Supporting Information

Structure-based stabilization of non-native protein-protein interactions of coronavirus nucleocapsid proteins in antiviral drug design

Shan-Meng Lin, Shih-Chao Lin, Jia-Ning Hsu, Chung-ke Chang, Ching-Ming Chien, Yong-Sheng Wang, Hung-Yi Wu, U-Ser Jeng, Kylene Kehn-Hall, Ming-Hon Hou

Table of Contents

Supplemental figures

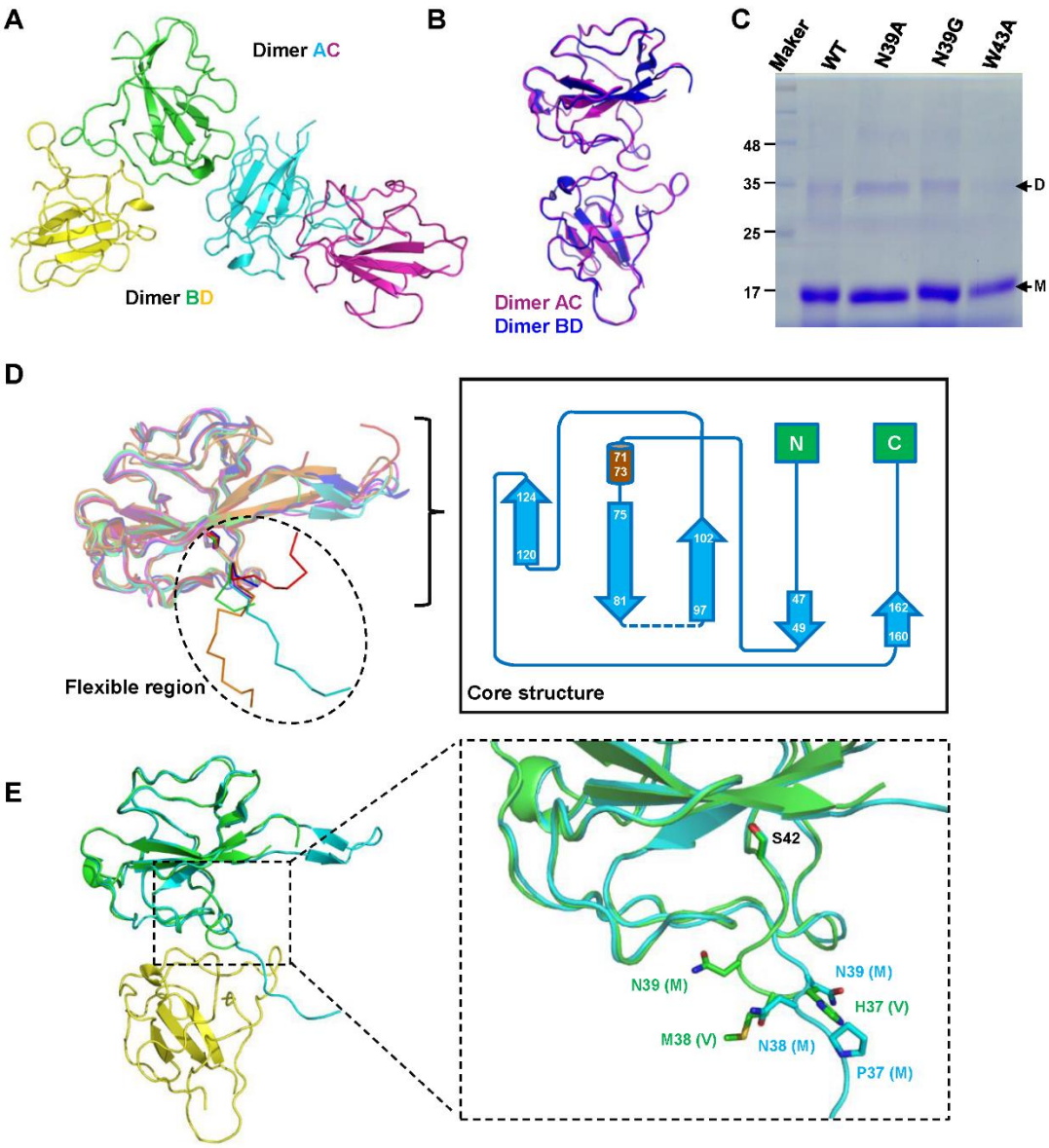
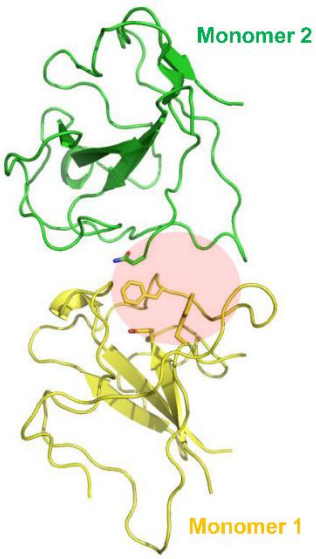


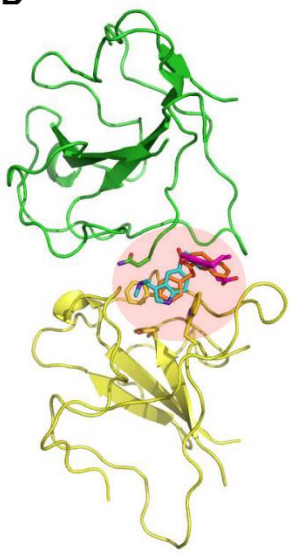
Figure S1. Dimeric conformation of MERS-CoV N-NTD. (A) Overall structure of MERS-CoV N-NTD. Four molecules (A-D) form two dimers per asymmetric unit. (B) Superimposition of two dimers in the same asymmetric unit showing high structural identity and RMSD = 0.303 Å. (C) Cross-linking analysis of wild-type (WT) MERS-CoV N-NTD and various MERS-CoV N-NTD mutants. Black arrows indicated the expected positions of the protein dimer and monomer. (D) Structural superposition was performed by aligning the reported CoVs N proteins to a monomer of the MERS-CoV structure. All reported CoV NTD were monomeric and contained a core structure with a conformation similar to that which is depicted in the topology diagram (right panel). (E) Superimposed structure was achieved by aligning the MERS-CoV N-NTD containing vector-fusion

residues (solved in this study; molecule color is the same as in Fig. 1A) to that containing native residues (PDB: 4ud1; shown in cyan). Black box highlights the flexible region. Residues from the vector and corresponding residues from the native MERS are shown as sticks and labeled with V and M in brackets, respectively. S42 indicates the starting point of the flexible area.

A



B



C

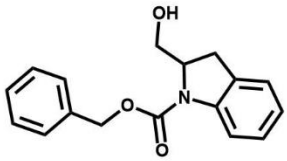
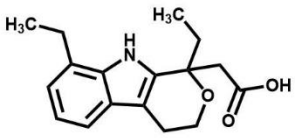
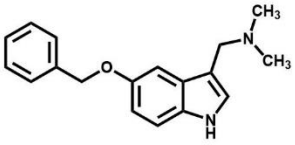
Compound name	Chemical structure	LibDock score	TPSA (Å²)	$S_{L/L}$ (Å²)
Benzyl 2-(Hydroxymethyl)-1-Indolinecarboxylate Name in this paper: P1 Compound		121.445	28.26	59.87
Etodolac Name in this paper: P2 Compound		85.4327	62.32	52.85
5-Benzyloxygramine Name in this paper: P3 Compound		113.82	49.77	59.24

Figure S2. LibDock pose of dimeric MERS-CoV N-NTD with potent compounds. (A) Cartoon representation of non-native dimer of MERS-CoV N-NTD showing the selected binding sphere used in the shape complementarity docking. The color of each monomer is the same as in Fig. 1A. The residues forming the W-43 pocket on monomer 2 and N39 on monomer 1 are shown as sticks. (B) Same as in (A) except the selected compounds bear the colors brown, purple and cyan for P1, P2, and P3, respectively. (C) Detailed information for each selected compound.

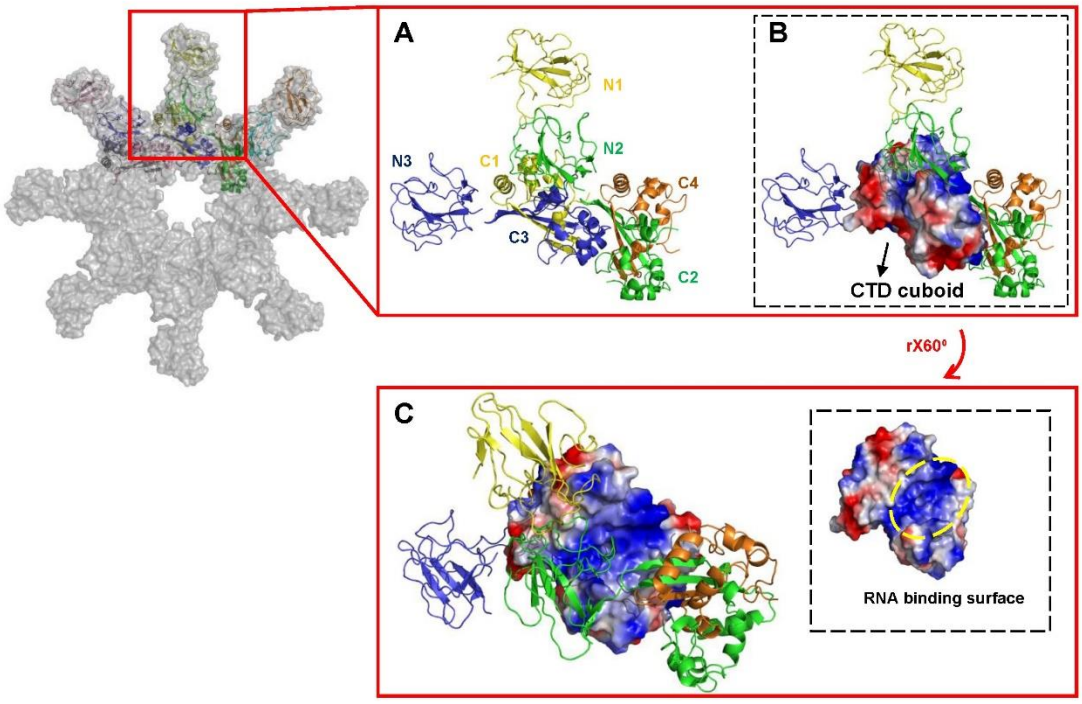


Figure S3. P3-induced aggregation of MERS-CoV N protein sequesters the RNA binding region on the N-CTD dimer. Red box indicates the contacts of each protein surrounding one CTD cuboid in the N protein aggregation described in detail in the right panel. (A) The indicated area includes three full-length N proteins and one N protein CTD labeled with N1, C1; N2, C2; N3, C3; and C4 for the N-and C-TD of each protein, respectively. Four molecules are shown in a cartoon and colored in yellow, green, blue, and orange for N proteins 1–4, respectively. (B) Same as (A) except the CTD cuboids formed by C1 and C3 are shown as an electrostatic surface to highlight the basic RNA binding region. Blue and red indicate positive and negative charge potential, respectively. (C) Same as (B) except the view is rotated 60° along the x-axis to clarify the sequestered RNA binding surface. (Insert) Proposed RNA binding surfaces are highlighted with yellow dotted circles.

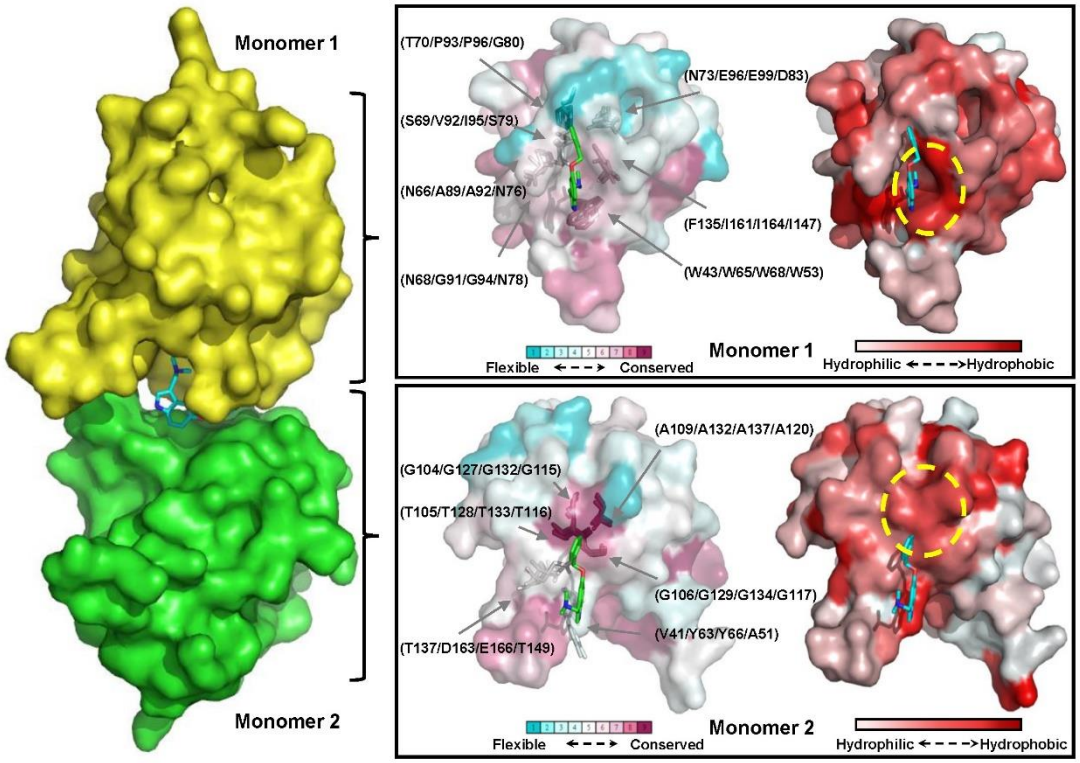


Figure S4. Potential target sites for broad-spectrum antiviral compound development. Left panel: surface representation of MERS-CoV N-NTD with P3 depicted as a stick structure. Right panel: maps of conserved and hydrophobic surfaces of monomer 1 (upper) and monomer 2 (lower). N protein structures of various CoV were aligned and colored according to sequence conservation based on alignment corresponding to Fig. 1D. Residues of the aligned proteins corresponding to the P3 binding pocket on MERS CoV are shown as sticks and labeled in the order HCoV-OC43 N-NTD (PDB ID: 4J3K), SARS-CoV N-NTD (PDB ID: 2OFZ), MHV N-NTD (PDB ID: 3HD4), and IBV N-NTD (PDB ID: 2GEC). Surface hydrophobicity colors range from the most hydrophobic (red) to the most hydrophilic (white) according to the surface residues of MERS CoV N-NTD. Potential target sites on N-NTD are highlighted with yellow circles.

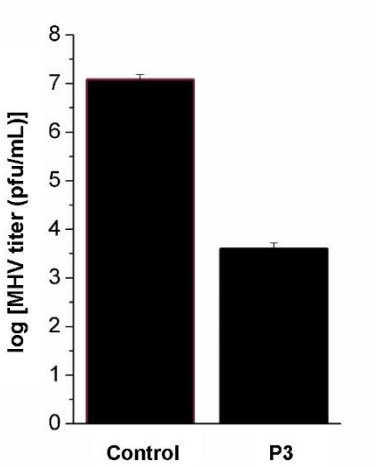


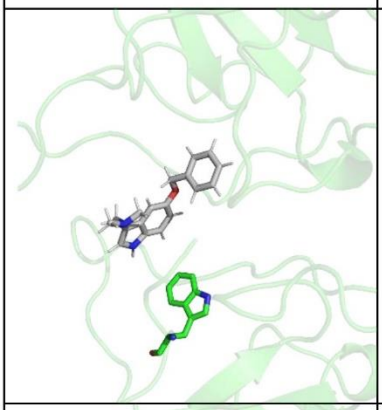
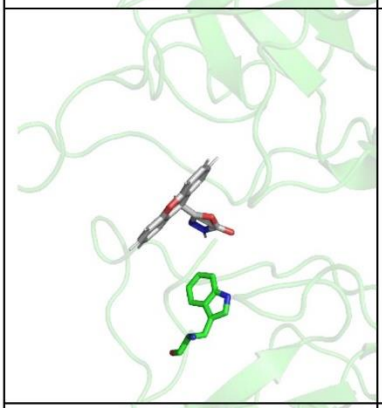
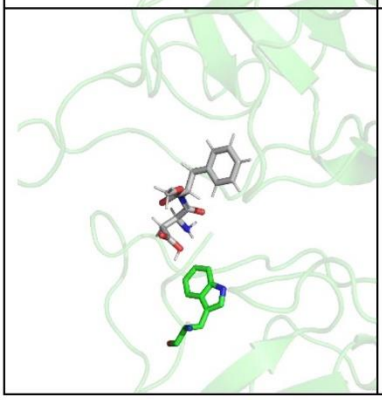
Figure S5. Antiviral activity of P3 against MHV (mouse hepatitis virus). MHV was pre-incubated with 100 μM P3 or control for 1 h and a plaque assay was performed to determine the virus titer in Vero E6 cells. All values are presented as mean ± SE (standard error of mean).

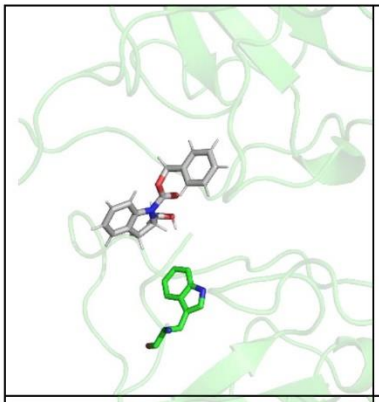
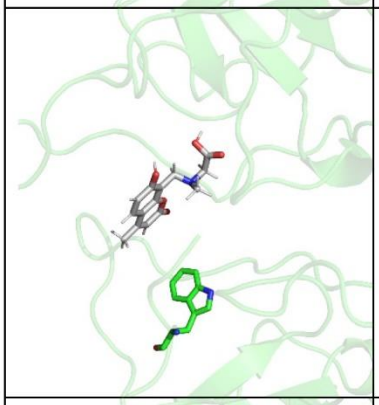
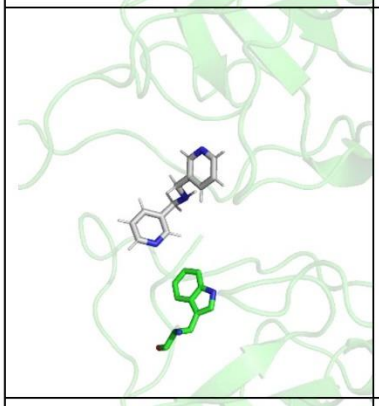
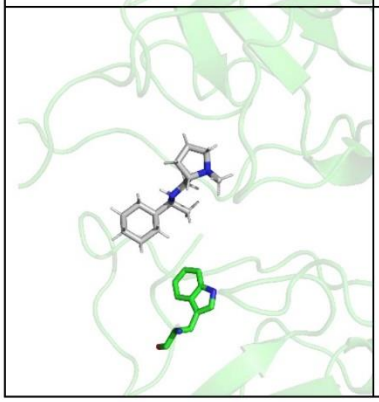
Supplemental tables

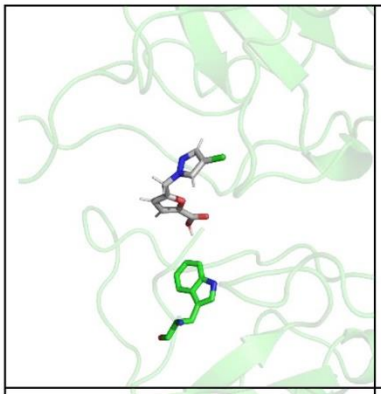
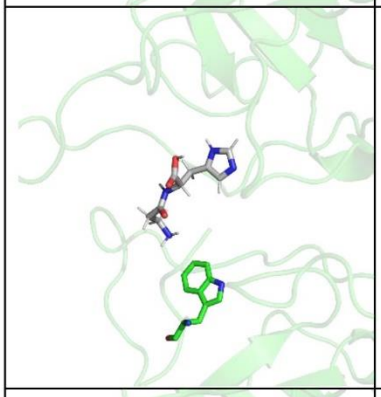
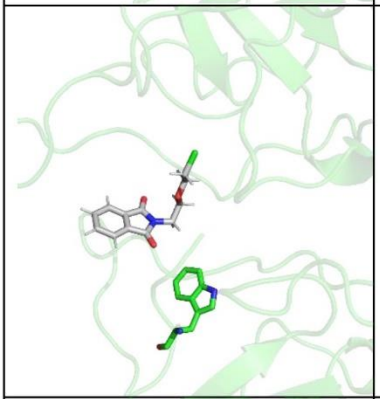
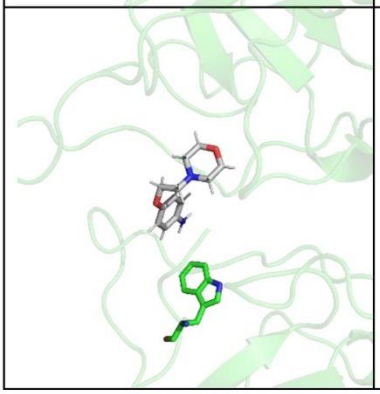
Table S1. Crystallographic data collection and refinement statistics.

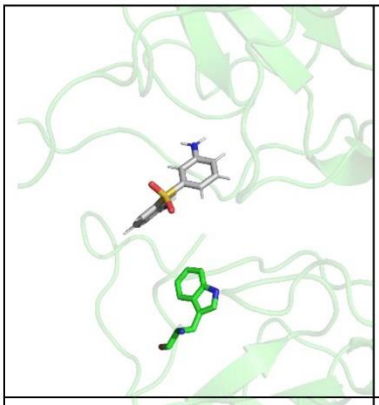
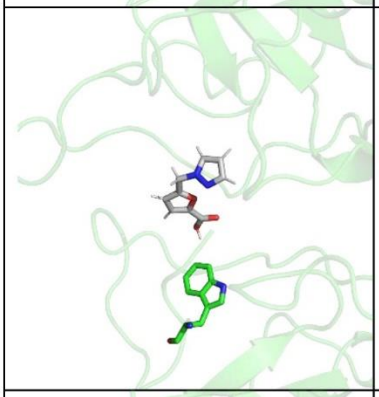
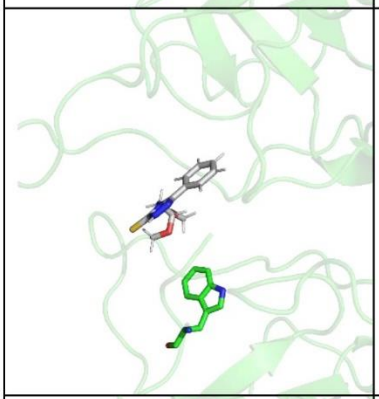
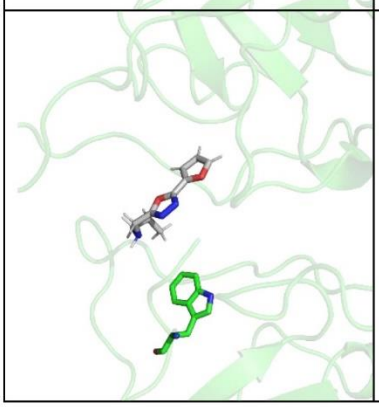
Data collection	NSRRC BL13B1	NSRRC BL13C1	SP44XU, Spring8
Crystal	Native	MERS:P1	MERS:P3
Wavelength (Å)	1	0.97622	0.9
Space group	P 21	P21	P21
Cell dimensions			
a, b, c (Å)	35.601, 109.645, 91.993	35.440, 109.897, 91.526	35.489, 108.477, 91.595
α, β, γ (º)	90, 101.23, 90	90, 90, 101.103	90, 90, 101.160
Resolution range (Å)	30.00-2.63 (2.73-2.63)	30.00-3.09 (3.20- 3.08)	30.00-2.77 (2.87-2.77)
Unique reflections	19199 (1597)	11929 (1009)	17364
Completeness (%)	94.5 (91.4)	96.1 (96.9)	99.5 (99.4)
Mean I/σ(I)	18.897 (3.304)	9.143 (2.500)	31.387 (99.4)
R-merge	0.096 (0.494)	0.141 (0.482)	0.072 (0.477)
Redundancy	7.1 (5.9)	3.5 (3.4)	6.3 (7.3)
Refinement			
R-work/ R-free	28/28.98	28.18/29.93	23.15/27.07
Number of atoms	3546	3498	3680
macromolecules	3445	3477	3549
ligands	n/a	21	21
water	101	0	110
Protein residues	543	445	453
Average B-factor	40.6	54.09	35.33
macromolecules	40.7	54.14	35.41
ligands	n/a	45.41	38.46
solvent	37.9	n/a	32.26
RMSD			
Bonds lengths (Å)	0.013	0.012	0.008
Angles (º)	1.59	1.67	1.13
Ramachandran plot			
Favored (%)	98.35	98.14	98.84
Outliers (%)	0	0	0
Clash score	15.7	15.01	25.53

Table S2. W43 docking pose with chemical structures, docking scores, and biochemical properties of 17 potential hits.

LibDock pose	Name	LibDock Score	TPSA (Å^2)	$S_{L/L}$ (Å^2)
	5-Benzyloxygramine (P3)	121.445	28.26	59.87
	5-(9H-xanthen-9-yl)-1,3,4-oxadiazol-2-ol	120.754	68.13	34.3
	Aspartame	120.453	118.72	35.38

	Benzyl 2-(Hydroxymethyl)-1-Indolinecarboxylate (P1)	113.82	49.77	59.24
	2-[(7-hydroxy-4-methyl-2-oxo-chromen-8-yl)methyl-methyl-amino]acetic	112.677	90.98	32.02
	3,3'-Dipicolylamine	104.66	37.81	10.89
	1-cyclohexyl-N-((1-methylpyrrolidin-2-yl)methyl)ethanamine	104.575	15.27	55.13

	5-(4-Chloro-pyrazol-1-ylmethyl)-furan-2-carboxylic acid	103.727	68.27	43.54
	L-Carnosine	102.696	121.10	0
	2-[2-(2-chloroethoxy)ethyl]isoindole-1,3-dione	101.784	48.31	33.29
	4-(2-morpholin-4-ylethoxy) aniline	101.191	47.73	28.86

	3,3'-Sulfonyldianiline	100.829	86.19	36.52
	5-(1H-pyrazol-1-ylmethyl)-2-furoic acid	99.7464	68.27	41.71
	4-(2,2-dimethoxyethyl)-5-phenyl-4H-1,2,4-triazole-3-thiol	99.6227	52.09	41.9
	N-{[5-(2-furyl)-1,3,4-oxadiazol-2-yl]methyl}-N-propylamine	97.9875	64.09	45.89

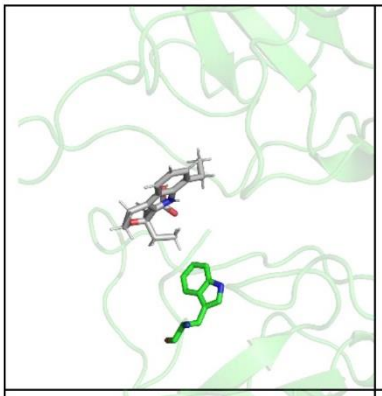
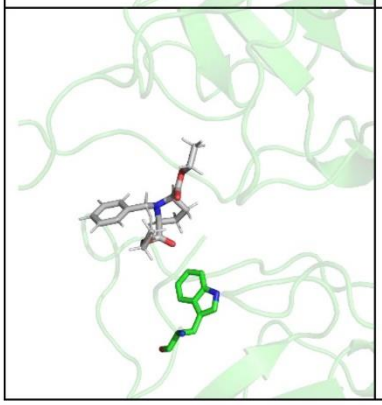
	Etodolac (P2)	85.4327	62.32	52.85
	Diethyl 1-benzylpyrrolidine-2,5-dicarboxylate	84.966	55.85	50.82

Table S3. Primers used for mutagenesis in this study.

N39A	Forward 5' CCGCGCGGCAGCCATATGGCGACCGTGAGCTGGTATACC3'
	Reverse 5' GGTATACCAGCTCACGGTCGCCATATGGCTGCCGCGCGG3'
N39G	Forward 5' CCGCGCGGCAGCCATATGGGCACCGTGAGCTGGTATACC 3'
	Reverse 5' GGTATACCAGCTCACGGTGCCCATATGGCTGCCGCGCGG3'
W43A	Forward 5' CATATGAACACCGTGAGCGCGTATACCGGCCTGACCCAG3'
	Reverse 5' CTGGGTCAGGCCGGTATACGCGCTCACGGTGTTCATATG3'