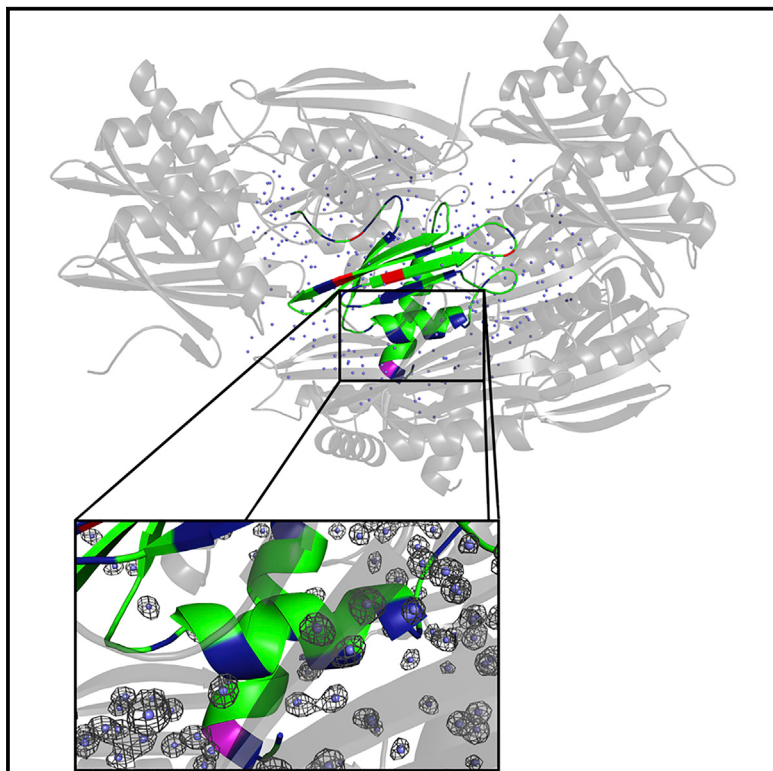


Structure

XFEL and NMR Structures of *Francisella* Lipoprotein Reveal Conformational Space of Drug Target against Tularemia

Graphical Abstract



Authors

James Zook, Mrinal Shekhar, Debra Hansen, ..., Vadim Cherezov, Abhishek Singharoy, Petra Fromme

Correspondence

asinghar@asu.edu (A.S.),
pfromme@asu.edu (P.F.)

In Brief

Using X-ray free-electron laser (XFEL) technology, Zook et al. describe an alternative protein conformation of the virulence determinant Flpp3 and use molecular dynamics (MD) to investigate conformational landscapes between these conformations. MD identifies several intermediate conformations that allow these conformational transitions to occur at physiologically relevant temperatures and provide insight to potential future drug therapies for the potentially deadly disease tularemia.

Highlights

- X-ray free-electron laser unveils alternative protein structure of lipoprotein
- Advanced molecular dynamics explore conformational space among solved structures
- Virtual ligand screening shows possible lead fragments for potential drug therapies

XFEL and NMR Structures of *Francisella* Lipoprotein Reveal Conformational Space of Drug Target against Tularemia

James Zook,¹ Mrinal Shekhar,^{1,2} Debra Hansen,¹ Chelsie Conrad,³ Thomas Grant,⁴ Chitrak Gupta,¹ Thomas White,⁵ Anton Barty,⁵ Shibom Basu,⁶ Yun Zhao,⁷ Nadia Zatsepin,¹ Andrii Ishchenko,⁸ Alex Batyuk,⁹ Cornelius Gati,⁵ Chufeng Li,⁷ Lorenzo Galli,⁵ Jesse Coe,¹⁰ Mark Hunter,⁹ Meng Liang,⁹ Uwe Weierstall,¹ Garret Nelson,⁷ Daniel James,¹¹ Benjamin Stauch,³ Felicia Craciunescu,¹ Darren Thifault,¹ Wei Liu,^{1,10} Vadim Cherezov,⁸ Abhishek Singharoy,^{1,10,*} and Petra Fromme^{1,10,12,*}

¹Center for Applied Structural Discovery, the Biodesign Institute, Arizona State University, Tempe, AZ 85287, USA

²Center for the Development of TherapeuticsThe Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

³Huntsman Cancer Institute, Salt Lake City, UT 84112, USA

⁴Hauptman Woodward Institute, Jacobs School of Medicine and Biomedical Science, SUNY, Buffalo, NY 14260, USA

⁵DESY, Center for Free Electron Laser Science, Hamburg 22607, Germany

⁶European Molecular Biology Laboratory, Grenoble, Grenoble Cedex 9 38042, France

⁷Department of Physics, Arizona State University, Tempe, AZ 85287, USA

⁸Department of Chemistry, Bridge Institute, University of Southern California, Los Angeles, CA 90089, USA

⁹Linac Coherent Light Source, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

¹⁰School of Molecular Sciences, Arizona State University, Tempe, AZ 85287, USA

¹¹Paul Scherrer Institute, 5232 Villigen, Switzerland

¹²Lead Contact

*Correspondence: asinghar@asu.edu (A.S.), pfromme@asu.edu (P.F.)

<https://doi.org/10.1016/j.str.2020.02.005>

SUMMARY

Francisella tularensis is the causative agent for the potentially fatal disease tularemia. The lipoprotein Flpp3 has been identified as a virulence determinant of tularemia with no sequence homology outside the *Francisella* genus. We report a room temperature structure of Flpp3 determined by serial femtosecond crystallography that exists in a significantly different conformation than previously described by the NMR-determined structure. Furthermore, we investigated the conformational space and energy barriers between these two structures by molecular dynamics umbrella sampling and identified three low-energy intermediate states, transitions between which readily occur at room temperature. We have also begun to investigate organic compounds *in silico* that may act as inhibitors to Flpp3. This work paves the road to developing targeted therapeutics against tularemia and aides in our understanding of the disease mechanisms of tularemia.

INTRODUCTION

Francisella tularensis strain SCHU S4 is among the most virulent pathogens on the planet—as few as ten colony-forming units can facilitate a tularemia infection (Hall et al., 2008). Because of its high virulence, *F. tularensis* was weaponized in the past and is considered a class A bioterrorism agent (Dennis et al.,

2001). The threat of bioterrorism, being endemic to the Northern Hemisphere, and frequency of outbreaks in underdeveloped regions warrants an imminent need for developing medical interventions. Vaccine development against *F. tularensis* has been ongoing for the last half century, and to date there is still no Food and Drug Administration-approved vaccine available (Elkins et al., 2016).

One reason for the vaccine's elusiveness is the complexity and number of pathogenic routes for tularemia infection and the many ways that *F. tularensis* manipulates the host immune response (Brock and Parmely, 2017; Krocova et al., 2017). Depending on how the disease is transmitted, a wide variety of proteins is utilized by the pathogen to induce virulence (Jones et al., 2014). Although a number of proteins have been identified as virulence determinants, elucidation of their structure and biological function remains a challenge (Elkins et al., 2016; Su et al., 2007). Consequently, there might be multiple mechanisms that cause the high virulence of tularemia and the challenge to develop a vaccine is still unresolved (Roberts et al., 2018).

The results described in this study shed light on the structure, dynamics, and druggability of a key virulence determinant, Flpp3. Flpp3 is an immunogenic (Eyles et al., 2007; Parra et al., 2010) outer membrane lipoprotein (Clinton et al., 2010; Parra et al., 2010) that is essential for virulence (Su et al., 2007), stimulates Toll-like receptor 2 (TLR2) (Kinkead et al., 2018; Parra et al., 2010), binds plasminogen (Clinton et al., 2010), and may contribute to protection of the bacterium in animal vaccine studies (Post et al., 2017). Flpp3 is anchored to the membrane by a palmitoyl group (Parra et al., 2010) and may also undergo glycosylation (Balonova et al., 2010). Noting that inhibition of the virulence determinant proteins is an effective antimicrobial strategy and that Flpp3 is unique to the *Francisella* genome,

Table 1. Flpp3solXa SFX Data Statistics—PDB: 6PNY

Beam energy (keV)	8.7
Space group	P2 ₁ 2 ₁ 2 ₁
Resolution (Å)	19–1.65 (1.71–1.65)
Unit cell parameters	
a, b, c (Å)	28.0, 51.9, 68.1
α , β , γ (°)	90, 90, 90
No. of crystal hits	82,157
No. of indexed patterns	52,699
Duration of data collection (min)	57
Unique reflections	16,570
Reflections used in refinement	14,250
Multiplicity	217.05 (42.6)
CC*	0.996
R _{split}	0.1073
R _{work} /R _{free}	0.1975 (0.3479)/0.2349 (0.3300)
Completeness (%)	100 (99.9)

Flpp3 serves as an attractive target for small-molecule interventions (Brannon and Hadjifrangiskou, 2016). Such therapeutic developments are normally driven by existing structural or biochemical information on related protein families. However, little is known about the precise function of Flpp3, and it shares no significant sequence homology to any other protein, rendering the traditional bioinformatics approaches for drug discovery ineffective. Thus, structure determination of Flpp3 may form the basis for finding additional therapeutics against tularemia.

In this article, we report on the structure of the soluble form of Flpp3 (denoted Flpp3sol) resolved for the first time at room temperature by serial femtosecond crystallography (SFX), and we highlight additional conformations that were not captured by our past NMR experiments (Zook et al., 2015). The experimental structure determination is bolstered with molecular dynamics (MD) and free energy simulations that further reveal the multiple possible conformations of Flpp3. These simulations highlight a functionally relevant conformational transition that encompasses SFX- and NMR-derived structures that correspond to the conformational space that Flpp3 may sample *in vivo*. Access to these conformations and combining virtual screening of small-molecule binding partners has the potential to uncover tularemia-specific antibiotics. Although it remains unknown whether lead-based drug design against this protein will result in a potent antibiotic, a pyrimidine-based molecule is identified as a potential lead. Future studies will concentrate on determining the antimicrobial effects of compounds that bind to Flpp3.

RESULTS

Construct Design Strategy

Using the data acquired from the NMR structure (Zook et al., 2015), Flpp3_{nmr}, a 119-residue protein with a molecular weight of 13.3 kDa and mean radius of 1.36 nm, was engineered to facilitate crystallization (Flpp3_{xtal}). Similar to Flpp3_{nmr}, Flpp3_{xtal} lacks the first 24 residues of the N terminus that constitutes the membrane component of Flpp3 and makes it water insoluble (see Figure S1). The new Flpp3_{xtal} construct was designed to

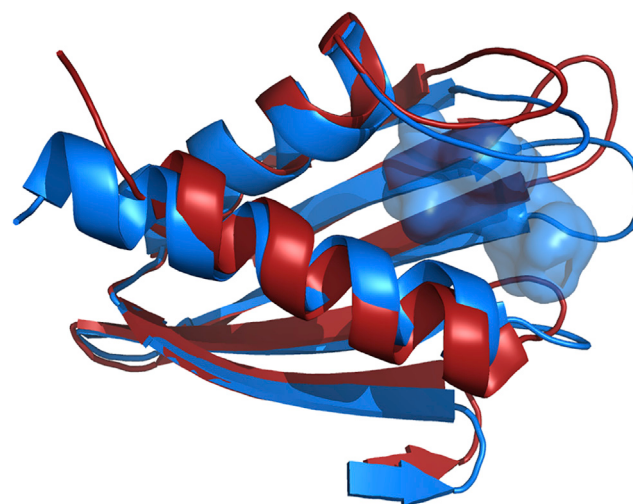


Figure 1. Flpp3 Structures from NMR (blue; PDB: 2MU4) and SFX (red; PDB: 6PNY)

There is no cavity observed in the SFX structure which would appear as a red, semi-transparent surface, while it is quite apparent by NMR (shown by the blue, semi-transparent surface).

decrease surface entropy based on structural and dynamics information derived from NMR data. Nuclear Overhauser effect (NOE) relaxation data revealed flexible protein termini, which were hypothesized to hamper crystallization. Furthermore, the presence of flexible termini was corroborated by NMR structural data, including secondary chemical shift analysis and absence of NOE cross-peaks of the terminal regions of the protein (Zook et al., 2015). The NMR structure had been solved based on a protein construct that contains a hexahistidine tag at the C terminus. For crystallization a new construct was designed, Flpp3_{xtal}, where the His-tag was moved to the N terminus, followed by a Factor Xa cleavage site (IEGR), which allows for removal of the His-tag before crystallization, resulting in a final construct consisting of 118 residues with a molecular weight of 13.2 kDa. For reference, the wild-type construct is 137 residues with a molecular weight of 14.9 kDa. See Figure S1 for a comparison of the sequence alignment of Flpp3_{nmr} and Flpp3_{xtal}.

SFX Structure of Flpp3_{xtal}

In SFX with X-ray free-electron lasers (XFELs), crystals are typically delivered to the XFEL beam in a jet at room temperature, and their structure is probed by the XFEL shots using the diffraction-before-destruction principle (Chapman et al., 2011). While up to 100 mg of protein crystals is required for liquid injection using a gas dynamic virtual nozzle (Chapman et al., 2011), the sample amount for a dataset can be dramatically reduced by the use of a viscous injector originally developed for membrane proteins grown in lipidic cubic phases to 0.3–0.5 mg (Weierstall et al., 2014). The use of viscous media for sample delivery has been extended to allow for the embedding and delivery of both soluble and membrane proteins in agarose (Conrad et al., 2015). The small microcrystals (~10 μ m) of Flpp3_{xtal} were embedded into agarose directly before sample delivery at LCLS. They remained both birefringent and visible by second-order nonlinear imaging of chiral crystals (Wampler et al., 2008)

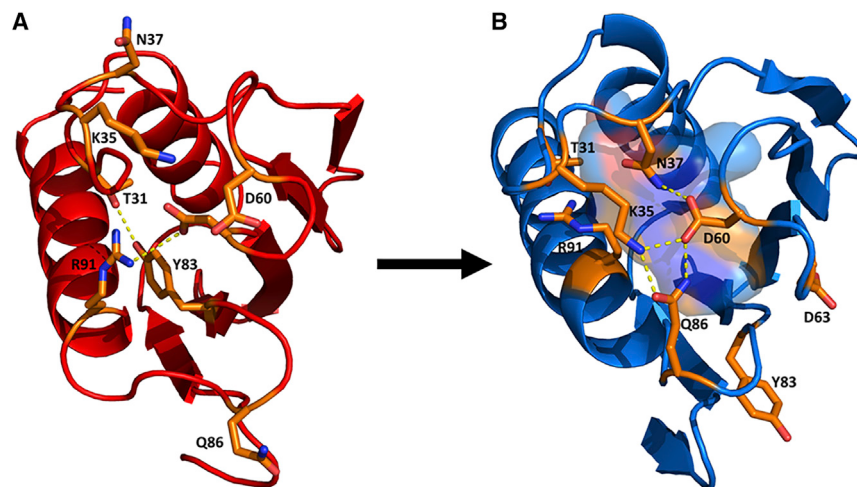


Figure 2. A Detailed Look at the Conformational Differences between Crystal and NMR Structures

One of the largest conformational changes of Flpp3 between the crystal state and the NMR state arises from Y83. Flpp3_{xtal} (A) shows Y83 within the interior of Flpp3, forming a single hydrogen bond with T31 and occupying most of the cavity observed in Flpp3_{nmr}. Flpp3_{nmr} demonstrates a 180° rotation of the Y83 side chain exposes the tyrosine to the protein surface, generating an interior cavity displayed by the surface representation (B). This transition is stabilized by a salt bridge forming between D63 and R91. Flpp3_{nmr} is further stabilized by a hydrogen bond network between K35, N37, D60, and Q86, and a salt bridge between K35 and D60.

after incubating the material in the agarose viscous media, indicating that Flpp3_{xtal} crystals are highly stable and suitable for SFX.

Protein yields and purity of Flpp3_{xtal} were similar to the Flpp3_{nmr} construct described in the previous NMR study (Zook et al., 2015). We embedded a total of 5 μ L of protein crystals in 10 μ L agarose medium (STAR Methods) and collected SFX data at the experimental station CXI. A complete dataset of Flpp3_{xtal} was collected during 1 h of beamtime at a resolution of 1.5 Å, with a 10% hit rate and 64% indexing rate, and with a total number of 52,699 indexed images. The merged data resulted in an R_{split} of 0.1073 with an overall CC* of 0.996. The NMR structure ensemble (Zook et al., 2015) was used as a starting model to phase the SFX diffraction data using molecular replacement employing ROSETTA-MR for loop rebuilding. The final model was refined to an R_{free} of 0.2349 and an R_{work} of 0.1975. A more detailed summary of the data statistics is shown in Table 1.

The SFX structure is shown in Figure 1 and has a comparable mean-radius with Flpp3_{nmr} of 1.39 nm. The most notable difference to the NMR structure ensemble is the very compact fold of Flpp3_{xtal}, where the main six-stranded β sheet are very tightly packed against two α helices, representing a “closed” conformation of Flpp3 that does not feature a cavity observed in most of the NMR models in the structure ensemble (Zook et al., 2015). Figure 1 illustrates side-by-side comparison of both structures (model 1 of the NMR ensemble was used, representing the lowest energy conformation). It was suggested that this cavity could bind a small organic molecule, either for a hitherto undescribed *in vivo* function of the protein, or as a possible drug target. The single largest determinant of the closed conformation of Flpp3 in the SFX structure is the position of Y83. This residue is solvent exposed in the NMR structures (referred to as the “Out” conformation), but in the SFX structure it is buried in the interior of the protein (“In” conformation). Therefore, for a cavity to be open, Y83 must be solvent exposed, as shown in Figure 2.

The SFX structure also revealed the position of 118 coordinating water molecules, where all but three (which form intramolecular hydrogen bonds) form a coordinated hydrogen bond

network at the interface of the protein (Figure 3A). In addition, the crystal packing is exceptionally tight, with a Matthew’s coefficient of 1.62, corresponding to 23.9% of the crystal volume occupied by uncoordinated solvent (Figure 3A). Crystal contacts are facilitated by four direct protein-protein intermolecular interactions through salt bridges, and 51 hydrogen bonds mediated via protein-water-protein (Figure 3B). There is only one hydrophobic crystal contact between F43 and I101.

MD Simulations Reveal the Conformational Space of Flpp3

The major structural difference between Flpp3_{nmr} and Flpp3_{xtal} is the more compact fold of Flpp3_{xtal} with a closed cavity that is determined by the conformation of Y83 (Figure 2). In the open conformation of Flpp3_{nmr} (Figure 2B), Y83 is exposed to the protein surface opening the cavity (illustrated in the surface representation of Figure 2B) and undergoing a 180° side chain rotation compared with the buried conformation in Flpp3_{xtal} (Figure 2A). We employed MD simulations and free energy methods (in particular, umbrella sampling) to characterize the conformational space of Flpp3 and energy changes that govern the transition between Flpp3_{nmr} and Flpp3_{xtal}.

The free energy plot (Figure 4) links the crystal structure (In conformation, state 0), and Out conformation (Flpp3_{nmr}, state D) via three intermediate states (A, B, and C) by using a root-mean-square deviation (RMSD) to monitor the reaction coordinate. Although RMSD is a degenerate coordinate, and cannot pick up subtle changes as noted in previous publications (Singharoy and Chipot, 2017; Singharoy et al., 2017), for relatively small systems, such as Flpp3, the variation RMSD ranges only about 5 Å. Therefore, an exhaustive sampling of the orthogonal space can be derived in microseconds, established in several articles by Harada and Shigeta (2017) and Harada et al. (2015).

The complete transition between state 0 and state D is summarized in Figure 5, emphasizing cavity formations, conformational changes, and removal in the surface representations of each model. State 0 and state A are characterized by the H-bonding interactions of Y83 with T31 (Figure 4, state 0 and state A). The conformational transition is summarized in Figure 5A and constitutes a total conformational RMSD of

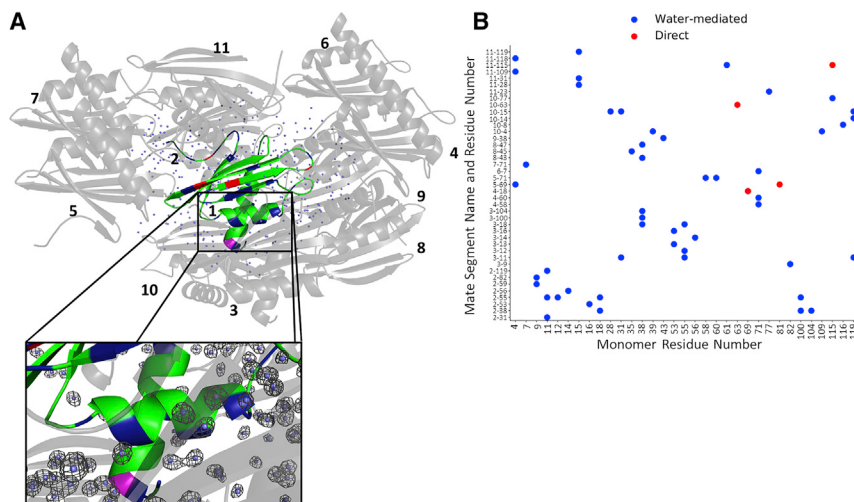


Figure 3. Flpp3_{xtal} Unit Cell with Symmetry Mates Illustrates a Very Dense Packing Configuration with Only 23% Volume Occupied by Uncoordinated Solvent

Most of the crystal packing is mediated via hydrogen bonding through water molecules. (A) This is shown as blue spheres representing water molecules (including the gray electron density shown in inset), connecting the central monomer (green) to its symmetry mates (gray); direct protein-protein interactions are highlighted in red, residues involved with crystal contacts mediated by a single water bridge are shown in dark blue, and a single residue, R115, contributes to both (colored purple). (B) A 2D plot of the molecular dynamics flexible fitting calculations used to analyze hydrogen bonds also illustrates the extent of water-mediated hydrogen bonding (blue), with 51 contacts facilitated by 1 water compared with 4 direct protein-protein interactions (red). With the exception of 3 waters (which coordinate solely an intramolecular interaction), the remaining waters contribute to an extensive hydrogen bonding network between Flpp3_{xtal} molecules in the crystal lattice (inset).

1.2 Å: loop 1 (L1) shifts away from the protein interior by 2.8 Å while at the same time loop 2 (L2) moves away from protein interior by 2.7 Å. This breaks the salt bridge between R91 and D63 while at the same time creating a cavity with a volume of approximately 170 Å³. In addition, the C terminus of the protein moves 9.8 Å to extend helix 2 (H2).

Thereafter, the protein overcomes a barrier of ~0.5 kcal/mol to visit the second local minimum corresponding to the Y83 buried NMR conformation, which lies at ~1.0 kcal/mol below the Y83 closed conformation of Flpp3 identified in the crystal structure (Figure 4, state B). As summarized in Figure 5B (with a total conformational RMSD = 1.1 Å), the transition from state A and state B includes L2 moving an additional 4.3 Å away from the protein interior, while the N terminus of H2 opens by 2.0 Å, and loop 3 (L3) pivots on G84 to further open the protein cavity to a volume of 350 Å³. This conformational change disrupts the hydrogen bond between T31 and Y83. The mean first passage time for this transition was calculated to be ~0.78 ms.

Subsequently, the protein overcomes another barrier of ~0.5 kcal/mol, as a consequence of 180° rotation of Y83 with respect to the loop backbone (Figure 4, state C), and partially exposes Y83 to solvent. The free energy differences between the three conformations of Y83 partially exposed in Flpp3_{nmr}, partially buried, and the closed conformation of Y83 in Flpp3_{xtal} are ~3 and 2 kcal/mol, respectively, and the mean free passage time for the transition was calculated to be 66 μs. This transition (total RMSD = 1.6 Å) from state B to State C is illustrated in Figure 5C. The N terminus of H2 moves toward the protein interior by about 2.8 Å; L1 also shifts toward the protein interior by 1.5 Å, while L2 moves away from the protein by 4.0 Å. In effect this reduces cavity volume to about 200 Å³.

Finally, the system samples the global minima corresponding to Y83 completely solvent exposed (Figure 4, state D). The mean free passage time for the transition from Y83 partially solvated to solvent exposed was calculated to be 5 μs. This conformation is

chiefly stabilized by the electrostatic interactions between R91, D60, and K35 (Figure 2B) and lies at ~4.0 kcal/mol below the crystal conformation. Globally (shown in Figure 5D), this final transition has an RMSD of 1.9 Å compared with state C and can be described as L1 moving toward the protein interior by 7.2 Å, while L2 closes by 4.1 Å, creating the final cavity depicted by the surface representation in Figure 5D with a size of about 180 Å³. Histograms for the umbrella sampling simulations demonstrating overlap of windows is demonstrated in Figure S2A. The convergence plot for the potential of mean force is detailed in Figure S2B.

These results suggest that the transitions between the closed and open conformation and their intermediate states are accessible at room temperature and may occur in solution in the μs to ms time range. This transition and the existence of intermediate states that lower the activation energy barrier would allow Flpp3 to tightly “wrap” around potentially substrates or drug molecules. Identification of potential substrates or drug molecules will be valuable.

Virtual Ligand Screening Requires Both Hydrophobic and Hydrophilic Moieties

The ZINC database was used for virtual screening of the drug binding of Flpp3_{nmr} using Autodock Vina (Trott and Olson, 2010). The molecule with the lowest calculated binding energy was N4-(cyclopentylmethyl)-N2,N4-trimethyl-pyrimidine-2,4-diamine (ZINC: 77213780), shown in Figure S2A. Several rounds of optimization were performed to improve binding affinities based on free energies calculated by MD, and ultimately the N2 dimethyl amide was replaced by a carboxylic acid moiety, a primary amide was added to the C6 position of the pyrimidine, and the N4 tertiary amide was exchanged for a ketone. In addition, an ethyl moiety was added to the C3 position of the cyclopentyl ring (Figure S2B). These modifications were made to increase the number of hydrogen bonds available to Flpp3_{nmr} as well as adding a negatively charged interaction with R91. To improve shape complementarity and hydrophobic interactions,

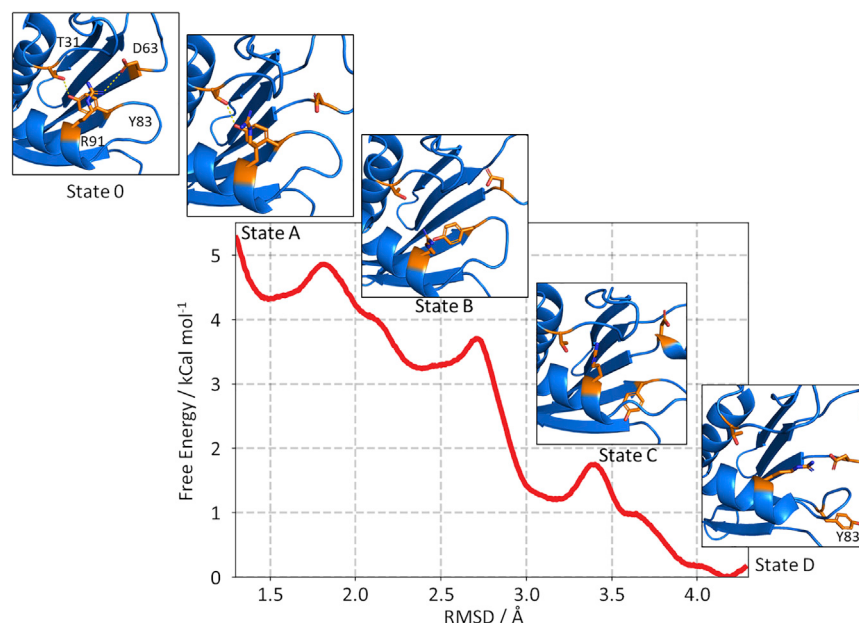


Figure 4. Free Energy Diagram Transitioning from Flpp3_{xtal} (state 0) to Flpp3_{nmr} (state D)

Y83 is shown breaking its hydrogen bond from T31 after the salt bridge dissociation between R91 and D63. Y83 rotates from the buried state, to the solvent-exposed state. Three intermediate states were observed in this calculation (states A, B, and C). Independently observed, state C is remarkably similar to other models calculated in the NMR ensemble.

an ethyl group was added to the C4 position of the cyclopentane ring. The result of these modifications is a ligand that inserts specifically deep into the protein cavity (Figure 6A). Five hydrogen bonds and a salt bridge provide polar-based stabilizing interactions, while the ethyl-cyclopentyl moiety provides hydrophobic support to nonpolar residues I26, Y39, and 95L (Figure 6B).

DISCUSSION

In this study the combination of serial room temperature femto-second crystallography at an XFEL, with the structure determination of Flpp3 by NMR and MD simulations, was used to unveil the structure and dynamics of Flpp3. Furthermore, we explore the full conformational space by using a hybrid approach, combining NMR, SFX, and MD, which unravels the transitions between the structures using molecular dynamics.

There are few solved lipoproteins to compare this structure with, but one such structure is bacterial lipocalin (RCSB: 1QWD), which is slightly larger, with 177 residues resulting in a 19.8-kDa protein with a similar mean radius of 1.50 nm (Campnacci et al., 2004). There are significant structural differences between the published NMR structure and this SFX structure. The greatest difference is the absence of an internal cavity in the structure of Flpp3_{xtal}. This finding was unexpected, as we observed a prominent cavity in the structure determined by NMR spectroscopy. Interestingly, the presence of 118 crystallographic water molecules contribute to the vast majority of crystal contacts: 51 directly bridging between monomers, 3 contributing solely to intramolecular hydrogen bonds, and the remaining 64 waters contribute to an intricate hydrogen bond network at the crystal interface. Only 5 direct protein-protein interactions (4 hydrophilic and 1 hydrophobic) are involved in crystal contacts (see Figure 3).

This strongly suggests that the protein itself orders water molecules surrounding it. Thereby the crystal structure represents

one of the native states of Flpp3, which is independently confirmed by the MD simulations, which show that the activation energy barriers between the different states are small and thereby thermally accessible to be overcome at room temperature.

The MD simulations revealed that the internal cavity from the NMR structure is transient. Recently, Flpp3 was identified as a component of a large (>50 kDa) extracellular complex that is released from the LVS strain of *F. tularensis*, and that actively delays the normal progression of apoptosis in isolated human neutrophils by a TLR-dependent mechanism (Kinkead et al., 2018). Our results suggest that Flpp3 undergoes a conformational change. As shown from MD calculations, the SFX structure exists in a significantly higher free energy state by 4.5 kCal/mol, suggesting that the NMR conformation may be more frequently observed in solution, with an equilibrium ratio (in the absence of potential binding partners) of 1 SFX conformer for every 1,800 NMR conformations. The high energy conformation observed from the SFX data is stabilized by energetically favorable intermolecular crystal contacts, facilitated by a large hydrogen bond network that includes 118 crystallographically visible water molecules. Therefore, while the closed Flpp3_{xtal} monomer conformation may be the less energetically favorable conformation, it is stabilized by hydrogen bonds in the crystals. Our MD simulations suggest that there is a low-energy conformational transition between the SFX conformer and NMR conformer. The intermediates resembling the NMR conformations, namely from states B and C were not imposed on MD, but rather were observed as an outcome of the simulated minimum free energy pathway. Sampling of such biologically relevant intermediates on one hand, adds credence to the simulated conformational changes of Flpp3, while on the other it allows application of previous knowledge on the NMR models as a cross-validation for the new structures derived by MD.

The next natural step is to begin investigating the potency of the ligand as both a binding partner to Flpp3 and a therapeutic agent against tularemia. Efforts are being directed toward the synthesis of likely candidates from the *in silico* studies described in this study followed by ligand binding assays in tandem with further structural work, including co-crystallization attempts and NMR studies. The importance of being prepared against a biological attack cannot be understated, and this work may form the basis for new solutions.

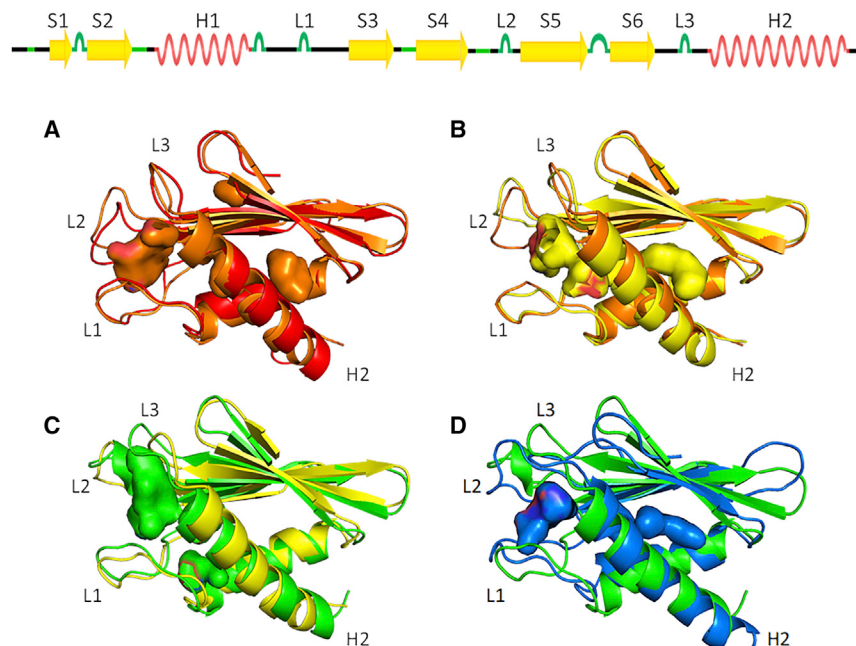


Figure 5. A Stepwise Global Summary of Protein Conformational Change Shows How Flipp3 Transitions Conformationally from the “In” state (Red) to the “Out” state (Blue), Emphasizing Cavity Conformations as Surface Representations

(A) To transition from state 0 (red) to state A (orange), loops 1 and 2 (L1 and L2) shift away from the protein interior by 2.8 and 2.7 Å, respectively, creating an internal cavity of about 170 Å³. In addition, the C terminus of the protein extends the C terminus of helix 2 (H2). This corresponds to a global RMSD change of 1.2 Å.

(B) To transition from state A to state B (yellow), L2 moves an additional 4.3 Å from the protein interior while the N terminus of H2 opens by 2.0 Å by pivoting on G84. This further opens the cavity to about 350 Å³ and changes the overall protein RMSD by 1.1 Å.

(C) Transitioning to state C (green) from state B, the cavity is reduced to 200 Å³ via the N terminus of H2 closing by about 2.8 Å, which also closes L3; L1 closes as well, shifting by 1.5 Å. L2 opens and pivots by 4.0 Å, contributing to a global conformation change of 1.6 Å.

(D) The final Out conformation (blue) transitions from state C by L1 closing by 7.2 Å, while L2 closes by 4.1 Å, creating a final cavity volume of 180 Å³. In addition, the C terminus of H2 is extended to the protein terminus, with a global conformational RMSD of 1.9 Å.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
 - Microbes
- METHOD DETAILS
 - Cloning, Expression, Purification, and Crystallization
 - SFX Data Collection and Structure Calculation
 - Conventional Molecular Dynamics
 - Steered Molecular Dynamics
 - Umbrella Sampling
 - Kinetics Calculation
 - Ligand Screening

- Hydrogen Bond Analysis
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.str.2020.02.005>.

ACKNOWLEDGMENTS

Use of LCLS, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under contract no. DE-AC02-76SF00515. Parts of the sample injector used at LCLS for this research were funded by the NIH, P41GM103393, formerly P41RR001209. This work was supported by the NIH Common Fund Grant R01 GM095583 (to P.F.) and the Center for Membrane Proteins in infectious diseases, MPID supported by the PSI:Biological Structural Biology initiative of the NIH U54 GM094599. We also acknowledge funding from the Science

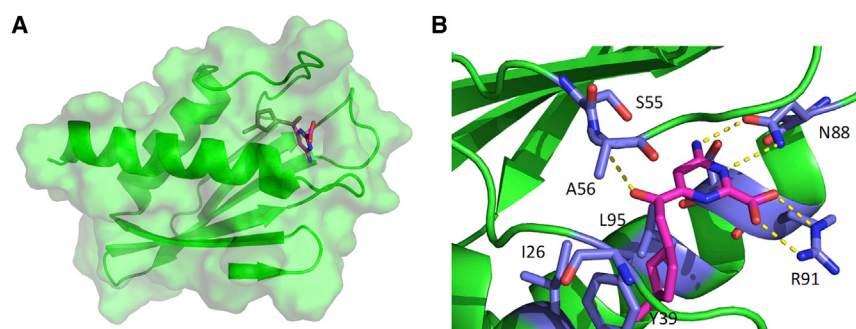


Figure 6. Key Residues for Ligand Binding to Flipp3

Polar residues, shown in blue and purple in the overview in (A) and in the close-up view in (B), that contribute to hydrogen bonds or electrostatic interactions include R91, N88, S55, and A56. Residues contributing to hydrophobic interactions include residues L95, I26, and Y39.

and Technology Center (STC) “BioXFEL” through award STC-1231306. The work was also supported by funds from the Biodesign Center for Applied Structural Discovery at Arizona State University. A.S. and C.G. acknowledge start-up funds from the SMS and CASD at Arizona State University, and the resources of the OLCF at the Oak Ridge National Laboratory, which is supported by the Office of Science at DOE under Contract No. DE-AC05-00OR22725, made available via the INCITE program. AS acknowledges CAREER award from NSF (MCB-1942763).

AUTHOR CONTRIBUTIONS

Conceptualization, J.Z., P.F., and A.S.; Methodology, J.Z., P.F., D.H., A.S., C.C., F.C., T.G., and M.S.; Software, T.W., A.Barty., M.S., J.Z., Y.Z., N.Z., A.I., A.Batyuk., and C.L.; Validation, T.G., S.B., J.Z., C.Gupta, M.S., and D.T.; Formal Analysis, T.G., J.Z., M.S., S.B., and A.Batyuk.; Investigation, J.Z., M.S., C.C., T.G., C.Gati., J.C., G.N., D.J., F.C., L.G., B.S., and D.T.; Resources, V.C., A.S., M.H., M.L., and U.W.; Data Curation, S.B., J.Z., A.Barty., and M.S.; Writing—Original Draft, J.Z., P.F., A.S., M.S., D.H., and F.C.; Writing—Review & Editing, J.Z., P.F., A.S., D.H., B.S., J.C., V.C., and W.L.; Visualization, J.Z., P.F., A.S., M.S., and C. Gupta.; Supervision, V.C., W.L., P.F., A.S., and J.Z.; Project Administration, J.Z., P.F., and A.S.; Funding Acquisition, P.F.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 29, 2019

Revised: December 23, 2019

Accepted: February 13, 2020

Published: March 5, 2020

REFERENCES

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol Crystallogr.* **66**, 213–221.
- Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol Crystallogr.* **68**, 352–367.
- Balonova, L., Herychova, L., Mann, B.F., Link, M., Bilkova, Z., Novotny, M.V., and Stulik, J. (2010). Multimethodological approach to identification of glycoproteins from the proteome of *Francisella tularensis*, an intracellular microorganism. *J. Proteome Res.* **9**, 1995–2005.
- Barty, A., Kirian, R.A., Maia, F.R.N.C., Hantke, M., Yoon, C.H., White, T.A., and Chapman, H. (2014). Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data. *J. Appl. Crystallogr.* **47**, 1118–1131.
- Brannon, J.R., and Hadjifrangiskou, M. (2016). The arsenal of pathogens and antivirulence therapeutic strategies for disarming them. *Drug Des. Devel. Ther.* **10**, 1795–1806.
- Brock, S.R., and Parmely, M.J. (2017). *Francisella tularensis* confronts the complement system. *Front. Cell. Infect. Microbiol.* **7**, 523.
- Campanacci, V., Nurizzo, D., Spinelli, S., Valencia, C., Tegoni, M., and Cambillau, C. (2004). The crystal structure of the *Escherichia coli* lipocalin B1c suggests a possible role in phospholipid binding. *FEBS Lett.* **562**, 183–188.
- Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R.J. (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688.
- Chapman, H.N., Fromme, P., Barty, A., White, T.A., Kirian, R.A., Aquila, A., Hunter, M.S., Schulz, J., DePonte, D.P., Weierstall, U., et al. (2011). Femtosecond X-ray protein nanocrystallography. *Nature* **470**, 73–77.
- Clinton, S.R., Bina, J.E., Hatch, T.P., Whitt, M.A., and Miller, M.A. (2010). Binding and activation of host plasminogen on the surface of *Francisella tularensis*. *BMC Microbiol.* **10**, 76.
- Conrad, C.E., Basu, S., James, D., Wang, D.J., Schaffer, A., Roy-Chowdhury, S., Zatsel, N.A., Aquila, A., Coe, J., Gati, C., et al. (2015). A novel inert crystal delivery medium for serial femtosecond crystallography. *IUCrJ* **2**, 421–430.
- Cormier, C.Y., Park, J.G., Fiocco, M., Steel, J., Hunter, P., Kramer, J., Singla, R., and LaBaer, J. (2011). PSI: biology-materials repository: a biologist’s resource for protein expression plasmids. *J. Struct. Funct. Genomics* **12**, 55–62.
- Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald—an N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.
- Dennis, D.T., Inglesby, T.V., Henderson, D.A., Bartlett, J.G., Ascher, M.S., Eitzen, E., Fine, A.D., Friedlander, A.M., Hauer, J., Layton, M., et al. (2001). Tularemia as a biological weapon—medical and public health management. *JAMA* **285**, 2763–2773.
- Elkins, K.L., Kurtz, S.L., and De Pascalis, R. (2016). Progress, challenges, and opportunities in *Francisella* vaccine development. *Expert Rev. Vaccines* **15**, 1183–1196.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501.
- Eyles, J.E., Unal, B., Hartley, M.G., Newstead, S.L., Flick-Smith, H., Prior, J.L., Oyston, P.C., Randall, A., Mu, Y., Hirst, S., et al. (2007). Immunodominant *Francisella tularensis* antigens identified using proteome microarray. *Proteomics* **7**, 2172–2183.
- Fiorin, G., Klein, M.L., and Henin, J. (2013). Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* **117**, 3345–3362.
- Hall, J.D., Woolard, M.D., Gunn, B.M., Craven, R.R., Taft-Benz, S., Frelinger, J.A., and Kawula, T.H. (2008). Infected-host-cell repertoire and cellular response in the lung following inhalation of *Francisella tularensis* Schu S4, LVS, or U112. *Infect. Immun.* **76**, 5843–5852.
- Harada, R., and Shigeta, Y. (2017). Temperature-shuffled parallel cascade selection molecular dynamics accelerates the structural transitions of proteins. *J. Comput. Chem.* **38**, 2671–2674.
- Harada, R., Takano, Y., Baba, T., and Shigeta, Y. (2015). Simple, yet powerful methodologies for conformational sampling of proteins. *Phys. Chem. Chem. Phys.* **17**, 6155–6173.
- Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph. Model.* **14**, 33–38.
- Irwin, J.J., and Shoichet, B.K. (2005). ZINC—a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **45**, 177–182.
- Jakalian, A., Jack, D.B., and Bayly, C.I. (2002). Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **23**, 1623–1641.
- Jones, B.D., Faron, M., Rasmussen, J.A., and Fletcher, J.R. (2014). Uncovering the components of the *Francisella tularensis* virulence stealth strategy. *Front. Cell. Infect. Microbiol.* **4**, 32.
- Kincaid, L.C., Whitmore, L.C., McCracken, J.M., Fletcher, J.R., Ketelsen, B.B., Kaufman, J.W., Jones, B.D., Weiss, D.S., Barker, J.H., and Allen, L.A.H. (2018). Bacterial lipoproteins and other factors released by *Francisella tularensis* modulate human neutrophil lifespan: effects of a TLR1 SNP on apoptosis inhibition. *Cell. Microbiol.* **20**, <https://doi.org/10.1111/cmi.12795>.
- Koerner, L.J., Philipp, H.T., Hromalik, M.S., Tate, M.W., and Gruner, S.M. (2009). X-ray tests of a pixel array detector for coherent X-ray imaging at the Linac Coherent Light Source. *J. Instrum.* **4**, P03001.
- Krocova, Z., Macela, A., and Kubelkova, K. (2017). Innate immune recognition: implications for the interaction of *Francisella tularensis* with the host immune system. *Front. Cell. Infect. Microbiol.* **7**, 446.
- Liang, M., Williams, G.J., Messerschmidt, M., Seibert, M.M., Montanez, P.A., Hayes, M., Milathianaki, D., Aquila, A., Hunter, M.S., Koglin, J.E., et al. (2015). The coherent X-ray imaging instrument at the Linac Coherent Light Source. *J. Synchrotron Radiat.* **22**, 514–519.
- Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713.

- Ortega, A., Amorós, D., and Garcia de la Torre, J. (2011). Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* **101**, 892–898.
- Parra, M.C., Shaffer, S.A., Hajjar, A.M., Gallis, B.M., Hager, A., Goodlett, D.R., Guina, T., Miller, S., and Collins, C.M. (2010). Identification, cloning, expression, and purification of *Francisella* lpp3: an immunogenic lipoprotein. *Microbiol. Res.* **165**, 531–545.
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802.
- Post, D.M.B., Slutter, B., Schilling, B., Chande, A.T., Rasmussen, J.A., Jones, B.D., D'Souza, A.K., Reinders, L.M., Harty, J.T., Gibson, B.W., et al. (2017). Characterization of inner and outer membrane proteins from *Francisella tularensis* strains LVS and Schu S4 and identification of potential subunit vaccine candidates. *MBio* **8**, <https://doi.org/10.1128/mBio.01592-17>.
- Roberts, L.M., Powell, D.A., and Frelinger, J.A. (2018). Adaptive immunity to *Francisella tularensis* and considerations for vaccine development. *Front. Cell. Infect. Microbiol.* **8**, 115.
- Singharoy, A., and Chipot, C. (2017). Methodology for the simulation of molecular motors at different scales. *J. Phys. Chem. B* **121**, 3502–3514.
- Singharoy, A., Chipot, C., Moradi, M., and Schulten, K. (2017). Chemomechanical coupling in hexameric protein-protein interfaces harnesses energy within V-type ATPases. *J. Am. Chem. Soc.* **139**, 293–310.
- Su, J.L., Yang, J., Zhao, D.M., Kawula, T.H., Banas, J.A., and Zhang, J.R. (2007). Genome-wide identification of *Francisella tularensis* virulence determinants. *Infect. Immun.* **75**, 3089–3101.
- Szabo, A., Schulten, K., and Schulten, Z. (1980). First passage time approach to diffusion controlled reactions. *J. Chem. Phys.* **72**, 4350–4357.
- Terwilliger, T.C., Dimaio, F., Read, R.J., Baker, D., Bunkoczi, G., Adams, P.D., Grosse-Kunstleve, R.W., Afonine, P.V., and Echols, N. (2012). phenix.mr_rosetta: molecular replacement and model rebuilding with Phenix and Rosetta. *J. Struct. Funct. Genomics* **13**, 81–90.
- Trott, O., and Olson, A.J. (2010). Software news and update AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461.
- Tu, K., Tobias, D.J., and Klein, M.L. (1995). Constant pressure and temperature molecular dynamics simulation of a fully hydrated liquid crystal phase dipalmitoylphosphatidylcholine bilayer. *Biophys. J.* **69**, 2558–2562.
- Wampler, R.D., Kissick, D.J., Dehen, C.J., Gualtieri, E.J., Grey, J.L., Wang, H.-F., Thompson, D.H., Cheng, J.-X., and Simpson, G.J. (2008). Selective detection of protein crystals by second harmonic microscopy. *J. Am. Chem. Soc.* **130**, 14076–14077.
- Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., and Case, D.A. (2004). Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174.
- Wang, J., Wang, W., Kollman, P.A., and Case, D.A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260.
- Weierstall, U., James, D., Wang, C., White, T.A., Wang, D., Liu, W., Spence, J.C., Bruce Doak, R., Nelson, G., Fromme, P., et al. (2014). Lipidic cubic phase injector facilitates membrane protein serial femtosecond crystallography. *Nat. Commun.* **5**, 3309.
- White, T.A., Kirian, R.A., Martin, A.V., Aquila, A., Nass, K., Barty, A., and Chapman, H.N. (2012). CrystFEL: a software suite for snapshot serial crystallography. *J. Appl. Crystallogr.* **45**, 335–341.
- White, T.A., Mariani, V., Brehm, W., Yefanov, O., Barty, A., Beyerlein, K.R., Chervinskii, F., Galli, L., Gati, C., Nakane, T., et al. (2016). Recent developments in CrystFEL. *J. Appl. Crystallogr.* **49**, 680–689.
- Zook, J., Mo, G., Sisco, N.J., Craciunescu, F.M., Hansen, D.T., Baravati, B., Cherry, B.R., Sykes, K., Wachter, R., Van Horn, W.D., et al. (2015). NMR structure of *Francisella tularensis* virulence determinant reveals structural homology to Bet v1 allergen proteins. *Structure* **23**, 1116–1122.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
<i>E. coli</i> Lemo21(DE3)	New England Biolabs	NA
Chemicals, Peptides, and Recombinant Proteins		
Trizma® base	Sigma-Aldrich	Cat# T1503-1KG
Sodium Chloride	Sigma-Aldrich	Cat# S7653-1KG
Imidazole	Sigma-Aldrich	Cat# 56750-1KG
Lysozyme	Sigma-Aldrich	Cat# L6876-10G
Calcium chloride	Sigma-Aldrich	Cat# C1016-500G
Restriction Protease Factor Xa	Sigma-Aldrich	Cat# 11585924001
Ultra-low temperature gelling agarose	Sigma-Aldrich	Cat# A5030-5G
Deposited Data		
Flpp3 _{xtal} Structure	This Paper	PDB: 6PNY
Flpp3 _{nmr} Structure	Zook et al., 2015	PDB: 2MU4
Recombinant DNA		
Plasmid: <i>pRSET-His6-Xa-Flpp3sol</i>	DNASU	Accession# FtCD00697202
Software and Algorithms		
NAMD	Phillips et al., 2005	http://www.ks.uiuc.edu/Research/namd/
VMD	Humphrey et al., 1996	http://www.ks.uiuc.edu/Research/vmd/
HYDROPRO	Ortega et al., 2011	http://leonardo.inf.um.es/macromol/programs/hydropro/hydropro.htm
DNASU	Cormier et al., 2011	https://dnasu.org/DNASU/
CHEETAH	Barty et al., 2014	http://www.desy.de/~barty/cheetah/Cheetah/Welcome.html
CrystFEL	White et al., 2016	http://www.desy.de/~twhite/crystfel/
PHENIX	Adams et al., 2010	https://www.phenix-online.org/
ROSETTA-MR	Terwilliger et al., 2012	https://www.rosettacommons.org/
Amber	Case et al., 2005	http://ambermd.org/
phenix.refine	Adams et al., 2010; Afonine et al., 2012	https://www.phenix-online.org
Coot	Emsley et al., 2010	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
PyMOL	Schrödinger	http://www.pymol.org

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Petra Fromme (pfromme@asu.edu). DNA constructs generated in this study have been deposited to DNASU (<http://dnasu.org>).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Microbes

E. coli cells were cultured in LB medium.

METHOD DETAILS

Cloning, Expression, Purification, and Crystallization

All vectors and their sequences are available from the PSI:Biology-Materials Repository at DNASU (<http://dnasu.org>) (Cormier et al., 2011). The expression plasmid *pRSET-His6-Xa-Flpp3sol* (DNASU accession #FtCD00697202), encoding for the protein sequence Flpp3_{xtal} was generated by ligation-independent cloning (Clontech In-Fusion HD Cloning Plus) of an insert that was PCR-generated

from *pRSET-FTT1416sol-6xHis* (Zook et al., 2015) into BseRI-digested *pRSET-C8xHis* (DNASU #EvNO00629010). The final open reading frame consisted of an N terminal methionine (ATG), a hexahistidine tag (CATCACCACCATCAC), the Factor Xa cleavage recognition site (amino acids IEGR, encoded by ATTGAAGGTCGC), and amino acid residues Q30-T137 of gene *FTT1416* from SCHU S4 (UniProtKB Accession Number: Q5NF33), followed by a strong double-stop codon (TAATAG). The *pRSET-C8xHis* parent vector had been constructed by modification of the small BseRI/EcoRI fragment of *pRSET-natGFPHis* (DNASU #EvNO00623704). DNA sequence confirmation was performed at the School of Life Sciences DNA Laboratory at Arizona State University.

Transformation, expression, and purification of Flpp3_{xtal} was performed similarly to Flpp3_{nmr} (Zook et al., 2015). The IMAC purified protein underwent buffer exchange by size exclusion chromatography into a low-salt buffer containing 50 mM NaCl, 1 mM CaCl₂, and 20 mM Sodium Phosphate, pH 6.5; the optimal solution conditions for Factor Xa protease activity. Protein was then concentrated to 20 mg/mL, and Factor Xa, from bovine plasma (Sigma) was added to a concentration of 1 U/mL from a stock concentration of 10 U/mL in nanopure water. The protein was vortexed for 5 seconds and then incubated at room temperature to crystallize, without removal of Factor Xa or the cleaved tag.

Unexpectedly, we observed that nanocrystals of Flpp3 could be grown very rapidly after addition of Factor Xa protease. A significant formation of nanocrystals was observed already between 20 seconds and 5 minutes after addition of the Factor Xa protease, without the introduction of precipitating agent or significant time for Factor Xa protease to cleave the affinity tag. Crystal growth was completed 10 minutes after vortexing.

The question if Flpp3 crystallization was induced by the cleavage of the His-tag was answered by the fact that electron density was well defined at the C-terminus which allowed us to model four histidines and the Factor Xa cleavage sequence into the electron density map. Thereby we conclude that that His-tag cleavage is not required for crystallization, instead crystallization is induced via vortexing the concentrated Flpp3_{xtal} sample. This was tested via vortexing a 15 mg/mL sample of Flpp3_{xtal} in cleavage buffer in the absence of Factor Xa protease, which resulted in crystal formation.

SFX Data Collection and Structure Calculation

Flpp3_{xtal} crystals were mixed with 7% (w/v) ultralow-gelling-temperature agarose in the low-salt buffer to a final agarose concentration of 5.6% as described by Conrad et al. (Conrad et al., 2015). The viscous sample was injected transverse to the XFEL beam at the CXI (Liang et al., 2015) instrument of LCLS using an ASU viscous injector (Weierstall et al., 2014). The crystal-embedded agarose sample was delivered as a continuous stream extruded into a vacuum chamber from a 50 μ m capillary at a flow rate of 160 nl min⁻¹. Diffraction data were collected with a CSPAD detector (Koerner et al., 2009) at a repetition rate of 120 Hz, at 8.7 keV, with a pulse length of the x-ray pulses of 35 fs with 2.49 mJ pulse energy. Images were sorted using CHEETAH to find crystal hits and resulting diffraction patterns were indexed, merged, and scaled using CrystFEL (Barty et al., 2014; White et al., 2012, 2016).

Merged diffraction data were phased by ROSETTA-MR molecular replacement program as part of the PHENIX suite employing the NMR-solved Flpp3sol (PDBID: 2MU4) ensemble as starting models (Terwilliger et al., 2012). Model building utilized PHENIX software (Adams et al., 2010; Afonine et al., 2012) and rebuilt in COOT and further refined with phenix.refine. The refined structures were visualized with PyMOL (<http://pymol.org>).

Conventional Molecular Dynamics

NMR and X-ray conformations were solvated with TIP3P water and neutralized with 100 mM NaCl. Initial equilibration was performed with NAMD2 (Phillips et al., 2005) in an NPT ensemble with periodic boundary conditions. The simulations were performed at 310 K using Langevin dynamics (Tu et al., 1995) with a damping constant of 0.5 ps⁻¹ running for a total of 100 ns. Nosé-Hoover Langevin piston method (Tu et al., 1995) was used to maintain constant pressure at 1 atm. The cutoff used for the short-range interactions were 12 Å with the switching applied at 10 Å. Particle mesh Ewald (PME) algorithm (Darden et al., 1993) was utilized to calculate long-range electrostatic force. Bonded, non-bonded, and PME calculations were performed at 2-, 2-, and 4-fs intervals, respectively.

Steered Molecular Dynamics

Initial conformations for the free energy calculations were generated by performing steered molecular dynamics simulation (SMD) with the NMR and X-ray as initial and target conformations respectively. SMD was performed in the constant velocity mode employing a force constant of $k = 100$ kcal/mol/Å² running for 100 ns. The collective variable used in the SMD was the RMSD difference between the residues from the target and the initial conformations, utilizing the colvar module (Fiorin et al., 2013) as implemented in NAMD2.

Umbrella Sampling

Initial conformations were sampled from the SMD trajectory in such a way that the RMSD difference between the consecutive conformations was less than 0.1 Å, thereby resulting in 49 conformations. Umbrella Sampling (US) simulations were performed by adjusting the harmonic restraints, ensuring an optimum overlap in the distributions of the neighboring windows. US calculations were performed on 40 windows each simulated for 60 ns, resulting in a cumulative simulation time of 2.4 μ s. The force constant for the harmonic restraints ranged from 10-100 kcal/mol/Å², ensuring an optimum overlap between the neighboring windows.

Kinetics Calculation

Kinetics calculations provide an estimate of the mean free passage time τ , which is the inverse of the rate constant k . Assuming a diffusion dominated motion of the protein, the free energy profile obtained from the umbrella sampling simulations can be connected to the mean free passage time using the following expression derived by Szabo et. al (Szabo et al., 1980).

$$\tau = k^{-1} = \sum_{l=i}^j \left\{ D(x_l) \frac{e^{-\beta F(x_l)}}{\sum_{k=i}^j e^{-\beta F(x_k)}} \right\}^{-1}$$

Here, the transition pathway is discretized into points $i, i+1, \dots, k, \dots, j$. $D(x_l)$ and $F(x_l)$ are the diffusion coefficient and free energy at point x_l in the pathway. $D(x)$ was calculated using the opensource software Hydropro (Ortega et al., 2011). A specially homogeneous diffusion model was assumed. The quantity $\frac{e^{-\beta F(x_l)}}{\sum_{k=i}^j e^{-\beta F(x_k)}}$ represents the probability of finding the system at the point x_l along the transition pathway.

Ligand Screening

A virtual ligand library was used to screen possible small-molecule binding partners to the Flpp3_{nmr} interior binding cavity observed previously in the NMR structure. The library was generated using the standard fragment-based subset of the ZINC database (Irwin and Shoichet, 2005) which contains 847,909 possible ligand fragments. These molecules were docked to the cavity of Flpp3_{nmr} using Autodock Vina and scored by lowest binding energy (Trott and Olson, 2010).

Using the Amber software suite, the best scoring ligands were further studied by MD simulation followed by Generalized Born Surface Analysis (GBSA) to estimate binding energy of the potential ligand (Case et al., 2005). Initial equilibration of the receptor ligand-complex was performed using Amber with ff14SB (Maier et al., 2015) forcefield being used for the receptor and Amber GAFF forcefield (Wang et al., 2006) as determined by the Antechamber program (Wang et al., 2004) in the Amber package. Partial charges of ligands are calculated using the AM1-BCC method (Jakalian et al., 2002). The system was equilibrated for a total of 500 ns and the MM-GBSA calculation was performed for the last 100 ns. The conformations from the last 100 ns of the trajectory were sampled at every 10 ps, thereby resulting in a total of 10,000 snapshots. The various free energy terms of complexes, receptors, and ligands derived from the snapshots are calculated by the MM/GBSA methods. The final binding free energy is an average of the binding free energies of the aforementioned conformations.

After potential fragments were identified by MD/GBSA, additional functional groups were added to the fragment to improve ligand binding. Functional groups were chosen based on likely hydrophobic interactions, hydrogen bonds, and salt bridges.

Hydrogen Bond Analysis

Protons were added to the X-ray structure using autopsf plugin of VMD (Humphrey et al., 1996). This process only takes into consideration the X-H bond length (where X is a non-hydrogen atom), ignoring the many possible orientations of the X-H bond. Traditionally, such bond orientations are fixed by performing an energy minimization on the structure followed by standard molecular dynamics simulation. However, this would cause the atoms to shift from their positions in the X-ray structure. Specifically, the water molecules in the structure are expected to drift significantly from their positions. Since we were interested in identifying water-mediated hydrogen bonds in the X-ray structure, this was not a feasible option. Instead, we performed molecular dynamics flexible fitting (MDFF) of the X-ray structure using the density map. The map resolution was high enough to resolve the water-oxygens as illustrated in Figure 3C. We coupled all the heavy atoms of the protein as well as these water-oxygens to the map, ensuring they do not drift from their initial position. The protons were free to reorient. We performed 50 ps of MDFF simulation, saving structures every 1 ps.

Hydrogen bonds were identified using the hbonds plugin of VMD, using a donor-acceptor distance of 3.0 Å and donor-H-acceptor angle of 25 degrees. For direct protein-symmetry mate hydrogen bonds, we only considered charged residues. Identification of water-mediated hydrogen bonds between the protein and the symmetry mates was done in two steps. First, hydrogen bonds were identified between monomer and water, then between symmetry mates and water. Then, these two lists were compared to identify the water residues that were common to both, yielding the list of hydrogen bonds mediated by a single water residue.

DATA AND CODE AVAILABILITY

The accession numbers for the refined structure of Flpp3_{xtal} is PDB: 6PNY. The Flpp3_{nmr} structure referenced for use in MD studies is under accession number PDB: 2MU4. The UniProt accession code Q5NF3 for *F. tularensis* Flpp3 was used in this study. All other data are available from the corresponding author on reasonable request.

Supplemental Information

XFEL and NMR Structures of *Francisella*

Lipoprotein Reveal Conformational Space

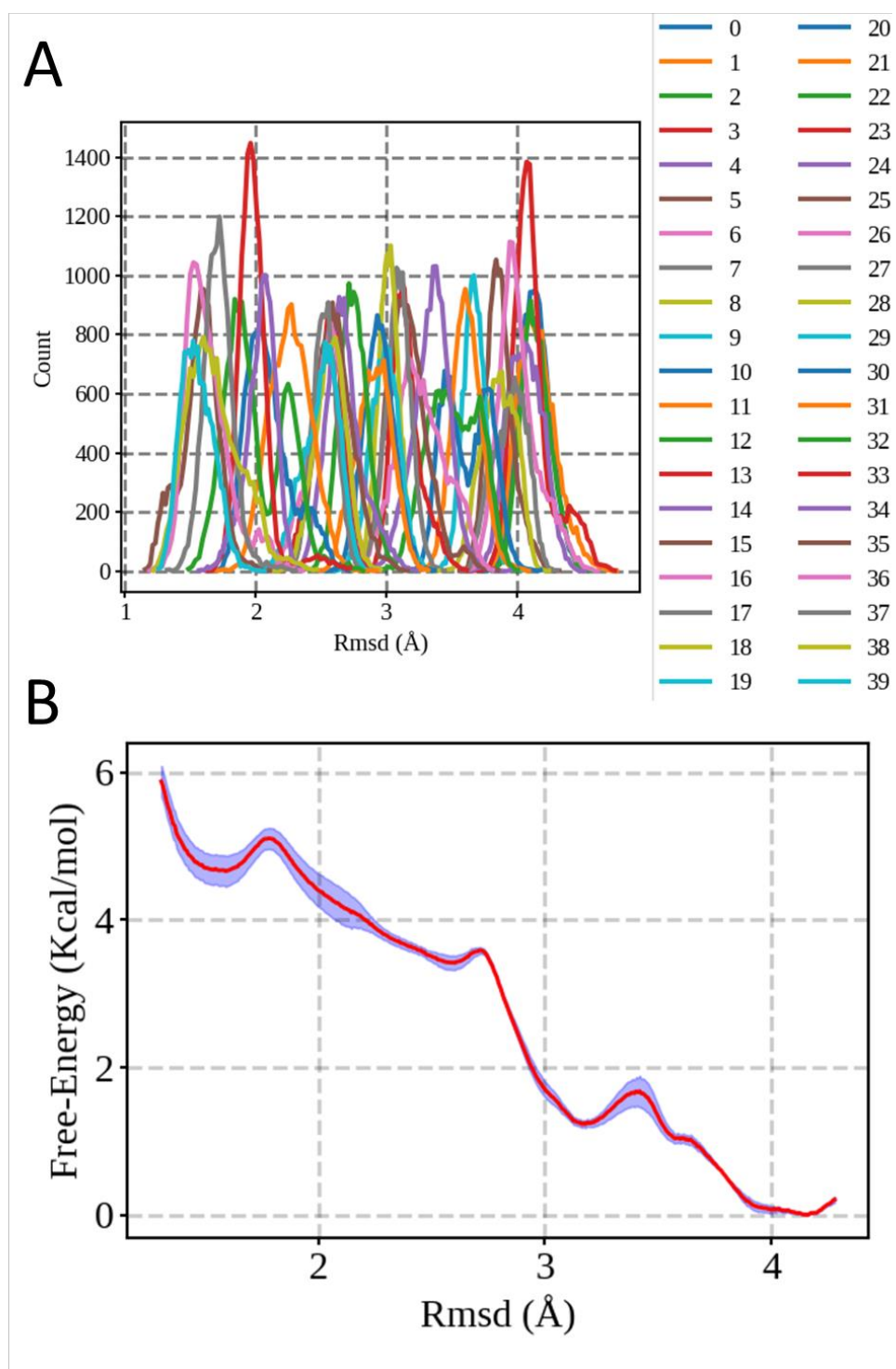
of Drug Target against Tularemia

James Zook, Mrinal Shekhar, Debra Hansen, Chelsie Conrad, Thomas Grant, Chitrak Gupta, Thomas White, Anton Barty, Shibom Basu, Yun Zhao, Nadia Zatsepin, Andrii Ishchenko, Alex Batyuk, Cornelius Gati, Chufeng Li, Lorenzo Galli, Jesse Coe, Mark Hunter, Meng Liang, Uwe Weierstall, Garret Nelson, Daniel James, Benjamin Stauch, Felicia Craciunescu, Darren Thifault, Wei Liu, Vadim Cherezov, Abhishek Singharoy, and Petra Fromme

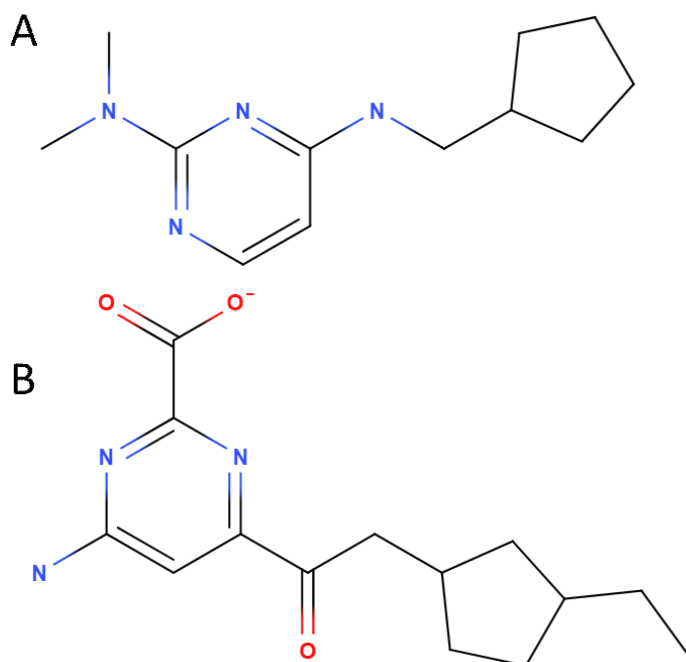
Supplemental Figures

```
Flpp3-wt      MKGLKAKIYIIIFLA AVLAVISGCATDKGTQYKDGYYITTLNYNFNTVYNATLQAIQNGQT
Flpp3-nmr     -----MDKGTQYKDGYYITTLNYNFNTVYNATLQAIQNGQT
Flpp3-xtal    -----MHHHHHIEGRQYKDGYYITTLNYNFNTVYNATLQAIQNGQT
                  : * *****
Flpp3-wt      FDYKSNPYDISVNKNNGTDAEIVSASDSDSTD SLQVAMKKLPNNATRISIKYGSQGNSIR
Flpp3-nmr     FDYKSNPYDISVNKNNGTDAEIVSASDSDSTD SLQVAMKKLPNNATRISIKYGSQGNSIR
Flpp3-xtal    FDYKSNPYDISVNKNNGTDAEIVSASDSDSTD SLQVAMKKLPNNATRISIKYGSQGNSIR
                  *****
Flpp3-wt      SSALIGIIEGNIRYANT-----
Flpp3-nmr     SSALIGIIEGNIRYANTHHHHHH
Flpp3-xtal    SSALIGIIEGNIRYANT-----
                  *****
```

Supplementary Figure 1, relating to Figure 1: Sequence alignment of Flpp3 constructs studied. Flpp3-wt (red) corresponds to the wild-type amino acid sequence encoded by gene FTT1416c from *F. tularensis* strain Schu S4. Flpp3-nmr (blue) is the protein sequence used to solve the NMR structure of Flpp3 (2mu4). This sequence omits the transmembrane signal peptide and palmitoylated cysteine at the N-terminus and adds a hexahistidine tag at the C-terminus. Flpp3-xtal (black) refers to the modified sequence used to crystallize Flpp3 for SFX studies in this study. The N-terminus has been truncated by six residues compared to Flpp3-nmr with the hexahistidine tag moved to the N-terminus followed by the Factor Xa cleavage site (cleaving after IFGR).



Supplementary Figure 2, relating to Figure 4: Umbrella Sampling results showing (A) The histogram of 40 overlapping windows and (B) The convergence plot for the PMF.



Supplementary Figure 3, relating to Figure 6: Virtual ligand binding studies.

A) N4-(cyclopentylmethyl)-N2,N2,N4-trimethyl-pyrimidine-2,4-diamine (ZINC ID: 77213780) was the fragment identified as a potential binding molecule to Flpp3 with an estimated binding affinity to a K_i of 4.1 μM . B) The fragment was modified by replacing the N2 dimethyl moiety to a carboxylate and adding a primary amine at the C6 pyrimidine position; the N4 amine was converted to a ketone and an ethyl moiety was added to the C3 position of the cyclopentyl group, increasing the binding affinity to a K_i of 277 nM.