Supporting Information

Probing the dynamic structure-function and structure-free energy relationships of the corona virus main protease with Biodynamics theory

Hongbin Wan, Vibhas Aravamuthan, Robert A. Pearlstein*

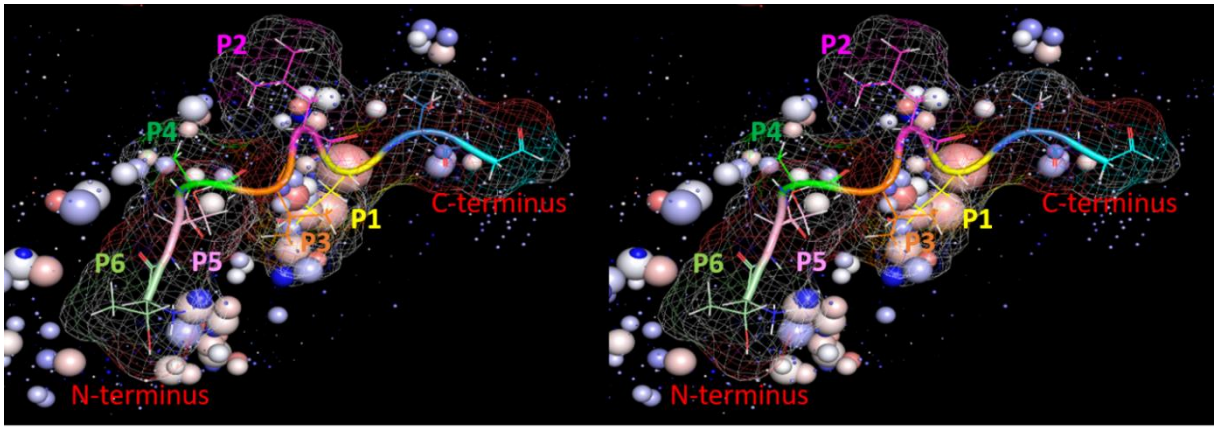
Global Discovery Chemistry, Computer-Aided Drug Discovery, Novartis Institutes for BioMedical Research, 181 Massachusetts Avenue, Cambridge, MA 02139, USA

* To whom correspondence should be addressed

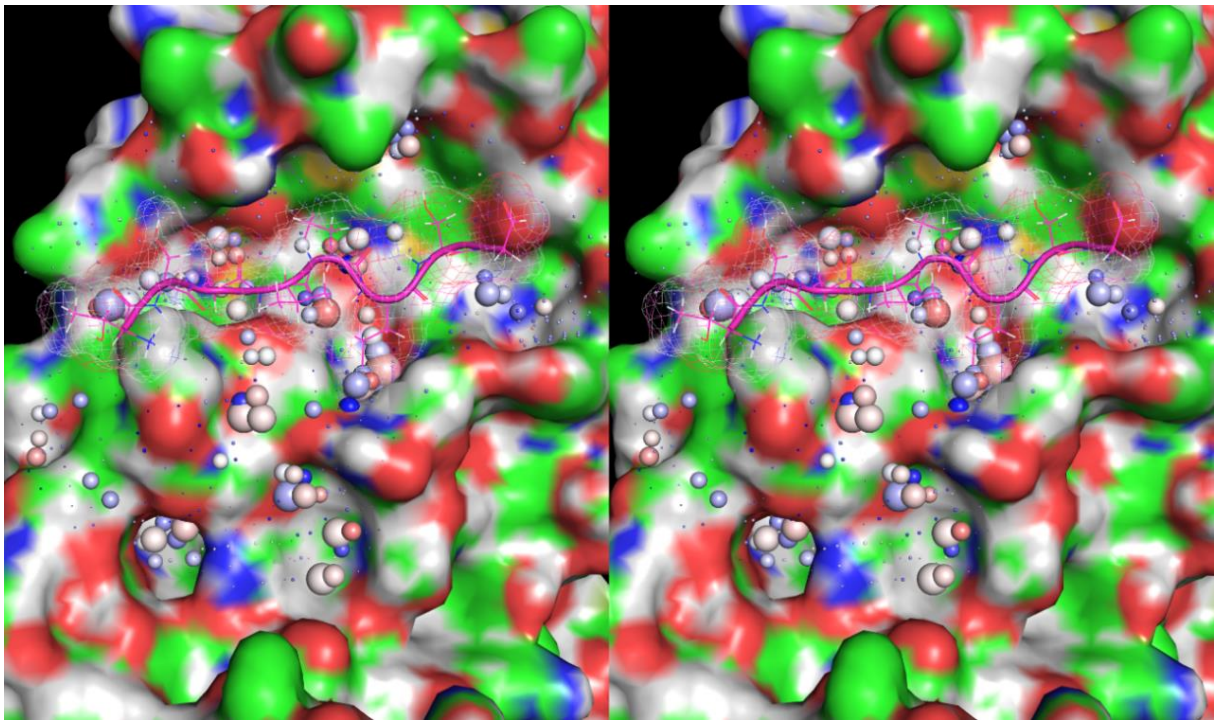
E-mail: robert.pearlstein@novartis.com

TABLE OF CONTENTS

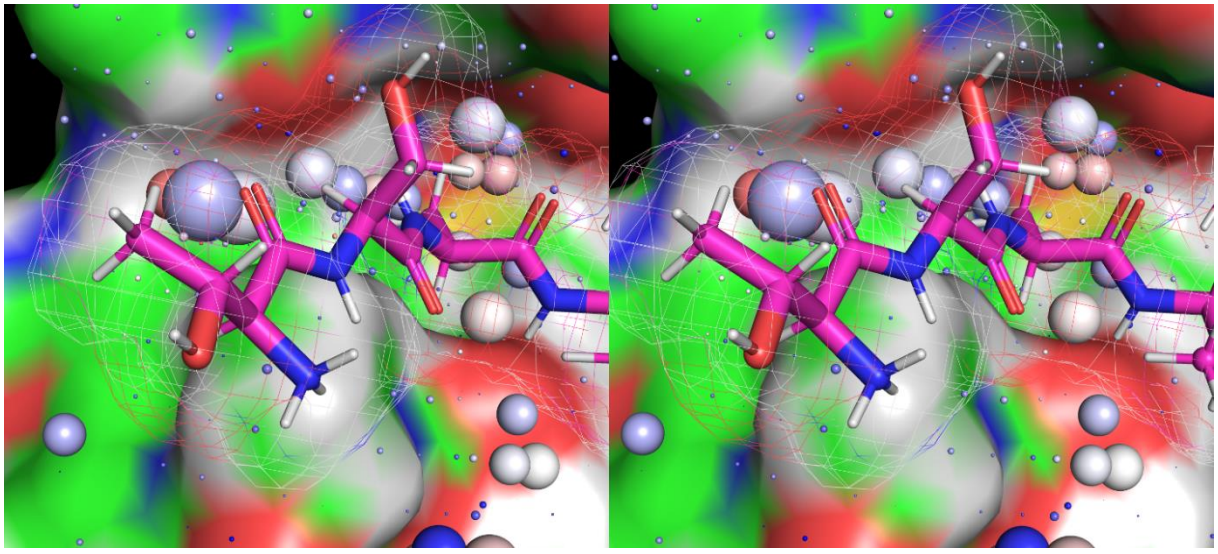
Figure S1. Stereo views of the substrate extracted from PDB code = 2Q6G (shown as a cartoon color-coded by sub-pocket) overlaid on the active site (AS) solvation structure of time-averaged 2QCY (shown as a surface, color-coded by element). HOV overlaps are defined as those falling within, on, or near the mesh surface of the substrate. Since induced-fit is not accounted for in the overlay, the spatial relationships between substrate groups and voxels is approximate (the tightness of which varies among the HOV clusters). In particular, we assume that the H-bonds of water occupying HOVs in reasonable proximity to polar substrate groups would be replaced to one degree or another by those groups, resulting in favorable de-solvation of such positions.



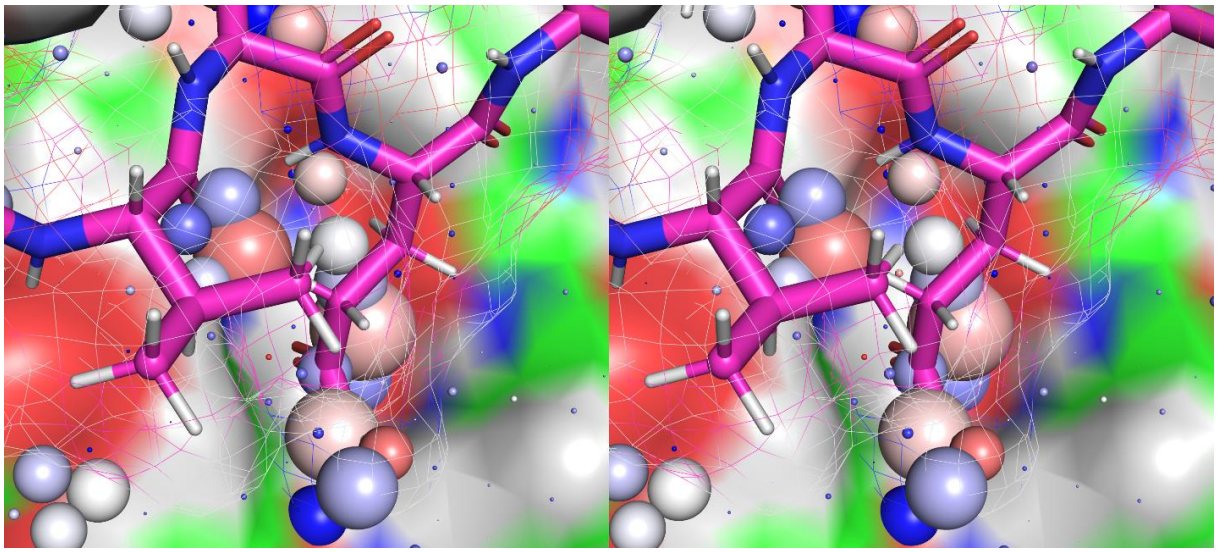
S1-A. Overlay of the substrate (shown with a mesh surface) and AS solvation structure. The HOVs, which serve as the gatekeepers of the pocket, effectively trace out the substrate backbone between the P1 and P4 residues, as well as the P1 side chain. ULOVs are distributed throughout the sub-pockets of the AS, most of which are expelled by substrates and inhibitors to varying degrees.



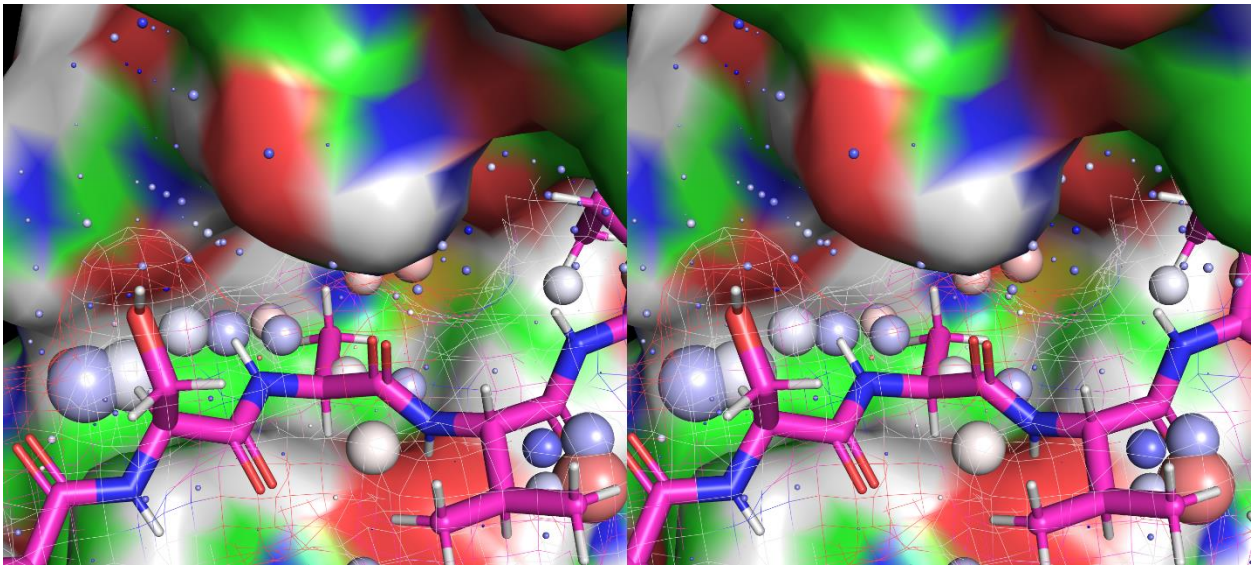
S1-B. Zoomed-out view of the substrate overlaid on the protein surface. The sub-pocket-substrate residue correspondences are denoted in Figure S1-A.



S1-C. Close-up view of the substrate N-terminus, which overlaps with several ULOVs, but no HOVs (predicted to slow k_{-1} of the substrate without affecting k_1).

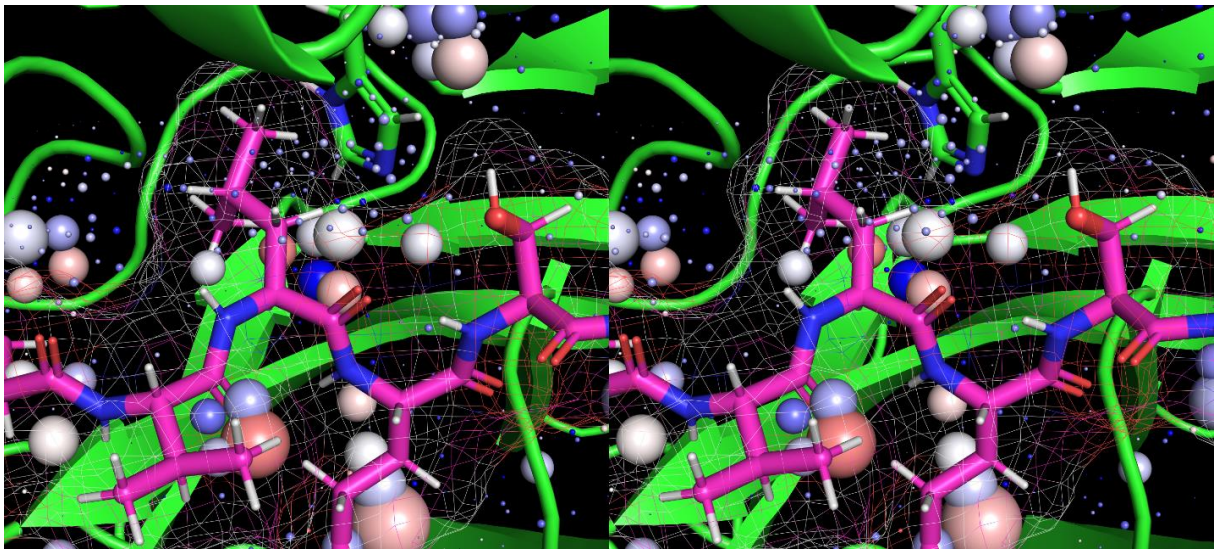


S1-D. Close-up view of the P3 (left) and P1 (right) residues, the intervening backbone of which forms a turn. The backbone O (left) overlaps favorably with a HOV cluster representing a single well-ordered water molecule (based on the strong acceptor/donor coloring of the voxels). The backbone NH overlaps favorably with a weaker HOV (faint pink). The P1 Gln amide overlaps with two HOV clusters representing one water molecule each. The amide NH2 group and amide O overlap with the lower and upper HOV cluster, respectively.



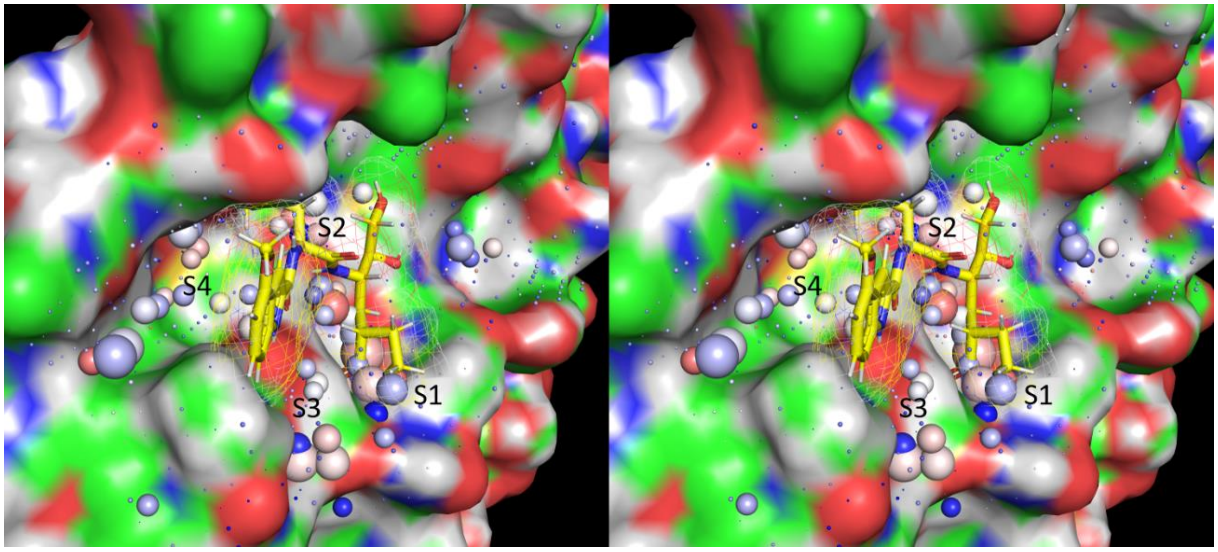
S1-E. Close-up view of the P3-P5 region. Three HOV clusters, representing three water molecules located at the rear of the AS, reside within the mesh surface of the substrate. The P4 Ala side chain

overlaps unfavorably with the first cluster at the rear of the S4 sub-pocket, the backbone NH in this region overlaps favorably with the second cluster, whereas the third cluster is not overlapped (i.e. the corresponding water molecule is possibly conserved in the bound state).

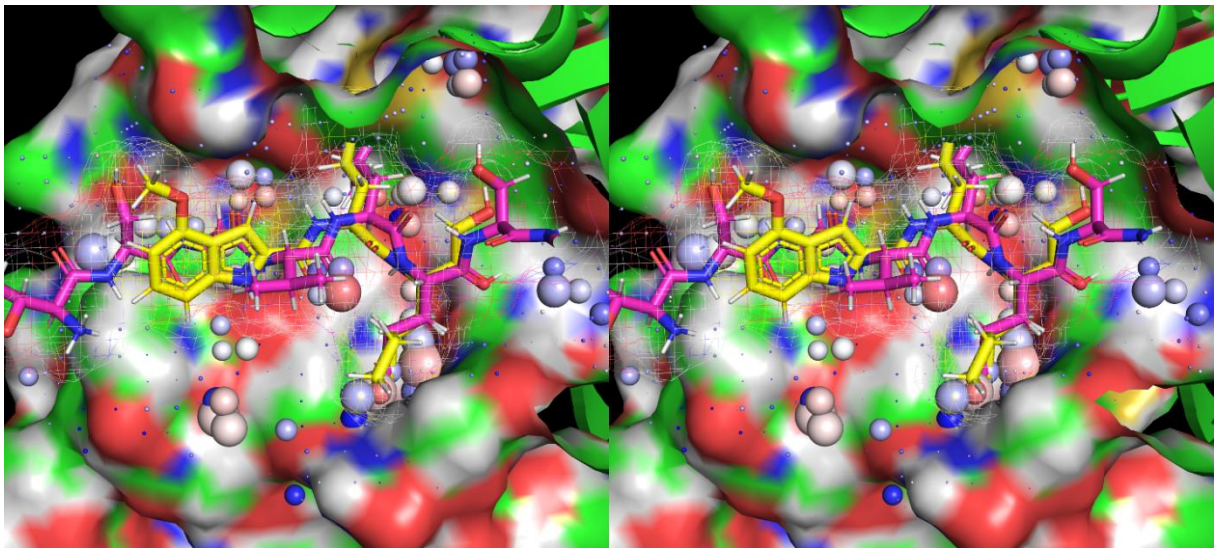


S1-F. Close-up view of the Leu P2 side chain within the S2 sub-pocket (noting that the substrate was crystallized in the inactive His41Ala mutant M^{pro} (2Q6G), the S2 sub-pocket of which necessarily differs from that of the wild type protein). The HOVs located just below His41 are quite possibly mis-registered with the P2 side chain in the overlay, and may instead overlap with the P2 backbone O in wild type M^{pro} (mirroring the HOV overlap with the backbone O of the P3 residue). The P2 Leu side chain overlaps with several ULOVs. Activation of the catalytic His has been attributed to de-solvation of the S2 sub-pocket by this side chain in NS3 protease (cited in the text).

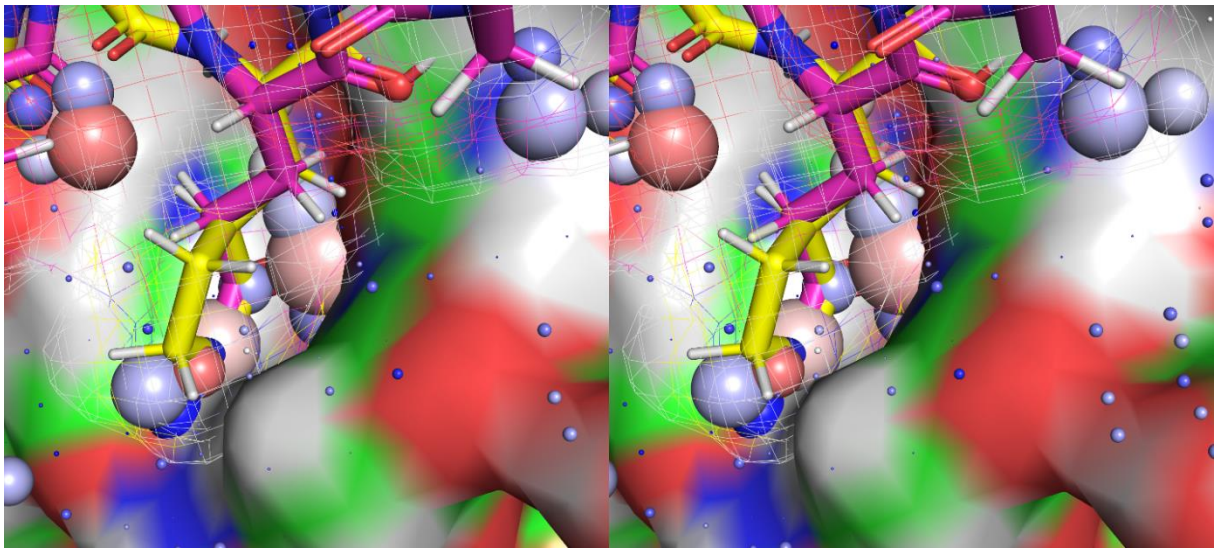
Figure S2. Same as Figure S1, except for PF00835321 (yellow) extracted from PDB code = 6XHM.



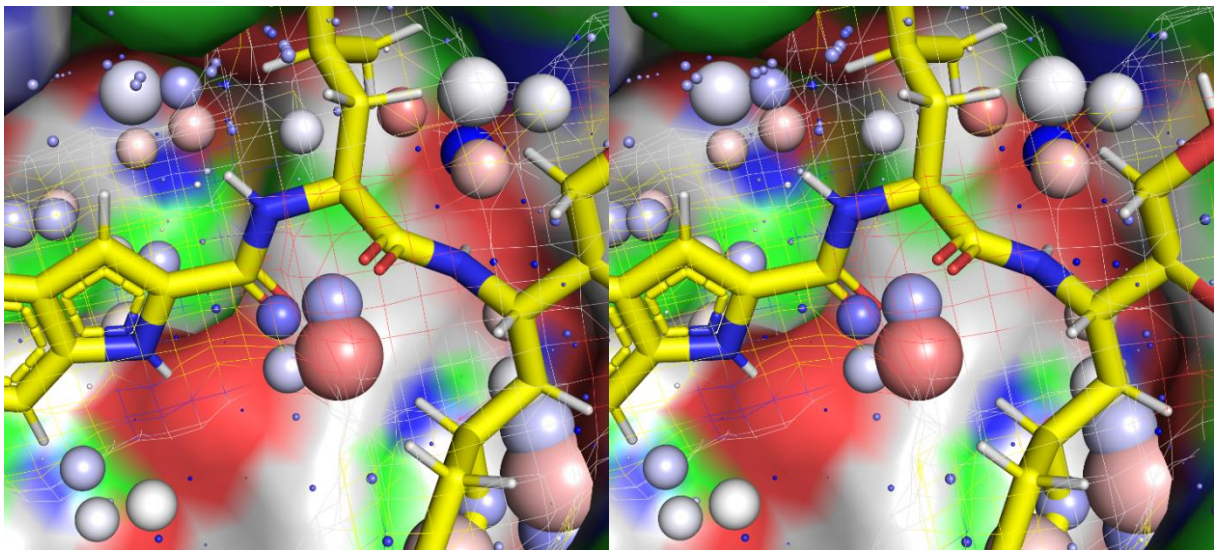
S2-A. Zoomed-out view of the M^{pro}-inhibitor-solvation structure overlay (AS sub-pockets S1-S4 are labeled).



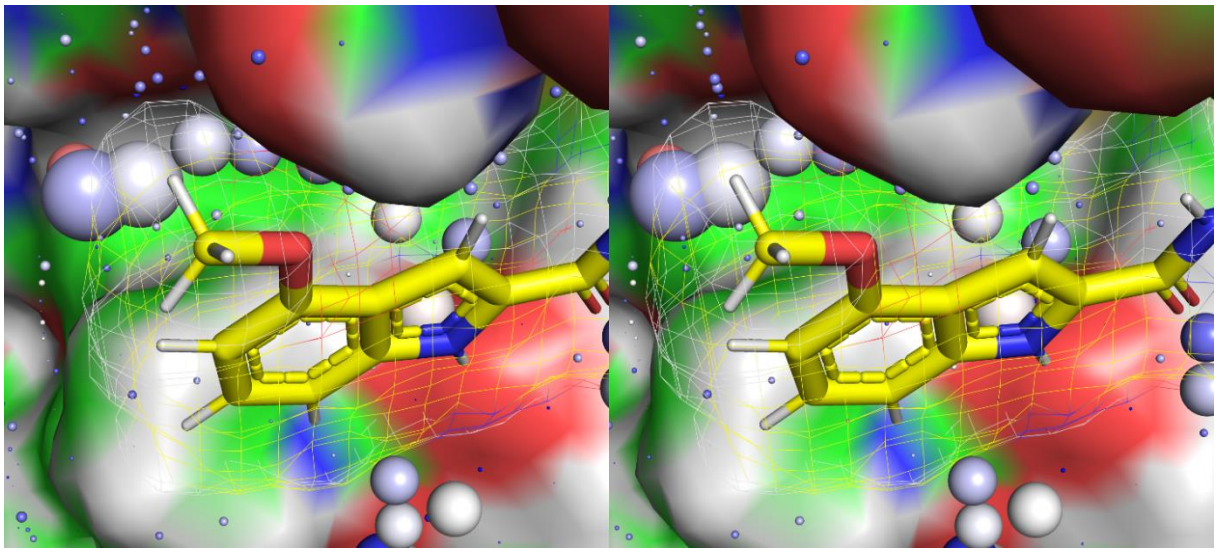
S2-B. Same as Figure S2-A, except with the substrate (magenta) included in the overlay for comparison.



S2-C. Close-up view of the S1 sub-pocket with the substrate P1 Gln included for comparison. The cyclic amide group of the inhibitor projects into this pocket. The binding geometry and HOV overlaps are highly similar for both molecules.



S2-D. Close-up view of the P2 and P3 regions. The two amide O groups of the inhibitor (corresponding to the backbone O groups of the substrate) overlap favorably with a HOV cluster corresponding to a single ordered water molecule (suggested by the relatively strong preference for H and O at well-defined positions).



S2-E. Close-up view of the N-terminal region of the inhibitor. The methoxyindole group of the inhibitor threads between HOV clusters in the S3 and S4 sub-pockets. The indole NH of the inhibitor overlaps favorably with a donor/acceptor-agnostic HOV cluster corresponding to a single water molecule.

Figure S3. Same as Figure S1, except for SID 24808289 (yellow) extracted from PDB code = 4MDS.

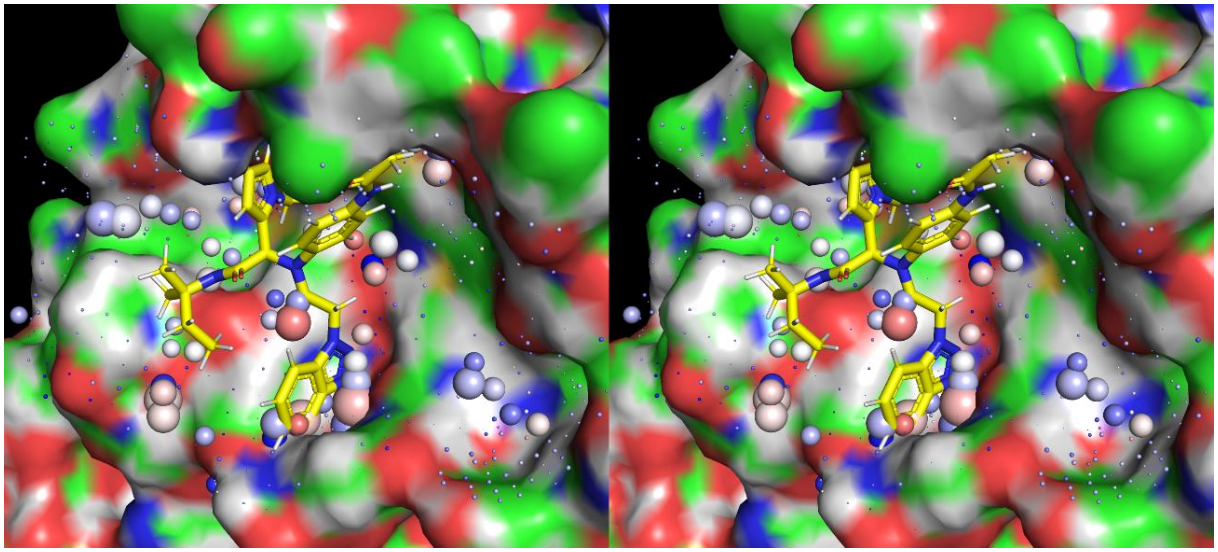


Figure S3-A. Zoomed-out view of the overlaid inhibitor (sub-pocket numbering shown in Figure S2-A). Entropic benefits may be realized by connecting the P1 and P3 substituents.

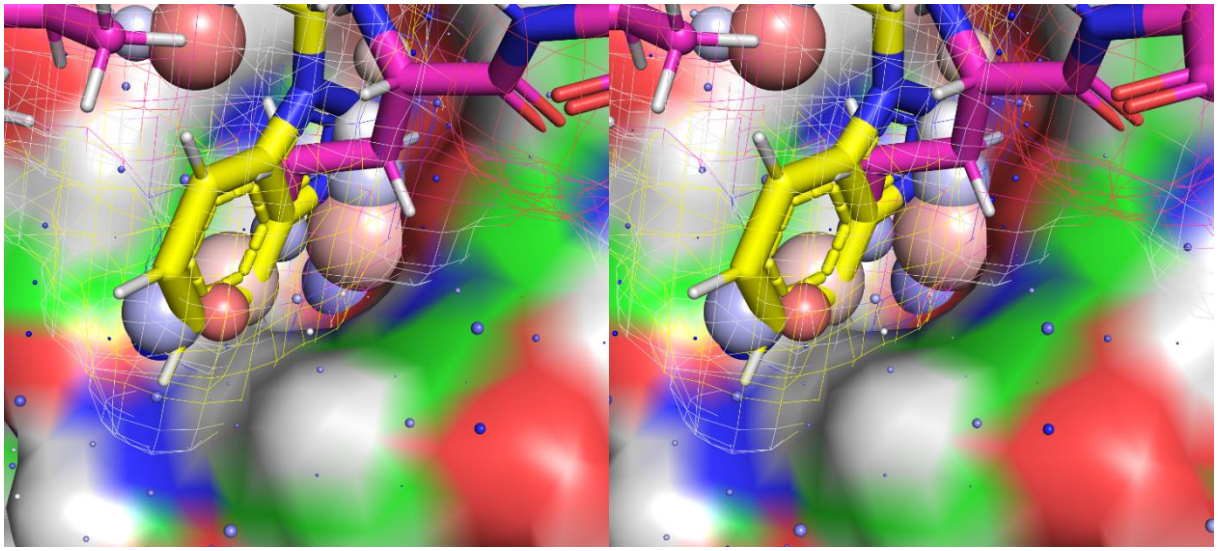


Figure S3-B. Close-up view of the S1 sub-pocket, showing the binding modes of the benzotriazole group of the inhibitor compared with the substrate Gln. The triazole nitrogens and aromatic C overlap favorably and unfavorably with the two HOV clusters in this sub-pocket, respectively. The

benzotriazole group lacks an H-bond donor at the lower HOV position, which instead is occupied by an aromatic C. Replacing this group with a donor/acceptor group is predicted to speed k_{on}.

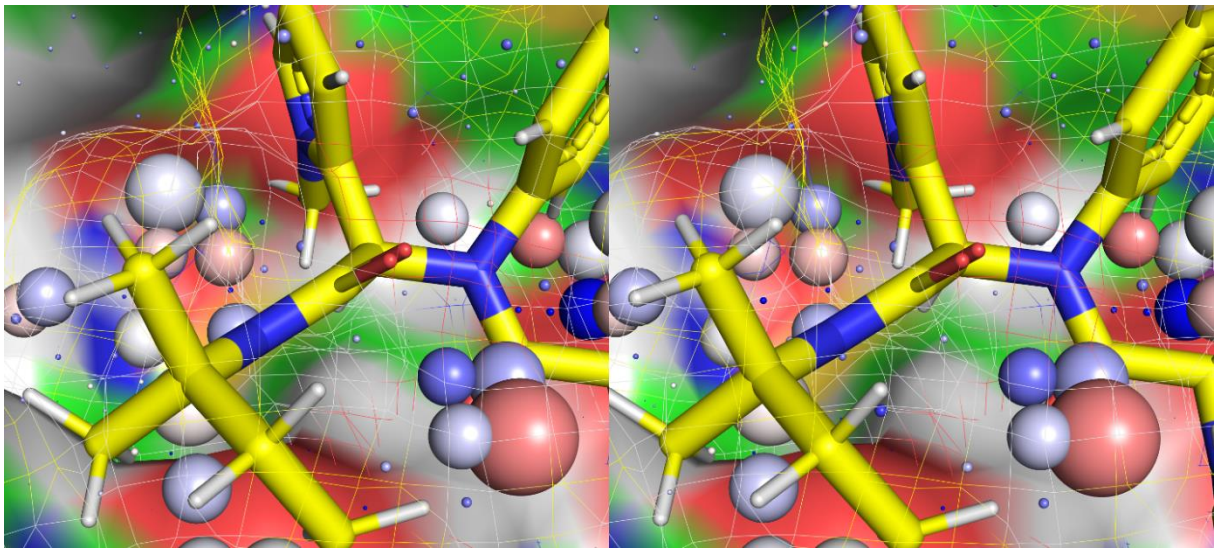


Figure S3-C. Close-up view of the S3 sub-pocket. As for the substrate P3 side chain, the isopentyl and methylpyrazole substituents overlap unfavorably with HOVs in the S3 and S4 sub-pockets. Replacement of these groups with more polar groups is predicted to speed k_{on}. The amide O of the inhibitor on the right (corresponding to the P2 amide O of the substrate) and the amide NH of the inhibitor on the left (corresponding to the NH at the P3 position of the substrate) overlaps favorably with HOV clusters in those positions.

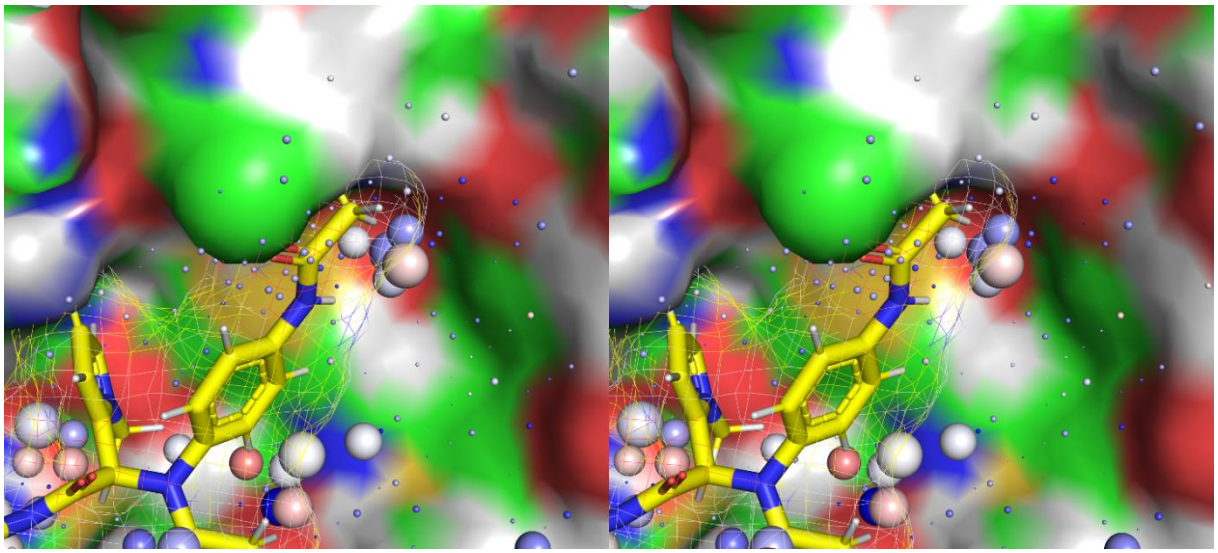


Figure S3-D. Close-up view of the S1' sub-pocket. The methylamide substituent of the inhibitor projects into this sub-pocket, which overlaps unfavorably with two HOV clusters. The phenyl analog (17a) of the inhibitor is predicted to relieve the unfavorable overlap with the upper HOV cluster. A pyridyl or polar substituent at the 2-position of the aryl ring may relieve the second unfavorable overlap with the lower HOV cluster and reduce the induced-fit requirement.

Figure S4. Same as Figure S1, except for boceprevir (yellow) extracted from PDB code = 6WNP.

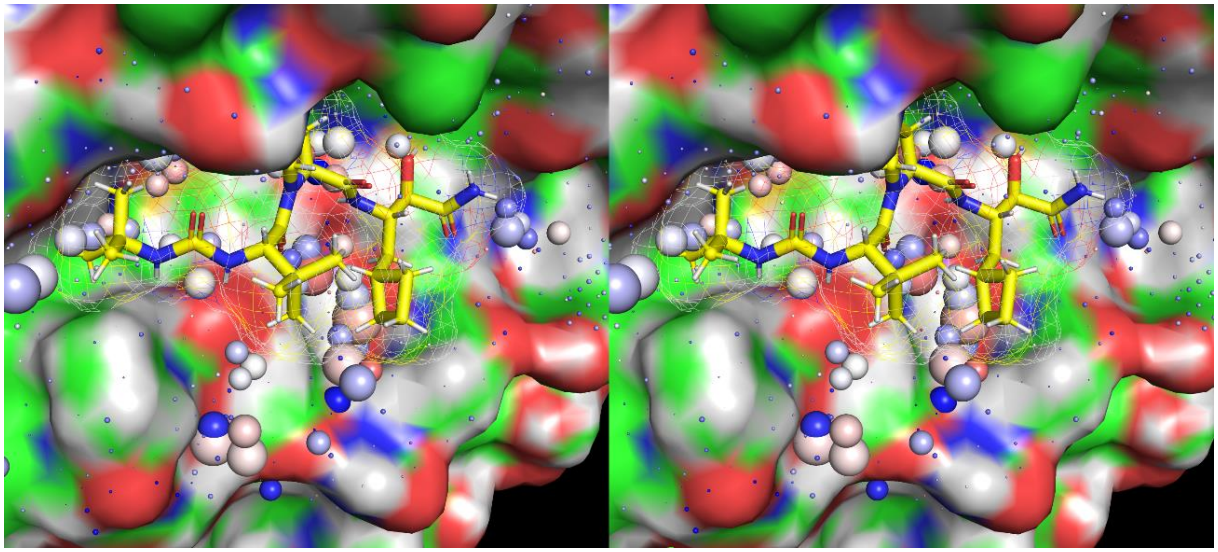


Figure S4-A. Zoomed-out view of the overlaid inhibitor (sub-pocket numbering shown in Figure S2-A).

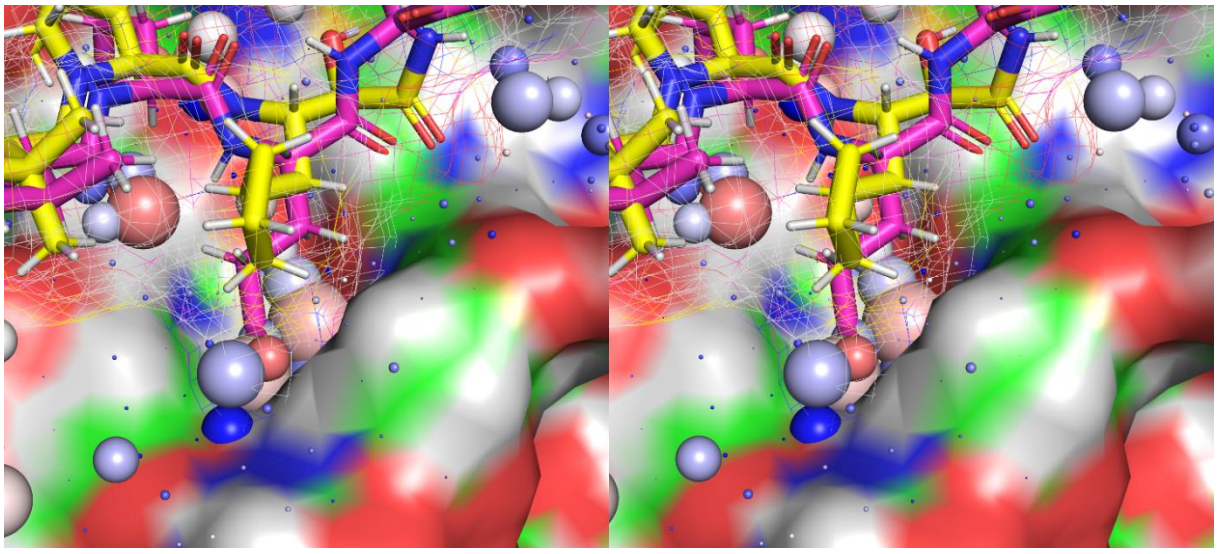


Figure S4-B. Close-up view of the S1 sub-pocket, showing the binding modes of the methyl-cyclobutyl group of the inhibitor compared with the substrate Gln. The inhibitor group resides outside of this sub-pocket, anterior to the Gln side chain of the substrate.

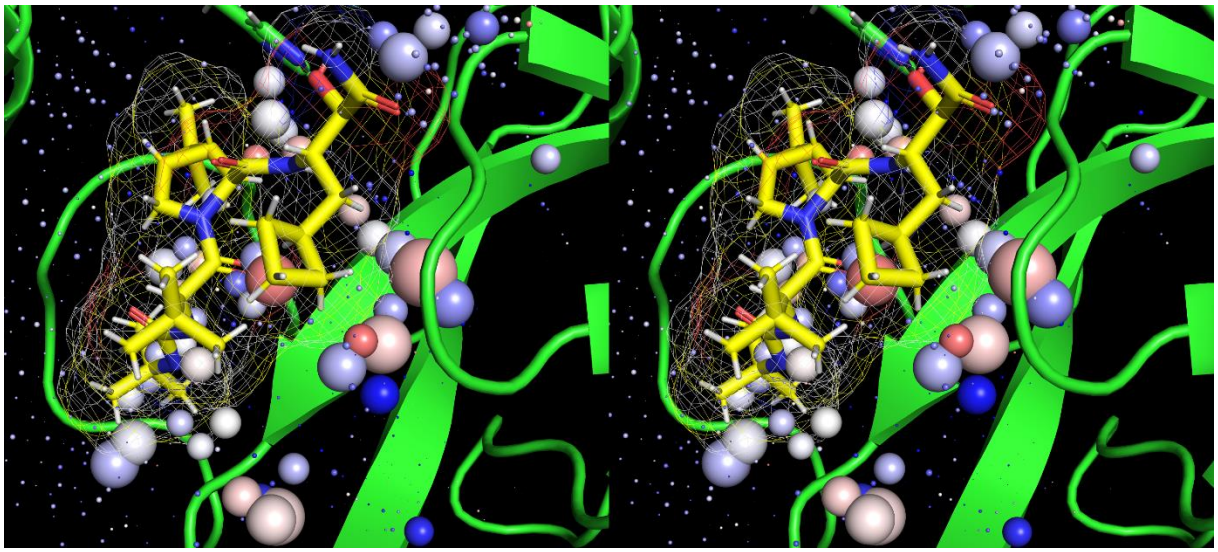


Figure S4-C. H-bond enriched water occupying the two HOV clusters in the S1 sub-pocket is trapped unfavorably behind the methylcyclobutyl group of the bound inhibitor.

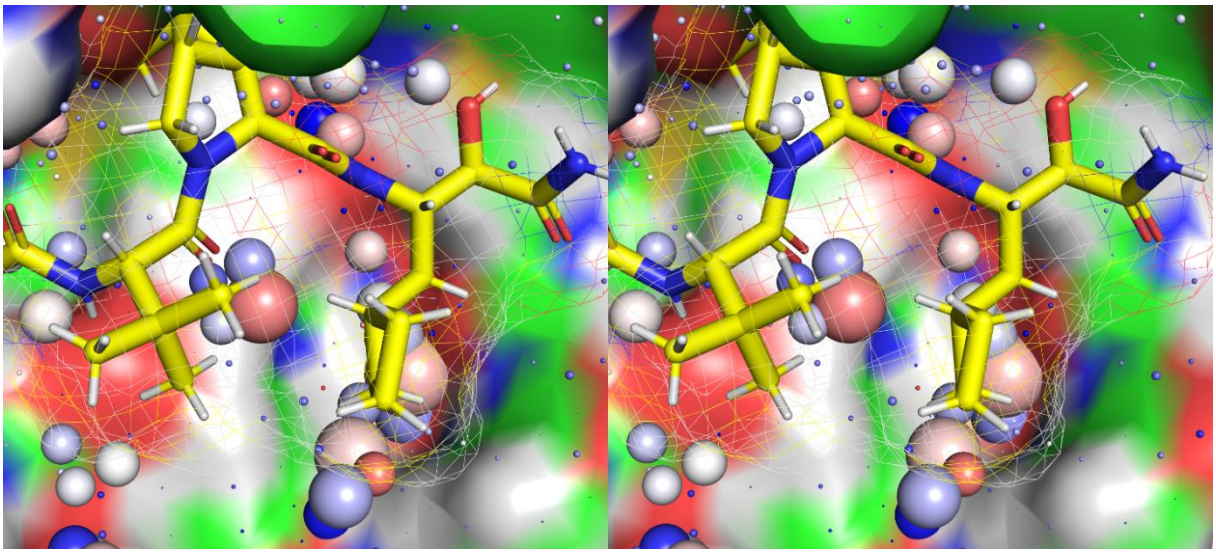


Figure S4-D. Close-up view of the S3 pocket. As for the substrate P3 side chain, the t-butyl substituent of the inhibitor overlaps unfavorably with the HOV cluster in the S3 sub-pocket. The amide O of the inhibitor on the right (corresponding to the P2 amide O of the substrate) and the amide NH of the inhibitor on the left (corresponding to the NH at the P3 position of the substrate) overlaps favorably with HOV clusters in those positions.

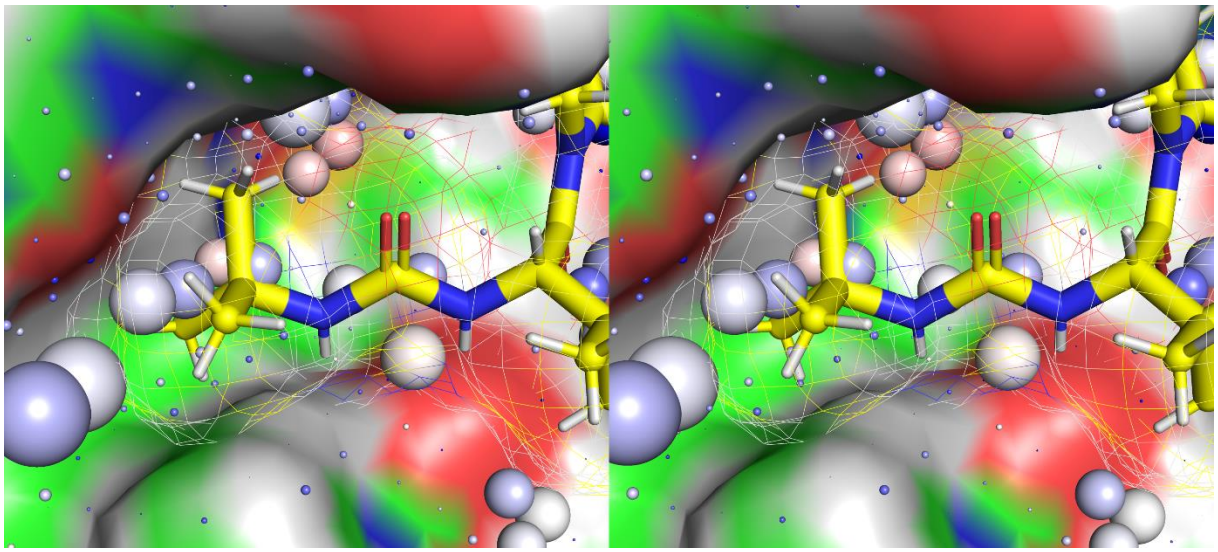


Figure S4-D. Close-up view of the N-terminal region of the inhibitor. The urea group of the inhibitor mimics the substrate backbone peptide group spanning between the P3 and P4 sub-pockets. The terminal t-butyl group of the inhibitor overlaps unfavorably with the HOV cluster in this region (in contrast with PF00835321, which does not).

Figure S5. Same as Figure S1, except for N3 (yellow) extracted from PDB code = 6LU7.

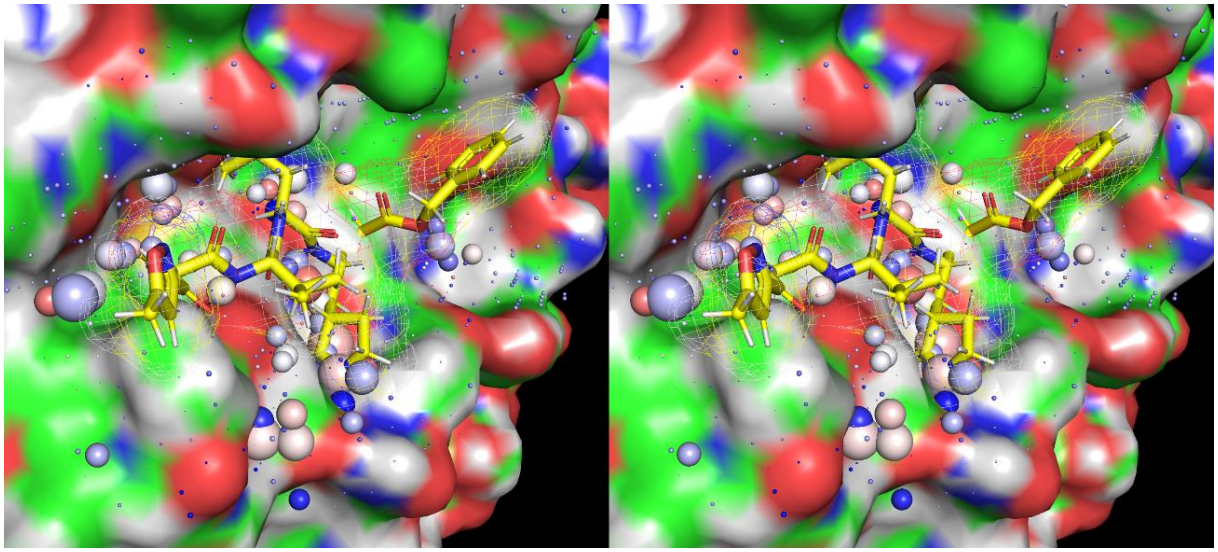


Figure S5-A. Zoomed-out view of the overlaid inhibitor (sub-pocket numbering shown in Figure S2-A).

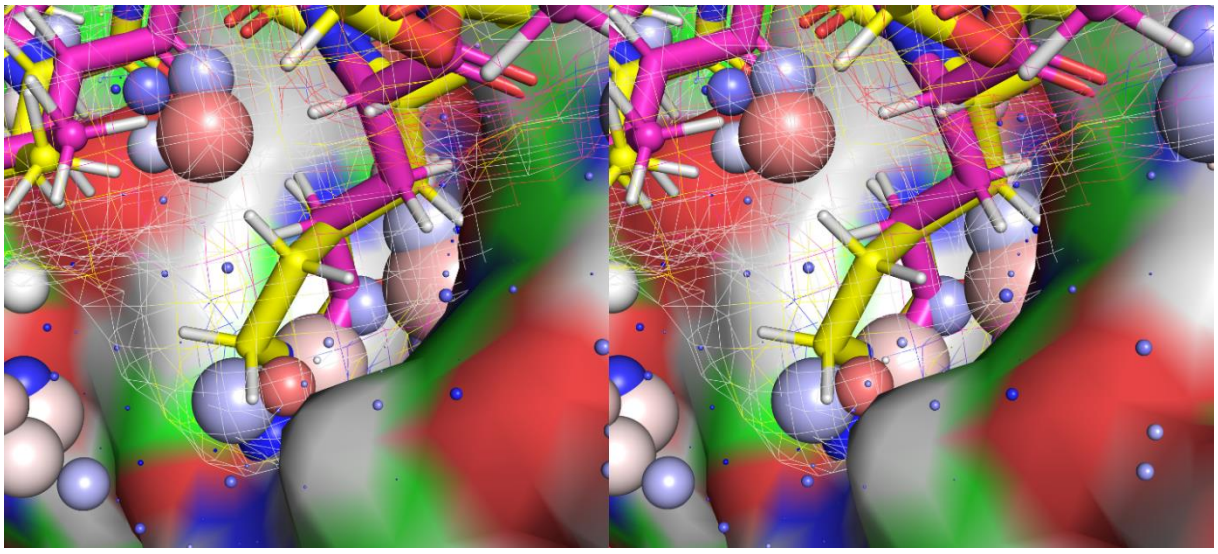


Figure S5-B. Close-up view of the S1 sub-pocket, showing the binding modes of the substrate P1 side chain compared with the methyl-cyclobutyl group of the inhibitor.

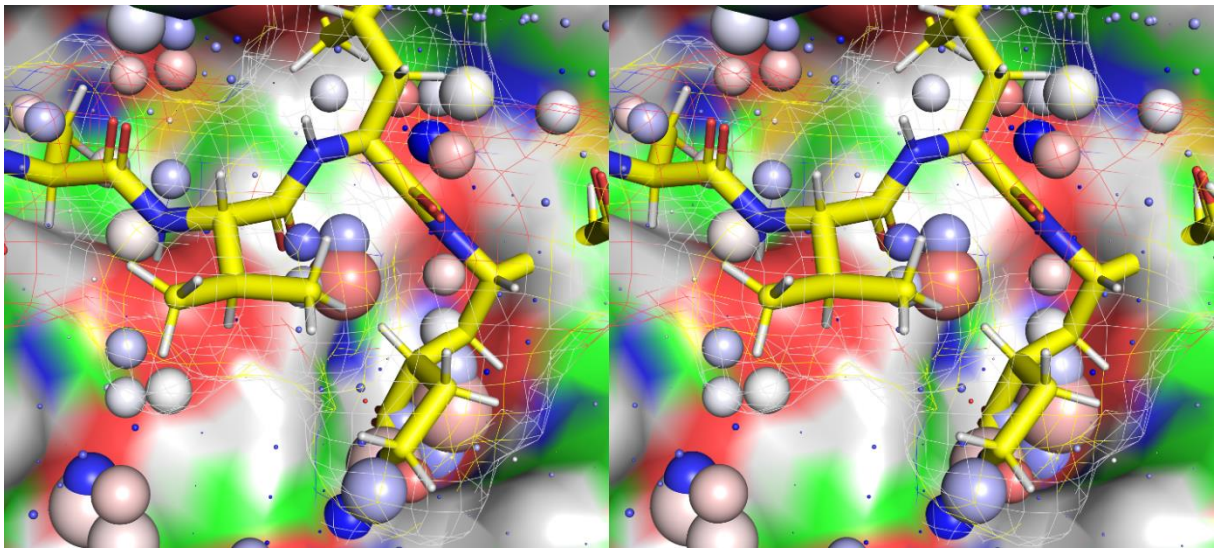


Figure S5-C. Close-up view of the S3 and S4 sub-pockets. As for the substrate P3 side chain, the isopropyl and methyl substituents of the inhibitor overlap unfavorably with the HOV clusters in the S3 and sS4 sub-pockets, whereas the amide O overlaps favorably with the HOV at the P3 backbone O position. The amide NH group of the inhibitor overlaps favorably with a HOV at the substrate P3 NH position.

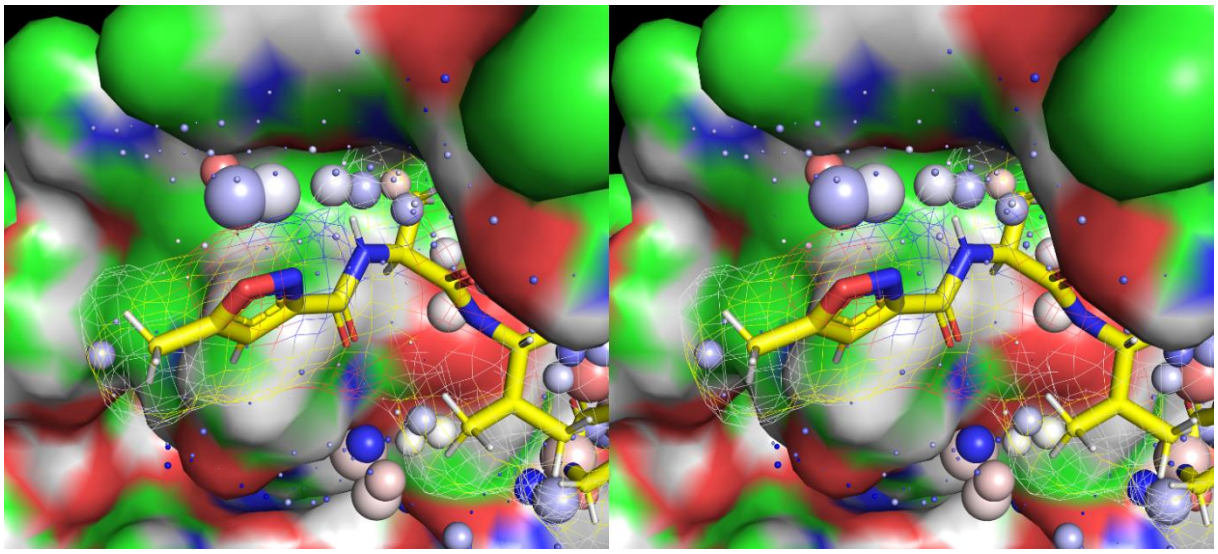


Figure S5-D. Close-up view of the N-terminal region of the inhibitor. The amide NH group overlaps favorably with the adjacent HOV cluster at the rear of the pocket.

Figure S6: 2QCY

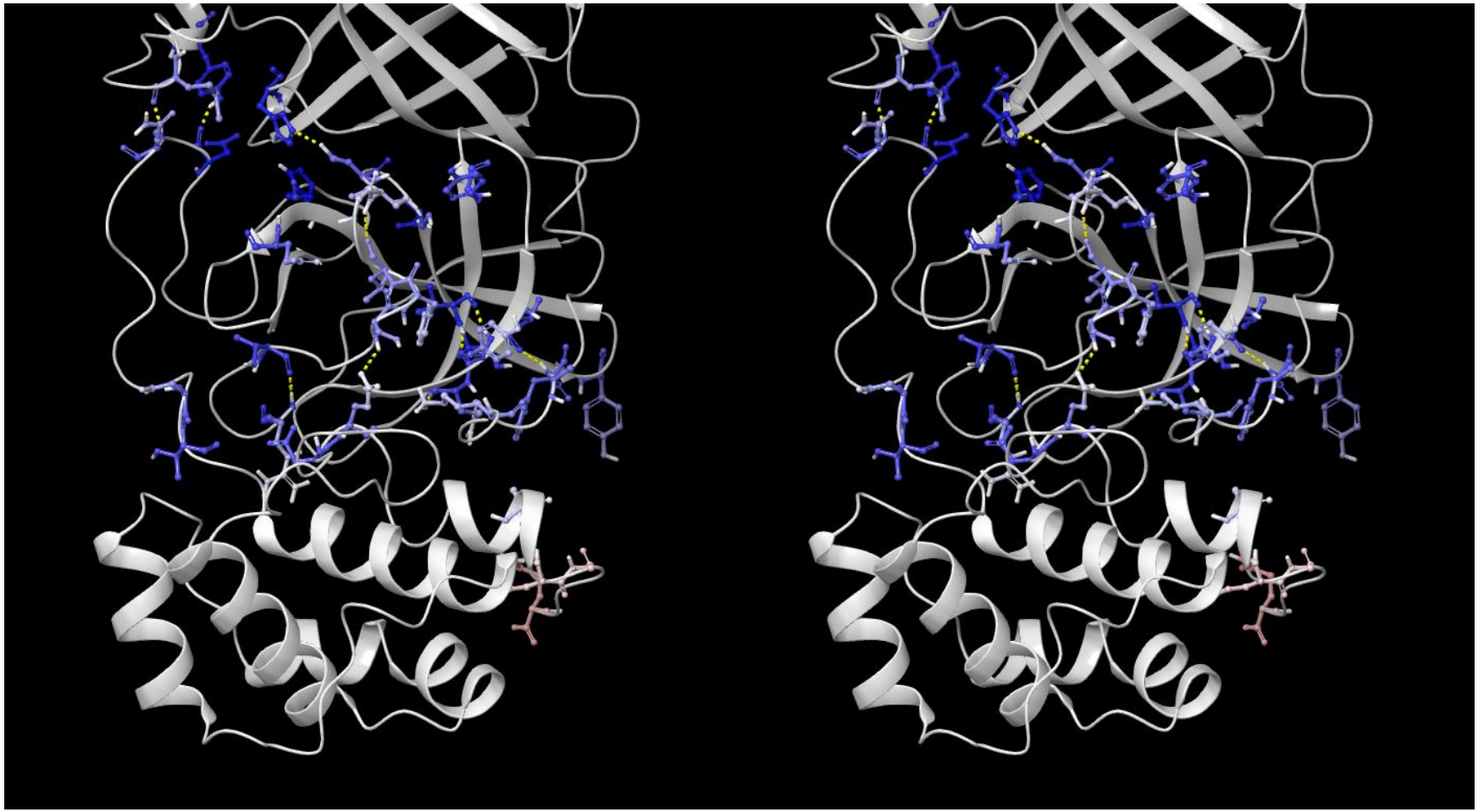


Figure S7: 6WNP

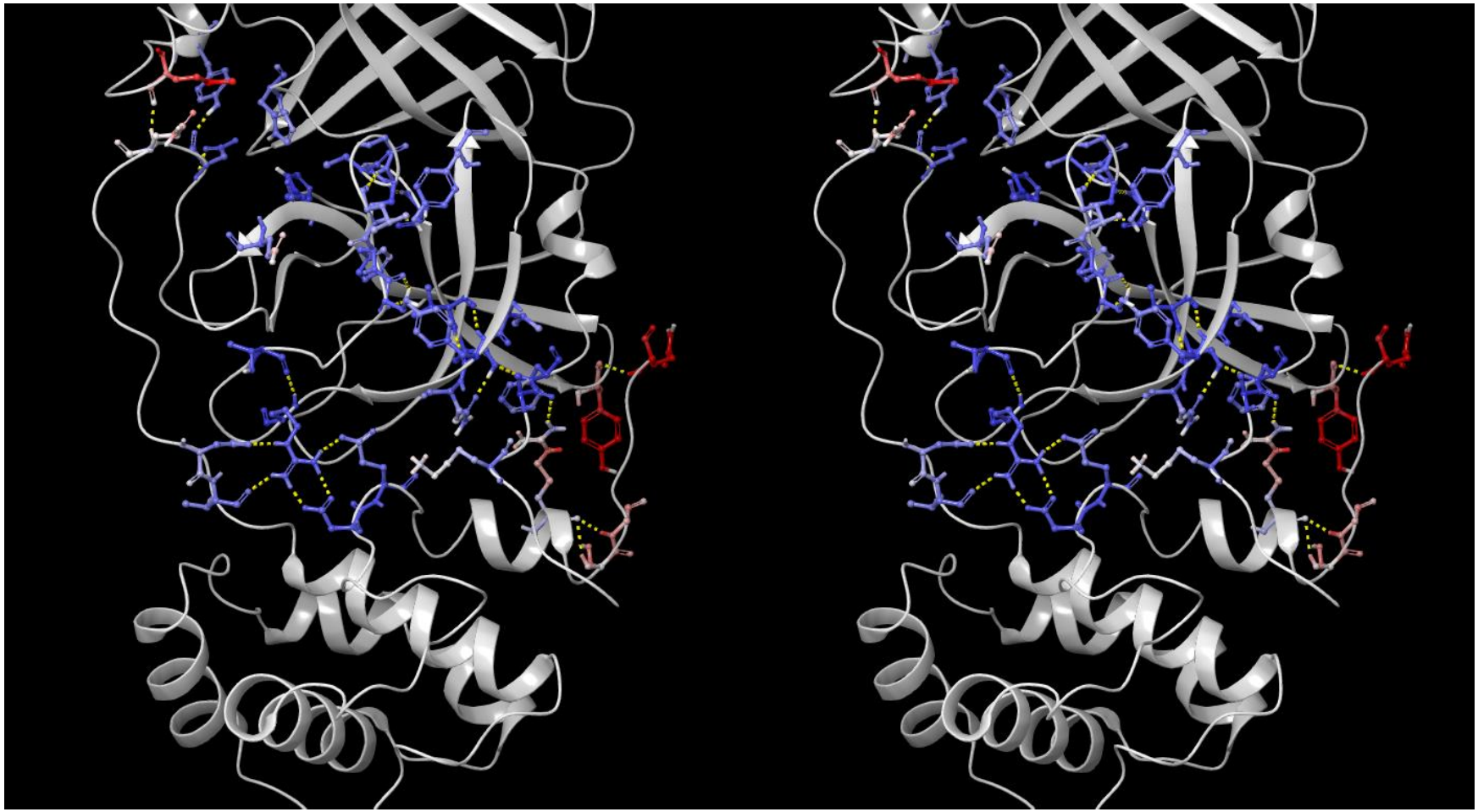


Figure S8: 6M03

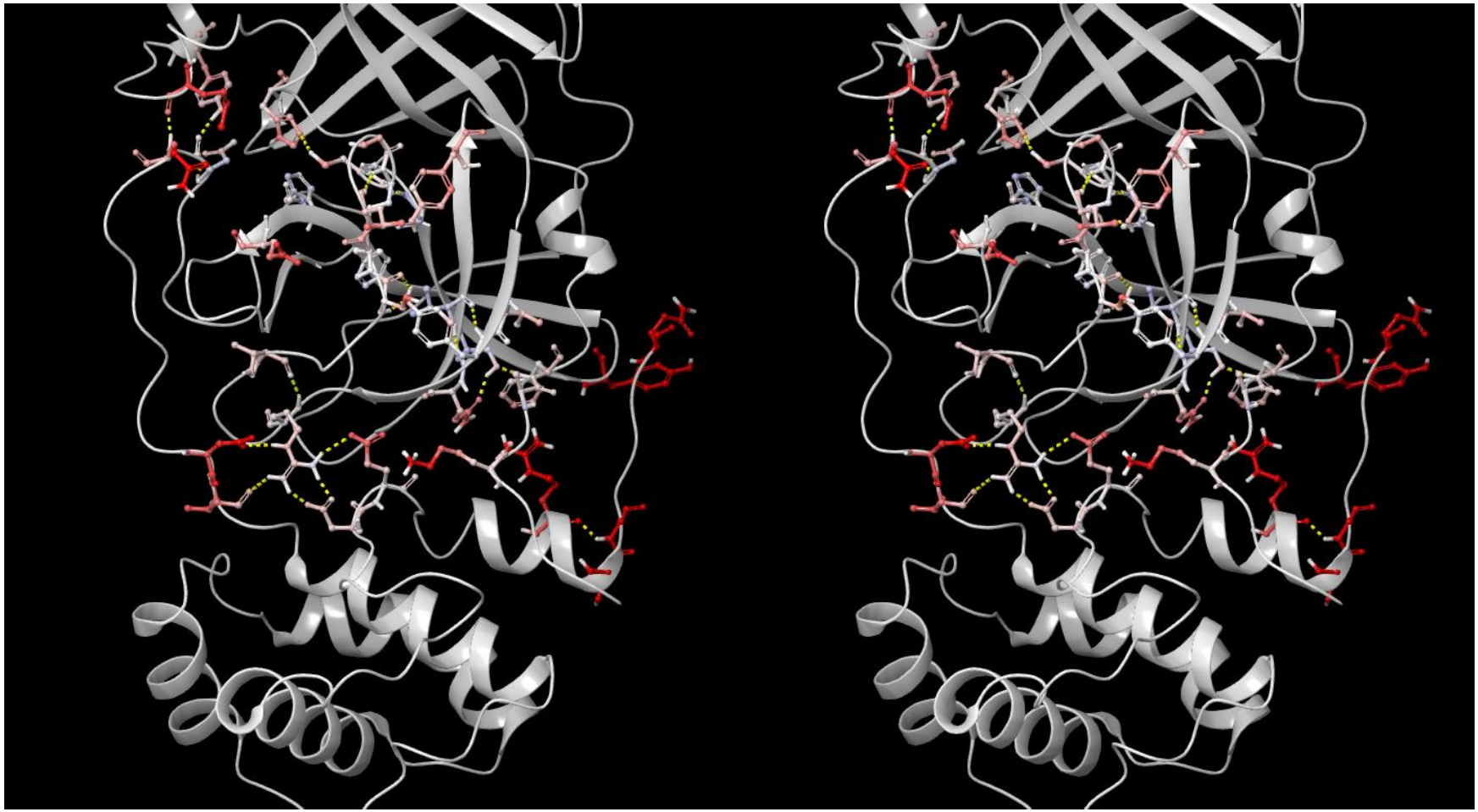


Figure S9: 2QCY

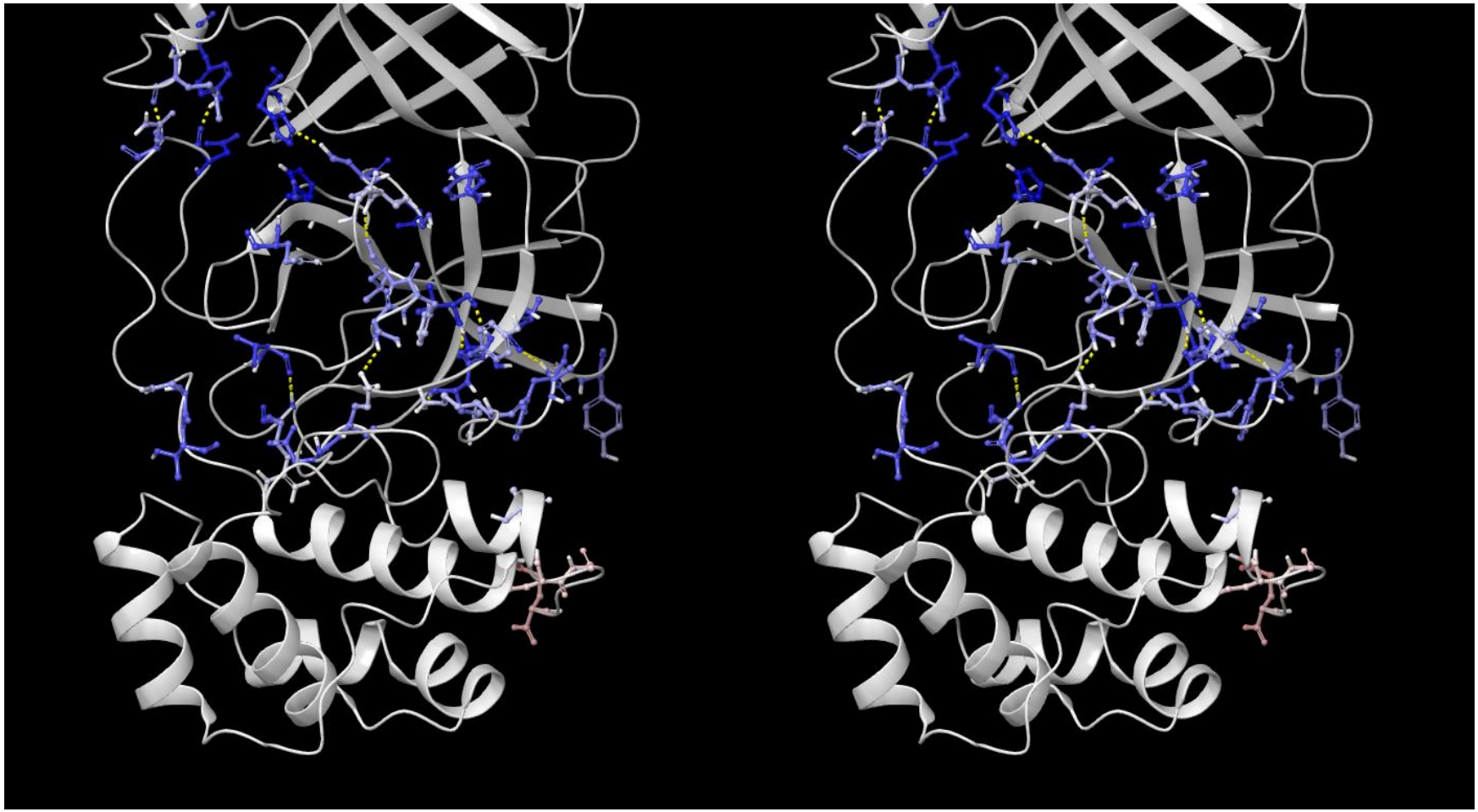


Figure S10: 2BX3

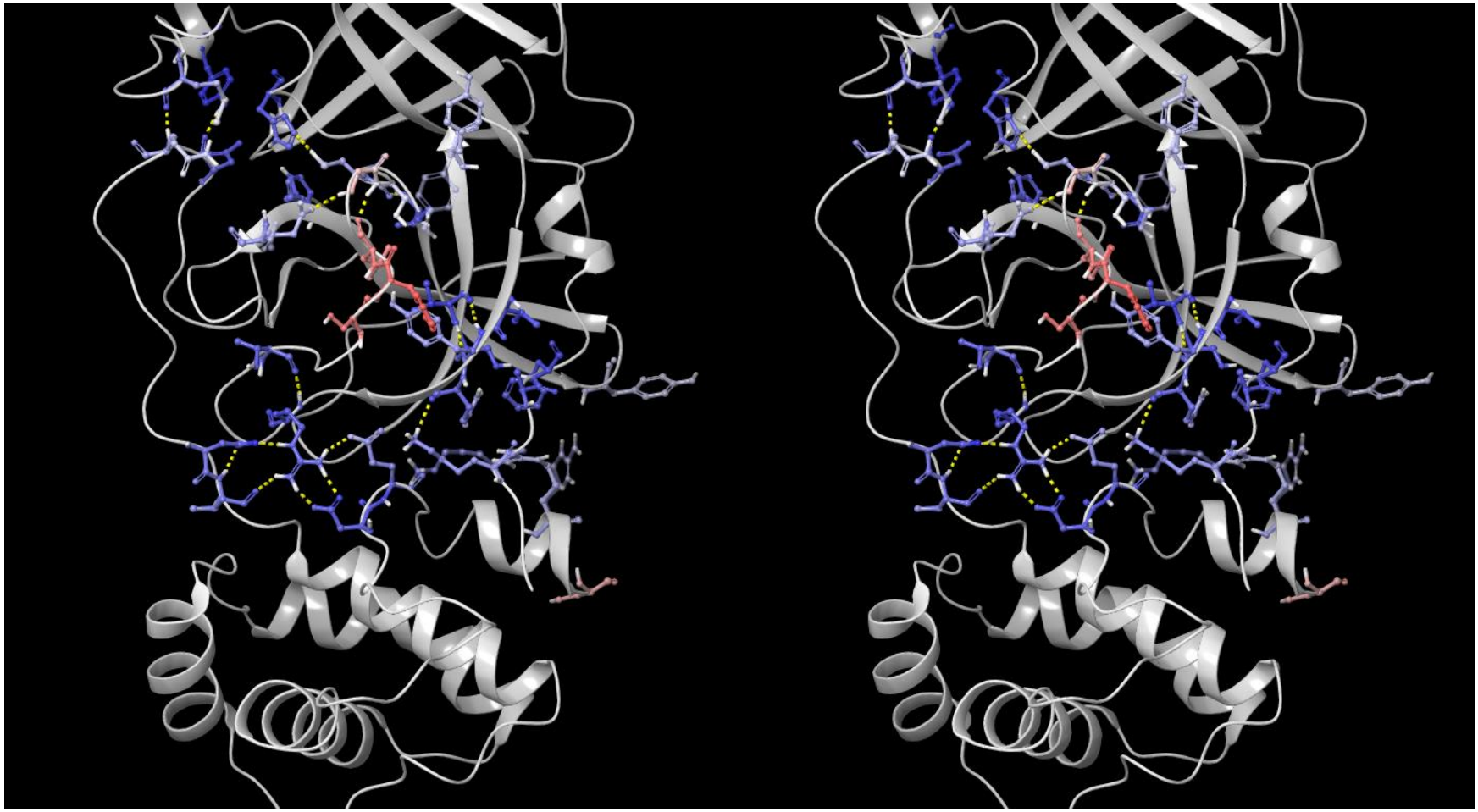


Figure S11: 2Q6G

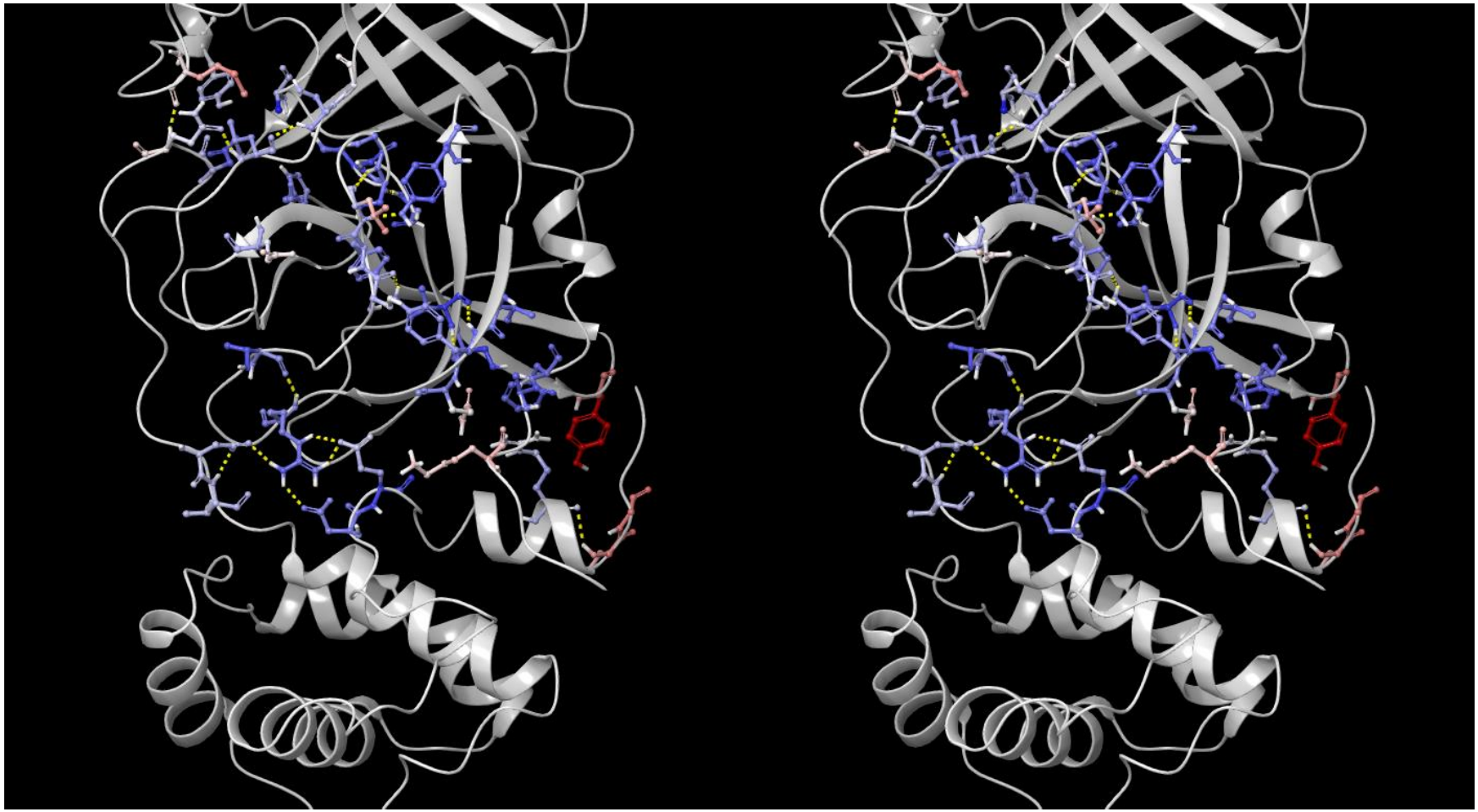


Figure S12: 6LU7