



HAL
open science

A comprehensive analysis of the protein-ligand interactions in crystal structures of *Mycobacterium tuberculosis* EthR

Abdalkarim Tanina, Alexandre Wohlkönig, Sameh H. Soror, Marion Flipo, Baptiste Villemagne, Hugues Prevet, Benoit Déprez, Martin Moune, Hélène Perée, Franck Meyer, et al.

► To cite this version:

Abdalkarim Tanina, Alexandre Wohlkönig, Sameh H. Soror, Marion Flipo, Baptiste Villemagne, et al.. A comprehensive analysis of the protein-ligand interactions in crystal structures of *Mycobacterium tuberculosis* EthR. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 2018, 1867 (3), pp.248-258. 10.1016/j.bbapap.2018.12.003 . hal-03177329

HAL Id: hal-03177329

<https://hal.science/hal-03177329>

Submitted on 23 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 **A comprehensive analysis of the protein-ligand interactions in**
2 **crystal structures of *Mycobacterium tuberculosis* EthR**

3
4 Abdalkarim Tanina^{a,1}, Alexandre Wohlkönig^{b,1}, Sameh H. Soror^c, Marion Flipo^d, Baptiste
5 Villemagne^d, Hugues Prevet^d, Benoit Déprez^d, Martin Moune^e, Hélène Perée^a, Franck Meyer^a,
6 Alain R. Baulard^e, Nicolas Willand^{d*} & René Wintjens^{a*}

7
8 ^a Unité Microbiologie, Bioorganique et Macromoléculaire (CP206/04), département R3D,
9 Faculté de Pharmacie, Université Libre de Bruxelles, B-1050 Brussels, Belgium.

10 ^b Center for Structural Biology, Vlaams Instituut voor Biotechnology (VIB), B-1050 Brussels,
11 Belgium.

12 ^c Center of scientific excellence, Helwan Structural Biology Research, Faculty of Pharmacy,
13 Helwan University, Ain helwan, 11792 Helwan, Cairo, Egypt.

14 ^d Univ. Lille, Inserm, Institut Pasteur de Lille, U1177 - Drugs and Molecules for living Systems,
15 F-59000 Lille, France.

16 ^e Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019-UMR 8204-CIIL-
17 Center for Infection and Immunity of Lille, F-59000 Lille, France.

18
19 ¹ The two authors contributed equally to this work

20 * Corresponding authors: Nicolas.Willand@univ-lille2.fr or Rene.Wintjens@ulb.ac.be

21

1 **Abstract**

2 The *Mycobacterium tuberculosis* EthR is a member of the TetR family of repressors, controlling
3 the expression of EthA, a mono-oxygenase responsible for the bioactivation of the prodrug
4 ethionamide. This protein was established as a promising therapeutic target against tuberculosis,
5 allowing, when inhibited by a drug-like molecule, to boost the action of ethionamide. Dozens of
6 EthR crystal structures have been solved in complex with ligands. Herein, we disclose EthR
7 structures in complex with 18 different small molecules and then performed in-depth analysis on
8 the complete set of EthR structures that provides insights on EthR-ligand interactions. The 81
9 molecules solved in complex with EthR show a large diversity of chemical structures that were
10 split up into several chemical clusters. Two of the most striking common points of EthR-ligand
11 interactions are the quasi-omnipresence of a hydrogen bond bridging compounds with Asn¹⁷⁹ and
12 the high occurrence of π - π interactions involving Phe¹¹⁰. A systematic analysis of the protein-
13 ligand contacts identified eight hot spot residues that defined the basic structural features
14 governing the binding mode of small molecules to EthR. Implications for the design of new
15 potent inhibitors are discussed.

16

1 **Highlights**

- 2 • Eighteen complex structures of EthR have been added to PDB;
- 3 • All available structural information on EthR has been compiled, analyzed and rationalized;
- 4 • Common ligand binding hot spots of EthR are defined;
- 5 • Discrepancies between ligand model and electron density are observed for some structures.

6

7

8 **Keywords:** tuberculosis; ethionamide; x-ray crystallography; binding interaction; small
9 molecule; drug design.

1 **1. Introduction**

2 Tuberculosis (TB) still represents one of the most important concern for global public
3 health. The situation is now more worrying with the growing emergence of multidrug-resistant
4 TB (MDR-TB); globally in 2016 MDR strains of *Mycobacterium tuberculosis* (*Mtb*) caused ca.
5 5% of the new TB cases (490,000 cases) [1], whereas in 2000 this figure was only 3.2% [2]. Our
6 inability to control the outbreak is largely due to the use of a drug regimen that generally requires
7 patients to undergo 6-9 months of treatment with a combination of several first-line drugs [3].
8 Although the recent approval of bedaquiline and delamanid, two drugs preconized in cases of
9 MDR-TB, the lack of new antitubercular therapies over the last 50 years is also a cause of
10 increasing concern about the disease.

11 The discovery and the development of a new class of anti-TB drugs is a long, arduous
12 and expensive process, thereby greatly circumscribing the rapid introduction of new anti-TB
13 therapy. Hence, for the last 5 decades only a few molecules have successfully entered the TB
14 clinical pipeline, and of course much less have reached the market [4]. As an alternative to the
15 search for new drugs, improving the efficacy of existing drugs may thus represent an effective
16 and pragmatic choice.

17 Ethionamide (ETH), a second-line anti-TB drug essentially prescribed for MDR
18 infections, is a pro-drug requiring *in vivo* bio-activation to exert its inhibitory action on the 2-
19 *trans*-enoyl reductase InhA, an essential enzyme involved in the mycolic acid synthesis pathway
20 [5]. Early experimental studies point to the Baeyer-Villiger monooxygenase EthA as the
21 mycobacterial enzyme mainly responsible for ETH bio-activation [6-8]. Recently, another
22 Baeyer-Villiger monooxygenase, namely MymA (*Rv3083*), was found to take part in the
23 activation of ETH [9]. Furthermore, the oxydoreductase protein Rv0077c was reported to play a

1 role in the sensitivity of *Mtb* to ETH [10], possibly in conjunction with MymA, EthA, or
2 additional factors [11, 12].

3 The transcriptional repressor EthR belongs to the tetracycline repressor family of
4 regulators (TFR) [13, 14], that was named according to the first-discovered member [15]. TFR is
5 one of the most widely distributed families of transcriptional regulators in bacteria. Notably,
6 whereas 15 TFRs have been identified in *Escherichia coli*, as many as 51 are annotated in the
7 *Mtb* genome [13]. EthR plays a key role in the control of the ETH activity as it represses the
8 expression of EthA, then reducing the effectiveness of ETH [16]. Conversely, abolishing EthR
9 DNA-binding function with small molecules was shown to improve ETH potency [17]. To date,
10 an important number of EthR inhibitors were described as boosters of the ETH activity in whole
11 cell *Mtb* assays, as well as in *in vivo* models of TB [18-23]. By allowing efficient therapy at
12 lower ETH concentration, this approach could drastically limit the side effects of the treatment
13 [17].

14 Hence, based on the ability of specific ligands to impede the interaction between EthR
15 and its DNA operators, drug-like small molecules have been designed and synthesized for a co-
16 administration with ETH. In this process, X-ray crystallography was instrumental in driving
17 fragment-based approaches and structure-guided drug design [20-27]. Many crystal structures of
18 *Mtb* EthR have been therefore solved. Since the three first crystal structures, determined in
19 complex with two fortuitous ligands, one with two putative 1,4-dioxane molecules (PDB id 1T56
20 [28]) and the two others with a molecule identified as hexadecyl octanoate (PDB ids 1U9N and
21 1U9O [29]), 62 other EthR structures have been solved afterwards, most of them in complex
22 with drug-like molecules [17-27, 30, 31].

1 No systematic analysis of all available structural information on EthR has yet been made,
2 although there is much to be learned for further drug design studies through such analysis. Here,
3 we compared and rationalized the observed interactions in crystal complex structures between
4 ligand molecules and EthR, with the aim of defining conserved features and rules that may be
5 useful for further drug design studies. Furthermore, in addition to the EthR structures available,
6 the present study integrates 18 unpublished ligand-EthR co-crystals structures that have driven
7 our drug-design screening campaigns. Finally, to assess the reliability of EthR structural data, an
8 analysis of the goodness for the fit of the ligand model to the observed electron density was also
9 carried out using different metrics.

10

1 **2. Material and Methods**

2 **2.1. EthR structural database**

3 Fifty-six EthR crystal structures with the corresponding structure factors were retrieved
4 from Protein Data Bank [32], whereas nine other structures in complex with ligands were
5 directly obtained from authors [25]. Eighteen unpublished EthR structures here solved in
6 complex with ligands were added to the database. In total, the EthR structural database includes
7 83 crystal structures, which were denoted according the PDB id when available otherwise
8 according the ligand name found in the associated publication. The structures are listed in the
9 supplementary Table S1. Throughout our analysis, in case of amino acid residue or ligand
10 molecule with alternate conformations, only the conformation A was considered. Similarly, in
11 case of multiple protein molecules in the asymmetric unit cell, only chain A and its bound ligand
12 was taken into account, except for the PDB id 5MYM where only chain B contains a ligand
13 molecule. Water molecules present in crystal structures were not considered in the analysis.
14 Finally, ligand in the PDB id 1U9O was not analyzed due to the lack of atomic coordinates.

15 **2.2. Ligand structure analysis**

16 Physico-chemical properties of small ligands bound to EthR were computationally
17 estimated by openBABEL2.4.1 [33]. The volume and the sphericity were calculated with the
18 “3V” web tool [34]. To compare ligand molecules, their 2D structures were transformed to
19 1,024-bit molecular fingerprints by applying the Morgan algorithm with a radius of two bonds in
20 the open-source Python cheminformatics RDKit (<http://www.rdkit.org>). The obtained
21 fingerprints were pair-wisely compared using the Jaccard/Tanimoto coefficients (Tc) as
22 similarity metrics, i.e. for two fingerprints F_A and F_B , $Tc = (F_A \cap F_B) / (F_A \cup F_B)$, where symbols
23 \cap and \cup are the algebraic intersection and union of non-zero bits, respectively. The so-produced

1 fingerprints approximate the extended-connectivity circular fingerprints with a radius of four
2 bonds (ECFP4 [35]) that were estimated among the best performing fingerprints [33]. Note also
3 that the Tc was found as an excellent metrics when ranking diverse small molecule structures
4 [36]. The function hclust of the statistical package R (version 3.5.0) was further used to generate
5 a hierarchical agglomerative clustering using the Ward's minimum variance method and using a
6 measure of dissimilarity obtained by subtracting the Tc score from 1. The resulting tree was
7 displayed with FigTree program [37] and with the heatmap.2 function of the R package.

8 **2.3. Crystal structure determination of EthR in complex with ligand**

9 EthR protein containing a N-terminal 6-His tag was produced in *E. coli* via a pET-15b
10 plasmid and purified as previously described [16]. EthR crystals were produced by the vapor
11 diffusion method using 1.4-1.65 M ammonium sulfate (using 0.05 M increment), 15% (v/v)
12 glycerol and 100 mM MES pH 6.7, as crystallization buffer [29]. The crystal complexes obtained
13 with the 18 ligands were prepared by mixing 1 μ l of ligand (33 mM in 100% DMSO) with 9 μ l
14 of purified protein (9 mg/ml) and equilibrated for 30 minutes at room temperature prior
15 submitting to crystallization trials.

16 X-ray diffraction data were collected on a Pilatus 6M detector using synchrotron
17 radiation at SLS-X06SA beamline (PSI, Switzerland). All crystals belonged to the tetragonal
18 space group P4₁2₁2 with Matthews coefficient of ~2.4 (solvent content ~48.9%), implying one
19 protein molecule in the asymmetric unit. Diffraction data indexation was performed with
20 iMOSFLM [38]. Data collection statistics are summarized in Table S2.

21 Structures were solved by molecular replacement using the pdb id 1U9N [29] as search
22 template, and MOLREP [39] as program. Structure refinements were performed by iteratively
23 cycling through REFMAC5 [40] and COOT [41]. Ligand geometry and stereochemical restraints

1 were generated using AceDRG [42]. The ligands were firstly automatically modeled into the
2 electron density using the FINDLIGAND tool of the CCP4 software suite [43] and then
3 manually adjusted during refinements steps. Final refinement statistics are given in Table S3.
4 OMIT density maps were generated by removing ligand from the model, then randomly shaking
5 all coordinates by a small amount (between 0.0 and 0.3 Å) using the PDBSET program of CCP4
6 package and finally running 20 cycles of refinement with REFMAC5. Structure qualities were
7 assessed with MolProbity [44].

8 **2.4. Structure analysis and binding interaction**

9 The ligand binding cavity of EthR was defined using CASTp with a probe radius of 1.4 Å
10 [45]. Dimeric structures were generated *via* the PISA server [46] and distances between Helix-
11 Turn-Helix (HTH) motifs were measured between C α atoms of the Tyr⁶² residue of each
12 monomer. Protein structure superimpositions were performed with DaliLite [47]. Side chain
13 displacements between EthR structures were computed by superimposing C α backbone atoms
14 with DaliLite program and by calculating the root-mean-square deviation (*RMSD*) on side chain
15 heavy atoms by:

$$RMSD = \sqrt{\frac{1}{n} \sum_{i=1}^n d_i^2}$$

16 where the averaging was performed over the n pairs of equivalent side chains atoms and
17 d_i is the distance between the two considered atoms in the i -th pairs. For that purpose, and to
18 ensure that comparison was performed on equivalent atoms, the nomenclature of the side chain
19 heavy atoms was standardized over the EthR structure database. Side chain conformations were
20 clustered using the complete linkage clustering in R package. Only large-scale side chain
21 displacements, *RMSD* value greater than 1.5 Å, were considered for attributing changes in side

1 chain orientations. Ligand binding interactions were analyzed with BINANA program [48] using
2 structures in which hydrogen atoms were automatically built using UCSF Chimera [49] and
3 charges added using AutoDockTools [50]. Structural images were drawn with UCSF Chimera or
4 with MolScript [51] and then rendered with Raster3D [52].

5 **2.5. *Quality of fit between ligand model and observed electron density***

6 In order to evaluate whether electron density covers all ligand atoms, EDSTATS was
7 used to compute four density-based validation metrics: the real-space R factor (RSR [53]), the
8 real-space correlation coefficient (RSCC [54]), the real-space Z-observed score (RSZO [55]),
9 and the real-space Z-difference score (RSZD [55]). Note that the conventional RSR and RSCC
10 metrics depend on both accuracy and precision of the structural model, whereas the RSZO or
11 RSZD metrics are only correlated to precision or accuracy, respectively [55]. The cut-off values
12 for considering a fit as valid were defined as $RSR < 0.1$, $RSCC > 0.9$, $RSZO > 2.0$ and $RSZD <$
13 2.0 .

1 **3. Results and discussion**

2 **3.1. *Eighteen new EthR complex structures***

3 In recent years, we worked on the drug design of *Mtb* EthR inhibitors, for which X-ray
4 crystallography was used as a high-throughput tool to determine at atomic level the binding
5 mode of selected ligands, and thereby to guide further chemical modifications for optimizations
6 in binding potency. Crystal structure of 18 compounds from our previous drug-design screening
7 campaigns were obtained in complex with EthR (Tables S2 and S3). The resolution of these
8 crystal structures, ranging from 2.5 to 1.5 Å, is high enough to identify the ligand orientation and
9 the binding interactions (Figure 1).

10 **3.2. *EthR fold and dimeric configuration***

11 Liganded EthR structures have been extensively described in many publications [17-22,
12 24-31]. EthR fold consists of nine α -helices, where the second and the third N-terminal α -helices
13 form the HTH DNA-binding motif, whereas the last six α -helices at C-terminal compose the
14 ligand-binding domain (LBD) (Figure S1). The vast majority of the structures (96%) were solved
15 in the tetragonal space group P4₁2₁2 with one protein molecule in the asymmetric unit cell. Only
16 three structures (PDB ids: 1U9O, 3Q0V and 5MYM) were solved in other space groups, in each
17 case with one or two dimers in the unit cell. EthR dimer serves as the minimal functional unit
18 [16]. In all crystal structures, the two HTH motifs of the EthR dimer, foreseen to bind to two
19 consecutive major grooves of the DNA operator, were systematically separated by a distance
20 greater than 34 Å, the typical distance between two consecutive major grooves of a B-form
21 DNA.

22 Over the 83 EthR structures, the distance between the two HTH motifs ranged from 40 to
23 48 Å, with a mean of 43.1 ± 1.5 Å (Table S1). This observation is not surprising, knowing that

1 most of the 83 structures contain at least one small molecule in their LBD. Indeed, as a common
2 mechanism in the TetR family of transcriptional regulators, the binding of a small molecule in
3 the cavity of LBD triggers extra spacing between the HTH motifs, responsible of their inability
4 to bind to the DNA operator [56]. Note that two structures (structures 3QPL and 3TP3) do not
5 contain any ligand in the cavity of LBD, but rather enclose mutation G106W that was found to
6 cause effects similar as ligands [57]. A wild-type EthR structure in unbound (apo) form has been
7 previously reported [25]. However, since the structural data are not available, we could not
8 analyze this structure in the present paper.

9 **3.3. *EthR ligand binding cavity: size, composition and hydrophobicity***

10 The LBD cavity of EthR forms a long L-shape tunnel with a diameter of ~ 6 Å and a
11 length of ~ 22 Å (Figure 2). Whereas the tunnel-like cavity arbors two entrances in the monomer,
12 one entrance is hidden by the dimeric assembly. Therefore, ligands are presumed to enter into the
13 cavity by the entrance located farthest from the HTH motif, on the top of the monomer (Figures
14 1 and S1). The mean cavity volume averaged on the 83 EthR structures is 871 ± 251 Å³ (Table
15 S1). This volume varied greatly according to the bound ligand, ranging from 476 Å³ (apo
16 structures 3QPL/3TP3) to 1518 Å³ (structure Cmpd1). Indeed, a moderate but significant
17 correlation was observed between the molecular volumes of ligands and cavities (Figure 3;
18 correlation coefficient of 0.58 with p-value of 1.0e-8), underlining the relationship between these
19 two parameters. This shows that the LBD of EthR has a high degree of conformational plasticity
20 by adjusting its cavity volume to accommodate ligands of different size.

21 Forty-three residues line the cavity at its biggest size (listed in the legend of Figure 2).
22 These residues are distributed so as the hydrophilic ones are predominantly at the two entrances
23 of the tunnel-like cavity while the hydrophobic residues are in the center. Noteworthy, the

1 hydrophobic residues are not distributed homogeneously: one side of the cavity is very
2 hydrophobic, lining with Leu⁷⁶, Leu⁸⁷, Leu⁹⁰, Ile¹⁰⁷, Phe¹¹⁰, Phe¹¹⁴, Leu¹³⁷, Leu¹⁸³, Phe¹⁸⁴ and
3 Phe¹⁸⁷, while another side, delimited by residues Gly¹⁰⁶, Thr¹²¹, Arg¹²⁸, Tyr¹⁴⁸, Thr¹⁴⁹, Asn¹⁷⁶,
4 Asn¹⁷⁹ and Glu¹⁸⁰, shows intermediate hydrophobic or hydrophilic features (according to Kyte-
5 Doolittle scale, i.e. values in the range -2 to 0 [58]) (Figure 2). The putative entrance of the
6 cavity located at the top of LBD is lined by Ala⁹¹, Pro⁹⁴, Ala⁹⁵, Asp⁹⁶, Thr⁹⁷, Gln¹²⁵, Val¹⁵²,
7 Arg¹⁵⁹ and Arg¹⁸¹.

8 **3.4. Chemical space of small molecules crystallized with EthR**

9 Small molecules co-crystallized with EthR do not absolutely represent all the possible
10 ligands of the protein. However, these are often recognized as promising fragments or even lead
11 compound for targeting EthR. Indeed, except 1,4-dioxane (structure 1T56) and hexadecyl
12 octanoate (structures 1U9N and 1U9O), all the other compounds were selected for their ability to
13 target EthR and their binding was measured by biophysical techniques such as thermal-shift
14 assay and/or surface plasmon resonance (see Table S4). Therefore, the analysis of these
15 molecules can provide insight on EthR ligands, especially considering the large number of
16 complex structures available.

17 The 81 ligands were pair-wisely compared using circular Morgan fingerprints and
18 Tanimoto coefficient (Tc) as metrics. The distribution of Tc scores over the 3,240 comparisons
19 indicated a large chemical diversity among EthR ligands (Figure 4a). A Tc score of 0.4 is
20 consistently used as a similarity threshold with circular fingerprints. Here, the observed mean Tc
21 score ($\langle Tc \rangle$) was 0.17 +/- 0.13 with only a few scores (~6%) higher than 0.4.

22 Despite of a low overall structural similarity, 36 chemical clusters were defined by a
23 hierarchical clustering when the tree was cut at a $\langle Tc \rangle$ higher than 0.48 (Figure 4b). Note that

1 the pattern of relationship network between structures is more important than the number of
2 clusters, as another choice of cut-off will give a different quantity of clusters but from the same
3 clustering tree. However, structure clustering clearly indicates a large chemical diversity within
4 EthR ligands. The heatmap of the clustering tree is provided in Supplemental data (Figure S2).

5 Eighteen clusters contained only one member, while the most populated ones were
6 clusters CC-17 and CC-23, with seven and eight members, respectively (Figures 4b and S4).
7 Five of new structures presented in this paper were distant from structures previously solved,
8 forming the two clusters CC-3 (BDM35133 and BDM44719) and CC-18 (BDM44814,
9 BDM44825 and BDM44847). The main feature of BDM35133 and BDM44719 (CC-3) is the
10 presence of a spirochromanone and benzamide moieties. Both structures are almost identical, the
11 N-phenyl group in BDM35133 being substituted by an N-benzyl group in BDM44719. As
12 concern the CC-18 cluster, a benzamide group is also present, here branched to a 3,3,3-
13 trifluoropropyl group. In addition, compounds BDM44825 and BDM44847 possess an azido or a
14 triazole group, respectively, at position 4 of the phenyl ring. Indeed, CC-17 and CC-18 are
15 structurally close, with the fragment BDM44814 as common scaffold. The rest of newly solved
16 structures were supplying clusters containing structures already published.

17 The EthR ligands showed a high degree of disparity in term of size and molecular weight
18 (Figure S4 and Table S5). Although primarily hydrophobic in nature with $\text{clogP} > 0$ (mean value
19 3.6 ± 1.3), most ligands contain several polar atoms giving a significant polar surface area (PSA)
20 of $\sim 87 \text{ \AA}^2$ on average. Furthermore, ligands comprise on average four hydrogen bond acceptors
21 and one donor (Table S5). Hence, the combination of hydrophobic and hydrophilic parts in
22 ligands was in agreement with the quasi-amphiphilic character of the binding cavity (see Figure
23 2).

1 **3.5. Side chain displacements in EthR binding cavity**

2 The displacements of side chains in the EthR cavity were examined by pair-wisely
3 computing *RMSD* over side chain atoms after structure backbone superimposition. Globally, side
4 chains of the binding cavity appeared to adopt same conformations in EthR complex structures.
5 Only five residues exhibited large-scale movements of *RMSD* > 1.5 Å, namely Met¹⁰², Trp¹³⁸,
6 Phe¹⁴¹, Trp¹⁴⁵ and Phe¹⁸⁴ (Figures 5 and S5).

7 The Phe¹⁸⁴ side chain adopts two distinct conformations, denoted “open” and “close”,
8 with an intermediate state found only in 4DW6 structure (Figure 5). Indeed, the flipping
9 movement of the Phe¹⁸⁴ side chain from “close” conformation leads to the extension of the cavity
10 and generally causes a second entry point, as seen in 1U9N (Figure S1d). This flipping
11 movement is likely due to the presence of a long ligand (structures 1U9N, 1U9O, 3O8H, 5F0C,
12 5F04, 5F27, 5EZG, 5EZH, Cmpd5, Cmpd6, Cmpd7, Cmpd8, Cmpd9, 5NIZ, 5MYM, 5MNO,
13 5IOZ, 5J1R, 5MXK, 3Q3S, BDM41848; with sphericity generally less than 0.8, or when a
14 second ligand molecule binds precisely to the second entry point site (structures Cmpd1, 5FJL,
15 5IOY, 5IPA) [25]. As mentioned previously, the second entry point is hidden by the dimeric
16 organization, with the both second entry points nearly face to face (Figure S1d).

17 The orientation of the pair of side chains Phe¹⁴¹/Trp¹⁴⁵ only differs in four structures, in
18 the mutated protein G106W (3QPL and 3TP3), and in structures 5F04 and 5F27 (Figure 5).
19 Moreover, the side chain conformations of Met¹⁰² and Trp¹³⁸ clustered in distinct groups (Figure
20 S5). However, whereas movements of the Phe¹⁸⁴ side chain can be figured out, those of the four
21 other side chains remain elusive.

22 **3.6. Binding interactions and EthR hot spot residues**

1 The binding interactions were analyzed for each ligand in the EthR dataset (Table S6). A
2 hydrogen bond with the side chain amide group of Asn¹⁷⁹ is found in 88% of structures (71/81),
3 while only 36% (29/81) present a ligand hydrogen-bonded with the Asn¹⁷⁶ side chain. As the
4 ligands have generally between one and two aromatic rings (see Table S4), one of the key
5 binding characteristics of EthR is the recurring presence of π - π interactions with one or several
6 of the aromatic rings of the cavity (Trp¹⁰³, Phe¹¹⁰, Phe¹¹⁴, Trp¹³⁸, Tyr¹⁴⁸ and Trp²⁰⁷). Eighteen
7 residues make on average more than one atomic contact with the ligand (listed in Table S6).
8 Phe¹¹⁰ is by far the most implicated residue in binding, with on average 16.6 contacts with
9 ligand, then Asn¹⁷⁹, Trp¹⁰³ and Trp¹³⁸, with on average 6.4, 4.6 and 4.1 contacts, respectively.

10 Using as criteria at least 3 contacts on average, eight residues were defined as basic hot
11 spots of EthR binding, namely Trp¹⁰³, Gly¹⁰⁶, Phe¹¹⁰, Trp¹³⁸, Tyr¹⁴⁸, Thr¹⁴⁹, Asn¹⁷⁶ and Asn¹⁷⁹. In
12 clear, since these hot spots correspond to side chains commonly found in interaction with small
13 molecules or fragments, they define the core binding region of EthR. The geometry and the
14 distances between hot spots reveal a cavity of near-rectangular shape, approximately 8 Å x 8 Å x
15 25 Å, with aromatic side chains at both extremities, Trp¹³⁸ on one side and the other side, Trp¹⁰³
16 and Tyr¹⁴⁸ (Figure 6). The trio formed by the side chains Phe¹¹⁰, Asn¹⁷⁶ and Asn¹⁷⁹, each
17 separated with a distance of 7-8 Å, is located near the center of the binding cavity.

18 3.7. *Ligand electron density fit validation*

19 Finally, we assessed the reliability of structural data here analyzed. Whereas metrics of
20 global model quality, such as R-factors, were broadly good, with a mean R_{free} of 0.226 over the
21 83 EthR crystal structures, local measures for electron density quality were poor with some
22 modeled ligands. Four density-based validation metrics, namely RSR, RSCC, RSZO and RSZD,
23 were calculated to assess the fit to the electron density for the set of ligand atoms (Table S7).

1 Based on pre-defined thresholds (see Material and Methods), several ligands matched poorly to
2 the density. In a correlated way, these ligands are also associated with higher B-factors, with
3 mean values generally greater than 70 \AA^2 . Indeed, high B-factors, high RSR/RSZO, or low
4 RSCC/RSZD may indicate that ligand has been modeled incorrectly, but also that there were a
5 large amount of motion or positional averaging within the crystal. Occupancy error could also be
6 responsible for the weak correlation between ligand model and observed electron density. Except
7 four structures in which ligands were modeled by two alternative conformations, each at half
8 occupancy (see Table S7), all EthR ligands were assumed *a priori* to bind at full occupancy in
9 the crystal. However, these ligands, as all non-covalent ligands, are always subject to a binding
10 equilibrium, so that even high-affinity ligands may not fill the binding site at 100% in the crystal.
11 An occupancy refinement of ligands in the crystal structure could be an option in the future [59].

12 From our analysis, 11 ligands in 9 structures present problematic agreements between
13 model and electron density, identified as those having three local metrics outside thresholds (all
14 illustrated in Figure S6). Fortunately, as no ligand has the four metrics non-valid, we worked
15 with the assumption that the EthR structural database is globally reliable. However, it can be
16 stressed that a special attention should be drawn to a perfect fit between modeled ligand and
17 electron density when solving a EthR complex crystal structure, or, to a general extent, when
18 solving any crystallographic complex structure.

1 **Conclusions**

2 Eighteen complex structures were solved, thereby increasing the number of available
3 EthR structures to 83. Ligands or chemical fragments co-crystallized with EthR, 81 in total,
4 presented a large set of diverse structures that could be grouped in several distinct clusters by
5 comparing chemical structures. Even if ligands differ in their chemical structures, they share
6 common features. Most ligands have hydrophobic and hydrophilic functions that seem to match
7 well with the quasi-amphiphilic character of the EthR binding cavity. Furthermore, the different
8 ligands showed binding modes with common interactions in term of hydrogen bond and π - π
9 interactions, defining hot spot side chains for the EthR binding cavity. The correlation between
10 observed molecular volumes of cavity and ligand underlined the plasticity of the cavity that can
11 accommodate small molecules of different size, and with diverse chemical properties.

12
13
14

1 **Protein Data Bank accession codes**

2 The experimental data and the coordinates of the refined crystal structures were deposited
3 to the RCSB Protein Data Bank under PDB ids: 6HNX (BDM35133), 6HNZ (BDM41231),
4 6HN0 (BDM41325), 6HN1 (BDM41974), 6HN2 (BDM43138), 6HN3 (BDM43265), 6HN4
5 (BDM44693), 6HN5 (BDM44719), 6HN6 (BDM44725), 6HN7 (BDM44814), 6HN8
6 (BDM44815), 6HN9 (BDM44825), 6HNA (BDM44830), 6HNB (BDM44831), 6HNC
7 (BDM44847), 6HND (BDM44848), 6HNE (BDM44850) and 6HNF (BDM44852).

8

9 **Competing interests**

10 The authors declare no conflict of interest associated with this manuscript.

11

12 **Acknowledgements**

13 We are indebted to the Swiss Light Source for diffraction data collection (Paul Scherrer
14 Institute, Villigen, Switzerland). The authors warmly thank Dr. Tom Blundell to kindly provide
15 X-ray coordinates. We are grateful to the institutions that support our laboratory (Inserm),
16 University of Lille, Institut Pasteur de Lille, CNRS, EU, Région Hauts de France and Pôle de
17 Recherche Interdisciplinaire du Médicament (PRIM). AT thanks to the association 'les Amis de
18 l'Institut Pasteur à Bruxelles' for the PhD fellowship. SS thanks the STDF for the support of
19 HSBR (5590). RW is Research Associate at the National Fund for Scientific Research FNRS-
20 FRS (Belgium).

21

1 **Figure 1. Close-up views of the 18 new EthR complex structures.** The protein is displayed in
2 grayscale ribbon representation, with atoms in lines representation and color by element
3 according CPK coloring scheme. Ligand are in sticks representation with carbon, nitrogen,
4 oxygen, sulfur and fluorine in green, blue, red, yellow and light blue, respectively. The PDB id
5 and the compound name are indicated for each structure. The OMIT density map contoured at
6 1.5σ level (blue mesh) is represented within 2.0 \AA around ligand atoms. The pictures were
7 drawn with PyMOL (2009-2014 Schrödinger, LLC, <http://pymol.org>).

8 **Figure 2. Two 180-degree rotation front views of the EthR structure showing the molecular**
9 **surface of ligand binding cavity colored according hydrophobicity scale.** The protein is
10 displayed in sand-colored ribbon representation. The 43 residues lining the binding cavity, i.e.
11 having at least one atom participating to the cavity surface, are showed in sticks representation
12 and labeled. There are Leu⁷⁶, Val⁷⁹, Val⁸⁰, Leu⁸⁷, Leu⁹⁰, Ala⁹¹, Asn⁹³, Pro⁹⁴, Ala⁹⁵, Thr⁹⁷, Asp⁹⁸,
13 Arg⁹⁹, Met¹⁰², Trp¹⁰³, Thr¹⁰⁵, Gly¹⁰⁶, Ile¹⁰⁷, Phe¹¹⁰, Phe¹¹⁴, Val¹²⁰, Thr¹²¹, Gly¹²⁴, Gln¹²⁵, Arg¹²⁸,
14 Val¹³⁴, Leu¹³⁷, Trp¹³⁸, Met¹⁴², Trp¹⁴⁵, Tyr¹⁴⁸, Thr¹⁴⁹, Val¹⁵², Ala¹⁵⁵, Glu¹⁵⁶, Arg¹⁵⁹, Asn¹⁷⁶, Asn¹⁷⁹,
15 Glu¹⁸⁰, Arg¹⁸¹, Leu¹⁸³, Phe¹⁸⁴, Phe¹⁸⁷ and Trp²⁰⁷. The coloring scheme for amino acids is carbon
16 in green, nitrogen in blue, oxygen in red and sulfur in yellow. The cavity molecular surface is
17 colored by amino acid Kyte-Doolittle hydrophobicity scale [58], with colors ranging from
18 dodger blue (value of -4.5) for the most hydrophilic to white at 0.0, to orange-red for the most
19 hydrophobic residues (value of +4.5). The surface was computed with CASTp [45] and
20 displayed with UCSF Chimera [49]. The EthR structure Cmpd1 was taken to illustrate the ligand
21 binding cavity because having the cavity with the highest molecular volume, i.e. 1518 \AA^3 (see
22 Table S1). Three other orientations of the picture are provided in Figure S3.

1 **Figure 3. Correlation between ligand and cavity volumes in EthR structures.** Plot of ligand
2 volumes obtained from 3V server [34], as a function of cavity volumes calculated using CASTp
3 server [45]. For the two structures of mutant G106W (pdb ids 3QPL and 3TP3), the van der
4 Waals volume of a Trp was considered, i.e. 163 Å³ according [60]. In case of two or three
5 ligands in the cavity, the considered ligand volume was doubled or tripled, respectively. The
6 Pearson's correlation coefficient and the p-value are indicated.

7 **Figure 4. (a) Distribution of Tc scores of small molecules found in EthR binding cavities.** Tc
8 scores from all-against-all comparison of the 81 ligands were used to generate the histogram.
9 Circular (Morgan) fingerprints were used for encoding chemical structures before comparisons
10 **(b) Hierarchical clustering tree of EthR ligands (polar representation).** The 36 chemical
11 clusters defined using a cutting tree at 0.48 in <Tc> are labeled by CC-1 to CC-36 and the value
12 of <Tc> characterizing the cluster indicated.

13 **Figure 5. Phe¹⁸⁴ side chain displacements in EthR binding cavity.** (a) Heatmap and
14 hierarchical clustering of Phe¹⁸⁴ side chain displacements. The conformations of the Phe¹⁸⁴ side
15 chain were pair-wisely compared over the 83 EthR structures by calculating *RMSD* values.
16 Conformations were clustered using hierarchical clustering and the obtained dendrogram is
17 displayed at the top and at the left of heatmap. The structure labels are shown at the bottom and
18 at the right of the heatmap. *RMSD* values are colored from red (low ~0 Å) to yellow (high ~3 Å).
19 Color key and *RMSD* histogram are shown at the left-top. (b) Stick representation showing the
20 three different conformations for the Phe¹⁸⁴ side chain: the “open”, “close” and “intermediate”
21 orientations are depicted with carbon atoms in green, magenta and orange, respectively. The
22 three conformations were taken from structures 1U9N (open), 1T56 (close) and 4DW6
23 (intermediate). The two orientations for the pair of side chains Phe¹⁴¹/Trp¹⁴⁵ are depicted with

1 carbons in green (structure 1U9N) and cyan (structure 3TP3). Protein structure is shown in light
2 gray ribbon.

3 **Figure 6. Hot spots in EthR binding cavity.** Two views, (a)-(b) and (c)-(d), of the 8 EthR
4 hotspots are represented. In order to locate the binding cavity, a ligand molecule (BDM44825) is
5 depicted in stick representation with carbon, nitrogen, oxygen and fluorine in magenta, blue red
6 and yellow, respectively. The side chains of hotspots are labeled and depicted in sticks
7 representation with carbon in green. The hot spots are showed by spheres colored in red, orange,
8 cyan, yellow and grey for Phe¹¹⁰, Gly¹⁰⁶, Asn¹⁷⁶/Asn¹⁷⁹, Trp¹⁰³/Trp¹³⁸/Tyr¹⁴⁸ and Thr¹⁴⁹,
9 respectively. On the left part (a) and (c), both side chains and hotspot spheres are displayed,
10 while only hotspots are displayed on the right part (b) and (d). Mean distance between hot spots
11 is indicated with standard deviation in parentheses.

12

13

1 **Supporting Information**

2 **Table S1. The database of EthR crystal structures.**

3 **Table S2. Data collection statistics of the 18 new complexes.**

4 **Table S3. Refinement statistics of the 18 new complexes.**

5 **Table S4. List of biophysical assays for EthR-compound binding and activities of studied**
6 **small molecules.**

7 **Table S5. Physical-chemical properties of small molecules co-crystallized with EthR.**

8 **Table S6. Detailed interactions between ligands and EthR.**

9 **Table S7. Electron density fit for EthR ligands.**

10 **Figure S1. EthR structure in biological assembly.** Front (**a, b**) and top (**c, d**) views of two
11 dimeric EthR structures, 1T56 and 1U9N. Proteins are shown in ribbon colored in cyan and in
12 magenta for 1T56 and 1U9N, respectively. The protein molecular surface is depicted in
13 transparency. HTH motif and LBD are labelled. The top views allow highlighting the putative
14 entrance of the LBD, whereas the front views clearly show the obstruction of the second entrance
15 due to the dimeric organisation. Ligand molecules are represented in stick representation with
16 carbon and oxygen in green and red, respectively. The pictures were obtained using PyMOL
17 (2009-2014 Schrödinger, LLC, <http://pymol.org>).

18 **Figure S2. Heatmap of Tc scores over the 81 EthR ligands.** The figure was produced by the
19 `heatmap.2` function in R package, using Ward's minimum variance method in hierarchical
20 agglomerative clustering and Tc as similarity metrics (see Material and Methods). The color key
21 and histogram (green line) of dissimilarity distances (i.e. obtained by subtracting the Tc from 1)
22 are shown in upper-left corner. The hierarchical tree is depicted on both sides of the heatmap
23 while the structure labels are denoted on the two other sides.

1 **Figure S3. EthR structure showing the molecular surface of ligand binding cavity colored**
2 **according hydrophobicity scale. (a)** Top view showing the putative entrance of LBD. **(b)** and
3 **(c)** Two front views rotated by 180°. These two views complement those of Figure 1 for a 360-
4 degree coverage of the front view. The legend is the same as in Figure 1.

5 **Figure S4. The 36 chemical clusters.** The eighteen new structures are labeled in blue. The mean
6 Tc of chemical clusters contained more than one member is given.

7 **Figure S5. Large-scale side chain movements for four residues of EthR binding cavity.**
8 Heatmap and hierarchical clustering of Met¹⁰², Trp¹³⁸, Phe¹⁴¹ and Trp¹⁴⁵, side chains are
9 presented. The dendrograms are displayed at the top and at the left of heatmap while structure
10 labels are showed at the bottom and at the right. *RMSD* values are colored from red (low ~0 Å)
11 to yellow (high ~2-3 Å). Color key and *RMSD* histogram are shown at the left-top. The figures
12 were produced by the *heatmap.2* function in R package, using complete linkage clustering
13 method and *RMSD* as dissimilarity metrics.

14 **Figure S6. Match between modelled ligand and electron density.** The 11 ligands of EthR
15 database displaying the worst agreements are shown. The PDB id and ligand label are indicated,
16 as well as values of the four calculated metrics. The map (2Fo – Fc) contoured at 1.5 σ level
17 (blue mesh) is represented within 1.5 Å around ligand atoms. Ligand molecules are in stick
18 representation with carbon, nitrogen, oxygen, sulfur and fluorine in green, blue, red, yellow and
19 light blue, respectively. The pictures were drawn with PyMOL (2009-2014 Schrödinger, LLC,
20 <http://pymol.org>).

21

1 **References**

- 2 [1] World Health Organization, Global tuberculosis report (2017)
3 http://www.who.int/tb/publications/global_report/gtbr2017_main_text.pdf
- 4 [2] C. Dye, M.A. Espinal, C.J. Watt, C. Mbiaga, B.G. Williams, Worldwide incidence of
5 multidrug-resistant tuberculosis, *J. Infect. Dis.* 85 (2002) 1197-1202.
- 6 [3] M. Lauzardo, C.A. Peloquin, Tuberculosis therapy for 2016 and beyond, *Expert Opin.*
7 *Pharmacother.* 17 (2016) 1859-1872.
- 8 [4] V. Singh, V. Mizrahi, Identification and validation of novel drug targets in *Mycobacterium*
9 *tuberculosis*, *Drug. Discov. Today* 22 (2017) 503-509.
- 10 [5] F. Wang, R. Langley, G. Gulten, L.G. Dover, G.S. Besra, W.R.Jr. Jacobs, J.C. Sacchettini,
11 Mechanism of thioamide drug action against tuberculosis and leprosy, *J. Exp. Med.* 204 (2007)
12 73-78.
- 13 [6] A.R. Baulard, J.C. Betts, J. Engohang-Ndong, S. Quan, R.A. McAdam, P.J. Brennan, C. Locht,
14 G.S. Besra, Activation of the pro-drug ethionamide is regulated in mycobacteria, *J. Biol. Chem.*
15 275 (2000) 28326-28331.
- 16 [7] A.E. Debarber, K. Mdluli, M. Bosman, L.G. Bekker, C.E. 3rd Barry, Ethionamide activation
17 and sensitivity in multidrug-resistant *Mycobacterium tuberculosis*, *Proc. Natl. Acad. Sci. USA*,
18 97 (2000) 9677-9682.
- 19 [8] T.A. Vannelli, A. Dykman, P.R. Ortiz de Montellano, The antituberculosis drug ethionamide is
20 activated by a flavoprotein monooxygenase, *J. Biol. Chem.* 277 (2002), 12824-12829.
- 21 [9] S.S. Grant, S. Wellington, T. Kawate, C.A. Desjardins, M.R. Silvis, C. Wivagg, M. Thompson,
22 K. Gordon, E. Kazyanskaya, R. Nietupski, N. Haseley, N. Iwase, A.M. Earl, M. Fitzgerald, D.T.
23 Hung, Baeyer-Villiger monooxygenases EthA and MymA are required for activation of

- 1 replicating and non-replicating *Mycobacterium tuberculosis* inhibitors, *Cell Chem. Biol.* 23
2 (2016) 666-677.
- 3 [10] N. Blondiaux, M. Moune, M. Desroses, R. Frita, M. Flipo, V. Mathys, M. Kiass, V. Delorme,
4 K. Djaout, V. Trebosc, C. Kemmer, R. Wintjens, A. Wohlkönig, R. Antoine, L. Huot, D. Hot,
5 M. Coscolla, J. Feldmann, S. Gagneux, C. Loch, P. Brodin, M. Gitzinger, B. Deprez, N.
6 Willand, A. R. Baulard, Reversion of antibiotic resistance in *Mycobacterium tuberculosis* by
7 spiroisoxazoline SMART-420, *Science* 355 (2017) 1206-1211.
- 8 [11] M.I. Samanovic, H.C. Hsu, M. B. Jones, V. Jones, M.R. McNeil, S.H. Becker, A.T. Jordan, M.
9 Strnad, C. Xu, M. Jackson, H. Li, K.H. Darwin, Cytokinin signaling in *Mycobacterium*
10 *tuberculosis*, *MBio.* 9 (2018) pii:e00989-18.
- 11 [12] M.L.T. Ang, S.Z. Rahim, P.F. de Sessions, W. Lin, V. Koh, K. Pethe, M.L. Hibberd, S.
12 Alonso, EthA/R-independent killing of *Mycobacterium tuberculosis* by ethionamide, *Front*
13 *Microbiol.* 8 (2017) 710.
- 14 [13] J.L. Ramos, M. Martinez-Bueno, A.J. Molina-Henares, W. Terán, K. Watanabe, X. Zhang,
15 M.T. Gallegos, R. Brennan, R. Tobes, The TetR family of transcriptional repressors, *Microbiol.*
16 *Mol. Biol. Rev.* 69 (2005) 326-356.
- 17 [14] L. Cuthbertson, J.R. Nodwell, The TetR family of regulators, *Microbiol. Mol. Biol. Rev.* 77
18 (2013) 440-475.
- 19 [15] W. Hillen, G. Klock, I. Kaffenberger, L.V. Wray, W.S. Reznikoff, Purification of the TET
20 repressor and TET operator from the transposon Tn10 and characterization of their interaction,
21 *J. Biol. Chem.* 257 (1982) 6605-6613.

- 1 [16] J. Engohang-Ndong, D. Baillat, M. Aumercier, F. Bellefontaine, G.S. Besra, C. Locht, A.R.
2 Baulard, EthR, a repressor of the TetR/CamR family implicated in ethionamide resistance in
3 mycobacteria, octamerizes cooperatively on its operator, *Mol. Microbiol.* 51 (2004) 175-188.
- 4 [17] N. Willand, B. Dirié, X. Carette, P. Bifani, A. Singhal, M. Desroses, F. Leroux, E. Willery, V.
5 Mathys, R. Déprez-Poulain, G. Delcroix, F. Frénois, M. Aumercier, C. Locht, V. Villeret, B.
6 Déprez, A.R. Baulard, Synthetic EthR inhibitors boost antituberculous activity of ethionamide,
7 *Nat. Med.* 15 (2009) 537-544.
- 8 [18] N. Willand, M. Desroses, P. Toto, B. Dirié, Z. Lens, V. Villeret, P. Rucktooa, C. Locht, A.
9 Baulard, B. Déprez, Exploring drug target flexibility using *in situ* click chemistry: application to
10 a mycobacterial transcriptional regulator, *ACS Chem. Biol.* 5 (2010) 1007-1013.
- 11 [19] M. Flipo, M. Desroses, N. Lecat-Guillet, B. Dirié, X. Carette, F. Leroux, C. Piveteau, F.
12 Demirkaya, Z. Lens, P. Rucktooa, V. Villeret, T. Christophe, H.K. Jeon, C. Locht, P. Brodin, B.
13 Déprez, A.R. Baulard, N. Willand, Ethionamide boosters: synthesis, biological activity, and
14 structure-activity relationships of a series of 1,2,4-oxadiazole EthR inhibitors, *J. Med. Chem.* 54
15 (2011) 2994-3010.
- 16 [20] M. Flipo, M. Desroses, N. Lecat-Guillet, B. Villemagne, N. Blondiaux, F. Leroux, C.
17 Piveteau, V. Mathys, M.P. Flament, J. Siepmann, V. Villeret, A. Wohlkönig, R. Wintjens, S.H.
18 Soror, T. Christophe, H.K. Jeon, C. Locht, P. Brodin, B. Déprez, A.R. Baulard, N. Willand,
19 Ethionamide boosters. 2. Combining bioisosteric replacement and structure-based drug design to
20 solve pharmacokinetic issues in a series of potent 1,2,4-oxadiazole EthR inhibitors, *J. Med.*
21 *Chem.* 55 (2012) 68-83.
- 22 [21] M. Flipo, N. Willand, N. Lecat-Guillet, C. Hounsou, M. Desroses, F. Leroux, Z. Lens, V.
23 Villeret, A. Wohlkönig, R. Wintjens, T. Christophe, H.K. Jeon, C. Locht, P. Brodin, A.R.

- 1 Baulard, B. Déprez, Discovery of novel N-phenylphenoxyacetamide derivatives as EthR
2 inhibitors and ethionamide boosters by combining high-throughput screening and synthesis, J.
3 Med. Chem. 55 (2012) 6391-6402.
- 4 [22] P.O. Nikiforov, M. Blaszczyk, S. Surade, H.I. Boshoff, A. Sajid, V. Delorme, N. Deboosere,
5 P. Brodin, A.R. Baulard, C.E. 3rd Barry, T.L. Blundell, C. Abell, Fragment-sized EthR inhibitors
6 exhibit exceptionally strong ethionamide boosting effect in whole-cell *Mycobacterium*
7 *tuberculosis* assays, ACS Chem. Biol. 12 (2017) 1390-1396.
- 8 [23] J. Costa-Gouveia, E. Pancani, A. Machelart, V. Delorme, G. Salzano, R. Iantomasi, C.
9 Piveteau, C.J. Queval, O.R. Song, M. Flipo, B. Déprez, J.P. Saint-André, J. Hureauux, L.
10 Majlessi, N. Willand, A. Baulard, P. Brodin, R. Gref, Combination therapy for tuberculosis
11 treatment: pulmonary administration of ethionamide and booster co-loaded nanoparticles, Sci.
12 Rep. 7 (2017) 7, 5390 doi: 10.1038/s41598-017-05453-3.
- 13 [24] B. Villemagne, M. Flipo, N. Blondiaux, C. Crauste S. Malaquin, F. Leroux, C. Piveteau, V.
14 Villeret, P. Brodin, B.O. Villoutreix, O. Sperandio, S.H. Soror, A. Wohlkönig, R. Wintjens, B.
15 Déprez, A.R. Baulard, N. Willand, Ligand efficiency driven design of new inhibitors of
16 *Mycobacterium tuberculosis* transcriptional repressor EthR using fragment growing, merging,
17 and linking approaches, J. Med. Chem. 57 (2014) 4876-4888.
- 18 [25] S. Surade, N. Ty, N. Hengrung, B. Lechartier, S.T. Cole, C. Abell, T.L. Blundell, A structure-
19 guided fragment-based approach for the discovery of allosteric inhibitors targeting the lipophilic
20 binding site of transcription factor EthR, Biochem J. 458 (2014) 387-394.
- 21 [26] P.O. Nikiforov, S. Surade, M. Blaszczyk, V. Delorme, P. Brodin, A.R. Baulard, T.L.
22 Blundell, C. Abell, A fragment merging approach towards the development of small molecule

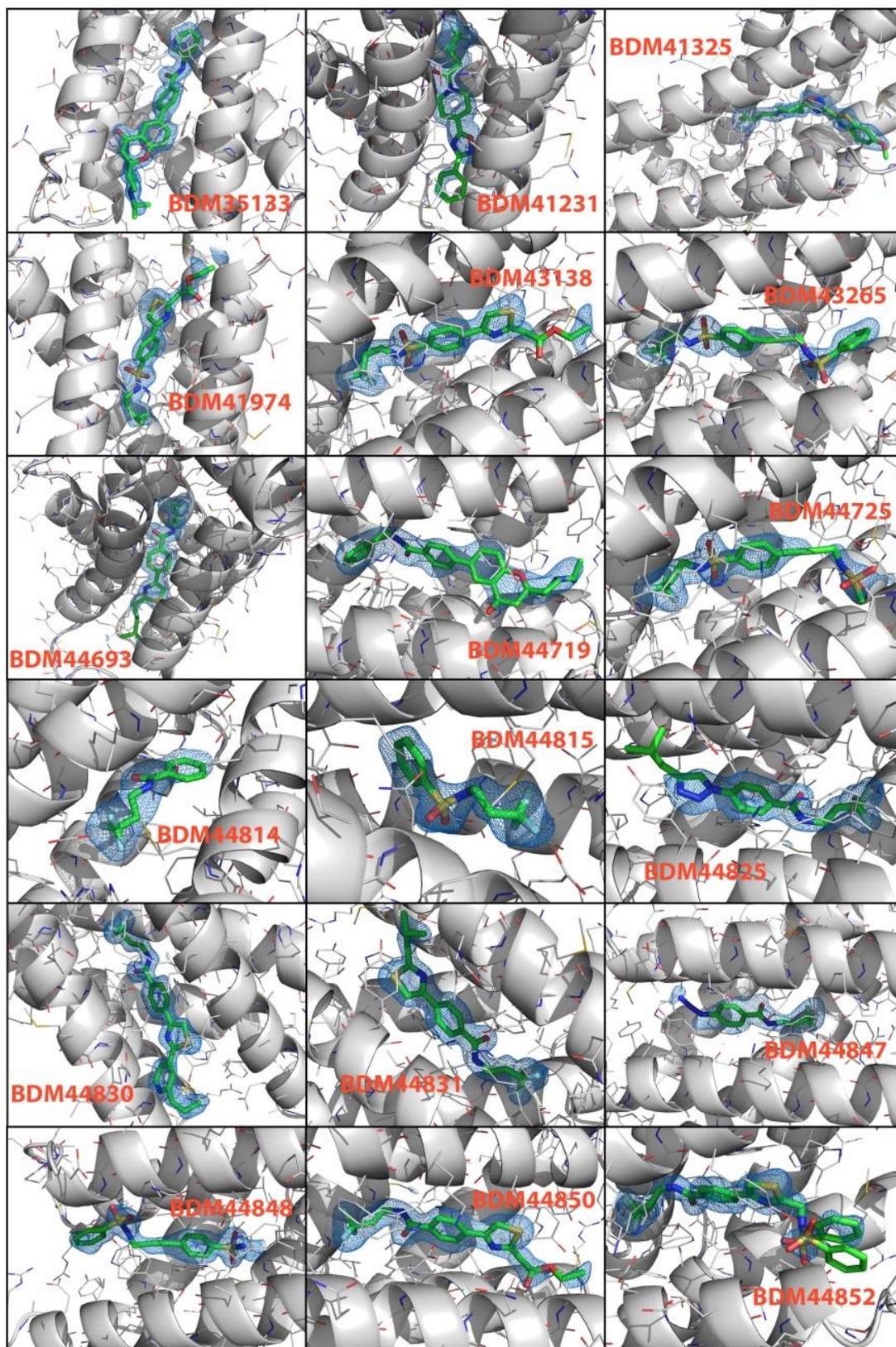
- 1 inhibitors of *Mycobacterium tuberculosis* EthR for use as ethionamide boosters, *Org. Biomol.*
2 *Chem.* 14 (2016) 2318-2326.
- 3 [27] D.S. Chan, V. Mendes, S.E. Thomas, B.N. McConnell, D. Matak-Vinković, A.G. Coyne, T.L.
4 Blundell, C. Abell, Fragment screening against the EthR-DNA interaction by native mass
5 spectroscopy, *Angew. Chem. Int. Ed. Engl.* 56 (2017) 7488-7491.
- 6 [28] L.G. Dover, P.E. Corsino, I.R. Daniels, S.L. Cocklin, V. Tatituri, G.S. Besra, K. Fütterer,
7 Crystal structure of the TetR/CamR family repressor *Mycobacterium tuberculosis* EthR
8 implicated in ethionamide resistance, *J. Mol. Biol.* 340 (2004) 1095-1105.
- 9 [29] F. Frénois, J. Engohang-Ndong, C. Locht, A.R. Baulard, V. Villeret. Structure of EthR in a
10 ligand bound conformation reveals therapeutic perspectives against tuberculosis, *Mol. Cell* 16
11 (2004) 301-307.
- 12 [30] N.J. Tatum, J.W. Liebeschuetz, J.C. Cole, R. Frita, A. Herledan, A.R. Baulard, N. Willand, E.
13 Pohl, New active leads for tuberculosis booster drugs by structure-based drug discovery, *Org.*
14 *Biomol. Chem.* 15 (2017) 10245-10255.
- 15 [31] G. Mugumbate, V. Mendes, M. Blaszczyk, M. Sabbah, G. Papadatos, J. Lelievre, L. Ballell, D.
16 Barros, C. Abell, T.L. Blundell, J.P. Overington, Target identification of *Mycobacterium*
17 *tuberculosis* phenotypic hits using a concerted chemogenomic, biophysical, and structural
18 approach, *Front Pharmacol.* 8 (2017) 681.
- 19 [32] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I.N. Shindyalov,
20 P.E. Bourne, The Protein Data Bank, *Nucleic Acids Res.* 28 (2000) 235-242.
- 21 [33] N.M. O'Boyle, R.A. Sayle, Comparing structural fingerprints using a literature-based
22 similarity benchmark, *J. Cheminform.* 8 (2016) 36.

- 1 [34] N.R. Voss, M. Gerstein, 3V: cavity, channel and cleft volume calculator and extractor, *Nucleic*
2 *Acids Res.* 38 (2010) W555-W562.
- 3 [35] D. Rogers, M. Han, Extended-connectivity fingerprints, *J. Chem. Inf. Model.* 50 (2010) 742-
4 754.
- 5 [36] D. Bajusz D, A. Rácz, K. Héberger, Why is Tanimoto index an appropriate choice for
6 fingerprint-based similarity calculations? *J. Cheminform.* 7 (2015) 20.
- 7 [37] A. Rambaut, Figtree – a graphical viewer of phylogenetic trees.
8 <http://tree.bio.ed.ac.uk/software/figtree>, 2016.
- 9 [38] T.G. Battye, L. Kontogiannis, O. Johnson, H.R. Powell, A.G. Leslie, iMOSFLM: a new
10 graphical interface for diffraction-image processing with MOSFLM, *Acta Crystallogr. D67*
11 (2011) 271-281.
- 12 [39] A. Vagin, A. Teplyakov, Molecular replacement with MOLREP, *Acta Crystallogr. D66* (2010)
13 22-25.
- 14 [40] G.N. Murshudov, P. Skubák, A.A. Lebedev, N.S. Pannu, R.A. Steiner, R.A. Nicholls, M.D.
15 Winn, F. Long, A.A. Vagin, REFMAC5 for the refinement of macromolecular crystal structures,
16 *Acta Crystallogr. D67* (2011) 355-367.
- 17 [41] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr.*
18 *D60* (2004) 2126-2132.
- 19 [42] F. Long, R.A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A.Vaitkus, G.N. Murshudov,
20 AceDRG: a stereochemical description generator for ligands, *Acta Crystallogr. D73* (2017) 112-
21 122.
- 22 [43] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan,
23 E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A.

- 1 Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and
2 current developments, *Acta Crystallogr. D67* (2011) 235-242.
- 3 [44] V.B. Chen, W.B. 3rd Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W.
4 Murray, J.S. Richardson, D.C. Richardson, MolProbity: all-atom structure validation for
5 macromolecular crystallography, *Acta Crystallogr. D66* (2010) 12-21.
- 6 [45] J. Dundas, Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, J. Liang, CASTp: computed atlas
7 of surface topography of proteins with structural and topographical mapping of functionally
8 annotated residues, *Nucleic Acids Res.* 34 (2006) W116-W118.
- 9 [46] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, *J.*
10 *Mol. Biol.* 372 (2007) 774-797.
- 11 [47] H. Hasegawa, L. Holm, Advances and pitfalls of protein structural alignment, *Curr. Opin.*
12 *Struct. Biol.* 19 (2009) 341-348.
- 13 [48] J.D. Durran, J.A. McCammon, BINANA: a novel algorithm for ligand-binding
14 characterization, *J. Mol. Graph. Model.* 29 (2011) 888-893.
- 15 [49] E.F. Petterson, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E.
16 Ferrin, UCSF Chimera – a visualization system for exploratory research and analysis, *J.*
17 *Comput. Chem.* 25 (2004) 1605-1615.
- 18 [50] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson,
19 Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility, *J.*
20 *Comp. Chem.* 30 (2009) 2785-2791.
- 21 [51] P.J. Kraulis, MOLSCRIPT: a program to produce both detailed and schematic plots of protein
22 structures, *J. Appl. Cryst.* 24 (1991) 946-950.

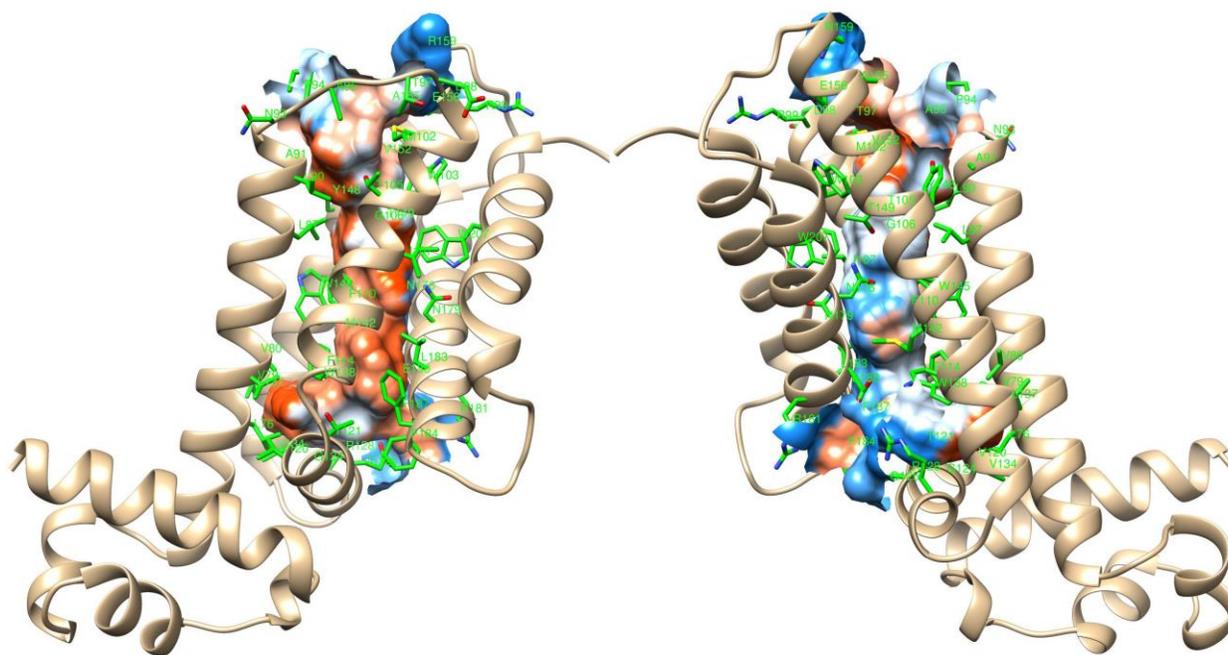
- 1 [52] E.A. Merritt, D.J. Bacon, Raster3D photorealistic molecular graphics, *Methods Enzymol.* 277
2 (1997) 505-524.
- 3 [53] C.I. Bränden, T.A. Jones, Between objectivity and subjectivity, *Nature* 343 (1990) 687-689.
- 4 [54] T.A. Jones, J.Y. Zou, S.W. Cowan, M. Kjeldgaard, Improved methods for building protein
5 models in electron density maps and the location of errors in these models, *Acta Crystallogr.*
6 A47 (1991) 110-119.
- 7 [55] I.J. Tickle, Statistical quality indicators for electron-density maps, *Acta Crystallogr.* D68
8 (2012) 454-467.
- 9 [56] Z. Yu, S.E. Reichheld, A. Savchenko, J. Parkinson, A.R. Davidson, A comprehensive analysis
10 of structural and sequence conservation in the TetR family transcriptional regulators, *J. Mol.*
11 *Biol.* 400 (2010) 847-864.
- 12 [57] X. Carette, N. Blondiaux, E. Willery, S. Hoos, N. Lecat-Guillet, Z. Lens, A. Wohlkonig, R.
13 Wintjens, S.H. Soror, F. Frénois, B. Dirié, V. Villeret, P. England, G. Lippens, B. Déprez, C.
14 Loch, N. Willand, A.R. Baulard, Structural activation of the transcriptional repressor EthR from
15 *Mycobacterium tuberculosis* by single amino acid change mimicking natural and synthetic
16 ligands, *Nucleic Acids Res.* 40 (2012) 3018-3030.
- 17 [58] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein,
18 *J. Mol. Biol.* 157 (1982) 105-132.
- 19 [59] N.M. Pearce, T. Krojer, F. von Delf, Proper modelling of ligand binding requires an ensemble
20 of bound and unbound states, *Acta Crystallogr.* D73 (2017) 256-266.
- 21 [60] N.J. Darby, T.E. Creighton (1993) *Protein Structure*. IRL Press, Oxford University Press.
- 22
- 23

1 Figure 1



2

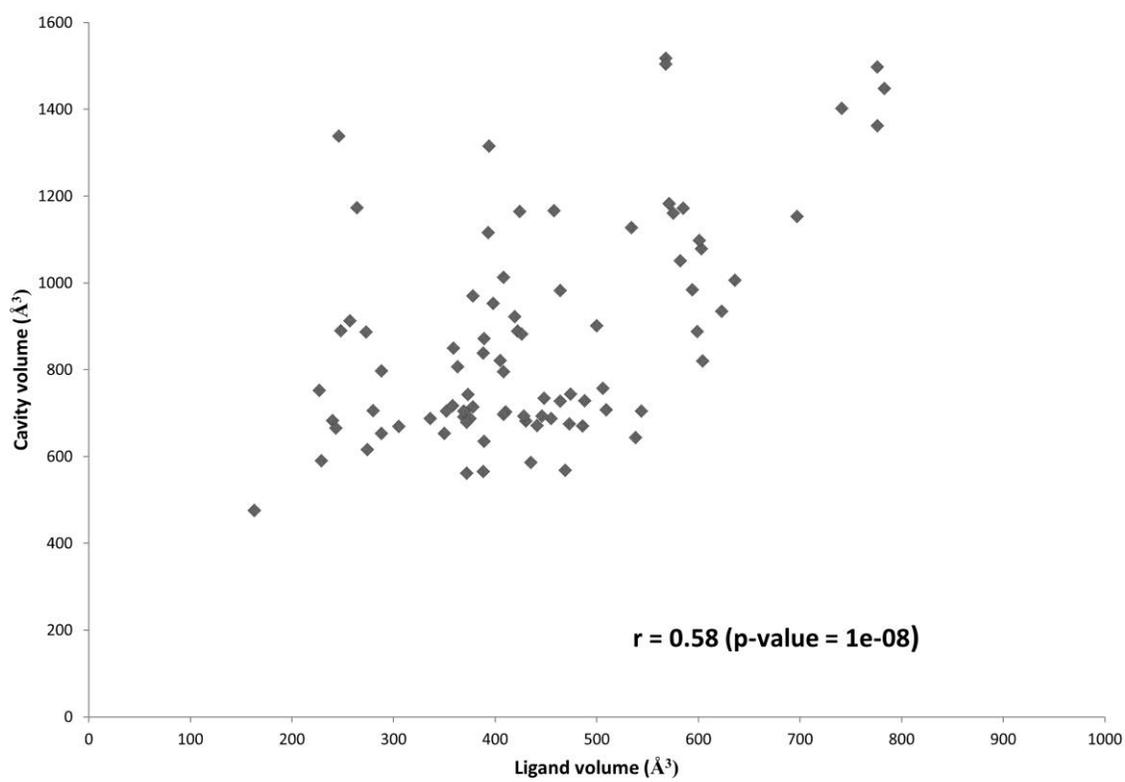
1 Figure 2



2

Author Manuscript

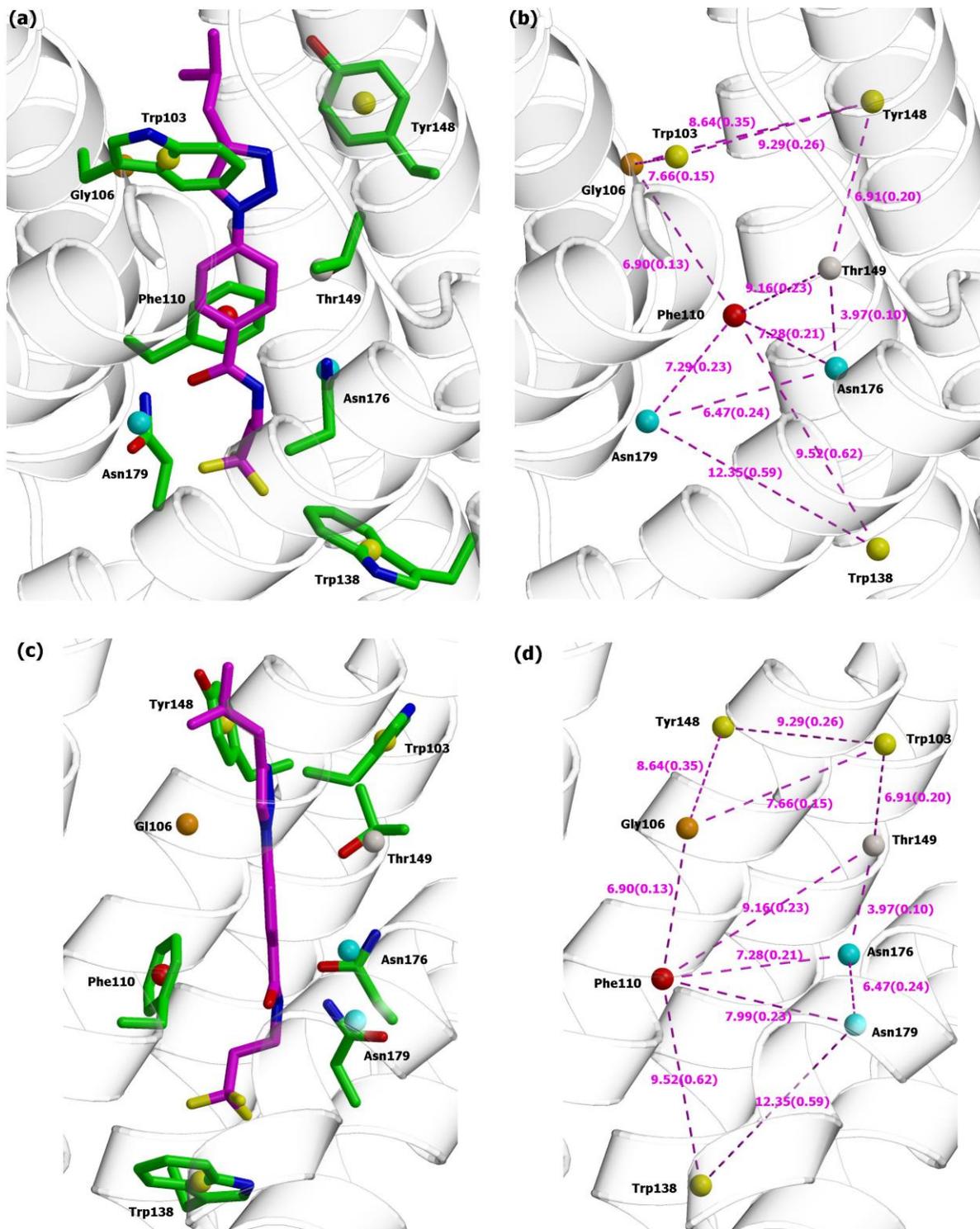
1 Figure 3



2

Author

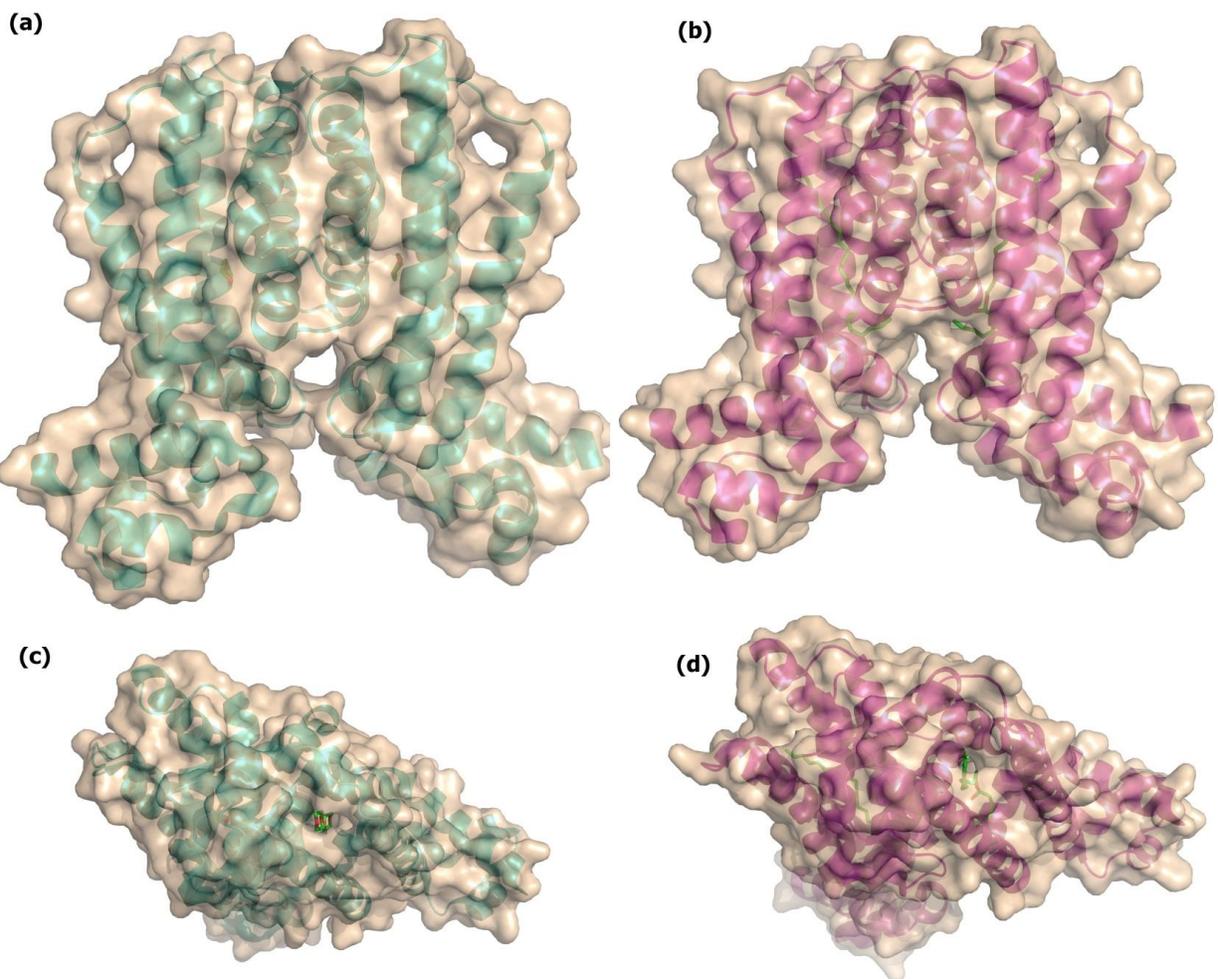
1 Figure 6



2

3

1 Figure S1

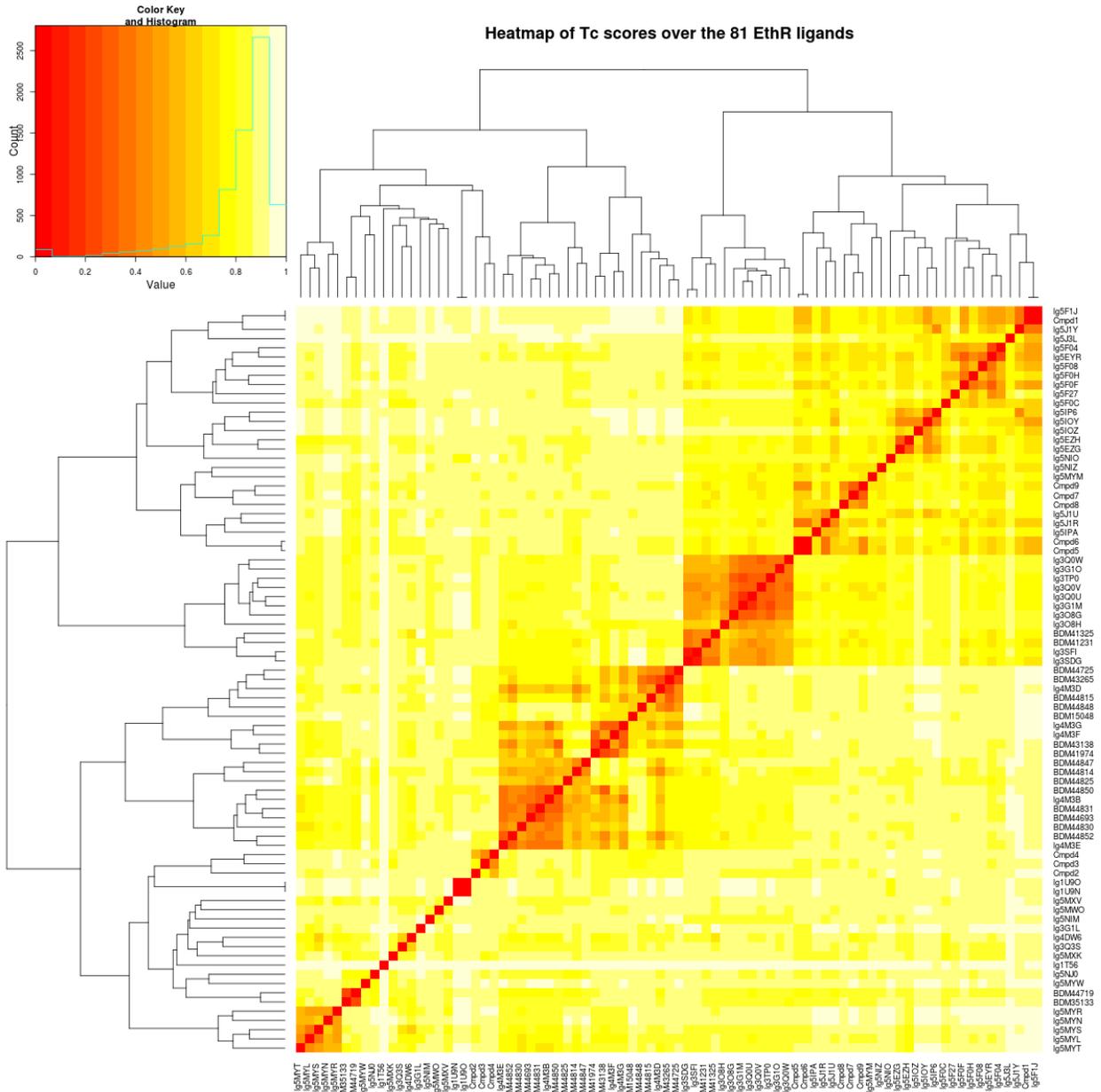


2

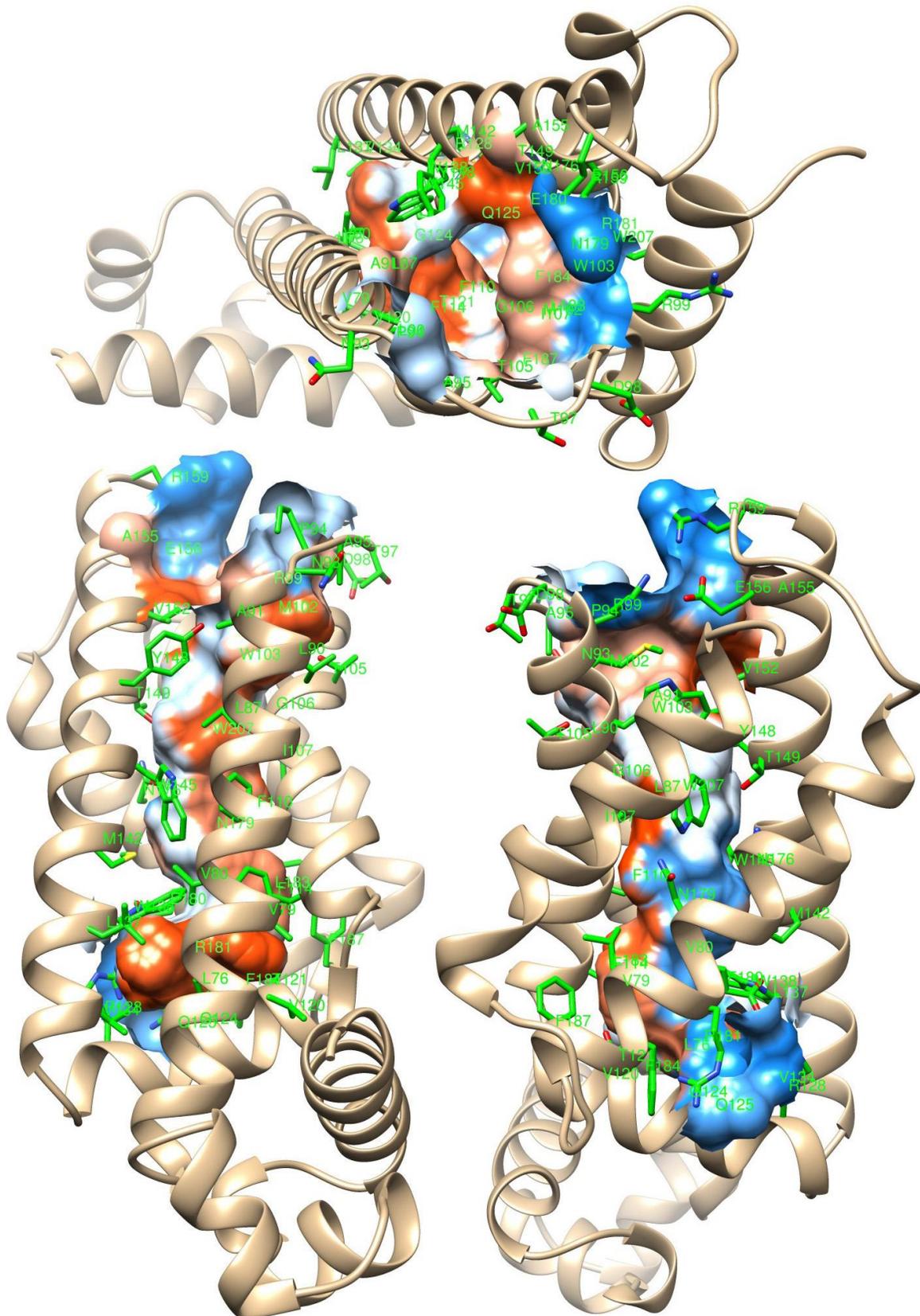
3

Auth

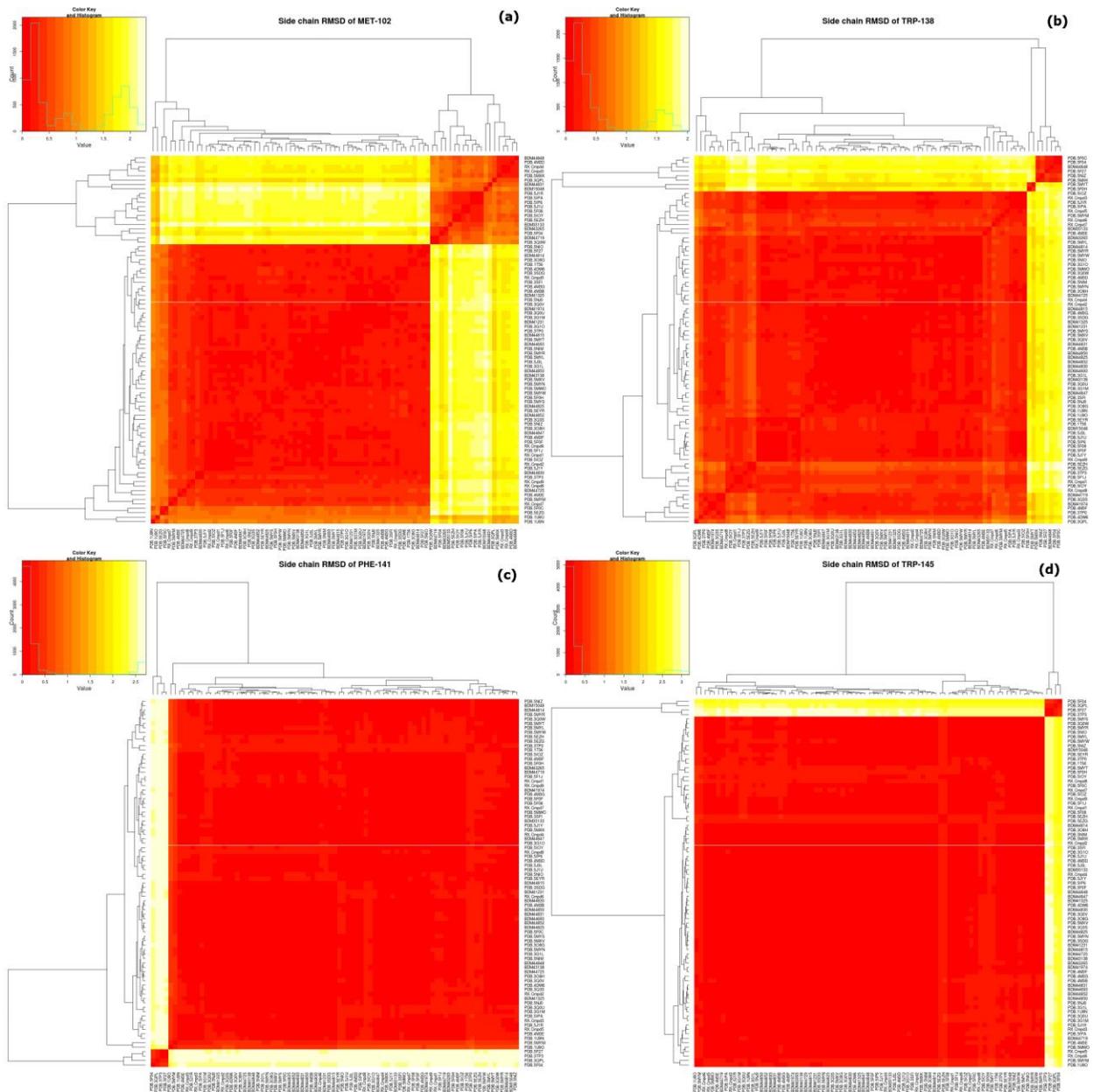
1 Figure S2



1 Figure S3



1 Figure S5



2