

Inhibitors of type III secretion in *Yersinia*: Design, synthesis and multivariate QSAR of 2-arylsulfonylamino-benzanilides

Anna M. Kauppi,^a C. David Andersson,^a Henrik A. Norberg,^b Charlotta Sundin,^b
Anna Linusson^a and Mikael Elofsson^{a,*}

^aDepartment of Chemistry, Umeå University, SE-90187 Umeå, Sweden

^bInnate Pharmaceuticals AB, Umestan Företagspark, SE-90347 Umeå, Sweden

Received 19 February 2007; revised 18 July 2007; accepted 25 July 2007

Available online 22 August 2007

Abstract—Compound **1**, 2-(benzo[1,2,5]thiadiazole-4-sulfonylamino)-5-chloro-*N*-(3,4-dichloro-phenyl)-benzamide, was identified as a putative type III secretion inhibitor in *Yersinia*, and the compound thus has a potential to be used to prevent or treat bacterial infections. A set of seven analogues was synthesized and evaluated in a type III secretion dependent reporter-gene assay with viable bacterial to give basic SAR. A second set of 19 compounds was obtained by statistical molecular design in the building block and product space and by subsequent synthesis. Evaluation in the reporter-gene assay showed that the compounds ranged from non-active to compounds more potent than **1**. Based on the data multivariate QSAR models were established and the final Hi-PLS model showed good correlation between experimentally determined % inhibition and the calculated % inhibition of the reporter-gene signal.

© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

Although mankind has developed a large arsenal of antibacterial agents, lethality in infectious diseases still is the 2nd leading cause of death worldwide.¹ Another worrisome fact is the constantly increasing number of microbial pathogens that develop or acquire resistance against antibiotics currently in use.^{2,3} To combat resistant bacteria, there is an obvious need for effective strategies such as development of new antimicrobial drugs against not yet exploited targets.^{4–6} Antibacterial drugs in use today act on a limited number of targets and a successful continuation of the antibiotic era will most likely rely on drugs with other modes of action. It is now clear that various pathogenic bacteria use related virulence systems and it has been shown that some components of certain virulence systems are conserved between different species. These findings offer a possibility to develop novel antibacterial agents that target virulence,^{4,5,7,8} that is, the bacteria's ability to cause disease.⁹ Targeting growth and virulence under in vivo

like conditions will likely identify completely new sets of molecules, as recently shown for *Vibrio cholerae* where a small molecule inhibitor of the transcriptional activator ToxT, virstatin, prevented both toxin and pili expression, protecting infant mice from colonization.¹⁰ The importance and potential of virulence mechanisms as drug targets is further underscored by a recent and extensive study of *Salmonella* that indicates that a strategy focusing on metabolic pathways is unlikely to be successful.¹¹ Virulence blocking agents can also be employed in a chemical biology approach to learn and understand more about protein function and the mechanisms underlying the complexity of bacterial virulence.¹²

The type III secretion (T3S) system is a virulence mechanism utilized by several Gram-negative pathogens such as *Yersinia* spp., *Salmonella* spp., *Shigella* spp., *Pseudomonas aeruginosa*, enteropathogenic *Escherichia coli*, enterohaemorrhagic *E. coli*, and *Chlamydia* spp.¹³ The clinical relevance of these pathogens suggests that T3S systems are potential targets for novel anti-infective drugs.^{14,15} The plasmid encoded Ysc (*Yersinia* secretion) T3S system in *Yersinia*¹⁶ is extensively studied and relatively well understood and thus serves as a good model system.¹⁷ Three of the eleven known species of *Yersinia*, *Y. pestis*, *Y. enterocolitica*, and *Y. pseudotuberculosis* are

Keywords: Type III secretion; *Yersinia*; Virulence; Inhibitors; Statistical molecular design; multivariate QSAR.

* Corresponding author. Tel.: +46 90 786 9328; fax: +46 90 138885; e-mail: mikael.elifsson@chem.umu.se

pathogenic to mammals of which *Y. pseudotuberculosis* and *Y. enterocolitica* are enteropathogens mainly causing gastroenteritis, whilst *Y. pestis* is the agent responsible for bubonic plague.^{17,18} The T3S system allows direct transfer of the effector proteins into the host cell cytosol, creating an environment supporting bacterial survival and proliferation by targeting specific host proteins. The Ysc T3S system is a complex machinery composed of specific Yop chaperones (Sycs) that deliver the effector proteins, *Yersinia* outer proteins (Yops) to the secretion channel through which the Yops are delivered into the host cell. The Ysc T3S secretion apparatus is essential for the bacteria to evade the host immune defense and compounds targeting this mechanism will attenuate bacterial virulence without affecting their growth. We hypothesize that compounds acting on the T3S will allow the host to clear the infection while development of resistance will be slow due to the high energy cost for the bacteria to mutate virulence genes. Moreover, microbial resistance against antibiotics is often obtained through horizontal gene transfer from the resident micro-flora and during antibiotic treatment of patients, a selection of already resistant micro-flora will occur.¹⁹ Compounds that target the T3S system will not affect the host micro-flora since it lacks T3S systems and hence development of resistance by target mutations cannot occur in the micro-flora.²⁰ Moreover, nine of the Ysc proteins have counterparts in almost all T3S systems and it has been demonstrated that some components of the secretion systems are interchangeable among different species,²¹ which provide evidence for evolutionary conservation. Given that the T3S systems are conserved among the Gram-negative bacteria, it is likely that compounds targeting T3S machinery in *Yersinia* might also affect the T3S system in other species. Up to date only a limited number of studies regarding virulence inhibitors and T3S as a potential target have been presented.^{22–27}

A general method for the identification of biologically active compounds is to screen large compound collections in whole-cell phenotypic assays or assays based on purified proteins, that is, high-throughput screening (HTS). Interesting compounds with desired biological activity are selected and must then be further developed by iterative design, synthesis, and biological evaluation. A systematic approach to investigate a singleton or a compound class is to characterize a large number of virtually possible compounds by theoretical molecular descriptors and subsequent selection of compounds for synthesis based on chemical diversity, for example by using statistical molecular design (SMD).^{28,29} It is possible to select a smaller subset of compounds by using SMD techniques without losing information contained in the full library. One major advantage of SMD is that the resulting information can be used for establishment of multivariate quantitative structure–activity relationship (QSAR) models which later on can be used for interpretation of important molecular properties and prediction of biological activity for new compounds belonging to the same class.^{29–33} SMD can be based either on the building blocks or on computer-generated products. The major advantages of design in building

block space are that the characterization of physiochemical properties will be made on a smaller set of compounds and the number of reactants will be reduced in a straightforward way.²⁸ This procedure also makes it possible to use the principal properties of building blocks as variables in a hierarchical partial least square regression to latent structures (Hi-PLS) analysis.³⁴ This regression technique makes the evaluation process more direct; indicating not only what molecular properties should be varied but also in what part of the molecule appropriate changes can be introduced.

Herein we present a SMD based approach for evaluation of the putative T3S inhibitor **1** (Fig. 1) that was identified as a singleton in a screening campaign in viable bacteria.²² Synthesis and evaluation of a focused library of analogues of **1** allowed establishment of a multivariate QSAR model of inhibitors of T3S in *Y. pseudotuberculosis*.

2. Results and discussion

2.1. Establishment of SAR

Compound **1** (Fig. 1) has earlier been identified as a single hit within its class in a screening campaign for T3S inhibitors. Further experiments confirmed that this compound selectively targeted the T3S without affecting bacterial growth although the exact bacterial target is unknown.²² A set of seven analogues with small variations in the structure was designed to examine this singleton's potential as a T3S inhibitor (Table 1, **2–8**). The 2-arylsulfonylamino-benzanilides were synthesized via a three-step synthesis from arylsulfonylchlorides, anilines, and 2-nitrobenzoic acids (Scheme 1, Section 2.5 and Section 4) and evaluated in a previously described reporter-gene assay.^{22,24} The chlorine in the 5-position of the benzoic acid is important since substitution to hydrogen (Table 1, **5** and **6**) resulted in non-active compounds. However, shifting the chlorine from 5- to 4-position did not affect activity (Table 1, cf. **1** and **2**) nor did the change to a 4, 5-difluoro derivative (Table 1, **8**). Substitution of both chlorines to hydrogens at the aniline part (Table 1, **4**) resulted in an almost non-active compound, while if the chlorine in 3-position was kept (Table 1, **7**), the activity was not affected to any large extent. The thiadiazole group on the arylsulfonyl-group could be removed without any detrimental effect on the activity (Table 1, **3**). In conclusion, the small variations made in the initial compound set affected the biological activity where the aniline and sulfonylamino groups appear to tolerate modifications while electron-withdrawing substituents seemed to be beneficial on the central benzoic acid. This basic knowledge obtained with compounds **1–8** was considered in the SMD when selecting building blocks for the next generation of a designed library.

2.2. Design by selection in principal property space

The SAR study of the eight analogues of **1** (**2–8**, Table 1) formed the basis for the design of the second set of com-

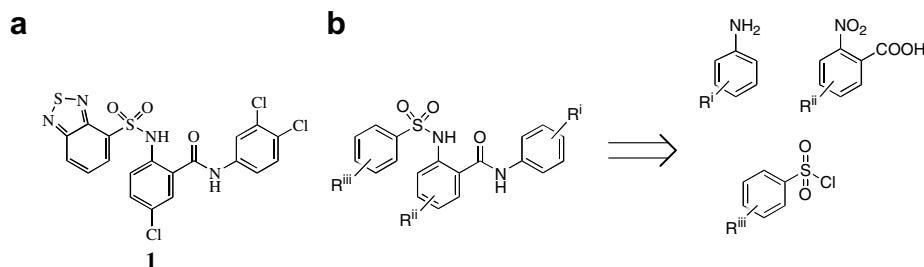


Figure 1. (a) Compound **1** identified as a single hit in its class and (b) a general representation of 2-arylsulfonylamino-benzanilides, constructed of anilines, 2-nitrobenzoic acids and arylsulfonylchlorides.

pounds. SMD was performed in a two-step process where a first selection was carried out in the building block space followed by a second design in the product space based on principal properties from a principal component analysis (PCA) of molecular descriptors (Fig. 2).

2.3. Selection of building blocks

ChemFinderACX2002Prod³⁵ was used to search for commercially available arylsulfonylchlorides, anilines, and 2-nitrobenzoic acids. Databases were generated and 1D and 2D descriptors were calculated in molecular operating environment (MOE)³⁶ and the data was compressed with PCA.^{37–39} Building blocks were selected by visual inspection of mainly the two first principal components, primarily describing size and lipophilicity/polarity (Supplementary material). The selection was made using the knowledge from the initial SAR study and also taking synthetic feasibility and diversity into consideration. For example, 2-nitrobenzoic acids with substituents in position three and six were excluded in order to avoid potential steric hindrance during the acylations and sulfonylations. To increase the probability to obtain a sufficient number of active compounds within a limited set, only 2-nitrobenzoic acids with electron-withdrawing substituents in position four and/or five were selected (Table 1, cf. **1–8**). Regarding the arylsulfonylchlorides, and anilines a more diverse selection of building blocks were made (Supplementary material) since they appeared to tolerate modifications (see above). The selection process resulted in 12 arylsulfonyl chlorides, 17 anilines, and three 2-nitrobenzoic acids (Supplementary material).

2.4. Library selection

The selected building blocks were combined to all possible virtual compounds resulting in 612 compounds for which 1D and 2D molecular descriptors were calculated. The data was compressed with PCA and descriptors with no contribution to the model were excluded resulting in a model with five significant principal components describing 89% of the variation (see Supplementary material for complete list of descriptors). The final compounds to synthesize were chosen by visual inspection of the score plots showing the first, second and third principal component ($R^2 = 0.74$), mainly describing size, polarity, and lipophilicity, respectively. The selection

was focused to include compounds in the vicinity to compounds known to be active according to the initial SAR study and to complement it with more diverse representatives. The total selection resulted in 19 derivatives of 2-arylsulfonylamino-benzanilides subjected for synthesis and biological evaluation (Table 1, **9–27**).

2.5. Chemistry

The 2-arylsulfonylamino-benzanilides (Table 1, **2–27**) were synthesized in three steps from arylsulfonylchlorides, anilines, and 2-nitrobenzoic acids (Scheme 1). The 2-nitrobenzanilide was formed by reacting the 2-nitrobenzoic acid with aniline in toluene with PCl_3 under microwave-assisted heating at 150 °C for 15 min or by reacting the 2-nitrobenzoic acid with aniline using HATU (*O*-(7-azabenzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluoro-phosphate) and DIEA (diisopropylethylamine) in DMF (dimethylformamide) overnight at room temperature.⁴⁰ Generally it was observed that for anilines with ionizable or basic atoms, the reaction was performed in DMF with HATU (Table 1, **19–22**). The nitro group was successfully reduced to the corresponding amine with $\text{H}_2\text{PO}_2\text{Na}$ ⁴¹ or hydrazine and FeCl_3 .⁴² Generally the hydrazine/ FeCl_3 method was used when reducing 2-nitrobenzanilides halogenated in position four on the benzoic acid and also later on due to the observation that the corresponding amine was obtained in higher purity. Reacting the resulting 2-amino-benzanilide with an arylsulfonylchloride in pyridine with DMAP (*N,N*-dimethylaminopyridine) under microwave-assisted heating at 100 °C for 20 min gave the final compound. After each step the products were purified with flash chromatography generally with CH_2Cl_2 as eluent. All intermediates were analyzed with LC–MS and brought forward without further characterization. The target compounds were analyzed and characterized with ¹H NMR, ¹⁹F NMR if relevant, and LC–MS (Section 4 and Supplementary material). The total yields were 5–74% over three steps. In total, a focused library composed of totally 27 compounds was generated (Table 1).

2.6. Biological evaluation

All compounds were tested for their ability to inhibit the T3S in *Y. pseudotuberculosis*. We have previously described a screening system based on viable *Y. pseudotuberculosis* for the identification of potential T3S

Table 1. Effect on luciferase light emission for strain YPIII-pIB29 (*yopE-luxAB*) in the presence of compounds 1–27 at different concentrations

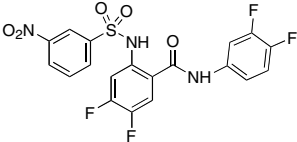
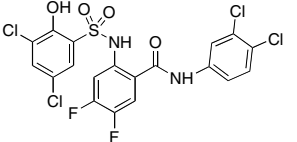
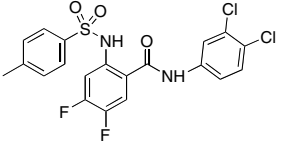
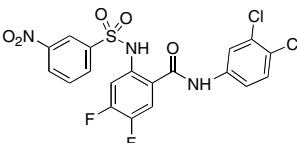
No.	Compound ^a	% Inhibition of light emission ^b			
		Concentration in μM			
		100	50	20	10
1		62 ± 4	69 ± 2	68 ± 8	48 ± 11
2		68 ± 5	71 ± 3	61 ± 2	35 ± 9
3		70 ± 8	73 ± 5	70 ± 8	58 ± 7
4		20 ± 1	9 ± 1	1 ± 2	-2 ± 4
5		13 ± 3	11 ± 1	10 ± 2	5 ± 4
6		18 ± 6	12 ± 3	5 ± 2	-1 ± 1
7		55 ± 2	47 ± 4	13 ± 2	3 ± 2
8		80 ± 0	66 ± 4	39 ± 2	9 ± 1
9		— ^c	—	—	—
10		88 ± 6	79 ± 3	43 ± 4	13 ± 2
11		87 ± 5	86 ± 4	50 ± 2	15 ± 4
12		63 ± 3	50 ± 6	17 ± 2	9 ± 5

(continued on next page)

Table 1 (continued)

No.	Compound ^a	% Inhibition of light emission ^b Concentration in μM			
		100	50	20	10
13		— ^c	—	—	—
14		72 ± 2	70 ± 3	33 ± 8	11 ± 3
15		51 ± 6	32 ± 12	9 ± 4	5 ± 0
16		72 ± 3	72 ± 1	67 ± 2	43 ± 3
17		15 ± 5	10 ± 4	7 ± 2	4 ± 3
18		69 ± 2	74 ± 1	66 ± 1	37 ± 2
19		1 ± 8	2 ± 7	1 ± 4	0 ± 0
20		7 ± 8	8 ± 7	4 ± 1	3 ± 1
21		10 ± 3	7 ± 1	2 ± 2	0 ± 2
22		52 ± 1	59 ± 1	51 ± 3	18 ± 2
23		75 ± 2	71 ± 3	51 ± 1	22 ± 3

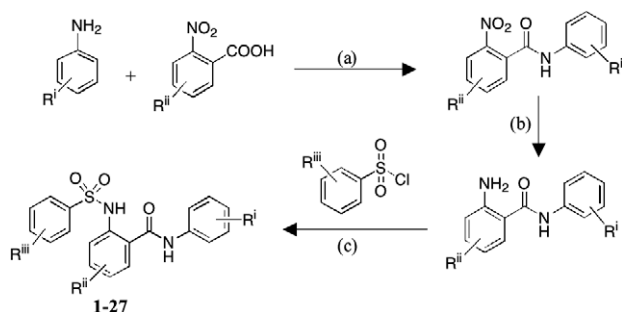
Table 1 (continued)

No.	Compound ^a	% Inhibition of light emission ^b Concentration in μM			
		100	50	20	10
24		49 \pm 2	23 \pm 5	2 \pm 1	-2 \pm 1
25		92 \pm 2	92 \pm 3	85 \pm 5	69 \pm 4
26		2 \pm 5	1 \pm 2	9 \pm 3	57 \pm 6
27		90 \pm 1	79 \pm 3	33 \pm 3	9 \pm 3

^a The compounds were prepared as described in Section 4.

^b Mean values and standard deviations (calculated with the Gauss approximation formula) are from triplicates, and experiments were reproduced at least twice.

^c Compounds showed insufficient solubility.



Scheme 1. Reagents and conditions: (a) PCl_3 , toluene, 150 $^\circ\text{C}$, 15 min, mwi or HATU, DMF, diisopropylethylamine, rt, o.n.; (b) $\text{H}_2\text{PO}_2\text{Na}$, Pd/C, THF/water or N_2H_4 , FeCl_3 , C (s), MeOH; (c) DMAP, pyridine, 100 $^\circ\text{C}$, 10 min, mwi.

inhibitors from chemical libraries in a high-throughput mode.^{22,24} The compounds (Table 1) were screened for the ability to inhibit light emission originating from the luciferase activity of the LuxAB protein produced from a *yopE-luxAB* transcriptional fusion.^{43,44} The luciferase activity is directly correlated to the amount of produced and secreted YopE protein.⁴⁵ The evaluation was carried out with the non-virulent YPIII-pIB29 (*yopE-LuxAB*) strain of *Y. pseudotuberculosis* lacking the effector protein YopH. Cultures were grown in growth medium lacking Ca^{2+} , conditions allowing for induction of Yop secretion by activation of T3S system when cultures are shifted from 26 to 37 $^\circ\text{C}$. Compounds were added to the medium in 96-well plates followed by medium with

bacteria and the plates were incubated for 1 h at room temperature followed by 2 h at 37 $^\circ\text{C}$ to allow full induction of the T3S system.^{43,46} The luciferase activity was measured within 30 min after the 2-h incubation and inhibition was calculated as % inhibition of the resulting light signal compared to control. Several compounds showed inhibition of the light emission and some of them proved to be more potent or equal to compound 1 (Table 1). We have previously shown that compound 1 (Fig. 1) does not affect growth in *Y. pseudotuberculosis*.²² In order to verify that this finding is general, a small set of the most potent compounds (Table 2) was tested for toxicity in a bacterial growth inhibition assay at different concentrations (10, 20, 50, and 100 μM). Wild type YPIII-pIB102 (*yopE-luxAB*) was grown in growth medium supplemented with Ca^{2+} in the presence of compounds (Table 2) and the absorbance at 600 nm was measured during six hours. In the presence of different compounds, the bacteria grew equally well at all concentrations compared to control with the growth curves at 50 μM shown in Figure 3 indicating that the compounds do not exert any general toxicity.

With the reporter-gene screening strategy, there is a risk ending up with compounds inhibiting bacterial growth and also that the compounds might influence luciferase instead of the secretion machinery. To reinforce our previously described screening assay for identification of potential T3S inhibitors in viable *Y. pseudotuberculosis*, an additional assay measuring the enzymatic activity of secreted YopH was added. For analysis of the enzymatic

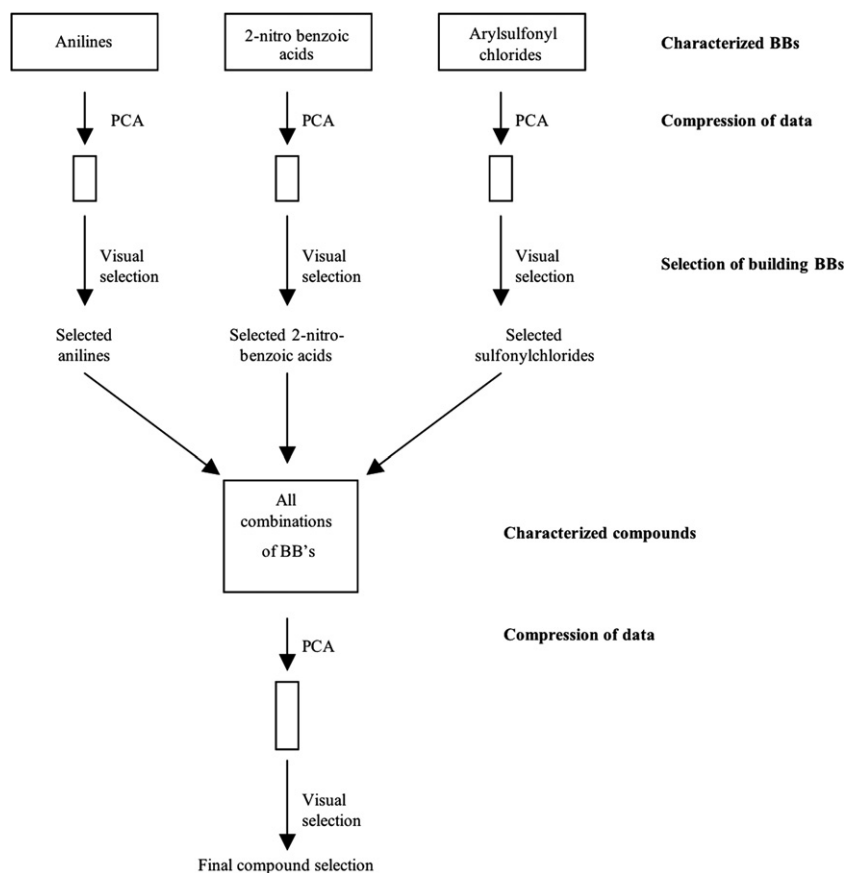


Figure 2. Overview over the design process, from the characterized building blocks to the final compound selections.

activity of YopH, a virulent wild type YPIII-pIB102 (*yopE-luxAB*) strain was used. The effector protein YopH is a protein tyrosine phosphatase that primarily targets focal adhesion proteins and kinases and its function is essential for virulence.⁴⁷ The N-terminal domain of YopH binds directly to tyrosine phosphorylated proteins and the catalytic activity of dephosphorylation can be measured in vitro. Ninety-six-well plates containing bacteria and compounds at different concentrations were grown at 37 °C for 2 h prior to the measurement of the enzymatic activity originating from secreted YopH. Inhibition of the reporter gene signal, and concomitant reduction of YopH activity at compound concentrations with no effect on bacterial growth strongly support that the compounds selectively target the T3S machinery. A representative set of compounds, including one inactive, (Table 2) was examined for the potential to inhibit YopH and it was found that the compounds inhibited both read-out signals to roughly the same extent (cf. Tables 1 and 2). Thus this set of inhibitors is selective for T3S system in *Y. pseudotuberculosis* and has no effect on bacterial growth.

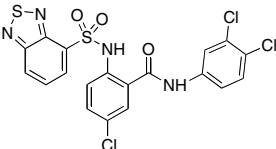
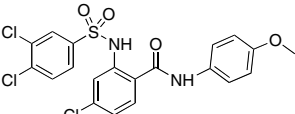
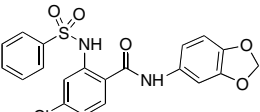
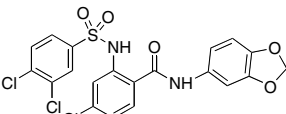
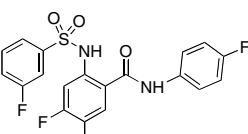
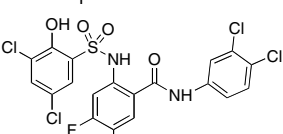
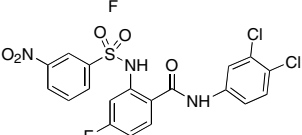
2.7. Multivariate QSAR for T3S inhibition

In total, 27 2-arylsulfonylamino-benzanilides were successfully synthesized and biologically evaluated (Table 1). The inhibitory effect of the compounds was determined as averages from triplicates in a luminescence assay at four concentrations (10, 20, 50, and

100 μM , Table 1) and the experiments were reproduced at least twice. The multivariate QSAR model of the compounds' T3S inhibitory effect was established by using a multi-Y Hi-PLS with the % inhibition at 20 and 50 μM as the response (Fig. 4). Due to insufficient solubility for some compounds, it was not feasible to use the response at 100 μM , and at concentrations below 20 μM several compounds showed low inhibitory effect. The low solubility might also be the underlying reason for compounds not reaching full inhibition despite a clear dose–response at 10, 20, and 50 μM . In total, 10 out of 27 compounds showed an inhibitory effect exceeding 40% at 20 μM concentration (Table 1).

PCA was performed for each class of building blocks, that is, eight arylsulfonylchlorides, 10 anilines, and three 2-nitrobenzoic acids that were present in the designed final compounds (Fig. 5). The three building block sets were investigated separately and structural descriptors representing molecular properties such as charge, size, polarity, and lipophilicity were used (Table 3). A two-component PCA model based on 56 descriptors with $R^2X = 0.84$ was obtained for the sulfonylchlorides where the first principal component described size and the second describe lipophilicity/polarity (Fig. 6). For the anilines, a two-component PCA model with $R^2X = 0.74$ based on 51 descriptors was achieved, where the first principal component described size and the second described polarity/lipophilicity (Fig. 6). Note that even though the second principal component described the same property

Table 2. Inhibition of the enzymatic activity of YopH for strain YPIII-pIB102 (*yopE-luxAB*) in the presence of selected compounds at different concentrations

No.	Compound	% Inhibition of YopH activity ^a			
		Concentration in μM			
		100	50	20	10
1		63 \pm 2	66 \pm 1	62 \pm 2	19 \pm 8
16		34 \pm 3	35 \pm 2	60 \pm 2	35 \pm 7
17		34 \pm 4	25 \pm 2	19 \pm 3	14 \pm 5
18		59 \pm 2	64 \pm 2	60 \pm 3	33 \pm 7
23		70 \pm 1	63 \pm 2	54 \pm 2	27 \pm 10
25		81 \pm 2	80 \pm 1	82 \pm 1	57 \pm 3
27		81 \pm 2	74 \pm 1	67 \pm 2	26 \pm 6

^a Mean values and standard deviations (calculated with the Gauss approximation formula) are from triplicates, and experiments were reproduced at least twice.

(lipophilicity/polarity) for both the anilines and sulfonylchlorides, it is rotated. In the score plots for sulfonylchlorides and anilines, three distinct groupings could be observed. Building blocks were grouped into small hydrophobic, chlorinated, and hydrogen bond accepting building blocks as indicated in score plots (Fig. 6a and c). For the three 2-nitrobenzoic acids, two indicator variables describing the substitution pattern of the halogens were added besides the one-component PCA model based on 39 descriptors with $R^2X = 0.99$.

These five new PCA score vectors describing the principal properties in the three building block sets and the two terms describing the substitution pattern on the benzoic acids were used as **X** variables in a multi-**Y** Hi-PLS with the % inhibition at 20 and 50 μM as the response (Fig. 4). In Hi-PLS the data matrix **X** is divided into sub-blocks, here the characterization of the building block sets, to each of which PCA is applied prior to the

PLS regression.⁴⁸ The main advantage with this approach is that it facilitates the interpretation of the structural influence of the different parts of the molecules on the biological activity. In addition, the PCA offers the possibility to reduce the noise level in **X** before the regression modeling and hence decreases the risk of chance correlations. The regression resulted in a two-component PLS model with an R^2Y of 0.78 for describing the % inhibition at 50 μM ($Q^2 = 0.38$) and an R^2Y of 0.72 for the response at 20 μM ($Q^2 = 0.29$) (Table 4; Model 1). The multivariate QSAR model was established with 22 compounds, from these, 16 were derived from the second set of compounds and six including the hit **1** from the primary SAR investigation. Compound **26** (Table 1) was excluded from the modeling process due to the observed reversed dose–response activity that could not be explained, that is, showing good activity at 10 μM and no activity at higher concentrations. Moreover, **5** and **6** (Table 1) were excluded

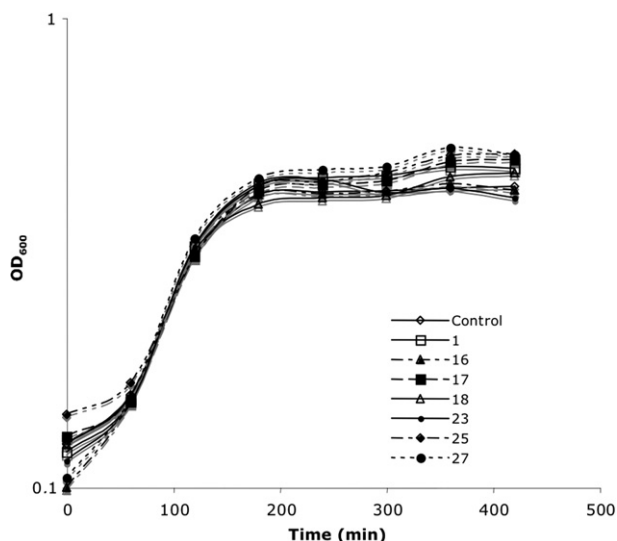


Figure 3. Growth curves for wt YPHII-pIB102 in the presence of different compounds as noted (Tables 1 and 2) at 50 μ M and control supplemented with 1% DMSO.

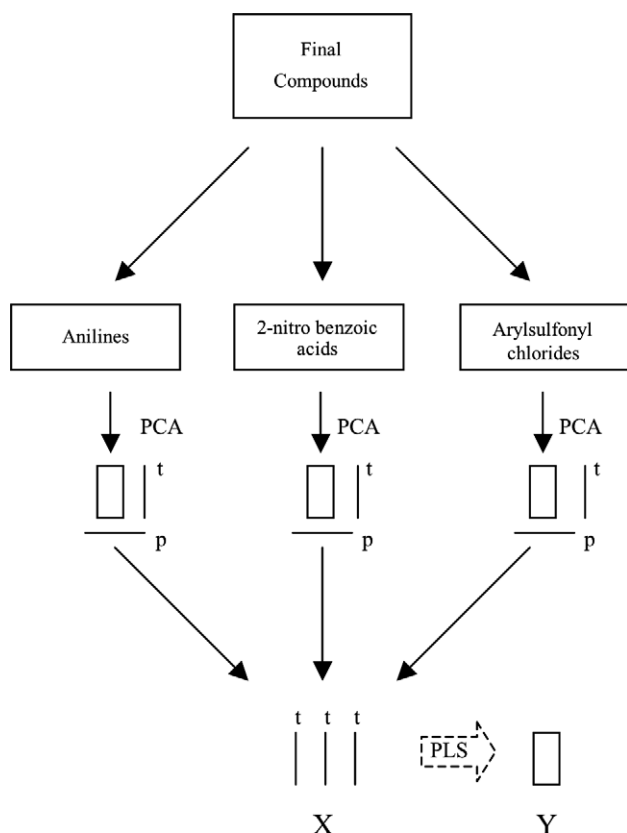


Figure 4. Schematic representation over the process for establishment of the multi-Y Hi-PLS MQSAR model.

from the modeling process in view of the fact that these compounds contained 2-nitrobenzoic acids that were not included in the design of the second set of compounds. Compounds showing insufficient solubility (Table 1, 9 and 13) were also excluded from the modeling process due to missing values.

The normal probability plots of the residuals and correlation plots of the experimentally determined and by the model calculated % inhibition indicated a non-linear behavior of the model (data not shown). Logit transformation of the % inhibition data could not compensate for the non-linear relationship (data not shown). Hence, to deal with non-linearity, square and interaction terms were added to the linear terms followed by exclusion of non-significant terms. This resulted in a two-component Hi-PLS model described by the seven linear terms and 12 interaction terms with $R^2 Y = 0.93$ and $Q^2 = 0.71$ for the response at 50 μ M and $R^2 Y = 0.91$ and $Q^2 = 0.64$ for the response at 20 μ M (Table 4, Model 2). The interpretation of the linear main terms did not differ between these two models but the model was significantly improved according to explained variance and cross-validation with the addition of the square and interaction terms (Table 4, Models 1 and 2). This final Hi-PLS model (Table 5, Model 2) gave a good linear correlation between the experimentally determined and by the model calculated % inhibition with a root mean square of error of estimation (RMSEE) of 8.54 for 50 μ M (Fig. 7) and an RMSEE of 8.91 for 20 μ M.

In all modeling it is important to validate the established models and here a combination of cross-validation and permutation was used for validation of the presented multivariate QSAR models. The permutation test with repetitive randomization of the order of the values in the Y matrix showed that the multivariate QSAR model

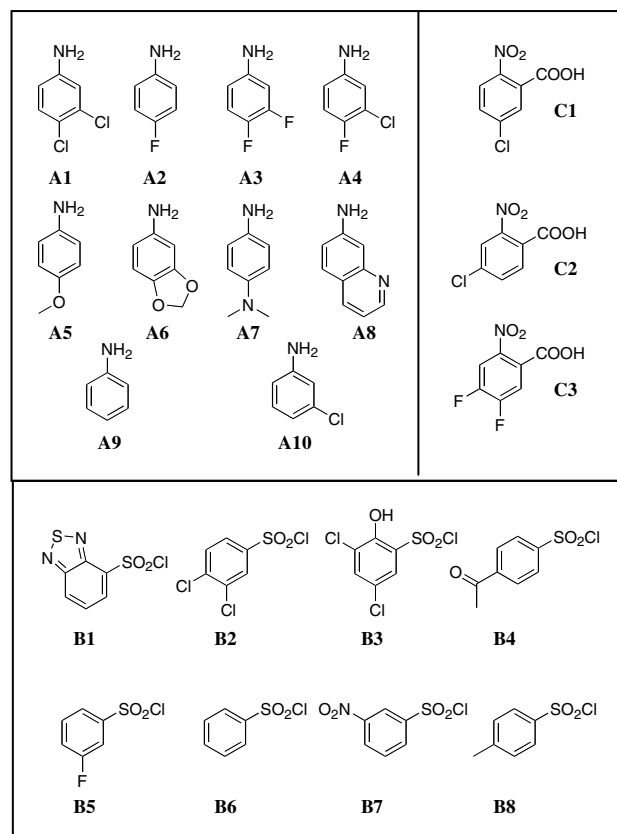


Figure 5. Arylsulfonylchlorides, anilines, and 2-nitrobenzoic acids that were present in the designed product collection and that were included in the Hi-PLS modeling process.

Table 3. Molecular descriptors used in the establishment of multi-Y PLS MQSAR model. Descriptors 1–45 represent 2D descriptors and 46–64 3D descriptors all calculated with MMFF94x force field and partial charges calculated according to Gasteiger–Marsili method

No.	Abbrev.	Explanation	No.	Abbrev.	Explanation
1	Diameter	Size	33	Q_VSA_FPOS	Electronic property/polarity
2	VDistEq	Size/shape	34	Q_VSA_FNEG	Electronic property/polarity
3	VDistMa	Size/shape	35	Q_VSA_FPPOS	Electronic property/polarity
4	weinerPath	Size/shape	36	Q_VSA_FPNEG	Electronic property/polarity
5	weinerPol	Size/shape	37	Q_VSA_FHYD	Electronic property/polarity
6	chi0v	Size/shape	38	Q_VSA_FPOL	Electronic property/polarity
7	chi0v_C	Size/shape	39	Slog P	Lipophilicity
8	chi1v	Size/shape	40	SMR	Size
9	Weight	Size	41	TPSA	Lipophilicity
10	chi0	Size/shape	42	density	Size
11	chi1	Size/shape	43	vdw_area	Lipophilicity
12	VAdjEq	Size/shape	44	vdw_vol	Lipophilicity
13	VAdjMa	Size/shape	45	log P(o/w)	Lipophilicity
14	zagreb	Size/shape	46	ASA	Lipophilicity
15	balabanJ	Size/shape	47	dens	Size
16	Kier1	Size/shape	48	pmi	Shape
17	Kier2	Size/shape	49	vol	Size/Polarity
18	Kier3	Size/shape	50	VSA	Lipophilicity
19	KierA1	Size/shape	51	ASA+	Lipophilicity/electronic property
20	KierA2	Size/shape	52	ASA–	Lipophilicity/electronic property
21	KierA3	Size/shape	53	ASA_H	Lipophilicity
22	KierFlex	Size/shape	54	ASA_P	Polarity
23	Q_PC+	Charge/polarity	55	DASA	Lipophilicity
24	Q_PC–	Charge/polarity	56	CASA+	Lipophilicity/electronic property
25	Q_RPC+	Charge/polarity	57	CASA–	Lipophilicity/electronic property
26	Q_RPC–	Charge/polarity	58	DCASA	Lipophilicity/electronic property
27	Q_VSA_POS	Charge/polarity	59	F ASA+	Lipophilicity/electronic property
28	Q_VSA_NEG	Charge/polarity	60	F ASA–	Lipophilicity/electronic property
29	Q_VSA_PPOS	Charge/polarity	61	FC ASA+	Lipophilicity/electronic property
30	Q_VSA_PNEG	Charge/polarity	62	FCASA–	Lipophilicity/electronic property
31	Q_VSA_HYD	Charge/polarity	63	FASA_H	Lipophilicity
32	Q_VSA_POL	Charge/polarity	64	FASA_P	Lipophilicity

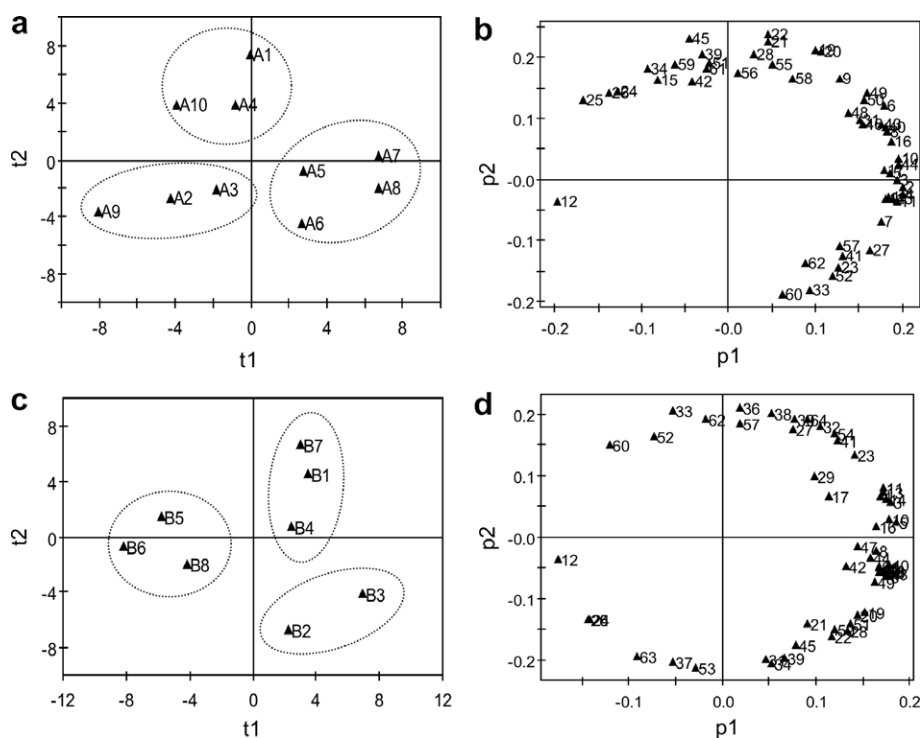
**Figure 6.** (a) Score plot for anilines A1–A10 (Fig. 5) with corresponding (b) loading plot. (c) Scoreplot for arylsulfonylchlorides B1–B8 (Fig. 5) with corresponding (d) loading plot. Explanation of the used molecular descriptors and their numbering in the loading plots (b) and (d) are given in Table 3.

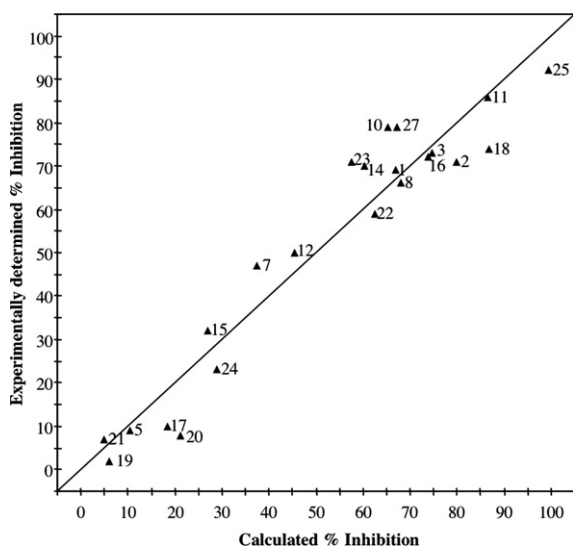
Table 4. Statistical data of the two-component multi-Y Hi-PLS models for T3S inhibitors with the % inhibition at 20 and 50 μM as the responses based on the principal properties of the individual building blocks

Compound concentration (μM)	Model 1 ^a			Model 2 ^b		
	R^2	Q^{2c}	Q^{2d}	R^2	Q^{2c}	Q^{2d}
50	0.78	0.38	0.43	0.93	0.71	0.73
20	0.72	0.29	0.34	0.91	0.64	0.63

^a Model 1: seven linear terms included.^b Model 2: seven linear terms and 12 interaction terms included.^c Cross-validation using eight rounds.^d Cross-validation with leave-one-out method.**Table 5.** Explanation of the Hi-PLS model derived and interpretation of the most important factors for anilines and sulfonylchlorides

Building block set	Score vector	Main property	High value	Low value
Anilines	$t1_A$	Size/shape	Large	Small
Anilines	$t2_A$	Lipophilicity/polarity	Lipophilic	Polar
Sulfonylchlorides	$t1_B$	Size/shape	Large	Small
Sulfonylchlorides	$t2_B$	Lipophilicity/polarity	Polar	Lipophilic

The abbreviation of score vectors refers to the coefficients in Figure 8 which can be correlated to the score and loading plots in Figure 6.

**Figure 7.** Experimentally determined % inhibition versus calculated % inhibition predicted by the two-component multi-Y Hi-PLS model (Model 2, Table 4) at 50 μM concentration of different compounds.

holds R^2 and Q^2 values significantly higher than the R^2/Q^2 values observed for the randomized models. The R^2 values of the permuted models had on average a value of 0.46 for non-correlated Y s and a negative Q^2 for both responses. The rather high R^2 value of the randomized models seen here has also been observed for others performing QSAR modeling^{49,50} where it was suggested that the Q^2 value is more appropriate when it comes to estimating the validity of the models.^{49,50} In addition, the use of eight cross-validation rounds and leave-one out method gave similar satisfying results (Table 4, cf. c and d) and no outliers were observed according to DModX. Moreover, a good correlation between the model at 50 and 20 μM was observed when comparing the regression coefficients to each other with a $R^2 = 0.96$, that is, the model has the same interpretation for the different responses. The correlation matrix of the terms included in the model showed that some of the

non-linear terms were confounded with each other and hence the interpretation was mainly focused on the linear terms.

Interpretation of the values of the regression coefficients showed that the variations made in the aniline and sulfonylchloride parts had the largest effect on the inhibition of T3S (Fig. 8). In particular, the size and lipophilicity of building blocks in the sulfonylchloride set are important for good activity as indicated with high absolute $t1$ and $t2$ values (Fig. 8 and Table 5). Traced back to the original descriptors, $t1$ was mainly described by size dependent descriptors such as molecular weight, surface area, and density (Fig. 6c and d). In the second component, $t2$, the sulfonylchlorides are mainly separated based on descriptors like $\log P$ describing lipophilicity (Fig. 6c and d). A subsequent inspection of the score plot (Fig. 6c) revealed that the building blocks located in the lower right quadrant were the ones that gave the highest activity, that is, sulfonylchlorides having two chlorine substituents on the phenyl ring (Fig. 5).

The structure–activity relationship for the aniline part was dominated by the hydrophobicity of the building blocks as indicated by the high positive value of the regression coefficient of $t2$ (Fig. 8 and Table 5). Building blocks A are separated by size dependent descriptors such as diameter, volume and surface area in the first component, $t1$, and in the second component, $t2$, by lipophilicity descriptors like $\log P$ and $S\log P$ (Fig. 6a and b). The score plot (Fig. 6a) reveals that it again is the halogenated building blocks that were responsible for the highest inhibition effect on the T3S (Fig. 5). The variation introduced on the benzoic acid part of the molecule did not seem to influence the inhibitory effect of the product, as implied by the small regression coefficient values for the variables describing this part (Fig. 8).

The modeling results show that the inhibitory effect of this class of molecules is strongly driven by the hydro-

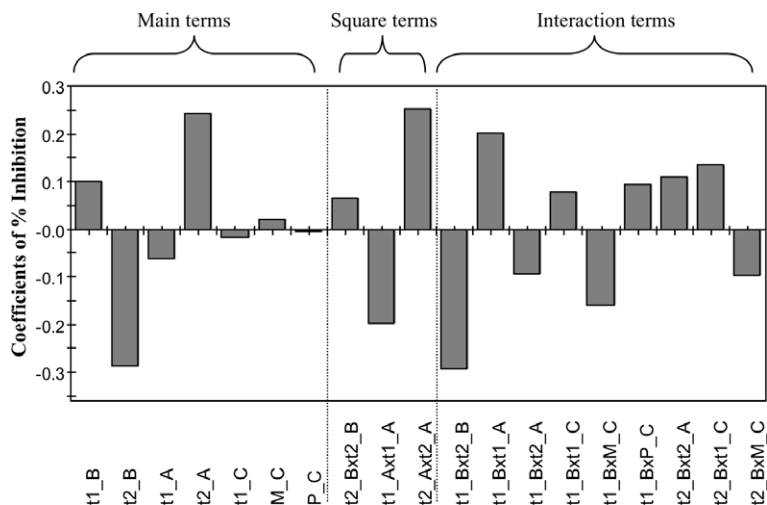


Figure 8. Coefficient plot for the two-component multi-Y Hi-PLS model at 50 μM with seven main terms and 12 interaction terms. A corresponds to the anilines, B to sulfonylchlorides, and C to the 2-nitrobenzoic acids.

phobicity of the molecules and that a high degree of halogenations, for example, chlorine substitution, is preferred.

3. Conclusions

Compound **1** was initially identified in a phenotype-based reporter-gene assay and although additional studies indicated T3S specificity, the target protein is unknown.²² Screening and evaluation using assays based on living bacteria have many significant advantages but also drawbacks. By employing a living bacterium as assay system, compound screening and evaluation are carried out with the agent that causes the disease and inhibitors with different modes of action can be identified. Hurdles including membrane penetration, efflux, and bacterial metabolism are directly addressed. However, establishment of QSARs becomes a challenge since structural changes might affect not only target interactions. In this study, we evaluated a compound set based on **1** in order to establish multivariate QSAR models to explore the potential for compound **1** as a starting point for development of T3S inhibitors as virulence blocking drugs. In total, 27 compounds were prepared and evaluated and the final Hi-PLS model showed good correlation between experimentally determined % inhibition and the calculated % inhibition of the reporter-gene signal at 50 μM (Fig. 7). However, it should not be ignored that the structural differences between the compounds are relatively small and the 2-phenylsulfonylaminobenzanilide scaffold was conserved in all molecules. The model indicates several factors that have a positive effect on activity. Hydrophobicity is important and substitution patterns including chlorines are beneficial. The fact that the compounds in general are lipophilic with $\log P$, 3.94–6.23 indicate limitations regarding the possibility to obtain potent compounds with sufficient solubility. However electron-withdrawing substituents like chlorine, in particular on the central aromatic ring, also influence the electronic distribution and thus the acidity of the sulfonylamide. At pH 7.0,

the sulfonylamide in compound **1** has a calculated pK_a ⁵¹ of 6.2 and thus a large fraction of the molecules might be deprotonated and thus charged at physiological pH. The pH of the BHI medium used in the assays is 7.0–7.4. Although substitution with halogens contributes to lipophilicity, it will also affect charge and thus solubility. Furthermore, it cannot be excluded that the molecules actually bind to the target in the charged form.

Despite the inherent challenges with data obtained from assays using live bacteria, a good model was obtained indicating that this strategy is powerful in development of virulence inhibitors targeting T3S.

4. Experimental

4.1. Characterization and selection of building blocks and synthetic targets

ChemFinderACX2002Prod³⁵ was searched for commercially available arylsulfonylchlorides, anilines, and 2-nitrobenzoic acids. The searches generated several thousands of anilines, approximately five hundred arylsulfonylchlorides and about hundred 2-nitrobenzoic acids. Compounds with additives and a molecular weight >200 were removed resulting in 512 anilines, 129 arylsulfonylchlorides, and 25 2-nitrobenzoic acids. All molecules were energy minimized with the implemented MMFF94 force field in MOE software³⁶ and characterized with 1D and 2D molecular descriptors.

PCA was used to compress the characterized building block sets into their respective main ‘principal’ structures, that is, to describe the information in the structural descriptors using a smaller number of new uncorrelated variables called principal components. These principal components consist of scores (T) that describe the main variation of the data and loadings (P), that express the relationship between scores and the original characterized building blocks.^{32,37–39} The

principal component analysis was performed in the Simca™ 10.5 and 11.0⁵² and Evince Beta software.⁵³

The molecular descriptors for the characterized compounds used in the PCA were selected based on their contribution to the model and descriptors with negative Q^2 values were excluded. The number of significant principal components used was selected based on their eigenvalues with a cut-off value >2. Building blocks were selected by cherry picking from the first and second component score plots. The first component described size and the second mainly hydrophobicity for their respective building blocks. The selection of building blocks was made considering the information gained from the primary SAR investigation but also with the purpose to make a diverse selection. From the initial building block libraries, a selection was made resulting in 12 arylsulfonylchlorides, 17 anilines, and three 2-nitrobenzoic acids (Supporting material). These were combined to 612 possible virtual compounds that were energy minimized in the MMFF94 force field in MOE³⁶ and characterized with 1D and 2D molecular descriptors. The data were compressed with PCA and variables with negative Q^2 values were excluded resulting in a five-component model with R^2 value of 0.89 (see Supporting material for complete list of variables). Compounds to synthesize were chosen by visual inspection of the score plots in the first, second, and third component mainly described by size, polarity, and lipophilicity, respectively. In the cherry picking of compound for synthesis, the selection was focused to include compounds in the vicinity of compounds known to be active from the SAR study but also taking diversity into account. The final selection resulted in 19 compounds for synthesis.

4.2. Characterization of building blocks and multivariate QSAR

The building blocks were created in the MOE software³⁶ and energy minimized with the implemented MMFF94 force field and partial charges were calculated with the Gasteiger–Marsali method.⁵⁴ The individual building blocks were characterized with actively chosen 1D, 2D, and 3D molecular descriptors known to describe important features like size, lipophilicity, flexibility, shape, and electronic properties.³⁶

The multivariate regression method Hi-PLS was used to relate the data matrix (X , descriptors) to the activity matrix (Y , biological response). PLS regression is a commonly used method for QSAR analysis that maximizes the covariance between the data matrix and the response.^{55,56} The PCA were performed in the Simca™ 10.5 and 11.0 and the PLS modeling were performed using the Simca software 10.5 and 11.0.⁵²

PCA was used to compress the characterized building block sets into a few number of orthogonal principal components or scores (T) describing the main variation of the data.^{37–39} The molecular descriptors for the characterized compounds used in the PCA were selected based on their contribution to the model and descriptors

with no contribution were excluded (see Supporting material for complete list). Two indicator variables were added to the score vectors to describe the halogens' substitution pattern (4- and 5-substitution) on the central fragment, the 2-nitrobenzoic acids.

In the multi- Y Hi-PLS, the extracted principal properties (scores, T) from the PCAs of the individual building blocks and the two indicator variables were used as the X matrix and % inhibition of the reporter gene signal as the response (Y). In the Hi-PLS modeling all linear terms were used, and square and interaction terms were added to handle non-linear behavior of the model. The use of a compression method prior to the PLS analysis, that is, Hi-PLS, led to substantially fewer, and hence more manageable number of square and interaction-terms than for standard PLS.

The number of interaction and square terms was reduced based on their coefficient values, with an absolute cut-off value <0.075. The number of significant components in Hi-PLS was decided by cross-validation using two independent cross-validation rounds (eight classes and leave-one out method). The Q^2 was calculated according to:

$$Q^2 = 1 - \frac{\text{PRESS}}{\text{SS}},$$

where PRESS is the predicted error sum of squares when all objects have been left out once and SS is the total sum of squares of Y corrected for the mean.

4.3. Model and design validation

The quality of Hi-PLS models was analyzed through their Q^2 values, calculated by internal validation using eight cross-validation rounds and the leave-one-out method.⁵² To study the significance of the models, their RMSEE values and the observations distance to the model in X space (DModX) were examined. Moreover, permutation tests were performed to investigate the significance of the Hi-PLS models. In these tests, the order of the responses, Y , was randomly permuted 50 times and the procedure was repeated 35 times allowing for calculation of mean values for the R^2 and Q^2 values intercepting zero.⁵⁰ The plot of the correlation coefficient between the original and permuted Y versus the cumulative R^2 and Q^2 gives a regression line where the intercept (R^2 and Q^2 when the correlation coefficient is zero) is an estimate of the significance of the model.^{50,52}

The quality of the SMD was evaluated in Modde⁵² by analysis of the correlation matrix for the coefficients in the final Hi-PLS model.

4.4. Model interpretation

Interpretation of the Hi-PLS model, i.e. what molecular properties that influence the biological activity, was made through two levels. First it was based on the PLS model and secondly through the PCA score vectors, T , with its corresponding loading vectors, P , that originate from the compression of the original structural descriptors. The influence of the variables' importance

for good activity was based on the regression coefficient values, where variables with large coefficient values (both positive and negative) contribute much to the activity whereas small do not. Here, this means that PCA score vectors with large positive coefficient values were positively correlated to the biological response, while those with large negative coefficient values had a negative effect on the % inhibition. To understand which molecular properties that were linked to the PCA score vectors, T , their corresponding loading vectors, P , were investigated in more detail, which is discussed in more detail in Section 2.

4.5. Biological evaluation, strains and growth conditions

The biological evaluation was performed with *Y. pseudotuberculosis* serotype III (YPIII) strains, pIB102 (*yopE-luxAB*) and pIB29 (*yopE-luxAB*). The different strains were grown at room temperature on LB plates as described elsewhere.^{22,44}

4.6. Reporter-gene assay

The experimental procedures were carried out essentially as described before.^{22,24} The *Y. pseudotuberculosis* strain YPIII-pIB29 (*yopE-luxAB*) was grown overnight in brain heart infusion (BHI) medium containing 5 mM EGTA and 20 mM MgCl₂ for calcium depletion at ambient temperature. The optical density at 600 nm was adjusted to 0.15–0.25 in BHI. In parallel, the compounds solved in DMSO to be tested were dispensed in four different concentrations (100, 50, 20, and 10 μM final concentration) to the plates with a pipetting robot (Hydra) in triplicates. All compounds were dissolved in DMSO and the final DMSO concentration in the biological assays was kept at 1% or below. To each well 50 μl of diluted bacteria was added with a dispenser instrument (Multi Drop, Thermo Lab Systems). The plates were incubated at ambient temperature on a rotary shaker for 1 h followed by incubation at 37 °C for 2 h. Within 30 min after incubation, 50 μl of fresh decanal solution (10 μl/100 ml water) was added to the wells and the luciferase activity was measured immediately in a micro-plate reader (TECAN GENios, Gain 150; Integration time 20 ms). Experiments were reproduced in at least three independent experiments.

4.7. YopH phosphatase assay

The *Y. pseudotuberculosis* strain YPIII-pIB102 (*yopE-luxAB*) was grown overnight in brain heart infusion (BHI) medium with 25 μg/ml chloramphenicol (Clm) on a rotary shaker at 26 °C. The optical density at 600 nm was adjusted to 0.08 in calcium depleted BHI (5 mM EGTA, 20 mM MgCl₂) and 100 μl of the diluted bacteria was added to white 96-well plate (FluoroMunc, Nunc) and different concentrations of the substances (100, 50, 20, and 10 μM) were added. The plate was incubated on a rotary shaker at 37 °C. After 2 h, 10 μl from every well was transferred to a new transparent plate containing 90 μl aqueous substrate mixture (25 mM *p*-nitrophenyl-phosphate, 40 mM 2-

(*N*-morpholino) ethanesulfonic acid pH 5.0, 1.6 mM DL-1,4-dithiothreitol). The plate was incubated for 15 min on a rotary shaker at 37 °C and the reaction was stopped by the addition of 20 μl aqueous 1 M NaOH. The inhibition of YopH secretion was quantified by absorbance measurement at 405 nm in a micro-plate reader (TECAN GENios, Gain 150). The experiment was reproduced twice in independent experiments.

4.8. Growth inhibition experiments

The experimental procedures were carried out essentially as described before.^{22,24} Growth inhibition was measured by growing bacteria at 37 °C in the presence of different compound concentrations in 96-well plates containing 100 μl bacterial culture medium diluted to an OD₆₀₀ = 0.05–0.1 in BHI medium with 2.5 mM CaCl₂. The experiment was carried out either in a Molecular Device Spectramax 340 plate reader with continuous shaking at 37 °C and a periodic reading of absorption at 600 nm or by reading of absorbance at 595 nm in a TECAN GENios plate reader once an hour and shaking carried out in a rotary shaker at 37 °C. Growth rates were followed for up to 24 h.

4.9. Chemistry

¹H NMR spectra were recorded on a Bruker DRX-400 in CDCl₃ [residual chloroform-*d*₁ (δ_H 7.26 ppm) as internal standard], ((CD₃)₂CO) [residual acetone-*d*₆ (δ_H 2.05 ppm) as internal standard] and ((CD₃)₂SO) [residual dimethylsulfoxide-*d*₅ (δ_H 2.50 ppm) as internal standard]. ¹⁹F NMR spectra were recorded on a Bruker DRX-400 in ((CD₃)₂CO) with 0.1% CFCl₃ (δ_H 0 ppm) as internal standard. Mass-spectra were recorded by detecting positive (ES⁺) and negative (ES⁻) molecular ions with an electro spray Waters Micromass ZG 2000 instrument using an XTerra[®] MS C₁₈ 5 μm 4.6*50 mm column and an H₂O/acetonitrile eluent system. The same LC-system was also used for purification with a preparative XTerra[®] Prep MS C₁₈ 5 μm 19*50 mm column and an H₂O/acetonitrile eluent system.

Methanol and pyridine were dried over 3 Å molecular sieves. All laboratory vessels were dried in oven (100 °C) before coming in contact with hydrolysis sensitive chemicals. Organic solutions were dried over Na₂SO₄ before being concentrated. Amberlite IR-120 was washed with water and methanol. Microwave heated reactions were carried out in Emrys[™] process vials (2–5 ml) using a SmithCreator[™] microwave instrument from Biotage. TLC was performed on Silica gel 60 F₂₅₄ (Merck) with detection by UV-light and staining with Indanetroin hydrate 95 parts 0.2% in butanol with 5 parts 10% acetic acid for detection of amines.

4.9.1. Typical procedures for the formation of 2-nitrobenzanilides. *Method A:* 2-Nitrobenzoic acid (1 equiv, 0.250 g) and aniline (1.2 equiv) were suspended in toluene (4 ml) and stirred for 5 min followed by addition of PCl₃ (0.7 equiv) and the reaction was carried out in a microwave instrument at 150 °C for 15 min. The

resulting solution was diluted with methanol (25 ml) and toluene (25 ml). Excess aniline was removed from the solution by treatment with Amberlite IR-120 (40 ml) for 40 min followed by filtration through a layer of silica. To remove excess 2-nitrobenzoic acid, the solution was washed with a 20% NaHCO₃ solution once and H₂O twice. The organic phase was dried over MgSO₄ (s), filtrated, and concentrated. The products were purified by flash chromatography typically with CH₂Cl₂ as eluent.

Method B: 2-Nitrobenzoic acid (1 equiv, 0.200 g) and HATU (*O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) (1 equiv) were dissolved in DMF (10 ml) followed by diisopropylethylamine and the resulting solution was stirred for 15 min at rt. The aniline (1 equiv) was added and stirring was continued overnight at rt. Solution was transferred to a separating funnel and diethyl ether (20 ml) was added. The organic phase was washed with H₂O and NaHCO₃ (10%, aq) to remove DMF and unreacted 2-nitrobenzoic acid. The organic phase was dried over MgSO₄ (s), filtrated, and concentrated. The product was purified by flash chromatography, typically with heptane/EtOAc as eluent.

4.9.2. Typical procedures for reduction of 2-nitrobenzanilides. **Method C:** 2-Nitrobenzanilide (1 equiv) was dissolved in THF (20 ml) and Pd/C (0.100 g) was added followed by H₂PO₂Na (8 equiv) dissolved in H₂O (10 ml). The reaction mixture was refluxed, and the reduction was run to completion within 20 min. Pd/C was filtered off and the product was extracted with CH₂Cl₂. The organic phase was dried over MgSO₄ (s), filtrated, and concentrated. The product was purified by flash chromatography, typically with CH₂Cl₂ as eluent.

Method D: 2-Nitrobenzanilide (1 equiv), FeCl₃ (0.3 equiv), active carbon (0.015 g per 0.100 g starting material) was suspended in MeOH (10 ml). The solution was heated to 65 °C and hydrazine monohydrate (21 equiv) was added dropwise. The reaction mixture was refluxed 3–6 h. Activated carbon was filtered off and solvent was removed by rotary evaporation. The product was dissolved in EtOAc (10 ml) and washed with H₂O. The organic phase was dried over MgSO₄ (s), filtrated, and concentrated. The product was purified by flash chromatography, typically with CH₂Cl₂ as eluent.

4.9.3. Typical procedure for the formation of 2-arylsulfonylamino-benzanilides. **Method E:** The 2-aminobenzanilide (1 equiv) and sulfonyl chloride (1.3 equiv) were dissolved in pyridine (4 ml) followed by 4-dimethylaminopyridine (1 equiv) and the reaction was performed in a microwave instrument at 100 °C for 20 min. Reaction mixture was transferred to a vessel using toluene (25 ml) and MeOH (25 ml) and Amberlite IR-120 (40 ml) was added to remove pyridine, DMAP and unreacted 2-aminobenzanilide. The solution was stirred for 40 min, filtered and the organic phase was dried over MgSO₄ (s), filtrated, and concentrated. The product was purified by flash chromatography resulting in products with generally >90% purity if not otherwise stated.

4.9.3.1. 2-(Benzo[1,2,5]thiadiazole-4-sulfonylamino)-5-chloro-*N*-(3,4-dichloro-phenyl)-benzamide (1). ¹H NMR ((CD₃)₂CO) δ 10.55 (br s, 1H), 9.74 (br s, 1H), 8.28–8.33 (m, 2H), 8.07 (d, 1H, *J* = 2.0 Hz), 7.86 (dd, 1H, *J* = 8.8, 7.2 Hz), 7.77 (d, 1H, *J* = 8.8 Hz), 7.64–7.70 (m, 3H). LC–MS [M–H][–] calcd *m/z* 506.69, found: 506.98.

4.9.3.2. 2-(Benzo[1,2,5]thiadiazole-4-sulfonylamino)-4-chloro-*N*-(3,4-dichloro-phenyl)-benzamide (2). The substance was synthesized according to procedure B, D, and E. ¹H NMR ((CD₃)₂CO) δ 10.88 (br s, 1H), 9.71 (br s, 1H), 8.41 (d, 1H, *J* = 7.2 Hz), 8.32 (d, 1H, *J* = 9.2 Hz), 8.10 (br s, 1H), 7.89 (dd, 1H, *J* = 8.4 and 7.2 Hz), 7.79 (d, 1H, *J* = 2.0 Hz), 7.69–7.74 (m, 2H), 7.65 (d, 1H, *J* = 8.8 Hz), 7.08–7.17 (m, 1H). LSMS [M–H][–] calcd *m/z* 511.93, found 511.63.

4.9.3.3. 2-Benzenesulfonylamino-5-chloro-*N*-(3,4-dichloro-phenyl)-benzamide (3). The substance was synthesized according to procedure B, C, and E. ¹H NMR ((CD₃)₂CO) δ 10.33 (s, 1H), 10.10 (br s, 1H), 8.06 (d, 1H, *J* = 2.4 Hz), 7.87 (d, 1H, *J* = 2.4 Hz), 7.80–7.82 (m, 2H), 7.67–7.73 (m, 2H), 7.57–7.62 (m, 3H), 7.47–7.51 (m, 2H). LC–MS [M–H][–] calcd *m/z* 453.97, found: 453.66.

4.9.3.4. 2-(Benzo[1,2,5]thiadiazole-4-sulfonylamino)-5-chloro-*N*-phenyl-benzamide (4). The substance was synthesized according to procedure A, C, and E. ¹H NMR ((CD₃)₂CO) δ 10.76 (s, 1H), 9.49 (s, 1H), 8.33 (dd, 1H, *J* = 7.2, 0.8 Hz), 8.29 (d, 1H, *J* = 8.8 Hz), 7.85 (dd, 1H, *J* = 8.8, 7.2 Hz), 7.74–7.79 (m, 3H), 7.67–7.69 (m, 1H), 7.45–7.75 (m, 3H), 7.24–7.28 (m, 1H). LC–MS [M–H][–] calcd *m/z* 444.01, found 443.74.

4.9.3.5. 2-(Benzo[1,2,5]thiadiazole-4-sulfonylamino)-*N*-(3,4-dichloro-phenyl)-benzamide (5). The substance was synthesized according to procedure A, C, and E. ¹H NMR (CDCl₃) δ 10.64 (s, 1H), 9.56 (s, 1H), 8.29 (d, 1H, *J* = 6.8), 8.24 (d, 1H, *J* = 9.2), 8.06 (d, 1H, *J* = 2.4), 7.81 (dd, 1H, *J* = 8.78, 1.74), 7.74 (d, 1H, *J* = 8.23), 7.66–7.70 (m, 1H), 7.59–7.66 (m, 2H), 7.44 (t, 1H, *J* = 7.96), 7.08 (t, 1H, *J* = 7.31). LC–MS [M–H][–] calcd 476.96, found 477.02.

4.9.3.6. 2-(Benzo[1,2,5]thiadiazole-4-sulfonylamino)-*N*-(4-chloro-phenyl)-benzamide (6). The substance was synthesized according to procedure A, C, and E. ¹H NMR (CDCl₃) δ 10.69 (s, 1H), 9.47 (s, 1H), 8.29 (d, 1H, *J* = 7.13), 8.24 (d, 1H, *J* = 8.78), 7.91 (s, 1H), 7.88 (dd, 1H, *J* = 8.87, 1.83), 7.74 (d, 1H, *J* = 8.14), 7.65–7.56 (m, 2H), 7.47–7.40 (m, 2H), 7.25 (d, 1H, *J* = 7.77), 7.07 (t, 1H, *J* = 7.59). LC–MS [M–H][–] calcd *m/z* 443.01, found 443.00.

4.9.3.7. 2-(Benzo[1,2,5]thiadiazole-4-sulfonylamino)-5-chloro-*N*-(3-chloro-phenyl)-benzamide (7). The substance was synthesized according to procedure A, C, and E. ¹H NMR (CDCl₃) δ 10.90 (s, 1H), 9.57 (s, 1H), 8.37 (d, 1H, *J* = 7.0 Hz), 8.28 (d, 1H, *J* = 9.0 Hz), 7.90–7.83 (m, 2H), 7.76 (s, 1H), 7.67 (d, 1H, *J* = 8.3 Hz), 7.61 (d, 1H, *J* = 8.2 Hz), 7.45 (t, 1H, *J* = 8.1), 7.26 (d, 1H,

$J = 8.0$ Hz) 7.12 (dd, 1H, $J = 8.5, 1.8$ Hz). LC–MS $[M-H]^-$ calcd m/z 476.97, found 477.02.

4.9.3.8. 2-(Benzo[1,2,5]thiadiazole-4-sulfonylamino)-*N*-(3,4-dichloro-phenyl)-4,5-difluoro-benzamide (8). The substance was synthesized according to procedure A, C, and E. 1H NMR ($(CD_3)_2SO$) δ 10.62 (br s, 1H), 10.41 (br s, 1H), 8.34 (d, 1H, $J = 8.3$ Hz), 8.25 (d, 1H, $J = 6.8$ Hz), 7.75–7.85 (m, 3H), 7.64 (d, 1H, $J = 8.8$ Hz), 7.52–7.60 (m, 1H), 7.46 (dd, 1H, $J = 11.0$ and 8.8 Hz). ^{19}F NMR ($(CD_3)_2SO$) δ 141.4 (d, 1F, $J = 4.3$ Hz), –130.5 (d, 1F, $J = 18.7$ Hz). LC–MS $[M-H]^-$ calcd m/z 513.95, found 513.60. 85% pure according to ^{19}F NMR.

4.9.3.9. 4-Chloro-2-(3,4-dichloro-benzenesulfonylamino)-*N*-(4-fluoro-phenyl)-benzamide (9). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 6.68 (d, 1H, $J = 2.0$ Hz), 6.63 (d, 1H, $J = 8.4$ Hz), 6.46 (dd, 1H, $J = 8.4, 2.0$ Hz), 6.39 (dd, 1H, $J = 8.8, 4.8$ Hz), 6.30 (d, 1H, $J = 8.4$ Hz), 6.11 (d, 1H, $J = 2.0$ Hz), 5.80 (t, 2H, $J = 8.8$ Hz), 5.54 (dd, 1H, $J = 8.4, 2.0$ Hz). ^{19}F NMR ($(CD_3)_2SO$) δ 119.7 (s, 1F). LC–MS $[M-H]^-$ calcd m/z 470.95, found 470.89.

4.9.3.10. 5-Chloro-2-(3,5-dichloro-2-hydroxy-benzenesulfonylamino)-*N*-(4-fluoro-phenyl)-benzamide (10). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 10.79 (br s, 1H), 9.84 (br s, 1H), 7.98 (s, 1H), 7.84 (d, 1H, $J = 2.4$ Hz), 7.74–7.77 (m, 2H), 7.69 (t, 1H, $J = 2.4$ Hz), 7.66 (s, 1H), 7.62 (d, 1H, $J = 2.4$ Hz), 7.52 (dd, 1H, $J = 8.9, 6.5$ Hz), 7.13–7.18 (m, 1H). ^{19}F NMR ($(CD_3)_2CO$) δ 118.1 (s, 1F). LC–MS $[M-H]^-$ calcd m/z 486.95, found 486.67.

4.9.3.11. 4-Chloro-2-(3,5-dichloro-2-hydroxy-benzenesulfonylamino)-*N*-(3,4-difluoro-phenyl)-benzamide (11). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 11.28 (br s, 1H), 10.02 (br s, 1H), 7.97–8.04 (m, 1H), 7.95 (d, 1H, $J = 8.8$ Hz), 7.50–7.57 (m, 2H), 7.35–7.43 (m, 2H), 7.25–7.34 (m, 1H), 6.95 (dd, 1H, $J = 8.4, 2.0$ Hz). ^{19}F NMR ($(CD_3)_2CO$) δ 145.0 (d, 1F, $J = 22.8$ Hz), –137.5 (d, 1F, $J = 20.8$ Hz). LC–MS $[M-H]^-$ calcd m/z 504.94, found 504.74.

4.9.3.12. 2-(Benzo[1,2,5]thiadiazole-4-sulfonylamino)-4-chloro-*N*-(3,4-difluoro-phenyl)-benzamide (12). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 10.93 (s, 1H), 9.85 (br s, 1H), 8.42 (d, 1H, $J = 6.8$ Hz), 8.31 (d, 1H, $J = 8.8$ Hz), 7.84–7.94 (m, 2H), 7.67 (s, 1H), 7.63 (d, 1H, $J = 8.4$ Hz) 7.47–7.56 (m, 1H), 7.33–7.44 (m, 1H), 7.04–7.14 (m, 1H). ^{19}F NMR ($(CD_3)_2CO$) δ 143.3 (br s, 1F), –137.0 (d, 1F, $J = 22.8$ Hz). LC–MS $[M-H]^-$ calcd m/z 478.91, found 478.82.

4.9.3.13. 5-Chloro-*N*-(3-chloro-4-fluoro-phenyl)-2-(3,4-dichloro-benzenesulfonylamino)-benzamide (13). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 10.16 (br s, 1H), 9.81 (br s, 1H), 7.94 (dd, 1H, $J = 6.8, 2.4$ Hz), 7.87

(s, 1H), 7.82 (d, 1H, $J = 2.4$ Hz), 7.61–7.69 (m, 5H), 7.34 (t, 1H, $J = 8.8$ Hz). ^{19}F NMR ($(CD_3)_2CO$) δ 120.7 (s, 1F). LC–MS $[M-H]^-$ calcd m/z 504.92, found 504.64.

4.9.3.14. 2-(4-Acetyl-benzenesulfonylamino)-4-chloro-*N*-(3-chloro-4-fluoro-phenyl)-benzamide (14). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 11.03 (br s, 1H), 10.66 (s, 1H), 7.93–8.01 (m, 5H), 7.84 (d, 1H, $J = 8.4$ Hz), 7.57–7.63 (m, 2H), 7.22 (t, 1H, $J = 9.2$ Hz), 6.98 (d, 1H, $J = 8.0$ Hz), 2.54 (s, 3H). ^{19}F NMR ($(CD_3)_2CO$) δ 121.8 (s, 1F). LC–MS $[M-H]^-$ calcd m/z 479.00, found 478.82.

4.9.3.15. 4-Chloro-2-(3-fluoro-benzenesulfonylamino)-*N*-(4-methoxy-phenyl)-benzamide (15). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 11.19 (br s, 1H) 9.62 (br s, 1H), 7.81–7.88 (m, 1H), 7.63–7.70 (m, 2H), 7.53–7.58 (m, 4H), 7.35–7.42 (m, 1H), 7.24 (dd, 1H, $J = 8.4, 2.0$ Hz), 6.95 (d, 2H, $J = 9.2$ Hz), 3.80 (s, 3H). ^{19}F NMR ($CDCl_3$) δ 111.6 (s, 1F). LC–MS $[M-H]^-$ calcd m/z 433.04, found 432.84.

4.9.3.16. 4-Chloro-2-(3,4-dichloro-benzenesulfonylamino)-*N*-(4-methoxy-phenyl)-benzamide (16). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 11.08 (br s, 1H), 9.62 (br s, 1H), 7.92 (d, 1H, $J = 2.0$ Hz), 7.85 (d, 1H, $J = 8.5$ Hz), 7.63–7.74 (m, 3H), 7.57 (d, 2H, $J = 9.0$ Hz), 7.28 (dd, 1H, $J = 8.5, 2.1$ Hz), 6.95 (d, 2H, $J = 9.1$ Hz), 3.81 (s, 3H). LC–MS $[M-H]^-$ calcd m/z 482.97, found 482.76.

4.9.3.17. 2-Benzenesulfonylamino-*N*-benzo[1,3]dioxol-5-yl-4-chloro-benzamide (17). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 11.06 (s, 1H), 9.68 (s, 1H), 7.84–7.87 (m, 3H), 7.73 (d, 1H, $J = 2.0$ Hz), 7.61–7.66 (m, 1H), 7.52–7.56 (m, 2H), 7.35–7.36 (m, 1H) 7.23 (dd, 1H, $J = 8.4, 2.0$ Hz) 7.09–7.12 (m, 1H), 6.89 (d, 1H, $J = 8.4$ Hz), 6.06 (s, 2H). LC–MS $[M-H]^-$ calcd m/z 429.03, found: 428.72.

4.9.3.18. *N*-Benzo[1,3]dioxol-5-yl-4-chloro-2-(3,4-dichloro-benzenesulfonylamino)-benzamide (18). The substance was synthesized according to procedure A, D, and E. 1H NMR ($(CD_3)_2CO$) δ 10.89 (br s, 1H), 9.65 (br s, 1H), 7.92 (d, 1H, $J = 2.0$ Hz), 7.82 (d, 1H, $J = 8.4$ Hz), 7.65–7.72 (m, 3H), 7.27–7.31 (m, 2H), 7.06 (dd, 1H, $J = 8.4, 2.0$ Hz), 6.84 (d, 1H, $J = 8.4$ Hz), 6.03 (s, 2H). LC–MS $[M-H]^-$ calcd m/z 496.95, found 496.70.

4.9.3.19. 2-Benzenesulfonylamino-5-chloro-*N*-(4-dimethylamino-phenyl)-benzamide (19). The substance was synthesized according to procedure B, C, and E. 1H NMR ($(CD_3)_2CO$) δ 10.89 (br s, 1H), 9.62 (br s, 1H), 7.85 (s, 1H), 7.81 (d, 2H, $J = 7.6$ Hz), 7.74 (d, 1H, $J = 8.8$ Hz), 7.51–7.64 (m, 4H), 7.45 (t, 2H, $J = 8.0$ Hz), 7.07 (br s, 2H), 3.02 (s, 6H). LC–MS $[M-H]^-$ calcd m/z 428.08, found 427.76.

4.9.3.20. 5-Chloro-*N*-(4-dimethylamino-phenyl)-2-(3-nitro-benzenesulfonylamino)-benzamide (20). The substance was synthesized according to procedure B, C, and E. ^1H NMR ($(\text{CD}_3)_2\text{CO}$) δ 10.93 (br s, 1H), 9.51 (br s, 1H), 8.45–8.48 (m, 1H), 8.36–8.44 (m, 1H), 8.12–8.20 (m, 1H), 7.70–7.83 (m, 3H), 7.57–7.64 (m, 1H), 7.48 (d, 2H, $J = 9.2$ Hz), 6.78 (d, 2H, $J = 9.2$ Hz), 2.98 (s, 6H). LC-MS $[\text{M}-\text{H}]^-$ calcd m/z 473.07, found 472.73. 75% pure according to ^1H NMR.

4.9.3.21. 2-Benzenesulfonylamino-5-chloro-*N*-quinolin-7-yl-benzamide (21). The substance was synthesized according to procedure B, C, and E. ^1H NMR ($(\text{CD}_3)_2\text{SO}$) δ 10.71 (br s, 1H), 10.50 (br s, 1H), 8.82–8.89 (m, 1H), 8.40–8.47 (m, 2H), 8.03 (d, 1H, $J = 9.1$ Hz), 7.85–7.94 (m, 2H), 7.75 (d, 2H, $J = 7.6$ Hz), 7.52–7.62 (m, 3H), 7.45–7.52 (m, 2H), 7.38 (d, 1H, $J = 8.8$ Hz). LC-MS $[\text{M}-\text{H}]^-$ calcd m/z 436.05, found 435.74.

4.9.3.22. 5-Chloro-2-(3,4-dichloro-benzenesulfonylamino)-*N*-quinolin-7-yl-benzamide (22). The substance was synthesized according to procedure B, C, and E. ^1H NMR ($(\text{CD}_3)_2\text{SO}$) δ 10.19–10.96 (m, 2H), 8.83 (dd, 1H, $J = 4.0, 1.6$ Hz), 8.33–8.43 (m, 2H), 8.00 (d, 1H, $J = 8.8$ Hz), 7.91 (d, 1H, $J = 1.9$ Hz), 7.80–7.87 (m, 2H), 7.68 (d, 1H, $J = 7.9$ Hz), 7.51–7.63 (m, 3H), 7.34 (d, 1H, $J = 8.8$ Hz). LC-MS $[\text{M}-\text{H}]^-$ calcd m/z 503.97, found 503.65.

4.9.3.23. 4,5-Difluoro-2-(3-fluoro-benzenesulfonylamino)-*N*-(4-fluoro-phenyl)-benzamide (23). The substance was synthesized according to procedure A, C, and E. ^1H NMR ($(\text{CD}_3)_2\text{CO}$) δ 10.72 (br s, 1H), 9.66 (br s, 1H), 7.89 (t, 1H, $J = 9.8$ Hz), 7.87–7.93 (m, 1H), 7.66–7.69 (m, 2H), 7.59–7.65 (m, 2H), 7.50–7.57 (m, 2H), 7.34–7.40 (m, 1H), 7.14–7.19 (m, 2H). ^{19}F NMR (CDCl_3) δ 111.6 (s, 1F), –119.3 (s, 1F), –131.6 (d, 1F, $J = 23.6$ Hz), –143.7 (d, 1F, $J = 23.2$ Hz). LC-MS $[\text{M}-\text{H}]^-$ calcd m/z 423.04, found 422.77.

4.9.3.24. *N*-(3,4-Difluoro-phenyl)-4,5-difluoro-2-(3-nitro-benzenesulfonylamino)-benzamide (24). The substance was synthesized according to procedure A, C, and E. ^1H NMR ($(\text{CD}_3)_2\text{CO}$) δ 10.40 (br s, 1H), 9.69 (br s, 1H), 8.48 (d, 1H, $J = 2.0$ Hz), 8.38 (ddd, 1H, $J = 8.0, 2.0, 0.8$ Hz), 8.16–8.18 (m, 1H), 7.70–7.87 (m, 3H), 7.58–7.67 (m, 1H), 7.29–7.36 (m, 2H). ^{19}F NMR (CDCl_3) δ 131.4 (d, 1F, $J = 22.8$ Hz), –138.6 (d, 1F, $J = 22.0$ Hz), –142.4 (d, 1F, $J = 22.4$ Hz), –144.6 (d, 1F, $J = 22.4$ Hz). LC-MS $[\text{M}-\text{H}]^-$ calcd m/z 468.03, found 467.74.

4.9.3.25. 2-(3,5-Dichloro-2-hydroxy-benzenesulfonylamino)-*N*-(3,4-dichloro-phenyl)-4,5-difluoro-benzamide (25). The substance was synthesized according to procedure A, C, and E. ^1H NMR ($(\text{CD}_3)_2\text{CO}$) δ 10.36 (br s, 1H), 9.91 (br s, 1H), 8.06 (d, 1H, $J = 2.8$ Hz), 7.85–7.93 (m, 1H), 7.73 (d, 1H, $J = 2.6$ Hz), 7.52–7.66 (m, 4H). ^{19}F NMR (CDCl_3) δ 126.2 (d, 1F, $J = 22.4$ Hz), –139.3 (d, 1F, $J = 22.4$ Hz). LC-MS $[\text{M}-\text{H}]^-$ calcd m/z 538.90, found 538.67.

4.9.3.26. *N*-(3,4-Dichloro-phenyl)-4,5-difluoro-2-(toluene-4-sulfonylamino)-benzamide (26). The substance was synthesized according to procedure A, C, and E. ^1H NMR ($(\text{CD}_3)_2\text{CO}$) δ 10.27 (br s, 1H), 9.69 (br s, 1H), 7.97 (d, 1H, $J = 2.2$ Hz), 7.81–7.89 (m, 1H), 7.55–7.65 (m, 5H), 7.25 (d, 2H, $J = 8.4$ Hz), 2.28 (s, 3H). ^{19}F NMR (CDCl_3) δ 131.5 (d, 1F, $J = 24.0$ Hz), –144.0 (d, 1F, $J = 23.2$ Hz). LC-MS $[\text{M}-\text{H}]^-$ calcd m/z 469.00, found 468.76.

4.9.3.27. *N*-(3,4-Dichloro-phenyl)-4,5-difluoro-2-(3-nitro-benzenesulfonylamino)-benzamide (27). The substance was synthesized according to procedure A, C, and E. ^1H NMR ($(\text{CD}_3)_2\text{CO}$) δ 10.37 (br s, 1H), 9.69 (br s, 1H), 8.49 (s, 1H), 8.39 (d, 1H, $J = 8.0$ Hz), 8.18 (d, 1H, $J = 8.0$ Hz), 7.93 (s, 1H), 7.87 (dd, 1H, $J = 10.9, 8.4$ Hz), 7.79 (t, 1H, $J = 8.1$ Hz), 7.64 (dd, 1H, $J = 11.6, 7.2$ Hz), 7.56 (s, 2H). ^{19}F NMR (CDCl_3) δ 131.6 (d, 1F, $J = 22.0$ Hz), –142.4 (d, 1F, $J = 24.0$ Hz). LC-MS $[\text{M}-\text{H}]^-$ calcd m/z 499.97, found 499.58.

Acknowledgments

We thank the Swedish National Research Council, the Foundation for Technology Transfer in Umeå, the Carl Trygger foundation and Umeå Biotech Incubator for support.

Supplementary data

HPLC chromatograms and ^1H and ^{19}F NMR spectra for compound 1–27. Structures of building blocks used in the design. Score and loading plots used for building block selection and descriptor set, score and loading plots used for selection of synthetic targets. Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2007.07.047.

References and notes

1. Projan, S. J. *Curr. Opin. Microbiol.* **2003**, *6*, 427–430.
2. Poole, K. *Curr. Opin. Microbiol.* **2001**, *4*, 500–508.
3. Levy, S. B. *Adv. Drug Deliv. Rev.* **2005**, *57*, 1446–1450.
4. Goldschmidt, R. M.; Macielag, M. J.; Hlasta, D. J.; Barrett, J. F. *Curr. Pharma. Des.* **1997**, *3*, 125–142.
5. Lee, Y. M.; Almqvist, F.; Hultgren, S. J. *Curr. Opin. Pharmacol.* **2003**, *3*, 513–519.
6. Walsh, C. *Nat. Rev. Microbiol.* **2003**, *1*, 65–70.
7. Alksne, L. E.; Projan, S. J. *Curr. Opin. Biotech.* **2000**, *11*, 625–636.
8. Alekshun, M. N.; Levy, S. B. *Drug Disc. Tod.* **2004**, *1*, 483–489.
9. Salyers, A. A.; Whitt, D. D. *Bacterial Pathogenesis A Molecular Approach*, 2 ed.; ASM Press: Washington, DC, USA, 2002.
10. Hung, D. T.; Shakhnovich, E. A.; Pierson, E.; Mekalanos, J. J. *Science* **2005**, *310*, 670–674.
11. Becker, D.; Selbach, M.; Rollenhagen, C.; Ballmaier, M.; Meyer, T. F.; Mann, M.; Bumann, D. *Nature* **2006**, *440*, 303–307.
12. Ward, G. E.; Carey, K. L.; Westwood, N. J. *Cell Microbiol.* **2002**, *4*, 471–482.

13. Hueck, C. J. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 379–433.
14. Muller, S.; Feldman, M. F.; Cornelis, G. R. *Expert Opin. Ther. Targets* **2001**, *5*, 327–339.
15. Gauthier, A.; Finlay, B. B. *Asm News* **2002**, *68*, 383–387.
16. Cornelis, G. R.; Boland, A.; Boyd, A. P.; Geuijen, C.; Iriarte, M.; Neyt, C.; Sory, M. P.; Stainier, I. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 1315–1352.
17. Wren, B. W. *Nat. Rev. Microbiol.* **2003**, *1*, 55–64.
18. Perry, R. D.; Fetherston, J. D. *Clin. Microbiol. Rev.* **1997**, *10*, 35–66.
19. Normark, S.; Nilsson, C.; Normark, B. H. *Science* **2005**, *307*, 1211–1212.
20. Henriques Normark, B.; Normark, S. *J. Int. Med.* **2002**, *252*, 91–106.
21. Rosqvist, R.; Hakansson, S.; Forsberg, A.; Wolf-Watz, H. *EMBO J.* **1995**, *14*, 4187–4195.
22. Kauppi, A. M.; Nordfelth, R.; Uvell, H.; Wolf-Watz, H.; Elofsson, M. *Chem. Biol.* **2003**, *10*, 241–249.
23. Gauthier, A.; Robertson, M. L.; Lowden, M.; Ibarra, J. A.; Puente, J. L.; Finlay, B. B. *Antimicrob. Agents Chemother.* **2005**, *49*, 4101–4109.
24. Nordfelth, R.; Kauppi, A. M.; Norberg, H. A.; Wolf-Watz, H.; Elofsson, M. *Infect. Immun.* **2005**, *73*, 3104–3114.
25. Kauppi, A. M.; Nordfelth, R.; Hagglund, U.; Wolf-Watz, H.; Elofsson, M. *Adv. Exp. Med. Biol.* **2003**, *529*, 97–100.
26. Muschiol, S.; Bailey, L.; Gylfe, Å.; Sundin, C.; Hultenby, K.; Bergström, S.; Elofsson, M.; Wolf-Watz, H.; Normark, S.; Henriques-Normark, B. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 14566–14571.
27. Wolf, K.; Betts, H. J.; Chellas-Géry, B.; Hower, S.; Linton, C. N.; Fields, K. A. *Mol. Microbiol.* **2006**, *61*, 1543–1555.
28. Linusson, A.; Gottfries, J.; Lindgren, F.; Wold, S. *J. Med. Chem.* **2000**, *43*, 1320–1328.
29. Linusson, A.; Gottfries, J.; Olsson, T.; Örnskov, E.; Folestad, S.; Norden, B.; Wold, S. *J. Med. Chem.* **2001**, *44*, 3424–3439.
30. Hansch, C. *Acc. Chem. Res.* **1969**, *2*, 232–239.
31. Wold, S.; Dunn, W. J., III. *J. Chem. Inform. Model.* **1983**, *23*, 6–13.
32. Eriksson, L.; Johansson, E. *Chemom. Intell. Lab. Syst.* **1996**, *34*, 1–19.
33. Larsson, A.; Johansson, S.; Pinkner, J. S.; Hultgren, S. J.; Almqvist, F.; Kihlberg, J.; Linusson, A. *J. Med. Chem.* **2005**, *48*, 935–945.
34. Eriksson, L.; Johansson, E.; Lindgren, F.; Sjöström, M.; Wold, S. *J. Comput. Aided Mol. Des.* **2002**, *16*, 711–726.
35. CambridgeSoft Cambridge, USA, 2002.
36. MOE; 2004.3 ed. Montreal, Quebec, Canada.
37. Wold, H. *Estimation of Principal Components and Related Models by Iterative Least Squares*; Academic Press: New York, 1966.
38. Wold, S.; Esbensen, K.; Geladi, P. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37–52.
39. Jackson, J. E. *A User's Guide to Principal Components Analysis*; Wiley: New York, 1991.
40. Montalbetti, C. A. G. N.; Falque, V. *Tetrahedron* **2005**, *61*, 10827–10852.
41. Entwistle, I. D.; Jackson, A. E.; Johnstone, R. A. W.; Telford, R. P. *J. Chem. Soc., Perkin Trans. 1* **1977**, 443–444.
42. Gowda, S.; Kempe Gowda, B. K.; Channe Gowda, D. *Synth. Comm.* **2003**, *33*, 281–289.
43. Olsson, O.; Koncz, C.; Szalay, A. A. *Mol. Gen. Genet.* **1988**, *215*, 1–9.
44. Forsberg, A.; Rosqvist, R. *Inf. Agents Dis.-Rev. Iss. Comm.* **1993**, *2*, 275–278.
45. Pettersson, J.; Nordfelth, R.; Dubinina, E.; Bergman, T.; Gustafsson, M.; Magnusson, K. E.; Wolf-Watz, H. *Science* **1996**, *273*, 1231–1233.
46. Forsberg, A.; Wolf-Watz, H. *Mol. Microbiol.* **1988**, *2*, 121–133.
47. Viboud, G. I.; Bliska, J. B. *Ann. Rev. Microbiol.* **2005**, *59*, 69–89.
48. Wold, S.; Trygg, J.; Berglund, A.; Antti, H. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 131–150.
49. Eriksson, L.; Verboom, H. H.; Peijnenburg, W. J. G. M. *J. Chemom.* **1996**, *10*, 483–492.
50. Lindgren, F.; Hansen, B.; Karcher, W.; Sjöström, M.; Eriksson, L. *J. Chemom.* **1996**, *10*, 521–532.
51. <http://www.chemaxon.com/demosite/marvin/index.html>.
52. Umetrics; SIMCA-P+ 10.0, 10.5, 11.0 ed. Box 7960, S-907 19 Umeå, Sweden.
53. UmBio UmBio AB, Umeå, Sweden.
54. Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219–3228.
55. Dunn, W. J., III; Wold, S.; Edlund, U.; Hellberg, S.; Gasteiger, J. *Quant. Struct.-Act. Relat.* **1984**, *3*, 131–137.
56. Höskuldsson, A. *J. Chemom.* **1988**, *2*, 211–228.