

The sequence and structure of the Mfd protein dictate its role during virulence

Delphine Cormontagne

Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Samantha Samson

Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Thomas Marino

Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Solène Albert

Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Seav-Ly Tran

Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Sylvain Marthey

Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Gwenaëlle André

`gwenaelle.andre@inrae.fr`

Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Nalini Ramarao

Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

Article

Keywords: Mfd, virulence factor, *Bacillus cereus*, insect assays, sequence/structure relationships, molecular dynamics

Posted Date: August 25th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10606469/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

The sequence and structure of the Mfd protein dictate its role during virulence

Delphine Cormontagne¹⁺, Samantha Samson^{1,2+}, Thomas Marino², Solène Albert^{1,2}, Seav-Ly. Tran¹, Sylvain Marthey², Gwenaëlle André^{2+*} and Nalini Ramarao^{1+*}

¹Université Paris-Saclay, IHU Sepsis, INRAE, Micalis Institute, 78350, Jouy-en-Josas, France

²Université Paris-Saclay, IHU Sepsis, INRAE, MaIAGE, 78350, Jouy-en-Josas, France

*Principal corresponding author gwenaelle.andre@inrae.fr

*Correspondence should be also addressed to nalini.ramarao@inrae.fr

+these authors contributed equally to this work

Key words

Mfd; virulence factor; *Bacillus cereus*; insect assays; sequence/structure relationships; molecular dynamics.

Abstract

The Mfd (Mutation Frequency Decline) protein is a ubiquitous RNA-polymerase (RNAP)-associated factor, responsible for displacing RNAP, stalled onto DNA lesions, and for recruiting UvrA of the NER (Nucleic Excision Repair) complex, ensuing a boosting in DNA repair. Recent studies in *E. coli* and *B. subtilis* showed that Mfd regulates transcription dynamics, particularly in temporally controlled and difficult-to-transcribe regions, including loci encoding structured regulatory RNAs. Mfd was identified as an evolvability factor, promoting mutagenesis and accelerating the development of antibiotic resistance through hypermutation. Mfd is a virulence factor in *B. cereus* and *S. flexneri*, conferring resistance to the host nitrogen immune response.

In the context of combating antimicrobial drug resistance, Mfd emerges as a highly strategic and a compelling therapeutic target. Our knowledge of Mfd, framed by biochemical and structural studies, misses information of Mfd behavior in living cells under bacterial virulence. Here we combine computational with cell biology to infer on the plausible correlation between Mfd sequence and virulence phenotype. We demonstrate that this relationship

32 intertwines and recapitulates the phenotype between clinical and non-clinical bacterial strains
33 from *Bacillus cereus*.
34 Evidencing that molecular Mfd dictates bacterial virulence enhances our understanding of
35 bacterial adaptation and pave the ways to its specific inhibition.
36

Introduction

As the second causative agent of food-born outbreaks (FBO) in Europe [1], and reported late 2025 in France, the virulence of the Gram positive bacteria *Bacillus cereus* is essential to be mastered both at cellular and at molecular levels. Despite the ubiquitous distribution of *B. cereus* in the environment, the pathogenicity of a strain varies widely, ranging from non-virulent to highly pathogenic phenotypes [2]. Indeed, *B. cereus* has been documented to cause severe localized and systemic infections in humans, particularly affecting immunocompromised individuals [3] and preterm neonates [4]. With its persistence over long periods of time in the environment, *B. cereus* is also a significant contributor to nosocomial infections [4, 5]. Numerous virulence-associated factors have already been identified and characterized, including the cytotoxin K (CytK), the non-haemolytic enterotoxin (Nhe), the haemolysin BL (Hbl) [6-9], as well as enzymatic proteins such as InhA1 and CwpFM [10, 11]. In addition to these well-established determinants, other factors contribute to *B. cereus* virulence [12, 13]. Notably, we have identified the Mutation Frequency Decline (Mfd) protein as a decisive determinant during bacterial pathogenesis [14].

Mfd is a highly conserved protein implicated in DNA repair through its interaction with the RNA polymerase (RNAP)[15]. Mfd is recruited at DNA damages sites where the RNAP is stalled. The protein has two main roles, timely regulated, the first one is to push the RNAP forward to dislocate it from the DNA template [15], the second role is subsequently to facilitate the recruitment of UvrA, thereby initiating DNA repair *via* the nucleotide excision repair (NER) pathway [16].

In addition to its role in NER, Mfd can associate with RNAP in the absence of exogenous DNA damage, and *mfd*-deficient cells do not exhibit heightened sensitivity to DNA-damaging agents. Moreover, although Mfd is generally regarded as a genome-protective factor, it has paradoxically been shown to enhance mutagenesis under specific conditions, including replication–transcription conflicts and stationery-phase mutagenesis. Consistent with this dual role, Mfd has been implicated in the emergence of antibiotic resistance in *Campylobacter jejuni* and *Helicobacter pylori* [17, 18] , and suggested as an evolvability factor that promotes hypermutation in bacteria, thus accelerating the evolution of antimicrobial resistance (AMR) [19-21].

Furthermore, Mfd has been demonstrated to contribute to virulence in Gram-positive *B. cereus* as well as in Gram-negative *S. flexneri*, and to confer resistance to nitric oxide (NO) stress, a critical component of the host immune response during infection [14, 22, 23].

The structure of the Mfd protein has been solved [16, 24]. Mfd is a multi-functional nanomachine of 1150 amino-acid residues, organized into eight structural domains that assembled into six functional modules. At the N terminal of the protein, the D1a-D2-D1b region contains the D2 domain, which is homologous to the DNA repair protein UvrB and mediates the recruitment and binding of UvrA [16, 25]. Accordingly, it is called the UvrB-homology module. The D3 domain, whose function remains unknown, is often grouped with the D1a–D2–D1b region and collectively designated as the N-terminal domain (NTD) [16]. The D4 domain is the RNAP Interacting Domain (RID), which specifically binds to the RNAP β subunit protrusion [16]. The D5 and D6 domains, respectively called TD1 and TD2 (Translocation Domain), hold the ATPase translocation module. This module is structurally homologous to the superfamily 2 (SF2) helicases, such as RecG [16], and shares both ATP binding and DNA binding sites. Despite harbouring a helicase motif VI, this module acts as a double strand DNA translocase to push the stalled RNAP forward [26]. Finally, the D7 domain is considered as a regulatory domain, able to maintain Mfd into a repressed and inactive conformation, thereby preventing any ATPase activity. Indeed, within this arrangement, D2 and D7 domains are locked in together, rendering D2-mediated binding between Mfd and UvrA impossible [27]. Consequently, D7 prevents premature association of Mfd with the UvrA₂UvrB complex.

We have previously characterized that Mfd confers bacterial resistance to the host nitrogen, through Nitric Oxid (NO) response [14]. Following infections, production of reactive nitrogen species is a key step in the immune response [23]. NO induces DNA lesions in bacteria, thus limiting bacterial proliferation within hosts. Mfd contributes to the repair of NO-induced DNA damage and is therefore essential for bacterial survival under nitrosative stress. Accordingly, a *B. cereus* mutant, deficient in *mfd*, was shown to be severely impaired in its virulence capacity in both cellular and animal models [14]. As Mfd is widely conserved in the bacterial kingdom, these data highlight a mechanism possibly ever-present in bacteria to overcome the mutagenic host immune response.

In this study, we report a correlation between the sequence/structure of Mfd and its role in *B. cereus* virulence. By leveraging a panel of strains displaying diverse pathogenic potentials

alongside non-pathogenic phenotypes, we evaluated the contribution of Mfd to virulence and investigated whether distinct Mfd sequence signatures correlate with specific virulence profiles. Accordingly, our data provide a better understanding of the structure/function/dynamic relationship of Mfd on bacterial virulence. Of note, Mfd undergoes extensive conformational rearrangements throughout its functional cycle, transitioning from an auto-repressed state to an active ATPase motor configuration, and ultimately to a form capable of recruiting the Nucleotide Excision Repair machinery. In this context, we identified point substitutions, located between functional modules, that could modulate by hooking or smoothing these structural transitions, with consequences on Mfd-dependent virulence.

Results

The sequences of Mfd are clustered according to the virulence of their strains

The prevalence of the *mfd* gene was evaluated using Blast, in a panel of *B. cereus* strains, either coming from environmental sources or isolated from infected patients. In each of the 22 strains, a *mfd* gene coding for a Mfd protein sequence was found as one copy (not shown). Next, the 22 Mfd protein sequences were retrieved and aligned according to their identity using MAFFT (Fig 1, Suppl. Fig 2). Overall, the Multiple Sequence Alignment (MSA) shows 98% of sequence identity, including 57 substitutions along the 1150 residues, with neither insertion nor deletion. This corresponds to a mutation ratio of 0.048 (Fig 1). Despite their overall high sequence identity, the multiple sequence alignment (MSA) resolves the sequences into three principal clusters: one comprising Mfd variants from environmental strains (highlighted in green), a second grouping Mfd variants from clinical strains (orange), and a third set of three sequences (grey), originating from both clinical and non-clinical strains. These latter sequences occupy a so-called intermediate “twilight zone” position, reflecting their strong similarity to members at the boundaries of the two main groups.

This three-part clustering, observed despite the overall high sequence similarity, points to a putative association between molecular sequence of Mfd and strain-specific virulence phenotype. Notably, the Mfd proteins from strains S50 and S53 are positioned at opposite ends of the MSA, representing the environmental and clinical groups, respectively, and differ by a total of 28 amino acid substitutions. To investigate on the assumption of sequence-

phenotype correlation, we focused on and selected S50 and S53 for experimental and computational assessment of their virulence.

Mfd is implicated in B. cereus virulence in vivo

First, the virulence of strains S50 and S53 was assessed in the insect *Bombyx eri* infection model (Fig 2A). At similar CFU injected, the infection with the S50 strain led to over 90% of insect survival, whereas infection with the S53 strain led to insect survival dropped down to 28%. Thus, in this context of infection, the S53 strain is virulent while the S50 is not.

Next, we aimed to evaluate the role of Mfd in strain virulence by comparing insect survival in the wild type (wt) strains and in the Δmfd mutants. No difference in virulence was observed between the S50 and the S50 Δmfd strains, with none of them inducing a significant decrease in insect survival (Fig 2B). By contrast, the S53 Δmfd mutant was strongly impaired in its virulence capacity compared to the wild type S53 strain, with insect survival reaching 51% and 15%, respectively (Fig 2C). **Overall, these data suggest that Mfd plays an essential role in promoting virulence of pathogenic strains.**

The sequence/structure of Mfd recapitulates virulence

The impact of the Mfd proteins from the two strains was assessed on the laboratory strain Bc407 (Fig 2D). A deletion of *mfd* in the Bc407 strain induces a decrease of virulence compared to the wild type, confirming the role of Mfd in the virulence of this strain, as previously described [12]. The complementation of the Bc407 Δmfd mutant with the *mfd* sequence of the S50 strain did not lead to an increased virulence. By contrast, the complementation of the Bc407 Δmfd mutant with the *mfd* sequence of the S53 strain induced an increase of virulence to a level similar to the one obtained with the wild type Bc407 strain. These findings demonstrate that the Mfd protein exhibits strain-specific sequence variability and strongly support a direct link between **Mfd sequence diversity and its contribution to strain pathogenicity.**

In silico analysis of Mfd sequences and structures

To go deeper, we led an *in-silico* sequence/structure analysis to extract plausible patterns of substitutions that associate molecular impact. The alignment of the Mfd sequences of strains S50 and S53 revealed 98% identity. Overall, the 28 counted differences seem unpredictably

disseminated along the 7 domains (Fig 3 upper panel). Markedly, none belongs to any variant of *E. coli* Mfd homolog already described in the literature as defective for Mfd activity [37]. At sequence level, there is no difference in positively charged residues between S50 and S53 strains, and the three more negatively charged residues in S53 finally do not impact the isoelectric point (pI) of 5.4 (Table 3). The most conserved domains are D4 and D5 with a single substitution per domain, then D3, D6 and D7 domains show 3, 2 and 3 differences, respectively. Reversely, the N-terminal domains D1a, D2 and D1b are more prone to substitutions, as they accumulate 60% of the residue differences, with 17 out of 28 mutations found between S50 and S53, including 6 substitutions into charged residues in S53 (Table 3). Because these substitutions, despite disseminated on the sequence, can cluster in 3D, they were mapped onto the 3D structures of Mfd_{S50} and Mfd_{S53} modeled with respect to their intensive reshaping as seen in Mfd_{E.coli} cycle. Markedly, all the substitutions insert at the interface between domains or between domains-forming-modules (Fig 4A). Three clusters were identified, which all located at the N-terminal module. Cluster #1 is formed by polar residues K51, H115, cluster #2 by D305, M131 and N292 and cluster #3 by S143, T152, F153 and E200 (S53 numbering). Because cluster#3 is facing domain D7, with residue 200 protruding out and participating to the binding surface to UvrA protein, we can hypothesize that these mutations could impact UvrA recognition. Overall, Clusters 1-3 share a flat surface, which is markedly accessible in every conformation of Mfd revolution (Fig 4B). Another feature is that residues L92, D210, T262 and S324 assemble together on the other side of the facet, and contact D6 of the translocase module throughout the remodeling (Fig 4B). Of note, an extra contact exists in L₀ where D210 forms salt bridges with K485 and K487 of the D3-D4 linker.

Because some of these substitutions could collide with the conformational reshaping of the protein or with the binding to partners, the impact of the mutations was investigated. They were reported on Mfd templates, homology modeled from the intensive structural rearrangement expected both in S50 and S53. Accordingly, five conformations of Mfd_{E. coli}, were accurately selected from solved structures using cryo-EM and used as conformational templates (Fig 4 B and suppl. Fig 3). The first conformation, called L₀ (load zero), describes the auto-repressed inactive structure (PDB code: 2EYQ)[38], the second -L1- refers to Mfd prone to accommodate the ATP, the third is L₂ and refers to the Mfd loaded onto RNAP and a double strand of DNA (PDB code 6X2F), the fourth and fifth are called C₁ and C₅, respectively, and

correspond to the first and last steps of the nucleotide hydrolysis cycle, respectively (PDB code 6X2N and 6X50)[39]. ***In fine the entire functional cycle of Mfd_{S53} was homology modeled.***

Study of relation sequence/structure in the virulence of S53

Complementation of the S53Δ*mfd* strain with its native full-length *mfd* partially restored virulence, reaching an intermediate level between the S53Δ*mfd* mutant and the wild-type strain (Fig 2C). By contrast, complementation with the *mfd*_{S50} did not result in any detectable restoration of virulence, thus weighting that Mfd_{S50} does not drive virulence while Mfd_{S53} does. To refine our analysis, we sought to assess whether sequence differences at the N- or C-terminal regions of Mfd may contribute to the phenotypic modulation of pathogenicity through molecular and functional fine-tuning. Indeed, the D1a, D1b and D2 domains, which form the UvrB-like module and locate at the N-terminal end, harbor 17 substitutions over 370 residues, whilst D3-D7 domains, located at the C-terminal end, display 11 mutations over 770 amino-acids. In short, the region D3-D7 encompasses one third less substitutions over twice more residues. Accordingly, the S53Δ*mfd* strain was complemented with chimeric *mfd* constructs, containing either the D1a/b–D2 domains from S50 and the D3–D7 domains from S53, or the reciprocal combination, the D1a/b–D2 domains from S53 and the D3–D7 domains from S50 (Supp Fig 1). Both chimeric constructs restored virulence to a comparable extent. However, complementation with *mfd*_{D^{1a/b-2}S50_D³⁻⁷S53} resulted in slightly higher virulence than *mfd*_{D^{1a/b-2}S53_D³⁻⁷S50}, with the former showing a statistically significant increase of virulence compared to the S53Δ*mfd* strain (Fig 2C). Overall, these results suggest that the presence of either the N-terminal or C-terminal region of S53 Mfd is sufficient to enhance virulence (Fig 2C). **Collectively, these findings support that Mfd varies across strains and that its contribution to pathogenicity is directly shaped by underlying sequence variation.**

In silico analysis of Mfd dynamics

Markedly, mutations all locate at the interfaces of domains or domains-forming-modules, where they could articulate Mfd flexibility and play a role in the lubrication of its remodeling. To document this hypothesis and to explore the impact of substitutions, particularly at the onset of the functional cycle, simulations of molecular dynamics were run on Mfd of strains S50 & S53 in conformations L₀ and L₁. The RMSF -Root Mean Square Fluctuations- was computed for each trajectory. RMSF is a metric that quantifies the fluctuation of the position

of a residue over time. Regions with high RMSF values indicate high structural mobility, while regions with low RMSF values are more rigid. We used the RMSF to quantify and compare the flexibility of L_0 and L_1 over the course of a simulation and between the two strains. The plotted RMSF shows no much variation between L_0 and L_1 in S50, but a drastic change in S53 (Fig 5 A). Additionally, the RMSF profile of Mfd in S50 indicates D1b, D3, and D7 as domains of increased flexibility compared to others in L_0 and L_1 . Reversely, the flexibility measured for L_1 in S53 encloses the entire Mfd and not only specific domains. Also, Mfd tends to unfold with some secondary structures becoming coils (data not shown). These simulations emphasized that the substitutions in S53 could largely impact the fold and dynamics of Mfd, possibly by greasing the hinges between the modules, favouring motion coupling and Mfd unlocking. This is in line with our experimental data that shows that alterations of N- or C-terminal moieties in Mfd S53 profile an increase in virulence. Indeed, N-terminal and C-terminal ends are in close interaction at L_0 and even if most of the mutations are located in N-terminal, the C-terminal end contains the ATP site that is mandatory to navigate onto DNA and RNAP binding, so that **any variation at one extremity impacts the other extremity and more importantly, disturb the entire Mfd, possibly challenging its functional remodelling** (Fig 5 B-C).

Discussion

This study highlights the importance of the Mfd sequence variation in the virulence of *B. cereus*. Despite the high homology degree, exchanging domains between proteins originating from environmental and clinical strains was sufficient to alter virulence, indicating that molecular differences can have phenotypic consequences. To investigate this hypothesis, we did an in-depth study of Mfd in order to identify if their sequences could encode such change in virulence. Hence, we combined *in vitro* sequence investigation with *in silico* sequence/structure and dynamics analysis.

In silico, our analysis showed that there is a clear distinction in the Mfd sequences between environmental and clinical strains, as sequences from environmental isolates cluster together and separately from the clinical isolates. To gain molecular insight of this trait, we generated homology models of Mfd from the S50 and S53 strains, which display the most divergent sequence alignment and thus harbour the highest number of substitutions. The homology models, with the mapped mutations, showed that these substitutions are predominantly

located at the junctions between domains and domain-forming modules. Indeed, three clusters of variants were identified at the N-terminal domain that could impact its interaction with the C-terminal domain and thus dysregulate transition between active and auto-repressed forms.

Thus, we hypothesized that such modifications positioned at joint sites may affect the functional remodelling of Mfd, lubricating the articulation of modules, promoting Mfd opening towards active conformations, thereby influencing virulence. MD simulations showed that Mfd of S53 meets with increased flexibility during the L_0 to L_1 transition. In line with that, the experimental substitution of domains D1a/b and D2 in Mfd_{S50} by those in Mfd_{S53} resulted in increased virulence, whereas the opposite substitution did not alter much virulence relative to complementation with the native Mfd. This suggests that the D1a/b-D2 domains can contribute to virulence variation. Additionally, the major determinant of virulence appears to reside within the C-terminal region of Mfd since the chimera with D3 to D7 of Mfd_{S53} displayed virulence levels comparable to the native Mfd_{S53}.

Of note, we performed a multiple sequence alignment of the sequences of UvrA, retrieved from the proteomes of environmental and clinical isolates, and we built a distance-based clustering tree. Sequences of UvrA aligned in a similar bimodal clustering between non-pathogenic and pathogenic strains, despite a sequence identity above 92 % over 960 amino-acid residues (Suppl. Fig 4). The binding between the N-terminal end and the UvrA moiety was modeled in S53 with respect to the crystal structure of the core complex of UvrA with Mfd in *E. coli* (PMID: 22331906) (Suppl. Fig 5). It emphasized the competition with D7. Particularly, the interface of interaction between D2 and UvrA or D7 is conserved, with the preserved salt bridges E167_{UvrA}/R201_{Mfd}, R176_{UvrA}/D183_{Mfd}, and R206_{UvrA}/D190_{Mfd} that align with K151_{UvrA}/E200_{Mfd}, E171_{UvrA}/R213_{Mfd}, R182_{UvrA}/E195_{Mfd}, R210_{UvrA}/D202_{Mfd}, in *E. coli* and in *B. cereus* strain S53, respectively. What could be underlined is that substitution E200Q occurs in S50 that could impact the binding to UvrA. Also, the connecting points between D2 and D7 differ in Mfd between *E. coli* and *Bacillus* strains, as interactions between A167/G1051 and H527/A1031 in *E. coli* are replaced by a mild connection with G178/G1070 and a strong salt bridge between D550 and R1050 (data not shown). This could emphasize that binding to cognate partner such as UvrA or RID has to be maintained over bacteria, while remodelling accepts some tuning. This is in line with R165 in D2 domain of *S. typhimurium* and L499 in RID of *E. coli*, whose mutations into Alanine and Arginine, abrogate their interaction with D7 and

291 RNAP, respectively [17]. In the light of our work, the C-terminal region of the Mfd appears to
292 play an important role in virulence and may account for a substantial part of the virulence-
293 associated activity of Mfd.

294 In parallel, we demonstrated *in vivo* that deleting *mfd* from a clinical strain induces a clear
295 increase of the insect survival while that same mutation has no impact on the environmental
296 strain. Hence, our study confirms Mfd as a virulence factor and further emphasizes that its
297 contribution to pathogenicity depends on its intrinsic activity, which can be altered at
298 exquisite points in sequence and structure. Markedly, we observed a similar clustering
299 between Mfd from virulent *Listeria monocytogenes* as compared to non-pathogenic *Listeria*
300 *innocua* (Suppl. Fig 6). Together, these observations point toward a potentially conserved
301 distinction in Mfd sequences between pathogenic and non-pathogenic bacteria, raising the
302 possibility that Mfd could serve as a biomarker for virulence-associated phenotypes.

303 304 **Methods**

305 **Sequence access and alignment**

306 The 22 sequences of Mfd used in this study were previously collected and de-identified from
307 bacterial isolates obtained from new-born patients. They are accessible at
308 <http://www.ebi.ac.uk/ena/> within projects PRJEB18787 or PRJNA718492, for all but S50 and
309 S53 which can be retrieved at <https://www.ebi.ac.uk/ena/browser/view/SAMEA4361354>
310 and at <https://www.ebi.ac.uk/ena/browser/view/SAMEA41933668>, respectively. All the
311 sequences were aligned using MAFFT (Multiple Alignment using Fast Fourier Transform) with
312 default parameters and ClustalW output format, from the EMBL-EBI web server
313 (<https://www.ebi.ac.uk/jdispatcher/msa/mafft>). The multiple sequence alignment (MSA)
314 grouped Mfd sequences based on their identity, that was visualised using Jalview [28].

315 316 ***In silico* analysis and molecular dynamics simulations of Mfd**

317 The Mfd protein structures of strains of *B. cereus* S50 and S53 were homology modelled from
318 the sequence of Mfd from *E. coli* strain K12, aligned with *E. coli* template, using MODELLER
319 software version 10.8 [30]. Five 3D templates were selected as relevant conformations within
320 the functional cycle, namely the conformations L₀ for "load zero state" (PDB code 2EYQ) [14],
321 L₁ (PDB code 6X26), L₂ (PDB code 6X2F), C₁ (PDB code 6X2N) and C₅ (PDB code 6X50), for first
322 and last steps of the "nucleotide hydrolysis Cycle", respectively [31]. Per template/strain,

hundred models were computed, then the five best models were selected according to the DOPE scores. Eventually, one model was kept after stereochemistry check and visual inspection, using PROCHECK v.6 (<https://saves.mbi.ucla.edu/>) [32] and PyMOL v.2.5.2 (Schrödinger, LLC.: <https://pymol.org/2/>), respectively. To quantify structural distortions between the Mfds, the align command in PyMOL that superimposes the C α trace of the proteins was employed and the models were aligned with their D4 superimposed onto the D4 of L₀. D4 was chosen as the reference because it is central, small and highly structured. Next, amino acid variants of S50 and S53 were mapped onto the structure and highlighted through space-filling rendering. Of note, AlphaFold3 (AF3) could not be used to model each conformation with respect to the reshaping because only the L₀ conformation is the template available in the AF3 database (code 2eyq). AF3 was used to model UvrA from S53. Next, UvrA model was superimposed onto the crystal structure of its core in complex with D2_{Mfd} in *E. coli* (pdb code 4DFC)[29].

Next, molecular dynamics (MD) simulations of 200 ns each were run in triplicates for Mfd from S50 and S53 modeled in L₀ and L₁, and using GROMACS [30]. Of note, for MD in L₁, the topology of the ATP ligand was created using acpype script [31]. The RMSF -Root Mean Square Fluctuation- was computed using the adequate command line in GROMACS and plotted with XmGrace (Turner PJ. XMGRACE, Version 5.1.19. Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, OR; 2005).

Bacterial strains and mutant construction.

The bacterial strain *B. thuringiensis* 407 Cry- (Bc 407) was used as a model for *B. cereus* [32]. The Bc 407 Δmfd was obtained through double homologous recombination using the thermosensitive vector pMAD and described in [14]. Two strain libraries were used for the study. The first one is composed of 10 strains, collected from various sources that did not induced any infection to date [33]. The second is composed of 12 clinical strains, isolated from 12 infected patients [34](Tab 1, 2). The S50 Δmfd and S53 Δmfd strains were obtained by double homologous recombination using the pAT113 plasmid [35] since it is not replicative in *B. cereus*. The upstream and downstream fragments of *mfd* from each strain and a tetracycline resistance cassette were cloned inside pAT113 by Genecust (Boynes, France). The two plasmids (pAT113_ Δmfd _S50 and pAT113_ Δmfd _S53) were transformed into the *E. coli* HB101 strain and then transferred by conjugation into the S50 and S53 strains. The donor and

the receptor strains were mixed together at a ratio of (5:1) on a sterile membrane filter with a pore size of 0.45 µm (HAWP02500 Milipore), placed on BHI plates and incubated overnight at 37°C. After mating, the bacteria were resuspended and incubated at least 30' in 1.5 mL of colicine, centrifuged 5' at 3000 g, the supernatant discarded and resuspended again in 1.5 mL of colicine. The mutants were selected on BHI plates, supplemented with 150 µg/mL of tetracycline and verified by PCR using primers upstream and downstream from the cloning region:

For S50: (5' – GTATGGGAATCGATCGTCCG – 3') and (5' – TTACAGACTTGCGCTCTTCC – 3')

For S53: (5' – GCATGGGAATTGATCGTCCG – 3') and (5' – AACTGGCAGCAGTATTGACC – 3').

All the complemented strains were obtained using the pAT28 plasmid [36] carrying either the native *mfd* sequences from strains S50 or S53, or chimeric *mfd* variants in which domains D1a/b and D2 were exchanged between the S50 and S53 sequences (Suppl. Fig 1). These constructs were cloned into pAT28 by Genecust. The resulting plasmids (pAT28_*mfd*_S50, pAT28_*mfd*_S53, pAT28_*mfd*_D^{1&2}S50_D³⁻⁷S53 and pAT28_*mfd*_D^{1&2}S53_D³⁻⁷S50) were transformed into the *E. coli* HB101 strain and transferred into S53 strain by conjugation as described above. The selection was done on BHI plates supplemented with 300 µg/mL of spectinomycin and verified by PCR using primers located on the plasmid:

(5' – CCCAGTCACGACGTTGTAAAA CG – 3') and (5' – AGCGGATAACAATTTACACAGG – 3')

Insect assay

Fourth instar silkworm larvae *Bombyx eri* were purchased from “L'office pour les insectes et leur environnement” (OPIE, Guyancourt, France). Larvae were reared with *Ligustrum vulgare* or *Ligustrum japonicum* until they reached 0.6 – 0.9 g. Prior to the assays, *B. eri* larvae were starved overnight and randomly distributed in groups of 10 to 15 per conditions. *B. cereus* strains were grown in Lysogeny broth (LB) medium at 37°C under agitation at 180 rpm until early exponential growth phase. Bacterial cultures were diluted in Phosphate Saline Buffer (PBS) to obtain 1.10³ – 3.2.10³ CFU/injection (Fig 2A and 2B) or 1.10⁴-3.2.10⁴ CFU/injection. Larvae were injected with 20 µL of bacterial suspension in the haemolymph *via* the last pro-leg. The injections were performed using a 1 mL hypodermic syringe with a 0.5 x 25 mm needle and an automated syringe pump (KD Scientific KDS 100). Post-injection, larvae were placed at 29°C for 24h where the survival rate was monitored.

Data availability statement

The data that support the findings of this study are available at <http://www.ebi.ac.uk/ena/> under projects # PRJEB18787 and PRJNA718492. Additionally, Strains 50 and S53 are also available here, respectively:

S50 : <https://www.ebi.ac.uk/ena/browser/view/SAMEA4361354>

S53 : <https://www.ebi.ac.uk/ena/browser/view/SAMEA41933668>

References

1. Authority, E.F.S. The Community Summary Report on Food-Borne Outbreaks in The European Union in 2007. *EFSA Journal*. **7**, 1-410 (2009).
2. Stenfors Arnesen, L.P., A. Fagerlund, and P.E. Granum, From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol Rev*. **32**(4), 579-606 (2008).
3. Bottone, E.J. *Bacillus cereus*, a volatile human pathogen. *Clin Microbiol Rev*. **23**(2), 382-398 (2010).
4. Cormontagne, D., et al. *Bacillus cereus* Induces Severe Infections in Preterm Neonates: Implication at the Hospital and Human Milk Bank Level. *Toxins* (Basel). **13**, 123-149 (2021).
5. Glasset, B., et al. *Bacillus cereus*-induced food-borne outbreaks in France, 2007 to 2014: epidemiology and genetic characterisation. *Euro Surveill*. **21**(48), 1-32 (2016).
6. Lund, T., M.L. De Buyser, and P.E. Granum. A new cytotoxin from *Bacillus cereus* that may cause necrotic enteritis. *Mol Microbiol*. **38**, 254-261 (2000).
7. Fagerlund, A., et al. *Bacillus cereus* Nhe is a pore-forming toxin with structural and functional properties similar to the ClyA (HlyE, SheA) family of haemolysins, able to induce osmotic lysis in epithelia. *Microbiology* (Reading), **154**, 693-704 (2008).
8. Lund, T. and P.E. Granum. Characterisation of a non-haemolytic enterotoxin complex from *Bacillus cereus* isolated after a foodborne outbreak. *FEMS Microbiol Lett*. **141**, 151-156. (1996).
9. Beecher, D.J., et al. Extracellular virulence factors in *Bacillus cereus endophthalmitis*: methods and implication of involvement of hemolysin BL. *Infect Immun*. **63**, 632-639 (1995).
10. Guillemet, E., et al. The InhA metalloproteases of *Bacillus cereus* contribute concomitantly to virulence. *J Bacteriol*. **192**, 286-294 (2010).
11. Tran, S.L., et al., Structural Modeling of Cell Wall Peptidase CwpFM (EntFM) Reveals Distinct Intrinsically Disordered Extensions Specific to Pathogenic *Bacillus cereus* Strains. *Toxins* (Basel). **12** (593), 1-28 (2020).
12. Kavanaugh, D.W., et al. The pathogenic biomarker alcohol dehydrogenase protein is involved in *Bacillus cereus* virulence and survival against host innate defence. *PLoS One*. **17**, 10.1371; 10.1371/journal.pone.0259386. (2022).
13. Porrini, C., et al. Implication of a Key Region of Six *Bacillus cereus* Genes Involved in Siroheme Synthesis, Nitrite Reductase Production and Iron Cluster Repair in the Bacterial Response to Nitric Oxide Stress. *Int J Mol Sci*. **22**, 10.3390; 10.3390/ijms22105079 (2021).

- 431 14. Guillemet, E., et al. The bacterial DNA repair protein Mfd confers resistance to the
432 host nitrogen immune response. *Sci Rep.* **6**, 29349. (2016).
- 433 15. Roberts, J. and J.S. Park. Mfd, the bacterial transcription repair coupling factor:
434 translocation, repair and termination. *Curr Opin Microbiol.* **7**(2) 120-125. 2004
- 435 16. Deaconescu, A.M., et al. Structural basis for bacterial transcription-coupled DNA
436 repair. *Cell.* **124**, 507-520. (2006).
- 437 17. Han, J., et al. Key role of Mfd in the development of fluoroquinolone resistance in
438 *Campylobacter jejuni*. *PLoS Pathog.* **4**(6): p. e1000083. (2008).
- 439 18. Lee, G.H., et al. The *Helicobacter pylori* Mfd protein is important for antibiotic
440 resistance and DNA repair. *Diagn Microbiol Infect Dis.* **65**, 454-456. (2009).
- 441 19. Ragheb, M.N., et al. Inhibiting the Evolution of Antibiotic Resistance. *Mol Cell.* **73**
442 157-165 e5. (2019).
- 443 20. Carvajal-Garcia, J., et al. A small molecule that inhibits the evolution of antibiotic
444 resistance. *NAR Mol Med.* **1**(1): p. ugae001 (2024).
- 445 21. Tran, S.L., et al., An anti-virulence drug targeting the evolvability protein Mfd
446 protects against infections with antimicrobial resistant ESKAPE pathogens. *Nat*
447 *Commun.* **16**(1) 3324-3343 (2025).
- 448 22. Darrigo, C., et al. The Bacterial Mfd Protein Prevents DNA Damage Induced by the
449 Host Nitrogen Immune Response in a NER-Independent but RecBC-Dependent
450 Pathway. *PLoS One.* **11**(10): p. e0163321. (2016).
- 451 23. Porrini, C., N. Ramarao, and S.L. Tran, Dr. NO and Mr. Toxic - the versatile role of
452 nitric oxide. *Biol Chem.* **401**, 547-572. (2020).
- 453 24. Brugger, C., et al. Molecular determinants for dsDNA translocation by the
454 transcription-repair coupling and evolvability factor Mfd. *Nat Commun.* **11**, 3740.
455 (2020).
- 456 25. Assenmacher, N., et al. Structural basis for transcription-coupled repair: the N
457 terminus of Mfd resembles UvrB with degenerate ATPase motifs. *J Mol Biol.* **355**,
458 675-683. (2006).
- 459 26. Park, J.S., M.T. Marr, and J.W. Roberts. E. coli Transcription repair coupling factor
460 (Mfd protein) rescues arrested complexes by promoting forward translocation. *Cell.*
461 **109**, 757-767. (2002).
- 462 27. Deaconescu, A.M., N. Savery, and S.A. Darst. The bacterial transcription repair
463 coupling factor. *Curr Opin Struct Biol.* **17** 96-102. (2007).
- 464 28. Waterhouse, A.M., et al. Jalview Version 2--a multiple sequence alignment editor and
465 analysis workbench. *Bioinformatics.* **25**, 1189-1191. (2009).
- 466 29. Deaconescu, A.M., et al. Nucleotide excision repair (NER) machinery recruitment by
467 the transcription-repair coupling factor involves unmasking of a conserved
468 intramolecular interface. *Proc Natl Acad Sci U S A.* **109**, 3353-3358. (2012).
- 469 30. Van Der Spoel, D., et al. GROMACS: fast, flexible, and free. *J Comput Chem,* **26**
470 1701-1718. (2005).
- 471 31. Sousa da Silva, A.W. and Vranken, W.F. ACPYPE - AnteChamber PYthon Parser
472 interfacE. *BMC Res Notes.* **5**, 367-375 (2012).
- 473 32. Lereclus, D., et al. Transformation and expression of a cloned delta-endotoxin gene in
474 *Bacillus thuringiensis*. *FEMS Microbiol Lett.* **51**, 211-217. (1989).
- 475 33. Kamar, R., et al. Pathogenic potential of *Bacillus cereus* strains as revealed by
476 phenotypic analysis. *J Clin Microbiol.* **51**, 320-323. (2013).
- 477 34. Glasset, B., et al. *Bacillus cereus*, a serious cause of nosocomial infections:
478 Epidemiologic and genetic survey. *PLoS One.* **13**(5): p. e0194346. 2018.

35. Trieu-Cuot, P., et al. An integrative vector exploiting the transposition properties of Tn1545 for insertional mutagenesis and cloning of genes from gram-positive bacteria. *Gene*. **106**, 21-27. (1991).
36. Trieu-Cuot, P., et al. A pair of mobilizable shuttle vectors conferring resistance to spectinomycin for molecular cloning in *Escherichia coli* and in gram-positive bacteria. *Nucleic Acids Res.* **18**, 4296 (1990).
37. Deaconescu, A.M. Mfd - at the crossroads of bacterial DNA repair, transcriptional regulation and molecular evolvability. *Transcription*. **12**, 156-170. (2021).
38. Sali, A. and T.L. Blundell. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol.* **234**, 779-815. (1993).
39. Kang, J.Y., et al. Structural basis for transcription complex disruption by the Mfd translocase. *Elife* **10**, 10.7554; 10.7554/elife.62117 (2021).

Acknowledgements

This work was supported by TWB biosciences and bioproduction (N.R., G.A), the Mécénat des Mutuelles AXA (laureate 2022, N.R.) and l'Agence Nationale de la Recherche (ANR) project NO ESKAPE ANR-23-CE18-0028-01 (S.M., N.R., G.A) within the frame of the IHU SEPSIS.

Author's Contributions

D.C performed microbiology experiments, insects' analysis, prepared figures 2, Suppl1 and tables 1-2, wrote the first version of the manuscript. S.S performed sequences alignments, started molecular dynamics simulations, prepared figures 1, Suppl2-6 and wrote the first version of the manuscript. T. M performed the homology modeling, followed up the molecular dynamics simulations and performed their analysis. S.A aligned the sequences, did the 3D mapping of the substitutions and prepared Table3. S-L.T supervised the insects' experiments. S.M ran the AlphaFold3 modeling part, interpreted the data, and prepared Suppl Fig 5. NR designed, supervised the experimental work, interpreted the data, corrected the main manuscript, and provided materials and funding. G.A designed, supervised the *in silico* work, interpreted the data, corrected the manuscript and prepared the figures 3-4, and provided materials and funding. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Figure legends

Fig 1: Mfd sequences alignment from *B. cereus* strains

The Mfd sequence of 10 environmental and 12 clinical strains of *B. cereus* were aligned using MAFFT algorithm available from EMBL-EBI webtools, with Clustal output format. The sequences, aligned with respect to their residue differences, are grouped into three main clusters: environmental (green), clinical (orange), and "twilight zone" strains (grey). The residue discrepancies are highlighted according to the conservation score with a gradient of color from brown (below 6) to yellow (above 8).

Fig 2: Implication of Mfd in the difference of virulence between S50 and S53 strains and impact of *mfd* complementation in the virulence of Bc407 strain in insect

Strains were grown to early exponential phase in LB medium at 37 °C and injected into forth instar *B. eri* larvae. Survival was monitored at 24h post infection. **A:** Between 1.10^3 to $3.2.10^3$ CFU of S50 and S53 were injected (N= 6-8). **B:** Between 1.10^3 to $3.2.10^3$ CFU of S50 or S50 Δ *mfd* were injected (N= 5-6). **C:** Between 1.10^4 to $3.2.10^4$ CFU of S53, S5 Δ *mfd*, S53 Δ *mfd* Ω *mfd*_53, S53 Δ *mfd* Ω *mfd*_5, S53 Δ *mfd* Ω *mfd*_D^{1&2}50_D³⁻⁷53 or S53 Δ *mfd* Ω *mfd*_D^{1&2}53_D³⁻⁷50 were injected (N= 7-12). **D:** *B. eri* larvae were injected between 1.10^3 to $3.2.10^3$ CFU of Bc407, Bc407 Δ *mfd*, Bc407 Δ *mfd* Ω *mfd*_50 or Bc407 Δ *mfd* Ω *mfd*_53 in early exponential phase.

Error bars show SEM and P value is calculated using Mann-Whitney test (* p< 0.05; ** p < 0.01; *** p< 0.001; **** p < 0.0001).

Fig 3: Mapping of the mutations in Mfd of S50 and S53 and L₀ and L₁ homology models of S53.

Upper panel: the mutations are evidenced on Mfd described with respect to their numbering and highlighted on the functional modules. The modules are colored from N-terminal to C-terminal ends, D1a in dark blue, D2 in cyan, D1b in blue, D3 in orange, RID D4 in pink, TD1-D5 in green, TD2-D6 in yellow, D7 in red. Lower panel: the homology models of Mfd in conformations L₀ and L₁, built on sequence of S53 representative of the clinical group. L₀ (load zero) and L₁ conformations describe the auto-repressed inactive structure (template Mfd in E.

coli PDB code: 2EYQ) and the one able to bind ATP (template Mfd *E. coli* PDB code: 6X26), respectively. The two models are shown as cartoons, colored with the code described above, with the ATP ligand shown in space-filling mode. The color code is used across the paper. The images were prepared using PyMOL Schrödinger LLC.

Fig 4: Mapping of the mutations between S50 and S53.

A: Mutational imprinting and identification of clusters in the 3D space.

The amino acid residues that differ between S50 and S53 are mapped onto the 3D structure of L₀, represented as a cartoon and coloured according to the module-specific scheme (left panel). The positions of these substitutions are highlighted as dots using the corresponding colour code. The residues that are at the junction of modules are circled in black boxes, while the 3 clusters of the N-terminal module are circled in cyan (right panel).

B: Conformational revolution of Mfd of S53 between L₀ and C₅ with mapping of the mutations.

Mfd from homology modeled in key conformations of the functional cycle: L₀, L₂ which refers to the Mfd loaded onto RNAP and double strand DNA (Mfd in *E. coli* PDB code 6X2F), C₁ and C₅ which correspond to the first and last steps of the nucleotide hydrolysis cycle, respectively (Mfd in *E. coli* PDB code 6X2N and 6X50). The extensive remodeling, which is also expected for Mfd in *Bacillus* strains, is highlighted. The conformations follow the color code, they were superimposed on RID-D4, chosen as reference module and hidden here for sake of clarity. The flat and anchorage facets are shown with arrows.

B: Evidence of rigid body motion for the functional modules. Superposition on the corresponding module in L₀ for each functional module of the modeled conformations in L₂, C₁ and C₅ as described in panel A. Except for the translocation module that shows a RMS deviation above 6 Å between L₀ and L₂, all the modules display a RMS deviation below 2 Å.

Fig 5: Root-mean-Square-Fluctuations (RMSF) of Mfd residues along MD simulations

A- RMSF plotted vs residue on Mfd full length in L₀ and L₁ for S50 (left) and S53(right)

B- RMSF vs residue with a focus on N and C-terminal ends on Mfd in L₀ for both S50 and S53

C- RMSF vs residue with a focus on N and C-terminal ends on Mfd in L₁ for both S50 and S53

Suppl. Fig 1: Schematic representation of the *Bacillus cereus* strains and Mfd complementation constructs used in this study

Two *B. cereus* parental strains were used in this study: the environmental isolate S50 and the clinical isolate S53. An *mfd* deletion mutant (Δmfd) was generated from both strains and S53 Δmfd was then complemented with different *mfd* constructs. Complementation strategies included: (1) homologous complementation using the native *mfd* gene from S53 (Ωmfd_S53), (2) heterologous complementation using the *mfd* gene from S50 (Ωmfd_S50), and (3–4) two chimeric *mfd* constructs combining domains from both strains. In chimeric construct 1, domains D1a–D2 originated from S53, whereas domains D1b and D3–D7 originated from S50. In chimeric construct 2, domains D1a–D2 originated from S50, whereas domains D1b and D3–D7 originated from S53. The domain organization of Mfd is shown at the bottom of the figure.

Suppl. Fig 2: Distance-based clustering tree of aligned Mfd sequences from *B. cereus* strains

The average-distance tree was constructed from the MSA of Mfd (Fig. 1) using the BLOSUM62 substitution matrix implemented in Jalview and the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) clustering method. Pairwise distances calculated from the alignment are indicated on the corresponding branches. The aligned sequences are grouped into three main clusters: environmental (green), clinical (orange), and ambiguous strains (grey).

Suppl. Fig 3: Conformational reshaping of Mfd of S53 between L₀ and C₅ with a mapping of the mutations.

This figure mirrors Fig 4 with all the modules of Mfd of S53 highlighted in a cartoon and the amino acid variants shown as dots with respect to the color code. Apart and only for L2, the double DNA helix is displayed in cartoon.

Suppl. Fig 4: Distance-based clustering tree of aligned UvrA sequences from *B. cereus* strains

The average-distance tree was constructed from a previous UvrA alignment using BLOSUM62 and the UPGMA clustering method. Pairwise longest distances calculated from the alignment are indicated on the corresponding branches. The aligned sequences are grouped into three main clusters: environmental (green), clinical (orange), and ambiguous strains (grey).

Suppl. Fig 5: Binding competition at D2 of Mfd S53 L₀ between D7 and UvrA.

Left panel: D2 and D7 homology modeled are shown as cyan and red cartoons, respectively. UvrA modeled by AF3 and superimposed on the core complex of UvrA/Mfd D2 in *E. coli* (4DFC crystal structure (hidden) is shown in wheat cartoon (lower left panel). Residues responsible for the binding on both sides are shown in dots. Right panel: close view of the binding at the core complex between UvrA and D2 with the residues displayed as ball and sticks and dots.

Suppl. Fig 6: Distance-based clustering tree of aligned Mfd from *L. monocytogenes* and *L. innocua* strains

The average-distance tree was constructed from the previous Mfd alignment (Suppl. Fig. 1) using BLOSUM62 and the UPGMA clustering method. Pairwise longest distances calculated from the alignment are indicated on the corresponding branches. The aligned sequences are grouped according to the species.

Table legends

Table 1: Environmental strains of *Bacillus cereus* used in this study

Table 2: Clinical strains of *Bacillus cereus* used in this study

Table 3: Distribution of substitutions per domains between Mfd from *Bacillus cereus* strains S50, S53 and *E. coli*. Inserted are the number of mutations per domain and its rate of mutations. Each domain respects the color code given for Mfd in Fig.3.

Table 1 : Environmental strains used in this study

Strains	Origin
INRA PF (S9)	Milk protein
I13 (S10)	Cooked rice
INRA 5 (S11)	Pasteurized zucchini purée
INRA C64 (S12)	Pasteurized vegetables
ADRIA I3 (S13)	Cooked foods
INRA A3 (S38)	Startch
INRA SL' (S44)	Soil
I2 (S46)	Dried fruit
ADRIA I21 (S50)	Cooked foods
WSBC 10204 (S52)	Pasteurized milk

Table 2 : Clinical strains used in this study

Strains	Origin	Outcome
09CEB13BAC (S6)	Blood culture, New born	ND
09CEB14BAC (S7)	Blood culture, New born	ND
09CEB33BAC (S8)	Cutaneous, New born	ND
12CEB31BAC (S14)	Blood culture, New born	Death
13CEB06BAC (S15)	Arterial catheter, New born	ND
09CEB11BAC (S16)	Blood culture, New born	Death
09CEB31BAC (S17)	Central catheter, New born	ND
09CEB12BAC (S53)	Meninge, New born	Death
09CEB17BAC (S55)	Neo-natal unit	ND
09CEB34BAC (S59)	Gastric tube, New born	ND
09CEB36BAC (S61)	Central catheter, New born	ND
09CEB37BAC (S62)	Neo-natal intensive care unit	ND

*ND : Not disclosed

638
639

Table 3

Domains	S50	S53	<i>E. coli</i>	Mutations Per domain	Mutation rate
D1a 1-127	Q51 Y67 V92 N115	K51 (l) H67 (h) L92 (l) H115 (l)	G40 H56 D79 R103	4/127	0.03
D2 128-228	I131 N143 A152 L153 Q200 E210 R218	M131 (l) S143 (l) T152 (h) F153 (h) E200 (l) D210 (l) K218 (b)	C118 K131 A140 L141 E188 D198 V206	7/100	0.07
D1b 229-366	L238 A262 K292 N304 T324 N355	S238 (h) T262 (l) N292 (l) D304 (l) S324 (h) H355 (h)	L226 V237 S266 D279 T299 E330	6/137	0.04
D3 378-478	A439 V442 F466	G439 (l) I442 (l) L466 (b)	R414 E417 L441	3/100	0.03
D4 RID 502-568	H505	Y505 (l)	P482	1/166	0.006
D5 601-804	G615	E615 (h)	L592	1/203	0.005
D6 804-988	I835 G871 R951	V835 (l) S871 (h) K951 (h)	V812 R848 K928	3/184	0.01
D7 1022-1076	I1052 F1106 S1126	V1052 (l) I1106 (b) N1126 (l)	K1034 I1087 L1103	3/154	0.02

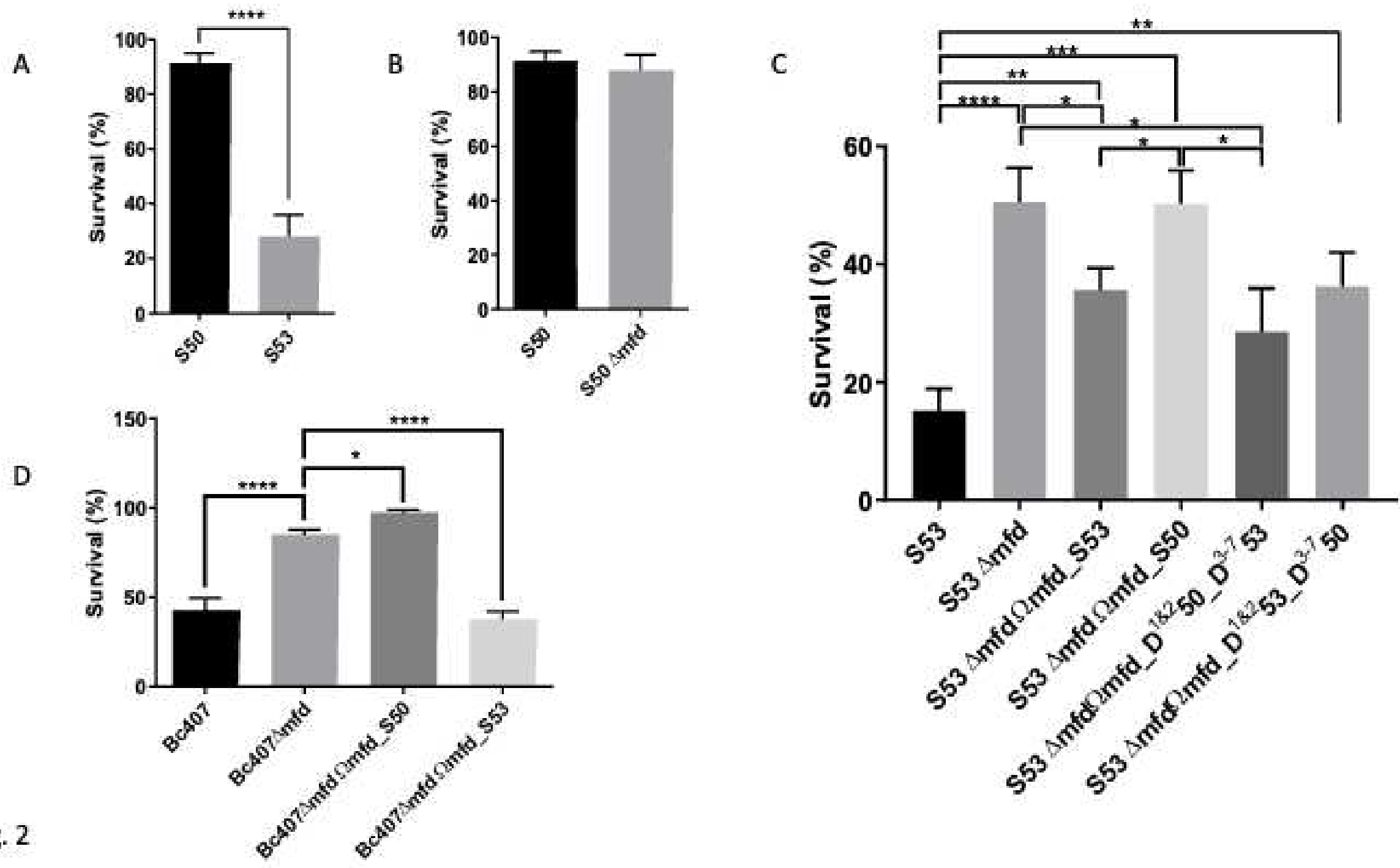


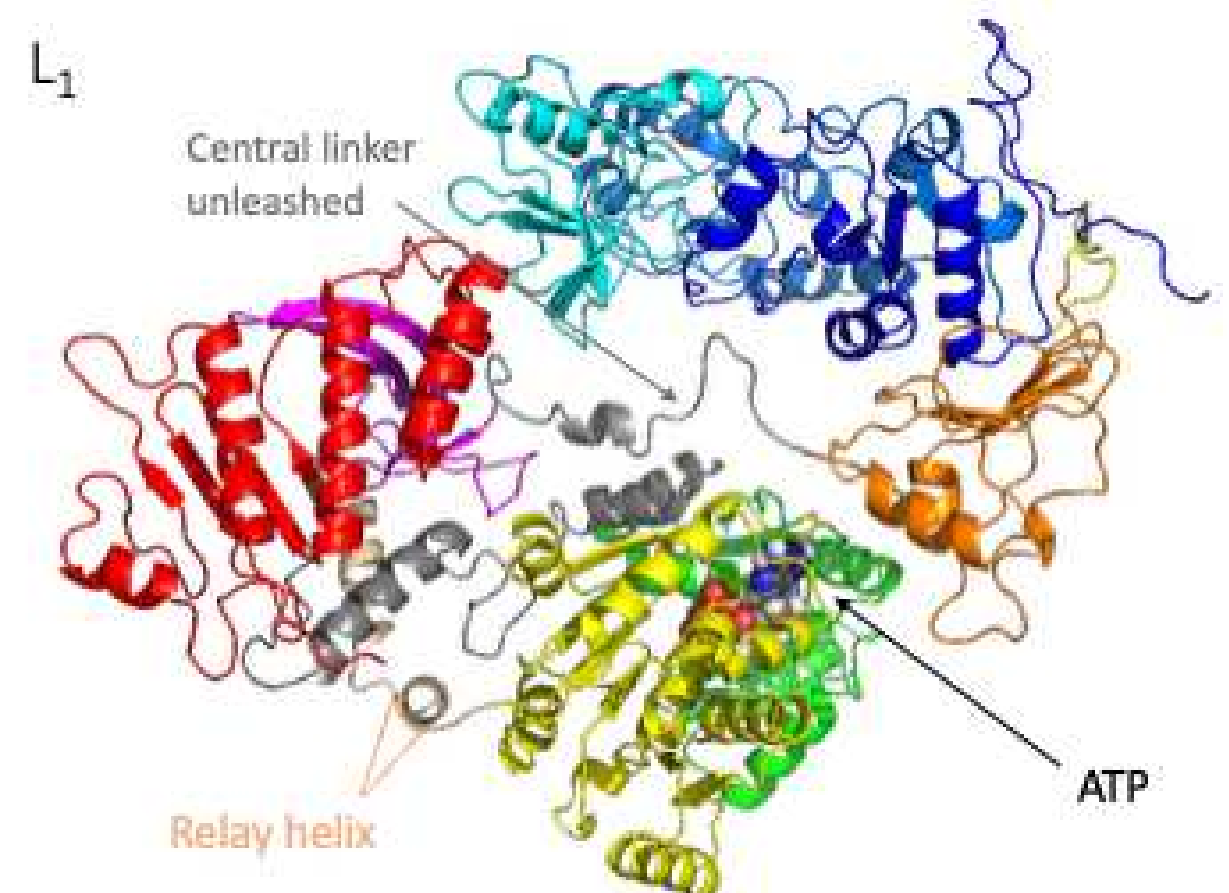
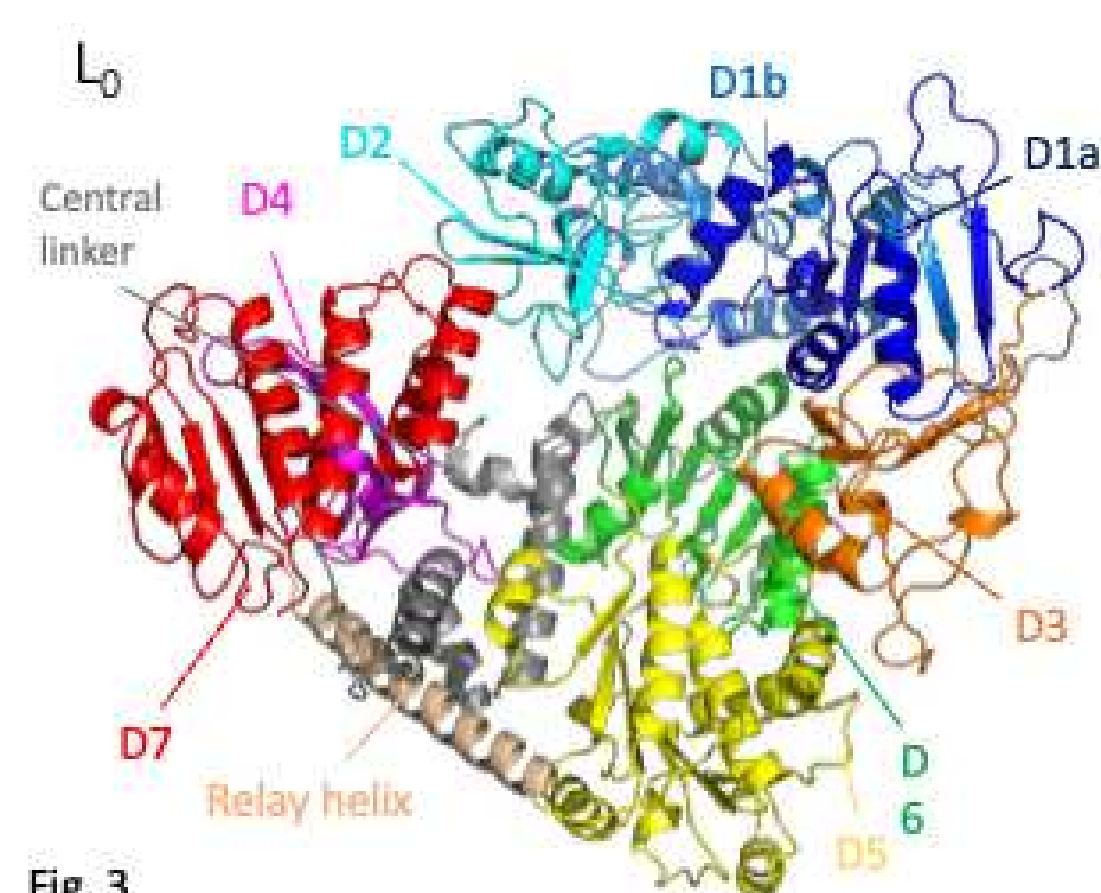
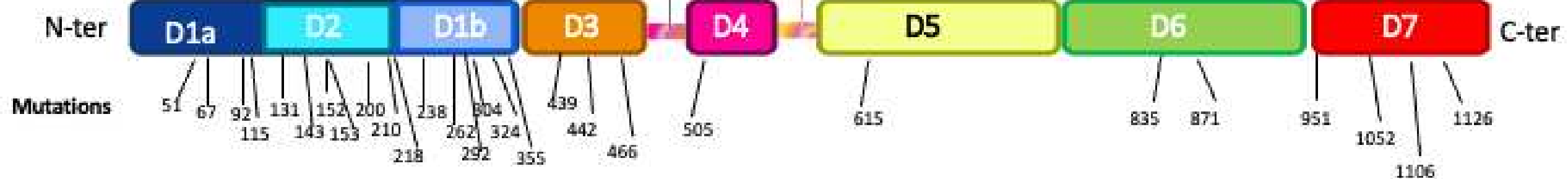
Fig. 2

Mfd Partner

UvrA

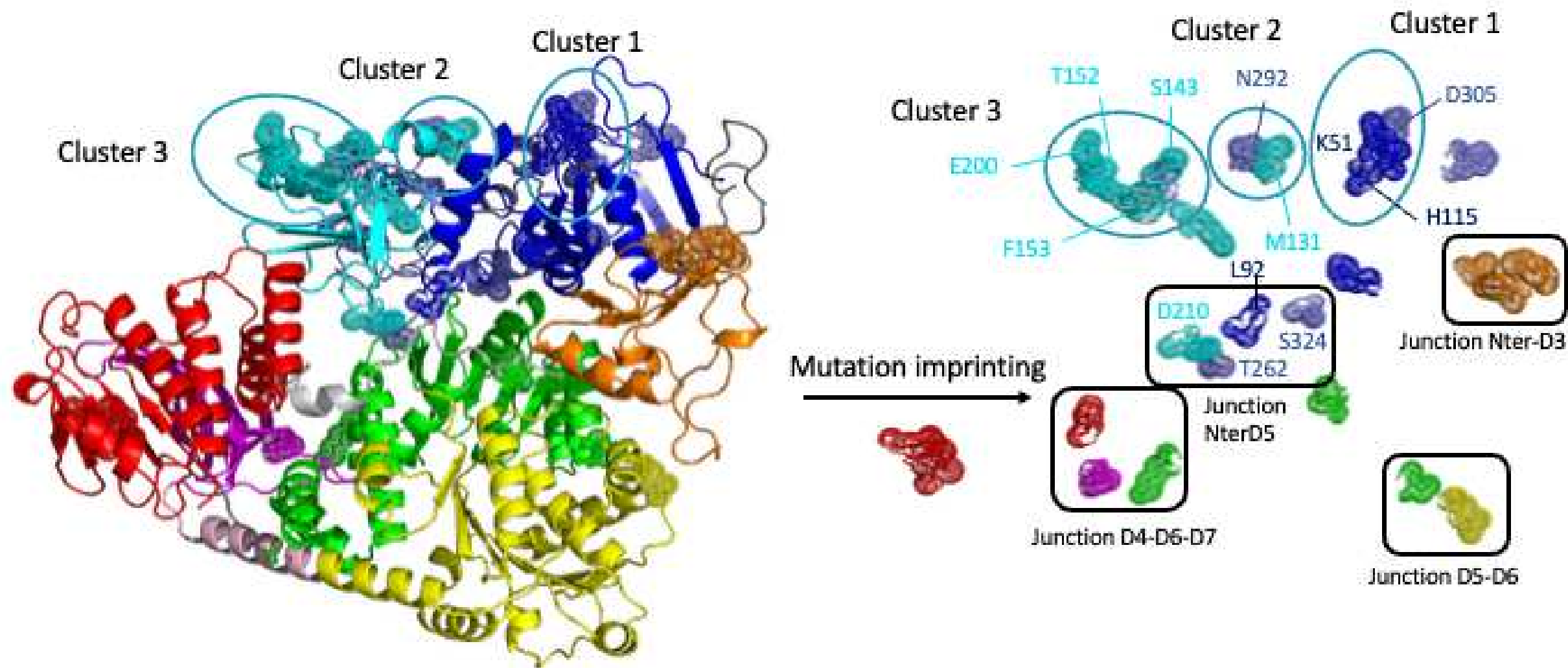
RNAP

DNA



S53

Fig. 3



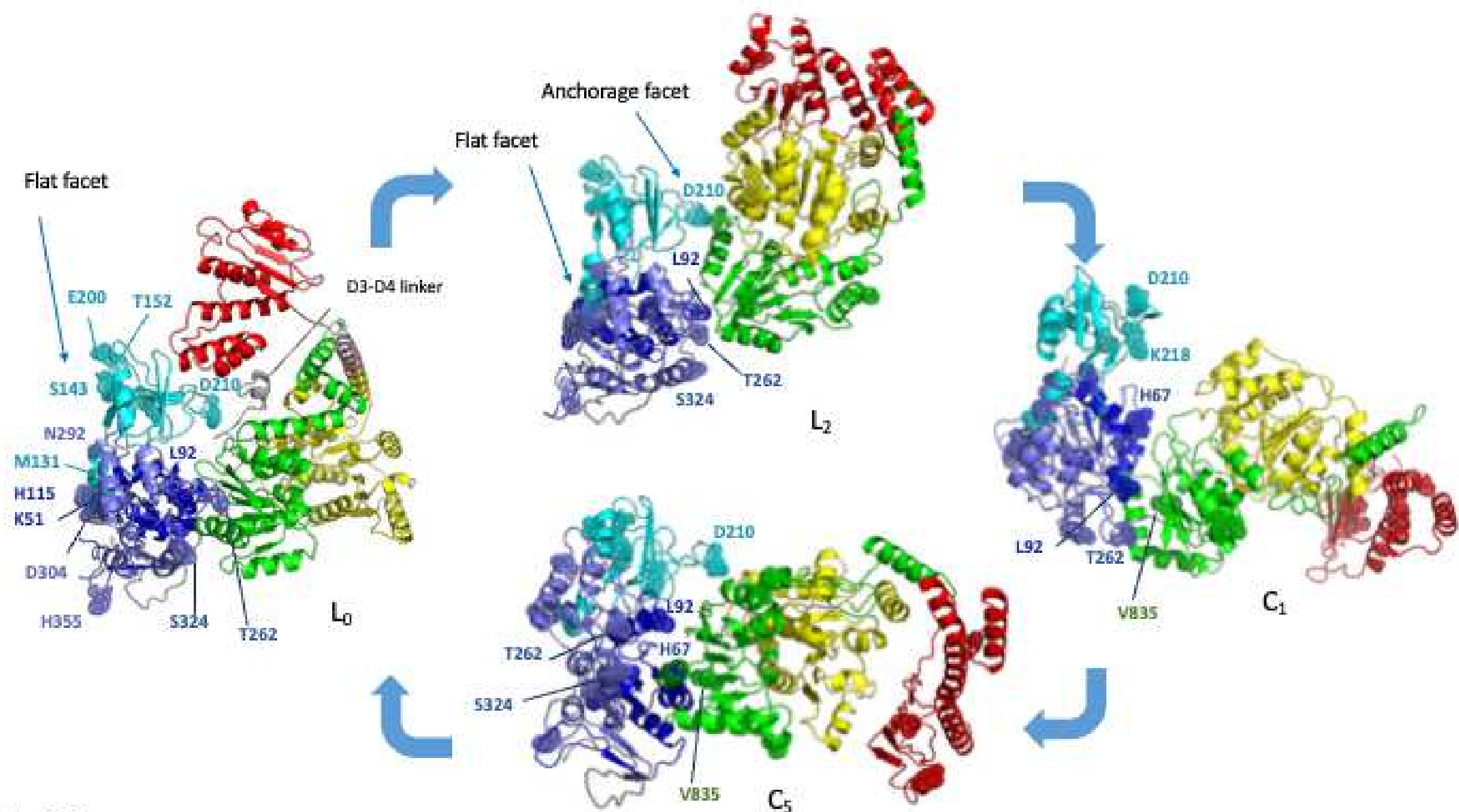


Fig. 4-B

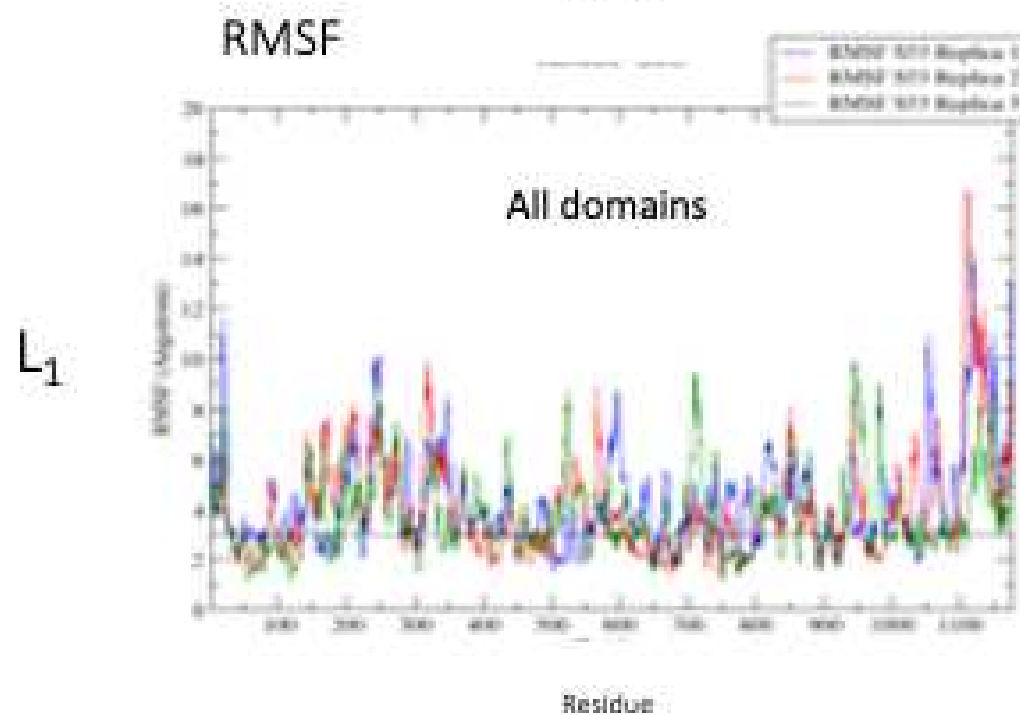
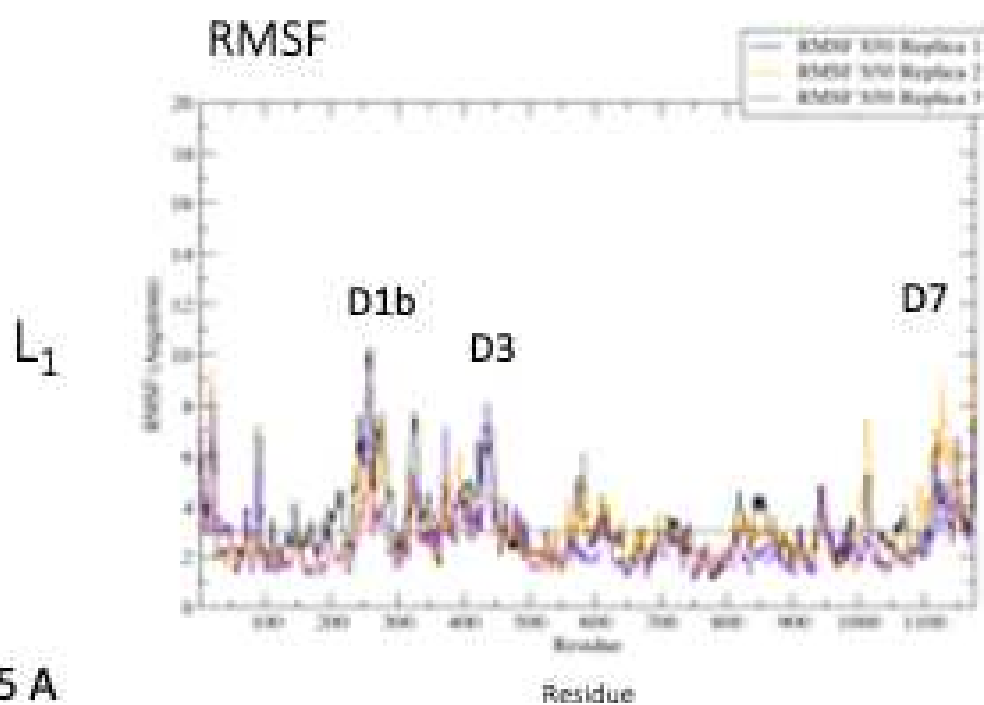
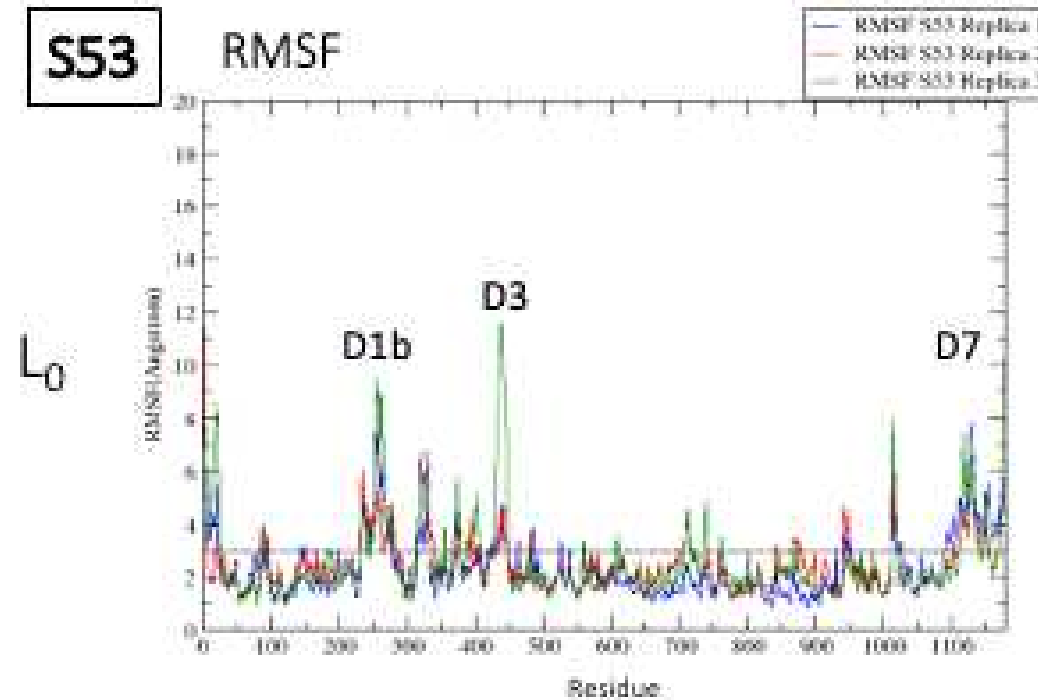
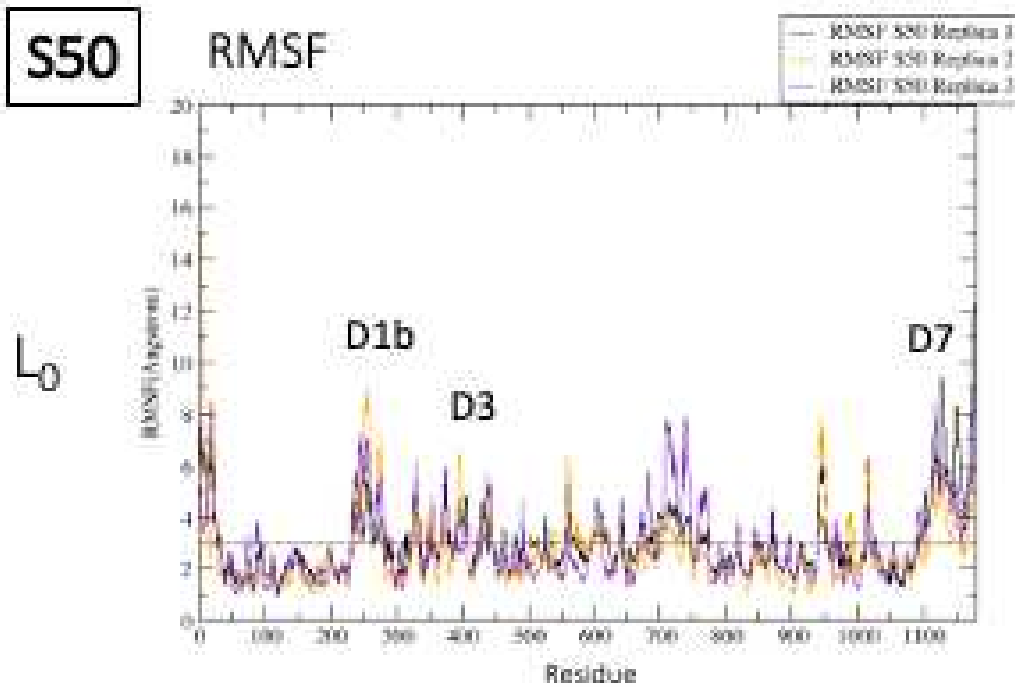


Fig. 5 A

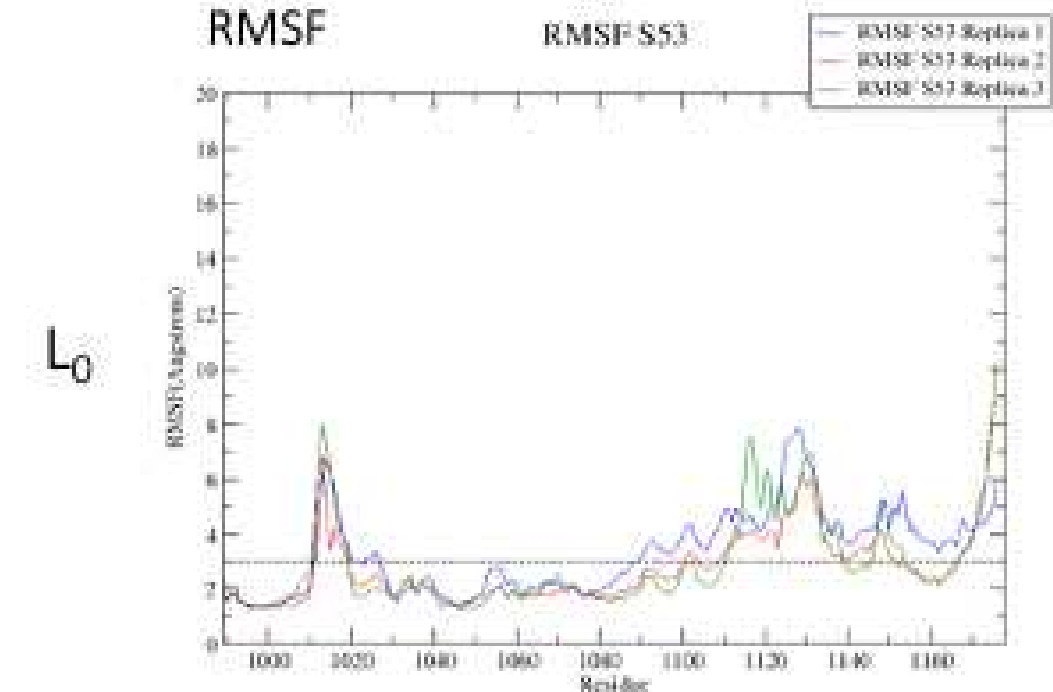
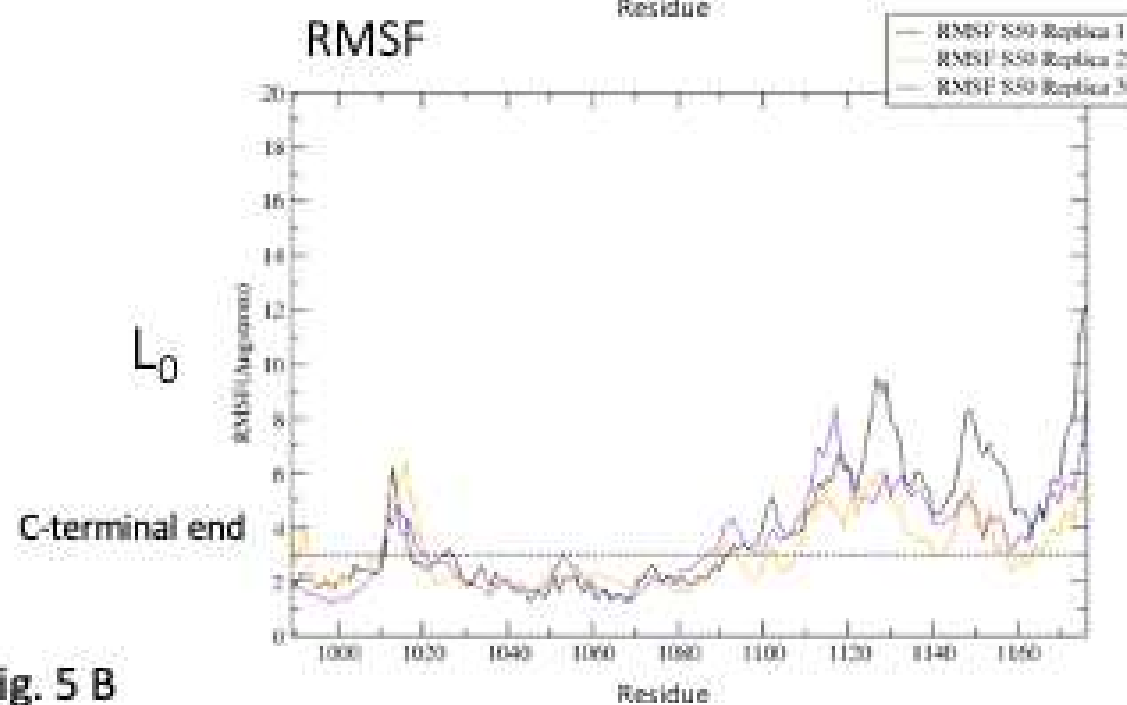
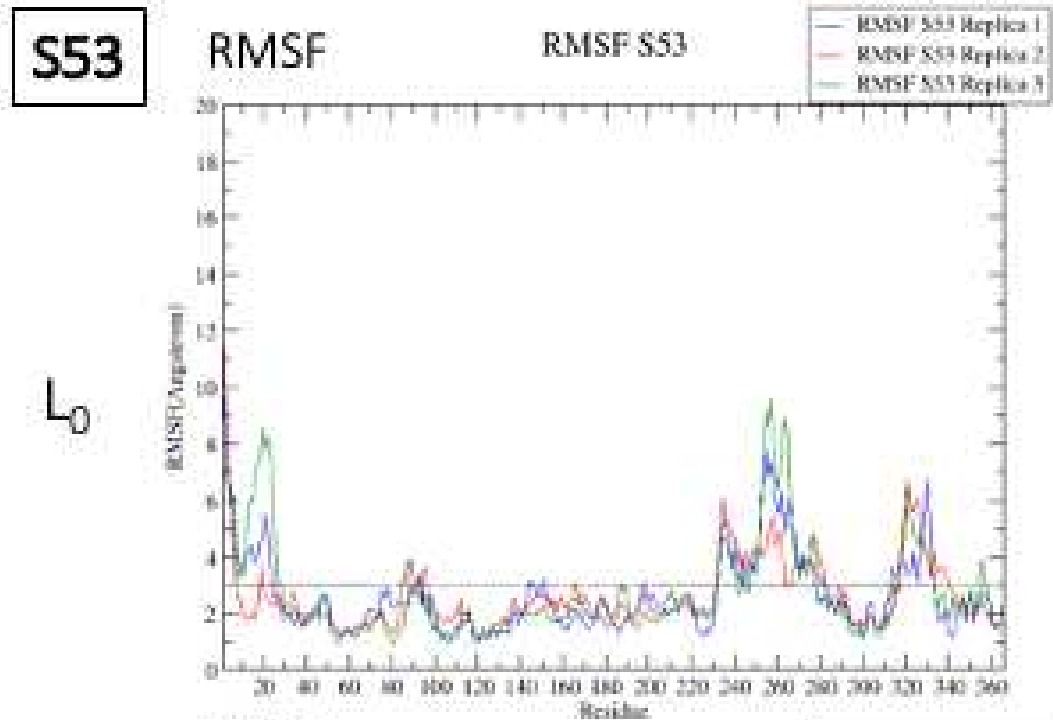
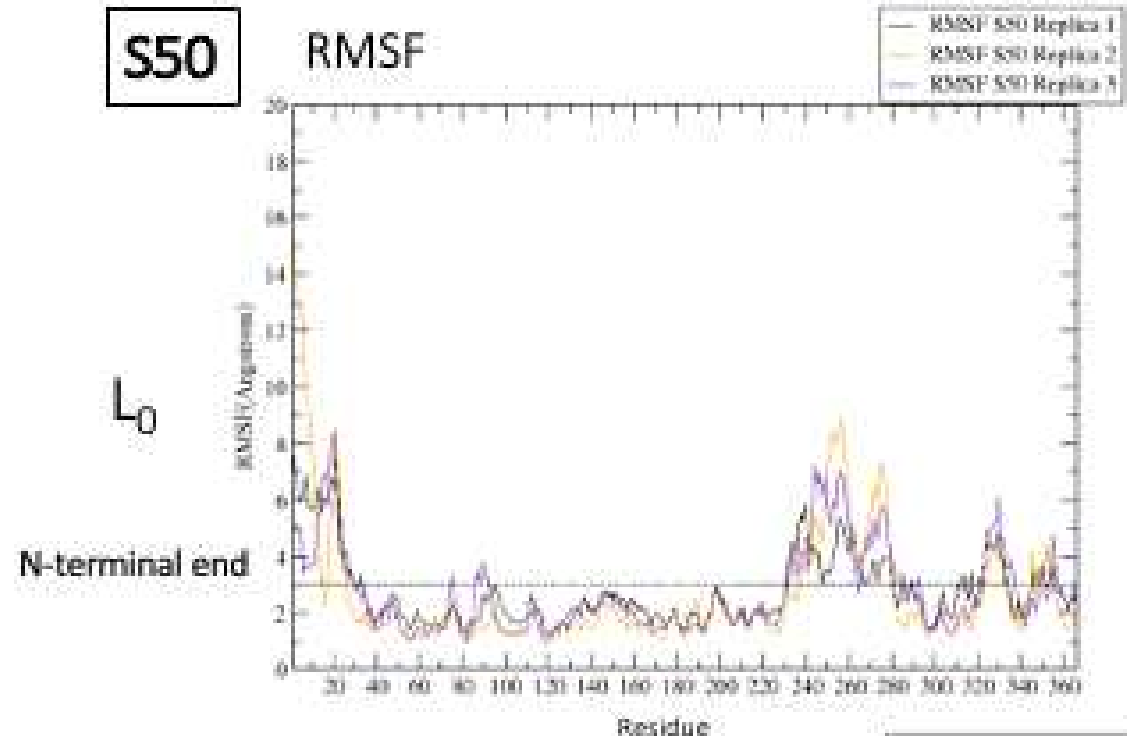
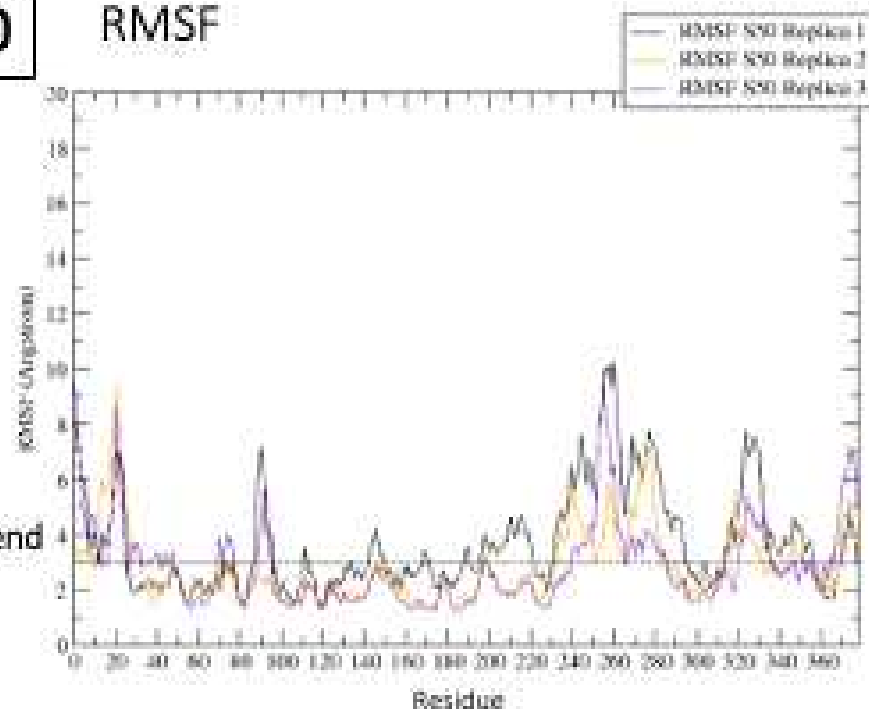


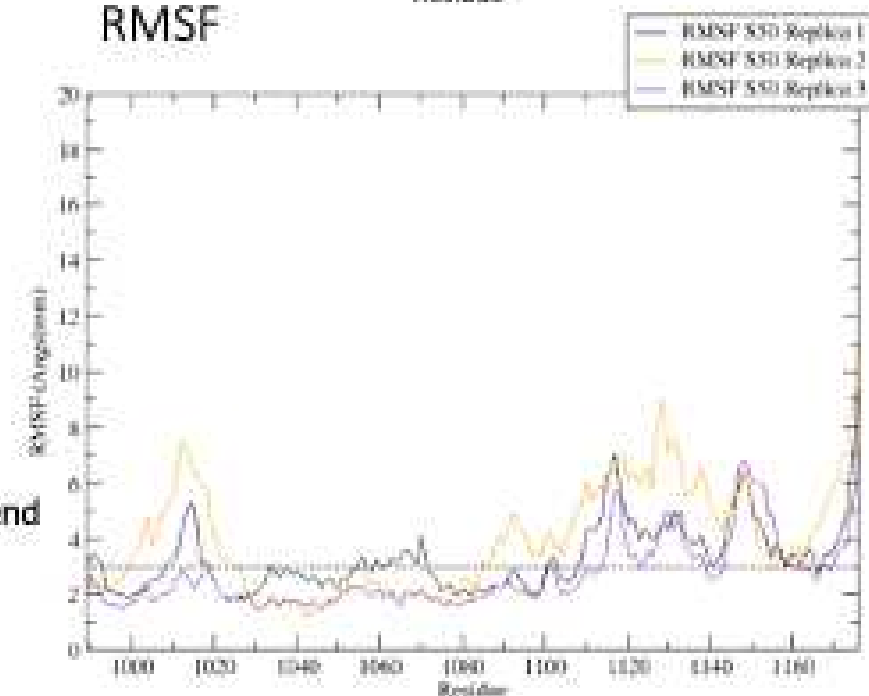
Fig. 5 B

S50

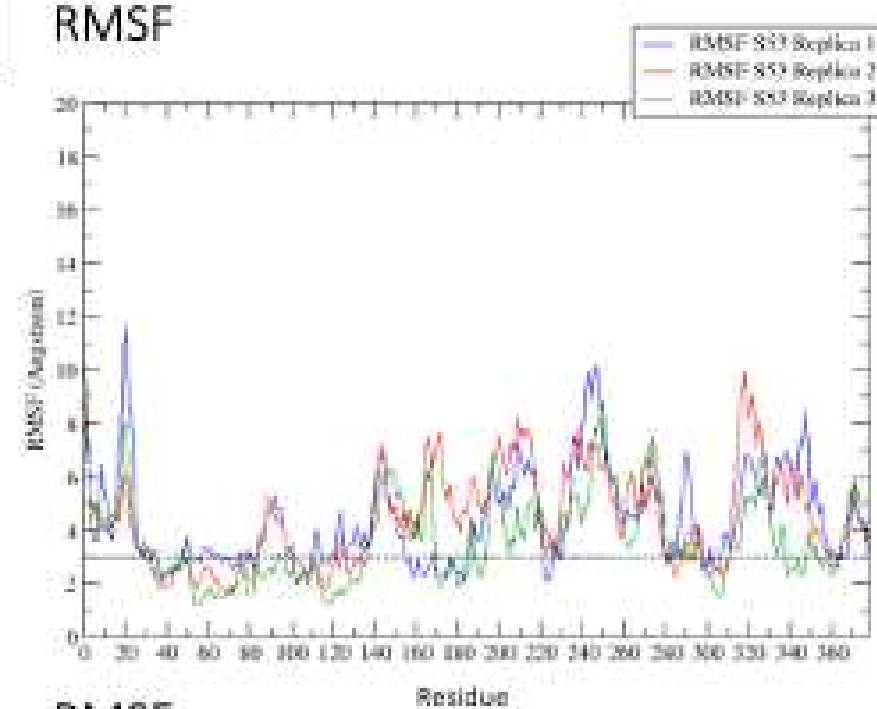
RMSF

 L_1
N-terminal end

RMSF

 L_1
C-terminal end**S53**

RMSF

 L_1 

RMSF

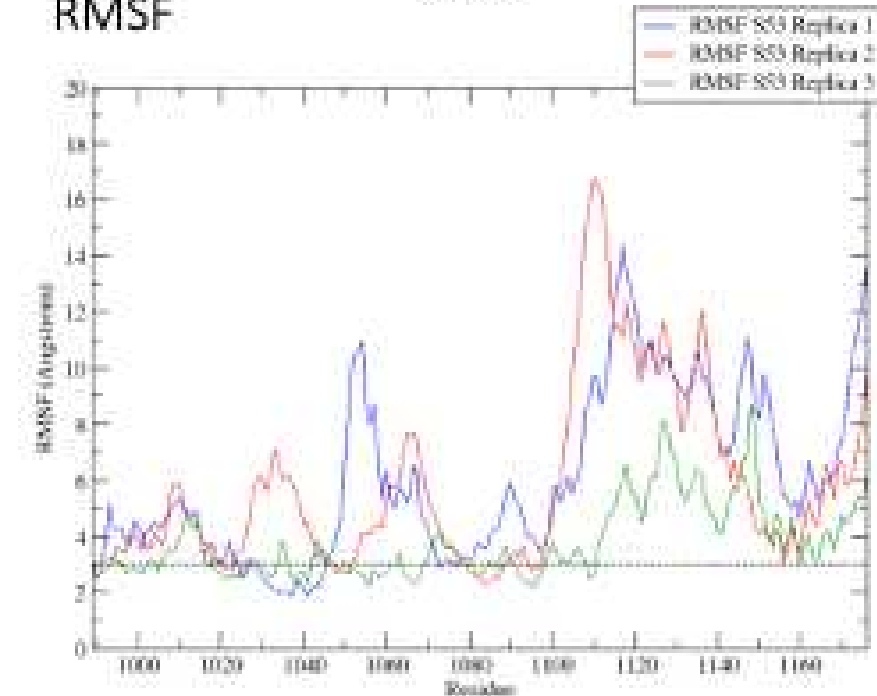
 L_1 

Fig. 5 C

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Suppl1.tiff](#)
- [Suppl2.tiff](#)
- [Suppl3.tiff](#)
- [Suppl4.tiff](#)
- [Suppl5.tiff](#)
- [Suppl6.tiff](#)