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

Supplementary Materials

Figure S1. Catalytic triad of trypsin in 3FP6 (PDB ID). Pictures A and B correspond to alternative locations A and B of the catalytic Ser195, respectively. Hydrogens are not shown.

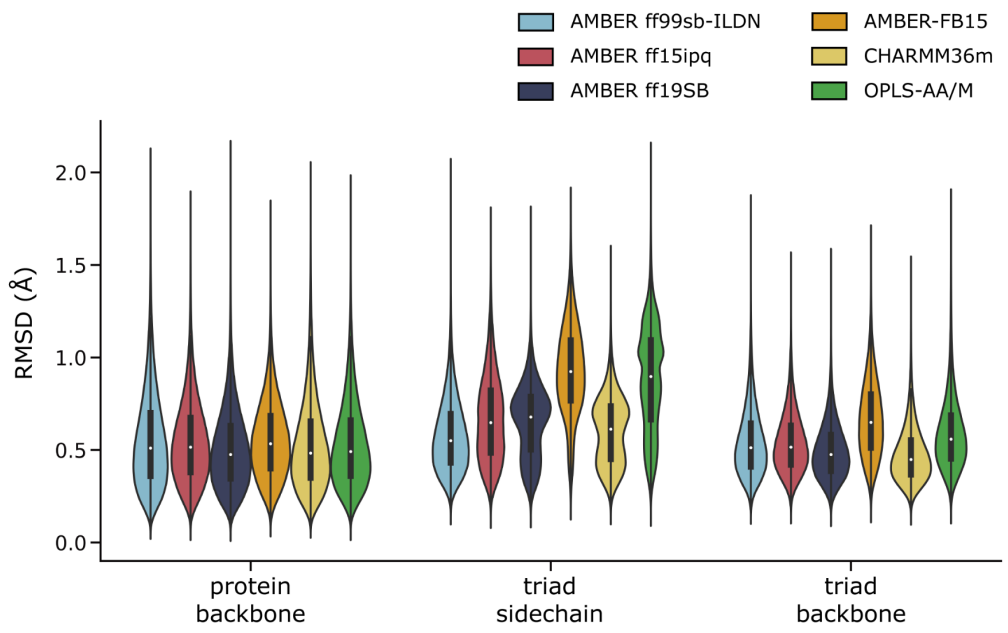

Figure S2. Distribution of RMSD values, calculated for (1) backbone atoms of proteins, (2) sidechain heavy atoms and (3) backbone atoms of catalytic triad residues.

Table S1. Distances that had been selected as a descriptors for the clusterization process

1. His57 N – Asp102 $O\delta_1$
2. His57 O – His57 Nδ
3. His57 Nδ – Asp102 $O\delta_2$
4. His57 Nδ – Asp102 $O\delta_1$
5. His57 Nε – Ser195 Oγ
6. Asp102 N – Leu99 O
7. Asp102 N – Asp102 $O\delta_1$
8. Asp102 O – Thr229 Oγ
9. Asp102 O – Asp102 $O\delta_1$
10. Asp102 $O\delta_1$ – Ala56 N
11. Asp102 $O\delta_2$ – Ser214 Oγ
12. Ser195 N – Lys15 (P1) O
13. Ser195 Oγ – Lys15 (P1) N
14. Ser195 Oγ – Lys15 (P1) O
15. Ser195 Oγ – Ala16 (P2) N
16. Ser195 Oγ – Ser195 O
17. Ser195 O – Gly43 N
18. Ser214 O – Lys15 (P1) N
19. Gly36 O – Ala16 (P2) N
20. Gly193 N – Lys15 (P1) O
21. Ser214 Oγ – Ser214 O

Table S2. Productive and non-productive clusters

Cluster 1	non-productive
Cluster 2	non-productive
Cluster 3.1	non-productive
Cluster 3.2	non-productive
Cluster 4.1	non-productive
Cluster 4.2	probably productive
Cluster 5	productive
Cluster 6	productive

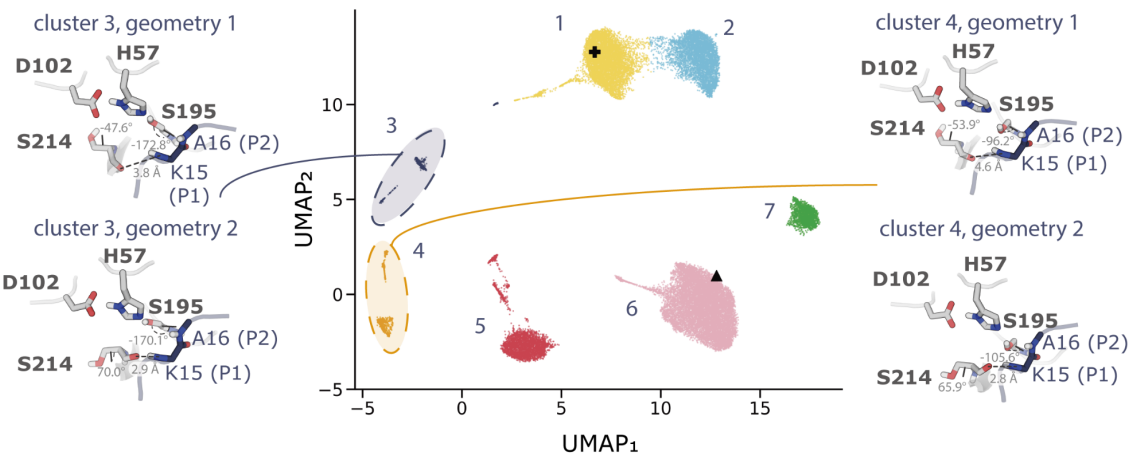

Figure S3. Major geometries of Michaelis complex presented in cluster 3 and cluster 4. Both clusters 3 and 4 are presented by frames simulated with the OPLS-AA/M force field.

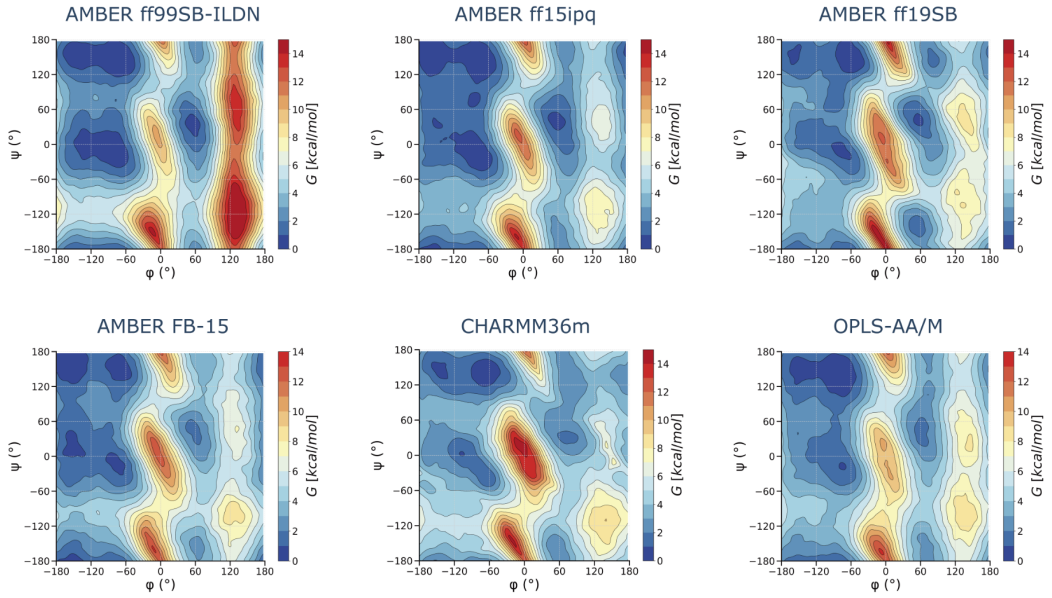

Figure S4. Free energy surfaces projected on φ-ψ axes. The profiles were obtained for Ace-Ser-NME peptide via 3D simulation of metadynamics with different force fields. One color level corresponds to 1 kcal/mol step.

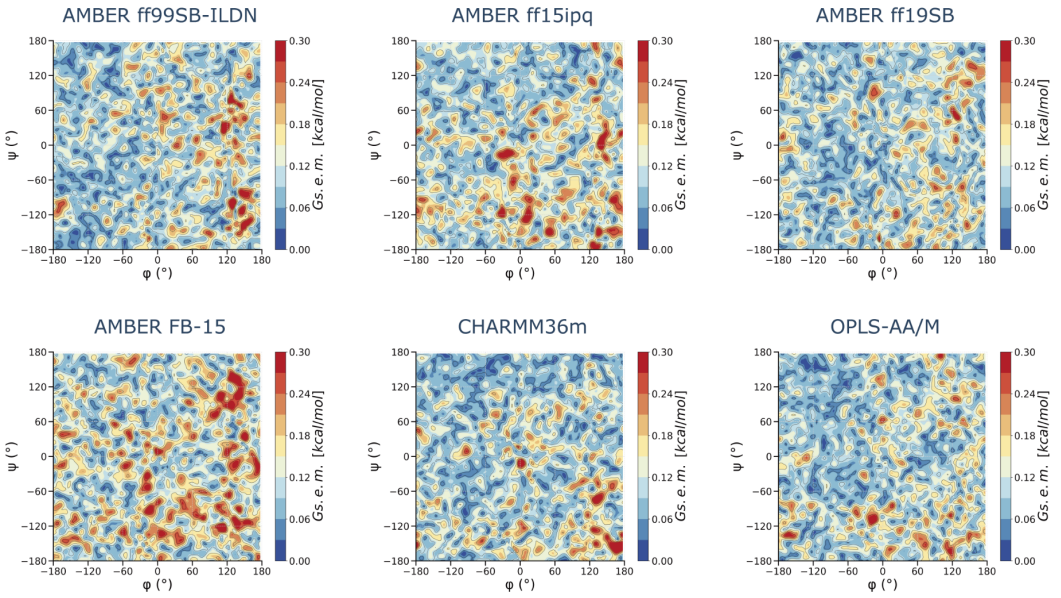

Figure S5. Reproducibility in metadynamics simulations. Reproducibility is shown as free energy profile standard errors of mean from 5 profiles obtained from 5 independent replicas.

Table S3. Atomic charges derived from six non-polarizable force fields

		Amber ff99sb-ildn	Amber ff15ipq	Amber ff19sb	Amber FB-15	CHARMM 36m	OPLS-AA/M
Ser	N	-0.4157	-0.51112	-0.4157	-0.4157	-0.47	-0.5
Ser	O	-0.5679	-0.58991	-0.5679	-0.5679	-0.51	-0.5
Ser	Oγ	-0.6546	-0.56858	-0.6546	-0.6546	-0.66	-0.683
Ser	Hγ	0.4275	0.3849	0.4275	0.4275	0.43	0.418
His	Nδ	-0.3811	-0.35736	-0.3811	-0.3811	-0.36	-0.57
His	Hδ	0.3649	0.37921	0.3649	0.3649	0.32	0.418
His	Nε	-0.5727	-0.62533	-0.5727	-0.5727	-0.7	-0.49
Lys	N	-0.3479	-0.3908	-0.3479	-0.3479	-0.47	-0.5
Lys	H	0.2747	0.31584	0.2747	0.2747	0.31	0.3
Lys	O	-0.5894	-0.63388	-0.5894	-0.5894	-0.51	-0.5
Asp	O	-0.5819	-0.63086	-0.5819	-0.5819	-0.51	-0.5
Asp	$Oδ_2$	-0.8014	-0.85227	-0.8014	-0.8014	-0.76	-0.8
Thr	Oγ	-0.6761	-0.57965	-0.6761	-0.6761	-0.66	-0.683
Thr	Hγ	0.4102	0.37731	0.4102	0.4102	0.43	0.418

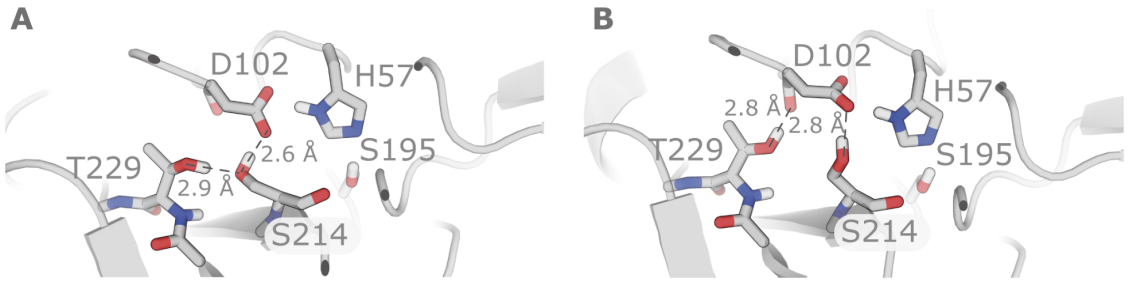

Figure S6. Competition for being acceptor of h-bond from Thr229 Oγ between Asp102 O and Ser214 Oγ. Here represented geometries corresponding to the (A) cluster 5, and (B) cluster 6.

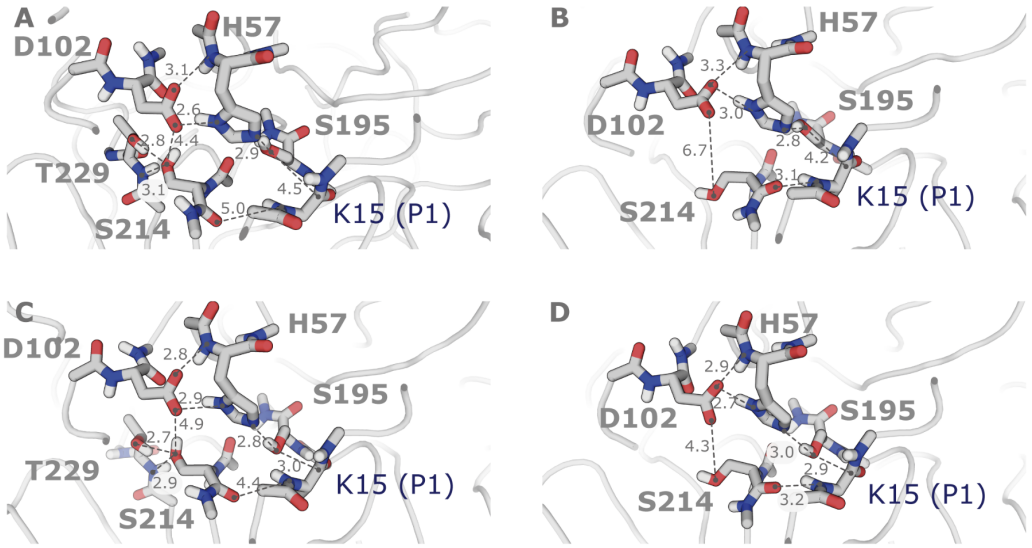

Fig. S7. Representative geometries of the active site produced by OPLS-AA/M. Subclusters 3.1 (A), 3.2 (B), 4.1 (C), 4.2 (D) are presented. Interatomic distances (Å) shown by dark-gray dashed lines. Enzyme labels are dark-gray (bold font), while substrate labels are dark-blue (regular font).

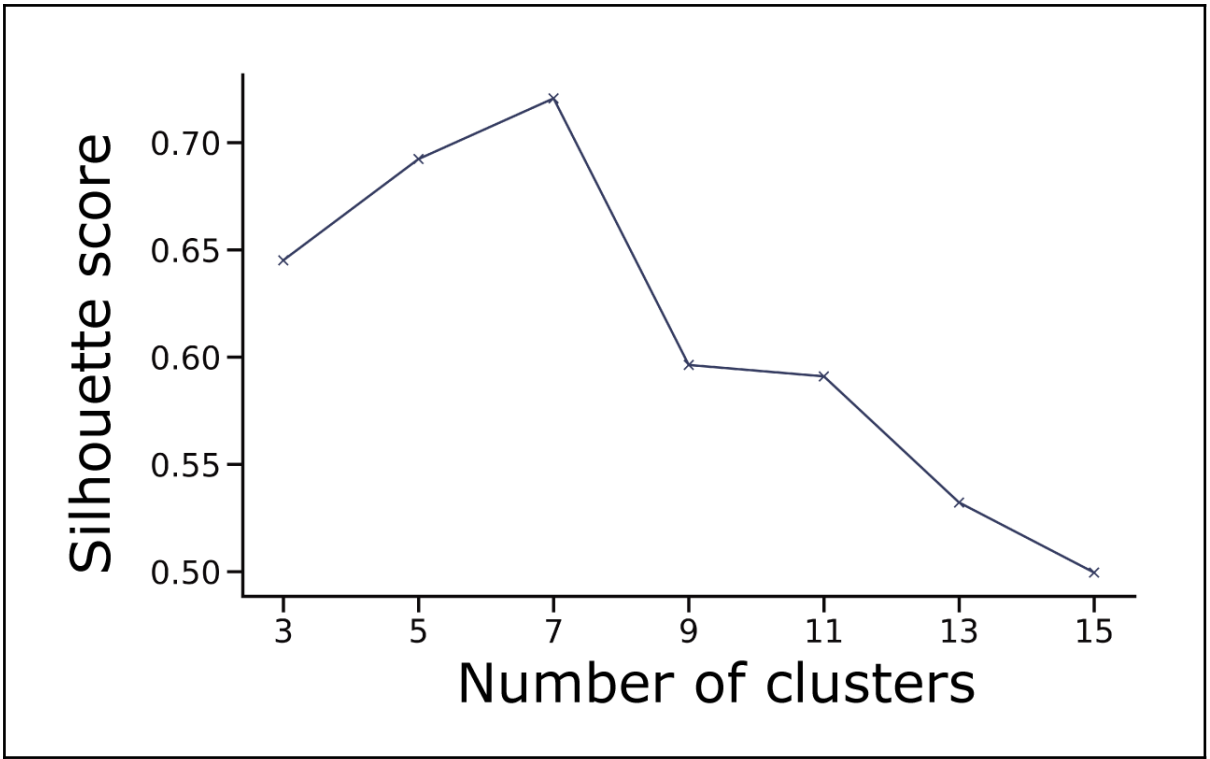

Figure S8. A change of Silhouette score with the number of clusters in the K-means algorithm. The best value is achieved when the number of clusters is equal to 7.

Table S4. Three-point water models used in simulations with every force field

force field	σ_{O2} (Å)	ε_{O2} (kJ/mol)	r_{OH} (Å)	r_{HH} (Å)
Amber ff99sb-ildn, Amber ff19sb, OPLS-AA/M (TIP3P)	3.15061	0.636386	0.9572	1.5139
Amber ff 15 ipq (SPC/Eb)	3.16557	0.650629	1.01	1.64933
Amber FB-15 (TIP3P-FB)	3.17796456355	0.652143528104	1.01181082494	1.63868385147
CHARMM36m (TIPS3P)	3.15057422683	0.6363864	0.9572	1.5139