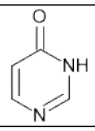
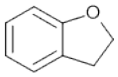
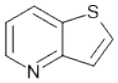
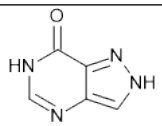
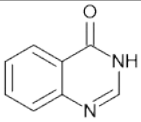
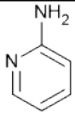
Supplementary information for: Computational analysis of energetic features and intermolecular interactions in protein-inhibitor USP7 complexes

M. Ziemniak[1*], U. Budniak[1], P. M. Dominiak[1], K. Woźniak[1*]

Table S1. List of crystallized structures of USP proteins complexed with a small molecule grouped according to their chemical scaffold used in the study. Bolded entries indicated USP7 complexes selected for detailed studies.

Chemical scaffold	Ligand	PDB entry	Ligand entry	IC_{50}	K_d	Reference
	**compound 1** **compound 16** **compound 46**	**6VN4** 5N9R **6F5H**	**R4D** 8RN **CQ5**	**8.2 µM** 0.3 µM **-**	**-** 2.6 µM **90 nM**	[1] [2] [3]
	compound 7	6VN5	R41	5.6 µM	-	[1]
	compound 14 compound 18 **compound 23**	6VN6 6VN2 **6VN3**	R4J R44 **R3Y**	9 nM 0.2 nM **1.2 nM**	- - -	[1] [1] [1]
	**L55** FT671 **compound 5**	**6M1K** 5NGE **5N9T**	**EZF** 8WK **8QQ**	**40.8 nM** 52 nM **22 nM**	**78 nM** 65 nM -	[4] [5] [2]
	**XL188** nd	**5VS6** 5VSK	**9QD** 9HS	**90 nM** 1.3 µM	**104 nM** -	[6] [6]
	**GNE6640**	**5UQV***	**8JM**	**0.75 µM**	-	[7]
	**compound 2**	**5WHC***	**AJJ**	**79 µM**	-	[8]

[a]lack of detailed information, only mentioned that estimated IC_{50} is ten times lower than for IU1

Table S2. Interatomic interactions from Hirshfeld surface analysis. Vey close contacts are highlighted in red, moderately-close in green and more distant, pause indicates lack of important contacts of a given type.

Structure	Interatomic contacts								Affinity range [nM]
	H···H	H···O	H···N	H···C	H···F	H···Cl	H···S	C-O	
6VN2	- | 2 | -	3 | 2 | 4	1 | - | -	- | - | -	ND	- | - | -	- | - | -	- | - | -	0.2
6VN3	- | 2 | -	3 | 2 | 4	1 | - | -	- | - | -	ND	- | - | -	- | - | -	- | - | -	1.0
6VN4	- | 2 | 1	3 | 2 | -	- | 2 | -	- | - | 5	ND	ND	- | - | -	- | - | -	10000
6VN5	- | 2 | -	- | 5 | 3	1 | - | -	- | - | -	ND	- | - | -	- | - | -	- | - | -	10000
6VN6	1 | 2 | 1	3 | 1 | 2	1 | - | -	- | - | 1	ND	- | - | -	- | - | -	- | - | -	10
5NGE	- | 1 | -	3 | 2 | 2	1 | 2 | 3	- | - | 1	- | 2 | -	ND	- | - | -	- | - | -	50
5N9T	5 | 1 | -	3 | 2 | 2	1 | 3 | -	- | - | 1	- | 2 | -	ND	- | - | -	- | - | -	20
6M1K	-| 7 | -	- | 4 | -	- | 1 | -	- | 2 | 2	ND	- | 1 | -	- | - | -	- | - | -	40
6F5H	2 | 10 | -	2 | 3 | 1	- | 2 | -	- | - | 3	ND	- | - | -	- | - | -	- | - | 1	100
5VSK	6 | 8 | -	- | 4 | -	- | 1 | -	- | - | 3	ND	- | - | -	- | - | -	- | - | -	1000
5WHC	3 | 1 | -	1 | - | -	- | - | 1	- | - | 3	ND	ND	- | - | -	- | - | -	100000
5UQV	3 | - | -	1 | 3 | -	- | - | -	- | - | 8	ND	ND	- | - | -	- | - | -	1000
5N9R	- | 2 | -	3 | 2 | -	- | 1 | -	- | - | 3	ND	- | - | 1 **(Br)**	- | - | -	- | - | -	200
5VS6	5 | 1 | -	1 | 1 | 2	- | 3 | -	- | - | 6	ND	ND	- | - | -	- | - | -	100

Table S3. Flexibility of selected aminoacid residues for USP7 catalytic domain in *apo* form.

Residue	RMSF [Å]
F291	0.62
M292	1.00
Q293	0.73
D295	1.10
V296	0.66
Q297	0.65
E298	0.63
R301	0.56
D305	0.47
D308	0.53
K312	0.82
Y348	0.23
D349	0.49
Q351	0.19
H403	0.19
R407	0.39
R408	0.34
K420	0.59
H456	0.29
D459	0.84
N460	1.05
H461	1.22
Y465	0.51
Y514	0.10

Table S4. Comparison between pair-wise electrostatic interactions between TAAM and DFT methods

6VN3	Energy [kJ mol^{-1}]		**6M1K**	Energy [kJ mol^{-1}]		**6FH5**	Energy [kJ mol^{-1}]		**5N9T**	Energy [kJ mol^{-1}]	
	DFT	TAMM		DFT	TAMM		DFT	TAMM		DFT	TAMM
292	-117.62	-101.65	292	-0.07	-1.88	292	1.84	-5.50	292	0.83	-3.43
295	-422.92	-346.23	295	-259.19	-206.79	295	-538.47	-316.48	295	-355.56	-261.08
296	-15.98	-21.26	296	-16.43	-37.63	296	-18.82	-18.07	296	-12.88	-19.75
298	-207.29	-160.55	298	-191.79	-148.08	298	-201.21	-146.23	298	-24.45	1.37
349	-95.30	-95.63	349	-217.54	-179.88	349	-117.32	-94.69	349	-166.23	-147.10
351	-0.21	4.10	351	-70.50	-26.82	351	-3.28	0.54	351	-225.25	-187.79
407	-26.44	-22.85	407	**-10.13**	**-54.62**	407	-29.49	-32.56	407	**-102.10**	**-45.06**
409	-38.23	-35.67	409	-34.33	-31.11	409	**-36.42**	**-91.77**	409	-23.08	-31.86
420	148.34	134.76	420	50.67	34.39	420	68.83	155.61	420	-29.04	-42.05
456	-19.74	-3.22	456	-15.27	7.07	456	-16.13	-13.35	456	34.60	86.65
459	-100.14	-123.31	459	-138.64	-34.91	459	-131.57	-132.56	459	-14.52	-13.86
460	-16.33	-5.88	460	-13.56	-44.85	460	-21.68	-19.45	460	-22.53	-26.29
465	**-15.21**	**2.19**	465	-6.41	4.35	465	**-29.89**	**-79.91**	465	-29.94	-32.83
514	-23.62	-22.45	514	-15.60	-27.81	514	-16.83	-23.71	514	-17.29	-18.12

Table 5. Regression analysis of DFT/TAMM energies using four distinctive methods.

Structure	Pearson		MM-estimation		Theil-Sen		Quantile	
	slope	R^2	slope	robust R^2	slope	R^2	slope	R^2
6VN3	1.18	0.98	1.18	0.98	1.20	0.98	1.18	0.98
6MK1	1.22	0.87	1.22	0.87	1.23	0.86	1.21	0.87
6FH5	1.19	0.82	1.21	0.82	1.17	0.82	1.20	0.82
5NT9	1.24	0.93	1.23	0.93	1.25	0.93	1.23	0.93

Table S6. ESP surface analysis of **6VN3** ligand.

local minima [kcal mol^{-1}]	local maxima [kcal mol^{-1}]
50.15	64.53
7.35	43.19
18.91	57.83
49.25	44.06
54.51	65.92
20.16	33.51
59.83	172.89
0.41	178.36
61.88	63.80
1.94	63.31
	99.72
	65.43
	102.63
	76.44
	70.79

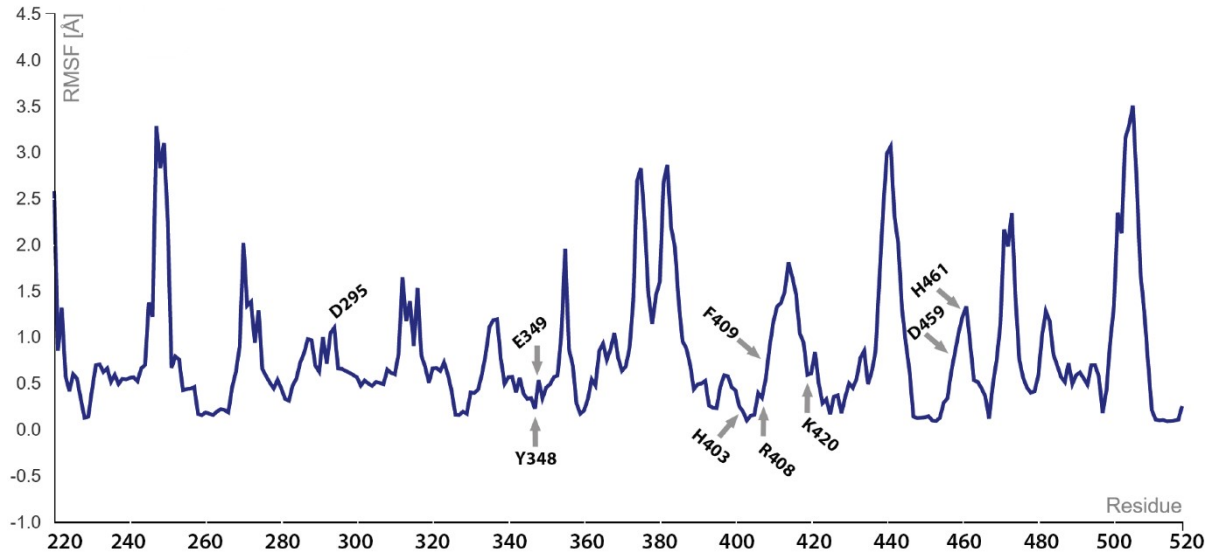


Fig. S1. Flexibility chart of USP7 catalytic domain in *apo* state obtained from CABS-Flex method.

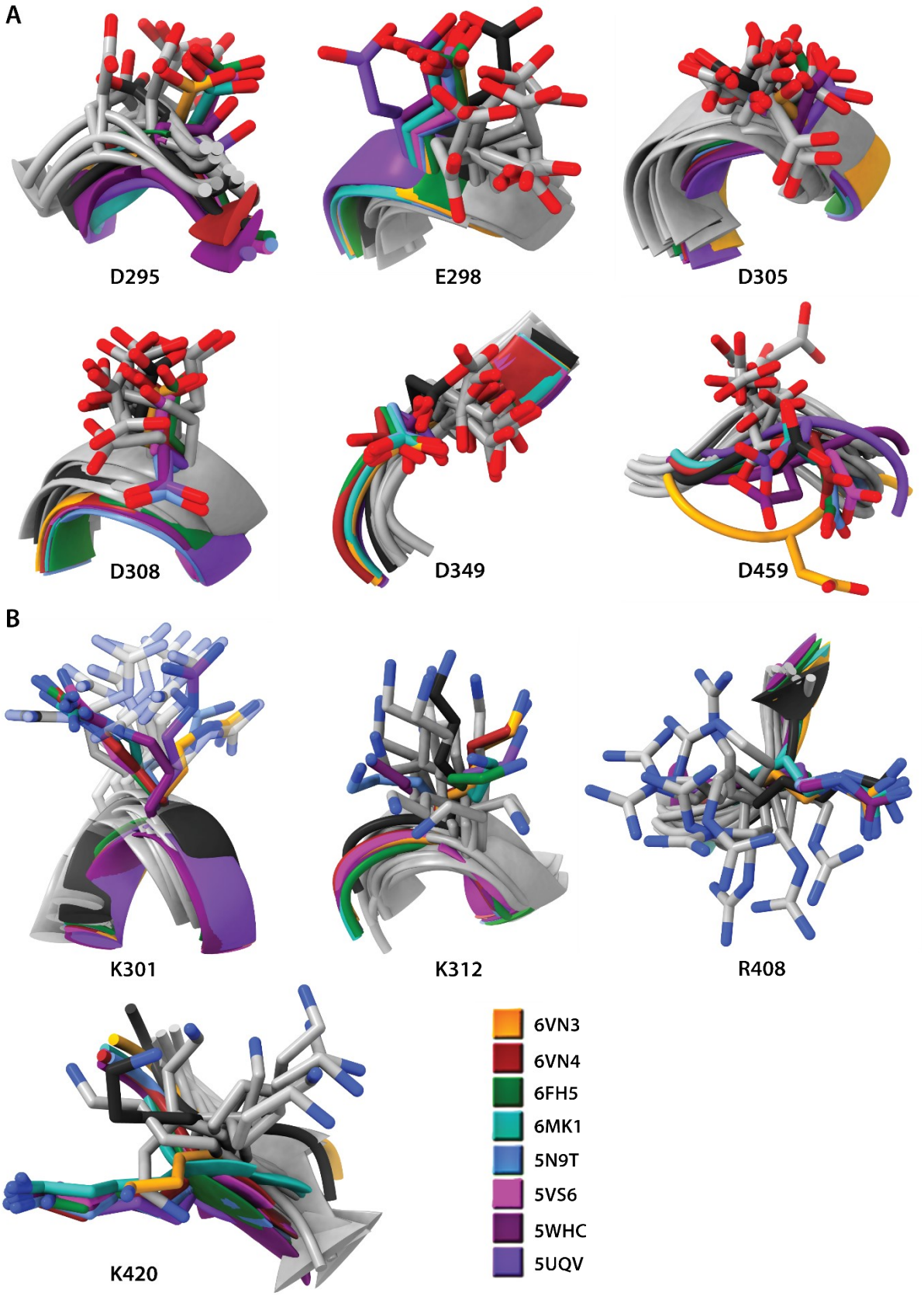


Fig. S2. Flexibility of selected charged residues in USP7. *Apo* form is coloured black, models from CABS-Flex are coloured grey and structures of protein-ligand complexes are depicted in various colours indicated in the legend.

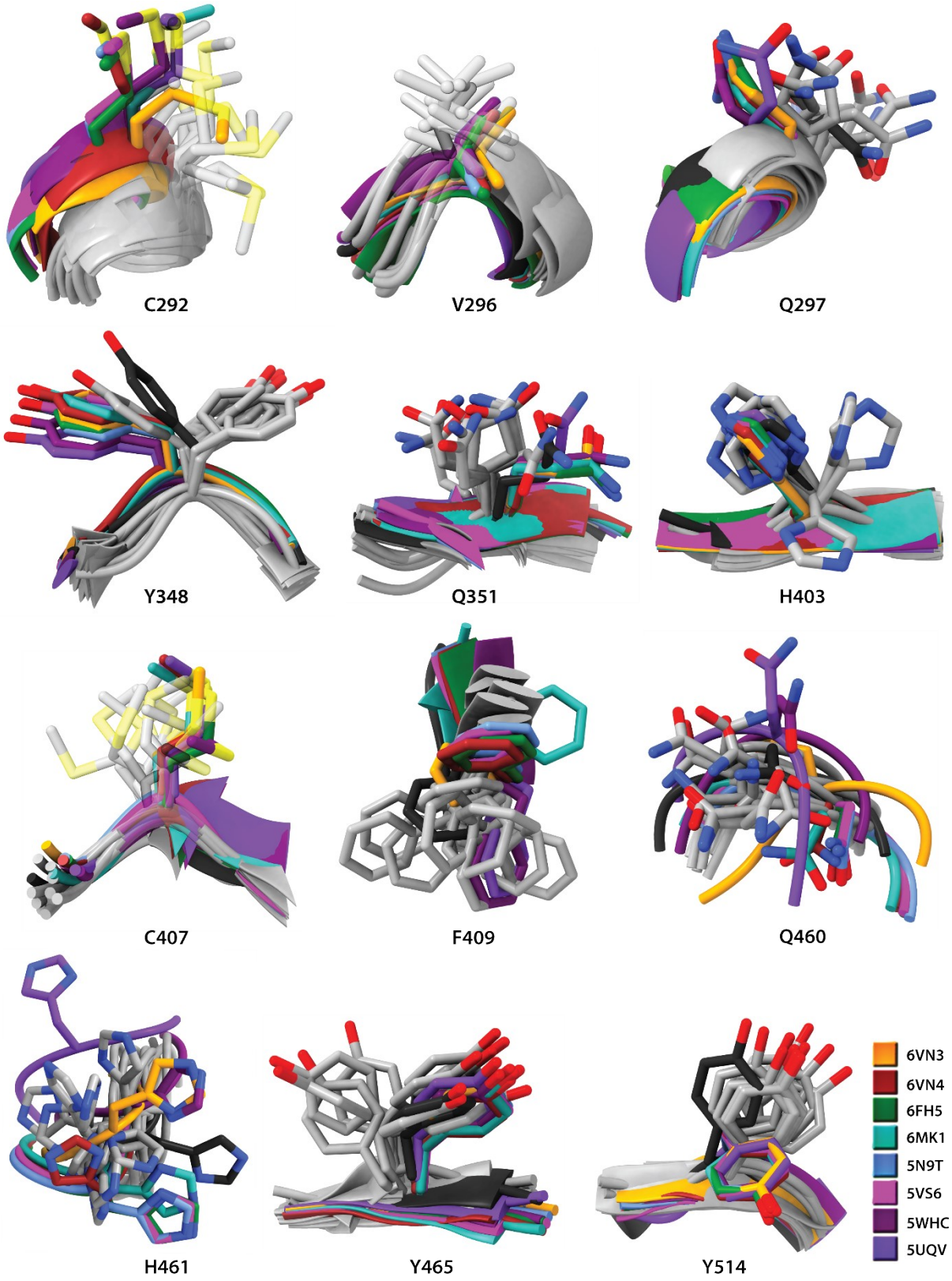


Fig. S3. Flexibility of selected neutral residues in USP7. *Apo* form is coloured black, models from CABS-Flex are coloured grey and structures of protein-ligand complexes are depicted in various colours indicated in the legend.

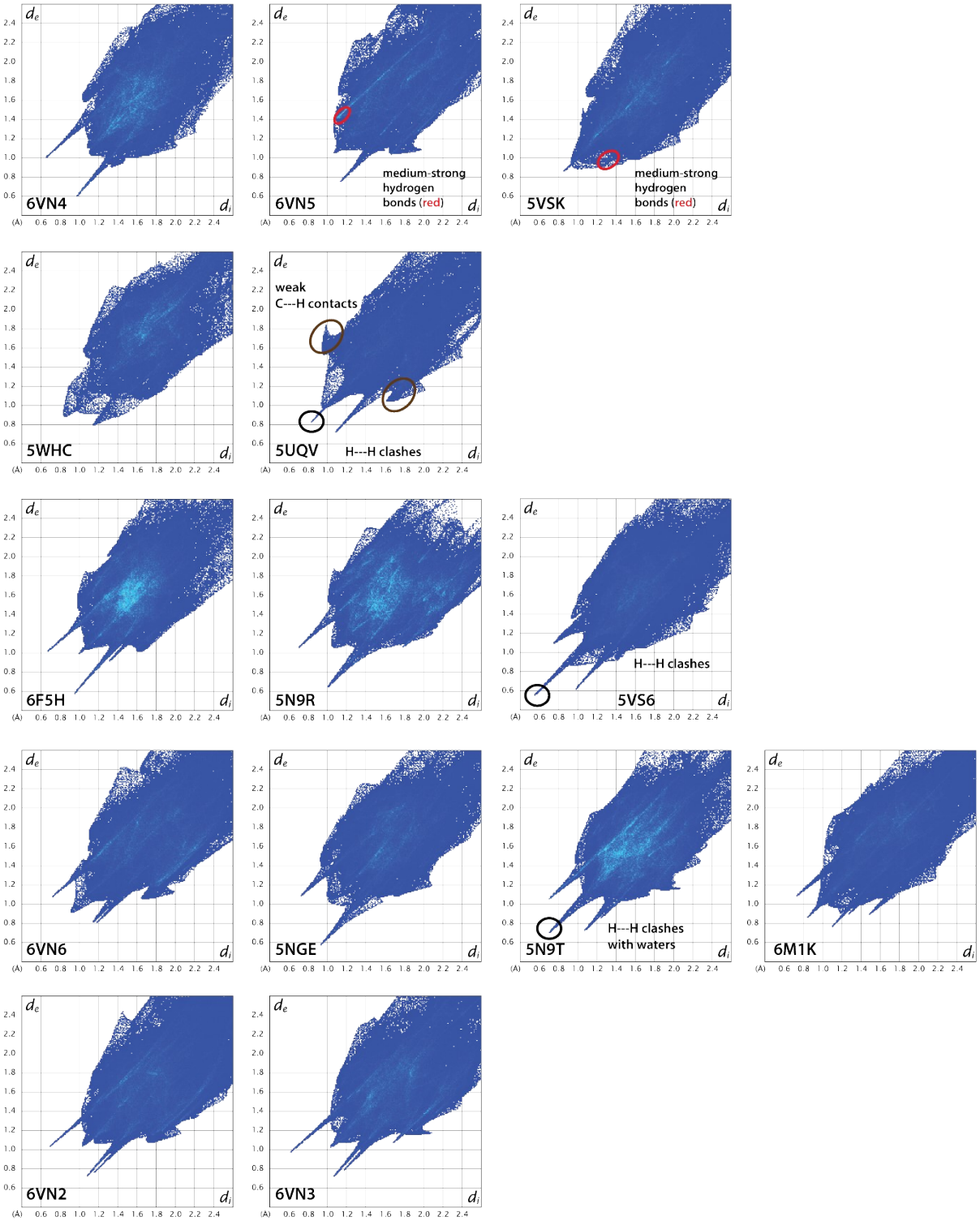


Fig. S4. Fingerprint plots for US7-inhibitor complexes. Some important interatomic contacts are highlighted on the figure.

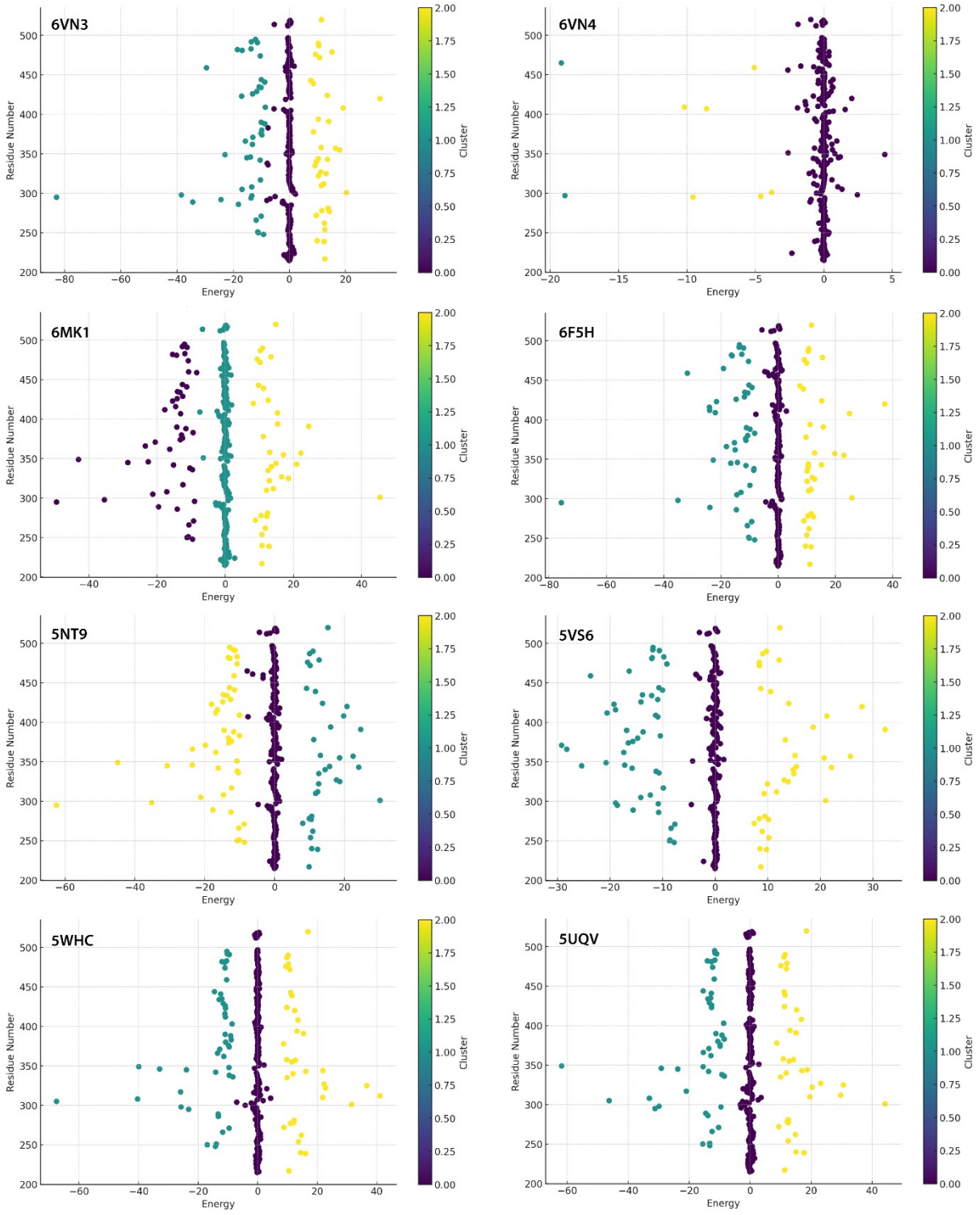


Fig. S5. K-clustering of pairwise residue-ligand energies for USP7 complexes, energies values are in kcal mol^{-1}. Cluster 0 represents residues with small positive or negative contribution to overall binding energy (-5 to +5 kcal mol^{-1}). Cluster 1 represent residues with high positive contribution to binding energy whilst cluster 2 represent residues with negative contribution to binding.

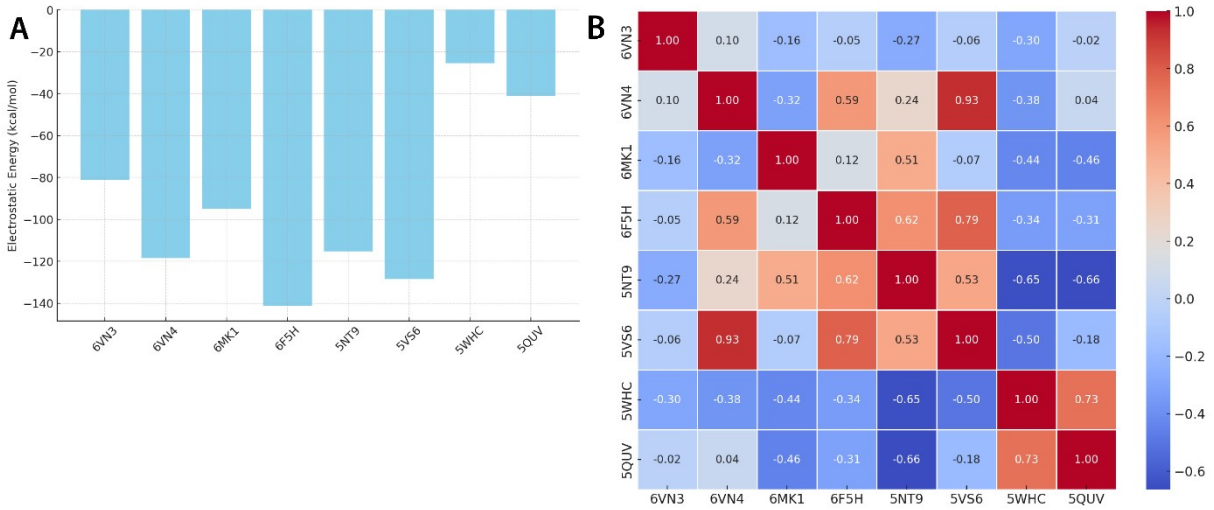


Fig. S6. **A.** Values of $E_{el,S}$ after exclusion of charged residues for USP7 complexes. **B**. Heat map for correlation matrices of $E_{el,S}$ after exclusion of charged residues. Positive Pearson correlation is coloured red and negative is blue.

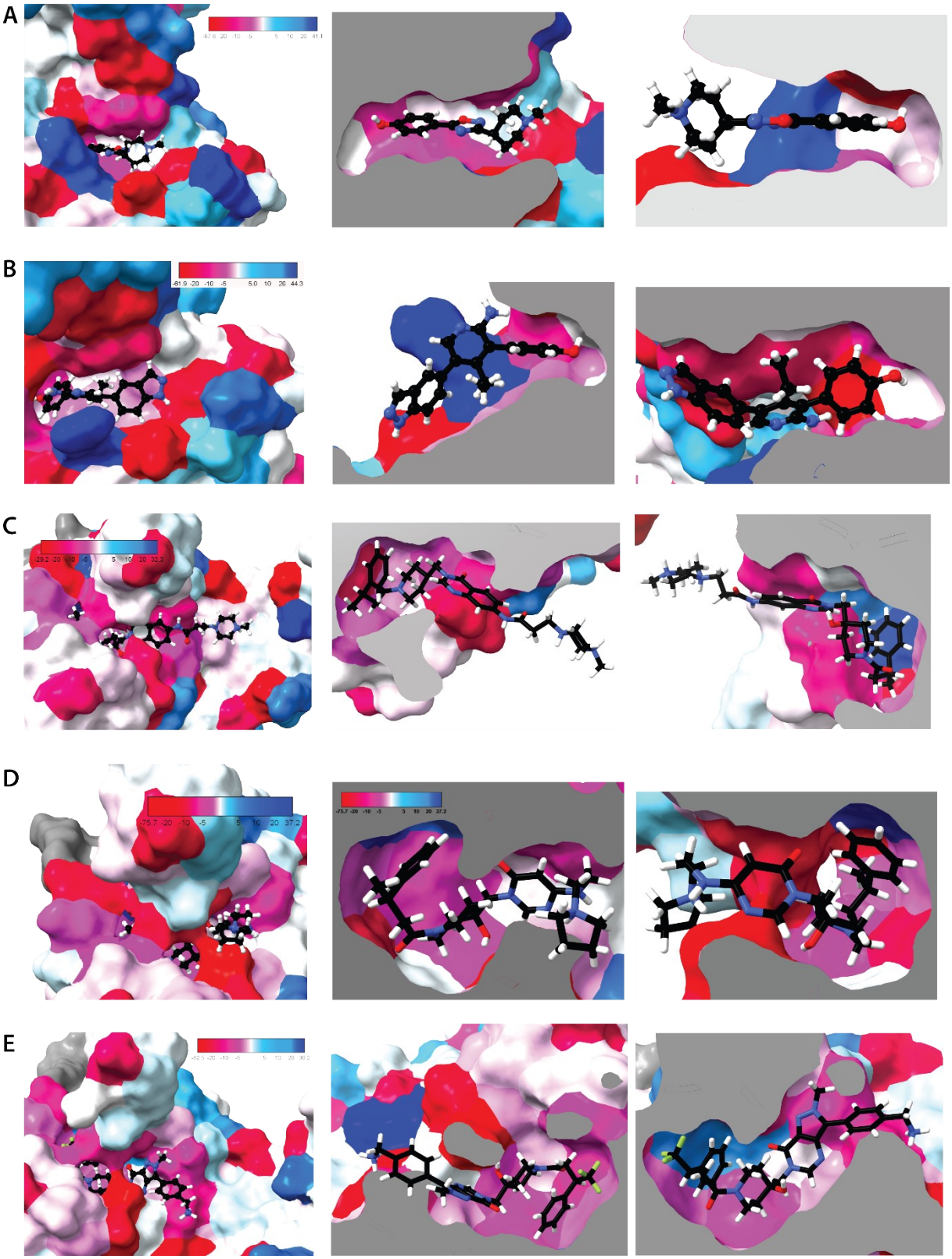


Fig. S7. Electrostatic pairwise residue-ligand binding energies mapped onto molecular surface of USP7 catalytic domain. Some residues missing in the experimental structures were added to show whole molecular surface of the protein. A red-white-blue gradient scheme is applied, where residues with highest stabilising contribution to (negative energy sign) are coloured red while residues with highest destabilising contribution (positive energy sign) are deep blue. Residues excluded in the calculations (missing in the experimental structures) are coloured grey. Pairwise interactions in the binding cavity of **5WHC** (**A**), **5UQW** (**B**), **5VS6** (**C**), **6F5H** (**D**) and **5N9T** (**E**) complexes.

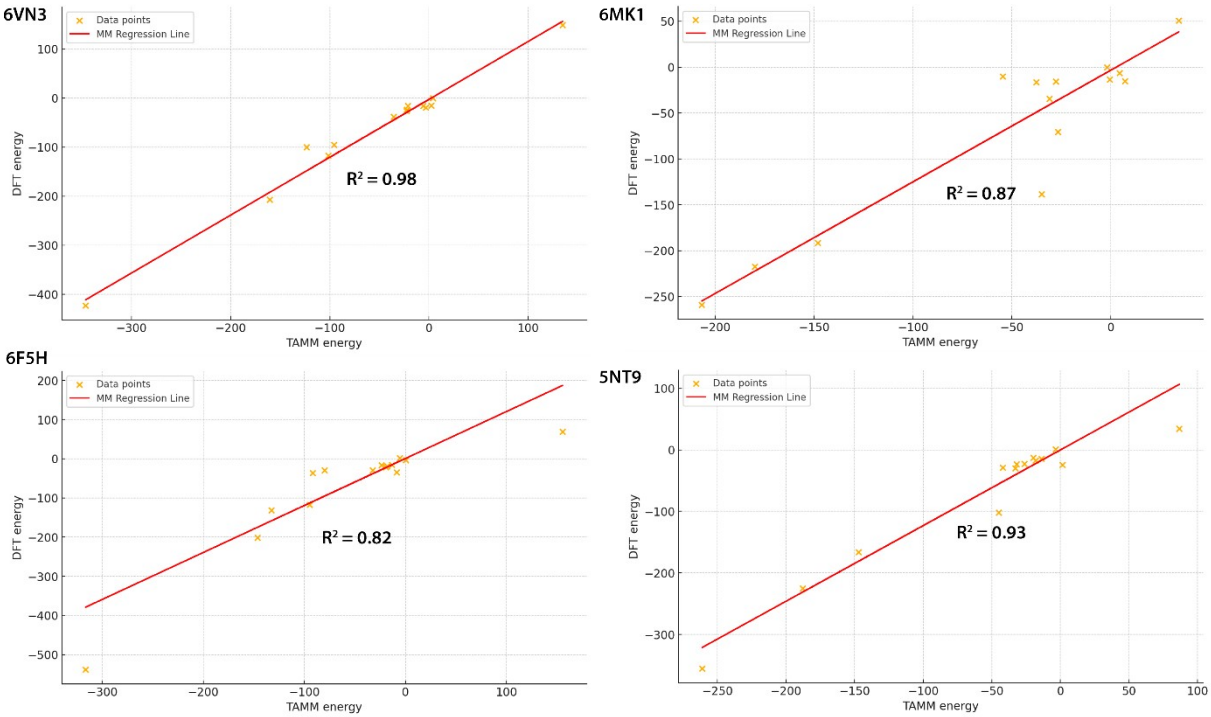


Fig. S8. MM-estimate regression analysis of correlation DFT and TAAM energies for selected USP7 complexes.

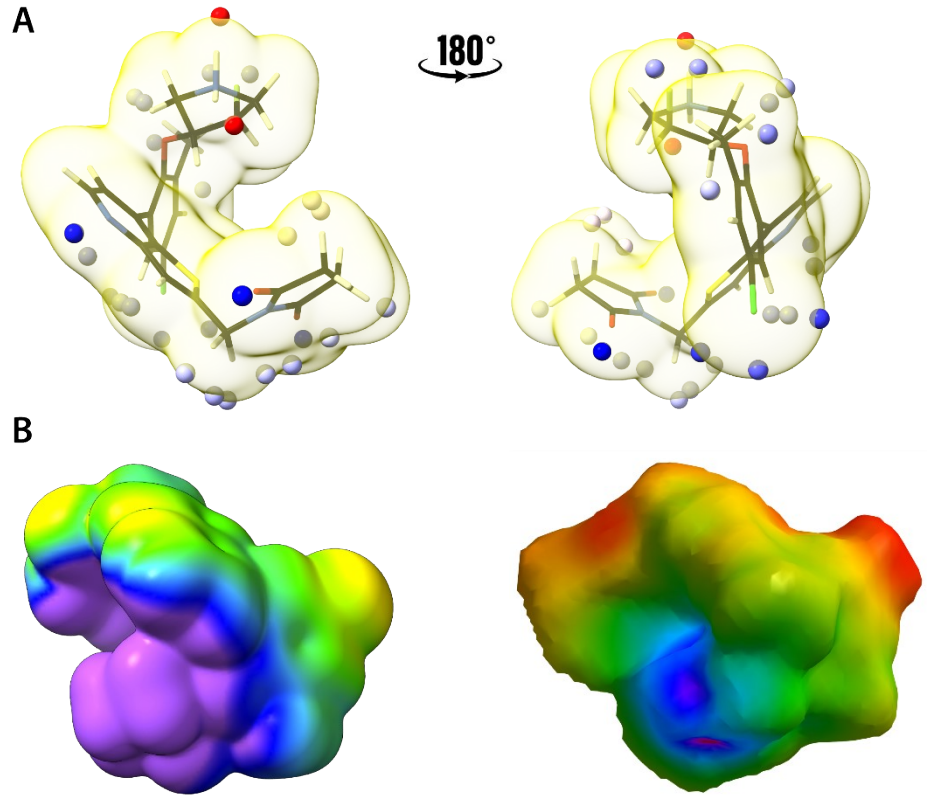


Fig. S9. ESP analysis for **6VN3** ligand. **A.** ESP local extrema on the ED isosurface (0.001 a.u.) depicted as spheres embedded onto the isosurface. A red-white-blue gradient scheme is applied, where maxima are coloured red and minima are deep blue. **B.** Comparison between ESP calculated from wavefunction mapped onto ED isosurface (0.001 a.u., left) and ESP calculated from MM mapped onto Hirshfeld surface (right). In both cases the same colouring scheme was applied, where violet indicate most positive (0.1 a.u.) and red least positive (0.01 a.u.) ESP value.

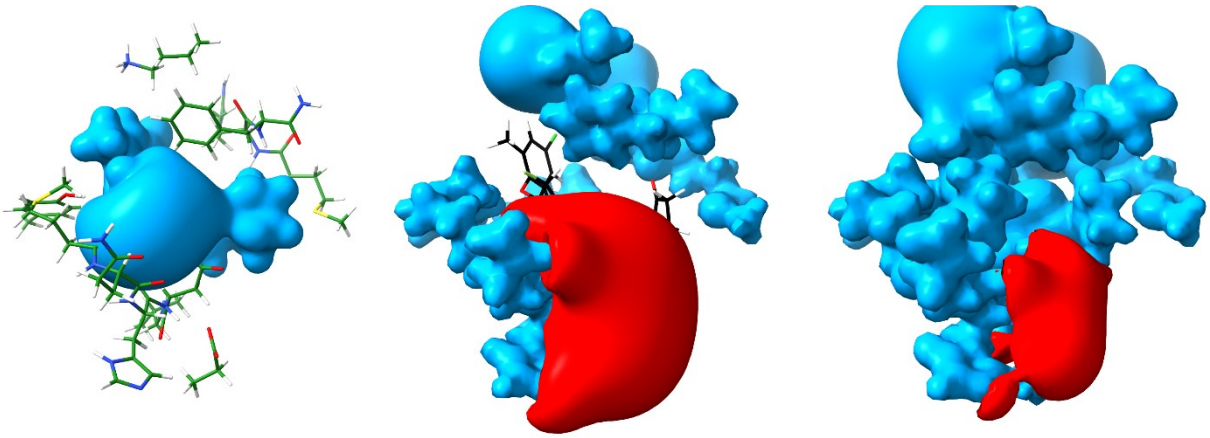


Fig. S10. ESP maps, where positive isosurfaces (+0.15 a.u.) are coloured in blue and negative isosurfaces (-0.15 a.u.) are coloured in red . The maps are as follow free ligand (left), free protein (centre) and protein-ligand complex (right).

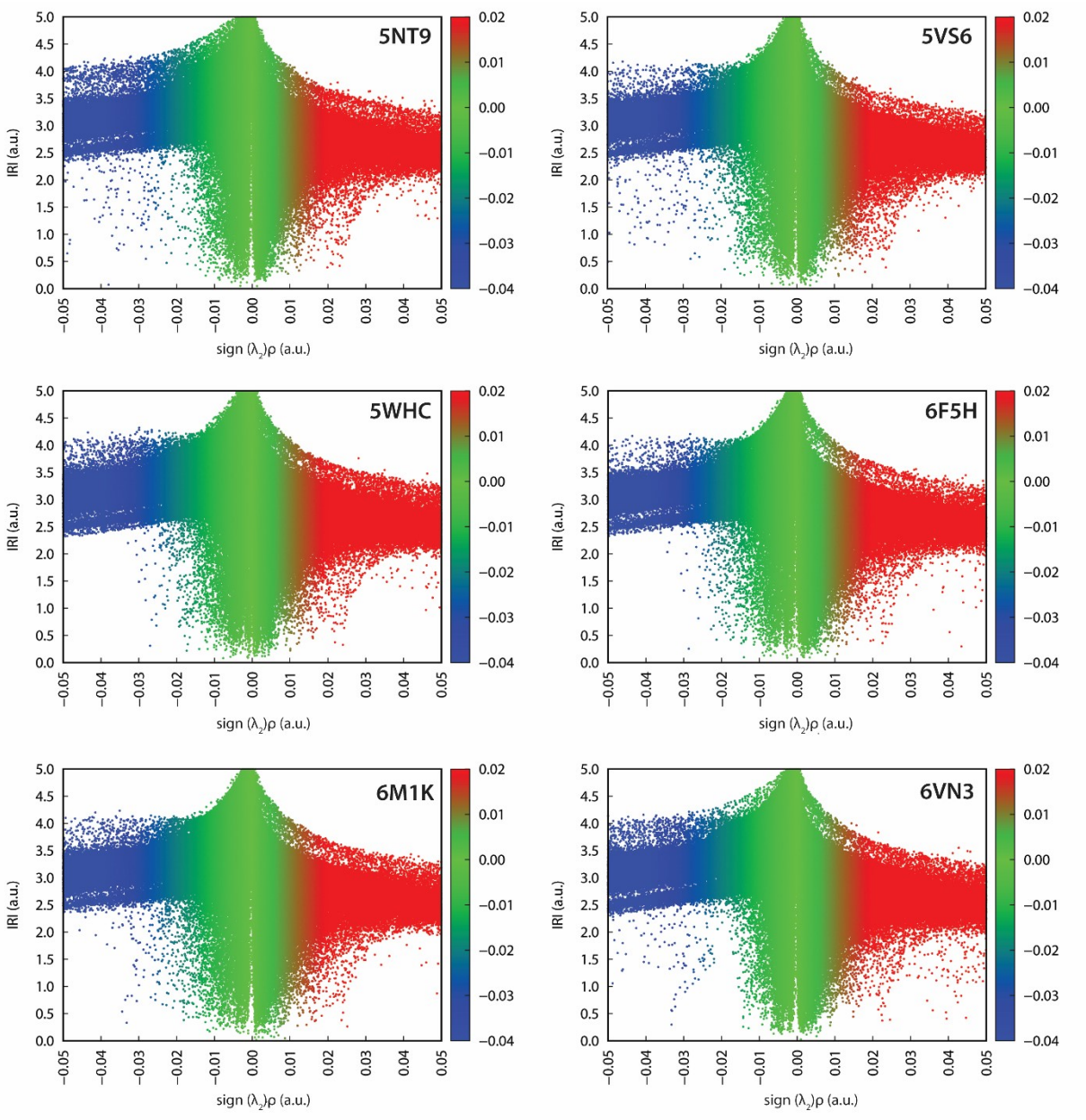


Fig. S11. IRI scatter plots of USP7 complexes. Values of IRI itself are in a.u., lower values indicate meaningful interactions. The sign(λ_2) denotes the sign of the second largest eigenvalue of ED Hessian. An RGB colour gradient is applied, where blue indicated notable interactions (i.e., H-bond or halogen bonds), green vdW interactions, and red repulsion (i.e., clashes, steric effects in rings). Large magnitude of sign(λ_2)ρ implies a relatively strong interaction, while regions with low ρ hence small sign(λ_2)ρ does not participate in considerable interactions.

IRI scatter plots reveal consistent interaction patterns across USP7 complexes (**Fig. 10**), characterised by abundant weak van der Waals interactions (green points) across binding interfaces, clustered strong interactions (blue points) indicating localised binding hotspots, and dispersed repulsive interactions (red points) varying across interfaces. Both **5VS6** and **5N9T** show higher blue point density, suggesting stronger hydrogen bonding compared to **6FH5** and **6M1K**. Complex **6F5H** exhibits concentrated repulsive interactions at higher sign(λ_2) values with reduced stabilising interactions, reflecting localised steric clashes from its strained five-membered ring rather than interface destabilisation. The relatively low density of stabilising interactions in **5WHC** compared to other ligands correlates with its much lower affinity to USP7 (micromolar vs. nanomolar). Complex **6VN3** exhibits a distinctive interaction profile with bimodal distribution of strong interactions (blue points): one cluster at lower IRI values and another at slightly higher values, indicating two distinct interaction types. The primary cluster represents various hydrogen bonds, while the secondary cluster likely originates from thieno[3,2-b]bipyridine ring interactions with adjacent residues through S···O and S···π contacts.

References:

1 P. R. Leger, D. X. Hu, B. Biannic, M. Bui, X. Han, E. Karbarz, J. Maung, A. Okano, M. Osipov, G. M. Shibuya, K. Young, C. Higgs, B. Abraham, D. Bradford, C. Cho, C. Colas, S. Jacobson, Y. M. Ohol, D. Pookot, P. Rana, J. Sanchez, N. Shah, M. Sun, S. Wong, D. G. Brockstedt, P. D. Kassner, J. B. Schwarz and D. J. Wustrow, *J. Med. Chem.*, 2020, **63**, 5398–5420.

2 G. Gavory, C. R. O'Dowd, M. D. Helm, J. Flasz, E. Arkoudis, A. Dossang, C. Hughes, E. Cassidy, K. McClelland, E. Odrzywol, N. Page, O. Barker, H. Miel and T. Harrison, *Nat. Chem. Biol.*, 2018, **14**, 118–125.

3 C. R. O'Dowd, M. D. Helm, J. S. S. Rountree, J. T. Flasz, E. Arkoudis, H. Miel, P. R. Hewitt, L. Jordan, O. Barker, C. Hughes, E. Rozycka, E. Cassidy, K. McClelland, E. Odrzywol, N. Page, S. Feutren-Burton, S. Dvorkin, G. Gavory and T. Harrison, *ACS Med. Chem. Lett.*, 2018, **9**, 238–243.

4 M. Li, S. Liu, H. Chen, X. Zhou, J. Zhou, S. Zhou, H. Yuan, Q.-L. Xu, J. Liu, K. Cheng, H. Sun, Y. Wang, C. Chen and X. Wen, *Eur. J. Med. Chem.*, 2020, **199**, 112279.

5 A. P. Turnbull, S. Ioannidis, W. W. Krajewski, A. Pinto-Fernandez, C. Heride, A. C. L. Martin, L. M. Tonkin, E. C. Townsend, S. M. Buker, D. R. Lancia, J. A. Caravella, A. V. Toms, T. M. Charlton, J. Lahdenranta, E. Wilker, B. C. Follows, N. J. Evans, L. Stead, C. Alli, V. V. Zarayskiy, A. C. Talbot, A. J. Buckmelter, M. Wang, C. L. McKinnon, F. Saab, J. F. McGouran, H. Century, M. Gersch, M. S. Pittman, C. G. Marshall, T. M. Raynham, M. Simcox, L. M. D. Stewart, S. B. McLoughlin, J. A. Escobedo, K. W. Bair, C. J. Dinsmore, T. R. Hammonds, S. Kim, S. Urbé, M. J. Clague, B. M. Kessler and D. Komander, *Nature*, 2017, **550**, 481–486.

6 I. Lamberto, X. Liu, H.-S. Seo, N. J. Schauer, R. E. Iacob, W. Hu, D. Das, T. Mikhailova, E. L. Weisberg, J. R. Engen, K. C. Anderson, D. Chauhan, S. Dhe-Paganon and S. J. Buhrlage, *Cell Chem. Biol.*, 2017, **24**, 1490-1500.e11.

7 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. McCleland, 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 and I. E. Wertz, *Nature*, 2017, **550**, 534–538.

8 P. Di Lello, R. Pastor, J. M. Murray, R. A. Blake, F. Cohen, T. D. Crawford, J. Drobnick, J. Drummond, L. Kategaya, T. Kleinheinz, T. Maurer, L. Rougé, X. Zhao, I. Wertz, C. Ndubaku and V. Tsui, *J. Med. Chem.*, 2017, **60**, 10056–10070.