RSC Advances

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances
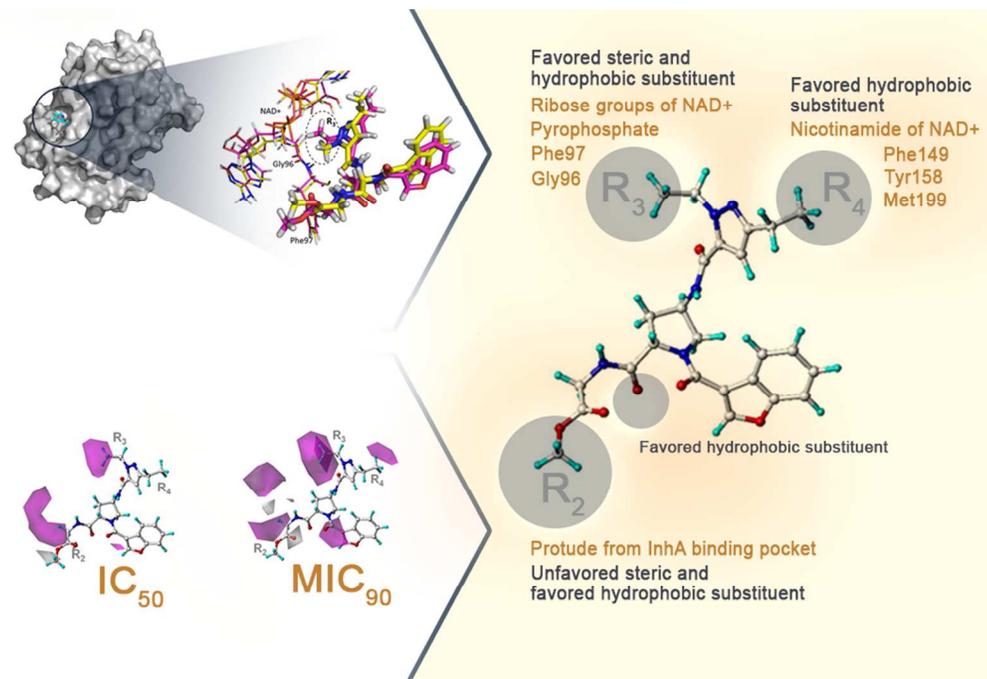
The structural concept for enhancing both IC50 and MIC90 activities summarized from MD simulations and CoMSIA results
90x60mm (300 x 300 DPI)

1   Structural Requirements of Benzofuran Pyrrolidine Pyrazole

2   Derivatives as Highly Potent Anti-tuberculosis Agents for

3   Good Correlation of IC_{50}-MIC Based on Integrated Results

4   from 3D-QSAR and MD Simulations

5

6   Pharit Kamsri[a], Auradee Punkvang[b], Supa Hannongbua[c], Patchreenart

7   Saparpakorn[c] and Pornpan Pungpo[a,*]

8

9   [a]*Department of Chemistry, Faculty of Science, Ubon Ratchathani University,*

10   *Ubonratchathani, Thailand*

11

12   [b]*Faculty of Science, Nakhon, Phanom University, Nakhon Phanom, Thailand*

13

14   [c]*Department of Chemistry, Faculty of Science, Kasetsart University, Bangkok, Thailand*

15

16   * All correspondence should be addressed to P. Pungpo: E-mail: pornpan_ubu@yahoo.com*

17   *Tel.: +664 535 3400 ext.4124, Fax: +664 528 8379*

18
19
20
21
22
23
24
25
26
27
28

1 **ABSTRACT**

2

3   A 2-*trans* enoyl-acyl carrier protein (ACP) reductase or InhA of *M. tuberculosis* is a

4 drug target of isoniazid (INH), the first-line drug for tuberculosis treatment. Many series of

5 compounds have been developed as novel inhibitors of this enzyme. However, they lack good

6 potency against purified InhA and activity against intact *M. tuberculosis* cells. Benzofuran

7 pyrrolidin pyrazole derivatives are potent direct InhA inhibitors. These compounds show high

8 potency for InhA inhibition with IC_{50} values at nanomolar levels. However, their activities

9 against *M. tuberculosis* cells in terms of MIC_{90} were about one-thousand fold than IC_{50}.

10 Accordingly, in this work, IC_{50} and MIC_{90} values of benzofuran pyrrolidin pyrazole

11 derivatives were subjected to CoMFA and CoMSIA studies in order to investigate the

12 structural basis required for good activity against both purified InhA and *M. tuberculosis*

13 cells. Moreover, MD simulations were employed to evaluate key interactions for binding

14 benzofuran pyrrolidin pyrazole derivatives in InhA. Based on MD results, the core structure

15 of these compounds is the key portion for binding in the InhA pocket. Alternatively, **R**

16 substituents showed weak interactions with the InhA pockets. Interpretation of IC_{50} and

17 MIC_{90} CoMSIA contour maps revealed the structural requirements in terms of steric,

18 electrostatic, hydrophobic and hydrogen donor and acceptor for IC_{50} and MIC_{90} values of

19 InhA inhibitors. Finally, the integrated results obtained from MD simulations and graphic

20 interpretation of CoMSIA models provided a structural concept for rational design of novel

21 InhA inhibitors with better potency against both the InhA enzyme and intact *M. tuberculosis*

22 cells.

23

24 **Keyword**: *M. tuberculosis*; InhA; MD simulations, 3D-QSAR

25

26

27

28

29

30

31

32

2

1   **1. Introduction**

2   Tuberculosis (TB) is an infectious disease caused by *Mycobacterium tuberculosis*

3   (*M. tuberculosis*) and remains one of the world's deadliest infectious diseases. The World

4   Health Organization (WHO) reported that an estimated 9.0 million people developed new TB

5   cases and 1.5 million people died from this disease in 2013. Moreover, the incidence of new

6   TB cases and deaths in 2013 was higher than those reported previously.[1] The high mortality

7   rate of TB is caused by multi drug-resistant tuberculosis (MDR-TB),[2-7] extensively drug-

8   resistant tuberculosis (XDR-TB),[8-9] totally drug-resistant tuberculosis (TDR-TB)[10-11] and

9   human immunodeficiency virus (HIV) co-infection.[1] A NADH-dependent 2-*trans* enoyl-acyl

10   carrier protein (ACP) reductase or InhA has been identified as potential drug target for

11   tuberculosis treatment.[12] This enzyme catalyzes the reduction of α,β-unsaturated fatty acids,

12   the last step in fatty acids biosynthesis in *M. tuberculosis*.[12-14] InhA was reported as the drug

13   target of isoniazid (INH), the first-line drug against tuberculosis.[15-23] Since INH is a prodrug,

14   it requires the activation process of catalase-peroxidase (KatG) to generate the acyl radical

15   active form. This radical is then covalently bound to nicotinamide adenine dinucleotide

16   (NAD$^+$) to produce an active INH-NAD adduct acting as a potent InhA inhibitor.[18-23] The

17   high potency of INH against InhA was lost by mutations in KatG. Therefore, many

18   researchers aimed to discover novel inhibitors that can directly inhibit InhA without the KatG

19   activation process. Inhibitors that can act like this are called direct InhA inhibitors. A class of

20   *N*-((3R,   5*S*)-1-(benzofuran-3-carbonyl)-5-carbamoylpyrrolidin-3-yl)-1*H*-pyrazole-5-

21   carboxamide derivatives (benzofuran pyrrolidin pyrazole derivatives) have been identified as

22   potent direct InhA inhibitors.[24] The majority of benzofuran pyrrolidine pyrazole derivatives

23   show high potency against purified InhA with inhibitory concentration of compound required

24   to inhibit InhA at 50% (IC$_{50}$) values at the nanomolar level. However, these compounds show

25   weak cellular activity against *M. tuberculosis*, with the minimum inhibitory concentration of

26   compound that resulted in complete inhibition in growth of *M. tuberculosis* 90% (MIC$_{90}$) at

27   the micromolar level. These results show poor correlation between IC$_{50}$ and MIC$_{90}$ values of

28   benzofuran pyrrolidine pyrazole derivatives. In this work, IC$_{50}$ and MIC$_{90}$ values of

29   benzofuran pyrrolidine pyrazole derivatives were used for comparative molecular field

30   analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) studies

31   in order to investigate the structural basis of these compounds for good activity against both

32   InhA and *M. tuberculosis*. Moreover, molecular dynamics (MD) simulations were employed

33   to evaluate the key interactions for binding of benzofuran pyrrolidin pyrazole derivatives in

1   InhA. Therefore, the integrated results obtained from MD simulations and graphic

2   interpretation of quantitative structure activity relationship (QSAR) models should provide

3   crucial structural concepts for improving the correlation between IC_{50} and MIC_{90} values of

4   benzofuran pyrrolidin pyrazole derivatives.

5

6   **2. Material and Methods**

7   *2.1 Data sets and biological activities*

8   Thirty-four benzofuran pyrrolidin pyrazole derivatives used for CoMFA and CoMSIA

9   studies were identified from the published literature.[24] Chemical structures and experimental

10   biological activities in terms of MIC_{90} and IC_{50} values of these compounds are shown in

11   **Table 1**. MIC_{90} and IC_{50} values were nominally converted into $\log(1/MIC_{90})$ and $\log(1/IC_{50})$

12   values for CoMFA and CoMSIA studies. Based on the diversity of structures and wide range

13   of activities, the data set of compounds was divided into 30 training set compounds for final

14   model development and 4 test set compounds for model validation. All chemical structures of

15   benzofuran pyrrolidin pyrazole derivatives were constructed using the standard tools

16   available in the GaussView 3.07 program and were then fully optimized using the HF/6-31G

17   method implemented in the Gaussian 09 program.[25] The harmonic vibrational frequencies of

18   the optimized geometries have also been calculated. All elements in the calculated Hessian

19   matrix are positive, which indicate that the structures are true minima on the potential energy

20   surface.

21

22   **Table 1** The chemical structures and activities against InhA and *M. tuberculosis* of

23   thirty-four benzofuran pyrrolidin pyrazole derivatives.

24

25
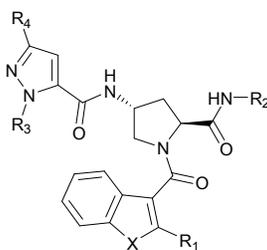
Cpd.	X	R_1	R_2	R_3	R_4	IC_{50} (μM)	MIC_{90}(μM)	$\log(1/IC_{50})$	$\log(1/MIC_{90})$
1	O	H	Et	Me	Et	0.034	8.00	7.47	5.10
2*	O	H	(CH2C(=O)OMe)	Me	Et	0.005	0.50	8.30	6.30
3	O	H	H	Me	Et	0.012	3.00	7.92	5.52

4

4	O	H	CH$_2$CF$_3$	Me	Et	0.046	4.00	7.34	5.40
5	O	H	CH$_2$CH$_2$CH$_3$	Me	Et	0.021	15.60	7.68	4.81
6	O	H	CH$_2$CH$_2$OMe	Me	Et	0.014	4.00	7.85	5.40
7*	O	H	CH$_2$CH$_2$COOEt	Me	Et	0.022	4.00	7.66	5.40
8	O	H	(tetrahydrofuran-2-yl)ethyl	Me	Et	0.045	4.00	7.35	5.40
9	O	H	cyclopentyl	Me	Et	0.040	4.00	7.40	5.40
10	O	H	4-(CF$_3$)benzyl	Me	Et	0.042	16.00	7.38	4.80
11*	O	H	3,4-difluorobenzyl	Me	Et	0.009	2.00	8.05	5.70
12	O	H	benzoxaborole methyl	Me	Et	0.035	3.00	7.46	5.52
13	O	H	2-benzyl-3-hydroxypropyl	Me	Et	0.112	1.00	6.95	6.00
14	O	H	2-(pyridin-2-yl)ethyl	Me	Et	0.025	1.00	7.60	6.00
15	O	H	(5-methylisoxazol-3-yl)methyl	Me	Et	0.018	16.00	7.74	4.80
16	O	H	(4-methyloxazol-2-yl)methyl	Me	Et	0.009	8.00	8.05	5.10
17	O	H	(2-ethyl-2H-1,2,3-triazol-4-yl)methyl	Me	Et	0.003	4.00	8.52	5.40
18	O	H	(1-ethyl-1H-pyrazol-3-yl)methyl	Me	Et	0.032	4.00	7.49	5.40
19	O	H	(oxazol-2-yl)methyl	Me	Et	0.005	1.00	8.30	6.00
20	O	H	(5-methylfuran-2-yl)methyl	Me	Et	0.021	1.50	7.68	5.82
21	O	H	CH$_2$COOMe	Me	Cyclopropyl	0.015	1.00	7.82	6.00
22	O	H	CH$_2$COOMe	Et	Et	0.003	0.05	8.52	7.30
23	O	H	H	Et	Et	0.004	0.50	8.40	6.30
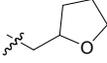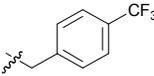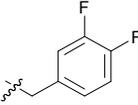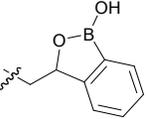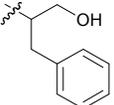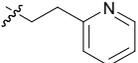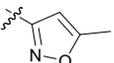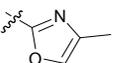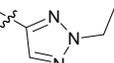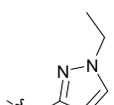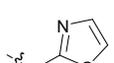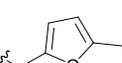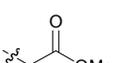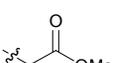
5
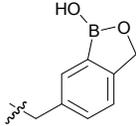
24	O	H	(boronic benzoxaborole group)	Et	Et	0.002	0.20	8.70	6.70
25	O	H	(3-methyl-1,2,4-oxadiazole group)	Et	Et	0.004	0.50	8.40	6.30
26	O	H	(pyridin-2-yl ethyl group)	Et	Et	0.004	0.50	8.40	6.30
27	O	H	CH$_2$CH$_2$OH	Et	Et	0.003	1.00	8.52	6.00
28	O	H	(pyridin-2-yl methyl group)	Et	Et	0.002	0.70	8.70	6.15
29	O	Et	(CH$_2$C(O)OMe group)	Me	Et	0.005	0.70	8.30	6.15
30	S	H	(CH$_2$C(O)OMe group)	Me	Et	0.029	1.00	7.54	6.00
31*	O	Ph	(CH$_2$C(O)OMe group)	Me	Et	0.003	1.50	8.52	5.82
32	O	H	(CH(CH$_2$OH)C(O)OMe group)	Me	Et	0.018	2.00	7.74	5.70
33	O	H	(CH$_2$C(O)O-tBu group)	Me	Et	0.007	2.00	8.15	5.70
34	O	H	(CH$_2$C(O)NHMe group)	Me	Et	0.008	2.00	8.10	5.70
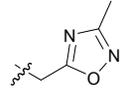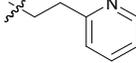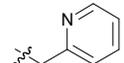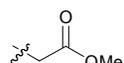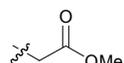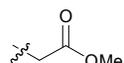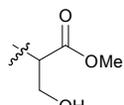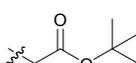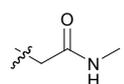
1   *test set

2

3   *2.2 Molecular docking calculations*

4   In this study, molecular docking calculations using the GOLD Program[26-30] were

5   employed with the aims of generating the initial structure for MD simulations and performing

6   molecular alignment to set up CoMFA and CoMSIA models. The available X-ray structure of

7   InhA in a complex with compound **1** (PDB code 4COD) was used as an initial structure for

8   molecular docking calculations. All atoms of the protein were kept rigid, whereas ligand was

9   flexible during the molecular docking calculations. The number of Genetic Algorithm (GA)

10   runs was set to 15 runs with the default search algorithm parameters. The docking

11   calculations were validated using the root-mean-square deviation (RMSD) value between the

12   docked and observed X-ray conformations of compound **1** in its pocket. A RMSD value

13   lower than 1 Å was acceptable. Then, molecular docking calculations with validated

1 parameters were used to dock all remaining compounds into the InhA binding pocket. The

2 binding mode that showed the lowest binding energy was selected for each compound and

3 was used to set up CoMFA and CoMSIA models. It was then used as the initial structure for

4 MD simulations of compounds **2**, **22**, **23** and **28**.

5

6 *2.3 Molecular dynamics simulations*

7 Compound **28**, with the best IC_{50} value, was selected to investigate its binding mode

8 in InhA. Moreover, the binding modes of compounds **2**, **22** and **23** were modelled by MD

9 simulations in order to investigate the effect of R_2 and R_3 substituents on the IC_{50} value. The

10 AMBER12 program[31] was employed to perform molecular dynamics simulations. The

11 complex structures of compounds **2**, **22**, **23** and **28** in InhA obtained from molecular docking

12 calculations were used as the initial structure in MD simulations. The Amber *ff03* force field

13 was used for the physical description of InhA[32]. The general Amber force field (GAFF)[33-34]

14 and restrained electrostatic potential (RESP) partial charges[35-38] of ligands and NAD^+ were

15 generated by the *antechamber* module implemented in the AMBER12 package. To generate

16 the system for MD simulations, the initial complex structure was solvated by TIP3P water[39]

17 in a truncated octahedral box extending up to 10 Å from the solute species. Five Na^+ ions

18 were added to neutralize the system charge. Initially, the energy of system was minimized

19 using a steepest decent method followed by the conjugate gradient method. Then, the system

20 was gradually warmed from 0 K to 300 K in 30 ps by restraining all atoms of the complex

21 with a restraint weight of 2 kcal/molÅ^2. This was followed by 70 ps of the position-restrained

22 dynamics simulations with a restraining weight of 2 kcal/molÅ^2 at 300 K under an isobaric

23 condition. Finally, 10 ns MD simulations without any restraints were performed using the

24 same conditions. Long-range electrostatic interactions were applied using the Particle Mesh

25 Ewald method (PME)[40] during the simulations. The cut-off distance for the long-range van

26 der Waals interaction was set to 8 Å. The SHAKE method [41] was applied to constrain the

27 bond lengths of hydrogen atoms attached to heteroatoms. Coordinates and energy outputs

28 during MD simulations were recorded at 2 ps intervals.

29

30 *2.4 Binding free energy calculations*

31 The Molecular Mechanics/Poisson–Boltzmann Surface Area (MM-PBSA) method[42-45]

32 was employed for calculating the binding free energy of compounds **2**, **22** and **23** in InhA. In

33 this calculation, 250 snapshots of the complex, receptor and ligand were extracted every 8 ps

1   from the last nanosecond of the MD trajectory, which represents the equilibrium state. The

2   binding free energy (ΔG_{bind}) of compounds **2**, **22** and **23** complexed with InhA were

3   estimated from equation **1,** where ΔG_{vacuum} and ΔG_{solv} were the binding free energy of the

4   complex in vacuum and the solvation free energy, respectively. In the MM-PBSA approach,

5   the solvation free energy was calculated by solving a linearized Poisson-Boltzman equation.

6   ΔG_{vacuum} was obtained by calculating the interaction energy between InhA and compounds **2**,

7   **22** and **23** (ΔE_{MM}) and taking the entropy change (TΔS) as shown in equation **2**. ΔE_{MM} is

8   divided into three components, non-covalent van der Waals energy (ΔG_{vdW}), electrostatic

9   energy (ΔG_{ele}) and internal energy (ΔG_{int}), as shown in equation **3**. ΔE_{MM} and ΔG_{solv} were

10   calculated using the SANDER module and a PBSA program of the AMBER suite,

11   respectively. The entropy contribution was estimated using normal mode analysis with the

12   NMODE module.[46] The entropy contribution was estimated using 250 snapshots for the

13   binding free energy calculation.

14

15   $$\Delta G_{bind} = \Delta G_{vacuum} + \Delta G_{solv} \qquad\qquad \textbf{1}$$

16   $$\Delta G_{vacuum} = \Delta E_{MM} - T\Delta S \qquad\qquad \textbf{2}$$

17   $$\Delta E_{MM} = \Delta G_{vdW} + \Delta G_{ele} + \Delta G_{int} \qquad\qquad \textbf{3}$$

18

19   ***2.5 CoMFA and CoMSIA methods***

20   IC$_{50}$ and MIC$_{90}$ values of compounds were used to set up CoMFA[47] and CoMSIA[48]

21   models in order to evaluate the key structural features relating to the activity against both

22   InhA and *M. tuberculosis*. The predicted binding modes of training set compounds obtained

23   from molecular docking calculations were used for molecular alignment to set up CoMFA

24   and CoMSIA models. SYBYL 8.0 molecular modelling software was used to run CoMFA

25   and CoMSIA models. Partial least square (PLS) analysis was employed to derive a linear

26   relationship between CoMFA and CoMSIA descriptor fields and activities. The PLS analysis,

27   using the leave-one-out (LOO) cross-validation method, was performed to determine the

28   optimal number of components. Sequentially, a final analysis with the optimal number of

29   components was performed to construct CoMFA and CoMSIA models that were not

30   cross-validated. The non-cross-validated correlation coefficient (r^2) and the leave-one-out

31   cross-validated correlation coefficient (q^2) were used to evaluate the predictive ability of

32   CoMFA and CoMSIA models. Selected CoMFA and CoMSIA models were employed to

1   predict IC_{50} and MIC_{90} values of test set compounds that were not used to construct models.

2   This was done to evaluate the external predictive ability of these models.

3

4   **3. Results**

5   *3.1 Stability of the complex models*

6   To reveal the structural stability of simulation system, the RMSD values for the

7   position of all solute species were separately analyzed. The RMSD plots for the four

8   simulation systems over 10 ns are shown in **Figure 1**. Convergent RMSD plots indicated that

9   the equilibrium state was reached for each system during this simulation period. As shown,

10 the RMSDs for compounds **2**, **22**, **23** and **28** in InhA converged after approximately 2 ns.
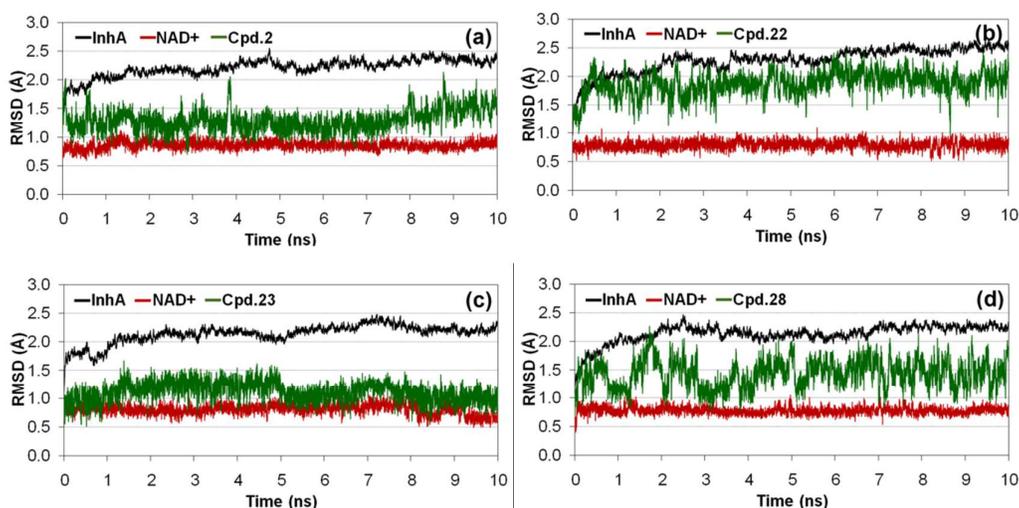
11

12
13

14 **Figure 1.** RMSD plots of compounds **2** (a), **22** (b), **23** (c), and **28** (d) complexed with InhA.

15

16 *3.2 Reliability of the calculation methods*

17   MD simulations were employed to model the binding modes of compounds **2**, **22**, **23**

18 and **28** in the InhA pocket. The experimental binding free energy (ΔGexp) lying within the

19 experimental error of the calculated values (ΔGbind) considered as the correlation between

20 the experimental binding free energy and the calculated values was used to indicate the

21 reliability of the modelled binding modes of these compounds. ΔGbind values of compounds

22 2, 22, 23 and 28 were close to their ΔGexp values (Table 2). Therefore, we concluded that

23 MD simulations reliably modelled binding modes of compounds 2, 22, 23 and 28 in the InhA

24 pocket.

25

1   **Table 2** ΔG_{bind} and ΔG_{exp} of compounds **2**, **22**, **23** and **28** in InhA (kcal/mol).

2

Cpd.	ΔH	$-T\Delta S$	ΔG_{bind}	ΔG_{exp}
2	-46.91±5.08	-31.03±6.06	-15.88±5.14	-15.52
22	-49.69±3.87	-33.15±6.41	-16.54±4.80	-15.82
23	-49.61±3.71	-32.79±5.57	-16.82±4.79	-15.65
28	-49.26±4.45	-32.52±6.58	-16.74±5.34	-16.07

3

4   *3.3 Binding mode of compound 28*

5   The binding mode of compound **28** complexed with InhA obtained from MD

6   simulations is shown in **Figure 2**. Residues located near each substituent and the core

7   structure are listed in **Figure 3**. A hydrogen atom (the R_1 substituent) is near the carbonyl

8   backbone of Met103. 2-pyridinyl methyl (the R_2 substituent) protrudes from the InhA pocket

9   and interacts with the solvent (**Figure 2**). The ethyl moiety (the R_3 substituent) is located near

10   backbones of Gly96, Phe97 and pyrophosphate and ribose groups of NAD^+. The ethyl group

11   (the R_4 substituent) was located in the hydrophobic side chains of Phe149, Tyr158, Met199

12   and nicotinamide of NAD^+. With regard to the core structure, the pyrazole ring in the core

13   structure was sandwiched between two hydrophobic side chains of Met161 and Ala198. CO

14   and NH of pyrazole amide formed hydrogen bonds with the backbones of Met98 and Ala198,

15   respectively. The benzofuran core was buried in the hydrophobic side chains of Ile215,

16   Ala157, Ile202 and Ala201, and was sandwiched between the hydrophobic side chains of

17   Leu207 and Met103. The carbonyl of benzofuran core formed a hydrogen bond with the NH

18   backbone of Ala201. NH of pyrrolidine amide formed a hydrogen bond with the CO

19   backbone of Leu197.
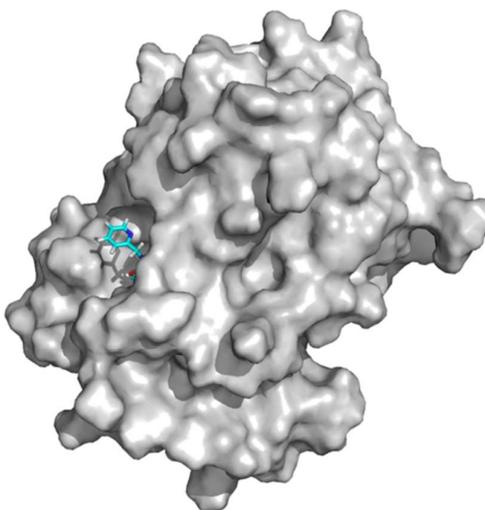
1
2 **Figure 2**.   Compound **28** (cyan) in its complex with whole InhA (grey) obtained from MD
3   simulations.
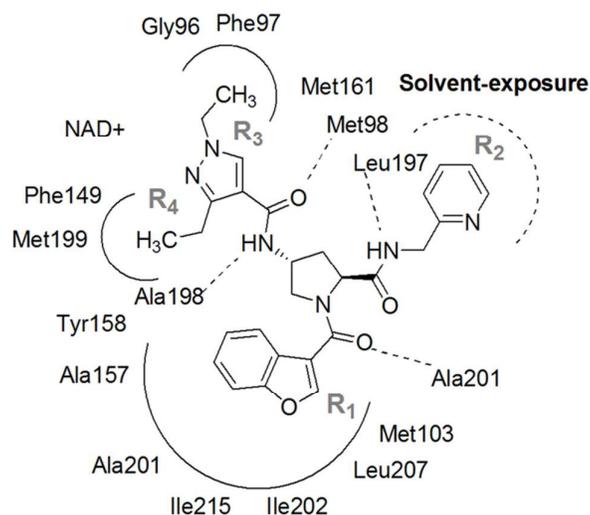
4
5 **Figure 3**.   List of residues surrounding within 4 Å from compound **28.**
6

7 *3.4 Interaction energy*

8   Free-energy decomposition calculations were used to investigate the interaction
9 energies between compound **28** and each residue in the InhA pocket. **Figure 4** shows these
10 interaction energies obtained from free-energy decomposition calculations. The lowest
11 interaction energy (-7.42 kcal/mol) was observed for Met103, indicating that this residue had
12 the largest contribution to binding of compound **28** in the InhA pocket. As previously
13 mentioned, Met103 and Leu207 were sandwiched in the benzofuran core. Another
14 remarkable interaction energy (-7.06 kcal/mol) was found for NAD^+. This was responsible

1   for van der Waal and electrostatic interactions with the R_3 and R_4 substituents of compound

2   **28** (**Figure 3**). Ala198 showed an interaction energy (-6.16 kcal/mol), comparable with those

3   of Met103 and NAD^+. This residue formed hydrogen bonds with the NH of pyrazole amide

4   and sandwiched the pyrazole ring (Figure 3). Met98, Leu197 and Ala201 formed other

5   hydrogen bonds with the core structure with interaction energies of -2.94, -3.27 and -5.33

6   kcal/mol, respectively. Based on interaction energy profile of compound **28**, the core

7   structure formed more attractive interactive energies with surrounding residues than **R**

8   substituents (**Figure 4**). This result indicates that the core structure is the key fragment for

9   binding of this compound in the InhA pocket.
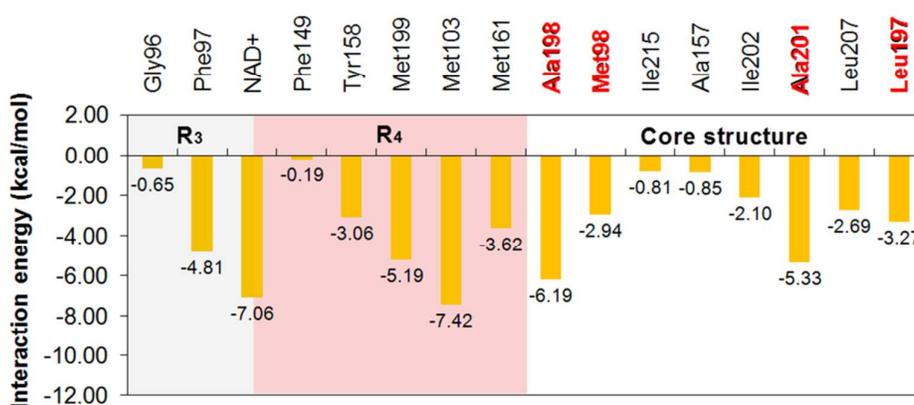
10

11 **Figure 4.**   Interaction energy profile of compound **28** and surrounding residues within 4 Å.

12

13 *3.5 The effect of the R_2 substituent on IC_{50} and MIC_{90} values*

14 As compared with the positions of other **R** substituents, the R_2 position had the most

15 varied substituents (**Table 1**). Compound **28** exposing the 2-pyridylmethyl at the R_2 position

16 showed the best activity for InhA inhibition with an IC_{50} of 0.002 μM. When the R_2

17 substituent of this compound was replaced by CH_2COOMe (compound **22**), the IC_{50} value

18 was slightly changed to 0.003 μM.   In contrast, the MIC_{90} value against whole

19 *M. tuberculosis* cell was greatly changed from 0.7 μM to 0.05 μM (**Table 1**). To reveal the

20 effect of the R_2 substituent on the IC_{50} value, the binding modes of compounds **28** and **22**

21 were compared (**Figure 5**). The binding modes of these compounds in the InhA pocket were

22 similar, and the R_2 substituents occupied in the same positions. Moreover, the interaction

23 energy profiles of compounds **28** and **22** with residues in InhA pocket were similar (Figure

24 6). As discussed above, the R_2 substituent of compound **28** protruded from the InhA pocket

25 leading to weak interaction of this substituent with the pocket. Therefore, the IC_{50} value

against InhA was not significantly changed when the R_2 substituent was varied. When the R_2
substituent was replaced by a hydrogen atom (compound **23**), the binding mode and
interaction energy profile of this compound were similar to those of compounds **22** and **28**
(**Figures 5** and **6**). With regard to IC_{50} values, compound **23** showed a comparable IC_{50} value
with those of compounds **22** and **28**. However, the MIC_{90} value of this compound (0.5 µM)
was largely increased over that that of compound **22** (0.05 µM). These results indicate that
the R_2 substituent had a small effect on the IC_{50} value against InhA due to its weak
interaction with the InhA pocket. Alternatively, this substituent is crucial to controlling the
MIC_{90} against intact *M. tuberculosis* cells.
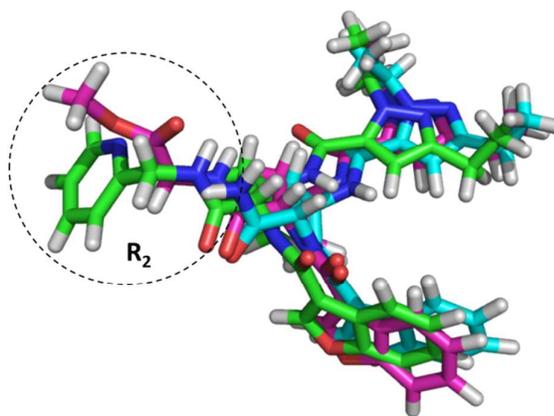
Figure 5. The superimposition of binding modes of compounds **22** (pink), **23** (cyan) and **28** (green).
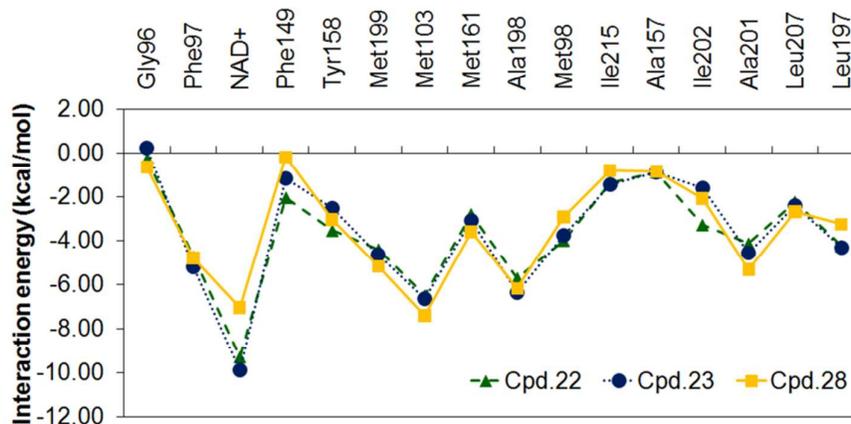
Figure 6. Comparison of the interaction energy profiles of compounds **22** (green), **23** (blue) and **28** (yellow) with surrounding pocket within 4 Å.

1 *3.6 The effect of the R_3 substituent on IC_{50} and MIC_{90} values*

2 The R_3 substituent of compounds in the data set was varied as ethyl (Et) or methyl

3 (Me) groups (**Table 1**). Compounds **2** and **22** with structural differences at the R_3 substituent

4 were selected to show the effect of the R_3 substituent on IC_{50} and MIC_{90} values. IC_{50} values

5 of these compounds (0.005 and 0.003 μM, respectively) were not significant different, but

6 their MIC_{90} values were tenfold different (0.5 and 0.05 μM, respectively). **Figure 7** shows

7 the binding modes of compounds **2** and **22** in InhA obtained from MD simulations. The R_3

8 substituents of these compounds were located in the same position and surrounded by

9 backbones of Gly96, Phe97 as well as pyrophosphate and ribose groups of NAD^+. The ethyl

10 group (The R_3 substituent) of compound **22** is close to Phe97 and pyrophosphate and ribose

11 groups of NAD^+ more than the methyl group of compound **2**. Therefore, interaction energies

12 of compound **22** with Phe97 and NAD^+ had greater attraction than those of compound **2**

13 (**Figure 8**). Moreover, the presence of a methyl group at the R_3 position of compound **2**

14 shifted the position of benzofuran core surrounded by Met103 and Ile202, and disrupted

15 hydrogen bond interaction with Met98. Accordingly, interaction energies of compound **2**

16 with Met98, Met103 and Ile202 showed less attraction than those of compound **22** (**Figure**

17 **8**). These results indicate that compound **22** should have a better IC_{50} against InhA compared

18 to compound **2**. However, other than the interaction energies of Met98, Met103, Ile202,

19 Phe97 and NAD^+, compounds **2** and **22** are comparable. The IC_{50} value for InhA inhibition

20 by compound **22** was slightly better than that of compound **2**. However, its MIC_{90} value was

21 tenfold better than that of compound **2**. The results indicated that the ethyl group at the R_3

22 position is more conducive to favorable IC_{50} and MIC_{90} values than the methyl group.
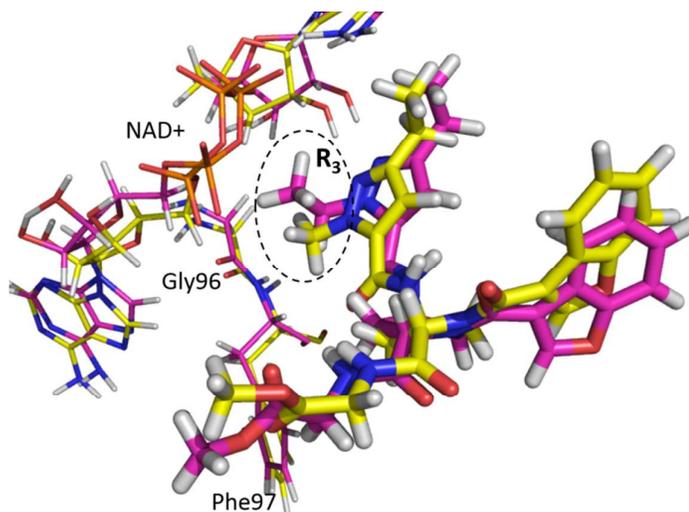
23

24 **Figure 7.** The superimposition of binding modes of compounds **2**(yellow) and **22** (pink).
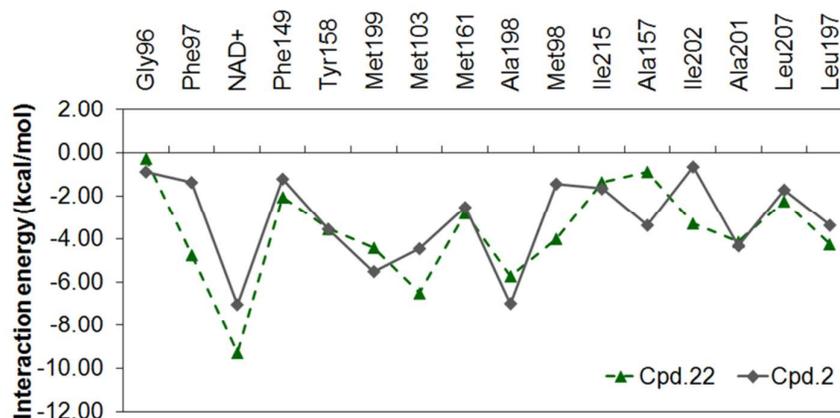
Figure 8. Comparison of the interaction energy profiles of compounds **2** (gray) and **22** (green) with surrounding pocket within 4 Å.

3.7 CoMFA and CoMSIA models

In this study, CoMFA and CoMSIA models were constructed from IC_{50} and MIC_{90} where prefixed with IC_{50} and MIC_{90}, respectively. IC_{50} and MIC_{90} CoMSIA models were constructed based on various combinations of molecular descriptor fields, in order to develop a highly predictive CoMSIA model (**Tables 3** and **4**). An IC_{50} CoMSIA model constructed from the combination of steric (S), electrostatic (E), hydrophobic (H) and hydrogen acceptor (A) fields [48] gave the highest q^2 (0.646), whereas an MIC_{90} CoMSIA model including steric, electrostatic, hydrophobic and hydrogen donor (D) fields [48] showed the highest q^2 (0.639). Therefore, these models were selected for graphical interpretation of IC_{50} and MIC_{90} CoMSIA contour maps. In order to assess the predictive abilities of IC_{50} and MIC_{90} CoMSIA models, IC_{50} and MIC_{90} values of the test set were predicted. Both IC_{50} and MIC_{90} CoMSIA models showed good ability to predict IC_{50} and MIC_{90} values of the test set data as shown in **Figure 9**. In case of IC_{50} and MIC_{90} CoMFA models, they had poor predictive ability with q^2 values of 0.464 and 0.432, respectively. Accordingly, these CoMFA models were not used further in this work.

15

1   **Table 3** Statistical results of IC_{50} CoMFA and CoMSIA models.

2

Models	Statistical parameters						Fraction
	q^2	r^2	s	SEE	N	F	
CoMFA							
S/E	0.464	0.996	0.392	0.035	6	909.618	60.3/39.7
CoMSIA							
S/E	0.084	0.977	0.512	0.081	6	162.845	32.1/67.9
S/H	0.465	0.950	0.383	0.118	5	90.431	29.1/70.9
S/D	0.624	0.923	0.321	0.145	5	57.579	54.3/45.7
S/A	0.146	0.970	0.495	0.093	6	123.724	39.7/60.3
S/E/H	0.260	0.981	0.460	0.074	6	194.704	16.6/44.5/38.9
S/E/D	0.592	0.980	0.342	0.076	6	185.576	21.0/53.7/25.3
S/E/A	0.281	0.975	0.454	0.085	6	149.701	22.5/42.8/34.7
S/E/H/D	**0.646**	**0.990**	**0.318**	**0.055**	**6**	**363.962**	**13.1/35.8/28.5/22.6**
S/E/H/A	0.336	0.983	0.436	0.070	6	222.520	12.3/31.5/29.4/26.8
S/E/H/D/A	0.610	0.991	0.334	0.050	6	437.341	10.0/25.4/22.6/20.7/21.4

3

4   Bold values indicate the best CoMSIA model

5   *N* optimum number of components; *s* standard error of prediction; *SEE* standard error of

6   estimate; *F* F-test value; *S* steric field; *E* electrostatic field; *H* hydrophobic field; *D* hydrogen

7   donor field; *A* hydrogen acceptor field

8

9

10

11

12

13

14

15

16

17

18

1   **Table 4** Statistical results of MIC$_{90}$ CoMFA and CoMSIA models.

2

Models	Statistical parameters						Fraction
	q^2	r^2	s	SEE	N	F	
CoMFA							
S/E	0.432	0.853	0.442	0.225	2	78.451	53.2/46.8
CoMSIA							
S/E	0.456	0.949	0.469	0.143	6	71.455	25.1/74.9
S/H	0.459	0.780	0.432	0.275	2	47.970	34.4/65.6
S/D	0.261	0.732	0.514	0.310	3	23.642	52.7/47.3
S/A	0.602	0.978	0.401	0.093	6	174.060	46.3/53.7
S/E/H	0.477	0.961	0.460	0.126	6	93.558	13.8/52.8/33.4
S/E/D	0.210	0.912	0.553	0.184	5	49.990	17.7/64.4/18.0
S/E/A	0.550	0.955	0.426	0.134	6	82.091	19.9/48.1/32.0
S/E/H/D	0.415	0.938	0.476	0.155	5	72.712	10.9/45.8/29.3/13.9
S/E/H/A	**0.639**	**0.973**	**0.382**	**0.105**	**6**	**136.014**	**12.5/35.6/42.2/27.7**
S/E/H/D/A	0.494	0.961	0.442	0.123	5	118.951	9.3/33.4/22.8/10.4/24.2

3

4   Bold values indicate the best CoMSIA model

5   *N* optimum number of components; *s* standard error of prediction; *SEE* standard error of

6   estimate; *F* F-test value; *S* steric field; *E* electrostatic field; *H* hydrophobic field; *D* hydrogen

7   donor field; *A* hydrogen acceptor field

8
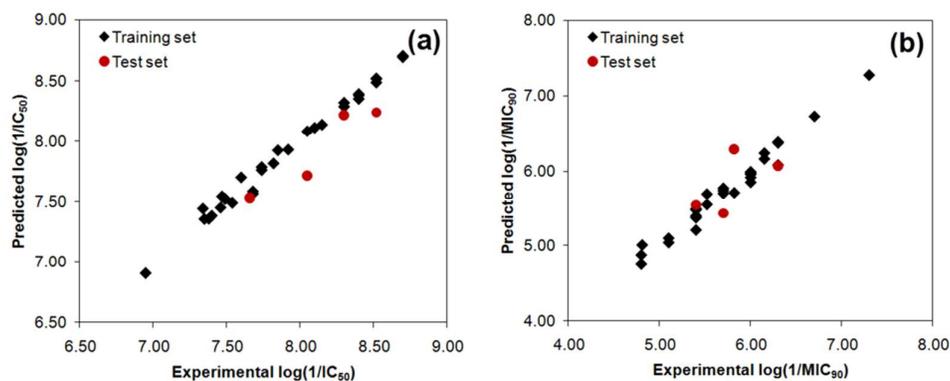
9

10   **Figure 9.**   The plot of experimental and predicted activities of the training and test data sets

11   derived from IC$_{50}$ (a) and MIC$_{90}$ (b) CoMSIA models.

17

1

3.8 CoMSIA contour maps

     To reveal the importance of molecular descriptor fields in both IC_{50} and MIC_{90} values of InhA inhibitors, CoMSIA contour maps were established. Compound **22** presented the best MIC value. Graphical interpretation of its IC_{50} and MIC_{90} CoMSIA contour maps was done. Interpretation of its IC_{50} and MIC_{90} CoMSIA contour maps revealed structural requirements in terms of steric, electrostatic, hydrophobic and hydrogen donor and acceptor fields for IC_{50} and MIC_{90} values of InhA inhibitors.

3.9 Steric requirements for IC_{50} and MIC_{90} values

     Figure 10 shows the CoMSIA steric contour maps obtained from selected IC_{50} and MIC_{90} CoMSIA models. These contours highlight the steric requirements for IC_{50} and MIC_{99} values of benzofuran pyrrolidine pyrazole derivatives. Both IC_{50} and MIC_{90} CoMSIA models show a green contour at the **R$_3$** substituent. These results indicated that a bulky **R$_3$** substituent is favourable for both IC_{50} and MIC_{90} values. Accordingly, an ethyl group is more preferred for the steric requirement of the **R$_3$** substituent than a methyl group. This is consistent with the MD simulations since an ethyl group can form more interactions with InhA. At the **R$_2$** position, IC_{50} and MIC_{90} CoMSIA models present a large yellow contour. However, IC_{50} CoMSIA model shows a favorable green steric contour at the terminal of the **R$_2$** substituent (**Figure 10a**). Based on MD simulations results, the **R$_2$** substituent had weak interaction with the InhA pocket leading to less influence on the IC_{50} value. Therefore, the steric requirement of **R$_2$** substituent should be based on the MIC_{90} CoMSIA steric contour that presented only a yellow contour near this substituent (**Figure 10b**).
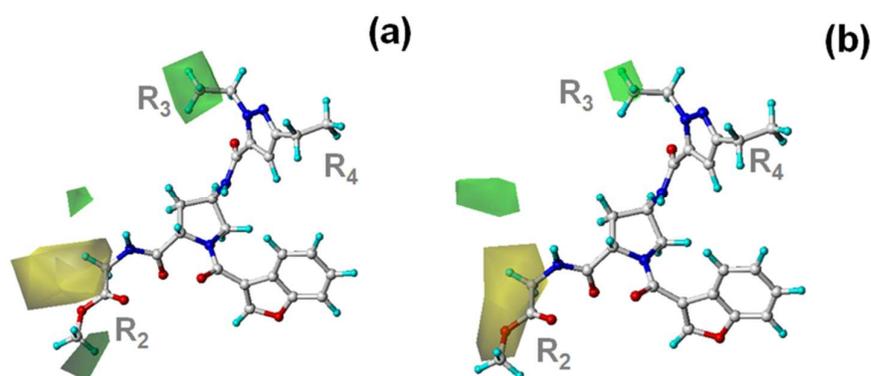
Figure 10. Steric contour maps of IC_{50} (a) and MIC_{90} (b) CoMSIA models in combination with compound **22.**

1   *3.10 Electrostatic requirements for IC_{50} and MIC_{90} values*

2   Electrostatic requirements for IC_{50} and MIC_{90} values of benzofuran pyrrolidine

3   pyrazole derivatives are visualized in **Figure 11**. Both IC_{50} and MIC_{90} CoMSIA contours

4   show only an electrostatic requirement at the **R_2** substituent. The IC_{50} CoMSIA shows a red

5   contour at the ester moiety of **R_2** substituent, whereas MIC_{90} CoMSIA presents a blue

6   contour at this position. These results show different electrostatic requirements for IC_{50} and

7   MIC_{90} values of benzofuran pyrrolidin pyrazole derivatives. However, the **R_2** substituent has

8   weak influence on the IC_{50} value. Therefore, the electrostatic requirement of **R_2** substituent

9   for MIC_{90} values should take more priority.

10
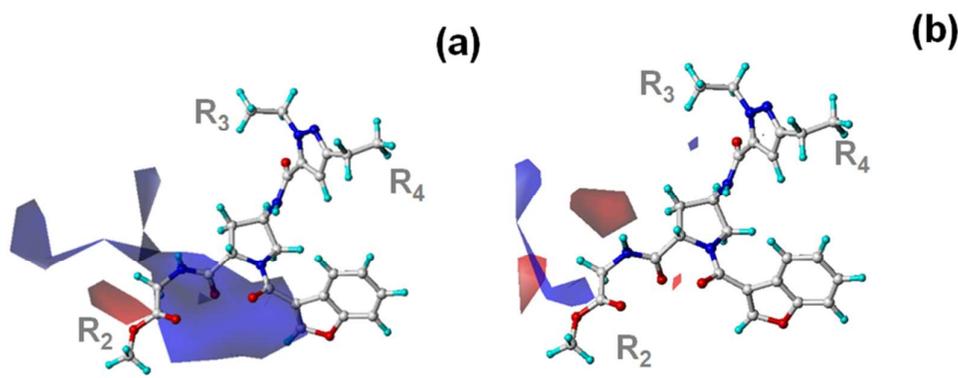
11   **Figure 11.** Electrostatic contour maps of IC_{50} (a) and MIC_{90} (b) CoMSIA models in

12   combination with compound **22.**

13

14   *3.11 Hydrophobic requirements for IC_{50} and MIC_{90} values*

15   Both IC_{50} and MIC_{90} CoMSIA contours show a purple contour at the **R_3** substituent of

16   compound **22** (**Figure 12**). This shows that the hydrophobic requirements of the **R_3**

17   substituent for both IC_{50} and MIC values were similar. The **R_3** substituent was either a methyl

18   or ethyl group. As seen in **Figure 12**, the terminal of ethyl group was buried in a purple **R_3**

19   contour. Therefore, the ethyl group was preferable for the hydrophobic requirement of the

20   substituent. IC_{50} and MIC_{90} values of compound **2** with the methyl group at the **R_3** substituent

21   were weaker than those of compound **22** containing an ethyl group. At the **R_2** substituent,

22   both IC_{50} and MIC_{90} CoMSIA contours display a purple contour at this position (**Figure 12**).

23   Therefore, the presence of a hydrophobic substituent at this purple region should enhance

24   both IC_{50} and MIC_{90} values. The grey contour located at the carbonyl moiety of the **R_2**

25   substituent in both IC_{50} and MIC_{90} CoMSIA contours indicated that this moiety is important

26   for both IC_{50} and MIC_{90} values. Another important hydrophobic contour is located at the **R_4**

27   substituent. The MIC_{90} CoMSIA shows a purple region near the **R_4** substituent (**Figure 12b**),

1   but this contour disappeared in the IC_{50} CoMSIA contour (**Figure 12a**). Therefore, a

2   hydrophobic moiety could be presented at purple region to enhance the MIC_{90} value without

3   a negative contribution to the IC_{50} value.
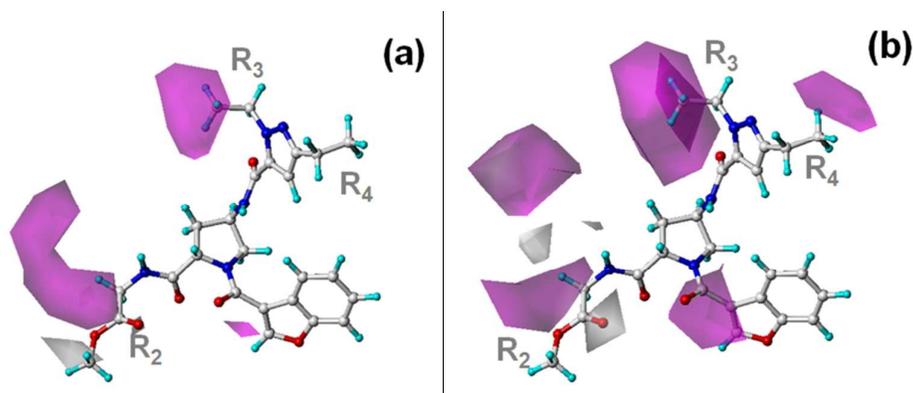
4

5   **Figure 12.** Hydrophobic contour maps of IC_{50} (a) and MIC_{90} (b) CoMSIA models in

6   combination with compound **22.**

7

8   *3.12 Hydrogen donor and acceptor requirements for IC_{50} and MIC_{90} values*

9   The hydrogen donor field was included in the selected IC_{50} CoMSIA model, but this

10   molecular descriptor was instead changed to a hydrogen acceptor field in the selected MIC_{90}

11   CoMSIA model. The IC_{50} CoMSIA model did not show any hydrogen donor contour near

12   any **R** substituents. However, this model showed a favourable hydrogen donor contour at the

13   amide moiety of the core structure. The amide moiety appears to impact the IC_{50} value.

14   Consistent with the MD simulations results, this moiety can form hydrogen bonds with

15   Leu197. The MIC_{90} CoMSIA model shows a favourable hydrogen acceptor contour at the

16   carbonyl moiety of **R$_2$** substituent, indicating that this moiety is essential to a good MIC_{90}

17   value.
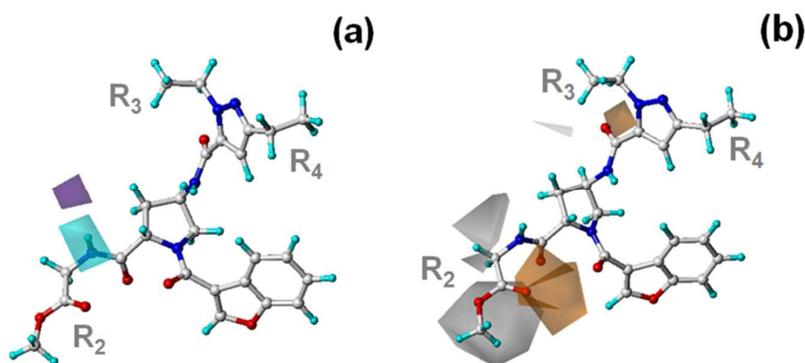
18

19   **Figure 13.** Hydrogen donor contour of IC_{50} CoMSIA model (a) and hydrogen acceptor

20   contour MIC_{90} CoMSIA model (b) in combination with compound **22.**

1

2   *3.13 The structural concept for good IC$_{50}$ and MIC$_{90}$ correlation*

3   Based on the MD simulations results, the core structure of benzofuran pyrrolidine

4   pyrazole derivatives is of key importance for binding in the InhA pocket. Therefore, this

5   fragment is crucial for favorable IC$_{50}$ values. Among all **R** substituents, the **R$_2$** substituent has

6   the least interaction with the InhA pocket because it protrudes from the pocket. Modifications

7   of the **R$_2$** substituent did not significantly change IC$_{50}$ values, but rather produced a tenfold

8   increase in MIC$_{90}$ values (compounds **22** and **23**). Accordingly, the **R$_2$** substituent is a key

9   group that can be used to adjust the MIC$_{90}$ value without negative contribution to the IC$_{50}$

10   value. Based on the results obtained from our MD simulations and CoMSIA studies, the

11   structural concept to correctly balance IC$_{50}$ and MIC$_{90}$ values of benzofuran pyrrolidin

12   pyrazole derivatives is summarized in **Figure 14**. New compounds designed based on this

13   concept should show better IC$_{50}$ and MIC$_{90}$ values.
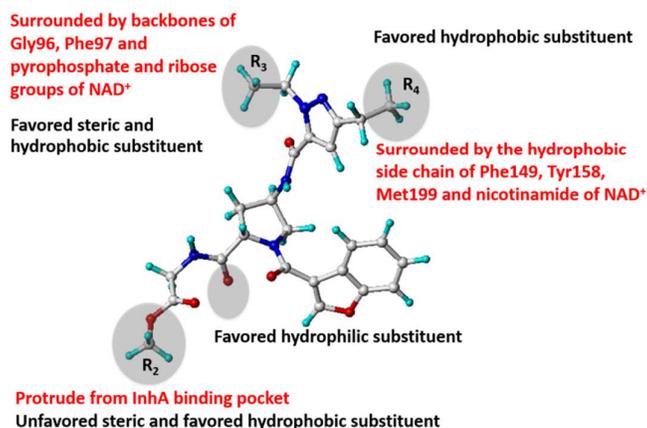
14

15   **Figure 14.** The structural concept for good IC$_{50}$ and MIC$_{90}$ correlation summarized from MD

16   simulations and CoMSIA results. Red and black letters indicate the results

17   obtained from MD simulations and CoMSIA results, respectively.

18

19   **4. Conclusion**

20   The combination of MD simulations and graphical interpretation of IC$_{50}$ and MIC$_{90}$ CoMSIA

21   models highlight the structural concept to correctly balance IC$_{50}$ and MIC$_{90}$ values of

22   benzofuran pyrrolidin pyrazole derivatives. The core structure of template compound is

23   crucial to attaining favorable IC$_{50}$ values, whereas the **R$_2$** substituent is a key group to

24   enhance MIC$_{90}$ values without negative effects on IC$_{50}$ values. Modifications of **R**

25   substituents following the structural concept suggested here should allow design of novel

1   InhA inhibitors with better potency against both the InhA enzyme and intact *M. tuberculosis*

2   cells.

3

4   **Acknowledgements**

10

11   **References**

12   1.   Global Tuberculosis Report; World Health Organization: Geneva, 2014.

13   2.   M. C. Becerra, J. Bayona, J. Freeman, P. E. Farmer, J. Y. Kim, *Int. J. Tuberc. Lung Dis*.,

14   2000, **4**, 387−394.

15   3.   C. Dye, M. A. Espinal, C. J. Watt, C. Mbiaga, B. G. Williams, *J. Infect. Dis*., 2002, **185**,

16   1197−1202.

17   4.   M. A. Espinal, *Tuberculosis (Edinburgh)*, 2003, **83**, 44−51.

18   5.   A. Wright, M. Zignol, A. Van Deun, D. Falzon, S. R. Gerdes, K. Feldman, S. Hoffner, F.

19   Drobniewski, L. Barrera, D. van Soolingen, F. Boulabhal, C. N. Paramasivan, K. M.

20   Kam, S. Mitarai, P. Nunn, M.Raviglione, *Lancet*, 2009, **373**, 1861−1873.

21   6.   C. Y. Chiang, R. Centis, G. B. Migliori, *Respirology*, 2010, **15**, 413−432.

22   7.   Multidrug and Extensively Drug-Resistant TB (M/XDR-TB):2010 Global Report on

23   Surveillance and Response (WHO/HTM/TB/2010.3). World Health Organization:

24   Geneva, 2010.

25   8.   N. S. Shah, A. Wright, G. H. Bai, L. Barrera, F. Boulahbal, N. Martin-Casabona, F.

26   Drobniewski, C. Gilpin, M. Havelkova, R. Lepe, R. Lumb, B. Metchock, F. Portaels, M.

27   F. Rodrigues, S. RüschGerdes, A. Van Deun, V. Vincent, K. Laserson, C. Wells, J. P.

28   Cegielski, *Emerging Infect. Dis*., 2007, **13**, 380−387.

29   9.   M. Berry, O. M. Kon, *Eur. Respir. Rev*., 2009, **18**, 195−197.

30   10. A. A. Velayati, P. Farnia, M. R. Masjedi, T. A. Ibrahim, P. Tabarsi, R. Z. Haroun, H. O.

31   Kuan, J. Ghanavi, M. Varahram, *Eur. Respir. J*., 2009, **34**, 1202−1203.

32   11. Z. F. Udwadia, R. A. Amale, K. K. Ajbani, C. Rodrigues, *Clin. Infect. Dis*., 2012, **54**,

33   579−581.

12. A. Quemard, J. C. Sacchettini, A. Dessen, C. Vilcheze, R. Bittman,ˋ W. R. Jacobs, J. S. Blanchard, *Biochemistry*, 1995, **34**, 8235−8241.

13. C. Vilcheze, H. R. Morbidoni, T. R. Weisbrod, H. Iwamoto, M. Kuo, J. C. Sacchettini, W. R. Jacobs Jr., *J. Bacteriol.*, 2000, **182**, 4059−4067.

14. D.A. Rozwarski, C. Vilchèze, M. Sugantino, R. Bittman, J.C. Sacchettini, *J. Biol. Chem.*, 1999, **274**, 15582–15589.

15. D.A. Rozwarski, G.A. Grant, D.H. Barton, W.R. Jacobs Jr., J.C. Sacchettini, *Science*, 1998, **279**, 98–102.

16. C. Vilcheze, F. Wang, M. Arai, M.H. Hazbon, R. Colangeli, L. Kremer, T.R. Weisbrod, D. Alland, J.C. Sacchettini, W.R. Jacobs Jr., *Nat. Med.*, 2006, **12**, 1027–1029.

17. A. Dessen, A. Quemard, J.S. Blanchard, W.R. Jacobs Jr., J.C. Sacchettini, *Science*, 1995, **267**, 1638–1641.

18. K. Johnsson, P.G. Schultz, *J. Am. Chem. Soc.*, 1994, **116**, 7425–7426.

19. B. Lei, C.J. Wei, S.C. Tu, *J. Biol. Chem.*, 2000, **275**, 2520–2526.

20. K. Johnsson, D.S. King, P.G. Schultz, *J. Am. Chem. Soc.*, 1995, **117**, 5009–5010.

21. Y. Zhang, B. Heym, B. Allen, D. Young, S; Cole, *Nature*, 1992, **358**, 591–593.

22. A. Quemard, A. Dessen, M. Sugantino, W.R. Jacobs Jr., J.C. Sacchetini, J.S. Blanchard, *J. Am. Chem. Soc.*, 1996, **118**, 1561–1562.

23. A. Banerjee, E. Dubnau, A. Quemard, V. Balasubramanian, K.S. Um, T. Wilson, D. Collins, G. de Lisle, W.R. Jacobs Jr., *Science*, 1994, **263**, 227–230.

24. L. Encinas, H. O'Keefe, M. Neu, M.J. Remuiñán, A.M. Patel, A. Guardia, C.P. Davie, N. Pérez-Macías, H. Yang, M.A. Convery, J.A. Messer, E. Pérez-Herrán, P.A. Centrella, D. Alvarez-Gómez, M.A. Clark, S. Huss, G.K. O'Donovan, F. Ortega-Muro, W. McDowell, P. Castañeda, C.C. Arico-Muendel, S. Pajk, J. Rullás, I. Angulo-Barturen, E. Alvarez-Ruíz, A. Mendoza-Losan, L. Ballell Pages, J. Castro-Pichel, G. Evindar, *J. Med. Chem.*, 2014, **57**, 1276-1288.

25. M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. Hratchian, A. Izmaylov, J. Bloino, G. Zheng, J. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. Peralta, F. Ogliaro, M. Bearpark, J. Heyd, E. Brothers, K. Kudin, V. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. Burant, S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. Millam, M. Klene, J. Knox, J. Cross, V.

Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. Stratmann, O. Yazyev, A. Austin, R. Cammi, C. Pomelli, J. Ochterski, R. Martin, K. Morokuma, V. Zakrzewski, G. Voth, P. Salvador, J. Dannenberg, S. Dapprich, A. Daniels, Farkas, J. Foresman, J. Ortiz, J. Cioslowski and D. Fox, Gaussian 09, Gaussian, Inc., Wallingford CT, 2009.

26. G. Jones, P. Willett, R.C. Glen, *J. Mol. Biol.*, 1995, **245**, 43-53.

27. G. Jones, P. Willett, R.C. Glen, A.R. Leach, R. Taylor, *J. Mol. Biol.*, 1997, **267**, 727-748.

28. J.W. Nissink, C. Murray, M. Hartshorn, M.L. Verdonk, J.C. Cole, R. Taylor, *Proteins*, 2002, **49**, 457-471.

29. M.L. Verdonk, J.C. Cole, M.J. Hartshorn, C.W. Murray, R.D. Taylor, *Proteins*, 2003, **52**, 609-623.

30. M.L. Verdonk, G. Chessari, J.C. Cole, M.J. Hartshorn, C.W. Murray, J.W. Nissink, R.D. Taylor, R. Taylor, *J. Med. Chem.*, 2005, **48**, 6504-6515.

31. D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P.A. Kollman, AMBER 12, University of California, San Francisco.

32. Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P.A. Kollman, *J. Comput. Chem.*, 2003, **24**, 1999–2012.

33. J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.

34. J. Wang, W. Wang, P. A. Kollman, D.A. Case, *J. Mol. Graphics Model.*, 2006, **25**, 247-260.

35. P. Cieplak, W.D. Cornell, C. Bayly, P.A. Kollman, *J. Comput. Chem.*, 1995, **16**, 1357–1377.

36. W.D. Cornell, P. Cieplak, C.I. Bayly, P.A. Kollmann, *J. Am. Chem. Soc.*, 1993, **115**, 9620–9631.

37. C.I. Bayly, P. Cieplak, W. Cornell, P.A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.

38. J. Wang, P. Cieplak, P.A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049–1074.

1   39. M.W. Mahoney, W.L. Jorgensen, *J. Chem. Phys*., 2000, **112**, 8910–8922.

2   40. T. Darden, D. York, L. Pedersen, *J. Chem. Phys*., 1993, **98**, 10089–10092.

3   41. J.P. Ryckaert, G. Ciccotti, H.J.C. Berendsen, *J. Comp. Phys.,* 1977, **23**, 327–341.

4   42. N. Homeyer, Gohlke, H. *Mol. Inf*., 2012, **31**, 114–122.

5   43. J. Wang, T. Hou, X. Xu, *Curr. Comput. Aided Drug Des*., 2006, **2**, 95–103.

6   44. J. Wang, P. Morin, W. Wang, P.A. Kollman, *J. Am. Chem. Soc*., 2001, **123**, 5221–5230.

7   45. T. Hou, J. Wang, Y. Li, W. Wang, *J. Chem. Inf. Model*., 2011, **51**, 69–82.

8   46. M. Kaledin, A. Brown, A.L. Kaledin, J.M. Bowman, *J. Chem. Phys*., 2004, **121**, 5646–
9   5653.

10 47. R. D. Cramer, III, D. E. Patterson, J. D. Bunce, *J. Am. Chem. Soc*., 1988, **110**, 5959
11   5967.

12 48. G. Klebe , U. Abraham , T. Mietzner, *J. Med. Chem*., 1994, **37**, 4130–4146.