ORIGINAL PAPER

Insight into crucial inhibitor–enzyme interaction of arylamides as novel direct inhibitors of the enoyl ACP reductase (InhA) from *Mycobacterium tuberculosis*: computer-aided molecular design

Auradee Punkvang · Patchreenart Saparpakorn · Supa Hannongbua · Peter Wolschann · Heinz Berner · Pornpan Pungpo

Abstract The enoyl ACP reductase enzyme (InhA) involved in the type II fatty acid biosynthesis pathway of *Mycobacterium tuberculosis* is an attractive target enzyme for antitubercular drug development. Arylamide derivatives are a novel class of InhA inhibitors used to overcome the drug-resistance problem of isoniazid, the frontline drug for tuberculosis treatment. Their remarkable property of inhibiting the InhA enzyme directly without requiring any coenzyme, makes them especially appropriate for the design of new antibacterials. In order to find a sound binding conformation for the different arylamide analogs, molecular docking experiments were performed with subsequent QSAR investigations. The X-ray conformation of one arylamide within its cocrystallized complex with InhA was used as a starting conformation for the docking experiments. The results thus obtained are perfectly consistent (rmsd = 0.73 Å) with the results from X-ray analysis. A thorough investigation of the arylamide binding modes with InhA provided ample information about structural requirements for appropriate inhibitor–enzyme interactions. Three different QSAR models were established using two three-dimensional (CoMFA and CoMSIA) and one two-dimensional (HQSAR) techniques. With statistically ensured models, the QSAR results obtained had high correlation coefficients between molecular structure properties of 28 arylamide derivatives and their biological activity. Molecular fragment contributions to the biological activity of arylamides could be obtained from the HQSAR model. Finally, a graphic interpretation designed in different contour maps provided coincident information about the ligand–receptor interaction thus offering guidelines for syntheses of novel analogs with enhanced biological activity.

Keywords Molecular modeling · Molecular docking · CoMFA · CoMSIA · HQSAR

Introduction

Tuberculosis (TB) caused by *Mycobacterium tuberculosis* (*M. tuberculosis*) remains a serious public health problem. Globally, there were estimated 9.27 million incident cases of TB in 2007. This is an increase from 9.24 million cases in 2006, compared with 8.3 million cases in 2000 and 6.6 million cases in 1990 [1]. Multidrug resistant tuberculosis (MDR-TB), widespread extensive drug-resistance TB (XDR-TB), and co-infection between *M. tuberculosis* and HIV (TB/HIV) are rendering tuberculosis treatment complicated, and cause severe financial strain [2–11]. Accordingly, to address these problems the design of novel and more potent antitubercular agents has to be regarded as highly important.

The enzymes involved in the bacterial fatty acid biosynthetic pathway, the type II fatty acid synthase (FAS II) system, are attractive targets for designing novel antibacterial agents and improving existing antibacterial agents [12–17]. Bacterial FAS-II organization is distinct from its mammalian counterpart; thus, the FAS-II pathway offers

A. Punkvang · P. Pungpo (✉)
Faculty of Science, Ubonratchathani University,
Ubon Ratchathani, Thailand
e-mail: pornpan_ubu@yahoo.com

P. Saparpakorn · S. Hannongbua
Faculty of Science, Kasetsart University, Bangkok, Thailand

P. Wolschann · H. Berner
Institute for Theoretical Chemistry, University of Vienna,
Vienna, Austria

several targets for potential selective interactions with antibacterial agents [15]. One of these targets in the FAS-II pathway is enoyl-acyl ACP reductase (InhA). This enzyme, involved in mycolic acid biosynthesis, catalyzes the NADH-specific reduction of 2-*trans*-enoyl-ACP, which is essential for fatty acid elongation [18]. Subsequently, InhA has been identified as the primary target of isoniazid (INH), the frontline drug for tuberculosis chemotherapy [19–27]. Being a prodrug, INH does not inhibit InhA enzyme directly. It must first be activated by catalase-peroxidase (KatG) to generate the reactive acyl radical [28–34]. Thereupon, the reactive species binds covalently to nicotinamide adenine dinucleotide (NAD+) to form the active adduct (INH-NAD adduct) that functions as a highly potent inhibitor of InhA [18–20, 23, 30, 35, 36]. However, the high potency of INH for tuberculosis treatment is diminished by drug resistance. High levels of resistance to INH are caused by mutations in KatG, commonly found in *M. tuberculosis* clinical isolates [37]. The predominant mutation of INH-resistant strains, KatG[S315G], has been found and examined in detail [38]. These results reveal that catalase and peroxidase activities of KatG were moderately reduced (50 and 35%, respectively). The INH binding affinity for the residuary enzyme was unchanged, whereas INH activation was reduced by 30% compared with WT KatG, which would be expected to correlate with a moderate increase of the minimum inhibitory concentrations (MIC) of INH. In order to reduce the resistance against INH associated with mutations in the KatG enzyme, compounds which directly inhibit the InhA enzyme without requiring activation by KatG are seen as very promising new agents against tuberculosis [39–43]. A series of arylamides have already been identified as a novel class of potent InhA inhibitors [44]. In addition some crystal structures of InhA–arylamide inhibitor complexes incubated with NADH are already available [44]. On the other hand it must be taken into account that most arylamides have high MIC values only against *M. tuberculosis* strain H37Rv. However, it can be reasonably assumed that these compounds are extruded from the bacterial cell by efflux pumps.

This information, especially the remarkable property of arylamides which directly inhibit the InhA enzyme—possibly the crucial point for new therapy—justifies more detailed examination of their structural requirements for suitable therapeutic activity against tuberculosis.

Molecular modeling and computer-aided molecular design approaches are powerful tools for developing new and more potent InhA inhibitors. Recently, structure-based and ligand-based approaches to drug design have been used to identify important features of InhA inhibitors [45–48]. In this study, molecular docking calculations were used to investigate the important drug–enzyme interactions of

arylamides in the InhA binding pocket. Additionally, the relationship between the structure and activity of these compounds was elucidated by the CoMFA, CoMSIA, and HQSAR methods [49–52] giving us new and more detailed information for the design of highly active antibacterials.

Results and discussion

Structural comparison of ligand-bound and ligand-free InhA

The X-ray structures of arylamide-bound (pdb code 2NSD) and arylamide-free (pdb code 1ENY) InhA were used for comparison. Because of this binding, the two structures are different within the region of residues Leu197–Arg225 including two α-helixes and one loop labeled in red in Fig. 1. NADH in two complexes was held at the same position implying that ligand binding in these cases has no effect on the binding of NADH cofactor. The effect of arylamide binding on the surrounding residues is shown in Fig. 2. The positions of residues near the NADH binding site including Gly96, Phe97, Phe149, and Pro193 are insignificantly changed whereas the sidechain of Tyr158 flips away from the reference position to form a hydrogen bond with the amide carbonyl oxygen of the inhibitor. The altered position of the Tyr158 sidechain, in turn, also induces a sidechain shift of Met155 and Met161. Likewise
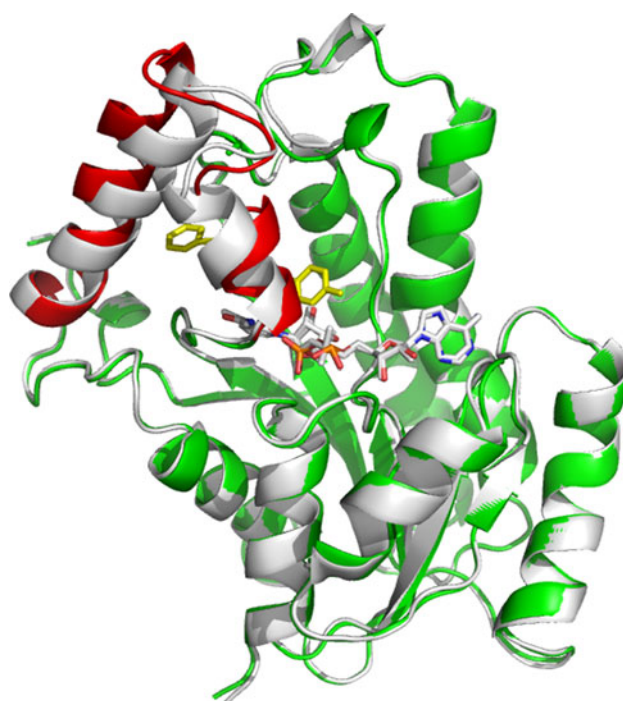
Fig. 1 Superimposition of InhA of pdb code 1ENY (*green*) and 2NSD (*grey* and *red*); arylamide is labeled in *yellow*
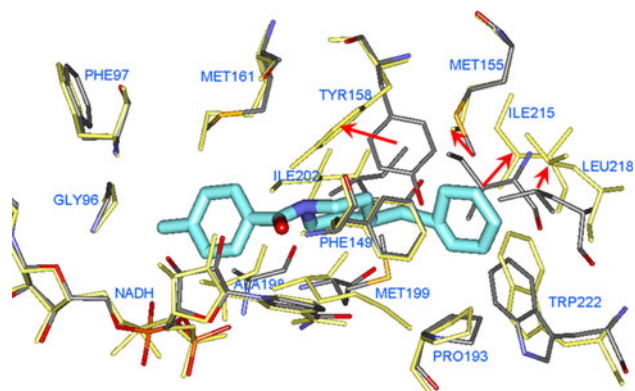
Fig. 2 Arylamide (*blue*) in the InhA binding site (*yellow*) and ligand-unbound InhA (colored by atom type). *Red arrows* show position shifts of residues

the hydrophobic residues of Met199, Ile202, Ile215, Leu218, and Trp222 shift their position. In another arylamide–InhA complex (pdb code IP44), the corresponding shifts of these hydrophobic residues are also observed. These results suggest that the hydrophobic pocket is a highly flexible region capable of binding arylamides most favorably. On the other hand, NADH, Gly96, Phe97, Phe149, and Pro193 are not involved in ligand binding.

Validation of the molecular docking calculations

As already mentioned, the potential binding modes of arylamide derivatives in the InhA binding pocket were carried out by molecular docking calculations with the software Glide (see the section "Data sets and calculation methods"). To assess the reliability of the binding modes obtained from molecular docking calculations, compound **b3** in the InhA X-ray crystal structure was extracted and docked back into the binding pocket. The root mean-square deviation (rmsd) between the docked and crystallographic conformation of compound **b3** is 0.73 Å, indicating that molecular docking calculations with the software Glide enable highly reliable reproduction of the binding mode of compound **b3** in the InhA binding pocket.
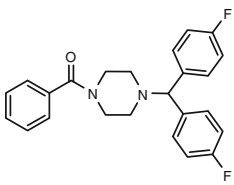
Molecular docking analysis of arylamide derivatives

All the arylamide derivatives in the data set have structural differences in ring moiety B. Accordingly, the arylamide derivatives were classified into three series, a, b, and p which are given in Table 1. The top ranking poses produced by Glide docking were selected as potential interaction modes of arylamide derivatives in the InhA binding pocket. All predicted binding modes of these compounds are consistent with the crystallographic conformation of compound **b3** as shown in Fig. 3.

Table 1 Chemical structures and experimental biological activities for InhA inhibition of arylamide derivatives

Compound	X	n	R^1	R^2	$\log(1/IC_{50})$
a1[a]	N	0	H	H	4.41
a2	N	0	4-CH_3	H	4.78
a3	N	0	4-CH_3	3-CF_3	5.20
a4	N	0	4-CH_3	3-Cl	5.51
a5	N	0	3-CH_3	3-Cl	5.03
a6	N	0	3-CH_3	4-NO_2	4.81
a7	N	0	3,4-Me_2	3-Cl	6.00
a8	N	0	3,4-Me_2	3-CF_3	5.73
a9	N	0	4-i-Pr	3-Cl	<4.00
a10	N	0	4-t-Bu	3-Cl	<4.00
a11	N	0	4-t-Bu	3-CF_3	<4.00
a12	N	0	4-t-Bu	4-CH_3, 3-Cl	<4.00
a13	N	0	2-F	3-Cl	4.86
a14[a]	N	0	4-F	3-Cl	5.01
a15	N	0	3-Cl	3-Cl	5.17
a16	N	0	3,4-Cl_2	3-Cl	5.22
a17	N	0	3,4-Cl_2	H	4.75
a18	N	1	H	H	4.50
b1	C	1	3-Cl	H	5.11
b2	C	1	2-F	H	4.85
b3[a]	C	1	4-CH_3	H	5.29
b4	C	1	3-CH_3	H	5.13
p1					6.40
p2					7.05
p3[a]					6.70
p4					5.98
p5					5.72

Table 1 continued

| Compound	X	n	R^1	R^2	log($1/IC_{50}$)
p6					5.69

^{a} The test set compound for QSAR studies
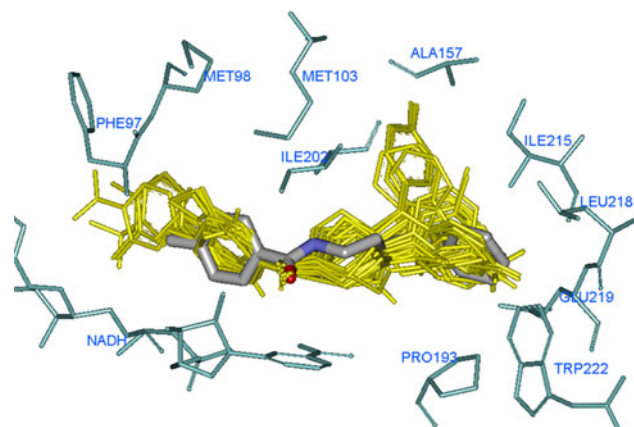
Fig. 3 Docked conformations of arylamide derivatives (*yellow*) and the crystallographic conformation of compound **b3** (colored by atom type) in the InhA binding pocket obtained from Glide docking

Molecular docking analysis of arylamide derivatives in the series a

Compound **a7** has the highest InhA inhibitory activity among compounds in the series a (log($1/IC_{50}$) = 6.00). The binding mode of this compound predicted by molecular docking calculations is shown in Fig. 4. The aryl ring B forms hydrophobic interactions with Phe149, Pro193, Val203, Leu218, and Ile215. The 3-Cl substituent on B forces the aliphatic sidechain of Ile215 and Leu218 to undergo hydrophobic interactions and the nearby located Glu219 forms van der Waals interactions. The ring C of this compound participates in van der Waals interactions with Phe149, Met199, Tyr158, Ile202, and the nicotinamide part of NADH. The amide carbonyl oxygen present in all arylamide derivatives has two hydrogen bond interactions with the hydroxyl groups of the nicotinamide ribose and Tyr158. The aryl ring A and its substituents can form van der Waals interactions with Gly96, Phe97, Met103, Tyr158, Met161, Thr196, Ala198, Met199, and Ile202. Additionally, hydrogen atoms at the ortho and meta positions of the aryl ring A are located near the pyrophosphate oxygen and the oxygen linker of NADH with short
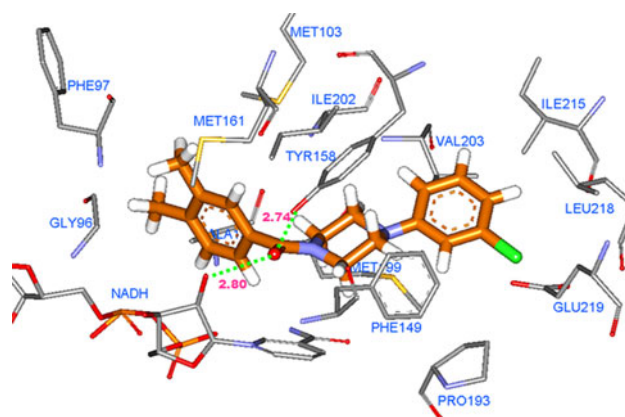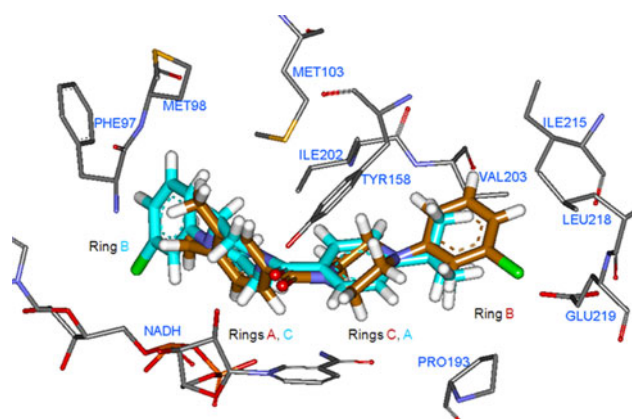

Fig. 4 The docked conformation of compound **a7** in the InhA binding pocket obtained from Glide docking

distances of 3.03 and 2.72 Å, respectively. Accordingly, hydrogen bond interactions may occur among them.

The replacement of both 3,4-Me$_2$ groups on ring A of compound **a7** by two chlorine atoms (3,4-Cl$_2$) results in complete loss of activity, as shown for compound **a16**. Because of the smaller chlorine substituents, the aryl ring A of compound **a16** forms less strong van der Waals interactions with Gly96, Phe97, Met103, Tyr158, Met161, Thr196, Ala198, Met199, and Ile202 than that in compound **a7**. Substituting the para position of the ring A with the bulkier 4-*t*-butyl group (compounds **a10**, **a11**, and **a12**) brings about severe loss of activity (log($1/IC_{50}$) < 4). The docking analysis reveals that these lowest activity compounds have very different binding modes compared with that of compound **a7** as presented in Fig. 5. Because of steric hindrance of the 4-*t*-butyl substituent, ring A of compounds **a10**, **a11**, and **a12** cannot occupy the same pocket region as occupied by ring A of compound **a7**. This result is in accordance with experimental data. Ring A surrounding residues, including Phe97, Gly96, and NADH, prove to be insufficiently flexible for ligand binding as described above. The 4-*t*-butyl groups of these compounds are placed at the same position as aryl moiety B of compound **a7**, because these sites are big enough to accommodate the ring A and its sterically demanding substituents. Even though these compounds can adapt their binding modes in the InhA binding pocket, the important hydrogen-bonding interactions with NADH and Tyr158 are not maintained at the same level as those of compound **a7**. Moreover, the 4-*t*-butyl group loses hydrophobic contact with the hydrophobic residues of Met199, Pro193, and Ile215 comparable with the aryl ring moiety B of compound **a7**. Accordingly, compounds **a10**, **a11**, and **a12** have the lowest activity.

The size of a substituent on B is to some extent related to the size of the substituent on ring A. If A contains a small substituent, a bulky group on B exerts its inhibitory activity without restriction, because of the hydrophobic interactions

Fig. 5 The docked conformations of compound **a7** (carbon atoms colored *orange*) and compound **a10** (carbon atoms colored *blue*) in the InhA binding pocket

of B, as in compounds **a16** and **a17**. If, on the other hand, A bears a bulky substituent, the inhibitory activity of a bulky group on B is diminished, because of alteration of its alignment; this is shown for compounds **a7** and **a8**. The docking analysis reveals that the position of compound **a8** is slightly shifted compared with compound **a7**, because of the 3-CF$_3$ group on ring B. As mentioned above, both 3-Cl and 3-CF$_3$ substituents on ring B situated near Pro193 are not responsible for ligand binding. If compound **a8** lays in the same position as compound **a7**, a steric clash of the 3-CF$_3$ group with this residue may occur. The altered position of compound **a8** leads to a change of the hydrogen bond pattern of this compound compared with compound **a7**; this is shown in Table 2. Because the hydrogen bond distances of compound **a8** between the amide carbonyl oxygen and Tyr158, and the hydrogen atom at the meta position of ring A and the oxygen linker of NADH are increased, the activity of compound **a8** is lower. This situation is obviously responsible for the different activity of compounds **a3** and **a4**. Regarding compounds **a4** and **a5**, the main difference between these two compounds is the methyl group at para

and meta positions. The methyl group at the meta position reduces the inhibitory activity. If the methyl group at the para position is shifted to the meta position, the length of the structural moiety between A and B is reduced, which mainly affects the hydrogen bond interactions of compounds **a4** and **a5** with NADH and Tyr158, as presented in Table 2.

Molecular docking analysis of arylamide derivatives in series b

All compounds in series b have moderate activity, with log(1/IC_{50}) values ranging between 4.85 and 5.29. The binding modes of these compounds, evaluated by means of molecular docking calculations, are similar, as shown in Fig. 6. The ring B and the CH$_2$ linker are located in the neighborhood of the hydrophobic residues Met199, Pro193, Phe149, Ile215, Val203, Trp222, and Leu218. Moreover, van der Waals interactions with Glu219 are observed. The ring C of these compounds forms van der Waals interactions with Phe149, Met199, Tyr158, and Ile202, and nicotinamide of NADH. The important hydrogen bond interactions of the amide carbonyl oxygen found for compounds in series a are still maintained for all compounds of series b.

Aryl ring A of the compounds in series b is oriented similarly to that of the compounds in series a. Hydrogen bond interactions can occur between the hydrogen atoms at the ortho and meta positions of aryl ring A and the pyrophosphate oxygen and the oxygen linker of NADH. The substitution pattern of aryl ring A of the compounds in series b leads to altered compound extensions, which affect the lengths of the hydrogen bonds as shown in Table 3. Moreover, they also affect the van der Waals interactions of the aryl ring A. Compared with the most active compound of series b, compound **b3**, compound **b2**, the least active compound, has shorter hydrogen bond distances with the pyrophosphate oxygen and the oxygen linker of

Table 2 Hydrogen bond distances between compounds **a7**, **a8**, **a4**, **a5**, and NADH and Tyr158 in the InhA enzyme

Residue	Residue group	Ligand atom	Distance (Å)			
			a7	**a8**	**a4**	**a5**
Tyr158	OH	C=O	2.74	2.92	2.95	3.16
			159	163	161	158
NADH (nicotinamide ribose)	OH	C=O	2.80	2.75	2.69	2.72
			179	173	173	167
NADH (oxygen linker)	O	*meta*-H	2.72	2.91	3.64	2.90
			134	123	99	130
NADH (pyrophosphate oxygen)	P–O	*ortho*-H	3.03	2.98	3.34	2.93
			116	121	111	129
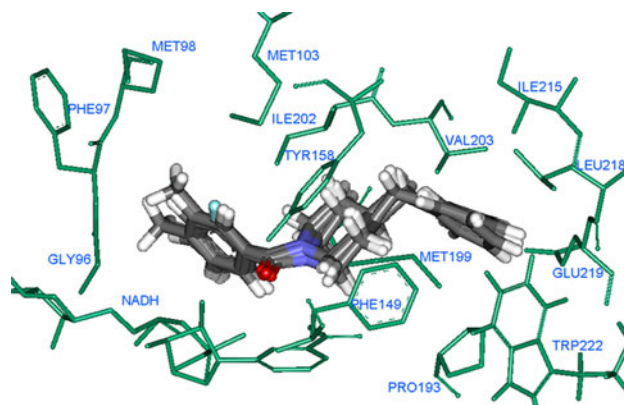
Fig. 6 The docked conformations of compounds **b1**, **b2**, **b3**, and **b4** in the InhA binding pocket

Table 3 Hydrogen bond distances between compounds **b1**, **b2**, **b3**, and **b4** and NADH and Tyr158 in InhA enzyme

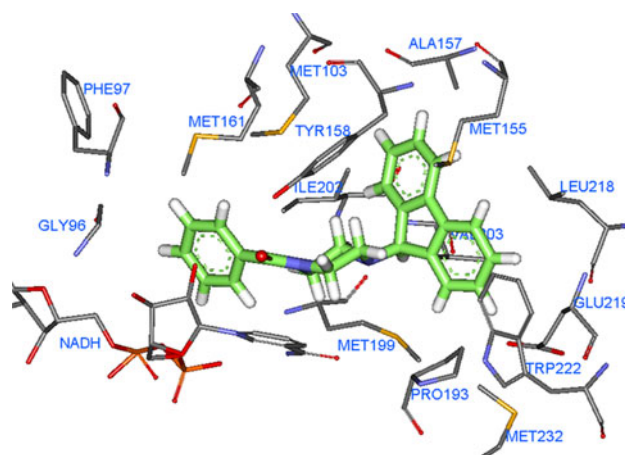| Residue	Residue group	Ligand atom	Distance (Å)			
			b1	b2	b3	b4
Tyr158	OH	C=O	3.04	3.04	2.99	2.81
			162	166	162	161
NADH (nicotinamide ribose)	OH	C=O	2.95	2.73	2.72	2.81
			170	166	172	177
NADH (oxygen linker)	O	*meta*-H	2.63	2.39	2.45	2.47
			127	125	139	137
NADH (pyrophosphate oxygen)	P–O	*ortho*-H	2.86	2.84	2.87	2.79
			112	112	123	122

NADH. The distance to Tyr158 is also shifted. Moreover, compound **b2** loses van der Waals interactions with Gly96 and NADH.

Molecular docking analysis of the arylamide derivatives in series p

Among all the series of arylamide derivatives shown in Table 1, compounds in series p are highly active with $\log(1/IC_{50})$ values ranging from 5.69 to 7.05. Compound **p2** has the highest activity of all the compounds studied. The potential binding mode of this compound within the InhA binding pocket is obtained by flexible ligand docking simulations as presented in Fig. 7. The bulky ring B interacts with hydrophobic residues of Ala157, Met199, Pro193, Val203, Leu218, Ile202, Trp222, and Met232, also forming van der Waals interactions with Tyr158, Met103, Met155, and Glu219. The ring B moieties of the compounds in series p are bulkier than those of the compounds in series a and b, resulting in stronger hydrophobic and van der Waals interactions with the surrounding amino acids. These bulkier rings B are obviously responsible for the higher activity of the compounds in the series p. Hydrogen bond interactions with NADH and Tyr158 as are found for all compounds in series a and b are also observed in series p.

Substitution of the bulky group on ring A of compound **p2** reduces the compound's activity, as shown for compounds **p1** and **p3**. Because of the bulky group on ring A, the binding modes of compounds **p1** and **p3** are slightly shifted compared with the binding mode of compound **p2**, thus affecting the hydrogen bonding pattern of compounds **p1** and **p3**. Some hydrogen bond distances of these compounds are slightly increased compared with those of compound **p2**, as shown in Table 4. These differences obviously bring about the activity loss of compounds **p1** and **p3**.

In another case, compound **p6**, the fluorene ring B of compound **p2** is replaced by a bis(4-fluorophenyl)methyl

Fig. 7 The docked conformation of compound **p2** in the InhA binding pocket

Table 4 Hydrogen bond distances between compounds **p1**, **p2**, **p3**, and **p6** and NADH and Tyr158 in InhA enzyme

Residue	Residue group	Ligand atom	Distance (Å)			
			p1	p2	p3	p6
Tyr158	OH	C=O	2.84	3.05	2.67	3.05
			162	161	170	165
NADH (nicotinamide ribose)	OH	C=O	2.8	2.74	2.93	2.84
			176	170	168	169
NADH (oxygen linker)	O	*meta*-H	2.44	2.38	2.30	2.77
			144	126	149	124
NADH (pyrophosphate oxygen)	P–O	*ortho*-H	2.7	2.79	3.06	2.93
			133	117	116	121

substituent. Because of the presence of a bis(4-fluorophenyl)methyl substituent, the interacting distances of compound **p6** with NADH are strongly increased compared with compound **p2**, as shown in Table 4. The lower activity of compound **p6** can thus be explained.

Structural conditions for favorable interaction of arylamide derivatives with the InhA binding pocket

1. Two hydrogen bonds between the amide carbonyl oxygen and the hydroxyl group of the nicotinamide ribose and the hydroxyl group of Tyr158.
2. Hydrogen bond-type interactions of ortho and meta hydrogens on aryl ring A with the pyrophosphate oxygen and oxygen linker of NADH.
3. Hydrophobic interactions of ring B with side chains of Ala157, Met199, Pro193, Val203, Ile215, Leu218, Ile202, Trp222, and Met232.
4. The size of substituents on ring A and ring B should be checked against one another to keep the optimum

distance for the required hydrogen bond interactions of arylamides with Tyr158 and NADH.

CoMFA and CoMSIA models

 Results from use of CoMFA and CoMSIA methods, which use pharmacophore-based alignments, on 40 arylamide derivatives from two different laboratories have recently been reported [48]. To ensure that all experiments were performed under comparable conditions, only 28 arylamide derivatives taken from one laboratory were used in our work. This smaller data set gave statistically sound results indicating its adequate size for construction of reasonable QSAR models. Statistical data from PLS analyses of CoMFA and CoMSIA models are given in Table 5. The PLS procedure extracts from the steric and the electrostatic fields of CoMFA six relevant components with correlation coefficients $q^2 = 0.68$ and $r^2 = 0.99$. The contribution of steric and electrostatic fields is 67.6 and 32.4%, respectively. On the basis of the better statistical values and more field descriptors, a model including steric, electrostatic, and hydrophobic fields was selected as the best CoMSIA model. The PLS results for this model have correlation coefficients of $q^2 = 0.64$ and $r^2 = 0.95$. Compared with the CoMFA model, the predictive ability of the best CoMSIA model is insignificantly inferior. For the CoMSIA model, the contribution of steric, electrostatic, and hydrophobic fields is 19.0, 39.5, and 41.5%, respectively,

indicating that the additional hydrophobic field has greater effect on inhibitory activity than the others.

HQSAR models

Numerous HQSAR models with a default fragment size (4–7) were generated on the basis of various combinations of the different fragment types which constitute the hologram. Statistical data for all generated HQSAR models are shown in Table 6. Chirality plays a less important role in inhibitory activity of arylamide derivatives, because incorporation of a chirality fragment into molecular hologram models did not improve the q^2 and r^2 values of these models. On the basis of the better statistical values and more fragment distinction parameters, model 10, which includes atoms, bonds, connections, and hydrogen atoms (A/B/C/H) was chosen as the best HQSAR model with q^2 of 0.74, r^2 of 0.95, standard error of 0.48, hologram length 53, and an optimum number of components of six. These statistical results demonstrate the reliability and good predictive power of the best HQSAR model.

Validation of the QSAR models

The predicted activities of the training and test sets derived from the best CoMFA, CoMSIA, and HQSAR models are listed in Table 7. The graphs of correlations between experimental and predicted activities are depicted in Fig. 8.

Table 5 Summary of statistical results for the CoMFA model and various CoMSIA models with different combined fields

Models	Statistical data						Fraction
	q^2_{cv}	r^2	N	s	SEE	F	
CoMFA	0.68	0.99	6	0.53	0.09	287.88	67.6/32.4 (S/E)
CoMSIA							
S/E	0.65	0.91	3	0.51	0.26	60.78	36.7/63.3
S/H	0.61	0.96	5	0.57	0.17	88.60	31.2/68.8
S/HA	0.62	0.83	2	0.52	0.35	47.13	76.0/24.0
S/HD	0.62	0.88	4	0.54	0.31	31.68	94.3/5.7
S/E/H	**0.64**	**0.95**	**4**	**0.53**	**0.20**	**85.44**	**19.0/39.5/41.5**
S/E/HA	0.62	0.86	2	0.51	0.31	61.74	30.6/57.3/12.1
S/E/HD	0.62	0.92	4	0.54	0.24	53.61	36.8/59.5/3.7
S/H/HA	0.56	0.82	2	0.55	0.35	46.58	28.3/59.5/12.2
S/H/HD	0.61	0.97	6	0.58	0.16	85.05	29.9/67.0/3.1
S/HA/HD	0.63	0.83	3	0.52	0.35	31.34	74.2/23.3/2.5
S/E/H/HD	0.61	0.96	5	0.56	0.19	75.56	18.7/38.3/40.6/2.4
S/E/H/HA	0.57	0.91	3	0.56	0.25	65.23	17.8/35.3/38.1/8.8
S/E/H/HD/HA	0.57	0.89	3	0.56	0.29	48.53	18.5/35.1/38.0/0.8/7.6

Bold values indicate the best CoMSIA model

q^2_{cv}, leave-one-out (LOO) cross-validated correlation coefficient; r^2, non-cross-validated correlation coefficient; N, optimum number of components; s, standard error of prediction; SEE, standard error of estimate; F, F test value; S, steric field; E, electrostatic field; H, hydrophobic; HD, hydrogen donor field; HA, hydrogen acceptor field

Table 6 The statistical results of HQSAR models combined with various fragment distinction parameters

Model	Fragment type	q_{cv}^2	r^2	S	SEE	HL	N
1	A/B	0.76	0.93	0.44	0.24	199	5
2	A/C	0.76	0.93	0.46	0.24	83	6
3	A/DA	0.54	0.87	0.63	0.35	97	6
4	A/H	0.73	0.90	0.46	0.28	257	4
5	A/Ch	0.70	0.85	0.50	0.35	61	5
6	A/B/C	0.77	0.93	0.44	0.24	61	6
7	A/B/H	0.73	0.96	0.49	0.20	97	6
8	A/B/DA	0.72	0.92	0.50	0.28	307	6
9	A/B/Ch	0.76	0.89	0.42	0.28	83	3
10	**A/B/C/H**	**0.74**	**0.95**	**0.48**	**0.20**	**53**	**6**
11	A/B/C/DA	0.72	0.92	0.50	0.27	59	6
12	A/B/DA/H	0.69	0.94	0.52	0.24	97	6
13	A/B/C/Ch	0.77	0.93	0.44	0.24	61	6
14	A/B/DA/Ch	0.72	0.91	0.49	0.29	71	6
15	A/B/H/Ch	0.73	0.95	0.49	0.21	53	6
16	A/B/C/DA/H	0.68	0.92	0.51	0.26	97	5
17	A/B/C/DA/Ch	0.73	0.92	0.48	0.27	53	6
18	A/B/C/H/Ch	0.74	0.95	0.48	0.20	53	6
19	A/B/DA/H/Ch	0.66	0.93	0.54	0.25	97	6
20	A/B/C/DA/H/Ch	0.65	0.93	0.56	0.25	97	6

Bold font indicates the best HQSAR model

HL, hologram length; A, atoms; B, bonds; C, connections; H, hydrogen atoms; Ch, chirality; DA, donor and acceptor

The predicted activities of the training set are close to the experimental activities with deviation values <0.19, 0.37, and 0.37 logarithmic units for CoMFA, CoMSIA, and HQSAR models, respectively, in agreement with experimental and predicted activities. By examining the statistical results, **a6** could be identified as an outlier within the best CoMFA, CoMSIA, and HQSAR models. In order to assess the predictive ability of these QSAR models, the biological activities of the test set compounds were predicted. On the basis of the best QSAR models, all test set compounds have predicted values within one logarithmic unit of the experimental values presented in Table 7. The r^2 values from the best CoMFA, CoMSIA, and HQSAR models are 0.86, 0.81, and 0.92, respectively. These results reveal that all selected QSAR models are reliable, with high predictive power. Therefore, the best CoMFA, CoMSIA, and HQSAR models could be used to design new arylamide derivatives with improved properties.

CoMFA and CoMSIA contour maps

To easily visualize the importance of steric, electrostatic, and hydrophobic fields on the inhibitory activity of arylamide derivatives, CoMFA and CoMSIA contour maps were

established. Favorable and unfavorable steric regions are represented by green and yellow contours, respectively, whereas blue and red contours characterize regions which favor positive and negative charges, respectively. For CoMSIA magenta and white contours represent favorable and unfavorable hydrophobic regions, respectively. The predicted binding modes obtained from Glide docking of compounds **a7**, **b3**, and **p2**, the most active compounds in series a, b, and p, respectively, were used for the contour map interpretation.

From CoMFA steric contours as shown in Fig. 9a, two large yellow contours located near the meta and para substituents on the aryl ring A of compounds **a7**, **b3**, and **p2** indicate that these regions must not contain bulky substituents. These findings explain why compounds **a9**, **a10**, **a11**, and **a12** which bear bulky *i*-propyl and *t*-butyl substituents at the para position are less potent than compound **a7**. As described in the docking analysis, these bulky substituents change the binding modes of these inhibitors leading to loss of the main binding interactions. Another interesting large green contour is located near ring B of compounds **a7**, **b3**, and **p2**. As seen from Fig. 9a, only the bulky ring B of compound **p2** is buried in this green contour. This finding explains why compounds in series a and b containing the smaller ring B are less potent than compound **p2** bearing the bulkier ring B. These results agree well with docking results showing that a bulky ring B interacts with the hydrophobic residues of Ala157, Met199, Pro193, Val203, Leu218, Ile202, Trp222, and Met232 to a greater extent than a small ring B. Therefore, introducing a bulkier group to ring B of compounds **a7** and **b3** within the large green region would enhance the activity of these compounds.

Figure 9b presents the electrostatic CoMFA contour map. A blue contour appears near the hydrogen atoms at the ortho position of aryl ring A of compounds **a7**, **b3**, and **p2** indicating that a more positively charged substituent is favored in this position for increased affinity. In good agreement with the docking results, the presence of positively charged substituents at this position favors hydrogen bonding interactions with the pyrophosphate oxygen of NADH. Additionally, two large blue contours situated near the meta and para substituents on aryl ring A of these compounds imply that electron-deficient substituents were preferred in these regions. In another area, CoMFA model shows that a large electronegatively favored region is located above the aromatic fluorene ring B of compound **p2**, implying that the aromatic ring at this position would enhance the inhibitory potency.

Steric and electrostatic CoMSIA contours (not shown) are similar to the corresponding CoMFA contours. Therefore, only the hydrophobic contour is discussed in this study. The CoMSIA model indicating hydrophobicity is

Table 7 Experimental and predicted activities of training and test sets from CoMFA, CoMSIA, and HQSAR models

| Compound	Experimental	log $1/IC_{50}$					
		CoMFA		CoMSIA		HQSAR	
		Predicted	Residue	Predicted	Residue	Predicted	Residue
a1[a]	4.41	4.87	−0.46	5.05	−0.64	4.82	−0.01
a2	4.78	4.86	−0.07	4.73	0.05	5.03	−0.25
a3	5.20	5.26	−0.06	5.38	−0.18	5.15	0.05
a4	5.51	5.53	−0.02	5.30	0.21	5.14	0.37
a5	5.03	5.05	−0.02	4.95	0.08	5.19	−0.15
a7	6.00	5.91	0.09	5.63	0.37	5.76	0.24
a8	5.73	5.74	−0.01	5.57	0.16	5.77	−0.04
a9	4.00	3.97	0.04	3.98	0.02	3.95	0.05
a10	4.00	4.01	−0.01	3.99	0.01	3.97	0.03
a11	4.00	3.98	0.02	4.06	−0.06	3.97	0.03
a12	4.00	4.05	−0.05	3.99	0.01	4.07	−0.07
a13	4.86	4.86	0.00	4.87	−0.01	4.97	−0.09
a14[a]	5.01	4.64	0.37	4.83	0.18	4.97	0.04
a15	5.17	5.29	−0.12	5.48	−0.31	5.18	−0.01
a16	5.22	5.18	0.04	5.36	−0.14	5.14	0.08
a17	4.75	4.65	0.10	4.64	0.11	5.04	−0.29
a18	4.50	4.51	−0.01	4.65	−0.15	4.39	0.11
b1	5.11	4.92	0.19	5.06	0.05	5.10	0.01
b2	4.85	4.83	0.02	4.94	−0.09	4.88	−0.03
b3[a]	5.29	5.10	0.19	5.07	0.22	5.08	0.21
b4	5.13	5.29	−0.16	5.21	−0.08	5.10	0.03
p1	6.40	6.46	−0.06	6.67	−0.27	6.75	−0.35
p2	7.05	6.96	0.09	6.77	0.28	6.68	0.37
p3[a]	6.70	6.61	0.09	6.11	0.59	7.06	−0.36
p4	5.98	6.03	−0.04	5.97	0.01	6.04	−0.06
p5	5.72	5.75	−0.03	5.51	0.21	5.74	−0.02
p6	5.69	5.64	0.06	5.99	−0.30	5.68	0.01

[a] The test set compound

shown in Fig. 9c. Two hydrophobically unfavorable white contours appear near the meta and para substituents on aryl ring A of compounds **a7**, **b3**, and **p2**. The combined results from CoMFA and CoMSIA contour maps indicate that electron-donating groups with low steric demand and high hydrophilicity in these regions would be helpful to enhance the binding affinity of arylamide derivatives with the InhA pocket. Two hydrophobically favorable magenta contours are located near ring B of compounds **a7**, **b3**, and **p2**. Only the bulky fluorene substituent of compound **p2** is buried in both magenta contours. These findings clearly show that the bulky fluorene substituent of compound **p2** favors hydrophobic interactions to a greater extent than aryl ring B of compounds **a7** and **b3**. Therefore, compounds that contain the bulky fluorene ring as the B substituent have higher inhibitory activity than the others, as shown for compounds **p1**, **p2**, and **p3**. With regard to compounds **a7** and **b3**, the hydrophobic group matches two favorable

hydrophobic magenta contours of ring B and hence should enhance their activity.
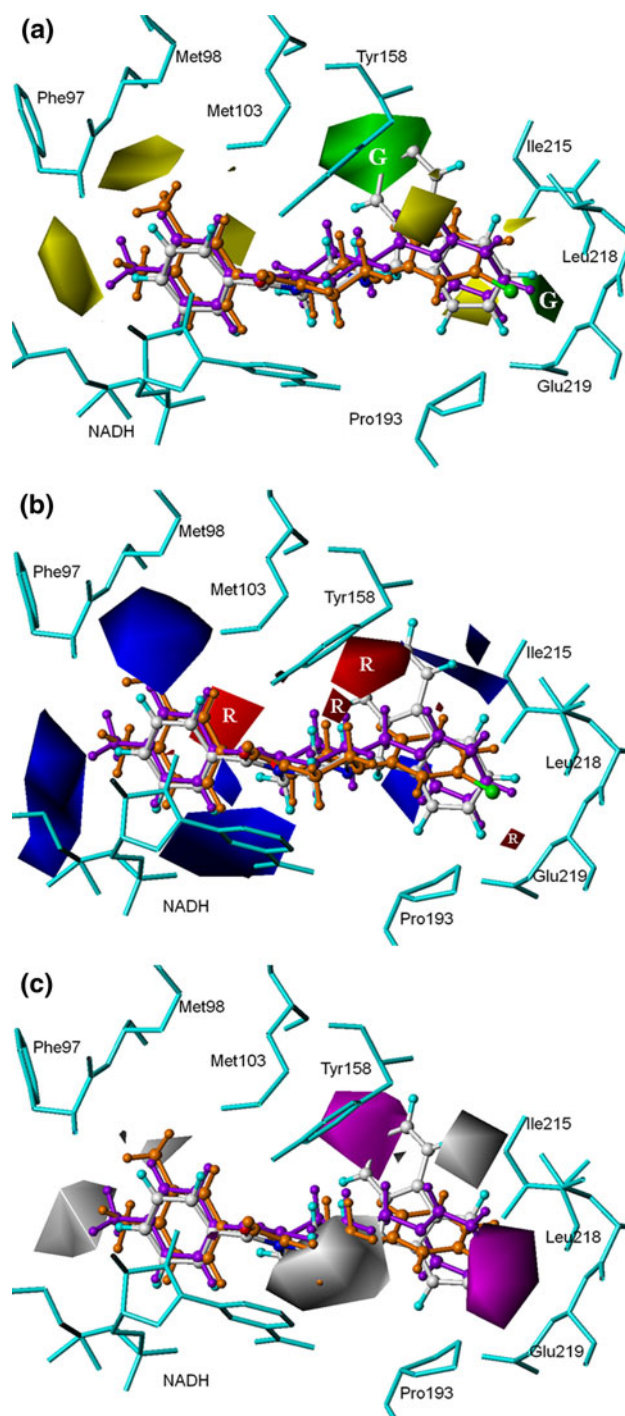
HQSAR contribution maps

Molecular fragments of arylamides which contribute directly to biological activity can be visualized through HQSAR contribution maps. The different contributions of all atoms in a molecule to the biological activity are discriminated by a color code. Atoms with negative contributions are represented at the red end of the spectrum whereas atoms with positive contributions are presented by the colors at the green end of the spectrum. The white colored atoms make intermediate contributions. Figure 10 depicts the individual atomic contributions to the activity of the highly active compounds **a7**, **b3**, and **p2**.

As seen from Fig. 10, the fragments of the fluorene ring B of the most active compound **p2** are positively correlated
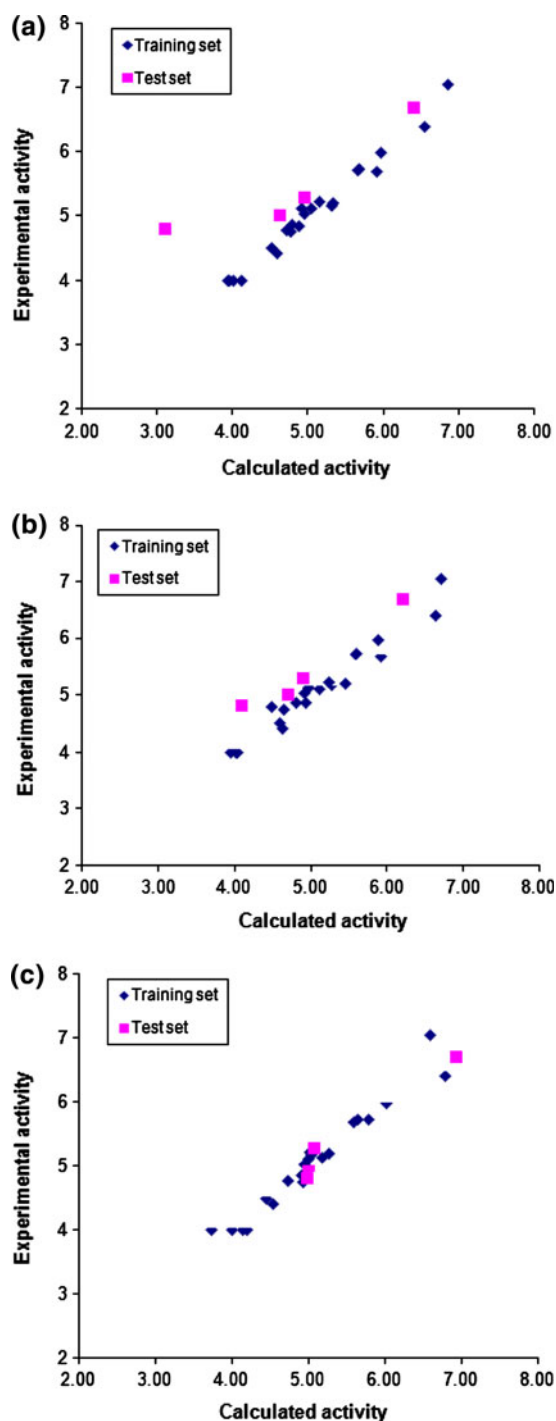
Fig. 8 Plots of experimental activity against predicted activity for the training and test sets derived from CoMFA (**a**), CoMSIA (**b**), and HQSAR (**c**) models

Fig. 9 CoMFA steric contour (**a**), CoMFA electrostatic contour (**b**), and CoMSIA hydrophobic contour (**c**) in combination with compounds **a7** (*orange*), **b3** (*violet*), and **p2** (colored by atom type) in the InhA binding pocket (*cyan*). *Green* and *yellow* contours represent favorable and unfavorable steric regions, respectively. *Blue* and *red* contours are favored for electron-positive groups and electron-negative groups, respectively. *Magenta* and *white* contours show favorable and unfavorable hydrophobic regions, respectively

with the biological activity of this compound. Compounds **p1** and **p3** containing a fluorene ring also have high biological activity. If the fluorene ring of compound **p2** is replaced by the smaller aryl ring of compounds **a7** and **b3**,

the individual atoms of the aryl ring make less positive contributions than the individual atoms of the fluorene ring; particularly the aryl ring of compound **b3** as shown in Fig. 10. These observations confirm the previous CoMFA and CoMSIA results which suggest that the bulky volume of ring B would enhance the biological activity of arylamide derivatives. Consistent with the docking results also, a bulky ring B is favorable for hydrophobic interactions leading to increased binding affinity. With regard to the aryl ring A, this fragment of compounds **b3** and **p2** is positively related to the biological activity. As described in the docking analysis, yellow and green colored hydrogen atoms at the ortho and meta positions of aryl ring A favor the formation of hydrogen bond interactions with NADH. This HQSAR result corroborates the importance of these hydrogen bond interactions in the biological activity of the arylamide derivatives. In compound **a7**, all substituents on the aryl ring A make no contributions to the biological

Fig. 10 The HQSAR contribution maps for compounds **a7** (**a**), **b3** (**b**), and **p2** (**c**)

activity of this compound, indicating that modification of these substituents based on CoMFA and CoMSIA suggestions should enhance the potency. Importantly, the 3-Cl substituent on the aryl ring B of compound **a7** is colored green, implying that this substituent is necessary for the binding affinity of this compound and others in series a.

Data sets and calculation methods

Data sets and biological activities

Some of the 28 arylamide derivatives were found in the literature [44]. Chemical structures and experimental biological activity of these compounds are listed in Table 1. All the arylamide derivatives in the data set are structurally different. Compounds bearing a nitrogen atom at position X and an aryl ring as the ring moiety B are defined as series a, whereas compounds that contain a carbon atom at position X are defined as series b. Another series of compounds bearing a fluorene or bis(4-fluorophenyl)methyl ring at position B are defined as series p. The biological activity of these compounds for InhA inhibition were expressed in terms of IC_{50} values. For QSAR studies, IC_{50} values were converted as usual to the corresponding $\log(1/IC_{50})$. All chemical structures of arylamide derivatives were constructed using the standard tools available in GaussView 3.07 software [52] and were then fully optimized using the HF/3-21G method implemented in the Gaussian 03 software [53].

Molecular docking calculations

The X-ray crystal structure of compound **b3** complexed with InhA (pdb code 2NSD) was used for molecular docking experiments. Docking of arylamide derivatives was carried out by use of the software Glide, version 5 [54]. The InhA protein for the docking study was prepared by using the Protein Preparation Wizard tool of Glide. All hydrogen atoms were added to the protein structure and all crystallographic water molecules were excluded. To keep the original X-ray crystal coordinates unaltered, the entire protein was not minimized; only the hydrogen bond network between the ligand and the protein was optimized. The receptor grid was generated using the receptor grid generation tool in the Glide software. The scaling of the van der Waals radius was set as default, a scaling factor of 1.0 and a partial charge cutoff of 0.25. The grid box with size of 14 Å was centered on the workspace ligand. For subsequent molecular docking of the ligand into the InhA binding site, standard precision (SP) with dock flexibly option was selected and ten distinct ligand poses per ligand were written out. No constraints were selected in the docking
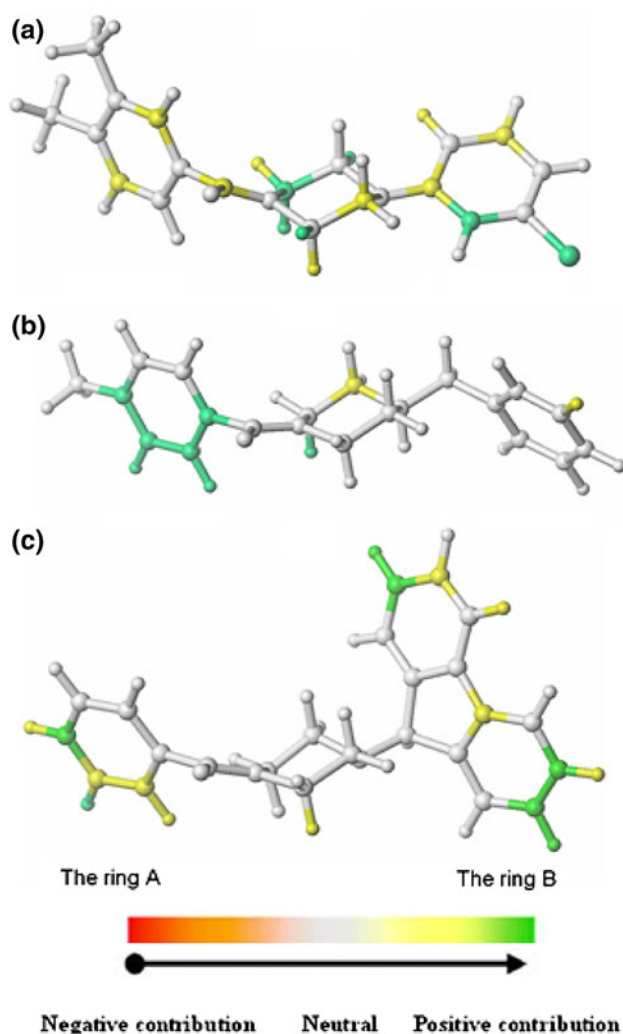
runs. Post-docking minimization was performed with the default setting. Compound **b3** was docked back into the InhA binding pocket to validate the docking method, and subsequently all derivatives were docked. The ligand pose with highest G score was selected as the best binding mode of arylamide derivatives in the InhA binding pocket.

Training and test sets

The chemical structures and biological activities of arylamide derivatives used to set up CoMFA, CoMSIA, and HQSAR models are listed in Table 1. Experimental activities of these compounds are spread over a range of $3.05 \log(1/IC_{50})$ units thus providing a broad and homogenous data set for CoMFA, CoMSIA, and HQSAR studies. The 28 arylamide derivatives were divided into a training set of 24 compounds and a test set of four compounds for final model development and model validation, respectively. The representatives of the test set were manually selected and cover the entire range of activity and structural diversity of the arylamides in the data set.

CoMFA and CoMSIA techniques

Structural alignment of compounds according to their bioactive conformations is an important prerequisite for the set up of appropriate CoMFA and CoMSIA models. In this study, the reasonable binding modes of compounds in the InhA binding pocket obtained from Glide docking were used for molecular alignment. SYBYL 8.0 molecular modeling software [55] was used to calculate the CoMFA and CoMSIA models. A sp^3 carbon atom with a formal charge of $+1$ was selected as the probe atom to generate the steric (Lennard–Jones potential) and the electrostatic (Coulomb potential) fields. The probe atom was placed at all intersections in a grid with spacing of 2 Å. The steric and electrostatic fields around the aligned compounds were all calculated with CoMFA standard scaling. The maximum steric and electrostatic energies were truncated at 125 kJ/mol.

Five CoMSIA similarity index descriptors of steric, electrostatic, hydrophobic, hydrogen bond donor, and hydrogen bond acceptor fields were derived with the same grid as used for the CoMFA field calculation. There are no energy cutoffs for CoMSIA calculations because a distance-dependent Gaussian type was used, differing from the procedure for CoMFA calculations. To generate a contour map with prominent molecular features in the CoMSIA study, an attenuation factor of 0.3 was used.

To derive a linear relationship between molecular descriptors and activities, the partial least square (PLS) approach was used, in which CoMFA and CoMSIA descriptors were set as independent variables and $\log(1/IC_{50})$

values were used as dependent variables. The cross-validation was performed using the leave-one-out method with an 8.4 kJ/mol column filter to minimize the influence of noisy columns. A final non-cross-validated analysis with the optimum number of components was sequentially performed and was then used to analyze the results. The r^2 and q^2 values were used to evaluate the predictive ability of the CoMFA and CoMSIA models.

HQSAR

Hologram QSAR (HQSAR) does not require information about the 3D structure of inhibitors, because this method uses only 2D structural information. Hence, in contrast with the CoMFA and CoMSIA methods, HQSAR needs no molecular alignment. For the HQSAR study, the HQSAR module of SYBYL 8.0 was used. The same training and test sets as for the CoMFA and CoMSIA studies were used in the HQSAR study. Each compound in the training set was converted into all possible molecular fragments including linear, branched, cyclic, and overlapping fragments in the size range 4–7 atoms. Molecular fragment generation utilizes fragment-distinction properties including atoms (A), bonds (B), connections (C), hydrogen atoms (H), chirality (Ch), and donor and acceptor (DA). The generated molecular fragments are counted in bins of a fixed length array to produce a molecular hologram. The hologram length was set with 12 as a default length ranging from 53 to 401. The PLS method was used to establish a correlation of the molecular hologram descriptors with the biological data. The best model was selected on the basis of the best cross-validated r^2. To develop robust HQSAR models, numerous models with various combinations of the fragment-distinction properties were built.

References

1. World Health Organization (WHO) (2009) Global tuberculosis control—epidemiology, strategy, financing. http://www.who.int/tb/publications/global_report/2009/en/index.html
2. Aziz MA, Wright A, Laszlo A, De Muynck A, Portaels F, Van Deun A, Wells C, Nunn P, Blanc L, Raviglione M (2006) Lancet 368:2142
3. Raviglione MC, Smith IM (2007) N Engl J Med 356:656
4. Morlock GP, Metchock B, Sikes D, Crawford JT, Cooksey RC (2003) Antimicrob Agents Chemother 47:3799

5. Ormerod LP (2008) Thorax 63:440
6. Ormerod LP (2005) Br Med Bull 73–74:17
7. Dravniece G, Cain KP, Holtz TH, Riekstina V, Leimane V, Zaleskis R (2009) Eur Respir J 34:180
8. Suchindran S, Brouwer ES, Van Rie A (2009) PLoS One 4:e5561
9. Chakaya J, Getahun H, Granich R, Havlir D (2008) J Int AIDS Soc 11:6
10. El-Sadr WM, Tsiouris SJ (2008) Semin Respir Crit Care Med 29:525
11. Nunn P, Reid A, De Cock KM (2007) J Infect Dis 196:S5
12. Campbell JW, Cronan JE Jr (2001) Annu Rev Microbiol 55:305
13. Heath RJ, Rock CO (2004) Curr Opin Invest Drugs 5:146
14. White SW, Zheng J, Zhang YM, Rock CO (2005) Annu Rev Biochem 74:791
15. Zhang YM, Lu YJ, Rock CO (2004) Lipids 39:1055
16. Wen L, Chmielowski JN, Bohn KC, Huang JK, Timsina YN, Kodali P, Pathak AK (2009) Protein Expr Purif 65:83
17. Wright HT, Reynolds KA (2007) Curr Opin Microbiol 10:447
18. Quémard A, Sacchettini JC, Dessen A, Vilcheze C, Bittman R, Jacobs WR Jr, Blanchard JS (1995) Biochemistry 34:8235
19. Rozwarski DA, Grant GA, Barton DH, Jacobs WR Jr, Sacchettini JC (1998) Science 279:98
20. Vilchèze C, Wang F, Arai M, Hazbón MH, Colangeli R, Kremer L, Weisbrod TR, Alland D, Sacchettini JC, Jacobs WR Jr (2006) Nat Med 12:1027
21. Dessen A, Quemard A, Blanchard JS, Jacobs WR Jr, Sacchettini JC (1995) Science 267:1638
22. Johnsson K, Schultz PG (1994) J Am Chem Soc 116:7425
23. Lei B, Wei CJ, Tu SC (2000) J Biol Chem 275:2520
24. Johnsson K, King DS, Schultz PG (1995) J Am Chem Soc 117:5009
25. Zhang Y, Heym B, Allen B, Young D, Cole S (1992) Nature 358:591
26. Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um KS, Wilson T, Collins D, de Lisle G, Jacobs WR Jr (1994) Science 263:227
27. Quemard A, Dessen A, Sugantino M, Jacobs WR Jr, Sacchettini JC, Blanchard JS (1996) J Am Chem Soc 118:1561
28. Lei B, Wei C, Tu SC (2000) J Biol Chem 275:2520
29. Saint-Joanis B, Souchon H, Wilming M, Johnsson K, Alzari PM, Cole ST (1999) Biochem J 338:753
30. Zhao X, Yu H, Yu S, Wang F, Sacchettini JC, Magliozzo RS (2006) Biochemistry 45:4131
31. Metcalfe C, Macdonald IK, Murphy EJ, Brown KA, Raven EL, Moody PC (2008) J Biol Chem 283:6193
32. Sinha BK (1983) J Biol Chem 258:796
33. Nguyen M, Claparols C, Bernadou J, Meunier B (2001) ChemBioChem 2:877
34. Heym B, Zhang Y, Poulet S, Young D, Cole ST (1993) J Bacteriol 175:4255
35. Timmins GS, Deretic V (2006) Mol Microbiol 62:1220
36. Johnsson K, Froland WA, Schultz PG (1997) J Biol Chem 272:2834
37. De La Iglesia AI, Morbidoni HR (2006) Rev Argent Microbiol 38:97
38. Suarez J, Ranguelova K, Schelvis JP, Magliozzo RS (2009) Biol Chem 284:16146
39. Freundlich JS, Wang F, Vilchèze C, Gulten G, Langley R, Schiehser GA, Jacobus DP, Jacobs WR Jr, Sacchettini JC (2009) ChemMedChem 4:241
40. Ende CWA, Knudson SE, Liu N, Childs J, Sullivan TJ, Boyne M, Xu H, Gegina Y, Knudson DL, Johnson F, Peloquin CA, Slayden RA, Tonge PJ (2008) Bioorg Med Chem Lett 18:3029
41. Boyne ME, Sullivan TJ, am Ende CW, Lu H, Gruppo V, Heaslip D, Amin AG, Chatterjee D, Lenaerts A, Tonge PJ, Slayden RA (2007) Antimicrob Agents Chemother 51:3562
42. Tonge PJ, Kisker C, Slayden RA (2007) Curr Top Med Chem 7:489
43. He X, Alian A, Stroud R, Ortiz de Montellano PR (2006) J Med Chem 49:6308
44. He X, Alian A, Ortiz de Montellano PR (2007) Bioorg Med Chem 15:6649
45. Lu XY, Chen YD, Jiang YJ, You QD (2009) Eur J Med Chem 44:3718
46. Kumar A, Siddiqi MI (2008) J Mol Model 14:923
47. Andrade CH, Salum Lde B, Castilho MS, Pasqualoto KF, Ferreira EI, Andricopulo AD (2008) Mol Divers 12:47
48. Lu XU, Chen YD, You QD (2010) Chem Bio Drug Des 75:195
49. Cramer RD III, Patterson DE, Bunce JD (1988) J Am Chem Soc 110:5959
50. Klebe G, Abraham U, Mietzner T (1994) J Med Chem 37:4130
51. Tong W, Lowis DR, Perkins R, Chen Y, Welsh WJ, Goddette DW, Heritage TW, Sheehan DM (1998) J Chem Inf Comput Sci 38:669
52. GaussView 03, Revision 3.07 (2006) Gaussian, Inc., Wallingford
53. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03, Gaussian Inc., Wallingford
54. Glide, version 5.0 (2008) Schrödinger, LLC, New York, NY
55. SYBYL 8.0 (2007) Tripos, Inc