Supporting information

DPP1 inhibitors: Exploring the role of water in the S2 pocket of DPP1 with substituted pyrrolidines

Helena Käck[1]*, Kevin Doyle[2§], Samantha J. Hughes[2,3†], Michael S. Bodnarchuk[3], Hans Lönn[4], Amanda Van De Poël[2], Nicholas Palmer[2]

[1] Structure, Biophysics and Fragments, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden

[2] Charles River Discovery Services, Cambridge, UK

[3] Medicinal Chemistry, Oncology R&D, AstraZeneca, Cambridge, UK

[4] Medicinal Chemistry, Respiratory, Inflammation and Autoimmune (RIA), BioPharmaceuticals R&D, AstraZeneca (RIA), Gothenburg, Sweden

*Corresponding author: helena.kack@astrazeneca.com

Content

Chemistry

General Procedures

All chemicals purchased from commercial suppliers where used as received. Flash chromatography was carried out with prepacked SiO2 SNAP cartridges (KP-SIL) from Biotage using a Biotage Isolera Four system using gradient elution. Analytical thin-layer chromatography (TLC) was performed on silica using PolygramSIL G/UV254 with fluorescent indicator (200 μm thickness) and visualized under UV light. ^{1}H NMR spectra were recorded on a Bruker AV 400 (^{1}H = 400.13 MHz) instrument and referenced in $CDCl_3$ to tetramethylsilane (0.00 ppm) and in DMSO-d6 referenced to DMSO-d6 (2.50 ppm). The following abbreviations are used: s = singlet, d = doublet, dd = doublet of doublets, dt = doublet of triplets, t = triplet, q = quartet, m = multiplet. Preparative HPLC was performed on a Waters Sunfire column, eluting with a gradient of acetonitrile in aqueous sodium bicarbonate or trifluoroacetic acid solution. All final compounds were purified to >95% chemical purity as assayed by HPLC/MS. HRMS experiments were performed on a Waters Acquity, Waters 2777, or Waters 2700 UV−HPLC system and a Waters Xevo G2 TOF, Waters LCT Premiere, Waters LCT, or Waters QTOFmicro mass spectrometer. Compounds were named with the aid of the Cambridgesoft Chemistry Cartridge (version 15.1.0.144) software. All reactions involving air- or moisture-sensitive reagents were performed under a nitrogen atmosphere using dried solvents and glassware. Calculated pKa values were obtained using ACD software, version 12.0.

(2*S*,4*R*)-4-(*tert*-butoxy)-1-(*tert*-butoxycarbonyl)pyrrolidine-2-carboxylic acid (11a)

Di-*tert*-butyl dicarbonate (1.282 g, 5.87 mmol) was added to a solution of (2*S*,4*R*)-4-*tert*-butoxypyrrolidine-2-carboxylic acid (1 g, 5.34 mmol) and triethylamine (2.08 mL, 14.95 mmol) in THF (10 mL) at room temperature and stirred overnight. The reaction mixture was

poured into water and some of the tetrahydrofuran removed under reduced pressure. The mixture was adjusted to pH~1 with dilute hydrochloric acid and extracted with diethyl ether (3 x 20mL). The combined organic extracts were dried (magnesium sulfate) and evaporated *in vacuo* to afford the title compound (1.20 g, 78 %).

***tert*-butyl (2*S*,4*R*)-2-(((*S*)-1-amino-3-(4'-cyano-[1,1'-biphenyl]-4-yl)-1-oxopropan-2-yl)carbamoyl)-4-(*tert*-butoxy)pyrrolidine-1-carboxylate (15)**

Triethylamine (0.197 mL, 1.41 mmol) followed by TBTU (272 mg, 0.85 mmol) were added sequentially to a solution of (*S*)-2-amino-3-(4'-cyanobiphenyl-4-yl)propanamide (**14**) (150 mg, 0.57 mmol, **WO 2011/154677**) and (2*S*,4*R*)-4-*tert*-butoxy-1-(*tert*-butoxycarbonyl)pyrrolidine-2-carboxylic acid (162 mg, 0.57 mmol) in DMF (5 mL) at room temperature with stirring. After 2 hours the mixture was poured into water and diethyl ether (20 mL) and the layers separated. The aqueous layer was further extracted with diethyl ether and the combined organic extracts dried (magnesium sulfate). The residue was purified by silica gel column chromatography eluting with methanol/CH_2Cl_2/Et_3N 10/90/1 to afford the title compound (60 mg, 20 %).

***tert*-butyl (2*S*,4*R*)-4-(*tert*-butoxy)-2-(((*S*)-1-cyano-2-(4'-cyano-[1,1'-biphenyl]-4-yl)ethyl)carbamoyl) pyrrolidine-1-carboxylate (13a)**

Burgess' reagent (167 mg, 0.70 mmol) was added to a solution of (2*S*,4*R*)-*tert*-butyl 2-((*S*)-1-amino-3-(4'-cyanobiphenyl-4-yl)-1-oxopropan-2-ylcarbamoyl)-4-*tert*-butoxypyrrolidine-1-carboxylate (250 mg, 0.47 mmol) in dichloromethane (3 mL). After stirring overnight the mixture was poured into water (10 mL) and extracted with dichloromethane (3x10mL). The combined organic extracts were dried (magnesium sulfate) and evaporated. The resultant oil was purified by silica gel column chromatography eluting with 3% methanol in dichloromethane to afford the title compound as a colourless film (180 mg, 74 %).

(2*S*,4*R*)-*N*-((*S*)-1-cyano-2-(4'-cyano-[1,1'-biphenyl]-4-yl)ethyl)-4-hydroxypyrrolidine-2-carboxamide trifluoroacetate salt (4)

tert-Butyl (2*S*,4*R*)-4-(*tert*-butoxy)-2-(((*S*)-1-cyano-2-(4'-cyano-[1,1'-biphenyl]-4-yl)ethyl)carbamoyl)pyrrolidine-1-carboxylate (180 mg, 0.35 mmol) was dissolved in formic acid (3mL, 78.2 mmol) and the mixture heated at 50°C for 4h. The mixture was allowed to cool to room temperature and evaporated under reduced pressure. The residue was dissolved in methanol (3 mL) and purified by reverse phase preparative HPLC to afford the title compound as the TFA salt (28 mg, 17 %). 1H NMR (400 MHz, DMSO-*d6*) δ 9.67 - 9.55 (m, 1H), 9.41 (d, J = 7.2 Hz, 1H), 8.85 - 8.69 (m, 1H), 7.98 - 7.85 (m, 4H), 7.78 - 7.72 (m, 2H), 7.50 - 7.41 (m, 2H), 5.53 (s, 1H), 5.07 (q, J = 7.5 Hz, 1H), 4.48 - 4.40 (m, 1H), 4.31 - 4.20 (m, 1H), 3.24 - 3.17 (m, 3H), 3.13 - 3.05 (m, 1H), 2.26 (dd, J = 13.1, 7.2 Hz, 1H), 1.93 - 1.83 (m, 1H). HRMS $C_{21}H_{20}N_4O_2$ $(M+H)^+$ calculated mass 361.1664, found 361.1670

***tert*-Butyl (2*S*, 4*S*)-2-[[(1*S*)-1-cyano-2-[4-(4-cyanophenyl)phenyl]ethyl]carbamoyl]-4-hydroxy-pyrrolidine-1-carboxylate (13c)**

2-Hydroxypyridine-*N*-oxide (0.068 g, 0.61 mmol) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (0.117 g, 0.61 mmol) were added to a

solution of (2*S*,4*S*)-1-*tert*-butoxycarbonyl-4-hydroxy-pyrrolidine-2-carboxylic acid (**11c**) (0.128 g, 0.50 mmol) in dichloromethane (5 mL) at room temperature with stirring. After 30 minutes 4-[4-[(2*S*)-2-amino-2-cyano-ethyl]phenyl]benzonitrile (**12**) (0.123 g, 0.5 mmol, **WO 2015/110826**) and *N*,*N*-diisopropylethylamine (0.137 µL, 0.76 mmol) were added and the mixture continued to stir at room temperature. After 18 hours the reaction mixture was diluted with dichloromethane (10 mL) and washed sequentially with 2.0 M aqueous hydrochloric acid, saturated aqueous solution of sodium hydrogen carbonate and saturated sodium chloride solution. The organic extracts were dried (phase separator) and evaporated under reduced pressure. The resultant foaming yellow oil was used without further purification in the next step (0.232 g, >100%).

(2*S*,4*S*)-*N*-[(1*S*)-1-cyano-2-[4-(4-cyanophenyl)phenyl]ethyl]-4-hydroxy-pyrrolidine-2-carboxamide (6)

tert-Butyl (2*S*, 4*S*)-2-[[(1*S*)-1-cyano-2-[4-(4-cyanophenyl)phenyl]ethyl]carbamoyl]-4-hydroxy-pyrrolidine-1-carboxylate (0.50 mmol) was dissolved in formic acid (3 mL) and heated at 50°C for 10 minutes on a pre-heated stirrer hotplate. After this time the reaction was concentrated under reduced pressure, dissolved in dimethylsulfoxide and submitted to reverse phase preparative HPLC for final purification. The resultant solid was dissolved in dichloromethane (4 mL) and methanol (1 mL). Polymer supported carbonate (4 eq.) was added and the mixture stirred for 20 minutes. The reaction mixture was filtered and evaporated under reduced pressure to afford *title compound* as a white solid (71 mg, 39%). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.24 (d, J = 9.2 Hz, 1H), 7.74 (d, J = 8.4 Hz, 2H), 7.68 (d, J = 8.5 Hz, 2H), 7.59 (d, J = 8.3 Hz, 2H), 7.47 (d, J = 8.3 Hz, 2H), 5.24 - 5.17 (m, 1H), 4.44 - 4.43 (m, 1H), 3.87 - 3.83 (m, 1H), 3.21 - 2.98 (m, 4H), 2.28 - 2.21 (m, 1H), 2.18 - 2.12 (m, 1H), 2.01 - 2.01 (m, 1H), 1.42 (d, J = 2.9 Hz, 1H). HRMS $C_{21}H_{20}N_4O_2$ $(M+H)^+$ calculated mass 361.1664, found 361.1632

(2*S*)-*N*-[(1*S*)-1-Cyano-2-[4-(4-cyanophenyl)phenyl]ethyl]pyrrolidine-2-carboxamide (5)

Prepared following the method for **compound 6** starting from Boc-L-proline (**11b**) (108 mg, 0.5 mmol) to afford the *title compound* as a white solid (52 mg, 30% over two steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.21 (d, J = 8.9 Hz, 1H), 7.74 (d, J = 8.5 Hz, 2H), 7.67 (d, J = 8.4 Hz, 2H), 7.58 (d, J = 8.3 Hz, 2H), 7.38 (d, J = 8.3 Hz, 2H), 5.20 - 5.13 (m, 1H), 3.78 (dd, J = 5.0, 9.3 Hz, 1H), 3.17 - 3.14 (m, 2H), 3.05 - 2.97 (m, 1H), 2.84 - 2.77 (m, 1H), 2.16 - 2.06 (m, 1H), 1.85 - 1.76 (m, 1H), 1.73 - 1.52 (m, 3H). HRMS $C_{21}H_{20}N_4O$ $(M+H)^+$ calculated mass 345.1715, found 345.1696

(2*S*,4*S*)-*N*-[(1*S*)-1-cyano-2-[4-(4-cyanophenyl)phenyl]ethyl]-4-methylsulfanyl-pyrrolidine-2-carboxamide (10)

Prepared following the method for **compound 6** starting from (2*S*,4*S*)-1-*tert*-butoxycarbonyl-4-methylsulfanyl-pyrrolidine-2-carboxylic acid (**11g**) (0.20 g, 0.77 mmol). The crude material was purified by silica gel column chromatography eluting with 0-2% methanolic ammonia (7N) in dichloromethane to afford the *title compound* as a white solid (35 mg, 11% over two steps). ^{1}H NMR (400 MHz, CDCl3) δ 8.14 (d, J=9.3 Hz, 1H), 7.74-7.68 (m, 4H), 7.60 (d, J=8.2 Hz, 2H), 7.42 (d, J=8.2 Hz, 2H), 5.22 - 5.15 (m, 1H), 3.87 - 3.82 (m, 1H), 3.40 - 3.34 (m, 1H), 3.18 - 3.10 (m, 3H), 2.72 - 2.66 (m, 1H), 2.60 - 2.51 (m, 1H), 2.09 (s, 3H), 1.84 - 1.75 (m, 1H), NH not observed. HRMS $C_{22}H_{22}N_4OS$ $(M+H)^+$ calculated mass 391.1592, found 391.1596

***tert*-Butyl (2*S*)-2-[[(1S)-1-cyano-2-[4-(4-cyanophenyl)phenyl]ethyl]carbamoyl]-4,4-difluoro-pyrrolidine-1-carboxylate (13f)**

Triethylamine (318 µL, 2.28 mmol) and dimethylformamide (3 mL) were added to (2*S*)-1-*tert*-Butoxycarbonyl-4,4-difluoro-pyrrolidine-2-carboxylic acid (**11f**) (127 mg, 0.506 mmol) and 4-[4-[(2*S*)-2-Amino-2-cyano-ethyl]phenyl]benzonitrile (**12**) (100 mg, 0.405 mmol, **WO 2015/110826**) in a reaction tube at room temperature with stirring. 2,4,6-Tripropyl-1,3,5,2,4,6-trioxatriphosphorinane-2,4,6-trioxide solution (T3P, 348 mg, 50% solution in DMF) was added

and the reaction stirred at room temperature for 2 hours. After this time the reaction mixture was diluted with ethyl acetate and washed successively with 0.5 M aqueous hydrochloric acid, saturated aqueous solution of sodium hydrogen carbonate (1 M) and saturated sodium chloride solution. The organic extracts were dried (magnesium sulfate), filtered and concentrated under reduced pressure. The crude material was purified by reverse phase preparative HPLC to afford the *title compound as* a yellow oil which was used without further purification in the next step.

(*S*)-*N*-((*S*)-1-cyano-2-(4'-cyanobiphenyl-4-yl)ethyl)-4,4-difluoropyrrolidine-2-carboxamide (9)

tert-Butyl (2*S*)-2-[[(1*S*)-1-cyano-2-[4-(4-cyanophenyl)phenyl]ethyl]carbamoyl]-4,4-difluoro-pyrrolidine-1-carboxylate (0.41 mmol) was dissolved in formic acid (3 mL) and heated at 50°C for 10 minutes on a pre-heated stirrer hotplate. After this time the reaction was concentrated under reduced pressure, dissolved in dimethylsulfoxide and submitted to reverse phase preparative HPLC for final purification to afford the *title compound* as a white solid (29 mg, 19%). ^{1}H NMR (400 MHz, $CDCl_3$) 7.95 (d, J = 9.2 Hz, 1H), 7.74 (d, J = 8.5 Hz, 2H), 7.67 (d, J = 8.4 Hz, 2H), 7.59 (d, J = 8.3 Hz, 2H), 7.39 (d, J = 8.2 Hz, 2H), 5.24 - 5.17 (m, 1H), 4.00-3.94 (m, 1H), 3.34 - 3.13 (m, 3H), 3.00 - 2.87 (m, 1H), 2.67 - 2.53 (m, 1H), 2.39 - 2.23 (m, 2H).

(2*S*,4*R*)-*N*-((*S*)-1-Cyano-2-(4'-cyanobiphenyl-4-yl)ethyl)-4-fluoropyrrolidine-2-carboxamide (7)

Prepared following the method for **compound 9** starting from (2*S*,4*R*)-1-(*tert*-butoxycarbonyl)-4-fluoropyrrolidine-2-carboxylic acid (**11d**) (118 mg, 0.51 mmol) to afford the title compound as a white solid (21 mg, 11% over two steps). ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 8.11 (d, J = 9.2 Hz, 1H), 7.74 (d, J = 8.5 Hz, 2H), 7.67 (d, J = 8.5 Hz, 2H), 7.58 (d, J = 8.3 Hz, 2H), 7.37 (d, J = 8.3 Hz, 2H), 5.21 - 5.04 (m, 2H), 4.06 - 4.00 (m, J = 8.3, 8.3 Hz, 1H), 3.35 - 3.18 (m, 1H), 3.21 - 3.13 (m, 2H), 2.64 - 2.38 (m, 3H), 1.97 - 1.79 (m, 1H).

(2*S*,4*S*)-*N*-((*S*)-1-Cyano-2-(4'-cyanobiphenyl-4-yl)ethyl)-4-fluoropyrrolidine-2-carboxamide (8)

Prepared following the method for **compound 9** starting from (2*S*,4*S*)-1-(*tert*-butoxycarbonyl)-4-fluoropyrrolidine-2-carboxylic acid (**11e**) (118 mg, 0.51 mmol). The crude reaction mixture was purified by silica gel column chromatography eluting with 7N methanolic ammonia in dichloromethane to afford the *subtitled compound* as a white solid (40 mg, 22% over two steps). ^{1}H NMR (400 MHz, DMSO-d_6) δ ppm 8.73 (d, *J* = 8.4 Hz, 1H), 7.94 - 7.86 (m, 4H), 7.72 (d, *J* = 8.3 Hz, 2H), 7.43 (d, *J* = 8.3 Hz, 2H), 5.15 (d, J=53.0 Hz, 1H), 5.05 - 4.97 (m, 1H), 3.65 (d, *J* = 7.4 Hz, 1H), 3.19 (d, *J* = 7.5 Hz, 2H), 3.09 - 2.98 (m, 3H), 2.27 - 2.08 (m, 1H), 2.03 - 1.90 (m, 1H). HRMS $C_{21}H_{19}FN_4O$ $(M+H)^+$ calculated mass 363.1621, found 363.1636

Crystallography

Cathepsin C was prepared as described. [1] Crystals of DPP1 were grown by vapour diffusion at 20 °C from a mix of 2 µl of protein solution (10 mg/ml in 20 mM Na Acetate pH 4.5, 200 mM NaCl, 2 mM dithiothreitol and 0.5 mM EDTA) and 2 µl crystallisation solution containing 20-27 % PEG 3350, 0.2 M Ammonium Citrate and 10 mM dithiothreitol. Compound introduction was done by incubation in crystallization solution with 5 % increased PEG 3350 concentration and 40 mM compound, dissolved to 0.2 M in DMSO for 24-28 h. The crystals were cryo-cooled directly from the soaking solution. With one exception, X-ray diffraction data was collected to better than 1.9 Å resolution on a Rigaku A200 CDD detector mounted on a rotating anod (FRE+, Rigaku). X-ray diffraction data from a crystal incubated with compound **3** were collected at the European Synchrotron Radiation Facility, Grenoble, at beam line ID23. Data were integrated and processed using autoPROC.[2] The crystals belong to space group I222 with one molecule in the asymmetric unit and $P3_121$ with two molecules in the asymmetric unit. The structures were solved with molecular replacement using the Protein Data Bank

(PDB) entry 1K3B as search model.[3] Structural refinement was carried out using the Refmac[4] and AutoBuster[5]. Manual rebuilding was done using Coot.[6] All structure illustrations were prepared using the program Pymol.[7] Crystallographic coordinates have been deposited into the Protein Data Bank.

Computational Simulations

Protein-preparation

All structures were prepared using the 'Protein Preparation Wizard' tool in Maestro.[8]
All waters were retained, with the NAG groups covalently bonded to the receptor by manually inserting a bond between C and N. A double bond was created at the ligand imine, whilst surface GOL and DMS residues were retained. Protein and ligand protonation states assigned using PROPKA and Epik respectively at pH 7.0 ± 1. The hydrogen-bonding network was assigned for each complex, followed by a restrained minimisation using the OPLS(2) forcefield to a RMSD of 0.3 Å.

ProtoMS set-up

All simulations were performed using the ProtoMS software package (v3.4)[9] Owing to the high similarity between the proteins in each crystal structure, the protein crystallised with compound **5** (6RN7) was used in all subsequent simulations.
The protein, ligands, and water were modeled using Amber14SB, gaff14, and TIP4P force-fields, respectively. A scoop of 20 Å around the ligand was taken to reduce the system size, with side chain and backbone sampling in the inner 16 Å and side chain only in the remaining 4 Å. The DPP1 complex was solvated with TIP4P water using a half-harmonically restrained sphere of radius of 30 Å. Since ProtoMS cannot simulated covalently-bonded species, a harmonic restraint of 5 kcal/mol was applied between the ligand imine and Cys234 to prevent

disassociation. For the free simulation legs, the ligands were solvated in a cubic box with a padding distance of 10 Å between ligand and box edge. For grand canonical simulations, water molecules within the GCMC region are removed prior to the simulation.

GCMC simulations

A small box, measuring 3 Å x 6 Å x 3 Å, was used to cover the deep part of the S2 pocket. An initial GCMC equilibration of 5 M MC moves was performed, using a 1:1:1 ratio of insertion, deletion, and GC water sampling moves. A further 5 M equilibration moves were performed, followed by 40 M production MC moves across the entire system using the sampling rations in Table S2.

Bulk solvent is prohibited from entering the GCMC region whilst ligand and protein atoms can fully sample the region. GCMC was performed using 59 equally-spaced B values spanning B=0.0 to B=-29.0. Data was collected over the last 60 M MC moves.

Alchemical perturbation simulations

Single-topology alchemical transformations were performed on key pairs of DPP1 ligands. Perturbations were performed in two stages; considering the perturbation as taking place from the larger molecule to the smaller, the electrostatic parameters were first perturbed, followed by the van der Waals (vdW) interactions. Each simulation was split across 16 equally spaced λ windows, with perturbations performed both in the bound state and for the ligand in bulk solvent. 5 M MC equilibration steps are performed, followed by 40 M production steps. The ratio of MC moves for each system is shown in Table S2. The resultant free-energies were calculated using Multistate BAR (MBAR).

Table S1. Statistics from crystallographic data reduction and refinement.

Compound	**3**	**5**	**6**	**8**	**10**
PDB accession code	6RN6	6RN7	6RN9	6RNE	6RNI
Space group:	$P3_121$	I222	I222	I222	I222
Cell constants: a b c (Å)	84.55 84.55 223.3	87.21 87.82 114.9	87.35 87.77 114.5	87.54 87.74 114.6	87.37 87.46 114.8
Resolution range (Å)	73.2 - 2.4	69.78 - 1.66	27.74 - 1.9	28.28 - 1.65	69.53 - 1.54
Highest resolution shell (Å)	2.44 – 2.4	1.69 – 1.66	1.95 – 1.9	1.73 – 1.65	1.62 – 1.54
Completeness overall (%)	93.0 (88.8)	92.7 (49.3)	100 (100)	99.9 (99.5)	90 (55)
Total number of observations	135150	296053	236747	370946	402130
Reflections, unique	34455	48695	34994	53715	58771
Redundancy	3.9 (3.9)	6.1 (2)	6.8 (6.4)	6.9 (6.1)	6.8 (4.5)
R merge $_{overall}$ [1]	0.097 (0.769)	0.051 (0.676)	0.089 (0.417)	0.053 (0.481)	0.08 (0.966)
I/SigI[2]	11.6 (2.1)	24.4 (2.5)	14.7 (4.1)	28.3 (4.2)	18.8 (1.6)
R value $_{overall}$ (%)[3]	17.4	17.5	18.5	18.0	19.6
R value $_{free}$ (%)[4]	22.1	19.2	20.9	20.3	20.9
R.m.s. deviations from ideal values					
Bond lengths (Å)	0.011	0.010	0.009	0.009	0.008
Bond angles (°)	1.09	1.01	0.98	0.99	0.96
Φ, Ψ angle distribution for residues [5]					
In most favoured regions (%)	85.1	85.1	85.0	85.0	84.5
In additional allowed regions (%)	14.2	13.6	13.6	13.9	14.9
In generously regions (%)	0.2	0.7	0.7	0.0	0.3

In disallowed regions (%)	0.5	0.7	0.7	1.0	0.3

[1] $R_{merge} = \Sigma_{hkl} [(\Sigma_i |I_i - \langle I \rangle|)/ \Sigma_i I_i]$
[2] I/sigI avg is the mean I/sig for the unique reflections in the output file
[3] $R_{value} = \Sigma_{hkl} ||F_{obs}| - |F_{calc}|| / \Sigma_{hkl} |F_{obs}|$
[4] R_{free} is the cross-validation R factor computed for the test set of 5 % of unique reflections
[5] Ramachandran statistics as defined by PROCHECK[10]

Table S2: MC move rations for each simulation type performed

Simulation	Solvent	Protein	Ligand	Insertion	Deletion	GC solvent
GCMC	349	145	7	167	167	167
Alchemical perturbation	697	289	14	-	-	-

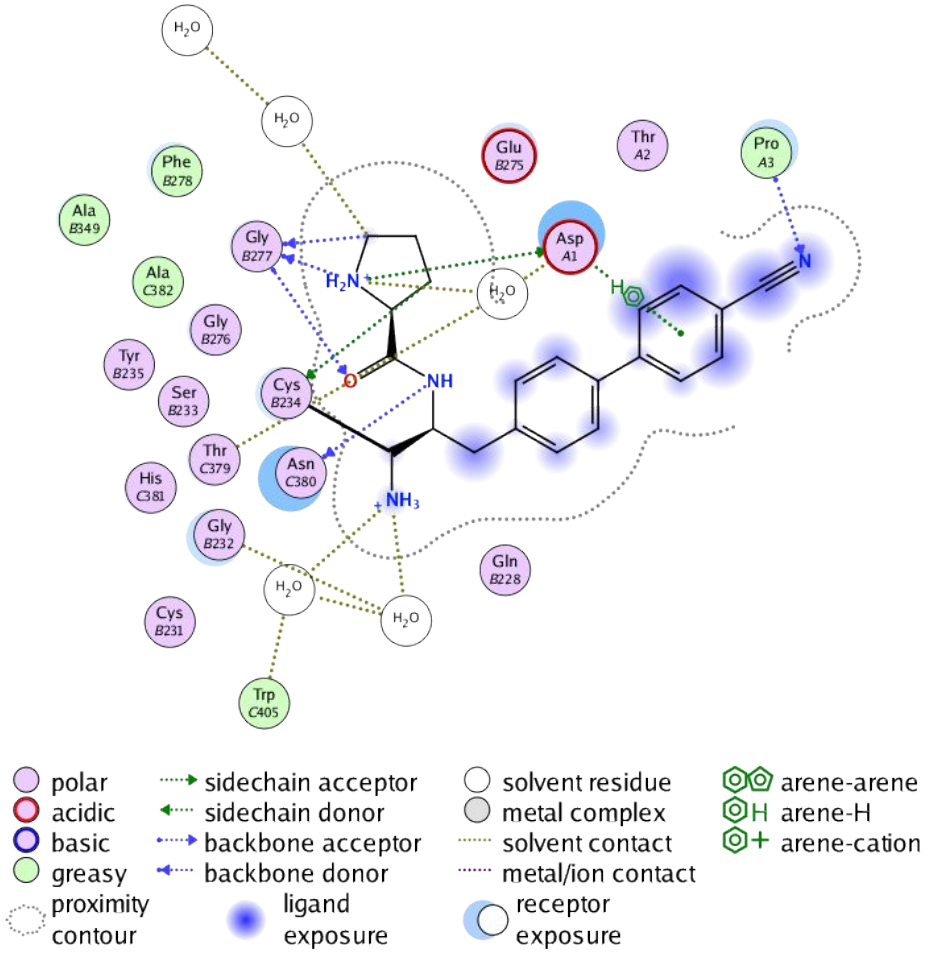

Figure S1. Schematic binding mode of compound **5** showing protein-ligand contacts generated with the software MOE.[11] The Cl^- ion is not included in this view.

References:

1. Furber, M.; Tiden, A.-K.; Gardiner, P.; Mete, A.; Ford, R.; Millichip, I.; Stein, L.; Mather, A.; Kinchin, E.; Luckhurst, C.; Barber, S.; Cage, P.; Sanganee, H.; Austin, R.; Chohan, K.; Beri, R.; Thong, B.; Wallace, A.; Oreffo, V.; Hutchinson, R.; Harper, S.; Debreczeni, J.; Breed, J.; Wissler, L.; Edman, K., Cathepsin C inhibitors: Property optimization and identification of a clinical candidate. Journal of medicinal chemistry 2014, **57**, 2357-2367.

2. Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.; Bricogne, G., Data processing and analysis with the autoPROC toolbox. Acta Crystallogr D Biol Crystallogr 2011, **67** (Pt 4), 293-302.

3. Turk, D.; Janjic, V.; Stern, I.; Podobnik, M.; Lamba, D.; Dahl, S. W.; Lauritzen, C.; Pedersen, J.; Turk, V.; Turk, B., Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. The EMBO journal 2001, **20**, 6570-6582.

4. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J., Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 1997, **53** (Pt 3), 240-55.

5. Bricogne G., Blanc E., Brandl M., Flensburg C., Keller P., Paciorek W., Roversi P, Sharff A., Smart O.S., Vonrhein C., Womack T.O. (2017). BUSTER version 2.11.7. Cambridge, United Kingdom: Global Phasing Ltd.

6. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of Coot. Section Title: Biochemical Methods 2010, **66** (4), 486-501.

7. The PyMOL Molecular Graphics System, Version 1.7: Schrödinger, LLC, New York, NY, 2010.

8. Schrödinger Release 2019-1: Maestro, Schrödinger, LLC, New York, NY, 2017.

9. ProtoMS 3.4, http://www.essexgroup.soton.ac.uk/ProtoMS/ (Accessed 09/04/2019), 2018

10. Roman A Laskowski, Malcolm W MacArthur, David S Moss and Janet M Thornton, PROCHECK: a program to check the stereochemical quality of protein structures J. App. Cryst. 1993, **26**, 283-291.

11. MOE 2017.01; Chemical Computing Group, 2017.