Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information

High-throughput Designing of Symmetrical Dimeric SARS-CoV-2 Main Protease: Structural and Physical Insights into Hotspots for Adaptation and Therapeutics

Aditya K. Padhi[1*], and Timir Tripathi[2*]

[1]Laboratory for Structural Bioinformatics, Center for Biosystems Dynamics Research, RIKEN, Yokohama, Kanagawa 230-0045, Japan

[2]Molecular and Structural Biophysics Laboratory, Department of Biochemistry, North-Eastern Hill University, Shillong- 793022, India

***Corresponding authors:**

adityapadhi.iitd@gmail.com

timir.tripathi@gmail.com

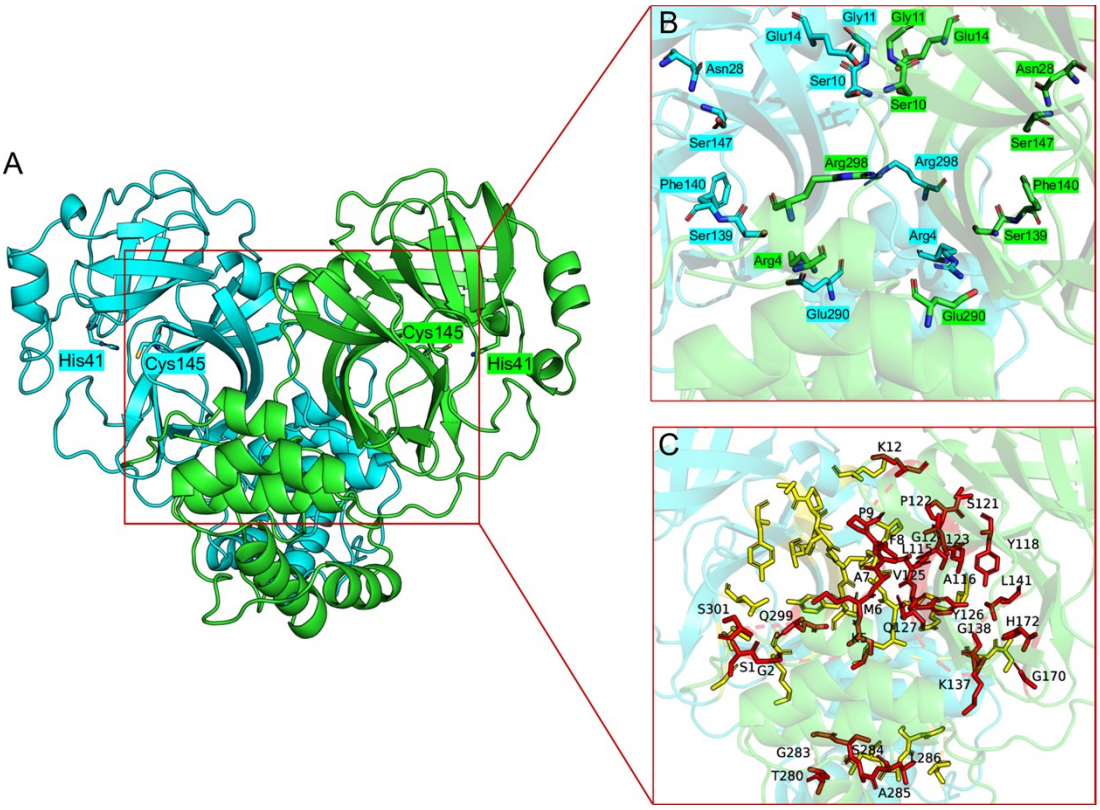

Fig. S1. Structure of the SARS-CoV-2 M^{pro} in dimeric form and crucial residues. (A) The dimeric structure of SARS-CoV-2 M^{pro} is shown as a cartoon, where the catalytic dyad residues (His41 and Cys145) are shown as sticks. In (B), the dimerization residues of M^{pro} are labeled and shown as sticks. In (C), the residues (Ser1, Gly2, Lys5, Met6, Ala7, Phe8, Pro9, Lys12, Leu115, Ala116, Tyr118, Ser121, Pro122, Ser123, Gly124, Val125, Tyr126, Gln127, Lys137, Gly138, Leu141, Gly170, His172, Thr280, Gly283, Ser284, Ala285, Leu286, Gln299, Ser301) that are part of the M^{pro}'s dimeric interface are labeled and shown as sticks.

Table S1. Strategies for symmetrical dimer design of M^{pro} and information of key residues involved in various functions of M^{pro}.

Design experiment 1 (With dimerization residues) *		
Residues	**Important function in M^{pro}**	**Design status**
His41, Cys145	Catalytic dyad	Not designed
Arg4, Ser10, Gly11, Glu14, Asn28, Ser139, Phe140, Ser147, Glu290, Arg298	For dimerization	Designed
Ser1, Gly2, Lys5, Met6, Ala7, Phe8, Pro9, Lys12, Leu115, Ala116, Tyr118, Ser121, Pro122, Ser123, Gly124, Val125, Tyr126, Gln127, Lys137, Gly138, Leu141, Gly170, His172, Thr280, Gly283, Ser284, Ala285, Leu286, Gln299, Ser301	Interface residues	Designed
Design experiment 2 (Without dimerization residues) *		
Residues	**Important function in M^{pro}**	**Design status**
His41, Cys145	Catalytic dyad	Not designed
Arg4, Ser10, Gly11, Glu14, Asn28, Ser139, Phe140, Ser147, Glu290, Arg298	For dimerization	Not designed
Ser1, Gly2, Lys5, Met6, Ala7, Phe8, Pro9, Lys12, Leu115, Ala116, Tyr118, Ser121, Pro122, Ser123, Gly124, Val125, Tyr126, Gln127, Lys137, Gly138, Leu141, Gly170, His172, Thr280, Gly283, Ser284, Ala285, Leu286, Gln299, Ser301	Interface residues	Designed
Design experiment 3 (7-residue designing) *		
Residues	**Important function in M^{pro}**	**Design status**
His41, Cys145	Catalytic dyad	Not designed
Arg4, Ser10, Gly11, Glu14, Asn28, Ser139, Ser147, Glu290, Arg298	For dimerization	Not designed
Met6, Phe8, Leu115, Tyr118, Val125, Tyr126, Phe140	Interface residues with relatively lower stability	Designed

*Residues of M^{pro} that are involved in substrate-binding (Met49, Gly143, Ser144, His163, His164, Met165, Glu166, Leu167, Asp187, Arg188, Gln189, Thr190, Ala191, Gln192) are not designed as they are not part of the dimeric interface and necessary for M^{pro}'s function.

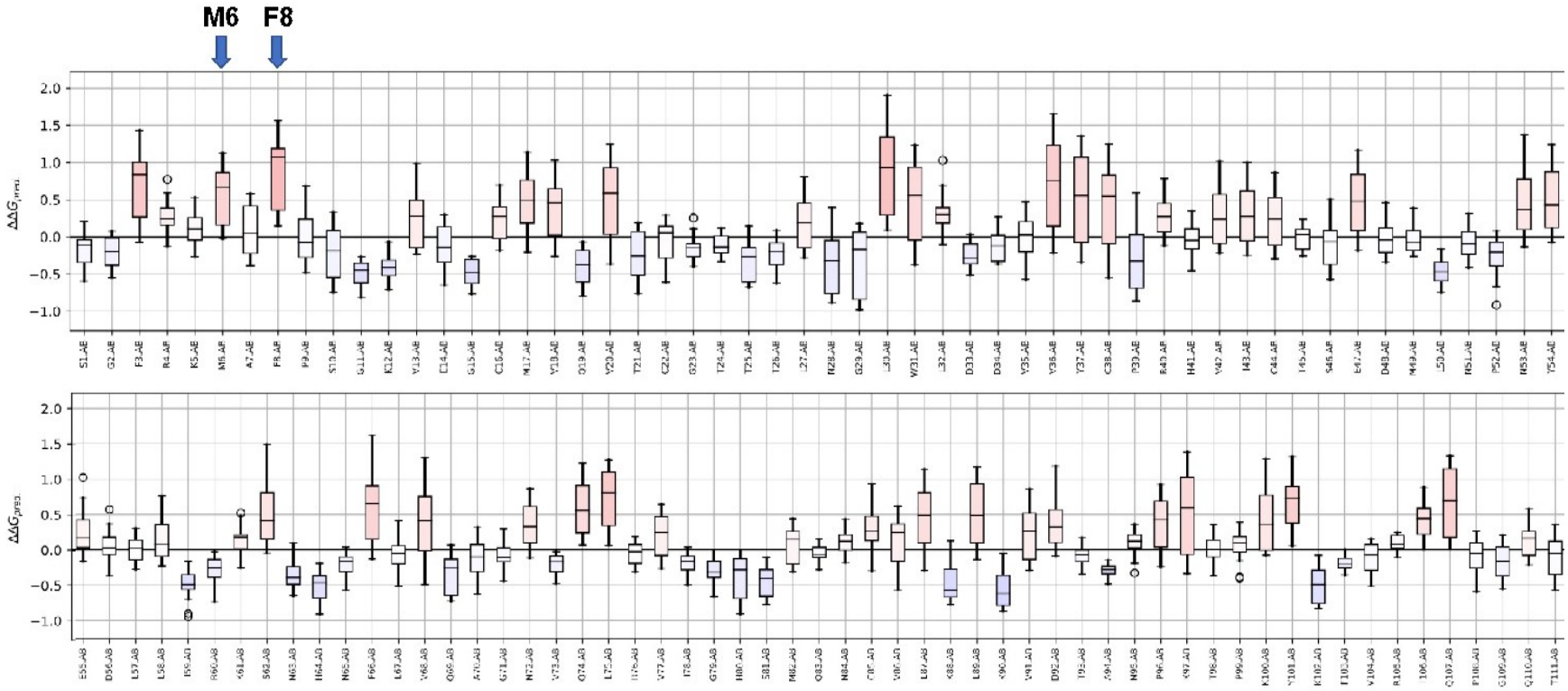

Fig. S2. Mutation sensitivity profile of dimeric native M^{pro}. The mutation sensitivity profile of dimeric native M^{pro} from Ser1 to Thr11 is shown, where the total predicted change in stability is determined as $\triangle\triangle G_{pred}$ ($\triangle\triangle G_{pred}<0$ indicates stabilizing and $\triangle\triangle G_{pred}>0$ indicates destabilizing mutation). M6 and F8 indicate destabilizing residues at the interface of dimeric M^{pro}.

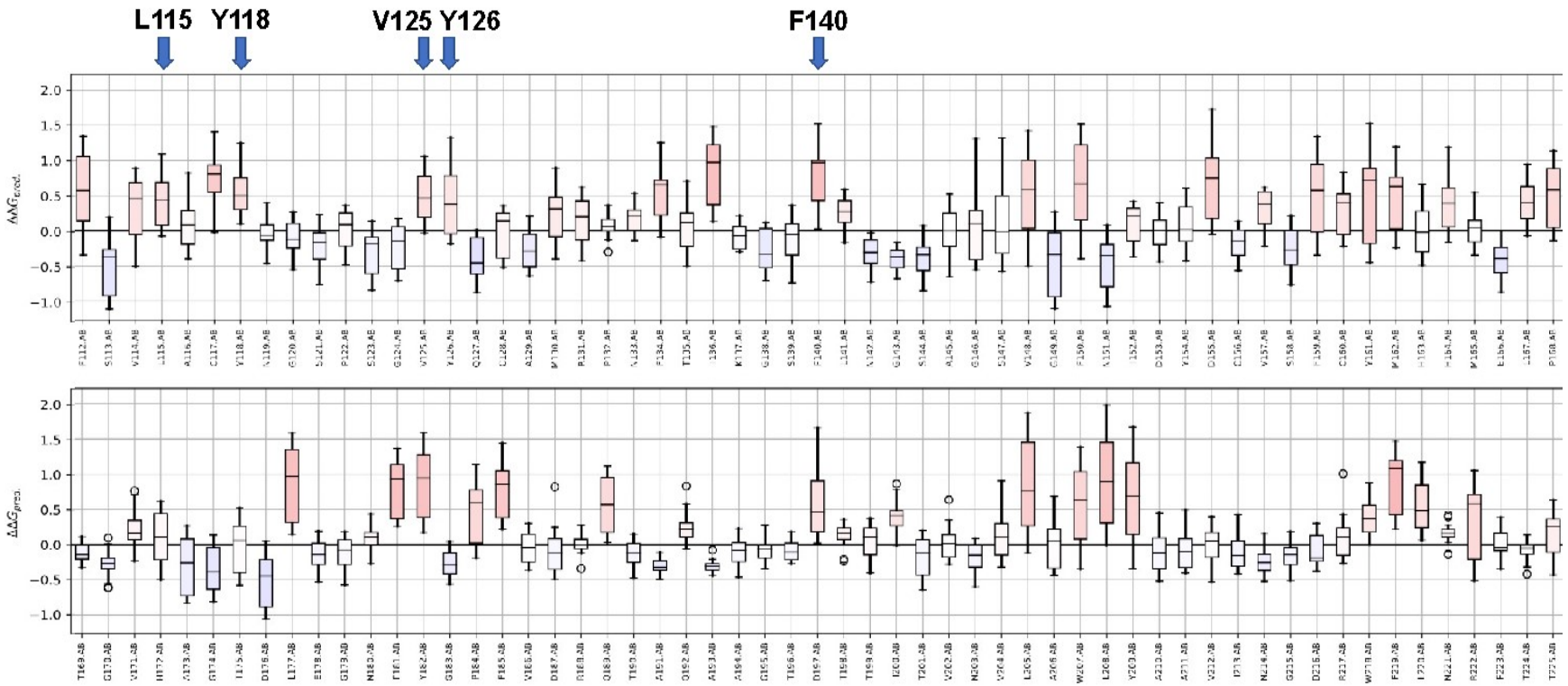

Fig. S3. Mutation sensitivity profile of dimeric native M^{pro}. The mutation sensitivity profile of dimeric native M^{pro} from Phe112 to Thr225 is shown, where the total predicted change in stability is determined as $\triangle\triangle G_{pred}$ ($\triangle\triangle G_{pred}$<0 indicates stabilizing and $\triangle\triangle G_{pred}$>0 indicates destabilizing mutation). L115, Y118, V125, Y126, and F140 indicate destabilizing residues at the interface of dimeric M^{pro}.

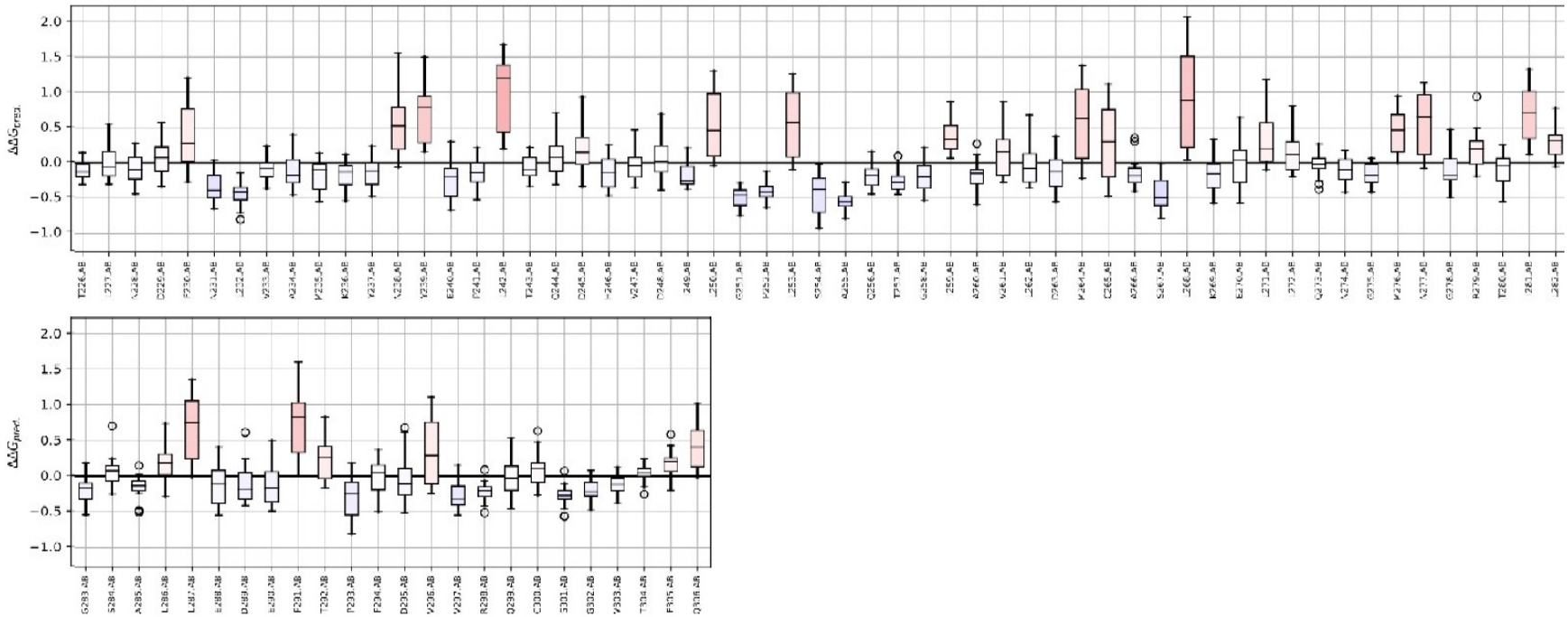

Fig. S4. Mutation sensitivity profile of dimeric native M^{pro}. The mutation sensitivity profile of dimeric native M^{pro} from Thr226 to Gln306 is shown, where the total predicted change in stability is determined as $\triangle\triangle G_{pred}$ ($\triangle\triangle G_{pred}$<0 indicates stabilizing and $\triangle\triangle G_{pred}$>0 indicates destabilizing mutation).

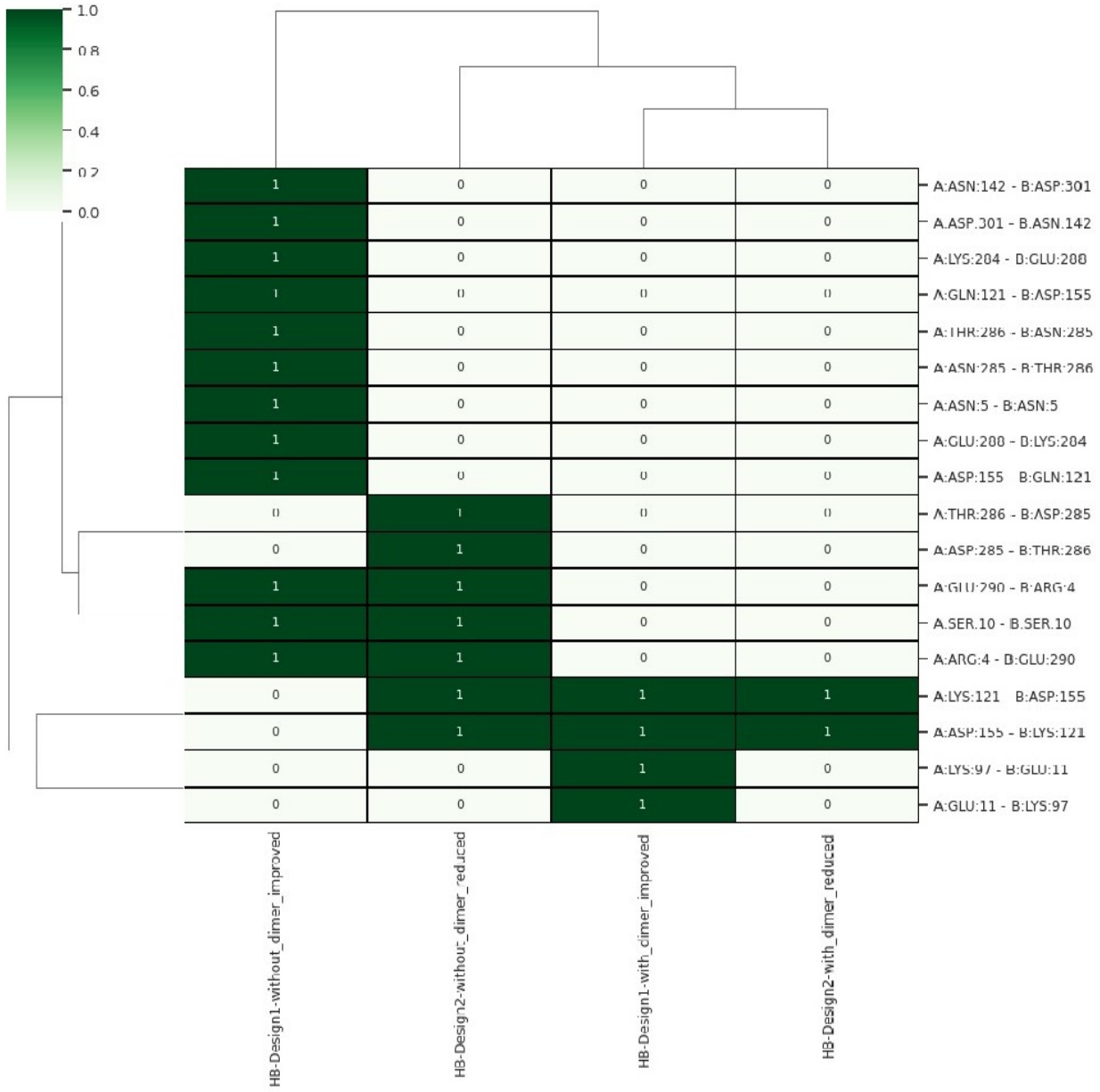

Fig. S5. Computed hydrogen bonds between the monomers of dimeric M^{pro} top-scored and low-scored designs. The hydrogen bonds obtained between the two monomers of top-scored and low-scored designed dimeric M^{pro} are shown. The with_dimer and without_dimer in the labels are from design experiment 1 and design experiment 2, respectively. The green color with '1' denotes the presence of hydrogen bond interactions. The contacts were shown as cluster gram to make the interpretation clear for visualization.

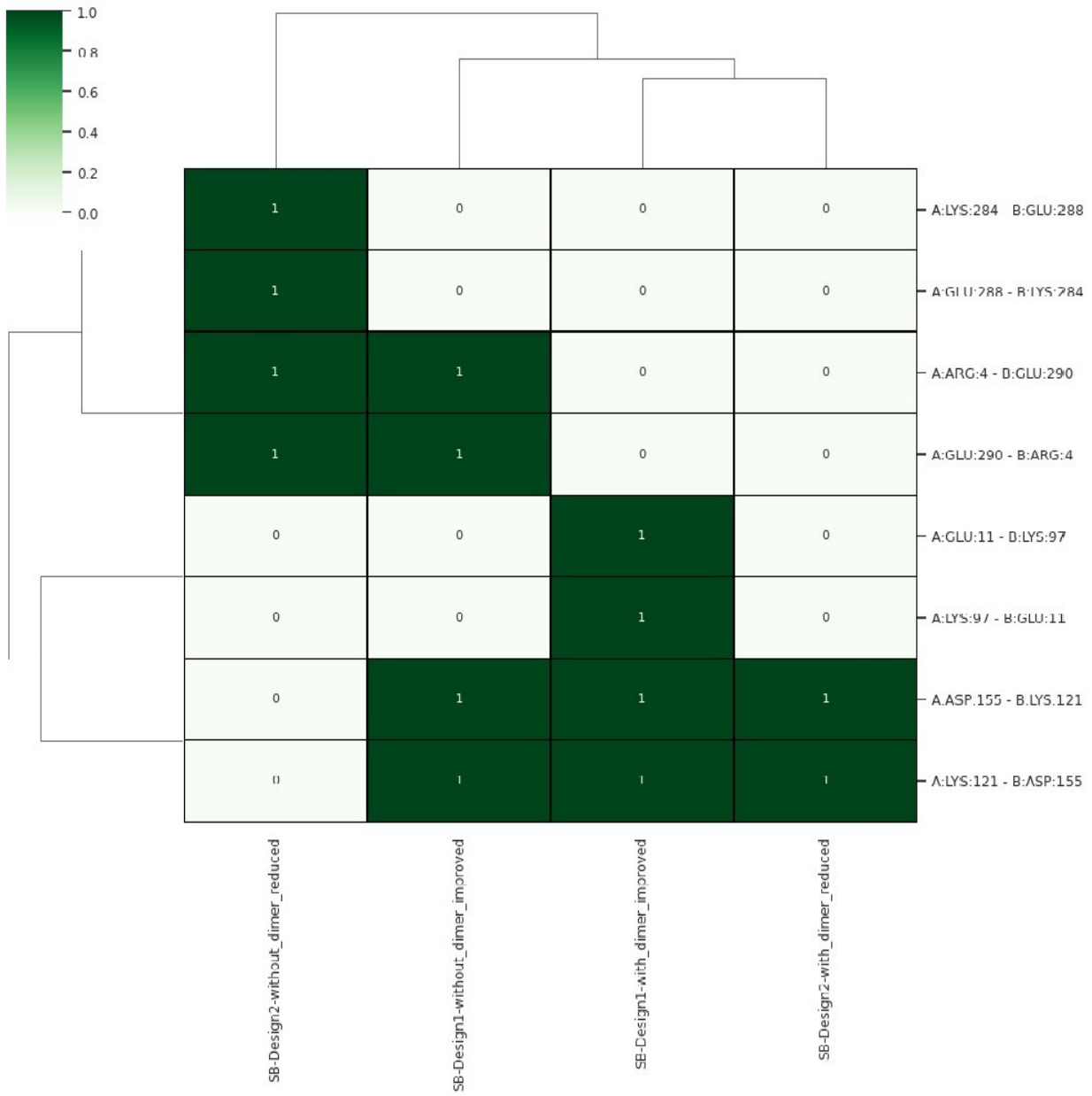

Fig. S6. Computed salt bridges between the monomers of dimeric M^{pro} top-scored and low-scored designs. The salt bridges obtained between the two monomers of top-scored and low-scored designed dimeric M^{pro} are shown. The with_dimer and without_dimer in the labels are from design experiment 1 and design experiment 2, respectively. The green color with '1' denotes the presence of salt bridges. The contacts were shown as cluster gram to make the interpretation clear for visualization.

Table S2. Intermolecular interactions formed between the two symmetric monomers of M^{pro} for the top-scored and low-scored designs from each design experiment.

Design experiment 1 (With dimerization residues)		
Types of interactions	**Top-scored design**	**Low-scored design**
Van der Waals interactions	25	12
Proximal interactions	1271	943
Polar contacts	40	44
Hydrogen bonds	27	34
Hydrophobic contacts	64	24
Carbonyl interactions	4	2
Total number of interactions	**1431**	**1059**
Design experiment 2 (Without dimerization residues)		
Types of interactions	**Top-scored design**	**Low-scored design**
Van der Waals interactions	14	24
Proximal interactions	1146	1056
Polar contacts	38	32
Hydrogen bonds	34	30
Hydrophobic contacts	44	39
Carbonyl interactions	4	0
Ionic interactions	6	2
Aromatic contacts	4	8
Total number of interactions	**1290**	**1191**
Design experiment 3 (7-residue designing)		
Types of interactions	**Top-scored design**	**Low-scored design**
Van der Waals interactions	22	12
Proximal interactions	1110	986
Polar contacts	34	36
Hydrogen bonds	27	32
Hydrophobic contacts	46	34
Carbonyl interactions	2	0
Ionic interactions	6	2
Aromatic contacts	6	4
Total number of interactions	**1253**	**1106**

Notes

Note S1. The designs were sampled using Monte-Carlo-simulated annealing in the Rosetta all-atom forcefield.

Note S2. The Rosetta total score (REU) is the weighted sum of various energy terms, including physical forces such as electrostatics and van der Waals' interactions, and several other statistical terms.

Note S3. The percentage sequence identity denotes the identity of the designed sequences to that of the native sequence.

Note S4. The predicted LDDT (pLDDT) score: AlphaFold produces a per-residue estimate of its confidence on a scale from 0-100. This confidence measure is called pLDDT. Regions with pLDDT between 70 and 90 are expected to be modeled well (a generally good backbone prediction).