Conformational stability of the bacterial adhesin, FimH, with an inactivating mutation

Running title: Stability of inactivating mutant of FimH

Authors: Jenny Liu[1], Luis Amaral[2], Sinan Keten[1, 3]

Affiliations:

[1] Mechanical Engineering Department, Northwestern University

[2] Chemical and Biological Engineering, Northwestern University

[3] Civil Engineering Department, Northwestern University

Corresponding author: Sinan Keten, s-keten@northwestern.edu, 847-491-5282, 2145 Sheridan Road, Tech A133, Evanston, IL 60208-3109

Acknowledgments

We are extremely grateful to Kerim Dansuk for his helpful discussions and to Martha Dunbar for her help with the MD simulations. Jenny Liu thanks the Paul and Daisy Soros Fellowship, the Achievement Rewards for College Scientists (ARCS) Illinois chapter, the Northwestern Quest High Performance Computing Cluster, and the National Institute of General Medical Sciences T32GM008152. This project was also supported by the Office of Naval Research N00014163175 and N000141512701 (Sinan Keten), as well as by the National Science Foundation 1764421-01 and the Simons Foundation 597491-01 (Luis Amaral).

1 Abstract and Keywords

Allostery governing two conformational states is one of the proposed mechanisms for catch-bond behavior in adhesion proteins. In FimH, a catch-bond protein expressed by pathogenic bacteria, separation of two domains disrupts inhibition by the pili domain. Thus, tensile force can induce a conformational change in the lectin domain, from an inactive state to an active state with high affinity. To better understand allosteric inhibition in two-domain FimH (H2 inactive), we use molecular dynamics simulations to study the lectin domain alone, which has high affinity (HL active), and also the lectin domain stabilized in the low-affinity conformation by an Arg-60-Pro mutation (HL mutant). Because ligand-binding induces an allostery-like conformational change in HL mutant, this more experimentally tractable version has been proposed as a "minimal model" for FimH. We find that HL mutant has larger backbone fluctuations than both H2 inactive and HL active, at the binding pocket and allosteric interdomain region. We use an internal coordinate system of dihedral angles to identify protein regions with differences in backbone and sidechain dynamics beyond the putative allosteric pathway sites. By characterizing HL mutant dynamics for the first time, we provide additional insight into the transmission of allosteric information across the lectin domain and build upon structural and thermodynamic data in the literature to further support the use of HL mutant as a "minimal model." Understanding how to alter protein dynamics to prevent the allosteric conformational change may guide drug development to prevent infection by blocking FimH adhesion.

MeSH Keywords: Molecular Dynamics Simulation; Protein Conformation; Protein Domains; Allosteric Regulation; Adhesins, Escherichia coli; Infections

2 Introduction

The bacterial adhesin FimH is one of the most well-characterized model proteins for catch-bond behavior, in which tensile force paradoxically increases ligand affinity and increases the adhesion time.[1,2] For many catch-bond proteins,[2] including FimH, P-selectin,[3] α-catenin,[4] and the $\alpha\beta$ T-cell receptor,[5] tensile force exposes previously buried regions and induces an allosteric conformational change.[6–8] A better understanding of the protein dynamics involved in the activation of FimH[9] could lead to treatments that target the allosteric site, which could prevent bacterial adhesion to the host during infections.[10]

FimH is a 30 kDa two-domain bacterial adhesin found on the fimbrial tips of uropathogenic *E. coli* (UPEC) that binds to mannosylated ligands on urothelial cells.[11] Because it enables bacterial adhesion, which is one of the first steps in urinary tract infections, FimH is a critical virulence factor.[11] During urination, shear forces introduce tension into the FimH-mannose interaction to produce catch-bond behavior, causing FimH to bind more tightly and for longer duration.[2,9,12,13] This catch-bond behavior distinguishes mannosylated ligands on the urinary tract from decoys and limits bacterial removal.[11,14] Several glycomimetic drugs (mannosides) have been developed that competitively inhibit FimH and thus prevent bacterial adhesion.[15,16] However, it is currently not known how to target the allosteric site(s) to block the conformational change.

The mechanism underlying catch-bond behavior involves an allosteric conformational change between an inactive state with low affinity and an active state with high affinity.[2] In the inactive state, the two domains of FimH are close together, which stabilizes the interdomain region between the ligand-binding lectin domain (HL) and the allosterically inhibitory pili domain (HP).[14,17,18] The transition to the active state is thought to occur after the ligand binds to the lectin domain and tensile force pulls the domains apart.[19] Exposing the interdomain region to water, which disrupts inhibition by the pili domain, is thought to induce a conformational change in the lectin domain.[18,20] The conformational change has been described by the width of the β-sandwich fold,[14,17,21] as well as the putative allosteric pathway sites connecting the interdomain region to the binding pocket.[12,20,22]

While two-domain FimH in the inactive conformation (H2 inactive) can undergo a force-induced, allosteric conformational change to the active state, a truncated protein consisting of the lectin domain alone (HL active) is constitutively active.[2,18,22]. However, while the lectin domain with a single Arg-60-Pro amino acid mutation (HL mutant) is stabilized in the inactive conformation,[12] HL mutant undergoes an allostery-like conformational change upon binding mannoside.[22] As a result, Rabbani *et al.*[22] have proposed HL mutant as a "minimal model" for FimH allostery that is smaller, consists of a single domain, and is more experimentally tractable than full-length FimH. Although the structure and function of the HL mutant have been well-characterized, its dynamics have not yet been investigated experimentally or computationally. Studying the dynamics of HL mutant will make it possible to identify protein regions with dynamical differences compared to H2 inactive, which can provide insight into the allosteric conformational change and help design additional mutations that lock HL mutant in the inactive state.

The impact of the Arg-60-Pro mutation on the structure and function of HL mutant has also been well-characterized. After Rodriguez *et al.*[12] selected the Arg-60-Pro mutation to energetically favor the inactive state from a set of trial mutations tested in RosettaDesign, Rabbani *et al.*[22] then confirmed that the backbone structure of HL mutant matches the lectin domain of H2 inactive using crystallography and chemical shift mapping from ^{1}H-^{15}N-HSQC NMR spectroscopy. The similarity in backbone structure, despite the absence of the inhibitory pili domain, suggests that allosteric pathway sites vary in coupling strength to the binding pocket,[20,22] while raising questions about the relative contributions of backbone and sidechain dynamics towards information transfer. Rabbani

et al.[22] also found that the addition of a mannoside ligand (n-heptyl α-D-mannopyranoside) induced a conformational change in HL mutant with NMR peaks that matched HL active at the binding pocket and the β-sandwich, but not near the mutation or the interdomain region. The HL mutant conformational change was hypothesized to be similar to that of full-length FimH due to mannoside binding.[17] However, HL mutant has more than seven times higher affinity for mannoside than H2 inactive.[22] The difference in binding affinity, despite the similarity in backbone structure, suggests further differences in the backbone dynamics or the sidechains.

The dynamics of HL mutant have not yet been studied using molecular dynamics (MD) simulations. However, Interlandi and Thomas[20] isolated the lectin domain from a crystal structure of H2 inactive and used nanosecond-timescale MD simulations in order to identify protein regions with different dynamics than HL active, focusing on the putative allosteric pathway. We seek here to identify the protein regions with dynamical differences across HL mutant, H2 inactive, and HL active, for both the backbone and sidechains. In contrast to the top-down approach that starts with identifying landmarks along the putative allosteric pathway, we employ a bottom-up without defining sites *a priori* by directly comparing the structure and dynamics across the sequence of the lectin domain. Moreover, because using external Cartesian coordinates introduces artifacts from rigid body motion and relative domain motion, we use a system of internal coordinates based on the dihedral angles.[23] To this end, we perform 20 nanosecond (ns)-long all-atomistic MD simulations of HL mutant, H2 inactive, and HL active to identify protein regions with differences in dynamics.

We find that structural differences between HL mutant and H2 inactive are not limited to the clamp segment in the binding pocket, the interdomain loops, or the β-bulge region with the mutation. For the backbone dynamics, we find the greatest differences in the interdomain region and the binding pocket. There are also differences in sidechain orientation and dynamics, beyond the localized effect of having Arg or Pro at position 60. Our analyses using similarity matrices, a common tool in data science, confirms that the differences in backbone dynamics are more distinct than the differences in backbone structure or in sidechain orientation and dynamics.

The implications of our study extend beyond FimH. Our investigation demonstrates the advantages of a system of internal coordinates over Cartesian coordinates for quantifying dynamics. The regions we identify with dynamical differences across HL mutant, H2 inactive, and HL active using dihedral angles are not restricted to the identified landmarks on the putative allosteric pathway and further highlights the power of a bottom-up approach, which may be applicable to proteins where these landmarks are not yet identified.

3 Materials and Methods

3.1 FimH structures

We retrieved crystal structures for HL mutant, H2 inactive, and HL active from the Protein Data Bank (RRID:SCR_012820), as detailed in Table 1. We denote the full-length FimH protein in the inactive conformation as H2 inactive and denote the truncated structures that only include the lectin domain as HL mutant and HL active. For H2 inactive, we focus on two structures (4XOD and 5JQI) that were crystallized by separate groups. Due to natural sequence variations, different strains of *E. coli* produce FimH with slightly different amino acid sequences. To directly compare sidechain dynamics with matching sequences, we mutated 4XOD at three positions *in silico* to match HL mutant and HL active (Table 1). In contrast, we made no changes to 5JQI. Due to the sequence mismatch, sidechain comparisons involving 5JQI were limited, but backbone comparisons were not affected.

HL mutant has a single Arg-60-Pro mutation. We also created and studied a modified HL mutant by changing back the Pro at position 60 to Arg *in silico* using MODELLER.[24] Similarly, for HL active, we studied both the wild-type crystal structure and a modified version with the Arg-60-Pro mutation. For all structures, we investigated systems with and without mannose.

3.2 Equilibrium simulations

We performed all-atomistic MD simulaions using NAMD[25], with the CHARMM force field.[26] Our NAMD simulation parameters and system details are listed in Table 2. We prepared all simulation systems using VMD (Visual Molecular Dynamics).[27] We solvated each protein with enough water molecules to prevent interactions with itself through the periodic boundary conditions.

We performed six replicate simulations for each system. Because we found that replicates starting from a single solvated system were too similar to each other, we solvated, minimized, and equilibrated our replicates separately.

3.3 Backbone and sidechain dihedral angles

We focus on dihedral angles to compare protein structures and dynamics instead of Cartesian coordinates because dihedral angles are better suited for identifying the critical regions driving collective motion in Cartesian space, such as hinges that displace distal elements.[28] Moreover, using dihedral angles avoids artifacts from structural alignment and better captures angular motions that affect large segments in hinged proteins, such as FimH.[23,28,29] We describe backbone motion using the dihedral angles ϕ and ψ (see Fig. S5). We describe sidechain motion using both rotamer angles (χ) and custom-defined pseudo-dihedral angles (υ). Each amino acid has one to five rotamer angles, except for Ala and Gly. To compare the wild-type with the Arg-60-Pro mutation, we omit the last three rotamer angles in Arg. We define the pseudo-dihedral angles, υ, to compare sidechain dynamics when there are differences in the primary sequence, due to either the Arg-60-Pro mutation or the two versions of H2 inactive. We define one pseudo-dihedral angle per amino acid using four atoms: N, CA, and the last two atoms from the most distal χ angle. For example, since Arg has χ_5 calculated by CD-NE-CZ-Nh1, we calculate υ from N-CA-CZ-Nh1 (Fig. S5). For Ala, we use N-CA-CB-H1, and we do not define υ for Gly.

We calculate the average, variance, and standard deviation using circular statistics, which account for the periodic wrapping from $-\pi$ to π.[23] We use the standard deviation to assess the range of backbone and sidechain fluctuations. We compare the standard deviation at each residue to identify protein regions with differences in fluctuation size among the starting structures. Because the backbone dynamics were similar in the presence or absence of mannose, as well as the amino acid at position 60, we pooled replicates for backbone dynamics to increase the statistical power. Specifically, we found no significant dynamical differences in the presence or absence of mannose, and extremely small, yet statistically significant, differences at isolated sites due to the amino acid at position 60 (Fig. S6). However, having Arg or Pro at position 60 affected the sidechain dynamics, so we did not pool replicates. For statistical significance testing, we use Mood's median test from the python package scipy (RRID:SCR_008058).[30] To correct for multiple comparisons, we use the two-stage Benjamini-Krieger-Yekutieli procedure from the python package statsmodels (RRID:SCR_016074),[31,32] with a false discovery rate of $\alpha = 0.05$. For robustness testing, we also implement the Bonferroni correction for multiple comparisons with $p = 0.05$.

3.4 Similarity matrix

After we identified differences between groups of HL mutant and H2 inactive trajectories, we quantify the extent of the difference by constructing similarity matrices that compare each pair of trajectories. As references, we choose one type of comparison where we expect low similarity and two types of comparisons where we expect high similarity. We expect low similarity, or large differences, when we compare HL active with both versions of H2 inactive. In contrast, we expect high similarity when we compare the backbone for both versions of H2 inactive, as well as when we compare replicates. Using this framework with high and low references, we use similarity matrices to compare the structure and dynamics of both the backbone and sidechains. Each similarity value is a non-parametric (rank-based) correlation. While we analyze dynamics with Spearman's correlation, we analyze structure using the equivalent correlation for circular variables described by Mardia and modified by Fisher and Lee.[33] Using these similarity matrices, we quantify the similarity between HL mutant, H2 inactive, and HL active and perform statistical testing using Mood's median test with a Bonferroni correction for multiple comparisons.

4 Results

4.1 Crystal structure comparisons

We use root mean squared displacement of the backbone Cα atoms (Cα-RMSD) to measure crystal structure differences and to broadly identify structural differences across HL mutant, H2 inactive, and HL active (Fig. 1 and S3B). As expected, our analyses show that the Cα-RMSD between HL mutant (5MCA) and the two versions of H2 inactive are quite small (0.76 Å to 4XOD and 0.93 Å to 5JQI). This similarity is consistent with the similarity between HL mutant and another version of H2 inactive (3JWN) reported by Rabbani *et al.*[22] While HL mutant and H2 inactive have similar crystal structures, they are both quite different from HL active (4AUU). The Cα-RMSD between HL active and HL mutant is 3.15 Å and between HL active and the two versions of H2 inactive are 3.30 Å (4XOD) and 3.25 Å (5JQI).

Even though the structural differences between HL mutant and H2 inactive are small, we nonetheless examine the backbone regions with the greatest contributions to the Cα-RMSD differences (Fig. 1B and Fig. S3). As described by Rabbani *et al.*,[22] we find the largest difference at the clamp segment in the binding pocket and few differences within the "β-sandwich." We also detect some differences in an unnamed region on the opposite side of the binding pocket, as well as in the swing and insertion loops in the interdomain region (see Fig. 2 and S1 for named regions).

The difference between HL mutant and H2 inactive at the clamp segment has been attributed to its flexibility in the absence of ligand.[22] In particular, the angular differences found in the backbone dihedrals at the C-terminal end of the clamp segment could act like a "hinge" that displaces distal elements (arrow in Fig. 2). This hinge motion may cause the large displacements at the tip of the clamp identified by Cα-RMSD (Fig. 1). Compared to the localized angular differences (Fig. S3 and S4), the overly broad protein regions identified by Cα-RMSD reveals the limitations of calculations involving structural alignment, which have previously been reported reported in the literature.[23,28] For this reason, regions with large dihedral angle motion may be more dynamically important than those identified by large Cα-RMSD.

4.2 Backbone dynamics using internal coordinates

To avoid the artifacts introduced by structural alignment, we thus consider an internal coordinate system that uses the average backbone dihedral angles[23] to identify differences between average structures obtained from the MD simulations (Fig. 3). Using the dihedral angles to compare HL mutant and H2 inactive, we mostly find small differences in backbone angle orientation across the lectin domains (Fig. 3B). In contrast, we find large differences at the insertion loop and the hinge at the C-terminal end of the clamp segment (arrow label in Fig. 2). Because we also find differences at the hinge in our crystal structure comparisons, we think that the difference in the orientation of the clamp segment is not resolved within the 20-ns simulation timescale. Since we find the largest angle change at the hinge-joint and the largest contribution to Cα-RMSD at the center of the clamp segment, we interpret these results to demonstrate that our approach enables us to detect the joints of regions displaying hinge-like motion in an automated fashion.

We next investigate backbone dynamics for the dihedral angles found in these hinge regions and use the standard deviation to compare the magnitude of the fluctuations (Fig. 3, S7, and S6). We first examine the effect of the Arg-60-Pro mutation alone on HL mutant by changing Pro to the wild-type Arg *in silico*. Analogously, we modified HL active by changing the Arg to Pro at position 60 *in silico* to mimic the mutation. We find that the differences in backbone structure and dynamics due to the Arg-60-Pro mutation are small and isolated for HL mutant and HL active (Fig. S6). In contrast to these small changes due to the amino acid at position 60, we find that HL mutant has larger fluctuations than H2 inactive in the insertion and swing loops of the interdomain region and in the clamp segment of the binding pocket (Fig. 3C). Thus, the backbone regions with larger dynamical differences are also the regions with larger structural differences.

Since it has been hypothesized that the difference in dynamics is related to the exposed interdomain region of HL mutant,[22] we next compare the dynamics of HL mutant and HL active, as both have the interdomain region exposed. Contradicting that hypothesis, we find that HL mutant also has larger fluctuations than HL active in the insertion and swing loops of the interdomain region (Fig. S7D). For the clamp segment, we find that HL mutant is more dynamic at the hinge-joint, while HL active is only slightly more dynamic at the tip of the clamp.

Surprisingly, we also find that HL active and H2 inactive display small differences in dynamics across the protein (Fig. S7E). HL active is only slightly more dynamic in the clamp segment and the C-terminus, with minute differences in dynamics across the protein. We interpret this finding to show that HL active and H2 inactive are more stable compared to HL mutant. The interdomain regions in H2 inactive and HL active may be stabilized by different mechanisms: the pili domain for H2 inactive, and a more rigid fold for HL active.

The presence or absence of the mannose ligand does not result in statistically significant differences for the backbone structure or dynamics. We interpret these results to indicate that the initial protein structure has the largest impact on both structure and dynamics, at least for the up to 20 ns duration of our MD simulations.

4.3 Comparing sidechain orientations and fluctuations

For many proteins, including FimH, ligand-binding also depend on sidechain flexibility and optimal sidechain-ligand interactions.[34] To separate sidechain motion from displacements due to backbone motion, we again use internal angles. We quantify sidechain structure and dynamics using two alternative sets of angles: χ dihedral angles (or rotamer angles) and the custom-defined pseudo-dihedral angles υ. We define one υ angle, using N, Cα, and two distal atoms, for each amino acid except for Gly (see Methods and Fig. S5). Each amino acid, except for Gly and Ala, also has one

to five χ rotamer angles. By construction, the υ angles have broader angle distributions than the χ angles and enable us to compare the two versions of H2 inactive in spite of their minor sequence variations.

While HL mutant and H2 inactive have similar backbone structures, sidechain orientations estimated from the crystal structure suggest that the geometries of their binding pockets are quite different (Fig. 1C). To separate sidechain motion from displacements due to backbone motion, we first compare the average sidechain orientations of HL mutant and H2 inactive using υ (Fig. S8) and χ (Fig. S10) angles. Our data show that the differences extend beyond the β-bulge segment with the Arg-60-Pro mutation. Indeed, in addition to the β-bulge segment, we find nine large differences in sidechain orientation across the protein, including the hinge-joint of the clamp segment, the swing loop, the parasteric site, and several unnamed regions. For example, we identify a difference in the sidechain orientation of Phe-144 in the β-sheet between the parasteric site and the linker loop. Phe-144 touches the tip of the clamp segment in HL mutant but not in H2 inactive. We find large differences in the orientation of Phe-43, Tyr-108, and Phe-144 using both υ and χ, but we only find a statistically significant difference at Arg-92 using υ, which combines all five χ angles for Arg into one measure between the base and the tip (Fig. S8 and S10).

We also find differences in the sidechain orientation and dynamics between HL mutant and H2 inactive that extend beyond the Arg-60-Pro mutation. To test the impact of the amino acid at position 60 on local and global sidechain dynamics, we modified HL mutant *in silico* to change Pro to Arg at position 60. The cyclic Pro at residue 60 in HL mutant has smaller fluctuations than the longer Arg, and this is the only statistically significant difference (Fig. S9B, C). However, when we compare HL mutant and H2 inactive, we find that the sidechain dynamics measured with υ are different at two positions in addition to position 60: at the edge of the clamp segment and in the swing loop (Fig. 4C). Using χ instead of υ, we find additional differences in the dynamics of the pocket zipper within the clamp segment (Fig. S11B). Overall, the protein regions identified using χ and υ are very similar.

As expected, we find that HL active is different from both HL mutant and H2 inactive in both sidechain structure and dynamics at many positions across the protein (Fig. 4D and S11C). In particular, many of these differences fall outside of the landmarks of the putative allosteric pathway[12] which are frequently the focus of attention and are primarily studied using backbone dynamics.[20,22]

4.4 Robustness checking using a similarity matrix approach

In the previous section, we identified protein regions on HL mutant and H2 inactive with statistically significant differences in structure and dynamics by comparing groups of trajectories. Nonetheless, it remains difficult to quantify how different HL mutant is from H2 inactive relative to variations that may occur across replicates. To more fully assess the dissimilarity between the dynamical behavior of HL mutant and H2 inactive, we directly compare pairs of trajectories by calculating rank-based correlations and constructing similarity matrices (Fig. 5). We use the Spearman correlation for angle fluctuation size and an equivalent correlation for circular variables for average angle orientation. The similarity matrices allow us to compare each pair of trajectories, complementing our previous comparisons across groups.

We define the reference similarity values as: the low similarity expected for HL active vs H2 inactive, very high similarities for replicates, and high similarity for backbone comparisons of the two versions of H2 inactive. We use these reference values to compare the differences between HL mutant and H2 inactive.

As expected, we find that HL active is very different from both HL mutant and H2 inactive for all similarity matrices, and HL active replicates are relatively similar to each other. We find that the two H2 inactive structures are very similar to each other for backbone comparisons, but different for sidechain comparisons. These differences are smaller than the ones between H2 inactive and HL active, but are slightly larger than the ones between replicates of each version of H2 inactive (Fig. S12 and S13).

We find the replicates for HL mutant are similar to each other, except for the average backbone structure of HL mutant. One possible explanation is that the larger fluctuations found in HL mutant make the replicates less similar to each other. We find that having separate solvation, minimization, and equilibration preparations accentuated the difference between technical replicates, compared to having replicate production runs following a single preparation. Thus, for backbone structure, we find the lower similarity values for HL mutant vs H2 inactive comparisons is about the same as those for replicates of HL mutant (Fig. S12). In comparison, for backbone dynamics, we find clear differences between HL mutant and H2 inactive that are larger than the differences between replicates, yet smaller than the differences for HL active compared to both H2 inactive and HL mutant (Fig. S13).

We also find that average sidechain orientations had lower similarity overall for both the rotamer (χ) and the custom pseudo-dihedral (υ) angles. Despite overall lower similarities, we are still able to use the pseudo-dihedral angles to distinguish the two versions of H2 inactive, which have amino acid sequence differences at three positions in the lectin domain (see Table 1). We also identify differences between HL mutant and H2 inactive (Fig. S12). For sidechain dynamics, we find slightly lower similarity values between HL mutant and H2 inactive relative to the replicates, which is comparable to the low similarity between the two versions of H2 inactive.

We find that the reduction in similarity for HL mutant and H2 inactive, relative to the high similarity between replicates, is smaller for sidechain dynamics than for backbone dynamics (Fig. S13). Across the four types of comparisons, we find that the differences primarily depend on whether the the initial protein structure is HL mutant, H2 inactive, or HL active (Fig. 5). We find larger differences due to the initial protein structure than those due to the amino acid at position 60 or to the presence of the ligand, which is consistent with our results in the previous section.

To compare the protein regions with differences in backbone and sidechain fluctuations, we show the regions identified with the variance as the volume of the sphere (Fig. 6). Different regions are identified for backbone and sidechain dynamics. For example, position 60 in the β-bulge and nearby regions showed a larger difference for Pro vs Arg in the sidechain than in the backbone. The regions where HL mutant had larger backbone fluctuations than H2 inactive include the interdomain region and the hinge of the clamp segment. Larger fluctuations at the binding and allosteric sites may explain why HL mutant can undergo an allostery-like conformational change without the tensile force required to separate the two domains of H2 inactive.

5 Discussion

HL mutant (Arg-60-Pro) has been proposed as a "minimal model" of FimH allostery because the single-domain structure is more experimentally tractable than H2 inactive. Moreover, HL mutant undergoes an allostery-like conformational change upon binding mannoside ligand in contrast to the wild-type HL active.[22] The HL mutant crystal structure was shown to match H2 inactive, except at the clamp segment of the binding pocket, and this difference was attributed to the flexibility of the clamp segment.[22] However, the dynamics of HL mutant have not yet been studied experimentally, and this is the first computational characterization using MD simulations.

Although HL mutant and H2 inactive were hypothesized to have matching structures and similar dynamics, our analysis shows that large differences in average backbone structure exist in the interdomain loops, in addition to the clamp segment of the binding pocket. In these regions, HL mutant also has larger backbone fluctuations than H2 inactive. We further used a technique from data science to show that these differences are robust compared to the variation among replicate simulations.

Remarkably, switching the Arg-60-Pro mutation in HL mutant to Arg *in silico* does not resolve the dynamical differences. This suggests that a combination of structural differences beyond position 60 and the exposure of the interdomain region to water are responsible for the structural and dynamical differences between HL mutant and H2 inactive.

These structural and dynamical differences likely contribute to the more than seven times higher affinity to mannoside for HL mutant compared to H2 inactive,[22] but their role in the mannoside-induced conformational change remains unknown. Rabbani *et al.*[22] has proposed studying additional mutations to HL mutant in order to prevent the conformational change. This would prevent the high-affinity state that strengthens bacterial adhesion during urination, which limits clearance and prolongs infections.[11,15] Studies of single mutations suggest that the allosteric coupling to the binding pocket is stronger at position 60 than at the interdomain loops.[12,20,22] This is further supported by our finding that HL mutant has larger fluctuations at the binding pocket and the interdomain loops than H2 inactive. Thus, our dynamics characterization further supports the need to determine the structure of the mannoside-bound HL mutant.[22] In addition, greater backbone flexibility in the insertion and swing loops of the interdomain region may be useful for predicting additional mutations. Mutations that stabilize the interdomain region may inhibit the conformational change upon mannoside-binding. On the computational side, we will perform eigenmode analyses on longer simulations to clarify the relationship between these dynamical differences and the mannoside-induced conformational change.

In addition to comparing HL mutant and H2 inactive, we also compared these structures to HL active. Interlandi and Thomas[20] found that differences between inactive and active FimH were isolated to the six sites defining the putative allosteric pathway in their study of backbone structure and dynamics using Cα-RMSD. In a later study, Rabbani *et al.*[22] focused on a subset of these sites when comparing inactive and and active FimH using Cα-RMSD. However, the alignment step in Cα-RMSD calculations results in artifacts since angular fluctuations displace distal elements.[23,28,29] Because we use an internal coordinate system of dihedral angles, we identify more localized differences. Thus, our backbone comparisons identify large differences at specific angles along the allosteric pathway.

When we extended our analyses from the backbone to the sidechains, we find that HL active is dynamically different from HL mutant and H2 inactive at multiple regions beyond the allosteric pathway using both χ and υ. Thus, our comparisons of HL mutant, H2 inactive, and HL active demonstrate that a bottom-up approach can identify differences in structure and dynamics that are not restricted to the putative allosteric pathway. Beyond our analyses of a "minimal model" for FimH allostery, our bottom-up approach and use of techniques from data science may be applicable to comparing the structure and dynamics of other proteins.

References

[1] E. V. Sokurenko, V. Vogel, and W. E. Thomas. Catch-bond mechanism of force-enhanced adhesion: counterintuitive, elusive, but ... widespread? Cell Host Microbe, 4(4):314–23, 2008. ISSN 1931-3128. doi: 10.1016/j.chom.2008.09.005. URL http://dx.doi.org/10.1016/j.chom.2008.09.005.

[2] T. T. Waldron and T. A. Springer. Transmission of allostery through the lectin domain in selectin-mediated cell adhesion. Proc Natl Acad Sci U S A, 106(1):85–90, 2009. ISSN 0027-8424 (Print)1091-6490 (Electronic). doi: 10.1073/pnas.0810620105. URL http://dx.doi.org/10.1073/pnas.0810620105.

[3] N. Ishiyama, R. Sarpal, M. N. Wood, S. K. Barrick, T. Nishikawa, H. Hayashi, A. B. Kobb, A. S. Flozak, A. Yemelyanov, R. Fernandez-Gonzalez, S. Yonemura, D. E. Leckband, C. J. Gottardi, U. Tepass, and M. Ikura. Force-dependent allostery of the alpha-catenin actin-binding domain controls adherens junction dynamics and functions. Nat Commun, 9(1):5121, 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-07481-7. URL http://dx.doi.org/10.1038/s41467-018-07481-7.

[4] D. K. Das, Y. Feng, R. J. Mallis, X. Li, D. B. Keskin, R. E. Hussey, S. K. Brady, J. H. Wang, G. Wagner, E. L. Reinherz, and M. J. Lang. Force-dependent transition in the t-cell receptor beta-subunit allosterically regulates peptide discrimination and pmhc bond lifetime. Proc Natl Acad Sci U S A, 112(5):1517–22, 2015. ISSN 0027-8424. doi: 10.1073/pnas.1424829112. URL http://dx.doi.org/10.1073/pnas.1424829112.

[5] R. Nussinov and C. J. Tsai. Allostery in disease and in drug discovery. Cell, 153(2):293–305, 2013. ISSN 0092-8674. doi: 10.1016/j.cell.2013.03.034. URL http://dx.doi.org/10.1016/j.cell.2013.03.034.

[6] N. V. Dokholyan. Controlling allosteric networks in proteins. Chem Rev, 116 (11):6463–87, 2016. ISSN 0009-2665. doi: 10.1021/acs.chemrev.5b00544. URL http://dx.doi.org/10.1021/acs.chemrev.5b00544.

[7] S. J. Wodak, E. Paci, N. V. Dokholyan, I. N. Berezovsky, A. Horovitz, J. Li, V. J. Hilser, I. Bahar, J. Karanicolas, G. Stock, P. Hamm, R. H. Stote, J. Eberhardt, Y. Chebaro, A. Dejaegere, M. Cecchini, J. P. Changeux, P. G. Bolhuis, J. Vreede, P. Faccioli, S. Orioli, R. Ravasio, L. Yan, C. Brito, M. Wyart, P. Gkeka, I. Rivalta, G. Palermo, J. A. McCammon, J. Panecka-Hofman, R. C. Wade, A. Di Pizio, M. Y. Niv, R. Nussinov, C. J. Tsai, H. Jang, D. Padhorny, D. Kozakov, and T. McLeish. Allostery in its many disguises: From theory to applications. Structure, 27(4): 566–78, 2019. ISSN 0969-2126 (Print)1878-4186 (Electronic). doi: 10.1016/j.str.2019.01.003. URL http://dx.doi.org/10.1016/j.str.2019.01.003.

[8] D. I. Kisiela, H. Avagyan, D. Friend, A. Jalan, S. Gupta, G. Interlandi, Y. Liu, V. Tchesnokova, V. B. Rodriguez, J. P. Sumida, R. K. Strong, X. R. Wu, W. E. Thomas, and E. V. Sokurenko. Inhibition and reversal of microbial attachment by an antibody with parasteric activity against the fimh adhesin of uropathogenic e. coli. PLoS Pathog, 11(5):e1004857, 2015. ISSN 1553-7366. doi: 10.1371/journal.ppat.1004857. URL http://dx.doi.org/10.1371/journal.ppat.1004857.

[9] C. N. Spaulding and S. J. Hultgren. Adhesive pili in uti pathogenesis and drug development. Pathogens, 5(1), 2016. ISSN 2076-0817 (Print)2076-0817. doi: 10.3390/pathogens5010030. URL http://dx.doi.org/10.3390/pathogens5010030.

[10] E. V. Sokurenko, V. Vogel, and W. E. Thomas. Catch-bond mechanism of force-enhanced adhesion: counterintuitive, elusive, but ... widespread? Cell Host Microbe, 4(4):314–23, 2008. ISSN 1931-3128. doi: 10.1016/j.chom.2008.09.005. URL http://dx.doi.org/10.1016/j.chom.2008.09.005.

[11] W. E. Thomas, V. Vogel, and E. Sokurenko. Biophysics of catch bonds. Annu Rev Biophys, 37:399–416, 2008. ISSN 1936-122X (Print)1936-122x. doi: 10.1146/annurev.biophys.37.032807.125804. URL http://dx.doi.org/10.1146/annurev.biophys.37.032807.125804.

[12] V. Kalas, J. S. Pinkner, T. J. Hannan, M. E. Hibbing, K. W. Dodson, A. S. Holehouse, H. Zhang, N. H. Tolia, M. L. Gross, R. V. Pappu, J. Janetka, and S. J. Hultgren. Evolutionary fine-tuning of conformational ensembles in fimh during host-pathogen interactions. Sci Adv, 3(2):e1601944, 2017. ISSN 2375-2548. doi: 10.1126/sciadv.1601944. URL http://dx.doi.org/10.1126/sciadv.1601944.

[13] L. K. Mydock-McGrane, T. J. Hannan, and J. W. Janetka. Rational design strategies for fimh antagonists: new drugs on the horizon for urinary tract infection and crohn's disease. Expert Opin Drug Discov, 12(7):711–731, 2017. ISSN 1746-0441. doi: 10.1080/17460441.2017.1331216. URL http://dx.doi.org/10.1080/17460441.2017.1331216.

[14] R. Hevey. Strategies for the development of glycomimetic drug candidates. Pharmaceuticals (Basel), 12(2), 2019. ISSN 1424-8247 (Print)1424-8247. doi: 10.3390/ph12020055. URL http://dx.doi.org/10.3390/ph12020055.

[15] M. M. Sauer, R. P. Jakob, J. Eras, S. Baday, D. Eris, G. Navarra, S. Berneche, B. Ernst, T. Maier, and R. Glockshuber. Catch-bond mechanism of the bacterial adhesin fimh. Nat Commun, 7:10738, 2016. ISSN 2041-1723. doi: 10.1038/ncomms10738. URL http://dx.doi.org/10.1038/ncomms10738.

[16] P. Aprikian, V. Tchesnokova, B. Kidd, O. Yakovenko, V. Yarov-Yarovoy, E. Trinchina, V. Vogel, W. Thomas, and E. Sokurenko. Interdomain interaction in the fimh adhesin of escherichia coli regulates the affinity to mannose. J Biol Chem, 282(32):23437–46, 2007. ISSN 0021-9258 (Print)0021-9258. doi: 10.1074/jbc.M702037200. URL http://dx.doi.org/10.1074/jbc.M702037200.

[17] O. Yakovenko, V. Tchesnokova, E. V. Sokurenko, and W. E. Thomas. Inactive conformation enhances binding function in physiological conditions. Proc Natl Acad Sci U S A, 112(32):9884–9, 2015. ISSN 0027-8424. doi: 10.1073/pnas.1503160112. URL http://dx.doi.org/10.1073/pnas.1503160112.

[18] G. Interlandi and W. E. Thomas. Mechanism of allosteric propagation across a beta-sheet structure investigated by molecular dynamics simulations. Proteins, 84(7):990–1008, 2016. ISSN 0887-3585. doi: 10.1002/prot.25050. URL http://dx.doi.org/10.1002/prot.25050.

[19] I. Le Trong, P. Aprikian, B. A. Kidd, M. Forero-Shelton, V. Tchesnokova, P. Rajagopal, V. Rodriguez, G. Interlandi, R. Klevit, V. Vogel, R. E. Stenkamp, E. V. Sokurenko, and W. E. Thomas. Structural basis for mechanical force regulation of the adhesin fimh via finger trap-like sheet twisting. Cell, 141(4):645–55, 2010. ISSN 0092-8674 (Print)1097-4172 (Electronic). doi: 10.1016/j.cell.2010.03.038. URL http://dx.doi.org/10.1016/j.cell.2010.03.038.

[20] S. Rabbani, B. Fiege, D. Eris, M. Silbermann, R. P. Jakob, G. Navarra, T. Maier, and B. Ernst. Conformational switch of the bacterial adhesin fimh in the absence of the regulatory domain: Engineering a minimalistic allosteric system. J Biol Chem, 293(5):1835–1849, 2018. ISSN 0021-9258. doi: 10.1074/jbc.M117.802942. URL http://dx.doi.org/10.1074/jbc.M117.802942.

[21] A. Altis, P. H. Nguyen, R. Hegger, and G. Stock. Dihedral angle principal component analysis of molecular dynamics simulations. J Chem Phys, 126(24):244111, 2007. ISSN 0021-9606 (Print)0021-9606. doi: 10.1063/1.2746330. URL http://dx.doi.org/10.1063/1.2746330.

[22] N. Eswar, B. John, N. Mirkovic, A. Fiser, V. A. Ilyin, U. Pieper, A. C. Stuart, M. A. Marti-Renom, M. S. Madhusudhan, B. Yerkovich, and A. Sali. Tools for comparative protein structure modeling and analysis. Nucleic Acids Res, 31(13):3375–80, 2003. ISSN 0305-1048. doi: 10.1093/nar/gkg543. URL http://dx.doi.org/10.1093/nar/gkg543.

[23] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. Scalable molecular dynamics with namd. J Comput Chem, 26(16):1781–802, 2005. ISSN 0192-8651 (Print)0192-8651. doi: 10.1002/jcc.20289. URL http://dx.doi.org/10.1002/jcc.20289.

[24] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmuller, and Jr. MacKerell, A. D. Charmm36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods, 14(1):71–73, 2017. ISSN 1548-7091. doi: 10.1038/nmeth.4067. URL http://dx.doi.org/10.1038/nmeth.4067.

[25] W. Humphrey, A. Dalke, and K. Schulten. Vmd: visual molecular dynamics. J Mol Graph, 14 (1):33–8, 27–8, 1996. ISSN 0263-7855 (Print)0263-7855. doi: 10.1016/0263-7855(96)00018-5. URL http://dx.doi.org/10.1016/0263-7855(96)00018-5.

[26] F. Sittel, A. Jain, and G. Stock. Principal component analysis of molecular dynamics: on the use of cartesian vs. internal coordinates. J Chem Phys, 141(1):014111, 2014. ISSN 0021-9606. doi: 10.1063/1.4885338. URL http://dx.doi.org/10.1063/1.4885338.

[27] Y. Mu, P. H. Nguyen, and G. Stock. Energy landscape of a small peptide revealed by dihedral angle principal component analysis. Proteins, 58(1):45–52, 2005. ISSN 0887-3585. doi: 10.1002/prot.20310. URL http://dx.doi.org/10.1002/prot.20310.

[28] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. URL http://www.scipy.org/. [Online; accessed Oct, 2018].

[29] Josef Perktold, Skipper Seabold, Jonathan Taylor, et al. statsmodels, 2009–. URL http://statsmodels.org/. Online; accessed Oct, 2018.

[30] Yoav Benjamini, Abba M. Krieger, and Daniel Yekutieli. Adaptive linear step-up procedures that control the false discovery rate. Biometrika, 93(3):491–507, 2006. doi: https://doi.org/10.1093/biomet/93.3.491.

[31] NI Fisher and AJ Lee. A correlation coefficient for circular data. Biometrika, 70(2):327–332, 1983.

[32] A. Wellens, M. Lahmann, M. Touaibia, J. Vaucher, S. Oscarson, R. Roy, H. Remaut, and J. Bouckaert. The tyrosine gate as a potential entropic lever in the receptor-binding site of the bacterial adhesin fimh. Biochemistry, 51(24):4790–9, 2012. ISSN 0006-2960. doi: 10.1021/bi300251r. URL http://dx.doi.org/10.1021/bi300251r.

[33] R. B. Best, X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, and Jr. Mackerell, A. D. Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. J Chem Theory Comput, 8(9):3257–3273, 2012. ISSN 1549-9618. doi: 10.1021/ct300400x. URL http://dx.doi.org/10.1021/ct300400x.

Tables

Table 1: **PDB structures studied.** For H2 inactive, both 4XOD and 5JQI had differences in the primary sequence. For 4XOD, we mutated three positions to match the primary sequence for HL active. For HL mutant, we studied the effect of changing Pro to Arg. For HL active, we studied the effect of introducing the Arg-60-Pro mutation.

Name	HL mutant[20]	HL active[32]	H2 inactive v1[15]	H2 inactive v2[12]
PDBID	5MCA	4AUU	4XOD	5JQI
Resolution (Å)	1.604	1.5	1.14	1.982
Residue 60	Pro & Arg	Arg & Pro	Arg	Arg
Residue 27	Val	Val	Ala-27-Val	Ala
Residue 70	Asn	Asn	Ser-70-Asn	Ser
Residue 78	Ser	Ser	Asn-78-Ser	Asn

Table 2: **Details of Molecular dynamics simulations** We set up systems using VMD

System dimensions	60 Åx 62 Åx 93 Åfor HL active
	60 Åx 72 Åx 123 Åfor H2 inactive
System sizes	>32,000 for HL active
	>60,000 for H2 inactive
Solvation	TIP3P water model with 16 Å padding on each side
Ionization	NaCl ions to neutralize and achieve 50 mM salt conc.
Simulation engine	NAMD[23]
Ensemble	NPT
Temperature	300 K
Pressure	1 atm
Non-bonded interactions	Lennard-Jones potential (12 Å cutoff[33])
Electrostatic interactions	Particle-Mesh Ewald sum method[23]
Timestep	1 fs
Coordinates saved every	1 ps
Energy-minimization	Conjugate gradient algorithm in NAMD
	10,000 steps with protein fixed
	10,000 steps with no atoms fixed
Equilibration	1 ns
Simulation	No ligand: 20 ns
	Ligand: disassociation or up to 20 ns

Figure Captions

Figure 1: **The Arg-60-Pro mutant has a backbone structure similar to that of inactive FimH but with a partially closed binding pocket.** (A) NewCartoon representation of the backbone structures for H2 inactive with the pili domain cropped (grey, PDBID 4XOD, low-affinity) overlaid by the lectin domain structures: HL active (PDBID 4AUU, orange, high-affinity) and HL mutant with Arg-60-Pro (PDBID 5MCA, purple). We highlight residue 60, the mutation site, with a bead. We use color to illustrate the displacement from H2 inactive for (B) the backbone and (C) the surface rendering of predicted sidechain orientation. The mannose ligand is shown in blue and was positioned after alignment with high-affinity, two-domain FimH (PDBID 1KLF).

Figure 2: **Angle differences between HL mutant and H2 inactive.** (A) NewCartoon representation of the HL mutant backbone shows large changes in backbone dihedral angle differences at the C-terminal end of the clamp segment that may act like a hinge-joint (arrow). (B) Putative allosteric pathway sites[10], the parasteric site[8], and residue 60 (bead).

Figure 3: **Localized differences in the structure between HL mutant and H2 inactive cause differences in dynamics.** (A) Secondary structure of H2 inactive (top) and HL mutant (bottom) with features. (B) Protein structure comparison between HL mutant (orange square) and H2 inactive (black diamond) used average backbone dihedral angles ϕ and ψ, which provide an internal frame of reference. Structural differences are localized to the hinge of the clamp loop in the binding pocket and the insertion loop in the interdomain region, which is exposed to water in HL mutant. (C) Backbone dynamics comparison used the standard deviations of ϕ and ψ, quantifying the magnitude of angle fluctuations. HL mutant has larger fluctuations (red lines) at the clamp loop and the insertion loop. For statistically significant differences, the median values are shown.

Figure 4: **Differences in the sidechain dynamics between HL mutant and H2 inactive affect the binding pocket, and both are different from HL active.** (A) Secondary structure of HL active (top), H2 inactive (middle) and HL mutant (bottom) with features. Sidechain dynamics were compared using the standard deviation of the custom dihedral angles υ. (B) Arg-60 in H2 inactive (blue diamonds) has larger fluctuations than Pro-60 in HL mutant (orange squares). HL mutant also has larger fluctuations in the exposed swing loop. (C, D) Comparisons of HL active (black circles) against HL mutant and H2 inactive show differences in sidechain dynamics beyond the allosteric pathway sites.

Figure 5: **Differences between HL mutant and H2 inactive in comparisons of backbone dynamics and sidechain orientation.** (top row, from left to right) Expected patterns based on the hypotheses that similarity depends on: the four PDB structures; the three protein states; grouping HL mutant with H2 inactive; the amino acid at position 60; and the presence of ligand. We construct similarity matrices by comparing backbone dihedral (upper) and sidechain pseudo-dihedral (lower) angles, using the average angle for structure (left) and the standard deviation, for dynamics (right). The greyscale for similarity is the same for all four matrices. The system setup is described using a three-layer color code for the initial protein structure, mutation at position 60, and ligand. The purpose is to compare trajectories from different initial protein structures. HL active is distinct from both H2 inactive and HL mutant for all measures. For backbone dynamics, HL mutant is distinct from H2 inactive. For sidechain structure, the two versions of H2 inactive are different, and HL mutant is distinct from H2 inactive.

Figure 6: **HL mutant has larger fluctuations in the interdomain region and clamp segment.** Backbone and sidechain angles with larger fluctuations for HL mutant (orange) or H2 inactive (blue) are shown as spheres on the protein backbone. The landmarks are shown on HL mutant, and the Arg-60-Pro mutation is indicated with a purple sphere at Cα. Fluctuation variance is shown as the sphere volume. For backbone dihedrals from Fig. 3, ϕ is shown on the N atom, while ψ is shown on the C atom. Sidechain pseudo-dihedrals (υ) from Fig. 3 are shown on the Cα atom. Sidechain dihedrals χ are shown on the second atom defining the dihedral angle, which was usually the Cα atom. For angles further from the backbone, the sidechain was drawn in the licorice representation if the difference in variance was larger than 0.2.

Figures

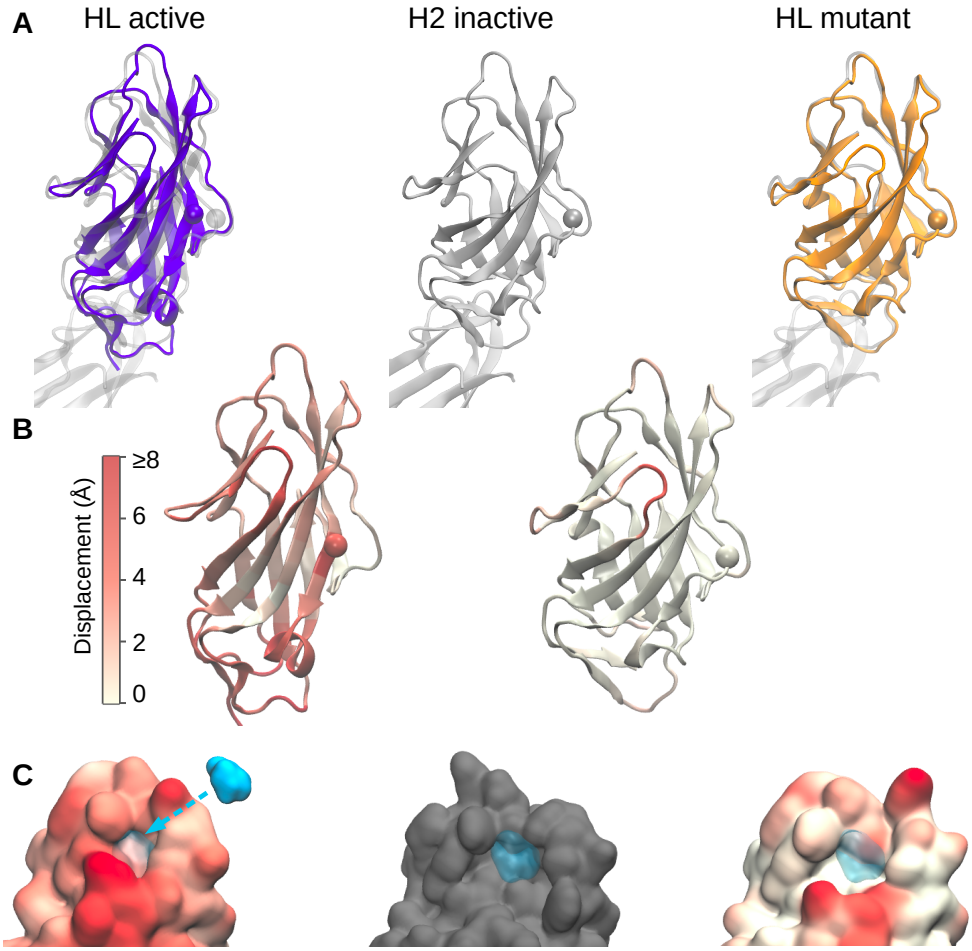

Figure 1: **The Arg-60-Pro mutant has a backbone structure similar to that of inactive FimH but with a partially closed binding pocket.** (A) NewCartoon representation of the backbone structures for H2 inactive with the pili domain cropped (grey, PDBID 4XOD, low-affinity) overlaid by the lectin domain structures: HL active (PDBID 4AUU, orange, high-affinity) and HL mutant with Arg-60-Pro (PDBID 5MCA, purple). We highlight residue 60, the mutation site, with a bead. We use color to illustrate the displacement from H2 inactive for (B) the backbone and (C) the surface rendering of predicted sidechain orientation. The mannose ligand is shown in blue and was positioned after alignment with high-affinity, two-domain FimH (PDBID 1KLF).

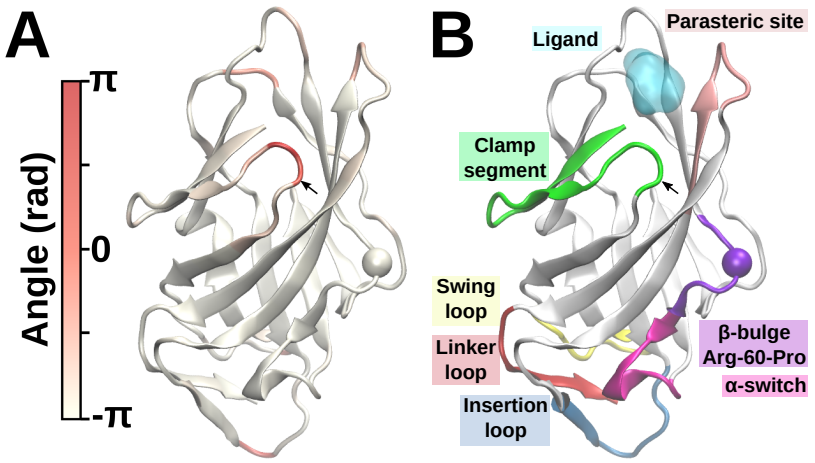

Figure 2: **Angle differences between HL mutant and H2 inactive.** (A) NewCartoon representation of the HL mutant backbone shows large changes in backbone dihedral angle differences at the C-terminal end of the clamp segment that may act like a hinge-joint (arrow). (B) Putative allosteric pathway sites[10], the parasteric site[8], and residue 60 (bead).

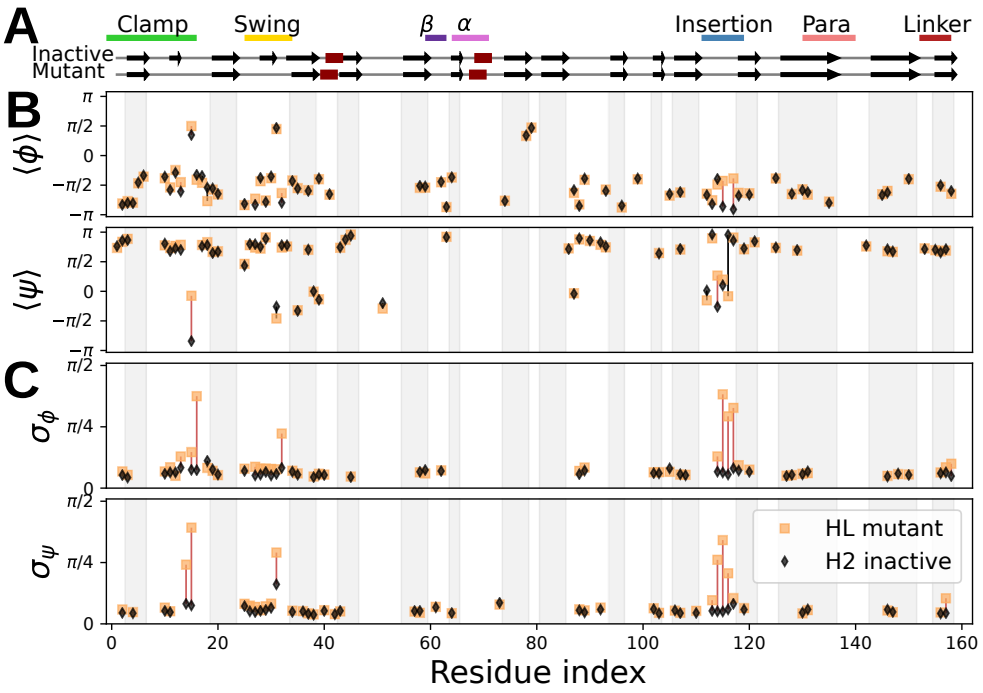

Figure 3: **Localized differences in the structure between HL mutant and H2 inactive cause differences in dynamics.** (A) Secondary structure of H2 inactive (top) and HL mutant (bottom) with features. (B) Protein structure comparison between HL mutant (orange square) and H2 inactive (black diamond) used average backbone dihedral angles ϕ and ψ, which provide an internal frame of reference. Structural differences are localized to the hinge of the clamp loop in the binding pocket and the insertion loop in the interdomain region, which is exposed to water in HL mutant. (C) Backbone dynamics comparison used the standard deviations of ϕ and ψ, quantifying the magnitude of angle fluctuations. HL mutant has larger fluctuations (red lines) at the clamp loop and the insertion loop. For statistically significant differences, the median values are shown.

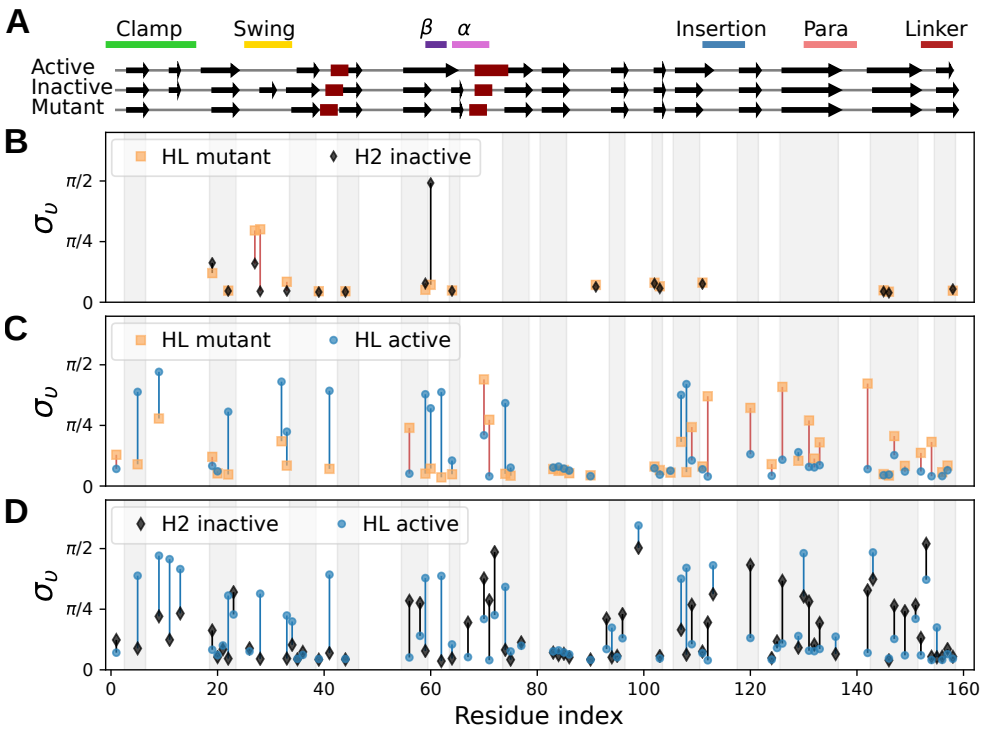

Figure 4: **Differences in the sidechain dynamics between HL mutant and H2 inactive affect the binding pocket, and both are different from HL active.** (A) Secondary structure of HL active (top), H2 inactive (middle) and HL mutant (bottom) with features. Sidechain dynamics were compared using the standard deviation of the custom dihedral angles υ. (B) Arg-60 in H2 inactive (blue diamonds) has larger fluctuations than Pro-60 in HL mutant (orange squares). HL mutant also has larger fluctuations in the exposed swing loop. (C, D) Comparisons of HL active (black circles) against HL mutant and H2 inactive show differences in sidechain dynamics beyond the allosteric pathway sites.

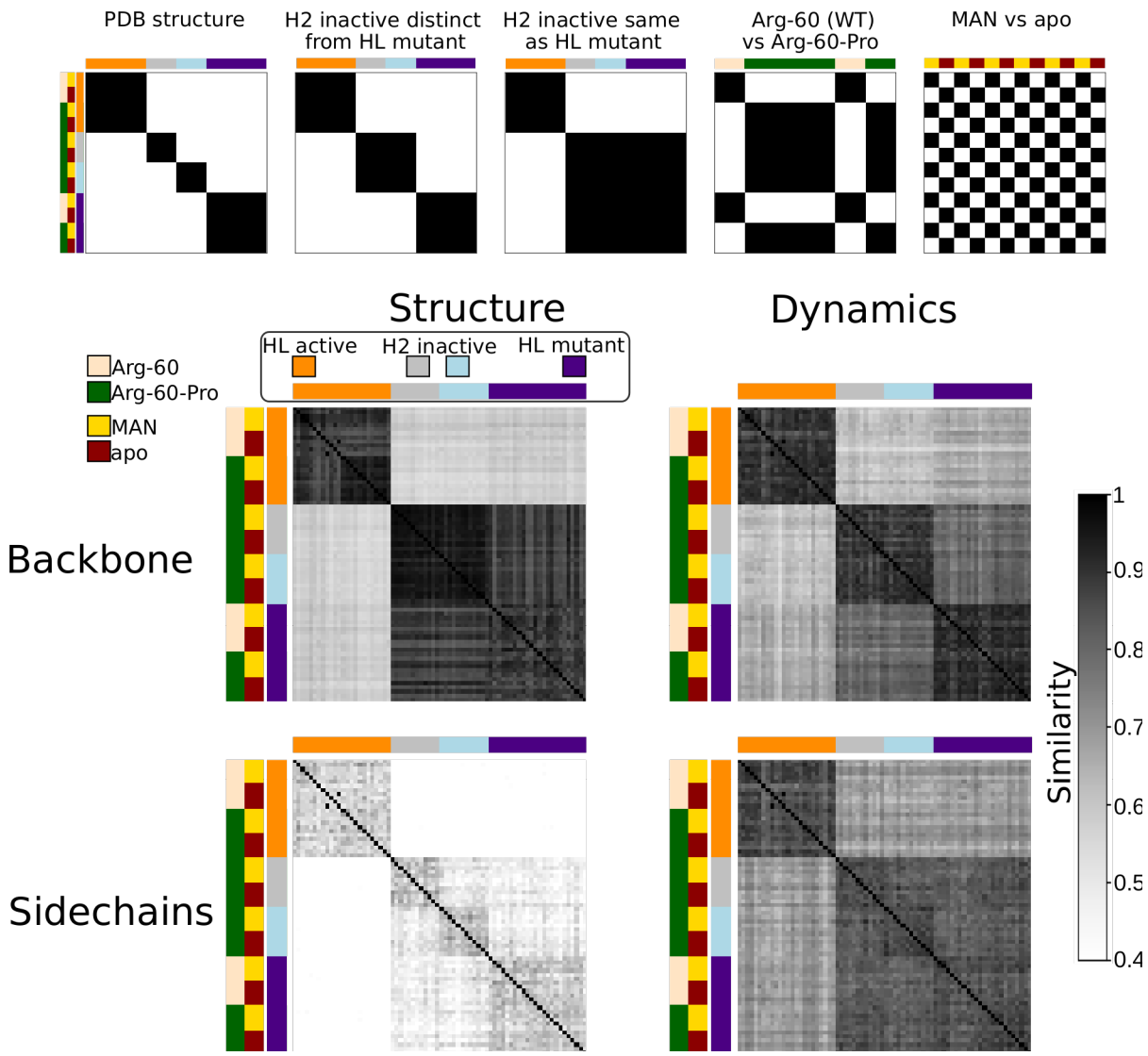

Figure 5: **Differences between HL mutant and H2 inactive in comparisons of backbone dynamics and sidechain orientation.** (top row, from left to right) Expected patterns based on the hypotheses that similarity depends on: the four PDB structures; the three protein states; grouping HL mutant with H2 inactive; the amino acid at position 60; and the presence of ligand. We construct similarity matrices by comparing backbone dihedral (upper) and sidechain pseudo-dihedral (lower) angles, using the average angle for structure (left) and the standard deviation, for dynamics (right). The greyscale for similarity is the same for all four matrices. The system setup is described using a three-layer color code for the initial protein structure, mutation at position 60, and ligand. The purpose is to compare trajectories from different initial protein structures. HL active is distinct from both H2 inactive and HL mutant for all measures. For backbone dynamics, HL mutant is distinct from H2 inactive. For sidechain structure, the two versions of H2 inactive are different, and HL mutant is distinct from H2 inactive.

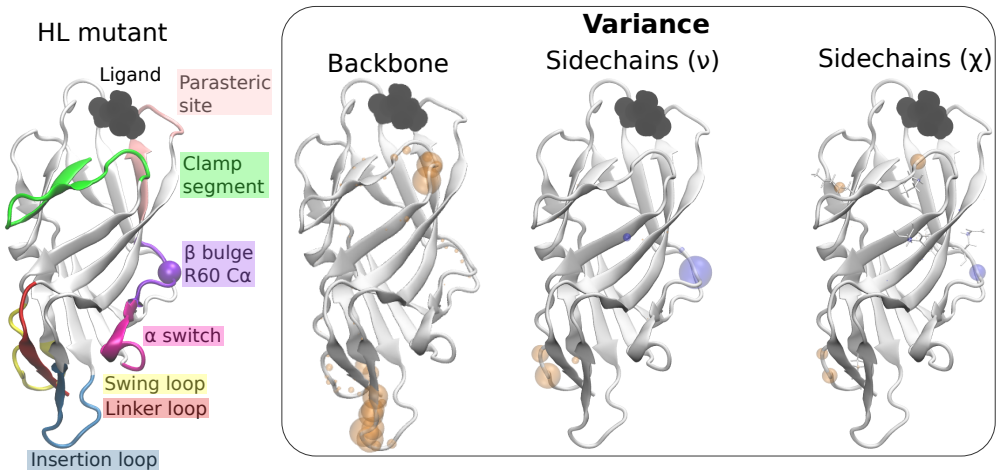

Figure 6: **HL mutant has larger fluctuations in the interdomain region and clamp segment.** Backbone and sidechain angles with larger fluctuations for HL mutant (orange) or H2 inactive (blue) are shown as spheres on the protein backbone. The landmarks are shown on HL mutant, and the Arg-60-Pro mutation is indicated with a purple sphere at Cα. Fluctuation variance is shown as the sphere volume. For backbone dihedrals from Fig. 3, ϕ is shown on the N atom, while ψ is shown on the C atom. Sidechain pseudo-dihedrals (υ) from Fig. 3 are shown on the Cα atom. Sidechain dihedrals χ are shown on the second atom defining the dihedral angle, which was usually the Cα atom. For angles further from the backbone, the sidechain was drawn in the licorice representation if the difference in variance was larger than 0.2.