Supplementary Material

EXPERIMENTAL

Cheminformatics – The *pre-docking* filtering was performed using KNIME (Konstanz Information Miner) pipeline software where nodes representing Java scripts are assembled in a graphical user interface to parse data into tables [1]. For pre-docking filtering 20 descriptors were calculated for all 6.06 million compounds in the DrugsNow subset from the ZINC database [2]. The same descriptors were calculated for the known EthR inhibitors to choose suitable parameters for the *pre-docking* step and select the compounds used in the decoy data set. The KNIME clustering pipeline used for *post-docking* filtering is given in the supplementary material (Figure S1).

Docking – Virtual screening was conducted using GOLD (v5.2) [3]. After extensive protocol development, a protein structure of EthR (PDB: 1U9N [4]) with native ligands removed, H-atoms added, and with residue Asn179 flipped relative to crystallographic position was selected for screening. There are no water molecules in the binding site. For *in silico* docking a search space with a radius of 10 Å was defined in the center of the ligand-binding channel. The search efficiency was set to 20%. All other settings were as default unless otherwise specified. The *diverse* solutions termination in GOLD was applied to increase conformational diversity in ligand pose generation, with five clusters of one pose each defined as 1 Å RMS difference apart. Thus, five poses were generated for each screened ligand, the product of five genetic algorithm runs. For *post-docking filtering*, the highest scoring pose of each ligand was extracted and the top scoring 10% of ligands carried forward into GoldMine. A variety of physicochemical descriptors, previously calculated in the KNIME platform, were also carried forward and GoldMine was used to calculate additional descriptors for filtering. Crucially, a hydrogen bonding contact angle filter was applied to pass only those ligands, which formed a H-bond at a minimum angle of 120°. Manual selection aided by Mogul evaluation, and removal of duplicate Murcko Scaffolds, resulted in 85 commercially available compounds for further experimental evaluation.

Protein expression, purification and characterisation - EthR from *M. tuberculosis* was overexpressed in *E. coli* and purified by affinity chromatography using the published protocol [5]. On additional purification step using a Superdex 200 gel filtration column (GE Healthcare) was included to ensure maximum purity. All protein samples were characterized by SDS PAGE gel, electron spray mass spectrometry and circular dichroism to ensure purity, homogeneity and correct folding. The protein sample was dialysed against 10 mM TRIS HCl pH 7.5, 200 mM NaCl and stored at 4°C.

Thermal Shift Assays – Fluorescent dye SYPRO Orange (Invitrogen) was used to monitor protein unfolding with an Applied Biosystems Fast 7500 qPCR system. Experiments were conducted in 96-well plates (ThermoScientific) and sealed with polyolefin film (ThermoScientific). Samples were heated from 24 to 95°C at a rate of 1°C/min with fluorescence measured at Ex/Em = 450/580 nm. Final sample concentrations in 20 µL were: 40 µM EthR, 5X SYPRO Orange, ligand concentrations of either 400, 320, 160 or 80 µM, and DMSO content adjusted to 1%. Experiments were performed in triplicate, and shifts relative to non-compound control calculated using NAMI [6]. Hits were selected as those compounds which demonstrated the ability to shift the melting temperature in a concentration-dependent manner, in a range of 1 to 9°C at highest assayed concentration.

Co-crystal Structure Determination – EthR-ligand complexes were obtained by incubating 9 µl freshly prepared protein sample at approximately 9 mg/ml in 10 mM Tris HCl, pH 7.5 and 200 mM NaCl with 1 µl of ligand at a concentration of 33mM in DMSO overnight. Crystallisation followed the published conditions mixing 1µl protein-ligand complex solution with 1 µl of crystallization solution (0.1-0.2 MES pH 6.0-6.5 with 1.4-1.6 M $(NH_4)_2SO_4$ and 10% glycerol) in hanging drop vapor diffusion crystallization trays. Needle-shaped crystals with dimensions of up to 1 mm in the longest direction and 0.05 mm in the other two directions appeared over the course of one to three days. Crystals were flash cooled in liquid nitrogen and mounted in micromesh or loops. All diffraction data were collected in remote access mode on the DLS beam lines IO3 and I04-1, respectively, equipped with a Pilatus pixel detector [7]. Initial data processing was

done with xia2 [8], and carefully reprocessed with XDS [9]. All structures were solved with Phaser [10], refined with Refmac5 [11] with 5% of reflections used to calculate R_{free} [12], validated using interactive computer graphics and tools implemented in Coot [13,14]. Target restraints for ligands were generated with PRODRG [15]. All crystal structures displayed clear unbiased electron density for the ligand and were consecutively refined to satisfactory R-factors and good geometry. All ligand geometries were validated using the tools available in Coot and MOGUL [16]. Further crystallographic details are summarized in Table S2.

Biological in-vivo assays – Growth inhibition of *Mycobacterium tuberculosis* GFP strains was performed as described earlier [5]. Briefly, A recombinant strain of *M. tuberculosis* H37Rv expressing the green fluorescent protein (H37Rv-GFP) was obtained by transformation of the integrative plasmid pNIP48. In this plasmid derived from the Ms6 mycobacteriophage, the *gfp* gene is cloned under the strong mycobacterial promoter pBlaF and the GFP is constitutively expressed. This plasmid also contains an hygromycin resistance gene. Bacterial stocks kept at -80°C are used to inoculate 5 ml of Middlebrook 7H9 medium supplemented with oleic acid-albumin-dextrose-catalase (OADC, Difco, Sparks MD, USA) and with 50 μg ml^{-1} hygromycin (Invitrogen, Carlsbad, CA USA) in 25 cm^2 tissue-culture flasks. Flasks are incubated at 37°C without shaking for 7 days. Cultures are then diluted with fresh culture medium to reach an OD_{600} of 0.1. Culture flasks (75 cm^2) are filled with 50 ml of this diluted culture, which are cultivated 7 days at 37°C without shaking. The microplates are prepared as follows. Ethionamide (Sigma, E6005) is diluted in DMSO at 0.1 mg/mL; aliquots are stored frozen at −20 °C. Compounds are resuspended in DMSO at a final concentration of 10 μM. Ethionamide and test-compounds are transferred to a 384-well low-volume polypropylene plate and used to prepare assay plates. Ten 3-fold serial dilutions of compounds (typically in the ranges of 30 to 4.5e-3 uM) are performed into black Greiner 384-well clear bottom polystyrene plates using an Echo 550 liquid Handler (Labcyte). DMSO volume are compensated so that the concentration across all wells is equal (0.3%). Ethionamide is then transferred to the 384-well plates, using Echo. The final concentration of ETH is 0.1 μg/ml (10 to 20 times lower than MIC). The final amount of DMSO in the

assay plate remains <1% v/v for each well. Controls in the assay plate include DMSO at 0.3% (negative control) and INH at 1 μg/ml (positive control). A reference plate includes rifampicin, INH and ETH ranging from 30 to 1.8e-3 μg/ml (15 points, 2x dilutions). Cultures of H37Rv-GFP to be added to assay plates are washed two times in PBS, resuspended in fresh culture medium (without Hygromycin), and grown for 5 days at 37°C. Finally, cultures are diluted to an OD600 nm of 0.02 (using fresh culture medium with no addition of Hygromycin) and 50 μL are transferred to each assay plate. Assay plates are incubated at 37ºC for 5 days. Fluorescent signal is acquired on a Victor 3 multilabel plate reader (Perkin Elmer), using exc=485nm/em=535nm.

Table S1. Summary of the 23 clusters of different ring chemistry

Cluster name	No of compounds	Screening cluster size
Imidazoles	121369	**36411**
Pyrazoles	133450	**50035**
Pyrroles	240174	**24017**
Oxazoles	39025	**11708**
Thiadiazoles	8805	**8805**
Pyrimidines	107658	**32297**
Pyridines	91590	**27477**
Pyridazines	16227	**17277**
Heterocycles	6091	**6091**
7-membered rings	10768	**10768**
Triazine	4150	**4150**
Thiazoles	9372	**9372**
Piperidines	182898	**18290**
Oxanes	29035	**14518**
Dioxanes	11814	**11814**
Morpholines	38142	**11443**
Thioles	22934	**11467**
Dioxalanes	28635	**28635**
Sulphonyls	32986	**32986**
Others	52297	**15690**
Total	1284064	**409201**

ZINC REFERENCE	COMPOUND	SCORE RANK	COMPANY	COMPANY REFERENCE	RAW SCORE	ΔTm 400μM/ °C
ZINC12201617	1	22	ChemBridge	86370596	89.1069	0.119
ZINC00616109	2	19	ChemBridge	7778963	89.2407	0.152
ZINC67692217	3	29	ChemBridge	58238331	87.9500	1.386
ZINC65406277	4	33	ChemBridge	50509137	87.2316	0.819
ZINC72158865	5	38	ChemBridge	75389916	86.8409	0.152
ZINC65507786	6	11	ChemBridge	89710978	91.3825	0.252
ZINC67973854	8	18	ChemBridge	94793758	89.3759	0.052
ZINC67654656	9	23	ChemBridge	26569393	88.5067	-0.014
ZINC67974892	10	45	ChemBridge	95652737	85.5077	1.986
ZINC71775561	11	46	ChemBridge	96176937	85.4536	-0.014
ZINC19952823	12	47	ChemBridge	68375090	85.4176	-0.114
ZINC72172695	13	49	ChemBridge	97057671	85.2338	0.202
ZINC65526448	14	55	ChemBridge	92870208	84.5794	0.152
ZINC00237604	15	59	ChemBridge	5304371	84.1084	3.486
ZINC67955631	16	72	ChemBridge	91757787	83.4029	-0.014
ZINC32576321	17	73	ChemBridge	64993287	83.3492	0.619
ZINC11817892	18	76	ChemBridge	73552030	83.2528	0.186
ZINC67673539	19	94	ChemBridge	27987503	81.5626	-0.214
ZINC20109834	20	95	ChemBridge	69520354	81.5585	0.386
ZINC67975203	21	24	ChemBridge	95916741	88.4118	0.052
ZINC00132569	22	9	ChemBridge	7618548	92.8909	0.119
ZINC67820979	23	12	ChemBridge	72213543	91.2764	0.119
ZINC00714237	24	56	ChemBridge	5652915	84.4833	-0.048
ZINC01216083	26	80	ChemBridge	6088631	82.9549	0.486
ZINC02438552	27	6	ChemBridge	9116022	95.2030	0.402
ZINC01226689	28	7	ChemBridge	5249858	93.3650	-0.098
ZINC04828017	29	10	ChemBridge	7962144	92.2616	0.319
ZINC20137454	30	43	ChemBridge	9228506	85.8673	0.686
ZINC24615603	31	44	ChemBridge	9283033	85.6908	0.719
ZINC00425482	32	67	ChemBridge	9331185	83.7565	0.086
ZINC09236877	33	91	ChemBridge	6985499	81.7739	0.052
ZINC17076365	34	25	Vitas M	STK470583	88.3717	-0.081
ZINC13758755	35	52	Vitas M	STL050026	84.9144	0.052
ZINC71775333	36	58	Vitas M	STL167978	84.1745	0.052
ZINC00924619	37	60	Vitas M	STK011909	84.1005	-
ZINC69462812	38	3	Enamine	Z1025693264	96.1781	-0.348
ZINC40151811	39	20	Enamine	Z596096640	89.2407	-0.248
ZINC58327836	40	21	Enamine	Z1101450797	89.2192	0.152
ZINC52487551	41	54	Enamine	Z646440516	84.7711	-0.281
ZINC53275627	42	62	Enamine	Z167023036	83.9472	0.819
ZINC46957685	43	51	Enamine	Z422902610	85.0971	-0.081
ZINC44936750	44	34	Enamine	Z415248090	87.0520	-0.014
ZINC71839930	45	27	Enamine	Z908777508	88.3083	-0.048

ZINC69461964	46	42	Enamine	Z1085724378	85.9954	-0.114
ZINC05187134	47	28	Enamine	Z54071047	88.0150	-0.348
ZINC69489717	48	85	Enamine	Z1139229987	82.3846	0.619
ZINC12876413	49	90	Enamine	Z102922270	81.8401	0.352
ZINC12794882	50	71	Enamine	Z119631558	83.5383	0.586
ZINC22831673	51	31	Enamine	Z90609697	87.4262	0.352
ZINC65575010	52	70	Enamine	Z558478778	83.5945	0.019
ZINC12609641	53	57	Enamine	Z279903378	84.4802	-0.048
ZINC14168858	54	26	Enamine	Z24677903	88.3478	0.286
ZINC10777920	55	2	Enamine	Z226510926	97.5458	0.286
ZINC08987568	56	68	Enamine	Z224851022	83.6657	0.819
ZINC09689958	57	86	Enamine	Z30508852	82.2755	6.352
ZINC71795591	58	30	Enamine	Z281523990	87.9420	0.752
ZINC71794585	59	16	Enamine	Z257202932	90.7126	-0.414
ZINC14848503	60	35	Enamine	Z237505970	86.9967	6.252
ZINC22281715	61	89	Enamine	Z225731946	81.8996	-0.514
ZINC45517734	62	36	Enamine	Z351674606	86.9565	-0.614
ZINC09725714	63	37	Enamine	Z197481400	86.8884	0.486
ZINC69416783	64	74	Enamine	Z1001812522	83.3350	1.252
ZINC24418394	65	32	Enamine	Z153679482	87.2489	-0.781
ZINC25187332	66	93	Enamine	Z25094450	81.5864	-0.548
ZINC23053302	67	81	Enamine	Z281178970	82.8968	-0.448
ZINC65490004	68	69	Enamine	Z225720568	83.6479	-0.314
ZINC69669548	69	87	Enamine	Z1139749166	82.0951	-0.381
ZINC69381663	70	78	Enamine	Z324839670	82.9982	-0.281
ZINC08706459	71	5	Enamine	Z226452786	95.2065	-0.514
ZINC03420970	72	8	Enamine	Z24107796	93.0210	0.519
ZINC58441335	73	82	Enamine	Z649796534	82.5989	-0.281
ZINC08980600	74	39	Enamine	Z226124130	86.5464	0.352
ZINC58145475	75	63	Enamine	Z846030592	83.9063	-0.381
ZINC18962750	76	61	Enamine	Z98973900	84.0079	-0.314
ZINC03905710	77	13	Ambinter	Amb16755507	91.1945	-1.114
ZINC04370962	78	92	Ambinter	Amb20448045	81.7662	-1.181
ZINC33004106	79	84	Ambinter	Amb16657739	82.4998	-0.781
ZINC01502258	80	40	Ambinter	Amb16766591	86.3118	3.452
ZINC00499610	81	77	Ambinter	Amb19773311	83.1186	-1.181
ZINC03903089	82	50	Ambinter	Amb6875571	85.1973	-1.448
ZINC04870694	83	66	Ambinter	Amb5371151	83.8413	-0.148
ZINC09087663	84	79	Ambinter	Amb5367264	82.9583	1.486
ZINC19234873	85	1	Ambinter	Amb13914491	99.5423	-1.548

This table denotes all compounds used for analysis. Note: NJT07 is not listed as it failed quality control at the supplier and could not be sourced elsewhere. Additionally, NJT25 arose from an ordering error and as such is not listed with other virtual screening results. Finally, NJT37 was completely insoluble, and so no assays were performed with it. ΔTm given is an average of three independent measurements.

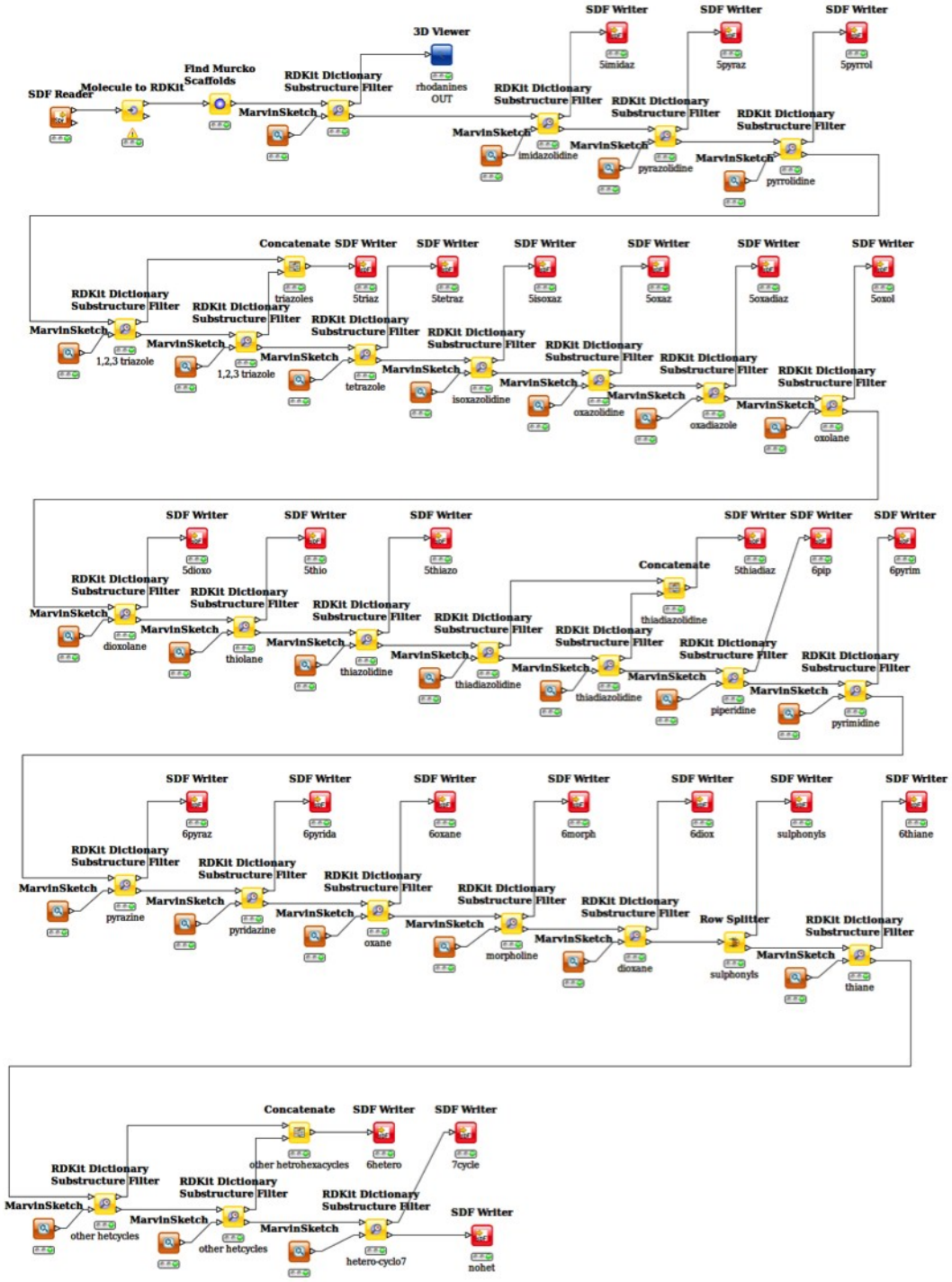
Table S3. Summary of Crystallographic Data

Structure	25	10	42	57
DLS beamline	I04-1	I03	I03	I04-1
Wavelength [Å]	0.9174	0.9600	0.9600	0.9174
a, b [Å]	120.08	122.54	120.48	121.41
c [Å]	33.66	33.82	33.65	33.69
α=β=χ [°]	90	90	90	90
Space group	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$
Resolution	1.4	1.5	1.95	2.1
No of refl.	498 518	521 487	155 537	156 134
Multiplicity	10.2	14.2	8.3	10.2
Wilson B [Å^2]	26.1	27.7	34.6	37.7
I/sigma#	22.6 (3.2)	12.5 (3.2)	13.6 (4.0)	14.3 (2.9)
R_{merge} [%]*	3.4 (40.2)	8.0 (57.3)	7.4 (48.0)	10.5 (58.8)
Compl. [%]*	99.2 (96.3)	99.9 (100)	99.7 (99.5)	99.4 (99.8)
No. of refl. in Refinement	46 522	39 920	17 755	14 449
R_{work} [%]	0.181	0.175	0.175	0.183
R_{free} [%]	0.205	0.204	0.229	0.222
No of residues	193	196	192	193
No of water	120	148	74.5	84
Protein B [Å^2]	26.1	27.3	34.4	37.2
Ligand B [Å^2]	22.3	32.3	51.5	36.8
Water [Å^2]	35.2	37.5	41.6	42.8
Rmsd bond lengths [Å]	0.018	0.017	0.015	0.015
Rmsd angles [°]	1.77	1.70	1.62	1.62
Ramach. outlier	1/193	0/196	0/192	1/193
PDB code	5NIO	5NIM	5NIZ	5NJ0

* calculated by XDS, Friedel pairs not merged [9].

numbers in brackets refer to last shell.

Figure S1: KNIME pipeline for post-docking clustering and selection.



1 Beisken, S. *et al.* KNIME-CDK: Workflow-driven cheminformatics. *BMC Bioinformatics* **14**, 257, doi:10.1186/1471-2105-14-257 (2013).

2 Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S. & Coleman, R. G. ZINC: a free tool to discover chemistry for biology. *J Chem Inf Model* **52**, 1757-1768, doi:10.1021/ci3001277 (2012).

3 Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W. & Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **52**, 609-623, doi:10.1002/prot.10465 (2003).

4 Frenois, F., Engohang-Ndong, J., Locht, C., Baulard, A. R. & Villeret, V. Structure of EthR in a ligand bound conformation reveals therapeutic perspectives against tuberculosis. *Mol Cell* **16**, 301-307, doi:10.1016/j.molcel.2004.09.020 (2004).

5 Willand, N. *et al.* Synthetic EthR inhibitors boost antituberculous activity of ethionamide. *Nat Med* **15**, 537-544, doi:10.1038/nm.1950 (2009).

6 Groftehauge, M. K., Hajizadeh, N. R., Swann, M. J. & Pohl, E. Protein-ligand interactions investigated by thermal shift assays (TSA) and dual polarization interferometry (DPI). *Acta Crystallogr D Biol Crystallogr* **71**, 36-44, doi:10.1107/S1399004714016617 (2015).

7 Broennimann, C. *et al.* The PILATUS 1M detector. *J Synchrotron Radiat* **13**, 120-130, doi:10.1107/S0909049505038665 (2006).

8 Winter, G., Lobley, C. M. & Prince, S. M. Decision making in xia2. *Acta Crystallogr D Biol Crystallogr* **69**, 1260-1273, doi:10.1107/S0907444913015308 (2013).

9 Kabsch, W. Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).

10 Bunkoczi, G. *et al.* Phaser.MRage: automated molecular replacement. *Acta Crystallogr D Biol Crystallogr* **69**, 2276-2286, doi:10.1107/S0907444913022750 (2013).

11 Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D* **67**, 355-367, doi:10.1107/S0907444911001314 (2011).

12 Brunger, A. T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472-475 (1992).

13 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).

14 Debreczeni, J. E. & Emsley, P. Handling ligands with Coot. *Acta Crystallogr D Biol Crystallogr* **68**, 425-430, doi:10.1107/S0907444912000200 (2012).

15 Schuttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D* **60**, 1355-1363, doi:10.1107/S0907444904011679 (2004).

16 Cottrell, S. J., Olsson, T. S., Taylor, R., Cole, J. C. & Liebeschuetz, J. W. Validating and understanding ring conformations using small molecule crystallographic data. *J Chem Inf Model* **52**, 956-962, doi:10.1021/ci200439d (2012).