

What strengthens Protein-Protein Interactions: Analysis and Applications of Residue Correlation Networks

Ta I Hung^{1,2†}, Yun-Jung Hsieh^{3,4†}, Wei-Lin Lu³, Kuen-Phon Wu^{3,4*}, and Chia-en A. Chang^{1*}

1. Department of Chemistry, University of California, Riverside, United States
2. Department of Bioengineering, University of California, Riverside, United States
3. Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan
4. Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan

† These authors contributed equally to this work

* Correspondences

Kuen-Phon Wu: kpwu@gate.sinica.edu.tw

Chie-en A. Chang: chiaenc@ucr.edu

Keywords: Dihedral angle correlation network, protein-based inhibitors, ubiquitin variant, papain-like protease, nanomolar binder

Abstract

Identifying residues critical to protein-protein binding and efficient design of stable and specific protein binders are challenging tasks. Extending beyond the direct contacts in a protein-protein binding interface, our study employs computational modeling to reveal the essential network of residue interactions and dihedral angle correlations critical in protein-protein recognition. We hypothesized that mutating residues exhibiting highly correlated dynamic motion within the interaction network could efficiently optimize protein-protein interactions to create tight and selective protein binders. We tested this hypothesis using the ubiquitin (Ub) and MERS coronavirus papain-like protease (PLpro) complex, since Ub is a central player in multiple cellular functions and PLpro is an antiviral drug target. Our designed ubiquitin variant (UbV) hosting three mutated residues displayed a ~3,500-fold increase in functional inhibition relative to wild-type Ub. Further optimization of two C-terminal residues within the Ub network resulted in a K_D of 1.5 nM and IC_{50} of 9.7 nM for the five-point Ub mutant, eliciting 27,500-fold and 5,500-fold enhancements in affinity and potency, respectively, as well as improved selectivity, without destabilizing the UbV structure. Our study highlights residue correlation and interaction networks in protein-protein interactions, and introduces an effective approach to design high-affinity protein binders for cell biology research and future therapeutics.

Introduction

Understanding what drives protein-protein binding and selecting appropriate protein residues for modification to strengthen protein-protein interactions (PPIs) are crucial to designing a protein binder that targets its binding partner [1,2]. Strategies that can efficiently and accurately identify residues to enhance PPIs have broad applications in therapeutics and studies of cell biology. Knowledge-based, physics-based, and data-driven methods have all been developed previously to explore PPIs and to select mutations that enhance them [3–9]. Computational and combinatorial libraries or *in vitro* evolutionary approaches also represent popular protein engineering strategies to design stable and specific protein binders [10–12]. Importantly, minimizing the number of residues mutated to significantly enhance PPIs lowers the possibility of engineering unstable proteins. Nevertheless, protein engineering remains challenging, as highly integrated molecular modeling and experimental techniques are needed to understand PPIs in order to re-engineer a protein to increase its binding affinity.

PPI networks are extremely complex, so selecting an appropriate target system for experimental modification requires specialized expertise. Here, we chose Ubiquitin (Ub) as our target system, as it plays critical roles in numerous biological functions [13]. Ub is a small 76-residue protein associated with post-translational modifications. This regulatory protein canonically binds to its cascade E1-E2-E3 enzymes to drive ubiquitination and Ub chain formation, thereby modifying nearly half of the human proteome [14,15]. Conversely, deubiquitinases (DUBs) [16] cleave the covalent isopeptide bonds from Ub chains or substrates to release Ub and its substrates. The interactome of Ub and cellular proteins have been assessed [15] revealing that precise Ub network regulation governs cellular fates. Misregulation of the responsible enzymes significantly impacts cellular functions, leading to diseases such as cancers [17]. Moreover, viral DUBs have been found to interfere with host antiviral defenses. For example, the Papain-like protease (PLpro) of coronaviruses (CoVs) is classified as a viral DUB specific to Ub and Ub-like ISG15 [18,19]. Previous studies have shown that PLpro alters host innate immune responses, which contributes to the rapid spread of CoVs (such as MERS, SARS-CoV, SARS-CoV-2) [20–24], thus causing pandemics, mortality and perturbing the global economy [25].

The PLpro proteins of MERS and SARS-CoV-2 are crucial for viral replication through their role in proteolytic cleavage of viral nonstructural proteins (NSPs). The PLpro domain resides in NSP3, which drives viral genome replication and subgenomic RNA synthesis [26,27]. PLpro recognizes and cleaves the NSP1-2, NSP2-3 and NSP3-4 junctions after the amino acid sequence LXGG to yield functional viral proteins, as well as to perform deubiquitination and deISGylation [23,24,28]. Deubiquitination and deISGylation alter host signaling pathways critical to induction of cellular antiviral and pro-inflammatory innate immune responses, ultimately suppressing the host antiviral response [28,29]. Therefore, inhibiting PLpro simultaneously disrupts viral replication and prevents PLpro from impairing the innate immune response. Given both these properties, PLpro represents an ideal antiviral drug target.

Importantly, wild-type Ub (wtUb) exhibits high thermostability ($T_m > 90^\circ\text{C}$), so it is an ideal template for protein design. Re-engineered Ub also has potential advantages, such as enhanced binding specificity to PLpro and easier synthesis compared to chemical compounds. Screening of phage-displayed Ub variants (UbVs) against cognate enzymes, including MERS PLpro, has previously demonstrated the feasibility of regulating the

activities of E3 ligases and DUBs [30–32]. The phage-display screening technique focused on three surface patches in Ub to iteratively mutate and select variants that displayed tight binding. The resulting DUB UbVs proved to be strong inhibitors, exhibiting IC_{50} values in the range of 1–30 nM [31–33]. As an alternative approach, computational data were used to rationally design a screening library for the identification of tightly binding regulatory UbVs for USP7 [34] and USP21[35]. A combined computational and phage-display screening of UbVs targeting USP7 resulted in an equilibrium constant (K_D) for the U7Ub25.2540 variant of 56 nM, whereas for wtUb-USP7 it was $>200 \mu M$. A pool of 6,000 designed UbVs for USP21 revealed that $\sim 10\%$ of the variants tightly bound USP21 consistently between experimental and computational screenings. However, *in silico* screenings of such large UbV-USP21 pairings requires intensive computational resources. Such expensive and time-consuming empirical screenings impede rational design of protein-based inhibitors.

Here, we present an integrated computational and experimental approach to identify critical regions for protein-protein binding that display highly correlated dynamic motion. Specifically, we focus on side-chain dihedral angle correlations at the protein-protein contact interface where mutation of highly correlated residues resulted in both local and distal conformational changes. We demonstrate that mutating residues in these regions can efficiently optimize PPIs to create tight and selective protein binders. We show that our designed UbVs hosting two or three mutated residues achieved 3,500-fold inhibitory efficiency and binding affinity relative to wtUb for MERS PLpro (Table 1). MERS PLpro cleaves both K48- and K63-linked Ub chains [18,23], and it exhibits distinct inhibitor recognition specificity to that of the PLpro of SARS-CoV and SARS-CoV-2 [36]. We used non-covalent amino acid interaction and side-chain dihedral angle networks of the Ub and MERS PLpro (Ub-PLpro) complex to guide our design of UbVs that enhance UbV–PLpro binding affinity, thereby inhibiting PLpro activity. Initially, we designed two-point mutations for cost efficiency and to retain intact the overall complex structure. Integrating experimental data and computational analyses informed our experimental design to yield more UbVs (Fig. 1). Binding affinity K_D and IC_{50} measurements of our designed UbVs support that more extensively mutated UbV3, UbV4, and UbV5 represent strong inhibitors.

Results and Discussion

The process of optimizing PPIs involves identifying residues suitable for mutation, and then determining substitutions aimed at generating novel protein variants. We used Molecular Dynamics (MD) simulation to reveal residues in protein interfaces displaying strongly correlated dynamic motions. The central hypothesis underlying our rationale is that residues exhibiting a high degree of correlation play a pivotal role within the protein-protein interface. Moving just one protein side chain can influence the motion of other side chains. Thus, mutating highly correlated residues not only can strengthen or weaken local interactions, but can also maximize the impact of changes across all correlated residues in the interface. Mutating these dynamically correlated residues can enhance intermolecular attractions and thereby significantly improve PPIs. We employed a dihedral angle correlation network to identify important protein residues in the PPI interface of Ub and MERS PLpro. This network allowed us to understand how different parts of the proteins move in relation to one another. Then, we selected candidate residues for mutation by means of a pairwise force distribution analysis. For each of our designed UbVs, we computed their binding energy data using molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) and the dissociation time by means of PPI Gaussian-accelerated MD (PPI-GaMD). To further validate our results, we employed a fluorescence polarization (FP)-based inhibition assay to quantify their binding affinities and selectivities, and adopted circular dichroism (CD) spectroscopy to assess protein thermostability.

Identifying key residues in the PPI interface

Mutating highly correlated residues to increase overall binding affinity necessitates investigating localized attraction, structural dynamics, and the biological effects of protein residues. The process of mutant design involved initial adjustments of Ub residues within stable Ub-PLpro complex domains, such as the hydrophobic core for the S1 Ub (HCS1) and alpha helix in the canyon (AHC), before extending modifications to more flexible domains such as the finger-Ub interaction (FBI) and Ub chain cleavage (UCC) domains in the C-terminal tail (Fig. 2A).

First, we performed several 500-ns atomic MD simulations to model full protein flexibility of the wtUb-PLpro complex. We employed two principle techniques to study correlated protein motions, i.e., side-chain dihedral angle correlation (Fig. 2B) and force distribution

analysis (FDA) [37,38] (Fig. S1). Next, we selected highly interactive residues displaying strong correlation in the residue network and also in the Ub–PLpro interface as potential mutation targets in each region (Fig. 2B and C). Notably, previous studies have shown that altering residues near the C-terminus of Ub hampers its biological activity [39–41], so we hypothesized that mutating any residues in the UCC could prevent substrate ubiquitination. MERS PLpro favors binding and cleaving the K48- and K63-linked Ub chains [18,23], which implies that residues near the conjugation points interact frequently with PLpro. Notably, K48 and K63 lie in the HCS1 and FBI domains, respectively. Mutating residues within interactive hotspots may easily hamper PPI, but also offers the potential to strengthen them. Since the FBI and UCC are rather flexible [24], we began by modifying residues A46 and K48 in HCS1 and V70 and R42 in the AHC.

Selection of candidate residues for mutational analysis

Residues A46 and K48 in HCS1 exhibit suboptimal interactions, with weak attraction to surrounding PLpro residues. Specifically, A46 is flanked by Y208 and Y223. To maximize the hydrophobic effect and local nonpolar attraction, we selected a bulkier non-polar amino acid, PHE, for A46 substitution. TYR and TRP were excluded as substitute residues due to spatial constraints. Residue K48 is surrounded by the side-chain of K204. Both K48 and K204 are positively charged, so they are not favored to interact. Accordingly, we selected E48, L48, S48, and I48 as mutation candidates. These substitutions introduce negative, nonpolar, or polar properties, thereby augmenting the potential for enhanced interactions with K204 and Y208 (Fig. 3B). V70 is subjected to weak repulsion from neighboring residues, so the E70 substitution represented a promising mutation due to the ability of its longer and polar side-chain to enhance electrostatic interaction between the G247 backbone and G248 side-chain (Fig. S2).

In the AHC, R42 presents a strong compensatory effect through strong repulsive forces with the D164 side-chain and an attractive force with the S165 backbone. We generated a D42 mutant variant displaying a negative charge and a shorter side-chain with a view to minimizing the repulsive forces that mainly arise from crossing van der Waal (vdW) radii (Fig. S2).

In the FBI, our dihedral angle correlation network selected E64 and S65 as targets in different MD simulation runs because of the inherent flexibility of the FBI. MERS PLpro

unambiguously recognizes K63-linked Ub chains, so we postulated that modifying E64, i.e., between K63 and S65, may simultaneously enhance binding strength and specificity. E64 lies in close proximity to G208 and G209, both of which are nonpolar (Fig. 3C). To maximize nonpolar interactions, we generated 64Y and 64F substitution variants that have large hydrophobic rings.

We targeted R74 and G75 in the UCC for mutation, both of which lie in the LXGG recognition site. Mutating residues of the LXGG recognition site can block the proteolytic cleavage activity of PLpro. [24] We tested several dual mutations proposed previously [33], including R74P/R74N and G75R/G75S. From a structural point of view, steric hindrance within the UCC introduces strong vdW repulsion. Therefore, we assumed that the smaller and polar side-chain of the N74 substitution would minimize this repulsive force while retaining polar attraction. We also selected P74 mutation because it could generate localized steric hindrance to stabilize rotation of the surrounding side-chain. R75 and S75 are both polar residues and have the potential to increase local electrostatic interaction relative to nonpolar G75 (Fig. 3D).

To predict the intermolecular attractions between each of our UbV-PLpro variants, we used MM/PBSA to evaluate PPI energy (Fig. S3) and local structural analysis to investigate localized attraction (Fig. S2). Variants harboring the A46 and K48 mutations displayed enhanced UbV-PLpro attraction. However, the K48E-V70E variant did not yield good vdW interactions, and the R42D mutation that resulted in an opposing charge yielded poor predicted UbV-PLpro interaction energies (Table S1 and Fig. S2). These results indicated that mutating R42 or V70 may not promote tight binding, so thereafter we focused on altering the A46 and K48 residues. Additionally, for every designed UbV exhibiting UbV-PLpro attractions that were predicted to be stronger than that of wtUb (Table S1), we also performed PPI-GaMD in an explicit solvent model (to account for solvent effects) to examine their binding residence time [42] (Fig. S4). Since PPI-GaMD may enhance conformational sampling, the residence time cannot be compared directly to binding affinity or dissociation rate constant (k_{off}) data. Nevertheless, the dissociation times for these UbVs were invariably longer than that determined for wtUb, supporting that these variants likely exhibit tighter binding than wtUb.

Experimental characterization of inhibitory UbVs

Next, we evaluated the UbV-dependent inhibitory efficiencies for MERS PLpro by means of sensitive fluorescent polarization (FP) to monitor the dynamics of fluorescein conjugated to the C-terminus of ISG15 (Fig. 4A). We assessed four designed UbVs harboring the dual mutations of A46F and K48E/S/I/L, as well as the two single-mutation UbVs, i.e., A46F and K48E. These UbVs all efficiently inhibited PLpro enzyme function, displaying reduced IC_{50} values (Fig. 5B, Table 1 and Fig. S6), and all bound tightly to PLpro (K_D values shown in Table 1 and Fig. S7). Relative to wtUb, substitution of A46 in the HCS1 by non-polar PHE resulted in stronger vdW attractions with nearby PLpro residues (Fig. 6A and Table S1). More specifically, A46F induced local conformational arrangements to prompt Y208 and Y223 in PLpro into forming π - π interactions and R233 in PLpro into forming π -cation attraction with the UbV (Fig. 6A). Moreover, A46F substitution substantially stabilized overall Ub-PLpro interactions, increasing the K_D value 15-fold compared to wtUb (Table S1). The polarity of K48 in wtUb significantly affects both local interactions and network correlation due to the repulsive force with the nearby K204 residue of PLpro. We postulated that mutating K48 to a nonpolar or negatively charged residue such as LEU or GLU would enhance the attractive force between K48 and K204 (Fig. 6A). Notably, the single-point A46F or K48E mutations alone resulted in IC_{50} values of 1.6 and 3.9 μ M, respectively (52.9 μ M for wtUb). The respective dual mutant variant further elevated the IC_{50} value to $\sim$ 0.2 μ M (Table 1), i.e., approximately 250-fold greater PLpro inhibition compared to wtUb. Overall, we observed a synergistic inhibitory effect and enhanced binding affinity for A46F and K48E (or K48L/K48S/K48I) substitutions in terms of PLpro and UbV interactions (Fig. 3C, 4A and Fig. S7).

Above, we have demonstrated that mutating two residues in HCS1 represents an effective strategy for designing binding inhibitors, so we explored mutating additional residues based on local structural analysis. We focused on further modifying the A46F-K48E variant (termed UbV2 hereafter), which exhibited favorable IC_{50} and K_D values. First, we subjected UbV2 to E64Y mutation, representing a residue in the FBI. Experimental measurements of the A46F-K48E-E64Y variant (hereafter denoted UbV3) revealed remarkable binding specificity to and inhibition of PLpro, with an IC_{50} of 15 nM, i.e., in significant agreement with the strongest computed binding energy (Table 1). Both the root mean square fluctuation (RMSF) and dihedral entropy of PLpro (Fig. S5) were reduced upon interaction with UbV3 compared to UbV2, with the FBI of the UbV3-PLpro complex being more stabilized (Fig. 6B). We measured a K_D of 2.77 nM for UbV3-PLpro,

which is ~80-fold greater than that determined for UbV2-PLpro. Furthermore, the IC_{50} value of A46F-E64Y (34 nM) is 48-fold or 8-fold greater than that of A46F or E64Y alone, respectively. Thus, our inhibition assays and binding affinity data greatly support that the three residues at positions 46, 48 and 64 cooperatively stabilize the interaction of Ub with PLpro, hence UbV3 is a better inhibitor of PLpro than UbV2 (Fig. 4C, 5). Altering the surface charge (K48E and E64Y) or hydrophobicity (A46F and E64Y) of the small 76-residue Ub peptide may be structurally destabilizing. Therefore, we employed circular dichroism (CD) spectroscopy to measure the far-UV CD spectra of Ub and nine selected UbVs at 25 °C and 80 °C to validate their thermostability. This analysis revealed that UbV3 displays higher thermostability than the single (A46F, K48E, E64Y) or double-point (UbV2) mutants (Fig. S8).

Ub C-terminal mutations R74N and G75S further stabilize PPIs

Next, we altered two additional residues of UbV2, R74 and G75 of the UCC region based on information from our non-covalent correlation network, with these residues serving a role in preventing cells from recognizing the UbVs. The four-point A46F-K48E-R74N-G75S mutant variant (hereafter termed UbV4) yielded a reduced IC_{50} of 110 nM. Similar to the data for double mutants, the K48E/K48L/K48S/K48I mutations resulted in the same inhibitory effect (IC_{50} : 110-290 nM, Table S1). Therefore, we further included the E64Y substitution in UbV4 to encompass all three highly correlated regions, with the resulting designed five-point UbV5 mutant displaying an IC_{50} value of 9.7 nM (i.e., 5,500-fold that of wtUb) and a K_D of 1.5 nM (Table 1). The characterized K_D values for UbV3 and UbV5 are approximately 20-fold greater than those established for the phage-display-screened UbVs ME.2 and ME.4 generated in a previous study [33] (Table S1). In addition to their impressive attribute of functional PLpro inhibition, both UbV3 and UbV5 bound highly specifically to MERS PLpro (Fig. S9), which is also an essential attribute for a good inhibitor. Thus, our UbV3 and UbV5 variants exhibit strong inhibition, equivalent to reported variants for MERS PLpro [33]. In contrast to the preserved thermostability of the UbVs designed in this study, the previously reported ME.2 and ME.4 variants harboring 15 mutated residues denatured at 80 °C [34], and thus are significantly less thermostable than our UbVs (Fig. S8). Similarly, the T_m value of the nine-point U7Ub25.2540 mutant for USP7 is 64 °C lower than that of wtUb [34]. Consequently, mutating too many residues in Ub can easily elicit stability issues. Our study demonstrates the advantage and efficacy of modifying only a few residues of a protein template to enhance PPI.

In this study, we established a novel approach to computationally select key residues responsible for PPI based on dihedral angle networks. We found that the mapped dihedral angle networks are useful for identifying critical interactions between proteins. Then, we used this information to select residues for mutation to alter PPI strengths. Using Ub-PLpro as our model system, we have demonstrated that modifying only two or three residues within the correlation network of the Ub-PLpro interface successfully enhanced PPI and resulted in a 250- to 3,500-fold reduction of MERS PLpro activity. A combination of five mutated residues in the HCS1, FBI and UCC domains of the Ub-PLpro complex resulted in a 5,500-fold ($IC_{50} = 9.7$ nM) reduction in PLpro activity and a 27,500-fold enhancement in UbV-PLpro complex affinity. Our design platform can be used to computationally examine a correlation network of protein side-chains and local pair-wise forces to efficiently design UbVs for further experimental assessment. The considerable correlation between IC_{50} and K_D values ($R^2=0.94$) implies that our experimental IC_{50} data can be used to indirectly estimate K_D values for UbV-PLpro complexes. Integrating experimental measurements and structural analyses using MD simulations, together with appropriate post-hoc analysis, can iteratively inform new designs.

Our strategy is transformative and highly efficient for identifying key mutation sites and specific residues to guide rational design of many disease-linked DUBs and Ub-bound proteins [31,32], including USP4[43], USP7[44], USP11[45], and PLpro of SARS-CoV-2[31]. Apart from producing variant-specific antibodies/vaccines for diverse spike proteins of viruses, blocking the functions of viral nonstructural proteins represents an alternative therapeutic solution to tackle COVID-19 and other coronaviruses. Designing tightly binding and specific UbV inhibitors for the coronavirus PLpro complex is an elegant solution to retard viral replication and simultaneously rescue the host antiviral immune response. This same strategy can be applied straightforwardly to other protein-protein interacting systems related to signaling and enzymatic activity. As these cellular events frequently reoccur, the protein complexes involved in cell signaling and enzymatic reactions do not have perfectly optimized PPI as stably bound complexes. Therefore, mutating residues in their interacting regions can remarkably improve binding. Overall, modifying residues displaying highly correlated dynamic motion can be used to engineer tightly binding variants for various applications.

Materials and Methods

MD Simulation Protocol

The X-ray crystal structure of MERS-CoV-PLpro-wtUb was obtained from the RCSB protein data bank (PDBID: 4RF0)[26]. PLpro was extended by one residue at the N-terminus, and wtUb was extended by two residues to ensure consistent numbers of residues for comparison with preexisting UbVs such as ME.2[33]. PLpro comprises 319 residues and Ub comprises 78 residues. The Molecular Operating Environment platform [46] was used to perform all UbV mutations. All MD simulations were performed in the AMBER 20 package with GPU acceleration [47] Force Field ff14sb[48] was applied to proteins. First, we minimized the hydrogen atoms, amino acid side-chain, and the entire protein system for 500, 1000, and 5000 steps, respectively, in a generalized Born implicit solvent. All systems were then solvated in TIP3P water, with an extension of 12 Å from the solute edge. Two Cl⁻ counter-ions were added to neutralize the charge of the system. The solvated system contains approximately 72,000 atoms. The water molecules were minimized for 1000 steps, followed by minimization of the entire system for 2000 steps. The solvated system was equilibrated under constant pressure and temperature (NPT ensemble) from 50 K to 275 K with 25 K increments and 100 ps each, and finally at 298 K for 500 ps. Production runs were also performed in the NPT ensemble at 298 K using a Langevin Thermostat with 2-fs time-steps. The first 50 ns of the MD simulation were treated as equilibrium plus. Force Distribution Analysis (FDA) [37,38] and molecular mechanics Poisson-Boltzmann surface area (MM/PBSA) calculations were performed using data from the subsequent 450 ns. The cutoff for nonbonding interactions, which includes vdW and electrostatic components, was set to 12 Å. The particle Mesh Ewald Method was used to compute long-range electrostatic interactions.

It is possible that a large protein-protein system becomes trapped at a specific local minimum, resulting in suboptimal outcomes. We performed three independent 150-ns MD simulations for each conformation and selected the lowest energy trajectories by calculating MM/PBSA energies. The exterior dielectric constant was set to 15 to accommodate the polar protein residues at the protein-water interface. The trajectories with the lowest energy were extended to a 500-ns production run. Output trajectories were saved every 1 ps for further analysis.

Side-chain Dihedral Angle Correlation Network

The following steps were used as a rationale for selecting residues for substitution:

1. Construct a correlation network of side-chain dihedral angles for wtUb-PLpro

We used T-analyst [49] to calculate the side-chain dihedral rotations of each amino acid residue and their pairwise correlations. Each side-chain dihedral angle was recorded every 100 ps through 500 ns trajectories to generate 5000 different angles per dihedral selection. Pairwise correlations were computed using a Pearson correlation formula. We converted the side-chain dihedral angles to Cartesian coordinates by means of equations (1)-(3) to accurately capture their differences and means, thereby preventing erroneous computation of their correlation at the discontinuity margin ($\pm 180^\circ$ or $360^\circ/0^\circ$) [49]. Notably, positive correlation between two side-chains indicates that the two sides rotate similarly during MD simulation.

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

$$\bar{x} = \arctan \left(\frac{\sin(x_1) + \sin(x_2) + \dots + \sin(x_n)}{\cos(x_1) + \cos(x_2) + \dots + \cos(x_n)} \right) \quad (2)$$

$$x_i - \bar{x} = \arctan \left(\frac{\sin(x_i) \cos(\bar{x}) + \sin(\bar{x}) \cos(x_i)}{\cos(x_i) \cos(\bar{x}) - \sin(x_i) \sin(\bar{x})} \right) \quad (3)$$

r_{xy} = Dihedral Pearson Correlation, $\bar{x}$ = mean of dihedral angles,

x_i = side chain dihedral angles

From the pairwise correlation matrix, we generated a correlation network by using python library NetworkX [50], enabling us to visualize the correlation between each residue.

Specific side-chain rotation can generate a chain effect and impact the rotation of distal residues. We applied a correlation cutoff of 0.3 to eliminate less correlated residues. The Ub residues selected for further analysis were: I3, T14, T22, E24, K27, I30, E34, R42, K48, Q62, S65, T66, H68, V70, L71, L73, R74, and D77.

2. Define the contact interface and eliminate residues not at the contact interfaces

We adopted FDA to select the contact interface between wtUb and PLpro. A cutoff of 10 (pN) was applied to disregard areas of low interaction. Next, we generated a heatmap of pairwise forces between wtUb and PLpro (Fig. S1). From this heatmap, we identified four clear interface regions: HCS1, AHC, FBI, and UCC. Some highly correlated residues selected from the previous step are not located at the contact interface. Residues located outside the contact interface contribute little or no interaction to binding affinity, so they were discarded from further analysis. Ub residues selected for analysis: R42, K48, Q62, S65, V70, L71, L73, and R74.

3. Consider multiple residues at the FBI

The FBI region is highly dynamic, displaying high RMSF values and it has been shown previously to crystallize in different conformations [24]. Our FDA data also revealed differential patterns among randomly-seeded production runs. Notably, our dihedral angle correlation network selected residues S65 and T66, as well as K63 on the conjugated side of PLpro. We postulated that mutating E64, which lies between S65 and K63, could further optimize our results. Ub residues selected for analysis based on this step: R42, K48, Q62, E64, S65, T66, H68, V70, L71, L73, and R74.

4. Include ALA, GLY and PRO residues

ALA, GLY and PRO residues do not have side-chains, so they would have been excluded from dihedral angle selection. To avoid excluding potentially useful mutation sites, we included all ALA, GLY and PRO residues within 5 Å of our previously selected mutation sites. Ub residues selected for analysis from this step: R42, A46, K48, Q62, E64, S65, T66, H68, V70, L71, L73, R74, and G75.

5. Further select residues based on FDA data

Residues of interest displaying strong attraction to their surrounding residues will not have been selected for mutation in the previous steps. Accordingly, we sought residues that displayed strong repulsion or weak interactions with their surrounding residues. A46 interacts only weakly with its surrounding residues, and K48 has weak interactions with surrounding residues K204, Y208 and V209 of PLpro at the HCS1. Therefore, we felt that A46 and K48 represented ideal residues for mutation. In addition, R42, V70, and G75 presented repulsive forces with their surroundings and accordingly were chosen for mutation. In contrast, L71 and L73 in the UCC exert strong attractive forces on the BL2 and AHC domains, so they were excluded from mutational analysis. Although E64 and R74 present strong attractive forces to their surroundings, we still considered them as possible mutational sites because both the dihedral angle correlation network and FDA results from different random seeds resulted in large standard deviations at the FBI and UCC because of their highly flexible nature (Fig. S1). Moreover, R74 is part of the LRGG recognition site, with mutation of R74 further hindering proteolytic cleavage [24]. Accordingly, only seven residues were ultimately selected for mutational analysis: R42, A46, K48, E64, V70, R74, and G75.

PPI-GAMD Simulation Protocol

Starting from the last frame of our MD simulation, we performed 5 ns of classical MD simulation, followed by 5 ns of PPI-GAMD equilibration to obtain the boost parameters (ntcmdprep = 500,000, ntcmd=2,500,000, ntebprep = 500,000, and nteb=2,500,000 steps). Production runs were performed in the NPT ensemble at 350 K using a Langevin Thermostat with a 2-fs time-step. We applied both potential boost and dihedral boost (igamd=17) on Ub residues within 5 Å of the PLpro residues at the contact interfaces. Applying dual boost potentials on the entire Ub structure can result in denaturation of the protein structure, leading to suboptimal results. The threshold energy of potential boost was set to the upper bound limit (iEP=2), and the threshold energy for dihedral boost was set to the lower bound limit (iED=1). The upper limit of the standard deviation for dual boost potential was set to 10 kcal/mol (sigma0P and sigma0D=10). The production run continued until we observed Ub dissociating from PLpro. We defined dissociation according to a sudden increase in C α RMSD values. Production runs were repeated with three different random seeds, with the longest dissociation times being reported herein.

Protein expression and purification

Genes encoding SARS-CoV-2 PLpro, MERS PLpro, ISG15 and ubiquitin were synthesized by GenScript (NJ, USA). Ubiquitin variants were made by site-directed mutagenesis or directed amplification (C-terminal mutations). UbV genes *ME.2* and *ME.4* were synthesized and subcloned by Genomics (Genomics Inc., Taiwan). All genes were placed in pRSFDuet-1 vector with a N-terminal hexahistidine tag (his-tag) and a TEV cleavage sequence. All resulting plasmids were transformed into the BL21 RIL cell line for protein production. For PLpro and ISG15-AVTRYVDC, *E. coli* grown in LB medium at 37 °C to an OD₆₀₀ of 0.6-0.8 was induced by treatment with 0.6 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) overnight at 16 °C (MERS PLpro and SARS-CoV-2 PLpro) or at 25 °C (ISG15-AVTRYVDC). For all 26 UbVs assessed in the current study, *E. coli* was grown in autoinduction medium containing base broth (25 mM Na₂HPO₄, 25 mM KH₂PO₄ pH 7.2, 85 mM NaCl, 0.5% yeast extract, 2% tryptone) and a sugar mix (15% v/v glycerol, 1.25% w/v glucose, 5% w/v lactose) at a 25:1 volume ratio. The *E. coli* was cultured at 37 °C for 24 hours, with UbV proteins being automatically expressed once glucose had been depleted.

E. coli cell pellets were spun down and resuspended in Buffer A (25 mM Tris-HCl pH 7.6, 200 mM NaCl, 3 mM 2-Mercaptoethanol) with the addition of 1 mM PMSF for

sonication. Cell lysates were then centrifuged, before loading the supernatant on Roche cOmplete nickel resin and then washing and eluting it using 300 mM imidazole. The his-tag of PLpro was removed by means of TEV protease treatment, and the resulting proteins were further purified by size exclusion chromatography (SEC) on an Akta FPLC system (Cytiva). ISG15-AVTRYVDC was further crosslinked with fluorescein-5-Maleimide (Santa Cruz Biotechnology) at 4 °C for 1 hour. Excess fluorescein was removed by processing through desalting columns. All proteins were flash-frozen in liquid nitrogen, aliquoted and stored at -80 °C.

Fluorescence polarization (FP)-based inhibition assay

To detect and characterize the inhibitory effects on PLpro by UbVs, we used fluorescein-labeled ISG15-AVTRYVDC (denoted ISG15* hereafter), where AVTRYVD is the N-terminal sequence of SARS-CoV-2 and a CYS residue was introduced at the C-terminal end for crosslinking with the Fluorescein-5-maleimide (Ana Spec). Values of fluorescence polarization (FP) for ISG15* and the cleaved AVTRYVDC* are approximately 120 and 20, respectively (Fig. S6), representing a sensitive tool to unravel the activity of MERS PLpro. We measured 40-μl samples in multiple wells composed of 2 μM ISG15*, 50 nM MERS PLpro (or SARS-CoV-2 PLpro), and a wide range of UbV concentrations (i.e., from 0.06 nM – 150 μM) in a 384-well plate for 1,800 seconds or longer. ISG15* alone was used as a control. Assessments of individual UbV concentrations were conducted in triplicate to determine standard errors.

The reaction rate constants (k_{obs}) of ISG15* cleavages were obtained by curve fitting using one-phase decay (equation 4). Enzymatic activity was normalized according to the ratio of the k_{obs} values with or without UbV. To determine IC₅₀ values, normalized UbV concentration-dependent enzymatic activities were fitted according to the logistic non-linear regression model (equation 5).

$$y = (y_0 - y_i) * e^{((-K * x) + y_i)} \quad (4)$$

where y_0 is the Y value when X (time) is zero, y_i is a plateau Y value at infinite time, and K is the rate reaction constant.

$$y = 1 + \left(\frac{x}{IC_{50}} \right)^{Hill\ slope} \quad (5)$$

The Y values (representing the FP readouts) were normalized to between 0 and 1.0 for the fitting that included “Hill slope”, with all of the values being ~1.0.

Bio-Layer Interferometry (BLI)

BLI experiments were performed on an Octet RED 96 system (Sartorius) using anti-GST antibody biosensors for GST-tagged MERS PLpro and UbV (or Ub) as analytes at 25 °C. The ligand and analytes were diluted into reaction buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.1 mg/ml BSA, 0.01% Tween-20). Steady-state response wavelength shifts of analytes in multiple concentrations were used to fit a single-site binding system and determine the dissociation constant (K_D) according to equation 6.

$$R_{eq} = R_{max} \frac{[C]}{K_D + [C]} \quad (6)$$

where R_{eq} is the steady-state response shift of the sensorgram curve, $[C]$ is the concentration of analytes, R_{max} is the maximal response, and K_D is the dissociation constant. To determine R_{max} and K_D values, we applied a Levenberg–Marquardt algorithm to perform iterative non-linear least squares curve fitting.

The k_{on} and k_{off} values of UbV3 and UbV5 were globally fitted to the time-dependent response wavelength shifts in the association and dissociation well, respectively, using the Octet Data Analysis software (Sartorius). K_D values for UbV3 and UbV5 were calculated as “ k_{off}/k_{on} ”.

Circular Dichroism (CD) Spectroscopy

CD measurements were performed on a Jasco J-815 spectropolarimeter. Far-UV spectra were measured from 260 to 195 nm at 25 and 80 °C. UbV samples (10 μM) diluted in 25 mM Tris pH 7.6 and 50 mM NaCl buffer were measured in a 1 mm quartz cell (Hellma GmbH). Melting temperature experiments were performed at 25 and 80 °C at 1 °C intervals. The CD spectra were averaged from triplicates acquired at a scanning speed of 50 nm/min and a digital integration time of 1 second.

Acknowledgments

This study was supported by the US National Institutes of Health (GM-109045), the US National Science Foundation (MCB-1350401 to C.C), and an Academia Sinica (Taipei,

Taiwan) Career Development Award (AS-CDA-110-L03 to K.P.W.), and the National Science and Technology Council (110-2311-B-001-046 and 111-2113-M-001-026-MY2 to K.P.W.). We appreciate the technical and instrumental support from the Academia Sinica (AS) Biophysics Core Facility (AS-CFII-111-201) and the Biophysics Instrumentation Laboratory, Institute of Biological Chemistry, Academia Sinica.

Author contributions

Conceptualization: KPW, CC

Methodology: TIH, YJH, WLL, KPW, CC

Formal Analysis: TIH, YJH

Data Curation: TIH, YJH, WLL

Supervision: KPW, CC

Writing—original draft: TIH, YJH, KPW, CC

Code Availability:

Source code for the dihedral angle correlation network analysis is available upon request.

Data and materials availability:

All data used in the main text or the supplementary materials are available on request.

References

- [1] J. Zahradník, G. Schreiber, Protein Engineering in the Design of Protein–Protein Interactions: SARS-CoV-2 Inhibitors as a Test Case, *Biochemistry*. 60 (2021) 3429–3435. <https://doi.org/10.1021/acs.biochem.1c00356>.
- [2] A. Marchand, A.K. Van Hall-Beauvais, B.E. Correia, Computational design of novel protein–protein interactions – An overview on methodological approaches and applications, *Curr Opin Struct Biol*. 74 (2022) 102370. <https://doi.org/10.1016/j.sbi.2022.102370>.
- [3] D. Petrey, H. Zhao, S.J. Trudeau, D. Murray, B. Honig, PrePPI: A Structure Informed Proteome-wide Database of Protein–Protein Interactions, *J Mol Biol*. (2023) 168052. <https://doi.org/10.1016/j.jmb.2023.168052>.
- [4] L. Cao, B. Coventry, I. Goresnik, B. Huang, W. Sheffler, J.S. Park, K.M. Jude, I. Marković, R.U. Kadam, K.H.G. Verschueren, K. Verstraete, S.T.R. Walsh, N. Bennett, A. Phal, A. Yang, L. Kozodoy, M. DeWitt, L. Picton, L. Miller, E.-M. Strauch, N.D. DeBouver, A. Pires, A.K. Bera, S. Halabiya, B. Hammerson, W. Yang, S. Bernard, L. Stewart, I.A. Wilson, H. Ruohola-Baker, J. Schlessinger, S. Lee, S.N. Savvides, K.C. Garcia, D. Baker, Design of protein-binding proteins from the target structure alone, *Nature*. 605 (2022) 551–560. <https://doi.org/10.1038/s41586-022-04654-9>.
- [5] V. Frappier, A.E. Keating, Data-driven computational protein design, *Curr Opin Struct Biol*. 69 (2021) 63–69. <https://doi.org/10.1016/j.sbi.2021.03.009>.
- [6] R. Pearce, X. Huang, D. Setiawan, Y. Zhang, EvoDesign: Designing Protein–Protein Binding Interactions Using Evolutionary Interface Profiles in Conjunction with an Optimized Physical Energy Function, *J Mol Biol*. 431 (2019) 2467–2476. <https://doi.org/10.1016/j.jmb.2019.02.028>.
- [7] D. Rooklin, A.E. Modell, H. Li, V. Berdan, P.S. Arora, Y. Zhang, Targeting Unoccupied Surfaces on Protein–Protein Interfaces, *J Am Chem Soc*. 139 (2017) 15560–15563. <https://doi.org/10.1021/jacs.7b05960>.
- [8] S. Pahari, G. Li, A.K. Murthy, S. Liang, R. Fragoza, H. Yu, E. Alexov, SAAMBE-3D: Predicting Effect of Mutations on Protein–Protein Interactions, *Int J Mol Sci*. 21 (2020) 2563. <https://doi.org/10.3390/ijms21072563>.
- [9] S.T. Kudlacek, S. Metz, D. Thiono, A.M. Payne, T.T.N. Phan, S. Tian, L.J. Forsberg, J. Maguire, I. Seim, S. Zhang, A. Tripathy, J. Harrison, N.I. Nicely, S. Soman, M.K. McCracken, G.D. Gromowski, R.G. Jarman, L. Premkumar, A.M. de Silva, B. Kuhlman, Designed, highly expressing, thermostable dengue virus 2 envelope protein dimers elicit quaternary epitope antibodies, *Sci Adv*. 7 (2021). <https://doi.org/10.1126/sciadv.abg4084>.
- [10] L. Rosenfeld, M. Heyne, J.M. Shifman, N. Papo, Protein Engineering by Combined Computational and In Vitro Evolution Approaches, *Trends Biochem Sci*. 41 (2016) 421–433. <https://doi.org/10.1016/j.tibs.2016.03.002>.
- [11] T.-L. V Lite, R.A. Grant, I. Nosedal, M.L. Littlehale, M.S. Guo, M.T. Laub, Uncovering the basis of protein-protein interaction specificity with a combinatorially complete library, *Elife*. 9 (2020). <https://doi.org/10.7554/eLife.60924>.
- [12] J. Parkinson, R. Hard, W. Wang, The RESP AI model accelerates the identification of tight-binding antibodies, *Nat Commun*. 14 (2023) 454. <https://doi.org/10.1038/s41467-023-36028-8>.
- [13] K.N. Swatek, D. Komander, Ubiquitin modifications, *Nature Publishing Group*. 26 (2016) 399–422. <https://doi.org/10.1038/cr.2016.39>.
- [14] D. Komander, M. Rape, The Ubiquitin Code, *Annu Rev Biochem*. 81 (2012) 203–229. <https://doi.org/10.1146/annurev-biochem-060310-170328>.
- [15] W. Kim, E.J. Bennett, E.L. Huttlin, A. Guo, J. Li, A. Possemato, M.E. Sowa, R. Rad, J. Rush, M.J. Comb, J.W. Harper, S.P. Gygi, Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome, *Mol Cell*. 44 (2011) 325–340. <https://doi.org/10.1016/j.molcel.2011.08.025>.
- [16] J.A. Harrigan, X. Jacq, N.M. Martin, S.P. Jackson, Deubiquitylating enzymes and drug discovery: emerging opportunities, *Nature Reviews Drug Discovery* 2017 17:1. 17 (2017) 57–78. <https://doi.org/10.1038/nrd.2017.152>.
- [17] J.K. Morrow, H.K. Lin, S.C. Sun, S. Zhang, Targeting ubiquitination for cancer therapies, *Future Med Chem*. 7 (2015) 2333. <https://doi.org/10.4155/FMC.15.148>.
- [18] M. Békés, W. Rut, P. Kasperkiewicz, M.P.C. Mulder, H. Ovaa, M. Drag, C.D. Lima, T.T. Huang, SARS hCoV papain-like protease is a unique Lys48 linkage-specific di-distributive deubiquitinating enzyme, *Biochemical Journal*. 468 (2015) 215–226. <https://doi.org/10.1042/BJ20141170>.
- [19] B.A. Bailey-Elkin, R.C.M. Knaap, M. Kikkert, B.L. Mark, Structure and Function of Viral Deubiquitinating Enzymes, *J Mol Biol*. 429 (2017) 3441–3470. <https://doi.org/10.1016/j.jmb.2017.06.010>.
- [20] B.H. Harcourt, D. Jukneliene, A. Kanjanahaluethai, J. Bechill, K.M. Severson, C.M. Smith, P.A. Rota, S.C. Baker, Identification of Severe Acute Respiratory Syndrome Coronavirus Replicase

- Products and Characterization of Papain-Like Protease Activity, *J Virol.* 78 (2004) 13600–13612. <https://doi.org/10.1128/JVI.78.24.13600-13612.2004>.
- [21] D. Shin, R. Mukherjee, D. Grewe, D. Bojkova, K. Baek, A. Bhattacharya, L. Schulz, M. Widera, A.R. Mehdipour, G. Tascher, P.P. Geurink, A. Wilhelm, G.J. van der Heden van Noort, H. Ovaa, S. Müller, K.-P. Knobeloch, K. Rajalingam, B.A. Schulman, J. Cinatl, G. Hummer, S. Ciesek, I. Dikic, Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity, *Nature.* 587 (2020) 657–662. <https://doi.org/10.1038/s41586-020-2601-5>.
- [22] T. Klemm, G. Ebert, D.J. Calleja, C.C. Allison, L.W. Richardson, J.P. Bernardini, B.G. Lu, N.W. Kuchel, C. Grohmann, Y. Shibata, Z.Y. Gan, J.P. Cooney, M. Doerflinger, A.E. Au, T.R. Blackmore, G.J. Heden van Noort, P.P. Geurink, H. Ovaa, J. Newman, A. Riboldi-Tunnicliffe, P.E. Czabotar, J.P. Mitchell, R. Feltham, B.C. Lechtenberg, K.N. Lowes, G. Dewson, M. Pellegrini, G. Lessene, D. Komander, Mechanism and inhibition of the papain-like protease, PLpro, of SARS-CoV-2, *EMBO J.* 39 (2020). <https://doi.org/10.15252/embj.2020106275>.
- [23] Y.M. Báez-Santos, A.M. Mielech, X. Deng, S. Baker, A.D. Mesecar, Catalytic Function and Substrate Specificity of the Papain-Like Protease Domain of nsp3 from the Middle East Respiratory Syndrome Coronavirus, *J Virol.* 88 (2014) 12511–12527. <https://doi.org/10.1128/JVI.01294-14>.
- [24] Y.M. Báez-Santos, S.E. St. John, A.D. Mesecar, The SARS-coronavirus papain-like protease: Structure, function and inhibition by designed antiviral compounds, *Antiviral Res.* 115 (2015) 21–38. <https://doi.org/10.1016/j.antiviral.2014.12.015>.
- [25] United Nations, World Economic Situation and Prospects 2022, New York, 2022.
- [26] B.A. Bailey-Elkin, R.C.M. Knaap, G.G. Johnson, T.J. Dalebout, D.K. Ninaber, P.B. Van Kasteren, P.J. Bredenbeek, E.J. Snijder, M. Kikkert, B.L. Mark, Crystal Structure of the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) Papain-like Protease Bound to Ubiquitin Facilitates Targeted Disruption of Deubiquitinating Activity to Demonstrate Its Role in Innate Immune Suppression, *J Biol Chem.* 289 (2014) 34667. <https://doi.org/10.1074/JBC.M114.609644>.
- [27] J. Lei, Y. Kusov, R. Hilgenfeld, Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein, *Antiviral Res.* 149 (2018) 58–74. <https://doi.org/10.1016/j.antiviral.2017.11.001>.
- [28] D. Shin, R. Mukherjee, D. Grewe, D. Bojkova, K. Baek, A. Bhattacharya, L. Schulz, M. Widera, A.R. Mehdipour, G. Tascher, P.P. Geurink, A. Wilhelm, G.J. van der Heden van Noort, H. Ovaa, S. Müller, K.-P. Knobeloch, K. Rajalingam, B.A. Schulman, J. Cinatl, G. Hummer, S. Ciesek, I. Dikic, Papain-like protease regulates SARS-CoV-2 viral spread and innate immunity, *Nature.* 587 (2020) 657–662. <https://doi.org/10.1038/s41586-020-2601-5>.
- [29] X. Yang, X. Chen, G. Bian, J. Tu, Y. Xing, Y. Wang, Z. Chen, Proteolytic processing, deubiquitinase and interferon antagonist activities of Middle East respiratory syndrome coronavirus papain-like protease, *Journal of General Virology.* 95 (2014) 614–626. <https://doi.org/10.1099/VIR.0.059014-0/CITE/REFWORKS>.
- [30] W. Zhang, K.-P. Wu, M.A. Sartori, H.B. Kamadurai, A. Ordureau, C. Jiang, P.Y. Mercredi, R. Murchie, J. Hu, A. Persaud, M. Mukherjee, N. Li, A. Doye, J.R. Walker, Y. Sheng, Z. Hao, Y. Li, K.R. Brown, E. Lemichez, J. Chen, Y. Tong, J.W. Harper, J. Moffat, D. Rotin, B.A. Schulman, S.S. Sidhu, System-Wide Modulation of HECT E3 Ligases with Selective Ubiquitin Variant Probes, *Mol Cell.* 62 (2016) 121–136. <https://doi.org/10.1016/j.molcel.2016.02.005>.
- [31] V.J.E. van Vliet, N. Huynh, J. Palà, A. Patel, A. Singer, C. Slater, J. Chung, M. van Huizen, J. Teyra, S. Miersch, G.-K. Luu, W. Ye, N. Sharma, S.S. Ganaie, R. Russell, C. Chen, M. Maynard, G.K. Amarasinghe, B.L. Mark, M. Kikkert, S.S. Sidhu, Ubiquitin variants potently inhibit SARS-CoV-2 PLpro and viral replication via a novel site distal to the protease active site, *PLoS Pathog.* 18 (2022) e1011065. <https://doi.org/10.1371/journal.ppat.1011065>.
- [32] A. Ernst, G. Avvakumov, J. Tong, Y. Fan, Y. Zhao, P. Alberts, A. Persaud, J.R. Walker, A.-M. Neculai, D. Neculai, A. Vorobyov, P. Garg, L. Beatty, P.-K. Chan, Y.-C. Juang, M.-C. Landry, C. Yeh, E. Zeqiraj, K. Karamboulas, A. Allali-Hassani, M. Vedadi, M. Tyers, J. Moffat, F. Sicheri, L. Pelletier, D. Durocher, B. Raught, D. Rotin, J. Yang, M.F. Moran, S. Dhe-Paganon, S.S. Sidhu, A Strategy for Modulation of Enzymes in the Ubiquitin System, *Science* (1979). 339 (2013) 590–595. <https://doi.org/10.1126/science.1230161>.
- [33] W. Zhang, B.A. Bailey-Elkin, R.C.M. Knaap, B. Khare, T.J. Dalebout, G.G. Johnson, P.B. van Kasteren, N.J. McLeish, J. Gu, W. He, M. Kikkert, B.L. Mark, S.S. Sidhu, Potent and selective inhibition of pathogenic viruses by engineered ubiquitin variants, *PLoS Pathog.* 13 (2017). <https://doi.org/10.1371/journal.ppat.1006372>.
- [34] Y. Zhang, L. Zhou, L. Rouge, A.H. Phillips, C. Lam, P. Liu, W. Sandoval, E. Helgason, J.M. Murray, I.E. Wertz, J.E. Corn, Conformational stabilization of ubiquitin yields potent and selective inhibitors of USP7, *Nature Chemical Biology* 2012 9:1. 9 (2012) 51–58. <https://doi.org/10.1038/nchembio.1134>.

- [35] M.G.F. Sun, M.H. Seo, S. Nim, C. Corbi-Verge, P.M. Kim, Protein engineering by highly parallel screening of computationally designed variants, *Sci Adv.* 2 (2016).
https://doi.org/10.1126/SCIADV.1600692/SUPPL_FILE/1600692_TABLE_S3.XLSX.
- [36] H. Lee, H. Lei, B.D. Santarsiero, J.L. Gatuz, S. Cao, A.J. Rice, K. Patel, M.Ż. Szypulinski, I. Ojeda, A.K. Ghosh, M.E. Johnson, Inhibitor Recognition Specificity of MERS-CoV Papain-like Protease May Differ from That of SARS-CoV, *ACS Chem Biol.* 10 (2015) 1456–1465.
<https://doi.org/10.1021/cb500917m>.
- [37] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX.* 1–2 (2015) 19–25. <https://doi.org/10.1016/j.softx.2015.06.001>.
- [38] B.I. Costescu, F. Gräter, Time-resolved force distribution analysis, *BMC Biophys.* 6 (2013) 5.
<https://doi.org/10.1186/2046-1682-6-5>.
- [39] P.J. Duerksen-Hughes, X. Xu, K.D. Wilkinson, Structure and function of ubiquitin: evidence for differential interactions of arginine-74 with the activating enzyme and the proteases of ATP-dependent proteolysis, *Biochemistry.* 26 (1987) 6980–6987. <https://doi.org/10.1021/bi00396a019>.
- [40] T.J. Burch, A.L. Haas, Site-Directed Mutagenesis of Ubiquitin. Differential Roles for Arginine in the Interaction with Ubiquitin-Activating Enzyme, *Biochemistry.* 33 (1994) 7300–7308.
<https://doi.org/10.1021/bi00189a035>.
- [41] K.E. Sloper-Mould, J.C. Jemc, C.M. Pickart, L. Hicke, Distinct Functional Surface Regions on Ubiquitin, *Journal of Biological Chemistry.* 276 (2001) 30483–30489.
<https://doi.org/10.1074/jbc.M103248200>.
- [42] J. Wang, Y. Miao, Protein–Protein Interaction-Gaussian Accelerated Molecular Dynamics (PPI-GaMD): Characterization of Protein Binding Thermodynamics and Kinetics, *J Chem Theory Comput.* 18 (2022) 1275–1285. <https://doi.org/10.1021/acs.jctc.1c00974>.
- [43] M. Clerici, M.P.A. Luna-Vargas, A.C. Faesen, T.K. Sixma, The DUSP–Ubl domain of USP4 enhances its catalytic efficiency by promoting ubiquitin exchange, *Nat Commun.* 5 (2014) 5399.
<https://doi.org/10.1038/ncomms6399>.
- [44] L. Kategaya, P. Di Lello, L. Rougé, R. Pastor, K.R. Clark, J. Drummond, T. Kleinheinz, E. Lin, J.-P. Upton, S. Prakash, J. Heideker, M. McClelland, M.S. Ritorto, D.R. Alessi, M. Trost, T.W. Bainbridge, M.C.M. Kwok, T.P. Ma, Z. Stiffler, B. Brasher, Y. Tang, P. Jaishankar, B.R. Hearn, A.R. Renslo, M.R. Arkin, F. Cohen, K. Yu, F. Peale, F. Gnad, M.T. Chang, C. Klijn, E. Blackwood, S.E. Martin, W.F. Forrest, J.A. Ernst, C. Ndubaku, X. Wang, M.H. Beresini, V. Tsui, C. Schwerdtfeger, R.A. Blake, J. Murray, T. Maurer, I.E. Wertz, USP7 small-molecule inhibitors interfere with ubiquitin binding, *Nature.* 550 (2017) 534–538. <https://doi.org/10.1038/nature24006>.
- [45] H.-C. Wu, Y.-C. Lin, C.-H. Liu, H.-C. Chung, Y.-T. Wang, Y.-W. Lin, H.-I. Ma, P.-H. Tu, S.E. Lawler, R.-H. Chen, USP11 regulates PML stability to control Notch-induced malignancy in brain tumours, *Nat Commun.* 5 (2014) 3214. <https://doi.org/10.1038/ncomms4214>.
- [46] Chemical Computing Group ULC, Molecular Operating Environment (MOE), (2023).
- [47] D.A. Case Ross C Walker, A.M. Roitberg Kenneth Merz Pengfei Li, Amber 2020 Reference Manual Principal contributors to the current codes, n.d. <http://ambermd.org/contributors.html>.
- [48] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J Chem Theory Comput.* 11 (2015) 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>.
- [49] R. Ai, M. Qaiser Fatmi, C.A. Chang, T-Analyst: a program for efficient analysis of protein conformational changes by torsion angles, *J Comput Aided Mol Des.* 24 (2010) 819–827.
<https://doi.org/10.1007/s10822-010-9376-y>.
- [50] P. Swart, A.A. Hagberg hagberg, lanlgov -Los, D.A. Schult, P.J. Swart swart, Exploring Network Structure, Dynamics, and Function Using NetworkX, 2008.
http://conference.scipy.org/proceedings/SciPy2008/paper_2.

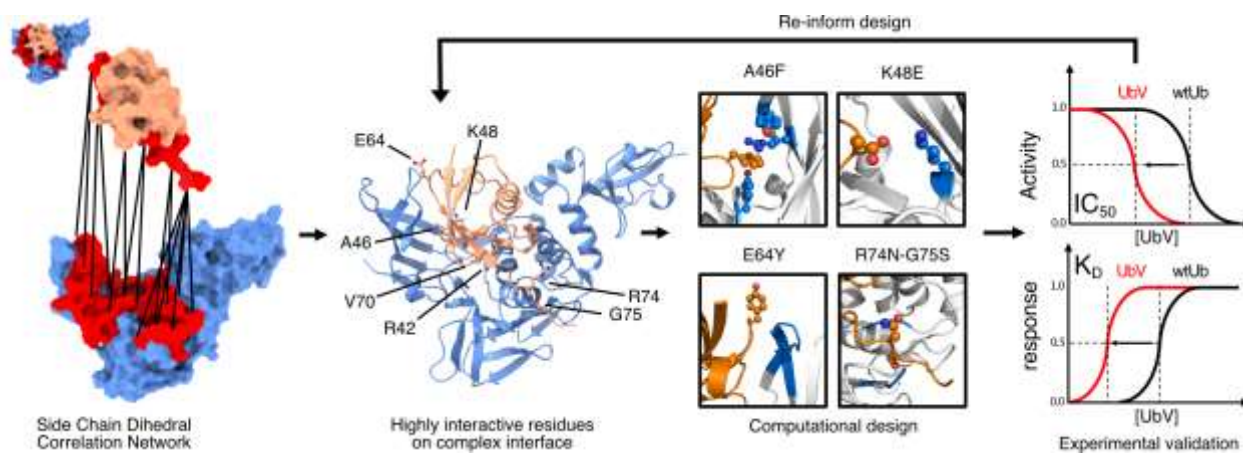


Figure 1. Workflow for the rational design of ubiquitin (Ub) variants. Molecular modeling by means of molecular dynamics (MD) simulations and post-hoc analysis reveals highly correlated regions in the interface (red) between PLpro (blue) and Ub (orange). The workflow shows how we integrated computational and experimental design, validation, and interpretation.

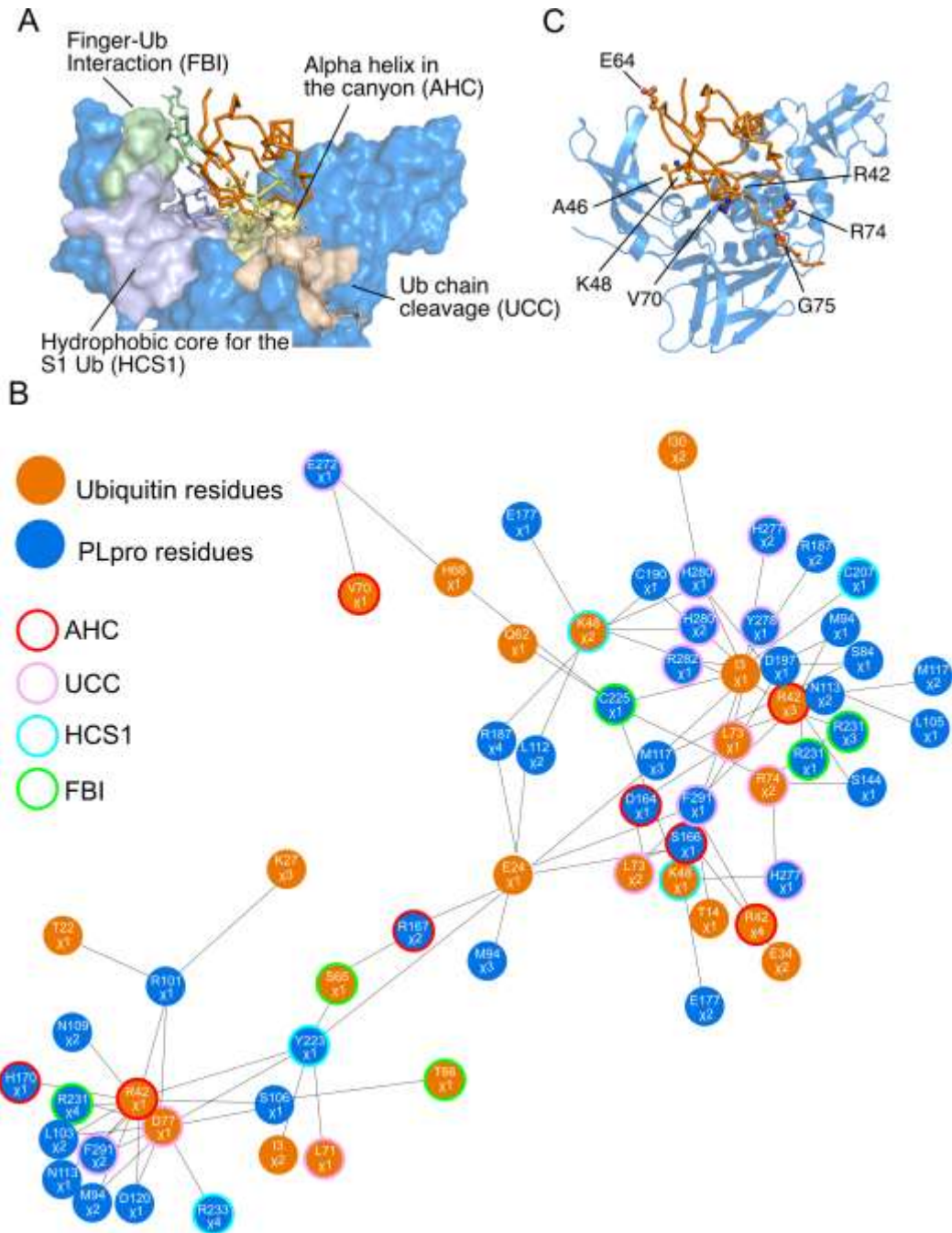


Figure 2. Highly correlated regions between Ub and MERS-CoV PLpro. **(A)** Four highly correlated regions at the contact interface of Ub (stick model) and PLpro (surface model) are shown: FBI (pale green), HCS1 (pale purple), AHC (pale yellow) and UCC (pale orange). **(B)** The side-chain dihedral angle correlation network showing how mutation leads to conformational changes in distal regions (see Methods for details). The specific dihedral angles are indicated as χ . **(C)** Residues targeted for mutation based on interaction networks: R42, A46, K48, E64, V70, R74, and G75. Note that R42 and V70 were not considered in further Ub variant (UbV) designs and experiments after computational prediction.

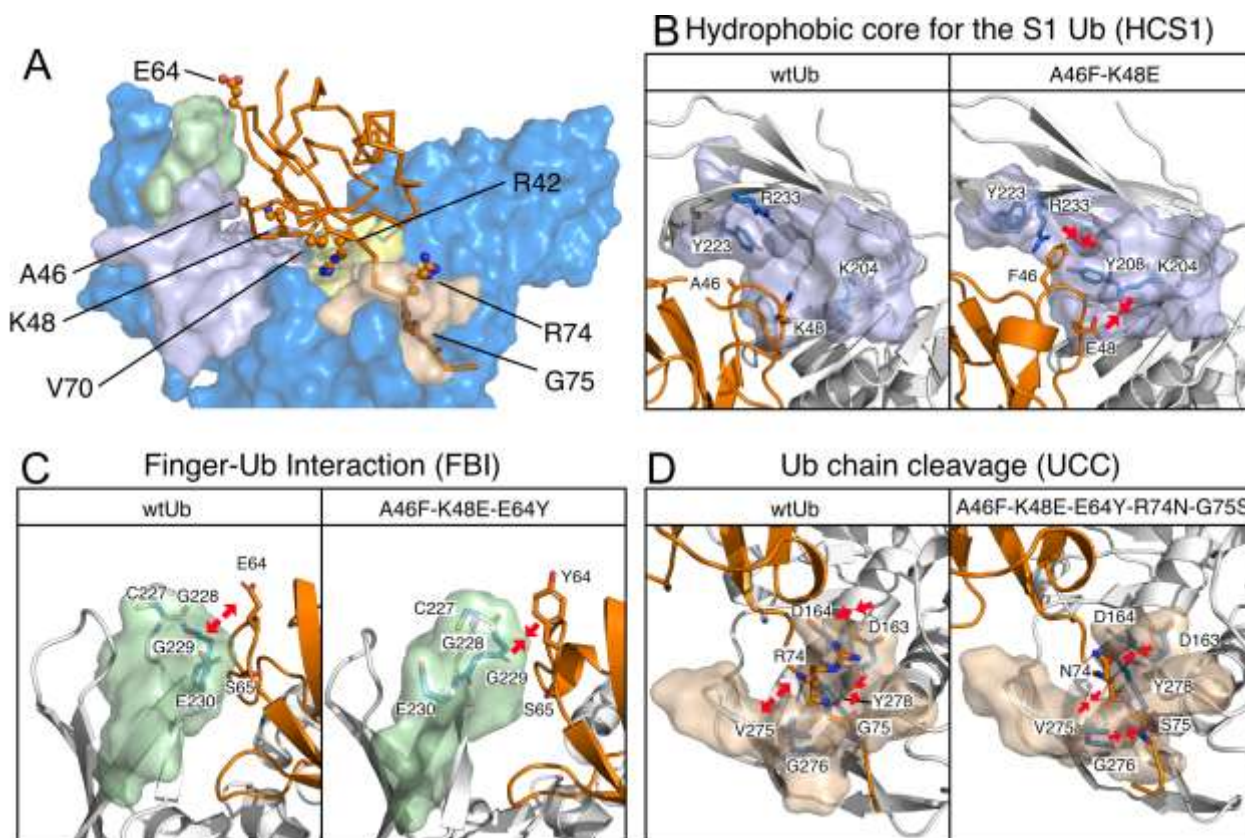


Figure 3. Selection of candidate residues suitable for mutational analysis to strengthen localized attraction. **(A)** Residues targeted for mutation based on interaction networks: R42, A46, K48, E64, V70, R74, and G75. **(B)** wtUb displays weak interaction with surrounding residues. The A46F and K48E mutations have the potential to increase both vdW and electrostatic interactions. **(C)** The E64Y mutation increases vdW attraction for G228 and G229. **(D)** The R74N mutation enhances electrostatic attraction for V275, whereas the G75S mutation increases the attraction for G276. Inward-pointing arrows indicate attraction, outward-pointing arrows represent repulsion. Color-coded regions: Ubiquitin (Ub, orange stick model), PLpro (blue surface model), FBI (pale green), HCS1 (pale blue), AHC (pale yellow), and UCC (pale orange).

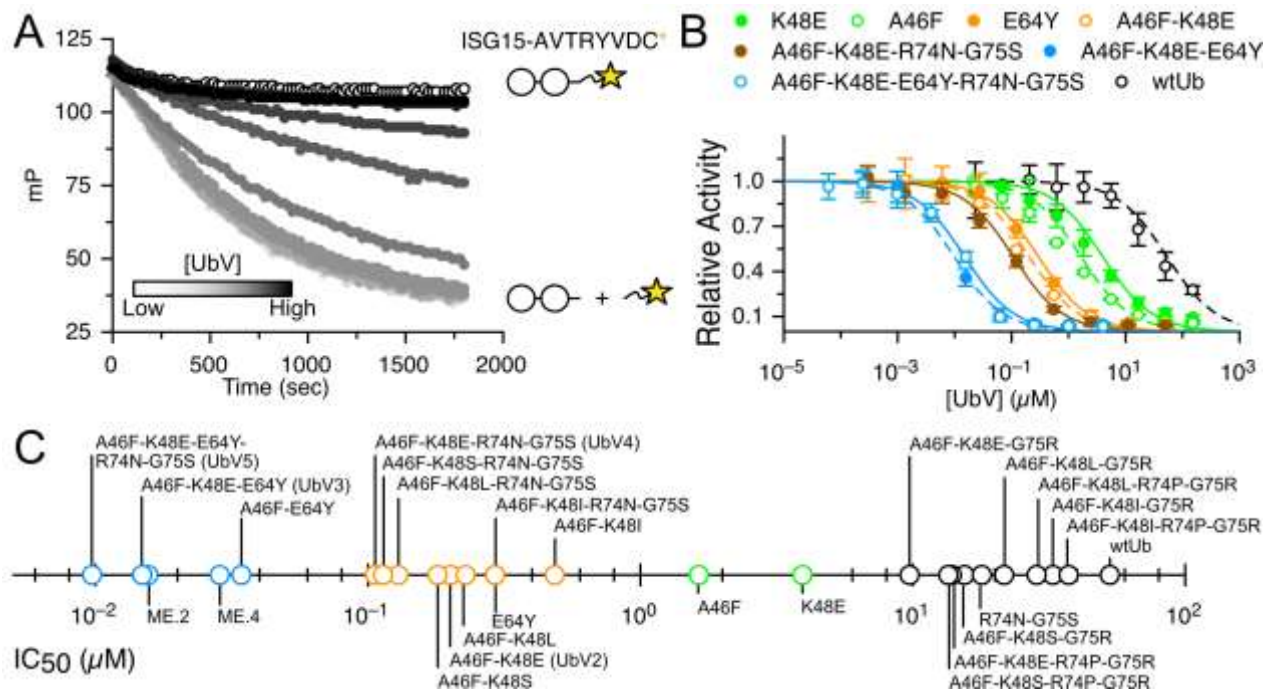


Figure 4. The inhibitory IC₅₀ values for designed UbVs. The schematics to the right in (A) represent the cleavage reaction for ISG15* substrate, where * stands for the AVTRYVDC sequence (part of SARS-CoV-2 NSP2) crosslinked with a fluorescein probe. When an UbV is added to the mixture, the substrate binding site is blocked and cleavage is retarded. The grayscale gradient reflects inhibition strength. (B) Monitoring the real-time cleavage reaction by detecting fluorescence polarization. Relative inhibitory activity (IC₅₀ curves) of seven selected UbVs and wtUb reveal how PLpro activity is progressively inhibited as the number of mutations increases from two to five. (C) Experimentally measured IC₅₀ values for 26 UbVs summarized on a logarithmic scale, with black, green, orange, and blue circles indicating weak, medium, strong and, very strong inhibitors, respectively.

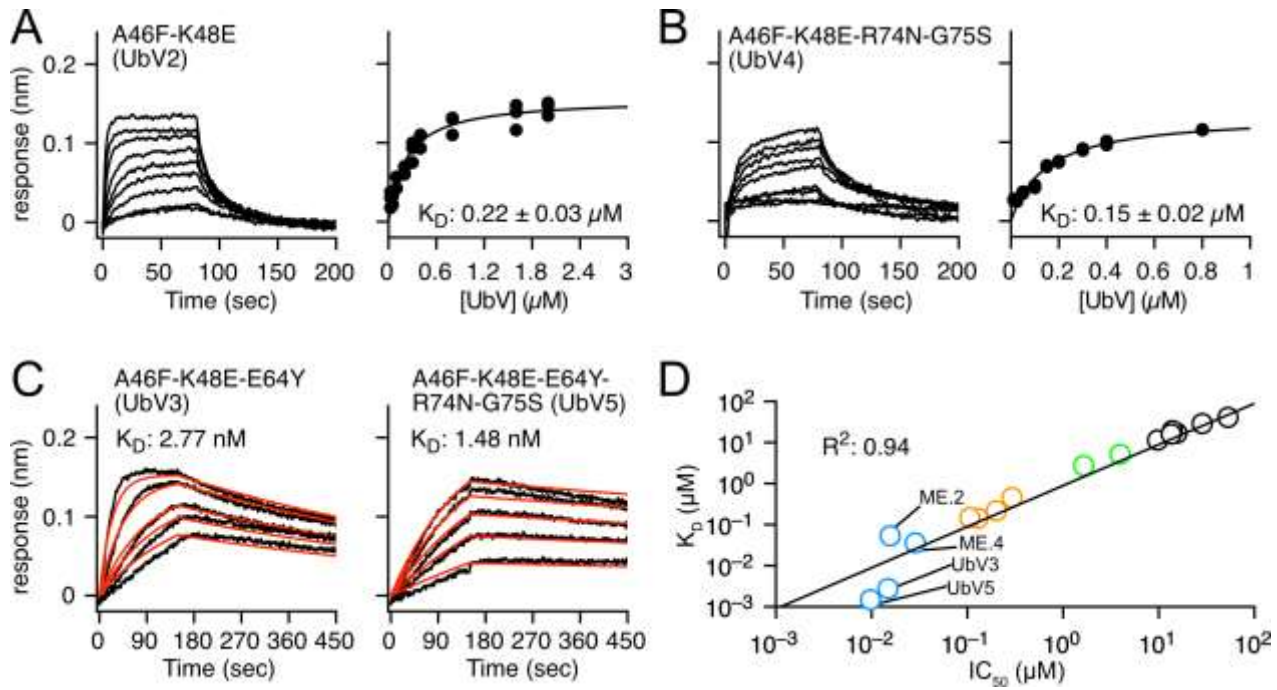


Figure 5. (A, B) BLI sensorgrams (left) and fitting curves (right) for UbV2 and UbV4 show K_D values of $0.22 \mu M$ and $0.15 \mu M$, respectively. (C) UbV3 and UbV5 bind tightly to PLpro, showing K_D values of 2.77 and 1.48 nM , respectively. The fitting values for k_{on} and k_{off} are colored red in the titrated BLI sensorgrams. (D) The IC_{50} and K_D values we measured are strongly correlated ($R^2 = 0.94$). The K_D values of ME.2 and ME.4 were published previously[33].

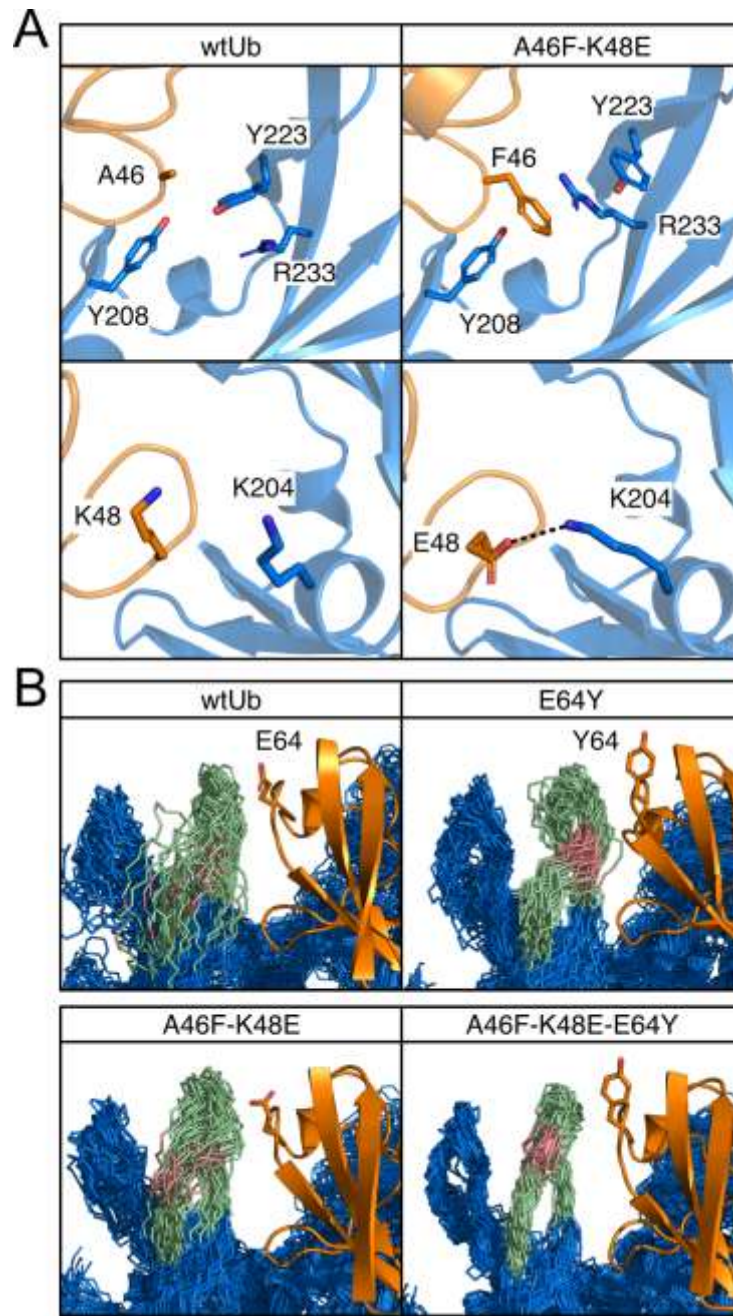


Figure 6. Conformational dynamics and interactions of wtUb and UbVs residues with the HCS1 and FBI regions of the Ub-PLpro complex. Ub and PLpro are shown in orange or blue, respectively. **(A)** The new attraction forces introduced by the A46F and K48E mutations in the HCS1 region are shown in the right panels. **(B)** Superimposition of 50 PLpro conformations from a 500-ns MD run. The FBI region (pale green) of the Ub-PLpro complex is highly stabilized by the E64Y mutation, with the hinging motion of backbone residue E230 (pink) in the interaction network contributing significantly to the loop conformations. Wild-type E64 or mutated Y64 of Ub (orange) is shown in stick format.

Table 1. Computational and experimental evaluation of the binding affinity between MERS PLpro and UbVs.

Ubiquitin and variants	Binding energy (kcal/mol)	Dissociation time (ns)	IC ₅₀	K _D	PLpro selectivity*	
					MERS	SARS-CoV-2
wtUb	-41.95 ± 2.56	32	52.91 ± 6.98 µM	40.75 ± 3.82 µM	-	-
A46F	-48.25 ± 1.09	67.5	1.64 ± 0.03 µM	2.74 ± 0.29 µM	+++	-
K48E	-49.52 ± 2.56	75	3.94 ± 0.42 µM	5.15 ± 0.71 µM	++	-
E64Y	-42.00 ± 1.49	68	0.29 ± 0.03 µM	0.46 ± 0.04 µM	++++	-
A46F-K48E (UbV2)	-50.54 ± 2.35	207	0.20 ± 0.00 µM	0.22 ± 0.03 µM	+++++	-
A46F-K48L	-49.03 ± 2.33	200	0.23 ± 0.01 µM	N/A	N/A	N/A
A46F-K48S	-47.13 ± 2.84	N/A	0.18 ± 0.01 µM	N/A	N/A	N/A
A46F-K48E-G75R	-45.24 ± 1.48	40	9.74 ± 0.15 µM	11.41 ± 1.26 µM	N/A	N/A
A46F-K48I	N/A	N/A	0.49 ± 0.04 µM	N/A	N/A	N/A
A46F-K48E-E64Y (UbV3)	-53.77 ± 1.37	>250	14.84 ± 1.44 nM	2.77 nM	+++++	-
A46F-K48E-R74N-G75S (UbV4)	-50.18 ± 0.88	65	0.11 ± 0.01 µM	0.15 ± 0.02 µM	+++++	-
A46F-K48L-R74N-G75S	-46.49 ± 0.98	>250	0.13 ± 0.01 µM	0.14 ± 0.02 µM	N/A	N/A
A46F-K48E-E64Y-R74N-G75S (UbV5)	-50.81 ± 3.25	88	9.71 ± 0.74 nM	1.48 nM	+++++	-
ME.2**	-47.98 ± 1.72	>250	15.62 ± 2.54 nM	53.2 ± 2.2 nM	+++++	-
ME.4**	N/A	N/A	28.57 ± 1.94 nM	35.9 ± 1.6 nM	+++++	-

* Each “+” indicates a 0-20% reduction of PLpro activity compared with UbV-free conditions.

** ME.2 and ME.4 are UbVs [33] both of which harbor 15 mutated sites. The K_D values and PLpro selectivity for the ME.2 and ME.4 UbVs were extracted from a study published previously [34].