

Chapter 3

Published Results

Gao N, Mazzeletti D et al. **"DnaB and DciA: mechanisms of helicase loading and translocation on ssDNA"** *Nucleic Acid Research*. 2025

Chapter 4

Unpublished Results pt.1:

Mazzeletti D et al. **"Unveiling unprecedented binding modes of DnaA and DciA to the *Vibrio cholerae* DnaB using single-particle cryo-EM"**

Unpublished Results pt. 2:

Integrative structural biology and biophysics approaches for the investigation of mycobacterial proteins involved in the DNA replication initiation process.

- The DH-LH element of DnaB is a unique target for designing new drugs inhibiting the DNA replication initiation in *Mycobacterium tuberculosis*.
- SAXS and NMR-based structural investigation of the mycobacterial helicase loader DciA.

Chapter 5

Discussion, Conclusions, and Future Perspectives

List of publications

Curriculum Vitae

Acknowledgements

DECLARATION AND AUTHORISATION TO ANTIPLAGIARISM DETECTION

The undersigned **Daniele Mazzeletti** student of the *Drug Innovation* Ph.D course
(XXXVII Cycle)

declares:

- to be aware that the University has adopted a web-based service to detect plagiarism through a software system called "Turnit.in",
- his/her Ph.D. thesis was submitted to Turnit.in scan and reasonably it resulted an original document, which correctly cites the literature;

acknowledges:

- his/her Ph.D. thesis can be verified by his/her Ph.D. tutor and/or Ph.D Coordinator in order to confirm its originality.

Date: **Novara, 20/01/2025**

Signature:

A handwritten signature in black ink, reading 'Daniele Mazzeletti', is written below the signature line.

Chapter 1

INTRODUCTION

TABLE OF CONTENTS

1. TUBERCULOSIS: FROM THE PHYSIOPATHOLOGY TO THE CLINICAL TREATMENT	1
2. SEEKING FOR NEW CLINICAL TARGETS IN THE TUBERCULOSIS TREATMENT: THE DNA METABOLISM	6
2.1 UNLOCKING THE DNA REPLICATION MECHANISM AS PHARMACOLOGICAL TARGET INFECTIOUS DISEASES	9
2.1.1 A BRIEF RECAP ON THE PRESENT-DAY SCENARIO	9
2.1.2 THE DNA REPLICATION INITIATION MECHANISM: A COMPARISON BETWEEN THE EUKARYOTIC AND PROKARYOTIC WORLD	11
2.2 THE INITIATION PHASE OF PROKARYOTIC DNA REPLICATION AS AN UNPRECEDENTED ANTITUBERCULAR CHEMOTHERAPY TARGET	16
2.2.1 DEVELOPING BIOACTIVE PEPTIDOMIMETICS AS POTENTIAL PHARMACOLOGICAL TOOL INHIBITING THE DNA REPLICATION INITIATION	18

INTRODUCTION

Owing to the COVID-19 pandemic soared in March 2020, the past few years have been strongly marked by the spreading awareness on infectious diseases, conditions elicited by harmful pathogenic microorganisms that get in contact with the human body through a number of different sources, such as food, water or, more in general, to organisms in the environment. Many infectious diseases can be directly or indirectly passed from one person to another. The World Health Organization (WHO) classified COVID-19 as third in the leading causes of death worldwide in 2020, list that includes five infectious diseases in the first ten positions when considering low-income countries, thus acknowledging these pathologies as extremely life-threatening.

1. TUBERCULOSIS: FROM THE PHYSIOPATHOLOGY TO THE CLINICAL TREATMENT

Among the most wide-spread infectious diseases and characterized by one of the highest health and social burden, Tuberculosis (TB) is a respiratory disorder caused by the pathogen *Mycobacterium tuberculosis* (MTB, *Mt*) related to 1.3 millions of death in 2022 (Figure 1A), making it the second leading cause of death only after COVID-19 (WHO, Global Tuberculosis Report 2023)¹. This human pathogenic bacterium is member of the wider family named "*Mycobacterium tuberculosis* complex" or MTBC, able to cause Tuberculosis in different organisms, and has been classified as Gram-positive bacillus with the ability of persisting in the human host up to transmission from an individual with active disease to another. Specifically, they intrude the lower pulmonary area, where alveolar macrophages

surround and phagocyte them, with the goal of annihilating MTB bacilli. However, this is the environment in which the pathogenic bacteria initiate the proliferative activity, building up a silent, asymptomatic and non-transmissible infection. In particular, the pulmonary TB is characterized by the so-called granuloma (Figure 2), a unique structure made of a necrotic core full of bacteria surrounded by layers of infected epithelioid macrophages and other host immune cells, aiming to contain the infection that, in fact, may persist for decades, defining the latency stage. Despite defending the human host from mycobacteria spreading, this immune cell coating creates a physical barrier to antitubercular drugs, altering the pathogen's susceptibility to the treatment. Nevertheless, a plethora of virulence factor is secreted by MTB bacilli to evade the host immune response, so that efficient bacteria replication surges again². And so, under altered health circumstances, like the insurgence of secondary diseases (HIV/AIDS, Figure 1B) or even advancing age, MTB bacilli can wake-up from latency stage to active infection favoring well-known clinical symptoms.

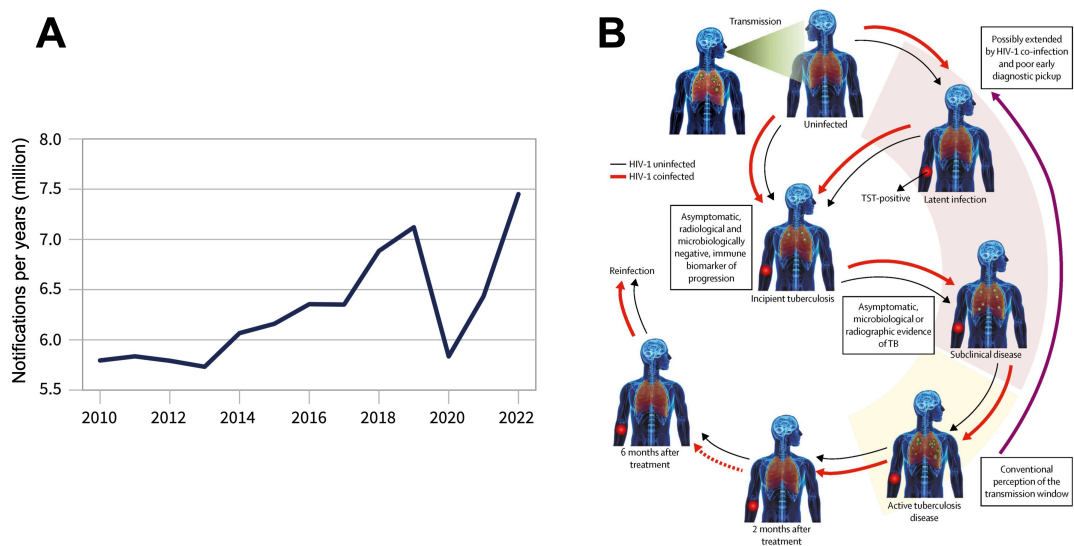


Figure 1. A) Global trend in TB case notifications in the 2010-2022 timeframe, highlighting the clinical drawbacks established over COVID-19 pandemic, started in 2020¹. **B)** Tuberculosis infection pathway and HIV-related coinfection, leading to severe disease progression³⁶.

The contagiousness is the outcome of *Mycobacterium tuberculosis* re-activation³. Prompt diagnosis and high-efficacy treatments adopting optimized administration regimens are functional to hamper the selection of drug-resistant MTB strains, that are, however, the major current obstacle in patients' survival. On one hand, the six-month long treatment with first-line antitubercular drugs, such as isoniazid, rifampicin, ethambutol, and pyrazinamide, used in combination, is key to halt the fast-growing bacilli and to minimize the progression of drug resistance. Yet, inappropriate adherence to the medication protocol leads to the selection of MTB strains defined rifampicin-resistant (RR), multidrug-resistant (MDR), and extensively-resistant (XDR). Then, treatment with second-line antibiotics, based on the drug susceptibility testing (DST) result, is enforced to manage harsh health conditions, but it is oftentimes associated with moderate to severe side effects. Common regimens include from fluoroquinolones, like moxifloxacin and levofloxacin, to cyclosporine, ethionamide and the more recent bedaquiline, a novel antitubercular medicinal product that proved to be significantly effective against MDR-TB and XDR-TB, finally allowing a major breakthrough in TB control, even though, once again, its inappropriate use may lead to future drug resistance.

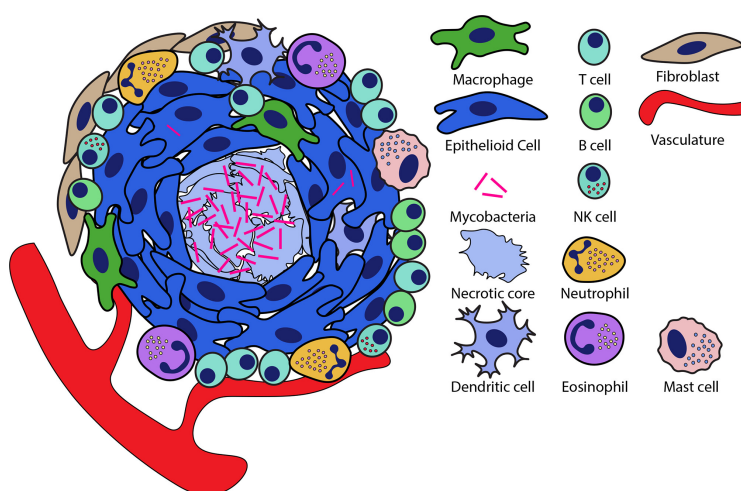


Figure 2. Detailed assembly of the granuloma structure, as a result of the host immune response for the tubercular bacilli restraint during the latent phase of Tuberculosis infection².

Second-line treatment efficacy is a matter of a balanced administration regime, able to lessen severe side effects, ensuring patient compliance⁴.

On these bases, advanced novel active compounds against TB are being developed and currently tested in clinical trials. As the linezolid, the first among the oxazolidinone family, is nowadays crucial for resistant TB treatment, second-generation bedaquiline analogs showcased activity against already-existing bedaquiline-resistant strains⁵. At the same time, in 2008, Singh *et al.* published a study on the efficacy of PA-824, also known as pretomanid, on both latent and active Tuberculosis, as monotherapy or in combination, introducing a new mode of action based on the nitric oxide release⁶. Valuable anti-TB candidates come also from the advancements in drug repurposing. Indeed, the tricyclic carbapenem sanfetrinem demonstrated its bactericidal effect in a Phase II clinical study. Besides, the seek for new modes of action for antitubercular drug leverages on phenotypic screening, MTB cell wall pathways remodeling, respiratory chain targeting for inhibiting the energy production, host-directed therapies to weaken the granuloma structure, and bacterial targeted protein degradation based on the new 'BacPROTACs', which involve the Clp protease system as antibacterial target⁴. And, all in all, this is a non-exhaustive list.

Despite great efforts employed and significant achievements reached in the TB control at the level of both the hospital care, with innovative drugs and treatment regimes, and the social containment measures and prevention, using inactivated vaccines, the seek for completely new druggable targets is of utmost importance. The reckless overuse of antibiotics to fight infectious diseases, or misuse against viral infections, in the last century, paved the way for the generation of new hyper drug-resistant "superbugs", whose treatment is calling for multitarget approaches⁷ and engineered phages⁸, given that even the most recent antibiotics,

based on the conventionally-known mechanisms of action, are ineffective. And MTB is no different. As the NIH expert in bacterial and fungal diseases Dr. Dennis Dixon said, "Bacterial infections that were treatable for decades are no longer responding to antibiotics, even the newer ones".

2. SEEKING FOR NEW CLINICAL TARGETS IN THE TUBERCULOSIS TREATMENT: THE DNA METABOLISM

Trying to keep ahead of additional multidrug-resistant bacteria strains requires focusing on pharmacological targets never exploited in antibiotic-based chemotherapy treatments. In this scenario, the DNA metabolism encompasses biological pathways, namely the nucleobases synthesis, the DNA replication, and the DNA repair, that are essential for the survival of every organism yet not exploited in clinic. Hence, this lays the foundation for the development of novel molecules able to treat drug-resistant forms of TB, both in monotherapy or in combination with already-available agents, to enhance their effect, shortening the administration regime and lessening potential side effects.

Within the tangled process of nucleobases production, the pyrimidine synthesis is vital for bacterial proliferation and, therefore, its malfunction leads to cell death. The *de-novo* pyrimidine production pathway leverages on orotate synthesis performed by the dihydroorotate dehydrogenase or DHODH, a rate-limiting FMN- and menaquinone-dependent enzyme. Thus, pathogen-directed therapies might target said enzyme, especially based on significant structural and functional divergence between the mycobacterial and the human DHODH. At the same time, the dual strategy blending pathogen-directed and host-directed therapies brings back the idea of a multitarget approach, overtaking present-day issues with conventional TB treatments⁹.

The horizontal gene transfer (HGT) let MTB emerge as human pathogen, but current evidence has it that its role in genome shaping is very limited if not completely absent, a substantial difference than other pathogenic bacteria. Chromosomal rearrangements and exposure to genotoxic stress induced by the

host immune system during active infection opened to the idea that mycobacterial viability is somehow maintained by critical DNA metabolic pathways, recognizing mechanisms like DNA replication, repair, and recombination, the so-called '3Rs', as key to the fixation of gene mutations, in the framework of genomic integrity maintenance, adaptive evolution and drug-resistance of TB bacilli. 3Rs pathways are, therefore, potential druggable targets to fight TB in an unprecedented fashion¹⁰. Pointing to those mechanisms at the basis of genetic diversity, source of antibiotic resistance, might give rise to a new class of chemotherapy agents for infectious diseases that, for the first time, may avoid the surge of drug-resistance. As previously mentioned, the host immune system enacts a DNA-damaging molecules-based response against TB bacilli using reactive oxygen species (ROS) and nitrogen products. In parallel, it seems that ineffective doses of antibiotics induce ROS production, resulting in an increased mutation rate¹¹. Taken together, these evidence state that MTB resistance to medicinal products might be prevented by forcing TB bacilli to mutate at a decreased rate.

As mutations can be spontaneous, in DNA replication, or induced, in the error-prone DNA repair, increasing the fidelity of DNA polymerase sounds surreal, while acting on the accuracy of SOS damage response seems encouraging. In fact, knockout studies on the error-prone DnaE2 polymerase associated with the RecA/LexA-dependent SOS response, already provided *in vivo* evidence that its deletion mitigates the spreading of drug resistant strains in relation to the wild-type control¹². Small molecules aimed at inhibiting LexA and DnaE2 active sites, seeking for selectivity on peculiar structural elements than DnaE1, or RecA DNA-binding ability, represent stimulating opportunities for basic research in preventing induced mutations. The so-called "anti-evolution" compounds family (Figure 3) might be included in treatment regimens with conventional anti-TB drugs, enhancing their

effect in a non-mutagenic environment and decreasing the chemotherapy failure¹⁰. On the other side, the stimulation of mutation rate, pushing the pathogenic bacteria over the 'viability threshold' is a biological theory named "error catastrophe" (Figure 3). This concept has been borrowed from the RNA virus world. Lethality seems to be mathematically plausible by embedding at least 3 to 6 mutations per genome per generation, even in MTB¹⁰. On these grounds, inhibiting the proofreading activity of the DNA polymerase DnaE1 showcased an increased fixation of genome mutations up to 11 per generation in *Mycobacterium smegmatis*, reaching the fatal limit¹³. Eventually, combining this approach to the administration of mutagenic nucleosides in TB treatment might provoke error catastrophe in MTB bacilli, leading to bacterial death.

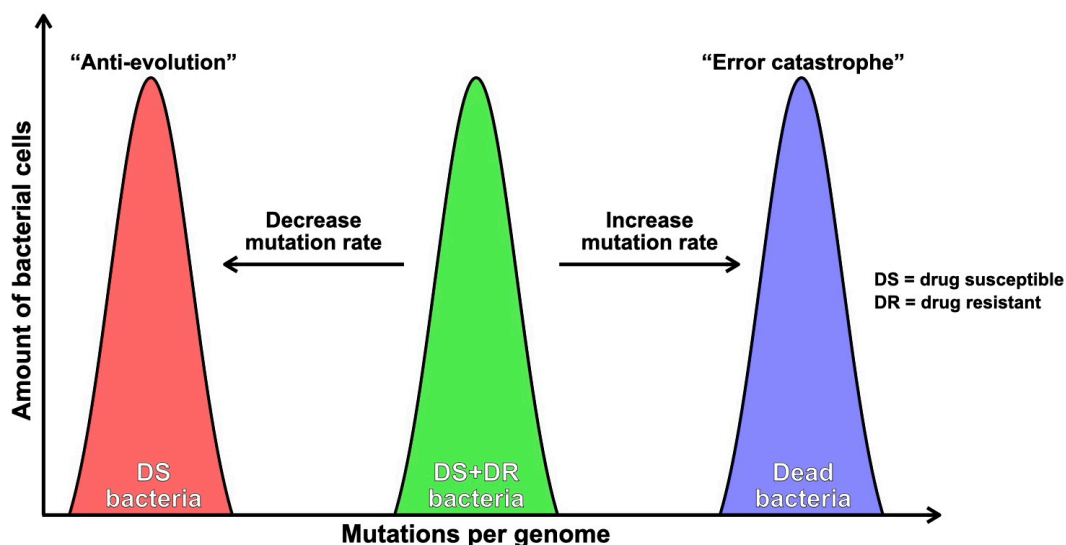


Figure 3. Graphical representation of the mutation rate control and outcome in the pharmacological chemotherapy for infectious diseases. While enforcing systems able to decrease the mutation rate leads to an increased bacterial susceptibility to already available treatments, increasing this rate seems to be a mathematically plausible method resulting in bacterial death. (Figure adapted from¹⁰).

2.1 UNLOCKING THE DNA REPLICATION MECHANISM AS PHARMACOLOGICAL TARGET INFECTIOUS DISEASES

2.1.1 A BRIEF RECAP ON THE PRESENT-DAY SCENARIO

Mycobacterium tuberculosis genome embeds more than 400 genes that have been defined as “essential” for the bacilli viability. Notably, a subgroup made of 15 genes encodes for DNA replication-related proteins, such as the initiator protein DnaA, the replicative helicase DnaB, the helicase loader DciA, the primase DnaG, the single-stranded DNA binding proteins SSBs, DNA polymerases I and III, the DnaN β -clamp and clamp loaders, the DNA ligase I, and both topoisomerase types. As of today, only fluoroquinolones are active compounds targeting the DNA replication machinery, in particular the DNA gyrase, approved in the TB therapy for the treatment of multidrug-resistant strains. Consequently, the leeway towards the development of innovative compounds targeting essential proteins part of the mycobacterial DNA replication machinery is wide open, yet quite ambitious. In this scenario, the validation of new targets as clinically relevant is challenging due to a series of factors. Notwithstanding the necessity of overcoming the compound metabolic inactivation by the TB bacteria or the host immune cells, the lack of evidence on how to biochemically express and purify most of the abovementioned mycobacterial proteins, and so enforcing *in-vitro* reconstitution for further studies, is restricting bioactivity test on newly developed molecules, hence requiring to infer the possible antimycobacterial action by relying on homology models, especially leveraging on the protein tertiary structure convergent evolution theory^{14,15}. However, to avoid interfering with the human DNA metabolism in host immune cells assembled in close vicinity to bacterial cells within the granuloma, antitubercular active compounds must exploit special features of MTB replisome.

Most attempts in targeting the mycobacterial DNA replication found many natural products being able to inhibit some of the replisome crucial players, subtly suggesting how this mechanism is clinically unexplored. As previously mentioned, fluoroquinolones, like moxifloxacin, are active stabilizers of the gyrase-DNA intermediate, leading to double-stranded DNA breaks. Moreover, griselimycin-like synthetic compounds retained the ability of binding to the DnaN β -clamp, impeding its interaction with the polymerase DnaE1 and halting the DNA replication, while showing improved pharmacokinetic properties. At the same time, the macrolide nargenicin showed a potent bactericidal power against *M. tuberculosis* by pointing to the replicative polymerase DnaE. Even though its precise pharmacodynamics has still to be unfolded, nargenicin proved to be ineffective on the human polymerases, confirming selectivity¹⁰.

Medicinal chemistry efforts have been implemented in order to target the mycobacterial DNA ligase LigA and primase DnaG. The former already displayed structural and functional uniqueness than the human equivalent¹⁶. As a result, pyridochromanones were identified as bactericidal molecules against *E. coli* and *S. pneumoniae*, at first, and then to *in-vitro* inhibit the *M. tuberculosis* LigA¹⁷. Yet, excessive LigA capability resulting in spontaneous high drug-resistance levels detected in *S. aureus*, also inferred in the *Mycobacterium smegmatis* system¹⁸, turned down the potentiality for DNA ligase to be classified as a clinically relevant target. On the other side, the most promising molecules against MtDnaG, fetched by high-throughput screening, are doxorubicin and suramin¹⁹, the former intercalating the DNA in a non-specific fashion, the latter inhibiting both the human polymerase α and primase. On top of this, neither showcased strong whole-cell activity towards *M. smegmatis*. Eventually, these results raised questions on the value of the mycobacterial DnaG as druggable target¹⁰.

2.1.2 THE DNA REPLICATION INITIATION MECHANISM: A COMPARISON BETWEEN THE EUKARYOTIC AND PROKARYOTIC WORLD

Let's step back to a more general overview on the initiation phase of the DNA replication mechanism, with a view to addressing similarities and differences between the prokaryotic and eukaryotic systems that, as discussed later on, will be key for unlocking this essential biological pathway as a new potential target for the treatment of infectious diseases and, concerning this dissertation, for TB.

Survival and propagation of any living organism implies the faithful duplication of the genetic information carried by the double-stranded DNA molecule contained in each cell. The DNA replication is a fundamental biological mechanism, result of the concerted action and interaction of many players forming a core machinery termed 'replisome'. This system was found to be highly conserved in all domains of life, namely archaea, bacteria, and eukaryotes. As the overall system complexity differs among them, so does the replisome machinery. For instance, on one side, bacteria often possess only one circular chromosome spanning from 3 to 5 Megabases and bearing only one origin of replication (*oriC*), whereas the eukaryotic genome is made of several linear chromosomes, carrying multiple origins of replication and commonly reaching a total dimension of thousands of Megabases. Moreover, the DNA replication in eukarya has to do with the coordination of the initiation phase throughout the whole genome, managing a series of basic proteins named histones, whose role is to structurally regulate the chromatin condensation state, modulating euchromatin and heterochromatin, in order to make it as more compact as possible. As a result, the eukaryotic DNA is wrapped up in the cell's nucleus in a length compression rate of about 2 million-fold²⁰.

DNA replication accuracy is mandatory for preventing the incorporation of mutations within the cell genome, otherwise leading to cell death or, worse, inheritance by daughter cells after the duplication occurs, most times associated with cancer insurgence or other severe diseases. Hence, based on the extreme complexity of this process, the replisome machinery is one of the most intriguing cell apparatuses known, yet utterly uncharacterized. Not only does this apply to the eukaryotic system, but also to the bacterial world: unfolding each phase would be crucial for infectious diseases containment.

When the bacterial DNA replication starts, two 'forks' are unzipped in opposing direction. Their convergence at the opposite chromosome location than *oriC* leads to the release of two daughter DNA molecules. Alternatively, the eukaryote DNA replication is fired at each origin within every chromosome, producing divergent/convergent replication forks, once per cell cycle, that gather and tether together to form two new chromosomes (Figure 4). If the bacterial *oriC* is well-defined, the eukaryotic one is usually addressed by different chromatin density states, that permit the exposure of DNA regions featuring transcriptional activity, therefore representing interaction sites for origin-binding factors²¹.

Triggering the DNA replication initiation phase proved to be highly conserved in the three domains of life. Indeed, it involves multimeric complexes composed of DNA binding proteins belonging to the AAA⁺ ATPase family. On the bacterial side, DnaA builds up helical filaments on the "less stable" *oriC* A/T-rich region, causing the DNA melting and single strands exposure, so that the replicative 'bubble' gets ready for the helicase loading mechanism. At the same time, the eukaryotic Origin Recognition Complex (ORC) is made of six protomers, that associates with an additional factor called Cdc6, forming a structure able to locate the chromosomal origin, with the goal of leading to the allocation of the

eukaryotic replicative helicase²². Subsequently, as just mentioned, the replicative helicase loading and activation is a mandatory step for the double-stranded DNA opening. Helicases are conserved machineries featuring a ring-shaped hexameric assembly, with the ability of encircling, threading and translocating onto single-stranded DNA ahead of the replisome in an ATP-dependent manner, tearing apart Watson-Crick-Franklin base pairs through steric hindrance. This accounts for the replication forks to move away from the origin of replication²³.

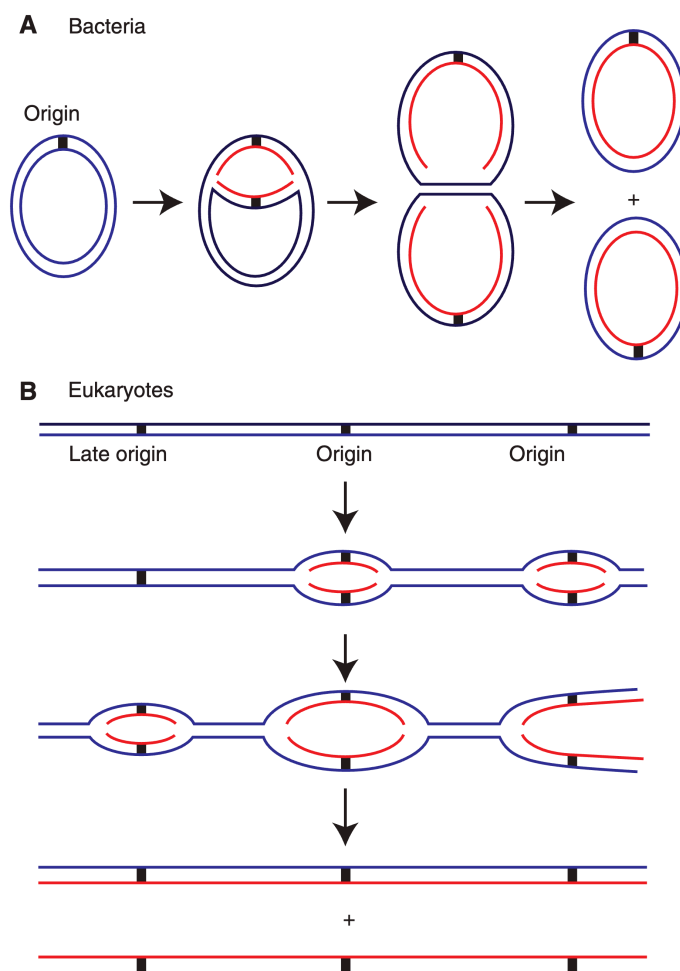


Figure 4. Chromosome duplication schemes. In the bacterial system (A), the circular chromosome presents one origin, and two replication forks move in opposite directions. Conversely, the eukaryotic cell displays long linear chromosomes bearing several origins of replication, where the replicative bubble forms (B).²⁰

Here, a non-exhaustive list of functional and structural differences between the prokaryotic helicase DnaB and the eukaryotic MCM2-7:

1. While the bacterial helicase is an homohexamer, the eukaryotic one is an heterohexamer.
2. DnaB's ATPase architecture is based on the RecA-like fold, whereas MCM2-7 is related to the AAA⁺ family.
3. DnaB translocates on ssDNA in the 5' – 3' direction, following an opposite polarity than MCM2-7, that works 3' – 5'.
4. The eukaryotic helicase is allocated onto double-stranded DNA in a head-to-head double hexamer configuration, while DnaB is loaded onto exposed single strands as single hexamer²⁰.

In addition, the transition from inactivity to a functioning state of MCM2-7 is permitted by the recruitment of Cdc45 and GINS protein factors, finally composing the CMG active complex²⁴.

The allocation of the replicative helicase onto the DNA is a finely tuned process that requires additional factors termed helicase loaders. From this perspective, a helicase loader should, at least, i) interact with the replicative helicase, ii) sustain the helicase loading onto the DNA, and iii) repress its ATPase activity to prevent the initiation of unwanted replication processes. However, as reported from O'Donnell *et al.*, "The precise mechanism by which these proteins load the helicase is unknown in any system"²⁰. In general, the eukaryotic Cdt1 helps MCM2-7 to be placed next to the ORC-Cdc6 complex bound to *oriC*, giving rise to the pre-replicative complex, or Pre-RC, that surrounds the dsDNA in an inactive form. Then, the ATP hydrolysis event within the loaded and activated helicase promotes the release of both Cdt1 and Cdc6, leading to translocation²¹. On the bacterial side, the well-studied DnaC from *Escherichia coli* is an AAA⁺ ATPase

engaging with the inactive DnaB and able to induce an overall structural change of the helicase, which leads the homohexamer to adopt a cracked open and tilted conformation from a closed flat ring shape²⁵. However, as thoroughly discussed further on in Chapter 4 (Unpublished Results pt. 2 – “SAXS and NMR-based Structural Investigation of the Mycobacterial Helicase Loader DciA”), DnaC does not represent the most widespread helicase loader among all the bacterial phyla. Indeed, a peculiar replication factor which displays key features of proteins associated with DnaB and operating during the DNA replication initiation, termed DciA (DnaC-DnaI Antecedent), was found in 23 out of 26 bacterial phyla²⁶. As of today, DciA has been described as a possible ‘ring breaker’, according to O’Donnell’s classification²⁷, based on how it liaises with DnaB proteins featuring a native hexameric assembly in solution. Moreover, DciA proteins are not defined by any ATPase activity, as opposite to DnaC. Nevertheless, structural and functional details on the DnaB-DciA complex are a deep lack in this biological field.

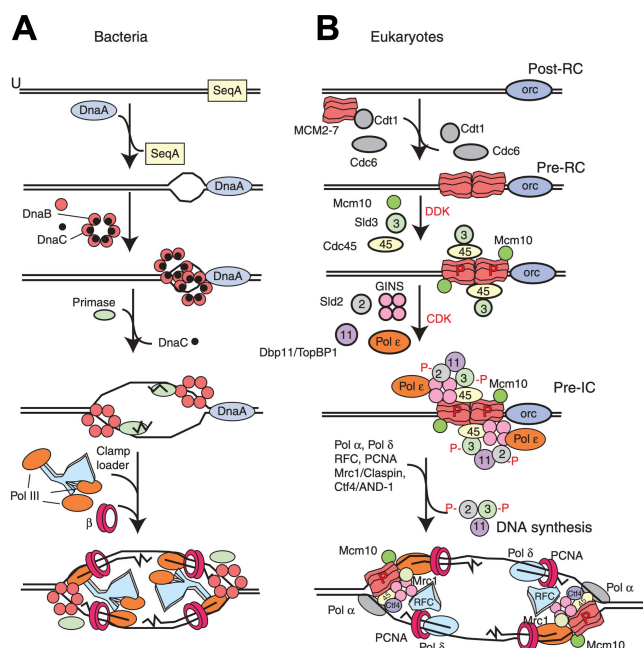


Figure 5. DNA replication initiation phase in bacteria (A) and eukaryotic cells (B). This panel intends to highlight how the outcome of the process is the same, but the players involved are different among the two domains of life. See text for details.²⁰

Recognition of *oriC*, opening of the DNA 'bubble', and loading of the replicative helicase are the three main steps occurring during the DNA replication initiation phase. Prokaryotes and eukaryotes feature equivalent processes, yet the number of factors involved is different (Figure 5), the concerted interplay among them is a key divergence point, and most of these protein factors are structurally unrelated. Altogether, these details open the door to a rationally-driven design of selective active molecules able to interfere and inhibit the DNA replication initiation in human pathogenic bacteria.

2.2 THE INITIATION PHASE OF PROKARYOTIC DNA REPLICATION AS AN UNPRECEDENTED ANTITUBERCULAR CHEMOTHERAPY TARGET

As previously mentioned, several efforts have been made on both the medicinal chemistry and the natural product side with the aim of targeting many of the bacterial DNA replication machinery components. Currently, only fluoroquinolones have been approved in clinic for the treatment of infectious diseases, TB included. All other molecular scaffolds identified opened to important improvements at the level of basic and translational research that, hopefully, will find clinical use in a short amount of time.

The ability of interfering with the high-fidelity duplication of the cell's chromosome reside in either the control of the DNA polymerase activity, but also in the inhibition of proteins involved in the structural management of the genetic material, and replicative helicases possess a well-known fundamental role. In 2013, Shadrick *et al.* reported and discussed on challenges and progress made in targeting replicative helicases as new pharmacological target, not only for the treatment of cancer disease, but also for viral and bacterial infections²⁸. Firstly, the

clearest way to freeze their activity was to inhibit the ATPase catalytic motor through the use of nucleobases analogues. Unfortunately, the extremely high conservation of ATP binding sites among ATPases led to the development of promiscuous compounds, therefore promptly halting this approach²⁹. Then, pointing to the DNA binding site was thought to be a cleverer approach, especially because helicases belonging to different organisms possess structural peculiarities within this specialized cleft. In this way, helicase unwinding activity would have been stopped notwithstanding its catalytic activity. However, most of the molecules identified were enacting their inhibitory activity by interacting with the double-stranded DNA, and not with the helicase itself. In any case, pharmacophores like anthracyclines, tropolones, and acridones were developed as antiviral drugs²⁸.

Considering that the screening of immense compound libraries aimed at directly inhibiting the bacterial replicative helicase DnaB produced few promising outcomes, among which aminocoumarin-like molecules³⁰, researchers turned to a different technique, intended for the disruption of the helicase interaction with fundamental partners in the DNA replication initiation phase. Boehringer Ingelheim was first in finding an active drug preventing the interaction between the helicase E1 and the loading factor E2 of the Human Papilloma Virus (HPV). And so, enforcing a structure-based drug design, the company developed the compound CHEMBL1207308, displaying an encouraging IC₅₀ value of 6 nM³¹. Other approaches in the antiviral field produced molecules capable of interfering with replicative helicase interactors that sustain its ATPase activity³² or engage with the enzyme to help packing the viral genome within the related capsid³³.

As of today, the seek for unveiling the multistep mechanism that brings the bacterial replicative helicase to be allocated onto the origin of replication produced flourish results, both from the kinetic and structural point of view, most of them

discussed further on in this dissertation. DnaA, DnaB and DciA have been defined as essential proteins, and engage in such a way that disrupting their interaction, by exploiting key docking interfaces, is predicted to be a turning point in the pharmacologic manipulation of the DNA replication in human pathogens like MTB.

2.2.1 DEVELOPING BIOACTIVE PEPTIDOMIMETICS AS POTENTIAL PHARMACOLOGICAL TOOL INHIBITING THE DNA REPLICATION INITIATION

The inhibition of protein-protein interaction is claimed to be particularly challenging, especially when involving small molecules. In general, enzymes' binding surfaces are defined by flat and large areas, lacking clear pockets that are usually exploited by conventional therapeutic molecules. Though, if specific features are found at the level of the protein-protein interaction interface, nothing is preventing small ligands to tightly bind even in these conditions. In addition, uniqueness leads to reduced off-target effects³⁴.

Nowadays, a frontier clinical approach leverages on peptide-based therapeutic agents, also known as peptidomimetics. Biological peptides usually consist of short amino acids sequences, covering a huge variety of functions in the three domains of life, spanning from physiological activities to pathogenic ones. As most of them are naturally able to hijack fundamental biological pathways, the pharmacological mimicry is anticipated to open to new therapeutic mechanisms of action. Yet, conventional small molecules are unable to imitate peptides' role, and so peptidomimetics have been developed, sustaining diverse physiological applications. They represent around the 10% of the current pharmaceutical market, and, among others, they serve as detection, prevention and treatment of infectious diseases. To name a few, some active peptides mimic natural antibiotic molecules, while some others are enforced as antigenic components in vaccines with an

immunostimulant role. Anyhow, many limitations like fast biodegradability, lack of oral activity and poor immunogenicity, together with complex pharmacokinetic profiles, called for the application of optimization strategies in order to overcome these flaws, including structural modifications and specific drug formulations³⁵.

This manuscript contains and summarizes the experimental workflow operated since 2018 by many colleagues and myself, with the goal of unfolding each step of the bacterial DNA replication initiation phase, especially in the TB pathogen. Our integrative approach is meant to pave the way to the development of new screening platforms able to identify innovative bioactive peptidomimetics that target the bacterial DnaA-DnaB-DciA complex, bearing in mind the ultimate purpose of discovering new antitubercular drugs, a therapeutic area now outlined by a severe clinical need.

REFERENCES

1. WHO. Global Tuberculosis Report. (2023).
2. Cronan, M. R. In the Thick of It: Formation of the Tuberculous Granuloma and Its Effects on Host and Therapeutic Responses. *Front. Immunol.* **13**, 820134 (2022).
3. Mohammadnabi, N. et al. Mycobacterium tuberculosis: The Mechanism of Pathogenicity, Immune Responses, and Diagnostic Challenges. *J. Clin. Lab. Anal.* **38**, e25122 (2024).
4. Sun, Q., Li, S., Gao, M. & Pang, Y. Therapeutic Strategies for Tuberculosis: Progress and Lessons Learned. *Biomed. Environ. Sci. : BES* **37**, 1310–1323 (2024).
5. Xu, J. et al. Comparative Efficacy of the Novel Diarylquinoline TBAJ-587 and Bedaquiline against a Resistant Rv0678 Mutant in a Mouse Model of Tuberculosis. *Antimicrob. Agents Chemother.* **65**, 10.1128/aac.02418-20 (2021).
6. Singh, R. et al. PA-824 Kills Nonreplicating Mycobacterium tuberculosis by Intracellular NO Release. *Science* **322**, 1392–1395 (2008).
7. Gray, D. A. & Wenzel, M. Multitarget Approaches against Multiresistant Superbugs. *ACS Infect. Dis.* **6**, 1346–1365 (2020).
8. Peng, H., Chen, I. A. & Qimron, U. Engineering Phages to Fight Multidrug-Resistant Bacteria. *Chem. Rev.* (2024) doi:10.1021/acs.chemrev.4c00681.
9. Alberti, M. et al. Biochemical characterization of Mycobacterium tuberculosis dihydroorotate dehydrogenase and identification of a selective inhibitor. *FEBS Lett.* **597**, 2119–2132 (2023).
10. Warner, D. F., Rock, J. M., Fortune, S. M. & Mizrahi, V. DNA Replication Fidelity in the Mycobacterium tuberculosis Complex. *Adv. Exp. Med. Biol.* **1019**, 247–262 (2017).
11. Cohen, N. R., Lobritz, M. A. & Collins, J. J. Microbial Persistence and the Road to Drug Resistance. *Cell Host Microbe* **13**, 632–642 (2013).
12. Boshoff, H. I. M., Reed, M. B., Barry, C. E. & Mizrahi, V. DnaE2 Polymerase Contributes to In Vivo Survival and the Emergence of Drug Resistance in Mycobacterium tuberculosis. *Cell* **113**, 183–193 (2003).
13. Rock, J. M. et al. DNA replication fidelity in Mycobacterium tuberculosis is mediated by an ancestral prokaryotic proofreader. *Nat. Genet.* **47**, 677–681 (2015).
14. Reiche, M. A., Warner, D. F. & Mizrahi, V. Targeting DNA Replication and Repair for the Development of Novel Therapeutics against Tuberculosis. *Front. Mol. Biosci.* **4**, 75 (2017).
15. Chase, J., Berger, J. & Jeruzalmi, D. Convergent evolution in two bacterial replicative helicase loaders. *Trends Biochem. Sci.* **47**, 620–630 (2022).

16. Dawes, S. S. et al. Ribonucleotide Reduction in Mycobacterium tuberculosis: Function and Expression of Genes Encoding Class Ib and Class II Ribonucleotide Reductases. *Infect. Immun.* **71**, 6124–6131 (2003).
17. Gong, C., Martins, A., Bongiorno, P., Glickman, M. & Shuman, S. Biochemical and Genetic Analysis of the Four DNA Ligases of Mycobacteria. *J. Biol. Chem.* **279**, 20594–20606 (2004).
18. Korycka-Machala, M. et al. Evaluation of NAD⁺-Dependent DNA Ligase of Mycobacteria as a Potential Target for Antibiotics. *Antimicrob. Agents Chemother.* **51**, 2888–2897 (2007).
19. Biswas, T., Resto-Roldán, E., Sawyer, S. K., Artsimovitch, I. & Tsodikov, O. V. A novel non-radioactive primase–pyrophosphatase activity assay and its application to the discovery of inhibitors of Mycobacterium tuberculosis primase DnaG. *Nucleic Acids Res.* **41**, e56–e56 (2013).
20. O'Donnell, M., Langston, L. & Stillman, B. Principles and Concepts of DNA Replication in Bacteria, Archaea, and Eukarya. *Cold Spring Harb. Perspect. Biol.* **5**, a010108 (2013).
21. Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N. & Oda, M. Eukaryotic Chromosome DNA Replication: Where, When, and How? *Annu. Rev. Biochem.* **79**, 89–130 (2010).
22. Costa, A., Hood, I. V. & Berger, J. M. Mechanisms for Initiating Cellular DNA Replication. *Annu. Rev. Biochem.* **82**, 25–54 (2013).
23. Gai, D., Chang, Y. P. & Chen, X. S. Origin DNA melting and unwinding in DNA replication. *Curr. Opin. Struct. Biol.* **20**, 756–762 (2010).
24. Ilves, I., Petojevic, T., Pesavento, J. J. & Botchan, M. R. Activation of the MCM2-7 Helicase by Association with Cdc45 and GINS Proteins. *Mol. Cell* **37**, 247–258 (2010).
25. Arias-Palomo, E., O'Shea, V. L., Hood, I. V. & Berger, J. M. The bacterial DnaC helicase loader is a DnaB ring breaker. *Cell* **153**, 438–448 (2013).
26. Brézellec, P., Vallet-Gely, I., Possoz, C., Quevillon-Cheruel, S. & Ferat, J.-L. DciA is an ancestral replicative helicase operator essential for bacterial replication initiation. *Nat. Commun.* **7**, 13271 (2016).
27. Davey, M. J. & O'Donnell, M. Replicative helicase loaders: Ring breakers and ring makers. *Current Biology* **13**, (2003).
28. Shadrick, W. R. et al. Discovering New Medicines Targeting Helicases. *J. Biomol. Screen.* **18**, 761–781 (2013).
29. Frick, D. & Lam, A. Understanding Helicases as a Means of Virus Control. *Curr. Pharm. Des.* **12**, 1315–1338 (2006).
30. Aiello, D. et al. Discovery, characterization and comparison of inhibitors of Bacillus anthracis and Staphylococcus aureus replicative DNA helicases. *Bioorg. Med. Chem.* **17**, 4466–4476 (2009).
31. Goudreau, N. et al. Optimization and determination of the absolute configuration of a series of potent inhibitors of human papillomavirus type-11 E1–E2 protein–protein interaction: A combined medicinal chemistry, NMR and computational chemistry approach. *Bioorg. Med. Chem.* **15**, 2690–2700 (2007).

32. Wright, C. M. et al. Inhibition of Simian Virus 40 replication by targeting the molecular chaperone function and ATPase activity of T antigen. *Virus Res.* **141**, 71–80 (2009).
33. Mousseau, G., Kota, S., Takahashi, V., Frick, D. N. & Strosberg, A. D. Dimerization-driven interaction of hepatitis C virus core protein with NS3 helicase. *J. Gen. Virol.* **92**, 101–111 (2011).
34. White, P. W., Faucher, A.-M. & Goudreau, N. Small-Molecule Inhibitors of Protein-Protein Interactions. *Curr. Top. Microbiol. Immunol.* **348**, 61–88 (2010).
35. Bojarska, J., Wang, X. & Skwarczynski, M. Editorial: Peptides against infectious diseases: from antimicrobial peptides to vaccines. *Front. Pharmacol.* **15**, 1522148 (2024).
36. Peters, J. S. et al. Advances in the understanding of Mycobacterium tuberculosis transmission in HIV-endemic settings. *Lancet Infect. Dis.* **19**, e65–e76 (2019).

Chapter 2

THESIS OUTLINE

THESIS OUTLINE

The investigation of the biological events at the base of the DNA replication initiation mechanism is a relatively new, yet well-structured scientific project, that belong to a wide-ranging experimental framework conceived to unfold and describe the DNA metabolism, including the pyrimidine synthesis and the DNA repair systems, in the human pathogen *Mycobacterium tuberculosis* (MTB), the etiologic agent of Tuberculosis. As part of the historical scientific portfolio of the Biochemistry and Structural Biology Unit at the Department of Pharmaceutical Sciences (UNIUPO), my supervisor Professor Riccardo Miggiano aims to gain a deep understanding of the role of enzymes involved in these pathways, considering their catalytic activity, concerted interactions and structural features, with a view to developing new libraries of active compounds able to interfere with targets identified as essential and the base of the pathogenicity of Tuberculosis bacilli.

Specifically, my PhD research project takes advantage of different techniques, in an integrative biophysics and structural biology approach to accurately characterize the key steps of DNA replication initiation phase, fundamental to load the replicative helicase DnaB onto the origin of replication, a finely tuned process that relies on a series of replication factors, whose biological relevance is yet to be uncovered, especially in MTB.

Since 2018, the fulfilling collaboration with Professor David Jeruzalmi from The City College of New York (USA), a long-time structural biology expert on DNA metabolism, paved our way towards a smarter experimental set up on the MTB system, thanks to the data Jeruzalmi Lab obtained by exploiting recombinant proteins from the model organism *Vibrio cholerae*. Hence, my PhD journey was

designed so that we promptly obtained robust data on the tuberculosis targets, letting us strengthen the so-called 'convergent evolution' concept on the bacterial DNA replication. At the same time, during my 1-year long period of time as a visiting PhD student at Professor Jeruzalmi's lab in New York, I had the chance of gaining expertise on what concerns the single-particle cryo-electron microscopy technique. And so, we acquired high-quality cryo-EM datasets that let us solve the 3D structure of DNA-free and DNA-bound multiprotein complexes, a further step in structurally outlining the events occurring at the initiation phase of the DNA replication.

This PhD dissertation contains both published and unpublished data obtained during the timeframe between November 2021 and November 2024, outcome of a stimulating and challenging work, handled mostly by me amid Italy and the United States of America. In particular, here I show and discuss:

- 1) Recombinant protein production, starting from *E. coli* as expression system to the high-yield and optimized purification using standard applied biochemistry approaches. The protocol development and optimization for *Mycobacterium tuberculosis* proteins occurred in the laboratory of Professor Miggiano at UNIUPO, while the production of *Vibrio cholerae* ones was part of the PhD program of Dr. Nicholas Gao, carried out in the laboratory of Professor Jeruzalmi at The City College of New York up to 2022.
- 2) Protein-protein interaction analyses, based on 'pull-down' assays and Surface Plasmon Resonance technology, carried out at the Miggiano Lab (DSF-UNIUPO), and native Mass Spectrometry, performed by Professor Brian T. Chait group at the Rockefeller University, New York.

- 3) Enzymatic catalytic activity determination, spanning from spectrophotometric ATPase assays to FRET-based systems using fluorescent DNA substrates, firsthand implemented and optimized in the laboratory of Professor Jeruzalmi at The City College of New York.
- 4) Cryo-EM-based 3D structure determination of *Vibrio cholerae* protein-DNA macro complexes, from the sample preparation to the three-dimensional modelling after data processing, that I personally implemented and conducted between the Jeruzalmi Lab at The CCNY and the New York Structural Biology Center.

Chapter 3

PUBLISHED RESULTS

DnaB and DciA: mechanisms of helicase loading and translocation on ssDNA

Nicholas Gao^{*1,2}, Daniele Mazzoletti^{*5}, Adele Peng¹, Paul Dominic B. Olinares⁶, Castrese Morrone^{§5}, Andrea Garavaglia^{^5}, Nourelhoda Gouda¹, Serge Tsoy¹, Albert Mendoza¹, Ahammad Chowdhury¹, Antonio Cerullo^{1,4}, Hrutvik Bhavsar¹, Franca Rossi⁵, Menico Rizzi⁵, Brian T. Chait⁶, Riccardo Miggiano^{#5}, and David Jeruzalmi^{#1,2,3,4}

¹ *Department of Chemistry and Biochemistry, City College of New York, New York, NY 10031, USA.*

Ph.D. Programs in ²Biochemistry, ³Biology, and ⁴Chemistry, The Graduate Center of the City University of New York, New York, NY 10016, USA.

⁵ *Department of Pharmaceutical Sciences, University of Piemonte Orientale, Novara, 28100, Italy*

⁶ *Laboratory for Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, New York 10065, USA*

[°] *Present address: Mirimus Inc, 760 Parkside Ave, Brooklyn, NY 11226, USA*

[§] *Present address: Telethon Institute of Genetics and Medicine (TIGEM) Via Campi Flegrei, 34 80078 Pozzuoli (NA) Italy.*

[^] *Present address: Department of Pharmaceutical Science (DISFARM), University of Milan, Milan, 20133 Italy*

[#] *Corresponding authors: dj@ccny.cuny.edu or riccardo.miggiano@uniupo.it*

^{*} *These authors contributed equally.*

ABSTRACT

Replicative helicases are assembled on chromosomes by helicase loaders before initiation of DNA replication. Here, we investigate mechanisms used by the bacterial DnaB replicative helicase and the DciA helicase loader. In the present structure of the DnaB-ssDNA•ATP γ S complex, the amino-terminal (NTD) tier, previously found as an open spiral in a GDP•AlF₄ complex, was observed to adopt a closed planar arrangement. Further, the DnaB subunit at the top of the carboxy-terminal spiral (CTD) tier is displaced by ~25 Å between the two forms. We suggest that remodeling the NTD layer between closed planar and open spiral configurations and migration of two distinct CTDs to the top of the DnaB spiral, which is repeated three times, mediates hand-over-hand translocation. Although DciA loaders from *Vibrio cholerae* and *Mycobacterium tuberculosis* feature an inverted domain arrangement, they use Lasso domains to contact DnaB near the Docking-Linker-Helix interface. Up to three copies of VcDciA bind to VcDnaB and suppress its ATPase activity during loading onto physiological DNA substrates. Our data suggest that DciA loads DnaB onto DNA using the ring-opening mechanism.

INTRODUCTION

Replisomes in all domains of life rely on replicative helicases to melt chromosomal DNA into single-stranded templates for DNA replication^{1–6}. Replicative helicases feature a toroidal ensemble comprised of six protein subunits and a topologically closed central chamber into which the active entity admits one strand of DNA while the second strand is excluded. DnaB leverages the energy of ATP binding and hydrolysis to translocate along single-stranded DNA (ssDNA) in its central chamber. Assembly of DnaB on ssDNA requires distinct mechanisms and factors (Figure 1).

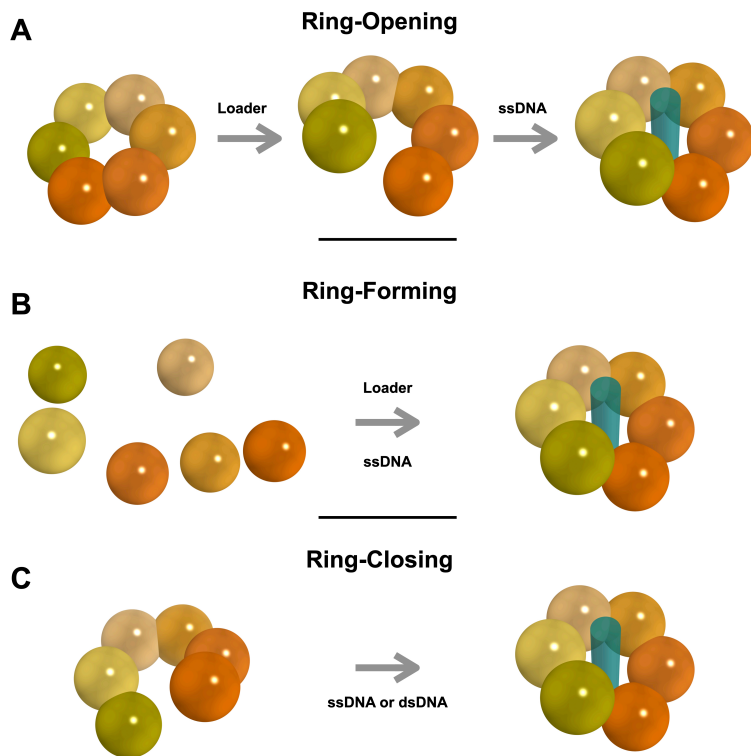


Figure 1. Mechanisms of Helicase Loading. Strategies used to load hexameric helicases onto nucleic acid substrates are ring-opening, ring-forming, and ring-closing^{22,55,56}; the terms refer to the initial state of the helicase. In the ring-opening scheme, closed ring hexamers are opened by helicase loaders and assembled on ssDNA. The hexameric helicase is assembled on ssDNA from monomers in the ring-forming scheme. Entry of the nucleic acid substrate and other factors mediate the ring-closing scheme. Hexameric helicase subunits are represented as spheres in shades of orange. For clarity, only the ATPase domains are shown. Nucleic acid substrates are depicted as deep teal cylinders.

In bacteria, replicative helicases are members of the DnaB family^{7–20} (henceforth: DnaB). The physiological assembly (reviewed in^{21–27}) presents as a homohexamer, with two tiers (termed the amino-terminal domain (NTD) and the carboxy-terminal domain (CTD)) that each circumscribes a central channel through which ssDNA will traverse. A short linker helix (LH) element connects each NTD and CTD; the LH packs against a neighboring CTD's docking helix (DH) in the complete hexamer. The finding that the translocating form of DnaB adopts a right-handed spiral configuration implied a hand-over-hand mechanism to explain translocation on ssDNA. Translocation occurs akin to how a human uses two hands to climb a rope, except that DnaB has six hands. Translocation initiates via ATP hydrolysis on a DnaB subunit at the bottom of the spiral (in the pose in Figure 2). This subunit peels away from its partner and from ssDNA^{10,24,28,29}, migrates to the top of the spiral, docks onto a DnaB subunit via its β -hairpin arginine finger, and binds ssDNA position 3' from the original location, further translocation occurs via progressive subunit migrations to the top of the DnaB spiral¹⁰. Crucial support for the hand-over-hand model emerged from a series of structures of the DnaB-family helicase Gp4 from T4^{24,29}.

Studies utilizing bacteria and phage have uncovered several helicase loaders and their interactions with the cognate DnaB helicase. These include *Escherichia coli* (*E. coli*) DnaC³⁰, bacteriophage λ P³¹, bacteriophage SPP1 G39P^{32,33}, *Bacillus subtilis* DnaI³⁴, T4 gene 59³⁵, and phage P2 B³⁶. Surprisingly, the genomes of most bacteria lack orthologs of these well-studied loaders. Instead, most bacteria employ members of the structurally, and perhaps mechanistically, distinct DciA or DopE families^{37–41}. Although the amino acid sequences of bacterial helicase loaders diverge significantly, they appear to mediate conserved mechanisms during helicase loading^{11,12,14,16,19,36,42–48}. Loaders display a conserved overall architecture: a globular domain adjacent to an intrinsically disordered segment that

often features an alpha helix at a terminus, termed the Lasso/grappling hook. Related arrangements can also be inferred for *B. subtilis* SPP1 phage G39P⁴⁹, *B. subtilis* DnaI^{16,50–52}, and *E. coli* phage T4 gp5953 helicase loaders. The DciA helicase loaders are organized around a type II K homology domain (KH) with Lasso elements at either the amino or carboxy terminus of a particular ortholog or both^{12,14,40,41,54}.

Cells apply three broad strategies to load hexameric helicases onto nucleic acid substrates: ring-opening, ring-forming, and ring-closing^{22,55,56} (Figure 1); the terms refer to the initial state of the helicase. These schemes encompass loading onto ssDNA, ssRNA and dsDNA. In the ring-opening scheme, closed ring hexamers, e.g., *E. coli* DnaB, are opened by helicase loaders, such as *E. coli* DnaC^{18–20} or phage λ P^{43,44,48} and assembled on ssDNA formed at the cognate replication initiator – origin DNA complex. In the ring-forming scheme, the hexameric helicase is assembled on ssDNA from monomers, either spontaneously as with the LT-antigen or E1 helicases or through the activity of a helicase loader such as DnaI, which loads the DnaB family helicase in *B. subtilis*^{34,119}. The ring-closing scheme is observed with the bacterial Rho and the eukaryotic MCM2-7 hexameric replicative helicases (reviewed in^{22,55,56}). In the Rho helicase, ring closure is mediated by the entry of the nucleic acid substrate. In the MCM2-7 ensemble, dsDNA is deposited via the chaperone origin recognition complex and associated factors. The scheme(s) the DciA loader uses to assemble DnaB onto the origin-DNA is unknown.

Although much is known about many bacterial loaders^{11,12,14,16,19,36,42–46}, relatively little is known about DciA helicase loaders and the mechanisms they implement during helicase loading. DciA orthologs from *Vibrio cholerae* (Vc), *Mycobacterium tuberculosis* (Mt), and *Caulobacter crescentus* (Cc) are all essential proteins^{37,42,57,58}, and they have been shown to bind to cognate DnaB helicases^{12,14,42,57,58}. Component KH and Lasso domains both contact DnaB^{12,14,58}.

The binding site on DnaB for DciA is proposed to encompass the DH-LH interface^{12,14,58}. DciA orthologs bind to DNA^{11,58,59}, a feature shared with other bacterial loaders^{31,35,47,60}. Suppression of DnaB's ATPase activity is a hallmark of the bacterial helicase loaders^{43,46,60}; however, Cc-DciA has been reported to promote ATP hydrolysis⁴². DciA activates cognate DnaB's DNA unwinding activity by mediating assembly on ssDNA, though the underlying mechanisms have not been established^{40,41}.

To gain insights into the mechanisms operating during loading of DnaB by the DciA loader and subsequent translocation, we studied orthologs from *Vibrio cholerae* (Vc) and *Mycobacterium tuberculosis* (Mt). The choice of organisms reflects relevance to human disease and a peculiar divergent DciA architecture. The cryogenic electron microscopy (cryo-EM) structure of the VcDnaB•ssDNA•ATP γ S complex, the end state of helicase loading, implies that cycles of opening/closing of the NTD layer and CTD remodeling are significant features in the mechanism of translocation on ssDNA. We suggest that remodeling of the NTD layer and distinct migrations of two subunits to the top of the DnaB spiral, repeated three times, mediates hand-over-hand translocation. Native mass spectrometry (nMS) indicates that DnaB is a hexamer and that the stoichiometry of the VcDnaB-DciA-ssDNA complex is 6:3:1. VcDciA features a KH domain at its amino terminus and a disordered segment (Lasso) at its carboxy terminus; by contrast, in MtDciA, this arrangement is inverted^{12,40,41,54,58}. Multiple binding experiments show that the Lasso segment specifies most of the binding energy in the interaction with DnaB in both orthologs. Furthermore, Vc and MtDciA bind to cognate DnaB at the interface between the docking helix (DH) and linker helix (LH) on adjacent subunits. VcDciA suppresses the intrinsic ATPase activity of DnaB. We show that VcDciA loads DnaB onto ssDNA and activates unwinding. Overall, our data suggests that DciA is likely a ring-opening helicase loader.

RESULTS

Cryo-EM Structure of VcDnaB Bound to ssDNA and ATP γ S

To understand the loading of the DnaB helicase by DciA, we applied cryogenic electron microscopy (cryo-EM) to VcDnaB helicase bound to ATP γ S, VcDciA, and ssDNA derived from the VcOriC-I replication origin. In assembling the complex, we relied on results from our nMS study (below).

Cryo-EM analysis (Table 1, Supplementary Methods, Supplementary Figures 1 and 2) produced a 3.37 Å structure of hexameric VcDnaB bound to ATP γ S and complexed to a 22-nucleotide ssDNA molecule (Figures 2 and Supplementary Figure 3). The resolution of the EM density corresponding to one of the chains (chain A) was of lower resolution than that of other chains. Notwithstanding the fourfold excess of VcDciA in the sample submitted to cryo-EM, our maps failed to show any density corresponding to the loader. Our efforts likely recovered the end-state structure of the VcDnaB-ssDNA complex after DciA-mediated loading or self-threading.

Compared to the *Bst*DnaB-ssDNA complex, determined with GDP•AlF₄ (PDB: 4ESV¹⁰), the ATP γ S form of the VcDnaB-ssDNA complex presents a distinct configuration with several significant changes (Figures 2 and 3). Both share the trimer of dimers and pseudo-sixfold architecture of the NTD and CTD tiers in extant structures (Figures 2 and 3). With diameters of ~55 Å and ~25 Å, the NTD and CTD layers in the VcATP γ S form are dilated and constricted⁸, respectively. However, the NTD from the VcATP γ S structure adopts a closed triangular planar and symmetric arrangement (Figure 3), resembling the NTD from other structures (PDB: 2R5U⁶¹, Root Mean Square Deviation (RMSD) = 2.6 Å; PDB: 7T20⁶², RMSD = 1.7 Å).

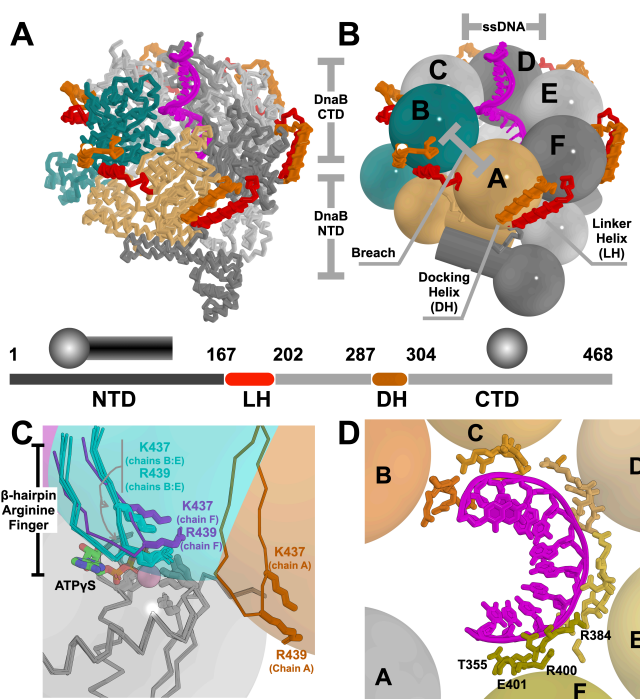


Figure 2. The VcDnaB-ssDNA-ATPyS Complex and its Interactions with Nucleotide and ssDNA. The VcDnaB-ssDNA-ATPyS is drawn in a ribbon representation (A) or a schematic using a design language wherein the NTDs are depicted as a mushroom shape and the CTDs as spheres (B). ssDNA, colored in magenta, is depicted in a cartoon and stick representation. The breach in the CTD layer is between chain A (light orange) and chain B (deep teal); other DnaB subunits are colored in alternating shades of gray. The Docking Helix (DH) and the Linker Helix (LH) are depicted as cylinders colored orange and red, respectively. The primary sequence of VcDnaB is depicted as a bar, with the salient domain boundaries indicated. (C) Close-up of the ATP catalytic center at the interface between each pair of DnaB CTDs. ATP is depicted in the stick representation. The six DnaB CTD pairs, depicted here as spheres, are superimposed on the subunit (colored in grey) that harbors the Walker A and B motifs (not indicated). For clarity, only a portion of the superimposed CTD is shown as a ribbon; the grey sphere shows the full extent of the superimposed CTD. Spheres corresponding to the CTDs excluded from the superposition are colored in purple (chain F), cyan (chains B:E), and orange (chain A). The β -hairpin arginine fingers from all six non-superimposed CTDs are shown in the ribbon format. Although excluded from the superposition, the β -hairpin four corresponding to four chains (chains B, C, D, and E, cyan) superimpose closely. In contrast, chains F (purple) and A (orange) are found ~ 3 Å and ~ 19 Å, respectively. (D) Schematic representation of the VcDnaB-ssDNA interface. The DnaB CTDs are represented as spheres, labeled by chain, and colored in shades of yellow/orange. The ssDNA is drawn in a cartoon, with a stick representation and colored cyan. The protein side chains from one subunit (chain F) that contact ssDNA are shown in the stick representation, colored by chain as above, and labeled. The complete list of protein-DNA contacts appears in Supplementary Table 2. Additional details on the VcDnaB-ssDNA complex can be found in Supplementary Figures 5 and 6.

This configuration differs from the open spiral in the GDP•AlF₄ form of the BstDnaB-ssDNA complex¹⁰. As with other DnaB structures, the VcNTD assembles around weak (head-to-head) and strong interfaces (tail-to-tail), defined based on

the buried surface area⁷; the NTD tier in the *Bst* structure is breached on the weak head-to-head interface. In both nucleotide forms, the constricted CTD tier is arranged as a closed spiral; however, the spirals are distinct.

The six DnaB subunits in the *Vc* and *Bst* structures are labeled A-F-E-D-C-B around the spiral; the breach in the CTD spiral lies between chains A and B, which are at the bottom and top of the spiral (as in the pose in Figures 2 and 3), respectively. Five of the six subunits of each structure (PDB: 4ESV: resid 181:441

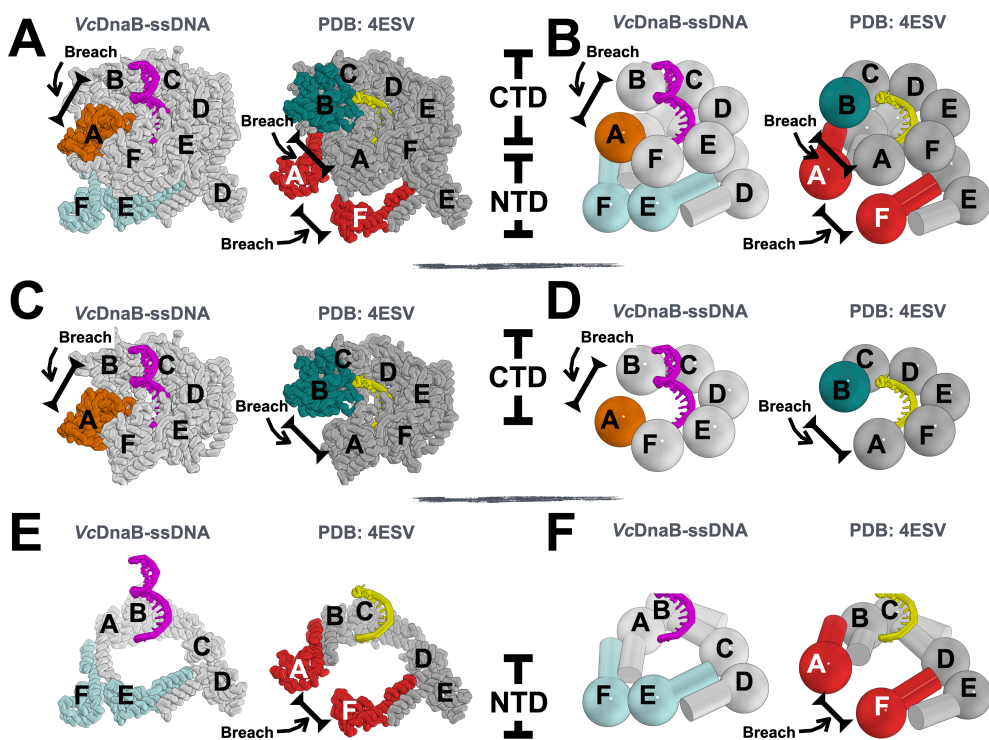


Figure 3. Nucleotide-Dependent Conformation-Switching between the *Vc*DnaB-ssDNA-ATPyS and *Bst*-GDP•AlF₄ complexes. The two complexes are superimposed in all panels, as described in the main text. A) Ribbon representation of both complexes. *Vc*DnaB NTD components chain F and E in the closed planar and triangular configuration are colored in pale cyan. *Bst*DnaB NTD components chain A and F in the open spiral configuration are colored in red. *Vc*DnaB chain A is colored orange. 4ESV chain B is colored deep teal. The ssDNA in the *Vc* and *Bst* complexes are drawn in the cartoon and stick representation and colored purple and yellow, respectively. All other chains of DnaB are colored in light gray (*Vc*) and dark gray (*Bst*). The breaches in the NTD and CTD layers are indicated. Chains corresponding to the NTD and CTD components are labeled. B) The depiction is identical to that in panel A, except that the NTD and CTD layers are represented using the sphere and mushroom-like shape design language. C) The exact representation as panel A, except that only the CTD layer is shown. D) Identical depiction as panel C except that the DnaB CTD layers are depicted as spheres. E) The exact representation as panel A, except that only the NTD layer is shown. F) Identical depiction as panel E except that the DnaB NTD layers are depicted using mushroom-like shapes.

and chain A, C, D, E, and F on VcDnaB: resid 204:441 and chain B, C, D, E, and F) chains could be superimposed with an RMSD of 1.7 Å (1016 Ca atoms). The most significant difference in CTD configurations is the distinct position occupied by VcDnaB chain A, which maps to chain B in the Bst 4ESV structure; we address this change in greater detail below. We also note that in the Vc ATP γ S form, chain A is located further away from its neighbors (distances to chain B: ~39.6 Å, to chain F: ~34.6 Å) than the average distance separating the remaining four subunits to their neighbors (Chains C, D, E, and F: ~33.4 Å). By contrast, in Bst 4ESV, Chain A is distant from one immediate neighbor (distance between chains B and A: 41.6 Å), while other chains display a distance of 32.5 Å to their neighbors. Lastly, in the VcATP γ S complex, the six Linker Helices (LH) occupy three positions relative to the NTD and CTD. The LH elements of chains B, C, D, and E occupy one position (LH1), with chains A (LH2) and F (LH3) diverging from this position by rotations/displacements of ~29.6°/12.3 Å and ~16.5°/~7.1Å, respectively. In the Bst 4ESV GDP•AlF₄ complex, the LH elements occupy two positions, one by chains B, C, D, E, and F (LH1), and a second position (LH4) by chain A (Supplementary Figure 4).

ssDNA in the complex traces a spiral path that passes through the internal chamber of VcDnaB (Figures 2 and 3). 12 of the 22 ssDNA nucleotides are bound to subunits of the VcDnaB hexamer; the remaining 10 are outside the DnaB perimeter (Supplementary Figure 5). In contrast to the GDP•AlF₄ 4ESV structure where each subunit contacts ssDNA, only five DnaB subunits in the current Vcstructure make contacts (less than 4.0 Å); chain A does not contact ssDNA. An inventory of protein-DNA contacts reveals the involvement of seven VcDnaB residues, though not every contact is made by each subunit (Supplementary Table 2 and Supplementary Figure 6).

Six ATP γ S/Mg²⁺ complexes are found at composite catalytic centers formed between adjacent VcDnaB subunits, wherein the Walker A/P-loop and Walker B motifs are provided by one subunit and the β -hairpin arginine finger by a second (Figure 2 and Supplementary Figure 7). In four nucleotide sites (chain F-E, E-D, D-C, and C-B), the β -hairpin/arginine finger makes a close approach to nucleotide on a neighboring subunit, with distances between the P-loop lysine (K234) and the arginine (R439) finger residues of ~ 11.3 Å. The corresponding distance at the A-F subunit interface is slightly longer (~ 11.9 Å), and the arginine finger has shifted by ~ 3 Å relative to other subunits. The β -hairpin/arginine finger at the nucleotide site that lines the breach (B-A) exhibits a considerably longer (~ 24.5 Å) distance. In the Bst GDP•AlF₄ complex, 5/6 sites (Chains A-F, F-E, E-D, D-C, and C-B) feature an average P-loop lysine (K216) to arginine (R420) finger distance of ~ 11.44 ; the chain B-A site, which lines the breach, distance is considerably longer at 32.81 Å.

Taken together, the significant changes between the GDP•AlF₄ and ATP γ S forms of DnaB, including distinct NTD configurations, the offset of a single DnaB CTD, and diversity in LH element conformations in sequential subunits around the ring, imply that our structure has captured a distinct intermediate in the translocation reaction.

VcDnaB Primarily Contacts the Phosphate Backbone of ssDNA

VcDnaB subunits use seven residues (Asn353, Thr355, Asn383, Arg384, Arg400, Glu401, and Gly403) to contact the ssDNA backbone, though not every chain uses every residue; one contact is seen to contact a nucleobase (chain D Asn353 and N2 of G5). The segment of ssDNA within DnaB's central chamber is arranged with 11-12 nucleotides per turn of ssDNA and a pitch of ~ 32 Å, a value that compares well with the pitch (~ 33.4 Å) displayed by the DnaB CTD spiral. The approximate match in pitch between protein and ssDNA spirals implies that

corresponding DNA binding residues on each subunit make equivalent contacts with the ssDNA backbone¹⁰ (Supplementary Figure 6 and Supplementary Table 2).

Remodeling of the NTD and CTD Tiers Accompany Translocation

To gain insights into the distinct ATP γ S form of DnaB and its place in the translocation reaction, we compared it against the *Bst* GDP•AlF₄¹⁰ entity (Figure 3) and a series of DNA complexes (PDB: 6N7I, 6N7S, 6N7T, and 6N7V) of the DnaB-family helicase Gp4 from phage T4^{24,29} (Comparisons to the VcDnaB structures appears in the Supplementary Information). Aligning the VcATP γ S and *Bst* GDP•AlF₄ structures by superimposing 5/6 CTDs (RMSD: 1.7 Å, a similar alignment could be obtained by aligning on the ssDNA (RMSD: 0.64 Å) reveals that Vc chain A, excluded from the alignment, is displaced by 25 Å and rotated by ~22° relative to the corresponding subunit in 4ESV (Chain B). Similar displacements/rotations are seen in the HelF subunits in the 6N7T and 6N7V DnaB-family Gp4 helicase structures, shifting positions by ~24 Å and ~13°^{24,29}. Further, the DnaB alignment juxtaposes all the components of the NTD layer save for the chains at the top (4ESV: chain A) and bottom (4ESV: chain B) of the *Bst* spiral. These two chains are offset owing to the distinct planar vs. spiral configurations adopted by the Vc ATP γ S and *Bst* GDP•AlF₄ complexes. Unexpectedly, the breaches in the two structures do not align (Figure 3). Instead, they are offset by one subunit.

Given the distinct states captured by each structure, we reasoned that the Vc ATP γ S structure might precede the transition state *Bst* complex in the DNA translocation reaction coordinate. Ordering the structures in this way specifies a direction for the observed structural changes. Accordingly, the NTD of chain F and chain E in the Vc ATP γ S complex will transition from the closed triangular planar configuration to the open spiral of the GDP•AlF₄ complex (mapped to 4ESV chains: A and F); concomitantly, the CTD of Vc chain A will move upward (as in the pose in

Figure 3) to assume the configuration seen in chain B of 4ESV (Supplementary Movie 1). With the migration of the Vc chain A CTD to the position specified in the 4ESV structure, the breach in the CTD spiral moves counterclockwise by one subunit. The LH element moves from the LH2 to the LH4 position in 4ESV chain A. Several lines of evidence in the Vc ATP γ S complex support the proposal that the chain A CTD migrates (Figure 3). First, the arginine finger/ β -hairpin element on chain F has peeled away by ~ 3 Å, relative to its position in chains B, C, D, and E, from the P-loop on chain A (Figure 2C). Second, chain A is farther away from its neighbors than any other DnaB subunit. Third, chain A makes no contact with ssDNA. Lastly, of all the CTDs, the EM density for chain A is the weakest. Together, these findings point to the chain A CTD as disengaged from DnaB and ssDNA and poised to transition.

Due to the NTD layer's asymmetry, the next subunit's motion, chain F, must follow a different trajectory. We propose that on ATP hydrolysis, chain F, now located at the bottom of the spiral, will also peel away from DNA and migrate upwards to a new binding site on ssDNA. In contrast to the motion of chain A, which results in the opening of the NTD layer, the motion of chain F CTD must be accompanied by the closure of the NTD ring from the open spiral to the closed planar configuration; this motion will also result in the adjustment of the position of the LH element, from the LH3 position to the LH1. Together, these two changes re-establish the NTD configuration in the Vc ATP γ S structure. Applying the above to chains E and D first and then to chains C and B allows each DnaB subunit to complete a catalytic cycle of ATP hydrolysis amid translocation on ssDNA. These findings highlight the importance of opening and closing the NTD layer, which had previously been suggested to specify the direction of unwinding⁶³ in the translocation reaction.

Comparison to other structures of VcDnaB

We compared our cryo-EM VcDnaB structure to three previously determined crystal structures: 6T66, 7QXM, and 8A3V. None feature ssDNA (6T66, 7QXM), and one includes DciA at a 6:6 ratio with DnaB (8A3V). We compared the structures for subunit proximity, ring breaches, nucleotide occupancy, disposition of the β -hairpin/arginine finger, and LH position.

The 6T66 structure¹² presents as a planar hexamer in the crystallographic asymmetric unit with a dilated NTD layer and a constricted CTD tier (Supplementary Figure 29A). The DnaB subunits in this structure are labeled A-B-C-D-E-F, counterclockwise down the CTD axis. Two of the CTDs (A – F: 35.1 Å and D – E: 36.4 Å) are 2-3 Å further away from their neighbors than the disposition of other subunit pairs (A – B: 32.8 Å; B – C: 32.7 Å; C – D: 32.7 Å; E – F: 33.1 Å), implying two breaches in the CTD layer; the NTD layer is unbreached. In the VcDnaB-ssDNA complex, chain A, at the bottom of the spiral, is further away (A – B: 39.6 Å; A – F: 34.6 Å) from its neighbors than other pairs (average: ~33.4 Å). The significance of two breaches in non-consecutive subunits in the CTD tier in the 6T66 structure remains to be understood.

The nucleotide-binding site of each subunit in 6T66 is populated with GDP•AlF₄/Mg²⁺. However, the β -hairpin/arginine finger elements that complete these binding sites populate three positions. The distance between the P-loop lysine 234 and arginine finger 439 at the sites between chains D to E (18.7 Å) and A to F (16.0 Å) is considerably longer than the corresponding values for other subunit pairs (A – B: 10.9 Å; B – C: 11.2 Å; C – D: 11.1 Å; E – F: 11.3 Å). Naturally, longer distances are seen in CTDs that are further apart (Supplementary Figure 29C).

We next examined the positions adopted by the LH element in the 6T66 hexamer. Alignment of the CTD onto the chain A CTD, but omitting the DH and

LH elements from the calculation (resid 204:465 and not (resid 287:304 or resid 178:205)) produces an average RMSD = 0.74 Å. In this alignment, the LH elements populate two positions (LH1 and LH2 above). The LH elements from chains A and E populate positions closely related to LH2; the rest are seen in the LH1 position (Supplementary Figure 29D). The subunit pairs with divergent dispositions and arginine fingers populate the LH2 position; the rest are in LH1.

The 7QXM VcDnaB crystal structure features a planar hexamer in the asymmetric unit and is unliganded¹³. The DnaB subunits in this structure are labeled A-F-E-D-C-B, counterclockwise down the CTD axis. The NTD layer is dilated and follows NTD layers from 6T66 and the VcDnaB-ssDNA complex. The CTD layer is constricted, with individual CTD further away from each other than in 6T66 or the VcDnaB-ssDNA complex (A – B: 33.7 Å; B – C: 37.0 Å; C – D: 33.7 Å, D – E: 36.3; E – F: 33.6 Å; F – A: 37.9 Å). Likewise, the P-loop lysine to arginine finger distances is variable (A – B: 21.1 Å; B – C: 16.8 Å; C – D: 20.4 Å, D – E: 15.7 Å; E – F: 20.9 Å; F – A: 17.3 Å). The observed variability in the disposition of CTDs in 7QXM may reflect the underlying lack of nucleotide to organize and stabilize inter-CTD contacts via the β -hairpin arginine finger. The LH elements in 7QXM populate two positions, LH2 (Chains B, D, and F) and another position (Chains A, C, and E), which are unique to this structure.

The 8A3V VcDnaB-DciA crystal structure has two DnaB subunits and two DciA molecules in the crystallographic asymmetric unit¹⁴. Applying crystallographic symmetry generates a 6:6 DnaB: DciA entity, revealing a dilated NTD layer and a constricted CTD layer. The nucleotide site of each VcDnaB CTD is filled with ADP. In the VcDnaB₂:VcDciA₂ asymmetric unit, the inter-CTD distance is ~32 Å, and the P-loop to arginine finger distance is 12.3 Å. The LH elements in 8A3V occupy the LH1 position.

Although parallels could be drawn between the 6T66, 7QXM, and 8A3V structures and the present VcDnaB-ssDNA structure, additional studies will be required to sequence them precisely in the reaction coordinate.

Structural Comparisons to *E. coli* DnaB-ssDNA Complexes

The RCSB PDB includes four unpublished entries corresponding to *E. coli* DnaB-ssDNA structures (PDB: 7T20, 7T21, 7T22, and 7T23^{62,120–122}). One structure was determined with the ATP analog AMP-PNP (7T20); the others harbor the ADP•AlF₄ transition state analog (7T21, 7T22, and 7T23). The ATP analog structure maps closely to the Vc structure, while the ADP•AlF₄ transition state DnaB structures map to the *Bst* complex. As such, conclusions made from comparisons between the Vc and *Bst* structures are fully compatible with the *E. coli* structures.

Divergent Solution Architecture of the DciA Helicase Loader

To understand the divergent solution architecture predicted by AlphaFold for Vc and *Mt* DciA (Supplementary Figure 28), we performed limited proteolysis mass spectrometry (MS) and native MS. Our proteolysis study encompassed two enzymes (Trypsin and Asp-N) with distinct substrate preferences (Supplementary Figures 12A and 12B). MS analysis of the proteolysis products revealed a series of cuts (red and magenta dots) that were concentrated on the Lasso regions of both DciA orthologs; in contrast, recognition sites for both proteases on the KH domains (gray dots) were protected from proteolysis. These data support the predicted inverted KH-Lasso arrangement in Vc and *Mt* DciA. We also showed using native MS that both DciA orthologs are monomeric in solution and do not bind ATP (Supplementary Figures 12C and 12D).

Oligomeric State and Stoichiometry of VcDnaB, DciA, ssDNA complexes

Next, we used high-resolution native mass spectrometry (nMS) to determine the oligomeric states and the stoichiometry of complexes in the DnaB-DciA system assembled from purified proteins (Supplementary Figure 8). We first investigated the oligomeric states of DnaB complexes from *Vc*, *E. coli*, and *B. subtilis* (Supplementary Figure 9A). nMS experiments on MtDnaB, with and without a SUMO tag, did not meet our quality control requirements and are not reported here. nMS analysis showed that *Vc* and *E. coli* DnaB helicases are hexameric, whereas *B. subtilis* DnaC is monomeric (Supplementary Figure 9A), as determined in previous studies^{34,64–66}. We observed a sizeable population of lower oligomeric states for Ec and VcDnaB (DnaB₅, DnaB₄, DnaB₃, and DnaB₂) in samples without Mg•ATP. The addition of a small amount of ATP (10 μM ATP for a 10 μM DnaB monomer sample) before nMS analysis yielded a predominant peak for the hexameric DnaB with minimal subcomplexes (Supplementary Figure 9A), indicating that ATP stabilizes the DnaB hexamer as previously suggested for the *E. coli* ortholog^{67,68}. All subsequent nMS experiments with DnaB-containing samples include 10 μM ATP and 0.5 mM magnesium acetate before nMS analysis. Close-up inspection of the nMS spectrum for the VcDnaB sample incubated with Mg•ATP showed a peak distribution corresponding to several Mg•ADP-bound states, indicating ATP binding and hydrolysis (Supplementary Figures 9B and Supplementary Table 5). Next, we examined the complexes between VcDnaB, VcDciA, and ssDNA derived from the OriC-I chromosomal origin. Our nMS analysis showed that VcDciA is monomeric (Supplementary Figure 10). Incubation of 10 μM (as monomer) VcDnaB with 7 μM full-length VcDciA, in the absence of ssDNA, revealed that the DnaB hexamer bound 1 – 2 copies of DciA, with small amounts of a third copy (Figure 4A). We attempted to analyze samples with higher relative amounts of DciA (> 10:7 DnaB: DciA) but encountered significant disassembly of

VcDnaB complexes with VcDciA upon initiation of electrospray during nMS analysis (Supplementary Figure 11). Stoichiometries of 6:3 or 6:6 for the VcDnaB-VcDciA complex have been estimated from small angle X-ray scattering, and the X-ray crystal structure^{12,14}.

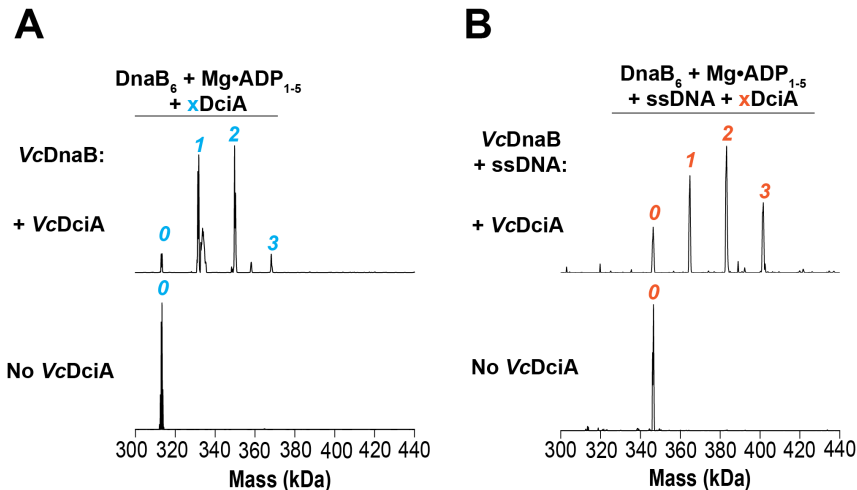


Figure 4. Up to Three Copies of VcDciA Bind to Isolated Hexameric VcDnaB and VcDnaB Complexed to Replication Origin-Derived ssDNA. The stoichiometry of reconstituted assemblies of VcDnaB, VcDciA, and origin-derived ssDNA. Deconvolved nMS spectra of 10 μ M VcDnaB (measured as the monomer) without and with 7 μ M full-length VcDciA (A) and in the absence or presence of 3 μ M origin-derived ssDNA (O4-VcF) (B). The number of VcDciA molecules observed bound to VcDnaB are indicated in blue (A) or red (B). The expected masses include DnaB₆ = 311.5 kDa, DciA = 18.4 kDa, and ssDNA = 32.5 kDa. All samples were exchanged into 10 μ M ATP, 500 μ M magnesium acetate, 500 mM ammonium acetate, and 0.01% Tween-20 before nMS analysis.

DnaB accepts ssDNA into its central chamber, the basis for the fork DNA unwinding assay^{12,42,69}. Accordingly, we found that VcDnaB alone could bind OriC-I-derived ssDNA (Figure 4B). nMS analysis of a mixture comprising VcDnaB, VcDciA, and the ssDNA showed that up to three copies of VcDciA could associate with ssDNA-bound DnaB (Figure 4B). Interestingly, there is a marked increase in the relative peak intensity for the 6:3 DnaB:DciA assembly when ssDNA is bound compared to the ssDNA-free sample, suggesting that ssDNA association promotes the formation or stabilization of this higher-order ensemble (Figure 4B).

We also investigated the binding of specific VcDciA domains to VcDnaB. We used limited proteolysis with nMS readout to validate the predicted arrangement of the KH and Lasso domains in VcDciA (Supplementary Information and Supplementary Figure 12). nMS analysis of VcDnaB titrated with increasing concentrations of the Lasso domain of DciA showed a progressive increase in the number bound to the DnaB hexamer with 1 – 4 copies bound when both proteins are present at equimolar amounts to 2 – 6 copies bound at four-fold excess of DciA-Lasso (Supplementary Figure 13A). In contrast, only one copy of the DciA KH domain was found associated with DnaB₆ at ~10% relative peak intensity even at six-fold molar excess, indicating that the KH domain of DciA weakly binds to DnaB (Supplementary Figure 13B). Incubation of VcDnaB with the two separate DciA domains at the highest concentrations studied did not change the maximum number of bound copies observed (i.e., one and six for the DciA KH and Lasso domains, respectively, Supplementary Figure 13B). Contrary to what was observed with DnaB and full-length DciA, no DnaB helicase disassembly was observed during the electrospray nMS analysis when either or both DciA domains were used in excess. Moreover, the distribution of DnaB:DciA-KH assemblies did not change (1 – 4 copies of DciA-KH remained bound to DnaB₆ at equimolar amounts of both proteins) in the absence or presence of ssDNA (Supplementary Figure 14).

Our nMS experiments revealed the stoichiometry of the isolated VcDnaB helicase and its complex with DciA and OriC-I-derived ssDNA; both estimates are crucial to deciphering the loading mechanism. The hexameric nature of VcDnaB implies that it is a closed protein ring and that the DciA helicase loader is likely a ring-breaker. The 6:3:1 stoichiometry of the VcDnaB-DciA-ssDNA complex suggests that although six DciA binding sites may exist on DnaB, only three are used; this value will be crucial to understanding how the DnaB:DciA will interact with other factors during replication initiation.

Vc and MtDciA Bind to Cognate DnaB Helicases with High Affinity

Next, we used orthogonal approaches to show that DciA binds to cognate DnaB and that the DciA binding is likely near the DH-LH interface (Figure 5). First, we conducted NiNTA ‘pull-down’ experiments to examine the binding between full-length and truncated constructs of DciA and DnaB (Figures 5A and 5B and Supplementary Figures 15 and 16). Full-length DciA (Vc and Mt) binds to full-length cognate DnaB. Vc and Mt DnaB constructs that span the LH and CTD domains also retain binding to full-length cognate DciA (a shorter construct spanning the CTD could not be expressed). No other DnaB construct (NTD or NTD-LH) showed binding to full-length cognate DciA. Parallel experiments with constituent KH (Vc: 1:98; Mt: 1:89) and Lasso domains (Vc: 98:157; Mt: 90:189) revealed that both bound to full-length cognate DnaB, although VcDciA KH domain bound weakly. Remarkably, although the KH and Lasso domain architecture of the Vc and Mt homologs of DciA exhibit is inverted, each Lasso domain binds to the cognate construct, including the LH-CTD domains of DnaB. No other truncated DnaB construct (NTD or NTD-LH) showed any interaction with either the KH or Lasso domains. These data point to the LH interface as a strong candidate for part of the DciA binding site, with another part comprised of the CTD itself.

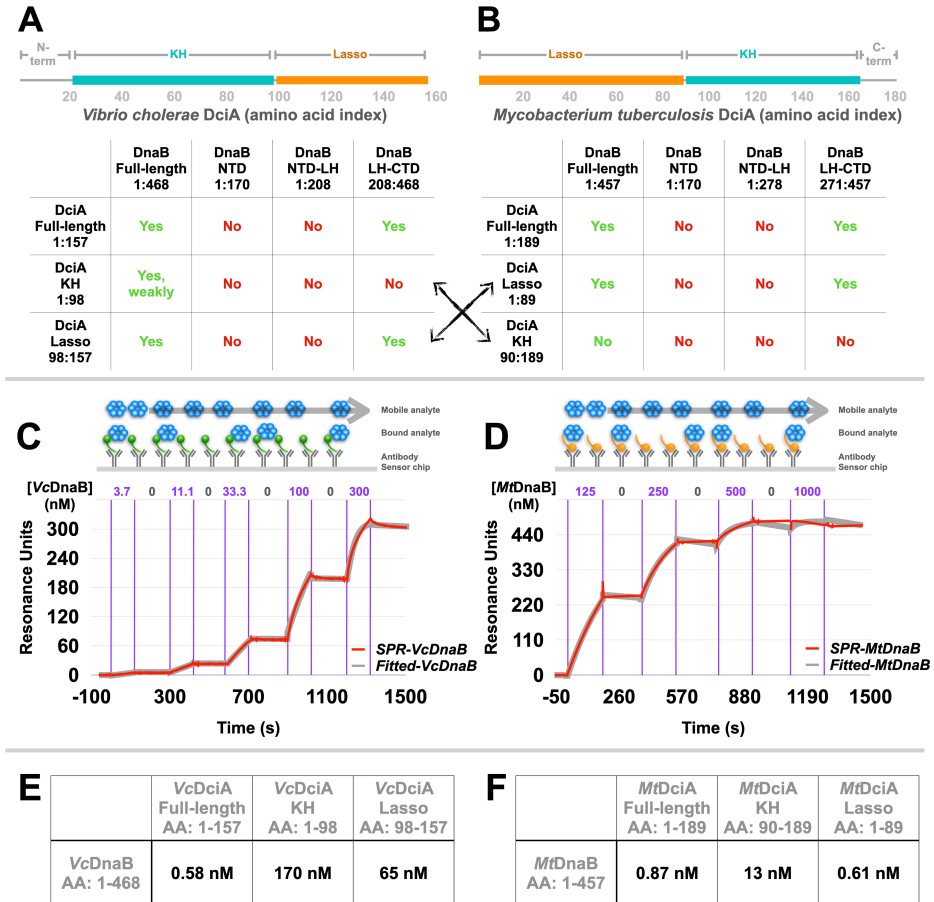


Figure 5. Although featuring distinct arrangements, the Vc and MtDciA loaders display comparable affinity to their cognate DnaB helicase. A) The primary sequence of VcDciA is arranged as a bar, and the locations of the KH and Lasso domains are indicated. The table below summarizes the NiNTA ‘pull-down’ experiments, wherein the green ‘yes’ and red ‘no’ indicate that the pull-down assay was scored to report binding or no binding, respectively (Supplementary Figure 15). B) Same as panel A except that the image and table capture studies with Mt proteins (Supplementary Figure 16). Strikingly, although the KH and Lasso domains in the two DciA orthologs feature an inverted arrangement, the Lasso domains of each specify binding to the LH-CTD construct of DnaB. The dissociation constants (K_D) between DnaB and DciA were evaluated through single-cycle kinetic analyses via Surface Plasmon Resonance (SPR). On and off rates of full-length DnaB at incremental concentrations (Vc: 3.7, 11.1, 33.3, 100, 300 nM; Mt: 125, 250, 500, 1000) were measured by immobilizing cognate C-term His-tagged full-length DciA, from Vc (C; $K_D = 0.59$ nM; $\chi^2 = 4.33$; $R_{max} = 312.3$) and Mt (D; $K_D = 0.87$ nM; $\chi^2 = 11.8$; $R_{max} = 487.8$) on the SPR sensor chip. A schematic of the SPR experiment accompanies the raw (in red) and fitted (in grey) SPR curves. Experimental data were fit as described in the methods. The same approach was applied to measure K_D s for the KH and Lasso constructs (Supplementary Figures 17 and 18). The K_D values emerging from the SPR analysis are summarized for the Vc (E) and Mt (F) systems.

Next, we applied surface plasmon resonance (SPR) to quantify the interaction between DnaB and DciA. We immobilized full-length DciA, the isolated

KH, and Lasso domains to the SPR chip, measured on and off rates for cognate full-length DnaB, and extracted dissociation constants (Figures 5C, 5D, 5E, and 5F). As our pull-down binding experiments anticipated, VcDnaB bound tightly ($K_D = 0.58$ nM) to full-length VcDciA. We found the isolated Vc Lasso domain bound to DnaB with a $K_D = 65$ nM, but the KH domain exhibited a weak affinity ($K_D = 170$ nM). Full-length MtDciA and the isolated Lasso and KH domains displayed similar K_D s to full-length MtDnaB (Supplementary Figures 17 and 18). The additional degrees of freedom to engage with MtDnaB enjoyed by the Mt Lasso and KH domains might explain why the truncated constructs bind with higher affinity than full-length.

The binding and SPR data suggest that, even though the Vc and MtDciA orthologues feature inverted KH and Lasso architectures, they bind to their cognate DnaB helicase using equivalent domains. Moreover, our data also reveal that the most significant contacts between DciA and DnaB are found at the LH and CTD domains of DnaB and that the Lasso domain is the locus of many of the most significant contacts to cognate DnaB. Given that both *E. coli* DnaC and phage λ P bind at the DH-LH interface of two adjacent DnaB subunits^{19,43,44,48}, the finding that the LH constitutes the DciA binding establishes this site as a conserved site with significance for DnaB helicase loading.

Suppression of DnaB's ATPase Activity by DciA

One of the hallmarks of a bacterial helicase loader is the suppression of ATPase activity of the cognate helicase^{43,46,60}. Our ATPase study first compared the wild-type and the Walker B mutant E259A VcDnaB at a single concentration (200 nM); we found robust activity in wild-type VcDnaB, which was lost in the mutant (Figure 6B). Next, we measured the influence of full-length and truncated constructs of DciA on the ATPase activity of DnaB. We found that titration of full-

length VcDciA revealed potent suppression of VcDnaB's ATPase activity (untagged DnaB: $IC_{50} \approx 26$ nM, NHis-DnaB: $IC_{50} \approx 26$ nM, Figure 6A).

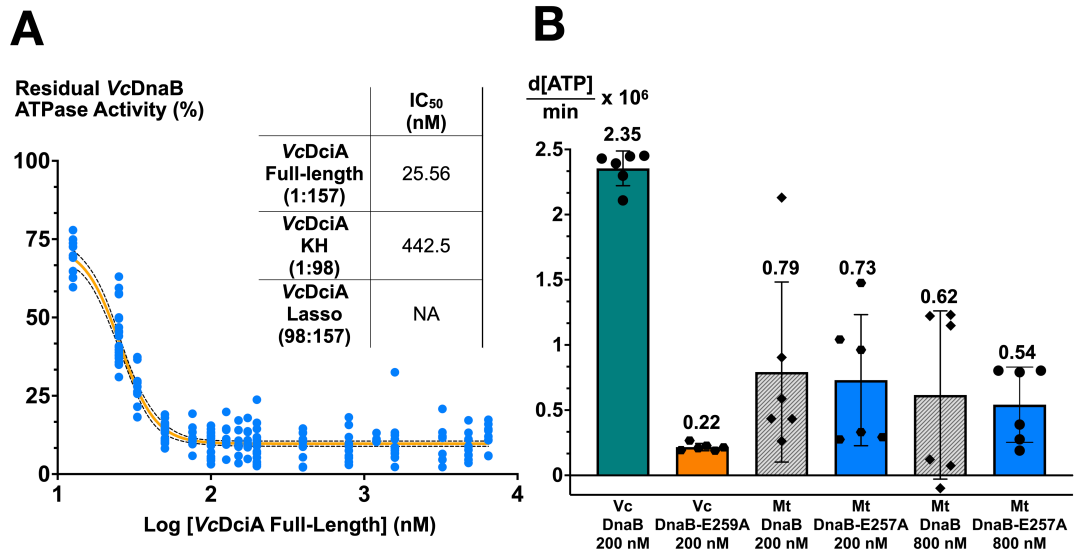


Figure 6. VcDnaB ATPase activity is suppressed by cognate full-length loader; MtDnaB is inactive. A) The effect of titrating VcDciA (0 - 6400 nM, log scale) on the residual ATPase activity of VcDnaB (200 nM) is plotted. For each experiment, the residual activity is the quotient of the ATPase rate of a particular DciA concentration and the rate for 0 nM DciA. The blue points represent seven experiments with different batches of VcDnaB and VcDciA on various days measured in triplicate. The light-orange line represents the non-linear fit of the residual activity points (Sigmoidal, 4PL, X is log (concentration) model in Prism. The dotted lines represent the 95% confidence interval. The fit encompassed 210 points, showed an R^2 value of 0.91, and produced an IC_{50} value of 25.56 nM (95% CI = 23.74 to 26.94 nM). Parallel analysis with the KH domain on four different batches revealed a 16-fold higher IC_{50} = 442.5 nM (R^2 = 0.9693; 95% CI: 380.5 to 520.0 nM). The Lasso domain was not found to inhibit; however, we observed activation of DnaB's ATPase at very high Lasso concentrations. The inset table summarizes the IC_{50} results. Control measurements with the E259A VcDnaB Walker B mutant exhibited no ATPase activity (panel B, orange). B) Investigation of the ATPase activity of MtDnaB. The ATPase rates for wild-type hardly exceed that measured from the Walker B mutant MtDnaB. The rate for 800 nM is lower than 200 nM MtDnaB. We conclude that MtDnaB, under the assay conditions, is not active. Positive (200 nM VcDnaB) and negative controls (200 nM VcDnaB-E259A) provide context for results with MtDnaB. Each experiment in panel B was performed two times in triplicate. The points represent the measured rates. The height of the bar represents the average of the measured points. The vertical bracket visually represents one standard deviation. Vc proteins are colored in deep teal and orange. Mt proteins are in blue and grey.

A construct encompassing the Vc KH domain (1:98) exhibited an $IC_{50} \approx 443$ nM (Supplementary Figure 19A); the weaker inhibitory constant is congruent with the weak interaction signaled by our nMS study and the NiNTA-based binding studies. Even though we found that the Vc Lasso domain (98:157) binds to full-

length VcDnaB, titration revealed no effect on VcDnaB's ATPase activity at low DciA concentrations (25 – 400 nM). However, at very high concentrations (800 – 6400 nM) of VcDciA, we observed reproducible stimulation (up to 2.5 x) of DnaB's ATPase activity (Supplementary Figure 19B).

Parallel analysis of *MtDnaB* and *MtDciA* unexpectedly revealed that ATPase activity could not be measured reliably and appeared inconsistent even when the concentration was increased fourfold. Our measurements showed that wild-type and Walker B mutant E257A *MtDnaB* exhibited similar catalytic rates (Figure 6B); these data suggested that *MtDnaB* was inactive in our hands (Supplementary Figure 20), and, as such, we could not test *MtDciA* for inhibitory activity towards *MtDnaB*.

VcDciA Stimulated Unwinding by VcDnaB Measured Using a Replication Fork-Mimic

Every bacterial helicase loader stimulates the DNA unwinding activity of its cognate DnaB helicase. To investigate the influence of DciA on cognate DnaB, we applied an ensemble FRET-based DNA unwinding assay wherein a DNA replication fork mimicking substrate²⁰ is fitted with a Cy3 labeled nucleobase Watson-Crick-Franklin paired with a nucleobase conjugated to the BHQ2 quencher; the unwinding of the fork substrate is readout by recording appearance of a fluorescent light signal at 570 nm (Supplementary Figure 21). This assay showed that 200 nM Vc DnaB exhibited measurable albeit low DNA unwinding activity. However, titration of full-length Vc DciA (0.11 to 6400 nM) revealed a 10-fold (15 to 150) increase in fluorescence at the highest concentration (6400 nM) (Supplementary Figure 21B). Unwinding depended on ATP and was wholly absent in Vc DnaB Walker B mutant E259A (Supplementary Figure 22). Parallel analysis of the isolated

DciA KH and Lasso domains implied retention of about 30% of the activity of the full-length VcDciA to activate the helicase. (Supplementary Figures 21B and 23).

DciA Loads DnaB on Physiological DNA Bubble Substrates

The nMS binding data and DNA unwinding experiments (Supplementary Information and Supplementary Figures 21, 22, 23, and 24) implicate DciA as a helicase loader since DciA enhances both underlying activities. However, the substrates employed in the above assays (ssDNA, 5' and 3' ssDNA tailed fork) do not permit self-threading into DnaB's central chamber to be distinguished from loading/assembly around ssDNA. Physiologically melted DNA by DnaA at a replication origin has no free termini, and, as such, assembly around ssDNA, rather than threading, is expected to be the predominant form of loading in vivo.

To distinguish the DciA-mediated assembly of VcDnaB on ssDNA from self-threading, we measured unwinding from the DNA bubble substrate, which features no free ends, unlike a tailed-fork substrate. Our assay employed DNA molecules #894 and #895 described in reference⁴², which we converted into a FRET-based readout by adding a Cy3 fluorophore at one 5' end and the Iowa Black RQ quencher at the corresponding 3' end (Figure 7). We also measured DnaB loading onto a doubly labeled DNA bubble (Supplementary Information and Supplementary Figure 25, 27C, and 27D).

No fluorescence signal was observed when DciA was omitted (Figure 7, red curve), leading us to conclude that self-threading of ssDNA into DnaB had been suppressed by using the bubble substrate. Titration of full-length VcDciA (0.11 to 6400 nM) revealed a 10-fold (~15 to ~150) increase in fluorescence at the highest concentration (6400 nM). The increase in fluorescence strictly depended on ATP since the VcDnaB E259A mutant showed no activity (Supplementary Figure 26). Substitution of the two DciA truncations – the NTD KH and the CTD Lasso – for full-

length VcDciA revealed essentially no loading activity (less than 5%) (Supplementary Figures 27A and 27B).

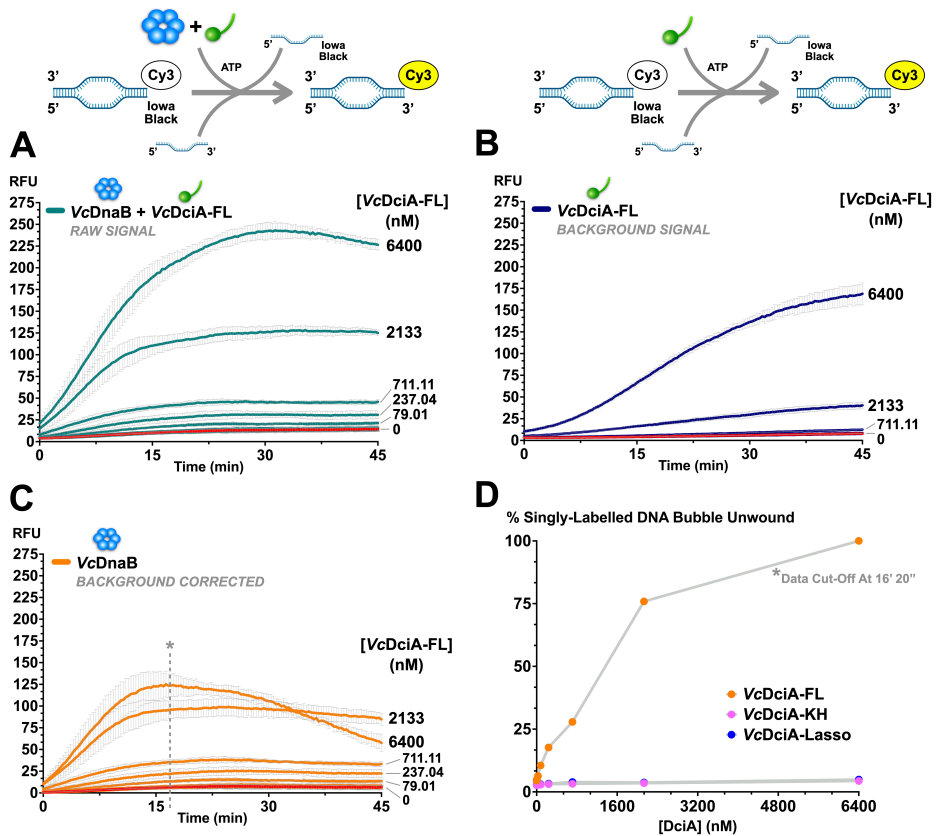


Figure 7. VcDnaB is loaded onto a physiological DNA bubble substrate by the full-length VcDciA loader. Helicase loading was measured using a DNA substrate that mimics unwound DNA at a replication origin as is diagrammed in the schematic above. A) Loading of full-length VcDnaB by full-length VcDciA. Loading is read out by measuring fluorescence (in relative fluorescence units, RFU) over 45 minutes at 570 nm from a DNA substrate after it had been unwound by DciA-loaded DnaB to relieve quenching. Data were taken by titrating VcDciA-FL (0.11-6400 nM) against VcDnaB (200 nM). Each labeled curve represents the average of 6 independent measurements taken from different batches of proteins on different days. The lack of loading activity in the curve that omits VcDciA (red curve) implies that DnaB's self-threading activity has been suppressed with the replication-origin mimicking DNA bubble substrate. The standard deviation of each point is plotted in light gray. B) Same as Panel A except that VcDnaB is omitted. Envisioned as a control, this measurement teaches that VcDciA appears to have the capacity to unwind bubble DNA on its own. C) Point-by-point subtraction of the curves in panels A and B estimates the bubble unwinding signal that arises solely from VcDnaB. The downward phase of the 6400 nM DciA curve after 16 minutes arises from unwinding rate differences between VcDnaB and VcDciA. These data suggest that VcDciA stimulates the loading of VcDnaB by 25-fold (~5 to ~125 RFU). D) The unwinding data in panel C is converted to a percentage unwound by dividing by the maximum RFU point and plotting the resulting values against VcDciA concentration. Parallel measurements reveal that the isolated VcDciA KH and Lasso domains (Supplementary Figure 27A and 27B) exhibit little (less than 5%) to no loading activity.

To better understand VcDciA's loading activity, we repeated the above titration without VcDnaB. Surprisingly, titration of DciA against the DNA bubble substrate revealed a substantial fluorescence signal (Figure 7B); these data imply that VcDciA has a previously undiscovered activity: to unwind dsDNA on its own. This activity did not require ATP. The finding that isolated VcDciA produced a fluorescence signal ~66% of that seen in the VcDnaB•DciA titration implies significance, though establishing biological relevance will require further experimentation. DciA's apparent unwinding activity is not observed when the DNA bubble is substituted with the fork substrate (Supplementary Figure 24), pointing to a role for DNA structure in this activity.

We subtracted the two signals to separate the unwinding activities of VcDnaB•DciA and isolated VcDciA (Figure 7C). The calculated signal from VcDnaB, when acted upon by VcDciA, showed a 25-fold (~5 to ~125 RFU, Figure 7C) increase in fluorescence at the highest concentration (6400 nM) (Figure 7D). These data strongly argue that VcDciA is a helicase loader for VcDnaB.

The congruence of the DNA unwinding and helicase loading curves during titration with full-length VcDciA suggests that, to a large degree, helicase loading rather than threading occurs with full-length DciA in both assays (Figure 7D and Supplementary Figure 21B). On the other hand, dissimilarity in the curves corresponding to the VcDciA truncations implies that loading is not retained in either construct or that the observed DNA unwinding activity in the truncated constructs may arise solely from self-threading. Our data establish a helicase-loading biochemical activity for the VcDciA protein. Moreover, the finding that loading onto the bubble substrate does not occur absent DciA supports a closed ring configuration for DnaB.

VcDciA Stimulated Unwinding by VcDnaB Measured Using a Doubly Labeled DNA Bubble Substrate

The bubble DNA substrate (Figure 7 and Supplementary Figure 25), in principle, permits the loading of DnaB onto both single strands, leading to unwinding in both directions; however, the bubble substrate above with the single FRET pair only reports on activity at one end of the DNA. To examine activity at both ends, we repeated the helicase loading assay with a bubble substrate that featured the above FRET pairs at both termini. Full-length DciA and the two truncated constructs displayed the same pattern of activity as seen with the singly labeled substrate (Supplementary Figures 25, 27C, and 27D). Full-length DciA produced 10% more fluorescence on the doubly labeled substrate instead of the expected doubling of the signal. This discrepancy may be due to unexpected quenching between the two ends of the molecule due to conformational switching that brings them closer together. Alternatively, loading two DnaB molecules may be more complex than previously thought.

DISCUSSION

Our analyses of the bacterial DciA loader and the DnaB helicase from *Vibrio cholerae* and *Mycobacterium tuberculosis* sheds new light on the underlying helicase-loading and activation pathway (Figure 8).

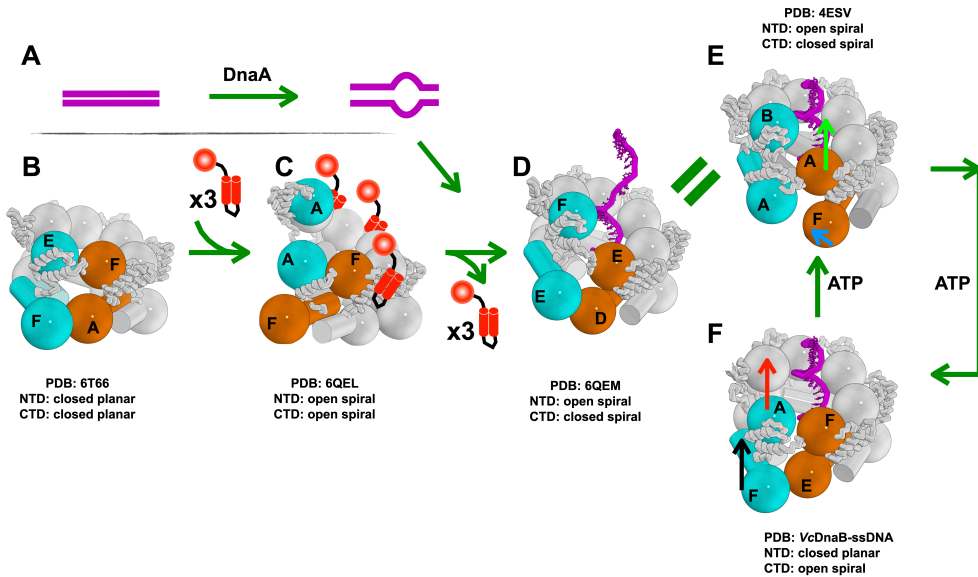


Figure 8. Model for Loading of DnaB by DciA and Translocation of DnaB on ssDNA. A) The DnaA protein binds to the replication origin and promotes melting into a DNA bubble. B) The hexameric closed planar DNA-free VcDnaB (PDB: 6T66¹²) provides a ground state model. C) Our data suggests that up to 3 DciA loaders (red) bind to VcDnaB; we suggest that DnaB adopts an open spiral configuration analogous to that seen with the DnaC/λP complexes^{19,43,44,48}. D) Remodeling of this complex accompanies the entry of DnaA-generated ssDNA into the central chamber of DnaB. The current structures suggest that the D and E conformations are identical (19). We suggest that DnaB populates the two states captured in the GDP•AlF₄ (E) and ATPγS (F) conformations during translocation. The first translocation step (panel F to panel E) involves opening the NTD tier and the migration of a CTD in the ATPγS complex to its position in the GDP•AlF₄ complex. The ATPγS-chain F NTD shifts (black arrow) in position to the location of the GDP•AlF₄-chain A. Opening of the NTD layer is accompanied by upward migration by 25 Å of the cyan CTD chain A in the ATPγS complex (panel F, red arrow) to the position occupied by chain B in the GDP•AlF₄ complex (panel E); this represents the first translocation step. In the second step (panel E to panel F), the GDP•AlF₄-NTD layer (chain F) returns (blue arrow) to the ATPγS-closed planar configuration (chain E). This change is accompanied by a 25 Å upward migration of the orange CTD chain A in panel E to a new position (chain F) in panel F (green arrow). Repetition of these two steps around the DnaB enables each subunit to translocate. We suggest that the DciA-loaded DnaB-ssDNA complex, which resembles the GDP•AlF₄ conformation, enters this cycle and then transitions to the ATPγS state. The DnaB models in panels B-F are colored in white except for certain blue and orange CTDs and NTDs that line the various breaches, except for the closed planar entity

In bacteria, two copies of the DnaB replicative helicase are assembled onto single-stranded DNA introduced into the replication origin by the DnaA initiator protein. It is now clear that most bacteria lack an ortholog for the well-studied DnaC helicase loader; instead, the DciA helicase loader appears likely to be the predominant loading factor^{37,40,41}. Vc and MtDciA bind to cognate hexameric DnaB with a low nM K_D . Although their KH and unstructured Lasso domains feature an inverted organization, the unstructured Lasso segment mediates significant binding energy to the interaction with DnaB. Also, irrespective of the divergent internal organization, an essential part of the DciA binding site lies at the nexus of DnaB's DH and LH elements. Prior studies suggested stoichiometries of 6:3 or 6:6 for the VcDnaB-VcDciA complex and a binding site at the DH-LH nexus^{12,14}. We find up to three copies of VcDciA and one copy of replication-origin-derived ssDNA bound to VcDnaB. The stoichiometry of the DnaB-DciA system distinguishes it from the *E. coli* DnaB-DnaC and DnaB• λ P complexes, which feature 6:6 and 6:5 stoichiometries, respectively^{19,43,44,48}. The unwinding activity of DnaB is significantly stimulated by DciA on fork and replication origin-mimicking DNA bubble substrates. The finding that isolated VcDnaB is a hexamer under physiological concentrations of ATP (Figure 4) and exhibits essentially no self-loading on the replication-origin mimicking DNA bubble substrate (Figure 7) leads us to conclude that DciA is a ring-opening-type helicase loader.

Our structure of the ATP γ S form of the DnaB-ssDNA complex provides new insights into the mechanisms of translocation. The topologically closed NTD and distinctly configured CTD layers distinguish the ATP γ S state from a structure previously determined bound to the transition state ATP analog GDP•AlF₄¹⁰ and provide a first view of nucleotide-dependent conformation switching during ssDNA translocation. Our model for ATP-dependent translocation by DnaB extends the hand-over-hand mechanism¹⁰ by integrating cycles of opening/closing of the NTD

layer, reconfiguration of the LH element, and dissociation of individual CTDs from ssDNA and neighboring subunits followed by their migration to the top of the DnaB spiral to new ssDNA binding sites. Migration to the top of the spiral is facilitated by the engagement of its β -hairpin arginine finger with a freshly bound ATP by the DnaB subunit that precedes it at the top of the spiral; ATP binding stabilizes contacts to ssDNA. We suggest that translocation encompasses three identical sub-steps. Each sub-step involves migration of the chain A CTD (or chain E or chain C) with the simultaneous opening of the NTD layer into an open spiral, as seen in the Bst4ESV structure. In the second step, migration of the next CTD (chain F, chain D, or chain B) in sequence, newly at the bottom of the spiral, is accompanied by the closure of the NTD layer to the triangular planar arrangement seen in the Vc ATP γ S structure. Our model specifies that two DnaB subunits move in sequence and that this program of motion is repeated thrice around the hexamer. Our structure may inform debate on whether the DnaB subunits' motion is sequential, stochastic, or concerted^{10,22,28,29,70–74}.

To link the significant findings of our work in DciA-mediated loading and subsequent translation by DnaB, we looked to another ring-opener, DnaC, to provide models for other steps in the helicase loading pathway (Figure 8). The configuration of DnaB in the *E. coli* DnaB-DnaC•ssDNA (PDB: 6QEM) complex¹⁹ closely resembles that of the 4ESV GDP•AlF₄ state¹⁰, with a closed spiral CTD and open spiral NTD and contacts to ssDNA by each subunit. Since the 6QEM structure has an unfilled nucleotide site at the top of the spiral (as in the pose in Figure 8), on the departure of the loader, it needs only bind ATP to align with the 4ESV structure and enter the two-DnaB subunit motion (Figure 8E-F) scheme proposed above. Notably, the departure of the DciA loader produces a DnaB configuration after it has taken the first of the two sequential steps described above (Figure 8D). We suggest that after the eviction of the loader, the subunit at the bottom of the

spiral will take the second of these steps to migrate to the top of the spiral (Figure 8E-F). Repetition of the two DnaB subunit sequenced motion around the ring accompanies further translocation along ssDNA. Notably, a crystal structure of the 6:6 VcDnaB-VcDciA complex implies two distinct DciA binding modes and a closed planar DnaB¹⁴. Additional studies will be required to sequence this complex in the helicase loading reaction.

An important unanswered question centers on the mechanism of expulsion of DciA from the DnaB-ssDNA complex. The DnaC^{19,60,75} and λ P^{76,77} systems feature biochemical programs to mediate loader dissociation. Our data reveals a tight interaction between helicase and loader (Figure 5) that suppresses DnaB's ATPase activity (Figure 6). Yet, DciA exhibits a dose-dependent stimulation of DnaB's unwinding activity (Figure 7 and Supplementary Figure 21), absent any other factor. These findings suggest that DciA may spontaneously dissociate to relieve inhibition of DnaB's ATPase activity and enable unwinding activities. It is also possible that interactions with other replisome components^{25,75,78} could contribute to the eviction of DciA. Future studies will have to clarify this question.

MATERIALS AND METHODS

Plasmids, Protein Expression, Purification, and Biochemistry

Standard techniques⁷⁹ were used to develop the 21 protein expression plasmids and express and purify the encoded proteins. Our studies encompassed efforts with full-length and designed truncations of DnaB and DciA from *Vibrio cholerae* and *Mycobacterium tuberculosis*; our collection included histidine and SUMO-tagged constructs. The DnaB constructs included the full-length sequence (Vc: 1-468; Mt: 1-457), the DnaB amino-terminal domain (NTD; Vc: 1-170; Mt: 1-170), the DnaB amino-terminal domain fused to the Linker Helix (NTD-LH; Vc: 1-208; Mt: 1-278), the Linker helix fused to the DnaB carboxy-terminal domain (CTD-LH; Vc: 208-468; Mt: 171-457) and the carboxy-terminal domain (CTD; Vc: 208-468; Mt: 279-457). The DciA constructs included the complete sequence (Vc: 1-157; Mt: 1-189), the KH domain (Vc: 1-98; Mt: 90-189), and the Lasso domains (Vc: 98-157; Mt: 1-89). Notably, the *MtDnaB* gene (1:874) contains an intein (residues 400:815)^{61,80,81}; our expression plasmids omitted the intein. To keep the *MtDnaB* numbering consistent with other orthologs that lack inteins, we numbered the sequence from 1-457, where the intein was excluded. The supplementary methods section (Supplementary Methods and Supplementary Tables 3 and 4) completes the details on the construction of the plasmids, expression, bacterial expression conditions, and purification.

Expression Plasmid Construction

Genes encoding full-length wild-type and Walker B point mutant (E257A) *Mycobacterium tuberculosis* *MtDnaB* (NCBI accession number CCP42780.1) were synthesized and cloned into the pET28a vector (Genscript) to generate an amino-terminal hexahistidine-SUMO-tagged *MtDnaB* (*MtDnaB*-SUMO_pET28a) or cloned

into the pET16b vector (Genscript) to express untagged *MtDnaB*. *MtDnaB* gene (1:874) contains an intein (400:815)^{61,80,81}, which was not included in the expression plasmid. To keep the *MtDnaB* numbering consistent with other orthologs that lack inteins, we numbered the sequence from 1-457, where the intein was excluded. Four truncated variants of *MtDnaB* (NTD: residues 1:170; CTD: residues 279-457; NTD-LH Residues 1:278; LH-CTD Residues 171:457) were subcloned into a pCDFDuet-1 plasmid (Genscript) for protein expression.

Genes encoding full-length *MtDciA* (NCBI accession number NP_214518.1) (*MtDciA*_pET22b), *MtDciA*-NTD (residues 1-89, *MtDciA*-NTD_pET22b), or *MtDciA*-CTD (residues 90-189, *MtDciA*-CTD_pET22b), as fusion proteins with a C-terminal 6xHis were cloned into pET22b vector for protein expression.

Protein Expression

MtDnaB. Cells bearing plasmids encoding NHis-SUMO-*MtDnaB* (*MtDnaB*-SUMO_pET28a) were grown in 2xYT medium¹²⁷, supplemented with kanamycin (50 µg/mL). The culture was permitted to grow at 37 °C to an OD600 ≈ 0.6, after which IPTG was added to 0.1 mM. Protein expression was allowed to continue for 20 hours at 17 °C (RPM = 210), after which the cells were harvested by centrifugation at 4 °C and frozen at -80 °C.

MtDnaB-untagged (*MtDnaB*-WT_pET16b) was expressed using the same procedure as adopted for the other *MtDnaB* constructs, except that the culture included 50 µg/mL ampicillin, was induced with 0.15 mM IPTG, and protein expression continued for 18 hours.

Untagged truncations here, *MtDnaB*-NTD (*MtDnaB*-NTD_pCDF-DUET-1), *MtDnaB*-NTD-LH (*MtDnaB*-NTD-LH_pCDF-DUET-1), *MtDnaB*-LH-CTD (*MtDnaB*-LH-CTD_pCDF-DUET-1) and *MtDnaB*-CTD (*MtDnaB*-CTD_pCDF-DUET-1) were expressed using the same procedure as for *MtDnaB*-untagged, except for the

inclusion of Streptomycin (50 µg/mL) during culturing and addition of 0.5 mM IPTG for protein expression.

MtDciA. MtDciA full-length (MtDciA_pET22b), and truncated variants MtDciA-NTD (MtDciA-NTD_pET22b), and MtDciA-CTD (MtDciA-CTD_pET22b), were expressed using the same protocol. Transformed bacterial cells were grown in a 2xYT medium supplemented with Ampicillin (50 µg/mL). Cells were grown to an OD₆₀₀ ≈ 0.6 at 37 °C; then 0.5 mM IPTG was added to the culture, and protein expression was allowed to continue for 3 hours at 30 °C (RPM = 210), after which the cells were harvested by centrifugation at 4 °C and frozen at -80 °C.

Protein Purification

MtDnaB. MtDnaB-NHis-SUMO containing bacterial cell pellet was thawed and resuspended in Buffer A containing 20 mM Tris-HCl pH = 8, 500 mM NaCl, 5 mM MgCl₂, 0.1 mM ATP, 20 mM imidazole, and 5% glycerol, supplemented with a protease inhibitor cocktail (ThermoScientific p/n 78439) at a volume ratio of 100 µL per 40 mL of lysis buffer. The resuspended cells were disrupted by sonication (total sonication ON time: 6 minutes). The soluble fraction from this procedure was incubated with Ni²⁺-nitrilotriacetic acid (Ni-NTA) beads equilibrated in Buffer A for 90 minutes. A 10-column-volume wash of Buffer A supplemented with 80 mM imidazole was followed by treatment of Buffer A containing 250 mM imidazole to elute MtDnaB-NHis-SUMO. Subsequent incubation (60 min, 22 °C) with 2 mM DTT and the SUMO protease ScUlp1 at a 1:250 protease:substrate molar ratio released MtDnaB-WT. The preparation was completed by concentrating the sample to 5 mL followed by chromatography over a HiLoad 16/600 Superdex 200 prep grade column equilibrated in 20 mM Tris-HCl pH = 8, 500 mM NaCl, 5 mM MgCl₂, and 5% glycerol. Fractions containing MtDnaB-WT were pooled, concentrated up to 1 mg/mL, frozen, and stored at -80 °C. MtDnaB point mutants Walker A (K232A), and

Walker B (E257A) were prepared in an identical manner as for the wild-type sequence.

Bacterial cell lysates with overexpressed truncated *MtDnaB* constructs (1-170, 1-278, 171-457, 279-457) were used in pull-down experiments but not purified.

MtDciA. The biomass obtained from 1 L of bacterial culture was thawed and resuspended in Buffer A containing 50 mM potassium phosphate buffer pH = 8, 500 mM NaCl, 5 mM imidazole, 5% glycerol, supplemented with a protease inhibitor cocktail (ThermoScientific p/n 78439) at a volume ratio of 100 μ L per 40 mL of lysis buffer. Cells were disrupted by sonication (total sonication ON time: 6 minutes). The soluble fraction was incubated with Ni-NTA resin beads for 90 minutes, previously equilibrated with buffer A. After washing the resin with 10 column volumes (CV) of Buffer A supplemented with 40 mM imidazole, *MtDciA* was eluted with Buffer A containing 300 mM imidazole. High concentration fractions were pooled, concentrated and loaded onto a HiLoad 16/600 Superdex 200 prep grade gel filtration column pre-equilibrated with 20 mM HEPES pH = 8, 300 mM NaCl. Fractions containing highly pure *MtDciA* were pooled, concentrated and stored at -80 °C. The same procedure was applied also to obtain truncated variants *MtDciA*-NTD and *MtDciA*-CTD.

Ni-NTA pull-down experiments (*Mt* proteins)

Interactions between *DnaB* and *DciA* and their respective truncations were interrogated by mixing and co-lysing bacteria that separately expressed each protein, where only one of the proteins was histidine-tagged. The lysate was passed over NiNTA agarose (QIAGEN), washed of contaminants, and histidine-tagged proteins were eluted. Interactions were scored based on the propensity of

the untagged protein in the experiment to be retained on Ni-NTA agarose (QIAGEN) compared to trials where the histidine-tagged protein was omitted.

A C-terminal 6xHis-tag *MtDciA* construct was used as bait, and one untagged *MtDnaB* construct was used as prey. The same protocol was exploited for every combination reported. Two bacterial pellets, independently expressing *DciA* and *DnaB*, harvested from 500 mL of bacterial cultures were thawed and mixed in Buffer A containing 20 mM Tris-HCl pH = 8, 300 mM NaCl, 5 mM MgCl₂, 5% glycerol. Cells were disrupted by sonication (total sonication ON time: 6 minutes). The lysate was clarified at 14,000g for 45 min at 4 °C, and the supernatant was incubated with Ni-NTA beads, previously equilibrated with Buffer A for 60 minutes, while shaking. Following the incubation, the beads were washed with 10 column volumes of Buffer A supplemented with 15 mM imidazole. Bound proteins were then eluted with Buffer A containing 250 mM imidazole. The collected protein fractions were analyzed using SDS-PAGE.

Surface Plasmon Resonance (SPR)

Dissociation constants (K_D) were estimated using surface plasmon resonance (SPR) from measurements of the on and off rates between *DnaB* and *DciA*. SPR data were acquired using a Biacore X100 instrument (GE Healthcare/Cytiva) and a CM5 sensor chip functionalized with an anti-histidine antibody according to the manufacturer's instructions. All measurements were conducted at 25 °C. Since SPR response is proportional to molecular weight, we chose to immobilize the smaller N-His-*DciA* (Vc: MW = 18.4 kDa; Mt: MW: 21 kDa) while positioning the much larger *DnaB* (Vc: MW = 312.3 kDa; Mt: MW: 301.5 kDa), in the mobile phase. *DciA* was immobilized on the surface of a single channel of the CM5 sensor chip via the interaction between the histidine tag and the anti-histidine antibody. The appropriate ligand density (RL) on the chip was calculated

according to the equation: $RL = \left(\frac{\text{ligand MW}}{\text{analyte MW}} \right) \times R_{\max} \times \left(\frac{1}{S_m} \right)$ where $R_{\max}$ is the maximum binding signal, and S_m corresponds to the binding stoichiometry. With the immobilization procedure complete, non-specifically bound proteins were removed by washing with buffer HBS-EP+ (Cytiva), until the resonance signal reached a constant baseline. The reference flow cell provided a control surface for refractive index change and non-specific binding. DnaB solution was injected five times at increasing concentrations over the DciA-functionalized sensor surface in a single-cycle kinetics experiment. A long dissociation phase and a single regeneration step followed the last sample injection without regeneration between each sample injection. The interaction between DnaB and DciA was assumed to display 1:1 kinetics; as such, the measured SPR data should display pseudo-first-order kinetics that is described by the following rate equation:

$$(1) \quad \frac{d[LA]}{d[t]} = k_a \cdot [L] \cdot [A] - k_d[LA]$$

where k_a is the association rate constant, which refers to the rate of complex formation (units: $M^{-1}s^{-1}$); the dissociation rate constant k_d describes the stability of the complex, i.e., the fraction of complexes that dissociate in the time frame (units: M). $[A]$ represents the analyte concentration; $[LA]$ is the complex concentration measured directly as R in Resonance Units (RU); $[L]$ is the ligand concentration on the surface.

To enable fitting of the SPR data, equation (1) was modified as follows:

$$(2) \quad \frac{d[R]}{d[t]} = k_a \cdot [C] \cdot [R_{\max} - R] - k_d[R]$$

where $R_{\max}$ is the maximum binding capacity, and $R_{\max} - R$ is the free concentration of the analyte. $R_{\max}$ and R were extracted from the experimental SPR curves. The fitted data proved estimates for k_d and k_a from which $K_D = k_d/k_a$ was calculated. Experimental binding kinetic data were fit to the '1:1 binding' model implemented in the Biacore X100 Control Software (Cytiva).

ATPase Rate Measurements

An NADH-coupled microplate spectrophotometric assay was used to measure rates of ATP hydrolysis^{86,87} by VcDnaB and how rates are affected by the presence of our suite of full-length and truncated constructs of VcDciA. In all experiments, VcDnaB was kept fixed at 200 nM, while titrations of DciA encompassed concentrations between 0 nM and 6400 nM.

NADH oxidation data were recorded in the presence of 2 mM ATP (Gold Biotechnology, Inc. p/n: A-081-1), 2 mM phosphoenolpyruvic acid monopotassium salt (PEP), 0.01 U/ μ L lactate dehydrogenase (Sigma-Aldrich, L1254), 0.002 U/ μ L pyruvate kinase (Sigma-Aldrich, p/n: P7768,) and 0.3 mM NADH in the following buffer: 50 mM potassium-HEPES – pH=7.5, 150 mM potassium acetate, and 8 mM magnesium acetate, 5 mM β -mercaptoethanol, and 0.25 mg/mL bovine serum albumin (Sigma-Aldrich p/n: A3912). The above components were combined, save for ATP, into wells of a Corning 96-well Clear Flat Bottom UV-Transparent Microplate (p/n: 3635), and the plate was warmed to 37 °C for 5 minutes. After the quick addition of 2 mM ATP, the plate was inserted into a SpectraMax M5 plate reader (Molecular Devices), and absorbance data was taken at 340 nm every 30 s over 30 min. Raw data were exported in text format and imported into Microsoft Excel for analysis. Each absorbance curve was manually inspected to identify a 10-minute period during the experiment wherein the slope of the 340 nm absorbance vs time was constant. The values in the above period were divided by the (-)6.22 mM⁻¹ cm⁻¹ extinction coefficient of NADH and the pathlength in cm calculated by the plate reader to result in the rate of ATP hydrolysis in mM/s; further division by the mg of DnaB in the assay provided the specific activity in mM/s/mg.

FRET-based DNA Strand Displacement Assay

DNA unwinding by DnaB and its stimulation by DciA was measured using a fluorescence resonance energy transfer (FRET) helicase strand displacement assay adapted from. The DNA substrate for the assay was identical to that in reference²⁰ and termed FRET-DNA-fork-top-strand, FRET-DNA-fork-bottom-strand, and FRET-DNA-fork-capture-oligo. The top strand is labeled with Cy3 at its 5' end; the bottom strand is labeled with the BHQ_2 quencher at its 3' end. The capture oligo is unlabeled (Supplementary Table 4 and Supplementary Figures 21, 22, 23, and 24). Strand displacement was measured by mixing annealed and fluorescence-quenched substrate at a concentration of 100 nM with 200 nM DnaB and DciA concentrations ranging from 0 mM to 6400 nM. An excess of the capture oligo (400 nM) was included to prevent the reannealing of the fork after strand displacement. FRET data was taken at 37 °C in the following buffer: 20 mM K-HEPES pH = 7.5, 5 mM Mg acetate, 50 mM K-glutamate, 5% glycerol, and 4 mM dithiothreitol (DTT). Reaction components were mixed into wells (volume: 100 µL/well) of a Corning black plate (p/n: 3880) with 96 half-area clear flat bottom wells. The working concentration of NaCl in the assay is 50 mM. The reaction was initiated by adding 1 mM ATP to each well; the plate was subsequently inserted in a SpectraMax M5 plate reader (Molecular Devices). FRET signals were recorded for 30 min at 37 °C. To avoid the crosstalk between Cy3's closely spaced excitation and emission signals, an excitation wavelength of 520 nm was used; at this wavelength, excitation is reduced to 50% of its maximal value; however, this wavelength does not overlap with the fluorescence emission peak, measured at 570 nm. Relative Fluorescence Unit (RFU) data taken in this way were exported for analysis in Microsoft Excel and GraphPad Prism.

FRET-based Replicative Helicase Loading Assay

To measure helicase loading by DciA, we adapted the unwinding assay described above; instead of a fork substrate with 5' and 3' unpaired segments, we employed substrates that featured single-stranded 'bubbles' (Figure 7 and Supplementary Figures 25, 26, and 27) whose sequences are taken from (#894 and #895, reference⁴²). Bubble substrates favor DnaB assembly on ssDNA instead of threading the free end of ssDNA end in the DnaB central chamber. In our first bubble substrate, the top strand (as drawn in Figure 7) was singly labeled at the 5' end with Cy3, and the bottom strand with the Iowa Black RQ quencher (IDT). To account for the possibility that DnaB helicase could, in principle, be loaded for translocation in either direction, we employed a second substrate wherein fluorophore – quencher pairs were installed at both ends of the 'bubble' (Supplementary Information and Supplementary Figures 25, 27C and 27D); such a substrate could inform on the loading and translocation on both strands and in both directions. As with the strand displacement assay, unlabeled capture strands forestalled the reannealing of the quencher-containing bubble strand.

FRET signals from helicase loading were recorded and analyzed as described for the strand displacement assay.

Sample and Grid Preparation for Single Particle Cryo-EM

To prepare complexes for cryo-EM, untagged VcDnaB-FL and VcDciA-FL-CHis were diluted in 20 mM HEPES pH = 7.5, 200 mM NaCl, 5 mM MgCl₂, 5 mM β -mercaptoethanol, and 5 mM ATP γ S to reach a final concentration of 18 μ M (DnaB monomer) and 12.7 μ M (DciA). A ssDNA oligonucleotide bearing the origin of the replication sequence of *Vibrio cholerae* serotype O1 (Supplementary Table 4, oligo: O1b-VcF) was added at a concentration of 5.4 μ M. 3 μ L of the VcDnaB-DciA-ssDNA sample were applied to a Quantifoil UltrAuFoil R1.2/1.3 Au300 grid,

previously plasma cleaned with a 3:1 ratio of Ar/O₂ at 15 W for 10 seconds in a Gatan Solarus II glow discharger. Liquid absorption was allowed for 30 seconds at 4 °C and 100% humidity. The sample was then blotted for 4 seconds with a blot force of 4 and plunge-frozen into liquid ethane. The procedure was conducted on a Vitrobot Mark IV (Thermo Fisher). The grid was stored in liquid nitrogen until data collection.

Cryo-EM Data Acquisition

Cryo-EM images were acquired on a Titan Krios (ThermoFisher) operating at 300 keV and fitted with a Gatan K3 direct electron detector. 12,265 movies were collected using Leginon^{88,89} at a magnification of 81,000, corresponding to a pixel size of 0.541 Å/pixel, and at a defocus of -1.57 µm. The total exposure dose was 51.37 e⁻/Å², and the exposure time per movie was 2 seconds. Additionally, 20° and 40° tilted images were collected to mitigate the expected preferred particle orientation that would otherwise compromise the EM reconstruction.

Cryo-EM Image Processing

The 12,265 images measured from the VcDnaB-DciA-ssDNA-ATPγS complex were processed using two approaches (Supplementary Figure 3): a) template-based picking using a prior DnaB structure (PDB entry = 7T20⁶²) as the initial template, and b) blob-picking. For facile computing, the blob-picking approach initially targeted 10% of the images, which produced 2.5 million particles that were 2D classified and submitted to *ab initio* reconstruction with C1 symmetry enforced. This reconstruction generated two volumes, C0 and C1 (~123K and ~151K particles), which were judged to be of low quality and influenced our decision not to pursue the blob-picking approach further.

For the PDB structure-based template-picking approach, for computational ease, we selected 100 frames at random, binned them by a factor of $\frac{1}{2}$, and submitted them for motion correction and Contrast Transfer Function (CTF) estimation as implemented in cryoSPARC⁹⁰⁻⁹⁶. Unless otherwise indicated, the following steps were performed within cryoSPARC. An EM map was calculated from PDB entry 7T20 in EMAN2⁹⁷ and used to derive templates for application to the above 100 images; this procedure netted 92,178 particles. From this collection, 79,820 particles were extracted using a box size of 360 pixels and 2D classification. Subsequently, *ab initio* reconstruction was performed with C1 symmetry, which produced two classes (A0, A1). Class A0 (18.5K particles) was carried forward to homogeneous refinement, which created an 8.72 Å map (B0). Inspection of Class A1 suggested a lower quality than A0. As such, further processing of Class A1 was not pursued.

The complete data set (12,265 images) was then submitted to motion correction (Fourier crop factor of $\frac{1}{2}$) and CTF estimation using standard settings. These calculations enabled manual curation leveraging ice thickness (0.82-1.1) and CTF fit resolution (2.17-9.98) parameters to eliminate images with low overall resolution. The resulting 11,817 images were submitted to template-based picking using the B0 map; this calculation produced ~4.3 million particles. These particles were extracted as above and submitted to 2D classification and *ab initio* reconstruction using C1 symmetry. The resulting volumes, D0 and D1 (~437K and ~291K particles), were submitted to heterogeneous refinement and produced two volumes, E0 and E1 (~526K and ~202K particles). Inspection of these two volumes implied that the 6.24 Å E0 volume was of higher quality.

Particles from the E0 Class were re-extracted using a box size of 360 pixels, with no Fourier crop, and submitted to non-uniform refinement, which yielded an EM map with a resolution of 3.36 Å (Volume F). To further improve the quality of

our reconstruction, volume F particles were submitted to global CTF refinement. The final non-uniform refinement job produced map G at 3.37 Å, which was further sharpened with DeepEMhancer⁹⁸ to produce volume H; this volume embodied our best map and was used for model building (below). EM density was evident for the protein, nucleic acid, and nucleotides. Indeed, estimates of the local resolution ranged from 2.8 Å to 4.8 Å, with the highest resolution density spanning the DnaB-CTDs (Supplementary Figure 4C). The quality of the protein density was sufficient to allow unambiguous assignment of side chains to the map. However, no density could be assigned to the DciA loader.

Next, we made several efforts to improve the resolution of Map H in hopes of locating the DciA loader. 3D classification of the particles that generated Map F produced ten classes (group I). Classes I1 and I9 retained the highest number of particles, 82K and 77K, respectively, and showed high similarity to Map H, but with no features that could be assigned to VcDciA. Non-uniform refinement was performed on volumes I4 (46K particles) and I6 (59K particles), producing maps at the resolution of 3.76 Å (class J4) and 3.78 Å (class J6). Automatic model building with ModelAngelo⁹⁹ on the sharpened map, once again, failed to show any evidence of VcDciA. Homogeneous refinement of the particles that produced the E0 volume created class L, followed by two consecutive rounds of non-uniform refinements (classes M and N) with different settings, using volume E0 as the input map. This procedure generated Map N, followed by Map O after DeepEMhancer sharpening, with both at resolutions of 3.45 Å. Neither map showed evidence of density better than Map H, let alone proof of the presence of VcDciA. Lastly, inspection of the 3D variability analysis of the particles that produced volume F (group K) produced no evidence of conformational changes in the VcDnaB-ssDNA complex.

Model Building and Structure Refinement

The DeepEMhancer-sharpened map H was selected to build a molecular model of the VcDnaB-ssDNA complex. Map H was submitted to ModelAngelo⁹⁹ to construct the initial protein and DNA model. The quality of the resulting model was high, even as it contained numerous chain breaks. Next, we superimposed the AlphaFold predicted structure of VcDnaB, which has been split into its three component domains (N-terminal Domain (NTD, residues 1:170), Linker Helix (LH, residues 171:211) and C-terminal Domain (CTD, residues 212:468)), onto the ModelAngelo created model to create a model of hexameric VcDnaB with excellent stereochemistry. The program COOT¹⁰⁰ was used to connect domains belonging to each monomer and to remove the following DnaB residues (chain A, B: 1-21 and 463:468, C-E-F: 1:21 and 464:468, D: 1:22 and 464:468) because of missing map density. Although VcDciA was present in a four-fold excess in our cryo-EM sample, we found no EM density that the VcDnaB model did not explain. As such, we conclude that VcDciA is not present in our reconstruction.

The EM density in Map H revealed an 11-mer ssDNA within the internal chamber of DnaB. This was sufficiently resolved to allow the DNA backbone to be distinguished from the nucleobases; however, it was not so resolved to permit assigning the nucleotide sequence to the map. In many places, purines could be distinguished from pyrimidines; however, we could not unambiguously position the 60-mer ssDNA in our cryo-EM sample. An 11-mer ssDNA molecule (TCCAGATACAC) was constructed in COOT^{100,101} and placed in density using its intrinsic real space refine function. In the lower resolution map E0 (6.24 Å), we observed density extending away from DnaB along the axis in which the 11-mer had been built (Supplementary Figure 5). This density quality did not permit the backbone to be resolved from nucleobase. However, model building suggested that an additional 11 nucleobases could explain this density. Since this density

provided no unique nucleobase information, we built this segment as polyadenine. As such, our final model contains a 22-mer of ssDNA. The extra 11 DNA bases visible in Map E0 were refined in COOT against this map. However, the complete 22-mer ssDNA was refined against two distinct maps. ssDNA residues 1-11 were refined against the high-resolution Map H while residues 12-22 were refined against Map E0. The separately refined ssDNA segments were combined for visualization.

Inspection of the six composite DnaB nucleotide binding sites around the Walker A/P Loop motif (DnaB residues G233, K234, T235) revealed occupancy by six ATP γ S molecules, each bound to a magnesium ion.

The VcDnaB-ssDNA- ATP γ S model was refined with Phenix^{102–109}. We used the 'real_space_refine' protocol, encompassing minimization global, morphing, NQH flips, and atomic displacement parameters (ADP). The refinement leveraged three rigid bodies (DnaB residues 1:171, 171:211, 212:468, 501 for each chain; ssDNA 1:11), Ramachandran, and secondary structure restraints. Refinement was carried out against the DeepEMhancer sharpened 3.37 Å density map H. A second round of real-space refinement encompassing minimization global, atomic displacement parameters (adp), NQH flips, with Ramachandran and secondary structure restraints completed the model (Supplementary Table 1). The final model includes Chain A (DnaB residues 22:461), chain B (DnaB residues 22:462), chain D (DnaB residues 23:463), chains C-E-F (DnaB residues 22:463), six ATP γ S•Mg²⁺ complexes, one per chain (chains A-B-C-D-E-F residue 501) and ssDNA (chain G, residues 1-22).

Model Analysis and Visualization

Structural analysis and visualization of the VcDnaB-ssDNA complex were performed in the CCP4 software package¹¹⁰, Coot^{100,101}, the Uppsala software suite

^{111–114}, UCSF-CHIMERA¹¹⁵, Phenix^{102–109}, PyMOL¹¹⁶, Nucplot¹¹⁷, and as described previously⁴⁴. Distances and angles within and between protein structures were calculated using the PyMol Python scripts titled “pairwise_dist” (https://pymolwiki.org/index.php/Pairwise_distances) and “angle_between_domains” (https://pymolwiki.org/index.php/Angle_between_domains). Molecular graphics and figures were generated using UCSF-CHIMERA¹¹⁵ and PYMOL¹¹⁶. The software used in this study was sourced from the SBGrid Consortium¹¹⁸.

DATA AVAILABILITY

Coordinates for the EM-derived model of the VcDnaB-ssDNA-ATPγS complex are available from the RCSB Protein Data Bank under accession codes 9DLS. Cryo-EM maps are available from the EMBD under the accession code EMD-46984.

FUNDINGS

This work was supported by the National Science Foundation (DJ: MCB 1818255), and the Department of Education (NG: PA200A150068). The nMS work is supported by funding from the National Institutes of Health P41 GM109824 and P41 GM103314 to BTC. Some of this work was performed at the Simons Electron Microscopy Center and National Resource for Automated Molecular Microscopy located at the New York Structural Biology Center, supported by grants from the Simons Foundation (SF349247), NYSTAR, and the NIH National Institute of General Medical Sciences (GM103310) with additional support from Agouron Institute (F00316), NIH (OD019994), and NIH (RR029300).

ACKNOWLEDGEMENT

We thank the Jeruzalmi and Miggiano lab members, the City College biophysics group, the CUNY Advanced Science Research Center, and the New York Structural Biology Center for their scientific and technical advice. We also thank Professor Anuradha Janakiraman for excellent scientific discussions and technical assistance.

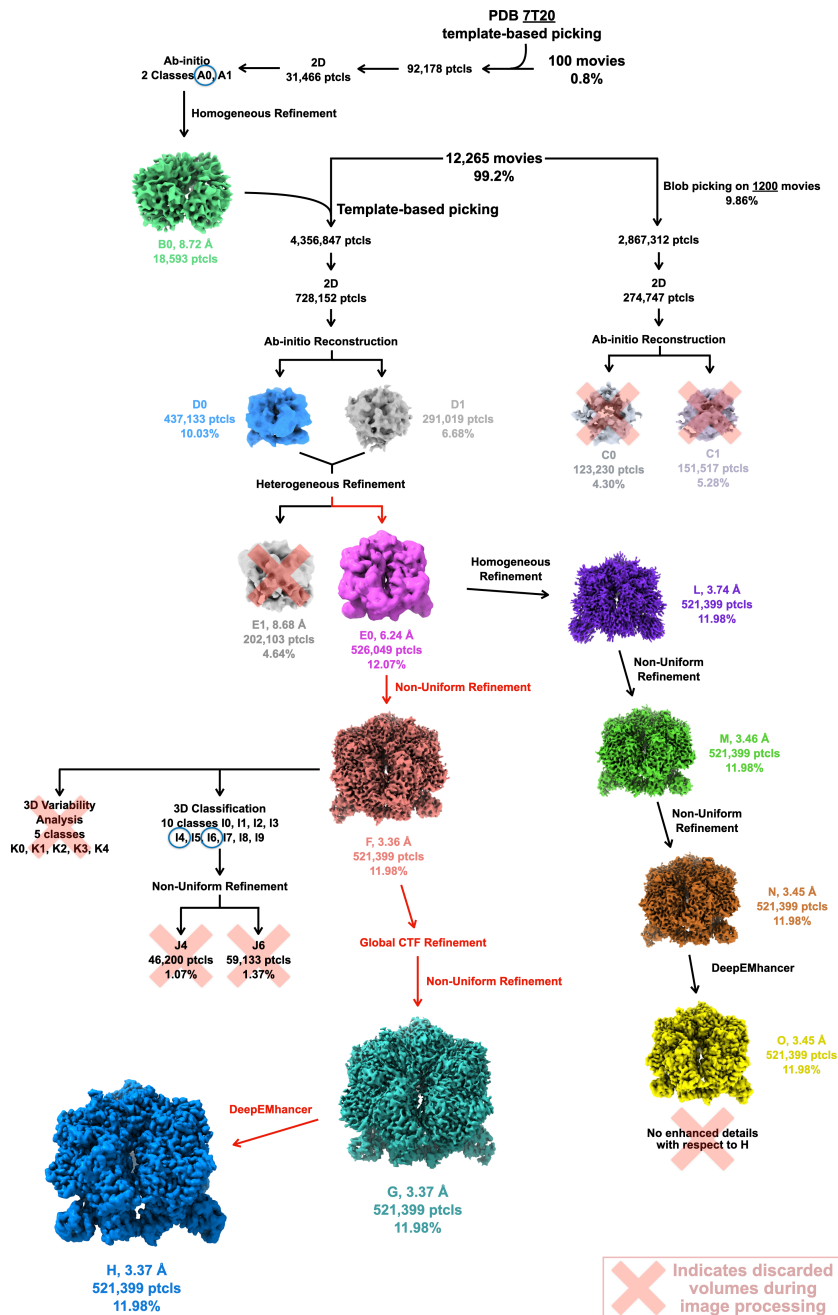
AUTHOR CONTRIBUTIONS

NGao, DM, RM, and DJ conceptualized the study. NGao designed and prepared all the proteins. NGao and DM performed the ATPase assays. NGao performed the pull-down assays. NGouda, ST, AM, AChowdhury, ACerullo, and HB assisted with various laboratory experiments. DM performed the fluorescence-based DNA unwinding and helicase loading assays. DM and CM performed the surface plasmon resonance analyses. All nMS analyses were performed by PDBO in the BTC lab. AP and NGao acquired the cryo-EM data and performed the initial data analysis. DM analyzed the cryo-EM data. DM built and analyzed the cryo-EM structure. NGao, DM, RM, and DJ drafted the manuscript, working with FR and MR.

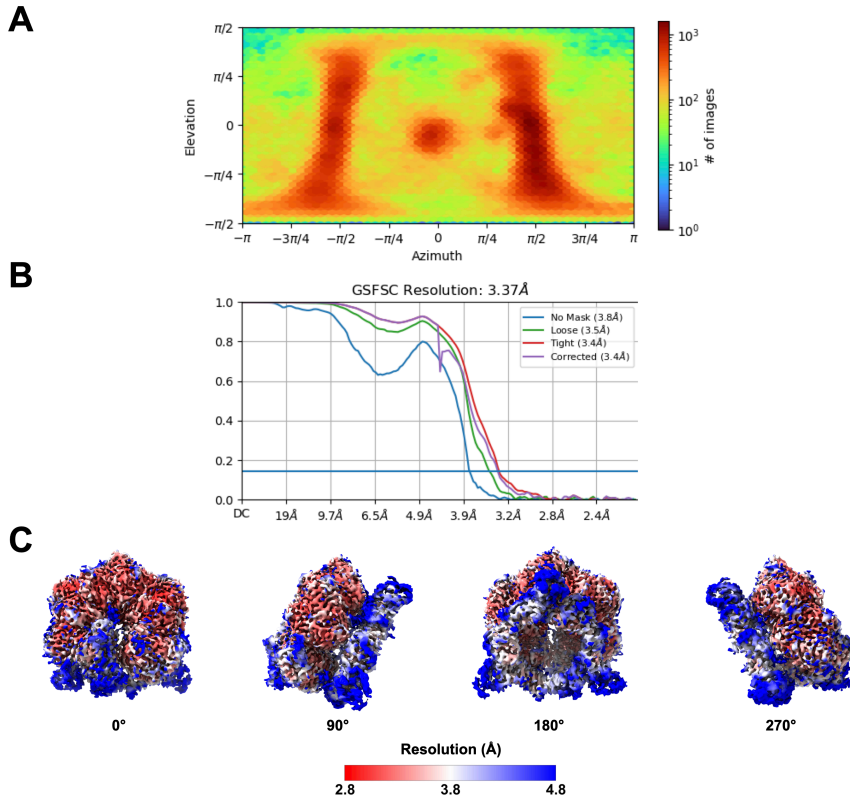
CONFLICTS OF INTEREST

The authors have no financial or non-financial conflicts of interest.

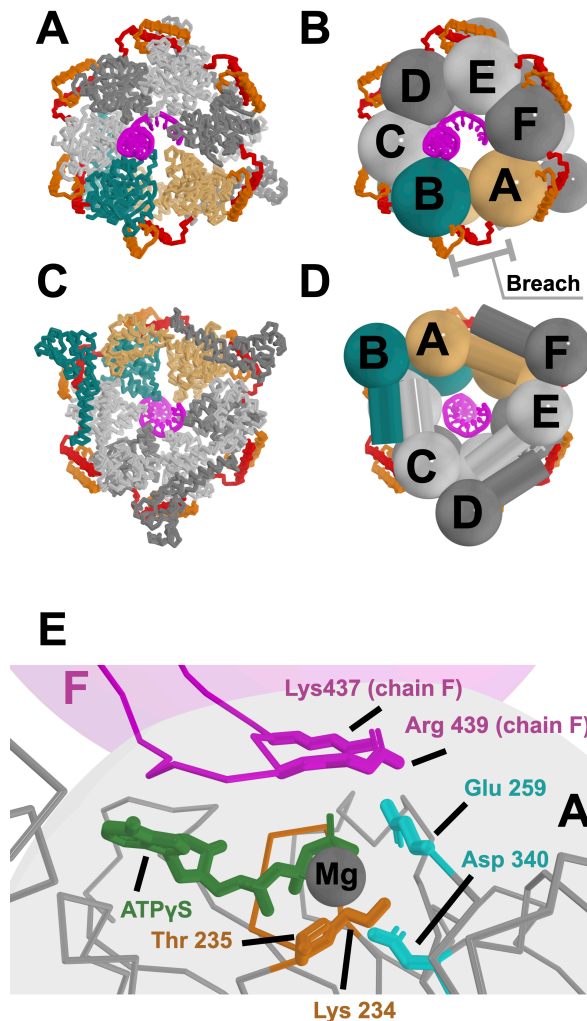
SUPPLEMENTARY FIGURES



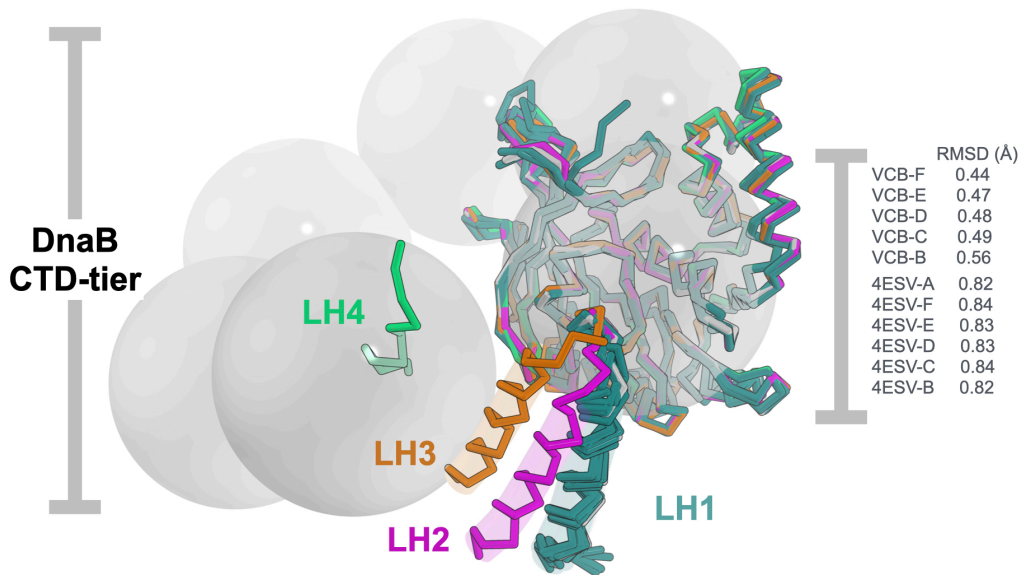
Supplementary Figure 1. Workflow Steps in the Processing of cryo-EM Data for the VcDnaB-ssDNA-ATPyS complex. Workflow steps are described in the Methods section. The percentages reported alongside the volumes are calculated against the number of starting picked particles. The various volumes/maps in the workflow are depicted with unique colors. Map H represents the highest quality map we obtained, against which the atomic model was built and refined. The extended length of ssDNA was built into map E0.



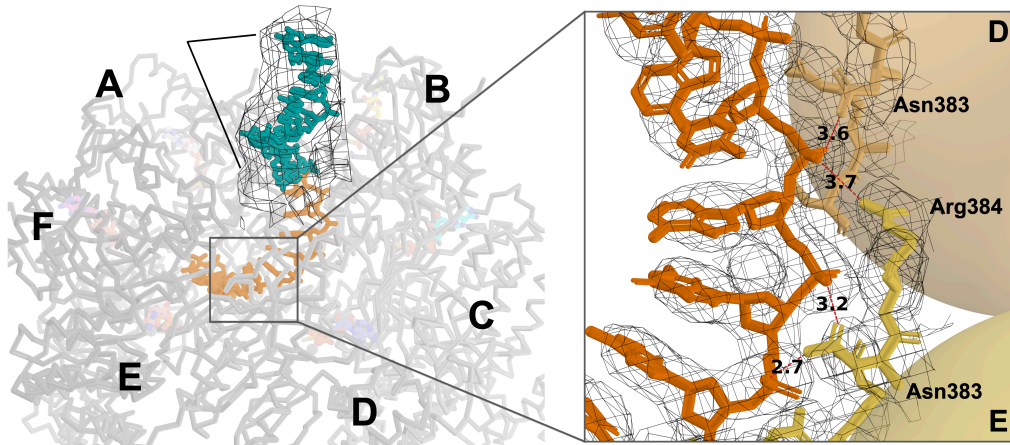
Supplementary Figure 2. Statistics, Fourier Shell Correlation, and Local Resolution. A) CryoSPARC Plot of particles' azimuthal and elevation angles in the 3.37 Å reconstruction of VcDnaB-ssDNA-ATPγS. The blue-to-red scale describes the number of particles, from 10^0 to 10^3 . The red clusters represent the preferred orientation of the particles in our data set; however, this distribution of particles did not impact the quality of the map reconstruction. B) Fourier Shell Correlation (FSC) plots taken from the non-uniform refinement job that produced the map G. Plots were generated by applying no mask, spherical mask, loose mask, tight mask, and corrected mask on the protein-DNA complex. The corrected tight mask plot shows a resolution of 3.37 Å at FSC = 0.143. C) The VcDnaB + ssDNA EM density map is colored (red: 2.8 Å to blue: 4.8 Å) according to the local resolution. The density map is depicted in four poses that differ by rotations of 90° along the Y axis.



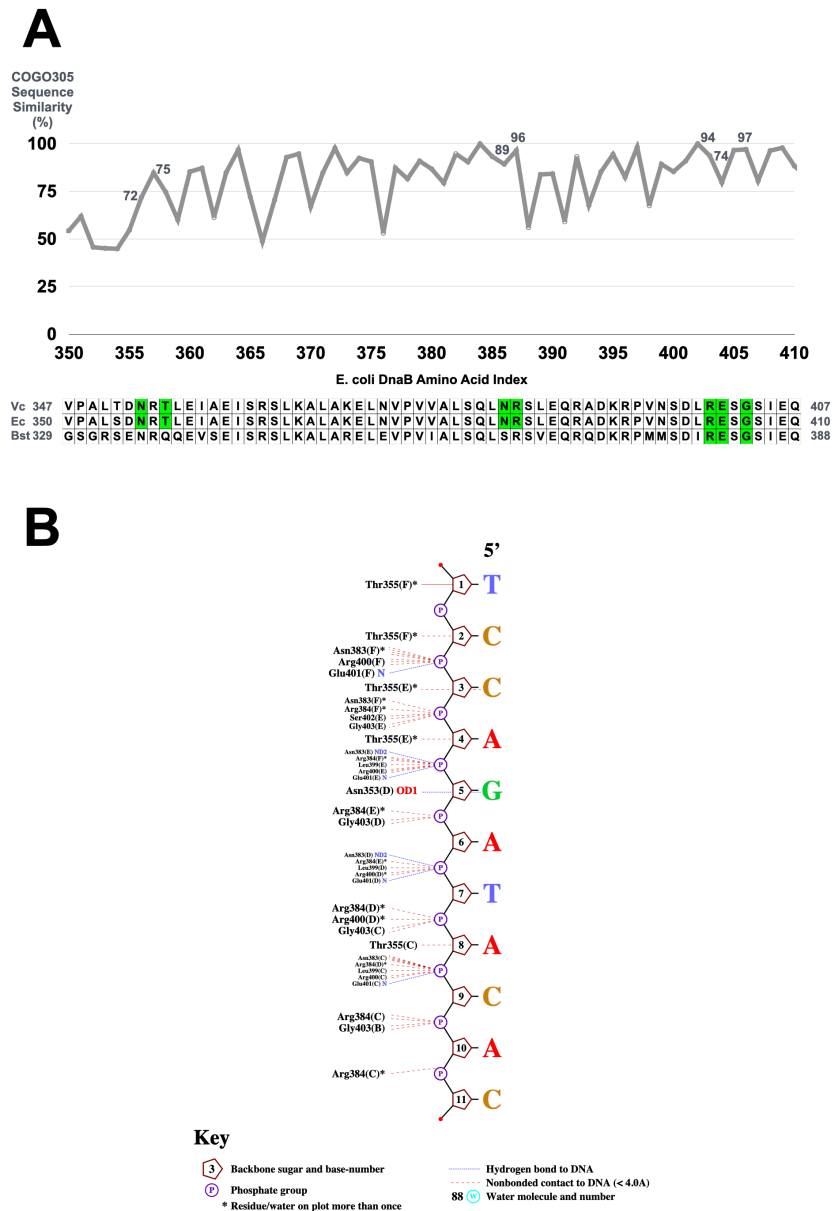
Supplementary Figure 3. The VcDnaB-ssDNA-ATPyS Complex. The VcDnaB-ssDNA-ATPyS is drawn in a ribbon representation (A or B) or a schematic using a design language wherein the NTDs are depicted as a mushroom shape and the CTDs as spheres (B or D). ssDNA, colored in magenta, is displayed in a cartoon and stick representation. The breach in the CTD layer is between chain A (light orange) and chain B (deep teal); other DnaB subunits are colored in alternating shades of gray. The Docking Helix (DH) and the Linker Helix (LH) are depicted as transparent cylinders colored orange and red, respectively. The pose in panels A and B is along the path of DNA. Panels C and D are in a pose that is rotated approximately 180° from that in panels A and B. E) Closeup of the composite nucleotide catalytic center in VcDnaB at the interface between subunits A (gray shape) and F (magenta shape). ATPyS (green) is bound at a composite site created by the Walker/P-Loop residues Thr 234 and Lys 235 (orange), Walker B residues Asp 340 and Glu 259 (cyan) from chain A, and the β -hairpin arginine finger residues Lys 437 and Arg 439 from chain F. Chains A (grey) and F (magenta) are depicted in the ribbon format.



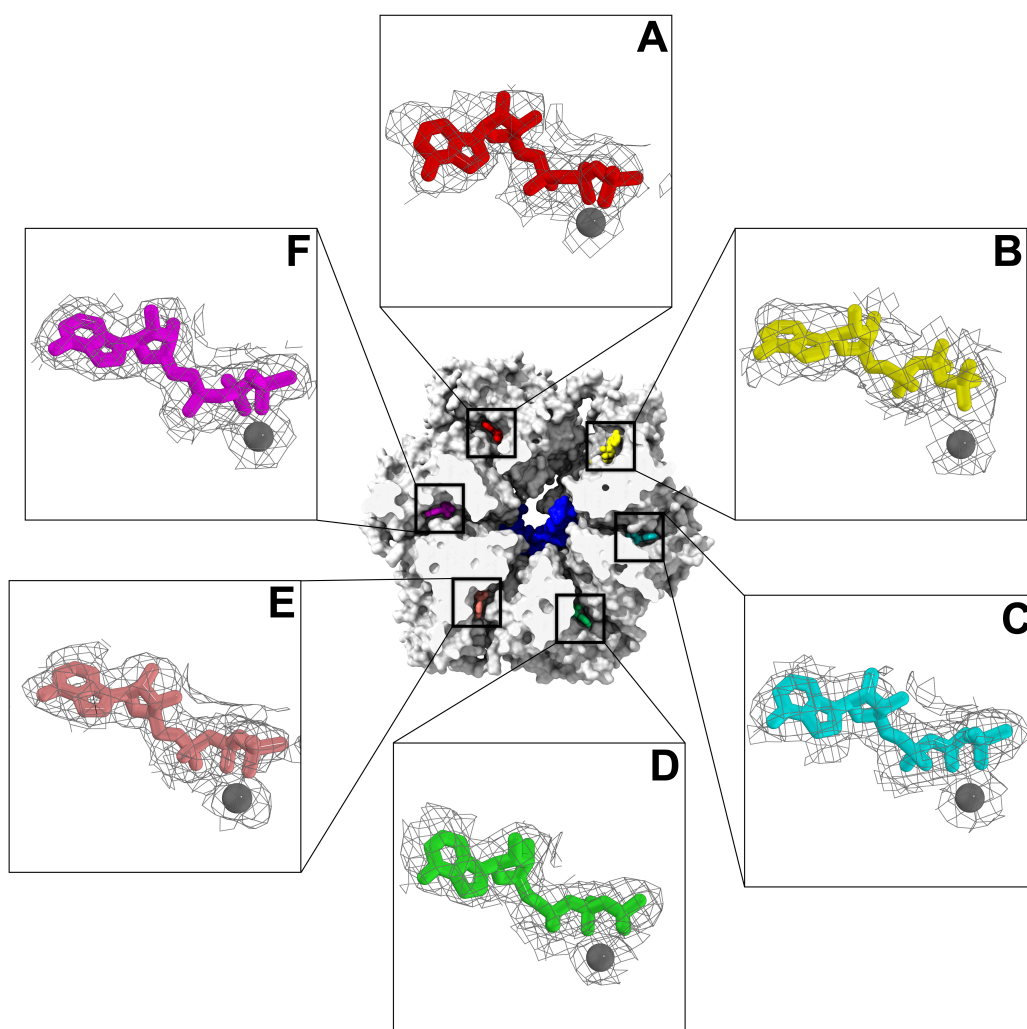
Supplementary Figure 4. **The DnaB Linker Helix (LH) Element is a Moving Part During Translocation on ssDNA.** Superposition analysis of the VC ATP γ S and GDP•AlF₄ structures reveals that the LH element of DnaB occupies at least four positions, termed LH1, LH2, LH3, and LH4). The CTDs (total of 11) of the Vc and Bst structures were superimposed on residues 204:465 of Vc chain A (residues corresponding to the LH 178:205 or DH 287:304 elements were omitted from the calculation). The four positions of the LH are colored in deep teal (LH1), magenta (LH2), orange (LH3), and lime-green (LH4). Except for chains A (LH3) and F (LH2), the Vc DnaB chains populate the LH1 position. Chain A of the Bst-4ESV structure populates the LH4 position; the others are seen in the LH1 position. The CTD layer of the Vc structure is shown as labeled spheres drawn at the centers of mass of each CTD. The superimposed CTDs are drawn in the ribbon representation, and the transparent cylinder indicates the LH.



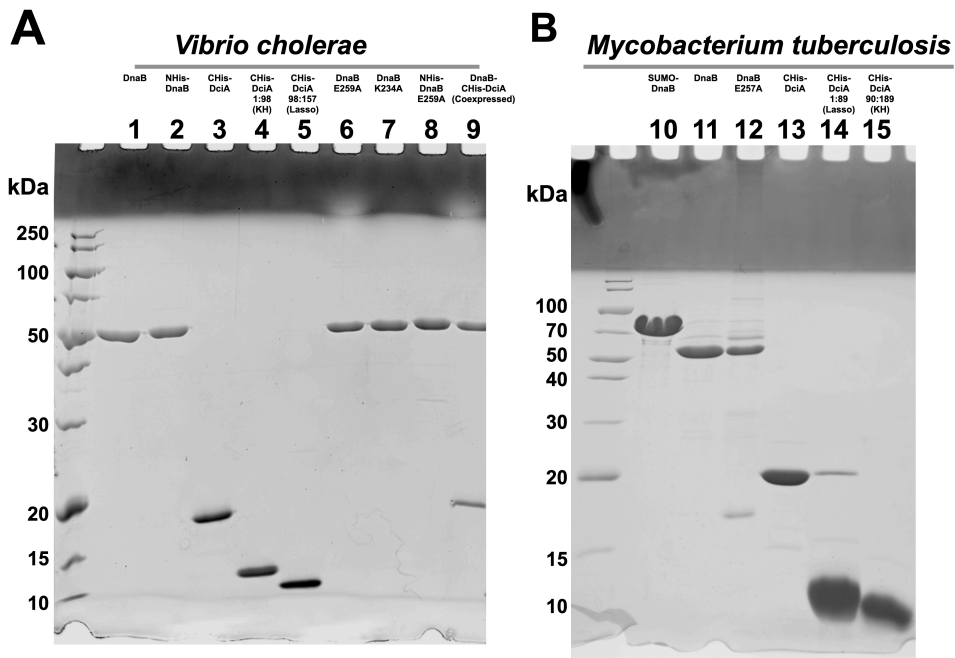
Supplementary Figure 5. 11 additional ssDNA positions outside the confines of VcDnaB are seen in the cryo-EM density map. Close inspection of the 6.24 Å map E0 (Supplementary Methods and Supplementary Figure 3) enabled an ssDNA model to be constructed that was 11 positions longer than that seen in map H. The map into which this segment was built is shown as a mesh (isomesh level: 2.5). The extra positions are colored in teal; the ssDNA built into map H is colored orange. VcDnaB is represented as a grey ribbon, and each monomer is labeled by chain (A-B-C-D-E-F). For reference, the six ATP γ S and six Mg $^{2+}$ atoms are displayed as sticks and spheres, respectively, applying transparency. As noted in the methods, neither map H nor map E0 was sufficiently resolved to permit the assignment of parts of the input 60-mer sequence to the map. As such, the sequence in our final model was built as follows: TCCAGATACACAAAAA. Inset: close-up view of the DnaB-ssDNA interface superimposed on the cryo-EM density. The inset was rotated $\sim 90^\circ$ counterclockwise with respect to the main panel and adjusted to align the nucleobases along the vertical axis.



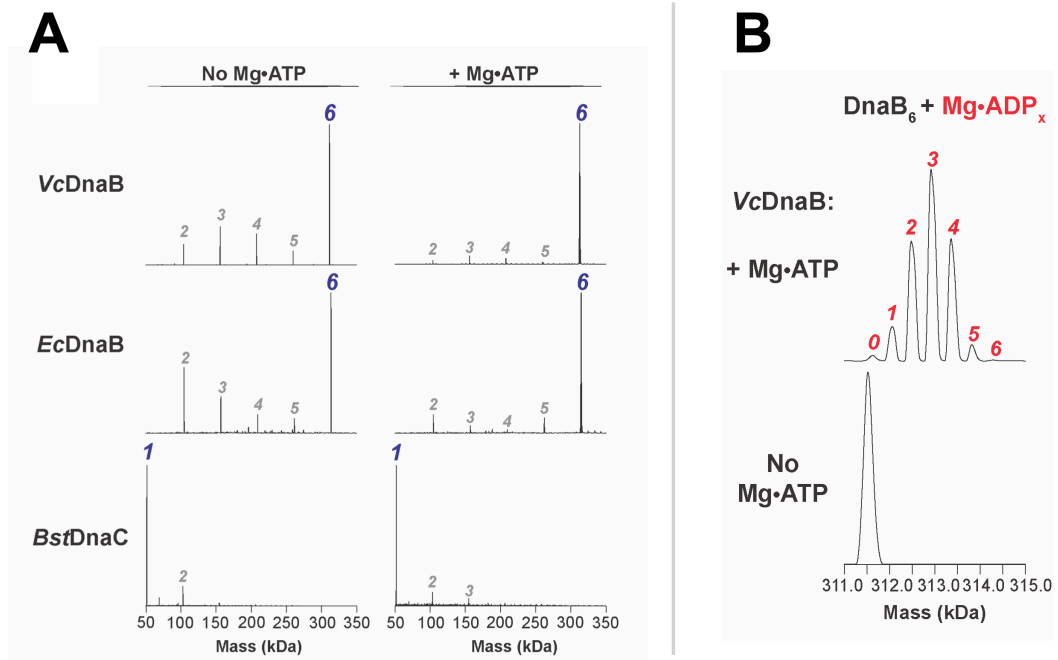
Supplementary Figure 6. **Protein – ssDNA interactions in the VcDnaB-ssDNA complex.** A) Sequence identity is plotted against the amino acid index of a portion of the sequence alignment derived from the DnaB COG0305 conserved protein domain family (100 sequences). Seven VcDnaB residues are in contact with ssDNA, though not every subunit uses each residue. These are N353, T355, N383, R384, R400, E401, and G403 (highlighted in green). The sequence alignment teaches that these residues are conserved in the COG0305 DnaB-family. The conservation value for each contact residue is listed in the plot. The residues used by *E. coli* (PDB: 7T20⁶²) and Bst (PDB: 4ESV¹⁰) are also highlighted. B) VcDnaB-ssDNA contacts are depicted schematically as drawn by Nucplot ¹¹⁷.



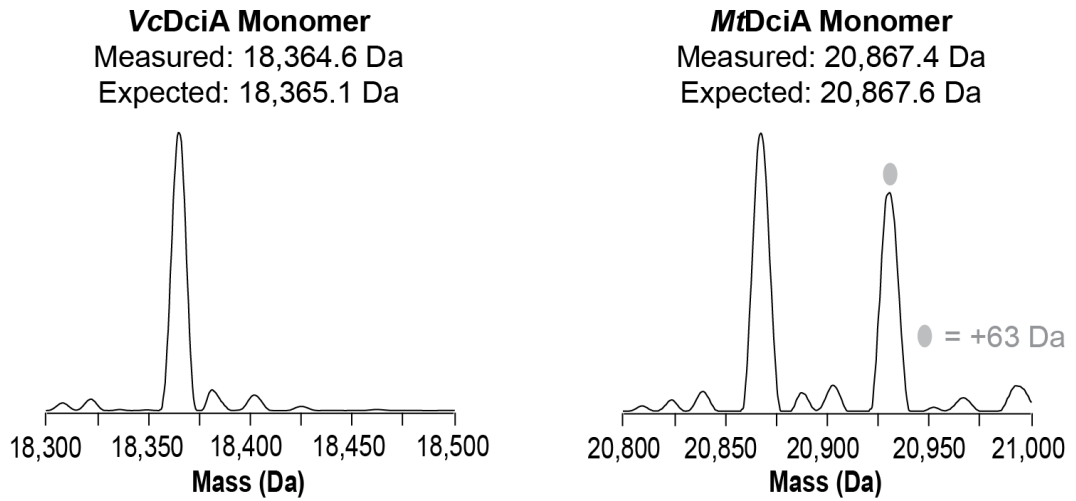
Supplementary Figure 7. The six nucleotide binding pockets in the VcDnaB-ssDNA structure are filled with ATP γ S coordinated to a Mg $^{2+}$ atom. The β and γ phosphates of each ATP γ S moiety are coordinated to an Mg $^{2+}$ atom. EM density is drawn as a mesh around each nucleotide, which supports our interpretation. The density is drawn in PyMol using the following parameters: isomesh level chain B: 0.5; chains A-C-D-E-F: 2 and isomesh carve: chains A-B-C-D-E-F: 2.8. The ATP γ S molecules, each colored uniquely are drawn in the stick representation, with the Mg $^{2+}$ ion depicted as a grey sphere. The position of the nucleotide pockets is indicated on the surface representation of VcDnaB sliced to expose the six pockets. ssDNA within DnaB is colored dark blue.



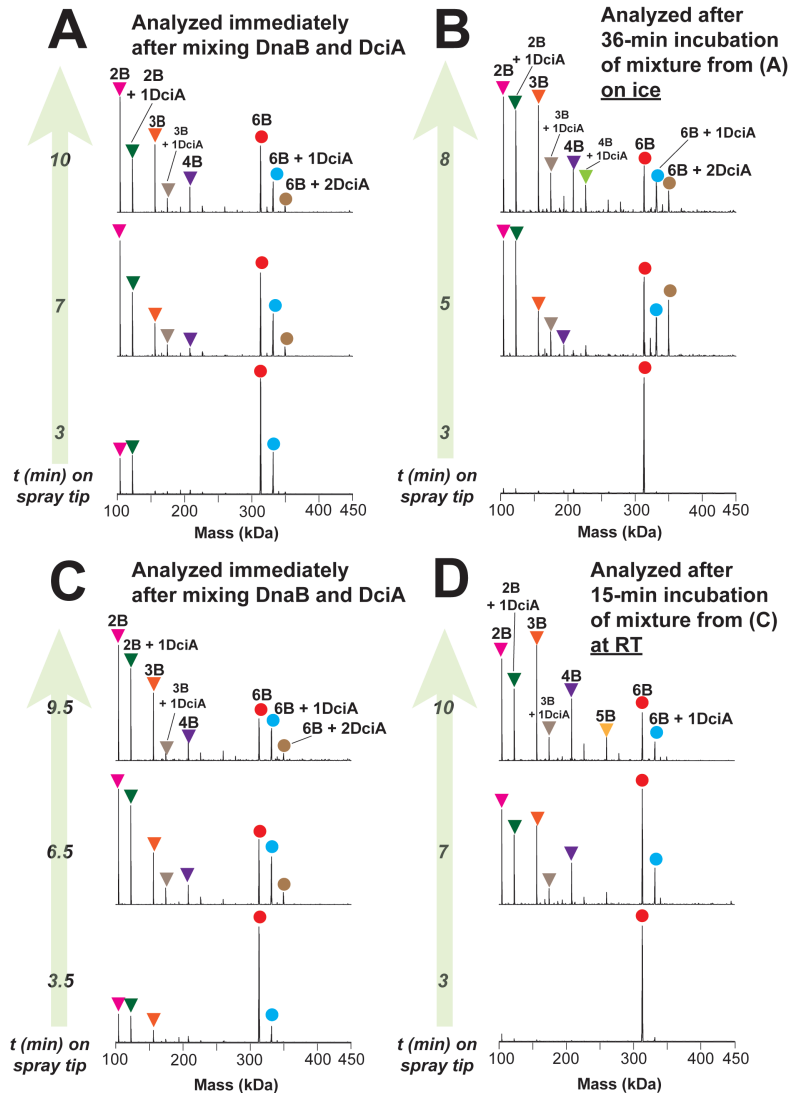
Supplementary Figure 8. Purified DnaB and DciA proteins from *Vibrio cholerae* and *Mycobacterium tuberculosis* used in this study. A) Gel electrophoretic analysis of purified Vc proteins. Lane 1: VcDnaB untagged; Lane 2: N-terminal His-tagged VcDnaB; Lane 3: C-terminal His-tagged VcDciA full-length; Lane 4: C-terminal His-tagged VcDciA KH domain (residues 1:98); Lane 5: C-terminal His-tagged VcDciA Lasso domain (residues 98:157); Lane 6: Untagged VcDnaB-E259A; Lane 7: Untagged VcDnaB-K234A; Lane 8: N-terminal His-tagged VcDnaB-E259A; Lane 9: Coexpressed VcDnaB untagged + C-terminal His-tagged Full-length VcDciA. B) Gel electrophoretic analysis of purified Mt proteins. Lane 10: SUMO-tagged-MtDnaB; Lane 11: Untagged MtDnaB; Lane 12: Untagged-MtDnaB-E257A; Lane 13: Full-length C-terminal His-tagged MtDciA; Lane 14: C-terminal His-tagged MtDciA Lasso Domain (1:89); Lane 15: C-terminal His-tagged MtDciA KH domain (90:189). The analyses in panels A and B are from 15% SDS-PAGE.



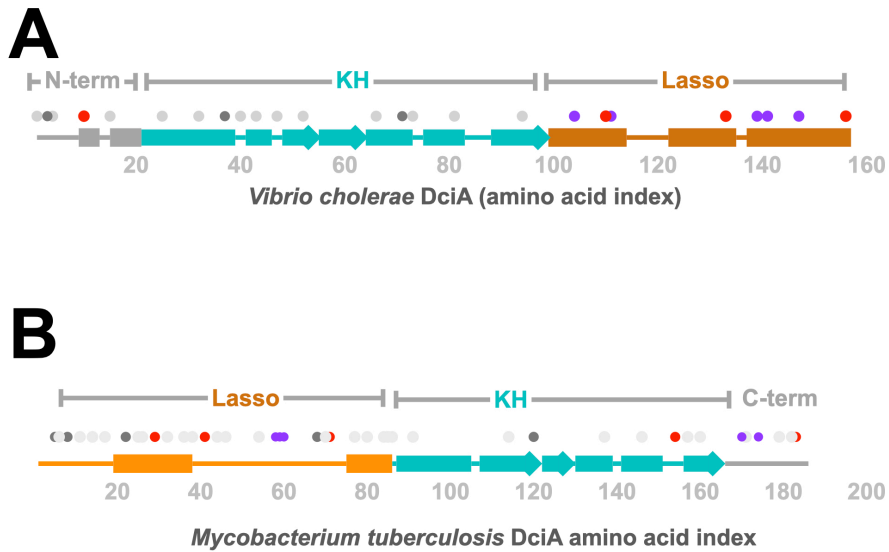
Supplementary Figure 9. Although highly conserved, particular DnaB helicases orthologs feature distinct oligomeric states. (A) Deconvolved nMS spectra of replicative helicases from *V. cholerae* (Vc), *E. coli* (Ec), and *B. subtilis* (Bst) with and without Mg•ATP. Unlike the *B. subtilis* ortholog, which is monomeric (with or without ATP), *V. cholerae* and *E. coli* DnaB are hexamers. Each measurement encompassed a monomeric protein concentration of 10 μ M. The BstDnaC sample refers to the DnaB family helicase in *B. subtilis*. (B) Close-up of panel A peaks corresponding to the hexameric VcDnaB complex. The average mass difference between adjacent peaks is 444 ± 20 Da; this value more closely matches the mass of ADP (Mass = 427 Da) than ATP (Mass = 507 Da). As such, we suggest that ATP hydrolysis has occurred, giving rise to distinct Mg•ADP occupied states. The numbers in panels A and B refer to the oligomeric state of VcDnaB (A, blue) and of Mg-ADP molecules bound to VcDnaB (B, red). A list of mass assignments of the various nucleotide-bound peaks in panels A and B appears in Table 5.



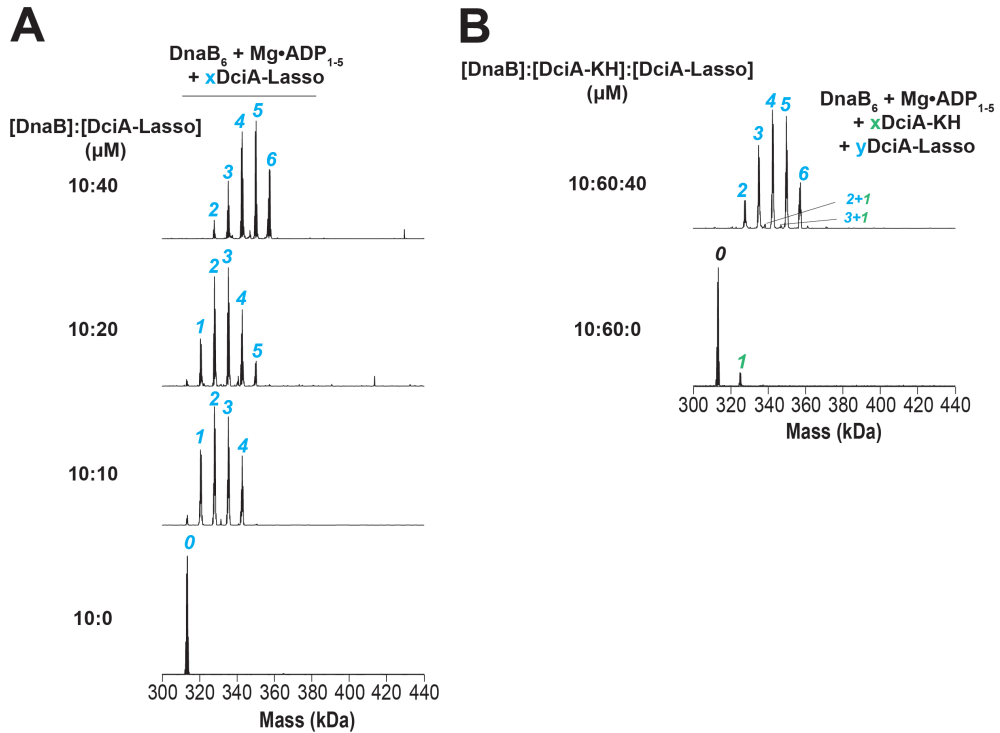
Supplementary Figure 10. Isolated Vc and MtDciA are monomeric. Deconvolved nMS spectra of Vc (A) and MtDciA (B). The additional +63 Da peak (grey oval) in the MtDciA nM spectrum most likely corresponds to a nonspecific metal ion adduct (Zn^{2+} or $\text{Cu}^{1+/2+}$) by the carboxy-terminal hexahistidine tag.



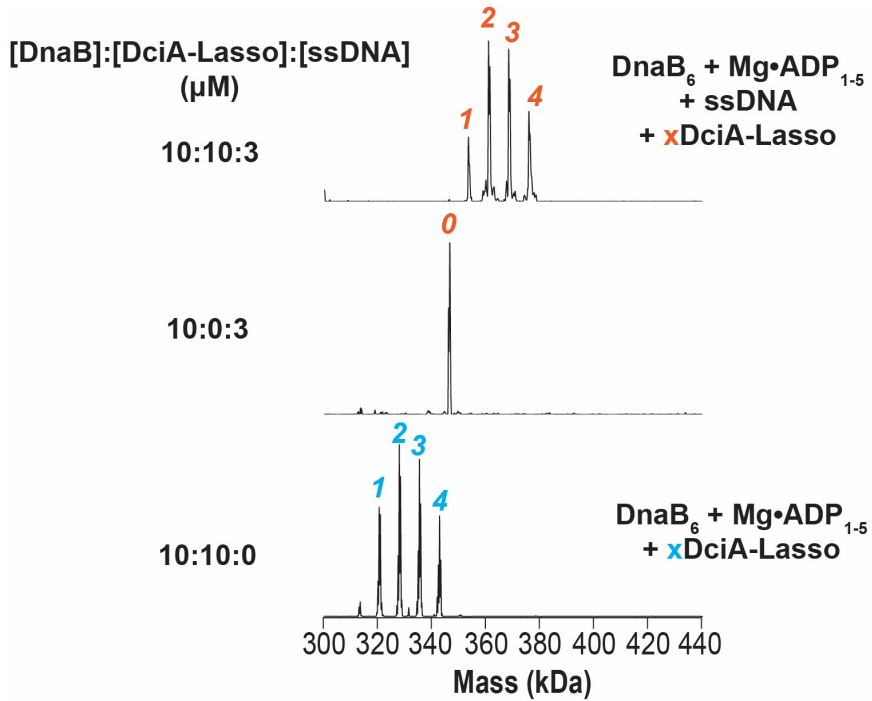
Supplementary Figure 11. Disassembly of VcDnaB upon incubation with VcDciA during nMS analysis. VcDnaB was mixed with VcDciA, and an aliquot of the mixture was loaded immediately on the electrospray tip for nMS analysis (A and C), and the rest of the mix was either incubated on ice for 36 min (B) or at room temperature for 13 min (D) before nMS characterization. The deconvolved spectra at specific time points during the nMS analysis (~100 scans averaged) and the corresponding peak assignments are shown for each sample. The overall delay time from loading the sample into the tip, setting up the electrospray, and initiating nMS data acquisition is typically 2 min. The sample mixture comprised of 10 μ M DnaB and 7 μ M DciA in 10 μ M ATP, 0.5 mM magnesium acetate, 500 mM ammonium acetate, 0.01% Tween-20. The peaks for DnaB₆ and DnaB₆ + DciA₁ predominate at the first 1 - 3 min of initiating the electrospray and then peaks with increasing intensities for DnaB₂, DnaB₂ + DciA₁ and other DnaB subcomplexes appear indicating progressive disassembly of DnaB₆. Analysis of the rest of the sample incubated on ice or at RT resulted in similar time-resolved MS profiles, suggesting that DnaB₆ dissociation only occurs during electrospray. Overall, the results indicate that DciA binding can destabilize DnaB₆, and this destabilization is intensified during the electrospray process, leading to DnaB₆ disassembly.



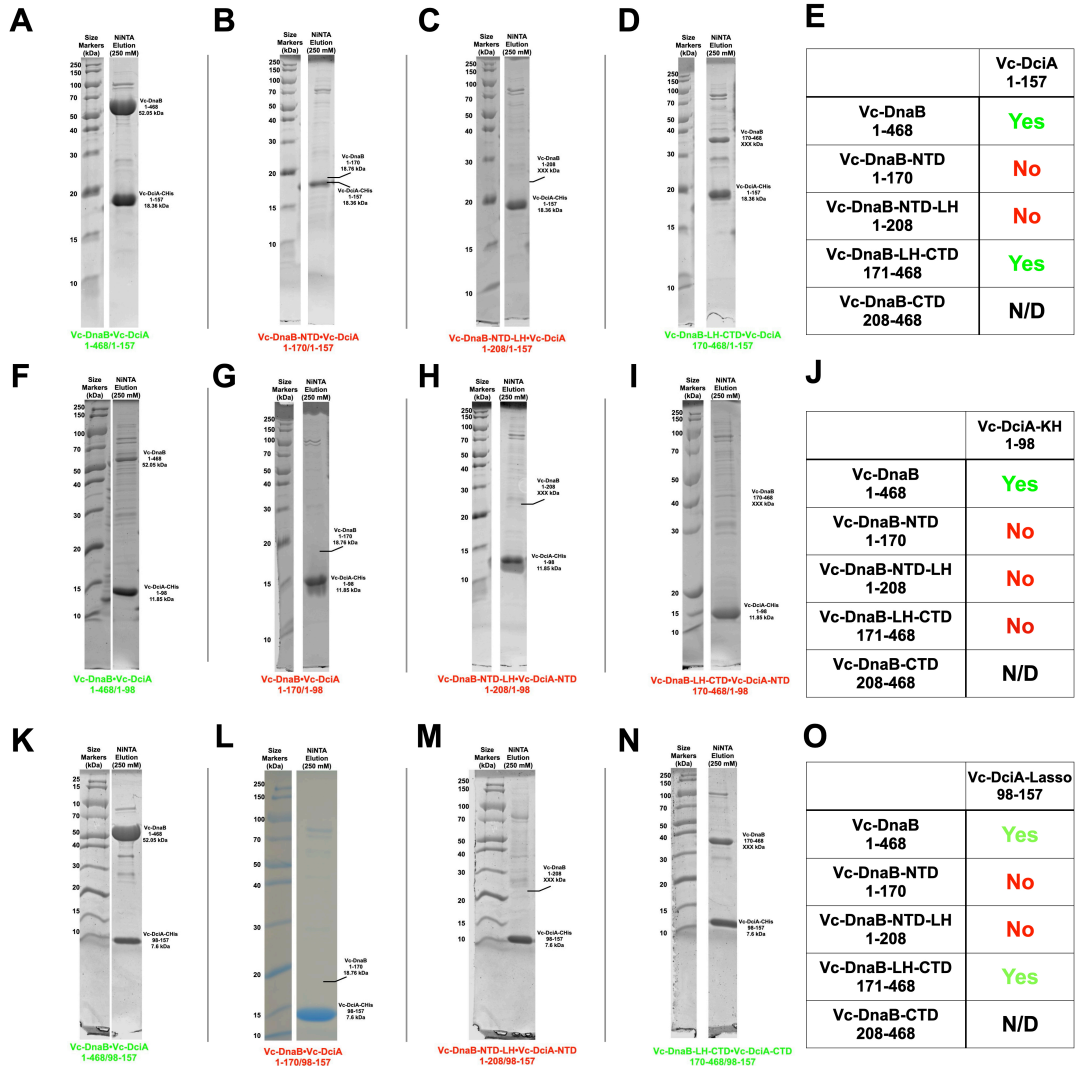
Supplementary Figure 12. The arrangement of KH and Lasso domains in the primary sequence DciA loaders from Vc and Mt is divergent. The domain architecture of DciA from Vc (A) and Mt (B) was inferred from prior analyses^{40,41,58} and limited proteolysis and mass spectrometry (Supplementary Figure 10). The red and purple circles represent proteolysis sites by the Asp-N and trypsin proteases, respectively. The dark (Asp-N) and light (trypsin) grey circles represent protease sites where cuts are not observed. The KH, Lasso, and amino and carboxy-terminal domains are labeled.



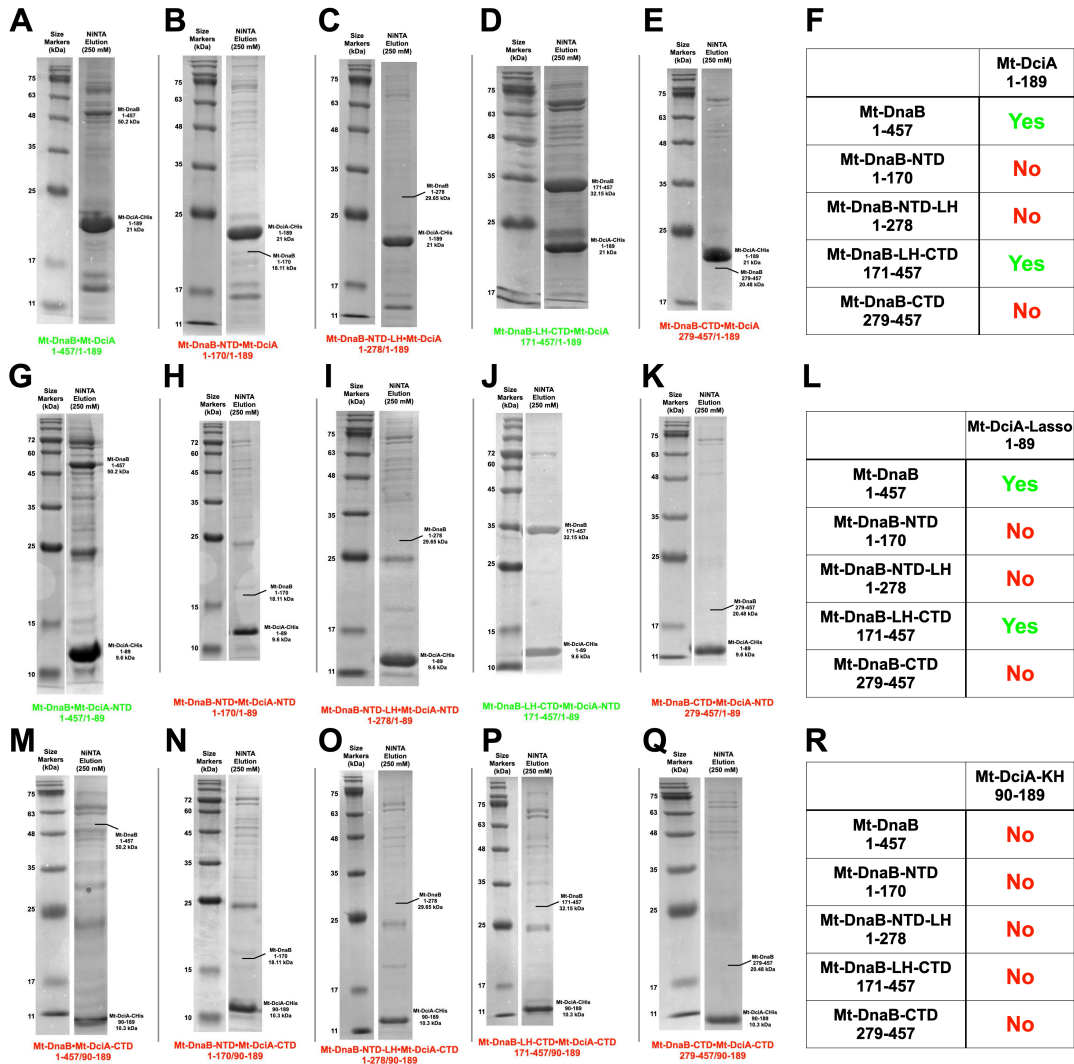
Supplementary Figure 13. Stoichiometry of binding the VcDciA KH domain, the VcDciA Lasso, and ssDNA to VcDnaB by nMS. Deconvolved nMS spectra of (A) DnaB titrated with increasing concentrations of DciA-Lasso, and (B) DnaB incubated with an excess DciA's KH domain in the absence or presence of DciA's Lasso domain. The expected masses of VcDnaB₆ = 311.5 kDa, DciA-KH = 11.8 kDa, DciA Lasso = 7.5 kDa, and ssDNA = 32.5 kDa. The colored numbers and Roman letters 'X' and 'Y' indicate the number of DciA-Lasso domains bound (A) or DciA-KH and Lasso bound (B).



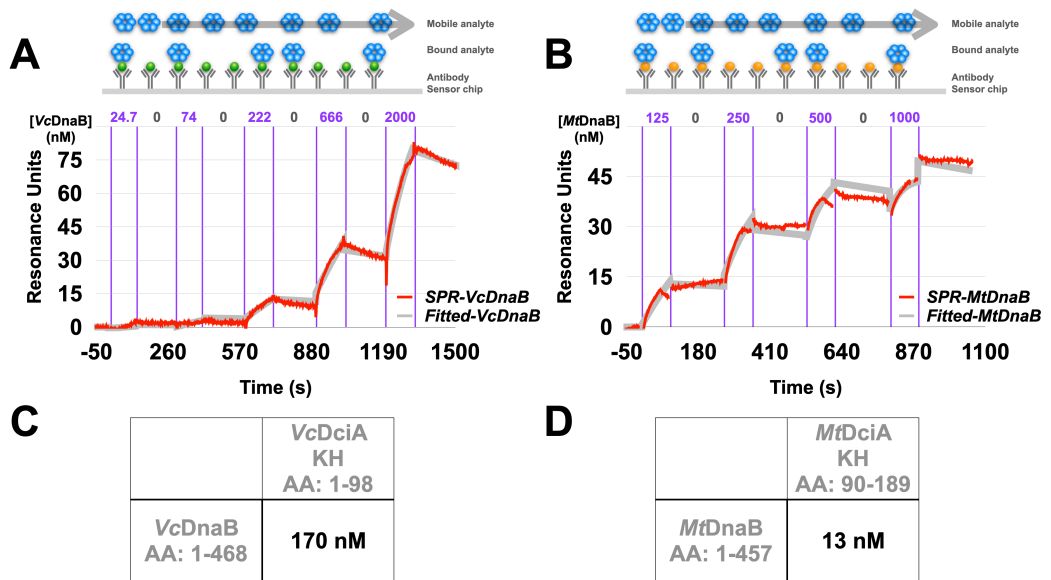
Supplementary Figure 14. The Vc Lasso domain forms complexes with VcDnaB and ssDNA. Up to 4 Vc Lasso domains are found in VcDnaB complexes, whether ssDNA is in the absence/presence of ssDNA. Deconvolved nMS spectra of DnaB and ssDNA (O4-VcF) incubated with and without an equimolar concentration of DciA-Lasso. All samples are in 10 μM ATP, 500 μM magnesium acetate, 500 mM ammonium acetate, and 0.01% Tween-20 before nMS analysis. Spectra were collected at the indicated DnaB: DciA-Lasso: ssDNA concentrations. The blue and red Roman letter 'X' indicates the number of copies of VcDciA revealed by nMS for a given sample. The expected masses are DnaB₆ = 311.5 kDa, DciA-KH = 11.8 kDa, DciA Lasso = 7.5 kDa, and ssDNA = 32.5 kDa.



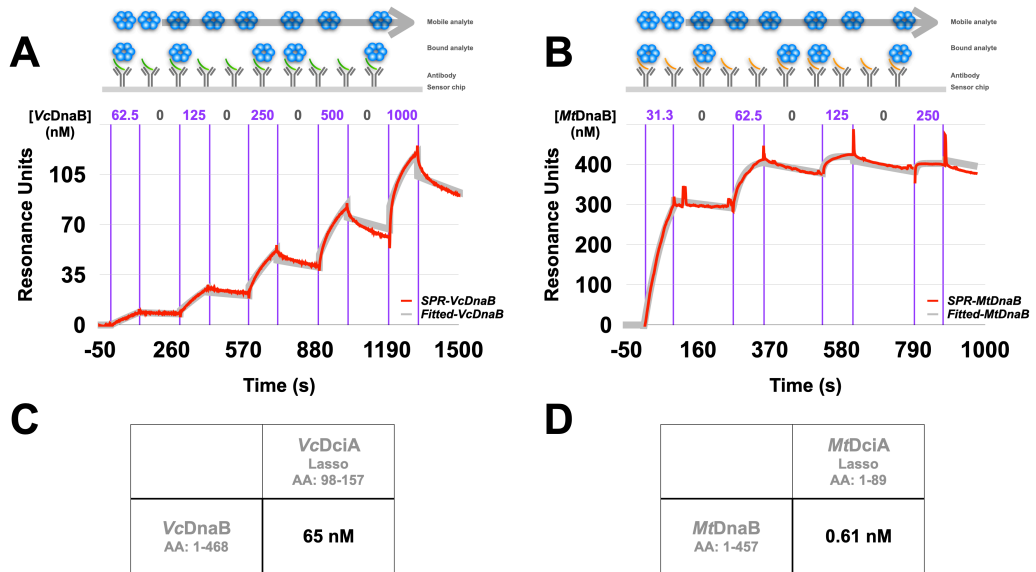
Supplementary Figure 15. VcDciA binds tightly to VcDnaB, and the VcLasso domain prefers the LH-CTD domain. Composite image of 12 NiNTA 'pull down' assay between VcDciA (full-length: 1:157; KH: 1:98; Lasso: 98:157) and DnaB (full-length: 1:468; NTD: 1:170; NTD-LH: 1:208; LH-CTD: 171:468; CTD: 208:468) constructs. A-D) Full-length VcDciA and various VcDnaB constructs. E) Summary of binding data from panels A-D; the VcDnaB CTD (residues 208:468) could not be stably expressed. F-I) VcDciA KH domain (1:98) and various VcDnaB constructs. J) Summary of binding data from panels F-I. K-N) VcDciA Lasso domain (98:157) and various VcDnaB constructs. O) Summary of binding data from panels K-N. The green text indicated a NiNTA pull-down scored as interacting, and the red indicated no interaction was observed. These data show that full-length VcDciA binds to full-length VcDnaB and that the Lasso domain binds to the LH-CTD, likely at the DH-LH interface, as with other helicase loaders^{12,14,19,43,44,48}.



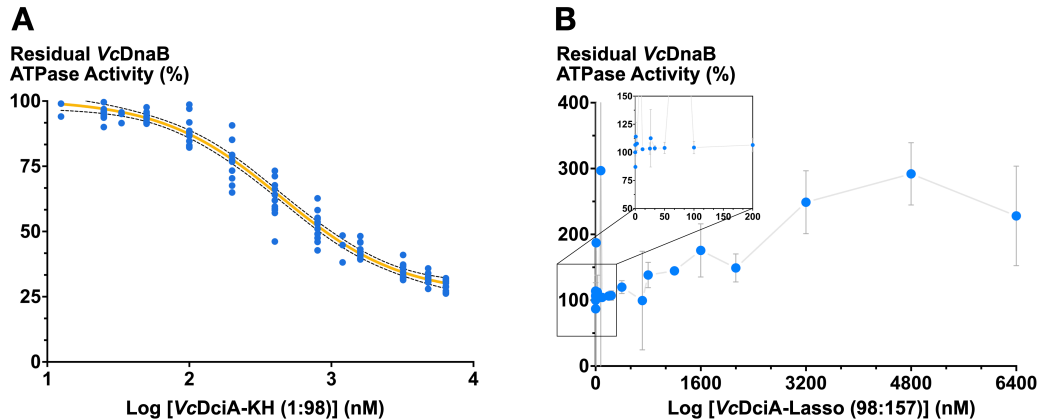
Supplementary Figure 16. The pattern of binding of MtDciA and its sub-domains to MtDnaB is identical even though the order of the domains is inverted in the primary sequence. Composite image of 15 NiNTA 'pull down' assay between MtDciA (full-length: 1:189; Lasso: 1:89; KH: 90:189) and DnaB (full-length: 1:457; NTD: 1:170; NTD-LH: 1:278; LH-CTD: 171:457; CTD: 279:457) constructs. A-E) Full-length MtDciA and various MtDnaB constructs. F) Summary of binding data from panels A-E. G-K) MtDciA KH domain (90:189) and various VcDnaB constructs. L) Summary of binding data from panels G-K. M-Q) MtDciA Lasso domain (1:89) and various MtDnaB constructs. R) Summary of binding data from panels M-Q. The green text indicated a NiNTA pull-down scored as interacting, and the red indicated no interaction was observed. These data show that full-length MtDciA binds to full-length MtDnaB and that the Lasso domain binds to the LH-CTD, likely at the DH-LH interface as with other helicase loaders^{12,14,19,43,44,48}.



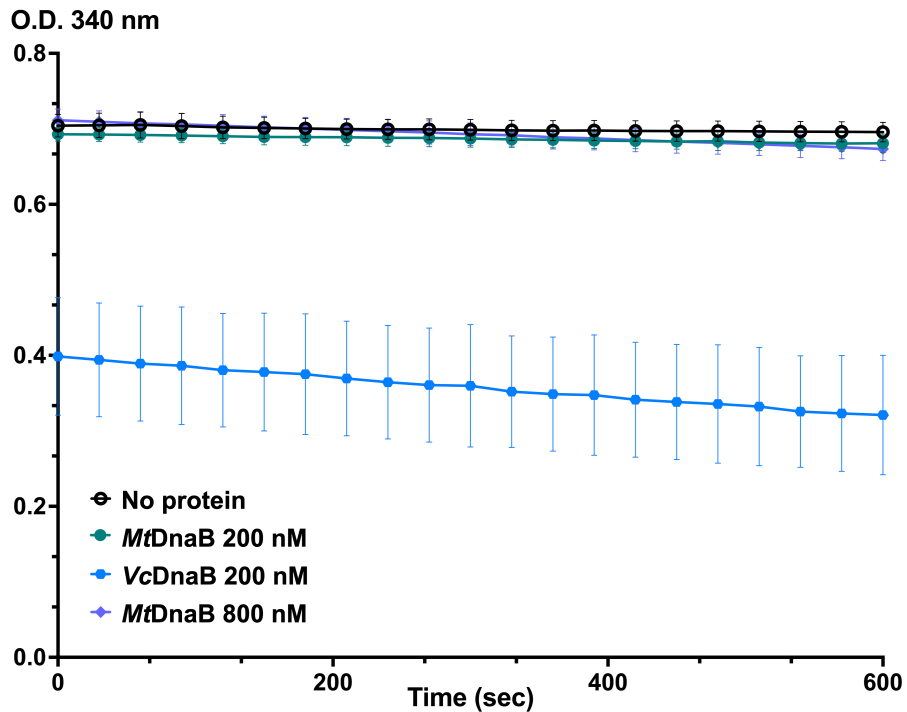
Supplementary Figure 17. Vc and MtDciA's KH domain retains affinity to cognate DnaB, albeit with a higher K_D . The SPR experiments are diagrammed schematically above each trace. On and off rates with full-length DnaB at incremental concentrations (Vc: 24.7, 74, 222, 666, 2000 nM; Mt: 125, 250, 500, 1000) were measured by immobilizing cognate C-terminal His-tagged DciA-KH, from Vc (A) and Mt (B) to the SPR sensor chip. The raw SPR signal is shown in red, and the fitted model is depicted in grey. Experimental data were fit as described in the methods. The single-cycle kinetic analyses through SPR revealed how DciA's KH domain retained a decreased affinity to cognate helicase. The Vc KH domain displays a $K_D = 170$ nM ($\chi^2 = 1.82$; $R_{max} = 135.2$), representing a ~300-fold reduction compared to full-length (A, C). This difference is less pronounced in the Mt system, where the K_D constant for binding to Mt KH is 13 nM ($\chi^2 = 4.18$; $R_{max} = 51.23$); this value is 15-fold lower than for full-length (B, D).



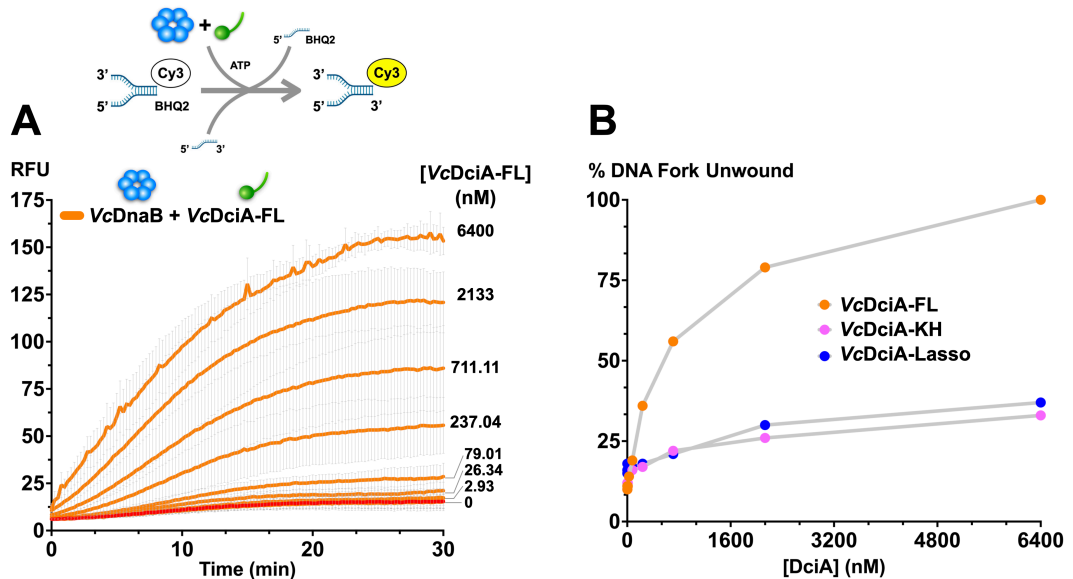
Supplementary Figure 18. The Vc and MtDciA Lasso domains form energetically significant contacts to cognate DnaB. The SPR experiments are diagrammed schematically above each trace. On and off rates with full-length DnaB at incremental concentrations (Vc: 62.5, 125, 250, 500, 1000 nM; Mt: 31.3, 62.6, 125, 250 nM) were measured as described in Methods and Supplementary Figure 17. The raw SPR signal is shown in red, and the fitted model is depicted in grey. Experimental data were fit as described in the methods. The Vc Lasso domain features a K_D of 65 nM ($\chi^2 = 5.89$; $R_{\max} = 126.3$), contrasting with the 0.58 nM value measured for full-length protein and the 170 nM value for the KH domain (A, C). In the Mt system, we measure a $K_D = 0.61$ nM ($\chi^2 = 123$; $R_{\max} = 411.5$), close to the value (0.87 nM) recorded for full-length. These data establish a central role for the DciA-Lasso domain in driving the helicase-loader interaction.



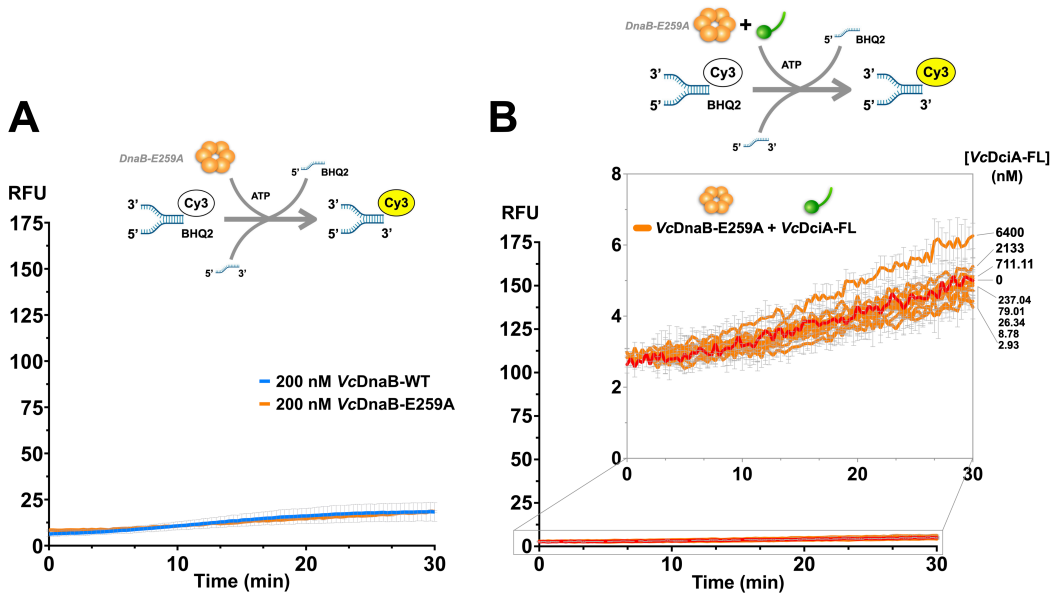
Supplementary Figure 19. The VcDciA KH, but not the Lasso, domain partially retains the ATPase suppression activity of full-length. A) The IC_{50} experiment and calculation described in Figure 6 were applied to the Vc KH and Lasso domains. Experimenting on four different protein batches, we find that the VcDciA KH domain represses VcDnaB ATPase activity with an IC_{50} of 442.5 nM ($R^2 = 0.9693$; 95% confidence interval: 380.5 to 520.0 nM), 16-fold weaker than the full-length loader. The blue dots represent individual measured ATPase rates plotted on a log scale. A non-linear fit (light-orange line), calculated using the Sigmoidal, 4PL model as implemented in PRISM, is superimposed on the experimental data. The dashed lines represent the 95% confidence interval (CI). B) The same analysis applied to the VcDciA Lasso domain revealed no effect on VcDnaB's ATPase activity at low concentrations (0.11 – 200 nM). However, the titration of a significant excess of the Lasso domain (800 – 6400 nM) revealed a 2-fold stimulation of DnaB's ATPase activity. The measured ATPase rates (blue dots) are plotted on a linear scale and connected with the grey line.



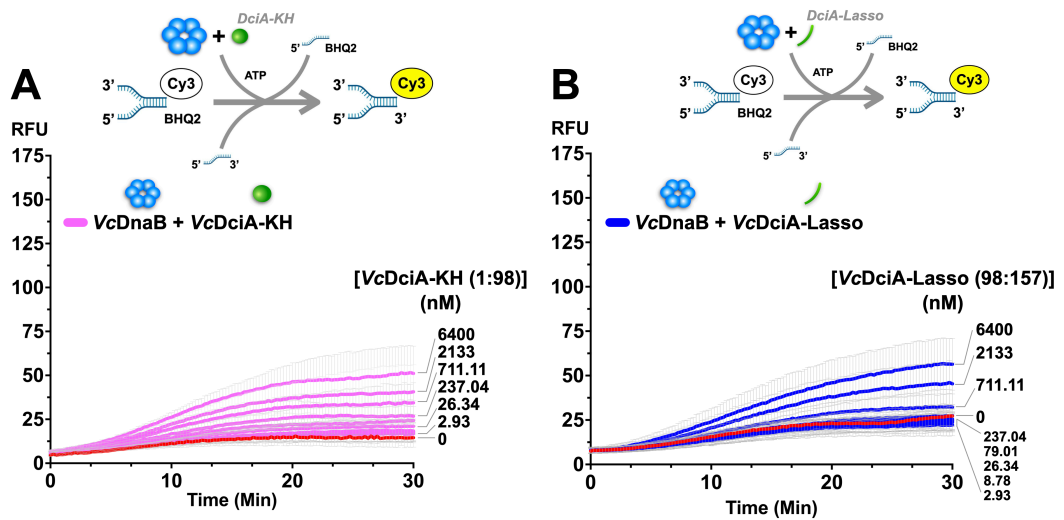
Supplementary Figure 20. MtDnaB is essentially inactive under the conditions tested. The basal ATPase activity of Vc and Mt DnaB were measured as described in the Methods. A comparison of the signals detected at 340 nM during the PK/LDH-coupled assay shows how *MtDnaB* is hardly more active than the measurements that lack protein. This finding is sustained at 200 nM (green) and 800 nM (violet) *MtDnaB* concentrations; both curves exhibit slopes nearly identical to the sample where the protein was omitted. By contrast, the rate produced by 200 nM *VcDnaB* (cyan) is ~10-fold higher.



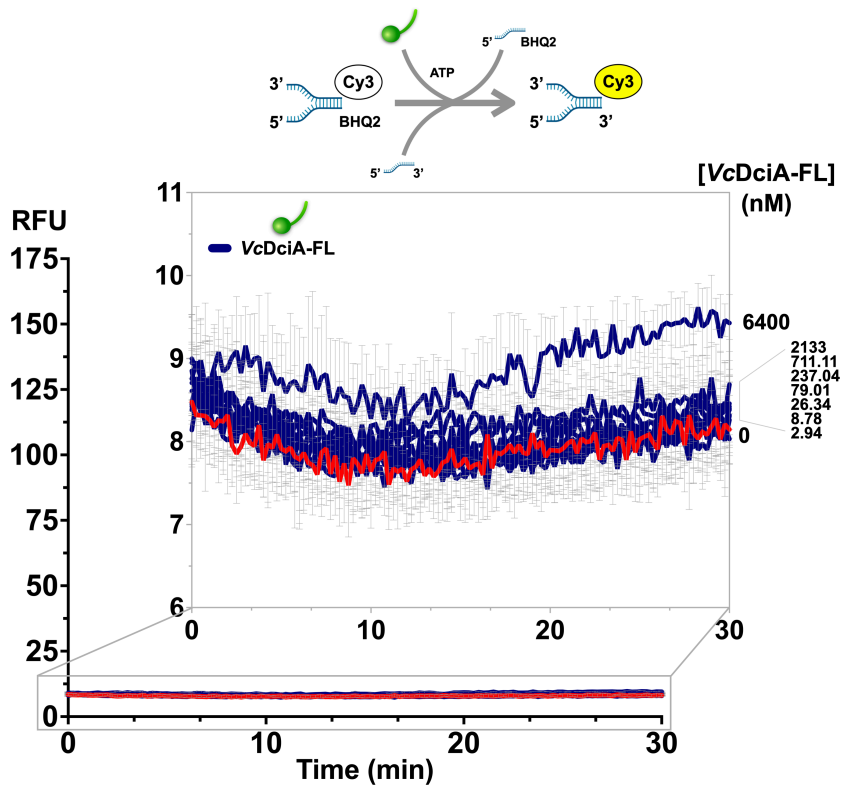
Supplementary Figure 21. VcDciA stimulates the DNA unwinding activity of VcDnaB in a concentration-dependent manner. A) The unwinding experiments employed a FRET-based 5' and 3' tailed replication fork mimic and are diagrammed schematically above. A time course of the evolution of fluorescence at 570 nm, due to relief of quenching at the Cy3/BHQ2 center during unwinding, is shown for unwinding experiments where the VcDnaB concentration was held fixed at 200 nM, and the VcDciA concentration varied from 0.11 to 6400 nM (orange curves). Shown in red is VcDnaB's unwinding when VcDciA was absent. The data imply a 10-fold stimulation of unwinding. B) The data in panel A are plotted as a function of %DNA unwound, where the highest RFU value was considered 100% unwound. By this measure, the VcDciA-KH (pink) and VcDciA-Lasso (blue) domains show 30% of the activity of full-length. Control measurements lacking VcDciA showed no unwinding, as was observed with the bubble substrate (Figure 7 and Supplementary Figure 25).



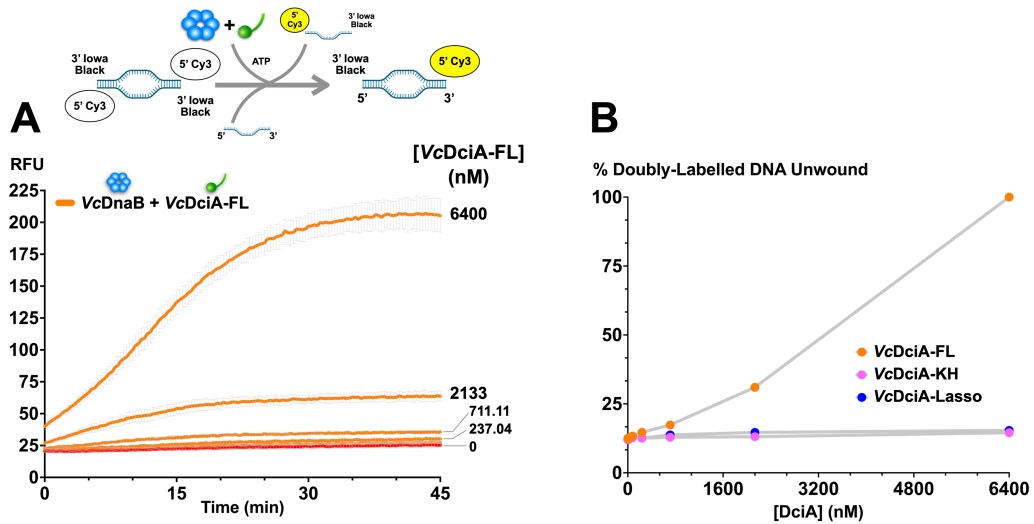
Supplementary Figure 22. The VcDnaB Walker B glutamate (E259A) mutation significantly reduces DNA unwinding activity and VcDciA-stimulated unwinding. The FRET-based DNA unwinding experiments used a 5' and 3' tailed replication fork mimic as diagrammed schematically above each panel. A) The time course of the 570 nm fluorescence signal due to the unwinding of the DNA substrate is plotted for wild-type (blue) and E259A mutant (orange) DnaB. In this experiment, the VcDnaB concentration was held at 200 nM. The blue line represents the average of 27 replicates measured from 200 nM VcDnaB-WT. Both wild-type and the E259A mutation show a slight basal fluorescence signal increase in the DNA fork unwinding assay. Although the E259A mutant behaves like wild-type in the fork unwinding assay, adding VcDciA (panel B) does not stimulate it as seen for wild-type B) The time course of the 570 nm fluorescence signal arising from titration of VcDciA (0.11-6400 nM) into an assay containing the Walker B VcDnaB-E259A mutant. Essentially, no unwinding activity was recovered. The inset shows the same data but expands a section of the plot.



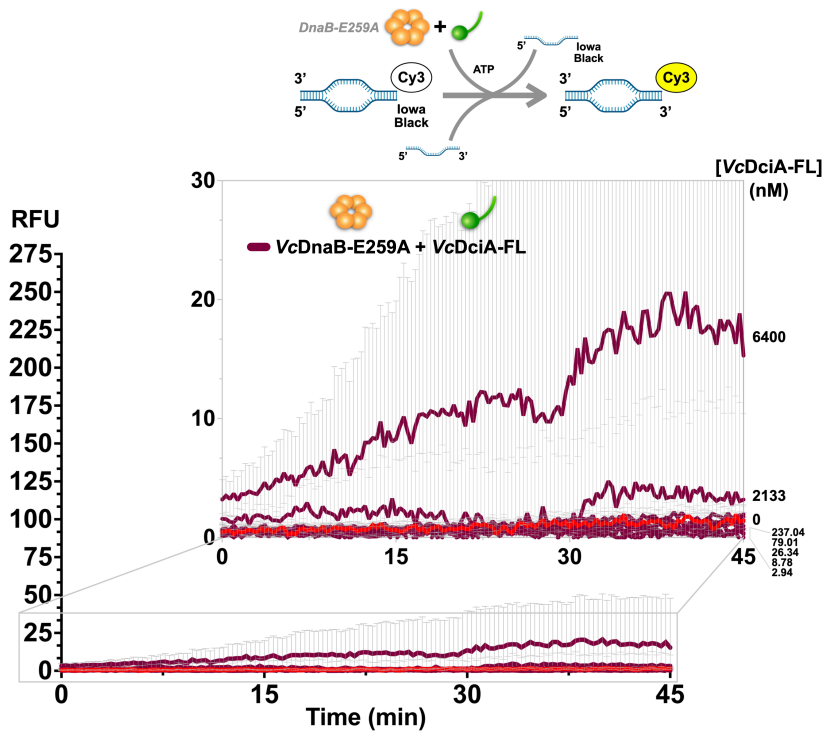
Supplementary Figure 23. The isolated KH and Lasso domains of VcDciA exhibit significantly reduced stimulatory action over the DnaB's unwinding activity. The FRET-based DNA unwinding experiments used a 5' and 3' tailed replication fork mimic as diagrammed schematically above each panel. Titration of VcDciA-KH (pink, A) or DciA-Lasso (blue) domains into the assay stimulated DnaB's unwinding activity by only 30% with respect to full-length DciA. The titration of appropriate DciA-construct encompassed concentrations from 0.11 to 6400 nM. At most, the isolated KH and Lasso domains stimulated DNA unwinding by 3.3-fold compared to a sample lacking DciA. By contrast, full-length DciA produces a 10-fold increase (Supplementary Figure 21).



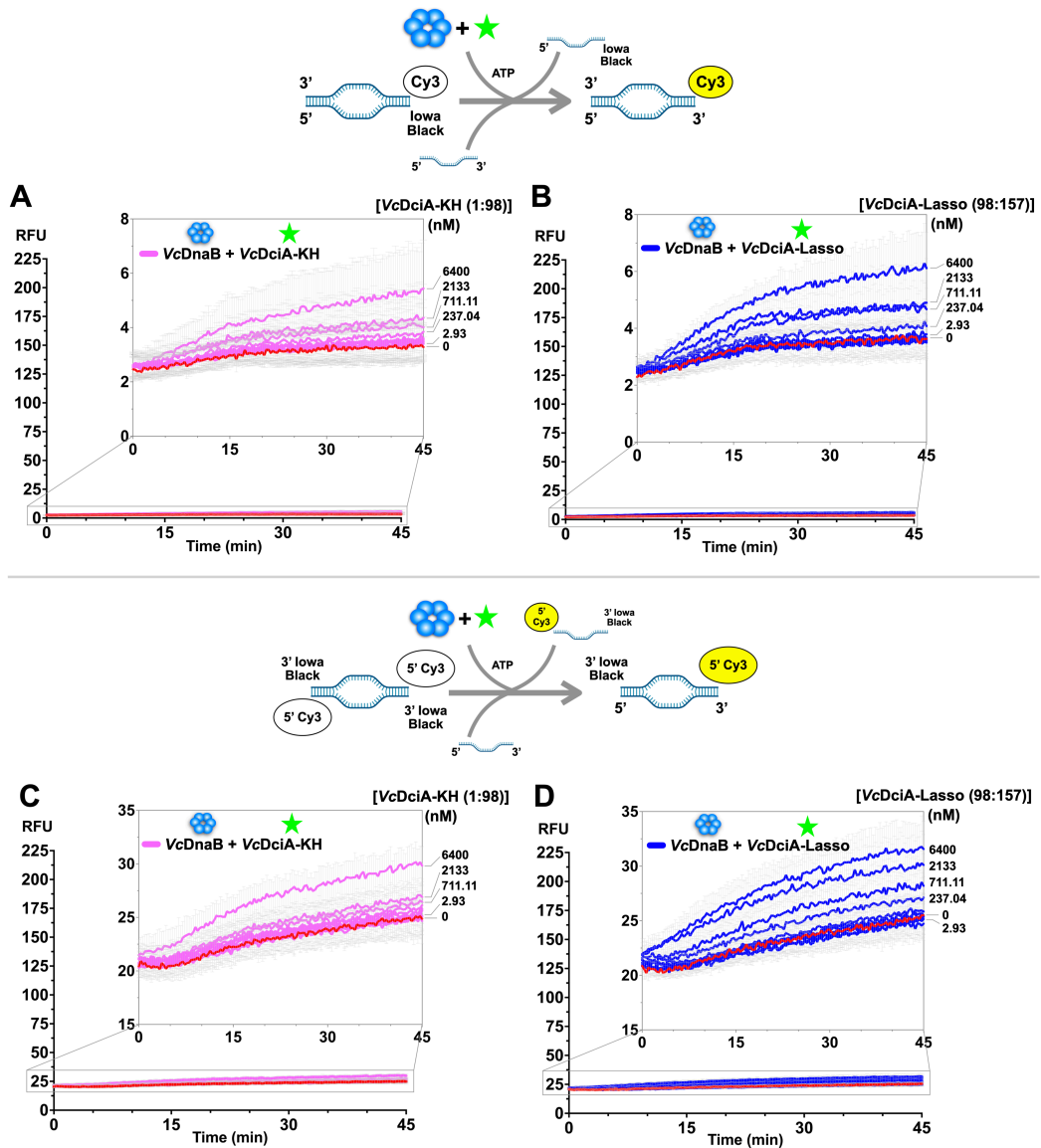
Supplementary Figure 24. Isolated VcDciA fails to unwind the 5' and 3' tailed replication fork-mimic. Titration from 0-6400 nM (blue curves) of VcDciA into an assay that lacks DnaB, as diagrammed schematically above, fails to show any 570 nm fluorescent signal arising from unwinding of the DNA. The inset shows a plot section expanded around RFU values of 6-11. The red curve describes fluorescent signals from an experiment that lacked DciA. These data suggest that DciA alone fails to unwind the fork DNA substrate; this finding contrasts with robust unwinding seen with the DNA bubble substrate; the reason for this difference in activity is unknown.



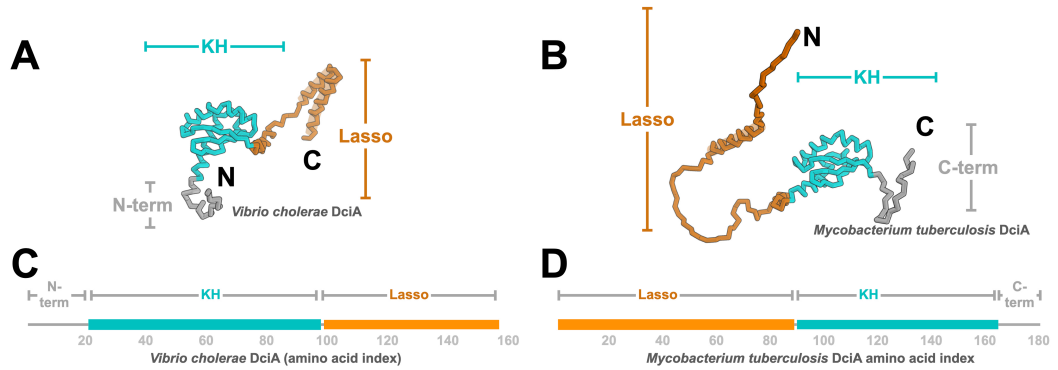
Supplementary Figure 25. Measurement of VcDciA-mediated loading of VcDnaB using a doubly labeled DNA bubble. The assay diagrammed above was performed precisely as the singly labeled DNA bubble substrate in Figure 7, except that FRET labels were included at both ends of the DNA bubble. A) A time course (orange curves) of fluorescence at 570 nm was recorded from an experiment wherein VcDciA was titrated from 0.11 to 6400 nM while the VcDnaB concentration was fixed at 200 nM. The red curve shows the signal from an experiment that lacked VcDciA. Surprisingly, the amplitude of the fluorescence signals did not increase two-fold compared to the singly labeled, as was expected, given the presence of two fluorescence centers; instead, the signal increased by 10%. B) Data from panel A plotted as a % DNA unwound as a function of VcDciA concentration. These data show no loading activity by the VcDciA-KH (pink) or Lasso domains (blue), even with a significant excess of each domain (Supplementary Figure 27).



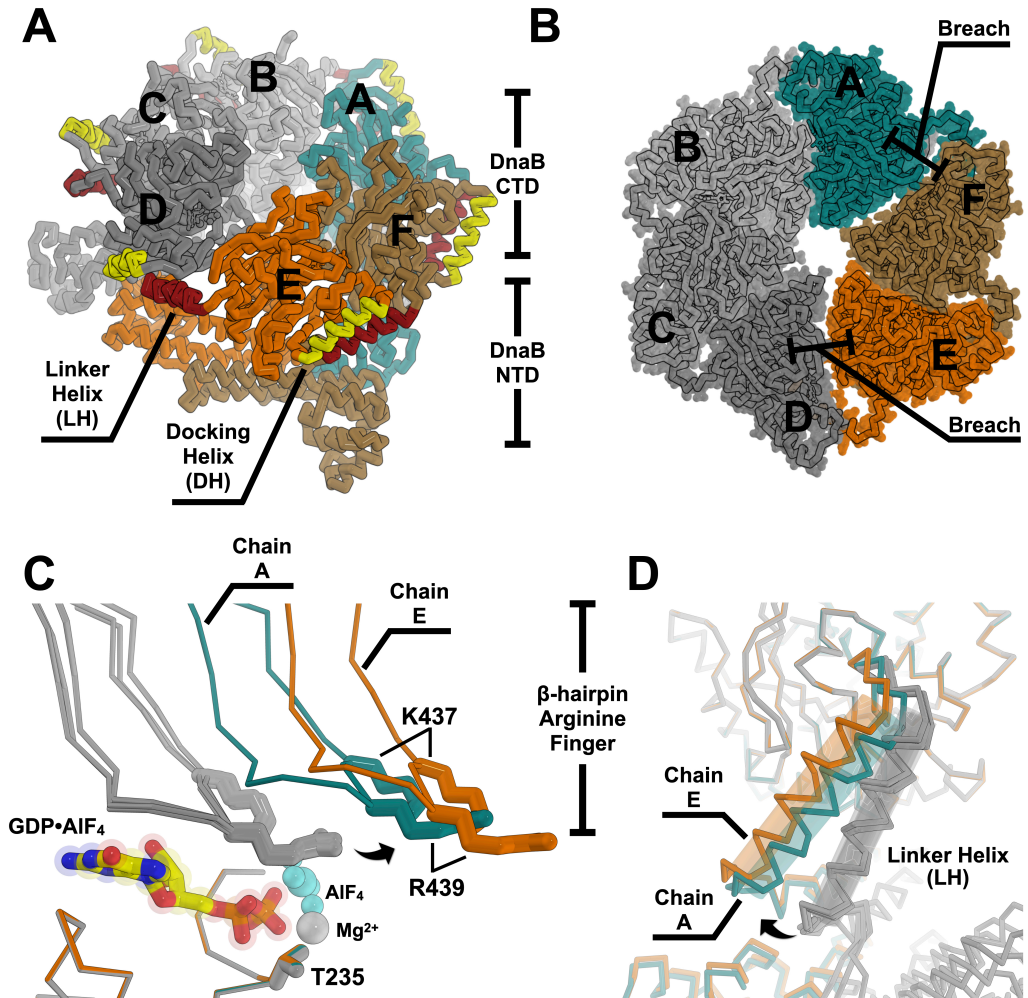
Supplementary Figure 26. The Vc Walker B glutamate (E259A) mutation is deficient in VcDnaB loading on a physiological bubble substrate. This experiment, diagrammed above the plot, is similar to that shown in Supplementary Figure 22, except that a DNA bubble substrate is used. The plot depicts a point-by-point subtraction of the fluorescence signals generated by VcDciA (taken from Figure 7B). The resulting curves (orange) show little loading activity above the basal rate measured without DciA (red). Only when DciA is present at 6400 nM is a curve with significant activity above the basal rate observed. However, the amplitude of this curve (~15) was 30% of the amplitude seen with wild-type VcDnaB. The inset shows the same data but expands a section of the plot.



Supplementary Figure 27. The isolated VcDciA KH and Lasso domains are inactive in the bubble-loading assay. These measurements diagrammed above each panel, represent controls for parallel studies with full-length VcDciA. Panels A and B employ the singly labeled substrate, and panels C and D employ the doubly labeled substrate. Titration of the appropriate DciA construct was carried out as described in Figure 7. Measurements with the VcDciA-KH domain are colored pink, and those with the VcDciA-Lasso domain are colored blue. The red curves in each plot represent the basal rate of loading measured in the absence of VcDciA. With the singly labeled substrate, a 1.5 to 2-fold increase in the fluorescence signal was seen at the highest concentration tested (6400 nM); in contrast, full-length VcDciA produces a 10-fold increase (Figure 7). With the doubly labeled substrate, a ~1.2-fold increase in the fluorescence signal was seen at the highest concentration tested (6400 nM); in contrast, full-length VcDciA produces an 8-fold increase (Supplementary Figure 25). The lack of differentiation of the DciA titration-derived curves from the basal rate (red) leads us to conclude that each VcDciA truncated construct is inactive in helicase loading.



Supplementary Figure 28. **Calculated structures of DciA loaders from *Vibrio Cholerae* and *Mycobacterium tuberculosis*.** Ribbon representation of the AlphaFold calculated structures of DciA from Vc¹²³ (A) and Mt¹²⁴ (B). Linear representation of VcDciA and MtDciA in panels C and D. The KH domain is colored cyan, the Lasso orange, and other segments in grey. VcDciA is also seen in the crystal structure of the VcDnaB-VcDciA complex (PDB: 8A3V)¹⁴.



Supplementary Figure 29. Crystal Structure of VcDnaB-GDP/AIF₄ (PDB: 6T66) and the disposition of β -hairpin arginine fingers and Linker Helix elements. A) VcDnaB is depicted as a ribbon with chains labeled. The NTD and CTD layers are indicated. The six Docking-Helix (DH) and Linker-Helix (LH) elements are shown in yellow and red, respectively. Chains A, F, and E are shown in blue, sand, and orange, and the other chains are in alternating shades of gray. B) The 6T66 DnaB structure depicts the two breaches in the CTD layer: the largest is between chains E (orange) and D (grey), and the narrowest is between chains A (deep teal) and F (sand). The protein surface is drawn with the PyMol solvent radius = 0.3 setting. C) The six CTD pairs (A-B, B-C, etc, residue 200:461) from 6T66 were superimposed (RMSD = 0.03 Å) on chain A, but where the DH elements (residue 280:306) and the second chain were excluded from the calculation. Each VcDnaB harbors a GDP•AIF₄ species, and only one is shown in the stick representation, which has transparent spheres overlaid. The β -hairpin arginine fingers are shown as ribbons, and K437 and K439 are shown as sticks. The β -hairpins from four chains (B, C, D, F) follow each other closely (grey), although not included in the alignment. The corresponding elements from chain A (deep teal) and E (orange) have moved away by ~9 Å and ~13 Å, respectively. D) The superposition from panel C was interrogated for the disposition of the LH elements represented as transparent cylinders. The LH elements from chains B, C, D, and F (grey) superimpose closely and occupy the LH1 position as defined in the main text and Supplementary Figure 4. In contrast, those belonging to chain A (deep teal) and chain E (orange) drift away by ~5-~9 Å and occupy position LH2.

SUPPLEMENTARY TABLE 1: Cryogenic Electron Microscopy Data Measurement and PDB Model Refinement

PDB	9DLS
EMDB	EMD-46984
Data Collection and Processing	
Microscope	Titan Krios
Voltage (kV)	300
Camera	Gatan K3
Magnification	81000
Defocus range (μm)	-0.7 to -2.9
Exposure time (ms)	2000
Electron dose ($\text{e}^-/\text{\AA}^2$)	51.37
Number of frames collected (no.)	40
Number of frames processed (no.)	40
Calibrated pixel size ($\text{\AA}$)	0.5413
Micrographs (no.)	12265
Total particle images (no.)	4315534
Refinement	
Particle per class (no.)	521399
Symmetry imposed	C1
Map resolution ($\text{\AA}$), FSC = 0.143	3.37
Map sharpening B-factor ($\text{\AA}^2$)	119.2
Map versus Model cross-correlation	0.7415
Model Composition	
Non-hydrogen atoms	21282
Protein residues	2648
Nucleotide residues	22
Ligand: ATP γ S	6

Ligand: Mg ²⁺	6
Model Refinement	
Resolution limit (Å)	3.37
Number of chains	14
Number of residues	2648
B factors (Å²)	
Protein	74.66
Nucleotide	145.76
Ligand (ATPγS)	55.68
RMS deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.426
Validation	
MolProbity score	1.48
Clashscore	4.88
Rotamer outliers (%)	2.00
Ramachandran plot (%)	
Favored (%)	98.07
Allowed (%)	1.93
Outliers (%)	0.00

REFERENCES

1. Parker, M. W., Botchan, M. R. & Berger, J. M. Mechanisms and regulation of DNA replication initiation in eukaryotes. *Crit Rev Biochem Mol* **52**, 1–41 (2017).
2. Burgers, P. M. J. & Kunkel, T. A. Eukaryotic DNA Replication Fork. *Annu Rev Biochem* **86**, 1–22 (2017).
3. Costa, A. & Diffley, J. F. X. The Initiation of Eukaryotic DNA Replication. *Annu Rev Biochem* **91**, 107–131 (2022).
4. O'Donnell, M., Langston, L. & Stillman, B. Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harbor perspectives in biology* **5**, a010108–a010108 (2013).
5. Greci, M. D. & Bell, S. D. Archaeal DNA Replication. *Annu Rev Microbiol* **74**, 1–16 (2020).
6. Li, H., Yao, N. Y. & O'Donnell, M. E. Anatomy of a twin DNA replication factory. *Biochem Soc T* **48**, 2769–2778 (2020).
7. Wang, G. et al. The structure of a DnaB-family replicative helicase and its interactions with primase. *Nat Struct Mol Biol* **15**, 94–100 (2008).
8. Strycharska, M. S. et al. Nucleotide and Partner-Protein Control of Bacterial Replicative Helicase Structure and Function. *Mol. Cell* **52**, 844–854 (2013).
9. Bailey, S., Eliason, W. K. & Steitz, T. A. Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science (New York, NY)* **318**, 459–463 (2007).
10. Itsathitphaisarn, O., Wing, R. A., Eliason, W. K., Wang, J. & Steitz, T. A. The Hexameric Helicase DnaB Adopts a Nonplanar Conformation during Translocation. *Cell* **151**, 267–277 (2012).
11. Cargemel, C. et al. Structural Insights of the DciA Helicase Loader in Its Relationship with DNA. *Int J Mol Sci* **24**, 1427 (2023).
12. Marsin, S. et al. Study of the DnaB:DciA interplay reveals insights into the primary mode of loading of the bacterial replicative helicase. *Nucleic Acids Res.* **49**, gkab463- (2021).
13. Cargemel, C. et al. The apo-form of the *Vibrio cholerae* replicative helicase DnaB is a labile and inactive planar trimer of dimers. *Febs Lett* (2022) doi:10.1002/1873-3468.14403.
14. Cargemel, C. et al. The LH-DH module of the bacterial replicative helicases is the common binding site for DciA and other helicase loaders. (2022) doi:10.1101/2022.07.08.499335.
15. Lo, Y.-H. et al. The crystal structure of a replicative hexameric helicase DnaC and its complex with single-stranded DNA. *Nucleic Acids Research* **37**, 804–814 (2009).
16. Liu, B., Eliason, W. K. & Steitz, T. A. Structure of a helicase-helicase loader complex reveals insights into the mechanism of bacterial primosome assembly. *Nature communications* **4**, 2495 (2013).
17. Bazin, A., Cherrier, M. V., Gutsche, I., Timmins, J. & Terradot, L. Structure and primase-mediated activation of a bacterial dodecameric replicative helicase. *Nucleic Acids Research* **43**, 8564–8576 (2015).
18. Nagata, K. et al. Crystal structure of the complex of the interaction domains of *E. coli* DnaB helicase and DnaC helicase loader: Structural basis implying a distortion-accumulation mechanism for the DnaB ring opening caused by DnaC binding. *Journal of biochemistry* **5**, a010108 (2019).

19. Arias-Palomo, E., Puri, N., Murray, V. L. O., Yan, Q. & Berger, J. M. Physical Basis for the Loading of a Bacterial Replicative Helicase onto DNA. *Mol Cell* **74**, 173-184.e4 (2019).
20. Arias-Palomo, E., O'Shea, V. L., Hood, I. V. & Berger, J. M. The Bacterial DnaC Helicase Loader Is a DnaB Ring Breaker. *Cell* **153**, 438-448 (2013).
21. O'Donnell, M. E. & Li, H. The ring-shaped hexameric helicases that function at DNA replication forks. *Nat. Struct. Mol. Biol.* **25**, 122-130 (2018).
22. Fernandez, A. J. & Berger, J. M. Mechanisms of hexameric helicases. *Crit Rev Biochem Mol* 1-19 (2021) doi:10.1080/10409238.2021.1954597.
23. Bleichert, F., Botchan, M. R. & Berger, J. M. Mechanisms for initiating cellular DNA replication. *Science* **355**, eaah6317 (2017).
24. Gao, Y. & Yang, W. Different mechanisms for translocation by monomeric and hexameric helicases. *Curr. Opin. Struct. Biol.* **61**, 25-32 (2020).
25. Spinks, R. R., Spenkelink, L. M., Dixon, N. E. & Oijen, A. M. van. Single-Molecule Insights Into the Dynamics of Replicative Helicases. *Frontiers Mol Biosci* **8**, 741718 (2021).
26. Perera, H. M., Behrmann, M. S., Hoang, J. M., Griffin, W. C. & Trakselis, M. A. Contacts and context that regulate DNA helicase unwinding and replisome progression. *Enzym.* **45**, 183-223 (2019).
27. Medagli, B. & Onesti, S. DNA Helicases and DNA Motor Proteins. *Adv. Exp. Med. Biol.* **767**, 75-95 (2012).
28. Hingorani, M. M. & Patel, S. S. Interactions of bacteriophage T7 DNA primase/helicase protein with single-stranded and double-stranded DNAs. *Biochemistry* **32**, 12478-12487 (1993).
29. Gao, Y. et al. Structures and operating principles of the replisome. *Science* **363**, eaav7003 (2019).
30. Kaguni, J. M. DnaA, DnaB, and DnaC. in 1-14 (Springer New York, 2014). doi:10.1007/978-1-4614-6436-5_142-1.
31. Learn, B. A., Um, S. J., Huang, L. & McMacken, R. Cryptic single-stranded-DNA binding activities of the phage lambda P and Escherichia coli DnaC replication initiation proteins facilitate the transfer of E. coli DnaB helicase onto DNA. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 1154-1159 (1997).
32. Bailey, S. et al. Structural Analysis of Bacillus subtilis SPP1 Phage Helicase Loader Protein G39P*. *J Biol Chem* **278**, 15304-15312 (2003).
33. Seco, E. M. et al. Bacteriophage SPP1 DNA replication strategies promote viral and disable host replication in vitro. *Nucleic Acids Res* **41**, 1711-1721 (2013).
34. Velten, M. et al. A Two-Protein Strategy for the Functional Loading of a Cellular Replicative DNA Helicase. *Mol Cell* **11**, 1009-1020 (2003).
35. Jones, C. E., Mueser, T. C., Dudas, K. C., Kreuzer, K. N. & Nossal, N. G. Bacteriophage T4 gene 41 helicase and gene 59 helicase-loading protein: A versatile couple with roles in replication and recombination. *Proc. Natl. Acad. Sci.* **98**, 8312-8318 (2001).
36. Odegrip, R., Schoen, S., Haggård-Ljungquist, E., Park, K. & Chatteraj, D. K. The Interaction of Bacteriophage P2 B Protein with Escherichia coli DnaB Helicase. *J. Virol.* **74**, 4057-4063 (2000).

37. Brézellec, P., Vallet-Gely, I., Possoz, C., Quevillon-Cheruel, S. & Ferat, J.-L. DciA is an ancestral replicative helicase operator essential for bacterial replication initiation. *Nature communications* **7**, 13271 (2016).
38. Brézellec, P. et al. Domestication of lambda phage genes into a putative third type of Replicative Helicase Matchmaker. *Genome Biol Evol* **9**, evx111 (2017).
39. Marsin, S. et al. Study of the DnaB:DciA interplay reveals insights into the primary mode of loading of the bacterial replicative helicase. *Nucleic Acids Res* gkab463- (2021) doi:10.1093/nar/gkab463.
40. Blaine, H. C., Burke, J. T., Ravi, J. & Stallings, C. L. DciA Helicase Operators Exhibit Diversity across Bacterial Phyla. *J Bacteriol* **204**, e00163-22 (2022).
41. Blaine, H. C., Simmons, L. A. & Stallings, C. L. Diverse Mechanisms of Helicase Loading during DNA Replication Initiation in Bacteria. *J. Bacteriol.* **205**, e00487-22 (2023).
42. Ozaki, S., Wang, D., Wakasugi, Y., Itani, N. & Katayama, T. The *Caulobacter crescentus* DciA promotes chromosome replication through topological loading of the DnaB replicative helicase at replication forks. *Nucleic Acids Res.* **50**, gkac1146- (2022).
43. Chase, J., Berger, J. & Jeruzalmi, D. Convergent evolution in two bacterial replicative helicase loaders. *Trends Biochem Sci* **47**, 620–630 (2022).
44. Chase, J. et al. Mechanisms of opening and closing of the bacterial replicative helicase. *Elife* **7**, 1822 (2018).
45. Bell, S. P. & Kaguni, J. M. Helicase loading at chromosomal origins of replication. *Cold Spring Harbor perspectives in biology* **5**, a010124 (2013).
46. Ayora, S., Stasiak, A. & Alonso, J. C. The *Bacillus subtilis* bacteriophage SPP1 G39P delivers and activates the G40P DNA helicase upon interacting with the G38P-bound replication origin. *J. Mol. Biol.* **288**, 71–85 (1999).
47. Ioannou, C., Schaeffer, P. M., Dixon, N. E. & Soultanas, P. Helicase binding to DnaI exposes a cryptic DNA-binding site during helicase loading in *Bacillus subtilis*. *Nucleic Acids Res* **34**, 5247–5258 (2006).
48. Brown, D. et al. An Autoinhibited Conformation of the DnaB-Replicative Helicase – phage λ P Loader Complex. *Biorxiv* 2022.12.30.522210 (2022) doi:10.1101/2022.12.30.522210.
49. AlphaFold-Protein-Structure-Database. Inhibitor_G39P domain-containing protein. <https://alphafold.ebi.ac.uk/entry/A0A0M0X576>.
50. Loscha, K. V. et al. A novel zinc-binding fold in the helicase interaction domain of the *Bacillus subtilis* DnaI helicase loader. *Nucleic Acids Research* **37**, 2395–2404 (2009).
51. Tsai, K.-L., Lo, Y.-H., Sun, Y.-J. & Hsiao, C.-D. Molecular interplay between the replicative helicase DnaC and its loader protein DnaI from *Geobacillus kaustophilus*. *Journal of Molecular Biology* **393**, 1056–1069 (2009).
52. AlphaFoldProteinStructureDatabase. *Bacillus subtilis* Primosomal protein DnaI. <https://alphafold.ebi.ac.uk/entry/P06567>.
53. Mueser, T. C., Jones, C. E., Nossal, N. G. & Hyde, C. C. Bacteriophage T4 gene 59 helicase assembly protein binds replication fork DNA. The 1.45 Å resolution crystal structure reveals a novel α -helical two-domain fold¹¹Edited by P. E. Wright. *J. Mol. Biol.* **296**, 597–612 (2000).
54. Chan-Yao-Chong, M., Marsin, S., Quevillon-Cheruel, S., Durand, D. & Ha-Duong, T. Structural ensemble and biological activity of DciA intrinsically disordered region. *J Struct Biol* **212**, 107573 (2020).

55. O'Shea, V. L. & Berger, J. M. Loading strategies of ring-shaped nucleic acid translocases and helicases. *Current Opinion in Structural Biology* **25**, 16–24 (2014).
56. Davey, M. J. & O'Donnell, M. Replicative helicase loaders: ring breakers and ring makers. *Curr Biol* **13**, R594–R596 (2003).
57. Chao, M. C. et al. High-resolution definition of the *Vibrio cholerae* essential gene set with hidden Markov model-based analyses of transposon-insertion sequencing data. *Nucleic Acids Res.* **41**, 9033–9048 (2013).
58. Mann, K. M. et al. Rv0004 is a new essential member of the mycobacterial DNA replication machinery. *PLOS Genetics* **13**, e1007115 (2017).
59. Quevillon-Cheruel, S. et al. DciA, the Bacterial Replicative Helicase Loader, is a DNA Condenser. (2024) doi:10.2139/ssrn.4880687.
60. Davey, M. J., Fang, L., McNerney, P., Georgescu, R. E. & O'Donnell, M. The DnaC helicase loader is a dual ATP/ADP switch protein. *The EMBO journal* **21**, 3148–3159 (2002).
61. Biswas, T. & Tsodikov, O. V. Hexameric ring structure of the N-terminal domain of Mycobacterium tuberculosis DnaB helicase. *The FEBS journal* **275**, 3064–3071 (2008).
62. Oakley, A. J. & Xu, Z. Q. *E. coli* DnaB bound to ssDNA and AMPPNP. (2022) doi:10.2210/pdb7t20/pdb.
63. Mesa, P., Alonso, J. C. & Ayora, S. Bacillus subtilis Bacteriophage SPP1 G40P Helicase Lacking the N-terminal Domain Unwinds DNA Bidirectionally. *J. Mol. Biol.* **357**, 1077–1088 (2006).
64. Reha-Krantz, L. J. & Hurwitz, J. The dnaB gene product of Escherichia coli. I. Purification, homogeneity, and physical properties. *J. Biol. Chem.* **253**, 4043–4050 (1978).
65. Yu, X., Jezewska, M. J., Bujalowski, W. & Egelman, E. H. The hexameric *E. coli* DnaB helicase can exist in different Quaternary states. *Journal of Molecular Biology* **259**, 7–14 (1996).
66. Bujalowski, W., Klonowska, M. M. & Jezewska, M. J. Oligomeric structure of Escherichia coli primary replicative helicase DnaB protein. *J. Biol. Chem.* **269**, 31350–31358 (2001).
67. Nakayama, N., Arai, N., Kaziro, Y. & Arai, K. Structural and functional studies of the dnaB protein using limited proteolysis. Characterization of domains for DNA-dependent ATP hydrolysis and for protein association in the primosome. *J. Biol. Chem.* **259**, 88–96 (1984).
68. Günther, E., Mikolajczyk, M. & Schuster, H. Stabilization by ATP and ADP of Escherichia coli dnaB protein activity. *J. Biol. Chem.* **256**, 11970–11973 (1981).
69. Kaplan, D. L. & Steitz, T. A. DnaB from Thermus aquaticus Unwinds Forked Duplex DNA with an Asymmetric Tail Length Dependence*. *J Biol Chem* **274**, 6889–6897 (1999).
70. Schlierf, M., Wang, G., Chen, X. S. & Ha, T. Hexameric helicase G40P unwinds DNA in single base pair steps. *eLife* **8**, e42001 (2019).
71. Notarnicola, S. M. & Richardson, C. C. The nucleotide binding site of the helicase/primase of bacteriophage T7. Interaction of mutant and wild-type proteins. *J. Biol. Chem.* **268**, 27198–27207 (1993).
72. Patel, S. S., Hingorani, M. M. & Ng, W. M. The K318A Mutant of Bacteriophage T7 DNA Primase-Helicase Protein Is Deficient in Helicase But Not Primase Activity and Inhibits Primase-Helicase Protein Wild-Type Activities by Heterooligomer Formation. *Biochemistry* **33**, 7857–7868 (1994).

73. Crampton, D. J., Mukherjee, S. & Richardson, C. C. DNA-induced switch from independent to sequential dTTP hydrolysis in the bacteriophage T7 DNA helicase. *Molecular Cell* **21**, 165–174 (2006).
74. Crampton, D. J., Guo, S., Johnson, D. E. & Richardson, C. C. The arginine finger of bacteriophage T7 gene 4 helicase: Role in energy coupling. *Proc. Natl. Acad. Sci.* **101**, 4373–4378 (2004).
75. Makowska-Grzyska, M. & Kaguni, J. M. Primase Directs the Release of DnaC from DnaB. *Mol Cell* **37**, 90–101 (2010).
76. Polissi, A., Goffin, L. & Georgopoulos, C. The Escherichia coli heat shock response and bacteriophage lambda development. *FEMS microbiology reviews* **17**, 159–169 (1995).
77. Zylicz, M. The Escherichia coli chaperones involved in DNA replication. *Philos. Trans. R. Soc. Lond. Ser. B: Biol. Sci.* **339**, 271–278 (1993).
78. Spinks, R. R. et al. DnaB helicase dynamics in bacterial DNA replication resolved by single-molecule studies. *Nucleic Acids Res* **49**, gkab493- (2021).
79. Green, M. R. & Sambrook, J. *Molecular Cloning: A Laboratory Manual*. (Cold Spring Harbor Laboratory Press, 2012).
80. Kelley, D. S. et al. Mycobacterial DnaB helicase intein as oxidative stress sensor. *Nat. Commun.* **9**, 4363 (2018).
81. Zhang, H., Zhang, Z., Yang, J. & He, Z. Functional characterization of DnaB helicase and its modulation by single-stranded DNA binding protein in Mycobacterium tuberculosis. *Febs J* **281**, 1256–1266 (2014).
82. Olinares, P. D. B. et al. Native Mass Spectrometry-Based Screening for Optimal Sample Preparation in Single-Particle Cryo-EM. *Structure* **29**, 186-195.e6 (2021).
83. Olinares, P. D. B. & Chait, B. T. The Eukaryotic RNA Exosome, Methods and Protocols. *Methods Mol. Biol.* **2062**, 357–382 (2019).
84. Marty, M. T. et al. Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).
85. Reid, D. J. et al. MetaUniDec: High-Throughput Deconvolution of Native Mass Spectra. *J. Am. Soc. Mass Spectrom.* **30**, 118–127 (2019).
86. Lindsley, J. E. DNA Topoisomerase Protocols. *Methods Mol Biology Clifton N J* **95**, 57–64 (2000).
87. Kiianitsa, K., Solinger, J. A. & Heyer, W.-D. NADH-coupled microplate photometric assay for kinetic studies of ATP-hydrolyzing enzymes with low and high specific activities. *Analytical Biochemistry* **321**, 266–271 (2003).
88. Cheng, A. et al. Leginon: New features and applications. *Protein Sci.* **30**, 136–150 (2021).
89. Cheng, A. et al. Fully automated multi-grid cryoEM screening using Smart Leginon. *IUCrJ* **10**, 77–89 (2023).
90. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods* **14**, 290–296 (2017).
91. Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat Methods* **17**, 1214–1221 (2020).
92. Punjani, A. & Fleet, D. J. 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J Struct Biol* **213**, 107702 (2021).

93. Punjani, A. & Fleet, D. J. 3DFlex: determining structure and motion of flexible proteins from cryo-EM. *Nat Methods* **1–11** (2023) doi:10.1038/s41592-023-01853-8.
94. Rubinstein, J. L. & Brubaker, M. A. Alignment of cryo-EM movies of individual particles by optimization of image translations. *J. Struct. Biol.* **192**, 188–195 (2015).
95. Stagg, S. M., Noble, A. J., Spilman, M. & Chapman, M. S. ResLog plots as an empirical metric of the quality of cryo-EM reconstructions. *J. Struct. Biol.* **185**, 418–426 (2014).
96. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION -3.1. *IUCrJ* **7**, 253–267 (2020).
97. Bell, J. M., Chen, M., Durmaz, T., Fluty, A. C. & Ludtke, S. J. New software tools in EMAN2 inspired by EMDatabank map challenge. *J. Struct. Biol.* **204**, 283–290 (2018).
98. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun Biology* **4**, 874 (2021).
99. Jamali, K. et al. Automated model building and protein identification in cryo-EM maps. *Nature* **628**, 450–457 (2024).
100. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta crystallographica Section D, Biological crystallography* **60**, 2126–2132 (2004).
101. Brown, A. et al. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta crystallographica Section D, Biological crystallography* **71**, 136–153 (2015).
102. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology* **75**, 861–877 (2019).
103. Williams, C. J. et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
104. Echols, N. et al. Automating crystallographic structure solution and refinement of protein-ligand complexes. *Acta crystallographica Section D, Biological crystallography* **70**, 144–154 (2014).
105. Zwart, P. H. et al. Automated structure solution with the PHENIX suite. *Methods in molecular biology (Clifton, NJ)* **426**, 419–435 (2008).
106. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr Sect D* **74**, 531–544 (2018).
107. Terwilliger, T. C. et al. Model morphing and sequence assignment after molecular replacement. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **69**, 2244–2250 (2013).
108. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta crystallographica Section D, Biological crystallography* **68**, 352–367 (2012).
109. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica Section D, Biological crystallography* **66**, 213–221 (2010).
110. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta crystallographica Section D, Biological crystallography* **67**, 235–242 (2011).

111. Kleywegt, G., Zou, J., Kjeldgaard, M. & Jones, T. *International Tables for Crystallography, Vol. F, Edited by MG Rossmann & E. Arnold, Ch. 17.1, Pp. 353–356, 366–367.* (2001).
112. Sierk, M. & Kleywegt, G. Deja Vu All Over Again: Finding and Analyzing Protein Structure Similarities. *Structure (London, England : 1993)* **12**, 2103–2111 (2004).
113. Kleywegt, G. J. Quality control and validation. *Methods in molecular biology (Clifton, NJ)* **364**, 255–272 (2007).
114. Kleywegt, G. & Jones, T. Model building and refinement practice. *Macromolecular Crystallography, Pt A* **277**, 208–230 (1997).
115. Pettersen, E. F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* **25**, 1605–1612 (2004).
116. LLC, S., . The PyMOL Molecular Graphics System, Version 2.1. (2017).
117. Luscombe, N. M., Laskowski, R. A. & Thornton, J. M. NUCPLOT: a program to generate schematic diagrams of protein-nucleic acid interactions. *Nucleic Acids Research* **25**, 4940–4945 (1997).
118. Morin, A. et al. Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).
119. Núñez-Ramírez, R. et al. Loading a ring: structure of the Bacillus subtilis DnaB protein, a co-loader of the replicative helicase. *Journal of Molecular Biology* **367**, 764–769 (2007).
120. Oakley, A. J. & Xu, Z. Q. *E. coli* DnaB bound to ssDNA and ADP-Alf4. (2023) doi:10.2210/pdb7t21/pdb.
121. Oakley, A. J. & Xu, Z. Q. *E. coli* DnaB bound to three DnaG C-terminal domains, ssDNA, ADP and Alf4. (2023) doi:10.2210/pdb7t22/pdb.
122. Oakley, A. J. & Xu, Z. Q. *E. coli* DnaB bound to two DnaG C-terminal domains, ssDNA, ADP and Alf4. (2023) doi:10.2210/pdb7t23/pdb.
123. DUF721 Domain-Containing Protein (VcDciA). <https://alphafold.ebi.ac.uk/entry/A0A5C2AYF9> (2024).
124. UPF0232 protein Rv0004 (MtDciA). (2024).
125. Tartoff, K. D. & Hobbs, C. A. Improved Media for Growing Plasmid and Cosmid Clones. *Bethesda Research Laboratories Focus* **9**, 12 (1987).
126. Wood, W. N., Smith, K. D., Ream, J. A. & Lewis, L. K. Enhancing yields of low and single copy number plasmid DNAs from Escherichia coli cells. *J. Microbiol. Methods* **133**, 46–51 (2017).
127. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning.* (Cold Spring Harbor Laboratory Press, 1989).

Supporting Information

DnaB and DciA: mechanisms of helicase loading and translocation on ssDNA

Nicholas Gao^{*1,2}, Daniele Mazzoletti^{*5}, Adele Peng¹, Paul Dominic B. Olinares⁶, Castrese Morrone^{§5}, Andrea Garavaglia^{^5}, Nourelhoda Gouda¹, Serge Tsoy¹, Albert Mendoza¹, Ahammad Chowdhury¹, Antonio Cerullo^{1,4}, Hrutvik Bhavsar¹, Franca Rossi⁵, Menico Rizzi⁵, Brian T. Chait⁶, Riccardo Miggiano^{#5}, and David Jeruzalmi^{#1,2,3,4}

¹ *Department of Chemistry and Biochemistry, City College of New York, New York, NY 10031, USA.*

Ph.D. Programs in ²Biochemistry, ³Biology, and ⁴Chemistry, The Graduate Center of the City University of New York, New York, NY 10016, USA.

⁴ *Department of Pharmaceutical Sciences, University of Piemonte Orientale, Novara, 28100, Italy*

⁵ *Laboratory for Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, New York, New York 10065, USA*

[°] *Present address: Mirimus Inc, 760 Parkside Ave, Brooklyn, NY 11226, USA*

[§] *Present address: Telethon Institute of Genetics and Medicine (TIGEM) Via Campi Flegrei, 34 80078 Pozzuoli (NA) Italy.*

[^] *Present address: Department of Pharmaceutical Science (DISFARM), University of Milan, Milan, 20133 Italy*

[#] *Corresponding authors: dj@ccny.cuny.edu or riccardo.miggiano@uniupo.it*

^{*} *These authors contributed equally.*

SUPPLEMENTARY METHODS

Plasmids, Protein Expression, Purification, and Biochemistry

Expression Plasmid Construction

Genes corresponding to full-length *Vibrio cholerae* (Vc) DnaB and VcDciA were amplified using polymerase chain reaction (PCR) and genomic DNA from *Vibrio cholerae* N16961 (ATCC # 39315D-5), and, variously, cloned into the following protein expression vectors (pET24a, pCDFDuet-1, or pRSFDuet-1). These efforts encompassed untagged, amino, or carboxy-terminal hexahistidine-tagged proteins. Point mutations were introduced using the Quickchange Site Directed Mutagenesis kit (Agilent).

E. coli DnaB was prepared as previously described of the *E. coli* DnaB• λ P complex, except that λ P was omitted ⁴⁴.

Protein Expression

For production-scale bacterial growths, colonies corresponding to bacterial transformants were picked and resuspended in LB media supplemented with the appropriate antibiotic concentration. Bacterial cultures were grown at 37 °C to an OD600 $\approx$ 0.6 – 0.7 (RPM = 180). Protein expression was induced by making the culture 0.5 mM in isopropyl-D-1-thiogalactopyranoside (IPTG), followed by a temperature shift of the culture to 16 °C; protein production was allowed to continue for 14-16 hours (RPM = 180). Cells were recovered by centrifugation, resuspended in 20 mM Tris-HCl pH 7.4, 500 mM NaCl, 5 mM MgCl₂, 10% sucrose, 5mM β -mercaptoethanol at a ratio of 5 mL buffer per gram of wet cells, and frozen at -80°C.

VcDnaB. VcDnaB-untagged (pNG066, 50 μ g/mL kanamycin), VcDnaB-NHis (pNG067, 50 μ g/mL kanamycin), VcDnaB-E259A-NHis (pNG072, 50 μ g/mL

kanamycin), VcDnaB-NTD-untagged (pNG114, 50 µg/ml kanamycin), VcDnaB-NTD_LH-untagged (pNG131, 50 µg/mL kanamycin) and VcDnaB-LH_CTD-untagged (pNG134, 50 µg/mL kanamycin) were expressed using the above procedure.

VcDciA. VcDciA-FL-CHis (pNG057, 50 µg/ml streptomycin) and VcDciA-CTD-CHis (pNG087, 50 µg/ml streptomycin) were expressed using the procedure above, except for three changes: a) the media was LB for pNG057 and SB^{125,126}; for pNG087, b). After adding 0.5 mM IPTG to the culture, protein expression was allowed to continue for 3 hours at 37 °C, and c) cells were resuspended after harvesting in 20 mM Tris-HCl pH 8, 300 mM NaCl, 5% glycerol, 5 mM β -mercapoethanol.

VcDciA-NTD-CHis (pNG078, 50 µg/ml streptomycin) was grown using the same procedure as full-length DciA, except that cells were resuspended in 20 mM Tris-HCl pH 8, 300 mM NaCl, 5% glycerol, and 5 mM β -mercapoethanol.

VcDnaB-VcDciA coexpression. VcDnaB-untagged (pNG066, streptomycin: 50 ug/mL) and VcDciA-FL-Chis (pNG057, Kanamycin: 50 ug/mL) were co-expressed using the above procedure altered to account for the co-expression.

Protein purification

Unless otherwise indicated, standard protein biochemistry techniques were used to purify all the proteins in this study (Supplementary Figure 8). All manipulations were carried out at 4 °C.

Unless otherwise indicated, 90 - 160 mls of cells expressing the protein of interest were lysed using a Sonic dismembrator (Model 500, Fisher Scientific) at the following settings: 3 min at 70% amplitude, with alternating 30s on and off cycles (Vc) and 15 min at 50% amplitude, with alternating cycles of 30s on and 60s off

(Mt). The protein of interest was recovered in the soluble fraction by centrifugation (RPM=16,000, JA-17, Beckman-Coulter).

VcDnaB. Untagged VcDnaB was purified by combining ammonium sulfate (AS) precipitation and heparin, hydroxyapatite, and size exclusion chromatography. VcDnaB was found in the supernatant after making the soluble fraction 20% of saturated AS (SAS), but in the pellet when treated with 45% SAS. The 45% SAS pellet was resuspended in heparin buffer A below and purified over a 50 mL Heparin Sepharose Fast Flow (Cytiva), with the following buffers: heparin buffer A: 20 mM Tris-HCl, pH = 7.4, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol; Heparin buffer B: A + 1 M NaCl and a linear gradient from 0% to 40% B. Heparin-purified VcDnaB was purified by hydroxyapatite chromatography (65 ml, CHT, Bio-Rad) in the following buffers: A: 5 mM potassium phosphate, pH = 7.5, 100 mM NaCl, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol; B: A + 250 mM potassium phosphate, pH = 7.5). CHT-purified VcDnaB was concentrated by 45% AS precipitation, redissolved in 6 mL of 20 mM potassium HEPES, pH=7.5, 500 mM NaCl, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol), and applied to a 320mL Superdex200 Prep Grade (Cytiva) column equilibrated in the above buffer. Size exclusion chromatography fractions free of nuclease activity were pooled, concentrated (Corning, 20 ml, 10 kDa molecular weight cutoff) to 55 mg/ml, aliquoted (~10-20 μ l aliquots), flash frozen in liquid nitrogen, and maintained at -80 °C until use.

VcDnaB-NHis and VcDnaB-E259A-NHis were purified using the protocol employed with untagged VcDnaB but with discrete changes. The 45% SAS pellet was resuspended in Ni²⁺-Nitrilotriacetic Acid (NiNTA) Buffer A (20 mM Tris-HCl, pH = 7.4, 50 mM NaCl, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol) and incubated with 15 mL of NiNTA agarose beads (Qiagen) according to the manufacturer's instructions. The beads were washed with 180 mL of NiNTA Buffer

B (20 mM Tris-HCl, pH=7.4, 50 mM NaCl, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol, 30 mM imidazole pH 8); VcDnaB-NHis and was eluted from the beads with 300 mL of NiNTA Buffer C (20 mM Tris-HCl pH=7.4, 50 mM NaCl, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol, 250 mM imidazole, pH = 8). NiNTA-purified VcDnaB-NHis was further purified over Heparin Sepharose Fast Flow as was done with untagged DnaB except that the elution gradient was modified; 4 column volumes of Heparin buffer A (above), followed by a linear gradient from 5% to 100% Heparin buffer B over six column volumes. VcDnaB-NHis fractions were recovered and pooled in the ~80 mM to ~600 mM NaCl positions on the gradient. The heparin pool was passed through a 23 mL SP Sepharose Fast Flow (Cytiva) equilibrated in Q buffer A: 20 mM Tris-HCl, pH=7.4, 100 mM NaCl, 5 mM MgCl₂, 5% glycerol, and 5 mM β -mercaptoethanol; VcDnaB-NHis was captured by a 50mL Q Sepharose Fast Flow column (Cytiva) attached in-line. After a 4-column volume wash with 10% Q buffer B: (20 mM Tris-HCl, pH = 7.4, 1 M NaCl, 5 mM MgCl₂, 5% glycerol, and 5 mM β -mercaptoethanol), a linear gradient from 10% to 100% Q buffer B was used to elute VcDnaB-NHis, which was recovered in the ~250 mM to ~800 mM NaCl positions on the gradient. The pooled VcDnaB-NHis containing fractions were further purified CHT and Superdex 200 Prep grade (Cytiva) as described above. Size exclusion chromatography fractions were pooled, concentrated as described above, to 40 mg/ml, aliquoted (~15-30 μ L aliquots), flash frozen in liquid nitrogen, and maintained at -80 °C until use.

VcDciA. VcDciA-FL-CHis, VcDciA-NTD-CHis, and VcDciA-CTD-CHis were purified using nearly identical protocols. Cells were lysed as above, except that Protease Inhibitor Mini Tablet, EDTA-free (Pierce, ThermoFisher Scientific p/n 32955), was added to the VcDciA-CTD-CHis sample at a ratio of one tablet for every 10 mL of lysate. All three VcDciA constructs were recovered in the soluble fraction. Each fraction was applied to ~10 mL of Ni-NTA agarose beads, incubated

for 40 min, and washed with 80 mL of NiNTA Buffer A (20 mM Tris-HCl, pH = 8, 300 mM NaCl, 5% glycerol, and 5 mM β -mercaptoethanol), supplemented with 15 mM imidazole, pH = 8 (VcDciA-NTD-CHis), or 30 mM imidazole, pH = 8 (VcDciA-FL-CHis and VcDciA-CTD-CHis). VcDciA constructs were eluted from the beads with NiNTA Buffer B (20 mM Tris-HCl, pH = 7.4, 50 mM NaCl, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol, and 250mM imidazole, pH = 8). Further purification was achieved by flowing the various VcDciA constructs through a 50mL Q Sepharose Fast Flow column for capture on a 23mL SP Sepharose Fast Flow column attached in-line; both columns had been equilibrated in SP buffer A: (20mM Tris-HCl, pH = 7.4, 50 mM NaCl, 5% glycerol, and 5 mM β -mercaptoethanol). After separating the two columns, weakly bound proteins were washed away with 4 columns volumes of 5% SP buffer B (20 mM Tris-HCl, pH = 7.4, 1 M NaCl, 5% glycerol, and 5 mM β -mercaptoethanol); this was followed by a linear gradient from 5% to 100% SP buffer B over 3 column volumes. Notably, we observed that VcDciA-FL-CHis eluted at the ~150 mM to ~350 mM NaCl position in the gradient, VcDciA-NTD-CHis between ~120mM and ~700mM NaCl, and VcDciA-CTD-CHis between ~180 mM and ~300 mM NaCl. Pooled fractions were concentrated (VcDciA-FL-CHis by 50% AS precipitation, VcDciA-NTD-CHis by 70% AS precipitation, and VcDciA-CTD-CHis using a 15 mL (2 kiloDalton molecular weight cutoff) spin concentrator (Sartorius Vivaspin). Each construct was then chromatographed over a 320mL Superdex200 Prep Grade column, equilibrated in 20 mM potassium-HEPES, pH = 7.5, 300 mM NaCl, 5% glycerol, and 5 mM β -mercaptoethanol, and processed for -80° storage as described above.

Coexpressed VcDnaB-VcDciA. A soluble fraction derived from coexpressed VcDnaB-untagged - VcDciA-FL-CHis, prepared as described above, was submitted to NiNTA and size exclusion chromatography (320mL Superdex200 Prep Grade). The NiNTA chromatography was performed as above with the

following buffers: NiNTA buffer A (20 mM Tris-HCl, pH = 7.4, 500 mM NaCl, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol, and 30 mM imidazole, pH = 8) and NiNTA buffer B (20 mM Tris-HCl, pH = 7.4, 500 mM NaCl, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol, and 250 mM imidazole, pH = 8). NiNTA-purified protein was concentrated by precipitation with 75% SAS and redissolved in 6 mL of the running buffer (below) for size exclusion chromatography. The size exclusion chromatography was carried out in (20 mM potassium HEPES, pH = 7.5, 500 mM NaCl, 5 mM MgCl₂, 5% glycerol, and 5 mM β -mercaptoethanol).

Ni-NTA pull-down experiments (Vc proteins)

Equal volumes (22.5 mL) of cells at similar densities expressing untagged VcDnaB (full-length, NTD, NTD-LH, or LH-CTD, each in 20mM Tris-HCl pH 7.4, 500mM NaCl, 5mM MgCl₂, 10% sucrose, 5mM β -mercaptoethanol) and his-tagged VcDciA (full-length, Lasso, or KH, each in 20 mM Tris-HCl pH 8, 300 mM NaCl, 5% glycerol, 5 mM β -mercaptoethanol), were mixed, diluted with 15mL of buffer (20mM Tris-HCl pH 8, 300mM NaCl, 5% glycerol 5 mM β -mercaptoethanol) were mixed to a total volume of 60 mL. The mixture was lysed using a Sonic dismembrator (Model 500, Fisher Scientific) at the following settings: 2 min at 70% amplitude, with alternating 30s on and off cycles. Sonication settings were: 30s on, 30s off, 2 min total on time, 70% amplitude. The lysate was centrifuged at 16k RPM in a JA-17 rotor (Beckman-Coulter), and a supernatant was recovered. The clarified lysate was incubated with Ni-NTA Agarose beads for 40 mins, washed with 90mL of NiNTA Buffer A (20 mM Tris-HCl pH 8, 300 mM NaCl, 5% glycerol, 5 mM β -mercaptoethanol, 15 mM Imidazole pH 8). Protein was eluted using 30 mL of NiNTA Buffer B (20 mM Tris-HCl pH 8, 300 mM NaCl, 5% glycerol, 5 mM β -mercaptoethanol, 250 mM Imidazole pH 8).

Native Mass Spectrometry

DnaB and DciA samples for native mass spectrometry (nMS) were exchanged into 500 mM ammonium acetate, pH 7.5, and 0.01% Tween-20 (nMS buffer) using a Zeba micro spin desalting column with a molecular cutoff of 40 kDa (ThermoFisher Scientific). Complexes containing DnaB and DciA in the above buffer were mixed in defined molar ratios before nMS measurements. To take data in the presence of nucleotide, ATP and magnesium acetate were added to final concentrations of 10 μ M and 0.5 mM, respectively. The list of DNA molecules used in the nMS study appears in Supplementary Table 4. Lyophilized synthesized ssDNAs (IDT) oligonucleotides were brought to 200 μ M in HPLC-grade water and diluted to a working concentration of 3 - 6 μ M in the above nMS buffer.

An aliquot (3 – 5 μ L) of an experimental sample was loaded into a gold-coated quartz capillary tip prepared in-house. The sample was then electrosprayed into an Exactive Plus EMR instrument (Thermo Fisher Scientific) using a modified static nanospray source^{82,83}. Spectra were continuously recorded from a single sample-loaded capillary. For analysis of DnaB and DnaB-containing assemblies, the general MS parameters used include spray voltage, 1.2 – 1.3 kV; capillary temperature, 125 – 150 °C; S-lens RF level, 200; resolving power, 8,750 or 17,500 at m/z of 200; AGC target, 1×10^6 ; number of micro scans, 5; maximum injection time, 200 ms; in-source dissociation (ISD), 0 – 10 V; injection flatapole, 8 V; interflatapole, 4 V; bent flatapole, 4 V; high energy collision dissociation (HCD), 200 V; ultrahigh vacuum pressure, $5 - 7 \times 10^{-10}$ mbar; total number of scans, 100. For analysis of full-length DciA, the MS parameters are similar as above except for capillary temperature, 100 °C in-source dissociation (ISD), 100 V; injection flatapole, 8 V; interflatapole, 7 V; bent flatapole, 6 V; higher-energy collisional dissociation (HCD), 0 V; ultrahigh vacuum pressure, 1.4×10^{-10} mbar. Mass calibration in positive EMR mode was performed using a cesium iodide solution. Raw nMS spectra were

visualized using Thermo Xcalibur Qual Browser (version 4.2.47). Data processing and spectra deconvolution was performed using UniDec version 4.2.0^{84,85} with the following general parameters: background subtraction (if applied), subtract curve, 10; smooth charge distribution, enabled; peak shape function, Gaussian; degree of Softmax distribution (beta parameter): 2 – 20. PDBO measured all mass spectrometric data in the laboratory of BTC.

Limited Proteolysis/Mass Spectrometry

Trypsin and Asp-N were selected for limited proteolysis to probe flexible domains and protease-accessible sites across the full-length DciA (Vc and Mt). DciA orthologs and the two proteases were buffer-exchanged into 500 mM ammonium acetate, 0.01% Tween-20, pH 7.5 using a Zeba micro spin desalting column with a molecular cutoff of 40 kDa. Proteolysis was initiated by mixing DciA with trypsin or AspN at protease-to-DciA molar ratios of 1:100 or 1:500. The sample was then loaded into a conductive emitter and electrosprayed into the nMS instrument. nMS data from the limited proteolysis reactions were acquired using the following parameters: spray voltage, 1.2 – 1.3 kV; capillary temperature, 125 °C; S-lens RF level, 200; resolving power, 8,750 at m/z of 200; AGC target, 1×10^6 ; number of microscans, 5; maximum injection time, 200 ms; in-source dissociation (ISD), 200 V; injection flatapole, 8 V; interflatapole, 7 V; bent flatapole, 6 V; higher-energy collisional dissociation (HCD), 0 – 10 V; ultrahigh vacuum pressure, $5 - 7 \times 10^{-10}$ mbar; total number of scans, 100.

The nMS RAW file for each time course was processed using the MetaUniDec software (v 4.4.1)^{84,85}. First, the MS RAW file was converted into a hierarchical data format 5 file (HDF5) using the Automatic Chromatographic Parsing function with auto import by time setting with a time window of every 0.5 min. This setting averages all the MS scans for every 0.5 min within the chromatogram. The

resulting HDF5 was then used for automatic deconvolution to extract mass and intensity information for each species detected. The deconvolution parameters used were m/z range: 500 – 4,000; mass range: 5,000 – 20,000 Da; sample mass every 0.5 Da; peak width FWHM: 0.8 Th; smooth charge state distribution, on; and peak shape function, Gaussian. The resulting time series were then plotted using Microsoft Excel.

SUPPLEMENTARY MOVIE

Supplementary Movie 1. Migration of Subunits from the Bottom to the Top of the DnaB spiral during translocation. This video shows the transition from the ATP γ S conformation of the present VcDnaB-ssDNA complex to the GDP-AlF₄ configuration of the Bst 4ESV DnaB ssDNA¹⁰ complex. The two complexes are aligned by superimposing the ssDNA molecules of each structure, as detailed in the main text. The motions depicted in the video are described in the main text. The DnaB-ssDNA-ATP γ S complex is drawn in the ribbon representation (left) or the sphere/cylinder/ribbon schematic (right). Both representations are colored by chain. The ssDNA is shown in cartoon and stick representation and is colored orange. To calculate the transition between structures, a version of the 4ESV coordinates was constructed that encompassed the *Vibrio cholerae* primary sequence. The two states were then submitted to PyMol to calculate the 90 intermediate structures between the ATP γ S and the GDP-AlF₄ conformers and an additional 90 structures between the GDP-AlF₄ and ATP γ S structures. Frames were rendered in PyMol and assembled in QuickTime player.

Supplementary Table 2. Contacts (<4 Å) in the VcDnaB-ssDNA complex.

Model1	Chain1	Res1	ResN1	Atom1	to	Model2	Chain2	Res2	ResN2	Atom2	Dist
VcB-ssDNA	B	THR	355	OG1	to	VcB-ssDNA	G	DC	9	O2	3.947
VcB-ssDNA	B	ASN	383	OD1	to	VcB-ssDNA	G	DC	11	C5'	3.716
VcB-ssDNA	B	ASN	383	ND2	to	VcB-ssDNA	G	DC	11	O5'	3.771
VcB-ssDNA	B	ASN	383	OD1	to	VcB-ssDNA	G	DC	11	O5'	3.727
VcB-ssDNA	B	ASN	383	CG	to	VcB-ssDNA	G	DC	11	OP1	3.107
VcB-ssDNA	B	ASN	383	ND2	to	VcB-ssDNA	G	DC	11	OP1	2.651
VcB-ssDNA	B	ASN	383	OD1	to	VcB-ssDNA	G	DC	11	OP1	2.814
VcB-ssDNA	B	ASN	383	ND2	to	VcB-ssDNA	G	DC	11	OP2	3.263
VcB-ssDNA	B	ASN	383	CG	to	VcB-ssDNA	G	DC	11	P	3.922
VcB-ssDNA	B	ASN	383	ND2	to	VcB-ssDNA	G	DC	11	P	3.259
VcB-ssDNA	B	ASN	383	OD1	to	VcB-ssDNA	G	DC	11	P	3.784
VcB-ssDNA	B	ASN	383	CA	to	VcB-ssDNA	G	DA	12	OP2	3.423
VcB-ssDNA	B	ASN	383	CB	to	VcB-ssDNA	G	DA	12	OP2	3.404
VcB-ssDNA	B	ASN	383	CG	to	VcB-ssDNA	G	DA	12	OP2	3.126
VcB-ssDNA	B	ASN	383	OD1	to	VcB-ssDNA	G	DA	12	OP2	2.483
VcB-ssDNA	B	ASN	383	OD1	to	VcB-ssDNA	G	DA	12	P	3.825
VcB-ssDNA	B	ARG	400	CA	to	VcB-ssDNA	G	DC	11	OP1	3.351
VcB-ssDNA	B	ARG	400	CD	to	VcB-ssDNA	G	DA	12	OP2	3.746
VcB-ssDNA	B	ARG	400	NH1	to	VcB-ssDNA	G	DA	12	OP2	3.77
VcB-ssDNA	B	GLU	401	CA	to	VcB-ssDNA	G	DA	10	C4'	3.997
VcB-ssDNA	B	GLU	401	CA	to	VcB-ssDNA	G	DA	10	C5'	3.812
VcB-ssDNA	B	GLU	401	CA	to	VcB-ssDNA	G	DA	10	O3'	3.636
VcB-ssDNA	B	GLY	403	CA	to	VcB-ssDNA	G	DA	10	OP1	3.887
VcB-ssDNA	C	THR	355	OG1	to	VcB-ssDNA	G	DA	8	C5'	3.886
VcB-ssDNA	C	ASN	383	ND2	to	VcB-ssDNA	G	DC	9	O5'	3.47
VcB-ssDNA	C	ASN	383	CG	to	VcB-ssDNA	G	DC	9	OP1	3.534
VcB-ssDNA	C	ASN	383	ND2	to	VcB-ssDNA	G	DC	9	OP1	3.136
VcB-ssDNA	C	ASN	383	OD1	to	VcB-ssDNA	G	DC	9	OP1	3.236
VcB-ssDNA	C	ASN	383	ND2	to	VcB-ssDNA	G	DC	9	OP2	2.885
VcB-ssDNA	C	ASN	383	ND2	to	VcB-ssDNA	G	DC	9	P	3.222
VcB-ssDNA	C	ARG	384	NH2	to	VcB-ssDNA	G	DA	10	O3'	3.553
VcB-ssDNA	C	ARG	384	NH2	to	VcB-ssDNA	G	DA	10	O5'	3.969
VcB-ssDNA	C	ARG	384	CB	to	VcB-ssDNA	G	DA	10	OP2	3.621
VcB-ssDNA	C	ARG	384	NH2	to	VcB-ssDNA	G	DC	11	OP1	3.334
VcB-ssDNA	C	ARG	384	NH2	to	VcB-ssDNA	G	DC	11	OP2	3.229
VcB-ssDNA	C	ARG	384	NH2	to	VcB-ssDNA	G	DC	11	P	3.476
VcB-ssDNA	C	ARG	400	NH2	to	VcB-ssDNA	G	DC	9	O3'	3.991
VcB-ssDNA	C	ARG	400	CA	to	VcB-ssDNA	G	DC	9	OP1	3.163
VcB-ssDNA	C	ARG	400	CB	to	VcB-ssDNA	G	DC	9	OP1	3.945
VcB-ssDNA	C	GLU	401	CA	to	VcB-ssDNA	G	DA	8	O3'	3.771
VcB-ssDNA	C	GLU	401	CA	to	VcB-ssDNA	G	DC	9	OP1	3.548
VcB-ssDNA	C	GLY	403	CA	to	VcB-ssDNA	G	DA	8	OP1	3.765
VcB-ssDNA	D	ASN	353	CG	to	VcB-ssDNA	G	DG	5	N2	3.974
VcB-ssDNA	D	ASN	353	OD1	to	VcB-ssDNA	G	DG	5	N2	3.119
VcB-ssDNA	D	ASN	383	ND2	to	VcB-ssDNA	G	DT	7	O5'	3.535
VcB-ssDNA	D	ASN	383	CG	to	VcB-ssDNA	G	DT	7	OP1	3.881
VcB-ssDNA	D	ASN	383	ND2	to	VcB-ssDNA	G	DT	7	OP1	3.317
VcB-ssDNA	D	ASN	383	ND2	to	VcB-ssDNA	G	DT	7	OP2	2.826
VcB-ssDNA	D	ASN	383	OD1	to	VcB-ssDNA	G	DA	8	OP2	3.144
VcB-ssDNA	D	ASN	383	ND2	to	VcB-ssDNA	G	DT	7	P	3.287
VcB-ssDNA	D	ARG	384	NH1	to	VcB-ssDNA	G	DC	9	OP1	3.975
VcB-ssDNA	D	ARG	384	CB	to	VcB-ssDNA	G	DA	8	OP2	3.817
VcB-ssDNA	D	ARG	384	NH1	to	VcB-ssDNA	G	DC	9	OP2	3.673
VcB-ssDNA	D	ARG	400	CA	to	VcB-ssDNA	G	DT	7	OP1	3.31
VcB-ssDNA	D	ARG	400	NH1	to	VcB-ssDNA	G	DA	8	OP1	3.523
VcB-ssDNA	D	ARG	400	NH1	to	VcB-ssDNA	G	DA	8	OP2	3.988
VcB-ssDNA	D	GLU	401	CA	to	VcB-ssDNA	G	DA	6	C5'	3.914
VcB-ssDNA	D	GLU	401	CA	to	VcB-ssDNA	G	DA	6	O3'	3.561
VcB-ssDNA	D	GLU	401	CG	to	VcB-ssDNA	G	DA	6	O3'	3.774
VcB-ssDNA	D	GLU	401	CA	to	VcB-ssDNA	G	DT	7	OP1	3.923
VcB-ssDNA	D	GLY	403	CA	to	VcB-ssDNA	G	DA	6	OP1	3.815
VcB-ssDNA	E	THR	355	OG1	to	VcB-ssDNA	G	DA	4	C5'	3.779
VcB-ssDNA	E	THR	355	OG1	to	VcB-ssDNA	G	DC	3	O2	3.494
VcB-ssDNA	E	THR	355	OG1	to	VcB-ssDNA	G	DA	4	O4'	3.947
VcB-ssDNA	E	ASN	383	OD1	to	VcB-ssDNA	G	DG	5	C3'	3.914
VcB-ssDNA	E	ASN	383	CG	to	VcB-ssDNA	G	DG	5	O5'	3.967
VcB-ssDNA	E	ASN	383	ND2	to	VcB-ssDNA	G	DG	5	O5'	3.376
VcB-ssDNA	E	ASN	383	CG	to	VcB-ssDNA	G	DG	5	OP1	3.957
VcB-ssDNA	E	ASN	383	ND2	to	VcB-ssDNA	G	DG	5	OP1	3.414
VcB-ssDNA	E	ASN	383	CG	to	VcB-ssDNA	G	DG	5	OP2	3.968
VcB-ssDNA	E	ASN	383	ND2	to	VcB-ssDNA	G	DG	5	OP2	2.769
VcB-ssDNA	E	ASN	383	OD1	to	VcB-ssDNA	G	DA	6	OP2	3.377
VcB-ssDNA	E	ASN	383	ND2	to	VcB-ssDNA	G	DG	5	P	3.268
VcB-ssDNA	E	ARG	384	CB	to	VcB-ssDNA	G	DA	6	OP2	3.604

DnaB and DciA: mechanisms of helicase loading and translocation on ssDNA

VcB-ssDNA	E	ARG	384	NH2	to	VcB-ssDNA	G	DT	7	OP2	3.483
VcB-ssDNA	E	ARG	400	CA	to	VcB-ssDNA	G	DG	5	OP1	3.17
VcB-ssDNA	E	ARG	400	CB	to	VcB-ssDNA	G	DG	5	OP1	3.856
VcB-ssDNA	E	GLU	401	CA	to	VcB-ssDNA	G	DA	4	C4'	3.884
VcB-ssDNA	E	GLU	401	CA	to	VcB-ssDNA	G	DA	4	C5'	3.87
VcB-ssDNA	E	GLU	401	CA	to	VcB-ssDNA	G	DA	4	O3'	3.897
VcB-ssDNA	E	GLY	403	CA	to	VcB-ssDNA	G	DA	4	OP1	3.625
VcB-ssDNA	F	THR	355	OG1	to	VcB-ssDNA	G	DC	2	C5'	3.719
VcB-ssDNA	F	THR	355	OG1	to	VcB-ssDNA	G	DT	1	O4'	3.647
VcB-ssDNA	F	ASN	383	CG	to	VcB-ssDNA	G	DC	3	O5'	3.901
VcB-ssDNA	F	ASN	383	ND2	to	VcB-ssDNA	G	DC	3	O5'	3.271
VcB-ssDNA	F	ASN	383	CG	to	VcB-ssDNA	G	DC	3	OP1	3.928
VcB-ssDNA	F	ASN	383	ND2	to	VcB-ssDNA	G	DC	3	OP1	3.269
VcB-ssDNA	F	ASN	383	ND2	to	VcB-ssDNA	G	DC	3	OP2	3.048
VcB-ssDNA	F	ASN	383	OD1	to	VcB-ssDNA	G	DA	4	OP2	3.653
VcB-ssDNA	F	ASN	383	ND2	to	VcB-ssDNA	G	DC	3	P	3.276
VcB-ssDNA	F	ARG	384	NH2	to	VcB-ssDNA	G	DA	4	C3'	3.931
VcB-ssDNA	F	ARG	384	NH2	to	VcB-ssDNA	G	DG	5	OP1	3.545
VcB-ssDNA	F	ARG	384	CB	to	VcB-ssDNA	G	DA	4	OP2	3.528
VcB-ssDNA	F	ARG	384	NH2	to	VcB-ssDNA	G	DG	5	OP2	3.591
VcB-ssDNA	F	ARG	384	NH2	to	VcB-ssDNA	G	DG	5	P	3.861
VcB-ssDNA	F	ARG	400	CA	to	VcB-ssDNA	G	DC	3	OP1	3.216
VcB-ssDNA	F	ARG	400	CB	to	VcB-ssDNA	G	DC	3	OP1	3.778
VcB-ssDNA	F	GLU	401	CA	to	VcB-ssDNA	G	DC	2	C4'	3.918
VcB-ssDNA	F	GLU	401	CG	to	VcB-ssDNA	G	DC	2	C4'	3.794
VcB-ssDNA	F	GLU	401	CA	to	VcB-ssDNA	G	DC	2	C5'	3.855
VcB-ssDNA	F	GLU	401	OE2	to	VcB-ssDNA	G	DC	3	C5'	3.76
VcB-ssDNA	F	GLU	401	CG	to	VcB-ssDNA	G	DC	2	O3'	3.81

Supplementary Table 3. Expression Plasmids Used in This Study

ExpressionPlasmid	Insert	Template	Primers	Plasmid vector
pNG066-pET24a-VcDnaB-unt	VcDnaB-untagged	<i>V. cholerae</i> N16961 genomic DNA	VcDnaB-NdeI-F, VcDnaB-XhoI-R	pET24a
pNG067-pET24a-VcDnaB-NHis	VcDnaB-NHis	pNG066	VcDnaB-pET24a-NHis-F, VcDnaB-pET24a-NHis-R	pET24a
pNG072-pET24a-VcDnaB-E259A-NHis	VcDnaB-E259A-NHis	pNG067	VcDnaB-E259A-F, VcDnaB-E259A-R	pET24a
pNG114-pET24a-VcDnaB-1to170-unt	VcDnaB-NTD-untagged	pNG066	VcDnaB_170-pET24a-F, VcDnaB_170-pET24a-R	pET24a
pNG131-pET24a-VcDnaB-1to208-unt	VcDnaB-NTD_LH-untagged	pNG066	VcDnaB_208-pET24a-F, VcDnaB_208-pET24a-R	pET24a
pNG134-pET24a-VcDnaB-170to468-unt	VcDnaB-LH_CTD-untagged	pNG066	pET24a-VcDnaB_170-F, pET24a-VcDnaB_170-R	pET24a
pNG057-pCDFDuet-1-VcDciA-CHis	VcDciA-CHis	<i>V. cholerae</i> N16961 genomic DNA	VcDciA-NdeI-F, VcDciA-XhoI-CHis-R	pCDFDuet-1
pNG078-pCDFDuet-VcDciA-1to98-CHis	VcDciA-NTD-CHis	pNG057	VcDciA_98-CHis_pCDF-F, VcDciA_98-CHis_pCDF-R	pCDFDuet-1
pNG087-pCDFDuet-VcDciA-98to157-CHis	VcDciA-CTD-CHis	pNG057	VcDciA_1-VcDciA_98-F, VcDciA_1-VcDciA_98-R	pCDFDuet-1
MtDciA_pET22b	MtDciA-FL-CHis	n/a, synthesized	n/a, synthesized	pET22b
MtDciA-NTD_pET22b (1-89)	MtDciA-NTD-CHis	n/a, synthesized	n/a, synthesized	pET22b
MtDciA-CTD_pET22b (90-189)	MtDciA-CTD-CHis	n/a, synthesized	n/a, synthesized	pET22b
MtDnaB-SUMO_pET28a	MtDnaB-untagged	n/a, synthesized	n/a, synthesized	pET28a
MtDnaB-SUMO_pET28a	MtDnaB-SUMO	n/a, synthesized	n/a, c synthesized	pET28a
MtDnaB-SUMO-E257A_pET28a	MtDnaB-E257A-untagged	n/a, synthesized	n/a, synthesized	pET28a
MtDnaB-SUMO-K232A_pET28a	MtDnaB-K232A-untagged	n/a, synthesized	n/a, synthesized	pET28a
MtDnaB-NTD_pCDF-DUET-1 (1-170)	MtDnaB-NTD_pCDF-DUET-1	n/a, synthesized	n/a, synthesized	pCDF-DUET-1
MtDnaB-NTD-LH_pCDF-DUET-1 (1-278)	MtDnaB-NTD-LH-untagged	n/a, synthesized	n/a, synthesized	pCDF-DUET-1
MtDnaB-LH-CTD_pCDF-DUET-1 (171-457)	MtDnaB-LH-CTD-untagged	n/a, synthesized	n/a, synthesized	pCDF-DUET-1
MtDnaB-CTD_pCDF-DUET-1 (279-457)	MtDnaB-CTD-untagged	n/a, synthesized	n/a, synthesized	pCDF-DUET-1
MtDnaB-WT_pET16b	MtDnaB-WT	n/a, synthesized	n/a, synthesized	pET16b

Supplementary Table 4. Primers and other DNAs Used in This Study

DNA Name	DNA Sequence (5' to 3')
VcDnaB-NdeI-F	GGGAGAGCATATGGCAGATCCACGTACTGATA
VcDnaB-XhoI-R	GGGAGAGCTCGAGTTACTCATCGTCAAATGCAGGT
VcDnaB-pET24a-NHis-F	CTTTAAGAAGGAGATATACATATGCATCACCATCACCATCACGCAGATCCACG TACTGATAATCG
VcDnaB-pET24a-NHis-R	CGATTATCAGTACGTGGATCTGCGTGATGGTGATGGTGATGCATATGTATATCT CCTTCTTAAAG
VcDnaB-E259A-F	GATCTTCTCGCTGGCGATGCCCGCCGAAC
VcDnaB-E259A-R	GTTCGGCGGGCATCGCCAGCGAGAAGATC
VcDnaB_170-pET24a-F	GCCGAAGCGCGTACCTAACTCGAGCACCAC
VcDnaB_170-pET24a-R	GTGGTGCTCGAGTTAGGTACGCGCTTCGGC
VcDnaB_208-pET24a-F	GAGTGAACACGGGCTTTAACTCGAGCACCACC
VcDnaB_208-pET24a-R	GGTGGTGCTCGAGTTAAAGCCCGTGTCTACTC
pET24a-VcDnaB_170-F	TAAGAAGGAGATATACATATGACCAGTGAAAATGAAGGCCCG
pET24a-VcDnaB_170-R	CGGGCCTTCATTTTCACTGGTCATATGTATATCTCCTTCTTA
VcDciA-NdeI-F	GGGAGAGCATATGCGTGATCACCGCCCTACC
VcDciA-XhoI-CHis-R	GGGAGAGCTCGAGTCAGTGATGGTGATGGTGATGATCGTCTTTCTGATTTGCT T
VcDciA_1-VcDciA_98-F	TAAGAAGGAGATATACATATGCCTGAGTTGTATCGGCAAAGC
VcDciA_1-VcDciA_98-R	GCTTTGCCGATACAACCTCAGGCATATGTATATCTCCTTCTTA
VcDciA_98-CHis_pCDF-F	GAAGTGCGAGTCAATCCTCATCACCATCACCATCAC
VcDciA_98-CHis_pCDF-R	GTGATGGTGATGGTGATGAGGATTGACTCGCACTTC
FRET-DNA-fork-top-strand	/5Cy3/TACGTAACGAGCCTGCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
FRET-DNA-fork-bottom-strand	TTTTTTTTTTTTTTTTTTTTTTTTTTTGCAGGCTCGTTACGTA/3BHQ_2/
FRET-DNA-fork-capture-oligo	GCAGGCTCGTTACGTA
895	TCCTTGTTATCCACAGGGCAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTGTGCGATATAGTTCTCCGA
894C	/5Cy3/TCGGAGAACTATATCGCACAATTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTCTGCCCTGTGGATAACAAGGA
895B:	TCCTTGTTATCCACAGGGCAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTGTGCGATATAGTTCTCCGA/3IAbRQSp/
894-C-B	/5Cy3/TCGGAGAACTATATCGCACAATTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTCTGCCCTGTGGATAACAAGGA/3IAbRQSp/
895-B-C	/5Cy3/TCCTTGTTATCCACAGGGCAGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TTTTTTGTGCGATATAGTTCTCCGA/3IAbRQSp/
O1b-VcF	GTC ATT AAA TAT ATA TAA AGA TCT ATA TAG AGA TCT TTT TAT TAG ATC TAC TAT TAA GGA

Supplementary Table 5. Native mass spectrometry mass measurements for predominant DnaB assemblies from three bacterial species with and without Mg-ATP.

Sample	Main Complex	+x MgADP	Mass/Da			Mass Error (%) ^b
			Measured	Expected ^a	D	
VcDnaB only	VcDnaB ₆	0	311,517	311,502	15	0.005
VcDnaB + Mg-ATP	VcDnaB ₆	0	311,612	311,502	110	0.04
		1	312,039	311,953	86	0.03
		2	312,476	312,405	71	0.02
		3	312,914	312,856	58	0.02
		4	313,354	313,307	47	0.01
		5	313,786	313,759	27	0.01
		6	314,274	314,210	64	0.02
EcDnaB only	EcDnaB ₆	0	313,714	313,553	161	0.05
EcDnaB + Mg-ATP	EcDnaB ₆	0	313,738	313,553	185	0.06
		1	314,183	314,004	179	0.06
		2	314,622	314,456	166	0.05
		3	315,070	314,907	163	0.05
		4	315,526	315,358	168	0.05
		5	316,055	315,810	245	0.08
		6	316,508	316,261	247	0.08
BsDnaC only ^c	BsDnaC ₁	0	51,398	51,399	-1	(0.002)
BsDnaC + Mg-ATP ^c	BsDnaC ₁	0	51,396	51,399	-3	(0.006)

Chapter 4

UNPUBLISHED RESULTS PART 1

Unveiling unprecedented binding modes of DnaA and DciA to the *Vibrio cholerae* DnaB using single-particle cryo-EM

Daniele Mazzeletti¹, Nicholas Gao^{o2}, Riccardo Miggiano^{#1}, and David Jeruzalmi^{#2}

¹*Department of Pharmaceutical Sciences, University of Piemonte Orientale, Novara, 28100, Italy*

²*Department of Chemistry and Biochemistry, City College of New York, New York, NY 10031, USA.*

^o*Present address: Mirimus Inc, 760 Parkside Ave, Brooklyn, NY 11226, USA*

[#]*Corresponding authors: dj@ccny.cuny.edu and riccardo.miggiano@uniupo.it*

Unpublished

This present work comes as part of the experimental framework set at the beginning of and produced during the 1-year long period of time spent in the laboratory of Professor David Jeruzalmi at The City College of New York (USA) as Visiting PhD student. As of this manuscript's drafting date, biophysics tests are being outlined, to further strengthen the structural biology detail discussed below. These data are part of an upcoming publication.

INTRODUCTION

In all domains of life, the initiation phase of DNA replication is a biological crucial step, yet not thoroughly understood. In the bacterial world, replicative helicases DnaB are recruited at the origin of replication to unwind genomic double-stranded DNA (dsDNA) and expose single strands, then used as template, while translocating ahead of the replication machinery in an ATP-dependent fashion¹. Loading replicative helicases onto single-stranded DNA (ssDNA) requires specialized factors found active throughout the initiation phase of the replication, namely the initiator protein DnaA² and the helicase loader DciA³. However, the sequence of events and the concerted interaction occurring among these players are still to be completely unfolded.

The bacterial origin of replication *oriC* mainly features two distinct functional portions: 1) conserved DNA repeats targeted by the initiator protein DnaA, therefore named DnaA-boxes; 2) the AT-rich region, known as DNA unwinding element or DUE. As the toroidal replicative helicase DnaB must be allocated onto ssDNA, in the first place DnaA binds to and assembles as pentamers on the DnaA-boxes, introducing a DNA distortion and facilitating the melting of dsDNA composing the DUE, rich in less stable A-T base pairs. All this leads to the exposure of single strands and the formation of the DNA fork, onto which DnaB can be loaded⁴ leveraging on three different strategies: ring-opening, ring forming and ring closing⁵ (see detailed description in Chapter 3). Secondly, DnaA seems to have a direct role in interacting and, hence, recruiting the helicase in close proximity to the DNA fork. In order to perform said activities, the genomic characterization of the initiator protein identified for domains with peculiar activities (Figure 1A):

1. Domain I: DnaB helicase interaction domain.
2. Domain II: linker.

3. Domain III: AAA⁺ ATPase site.
4. Domain IV: DNA binding domain.

The AlphaFold calculated structure of VcDnaA (Figure 1B) shows that smaller globular N-terminal Domain I is connected to the bigger C-terminal Domain, composed of Domains III and IV and bearing the ATPase active site, through a long, flexible and intrinsically disordered linker region, namely Domain II.

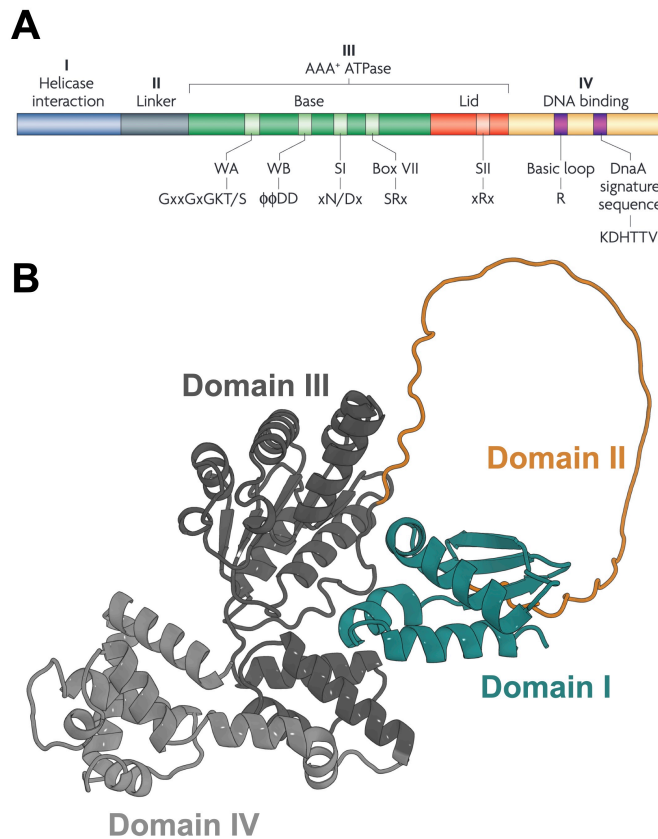


Figure 1. The initiator protein is composed of four distinct subdomains possessing specific biological roles.

A) Schematic representation of DnaA's domains. **B)** AlphaFold predicted structure of VcDnaA, shown as cartoon. The *Vibrio cholerae* ortholog was selected as representative of its group for describing subdomains arrangement. Domain I is colored in teal, Domain II in orange, Domain III in dark grey and Domain IV in light grey.

Overall, DnaA Domain I displays structural features common to the so-termed KH fold⁶. Mutagenesis analyses and studies using different truncated construct of DnaA and DnaB belonging to both the *M. tuberculosis* (MTB) and *E. coli* (*Ec*) systems

aimed to identify the key interaction sites letting the initiator protein liaising with the replicative helicase. Indeed, the binding interface was placed at the level of DnaB's amino acids 154:206, acting as counterpart to DnaA Domain I hydrophobic residues, previously identified via NMR to specifically bind to DnaB's Linker Helix element⁷. Moreover, while in the *Mt* system the presence of the helicase loader DciA positively affect the interaction between DnaA and DnaB³, on the other side EcDnaC seems to diminish the affinity between them, probably inducing structural changes in DnaA Domain I docking site on the replicative helicase⁸. Thus, a wider understanding of the DnaA-DnaB-DciA binding modes is required.

As already mentioned in Chapter 3 and widely discussed in the upcoming Chapter "SAXS and NMR-based Structural Investigation of the Mycobacterial Helicase Loader DciA", mounting DnaB onto the exposed ssDNA enforces replication initiation factors named helicase loaders. Most human pathogenic bacteria, including *Vibrio cholerae* (Vc) and MTB possess the ancestral replicative helicase loader DciA, which displays key features of peculiar helicase-associated proteins like *E. coli* DnaC, *B. subtilis* DnaI and bacteriophage λ P, operating during the initiation phase of the DNA replication⁹. In the last 10 years, efforts have been made for biologically and biochemically characterize DciA proteins. Indeed, since its identification in 2016 by Brézellec *et al.*, the ancestral helicase loader resulted being a fundamental member of the DNA replication initiation machinery even in the mycobacterial world³. DciA is essential for bacterial cells viability and its depletion is reported to stop the physiological cell cycle, lowering the DNA synthesis rate. Moreover, owing to our integrative biophysics experimental framework based on both the Vc and MTB, we proved how DciA strongly interact with the replicative helicase (Chapter 3, Figure 5), leading to the suppression of its ATPase activity (Chapter 3, Figure 6), thus controlling DnaB's activation in the early stages of the replication initiation. Besides, our FRET-based platform detected the

DciA-dependent DNA ‘bubble’ unwinding operated by DnaB (Chapter 3, Figure 7), finally suggesting that DciA loads DnaB onto ssDNA using the ring-opening mechanism.

However, what remains to be experimentally unveiled is the three-dimensional structure of DciA proteins. As of today, multicomponent machine-learning-based Artificial Intelligence systems like AlphaFold¹⁰ and RoseTTAFold¹¹, able to predict proteins’ tertiary structure, have been used to dig into the structural features and diversity of helicase loaders¹². In general, DciA features a globular domain flanked by extensions that can be either N-terminal or C-terminal or even both, named Lassos.

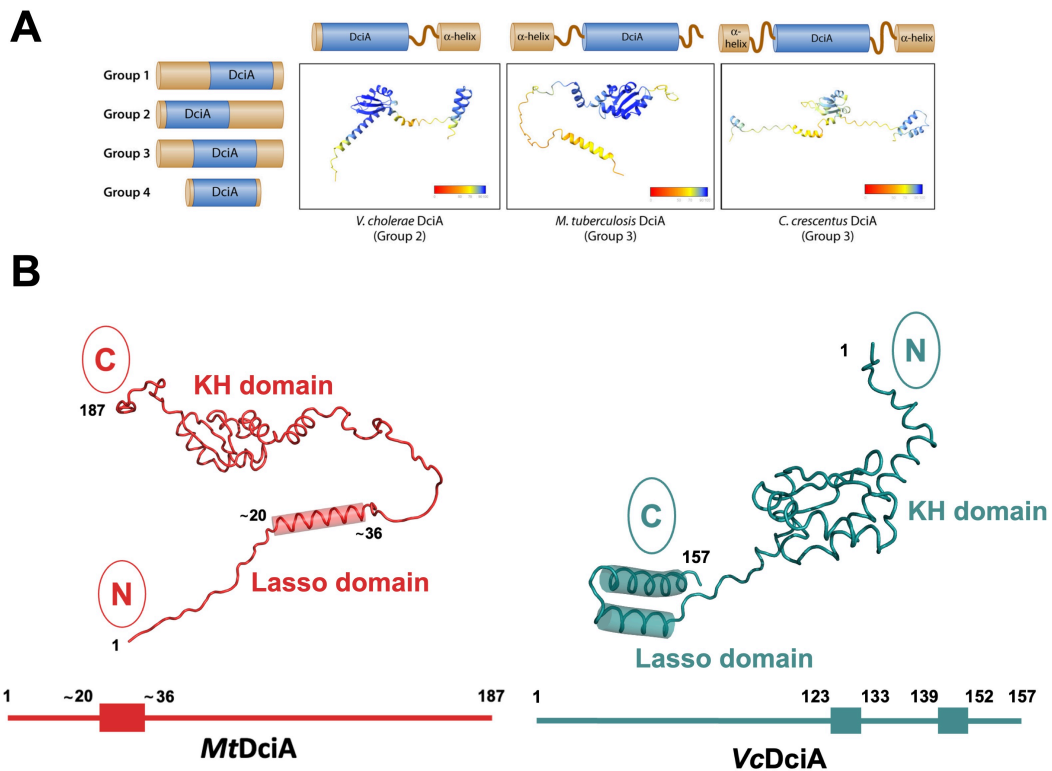


Figure 2. Structure-based convergent evolution of bacterial helicase loaders. The AlphaFold predicted models of helicase loaders can be classified into four major groups (A). As shown, α-helical extensions can be either at the N-terminus (*M. tuberculosis*) or at the C-terminus (*V. cholerae*) or at both (*C. crescentus*). Panel B depicts the inverted domain organization of Mt and Vc loaders, considering the amino acid chain polarity.

In particular, they can be partially or totally intrinsically disordered (Figure 2A). Taking into account Vc and MTB loaders (Figure 2B), notwithstanding the inverted amino acid chain polarity, causing a reversed domain organization, the structural features are retained, giving rise to the convergent evolution idea of bacterial loaders¹³. Besides, also the compact globular domain of DciA proteins features key secondary structure elements typical of the KH fold and, owing to the structural homology with DnaA Domain I, this region has been termed DnaA NTD-Like domain or DANL. Indeed, as represented in Figure 3, the superposition of the two *Vibrio cholerae* KH domains clarifies how the tertiary folding is maintained, providing a Root Mean Square Deviation (RMSD) value of 0.267 Å², while the sequence identity resulted to be irrelevant.

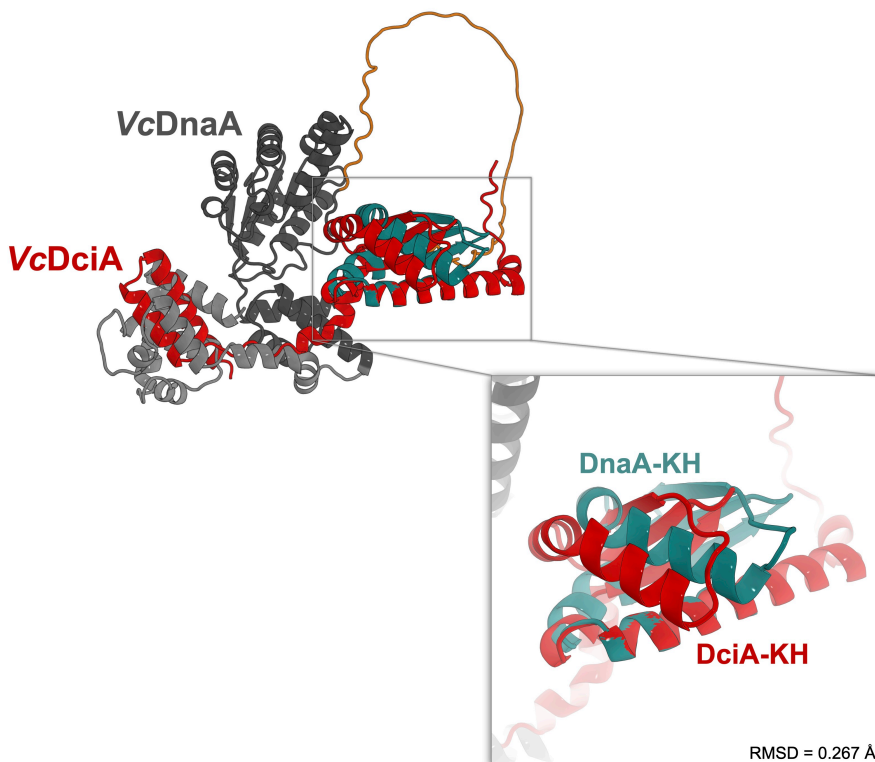


Figure 3. DnaA and DciA have in common a domain that shares structural homology to the KH fold. The globular domain of DciA is named after DnaA Domain I based on conserved structural features reminiscent of the KH fold. Here, AlphaFold calculated structure of the full-length Vc loader, colored in red, was superpose onto DnaA Domain I (teal), calculating a RMSD value of 0.267 Å. Despite the high structural homology, the sequence identity is not significant.

All in all, the extremely high structural similarity between DnaA's and DciA's KH domains and their coexistence during the early stages of DNA replication, in the face of an unclear role, give rise to the question of the biological implications and relevance of this homology. By leveraging on single-particle cryo-electron microscopy, here we present preliminary data on two multiprotein complexes bearing VcDnaA, DnaB and DciA in a ssDNA-free and ssDNA-bound forms, aiming to unveil unprecedented binding modes of both the initiator protein and the loader to the replicative helicase.

RESULTS AND DISCUSSION

With the goal of comprehending how the high structural homology between DnaA and DciA is reflected on their binding mode and competition when interacting with the replicative helicase DnaB, we leveraged on single particle cryo-electron microscopy exploiting *Vibrio cholerae* recombinant proteins, purified at high yield and degree of homogeneity. In particular, our investigation turned to a dual approach, by assembling both a ssDNA-free and a ssDNA-bound state of the DnaA-KH•DnaB•DciA protein complex. As described in the Material and Methods section, the sample and cryo-EM grid preparation of both macromolecular complexes was performed at the same time using identical parameters, except for the inclusion of a single-stranded DNA 45-mer exclusively in one of them.

After cryo-EM data collection, the obtained micrographs were processed independently. The analysis of the DNA-free DnaA-KH•DnaB•DciA complex (Supplementary Figures 1 and 2, Supplementary Table 1) produced two density maps of high value, based on the symmetry imposed during the last non-uniform refinement session. Indeed, deepEMhancer¹⁴ sharpened Map T at 3.87 Å was obtained setting C1 symmetry, while Map W at 3.63 Å was calculated through C3

symmetry (Figure 4). Both were used for reconstructing a three-dimensional model of the replicative helicase that turned out to be surrounded by additional EM density corresponding to six KH folds and six β -hairpin-like domains, to be thoroughly addressed.

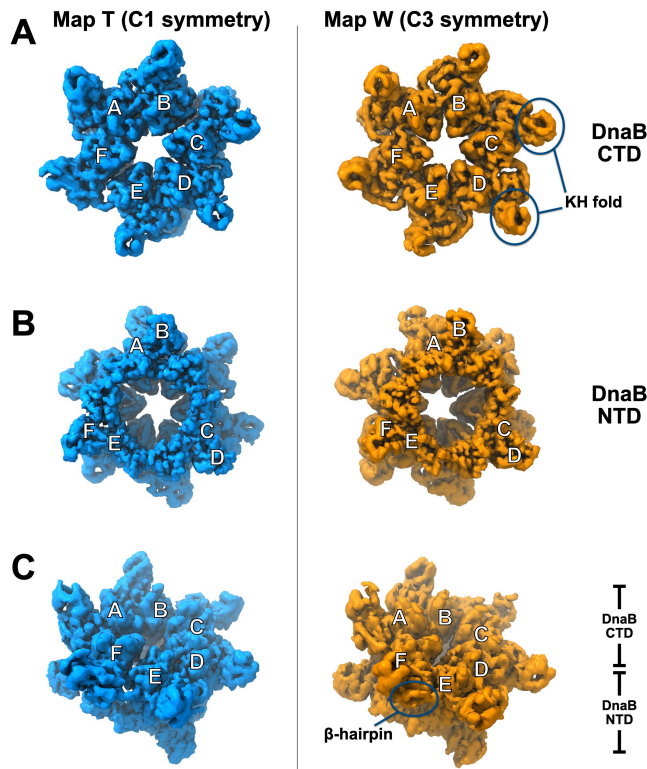


Figure 4. Surface representation of cryo-EM density maps relative to the DNA-free DnaA-KH•DnaB•DciA complex. DeepEMhancer sharpened map T (in blue) was obtained imposing C1 symmetry in the last non-uniform refinement trial; map W (in orange) was obtained setting C3 symmetry. **A)** DnaB is oriented to show the CTD tier; the blue circles highlight the density map of two out of six KH folds identified. **B)** DnaB is oriented to show the NTD tier. **C)** Overall view of the Vc protein complex. One β -hairpin-like domain out of the six detected is encircled in blue. White letters indicate DnaB monomers localization.

DnaB appears in the canonical hexameric assembly, where the N-terminal Domains (NTD) tier is a trimer-of-dimers and the C-terminal Domains (CTD) tier presents the usual pseudo-sixfold symmetry. The CTD layer (Figure 4A) is in the constricted form while the NTD one (Figure 4B) is in its dilated state. Overall, DnaB turned out to be

a flat closed ring (Figure 4C) and, as expected, no nucleic acid was found in the internal chamber as well as nucleotides in the ATP binding sites. Moreover, the careful inspection of the two maps, displaying same features at a different quality and resolution level, found globular densities adjacent to each DnaB Docking Helix (DH), as shown in Figure 5A. And so, we distinguished five DciA-KHs and one DnaA-KH, both in a software-based way and based on the primary sequence differences among the two domains, leveraging also on a manual inspection (Figure 5B).

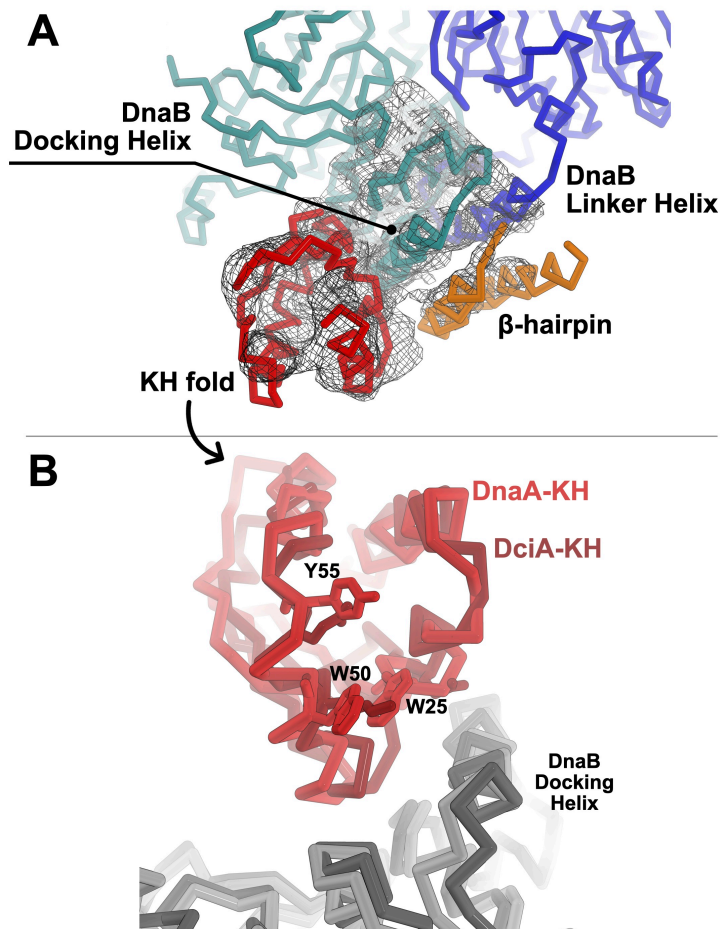


Figure 5. Structural features of the DNA-free DnaA-KH•DnaB•DciA complex. **A)** Focus on the DnaB interaction sites, highlighting the presence of a KH fold, in red, and a β -hairpin-like domain, in orange, in front of the DH helix and the DH-LH element, respectively. Only one out of six interfaces are displayed, composed of a DnaB dimer (teal to blue ribbon). The density map is shown as black mesh. **B)** KH folds belong either to DnaA (light red) or DciA (dark red). VcDnaB chains E and A (resid 200:500) were superposed to allow the neighboring KH domains to align (RMSD = 1.744 Å). Aromatic residues W25, W50 and Y55 are key to distinguish DnaA-KH to DciA-KH. In the present structure, one of six KH folds densities was justified through the docking of DnaA-KH.

Indeed, Trp25 and Trp50, involved in a π -stacking interaction, together with Tyr55 are aromatic residues characterizing DnaA's KH domain, which are replaced by lysines and glutamates in DciA-KH. However, the average resolution of the produced density maps, where amino acids secondary chains are invisible, leaves a certain degree of uncertainty in distinguishing KH domains from one another. Clashes between DnaA-KH and DnaB's associated monomer were considered also as a discriminatory factor, especially when DciA-KH's fitting displayed a higher confidence. On the other side, moving from C1 to C3 symmetry highlighted the existence of β -hairpin-like densities localized in front of DnaB's DH-LH element, at the inter-monomer interface (Figure 5A). Therefore, we successfully modelled six DciA's Lasso domains, specifically comprising residues 121:157. Unluckily, our maps failed in displaying any density corresponding to the full-length helicase loader, thus DciA's KHs and Lassos have not been connected (overall 3D model shown in Supplementary Figure 5).

The same level of detail was utilized for the cryo-EM data processing and analysis of the DNA-bound DnaA-KH•DnaB•DciA complex (Supplementary Figures 3 and 4, Supplementary Table 2). This calculated a 3.65 Å resolution density map that was further sharpened with DeepEMhancer¹⁴, producing volume M as reported in the cryoSPARC workflow below, which represented our best result, later used for three-dimensional model building (Figure 6). The resulting structure shows, once again, VcDnaB in its homohexameric assembly, featuring a constricted CTD tier (Figure 6A) and a dilated NTD one (Figure 6B), respectively in the canonical pseudo-sixfold symmetry and trimer-of-dimers configuration. By taking the advantage of the knowledge acquired while determining the structure of the DNA-free Vc protein complex, the examination of the EM volume map M quickly brought our attention to the identification of six KH folds, that were justified by modelling six DnaA-KH monomers, considering the previously discussed structural differences

with DciA-KH. Also, three density blobs reminiscent of DciA's Lasso β -hairpin-like shape were noticed, although noisy EM density can be visualized even in the three remaining binding sites. The model building procedure detailed in the Methods section yielded the 3D structure shown in Figure 7A.

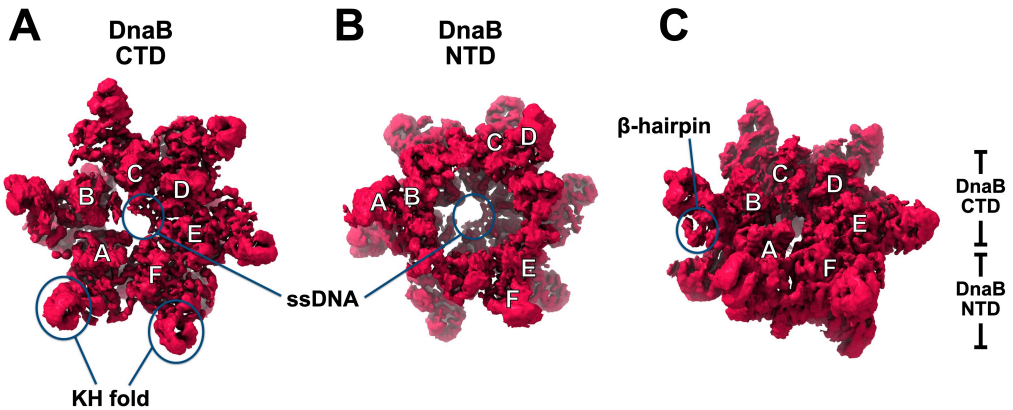


Figure 6. Surface representation of DNA-bound DnaA-KH•DnaB•DciA complex cryo-EM density map. DeepEMhancer sharpened map M in red was obtained following the procedure described in the method section. **A)** DnaB CTD tier as top view; the blue circles highlight the density map of two KH folds and ssDNA. **B)** Focus on DnaB NTD tier. **C)** Overall view of the DNA-bound Vc protein complex. One β -hairpin-like density is encircled in blue. White letters indicate DnaB monomers localization.

As previously discussed, DnaB's ATPase site is formed at the inter-monomer interface, hence six catalytic domains are present per hexamer. Four ADP molecules are found in our ssDNA-bound Vc DNA replication protein complex, specifically within sites shaped by chains C-D, D-E, E-F and F-A, even though the sample preparation did not include it. Here, Arg439 belonging to the β -hairpin arginine finger of one subunit (Walker B motif) and P-loop Lys234 coming from the neighboring monomer (Walker A motif) closely approach to the nucleotide. Besides, two free catalytic sites are found, and the Arg finger results being displaced respectively by ~ 7.6 Å and ~ 11.6 Å in catalytic centers A-B and B-C from the average position assumed in ADP-bound active sites (Figure 7B).

Based on the superpositions pattern applied to the comparison between our ATP γ S-bound VcDnaB-ssDNA and the GDP•AlF₄ complex-bound BstDnaB (PDB: 4ESV¹⁵) described in Chapter 3, we detected two different positions for LH elements also in the present EM structure. In particular, chains C, D, E and F occupy position termed LH1, while chains A and B sits in a rearranged location that, however, turned out to be in between LH2 and LH3 positions encountered in our VcDnaB-ssDNA-ATP γ S (Figure 7C).

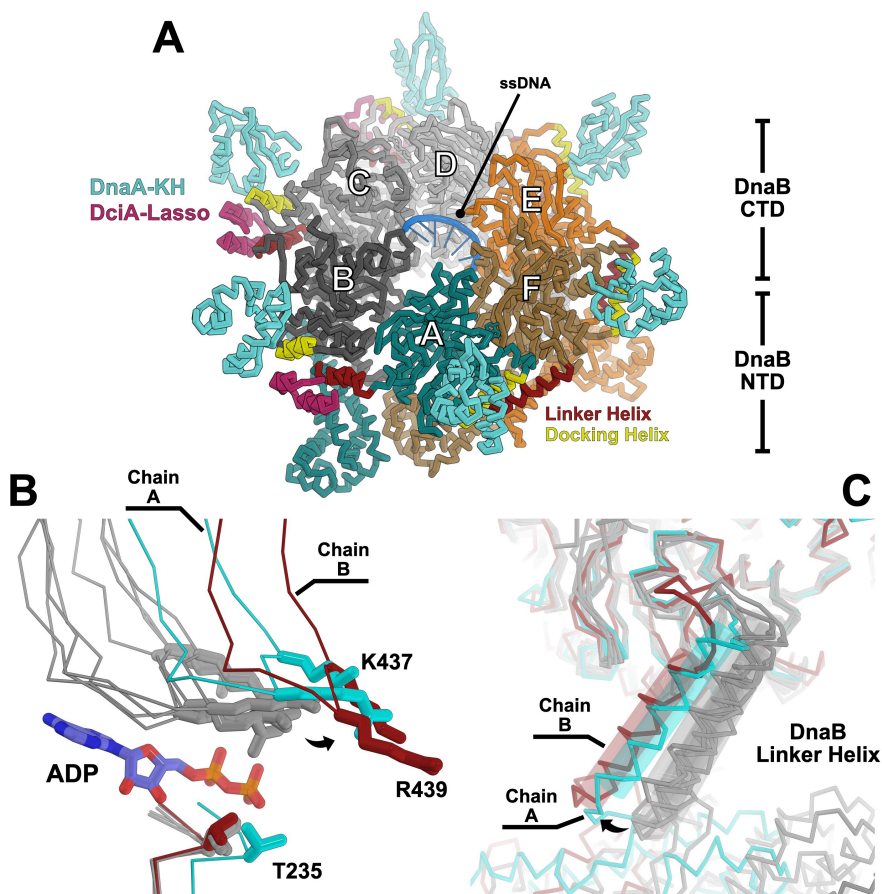


Figure 7. Cryo-EM structure of ssDNA-bound *Vibrio cholerae* DnaA-KH•DnaB•DciA complex and insights on the nucleotide binding site and Linker Helix arrangement. **A)** Vc complex is shown as ribbon; DnaB monomers are represented with white letters; Linker Helix (LH) and Docking Helix (DH) elements are colored in red and yellow, respectively; six DnaA-KH units are shown in cyan; three DciA-Lasso domains are represented in pink; ssDNA made of six nucleobases is displayed as cartoon and colored in blue. **B)** The superposition of six CTD pairs (A-B, B-C, etc., resid 200:468) showcased the displacement of catalytic center's Arg-finger in chain A (cyan) and B (red), compared to all other DnaB chains (grey). **C)** The superposition of each DnaB CTD (resid 200:468) allowed also to appreciate the rearrangement of LH element in chains A (cyan) and B (red).

As seen in our VcDnaB-ssDNA model (Chapter 3), conformational changes within the Walker A-Walker B catalytic site partnered with the adjustment of the Linker Helix from LH1 position to a more distal one in discreet DnaB units, namely chains A and B, are a symptom of an overall adaptation of the hexameric assembly, implying a closed spiral CTD tier, which also displays a breach between said monomers. As proof, Pro442 packing against ADP's adenine ring in chain C, D, E and F is found to be shifted by ~ 7.3 Å towards DnaB's inner chamber in chain A, as the whole Arginine finger exhibits a different configuration, leaving room for a larger gap between monomers A and B.

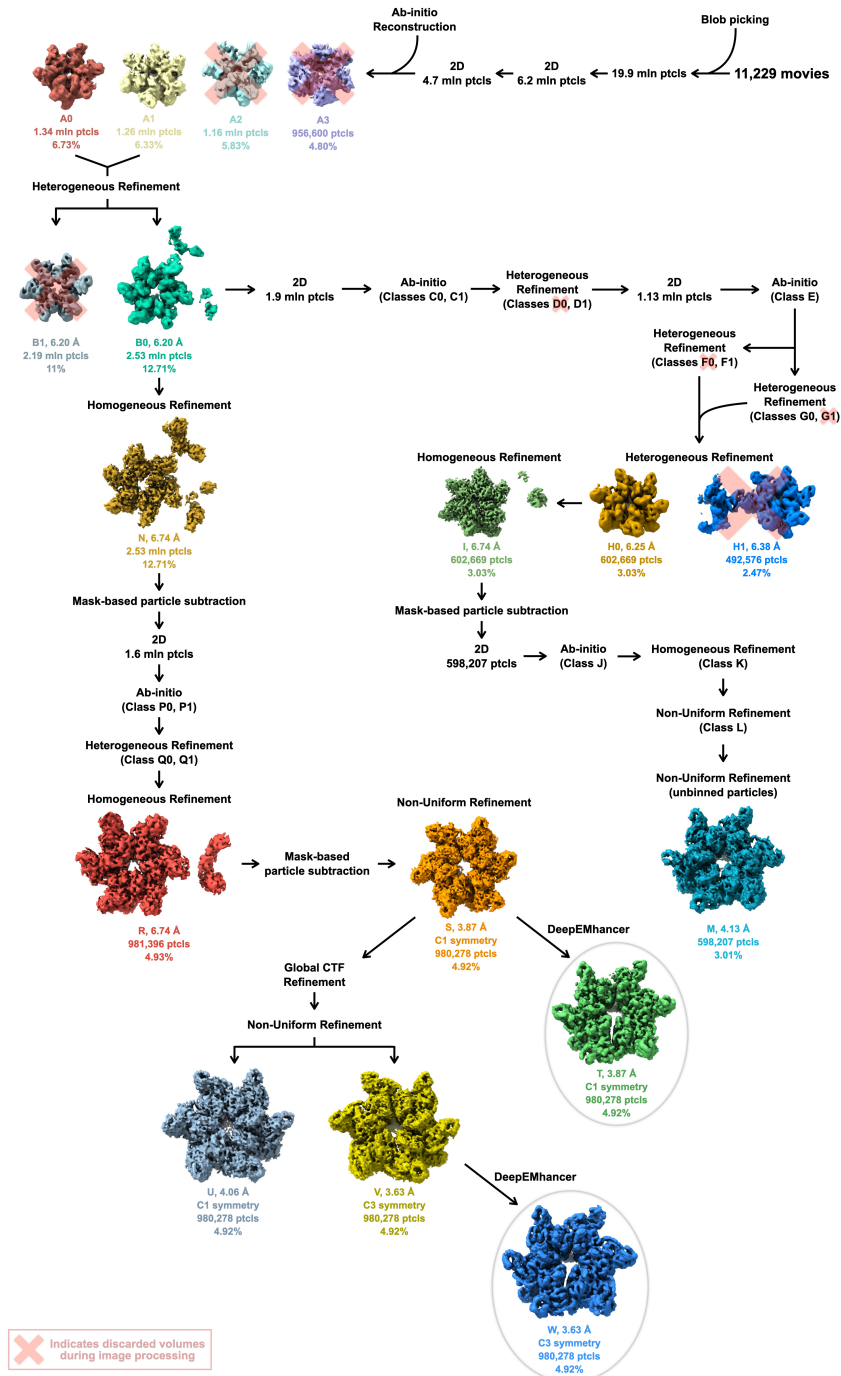
The six DnaB monomers encircle a chamber that allows the DNA to be hosted during the unwinding and translocation mechanism, when a single-strand is threaded in and the other is excluded. In our structure, we detected six nucleobases ascribable to the ssDNA poly-dT 45-mer included in the sample preparation. The local EM density was very poor and, therefore, thymine rings cannot be made out. However, the correct modelling of DNA was driven by the phosphate backbone density map, tracing a spiral path withing DnaB's CTD layer. Main contacts are registered between the nucleobases' phosphate groups and DnaB chains E, D and C, from ssDNA 5' to 3' end.

At this stage of 3D model building, analysis and interpretation, our efforts seem to have captured two different stages of the replicative helicase loading mechanism onto ssDNA in the *Vibrio cholerae* system, involving both the isolated KH domain of the initiator protein DnaA and the full-length ancestral helicase loader DciA. In the *in-vitro* reconstitution of a DNA-free specimen and a DNA-containing one, VcDnaA-KH and VcDciA were included in large and equimolar excess than VcDnaB, in order to let the 3-protein system assemble in the most physiological way, without strict conditions. Hence, two different protein

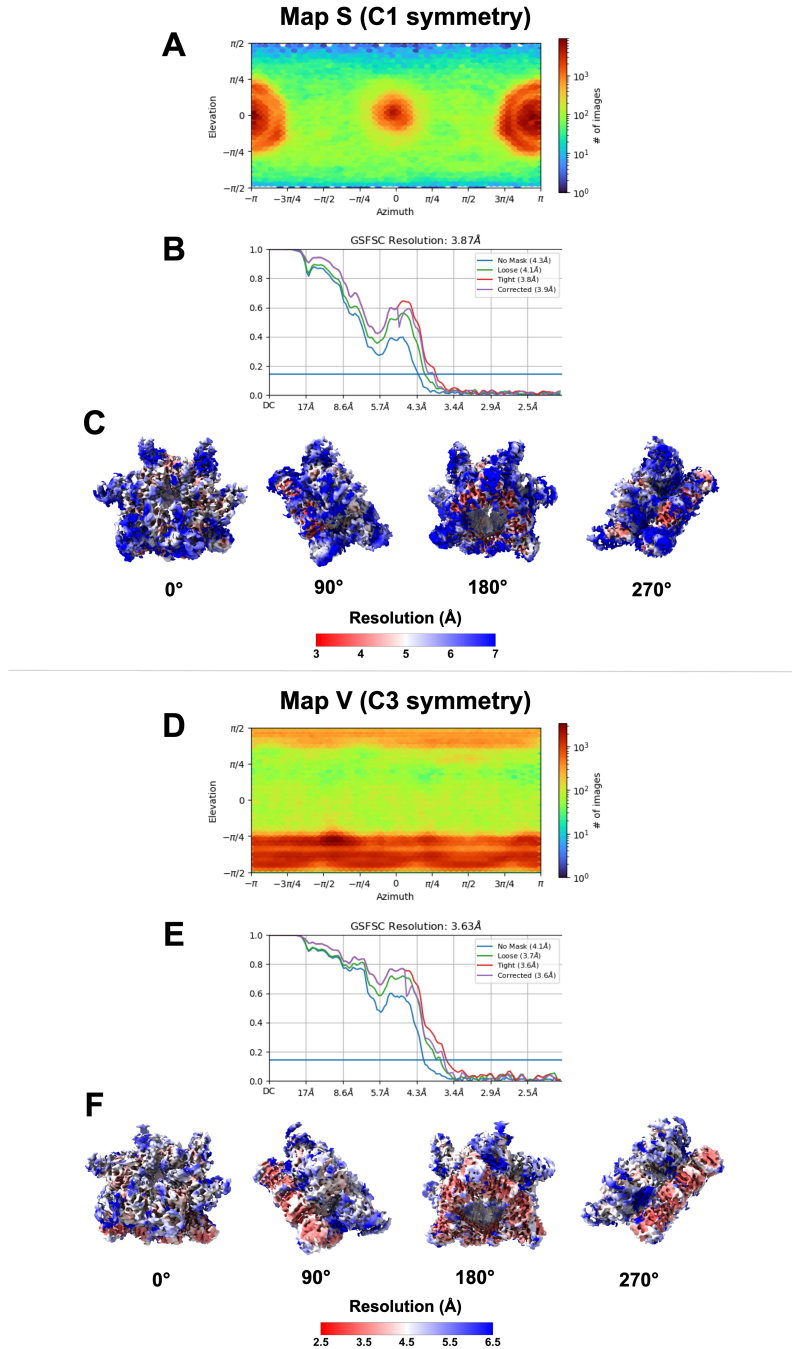
complexes have been capture using single-particle cryo-EM. The ssDNA-free sample provided a closed planar hexameric DnaB where six KH folds could be made out, five most likely being DciA-KH and one being DnaA-KH, together with six DciA-Lassos. On the other hand, adding ssDNA to the same protein system induced a perturbation so that a partially tilted yet close DnaB hexameric ring bound to four ADPs was found to encircle a six-nucleobases DNA strand, and to interact with six KH folds at each Docking Helix site, recognized as DnaA-KH thanks to big aromatic residues Trp25, Trp50, and Tyr55 (Supplementary Figure 6). At the same time, at least three β -hairpin-like DciA-Lasso domains were identified. Therefore, this boils down to the detection of a DNA-dependent protein network, in which the three main factors involved in the first stages of DNA replication initiation gather in different ratio based on the nucleic acid presence, and our structures most likely populate different transition states that allow the helicase to be correctly placed on the replication fork, finally assuming the structural configuration discussed in Chapter 3, switching from *Bst*DnaB (PDB: 4ESV¹⁵) and of our ATP γ S-bound *Vc*DnaB, permitting translocation.

As of today, given the unpublished nature of this Chapter, cryo-EM structure modelling, interpretation and comparison with already published DnaB structure is still ongoing. We aim to consolidate our preliminary ideas by means of native Mass Spectrometry, Cross-Linking Mass Spectrometry, Surface Plasmon Resonance and FRET-based DNA unwinding assays, with a view to integrating more and more the DNA replication initiation model shown and described in Chapter 3, Figure 8.

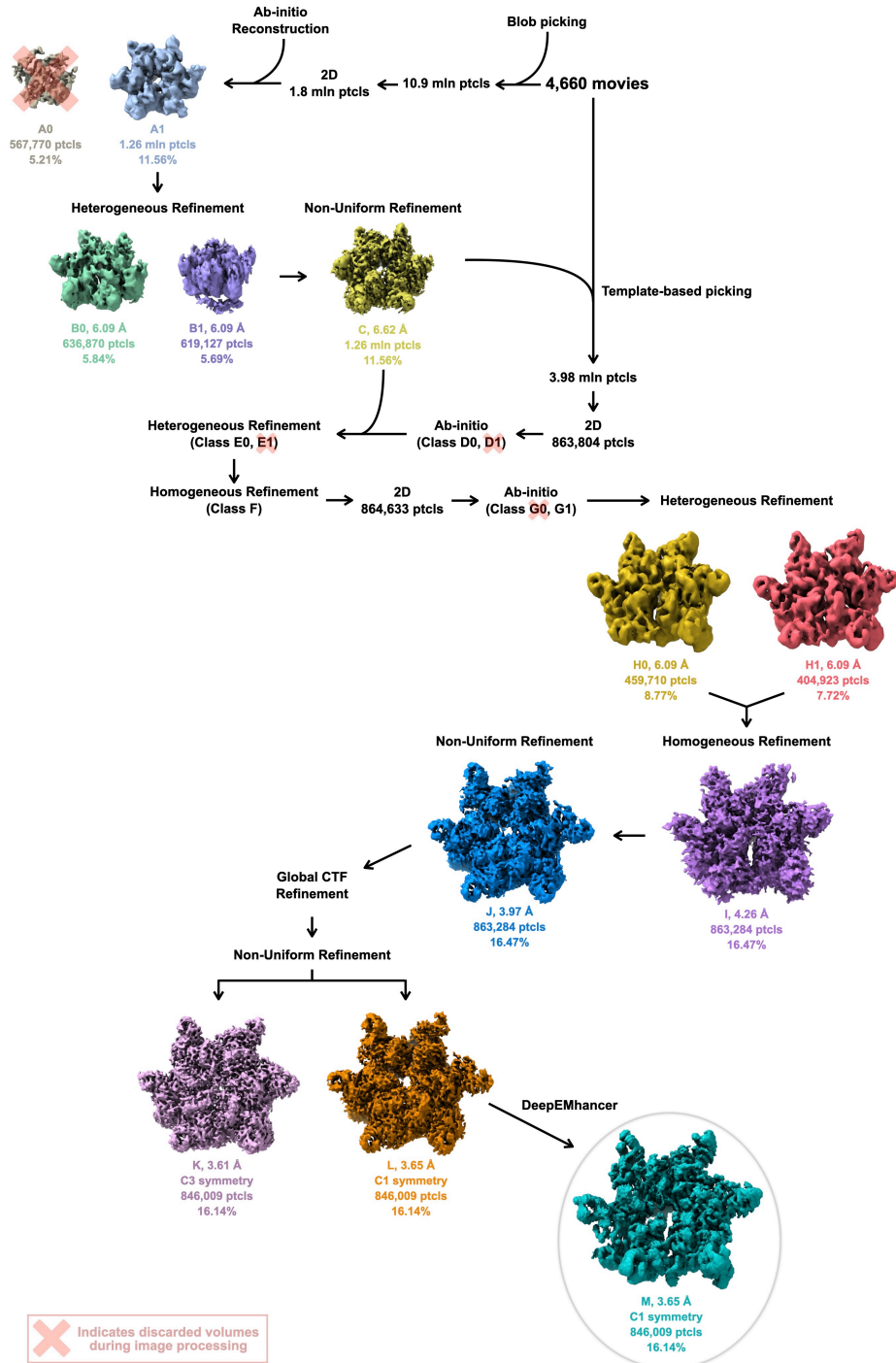
SUPPLEMENTARY FIGURES



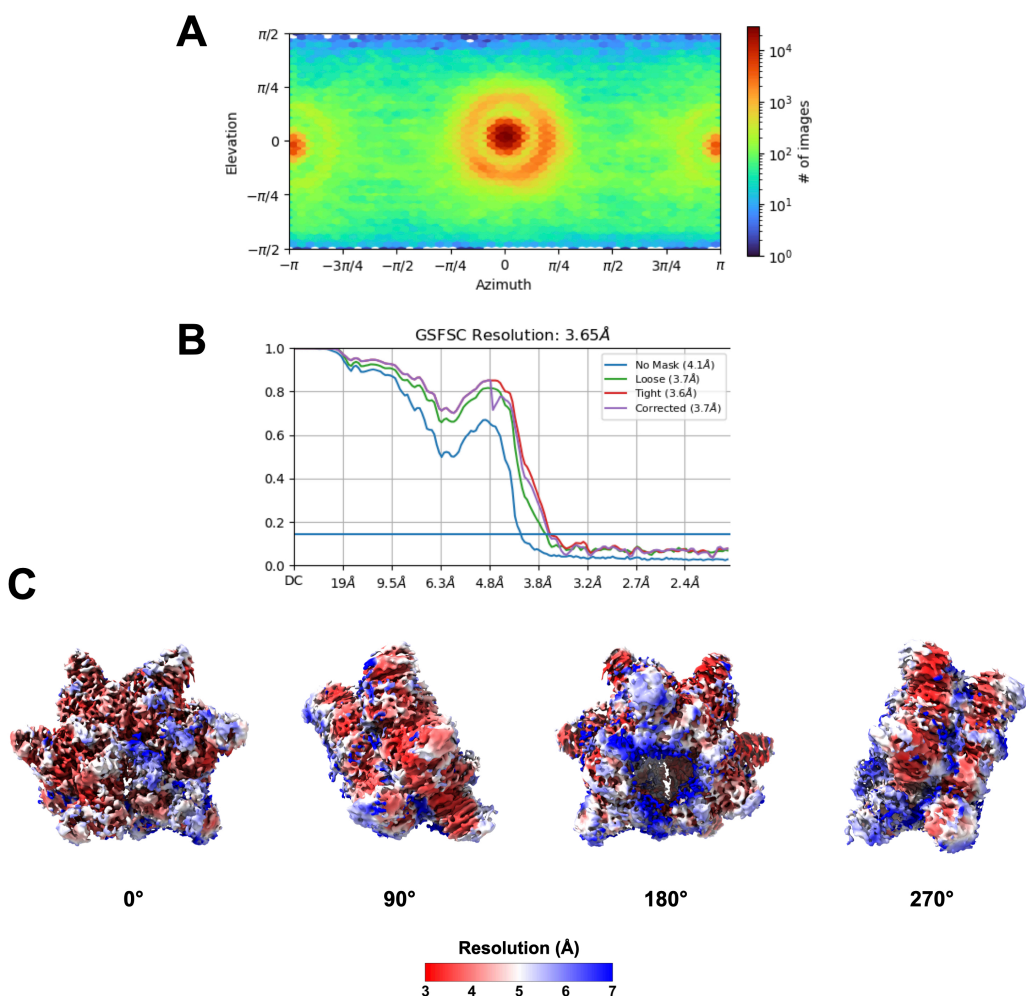
Supplementary Figure 1. Workflow steps in the processing of cryo-EM data for the DNA-free Vc DnaA-KH•DnaB-DciA complex. Data processing procedure as reported in the Methods section. Percentage of particle subsets composing each volume was calculated against the starting picked particles. Color are unique for representation. Maps T and W portray our best result and were used for 3D model building.



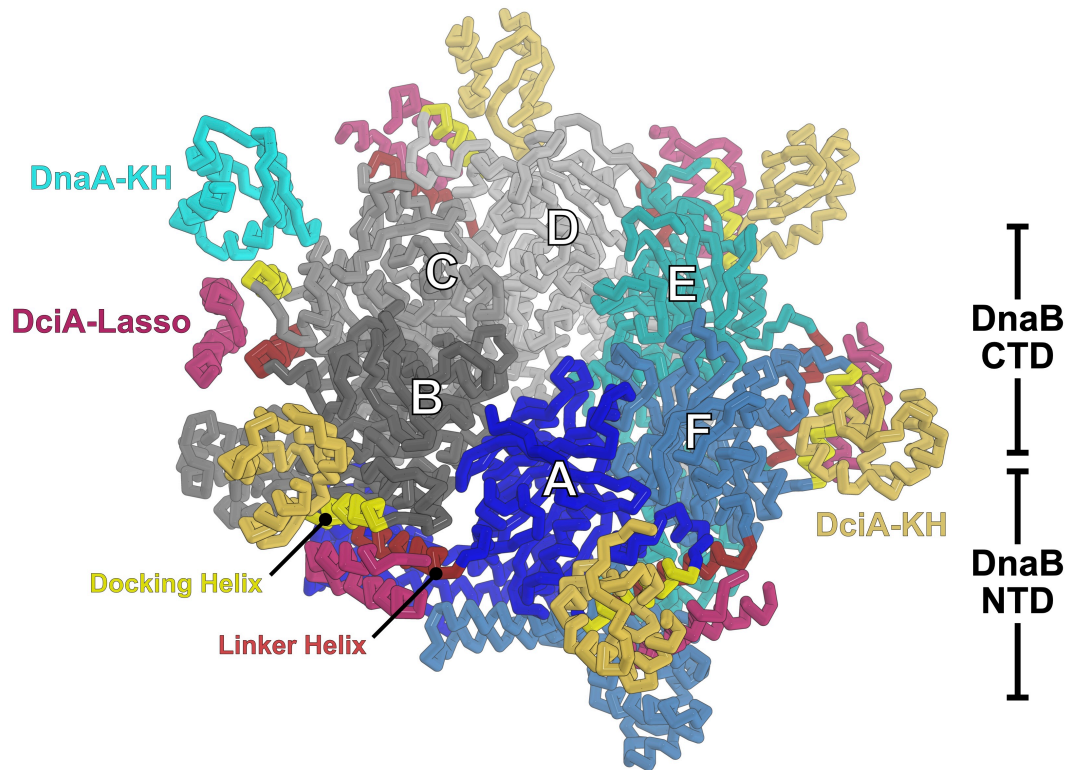
Supplementary Figure 2. Cryo-EM processing statistics and local resolution estimation for the DNA-free Vc DnaA-KH•DnaB-DciA complex. Data shown for Map S (above), refined imposing C1 symmetry, and Map V (below) refined with C3 symmetry imposed. CryoSPARC plot for particles distribution, where red clusters highlight particles preferred orientation (A, D). Fourier Shell Correlation plots from the Non-Uniform Refinement sessions for the calculation of density volume resolution at FSC = 0.143, namely 3.87 Å for Map S (B) and 3.63 Å for Map V (E). Local resolution estimation shown as surface color from red to blue. Each density map is represented in four poses, differing by 90° rotation along the Y axis (C, F).



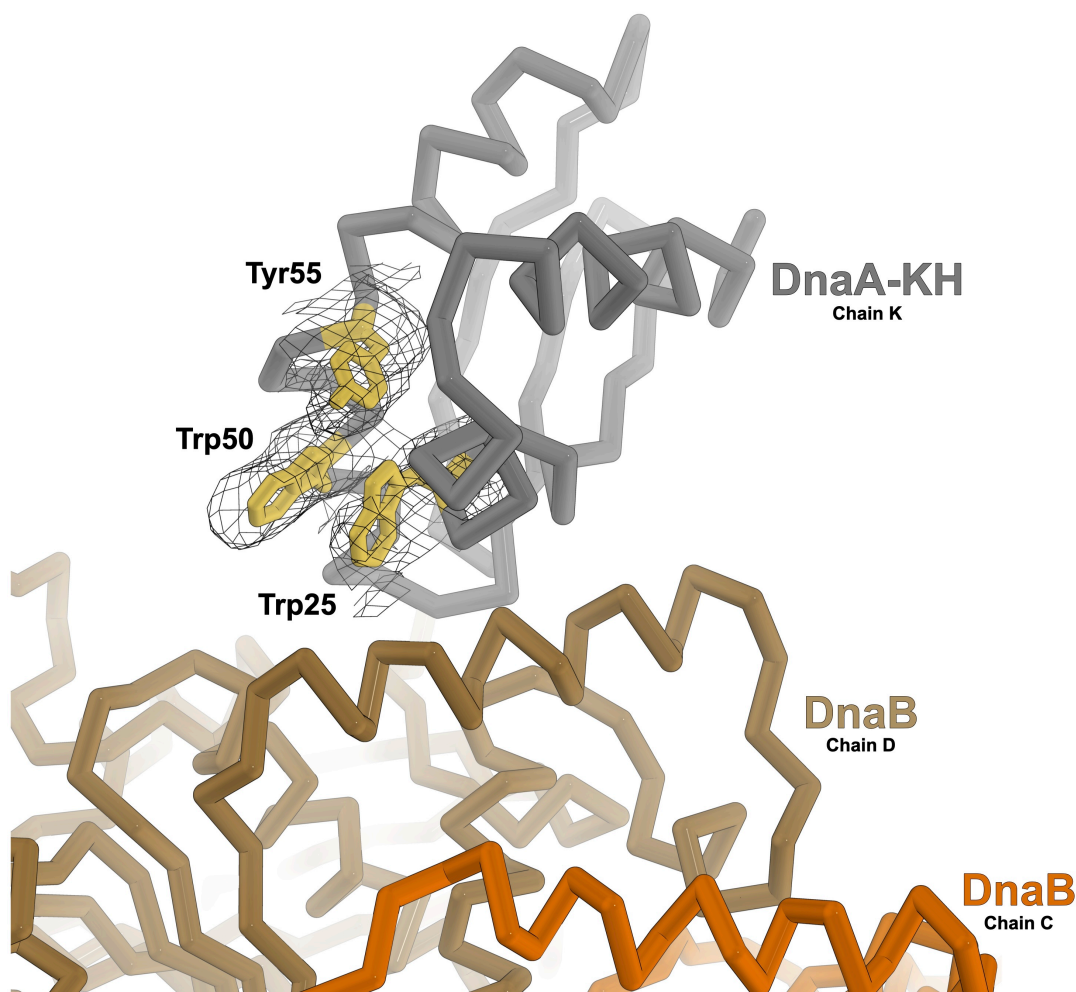
Supplementary Figure 3. Workflow steps in the processing of cryo-EM data for the DNA-bound Vc DnaA-KH•DnaB-DciA complex. CryoSPARC data processing scheme as described in the Methods section. Percentage of particle subsets composing each volume was calculated against the starting picked particles. Color are unique for representation. Map M represents the best result obtained from this dataset and was used as template for 3D model building.



Supplementary Figure 4. Cryo-EM processing statistics and local resolution estimation for the DNA-bound Vc DnaA-KH•DnaB-DciA complex. CryoSPARC plot for particles distribution (azimuth and elevation). The number of particles is colored from blue (10^0) to red (10^4), and the latter highlights particles preferred orientation (A). Fourier Shell Correlation plot from the Non-Uniform Refinement job producing Map L and calculating a resolution of 3.65 Å at FSC = 0.143 (B). Local resolution estimation shown as surface color from red to blue (3 Å to 7 Å). Each volume map is represented in four poses, imposing a rotation by 90° along the Y axis (C).



Supplementary Figure 5. Cryo-EM structure of ssDNA-free *Vibrio cholerae* DnaA-KH•DnaB•DciA complex. The Vc protein complex is shown as ribbon; DnaB monomers are labelled with white letters; Linker Helix (LH) and Docking Helix (DH) elements are colored in red and yellow, respectively; five DciA-KH units are colored in dark yellow while the only DnaA-KH monomer is in cyan; all six DciA-Lasso domains are represented in pink.



Supplementary Figure 6. **Big aromatic residues are unique to VcDnaA's KH fold.** The cryo-EM density map of aromatic amino acids Trp25, Trp50 and Tyr55 belonging to the KH domain of DnaA was key for 3D model building, as a distinguishing factor than DciA-KH. The density map is shown as black mesh. The illustrated structure belongs to the DNA-bound Vc complex. The protein backbone is shown as ribbon. Residues Trp25, Trp50 and Tyr55 of DnaA-KH (chain K) are displayed as yellow sticks.

MATERIALS AND METHODS

Plasmids, Protein Expression and Purification

For detailed plasmid design, expression and purification protocols on VcDnaB-untagged and VcDciA-FL-CHis refer to Materials and Methods section of Chapter 3.

VcDnaA-KH-NHis. The gene corresponding to the isolated KH fold-like Domain I of *Vibrio cholerae* DnaA (residues 1:79) was subcloned in the expression vector pACYCDuet1, bearing an amino-terminal hexahistidine purification tag.

For scaled-up bacterial growth, *E. coli* BL21(DE3) bacterial colonies obtained over-night at 37 °C on solid medium (LB supplemented with 34 µg/mL of chloramphenicol), were resuspended in 200 mL of liquid LB medium (containing 34 µg/mL of chloramphenicol) and let grow over-night at 37 °C under shaking. The derived culture was used as starter for the inoculation of 4 L of liquid LB medium. Bacterial growth was allowed at 37 °C, 180 rpm up to OD₆₀₀ = 0.6. Then, protein expression was induced with 0.5 mM IPTG and, and the bacterial culture was left at 16 °C over-night. The procedure yielded 7 g/L of bacterial biomass, that, after harvesting step, was resuspended in 50 mM K-HEPES pH 7.5, 400 mM NaCl, 1 mM MgCl₂, 5% glycerol, 5 mM β-mercaptoethanol (5 mL/g).

Protein purification was carried out at 4 °C or on ice. ~90 mL of resuspended cells were lysed by sonication (30 sec. On, 30 sec. Off. Total On time = 3 min. 70% amplitude) and centrifuged for 30 mins at 16,000 rpm. The supernatant was recovered and incubated with Ni²⁺-NTA agarose beads, previously equilibrated in Buffer A (20 mM Tris pH 8, 300 mM NaCl, 5% glycerol and 5 mM β-mercaptoethanol), for 40 minutes under shaking. Affinity chromatography was carried out by washing the resin with 12 column volumes (CVs) of Buffer A

supplemented with 15 mM imidazole and by eluting the His-tagged protein with Buffer A containing 250 mM imidazole. The protein preparation was then loaded onto a Q-Sepharose FF column for anion exchange, equilibrated using Buffer B (20 mM Tris pH 8, 50 mM NaCl, 5% glycerol and 5 mM β -mercaptoethanol). A 15 CVs gradient from 5% to 100% of Buffer C (20 mM Tris pH 8, 1 M NaCl, 5% glycerol and 5 mM β -mercaptoethanol) was enforced to elute VcDnaA-KH-NHis. AmSO₄ at the final concentration of 1 M was added to the obtained anion exchange column protein pool, subsequently loaded onto a Butyl HIC equilibrated in Buffer D (20 mM Tris pH 7, 1 M AmSO₄, 5% glycerol and 5 mM β -mercaptoethanol). 4 CVs were applied to wash the hydrophobic resin using Buffer E (20 mM Tris pH 7, 50 mM AmSO₄, 5% glycerol and 5 mM β -mercaptoethanol). A 15 CVs gradient from 0% to 100% of Buffer E was enforced to elute the recombinant protein, further precipitated in AmSO₄ (25% saturation, 10 minutes at 4 °C under stirring), recovered through centrifugation and resuspended in Buffer F (20 mM K-HEPES pH 7.5, 300 mM NaCl, 5% glycerol, 5 mM β -mercaptoethanol) for final gel-filtration step, carried out in a 320 mL Superdex 200 column. Around 16 mg of pure VcDnaA-KH-NHis were obtained at the final concentration of 13.6 mg/mL, stored at -80 °C for further use.

Sample and Grid Preparation for Single Particle Cryo-EM

The preparation of the DNA-free *Vibrio cholerae* protein complex for cryo-EM started with the dilution of VcDnaA-KH-NHis (residues 1:79), VcDnaB-untagged and VcDciA-FL-CHis in 20 mM K-HEPES pH=7.5, 200 mM NaCl, 5 mM MgCl₂, 5% glycerol, 5 mM β -mercaptoethanol, and 0.1 mM ATP, to reach the final concentration of 48 μ M (DnaA), 8 μ M (DnaB hexamer), 48 μ M (DciA). 3 μ L of sample were applied on a Quantifoil UltrAuFoil R1.2/1.3 Au300 grid, previously plasma cleaned in a Gatan Solarus II glow discharger with 3:1 ratio of Ar/O₂ at 15 W for 10

seconds. The grid plunge-freezing procedure was conducted on a Vitrobot Mark IV (ThermoFisher), setting a temperature of 4 °C and 100% humidity, a waiting time of 30 seconds to allow the liquid absorption, a blotting force of 4, and a blotting time of 4. The grid was plunge-frozen in liquid ethane and stored in liquid nitrogen up to the data collection time.

This procedure was followed also for the preparation of the protein-DNA complex, except that a single-stranded polydT45 oligonucleotide was added to the sample at the concentration of 15 µM prior to grid freezing.

Cryo-EM Data Acquisition

All the cryo-EM images were acquired at the New York Structural Biology Center. For the DNA-free Vc DnaA-KH•DnaB•DciA complex, 11229 cryo-EM movies were collected using Leginon¹⁶ on a Titan Krios (ThermoFisher), equipped with a Gatan K3 direct electron detector, operating at 300 KeV, at a magnification of 64,000x, with a pixel size of 1.076 Å/pixel and a defocus range of -0.6 to -1.7 µm. The grid was tilted by 0, 20 and 40 to alleviate the expected preferred orientation of the protein particles, that could compromise the image processing and the 3D volume reconstruction.

Similarly, for the DNA-bound Vc DnaA-KH•DnaB•DciA complex 4660 frames were collected on a 300 KeV Titan Krios (ThermoFisher) featuring a Gatan K3 direct electron detector, setting a magnification of 81,000x. The pixel size was 0.529 Å/pixel, and the defocus range applied was -1 to -2 µm. Once again, the grid was tilted by 0, 20 and 40 degrees to overcome the expected specimen preferred orientation.

Cryo-EM Image Processing

DNA-free Vc DnaA-KH•DnaB•DciA complex. The full dataset composed of 11229 movies and collected on the VcDnaA-DnaB-DciA complex was imported in cryoSPARC¹⁷ and submitted to motion correction and CTF estimation. Manual curation of micrographs based on ice thickness (0.94-1.1) and CTF fit resolution (2.16-10) allowed the selection of 9565 images, used as input for blob picking. About 23.7 million particles were identified, extracted (box size = 360; F-crop to box size = 120; ~19.9 mln particles) and submitted to two rounds of 2D classification, which ensured the selection ~4.7 million particles, suitable for the next steps. The *ab-initio* reconstruction with C1 symmetry enforced calculated four volumes (A0, A1, A2 and A3), where only maps A0 (~1.34 million particles) and A1 (~1.26 million particles) were carried on to heterogeneous refinement. This generated two volumes: B0 (~2.53 million particles), that showcased the higher quality despite the presence of extra density map, attributed to particles at very close contact, and B1 (~2.19 particles). With the aim of getting rid of contaminating density coming from neighboring particles and focusing on the 3D reconstruction of a single DnaB hexamer, two rounds of consecutive 2D classification, *ab-initio* reconstruction (C0, C1 and E) and heterogeneous refinement (D0, D1; F0, F1; G0, G1) were trialed. This procedure yielded a total of nine volumes, and the two best, namely F1 (~775K particles) and G0 (~1.07 million particles), were submitted to a further heterogeneous refinement round, obtaining classes H0 (~603K particles) and H1 (~492K particles), the former judged of higher overall quality, not bearing undesired extra density. Next, homogeneous refinement on H0 provided volume I at the resolution of 6.74 Å. Unfortunately, extra density map appeared once again. Hence, we moved on applying a mask-based particle subtraction approach. Newly calculated particles (~598K) were committed to *ab-initio* reconstruction (class J), homogenous refinement (class K) and non-uniform refinement (class L), producing

a volume at 6.74 Å resolution, free of any unwanted map contamination. The particles re-extraction at box size equal to 320 px and a final non-uniform refinement job let the achievement of a 4.13 Å resolution map (class M).

With the idea of involving more particles in the EM volume reconstruction, to improve the completeness and reduce the of anisotropy found in class M, a second mask-based approach was followed using map B0 as starting point. Homogeneous refinement produced volume N, suitable for mask-based particle subtraction. Subsequent *ab-initio* reconstruction (P0, P1), heterogeneous refinement (Q0, Q1) and homogenous refinement allowed us to generate class R (~981K particles). However, its close inspection revealed, once again, unwanted extra density map. Therefore, a new round of mask-based particle subtraction was enforced and the final non-uniform refinement on re-extracted particles without Fourier crop (pixel size = 2.14 Å) calculated map S at the resolution of 3.87 Å, featuring better details than class M and submitted to deep learning-based automatic sharpening with DeepEMhancer¹⁴, obtaining Class T. The quality of map T was high enough for allowing 3D model building.

Eventually, to cope with the volume anisotropy, global CTF refinement was used and resulting particles were employed in non-uniform refinement applying C1 or C3 symmetry, respectively producing map U (4.06 Å) and map V (3.63 Å). Owing to the enhanced resolution, map V was submitted to sharpening with DeepEMhancer, concluding the data processing obtaining map W, later used for 3D model building (described below) and comparison with class T.

DNA-bound Vc DnaA-KH•DnaB•DciA complex. The complete dataset of 4660 frames collected on the DnaA-DnaB-DciA-ssDNA sample was processed using cryoSPARC¹⁷. At first, micrographs were submitted to motion correction (Fourier crop factor of ½) and CTF estimation applying standard settings.

Leveraging on calculated ice thickness and CTF fit resolution, images at low overall quality were discarded, selecting a subset of 4419 frames, that underwent blob picking. About 10.9 million particles were identified, extracted (box size = 360; F-crop to box size = 120) and submitted to 2D classification and *ab-initio* reconstruction enforcing C1 symmetry. The obtained volumes A0 and A1 contained respectively ~568K and ~1.25 million particles. Class A1 was considered of higher quality and brought onward to heterogeneous refinement, which produced volumes B0 and B1. The former displayed the best features and was submitted to non-uniform refinement, yielding an EM map with a resolution of 6.62 Å (Volume C), exploited as template in the subsequent processing steps.

Template-based picking was performed against the manually curated dataset using map C, producing ~3.97 million particles. This collection was extracted (box size = 360; F-crop to box size = 120), 2D classified and submitted to *ab-initio* reconstruction (C1 symmetry) obtaining two classes, namely D0 and D1 (~600K and ~265K particles). EM density D0 featured the best details, whereas D1 is a junk volume, and underwent heterogeneous refinement together with volume C as job input. The inspection of the resulting volumes (E0, ~985K particles; E1, ~869K particles) implied that E0 was of higher quality and, therefore, carried on to particle re-extraction at full box size (360 px) and homogeneous refinement, that yielded map F (~979K particles) at the resolution of 4.28 Å. Particles composing volume F were further 2D classified, in order to remove bad ones that could affect the quality of the final result. ~865K particles were selected and submitted to *ab-initio* reconstruction (G0, G1) and heterogeneous refinement, that calculated maps H0 and H1 (~460K and ~405K particles), with H0 showcasing higher quality. Then, leveraging on all the particles included in the previous job, homogeneous refinement on volume H0 produced map I at 4.26 Å resolution, subsequently submitted to non-uniform refinement, obtaining map J (3.97 Å). Also in this case,

global CTF refinement was used to enhance 3D volume reconstruction, lowering the anisotropy level. Next, two non-uniform refinement trials were run in parallel imposing C3 and C1 symmetry, respectively calculating map K (3.61 Å) and L (3.65 Å). Despite the higher resolution, map K resulted in a loss of details around the single-stranded DNA present inside of DnaB's inner chamber. Therefore, class L was considered our best volume and finally sharpened with deepEMhancer, concluding the data processing with density map M.

Model Building and Structure Refinement

DNA-free Vc DnaA-KH•DnaB•DciA complex. The 3D model of *Vibrio cholerae* DnaA-DnaB-DciA protein complex was built based both on DeepEMhancer-sharpened map T and W. At first, the AlphaFold predicted model of VcDnaB was splitted into its N-terminal Domain (NTD, residues 22:170), Linker Helix (LH, residues 171:211) and C-terminal Domain (CTD, residues 212:468), employed as input in MolRep, part of the CCP4i2 suite¹⁸, in order to reconstruct the hexameric assembly of VcDnaB with excellent stereochemistry. Then, the COOT¹⁹ was used to connect each domain composing every DnaB monomer and to remove residues out of the density map. MolRep was also use to account for the density volume surrounding each DH-LH interface of the helicase, which was reminiscent of the KH fold that could belong to either VcDnaA or VcDciA. We successfully modelled five copies of DciA's isolated KH domain (residues 27:96), one copy of DnaA-KH (residues 1:79). In addition, six copies of DciA's lasso (residues 119:157) were included in the 3D reconstruction.

At last, the fitting of the model in the EM map was improved using Real Space Refinement in Phenix²⁰, relying on minimization global, morphing, NQH flips, atomic displacement parameters (ADP), Ramachandran and secondary structure restraints.

DNA-bound Vc DnaA-KH•DnaB•DciA complex. The 3D structure of the *Vibrio cholerae* protein-DNA complex was built exploiting an approach similar to that of the DNA-free state, previously described. Splitted N-terminal Domain (NTD, residues 1:170), Linker Helix (LH, residues 171:211) and C-terminal Domain (CTD, residues 212:468) of the AlphaFold¹⁰ predicted VcDnaB structure were used in MolRep¹⁸ and COOT¹⁹ to build the first hexameric assembly, based on the DeepEMhancer sharpened map M. Additional EM density reminiscent of both the KH fold and a β -hairpin domain was found adjacent to each DH-LH element also in this case, and we successfully modelled six VcDnaA-KH units (residues 1:79) and six VcDciA-Lassos (residues 121:157), exploiting both MolRep¹⁸ and PyMol.

Next, the examination of DnaB's nucleotide binding sites unveiled the presence of six ADP molecules, included in the model using COOT. At the same time, we found ssDNA inside DnaB's inner chamber. However, the resolution was enough only for making out the DNA backbone from the nucleobases, but not for assigning specific nucleotides unambiguously. The oligonucleotide composed of six thymines was modelled using COOT.

We relied on Phenix²⁰ to refine the ADP-bound Vc DnaA-DnaB-DciA •ssDNA complex. Real Space Refinement was utilized setting minimization global, NQH flips, atomic displacement parameters (ADP), Ramachandran and secondary structure restraints.

Model analysis and Visualization

Structural analysis and 3D model visualization of DNA-free and DNA-bound DnaA-DnaB-DciA complexes were performed using the CCP4 software package¹⁸, COOT¹⁹, UCSF-CHIMERA²¹, Phenix²⁰ and PyMol. Molecular graphics and figures were produced using UCSF-CHIMERA²¹ and PyMol.

Supplementary Table 1: Cryo-EM Data Measurement and PDB Model Refinement of DNA-free Vc DnaA-KH•DnaB•DciA complex

PDB	TBD
EMDB	TBD
Data Collection and Processing	
Microscope	Titan Krios
Voltage (kV)	300
Camera	Gatan K3
Magnification	64000
Defocus range (μm)	-0.6 to -1.7
Exposure time (ms)	2200
Electron dose (e ⁻ /Å ²)	57.00
Number of frames collected (no.)	44
Number of frames processed (no.)	44
Calibrated pixel size (Å)	1.076
Micrographs (no.)	11229
Total particle images (no.)	23714005
Refinement vs Map T (C1 symmetry)	
Particle per class (no.)	980278
Symmetry imposed	C1
Map resolution (Å), FSC = 0.143	3.87
Map sharpening B-factor (Å ²)	160.1
Map versus Model cross-correlation	0.54
Model Composition	
Non-hydrogen atoms	25208
Protein residues	3228

Model Refinement	
Resolution limit (Å)	3.87
Number of chains	15
Number of residues	3228
B factors (Å²)	
Protein	120.78
RMS deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.469
Validation	
MolProbity score	2.40
Clashscore	13.11
Rotamer outliers (%)	4.07
Ramachandran plot (%)	
Favored (%)	95.56
Allowed (%)	4.38
Outliers (%)	0.06
Refinement vs Map W (C3 symmetry)	
Particle per class (no.)	980278
Symmetry imposed	C3
Map resolution (Å), FSC = 0.143	3.63
Map sharpening B-factor (Å ²)	166.9
Map versus Model cross-correlation	0.53
Model Composition	
Non-hydrogen atoms	26105

Protein residues	3345
Model Refinement	
Resolution limit (Å)	3.63
Number of chains	18
Number of residues	3345
B factors (Å²)	
Protein	135.85
RMS deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.583
Validation	
MolProbity score	2.67
Clashscore	16.50
Rotamer outliers (%)	7.08
Ramachandran plot (%)	
Favored (%)	95.56
Allowed (%)	4.26
Outliers (%)	0.18

Supplementary Table 2: Cryo-EM Data Measurement and PDB Model Refinement of DNA-bound Vc DnaA-KH•DnaB•DciA complex

PDB	TBD
EMDB	TBD
Data Collection and Processing	
Microscope	Titan Krios
Voltage (kV)	300
Camera	Gatan K3
Magnification	81000
Defocus range (μm)	-1.0 to -2.0
Exposure time (ms)	494
Electron dose (e ⁻ /Å ²)	53.57
Number of frames collected (no.)	40
Number of frames processed (no.)	40
Calibrated pixel size (Å)	0.5291
Micrographs (no.)	4660
Total particle images (no.)	10894710
Refinement	
Particle per class (no.)	846009
Symmetry imposed	C1
Map resolution (Å), FSC = 0.143	3.65
Map sharpening B-factor (Å ²)	127.0
Map versus Model cross-correlation	0.58
Model Composition	
Non-hydrogen atoms	25975
Protein residues	3273
Nucleotide residues	6
Ligand: ADP	4

Model Refinement	
Resolution limit (Å)	3.65
Number of chains	20
Number of residues	3273
B factors (Å²)	
Protein	152.33
Nucleotide	61.93
Ligand (ADP)	103.77
RMS deviations	
Bond lengths (Å)	0.19
Bond angles (°)	0.648
Validation	
MolProbity score	2.53
Clashscore	15.53
Rotamer outliers (%)	6.42
Ramachandran plot (%)	
Favored (%)	96.58
Allowed (%)	3.42
Outliers (%)	0.00

REFERENCES

1. Fernandez, A. J. & Berger, J. M. Mechanisms of hexameric helicases. *Crit. Rev. Biochem. Mol. Biol.* **56**, 621–639 (2021).
2. Messer, W. et al. Bacterial replication initiator DnaA. Rules for DnaA binding and roles of DnaA in origin unwinding and helicase loading. *Biochimie* **83**, 5–12 (2001).
3. Mann, K. M. et al. Rv0004 is a new essential member of the mycobacterial DNA replication machinery. *PLoS Genetics* **13**, (2017).
4. Mott, M. L. & Berger, J. M. DNA replication initiation: mechanisms and regulation in bacteria. *Nat. Rev. Microbiol.* **5**, 343–354 (2007).
5. Davey, M. J. & O'Donnell, M. Replicative helicase loaders: Ring breakers and ring makers. *Current Biology* **13**, (2003).
6. Valverde, R., Edwards, L. & Regan, L. Structure and function of KH domains. *FEBS J.* **275**, 2712–2726 (2008).
7. Abe, Y. et al. Structure and Function of DnaA N-terminal Domains. *J. Biol. Chem.* **282**, 17816–17827 (2007).
8. Hayashi, C., Miyazaki, E., Ozaki, S., Abe, Y. & Katayama, T. DnaB helicase is recruited to the replication initiation complex via binding of DnaA domain I to the lateral surface of the DnaB N-terminal domain. *J. Biol. Chem.* **295**, 11131–11143 (2020).
9. Brézellec, P., Vallet-Gely, I., Possoz, C., Quevillon-Cheruel, S. & Ferat, J.-L. DciA is an ancestral replicative helicase operator essential for bacterial replication initiation. *Nat. Commun.* **7**, 13271 (2016).
10. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
11. Baek, M. et al. Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nat. Methods* **21**, 117–121 (2024).
12. Blaine, H. C., Burke, J. T., Ravi, J. & Stallings, C. L. DciA Helicase Operators Exhibit Diversity across Bacterial Phyla. *J. Bacteriol.* **204**, e00163-22 (2022).
13. Chase, J., Berger, J. & Jeruzalmi, D. Convergent evolution in two bacterial replicative helicase loaders. *Trends Biochem. Sci.* **47**, 620–630 (2022).
14. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).
15. Itsathitphisarn, O., Wing, R. A., Eliason, W. K., Wang, J. & Steitz, T. A. The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell* **151**, 267–277 (2012).
16. Cheng, A. et al. Leginon: New features and applications. *Protein Sci.* **30**, 136–150 (2021).

17. 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 (2017).
18. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **67**, 235–242 (2011).
19. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **60**, 2126–2132 (2004).
20. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D: Struct. Biol.* **74**, 531–544 (2018).
21. Pettersen, E. F. et al. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

UNPUBLISHED RESULTS PART 2

*INTEGRATIVE STRUCTURAL BIOLOGY APPROACHES
FOR THE INVESTIGATION OF MYCOBACTERIAL
PROTEINS INVOLVED IN THE DNA REPLICATION
INITIATION PROCESS*

The DH-LH element of DnaB is a unique target for designing new drugs inhibiting the DNA replication initiation in *Mycobacterium tuberculosis*

INTRODUCTION

The bacterial replicative helicase DnaB is commonly known as a two-layered ring-shaped hexameric protein, that assumes a right-handed spiral configuration during the translocation mechanism onto the ssDNA^{1,2}. As carefully depicted in Chapter 3, a peculiar feature is represented by two α -helices that pack against each other, composing the so-termed DH-LH module. Indeed, the Linker Helix (LH) of one protomer, usually spanning between residues ~170:215, positions itself in close vicinity to the Docking Helix (DH; residues ~285:300) of the adjacent unit.

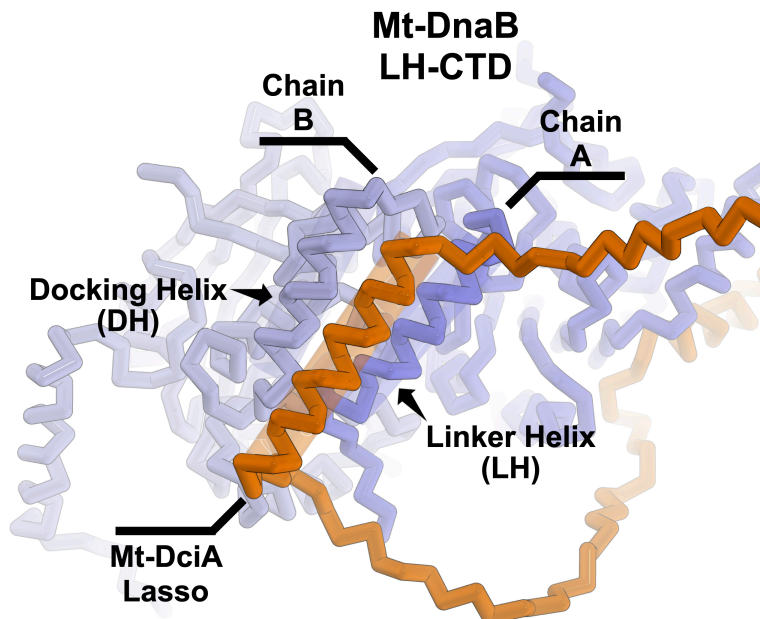


Figure 1. In the AlphaFold3 structure prediction, MtDciA's Lasso is facing the DH-LH module of DnaB. Ribbon representation of the AlphaFold3 calculated structure of two adjacent monomers of MtDnaB-LH-CTD (in purple shades) in complex with MtDciA-FL (orange). DciA-Lasso's α -helix is predicted to be docked in front of DnaB's DH-LH element, a step forward in confirming that this assembly is respected also in the mycobacterial helicase-loader system.

Besides, cryo-electron microscopy and crystallographic 3D structures on helicase-loader complexes from different homologs identified how DciA places its lasso/hook-like domain near the DH-LH motif to strongly interact with DnaB (PDB: 6BBM³, 6QEL⁴, 8A3V⁵). All this is sustained by a few biochemical complementary approaches⁵. Remarkably, even in *Mycobacterium tuberculosis* this system seems to be no different.

In absence of experimentally determined structures, we leveraged on AlphaFold3⁶ to predict how the short construct *MtDnaB*-LH-CTD at our disposal could interact with the full-length loader. As shown in Figure 1, by imposing a helicase-loader ratio of 2:1, we identified two fundamental pieces of evidence. Firstly, even though the input sequence for DnaB was the deleted one (residues 171:457), the dimeric assembly respects the canonical feature of bacterial helicases, giving rise to the abovementioned DH-LH element. Moreover, *MtDciA*'s Lasso was docked facing the DH-LH motif, as attended.

The inference from homolog systems and the computational reasoning over the mycobacterial complex need to be sustained by experimentally. Thus, while aiming to find the best conditions to solve the 3D structure of the mycobacterial target, we turned to 'pull-down' assays coupled to size-exclusion chromatography analysis, to biochemically determine the MTB helicase-loader complex ratio, which resulted being harder than expected to be solved via native Mass Spectrometry.

RESULTS AND DISCUSSION

As shown in Chapter 3, the series of so-called Ni-NTA 'pull-down' experiments based on the available *MtDnaB* and *MtDciA* truncated constructs, firstly allowed us to identify which were the minimum required elements so that helicase and loader could positively interact. Not only did the full-length proteins

co-elute, but the interaction with the C-terminal His-tagged loader, the “bait”, was also retained by utilizing untagged MtDnaB-LH-CTD (residues 171:457) as “prey”.

To further investigate the complex stability and the oligomeric assembly, we scaled-up the Ni²⁺-based affinity chromatography step in order to obtain a reasonable amount of stable protein complex, which subsequently underwent size-exclusion chromatography (SEC). The careful inspection of the SEC chromatogram and its comparison to the related calibration curve and SDS-PAGE showcased that the DnaB-LH-CTD:DciA-FL complex co-purifies at an elution volume compatible to a molecular weight of 75-85 kDa (Figure 2), a suitable range for a 2:1 ratio of the investigated protein complex (MW: DnaB-LH-CTD = 32.2 kDa; DciA-FL = 21 kDa).

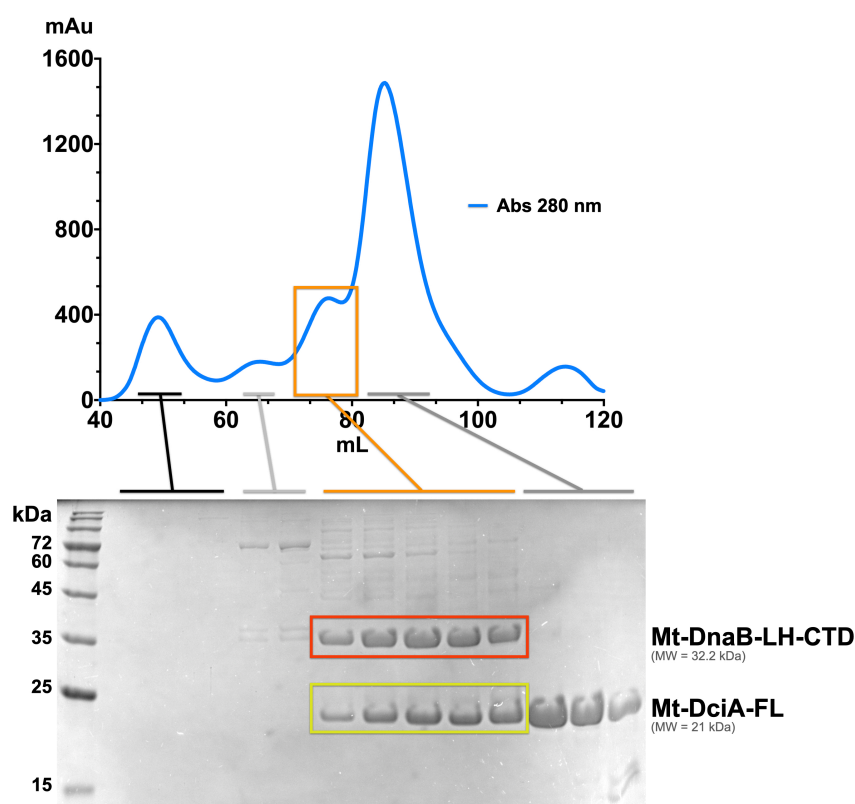


Figure 2. Full-length MtDciA tightly binds to DnaB’s truncation LH-CTD. Analysis of the co-purification through size-exclusion chromatography of the MtDnaB-LH-CTD:MtDciA-FL complex. The absorbance peak (blue line) corresponding to an elution volume of 70-80 mL (orange square) carries the investigated protein complex, confirmed by the SDS-PAGE below, which displays the simultaneous presence of pure MtDnaB-LH-CTD (red square) and MtDciA-FL (yellow square) bands, both at the expected molecular weight, in the same gel lanes.

Remarkably, this is in strong agreement with most of our native Mass Spectrometry (nMS) data obtained on the *Vibrio cholerae* system, stating that the correct assembly of the DH-LH interface of DnaB is required to allow DciA's docking. As in this peculiar structural domain the two α -helices belong to distinct DnaB monomers, the DH-LH element can take place only when two monomeric units are adjacent, allowing then DciA's Lasso to be engaged (Figure 1). Indeed, in our 'pull-down' series (see Chapter 3), even isolated DciA's Lasso retained a strong interaction with both the full-length and LH-CTD forms of DnaB. The DH helix is composed of amino acids 286:302, therefore our *MtDnaB*-LH-CTD construct bears all the features to let this interaction occur, hence proving how the N-terminal tier of the helicase is not involved in making any sort of contact with the loader. Moreover, this evidence helped in shedding more light on the protein complex formation ratio in the mycobacterial system, otherwise impossible to be established via nMS, given that the obtained results did not match our quality control.

The integration of this outcome in our scenario strengthens the overview on how fundamental DnaB's DH-LH interface is in the helicase loading mechanism, making it a strong druggable target. Finally, having optimized the procedure, in the next future we will exploit this protocol to screen new compounds able to interfere with the helicase-loader complex formation by displacing DciA from DnaB, aiming to inhibit the DNA replication initiation in *Mycobacterium tuberculosis* with a therapeutic purpose. The design of these molecules, most likely peptidomimetics, will be based on the DciA's isolated Lasso sequences that displayed strong and positive interaction with the cognate DnaB helicase.

MATERIALS AND METHODS

Plasmids and Protein Expression

For detailed plasmid design and protein expression conditions, see the Materials and Methods section of Chapter 3. BL21(DE3) *E. coli* cells transformed with the DnaB short construct LH-CTD (residues 271:457) were grown in 2xYT medium supplemented with Streptomycin (50 µg/mL) at 37 °C to an OD₆₀₀ ≈ 0.6, after which IPTG was added to 0.5 mM for culture induction. Protein expression was allowed for 18 hours at 17 °C (RPM = 210). MtDciA-FL was obtained using the same procedure, except that the protein expression was allowed for 3 hrs at 30 °C.

Protein co-purification

Specifically for this test, the scale-up was performed by co-purifying the bacterial biomass obtained from 1 L of culture independently expressing MtDnaB-LH-CTD and MtDciA-FL. Then, the Nickel-based affinity chromatography was performed based on the pull-down experiments described in the Materials and Methods section of Chapter 3. Highly concentrated fractions containing the desired protein complex were concentrated through ultracentrifugation using Vivaspin 10K. The preparation was then injected into a HiLoad 16/600 Superdex 200 prep grade column, previously equilibrated in 20 mM Tris-HCl pH = 8, 500 mM NaCl, 5 mM MgCl₂, and 5% glycerol. The protein composition of gel-filtration eluted aliquots was analyzed by SDS-PAGE (12.5% acrylamide gel).

REFERENCES

1. O'Donnell, M. E. & Li, H. The ring-shaped hexameric helicases that function at DNA replication forks. *Nature Structural and Molecular Biology* **25**, 122–130 (2018).
2. Fernandez, A. J. & Berger, J. M. Mechanisms of hexameric helicases. *Crit. Rev. Biochem. Mol. Biol.* **56**, 621–639 (2021).
3. Chase, J. et al. Mechanisms of opening and closing of the bacterial replicative helicase. *eLife* **7**, e41140 (2018).
4. Arias-Palomo, E., Puri, N., Murray, V. L. O., Yan, Q. & Berger, J. M. Physical Basis for the Loading of a Bacterial Replicative Helicase onto DNA. *Mol. Cell* **74**, 173-184.e4 (2019).
5. Cargemel, C. et al. The LH-DH module of bacterial replicative helicases is the common binding site for DciA and other helicase loaders. *Acta Crystallographica Section D: Structural Biology* **79**, 177–187 (2023).
6. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).

SAXS and NMR-based structural investigation of the mycobacterial helicase loader DciA

INTRODUCTION

DciA is a DNA replication initiation factor with an essential role for bacterial growth and viability¹, yet its structural characterization is quite limited, especially within the *Mycobacterium tuberculosis* complex (MTBC). Its ability consists in sustaining the loading of the replicative helicase DnaB onto the exposed single-stranded DNA (ssDNA) at the level of the origin of replication *oriC*, after the initiator protein DnaA had melted the AT-rich region, opening the replication bubble². As extensively discussed in Chapter 3, the helicase loading mechanism comprises the ring-opening, the ring-forming, and the ring-closing strategies, where *E. coli* DnaC, *B. subtilis* DnaI and the bacterial Rho are well-known representatives, respectively³⁻⁵. Although the widely studied DnaC and DnaI act as ATP-dependent factors, the investigation of DciA's role in diverse organisms never showcased a similar behavior. Notably, DnaC and DnaI are evidence of evolutionary divergence, indeed DciA was identified in most bacterial phyla (Figure 1), making it the common ancestral gene⁶.

To be validated as helicase loader, a DNA replication factor should, at least, i) interact with the replicative helicase, ii) sustain DnaB loading onto ssDNA, and iii) repress the helicase ATPase activity to suppress unwanted triggering of the replication initiation system⁷. Whereas some of these questions have been addressed on DciA belonging to different bacteria^{1,8,9}, with our work described in Chapter 3 joining and boosting this scenario, the experimental 3D reconstruction of full-length DciA-like protein and its reshaping when complexed to the replicative helicase DnaB are still to be fully unveiled.

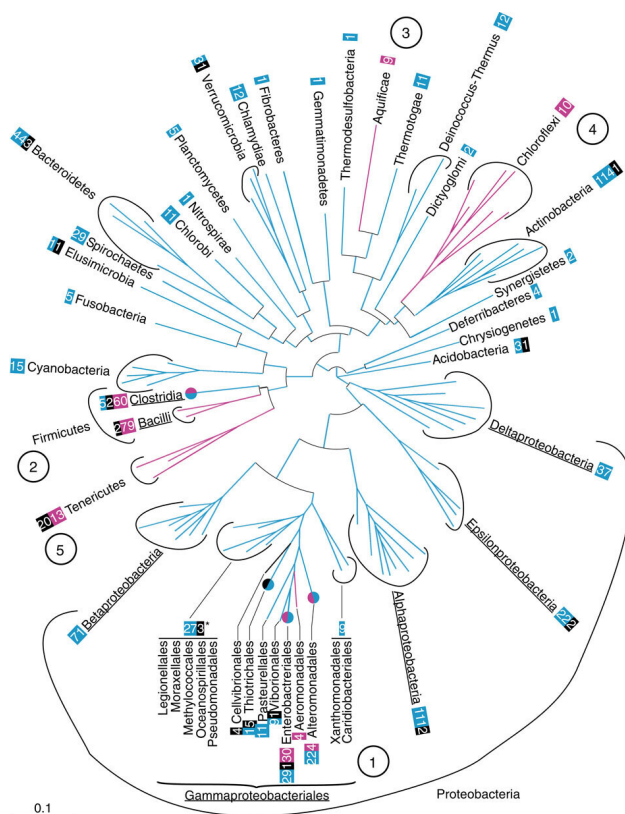


Figure 1. DciA is the most widespread bacterial replicative helicase loading factor. In 2016, Brézellec group identified the *dciA* gene in 23 out of 26 bacterial phyla, indicating it as the probable common ancestor from which DnaC- and DnaI-like loaders originated. In cyan, *dciA* gene-containing species, in magenta *dna[CI]* gene-containing species⁶.

In general, DciA proteins feature a globular, more compact domain, found to closely resemble the N-terminus of DnaA, named KH domain, and one or more intrinsically disordered regions (IDRs) at either the C-terminal or N-terminal extensions, named Lasso domain/s. In 2022, Stallings group took a computational approach to carefully deal with the structural diversity of DciA among different bacterial phyla¹⁰. They succeeded in segregating DciA homologs in different classes leveraging on the location of the IDRs, detecting lineage-specific trends that could shed light on how the DNA replication could be variously regulated in

different phyla. Notably, AlphaFold revealed structural tendencies that are common amid most homologs, regardless of their class (Figure 2).

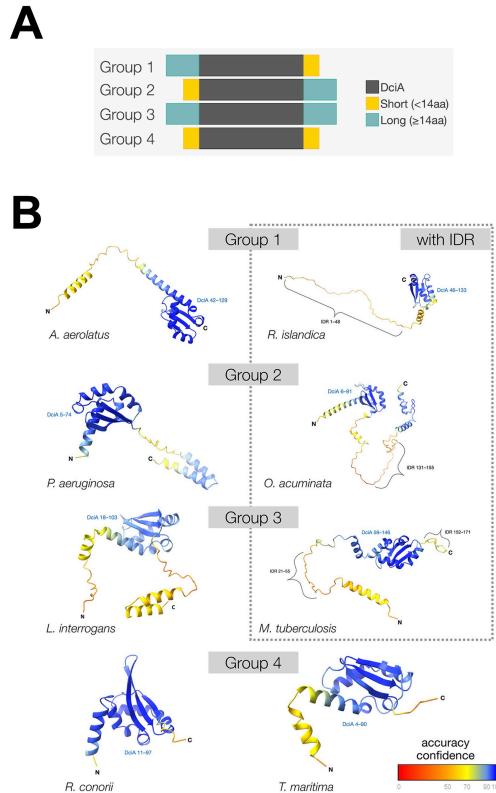


Figure 2. The helicase loader DciA retains common structural features among different bacterial phyla. The computational inspection of DciA's subdomains allowed to group the bacterial loaders into four classes, based on the length of flanking N- and C-terminal extensions (A). Panel B depicts DciA's representative homologs per class, which structure was predicted by AlphaFold. Notably, groups 1, 2 and 3 include members bearing intrinsically disordered regions (IDRs)¹⁰.

Despite all these computational efforts, many biological questions on the multifaceted role of DciA in maintaining the genome integrity, facilitating the efficient DNA replication in bacteria are still open. Above all, the lack of experimentally solved 3D structure is a crucial gap to fill in this specific research field.

Aiming to determine the mycobacterial DciA structure and its dynamics, we applied an experimental approach based on Small-Angle X-ray Scattering (SAXS) and solution Nucleic Magnetic Resonance (NMR) spectroscopy, in order to elucidate the molecular mechanisms underlying the early stages of the DNA replication in *Mycobacterium tuberculosis*.

RESULTS AND DISCUSSION

By means of the expression and purification protocol detailed in the Materials and Methods section below, we successfully obtained a batch of full-length MtDciA at high yield and degree of purity, suitable for Small-Angle X-ray Scattering (SAXS). With this structural biology technique, we aimed to define how the mycobacterial loader behaves in solution, namely its shape and oligomerization state, to account for some peculiarities encountered during the size-exclusion chromatography (SEC) step.

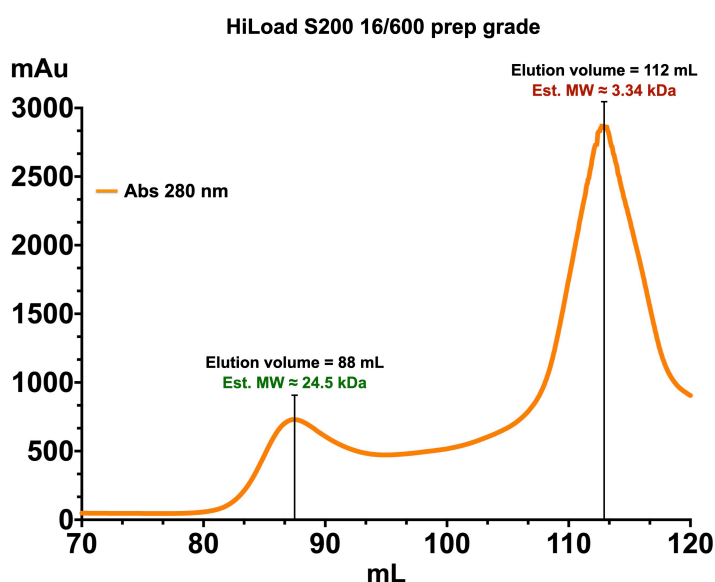


Figure 3. Size-exclusion chromatography of full length MtDciA. The gel-filtration analysis of a pure and concentrated MtDciA specimen produced a chromatogram showing two absorbance peaks at 280 nm (orange): the first seems to be related to a fully folded MtDciA, with an estimated molecular weight (MW) in accordance with the expected one (green), while the second corresponds to an unexpectedly small dimension (red).

In particular, as depicted in Figure 2, when DciA is injected into a 120 mL SEC column, two chromatographic peaks are produced. Based on the calibration curve, the estimated molecular weight of the sample corresponding to the first peak (~ 24.5 kDa) fits with the expected one (20.9 kDa). On the other side, the second peak can be related only to a protein with a molecular weight of about 3.34 kDa, giving rise to questions on *MtDciA*'s shape and dynamic behavior.

The SAXS analysis was performed at BM29 at the European Synchrotron Radiation Facility (ESRF) in Grenoble. Ten frames were acquired on a specimen containing *MtDciA* at the concentration of 233.4 μM , accompanied by 10 frames acquired on the clear protein buffer, necessary for the mandatory 'buffer subtraction' operation during the data processing step. The analysis of the averaged and subtracted signal operated in Scatter IV¹¹ generated the canonical Guinier plot, Kratky plot and Pair Distribution function $P(r)$ (Figure 4).

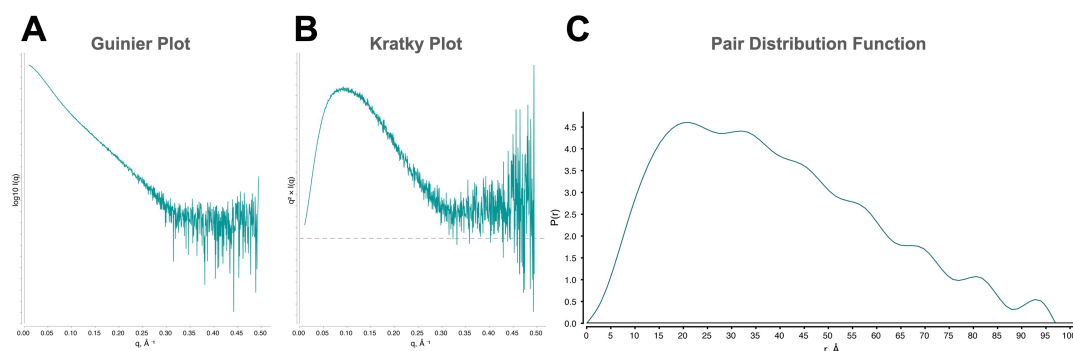


Figure 4. *MtDciA* is monodispersed, non-aggregated and folded in solution. Analysis of SAXS data obtained on full length *MtDciA*. **A)** Aggregation state identified by the Guinier plot, reporting $\log_{10}$ SAXS intensity versus scattering vector q . **B)** Folding state represented through the Kratky plot. **C)** Interatomic distance distribution or pair distribution function $P(r)$ for describing the solution shape of the protein.

Based on the linearity of the Guinier plot (Figure 4A) at high scattering intensity value (q , $\text{\AA}^{-1}$), we inferred how *MtDciA* is not aggregated and does not display interparticle interference under the condition tested. Then, the gaussian profile of the Kratky plot (Figure 4B) helped us in confirming that the protein is correctly

folded. Finally, by relating the position of the electrons constituting the scattering specimen¹², the pair distribution function $P(r)$ (Figure 4C) indicates that DciA features a globular shape, that could be accompanied by a rod-like tail.

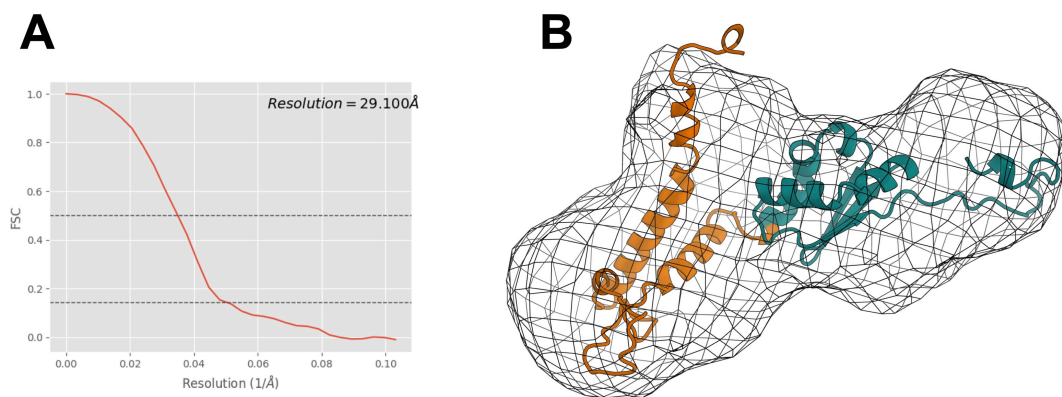


Figure 5. The RoseTTAFold predicted structure of MtDciA fits in the *ab-initio* electron density map determined from SAXS data. A) Fourier Shell Correlation (FSC) obtained in DENSS, for the estimation of the electron density map resolution. **B)** MtDciA RoseTTAFold predicted model fitted in the calculated electron density map from SAXS data.

Furthermore, this analysis laid the foundation for determining the electron density map using the *ab-initio* reconstruction tool in DENSS¹³ (Figure 5), which recalls most of the features described above. Due to lack of experimentally solved 3D structure of helicase loaders, we leveraged on RoseTTAFold deep learning method by Baker Lab's Robetta¹⁴ to predict MtDciA's one, then fitted into the electron density map using UCSF ChimeraX. As shown in Figure 5B, the quality of SAXS data was high enough to produce a reliable map, suitable for correctly docking DciA's calculated 3D assembly in a statistically relevant fashion.

Taken together, this data point out that our preparation of the mycobacterial loader renders a monodispersed, non-aggregated specimen displaying the expected tertiary folding. Hence, the source of the unusual retention time within the gel-filtration column could be attributed to the distinctive 'stickiness' of the intrinsically disordered N-terminal region. Indeed, SEC columns

allow to segregate protein conformers having different physical properties in solution, for instance compact or elongated shapes¹⁵.

Next, by taking advantage of the solution features and low-resolution structural details acquired on MtDciA through SEC and SAXS analyses, we turned to high resolution NMR spectroscopy in order to determine the three-dimensional structure of the helicase loader at atomic detail. The first step towards this procedure was the purification of isotopically enriched recombinant DciA, containing either Nitrogen 15 (^{15}N), for acquiring bidimensional spectra, or both Carbon 13 and Nitrogen 15 ($^{13}\text{C}^{15}\text{N}$), aimed for three-dimensional NMR analyses. We successfully prepped the abovementioned samples by following the expression and purification steps detailed in the Materials and Methods section below.

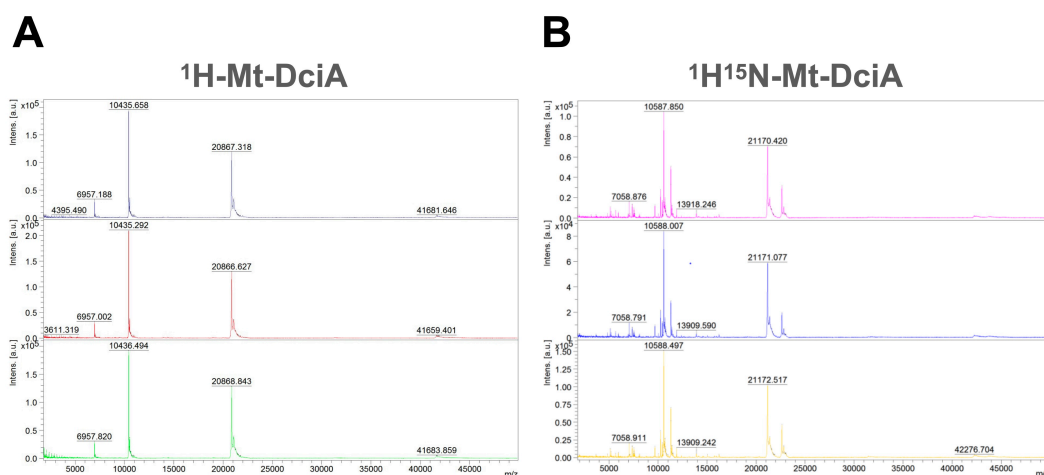


Figure 6. MtDciA properly incorporated Nitrogen 15 isotope in the optimized expression procedure. The triplicate MALDI-ToF analysis on the mycobacterial helicase loader confirmed a molecular weight of 20.87 kDa for ^1H -DciA (A) and for 21.17 kDa of $^1\text{H}^{15}\text{N}$ -DciA (B).

Secondly, we applied mass spectrometry to check on the correct incorporation of ^{15}N and confirm that the protein expression protocol was effective. According to the theoretical calculation, $^1\text{H}^{15}\text{N}$ -DciA should weight 303.663 Da more than the non-labelled counterpart. Indeed, the MALDI-ToF analysis initially

conducted on ^1H -DciA computed a molecular weight of 20,868.843 Da (Figure 6A), whereas the same analysis on $^1\text{H}^{15}\text{N}$ -DciA provided a value of 21,171.077 Da (Figure 6B). Thus, the calculated difference of 302.234 Da was enough for proving that Nitrogen 15 was accurately incorporated during the isotopically enriched bacterial growth procedure.

Being certain of the quality of our sample, we moved on with the acquisition of 2D NMR spectra. Firstly, with a 2D $^1\text{H}^{15}\text{N}$ HSQC at $T = 290\text{ K}$ and $\text{pH} = 8.0$ we were able to identify only 151 signals out of 176 expected (prolines are not visible in this test, and DciA contains 11 of them). By lowering the temperature to 278K, we recovered the missing signals, achieving 100% of completeness (Figure 7).

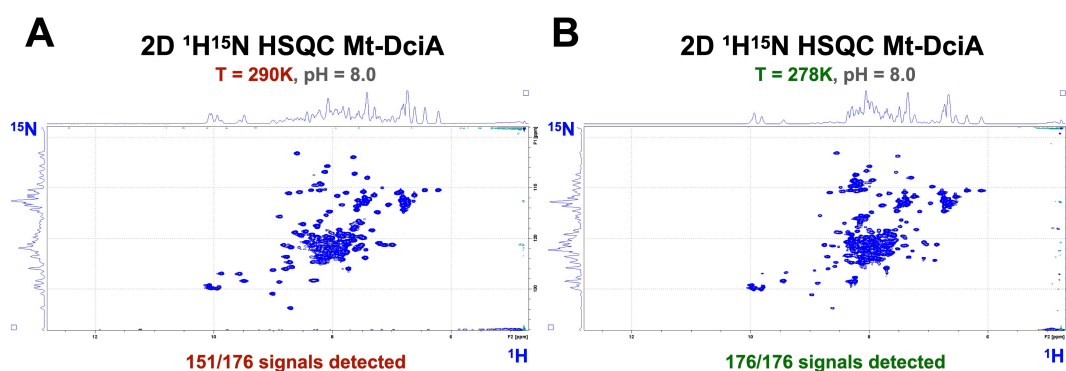


Figure 7. 2D $^1\text{H}^{15}\text{N}$ HSQC spectra of MtDciA acquired at 700 MHz at different temperatures. The signal detection completeness in 2D NMR experiments was achieved by lowering the temperature of analysis. Indeed, at 290K (A) 25 peaks were invisible, while at 278K (B) the missing signals were fully recovered.

This paved the way for performing a set of 3D NMR experiments, namely $\text{HN}(\text{CO})\text{CA}$, HNCA , $\text{HN}(\text{CA})\text{CO}$, HNCO , $\text{HN}(\text{CO})\text{CACB}$, HNCACB . The combination of these trials let us unambiguously assign the 61% of DciA's residues. Notably, the sequential walk for the protein backbone assignment identified amino acids belonging to the loader's IDRs. Therefore, the N-terminal α -helix together with the adjacent coils and the C-terminal loop have been characterized (Figure 8). In parallel, to confirm correct assignment in relation to the DciA's secondary

structure, Talos+¹⁶ was used to check on the relation between ¹³C-¹⁵N-¹H chemical shift and the backbone torsion angle.

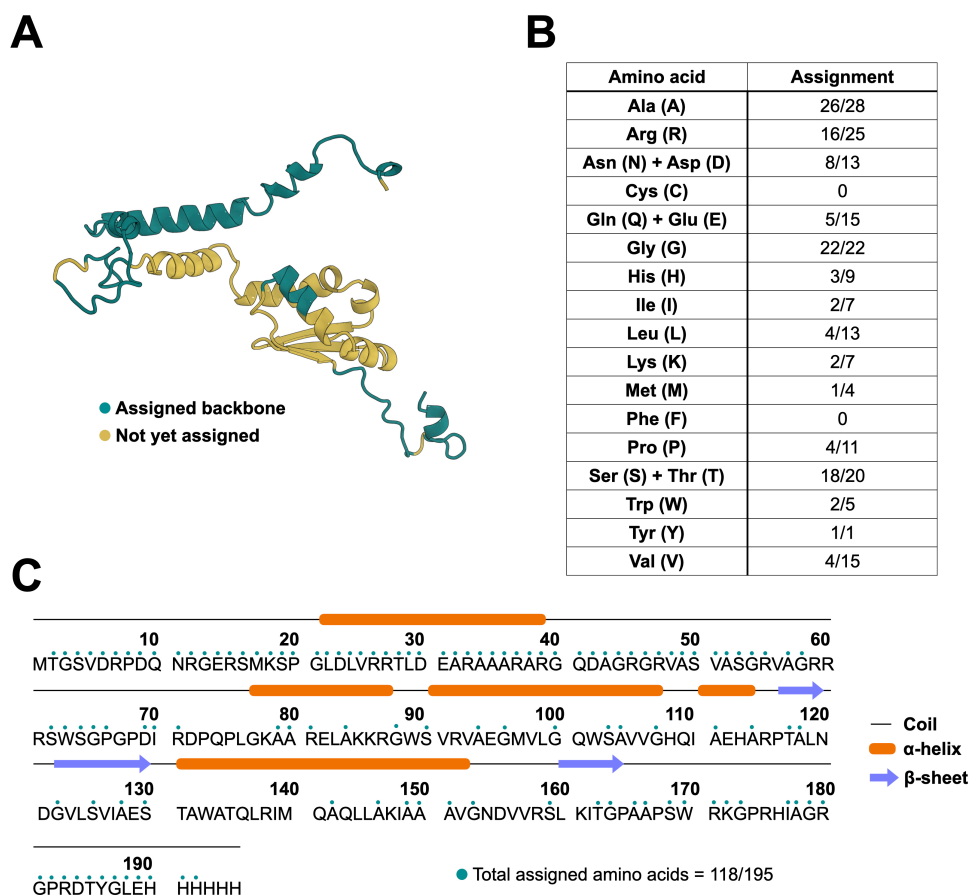


Figure 8. The more flexible domains of MtDciA have been characterized through solution NMR spectroscopy. The 3D NMR experiments unambiguously identified 118/195 MtDciA residues. A) RoseTTAFold calculated structure of MtDciA. In green, the assigned amino acids. In yellow, the protein backbone yet to be described. B) Table reporting the assigned amino acids. C) Primary sequence of full-length C-terminal His-tagged MtDciA with highlighted secondary structure features (coil, α -helix, β -sheet) and assigned residues (green dots).

As of today, based on the available results, we are not able to completely build the 3D structure of the mycobacterial loader. Ongoing bidimensional experiments aim to enhance the signal detection. By lowering the pH from 8 to 6 and increasing the temperature from 278K to 295K, α -helices and β -strands signal intensity soared from the 34% to the 58%. Hopefully, setting up 3D analyses based on such values, will improve and complete the protein backbone assignment.

MATERIALS AND METHODS

Plasmids and Protein Expression

For detailed plasmid design on *MtDciA*-FL see the Materials and Methods section of Chapter 3.

MtDciA for SAXS studies

Protein expression protocol for *MtDciA*-FL aimed for SAXS studies followed the steps detailed in the Materials and Methods section of Chapter 3.

MtDciA for NMR studies

To express ^{15}N -labelled C-terminal His-tagged full-length *MtDciA*, BL21(DE3) *E. coli* cells transformed with the engineered construct were grown at first on 2xYT/agar Petri plates supplemented with Ampicillin (50 $\mu\text{g}/\text{mL}$), over-night at 37°C, in order to obtain bacterial colonies. Then, a 9-hours pre-culture was conducted using 6 mL of 2xYT medium containing Ampicillin (50 $\mu\text{g}/\text{mL}$), incorporating a single *E. coli* colony grown on the solid medium. At the end of the set time, 2 mL of liquid culture were poured into 50 mL of M9 minimal medium (6 g/L Na_2HPO_4 , 3 g/L KH_2PO_4 , 1 g/L $^{15}\text{NH}_4\text{Cl}$, 4 g/L glucose, 2 mM MgSO_4 , 100 μM CaCl_2), comprising Ampicillin (50 $\mu\text{g}/\text{mL}$), and the cell growth was allowed once again for 12 hours at 37 °C under shaking. Next, this was used as starter of a scaled-up growth consisting of additional 950 mL of M9 medium for a total of 1 L. BL21(DE3) cells were grown at 37 °C to an $\text{OD}_{600} \approx 0.8$ and IPTG was subsequently included at 0.5 mM for culture induction. Expression of singly-labelled *MtDciA* was allowed for 4 hrs at 30 °C under shaking. At the end of the procedure, the bacterial biomass was harvested and stored at -80 °C.

The doubly-labelled protein for 3D NMR analysis, namely $^{13}\text{C}^{15}\text{N}$ -MtDciA-FL, was expressed following the same protocol except that the glucose used to prepare M9 medium was $^{13}\text{C}_6$ -enriched.

Protein purification

MtDciA for SAXS studies

The preparation of MtDciA for SAXS studies followed the procedure described in Materials and Methods section of Chapter 3, except for the purification buffers composition. Notably, the IMAC was based on Buffer A containing 20 mM Na-HEPES pH = 7.5, 500 mM NaCl, 10 mM imidazole, 2 mM β -mercaptoethanol, 5% glycerol. Ni-NTA resin was washed with Buffer A supplemented with 50 mM imidazole, whereas the recombinant protein was eluted with Buffer A containing 300 mM imidazole. Gel-filtration buffer was composed of 20 mM HEPES pH = 7.5, 300 mM NaCl, 2 mM β -mercaptoethanol, and 5% glycerol.

MtDciA for NMR studies

The purification protocol to obtain both the singly-labelled (^{15}N) and the doubly-labelled ($^{13}\text{C}^{15}\text{N}$) MtDciA followed the same steps. The bacterial cell pellet obtained from 1 L of M9 minimal medium was thawed and resuspended in Buffer A consisting of 50 mM potassium phosphate buffer pH = 7.5, 300 mM NaCl, and 10 mM imidazole, including protease inhibitor cocktail (ThermoScientific p/n 78439) at a volume ratio of 100 μL per 40 mL of lysis buffer. Sonication was used to disrupt *E. coli* cells (total sonication ON time: 6 mins), and the obtained soluble fraction was incubated for 90 minutes with Ni^{2+} -nitrilotriacetic acid (Ni-NTA) beads, previously equilibrated in Buffer A. Batch affinity chromatography was performed by applying 10-column-volume wash using Buffer A containing 40 mM imidazole, and the recombinant protein was eluted with Buffer A plus 300 mM imidazole.

Obtained sample was concentrated by ultrafiltration and injected into a HiLoad 16/600 Superdex 200 prep grade column equilibrated in 50 mM phosphate buffer pH = 6.0, and 150 mM NaCl, compatible with NMR spectroscopy analyses. Fractions containing highly pure and concentrated labelled MtDciA-FL (^{15}N or $^{13}\text{C}^{15}\text{N}$) were concentrated up to 6 mg/mL. Before stocking at $-80\text{ }^{\circ}\text{C}$, 1 mM EDTA and 1 $\mu\text{L/mL}$ of protease inhibitor cocktail were added to the preparation.

REFERENCES

1. Mann, K. M. et al. Rv0004 is a new essential member of the mycobacterial DNA replication machinery. *PLoS Genetics* **13**, (2017).
2. Chase, J., Berger, J. & Jeruzalmi, D. Convergent evolution in two bacterial replicative helicase loaders. *Trends Biochem. Sci.* **47**, 620–630 (2022).
3. Fernandez, A. J. & Berger, J. M. Mechanisms of hexameric helicases. *Crit. Rev. Biochem. Mol. Biol.* **56**, 621–639 (2021).
4. O'Shea, V. L. & Berger, J. M. Loading strategies of ring-shaped nucleic acid translocases and helicases. *Curr. Opin. Struct. Biol.* **25**, 16–24 (2014).
5. Davey, M. J. & O'Donnell, M. Replicative helicase loaders: Ring breakers and ring makers. *Current Biology* **13**, (2003).
6. Brézellec, P., Vallet-Gely, I., Possoz, C., Quevillon-Cheruel, S. & Ferat, J.-L. DciA is an ancestral replicative helicase operator essential for bacterial replication initiation. *Nat. Commun.* **7**, 13271 (2016).
7. Davey, M. J., Fang, L., McInerney, P., Georgescu, R. E. & O'Donnell, M. The DnaC helicase loader is a dual ATP/ADP switch protein. *The EMBO Journal* **21**, 3148–3159 (2002).
8. Ozaki, S., Wang, D., Wakasugi, Y., Itani, N. & Katayama, T. The *Caulobacter crescentus* DciA promotes chromosome replication through topological loading of the DnaB replicative helicase at replication forks. *Nucleic Acids Research* **50**, 12896–12912 (2022).
9. Chan-Yao-Chong, M., Marsin, S., Quevillon-Cheruel, S., Durand, D. & Ha-Duong, T. Structural ensemble and biological activity of DciA intrinsically disordered region. *Journal of Structural Biology* **212**, (2020).
10. Blaine, H. C., Burke, J. T., Ravi, J. & Stallings, C. L. DciA Helicase Operators Exhibit Diversity across Bacterial Phyla. *J. Bacteriol.* **204**, e00163-22 (2022).
11. Tully, M. D., Tarbouriech, N., Rambo, R. P. & Hutin, S. Analysis of SEC-SAXS data via EFA deconvolution and Scatter. *J. Vis. Exp.* (2021) doi:10.3791/61578.
12. Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **40**, 191–285 (2007).
13. Grant, T. D. Ab initio electron density determination directly from solution scattering data. *Nat. Methods* **15**, 191–193 (2018).
14. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
15. Uversky, V. N. Size-Exclusion Chromatography in Structural Analysis of Intrinsically Disordered Proteins. *Methods Mol. Biol.* **896**, 179–194 (2012).

16. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213–223 (2009).

Chapter 5

DISCUSSION, CONCLUSIONS, AND FUTURE PERSPECTIVES

DISCUSSION, CONCLUSIONS, AND FUTURE PERSPECTIVES

In all domains of life, the DNA replication is a fundamental biological pathway, essential for the survival of every organism, owing to the necessity of propagating the genetic material. Given the utmost complexity of the system, the biochemical processes underlying this event are strictly regulated and finely tuned. During the initiation phase of the DNA replication, peculiar enzymes, generally termed replicative helicases but specifically called DnaB within the bacterial world, are required to unwind the genomic double-stranded DNA and expose single strands, subsequently exploited by the DNA polymerase as template, in order to produce newborn DNA chromosomes¹. DnaB proteins feature a unique hexameric ring-shape like structure, hence their assembly onto the bacterial origin of replication underlies distinct mechanisms, and relies on the concerted action of replication factors, among them the initiator protein and the helicase loader. In particular, most pathogenic bacteria, including *Mycobacterium tuberculosis* and *Vibrio cholerae*, display the ancestral DciA loader², that proved to possess key structural and functional features of specific proteins associated with DnaB and operating during the DNA replication initiation, namely the helicase loaders DnaC from *E. coli* and λ P from phage λ . However, despite the central role of this event, structural and kinetic details in DciA-dependent bacteria still needs to be outlined.

In the context of a clinical world where infectious diseases like Tuberculosis are becoming more and more difficult to treat, due to emerging hyper resistant bacterial strains causing the surge of the mortality rate, especially in low-income countries, the seek for unprecedented pharmacological targets turns out to be imperative³. The biochemistry and structural biology unit where I have been working since the last three years is committed to dig into some of the biological

events that take part in the DNA metabolism, particularly the pyrimidine synthesis, the DNA repair and the DNA replication initiation, the latter being the topic of my PhD project. By doing so, we intend to discover new active compounds being employed in the treatment of the Tuberculosis disease with an “anti-evolution” approach⁴, as thoroughly described in Chapter 1. This would turn mycobacterial bacilli from being drug resistant to drug susceptible to already-available antibiotics. Our approach leverages on a wide number of experimental techniques, spanning from the applied biochemistry, to the biophysics and structural biology, in an integrative framework.

In Chapter 3, we shed new light on the helicase loading and activation pathway in the human pathogenic bacteria *Vibrio cholerae* (Vc) and *Mycobacterium tuberculosis* (Mt). Our cryo-EM structure of the ATP γ S-bound VcDnaB-ssDNA complex displays features such that it represents an alternative translocation state than the previously determined GDP•AlF₄-bound DnaB, providing proofs on the nucleotide-dependent conformational switching, reinforcing the proposed hand-over-hand model occurring during the DNA unwinding⁵. However, the expulsion of DciA from the DnaB-ssDNA complex remains unveiled. Our integrative experimental approach proved that up to three copies of loader tightly interact with the helicase, leading to the suppression of its ATPase activity, thus controlling DnaB’s activation in the early stages of replication initiation, as predicted and expected. Alongside, thanks to a FRET-based optimized platform, we were able to measure the DciA-dependent DNA fork unwinding activity operated by DnaB, and then we turned to a specifically designed DNA bubble to capture the loading-dependent helicase translocation onto a physiological substrate. Altogether, the data produced and published during my PhD program prove that DciA loads DnaB onto single-strand DNA using the ring-opening mechanism.

Furthermore, Chapter 4 contains unpublished results, and, in Part 1, it narrates about the prosecution of our single-particle cryo-EM-based studies on the *Vibrio cholerae* DnaA-DnaB-DciA protein complex. Indeed, two different configurations have been described, a DNA-free one and a DNA-bound one, supplying never seen before binding modes of KH domain of both the initiator protein and the helicase loader⁶. In particular, we addressed a possible DNA-dependent protein network, in which DciA and DnaA engage with DnaB in different ratio based on the presence of nucleic acid as fourth player in this system. As previously mentioned, the consolidation of these preliminary data calls for additional experiments, and we aim to apply native Mass Spectrometry, Cross-Linking Mass Spectrometry, and Surface Plasmon Resonance, at least, to further support our research, with the ultimate goal of populating more and more in detail the DNA replication initiation model (Chapter 3, Figure 8).

At the same time, Chapter 4 – Part 2 reports results obtained in ancillary projects, supporting our main evidence, and letting us strengthen our overview on the main proteins involved in the DNA replication initiation process. On one side, we proved how fundamental the Docking Helix-Linker Helix (DH-LH) interface is even at the level of the mycobacterial DnaB⁷. Notably, by means of co-purification techniques on recombinant proteins, we further sustained that the interaction between the replicative helicase and the helicase loader finds its hotspot exactly between DnaB's DH-LH element and DciA's lasso/hook-like domain, making it an important druggable target. On the other side, leveraging on techniques like Small-Angle X-Ray Scattering (SAXS) and Nuclear Magnetic Resonance (NMR), we investigated the three-dimensional structure of MtDciA, obtaining results in line with the expectations set, taking into account the AI-based 3D predictions.

At present days, three different yet parallel branches shape the outlook of this research project:

- a) Thanks to a strong collaboration with the medicinal chemistry group of Dr. Giannamaria Annunziato from the University of Parma, we are on our way to testing peptidomimetics against the DnaB-DciA interaction, both in the *Vibrio cholerae* and the *Mycobacterium tuberculosis* systems. They have been designed based on the isolated Lasso domain of both helicase loaders, and we plan to exploit already-optimized Surface Plasmon Resonance and FRET-based DNA unwinding protocols to collect proofs on their biological activity. Hopefully, this will lay the foundation for our drug discovery journey against the DNA replication initiation phase in human pathogens.
- b) Single-particle cryo-EM studies will be brought onward especially on the well-behaving Vc proteins. Specifically, we want to capture more and more intermediate states of the helicase loading and translocation mechanism by taking advantage of different non-hydrolyzable ATP analogs, such as ADP•AlF₄ and AMP-PMP, owing to how encouraging and trustworthy data ATPγS provided (Chapter 3). In parallel, the *in-vitro* reconstitution's complexity of the proteins assembling at the origin of replication will be increased step-by-step, through the inclusion of other replication initiation factors like DnaG, DnaK, and DnaJ. During the cryo-EM sample preparation, different ratios of both proteins, DNA substrates and nucleotides will be tested, to seize novel 3D structures of such protein-DNA macro complex.
- c) As reported in our published data contained in Chapter 3, *Mycobacterium tuberculosis* DnaB was found mostly inactive from the catalytic point of view, letting us collect protein-protein interaction data from one side, but

failing each and every enzymatic test on the other. Therefore, as of today, we have not yet confirmed that DciA possess equivalent features also in the mycobacterial system, putting on hold the parallelism with Vc proteins and, on top of it, our willingness of supporting the convergent evolution theory⁸. We hope to address this shortage of kinetic data by turning to *M. tuberculosis* homologs. Firstly, we aim to test *Mycobacterium smegmatis* DnaB and DciA. This is a well-known model organism that historically provided strong inference-based evidence on how antitubercular drugs act, owing to its utmost similarity with *M. tuberculosis* notwithstanding the absence of virulence. Then, we will analyze *Mycobacterium chelonae* helicase and loader, a member of the MTBC complex described as non-pathogenic and fast replicating. Nevertheless, it caught our attention given that its DnaB gene does not contain any intein, making it differ from most mycobacteria. In fact, usually the helicase gene is expressed, transcribed, and translated producing an inactive protein. The intein's enzymatic cleavage, under the control of different environment conditions like oxidative stress, is key to DnaB's action⁹. By undergoing the 'latency stage' during its life cycle, we predict that *M. tuberculosis* proteins activity is subjected to regulations that, at present, we are not yet aware of. Optimistically, leveraging on fast replicating mycobacteria will fill the gap in this biological field.

REFERENCES

1. Ditse, Z., Lamers, M. H. & Warner, D. F. DNA Replication in *Mycobacterium tuberculosis*. *Microbiol. Spectr.* **5**, 10.1128/microbiolspec.tb2-0027-2016 (2017).
2. Brézellec, P., Vallet-Gely, I., Possoz, C., Quevillon-Cheruel, S. & Ferat, J.-L. DciA is an ancestral replicative helicase operator essential for bacterial replication initiation. *Nat. Commun.* **7**, 13271 (2016).
3. WHO. Global Tuberculosis Report. (2023).
4. Warner, D. F., Rock, J. M., Fortune, S. M. & Mizrahi, V. DNA Replication Fidelity in the *Mycobacterium tuberculosis* Complex. *Adv. Exp. Med. Biol.* **1019**, 247–262 (2017).
5. Itsathitphaisarn, O., Wing, R. A., Eliason, W. K., Wang, J. & Steitz, T. A. The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell* **151**, 267–277 (2012).
6. Valverde, R., Edwards, L. & Regan, L. Structure and function of KH domains. *FEBS J.* **275**, 2712–2726 (2008).
7. Cargemel, C. et al. The LH-DH module of bacterial replicative helicases is the common binding site for DciA and other helicase loaders. *Acta Crystallographica Section D: Structural Biology* **79**, 177–187 (2023).
8. Chase, J., Berger, J. & Jeruzalmi, D. Convergent evolution in two bacterial replicative helicase loaders. *Trends Biochem. Sci.* **47**, 620–630 (2022).
9. Kelley, D. S. et al. Mycobacterial DnaB helicase intein as oxidative stress sensor. *Nature Communications* **9**, (2018).

LIST OF PUBLICATIONS

*Nicholas Gao, ***Daniele Mazzeletti**, Adele Peng, Paul Dominic B. Olinares, Castrese Morrone, Andrea Garavaglia, Nourelhoda Gouda, Serge Tsoy, Albert Mendoza, Ahammad Chowdhury, Antonio Cerullo, Hrutvik Bhavsar, Franca Rossi, Menico Rizzi, Brian T. Chait, Riccardo Miggiano, and David Jeruzalmi

“DnaB and DciA: Mechanisms of Helicase Loading and Translocation on ssDNA”

Nucleic Acid Research. **2 October 2024** – DOI: 10.1101/2024.11.09.622779
(under review)

Eleonora Panfili, Sarah Jane Rezzi, Annalisa Adamo, **Daniele Mazzeletti**, Alberto Minassi, Riccardo Miggiano, Silvia Fallarini, Sara Ambrosino, Alice Coletti, Pasquale Molinaro, Michele Milella, Salvatore Paiella, Antonio Macchiarulo, Stefano Ugel, Tracey Piralì, and Maria Teresa Pallotta

“Identification of a Compound Inhibiting Both the Enzymatic and Nonenzymatic Functions of Indoleamine 2,3-Dioxygenase 1”

ACS Pharmacology & Translational Science. Vol 7, 10 3056-3070. **12 September 2024** – DOI: 10.1021/acsptsci.4c00265

Edoardo L. M. Gelardi, Diego Caprioglio, Giorgia Colombo, Erika Del Grosso, **Daniele Mazzeletti**, Daiana Mattoteia, Stefano Salamone, Davide M. Ferraris, Eleonora Aronica, Giulia Nato, Annalisa Buffo, Menico Rizzi, Lorenzo Magrassi, Alberto Minassi, and Silvia Garavaglia

“Curcumin-based-fluorescent probes targeting ALDH1A3 as a promising tool for glioblastoma precision surgery and early diagnosis”

Communications Biology. Vol 5, 1 895. **1 September 2022** – DOI: 10.1038/s42003-022-03834-7

CURRICULUM VITAE

Date of birth: [REDACTED]

E-mail: daniele.mazzoletti@uniupo.it

ORCID: 0000-0002-0711-8269

LinkedIn: www.linkedin.com/in/daniele-mazzoletti/

WORK EXPERIENCE

01/05/2025 – Present

Project Manager @ IXTAL Srl – Understanding Proteins, Novara (NO), Italy.

04/11/2024 – 29/04/2025

Post-Doc Researcher (*Assegnista di Ricerca*) – University of Piemonte Orientale, Department of Pharmaceutical Sciences. Novara (NO), Italy.

Education field: Biological Science, BIO/10 – Biochemistry

Title: “Dynamics of recruitment of macromolecular complexes active on the mycobacterial orisome: mechanistic details and target validation”

Principal Investigator of the Project: Professor Riccardo Miggiano

Financial Support: Fondo PRIN 2022 (Progetto 2022LCN738)

02/11/2021 – 31/10/2024

PhD Student in Drug Innovation (XXXVII cycle) – University of Piemonte Orientale, Department of Pharmaceutical Sciences. Novara (NO), Italy.

Lab Head: Professor Menico Rizzi

Supervisor: Professor Riccardo Miggiano

*By taking advantage of single-particle cryo-Electron Microscopy (cryo-EM), native Mass Spectrometry, Surface Plasmon Resonance (SPR) and FRET-based analyses, we aim to provide new structural and biochemical insights on the DNA replication initiation process in the pathogenic bacteria *Mycobacterium tuberculosis* and *Vibrio cholerae*.*

Applied Techniques: recombinant protein expression (*E. coli* strains) and purification (FPLC system), spectrophotometry, spectrofluorimetry, Surface Plasmon Resonance, Differential Scanning Fluorimetry, Dynamic Light Scattering, Small-Angle X-ray Scattering, Nucleic Magnetic Resonance, native Mass Spectrometry, single-particle cryo-Electron Microscopy.

02/11/2021 – 31/10/2024

PhD Student in Drug Innovation (XXXVII cycle) – University of Piemonte Orientale, Department of Pharmaceutical Sciences. Novara (NO), Italy.

Lab Head: Professor Menico Rizzi

Supervisor: Professor Riccardo Miggiano

By taking advantage of single-particle cryo-Electron Microscopy (cryo-EM), native Mass Spectrometry, Surface Plasmon Resonance (SPR) and FRET-based analyses, we aim to provide new structural and biochemical insights on the DNA replication initiation process in the pathogenic bacteria Mycobacterium tuberculosis and Vibrio cholerae.

17/05/2023 – 10/02/2024, 06/07/2024 – 19/09/2024

Visiting PhD Student – The City College of New York (CUNY), Center for Discovery and Innovation. New York, New York, USA.

Supervisor: Professor David Jeruzalmi

Based on the PhD program, as a visiting PhD student I had the chance of exploiting single-particle cryo-EM at the New York Structural Biology Center (NYSBC) facility, to solve and investigate the 3D structure of protein-protein and protein-DNA complexes involved in the DNA replication initiation process both in Mycobacterium tuberculosis and Vibrio cholerae.

01/09/2021 – 31/10/2021

Post Graduate Fellow – University of Piemonte Orientale, Department of Pharmaceutical Sciences. Novara (NO), Italy.

Lab Head: Professor Menico Rizzi

Supervisor: Professor David Jeruzalmi

Design and optimization of expression and purification protocols for mycobacterial replicative helicase DnaB and helicase loader DciA.

11/05/2020 – 22/04/2021

Junior Scientist @ IXTAL Srl – Understanding Proteins, Novara (NO), Italy.

Supervisors: Professor Riccardo Miggiano, Professor Davide Ferraris

R&D of biochemical projects to reach the company prefixed targets, involving:

- *Expression and purification of recombinant proteins and antibodies fragments.*
- *High-throughput screening of potential enzyme inhibitors via spectrophotometry and spectrofluorimetry tests.*
- *Protein-protein and protein-ligand interaction analyses via SPR.*
- *Enzymatic quantitation of food components in complex food matrices using fluorimetry approaches.*

27/08/2018 – 27/02/2019

Pharmacist Intern – Farmacia Europea Snc di De Paola Paola Maria – Bareggio (MI), Italy.

13/05/2018 – 18/05/2018

Undergraduate Fellow – New York Structural Biology Center. New York, New York, USA.

In-depth analysis of state-of-the-art structural and molecular biology techniques: high-throughput protein purification, cryo-EM, NMR. Fellowship program in collaboration with the research team of Professor Filippo Mancia, Columbia University, New York.

EDUCATION

15/01/2021 – 12/12/2021

2nd level University Master in Clinical Trials of Drugs in Internal Medicine, Hematology and Oncology.

University of Pisa, Department of Clinical and Experimental Medicine. Pisa (PI), Italy.

Thesis title: “The Quality Management in Clinical Trials: an approach based on the identification, the evaluation and the management of risk”

Supervisor: Professor Romano Danesi

Co-supervisor: Dr. Danila Montinari

Final grade: "Ottimo"

01/10/2014 – 16/10/2019

Master's Degree in Medicinal Chemistry and Technologies (LM-13).

University of Piemonte Orientale, Department of Pharmaceutical Sciences. Novara (NO), Italy.

Experimental thesis title: "Characterization of fluorescent probes for human recombinant proteins targeting ALDH1A1, ALDH1A2, and ALDH1A3"

Supervisor: Professor Silvia Garavaglia

Co-supervisor: Dr. Edoardo Gelardi, PhD

Final grade: 110/110 cum laude and recommendation of publication

SCIENTIFIC CONTRIBUTION

Publication of Scientific Articles

Gao N, Mazzeletti D et al. "DnaB and DciA: Mechanisms of Helicase Loading and Translocation on ssDNA". *Nucleic Acid Research*. 02 Oct 2024, doi: 10.1101/2024.11.09.622779. Manuscript under review.

Panfili, Eleonora et al. "Identification of a Compound Inhibiting Both the Enzymatic and Nonenzymatic Functions of Indoleamine 2,3-Dioxygenase 1". *ACS Pharmacology & Translational Science*. Vol. 7,10 3056-3070. 12 Sep. 2024, doi:10.1021/acspstsci.4c00265

Gelardi, Edoardo L M et al. "Curcumin-based-fluorescent probes targeting ALDH1A3 as a promising tool for glioblastoma precision surgery and early diagnosis". *Communications biology*. Vol. 5,1 895. 1 Sep. 2022, doi:10.1038/s42003-022-03834-7

Conferences Participation, Poster Sessions, and Abstract Publications

29/06/2024 – 03/07/2024 – Allianz MiCo – Milano Convention Center, Milano, Italy

48th FEBS Congress – Mining Biochemistry for Human Health and Well-Being

“Bacterial DNA replication initiation: structural and kinetic insights into the replicative helicase loading mechanism in *Mycobacterium tuberculosis* and *Vibrio cholerae*”

27/07/2023 – Aaron Davis Hall, The City College of New York, New York, USA

New York Structural Biology Discussion Group

Summer Session 2023

06/02/2023 – 08/02/2023 – European Synchrotron Radiation Facility, Grenoble, France

ESRF User Meeting 2023

“Mycobacterial DNA replication initiation: structural insights into the replicative helicase loading mechanism”

12/09/2022 – 16/09/2022 – Institut Pasteur, Paris, France

EMBO Workshop on Tuberculosis 2022

“Mycobacterial DNA replication initiation: biochemical insights into the replicative helicase loading mechanism”

18/05/2022 – 20/05/2022 – Pisa, Italy

Italian Society of Biochemistry and Molecular Biology Meeting – Proteine 2022

“Mycobacterial DNA replication initiation: biochemical insights into the replicative helicase loading mechanism”

Awarded with a travel fellowship for young researchers

23/09/2021 – 24/09/2021

61st Italian Society of Biochemistry and Molecular Biology Meeting – Virtual Edition

20/05/2021 – 21/05/2021

WebPro – Proteins on the Web 2021

1st online meeting of the group “Proteine” of the Italian Society of Biochemistry and Molecular Biology

SCHOOLS AND WORKSHOPS

03/06/2024 – 07/06/2023 – Novara-Milano, Italy

2nd Advanced School on Biophysical Methods for Protein-Protein and Protein-Ligand Interaction Analysis

Hosting Institutions:

Department of Pharmaceutical Sciences, University of Piemonte Orientale

Department of Biosciences, University of Milan

Institute for Molecular and Translational Cardiology (IMTC), IRCCS Policlinico San Donato

Member of the organizing committee

01/06/2024 – 08/06/2024 – EMBL Grenoble, France

EMBO Practical Course: Structural Characterization of Macromolecular Complexes

Awarded the EMBO Travel Grant

29/09/2022 – 30/09/2022 – Human Technopole, Milano, Italy

Cryo-Electron Microscopy Symposium

31/08/2021 – 04/09/2021

AIC International School of Crystallography on Fundamentals of Crystallography

The theory behind crystal structure solution

CERTIFICATIONS

18/07/2024 – New York Structural Biology Center, New York, New York, USA

Vitrobot Independence Test – CryoEM Merit Badges

Supervising Staff: Slavic Serbynovski

Sponsored by: National Center for CryoEM Access and Training (NCCAT), Stanford SLAC CryoEM Facilities, Pacific Northwest National Laboratory (PNNL)

Supported by: NIH Common Fund Merit Badge supplement to 3U24 GM129539-02

SCIENTIFIC PROJECTS PARTICIPATION

BANDO PRIN 2022

Project n.: 2022LCN738

Position: Participant

Principal Investigator: Professor Riccardo Miggiano

Title: “Dynamics of recruitment of macromolecular complexes active on the mycobacterial orisome: mechanistic details and target validation”

Institution: University of Piemonte Orientale

BANDO “Ricerca biomedica condotta da giovani ricercatori”

Fondazione Cariplo

Project n.: 2020_3589

Position: Participant

Principal Investigator: Professor Riccardo Miggiano

Title: “DNA replicative helicase loading in *Mycobacterium tuberculosis*: understanding the molecular details of bacterial replication in active tuberculosis”

Institution: University of Piemonte Orientale

SCIENCE ENGAGEMENT

November 2021 – Present

Supervision of Master's Students – University of Piemonte Orientale, Department of Pharmaceutical Sciences. Novara (NO), Italy.

Training undergraduate students from Pharmaceutical Biotechnologies and Chemistry and Pharmaceutical Technologies programs during the development of experimental Master's theses.

November 2021 – Present

Third Mission Activities – University of Piemonte Orientale, Department of Pharmaceutical Sciences. Novara (NO), Italy.

ChimicAmica – In collaboration with Fondazione Novara Sviluppo
Development of basic molecular biology experiments with primary school students.

Universitario Per Un Giorno

Meeting with high school students to share my personal experience and journey from the graduation day up to the present day.

November 2021 – November 2023

Representative of PhD Students – University of Piemonte Orientale, Department of Pharmaceutical Sciences. Novara (NO), Italy.

Doctoral School in Drug Innovation

DIGITAL SKILLS

OS: Windows, MacOS, Linux, UNIX.

Microsoft Office: Word, Excel, PowerPoint.

Apple Keynote.

Adobe Suite: Lightroom, Illustrator.

Statistics and scientific graphing tools: SigmaPlot, GraphPad Prism, ImageLab.

Softwares for structural determination of macromolecules by X-ray crystallography, cryo-EM and SAXS: cryoSPARC, CCP4i2, Phenix, ATSAS, ScatterIV, COOT, PyMol, Chimera, ChimeraX.

LANGUAGE SKILLS

Mother tongue: Italian

Other languages: English (FCE GRADE B2 – REF n° B2927644)

REFERENCE PEOPLE

Professor Riccardo Miggiano

riccardo.miggiano@uniupo.it

Professor David Jeruzalmi

dj@ccny.cuny.edu

Professor Menico Rizzi

menico.rizzi@uniupo.it

ACKNOWLEDGMENTS

The work conducted within Professor Riccardo Miggiano unit at DSF-UNIUIPO was supported by Fondazione Cariplo (grant n. 2020_3589), and PRIN 2022 (grant n. 2022LCN738).

The work conducted together with Professor David Jeruzalmi (The City College of New York) and collaborators at The Rockefeller University was funded by the National Science Foundation (DJ: MCB 1818255), and the National Institute of Health (BTC: P41 GM109824; P41 GM103314).

The cryo-electron microscopy work was conducted at the New York Structural Biology Center, involving the Simons Electron Microscopy Center (MEMC, NCCAT staff) and the National Resource for Automated Molecular Microscopy. This work was supported by Simons Foundation (grant n. SF349247), NYSTAR, NIH National Institute of General Medical Sciences (grant n. GM103310), Agouron Institute (grant n. F00316), and the National Institute of Health (grant n. OD019994 and n. RR029300).