Supplementary Information for

Carbapenem antibiotic inhibiting a mammalian serine protease: structure of the acylaminoacyl peptidase meropenem complex

Anna J. Kiss-Szemán[1], Luca Takács[1], Zoltán Orgován[2], Pál Stráner[3], Imre Jákli[1,3], Gitta Schlosser[4], Simonas Masiulis[5], Veronika Harmat[1,3*], Dóra K. Menyhárd[1,3*] and András Perczel[1,3*]

[1] Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös Loránd University, Budapest – 1117, Hungary.
[2] Medicinal Chemistry Research Group, Research Centre for Natural Sciences, Budapest - 1117, Hungary
[3] ELKH-ELTE Protein Modelling Research Group, Eötvös Loránd Research Network, Budapest – 1117, Hungary.
[4] ELKH-ELTE Lendület Ion Mobility Mass Spectrometry Research Group, Institute of Chemistry, Eötvös Loránd University, Budapest, Hungary
[5] Materials and Structural Analysis Division, Thermo Fisher Scientific, Eindhoven, The Netherlands

Corresponding authors: Dóra K. Menyhárd, András Perczel, Veronika Harmat
*e-mail: dora.k.menyhard@ttk.elte.hu; perczel.andras@ttk.elte.hu; veronika.harmat@ttk.elte.hu

This file includes:

Figures S1–S14
Tables S1-S2

Figure S1. Mass spectrometric analysis of meropenem binding to AAP. A) Intact LC-MS analysis of the uncomplexed AAP (upper Figure, uncomplexed AAP) as well as the AAP protein complexed with meropemen (lower Figure, **AAP-MEPM**). UV detection was completed at 280 nm. **B**) The intact MS-spectra of AAP (upper Figure, uncomplexed AAP) as well as the AAP complexed with meropenem (lower Figure, AAP-MEPM). Detected average molecular masses were the following: AAP protein: 81300 Da and AAP complexed with meropenem: 81656 Da. Theoretical molecular mass of the AAP protein is 81286 Da.

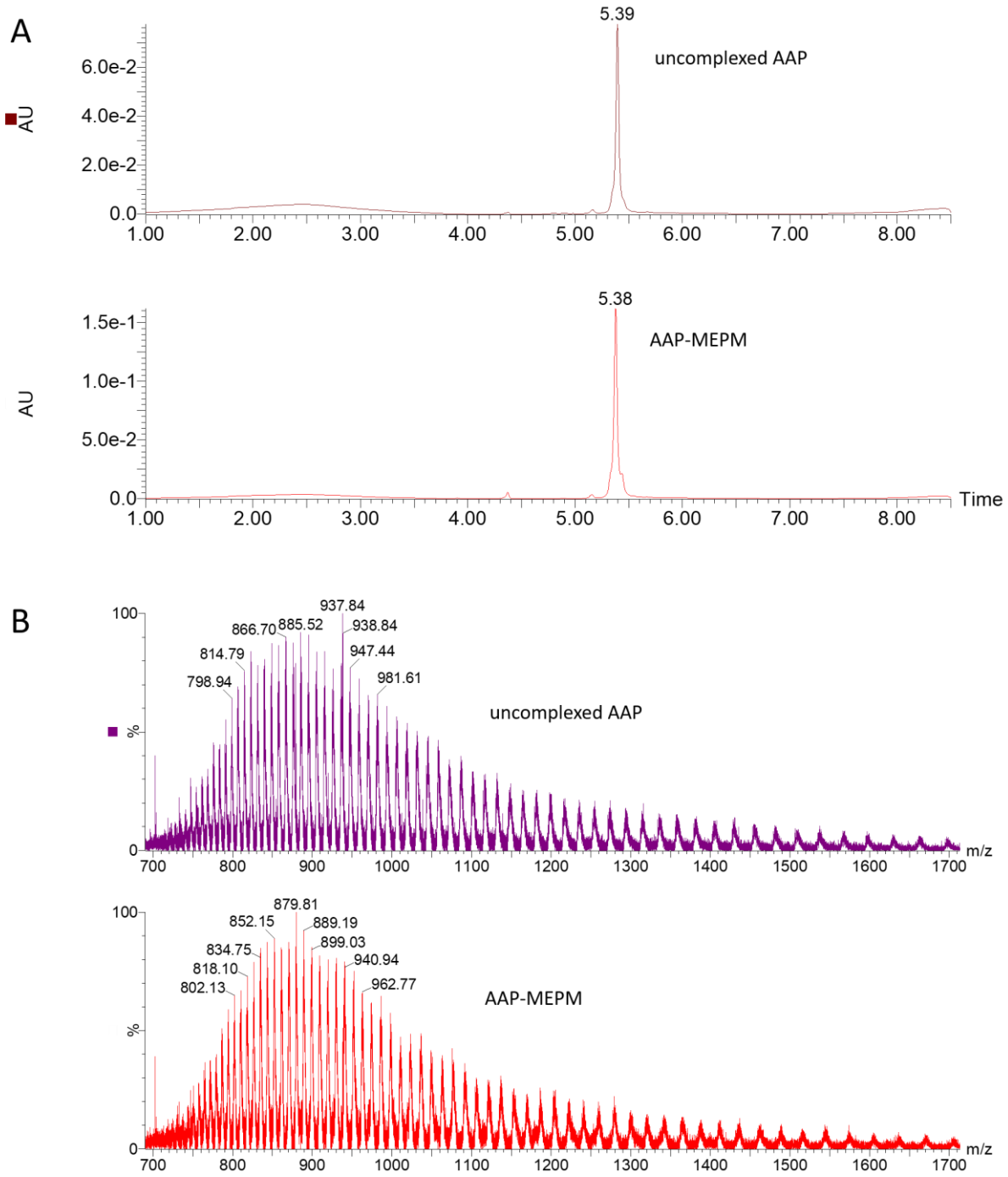


Figure S2. Workflow of cryo-EM data processing. A) Data processing workflow using cryoSPARC, showing 2D classes (left panel) and a typical micrograph for AAP-MEPM complex (right panel, particles circled in red). **B**) Maps obtained applying C1 and C2 point group symmetries in 3D refinement **C**) Local resolution maps contoured at a 0.12 threshold level (left), calculated using ResMap. Local resolution maps with the fitted models (right, derived using Phenix Dock) **D**) Local resolution of cryo-EM map (C2 symmetry, 0.12 threshold level) around meropenem (MEPM).

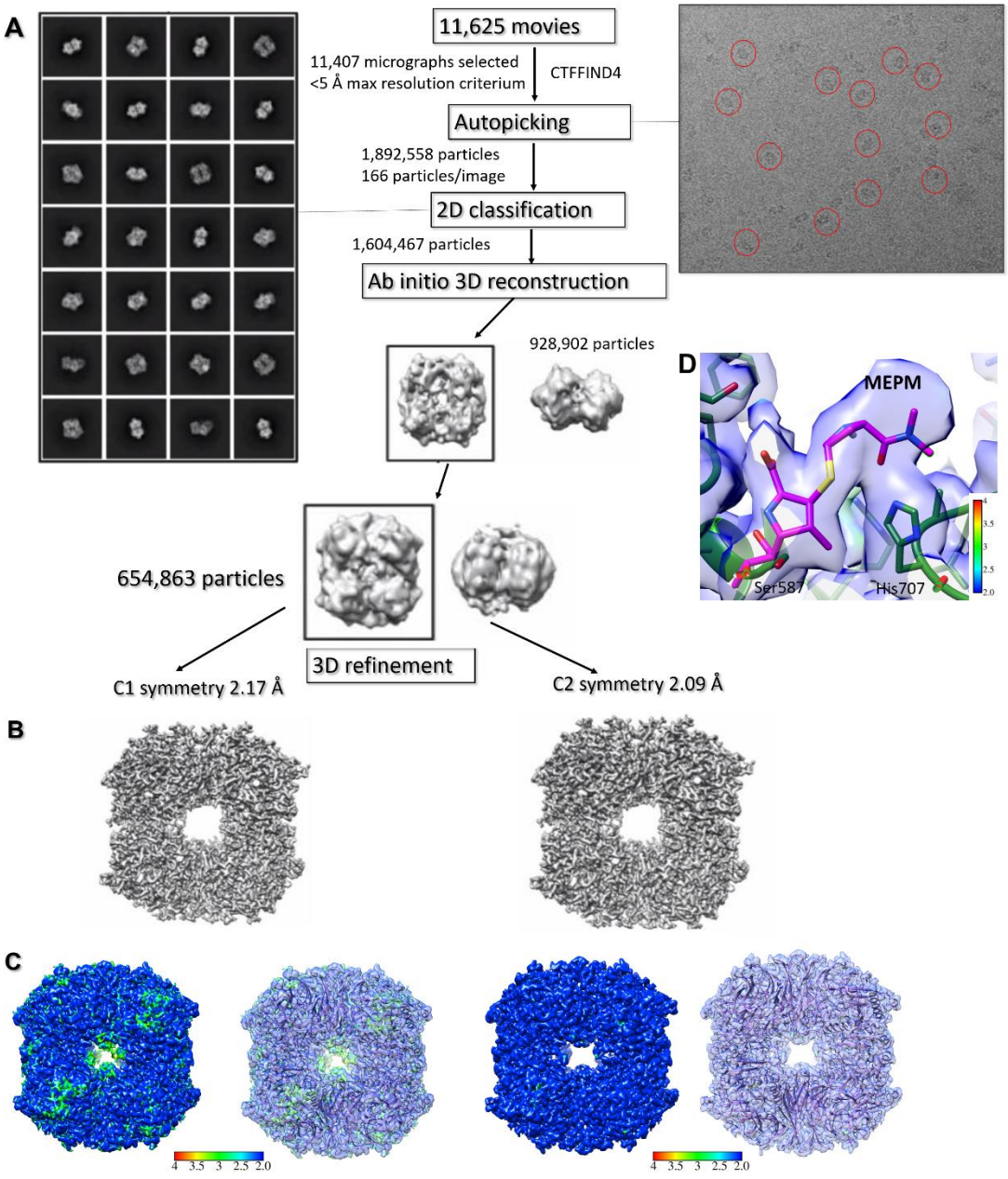


Figure S3. **Validation of the cryo-EM data processing**. **A**) The calculated map to map FSC curves (from cryoSPARC). **B**) The half-maps FSC curves were calculated using Phenix and **C**) the calculated model to map FSC curves.

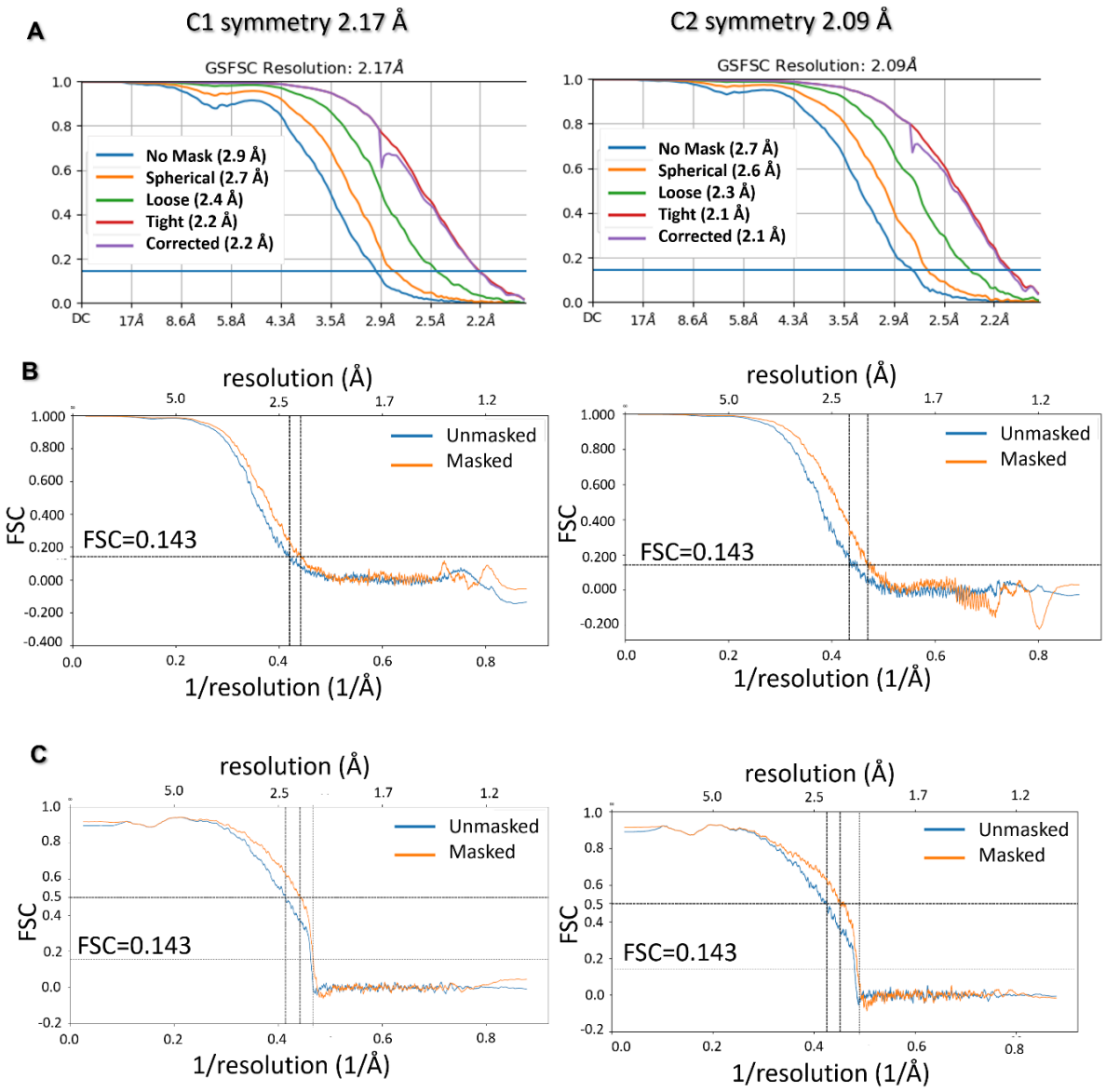


Table S1. **Cryo-EM data collection, refinement and validation statistics**

Data collection		
Magnification	x130,000	
Voltage (kV)	300	
Electron exposure (e/Å^2)	41	
Defocus range (μm)	0.5-1.5	
Pixel size (Å)	0.96	
Data processing		
Accession codes	**EMD- 14149, PDB id: 7qun**	*not deposited*
Symmetry imposed	C2	C1
Initial particle images (no.)	1,892,558	1,892,558
Final particle images (no.)	654,863	654,863
Map resolution (Å)	2.09	2.17
FSC threshold	0.143	0.143
Refinement		
Model resolution (Å)	2.10	2.17
Model composition		
Non-hydrogen atoms	21094	21094
Protein residues	2773	2773
B-factors (min/max/mean)		
Protein	11.72/44.21/19.09	9.33/83.03/32.62
ligand	19.67/20.75/20.21	33.41/56.09/43.01
r.m.s. deviations		
Bond length (Å)	0.003	0.005
Bond angles (°)	0.675	0.784
Validation		
MolProbility score	1.32	1.27
Clashscore	4.16	3.45
EMRinger score	6.15	5.46
Rotamer outliers (%)	16 (0.58%)	22 (0.79%)
CaBLAM outliers (%)	1.73	1.96
Ramachandran plot (%)		
Outliers	0	0
Allowed	72 (2.63%)	76 (2.74%)
Favored	2701 (97.37%)	2697 (97.26%)

Figure S4. Comparison of the bacterial β-lactamase (BlaC)-meropenem and L,D-transpeptidase-meropenem adducts. **a**) The catalytic site of uncomplexed *Pseudomonas aeruginosa* D,D-transpeptidase (PDB id: 3pbn, grey) and *Mycobacterium tuberculosis* D,D transpeptidase in covalent complex with meropenem (PDB id: 6kgs, blue) **b**) The catalytic site of uncomplexed BlaC enzyme, a serine-β-lactamase (PDB id 2gdn, grey) and meropenem bound structure of BlaC (PDB id: 3dwz, pink). The antibiotic acylates the catalytic Ser84. **c**) In the uncomplexed form of *Mycobacterium tuberculosis* (3,3) L,D-transpeptidase type 5 (Ldt_{Mt5}, PDB ids: uncomplexed form: 4z7a, yellow and meropenem adduct: 47fq, orange) two rotamers of Glu328 could be distinguished, one of these stabilizes (along with Asn362) an inactive conformation of His342. **d**)The long substrate-recognition loop of *Mycobacterium tuberculosis* (3,3) L,D-transpeptidase type 2 (uncomplexed Ldt_{Mt2} (yellow) is dislocated upon binding of meropenem (Ldt_{Mt2}-N140, green). **e**) Rotation of meropenem leads to the formation of three new H-bonds, stabilizing its binding conformation and slowing hydrolysis.

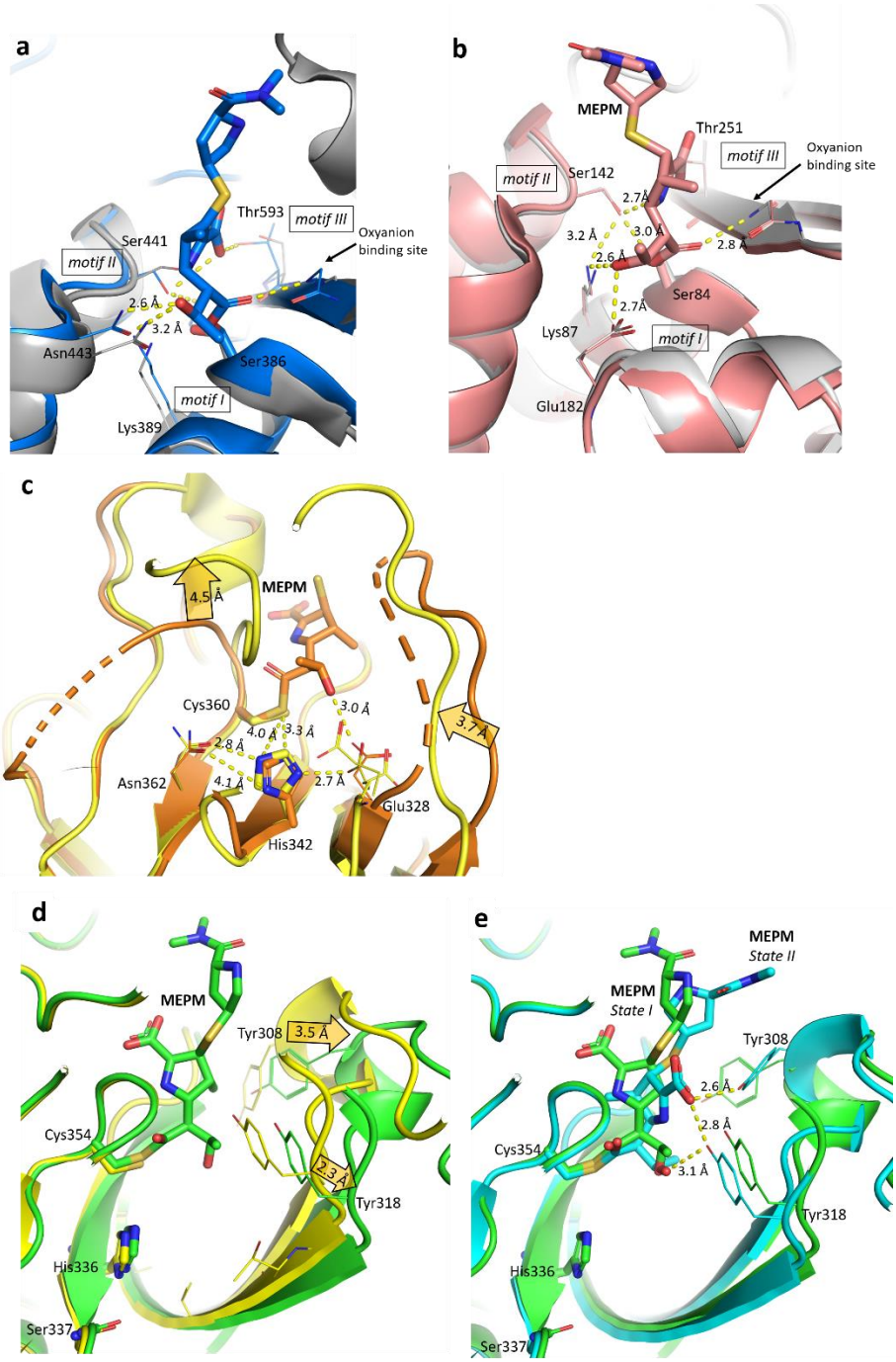


Figure S5. The chemical structure of Meropenem. a) Structure of intact meropenem (MEPM) with closed beta-lactam ring (atom labels according to PDB ligand id: DWZ). **b**) The two molecular structures of meropenem related to its binding to pAAP: the closed (MCMM-derived conformation, light pink) and open (cryo-EM experiment, magenta) forms. The structures are aligned with the dihydropyrrol ring. (IUPAC: ((4R,5S,6S)-3-(((3S,5S)-5-(Dimethylcarbamoyl)pyrrolidin-3-yl)thio)-6-((R)-1-hydroxyethyl)-4-methyl-7-oxo-1-azabicyclo[3.2.0]hept-2-ene-2-carboxylic acid).

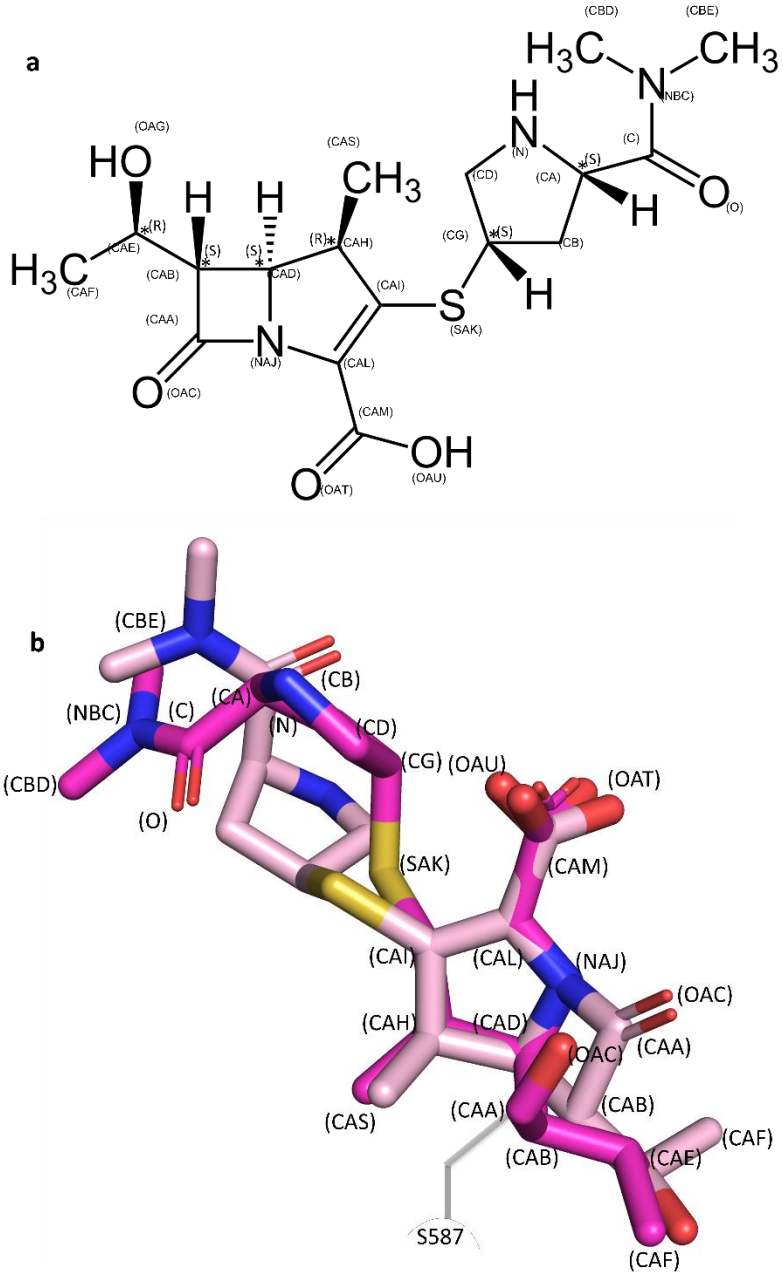


Table S2. Bond angles and dihedral angles of meropenem in its intact (closed β-lactam) and covalently-ligated (acyl-enzyme) states (with open β-lactam). MCMM-derived, enzyme-bound conformation of intact meropenem (second column) compared to the open β-lactam ring conformation of the covalently bound complex (cryo-EM experiment, third column) and their difference. Coloring codes the extent of the change (light to darker shades). (Atom names are compiled in Extended Data Figure 3)

a) Bond angles given in degrees

Atoms	Closed ring (unbound)	Open ring (covalently bound)	Δ
OAC-CAA-CAB	134.7	120.3	14.4
CAA-CAB-CAD	85.5	120.1	-34.6
CAA-CAB-CAE	114.9	100.7	14.2
CAB-CAE-CAF	118.1	112.4	5.7
CAB-CAE-OAG	109.4	110.1	-0.7
OAG-CAE-CAF	107.2	105.0	2.2
CAB-CAD-NAJ	90.0	113.3	-23.3
CAB-CAD-CAH	123.6	114.4	9.2
NAJ-CAL-CAI	111.2	108.4	2.8
NAJ-CAL-CAM	119.3	124.6	-5.3
CAL-CAM-OAT	115.5	119.2	-3.7
CAL-CAM-OAU	116.2	120.7	-4.5
CAM-CAL-CAI	129.5	125.6	3.9
CAL-CAI-CAH	109.1	105.6	3.5
CAD-CAH-CAI	103.6	104.3	-0.7
CAD-NAJ-CAL	113.0	113.9	-0.9
CAD-CAH-CAS	115.3	116.6	-1.3
CAI-CAH-CAS	111.4	112.1	-0.7
CAH-CAI-SAK	121.8	110.4	11.4
CAL-CAI-SAK	128.7	111.8	16.9
CAI-SAK-CG	102.3	117.1	-14.8
SAK-CG-CB	113.1	108.9	4.2
SAK-CG-CD	112.3	109.7	2.6
CG-CD-N	105.1	108.6	-3.5
CD-N-CA	105.8	106.6	-0.8
N-CA-CB	102.9	107.2	-4.3
CA-CB-CG	105.6	103.4	2.2
N-CA-C	110.2	108.4	1.8
CB-CA-C	114.0	108.6	5.4
CA-C-O	118.4	117.5	0.9
CA-C-NBC	119.6	120.2	-0.6
O-C-NBC	122.0	122.4	-0.4
C-NBC-CBD	121.6	122.2	-0.6
C-NBC-CBE	125.0	119.5	5.5
CBD-NBC-CBE	113.2	118.3	-5.1

b) Dihedral angles given in degrees

Atoms	Closed ring (unbound)	Open ring (covalently bound)	Δ
CAI-SAK-CG-CD	-60.6	-43.8	-16.8
CAI-SAK-CG-CB	-179.1	-163.5	-15.6
CAH-CAI-SAK-CG	-76.4	-163.9	87.5
CAL-CAI-SAK-CG	111.4	-46.4	157.8
CAE-CAB-CAD-NAJ	-114.9	-70.3	-44.6
CAB-CAD-NAJ-CAL	-134.5	-108.6	-25.9
CB-CA-C-O	113.0	85.7	27.3
SAK-CD-CG-N	-147.3	-124.6	-22.7
SAK-CG-CB-CA	123.6	114.6	9.0
CAF-CAE-CAB-CAD	63.6	179.9	-116.3
OAG-CAE-CAB-CAD	-173.4	-63.5	-109.9
CAF-CAE-CAB-CAA	-33.7	55.1	-88.8
OAG-CAE-CAB-CAA	89.3	171.7	-82.4
OAC-CAA-CAB-CAE	-71.6	65.9	-137.5
OAC-CAA-CAB-CAD	172.2	-46.7	218.9
CAA-CAB-CAD-NAJ	0.6	40.0	-39.4
CAA-CAB-CAD-CAH	-102.3	-78.9	-23.4
CAB-CAD-CAH-CAS	-13.8	-19.0	5.2
CAB-CAD-CAH-CAI	108.1	105.2	2.9
CAD-CAH-CAI-SAK	175.8	133.6	42.2
CAD-CAH-CAI-CAL	-10.7	12.6	-23.3
NAJ-CAL-CAI-SAK	177.5	-121.8	299.3
CAS-CAH-CAI-SAK	-59.8	-99.3	39.5
CAS-CAH-CAI-CAL	113.8	139.7	-25.9
CAS-CAH-CAD-NAJ	-109.9	-143.1	33.2
CAH-CAI-CAL-NAJ	4.6	-1.8	6.4
CG-CD-N-CA	39.3	18.5	20.8
CD-N-CA-CB	-38.0	-23.2	-14.8
CD-CG-CB-CA	0.8	-7.3	8.1
CG-CB-CA-N	21.9	18.3	3.6
CB-CA-C-NBC	-68.5	-93.2	24.7
CD-N-CA-C	88.0	94.3	-6.3
N-CA-C-NBC	172.5	150.5	22.0
N-CA-C-O	-6.0	-30.6	24.6
O-C-NBC-CBD	-179.5	-178.8	-0.7
O-C-NBC-CBE	-5.1	0.5	-5.6

Figure S6. Rotamer analysis of catalytic His707 side chain with EMRinger. In EMRinger model-to-map agreement method the C_γ atom is rotated around the χ_1 angle (upper panels) of a side chain, interpolating the density value in the map as it rotates. Properly modeled side chains should have χ_1 density peaks at or near rotameric positions (60°, 180° and 300°). The angle of peak density is interpreted as the correct position of the side chain in dihedral space. The peak in the distribution represents the most likely position of the C_γ atom of the side chain**. a**) AAP-MEPM tetramer fitted into cryo-EM map refined in C1 symmetry and **b**) in C2 symmetry calculated by phenix.emringer. Upper panels show the χ_1 angle, the lower panels show the χ_2 angle of His707 residues in chains A-D. Based on this His707 appears to be in an on-rotameric, low energy state and the side chain only rotates around its χ_2 angle.

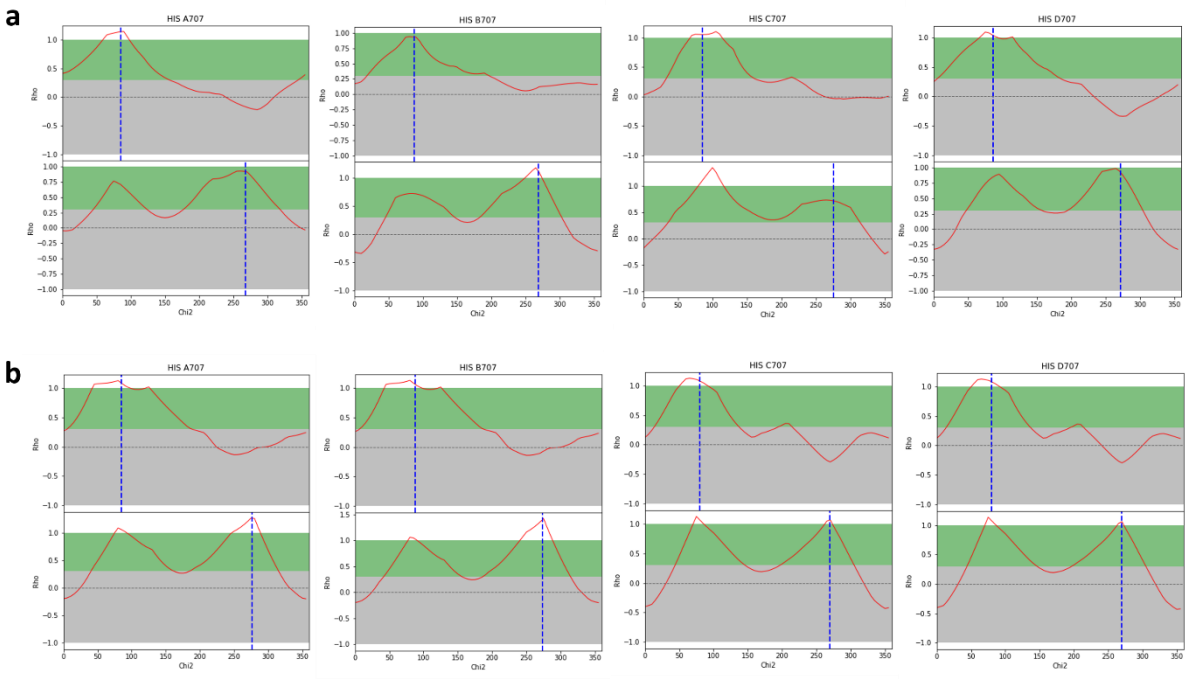
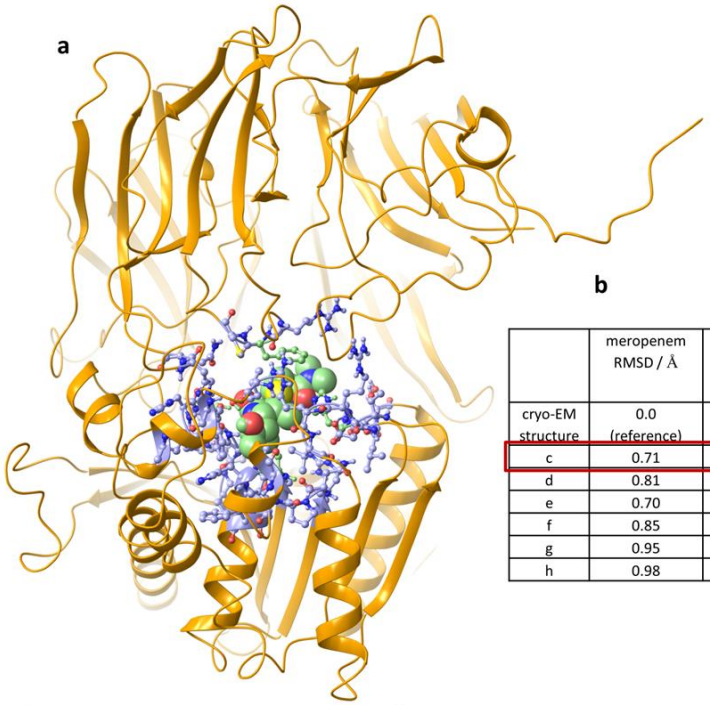


Figure S7. **Determining both the tautomerization and the protonation states of meropenem within the experimentally determined structure of the pAAP-meropenem complex using QM/MM calculations**. **a**) Overview of the QM/MM system (green: QM region (meropenem shown in space-filling representation): MEPM, Ser587, His707, Phe274, Ser512, 2 water molecules; lilac: freely moving MM region; gold: frozen MM region). **b**) Selecting the most likely state based on RMSD calculated for the non-hydrogen atoms of meropenem (2nd column), the distance between His707-$N_{\varepsilon 2}$ atom and the carbonyl oxygen of the amide substituent of the pyrrolidine ring of meropenem (3rd column) and the calculated conformational energy (4th column). **c-h**) The QM region of the optimized structure of the considered states (shown in dark green, overlaid on the experimental structure that is coloured according to **a**). The dihydro-pyrrole ring (indicated with cyan star) can assume imine (**c-e**) or enamine (**f-h**) form. Depending on the protonation state of its *N*-atom, the pyrrolidine ring (magenta star) can be neutral (single proton pointing toward Ser512 (**c, f**) or toward His707 (**d, g**)), or positively charged (doubly protonated: **e, h**).



b

	meropenem RMSD / Å	His707-NE2 ... O=C-meropenem distance / Å	conformational energy / kcal/mol
cryo-EM structure	0.0 (reference)	2.6	-
c	0.71	2.76	0
d	0.81	2.76	+1.18
e	0.70	3.45	+9.88
f	0.85	2.76	+1.02
g	0.95	2.75	+1.11
h	0.98	4.39	0

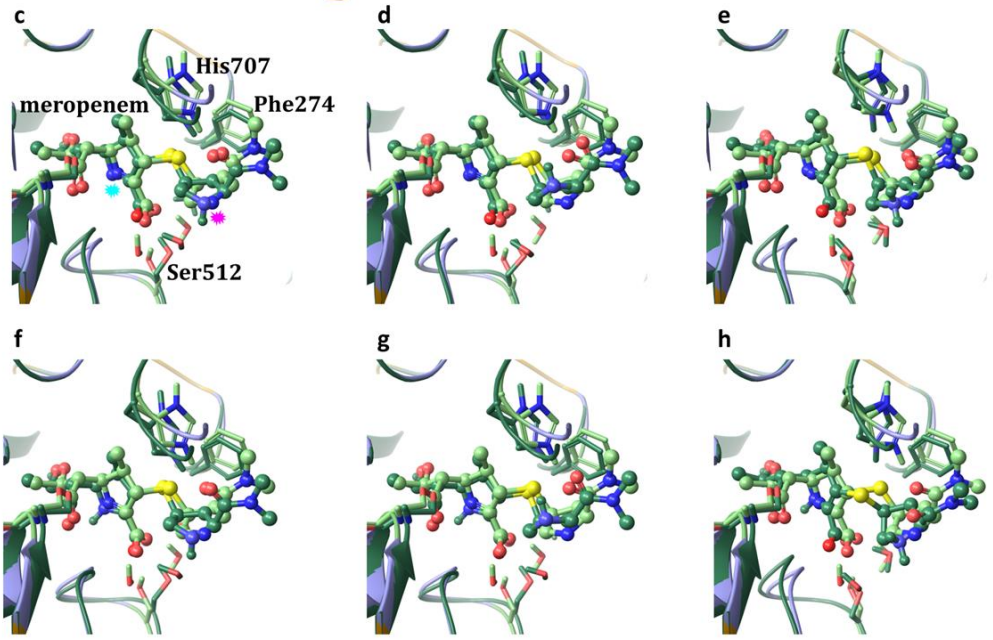


Figure S8. Monte Carlo Multiple Minimum conformational search conducted for intact and hydrolyzed (open) meropenem in water. a) The ensemble of low energy conformers of water solvated unhydrolyzed meropenem (global minimum arrangement and all those within a 3 kcal/mol energy window) as derived by MCMM calculation. **b**) Comparing the MCMM derived binding conformation of unhydrolyzed meropenem within the pAAP-meropenem complex derived by MCMM (pink) and a low energy conformer (+0.3 kcal/mol) from the water solvated ensemble, with high structure similarity (grey). **c**) The ensemble of low energy conformers of water solvated hydrolysed /open meropenem (global minimum arrangement and all those within a 3 kcal/mol energy window) as derived by MCMM calculation. **d**) Comparing the cryo-EM determined binding conformation of covalently bound meropenem within the pAAP-meropenem complex (pink) and the structurally most similar conformer from the low energy ensemble (+1.4 kcal/mol) of the water solvated system (grey).

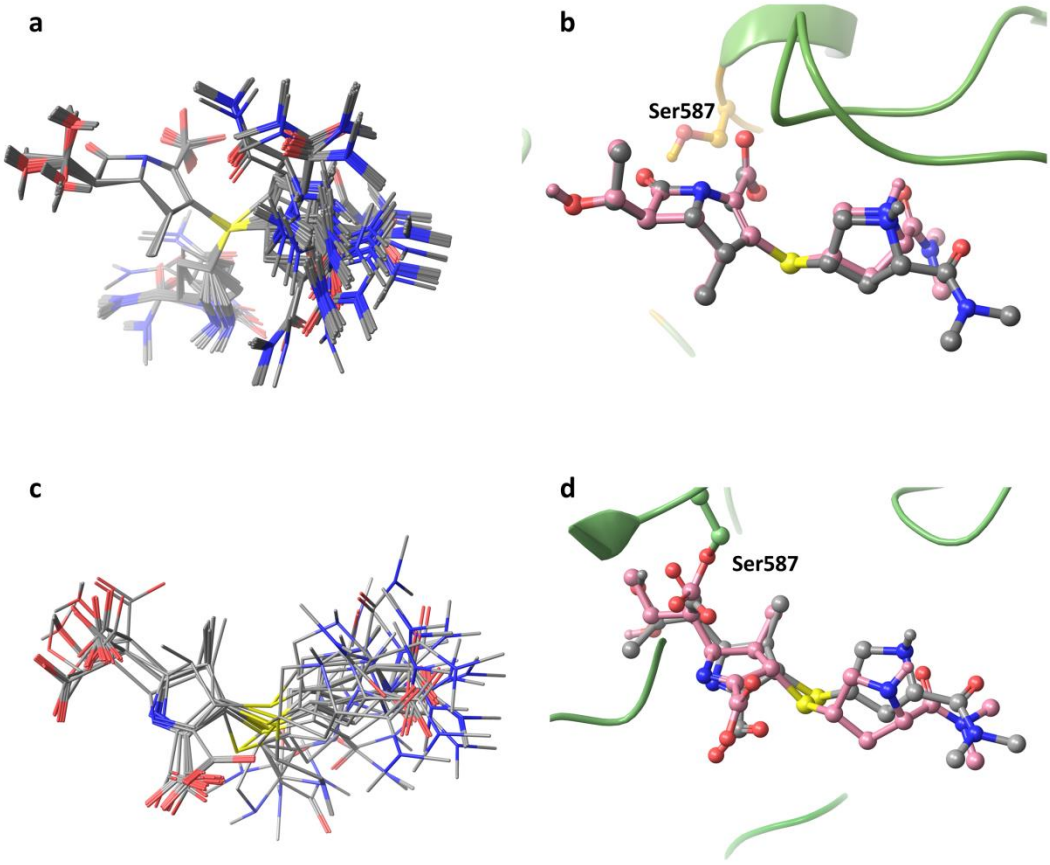


Figure S9. MCMM calculations concerning the non-covalent (pre-reaction) pAAP/VPA-G and pAAP/MEPM complexes. Overview of the minimum-energy conformers (a-b) and details of the local interactions (c-d). On all panels the hydrolase domain of pAAP is colored light green, the propeller dark green, while the residues of the catalytic triad in orange. VPA-G is shown in cyan, unhydrolyzed meropenem in pink.

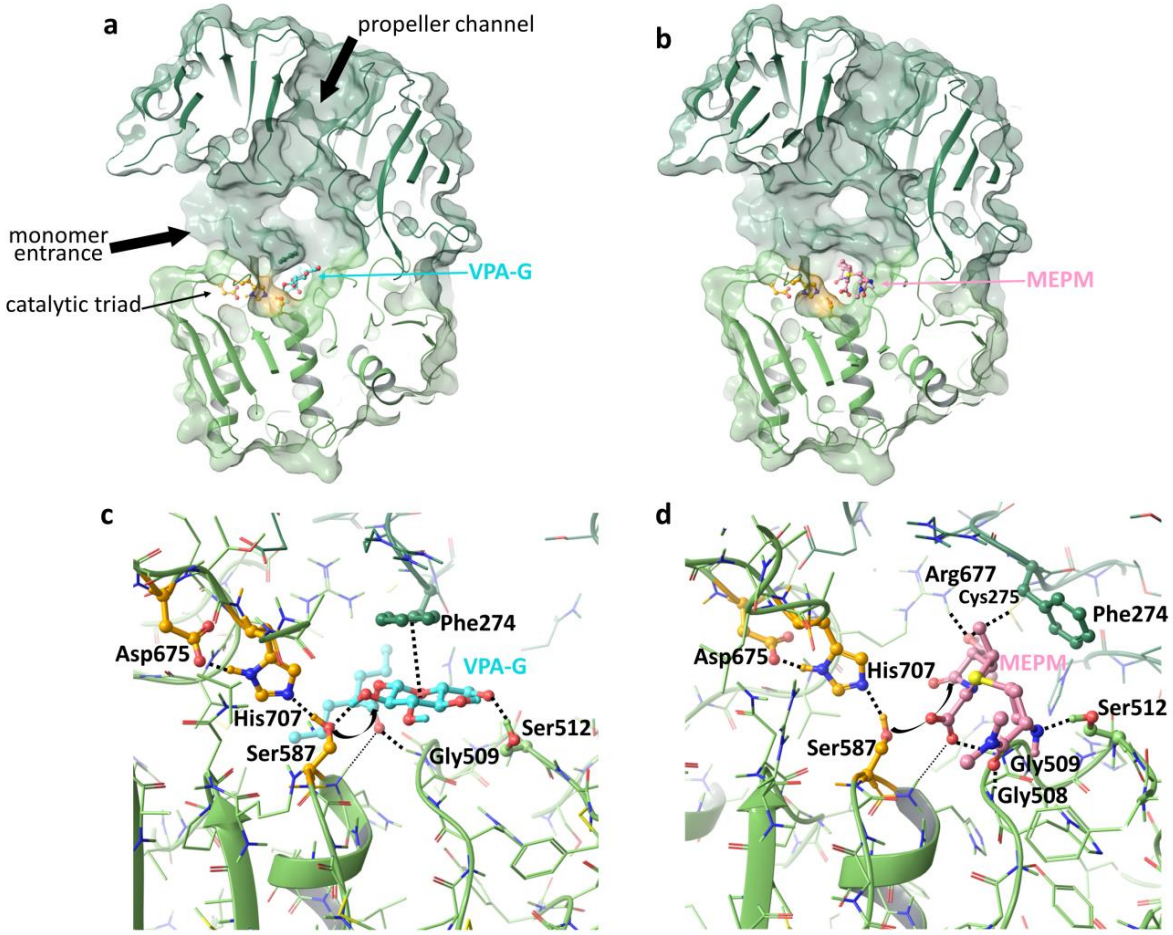


Figure S10. The proposed mechanism for VPA-G hydrolysis by AAP: nucleophilic attack of the carbonyl moiety of VPA-G, stabilization of the acyl-enzyme intermediate and subsequent deacylation by water (activated by His707) leading to the release of the products. The nucleophilic nature of Ser587 side chain OH is enhanced by the intact H-bond toward His707.

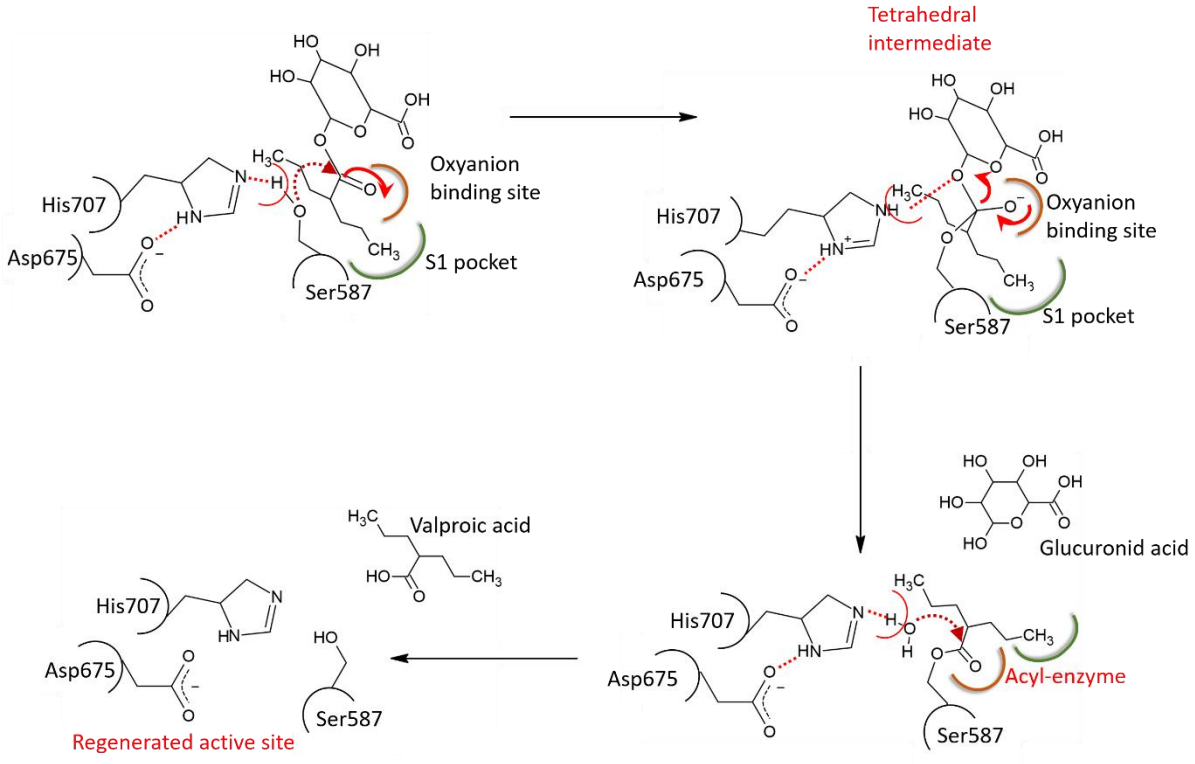


Figure S11. Sequence alignment of porcine and human AAP. (Uniprot proteinBLAST, codes P19205 and P13798) The sequence identity is 91.9% with 673 identical, 45 similar and 14 different amino acids (full length of the sequence is 732 amino acids). The active site (Ser587, Asp625, His707) is highlighted with red dots and rectangles. Blue dotted lines mark the borders of the hydrolase and the propeller domain, purple dotted lines mark the sequence of blade 3 insertion (Insert2).

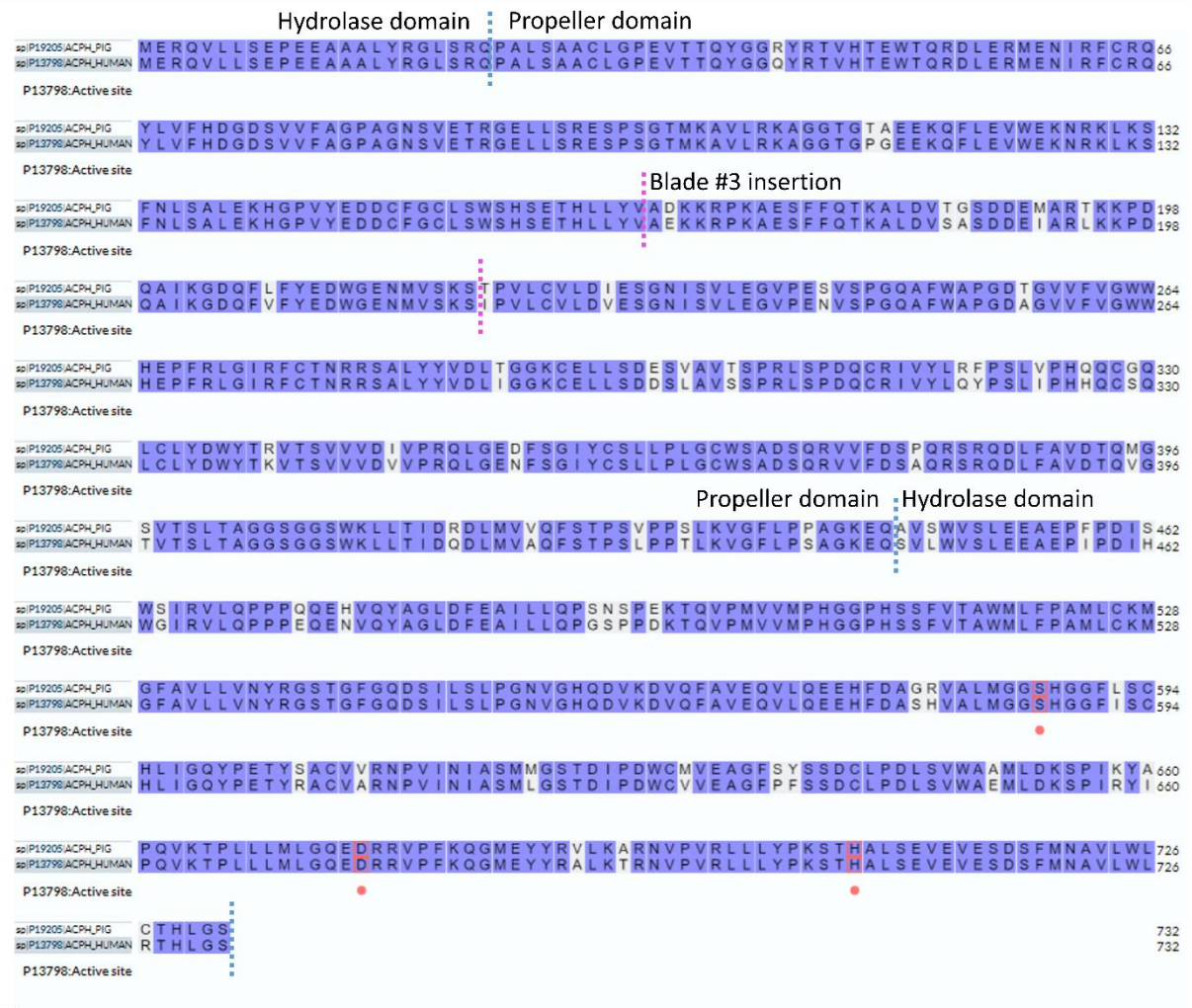


Figure S12 Structure of various human serine proteases showing well-accessible active sites. Chains are rainbow coloured from blue (*N*-term) to red (*C*-term). Catalytic serines are shown as black spheres. (PDB and AlphaFold codes are in parenthesis)

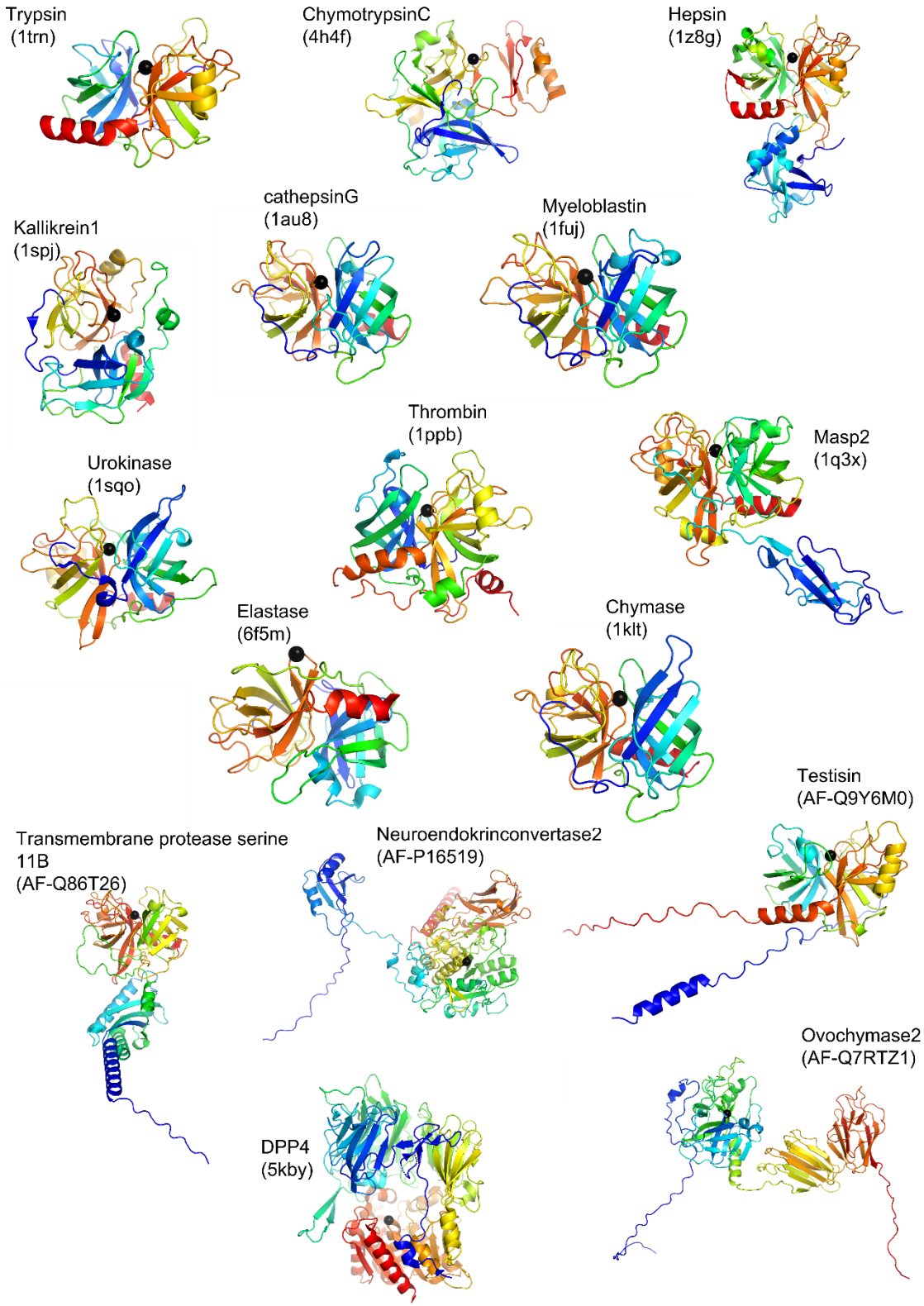
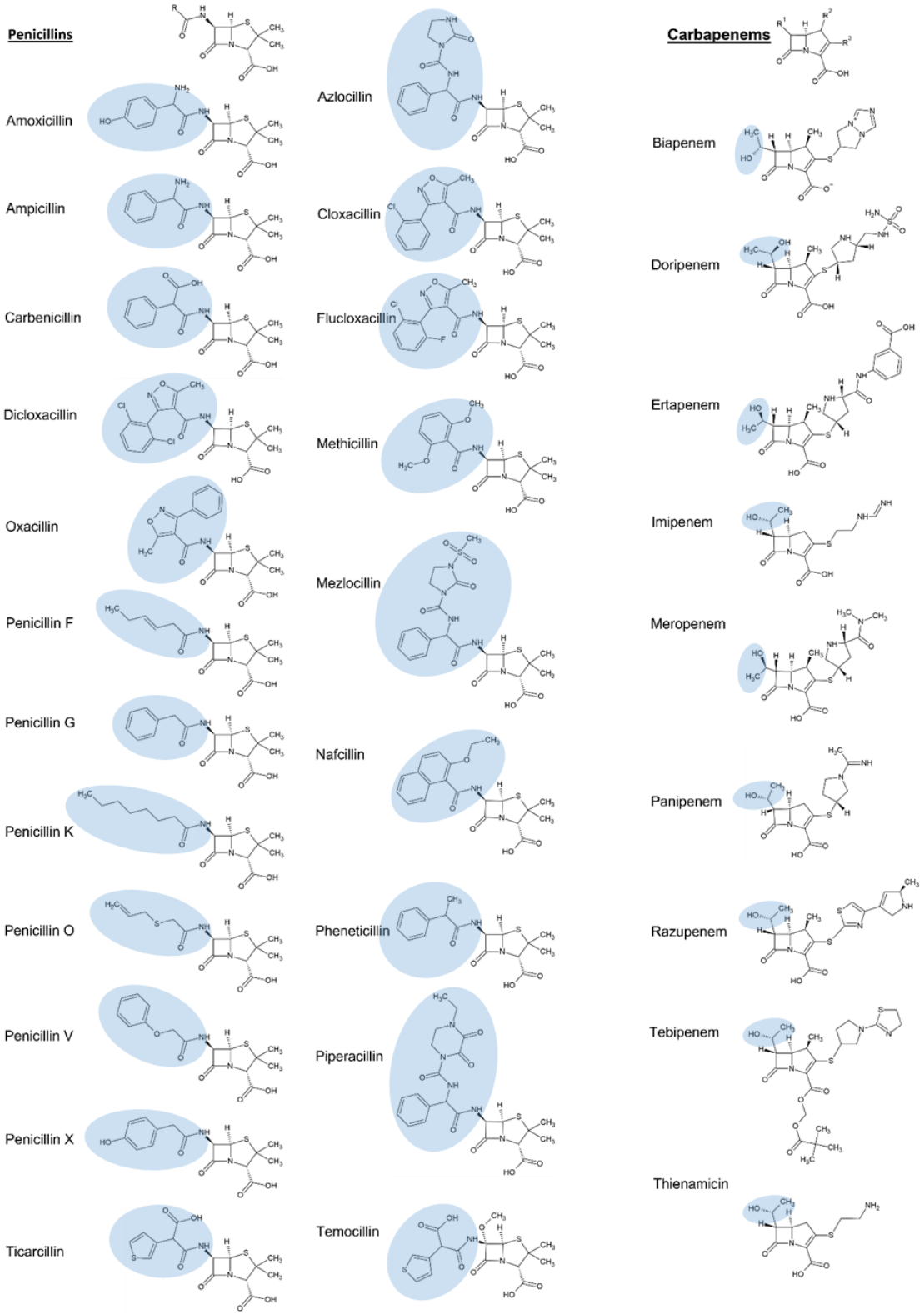
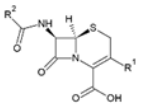


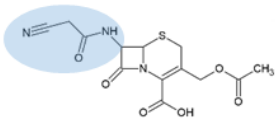
Figure S13. The ring structure of penicillin, carbapenem and cephalosporin type antibiotics. Carbapenems are notably different from the other β-lactams in the moderate size of their R^1 substituent (marked in blue) which is able to fit into the shallow S1 pocket of AAP.



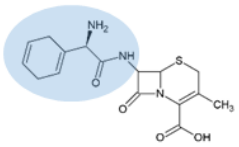
Cephalosporins



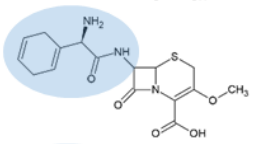
Cefacetril



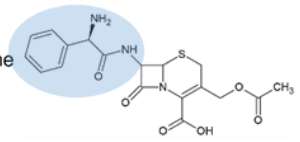
Cefradine



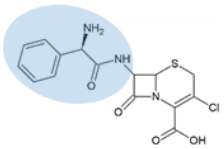
Cefroxadine



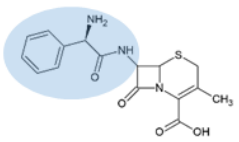
Cefaloglycine



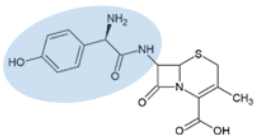
Cefaclor



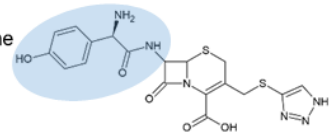
Cephalexine



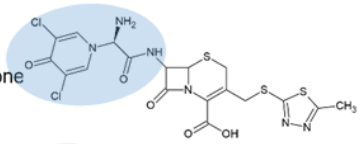
Cefadroxil



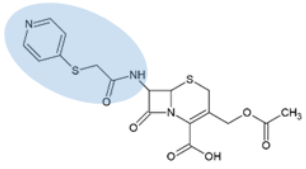
Cefatrizine



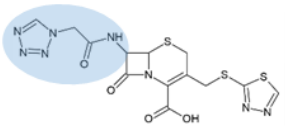
Cefazedone



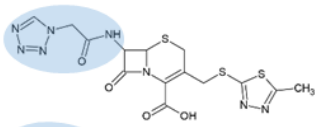
Cefapirin



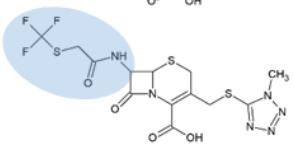
Ceftezole



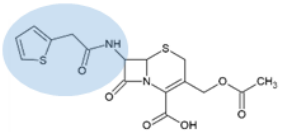
Cefazolin



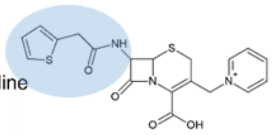
Cefazaflur



Cefalotin



Cephaloridine



Cefalonium

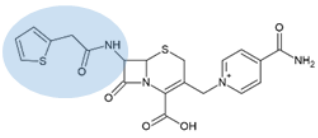


Figure S14. Interaction of meropenem with dithiothreitol monitored by mass spectrometry.

a) LC-UV-MS analysis of meropenem solution in water, **b**) 10 mM meropenem in buffer solution containing 5% dithiotheritol (5 mM NaOAc, pH=5, 44 mM EDTA, 0.8 w/w% PEG4000, 6.6 mM DTT) and **c**) 10 mM meropenem solution in buffer (5 mM NaOAc, pH=5, 44 mM EDTA, 0.8 w/w% PEG4000) used for crystallization (peaks corresponding to intact, unhydrolized meropenem are shown in green).