

Inhibitors of the cysteine synthase CysM with antibacterial potency against dormant *Mycobacterium tuberculosis*

Katharina Brunner, Selma Maric, Rudraraju Srilakshmi Reshma, Helena Almqvist, Brinton A. Seashore-Ludlow, Anna-Lena Gustavsson, Ömer Poyraz, Perumal Yogeewari, Thomas Lundbäck, Michaela Vallin, Dharmarajan Sriram, Robert Schnell, and Gunter Schneider

J. Med. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.6b00674 • Publication Date (Web): 05 Jul 2016

Downloaded from <http://pubs.acs.org> on July 6, 2016

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.

1
2
3
4
5
6 **Inhibitors of the cysteine synthase CysM with antibacterial potency against**
7
8 ***dormant Mycobacterium tuberculosis***
9
10

11
12
13 Katharina Brunner^{a#}, Selma Maric^{a#}, Rudraraju Srilakshmi Reshma^b, Helena Almqvist^c, Brinton
14 Seashore-Ludlow^c, Anna-Lena Gustavsson^c, Ömer Poyraz^a, Perumal Yogeeswari^b, Thomas
15
16 Lundbäck^c, Michaela Vallin^c, Dharmarajan Sriram^b, Robert Schnell^a
17
18
19
20
21 and Gunter Schneider^a
22
23
24
25

26 ^a Division of Molecular Structural Biology, Department of Medical Biochemistry & Biophysics,
27 Karolinska Institutet, S-171 77 Stockholm, Sweden
28
29

30 ^b Drug Discovery Research Laboratory, Department of Pharmacy, Birla Institute of Technology
31 & Science-Pilani, Hyderabad Campus, Shameerpet, R.R. District, Hyderabad-500078, Andhra
32 Pradesh, India
33
34
35
36

37 ^c Chemical Biology Consortium Sweden, Science for Life Laboratory, Division of Translational
38 Medicine & Chemical Biology, Department of Medical Biochemistry & Biophysics, Karolinska
39 Institutet, S-171 65 Solna, Sweden
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

ABSTRACT

Cysteine is an important amino acid in the redox defense of *Mycobacterium tuberculosis*, primarily as a building block of mycothiol. Genetic studies have implicated *de novo* cysteine biosynthesis in pathogen survival in infected macrophages, in particular for persistent *M. tuberculosis*. Here, we report on the identification and characterization of potent inhibitors of CysM, a critical enzyme in cysteine biosynthesis during dormancy. A screening campaign of 17312 compounds identified ligands that bind to the active site with μM affinity. These were characterized in terms of their inhibitory potencies and structure-activity relationships through hit expansion guided by three-dimensional structures of enzyme-inhibitor complexes. The top compound binds to CysM with 300 nM affinity and displays selectivity over the mycobacterial homologues CysK1 and CysK2. Notably, two inhibitors show significant potency in a nutrient-starvation model of dormancy of *Mycobacterium tuberculosis*, with little or no cytotoxicity towards mammalian cells.

INTRODUCTION

Nearly one-third of the global population is infected with *Mycobacterium tuberculosis*, the causative agent of tuberculosis (http://www.who.int/tb/publications/global_report/en/). In most infected individuals, the bacilli reside in macrophages where they evade the immune defense mechanisms of the host by adopting a dormant lifestyle.¹ In this non-contagious phase of the disease, the pathogen is contained in granuloma formed by macrophages, lymphocytes and fibroblasts within the lung tissue.² The large number of infected individuals represents a reservoir of bacilli that upon relapse into the acute state of tuberculosis can further spread the disease. Since dormant bacteria are less susceptible to standard antibiotic treatment new anti-mycobacterial agents are needed that are effective against persistent *M. tuberculosis*.³

Host macrophages use reactive oxygen (ROI) and nitrogen (RNI) intermediates to kill the engulfed bacteria and redox defense mechanisms are therefore crucial for intracellular survival of the pathogen.⁴ In mycobacteria, intracellular redox homeostasis is largely maintained by mycothiol, the functional analog of glutathione.^{5,6} Mycothiol contains a cysteine-derived sulfhydryl group and its biosynthesis is dependent on the availability of L-cysteine.⁷ *De novo* cysteine biosynthesis relies on the provision of reduced sulfur by the sulfate reduction pathway (APS/PAPS pathway). It is therefore not surprising that metabolic pathways related to sulfur assimilation and cysteine biosynthesis are upregulated in several different dormancy models.^{8,9} Enzymes from these pathways have therefore been suggested as potential targets for new anti-mycobacterial compounds^{10,11} and first potent inhibitors of adenosine-5'-phosphosulfate reductase with bactericidal activity have been described.¹²

1
2
3
4
5
6 To date three pyridoxal-phosphate (PLP) dependent cysteine synthases have been characterized
7
8 in *M. tuberculosis* (Figure 1).¹¹ CysK1, a *bona fide* O-acetylserine sulfhydrylase (OASS), uses
9
10 O-acetylserine and hydrogen-sulfide from the APS/PAPS pathway for the synthesis of
11
12 cysteine.^{13,14} CysM on the other hand utilizes O-phosphoserine (OPS) together with a small
13
14 sulfur delivery protein CysO, which contains a thiocarboxylated carboxy terminus.¹⁵⁻¹⁷ A third
15
16 putative cysteine synthase, CysK2, also utilizes OPS as acceptor substrate and is able to produce
17
18 cysteine with hydrogen-sulfide as sulfur donor. However, this enzyme shows a preference for
19
20 thiosulfate, leading to S-sulfocysteine as the primary product and is thought to be involved in
21
22 redox-signaling.¹⁸ *CysM*, *cysO* and *cysK2* have all been shown to be up-regulated under
23
24 conditions simulating dormancy^{8,9,19,20} and randomized transposon mutagenesis studies
25
26 demonstrated profound phenotypic effects for *cysM* and *cysO* by gene disruption.²¹ These studies
27
28 showed attenuation of *cysO* and *cysM* mutants of *M. tuberculosis* in *in vitro* models of dormancy
29
30 and in a mouse model of infection²², indicating the CysM/CysO pathway as the dominant route
31
32 to cysteine in dormant bacteria. These findings are consistent with the observation that CysM
33
34 protects the oxidation sensitive α -aminoacrylate intermediate from redox stress induced by small
35
36 molecules such as hydrogen peroxide to a larger extent than CysK1.²³ Inhibitors of CysM are
37
38 therefore expected to weaken the redox defense lines of *M. tuberculosis* and potentially allow
39
40 natural antibacterial actions of the host macrophages to eliminate the internalized mycobacteria.
41
42
43
44
45
46
47
48
49

50 Here we describe the identification and characterization of small molecules targeting *M.*
51
52 *tuberculosis* CysM. Active site binders were identified in a small-molecule screening campaign
53
54 based on ligand-induced changes in fluorescence of the PLP cofactor, previously observed in
55
56
57
58
59
60

1
2
3 other members of this enzyme family.²⁴⁻²⁶ The inhibitory effects of hit compounds were tested in
4
5 follow-up enzyme activity assays, which were also employed for examination of structure-
6
7 activity relationships following synthesis of closely related analogs. The binding modes of seven
8
9 of the strongest inhibitors were furthermore determined through high-resolution crystal structures
10
11 of protein-ligand complexes. The ligands were shown to bind to the CysM active site nearly
12
13 parallel with the PLP-cofactor, forming extensive interactions with the target both at the
14
15 substrate carboxylate binding pocket and at an adjacent cavity. This cavity is flexible such that it
16
17 adapts to variously sized substituents. The best inhibitors showed target-specific affinity in the
18
19 0.3-3 μM range and bactericidal activity in a nutrient-starvation model of dormancy of
20
21 *Mycobacterium tuberculosis*. Prevention of the bactericidal effect of the inhibitors by the
22
23 addition of cysteine points towards perturbation of cysteine biosynthesis as a significant
24
25 component of their mode of action.
26
27
28
29
30
31
32
33

34 RESULTS AND DISCUSSION

35
36
37
38 **Small-molecule screening and hit selection:** Binding of small-molecule ligands to the active
39
40 site of CysM from *M. tuberculosis* is accompanied by an increase in the fluorescence of the
41
42 covalently bound cofactor PLP (Figures 2a & b). We exploited this property in a 384-well
43
44 microplate-based screening assay to identify new small-molecule binders of this enzyme. To this
45
46 end 17312 chemically diverse and lead- to drug-like compounds²⁷ were screened against
47
48 recombinant CysM at a single concentration of 10 μM at Chemical Biology Consortium Sweden
49
50 (www.cbcs.se). Compounds that resulted in an arbitrarily chosen fluorescence increase by more
51
52 than 30% were considered as primary hits that warranted further investigation at multiple
53
54
55
56
57
58
59
60

1
2
3 concentrations and counter-assays to remove auto-fluorescent compounds. This selection
4
5 criterion gave 431 initial hits for follow-up experiments (Figure 2c, Supporting Figure S1).
6
7
8 Counterscreening at three compound concentrations in the absence of CysM effectively removed
9
10 a significant portion of these compounds as auto-fluorescent, leaving 20 compounds showing
11
12 concentration-dependent binding to CysM with minimal interference from background
13
14 fluorescence of the compound itself (Supporting Figure S2a). This corresponds to a hit rate of
15
16 0.1% in our library. Seven of these confirmed hits demonstrated a saturating CysM dependent
17
18 fluorescence signal in the interval between 2.5 to 40 μM . After examination of both pan-assay
19
20 interference compounds (PAINS)²⁸; removal of unwanted reactive groups, (REOS filter)²⁹ and
21
22 potential aggregator filters³⁰ (Supporting Figure S2b and Supporting Table 1) one hit compound,
23
24 2-(4-phenylbenzoyl)benzoyl chloride (CBK241255) was filtered out. All other compounds
25
26 passed the filters and were regarded as having low to medium risk of being aggregators
27
28 (Supporting table 1). Among these six remaining compound structures, two were closely related
29
30 urea derivatives of which none had previously appeared as hits in other in-house screens. As
31
32 further analogs of this scaffold were easily accessible to us through analog searches of the CBCS
33
34 compound collection and from synthesis, we chose to focus our next efforts on this compound
35
36 series.
37
38
39
40
41
42
43
44
45

46 The changes in PLP fluorescence were next used to assess the apparent affinity of the two initial
47
48 urea hits in full concentration-response curves (Figure 3a and Supporting Figure S3), giving
49
50 apparent K_d values of $0.95 \pm 0.09 \mu\text{M}$ for 3-(3-(3,4-dichlorophenyl)ureido)benzoic acid **1** and
51
52 $1.7 \pm 0.2 \mu\text{M}$ for 3-(3-([1,1'-biphenyl]-3-yl)ureido) benzoic acid **2** (Table 1), respectively.
53
54
55 Although the fluorescence-based assay was robust enough to support the screening campaign and
56
57
58
59
60

1
2
3 hit confirmation, it is associated with significant limitations when it comes to the characterization
4
5 of potent binders. This is because a CysM concentration of 2 μM is required to obtain a reliable
6
7 signal to background ratio in the assay, such that stoichiometry reasons require the addition of
8
9 corresponding concentrations of ligand to saturate all protein in the solution. Consequently,
10
11 affinities approaching the sub- μM range cannot be accurately measured with this assay, and
12
13 measurements result in underestimations of their true affinities. For this reason, compounds with
14
15 potentially underestimated affinities (those with low μM apparent affinities in Figure 3d and
16
17 Supporting Figure S3) were re-evaluated whenever possible by isothermal titration calorimetry
18
19 (ITC), which is not affected by the same stoichiometric limitations. For compound **1**, ITC gave a
20
21 K_d value of $0.32 \pm 0.01 \mu\text{M}$ (Table 1), whereas binding of compound **2** could not be evaluated by
22
23 ITC. We therefore regard the value of $1.7 \pm 0.2 \mu\text{M}$ from the PLP fluorescence experiment as an
24
25 upper apparent K_d for this compound.
26
27
28
29
30
31
32
33

34 **Hit characterization and expansion:** Hit expansion was carried out by several rounds of
35
36 compound selection/synthesis, where three-dimensional structures of CysM–inhibitor complexes
37
38 (see below) in combination with biochemical data informed the design of the analogs. The
39
40 identified hits had urea as a linker bridging two flanking phenyl rings (Table 1). Based on this
41
42 scaffold, 17 analogs were selected from the in-house compound collection and small sub-
43
44 libraries comprising an additional 54 molecules were synthesized. The synthetic pathways for the
45
46 synthesis of these urea derivatives are delineated in Scheme 1. A total of 71 analogues were thus
47
48 characterized to derive an initial understanding of the structure-activity relationships (SAR).
49
50 Data for the most potent inhibitors are included in Table 1, while Supplementary Table S2 gives
51
52 a summary of the activity data for all tested molecules. Collectively these span an affinity range
53
54
55
56
57
58
59
60

1
2
3 from 300 nM to completely inactive compounds. We next employed an orthogonal assay to
4 evaluate inhibition of the CysM enzymatic activity by monitoring the α -aminoacrylate
5 intermediate formed in the first half-reaction (Supporting Figure S4). A significant reduction of
6 the apparent first order rate constants (k_{obs}) was observed for compounds **1** (Figure 3) and **2**
7 (Table 1), while the corresponding data for other representatives mirrored the apparent affinities
8 from the PLP-fluorescence assay (Table 1 and Supporting Table S3).
9
10
11
12
13
14
15
16
17
18
19

20 The second half-reaction of the catalytic cycle of CysM (Supporting Figure S4) uses
21 thiocarboxylated CysO (CysO-SH) as sulfur donor, leading to the formation of a covalent CysO-
22 cysteine adduct which can be detected by mass spectrometry.^{15,16} In order to assess the ability of
23 the best inhibitors of the first half reaction to completely inhibit CysM (i.e. no formation of the
24 covalent CysO-cysteine adduct) the reactions were carried out in the absence and presence of hit
25 compounds and analyzed by ESI-MS (Figure 3d, Supporting Figure S5). The only peaks detected
26 in the presence of compounds **1** and **2** were at ~9571 Da corresponding to the mass of
27 thiocarboxylated CysO (CysO-SH) showing that these compounds completely inhibit CysO-
28 cysteine adduct (expected molecular weight of 9658 Da) formation, i.e. the overall reaction of
29 the enzyme. Compounds **3**, **5–7** showed peaks at both 9571 Da and 9658 Da in different ratios
30 (Supporting Figure S5) indicating partial inhibition, while **17** showed a main peak at 9658 Da
31 demonstrating no or very limited inhibition of CysO-cysteine adduct formation in the CysM
32 catalyzed reaction compared to the top compounds. For the tightest binding compound **1** we also
33 determined the IC_{50} value for the overall reaction (Figure 3e), which with 0.48 μM agrees well
34 with the K_d of 0.32 μM determined using ITC.
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 **Overall three-dimensional structures of the enzyme-ligand complexes.** The crystal structures
4 of the complexes of CysM with seven inhibitors, **1-7**, were determined by x-ray crystallography
5 to 1.64 – 2.7 Å resolution (Supporting Figure S6 and tables S3 and S4). Overall the three-
6 dimensional structures of the enzyme complexes are similar to the structure of the unliganded
7 enzyme with r.m.s.d. values for the C α atoms in the range of 0.7 -1.4 Å upon superimposition. In
8 chain A of the orthorhombic crystal form (inhibitors **2-4**) and chains A, B, and C of the triclinic
9 space group (inhibitors **1,5-7**) the ligand is bound to the enzyme in an open conformation
10 characterized by a disordered active site loop (residues 211 – 227) and a disordered C-terminus,
11 residues (310-323). The disorder of these peptide segments results in an open active site, which
12 renders the inhibitor partially accessible from the solvent. In the B chains of the orthorhombic
13 crystals the active site loop is however well defined in electron density and has folded back over
14 the active site thus shielding the bound ligand from the solvent. Also in chain D of the triclinic
15 space group both active site loop and the C-terminal residues are well defined in density and
16 contribute to ligand binding. The active site loop is folded over the ligand-binding site and the C-
17 terminal residues 320-323 are inserted into the active site cleft. In the following we describe
18 enzyme-ligand interactions for the closed conformations as in these structures ligand interactions
19 are fully developed.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

45 **Enzyme-inhibitor interactions and emerging SAR:** The confirmed small-molecule inhibitors
46 consist of three parts: a left-hand-side (LHS) substituted phenyl group, a central linker part and a
47 right-hand-side (RHS) aryl group (Table 1). In all seven co-crystal structures, the compounds
48 bind in the enzyme active site parallel to the plane of the pyrimidine ring of the PLP cofactor,
49 which forms stacking interactions with the urea linker (Supporting Figure S6A).
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6 The LHS carboxylate group is consistently positioned in the substrate binding pocket in a similar
7
8 way as the carboxylate group of the α -aminoacrylate intermediate, as previously observed in
9
10 CysK1 from *M. tuberculosis*¹⁴, while the RHS binds into a hydrophobic pocket that extends from
11
12 the PLP site toward the enzyme surface. The recognition between the enzyme and the linker and
13
14 RHS, respectively, differs significantly between compounds carrying a RHS phenyl (denoted
15
16 type I inhibitors) or biphenyl group (type II inhibitors), respectively, (Figure 4) as outlined in
17
18 detail below.
19
20

21
22
23
24 The LHS of type I inhibitors **1**, **2**, **4–7** show similar interactions to residues in the carboxylate
25
26 binding pocket, i.e. hydrogen bonds to main chain amides of Thr78, Ser79, Asn81 and Thr82
27
28 located at the N-terminal of an α -helix (Figure 4a & b). The interaction between the LHS
29
30 carboxylate substituent and the enzyme in the carboxylate binding pocket is found in all potent
31
32 inhibitors of CysM homologues described so far, i.e. CysK1 from *M. tuberculosis*^{14,25}, OASS
33
34 from *Haemophilus influenzae*^{31,32} and *Salmonella typhimurium*²⁶ and thus appears to be a
35
36 hallmark of this enzyme family. The exchange of the carboxylic acid group for the tetrazole
37
38 bioisostere further supports this notion, as this resulted in a complete loss of affinity (compounds
39
40 **1** vs **8**, Table 1). Further examples reveal that type I inhibitors showed little tolerance for
41
42 introduction of substituents at LHS. Modifications such as addition of a hydroxyl group and a
43
44 switch from *m*- to *p*-carboxylate led to loss of binding (compounds **9–11**). However, type II
45
46 inhibitors with the corresponding LHS modifications retained binding affinity and showed
47
48 considerable inhibitory power (compounds **3–4** vs **9–10**). This difference relates to
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 conformational changes in the protein induced by the linker and the RHS of type II inhibitors as
4
5 discussed below.
6
7

8
9
10 We next explored the role of the central linker region. In the crystal structures of CysM with type
11
12 I inhibitors the carbonyl group of the urea linker interacts with the side chain of Asn221 (Figure
13
14 4a). One of the amides interacts with the terminal carboxyl group of Ala323, inserted into the
15
16 active site and the second amide forms a hydrogen bond to an internal water molecule. In the
17
18 structures of type II inhibitors, the carbonyl oxygen of the linker instead forms a hydrogen bond
19
20 with Arg220 from the mobile active site loop and both amides interact with internal water
21
22 molecules (Figure 4b). To understand the SAR and explore the potential for further optimization
23
24 of the hit compounds, the urea motif was exchanged for thiourea, sulfonamide or amide linkers.
25
26
27 The linker was also extended by a one-carbon elongation, but all these modifications gave
28
29 inactive or poorly active compounds (compounds **12–15**, Table 1), confirming the importance of
30
31 the urea hydrogen bonds observed. The difference in activity between urea and thiourea analogs
32
33 (e.g. compound **7** vs **12**) may be due to the weaker hydrogen bonding ability of sulfur compared
34
35 to oxygen and the larger size of the sulfur atom that may affect the angle between the LHS and
36
37 RHS phenyl groups. Additional examples of thiourea analogs supporting these results are
38
39 provided in Supporting Table S2 (e.g. compound **20** and **21**).
40
41
42
43
44
45
46
47

48 Significant differences between type I and II inhibitors were also found for the interactions of the
49
50 enzyme with the RHS. The crystal structures of CysM with **1**, **5**, **6** or **7** (type I) are all very
51
52 similar, with r.m.s.d. values for C α atoms after superimposition of 0.25–0.3 Å and they also
53
54 display the same enzyme-ligand interactions. Binding of the type I inhibitors induces a disorder-
55
56
57
58
59
60

1
2
3 order transition involving the mobile active site loop (residues 211–227), which folds over the
4 inhibitor binding site, and the C-terminal residues that insert into the active site and form
5 important interactions with the linker (Figure 4c). The RHS phenyl group is bound in a
6 hydrophobic pocket denoted pocket A (Figure 4d). The size of this pocket is too small to
7 accommodate inhibitors with a biphenyl substituent, and binding of type II compounds induces a
8 different conformation of the active site loop that increases the size of cavity A by pocket B. The
9 different conformation of the active site loop at the same time prevents insertion of the C-
10 terminus into the active site leading to a change in the interactions of the linker with enzyme
11 residues. The expansion of the binding pocket A with pocket B thus provides sufficient space for
12 the larger biphenyl substituent of compounds **2**, **3** and **4** (Figure 4b and e).
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

28 The unsubstituted phenyl analog **16** showed an apparent K_d of 12.9 μM , similar to derivatives
29 with the strongly electron-withdrawing *p*-NO₂ substituent and the electron-donating *p*-OMe
30 group (compounds **17** and **18**) as well as the *p*-Me analog **5**. However, substitutions with
31 halogens (-Cl **7**, -Br **6**) resulted in a stronger binding, indicating that *p*-halogens are preferred.
32 Substitution with chloride in both the *para* and *meta*-position resulted in the best inhibitor **1** with
33 a 40 fold better K_d value than the parent compound **16**. The electron withdrawing acetyl group
34 had a negative effect on the activity as compound **19** was a poor inhibitor.
35
36
37
38
39
40
41
42
43
44

45 The terminal phenyl group of compound **2** binds in pocket B (Figure 4e) and the crystal
46 structures of CysM with **2**, **3** or **4** show that there is little additional space available in this pocket
47 (less than 4 Å in any direction) for further modifications of this phenyl ring. Consequently,
48 substitutions such as -Me, -CO₂H or -F all result in lower affinity or a complete loss of binding
49 (compounds **58**, **66**, and **69**, Supporting Table S2). Exchanging the RHS terminal phenyl ring for
50
51
52
53
54
55
56
57
58
59
60

1
2
3 a nitrogen heterocycle (pyridine or pyrimidine) (compounds **63**, **68**) led to less active compounds
4
5 (1–2 orders of magnitude difference in K_d), most likely due to less favorable interactions of these
6
7 more polar ring systems with the hydrophobic binding pocket.
8
9

10
11
12
13 **Anti-mycobacterial activity and cytotoxicity:** The most potent CysM inhibitors were further
14
15 tested against a growing culture of *M. tuberculosis* and all showed modest bactericidal effects,
16
17 with MIC values ranging from 2 – 20 μ M (0.7- 6.2 μ g/ml) (Table 2). Although these observed
18
19 MIC values are in the low μ M range, they do not reach the potency of clinically used drugs such
20
21 as isoniazid and rifampicin. As the gene coding for CysM is predominantly expressed during
22
23 dormancy^{8,9,19} this is also not to be expected. We therefore used the nutrient starvation model¹⁹ to
24
25 test the anti-mycobacterial activity of these inhibitors at 10 μ M concentration under conditions
26
27 simulating dormancy, including a comparison with the drugs isoniazid, rifampicin and
28
29 moxifloxacin (Figure 5A). As expected isoniazid showed only a marginal (~ 1 log) effect on
30
31 non-replicating cells, and compounds **6**, **7**, **17** and **18** displayed a similar lack of potency.
32
33 Compounds **1**, **2**, **3** and **5** performed considerably better, with **1** and **3** showing a 3-log decrease
34
35 in bacterial count, corresponding to similar or improved killing activities when compared to
36
37 rifampicin and moxifloxacin.
38
39
40
41
42
43
44
45

46
47 In an effort to initiate studies of the mechanism of action of these compounds also *in vivo* we
48
49 next examined the effect of cysteine complementation during the treatment of nutrient-starved
50
51 cultures of *M. tuberculosis* with compounds **1** and **3** (Figure 5B). Addition of cysteine (0.2 mM)
52
53 led to a complete loss of the bactericidal effect of the two inhibitors. We could not detect any
54
55
56
57
58
59
60

1
2
3 degradation of compound **1** during incubation with cysteine for 24 hrs using mass-spectrometry,
4
5 HPLC and spectroscopy, thus excluding inactivation of **1** by reaction with cysteine as the cause
6
7 for the loss of anti-mycobacterial activity. These data point towards cysteine biosynthesis as
8
9 important component of the action of compounds **1** and **3** also *in vivo*. Additional experiments
10
11 are required to further validate these data and to exclude contributions from other mechanisms.
12
13
14
15
16

17
18 To better understand whether these inhibitors are useful for later studies also in the context of
19
20 bacterial viability within mammalian cells we performed cell viability studies in a small panel of
21
22 human and mouse-derived cell lines. These cytotoxicity assays with the urea derivatives using
23
24 three different human cell lines and a mouse cell line did not reveal significant cytotoxic effects
25
26 at concentrations of 10 μM (Supporting Figure S7 and S8).
27
28
29
30
31

32 **Compound selectivity:** *De novo* cysteine biosynthesis in *M. tuberculosis* is rather complex, with
33
34 three potential pathways leading to cysteine. Each of these routes contains a specific PLP-
35
36 dependent cysteine synthase, CysK1, CysK2 and CysM (Figure 1).¹¹ In order to assess the
37
38 selectivity of the urea-derived inhibitors, we examined the binding and affinity of compound **1**
39
40 and **2** to the mycobacterial homologues using the fluorescence-based binding assay. Compounds
41
42 **1** and **2** bound to CysK1, albeit much weaker, returning K_d values of $241.1 \pm 25.3 \mu\text{M}$ and 96.8
43
44 $\pm 16.1 \mu\text{M}$, respectively, showing an eighty to several hundred fold stronger binding to CysM
45
46 than to CysK1. CysK2 also displayed weaker interactions with **1** ($30.1 \pm 3.4 \mu\text{M}$) and **2** (22.6 ± 2.4
47
48 μM), corresponding to a 15-100 fold preference of these compounds for CysM over CysK2.
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 As humans lack the *de novo* biosynthesis pathways to cysteine, these enzymes are particularly
4 attractive as potential targets for new antibiotics. A search of the human genome returned CBSA
5 (Cystathionine beta-synthase, Uniprot AccNr: P35520, 29% sequence identity) that uses serine
6 and homocysteine as substrates as the closest sequence homologue of CysM. CBSA binds
7 compound **1** with a K_d of 94 μ M, i.e. with approximately 300 times lower affinity compared to
8 the mycobacterial target protein and does not bind **2**. The low affinity of the inhibitors for the
9 closest human homologue and their low cytotoxicity suggest that these compounds are suitable
10 starting points for further development.
11
12
13
14
15
16
17
18
19
20
21
22
23

24 CONCLUSIONS

25
26 Activity assays of enzymes utilizing other post-translationally modified proteins, for instance
27 thiocarboxylated CysO, as (co-)substrates can be challenging and are not easily adapted to high
28 throughput format. Here we monitored changes in the fluorescence emission of the PLP cofactor
29 in 384-well format and successfully screened a chemically diverse substance library. The screen
30 resulted in the identification of new small-molecule binders of CysM following removal of false
31 positives that were disregarded based on their inherent fluorescence. We note, however, that our
32 approach does not capture those compounds that bind at sites far from the PLP cofactor and thus
33 potential allosteric inhibitors may have been missed.
34
35
36
37
38
39
40
41
42
43
44
45
46
47

48 One of the identified compound classes is based on an urea scaffold. This finding is in line with
49 results from a previous phenotypic high-throughput screen for inhibitors of *M. tuberculosis*
50 H37Rv that also included several urea-based hits with anti-mycobacterial activities.³³ We
51 explored the urea hits from our target-based screen in more detail, which led to the discovery of
52
53
54
55
56
57
58
59
60

1
2
3 potent CysM inhibitors that all displayed a similar binding mode to the enzyme active site.
4
5 However, active site loops of CysM showed a considerable flexibility in creating suitable
6
7 binding pockets to accommodate the differently substituted inhibitors. This conformational
8
9 flexibility of the enzyme is difficult to predict and illustrates present limitations in the scope of *in*
10
11 *silico* screening approaches.
12
13

14
15
16 A search of the Reaxys Medicinal Chemistry (www.reaxys.com) and the ChEMBL³⁴ databases
17
18 using the structures of compounds **1**, **2** and **3** regarding previously reported bioactivities
19
20 identified compound **1** in a patent³⁵ claiming inhibitors of the mammalian kinase p38, whereas
21
22 no known bioactivities could be found for **2** and **3**. Based on information from the co-crystal
23
24 structures and the emerging SAR we next expanded the search to include compounds displaying
25
26 80 % structural similarity to our top hits, retaining the crucial *m*-carboxylate on the LHS,
27
28 including the unsubstituted analog **16**. This search identified six compounds, for which a total of
29
30 147 bioactivities have been reported in ChEMBL. None of these is on targets of mycobacterial
31
32 origin. Activities include the mammalian targets Niemann-Pick C1, Ras-related protein Rab-
33
34 9A, luciferin 4-monooxygenase, GLP1-receptor, microRNA21, ATPase family AAA domain-
35
36 containing protein 5, nuclear factor erythroid 2-related factor 2, Ataxin-2, H4 and broadly on
37
38 Jurkat cells. Further optimization of our CysM top hits will have to address this issue, but the
39
40 lack of cytotoxicity of the top hits are encouraging to progress further.
41
42
43
44
45
46
47

48
49 In view of the large number of individuals infected with persistent *M. tuberculosis* worldwide
50
51 drugs active against the latent stage of this infectious disease are urgently needed. Indeed most of
52
53 the anti-mycobacterial drugs in clinical use are not effective in the dormant state of the disease.
54
55 Significantly, several of the top hits identified in this work display potent antibacterial activity,
56
57
58
59
60

1
2
3 with three compounds showing potency against non-replicating *M. tuberculosis* and thus are
4
5 promising leads for further development. First experiments towards the validation of CysM as
6
7 the *in vivo* target of the top hits showed that addition of cysteine during compound treatment led
8
9 to complete loss of bactericidal activity in a nutrient-starvation model of *M. tuberculosis*,
10
11 suggesting perturbed cysteine biosynthesis in the pathogen as a result of compound exposure.
12
13 Although these results are promising we note that unambiguous target validation requires more
14
15 elaborate efforts to firmly establish the *in vivo* mode of action of these molecules. Thus, at
16
17 present, we cannot exclude additional off-targets effects.
18
19
20
21
22
23

24
25 The genome of *M. tuberculosis* contains three genes that code for OASS homologues that have
26
27 been associated with three different *de novo* biosynthetic pathways to cysteine in this organism
28
29 (Figure 1). These enzymes belong to the same family, with sequence identities ranging from 26–
30
31 37%. While the precise role of these pathways and their regulation is not well understood, it
32
33 appears that cysteine synthesis under normal growth conditions relies predominantly on the
34
35 classical, CysK1-dependent pathway, whereas CysM/CysO dependent cysteine biosynthesis
36
37 occurs during intra-macrophage lifestyle and dormancy.^{21,22} Selective inhibitors of enzymes from
38
39 these pathways, for instance compound **1**, which inhibits CysM more potently than other family
40
41 members, are useful chemical tools to elucidate the regulation of cysteine biosynthesis in *M.*
42
43
44
45
46 *tuberculosis* under different metabolic conditions.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

MATERIALS & METHODS

Protein production. Recombinant CysM, CysO, CysK1 and CysK2 were produced and purified as described in the Supporting Information.

PLP-fluorescence-based screening assay. The CysM small-molecule screen was performed by monitoring the fluorescence of the PLP cofactor. A library of 17312 compounds was screened at a single concentration of 10 μM in the presence of 2 μM CysM. Compounds that showed an increase in fluorescence by $\geq 30\%$ were identified as hits. For details about the small molecule library, the screening assay and the follow-up hit confirmation as well as the counter assay experiments see Supporting Information.

Chemistry – General methods: The compounds in this study were either cherry-picked from the CBCS compound selection or synthesized. For synthesis, all commercially available chemicals and solvents were used without further purification. Analytical thin layer chromatography (TLC) was performed on alumina-backed silica gel 40 F254 plates (Merck, Darmstadt, Germany) that were visualized with UV light and KMnO_4 solution, and flash column chromatography was carried out on 60 \AA (40–63 μm) silica gel. HPLC-MS analyses were run on an Agilent series 6100B system with electrospray ionization (ESI+). ^1H and ^{13}C NMR spectra were recorded at 400 MHz and 100 MHz respectively on a Bruker AM spectrometer. Chemical shifts are reported in ppm (δ) with reference to TMS as internal standard, and coupling constants J are quoted in Hertz. Elemental analyses were carried out using a Vario MICRO cube in CHN mode. The experimental details of the most potent CysM inhibitors 1–7 are described below. A

1
2
3 full description of the synthetic protocol and spectroscopic analysis of the other compounds is
4
5 found in the Supporting information. All compounds show a purity > 95% as determined by
6
7 HPLC.
8
9

10
11
12 The seven hit compounds were examined using the PAINS²⁸ and REOS²⁹ filters in the software
13
14 Canvas³⁶. Also, the seven hits were evaluated as potential aggregators³⁰, using structural
15
16 similarity to known aggregators and SlogP calculations in a Knime work flow
17
18 (<http://www.myexperiment.org/workflows/4749.html>) assessing the aggregating risk.
19
20
21

22
23 **General procedure for the synthesis of 1, 5, 6 and 7:** Substituted phenyl isocyanate (1 mmol)
24
25 was added dropwise to a stirred ice-cooled solution of 3-amino benzoic acid, 4-amino salicylic
26
27 acid, 5-amino salicylic acid or 4-amino-3-hydroxy benzoic acid (1 mmol), respectively, and
28
29 trimethylamine (2 mmol) in dry ethanol. Reaction progress was monitored by TLC and HPLC-
30
31 MS. After 24 h, solvent was evaporated under reduced pressure and the resulting residue was
32
33 diluted with aq. sodium bicarbonate and the mixture was filtered. The filtrate was acidified with
34
35 2N hydrochloric acid and the precipitate was filtered and recrystallized in ethanol to yield the
36
37 respective urea derivatives.
38
39
40
41

42
43 **General procedure for the synthesis of 2, 3 and 4:** Substituted 3-(3-(4-
44
45 bromophenyl)ureido)benzoic acid (1 mmol), the corresponding boronic acid (2 mmol), sodium
46
47 carbonate (2M solution) (4 mmol) and Pd(PPh₃)₄ (0.06 mmol) were taken in acetonitrile and
48
49 irradiated in microwave at 100 °C for 1 h (monitored by TLC and LCMS for completion). The
50
51 reaction mixture was passed through celite and washed with methanol. The filtrate was
52
53
54
55
56
57
58
59
60

1
2
3 concentrated and purified by column chromatography using hexane:ethyl acetate as eluent to get
4
5 the desired product in good yield.
6
7
8
9

10
11 *3-(3-(3,4-Dichlorophenyl)ureido)benzoic acid (1)*. The compound was synthesized according to
12 the general procedure using 3-amino benzoic acid (0.10g, 0.73 mmol), 3,4-dichloro phenyl
13 isocyanate (0.14g, 0.73 mmol) and triethylamine (0.15g, 1.46 mmol) to afford **1** as white solid
14 (0.18g, 75% yield). M.P.:119 – 121 °C. ¹H NMR (DMSO-d₆): δ 12.96 (br s, 1H), 9.30 (s, 1H),
15 9.25 (s, 1H), 8.14 (s, 1H), 7.90 (s, 1H), 7.59 (d, *J* = 8 Hz, 1H), 7.54 (d, *J* = 8.8 Hz, 1H), 7.44 (t, *J*
16 = 8 Hz, 1H), 7.37 (m, 2H). ¹³C NMR (DMSO-d₆): δ 167.2, 152.3, 139.9, 139.6, 131.3, 131.0,
17 130.5, 129.0, 123.1, 123.0, 122.1, 119.3, 119.0, 118.3. ESI-MS *m/z* 324.01(M-H)⁺. Anal. Calcd
18 for C₁₄H₁₀Cl₂N₂O₃; C, 51.72; H, 3.10; N, 8.62; Found C, 51.82; H, 3.09; N, 8.64.
19
20
21
22
23
24
25
26
27
28
29
30
31

32
33 *3-(3-([1,1'-Biphenyl]-3-yl)ureido) benzoic acid (2)*. The compound was synthesized according
34 to the general procedure using 3-(3-(3-bromophenyl)ureido)benzoic acid (0.1 g, 0.24 mmol),
35 phenyl boronic acid (0.057 g, 0.47 mmol), Na₂CO₃ (0.1g, 0.96 mmol) and Tetrakis (0.019, 0.016
36 mmol) to afford **2** as white solid (0.17 g, 79% yield). M.P. 229–231 °C. ¹H NMR (DMSO-d₆): δ
37 12.8 (br s, 1H), 9.54 (s, 1H), 9.48 (s, 1H), 8.13 (s, 1H), 8.03 (s, 1H), 7.99 – 7.79 (m, 2H), 7.64 –
38 7.32 (m, 9H). ¹³C NMR (DMSO-d₆): δ 167.2, 152.6, 144.6, 141.2, 140.6, 135.8, 130.2 (2C),
39 129.8, 129.0, 128.5, 127.9, 127.6 (2C), 126.4, 126.3, 124.5, 122.2, 121.8, 120.6. EI-MS *m/z*:
40 331.01 (M-H)⁺. Anal. Calcd for C₂₀H₁₆N₂O₃: C, 72.28; H, 4.85; N, 8.43. Found: C, 72.46; H,
41 4.86; N, 8.41.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 *4-(3-([1,1'-Biphenyl]-3-yl)ureido)-2-hydroxybenzoic acid (3)*. The compound was synthesized
4
5 according to the general procedure using 4-(3-(3-bromophenyl)ureido)-2-hydroxybenzoic acid
6
7 (0.1 g, 0.28 mmol), phenyl boronic acid (0.069 g, 0.56 mmol), Na₂CO₃ (0.12g, 1.12 mmol) and
8
9 Tetrakis (0.021 g, 0.019 mmol) to afford **3** as white solid (0.042 g, 43% yield). MP: 210 – 212
10
11 °C. ¹H NMR (DMSO-d₆): δ 12.86 (br s, 1H), 9.68 (s, 1H), 9.57 (s, 1H), 7.87 (s, 1H), 7.66 (t, *J* =
12
13 8 Hz, 3H), 7.50 – 7.35 (m, 5H), 7.26 (d, *J* = 7.6 Hz, 1H), 7.03 (s, 1H), 6.84 (d, *J* = 7.2, 1H), 4.63
14
15 (s, 1H). ¹³C NMR (DMSO-d₆): δ 172.6, 162.9, 152.6, 143.6, 140.7, 140.6, 140.4, 130.6, 129.3,
16
17 128.9 (2C), 127.4, 126.6 (2C), 120.0, 117.2, 116.4, 113.2, 107.1, 104.5. EI-MS *m/z*: 347.01 (M-
18
19 H)⁺. Anal. Calcd for C₂₀H₁₆N₂O₄: C, 68.96; H, 4.63; N, 8.04; Found C, 69.13; H, 4.62; N, 8.03.
20
21
22
23
24
25
26

27 *5-(3-([1,1'-Biphenyl]-3-yl)ureido)-2-hydroxybenzoic acid (4)*. The compound was synthesized
28
29 according to the general procedure using 5-(3-(3-bromophenyl)ureido)-2-hydroxybenzoic acid
30
31 (0.1 g, 0.28 mmol), phenylboronic acid (0.069 g, 0.56 mmol), Na₂CO₃ (0.12g, 1.12 mmol) and
32
33 Tetrakis (0.021 g, 0.019 mmol) to afford **4** as brown solid (0.048 g, 48% yield). MP: 208 – 210
34
35 °C. ¹H NMR (DMSO-d₆): δ 12.79 (br s, 1H), 9.55 (s, 1H), 9.49 (s, 1H), 8.12 (s, 1H), 8.05 (s,
36
37 1H), 7.93 (d, *J* = 8.2 Hz, 1H), 7.65 – 7.01 (m, 9H), 5.03 (s, 1H). ¹³C NMR (DMSO-d₆): δ 172.3,
38
39 162.2, 152.5, 139.6, 139.0, 137.8, 137.5, 130.6, 130.2 (2C), 129.5, 128.3 (2C), 128.2, 124.6,
40
41 122.8, 121.6, 120.5, 118.7, 117.7. EI-MS *m/z*: 347.01 (M-H)⁺. Anal. Calcd for C₂₀H₁₆N₂O₄: C,
42
43 68.96; H, 4.63; N, 8.04; Found C, 69.10; H, 4.64; N, 8.02.
44
45
46
47
48
49

50 *3-(3-(p-Tolyl)ureido) benzoic acid (5)*. The compound was synthesized according to the general
51
52 procedure using 3-amino benzoic acid (0.10g, 0.73 mmol), p-tolylisocyanate (0.09g, 0.73 mmol)
53
54 and triethylamine (0.15g, 1.46 mmol) to afford **5** as pale yellow solid (0.14g, 71% yield). M.P:
55
56
57
58
59
60

1
2
3 242 – 244 °C. ^1H NMR (DMSO- d_6): δ 12.74 (br s, 1H), 9.32 (s, 1H), 8.91 (s, 1H), 8.23 (s, 1H),
4
5 7.78 (d, J = 8.2 Hz, 2H), 7.73 – 7.41 (m, 5H), 2.56 (s, 3H). ^{13}C NMR (DMSO- d_6): δ 167.5,
6
7 152.7, 142.7, 138.9, 138.1, 130.7 (2C), 126.7, 126.3, 124.4, 123.9, 119.8 (2C), 118.2, 22.6. ESI-
8
9 MS m/z 269.10(M-H) $^+$. Anal. Calcd for $\text{C}_{15}\text{H}_{14}\text{N}_2\text{O}_3$; C, 66.66; H, 5.22; N, 10.36; Found C,
10
11 66.48; H, 5.24; N, 10.39.
12
13
14
15
16
17

18 *3-(3-(4-Bromophenyl)ureido)benzoic acid (6)*. The compound was synthesized according to the
19
20 general procedure using 3-amino benzoic acid (0.10g, 0.73 mmol), 4-bromo phenyl isocyanate
21
22 (0.14g, 0.73 mmol) and triethylamine (0.15g, 1.46 mmol) to afford **6** as white solid (0.19g, 80%
23
24 yield). M.P: 275 – 277 °C. ^1H NMR (DMSO- d_6): δ 12.63 (br s, 1H), 9.00 (s, 1H), 8.93 (s, 1H),
25
26 8.12 (s, 1H), 7.65 (m, 1H), 7.57 (d, J = 8 Hz, 1H), 7.45 – 7.39 (m, 5H). ^{13}C NMR (DMSO- d_6): δ
27
28 167.3, 152.7, 145.8, 140.9, 133.8 (2C), 129.6, 128.4, 124.1, 122.8, 120.3, 119.6 (2C), 118.5. ESI-
29
30 MS m/z 334.01 (M-H) $^+$. Anal. Calcd for $\text{C}_{14}\text{H}_{11}\text{BrN}_2\text{O}_3$; C, 50.17; H, 3.31; N, 8.36; Found C,
31
32 50.32; H, 3.33; N, 8.38.
33
34
35
36
37
38

39 *3-(3-(4-Chlorophenyl)ureido)benzoic acid (7)*. The compound was synthesized according to the
40
41 general procedure using 3-amino benzoic acid (0.10g, 0.73 mmol), 4-chloro phenyl isocyanate
42
43 (0.11g, 0.73 mmol) and triethylamine (0.15g, 1.46 mmol) to afford **7** as white solid (0.17g, 78%
44
45 yield). M.P: 279 – 281 °C. ^1H NMR (DMSO- d_6): δ 12.88 (br s, 1H), 8.98 (s, 1H), 8.89 (s, 1H),
46
47 8.15 (s, 1H), 7.88 (d, J = 8 Hz, 2H), 7.68 – 7.36 (m, 5H). ^{13}C NMR (DMSO- d_6): δ 168.0, 152.7,
48
49 141.9, 139.3, 135.8, 130.3 (2C), 129.6, 128.5, 127.1, 123.4, 121.6, 118.6 (2C). ESI-MS m/z
50
51 289.05 (M-H) $^+$. Anal. Calcd for $\text{C}_{14}\text{H}_{11}\text{ClN}_2\text{O}_3$; C, 57.84; H, 3.81; N, 9.64; Found C, 57.67; H,
52
53 3.82; N, 9.62.
54
55
56
57
58
59
60

1
2
3
4
5
6 **Enzyme assays.** The catalytic cycle of CysM and the basis for the different assays used in this
7 study are shown in supporting Figure S4. The influence of the inhibitors on the first half-reaction
8 of CysM was monitored spectrophotometrically at 463 nm, the characteristic wavelength of the
9 aminoacrylate intermediate.¹⁶ The apparent k_{obs} values were derived from exponential fits to the
10 recorded absorbance data and compared to the non-inhibited reaction.
11
12
13
14
15
16
17
18
19

20 Electro spray ionization mass spectrometry was used to assess the formation of the covalent
21 CysO-cysteine adduct in the second half-reaction of the CysM catalytic cycle in the presence of
22 top HTS hits and their derivatives. Mass spectra of the reaction mixture products were compared
23 to the mass spectrum of the covalent CysO-cysteine adduct formed in the non-inhibited reaction
24 (9658.7 Da) as well as to the mass spectrum of the substrate CysO-SH (9571.6 Da). Multiple
25 turnover enzyme activity assays for IC_{50} determination for CysM with the natural substrates OPS
26 and thiocarboxylated CysO (CysO-SH) were based on the spectrophotometric determination of
27 the stoichiometric amount of phosphate ions released using the Malachite-green (MG) reagent.
28 For details of the assays, see Supporting Information.
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

44 **ASSOCIATED CONTENT**

45 **Supporting Information.**

46
47
48
49 Additional experimental details concerning isothermal titration calorimetry, compound synthesis,
50 crystallization and structure determination of enzyme-hit complexes, mammalian cell
51
52
53
54
55
56
57
58
59
60

1
2
3 cytotoxicity evaluation and anti-mycobacterial potency assays. This material is available free of
4 charge via the Internet.
5
6
7
8
9

10 **Accession Codes**

11
12 Crystallographic data for the enzyme-ligand complexes have been deposited with the Protein
13 Data Bank under accession codes 5I7A (CysM-1 complex), 5I7R (CysM-2 complex), 5IWC
14 (CysM-3 complex), 5IW8 (CysM-4 complex), 5I6D (CysM-5 complex), 5I7H (CysM-6
15 complex) and 5I7O (CysM-7 complex). The crystallographic data will be released upon article
16 publication.
17
18
19
20
21
22
23

24 **AUTHOR INFORMATION**

25 **Corresponding Author**

26
27
28 G. Schneider e-mail: Gunter.Schneider@ki.se; Phone:+4652487675 FAX:+468327626; R.
29
30

31
32 Schnell, e-mail: robert.schnell@ki.se; D. Sriram, email: dsriram@hyderabad.bits-pilani.ac.in
33
34
35

36 **Author contribution**

37
38 #: equally contribution authors
39
40
41
42

43 **Notes**

44
45 The authors declare no competing financial interest.
46
47
48
49

50 **ACKNOWLEDGMENTS**

51
52 We gratefully acknowledge access to synchrotron radiation at the ESRF, France and BESSY,
53 Germany and to instrumentation at the Protein Science Facility, Karolinska Institutet. We also
54
55
56
57
58
59
60

1
2
3 thank Jolanta Kopec, Structural Genomics Consortium (SGC-Oxford) for provision of a sample
4
5 of CBSA. This work was supported by the Swedish Research Council and the Department of
6
7 Biotechnology (DBT), India. Chemical Biology Consortium Sweden is funded by the Karolinska
8
9 Institutet, SciLifeLab and the Swedish Research Council. DS acknowledges DBT for the Tata
10
11 Innovation Fellowship.
12
13

14 15 16 17 **ABBREVIATIONS USED**

18
19 APS, adenosine-5'-phosphosulfate; CBSA cystathionine-beta-synthase; HTS high-throughput
20
21 screening; INH isoniazid; ITC, isothermal titration calorimetry; MIC, minimal inhibitory
22
23 concentration; MPN, most probable number; MOXI, moxifloxacin; OAS, O-acetyl-L-serine;
24
25 OASS, O-acetylserine sulfhydrylase, OPS, O-phospho-L-serine; RIF, rifampicin; RNI, reactive
26
27 nitrogen intermediates; PAPS, 3'-phosphoadenosine-5'-phosphosulfate; PLP, pyridoxal-
28
29 phosphate; PAINS, pan assay interference compounds; REOS, removal of unwanted reactive
30
31 groups; ROI, reactive oxygen intermediates; TMS, tetramethylsilane.
32
33
34
35
36

37 38 **REFERENCES**

- 39
40 (1) Gengenbacher, M.; Kaufmann, S.H. *Mycobacterium tuberculosis*: success through dormancy.
41
42 *FEMS Microbiol. Rev.* **2012**, *36*, 514-532.
43
44 (2) Guirado, E.; Schlesinger, L.S. Modeling the *Mycobacterium tuberculosis* granuloma - the
45
46 critical battlefield in host immunity and disease. *Front. Immunol.* **2013**, *4*, 98.
47
48 (3) Koul, A.; Arnoult, E.; Lounis, N.; Guillemont, J.; Andries, K. The challenge of new drug
49
50 discovery for tuberculosis. *Nature* **2011**, *469*, 483-490.
51
52
53
54
55
56
57
58
59
60

- 1
2
3 (4) Dorhoi, A.; Reece, S.T.; Kaufmann S.H. For better or for worse: the immune response
4 against *Mycobacterium tuberculosis* balances pathology and protection. *Immunol. Rev.* **2011**,
5
6 240, 235-251.
7
8
9
10 (5) Sareen, D.; Newton, G.L.; Fahey, R.C.; Buchmeier, N.A. Mycothiol is essential for growth of
11
12 *Mycobacterium tuberculosis*. *J. Bacteriol.* **2003**, 185, 6736–6740.
13
14
15 (6) Bhaskar, A.; Chawla, M.; Mehta, M.; Parikh, P.; Chandra, P.; Bhawe, D.; Kumar, D.; Carroll,
16
17 K.S.; Singh, A. Reengineering redox sensitive GFP to measure mycothiol redox potential of
18
19 *Mycobacterium tuberculosis* during infection. *PLoS Pathog.* **2016**, 10, e1003902.
20
21
22 (7) Jothivasan, V. K.; Hamilton, C. J. Mycothiol: synthesis, biosynthesis and biological functions
23
24 of the major low molecular weight thiol in actinomycetes. *Nat. Prod. Rep.* **2008**, 25,
25
26 1091-1117.
27
28
29 (8) Schnappinger, D.; Ehrt, S.; Voskuil, M. I.; Liu, Y.; Mangan, J. A.; Monahan, I. M.;
30
31 Dolganov, G.; Efron, B.; Butcher, P. D.; Nathan, C.; Schoolnik, G. K. Transcriptional adaptation
32
33 of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment.
34
35 *J. Exp. Med.* **2003**, 198, 693-704.
36
37
38 (9) Voskuil, M. I.; Visconti, K. C.; Schoolnik, G. K. *Mycobacterium tuberculosis* gene
39
40 expression during adaptation to stationary phase and low-oxygen dormancy. *Tuberculosis* **2004**,
41
42 84, 218-227.
43
44
45 (10) Bhawe, D. P.; Muse, W. B., 3rd; Carroll, K. S. Drug targets in mycobacterial sulfur
46
47 metabolism. *Infect. Disord. Drug Targets* **2007**, 7, 140-158.
48
49
50 (11) Schnell, R.; Sriram, D.; Schneider, G. Pyridoxal-phosphate dependent mycobacterial
51
52 cysteine synthases: structure, mechanism and potential as drug targets. *Biochem. Biophys. Acta*
53
54 **2014**, 1854, 1175-1183
55
56
57
58
59
60

- 1
2
3 (12) Palde, P.B.; Bhaskar, A.; Pedró Rosa, L.E.; Madoux, F.; Chase, P.; Gupta, V.; Spicer, T.;
4
5 Scampavia, L.; Singh, A.; Carroll K.S. First-in-class inhibitors of sulfur metabolism with
6
7 bactericidal activity against non-replicating *M. tuberculosis*. *ACS Chem. Biol.* **2016**, *11*, 172-184
8
9
10 (13) Schnell, R.; Sandalova, T.; Hellman, U.; Lindqvist, Y.; Schneider, G. Siroheme- and [Fe4-
11
12 S4]-dependent NirA from *Mycobacterium tuberculosis* is a sulfite reductase with a covalent Cys-
13
14 Tyr bond in the active site, *J. Biol. Chem.* **2005**, *280*, 27319-27328.
15
16
17 (14) Schnell, R.; Oehlmann, W.; Singh, M.; Schneider, G. Structural insights into catalysis and
18
19 inhibition of O-acetylserine sulfhydrylase from *Mycobacterium tuberculosis* - crystal structures
20
21 of the enzyme alpha-aminoacrylate intermediate and an enzyme-inhibitor complex. *J. Biol.*
22
23 *Chem.* **2007**, *282*, 23473-23481.
24
25
26 (15) Burns, K.E.; Baumgart, S.; Dorrestein, P.C.; Zhai, H.; McLafferty, F.W.; Begley, T.P.
27
28 Reconstitution of a new cysteine biosynthetic pathway in *Mycobacterium tuberculosis*. *J. Am.*
29
30 *Chem. Soc.* **2005**, *127*, 11602-11603.
31
32
33 (16) Ågren D.; Schnell R.; Oehlmann D.W.; Singh M.; Schneider G. Cysteine synthase (CysM)
34
35 of *Mycobacterium tuberculosis* is an O-phosphoserine sulfhydrylase: evidence for an alternative
36
37 cysteine biosynthesis pathway in mycobacteria. *J. Biol. Chem.* **2008**, *283*, 31567-31574.
38
39
40 (17) O'Leary, S.E.; Jurgenson, C.T.; Ealick, S.E.; Begley, T.P. O-phospho-L-serine and the
41
42 thiocarboxylated sulfur carrier protein CysO-COSH are substrates for CysM, a cysteine synthase
43
44 from *Mycobacterium tuberculosis*. *Biochemistry* **2008**, *47*, 11606-11615.
45
46
47 (18) Steiner, E.M.; Böth, D.; Lössl, P.; Vilaplana, F.; Schnell, R.; Schneider, G. CysK2 from
48
49 *Mycobacterium tuberculosis* is an O-phospho-L-Serine dependent S-sulfocysteine synthase. *J*
50
51 *Bacteriol.* **2014**, *196*, 3410-3420.
52
53
54
55
56
57
58
59
60

- 1
2
3 (19) Betts, J. C.; Lukey, P. T.; Robb, L. C.; McAdam, R. A.; Duncan, K. Evaluation of a nutrient
4 starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression
5 profiling. *Mol. Microbiol.* **2002**, *43*, 717-731.
6
7
8
9
10 (20) Hampshire, T.; Soneji, S.; Bacon, J.; James, B. W.; Hinds, J.; Laing, K.; Stabler, R. A.;
11 Marsh, P. D.; Butcher, P. D. Stationary phase gene expression of *Mycobacterium tuberculosis*
12 following a progressive nutrient depletion: a model for persistent organisms. *Tuberculosis* **2004**,
13 *84*, 228-238.
14
15
16
17
18
19 (21) Rengarajan, J.; Bloom B.R.; Rubin, E.J. Genome-wide requirements for *Mycobacterium*
20 *tuberculosis* adaptation and survival in macrophages. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*,
21 8327–8332.
22
23
24
25
26 (22) Sasseti, C. M.; Rubin E.J. Genetic requirements for mycobacterial survival during
27 infection. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12989–12994.
28
29
30 (23) Ågren, D.; Schnell, R.; Schneider, G. The C-terminal of CysM from *Mycobacterium*
31 *tuberculosis* protects the aminoacrylate intermediate and is involved in sulfur donor selectivity.
32 *FEBS Lett.* **2009**, *583*, 330-336.
33
34
35
36 (24) Campanini, B.; Speroni, F.; Salsi, E.; Cook, P.F.; Roderick, S.L.; Huang, B.; Bettati, S.;
37 Mozzarelli, A. Interaction of serine acetyltransferase with O-acetylserine sulfhydrylase active
38 site: evidence from fluorescence spectroscopy. *Protein Sci.* **2005**, *14*, 2115-2124
39
40
41
42 (25) Poyraz, O.; Jeankumar, V.U.; Saxena, S.; Schnell, R.; Haraldsson, M.; Yogeeswari, P.;
43 Sriram, D.; Schneider, G. Structure-guided design of novel thiazolidine inhibitors of O-acetyl
44 serine sulfhydrylase from *Mycobacterium tuberculosis*. *J. Med. Chem.* **2013**, *56*, 6457-6466.
45
46
47
48
49
50 (26) Pieroni, M.; Annunziato, G.; Beato, C.; Wouters, R.; Benoni, R.; Campanini, B.; Pertinhez,
51 T.; Bettati, S.; Mozzarelli, A.; Costantino, G. Rational Design, Synthesis and Preliminary
52
53
54
55
56
57
58
59
60

1
2
3 Structure-activity relationships of α -substituted-2-phenylcyclopropane carboxylic acids as
4 inhibitors of *Salmonella typhimurium* O-acetylserine sulfhydrylase. *J. Med. Chem.* **2016**, DOI:
5
6 10.1021/acs/jmedchem.5bo1775
7
8

9
10 (27) Lipinski, C.A. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discovery*
11
12 *Today* **2004**, *4*, 337-341
13

14
15 (28) Baell, J.B.; Holloway, G.A. New substructure filters for removal of pan assay interference
16
17 compounds (PAINS) from screening libraries and for their exclusion in Bioassays. *J. Med.*
18
19 *Chem.* **2010**, *53*, 2719–2740
20

21
22 (29) Walters, W.P.; Ajay; Murcko, M.A Recognizing molecules with drug-like properties. *Curr.*
23
24 *Opin. Chem. Biol.* **1999**, *3*: 384–387
25

26
27 (30) Irwin, J.J.; Duan, D.; Torosyan, H.; Doak, A.K.; Ziebart, K.T.; Sterling, T.; Tumanian, G.;
28
29 Shoichet, B.K. An aggregation advisor for ligand discovery. *J. Med. Chem.* **2015**, *58*, 7076-7087.
30

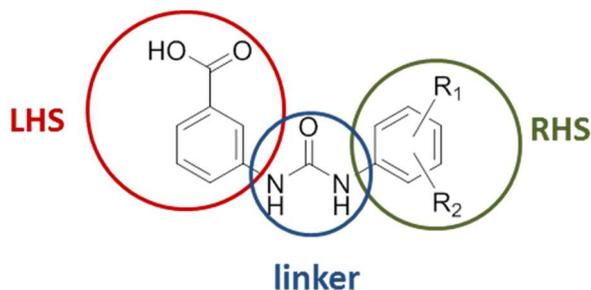
31
32 (31) Salsi, E.; Bayden, A.S.; Spyrakis, F.; Amadasi, A.; Campanini, B.; Bettati, S.; Dodatko, T.;
33
34 Cozzini, P.; Kellogg, G.E.; Cook, P.F.; Roderick, S.L.; Mozzarelli, A. Design of O-acetylserine
35
36 sulfhydrylase inhibitors by mimicking nature. *J. Med. Chem.* **2010**, *53*, 345-356
37

38
39 (32) Amori, L.; Katkevica, S.; Bruno, A.; Campanini, B.; Felici, P.; Mozzarelli, A.; Costantino,
40
41 G. Design and synthesis of trans-2-substituted-cyclopropane-1-carboxylic acids as the first non-
42
43 natural small molecule inhibitors of O-acetylserine sulfhydrylase. *MedChemComm* **2012**, *3*,
44
45 1111-1116
46

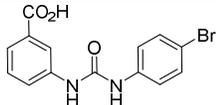
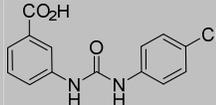
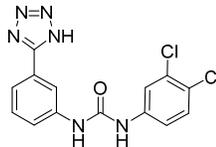
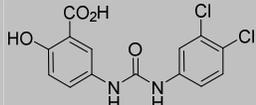
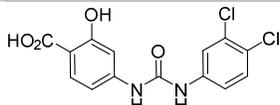
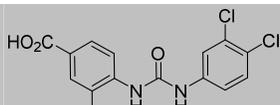
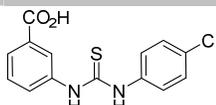
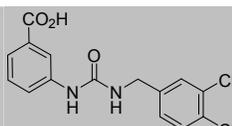
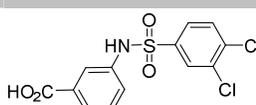
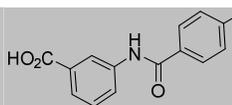
47
48 (33) Ananthan, S.; Faaleolea, E.R.M; Goldman, R.C.; Hobrath, J.V.; Kwong, C.D.; Laughon,
49
50 B.E.; Maddry, J.A.; Mehta, A.; Rasmussen, L.; Reynolds, R.C.; Secrist, J.A. 3rd; Shindo, N.;
51
52 Showe, D.N.; Sosa, M.I.; Suling, W.J.; White, E.L. High-throughput screening for inhibitors of
53
54 *Mycobacterium tuberculosis* H37Rv. *Tuberculosis* **2009**, *89*, 334-353.
55
56
57
58
59
60

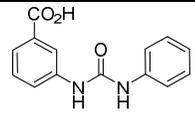
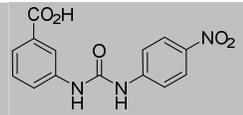
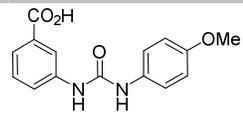
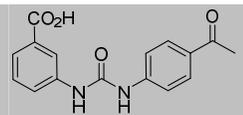
- 1
2
3 (34) Bento, A. P.; Gaulton, A.; Hersey A.; Bellis, L.J.; Chambers, J.; Davies, M.; Krüger, F.A.;
4
5 Light, Y.; Mak, L.; McGlinchey, S.; Nowotka, M.; Papadatos, G.; Santos, R.; Overington, J.P.
6
7 The ChEMBL bioactivity database: an update. *Nucleic Acids Res.* **2014**, *42*, D1083–D1090
8
9
10 (35) Salituro, F.G.; Bemis, G.W.; Green, J.; Kofron, J.L. Inhibitors of p38. US Patent
11
12 US6093742 A1, 2000.
13
14 (36) *Canvas*, version 2.6, Schrödinger, LLC New York, NY, 2015.
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1. Activities of selected hit compounds and synthesized derivatives. K_d values were determined using the fluorescence based binding assay, unless otherwise indicated. The general structure of the inhibitors with the LHS-linker-RHS fragment structure indicated is shown in the header.



Compound	Structure	K_d (μM)	k_{obs} (min^{-1})	% activity ^a
1		$0.32^b \pm$	0.10 ± 0.01	
		0.01	0.3	
2		$1.7 \pm$	0.24 ± 0.06	
		0.1	0.6	
3		$4.5 \pm$	0.48 ± 0.18	
		0.2	1.2	
4		$2.2^b \pm$	0.23 ± 0.05	
		0.1	0.6	
5		$8.0 \pm$	1.92 ± 1.08	
		0.6	4.9	

6		$1.4^b \pm$ 0.4	0.48 ± 0.12 1.2
7		$3.4^b \pm$ 0.1	0.24 ± 0.06 0.6
8		no signal ^d	$> 20^c$
9		no signal ^d	$> 20^c$
10		no signal ^d	$> 20^c$
11		no signal ^d	$> 20^c$
12		> 300	$> 20^c$
13		> 300	$> 20^c$
14		no signal ^d	$> 20^c$
15		no signal ^d	$> 20^c$

16		12.9 ±	1.02 ±0.42
		1.4	2.6
17		10.5 ±	2.22 ±0.54
		1.3	5.7
18		11.3 ±	3.84 ±1.32
		1.5	9.9
19		110.7 ±	> 20 ^c
		7.1	

a: k_{obs} is the apparent first order rate constant for the formation of the α -aminoacrylate intermediate. % activity is the remaining enzymatic activity when compared to the rate constant for the native, non-inhibited enzyme ($k_{\text{obs}} = 38.92 \text{ min}^{-1}$).

b: determined using ITC

c: due to limitations of the assay only remaining activities lower than 20% can be reliably determined. The term > 20% remaining activity indicates weak inhibition or no inhibition.

d: addition of the compound to CysM does not lead to a change in the PLP fluorescence, i.e. suggesting that the compound does not bind.

Table 2: MIC values against *M. tuberculosis* for urea-based inhibitors of CysM

Compound	MIC (μM)
1	9.6 ± 0.0
2	18.8 ± 0.0
3	2.2 ± 0.7
5	23.1 ± 3.8
6	18.6 ± 3.1
7	21.5 ± 0.0
17	20.7 ± 6.9
18	21.8 ± 7.3
Isoniazid	0.4 ± 0.0
Rifampicin	0.5 ± 0.0

Scheme and Figure legends

Scheme 1: Synthetic protocols to (A) *N,N*-diphenylureas (type I inhibitors) and *N,N*-diphenylthioureas and (B) type II inhibitors.

Figure 1: Cysteine biosynthetic pathways in *M. tuberculosis*. The classical pathway present in most bacteria, plants and fungi is based on the *bona fide* OASS CysK1, which utilizes O-acetylserine and sulfide ions. The Actinobacteria group (including Mycobacteria) harbors the genes for an alternative pathway based on CysM that utilizes OPS and thiocarboxylated CysO-SH as sulfur donor. The primary product of the reaction is a CysO-cysteine adduct that is hydrolyzed by the enzyme *mec*⁺ to release the L-cysteine. CysK2 also uses OPS as substrate and is able to synthesize L-cysteine with sulfide ions, however shows clear preference for thiosulfate leading to S-sulfocysteine. This compound may be converted to L-cysteine, but it has also been implicated in redox stress related signaling.

Figure 2: Illustration of the small-molecule screen of CysM. A. The biophysical basis of the screening assay is the increase in the fluorescence signal of the cofactor PLP upon binding of a ligand, here compound **3**, to the active site of the enzyme. Black line; compound **3** only, red: CysM only, blue: CysM in the presence of **3**. B: Dependence of the fluorescence signal on ligand (here compound **3**) concentration. C: Summary plot of the screening data after conversion to “% fluorescence increase” based on the average negative (0%) and positive (100%) controls on each individual plate. Magenta squares represent the negative controls (16 each on 50 plates),

1
2
3 turquoise triangles represent the positive controls (ditto) and the blue spheres represent the results
4
5 for the screening compounds.
6
7
8
9

10 **Figure 3:** *In vitro* characterization of hit compounds, exemplified for compound **1**. A.
11 Determination of K_d using the PLP fluorescence emission assay. The structure of **1** is inserted in
12 the graph. B. Inhibition of the first half reaction, monitored by following the formation of the α -
13 aminoacrylate intermediate in CysM spectrophotometrically at 463 nm. CysM in the presence
14 (red line) and absence (black line) of compound **1**. The insert shows the same reactions on a
15 longer time-scale. C: Inhibition of the overall CysM reaction by **1** monitored by detection of the
16 reaction product, the CysO-cysteine adduct (9658.7 Da) using ESI-MS mass-spectrometry. D:
17 Determination of K_d using ITC. The top panel illustrates the raw data from the ITC experiment,
18 whereas the bottom is based on the intergrated heats from each peak as a function of the molar
19 ratio of compound to protein. The solid line represents the best fit using the 1:1 binding site
20 model of the Origin software. E: Determination of IC_{50} of compound **1** based on a multiple
21 turnover enzyme assay using OPS and CysO-SH as substrates.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

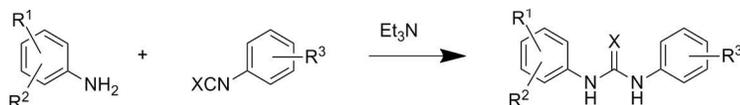
41 **Figure 4:** Structures of complexes of CysM with representative type I and type II inhibitors. A:
42 Enzyme-ligand interactions in the complex of CysM with compound **1**. B: Enzyme-ligand
43 interactions in the CysM-**2** complex. Color codes are red for oxygen, blue for nitrogen and
44 yellow for carbon atoms. For the bound inhibitors, carbon atoms are orange (**1**) and blue (**2**),
45 respectively, and chlorine atoms are green. Dotted lines indicate hydrogen bonds. C:
46 Superposition of the CysM - **5** and CysM - **2** complexes illustrating the different conformations
47 of the active site loop (residues 211-227) and the C-terminus in response to differences in
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 substitution of the RHS part of the inhibitor. CysM in complex with **2** is colored in light blue,
4
5 with active site loop and C-terminus in dark blue. CysM in complex with **5** is colored in yellow
6
7 and the corresponding active site loop and C-terminus in dark red. D: Differently sized binding
8
9 pockets for RHS in CysM-ligand complexes. Pocket A interacting with the RHS of type I
10
11 inhibitors. E: Extension of pocket A with pocket B (blue) to accommodate the large phenyl
12
13 substituent in type II inhibitors. Note the different position of Leu219 in the two structures,
14
15 acting as a gatekeeper to open up pocket B.
16
17
18
19
20
21

22 **Figure 5:** A. Biological activities of the lead compounds against *M. tuberculosis* in the nutrient
23
24 starvation model. The starved *M. tuberculosis* culture was treated with the compounds or anti-
25
26 TB drugs such as isoniazid (INH), rifampicin (RIF) or moxifloxacin (MOXI) at a concentration
27
28 of 10 μ M. Bacterial count estimation (given as means \pm standard deviations from three
29
30 independent experiments) for control and treated groups was carried out using the MPN (most
31
32 probable number) assay. Compounds **1** and **3** in particular gave significant inhibition of growth
33
34 of *M. tuberculosis* in this model as compared to the control ($p < 0.0001$, two tailed t-test using
35
36 GraphPad Prism Software). B. Effect of cysteine on bactericidal activities of compounds **1** and **3**.
37
38 Nutrient-starved cultures of *M. tuberculosis* were obtained as described above and split into three
39
40 groups, control (no compound added (left)), compound **1** (addition of 10 μ g/ml **1**) and compound
41
42 **3** (addition of 10 μ g/ml **3**). For each group, increasing cysteine concentration is indicated from
43
44 left to right: 0.0 mM, 0.05 mM, 0.2 mM, 0.5 mM. Bacterial count estimations were obtained as
45
46 described above. Experiments were carried out in triplicate.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

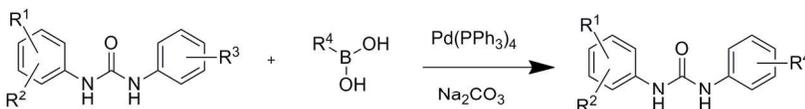
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

A. SCHEME FOR SYNTHESIS OF UREA AND THIOUREA DERIVATIVES (TYPE I INHIBITORS)



$R^1 = \text{COOH}$; $R^2 = \text{OH}$; $R^3 = 4\text{-NO}_2$; 4-OCH_3 ; 4-CH_3 ; 4-Acetyl ; 4-Br ; 4-Cl ; 3-Br ; Benzyl ; $3,4\text{-Dichloro}$; $X = \text{O}, \text{S}$.

B. SCHEME FOR SYNTHESIS OF SUZUKI DERIVATIVES (TYPE II INHIBITORS)



$R^1 = \text{COOH}$; $R^2 = \text{OH}$; $R^3 = 3\text{-Br}$ or 4-Br ; $R^4 = \text{Phenyl}$; 4-fluoro phenyl ; $2,3\text{-dichloro phenyl}$; I ; $2,5\text{-dimethyl}$; 3-benzoic acid ; 2-pyridyl ; 2-fluoro pyridine ; pyrimidyl .

Scheme 1

188x125mm (300 x 300 DPI)

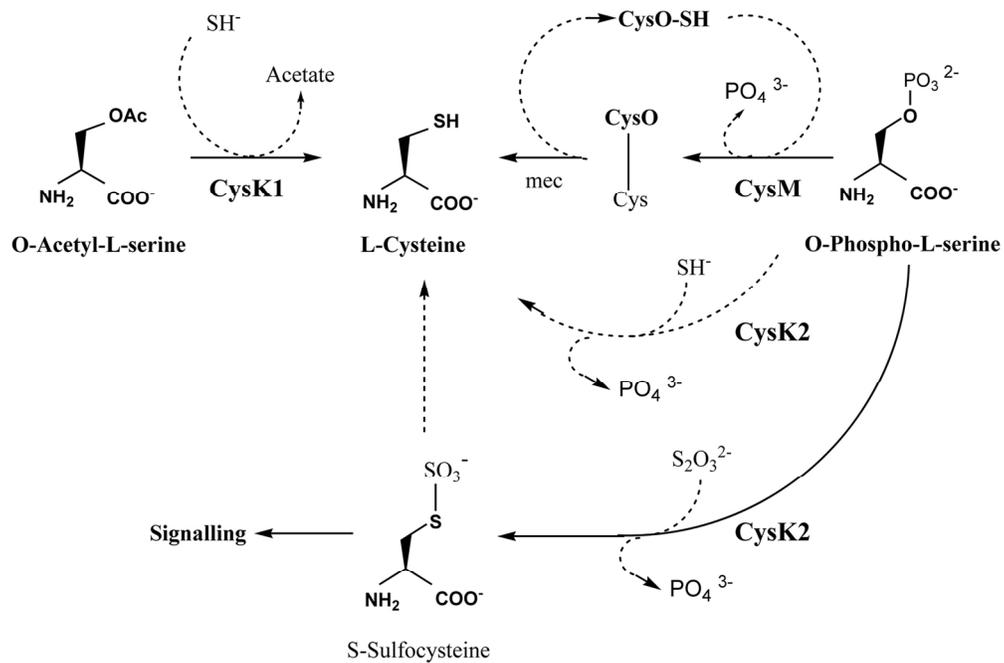


Figure 1

79x52mm (600 x 600 DPI)

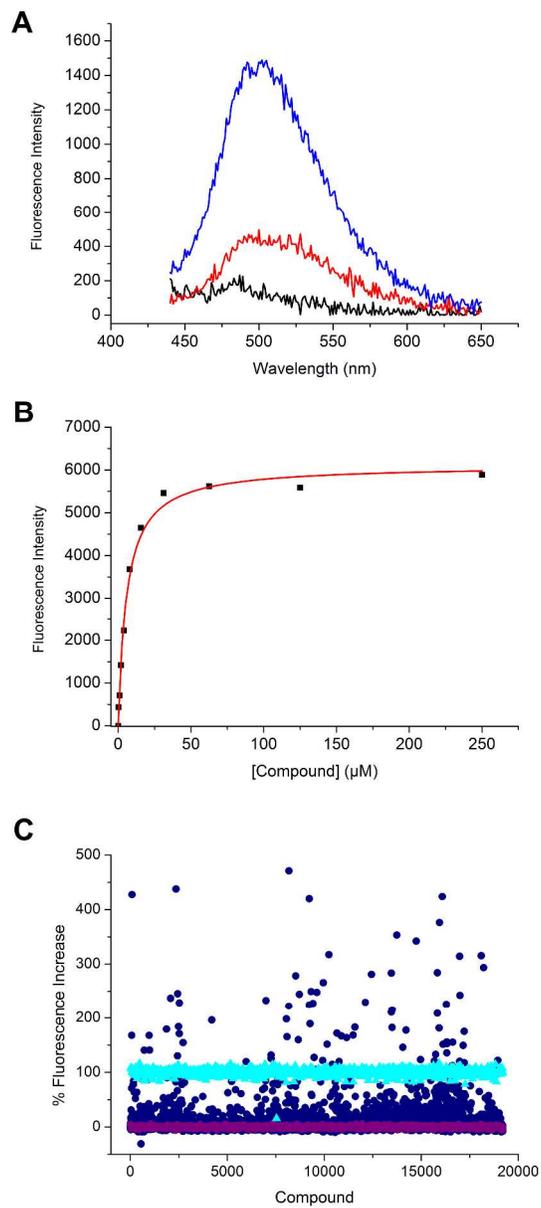


Figure 2

191x434mm (600 x 600 DPI)

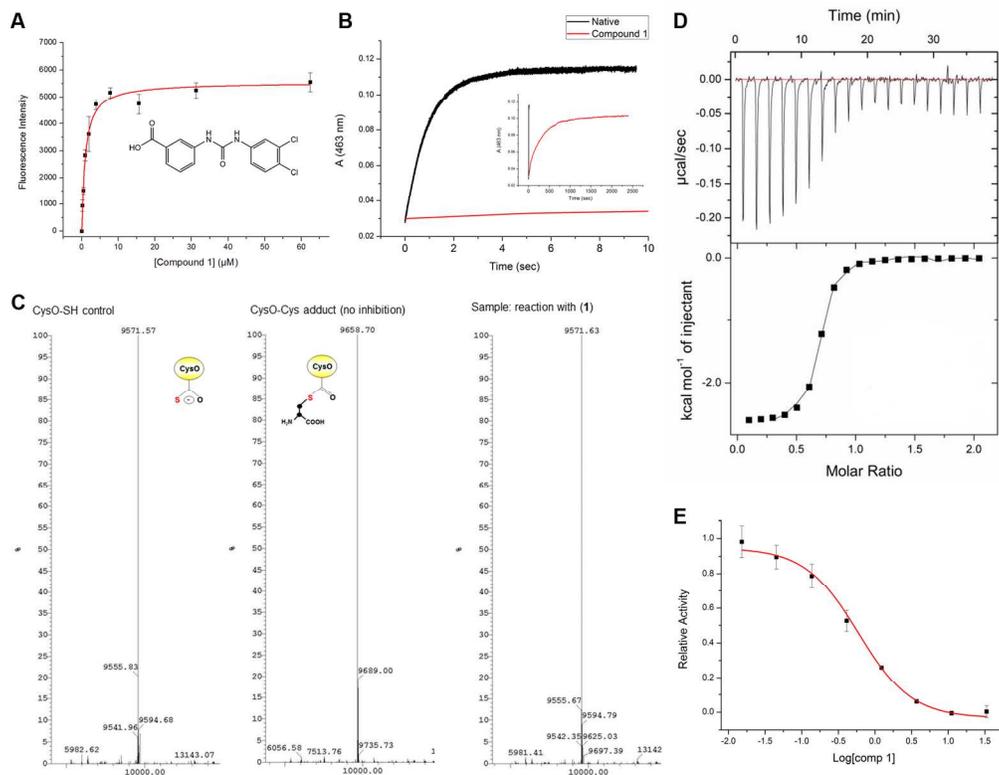


Figure 3

177x138mm (300 x 300 DPI)

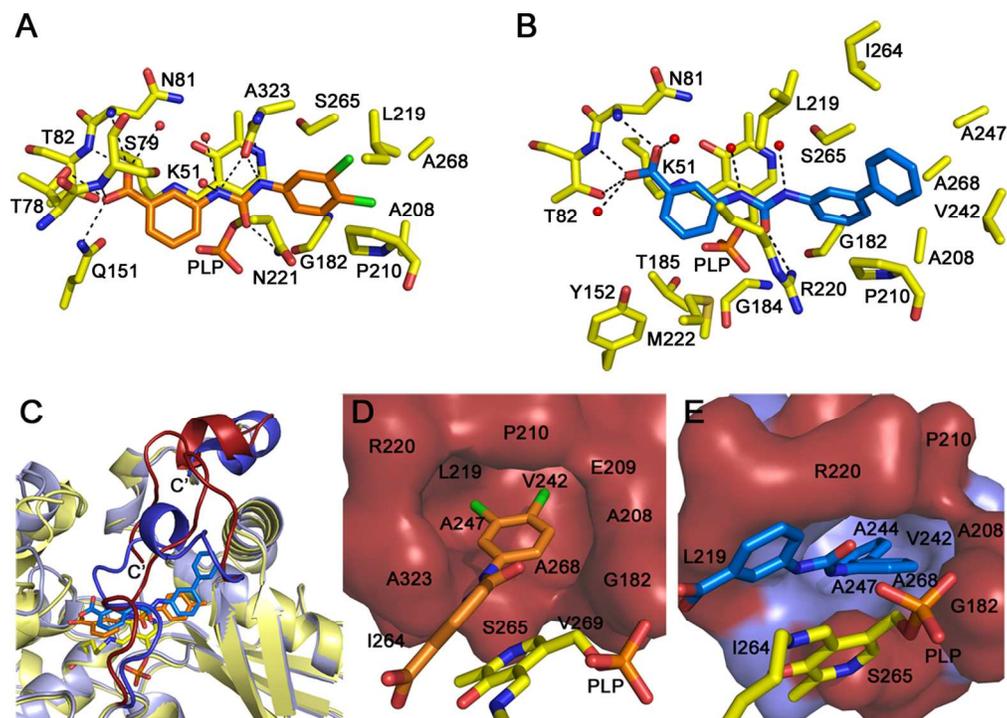


Figure 4

140x99mm (300 x 300 DPI)

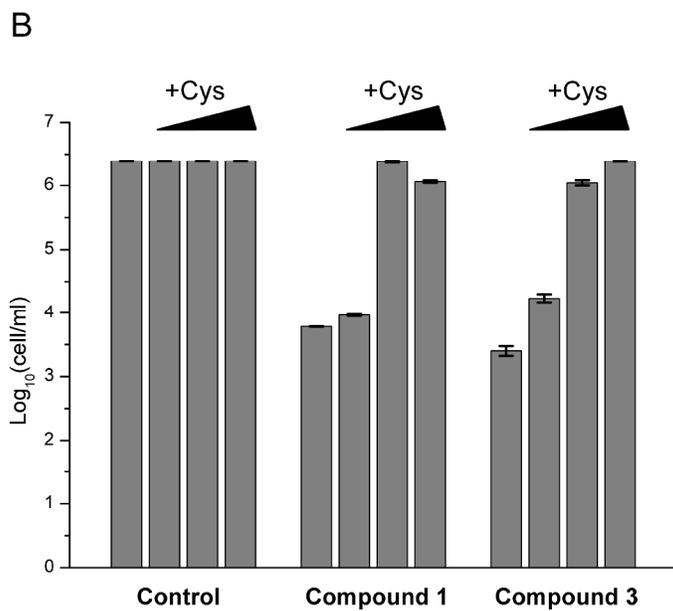
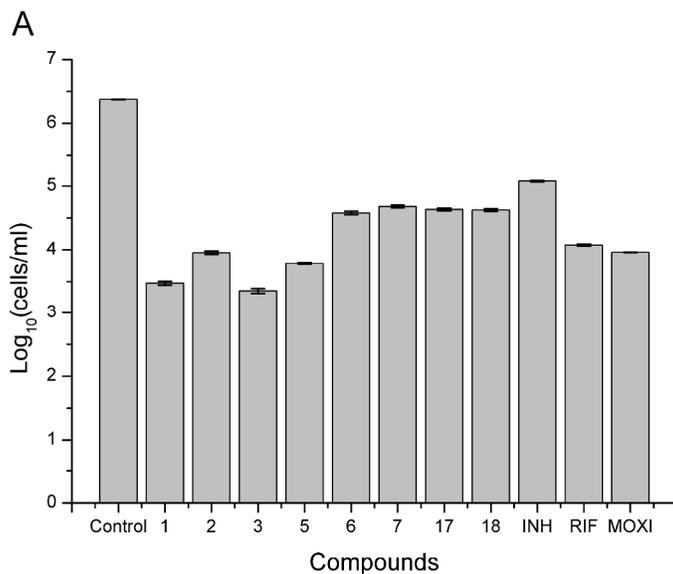
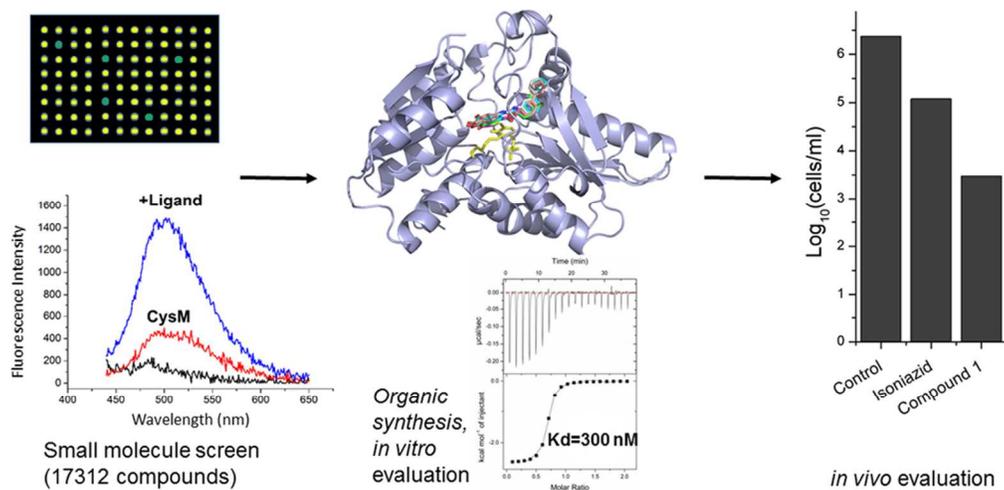


Figure 5

151x271mm (600 x 600 DPI)

60



Graphical abstract

112x55mm (300 x 300 DPI)