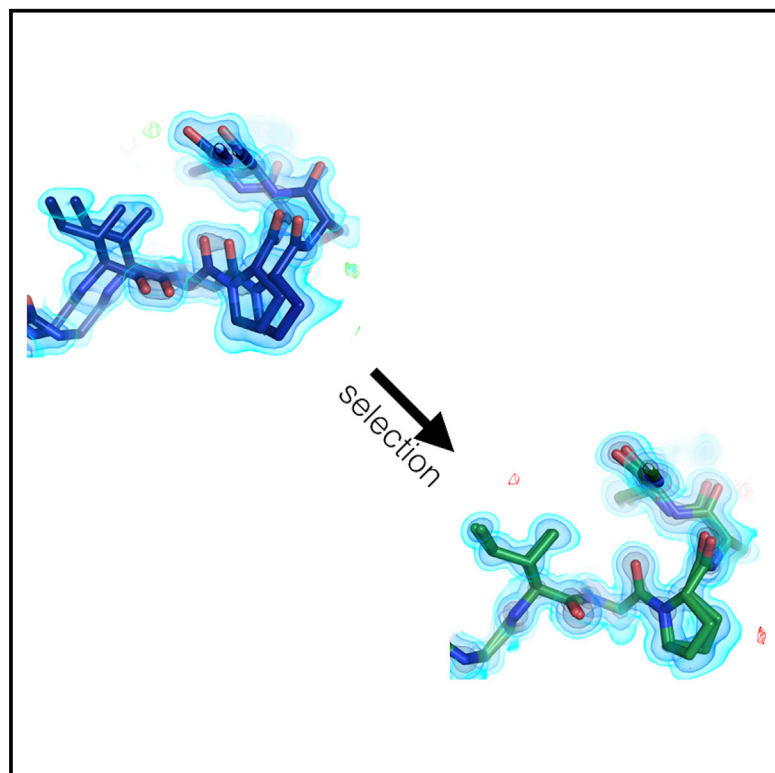


# Structure

## Flexibility and Design: Conformational Heterogeneity along the Evolutionary Trajectory of a Redesigned Ubiquitin

### Graphical Abstract



### Authors

Justin T. Biel, Michael C. Thompson, Christian N. Cunningham, Jacob E. Corn, James S. Fraser

### Correspondence

jfraser@fraserlab.com

### In Brief

Biel et al. use room-temperature X-ray crystallography and multiconformer modeling to uncover that a ubiquitin variant designed to inhibit USP7 shows high conformational heterogeneity. An improved variant previously selected by phage display reduces the observed heterogeneity and increases the binding affinity to USP7.

### Highlights

- Ubiquitin variant designed to inhibit USP7 shows high conformational heterogeneity
- Insights from multiconformer modeling of high-resolution room-temperature X-ray data
- Further optimization by phage display decreased heterogeneity
- New stabilizing interactions are not used in standard protein design calculations

# Flexibility and Design: Conformational Heterogeneity along the Evolutionary Trajectory of a Redesigned Ubiquitin

Justin T. Biel,<sup>1,2</sup> Michael C. Thompson,<sup>1</sup> Christian N. Cunningham,<sup>3</sup> Jacob E. Corn,<sup>3,4</sup> and James S. Fraser<sup>1,5,\*</sup>

<sup>1</sup>Department of Bioengineering and Therapeutic Sciences

<sup>2</sup>Graduate Group in Biophysics

University of California San Francisco, San Francisco, CA 94158, USA

<sup>3</sup>Department of Early Discovery Biochemistry, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080, USA

<sup>4</sup>Present address: Innovative Genomics Initiative, University of California Berkeley, Berkeley, CA 94702, USA

<sup>5</sup>Lead Contact

\*Correspondence: [jfraser@fraserlab.com](mailto:jfraser@fraserlab.com)

<http://dx.doi.org/10.1016/j.str.2017.03.009>

## SUMMARY

Although protein design has been used to introduce new functions, designed variants generally only function as well as natural proteins after rounds of laboratory evolution. One possibility for this pattern is that designed mutants frequently sample nonfunctional conformations. To test this idea, we exploited advances in multiconformer modeling of room-temperature X-ray data collection on redesigned ubiquitin variants selected for increasing binding affinity to the deubiquitinase USP7. Initial core mutations disrupt natural packing and lead to increased flexibility. Additional, experimentally selected mutations quenched conformational heterogeneity through new stabilizing interactions. Stabilizing interactions, such as cation- $\pi$  stacking and ordered waters, which are not included in standard protein design energy functions, can create specific interactions that have long-range effects on flexibility across the protein. Our results suggest that increasing flexibility may be a useful strategy to escape local minima during initial directed evolution and protein design steps when creating new functions.

## INTRODUCTION

One of the major aims of both protein engineering and computational design is to create proteins with new or enhanced functions (Butterfoss and Kuhlman, 2006). Proteins can be redesigned to promote new catalytic reactions or, in the simpler case, for binding to new partners. Recent years have seen successes including the design of new enzymes (Kiss et al., 2013), new small-molecule binders (Tinberg et al., 2013), and new protein-protein interfaces (Karanicolas et al., 2011). However, these redesigned proteins are generally optimized with many rounds of laboratory evolution to achieve functions approaching those of natural systems (Blomberg et al., 2013; Khersonsky et al.,

2010). There are several ways that the engineering or design process can fall short of optimal function: the protein may not fold in the conformation predicted; the intended structure may not imbue the desired function; or the protein may dominantly populate nonfunctional conformations. We (Bhabha et al., 2015), and others (Korendovych and DeGrado, 2014; Osuna et al., 2015), have speculated that the third explanation is a major reason why redesigned proteins are only marginally functional initially. The hypothesis is that initial mutations introduced during redesign disrupt the native packing of the parental wild-type (WT) protein, resulting in increased flexibility and sampling of nonfunctional conformations.

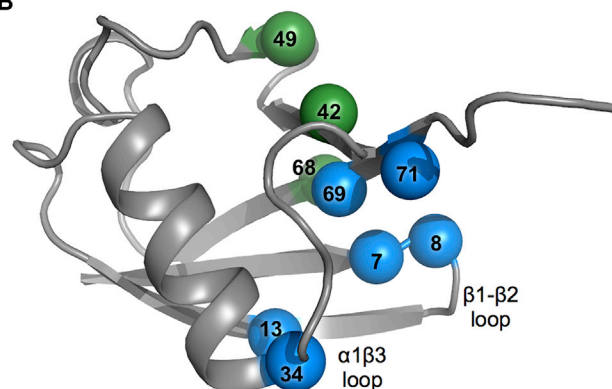
To improve function, subsequent directed evolution experiments fix mutations that act to stabilize the functionally important conformations found in the broadened ensemble of the redesigned protein (Frushicheva et al., 2014). This pattern of design and selection has been performed for functions including small-molecule binding (Tinberg et al., 2013) and protein-protein interaction (Karanicolas et al., 2011), the most developed examples of directed evolution changing dynamics and function that have emerged from enzyme design. Molecular dynamics simulations of designed enzymes indicate that initial designs suffer from poor preorganization (Frushicheva et al., 2010) and that mutations acquired during further selections reduce conformational dynamics near the active site (Osuna et al., 2015). Molecular dynamics simulations have also been used as a computational filter to screen out designs that suffer from conformational instability (Wijma et al., 2015). In principle, nuclear magnetic resonance (NMR) methods could be used to assess whether local dynamics are altered from WT by redesign and subsequent evolution (Johnson and Handel, 1999; Walsh et al., 2001). However, it has remained difficult to track changes in conformational heterogeneity during these processes because of the inability to resolve the relevant conformations in atomic detail for proteins that select for new binding or catalytic activities.

To overcome these limitations, X-ray crystallographic techniques to resolve alternative conformations with atomic resolution have recently been developed (Burnley et al., 2012; Keedy et al., 2015). The key insight enabling these techniques is the recognition that weak signals present in electron density maps represent alternative conformations (Lang et al., 2010, 2014).

A

	7	8	13	34	36	42	49	68	69	71
WT	T	L	I	E	I	R	Q	H	L	L
core	F	R	Y	L	I	R	Q	H	G	R
affinity enhanced	F	R	Y	L	I	W	R	R	G	R

B



**Figure 1. Locations and Identities of Mutations Made across Directed Evolution Trajectory of Ubiquitin**

(A) Table showing amino acid identities of mutation sites for the wild-type (WT), core, and affinity-matured mutants. Blue coloring represents residue identities first introduced in the core mutant. Green shows new residue identities for the affinity-matured mutations.

(B) WT ubiquitin (PDB: 1ubq) model shown in gray with spheres representing locations of mutation sites, colored as in (A) and labeled by residue number.

These signals can be interpreted with a multiconformer model whereby individual residues are built as a parsimonious set of alternative conformations with differing coordinates, occupancies, and B factors (Woldeyes et al., 2014). These features are often more noticeable in X-ray data collected at room temperature. Interpreting these signals in systems such as CypA (Fraser et al., 2009), Ras (Fraser et al., 2011), and DHFR (van den Bedem et al., 2013), has delivered new insights into the structural basis of correlated protein dynamics and their relationship to function. By monitoring differences in multiconformer models, it should be possible to observe how conformational heterogeneity changes as a function of redesign and laboratory evolution for novel enzymatic or binding activities. Here we investigated the changes to conformational heterogeneity of ubiquitin variants selected to bind tightly and specifically to the deubiquitinase, ubiquitin-specific protease 7 (USP7), also known as HAUSP in humans (Zhang et al., 2013).

Ubiquitin is a hub protein that binds partners with several different interfaces (Husnjak and Dikic, 2012). Previous studies have linked the conformational flexibility of ubiquitin, particularly in the  $\beta 1\beta 2$  loop, to its ability to bind these diverse binding partners (Lange et al., 2008). Unlike the enzyme design examples discussed above, USP7 is already a natural, albeit weak, binder of WT ubiquitin. Zhang et al. (2013) subjected ubiquitin to a com-

bination of Rosetta design and phage display to encode additional affinity and specificity for USP7. The premise of the project was to mutate the core of ubiquitin to stabilize the “down” conformation of the  $\beta 1\beta 2$  loop. An initial mutant was selected from a phage display library of seven sites selected based on Rosetta Design calculations. Originally referred to as u7ub25, this mutant is referred to herein as the “core” mutant. The final “affinity-matured” mutant, originally called u7ub25\_2540, contains an additional three surface mutations that were selected by additional phage display experiments (Figure 1). While structures of the variants bound to USP7 have not been possible to obtain, we reasoned that the structural and dynamic features of the unbound variants could provide new insights into the forces that stabilize binding. Here, we determined the structures of both the core and affinity-matured variants to high resolution using room-temperature X-ray diffraction. The structure of the affinity-matured variant showed that the  $\beta 1\beta 2$  loop adopted the same “up” conformation as WT ubiquitin (Zhang et al., 2013), and as the minor conformer of the core mutant. Thus, the mechanism by which these mutations lead to increased specificity remains unclear. We found that the heterogeneity of these mutants follow a trend whereby the initial mutations lead to increased flexibility when compared with previously determined WT ubiquitin crystal structures in different crystal forms, and that the final mutations then stabilize a dominant conformation. Our results show how characterizing the conformational landscapes of redesigned proteins could improve protein engineering and computational protein design.

## RESULTS

### Conformational Heterogeneity of the $\beta 1\beta 2$ Loop

During initial refinement and model building (Tables 1 and 2), we noticed electron density clearly motivating the need for multiconformer models (Figure 2A). For example, the  $\beta 1\beta 2$  loop displays many large difference density features that cannot all be accounted for by a single conformer model with isotropic B factors (Figure 2A), or even with anisotropic B factors. In fact, if added first, the anisotropic B factors can spread into the difference signal for alternative conformations, making the density much more difficult to interpret. Therefore, we chose to model alternative conformations first, then add anisotropic B factors if necessary. In an initial attempt to build a multiconformer model, we used the automated program qFit (Keedy et al., 2015; van den Bedem et al., 2009), which builds multiconformer models via a “sample-and-select” procedure. While qFit has been recently updated to accommodate more backbone flexibility (Keedy et al., 2015), it was not intended to model large displacements with separated backbone density peaks. Consequently, qFit is not able to model highly flexible areas, such as this  $\beta 1\beta 2$  loop (Figure 2B). For some residues, such as Phe7 in the core mutant, side chains are even moved by qFit into density that is unlikely to arise from that residue. Clearly the complex backbone shifts in this region are not well captured by these automated techniques.

Due to the limitations of automated model building for this region, we undertook extensive manual interpretation of these regions with alternative conformations. To ensure that the resultant manually built models were consistent with the data, we

**Table 1. Data Statistics at Different Resolution Cutoffs**

(A) Core Mutant											
Completeness (%)	CC <sub>1/2</sub> (%)	R <sub>merge</sub> (%)	I/σ	Refinement Resolution		1.16 Å	1.12 Å	1.08 Å	1.04 Å		
				R factor Calculation	Resolution (Å)						
100	96.8	17.9	8.46	1.16		14.4/16.3	14.2/16.0	15.1/16.2	15.0/16.1		
100	95.3	22.4	6.89	1.12			14.2/16.1	15.1/16.3	15.0/16.4		
97.2	89.4	29.9	4.47	1.08				15.1/16.5	15.1/16.5		
91.1	74.7	41.8	2.41	1.04					15.2/16.7		
(B) Affinity-Matured Mutant											
Completeness (%)	CC <sub>1/2</sub> (%)	R <sub>merge</sub> (%)	I/σ	Refinement Resolution		1.16 Å	1.12 Å	1.08 Å	1.04 Å	1.00 Å	0.96 Å
				R factor Calculation	Resolution (Å)						
100	98.1	24.8	9.24	1.16		11.0/12.5	11.1/12.4	11.1/12.3	11.3/12.4	11.4/12.4	11.6/12.6
100	96.9	32.5	7.01	1.12			11.0/12.5	11.1/12.4	11.2/12.5	11.3/12.5	11.5/12.6
96.8	90.9	47.9	4.22	1.08				11.2/12.5	11.3/12.5	11.3/12.6	11.5/12.7
87.8	76.7	66.6	2.28	1.04					11.5/12.8	11.5/12.8	11.7/12.9
68.7	41.3	118.3	1.00	1.00						11.8/13.1	12.0/13.3
34.3	7.7	233.1	0.38	0.96							12.2/13.5

(A) Table of values for core mutant structures at different maximal resolutions. Completeness, CC<sub>1/2</sub>, and I/σ are shown for the high-resolution bin at each resolution cutoff choice. Both R<sub>work</sub> and R<sub>free</sub> percentage values up to resolution cutoff shown for each row corresponding to the structure built using data to resolution cutoff shown for each column. Each column represents different structures built using data with different resolution cutoffs; however, structures can only be compared when the same reflections are compared, i.e., values in the same row. (B) Statistics for affinity-matured structures at different resolution cutoffs. As above, completeness, CC<sub>1/2</sub>, and I/σ are shown for the highest-resolution bin, and R/R<sub>free</sub> statistics are shown for all reflections within the designated resolution cutoff (rows). See also Figure S5.

performed two independent interpretations of the electron density maps (by J.T.B. and M.C.T., blinded from each other). The approach of comparing independent refinements has been previously used to assess the accuracy of structure determination under different purifications (Daopin et al., 1994), with different refinement software (Fields et al., 1994), and with the same data (Terwilliger et al., 2007). Although modeling alternative conformations at low signal levels is necessary to successfully interpret and minimize the local difference density, care must be taken not to interpret signal unless there is a stereochemically reasonable model that can be built (Richardson et al., 2013). The independent refinement procedure allowed us to check for consistent interpretation in regions of high disorder, such as the β1β2 loop, where relevant signals for alternative conformations frequently appear only at low electron density contour levels. The resulting models were almost identical, essentially varying only in the interpretation of rotamers for flexible side chains. For residues that had different rotamers or varied in number of conformations modeled (Figure S1), we evaluated the two models based on rotameric positions, steric clashes, plausible tertiary interactions (e.g., hydrogen bonds to nearby side-chain heteroatoms), consistency with the 2F<sub>o</sub>-F<sub>c</sub> map, and the extent to which local F<sub>o</sub>-F<sub>c</sub> difference density was explained. After making a consensus model based on these comparisons, we added anisotropic B factors to protein atoms and finalized the solvent placement. These additions improved the map quality, allowing extra signal for other features to be interpreted in the final model (Figure 2D). In the area of the β1β2 loop, the signal at the high and low contour levels clearly defines the molecular envelope, and the difference density in this region is largely

reduced (Figure 2D). When the final structure is overlaid with the difference density from the single conformer structure (Figure 2C), the new additional conformations clearly explain the previous difference peaks.

The heterogeneity of the β1β2 loop consists of large shifts of the C<sub>α</sub> atoms, separating alternative states by as much as 4.5 Å. Alternative conformations of backbone carbonyls are observed pointing in different orientations and the backbone takes two distinct paths for residues 9–11. These loop conformations differ from the expected conformations contained in NMR models of the WT protein (Figure 3C), in which the β1β2 loop moves in a hinge-like manner, with residues 8–10 moving in unison between “up” and “down” conformations (Lange et al., 2008).

In contrast to the core mutant (Figure 3A), the electron density for the affinity matured mutant is consistent with a much less heterogeneous conformational ensemble (Figures 3B and S2). Although the affinity-matured conformation is closer to the WT “up” conformation of previous crystal structures than the USP7-bound “down” conformation (Figure 3D), it is a more potent and selective binder than either the core mutant or WT (Zhang et al., 2013). Because the crystallization conditions and other considerations such as crystal lattice contacts and data collection temperature are equivalent between these two datasets, our results indicate that the addition of the final three surface mutations in the affinity-matured mutant are responsible for quenching this heterogeneity. Below we outline how specific mutations have acted to increase conformational heterogeneity from the WT to the core mutant and to decrease heterogeneity from the core to affinity-matured mutants.

**Table 2. Data Collection and Refinement Statistics**

	Core (u7ub25)	Affinity Matured (u7ub25_2540)
Wavelength (Å)	0.9537	0.9537
Resolution range (Å)	22.04–1.12 (1.16–1.12)	38.04–1.08 (1.119–1.08)
Space group	P 31 2 1	P 31 2 1
Unit cell dimensions		
a, b, c (Å)	42.72, 42.72, 54.88	43.92, 43.92, 55.44
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 120	90, 90, 120
Total reflections	130,992 (13,152)	311,084 (26,130)
Unique reflections	22,427 (2,242)	27,011 (2,620)
Multiplicity	5.8 (5.9)	11.5 (9.9)
Completeness (%)	0.98 (1.00)	1.00 (0.98)
Mean I/ $\sigma$ (I)	18.25 (7.13)	28.21 (4.70)
Wilson B factor (Å <sup>2</sup> )	9.91	8.95
R <sub>merge</sub>	0.0540 (0.224)	0.0469 (0.454)
R <sub>meas</sub>	0.05949 (0.2458)	0.0492 (0.4788)
CC <sub>1/2</sub>	0.998 (0.955)	1 (0.929)
CC*	1 (0.988)	1 (0.982)
Reflections used in refinement	22,426 (2,242)	27,011 (2,622)
Reflections used for R <sub>free</sub>	1,371 (137)	1,574 (155)
R <sub>work</sub>	0.154 (0.148)	0.103 (0.118)
R <sub>free</sub>	0.175 (0.204)	0.121 (0.158)
CC (work)	0.943 (0.946)	0.976 (0.973)
CC (free)	0.914 (0.892)	0.981 (0.953)
No. of nonhydrogen atoms		
Total	1,135	1,358
Macromolecules	1,093	1,247
Ligands	5	10
Solvent	37	101
Protein residues	73	75
Root-mean-square deviation		
Bonds (Å)	0.016	0.009
Angles (°)	1.57	1.05
Ramachandran favored (%)	99	100
Ramachandran allowed (%)	1.4	0
Ramachandran outliers (%)	0	0
Rotamer outliers (%)	1.7	0
Clashscore	2.71	1.18
B factor (Å <sup>2</sup> )		
Average	12.62	10.95
Macromolecules	12.27	10.02
Ligands	21.76	15.89
Solvent	21.56	21.92

Statistics for the highest-resolution shell are shown in parentheses.

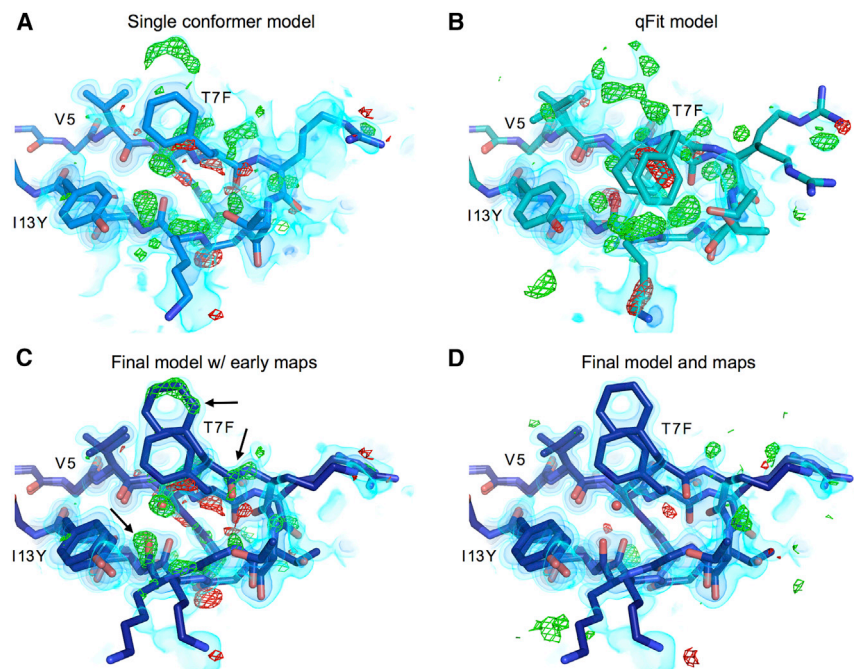
### Structural Changes of Mutations in and near the $\beta 1\beta 2$ Loop

The core variant has six mutations that were introduced with the goal of increasing the affinity to USP7, by changing the packing of the protein core to favor the “down” conformation of the  $\beta 1\beta 2$  loop. Two of the mutations are on the  $\beta 1\beta 2$  loop, namely T7F and L8R, which both mutate the side chains adjacent to the USP7

binding interface. Three of the other mutations are located in the core of the protein (I13Y, E34L, and L69G) and the final mutation, L71R, is on the edge of the hydrophobic core on the C terminus.

Many of the mutations made in the core variant are adjacent to each other, leading to compensatory effects on packing. At the base of  $\beta 1\beta 2$  loop the WT Thr7 is replaced by a bulkier Phe





**Figure 2.  $\beta 1\beta 2$  Loop of the Core Mutant Showing Residues 5–13, Models, and Maps at Different Points in the Refinement Procedure**

(A) Best single conformer model and corresponding  $2mF_o-DF_c$  map shown in volume representation at three contour levels:  $0.65 \text{ e}\text{\AA}^{-3}$ ,  $1.5 \text{ e}\text{\AA}^{-3}$ , and  $3.5 \text{ e}\text{\AA}^{-3}$  from lightest to darkest.  $mF_o-DF_c$  map shown in green and red mesh at  $3$  or  $-3 \text{ e}\text{\AA}^{-3}$ , respectively.

(B) Output model from qFit 2.0 built from single conformer model. While qFit is able to accurately model the side-chain heterogeneity at the more ordered base of chain for residue 5 valine, qFit was unable to capture the backbone heterogeneity in this loop, and in fact may be misled by the complex density as can be seen for the clearly misplaced Phe7 side chains.

(C) Final, manually built multiconformer model with final  $2mF_o-DF_c$  map, and single conformer difference map ( $mF_o-DF_c$ ). This shows how the newly built heterogeneity corresponds to the major difference peaks in the original map.

(D) The final manually built model with corresponding maps. While some difference features still exist, the difference signal in this region has largely been reduced in comparison with the original single conformer structure.

See also Figure S1.

residue, which is then compensated by a large-to-small mutation (L69G) on the adjacent  $\beta 5$  strand (Figures 4A and 4B). While steric packing is conserved, T7F can no longer make hydrogen bonds to the backbone carbonyl of Lys11 or to the side chain of Thr9. Both residues that previously participated in hydrogen bonds in the WT background show enhanced heterogeneity in the core mutant, as seen for the Lys11-carbonyl (Figure 4B). For residue Thr9, the lack of a hydrogen bond allows the threonine side chain to flip out relative to the WT conformation (Figures 4A and 4B). The lack of a side chain at residue 69 (L69G) removes packing interactions with residue 7 (T7F), leading to dramatically shifted conformations of Phe7 visible in the electron density. Collectively, these features likely stabilize the  $\beta 1\beta 2$  loop in WT ubiquitin, and their absence correlates with the increased heterogeneity of the core mutant.

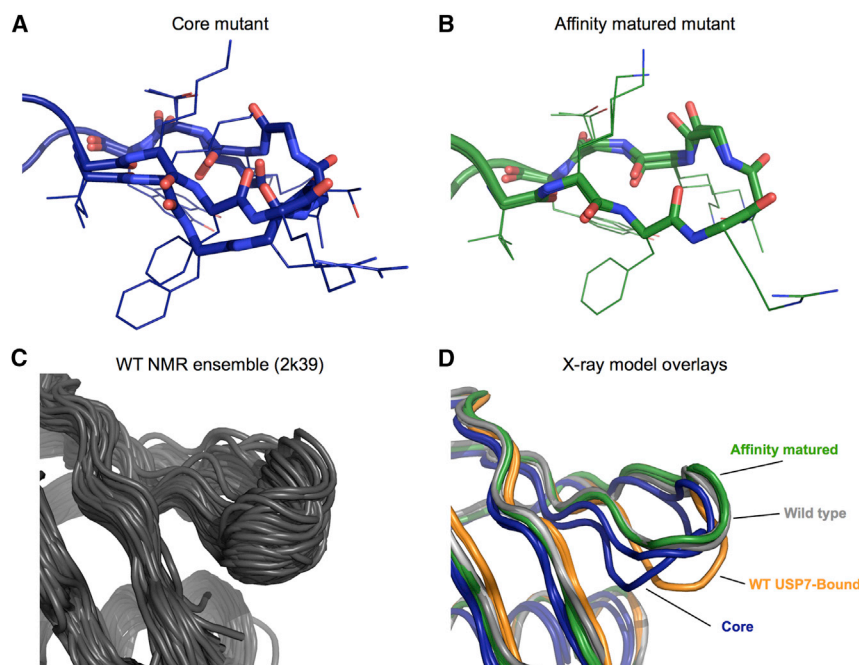
The affinity matured mutant has three additional mutations (R42W, Q49R, and H68R). H68R sits on the final  $\beta$  strand directly adjacent to the  $\beta 1\beta 2$  loop. A bridging water molecule links the carbonyl oxygen of Lys6 and the histidine side chain of residue 68 in the core mutant. This bridging water molecule is modeled in multiple previously determined WT ubiquitin crystal structures (PDB: 3ons and 1yiw) and there is a corresponding unmodeled electron density peak in the original WT ubiquitin dataset (PDB: 1ubq) (Figure S3). In the final affinity-matured mutant this water-mediated interaction is replaced by a direct hydrogen bond between the introduced arginine side chain and the Lys6 carbonyl directly, stabilizing the loop in one conformation (Figures 4C–4E).

### Conformational Heterogeneity Is Reduced in Affinity-Matured $\alpha 1\beta 3$ Loop and C Terminus

Two regions adjacent to the  $\beta 1\beta 2$  loop also follow the same trend as the loop itself, displaying increased heterogeneity

from the WT to the core mutant and decreased heterogeneity from the core to affinity-matured mutants. The first of these areas that display heterogeneity in the core mutant is the C-terminal tail (residues 73–76), which is disordered in the electron density map of the core mutant. The density for L71R, one of the core mutations located just before the C terminus, is not observable for the side chain beyond C $\beta$ , but is ordered in the affinity-matured structure. In the affinity-matured mutant, two of the three new mutations are involved in a new interaction with the C terminus. The R42W and Q49R mutations create a cation- $\pi$ -cation stacking interaction with R72 (Figures 5A–5C). This interaction further links the C terminus to the other  $\beta$  strands of the protein, ordering residue 72 near the C terminus.

A second area that displays significant conformational heterogeneity in the core mutant, but not in the affinity-matured mutant, is the loop region between the  $\alpha 1$  helix and the  $\beta 3$  strand (residues 32–41). In this loop the backbone displays heterogeneity where alternative conformations are shifted by a  $1.0$ – $2.2 \text{ \AA}$ , resulting from a subtle hinging at the ends of the loop (Figure 5D). In the most shifted region in the loop, the Asp39 backbone carbonyl, the density has two discrete peaks corresponding to two states, rather than a smooth continuum that could be modeled by a single conformation with an anisotropic B factor. Within the  $\alpha 1\beta 3$  loop, we observe another pair of compensatory mutations for residues 13 (I13Y) and 34 (E34L) in the core mutant. Specifically, Tyr13 buries Leu34 (Figure S4), maintaining the interaction between aliphatic groups at these side-chain positions. The hydrophilic hydroxyl group in Tyr13 is also retained in a nearly equivalent position to the carboxyl group of Glu34 in the WT protein. Although side-chain interactions appear to be maintained, this new interaction coincides with increased mobility of the  $\alpha 1\beta 3$  loop. Subtle differences in packing that



**Figure 3. Conformations of the  $\beta 1\beta 2$  Loop**

(A) Final multiconformer model of the  $\beta 1\beta 2$  loop for the core mutant. Backbone atoms in the loop are rendered in sticks, while the side chains are left as lines.

(B) Final model of  $\beta 1\beta 2$  loop for the affinity-matured mutant. Panel shown with electron density in [Figure S2](#).

(C) Backbone conformational ensemble from NMR relaxation dispersion experiments (PDB: 2k39). The majority of conformations exhibit a simple backbone shift, producing a hinge-like motion as observed in previous molecular dynamics simulations.

(D)  $\beta 1\beta 2$  loop conformations from different ubiquitin structures. WT ubiquitin apo structure (PDB: 1ubq) in gray, and bound to USP7 in orange (PDB: 1nbf). The core mutant is shown in dark blue, and the affinity matured in green.

See also [Figure S2](#).

result from these amino acid substitutions, and concomitant changes to adjacent residues, may introduce flexibility in this region. In contrast to the core mutant, the  $\alpha 1\beta 3$  loop in the affinity-matured mutant does not display significant heterogeneity ([Figure 5E](#)). Although none of the affinity-matured mutations are in this exact region, it is likely that the additional stabilization of the C terminus gained from the new cation- $\pi$ -cation interaction between residues W42, R49, and R71 in the affinity-matured mutant propagates to the  $\alpha 1\beta 3$  loop via backbone interactions. Notably, a network of hydrogen bonds connects residues 40 and 41 of the  $\alpha 1\beta 3$  loop with residues 70 and 72 of the C terminus. We therefore hypothesize that the new interactions observed in the affinity-matured mutant cooperate with native hydrogen bonding motifs to quench the dynamics of the  $\alpha 1\beta 3$  loop.

### Both States of Residue Asp52-Gly53 Peptide Flip Occur in Core and Affinity-Matured Structures

Residues Asp52 and Gly53 have been previously identified as a structural switch in ubiquitin that undergoes a discrete peptide flip ([Huang et al., 2011](#)) which exchanges on the microsecond timescale ([Sidhu et al., 2011](#); [Smith et al., 2016](#)). The original crystal structure of WT ubiquitin (1ubq), a standard in computational benchmarking studies, shows the peptide in the “NH-out” state with the Asp52 carbonyl making a hydrogen bond to the backbone of the  $\alpha 1$  helix starting at residue Glu24. In the crystal structure (3ons) from conditions similar to those used for solid-state NMR the “NH-in” state is seen, where the Glu24 side chain is swung down relative to the NH-out conformation to make a hydrogen bond with the Glu24 backbone and the G53 NH group. This observation has provided a structural rationale for slow NMR dynamics measurements in solution ([Majumdar and Ghose, 2004](#); [Massi et al., 2005](#)) and in solid state ([Tollinger et al., 2012](#)): the chemical shifts of backbone amides surround-

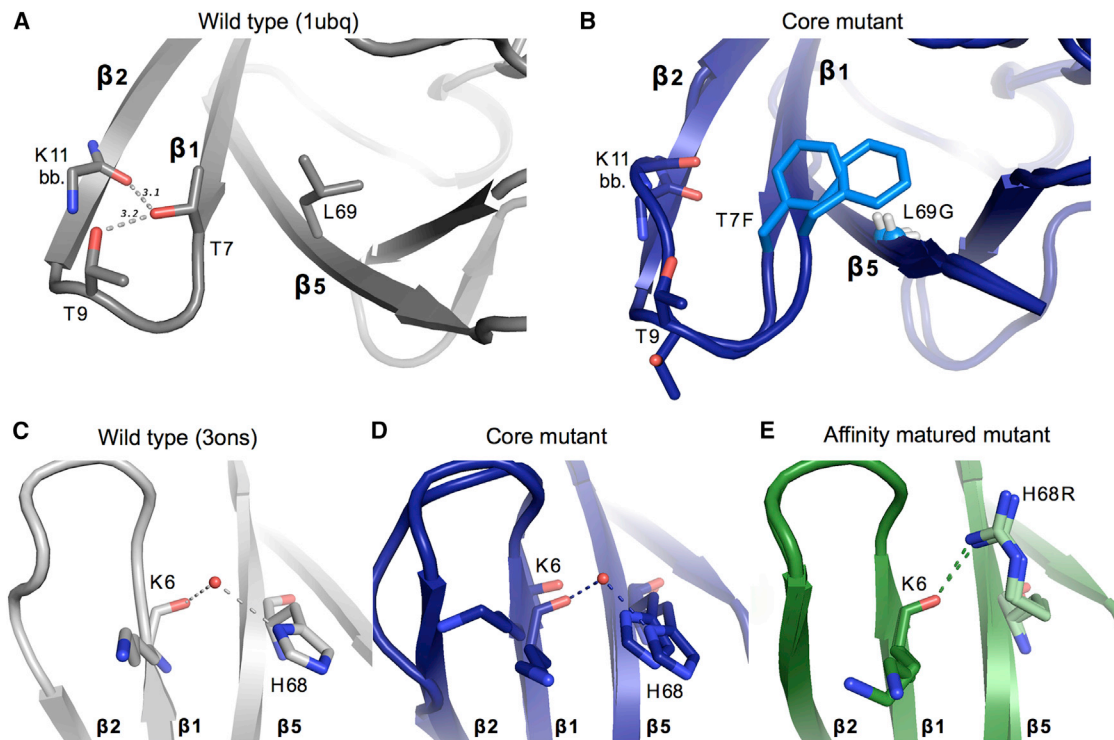
ing residue 24 are perturbed as the side chain transitions to a new rotamer. Moreover, this flip is thought to be a key structural switch between different states of ubiquitin. The flip state of the peptide can

be predicted from the backbone coordinates of other residues across the entire protein ([Smith et al., 2016](#)).

The major conformation in both of our structures corresponds to the NH-in conformation found in the 3ons structure. When modeling this region as a single conformer, we observed signals in the  $F_o - F_c$  map supporting an alternative conformation corresponding to the NH-out state ([Figure 6A](#)). In our final model, the peptide flip is modeled in both states ([Figure 6B](#)) for both structures at occupancies of 60%–70% for the major NH-in and 30%–40% for the NH-out. Interestingly, the NH-in state has been implicated in the binding mode of ubiquitin to the USP class of deubiquitinases which includes the target, USP7 ([Smith et al., 2016](#)). While the population of the NH-in state may have been increased relative to WT, it remains surprising that both peptide flip states are observed in our structures, given the strong association of this peptide flip with the binding of USP7. These structures indicate that it is possible for both of these states to exist within the same crystal form, and provide an additional example of how multiconformer X-ray models can be used to provide a structural basis for dynamics observed by NMR.

## DISCUSSION

We have structurally characterized two ubiquitin variants created via a combination of computational protein design calculations and phage display. These variants are poorly modeled by traditional single conformer models, or even with existing automated model-building tools for regions of high heterogeneity. Enabled by room-temperature X-ray data collection and manually built multiconformer models, we have described the emergence and quenching of conformational heterogeneity along a protein design and engineering trajectory. From these multiconformer models, we observed the interplay of computational protein



**Figure 4. Structural Changes upon Mutation near the  $\beta 1\beta 2$  Loop**

(A) Key packing and hydrogen bond interactions around mutation sites T7F and L69G. WT ubiquitin (1ubq) is shown in gray. Gray dashed lines show hydrogen bonds existent in WT ubiquitin between residues 7, 9, and 11. Sticks are shown for the side chains of residues 7 and 9, as well as for the backbone of residue 11. (B) Both conformations of the residues shown in (A) are shown for the core mutant. Mutated residues are colored in a lighter blue. The C $\alpha$  of Gly69 is shown as a small sphere for clarity.

(C–E) Interactions between residue 6 of  $\beta$ -strand 1 and residue 68 of strand 5 for the WT ubiquitin (gray, C), the core mutant (blue, D), and the affinity-matured mutant (green, E). A modeled water appears linking the backbone of residue 6 with the histidine 68 side chain in both the core mutant and WT (PDB: 3ons). This interaction is directly replaced in the affinity-matured mutant by a hydrogen bond between the new arginine side chain and the backbone carbonyl of residue 6. See also Figure S3.

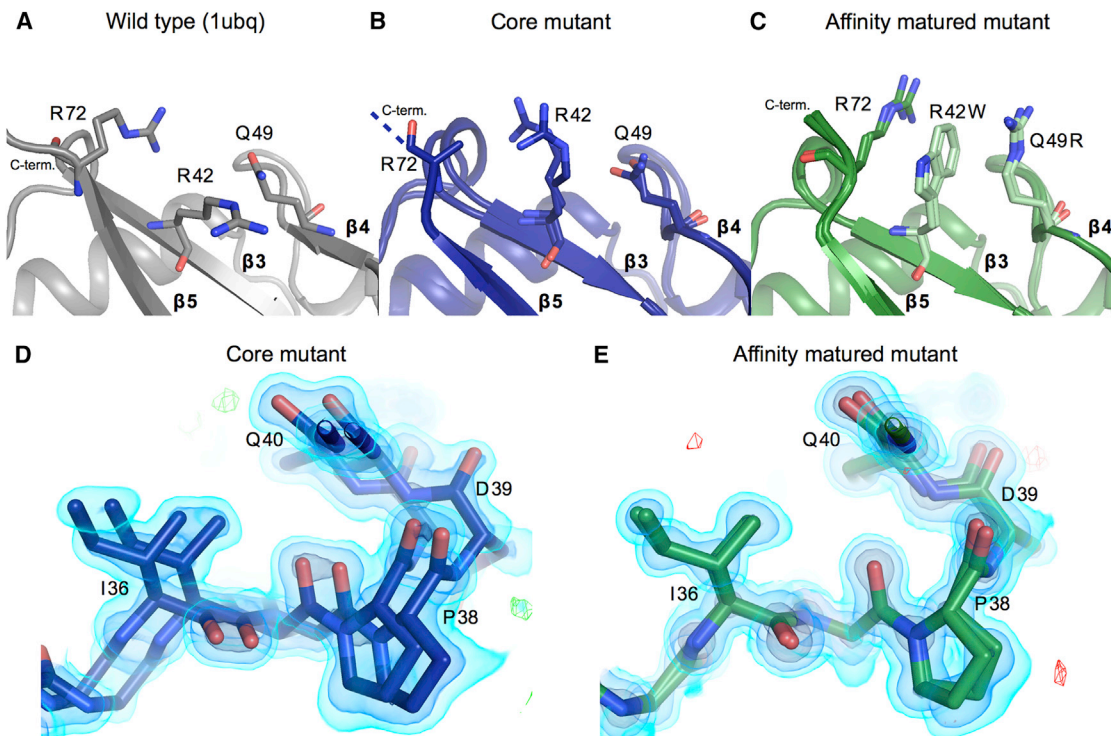
design and laboratory evolution, describing both the core and affinity-matured mutants that were developed to bind tightly to USP7. Notably, the intermediate core mutant displayed significant conformational heterogeneity across the majority of the protein, varying in magnitude. The large-scale backbone motions of the  $\beta 1\beta 2$  loop are in direct contrast to the original goal of stabilizing a single pro-USP7 binding conformation. Instead, the core mutations enhance the flexibility of the protein, creating  $\beta 1\beta 2$  loop conformations that are distinct from states in the hinge-like motion predicted by NMR analyses of the WT protein (Lange et al., 2008). In contrast, additional surface mutations introduced in the final affinity-matured mutant cooperate to reduce flexibility, and the  $\beta 1\beta 2$  loop can be modeled in a single conformation.

Our observations suggest that the process of rational protein design followed by directed evolution resembles simulated annealing procedures used to escape local conformational minima in X-ray refinement (Brünger et al., 1997; Korostelev et al., 2002). For these ubiquitin variants, the core mutations, selected based on Rosetta calculations and phage display, act like the heating step. These mutations disrupt the natural dynamics and packing of the WT protein, creating a large and diverse ensemble of states. Relative to the initial WT conformation, these dynamics

likely increase the sampling of states that have the desired function; however, many undesirable conformations are also sampled. Further affinity maturation acts like the cooling step in simulated annealing, selectively stabilizing the functional states. This final pattern of flexibility changes is similar to what has been hypothesized for designed enzymes subjected to directed evolution (Bhabha et al., 2015; Kiss et al., 2013) and observed in antibody maturation (Adhikary et al., 2015; Jimenez et al., 2004). The discordant mapping between the conformational landscape, which is most heterogeneous in the “core” mutant, and functional landscape, which shows the largest functional gain from the initial “core” phage display, demonstrates the complex interplay between conformational dynamics and function. Although our observations are made in the context of the evolution of new binding specificity, they can likely be translated to the more complicated challenge of evolving new catalytic activities (Campbell et al., 2016).

The ability for proteins to evolve new functions in this way relies on the fact that the hydrophobic cores of most proteins can accommodate many alternative sequences without compromising stability (Lim and Sauer, 1989). This permissiveness can be exploited to alter functional specificity (Koulechova et al., 2015). Here, we observed how changes in side-chain packing





**Figure 5. Reduced Conformational Heterogeneity from Core Mutant to Affinity-Matured Mutant**

(A–C) The packing of residues 42, 49, and 72 is shown. Panels and colors show the WT, core, and affinity-matured mutants in gray, blue, and green, respectively. Once mutated, residues are shown in a lighter shade of the same color. Residue Arg72 could not be fully built in the core mutant model and thus is truncated at the C $\beta$ . Residues mutated in the affinity-matured protein now show new cation- $\pi$  interactions, both between residues 72 and 42 and between residues 42 and 49. (D) Residues 36–39 are shown, highlighting heterogeneity in the affinity-matured mutant that spans residues 32–41. There is signal for two conformations that differ in this region by a shift of as much as 2.2 Å.  $2F_o - F_c$  map shown as a volume contoured to 3.5, 1.5, and 0.65 eÅ<sup>-3</sup> (light blue, blue, and black), with the difference  $F_o - F_c$  map contoured to 3 eÅ<sup>-3</sup>.

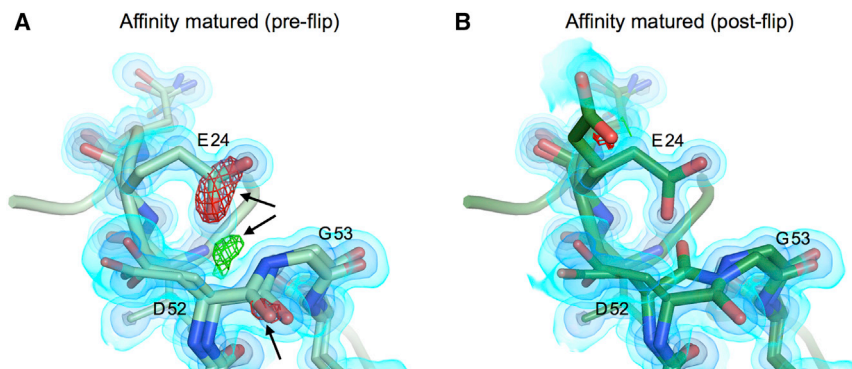
(E) The heterogeneity seen in the core mutant in (D) is not seen for the affinity-matured structure at the same residues. Maps are contoured as in (D). See also Figure S4.

between mutations in direct contact can lead to dramatic changes in backbone flexibility. Therefore, core mutations, even with multiple compensatory mutations, can disrupt the natural dynamic packing and may increase backbone flexibility. These changes lead to altered dynamics, which can be exploited for evolving new functions. The nearly global quenching of backbone heterogeneity in the affinity-matured mutant, both in regions directly adjacent to mutations and across the protein, point to the importance of the cooperativity in these dynamic packing interactions.

Surprisingly, the reduced dynamics in the affinity-matured variant are enabled by introduction of new surface interactions, including a cation- $\pi$ -cation interaction observed between residues Arg72, Trp42, and Arg49. These new side chains apparently cooperate to “freeze” motion on the backbone behind these new interactions. The importance of this cation- $\pi$ -cation motif in this structure suggests possibilities to improve rational design of new or altered protein function by incorporating rarer interactions in the design process. Despite efforts to develop potentials for cation- $\pi$  interactions in Rosetta (Misura et al., 2004), they are not currently modeled in the standard Rosetta energy function and thus would be completely missed in the design process. Cation- $\pi$  interactions have recently been used to stabilize a miniature designed protein (Craven et al., 2016) and are com-

mon in naturally occurring proteins (Dougherty, 2013). In addition, ordered waters, which are not directly accounted for by Rosetta (Jiang et al., 2005), play a key role in the stabilization of hydrogen bond networks such as between the WT His68 and the  $\beta$ 1 $\beta$ 2 loop. These overlooked features of protein structure play key roles in conformational stabilization and could be incorporated to improve the protein design process.

Interestingly, despite the dramatic increase in affinity for the binding partner USP7, the  $\beta$ 1 $\beta$ 2 loop in our apo-crystal structures does not adopt the “down” conformation seen in the WT ubiquitin USP7 complex. Instead, the affinity-matured loop state resembles the WT apo “up” state (Figure 3D). It is unclear whether the “down” conformation is stabilized by these mutations, albeit still as a minor conformation that cannot be detected crystallographically. Rather than disrupting the dynamics of the  $\beta$ 1 $\beta$ 2 loop in such a way that the “down” state is the only state accessed, one explanation for the enhanced affinity to USP7 may be that the mutations introduced key interfacial residues that produce a more complementary surface to the large binding cleft of USP7. Alternatively, the mutations may reduce the energetic penalty for reorganizing into the bound conformation in an induced fit mechanism. Repeated attempts to cocrystallize mutant ubiquitins with USP7 have not been successful. However, when our model is docked into holo-USP7 crystal



**Figure 6. Asp52-Gly53 Peptide Flip Is Seen in Both States**

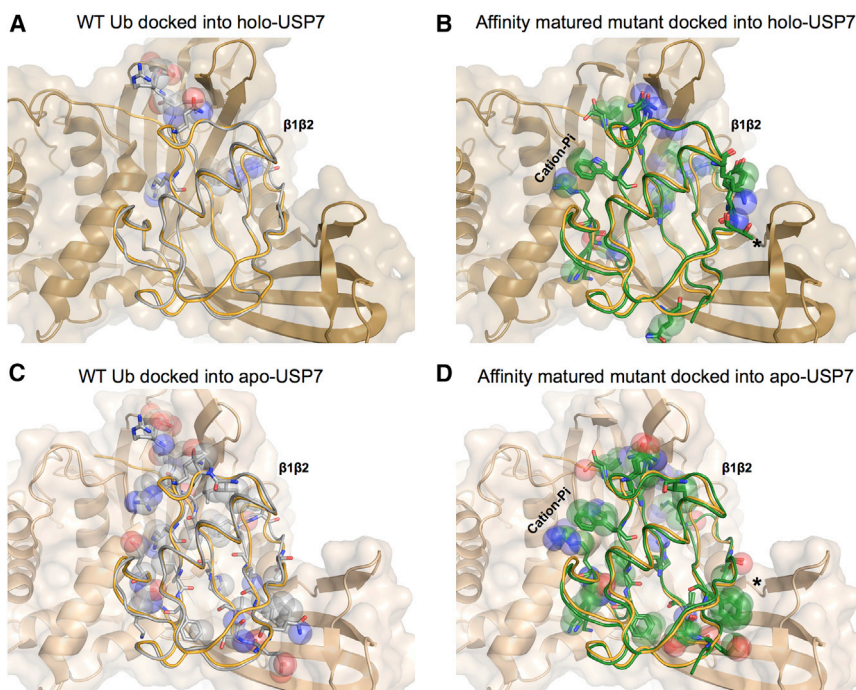
(A) Model and maps from the affinity-matured mutation prior to modeling a peptide flip in this region.  $2F_o - F_c$  map shown as a volume contoured to 3.5, 1.5, and  $0.65 \text{ eÅ}^{-3}$  (light blue, blue, and black), with the difference  $F_o - F_c$  map contoured to  $3 \text{ eÅ}^{-3}$ . There are clear difference features both positive (green) and negative (red), highlighted by black arrows.

(B) Modeled peptide flip in final structure of the core mutant. Maps are contoured as in (A). The difference features are now gone.

structures it has significant clashes, similar to docking a WT apo-ubiquitin crystal structure into this pocket (Figure 7). Interestingly, the pattern of clashes changes, with new clashes appearing near the  $\beta 1\beta 2$  loop (residues Tyr13 and Thr14) and for residues Lys33, 47–49, and Gln62. Also, the residues involved in the newly introduced cation- $\pi$  interaction show clashes when docked into USP7. The pattern of clashes also changes when docked into an apo crystal structure of USP7 without a ubiquitin in the binding site. Most notably there is a shift of clashes in the region of the  $\beta 1$  and  $\beta 2$  strands, shifting from residues Tyr13 and Thr14 to Thr12, Phe4, and the side chain of Glu64 (Figure 7), which is likely a result of the clam-shell shift of USP7 when binding to its partner. Collectively, these clashes argue against the ubiquitin mutants targeting the apo-like state of USP7. We speculate that the affinity-matured mutant also un-

dergoes rearrangement with USP7 upon binding, similar to what is seen based on WT crystal structures. Zhang et al. (2013) showed that while the specificity for the affinity-matured mutant was mostly specific to just USP7, the additional deletion of the two C-terminal glycine residues was necessary to obtain specificity over USP5. The C terminus makes extensive contacts with both USP7 and USP5, suggesting that the specificity is reliant on the folded part of the domain where new contacts enhance affinity for USP7 over USP5.

Future studies of the core and affinity-matured mutants bound to USP7 are needed to further answer these questions regarding the mechanism whereby these ubiquitin mutants obtained such high affinity for USP7. The relationship between protein function and heterogeneity and dynamics is complex; the creation of new alternative conformations does not necessarily imply reduced



**Figure 7. Wild-Type and Affinity-Matured Ubiquitin Have Distinct Patterns of Contacts to Apo and Holo USP7**

(A) Unbound WT Ub (gray, 1ubq) shown in ribbon overlaid with the bound Ub from the holo USP7 structure (orange, 1nbf). Ub-bound USP7 is shown in cartoon and surface (tan, 1nbf). Clashing atoms between the unbound WT Ub and USP7 are mainly concentrated at the  $\beta 1\beta 2$  loop (top) and shown as spheres, with the remainder of the residue shown as sticks.

(B) Affinity-enhanced mutant (green) shown overlaid with bound Ub from the holo USP7 structure. Although contacts are changed at the  $\beta 1\beta 2$  loop (top), additional clashes, indicative of an altered binding mode or receptor accommodation, are spread throughout the protein. An asterisk marks notable changes in clashes between apo and holo USP7 structures.

(C) Unbound WT Ub (gray, 1ubq) shown overlaid with the unbound apo-USP7 structure (lighter brown, PDB: 5j7t). Overlay was constructed by alignment of domains to the bound USP7 (1nbf). Increased clashes throughout the protein show the accommodation of the receptor in the holo form.

(D) Affinity-matured mutant (green) shown overlaid with the unbound apo-USP7 structure (5j7t). Although clashes are increased in some regions, be exploited by the affinity-matured mutant in

they are reduced in others, which suggests that the conformational flexibility of the receptor may be in the final binding pose. An asterisk marks notable changes in clashes between apo and holo USP7 structures.

affinity or performance. A protein's WT sequence is presumably nearly optimized for fitness and its conformational landscape is sufficiently tuned for function. Mutations can therefore introduce extra conformers, which might have a negative impact on fitness; however, these extra conformations can, in rare instances, provide the raw material for functional innovation or specialization. Directed evolution can stabilize the extra states introduced by the mutations selected through protein design, quenching heterogeneity. Given the recent interest in using ubiquitin variants for structural biology chaperones as in vitro modulators of the ubiquitin proteasome system (Canny et al., 2016; Ernst et al., 2013; Gorelik et al., 2016; Phillips et al., 2013; Zhang et al., 2016), there may be additional opportunities to use these proteins to test these hypotheses and learn about the importance of conformational dynamics in protein function (Phillips et al., 2013). Experimentally characterizing nearly iso-energetic states will improve our ability to evaluate the success of the design of protein ensembles. As design challenges move from stabilizing a single state to creating functional cycles, finding sequences that predictably modulate flexibility or rigidity will become increasingly important.

## STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - Protein Purification
  - Crystallization and data Collection
  - Optimizing the diffraction Resolution Cutoff
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

## SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and can be found with this article online at <http://dx.doi.org/10.1016/j.str.2017.03.009>.

## AUTHOR CONTRIBUTIONS

J.T.B. performed experiments; J.T.B. and M.C.T. performed refinements; J.T.B., C.N.C., J.E.C., and J.S.F. designed experiments; J.T.B. and J.S.F. wrote the manuscript; J.T.B., M.C.T., C.N.C., J.E.C., and J.S.F. edited the manuscript.

## ACKNOWLEDGMENTS

J.T.B. is supported by an NSF GRFP. M.C.T. is supported by NIH F32 HL129989. J.S.F. is supported by a Searle Scholar Award from the Kinship Foundation, a Pew Scholar Award from the Pew Charitable Trusts, a Packard Fellowship from the David and Lucile Packard Foundation, NIH DP5 OD009180, NIH R21 GM110580, and NSF STC-1231306. Data collection at BL831 at the Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231, UC Office of the President, Multi-campus Research Programs and Initiatives grant MR-15-32859, and the Program Breakthrough Biomedical Research, which is partially funded by the Sandler Foundation. We acknowledge helpful conversations with H. van den Bedem, T. Kortemme, N. Bethel, and members of the J.S.F. laboratory, espe-

cially D. Keedy. C.N.C. is a current employee of Genentech, Inc. and declares a competing financial interest.

Received: October 17, 2016

Revised: January 20, 2017

Accepted: March 13, 2017

Published: April 13, 2017

## REFERENCES

- Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* 58, 1948–1954.
- Adhikary, R., Yu, W., Oda, M., Walker, R.C., Chen, T., Stanfield, R.L., Wilson, I.A., Zimmermann, J., and Romesberg, F.E. (2015). Adaptive mutations alter antibody structure and dynamics during affinity maturation. *Biochemistry* 54, 2085–2093.
- Bhabha, G., Biel, J.T., and Fraser, J.S. (2015). Keep on moving: discovering and perturbing the conformational dynamics of enzymes. *Acc. Chem. Res.* 48, 423–430.
- Blomberg, R., Kries, H., Pinkas, D.M., Mittl, P.R.E., Grütter, M.G., Privett, H.K., Mayo, S.L., and Hilvert, D. (2013). Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature* 503, 418–421.
- Brünger, A.T., Adams, P.D., and Rice, L.M. (1997). New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure* 5, 325–336.
- Burnley, B.T., Afonine, P.V., Adams, P.D., and Gros, P. (2012). Modelling dynamics in protein crystal structures by ensemble refinement. *Elife* 1, e00311.
- Butterfoss, G.L., and Kuhlman, B. (2006). Computer-based design of novel protein structures. *Annu. Rev. Biophys. Biomol. Struct.* 35, 49–65.
- Campbell, E., Kaltenbach, M., Correy, G.J., Carr, P.D., Porebski, B.T., Livingstone, E.K., Afriat-Jurnou, L., Buckle, A.M., Weik, M., Hoffelder, F., et al. (2016). The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* 12, 944–950.
- Canny, M., Wan, L., Fradet-Turcotte, A., Orthwein, A., Moatti, N., Juang, Y.-C., Zhang, W., Noordermeer, S.M., Wilson, M.D., Vorobyov, A., et al. (2016). A genetically encoded inhibitor of 53BP1 to 1 stimulate homology-based gene editing. *bioRxiv*. <http://dx.doi.org/10.1101/060954>.
- Craven, T.W., Cho, M.-K., Traaseth, N.J., Bonneau, R., and Kirshenbaum, K. (2016). A miniature protein stabilized by a cation- $\pi$  interaction network. *J. Am. Chem. Soc.* 138, 1543–1550.
- Daopin, S., Davies, D.R., Schlunegger, M.P., and Grütter, M.G. (1994). Comparison of two crystal structures of TGF- $\beta$ 2: the accuracy of refined protein structures. *Acta Crystallogr. D Biol. Crystallogr.* 50, 85–92.
- Dougherty, D.A. (2013). The cation- $\pi$  interaction. *Acc. Chem. Res.* 46, 885–893.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.
- Ernst, A., Avvakumov, G., Tong, J., Fan, Y., Zhao, Y., Alberts, P., Persaud, A., Walker, J.R., Neculai, A.-M., Neculai, D., et al. (2013). A strategy for modulation of enzymes in the ubiquitin system. *Science* 339, 590–595.
- Fields, B.A., Bartsch, H.H., Bartunik, H.D., Cordes, F., Guss, J.M., and Freeman, H.C. (1994). Accuracy and precision in protein crystal structure analysis: two independent refinements of the structure of poplar plastocyanin at 173 K. *Acta Crystallogr. D Biol. Crystallogr.* 50, 709–730.
- Fraser, J.S., Clarkson, M.W., Degnan, S.C., Erion, R., Kern, D., and Alber, T. (2009). Hidden alternative structures of proline isomerase essential for catalysis. *Nature* 462, 669–673.
- Fraser, J.S., van den Bedem, H., Samelson, A.J., Lang, P.T., Holton, J.M., Echols, N., and Alber, T. (2011). Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc. Natl. Acad. Sci. USA* 108, 16247–16252.



- Frushicheva, M.P., Cao, J., Chu, Z.T., and Warshel, A. (2010). Exploring challenges in rational enzyme design by simulating the catalysis in artificial Kemp eliminase. *Proc. Natl. Acad. Sci. USA* **107**, 16869–16874.
- Frushicheva, M.P., Mills, M.J.L., Schopf, P., Singh, M.K., Prasad, R.B., and Warshel, A. (2014). Computer aided enzyme design and catalytic concepts. *Curr. Opin. Chem. Biol.* **21**, 56–62.
- Gorelik, M., Orlicky, S., Sartori, M.A., Tang, X., Marcon, E., Kurinov, I., Greenblatt, J.F., Tyers, M., Moffat, J., Sicheri, F., and Sidhu, S.S. (2016). Inhibition of SCF ubiquitin ligases by engineered ubiquitin variants that target the Cul1 binding site on the Skp1-F-box interface. *Proc. Natl. Acad. Sci. USA* **113**, 3527–3532.
- Huang, K.Y., Amodeo, G.A., Tong, L., and McDermott, A. (2011). The structure of human ubiquitin in 2-methyl-2,4-pentenediol: a new conformational switch. *Protein Sci.* **20**, 630–639.
- Husnjak, K., and Dikic, I. (2012). Ubiquitin-binding proteins: decoders of ubiquitin-mediated cellular functions. *Annu. Rev. Biochem.* **81**, 291–322.
- Jiang, L., Kuhlman, B., Kortemme, T., and Baker, D. (2005). A “solvated rotamer” approach to modeling water-mediated hydrogen bonds at protein-protein interfaces. *Proteins* **58**, 893–904.
- Jimenez, R., Salazar, G., Yin, J., Joo, T., and Romesberg, F.E. (2004). Protein dynamics and the immunological evolution of molecular recognition. *Proc. Natl. Acad. Sci. USA* **101**, 3803–3808.
- Johnson, E.C., and Handel, T.M. (1999). Effect of hydrophobic core packing on sidechain dynamics. *J. Biomol. NMR* **15**, 135–143.
- Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132.
- Karanicolas, J., Corn, J.E., Chen, I., Joachimiak, L.A., Dym, O., Peck, S.H., Albeck, S., Unger, T., Hu, W., Liu, G., et al. (2011). A de novo protein binding pair by computational design and directed evolution. *Mol. Cell* **42**, 250–260.
- Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. *Science* **336**, 1030–1033.
- Keedy, D.A., Fraser, J.S., and van den Bedem, H. (2015). Exposing hidden alternative backbone conformations in X-ray crystallography using qFit. *PLoS Comput. Biol.* **11**, e1004507.
- Khersonsky, O., Röthlisberger, D., Dym, O., Albeck, S., Jackson, C.J., Baker, D., and Tawfik, D.S. (2010). Evolutionary optimization of computationally designed enzymes: kemp eliminases of the KE07 series. *J. Mol. Biol.* **396**, 1025–1042.
- Kiss, G., Çelebi-Ölçüm, N., Moretti, R., Baker, D., and Houk, K.N. (2013). Computational enzyme design. *Angew. Chem. Int. Ed.* **52**, 5700–5725.
- Korendovych, I.V., and DeGrado, W.F. (2014). Catalytic efficiency of designed catalytic proteins. *Curr. Opin. Struct. Biol.* **27**, 113–121.
- Korostelev, A., Bertram, R., and Chapman, M.S. (2002). Simulated-annealing real-space refinement as a tool in model building. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 761–767.
- Koulechova, D.A., Tripp, K.W., Horner, G., and Marqusee, S. (2015). When the scaffold cannot be ignored: the role of the hydrophobic core in ligand binding and specificity. *J. Mol. Biol.* **427**, 3316–3326.
- Lang, P.T., Ng, H.-L., Fraser, J.S., Corn, J.E., Echols, N., Sales, M., Holton, J.M., and Alber, T. (2010). Automated electron-density sampling reveals widespread conformational polymorphism in proteins. *Protein Sci.* **19**, 1420–1431.
- Lang, P.T., Holton, J.M., Fraser, J.S., and Alber, T. (2014). Protein structural ensembles are revealed by redefining X-ray electron density noise. *Proc. Natl. Acad. Sci. USA* **111**, 237–242.
- Lange, O.F., Lakomek, N.-A., Farès, C., Schröder, G.F., Walter, K.F.A., Becker, S., Meiler, J., Grubmüller, H., Griesinger, C., and de Groot, B.L. (2008). Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science* **320**, 1471–1475.
- Lim, W.A., and Sauer, R.T. (1989). Alternative packing arrangements in the hydrophobic core of lambda repressor. *Nature* **339**, 31–36.
- Majumdar, A., and Ghose, R. (2004). Probing slow backbone dynamics in proteins using TROSY-based experiments to detect cross-correlated time-modulation of isotropic chemical shifts. *J. Biomol. NMR* **28**, 213–227.
- Massi, F., Grey, M.J., and Palmer, A.G., 3rd (2005). Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R1rho relaxation experiments. *Protein Sci.* **14**, 735–742.
- Misura, K.M.S., Morozov, A.V., and Baker, D. (2004). Analysis of anisotropic side-chain packing in proteins and application to high-resolution structure prediction. *J. Mol. Biol.* **342**, 651–664.
- Osuna, S., Jiménez-Osés, G., Noey, E.L., and Houk, K.N. (2015). Molecular dynamics explorations of active site structure in designed and evolved enzymes. *Acc. Chem. Res.* **48**, 1080–1089.
- Phillips, A.H., Zhang, Y., Cunningham, C.N., Zhou, L., Forrest, W.F., Liu, P.S., Steffek, M., Lee, J., Tam, C., Helgason, E., et al. (2013). Conformational dynamics control ubiquitin-deubiquitinase interactions and influence in vivo signaling. *Proc. Natl. Acad. Sci. USA* **110**, 11379–11384.
- Richardson, J.S., Prisant, M.G., and Richardson, D.C. (2013). Crystallographic model validation: from diagnosis to healing. *Curr. Opin. Struct. Biol.* **23**, 707–714.
- Sidhu, A., Suroli, A., Robertson, A.D., and Sundt, M. (2011). A hydrogen bond regulates slow motions in ubiquitin by modulating a  $\beta$ -turn flip. *J. Mol. Biol.* **411**, 1037–1048.
- Smith, C.A., Ban, D., Pratihar, S., Giller, K., Paulat, M., Becker, S., Griesinger, C., Lee, D., and de Groot, B.L. (2016). Allosteric switch regulates protein-protein binding through collective motion. *Proc. Natl. Acad. Sci. USA* **113**, 3269–3274.
- Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Adams, P.D., Moriarty, N.W., Zwart, P., Read, R.J., Turk, D., and Hung, L.W. (2007). Interpretation of ensembles created by multiple iterative rebuilding of macromolecular models. *Acta Crystallogr. D Biol. Crystallogr.* **63**, 597–610.
- Tinberg, C.E., Khare, S.D., Dou, J., Doyle, L., Nelson, J.W., Schena, A., Jankowski, W., Kalodimos, C.G., Johnsson, K., Stoddard, B.L., and Baker, D. (2013). Computational design of ligand-binding proteins with high affinity and selectivity. *Nature* **501**, 212–216.
- Tollinger, M., Sivertsen, A.C., Meier, B.H., Ernst, M., and Schanda, P. (2012). Site-resolved measurement of microsecond-to-millisecond conformational-exchange processes in proteins by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **134**, 14800–14807.
- van den Bedem, H., Dhanik, A., Latombe, J.C., and Deacon, A.M. (2009). Modeling discrete heterogeneity in X-ray diffraction data by fitting multi-conformers. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1107–1117.
- van den Bedem, H., Bhabha, G., Yang, K., Wright, P.E., and Fraser, J.S. (2013). Automated identification of functional dynamic contact networks from X-ray crystallography. *Nat. Methods* **10**, 896–902.
- Walsh, S.T., Lee, A.L., DeGrado, W.F., and Wand, A.J. (2001). Dynamics of a de novo designed three-helix bundle protein studied by  $^{15}\text{N}$ ,  $^{13}\text{C}$ , and  $^2\text{H}$  NMR relaxation methods. *Biochemistry* **40**, 9560–9569.
- Wijma, H.J., Floor, R.J., Bjelic, S., Marrink, S.J., Baker, D., and Janssen, D.B. (2015). Enantioselective enzymes by computational design and in silico screening. *Angew. Chem. Int. Ed.* **54**, 3726–3730.
- Woldeyes, R.A., Sivak, D.A., and Fraser, J.S. (2014). E pluribus unum, no more: from one crystal, many conformations. *Curr. Opin. Struct. Biol.* **28**, 56–62.
- Zhang, Y., Zhou, L., Rouge, L., Phillips, A.H., Lam, C., Liu, P., Sandoval, W., Helgason, E., Murray, J.M., Wertz, I.E., and Corn, J.E. (2013). Conformational stabilization of ubiquitin yields potent and selective inhibitors of USP7. *Nat. Chem. Biol.* **9**, 51–58.
- Zhang, W., Wu, K.-P., Sartori, M.A., Kamadurai, H.B., Ordureau, A., Jiang, C., Mercredi, P.Y., Murchie, R., Hu, J., Persaud, A., et al. (2016). System-wide modulation of HECT E3 ligases with selective ubiquitin variant probes. *Mol. Cell* **62**, 121–136.



## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
LB broth Miller Granulated	Fisher Scientific	CAT#BP97232
Ampicillin	Sigma-Aldrich	CAT#A9518 CAS: 69-53-4
IPTG	Goldbio	CAT#I2481C100 CAS: 367-93-1
Tris	VWR	CAT#71003-494 CAS: 77-86-1
NaCl	Fisher Scientific	CAT#BP358 CAS: 7647-14-5
Protease inhibitors	Sigma-Aldrich	CAT#11836170001
Imidazole	Acros Organics	CAT#122020020 CAS: 288-32-4
TEV	This paper	
Citric acid	Fisher Scientific	CAT#A104 CAS: 5949-29-1
Ammonium sulfate		CAT#A44118 CAS: 7783-20-2
Critical Commercial Assays		
BSA	Pierce	CAT#23227
Deposited Data		
Core mutant (u7ub25)	This paper	PDB: 5t0f SBGrid Dataset number 428 ( <a href="http://dx.doi.org/10.15785/SBGRID/428">http://dx.doi.org/10.15785/SBGRID/428</a> )
Affinity matured mutant (u7ub25_254)	This paper	PDB: 5t0g SBGrid Dataset number 429 ( <a href="http://dx.doi.org/10.15785/SBGRID/429">http://dx.doi.org/10.15785/SBGRID/429</a> )
Experimental Models: Cell Lines		
BL21(DE3)	New England Biolabs	CAT#:C2527H
Recombinant DNA		
Plasmid: pET vector	This paper	
Ubiquitin mutant genes	This paper	
Software and Algorithms		
XDS	Kabsch, 2010	<a href="http://xds.mpimf-heidelberg.mpg.de/">http://xds.mpimf-heidelberg.mpg.de/</a>
Phenix	Adams et al., 2002	<a href="https://www.phenix-online.org/download/">https://www.phenix-online.org/download/</a>
COOT	Emsley et al., 2010	In Phenix suite
PyMOL	The PyMOL Molecular Graphics System, Schrödinger, LLC.	<a href="https://www.pymol.org/">https://www.pymol.org/</a>
qFit 2.0	Keedy et al., 2015	<a href="https://simtk.org/home/qfit">https://simtk.org/home/qfit</a>
Other		
EmulsiFlex-C3	Avestin	<a href="http://www.avestn.com/English/c3page.html">http://www.avestn.com/English/c3page.html</a>
Ni-NTA column (HisTrap FF)	GE	CAT#11-0004-58
10 kDa concentrator	Millipore	CAT#UFC801024
30 kDa dialysis tubing	Fisher Scientific	CAT#68100
S75 column	GE	CAT#28989333
Nano Drop	Thermo Scientific	<a href="http://www.nanodrop.com/Productnd2000overview.aspx">http://www.nanodrop.com/Productnd2000overview.aspx</a>
Hanging drop crystal tray	Qiagen	CAT#132006

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, James Fraser ([jfraser@fraserlab.com](mailto:jfraser@fraserlab.com)).

## METHOD DETAILS

### Protein Purification

Protocol adapted from Zhang et al. ((Canny et al., 2016; Ernst et al., 2013; Gorelik et al., 2016; Phillips et al., 2013; Zhang et al., 2016)) and references therein. Each of the ubiquitin mutants were expressed from a pET derivative vector containing the protein gene with a TEV cleavable N-terminal His6 tag in *E. coli* BL21(DE3) cells. Cells were grown in LB media to an OD600 of 0.6-0.8, and were then induced with 1 mM IPTG at 37 °C for 4 hours. Cell pellets were resuspended into 25 mM Tris pH 8.0, 300 mM NaCl and protease inhibitors. The resuspended pellet was then lysed by EmuSiFlex. Lysate was cleared by centrifugation at 15,000 rpm. Supernatant was flowed over a 5 mL Ni-NTA column. Ni-NTA column was washed with 20 mM Tris pH 7.0, 300 mM NaCl, 20 mM imidazole and eluted with 20 mM Tris pH 7.0, 300 mM NaCl, 500 mM imidazole. Elution was dialyzed into 20 mM Tris pH 8.0, 150 mM NaCl and then cleaved with TEV overnight at 4 °C. The sample was then loaded back onto a Ni-NTA column, and the flow through was collected. Sample was then concentrated to less than 10 mL and loaded onto a S75 column equilibrated with the previous buffer (20 mM Tris pH 8.0, 150 mM NaCl). Protein concentration of the pooled fractions was assessed after concentration down to 1-2 ml via a BSA assay and monitored during further concentration via absorption. Protein was stored only overnight at 4 °C before concentration and use for crystallization.

### Crystallization and data Collection

Protein solution from was concentrated to 20 mg/ml for the core mutant and 10 mg/ml for the affinity matured mutant. Then 1  $\mu$ l of protein solution was mixed with 1  $\mu$ l of precipitation solution in hanging drop trays. Precipitation solution for core mutant contained 0.1 M citric acid pH 4.6, and 2.6 M ammonium sulfate. For the affinity matured mutant, the precipitation solution contained 0.1 M citric acid pH 4.2, and 2.2 M ammonium sulfate. Crystals formed overnight and grew to full size over the course of several days. Diffraction data were collected at the Advanced Light Source (ALS, Berkeley, CA), beamline 8.3.1.

### Optimizing the diffraction Resolution Cutoff

After optimizing crystals for the room temperature collection of high resolution datasets for the initial core and the final affinity matured mutants, we had to select an appropriate resolution cutoff for our data. Because the ability to detect and faithfully model conformational heterogeneity is resolution dependent, we aimed to push the resolution of our dataset as far as possible. High resolution reflections, despite their lower signal-to-noise, can contain meaningful information that can improve the map and model, even past more traditional methods of picking a resolution cutoff. Instead of simply using CC1/2, completeness and  $I/\sigma$ , we used the Karplus and Diederichs approach of monitoring Rfree in parallel refinements to determine the optimal resolution cutoff for our datasets (Karplus and Diederichs, 2012). When using parallel refinements to determine the optimal resolution cutoff, additional higher resolution reflections are judged to contain meaningful signals only if model agreement increases in lower resolution bins after refinement.

A molecular replacement solution was found with the 1ubq WT ubiquitin model. While, a previous lower resolution structure of the affinity enhanced mutant exists (PDB: 4hk2) this model was solved with a much larger unit cell containing extra non-crystallographic symmetry copies that were not justified by the diffraction images in our datasets. Due to these complications, we have not focused on comparisons with 4hk2 and restricted our analysis to the (high resolution) room temperature dataset.

For an initial conservative resolution cutoff, we chose 1.16 Å as a cutoff to begin our parallel refinement tests, and created additional bins of reflections to be added in 0.04 Å increments up to a high resolution cutoff of 0.96 Å. While a 0.04 Å resolution change may seem small, ~2000 unique reflections are added in each bin. The values for CC1/2 in the high resolution bins remained over 50% up to 1.04 Å resolution, while completeness began to drop at resolutions better than 1.12 Å (Table 1). To begin the resolution test, first we built a single conformer model using the reflections up to the first cutoff (1.16 Å). We then built alternative conformations into strong difference density signals corresponding to clear alternative conformations. At this point, we added reflections from each additional resolution bin ranging from 1.16 Å to 0.96 Å, and re-refined the model to convergence in parallel refinements. Based on the R-values from these parallel refinements, the determined optimal resolution cutoffs were 1.12 Å for the core mutant and 1.08 Å for the affinity matured mutant (Tables 1 and 2, Figure S5).

## QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical parameters are reported in Figure Legends and in Method Details.

## DATA AND SOFTWARE AVAILABILITY

XDS was used to process the raw diffraction data and different software found in the phenix suite was then used to determine, refine, and build the structural model. All software used are reported in [Method Details](#) and in the [Key Resources Table](#). The accession number for the coordinates and structure factors for the core mutant (u7ub25) is PDB: 5t0f, and for the affinity matures mutant (u7ub25\_2540) is PDB: 5t0g.