The enzyme DXS as an anti-infective target
Masini, Tiziana

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 6

Ligand-based virtual screening as a powerful tool for the discovery of DXS inhibitors with potent antituberculotic activity

In this chapter, we will present a ligand-based virtual screening (LBVS) campaign that we carried out using the structures of our modeled inhibitors (Chapter 2) as reference molecules. After performing several cycles of LBVS by screening part of the ZINC database, we identified low-micromolar inhibitors of both Deinococcus radiodurans and Mycobacterium tuberculosis DXS having diverse scaffolds and with potent activity in cell-based assays against multi- and extensively-drug resistant strains of M. tuberculosis. One of them also showed promising activity against the Plasmodium falciparum 3D7 strain in a preliminary cell-based assay. Moreover, we showed that the most promising hits have a high degree of selectivity of over a mammalian TDP-dependent enzyme.

Part of this chapter is being patented


Number of patent: P107389EP00

Manuscript in preparation
6.1 High throughput screening and virtual screening in drug discovery

High-throughput-screening (HTS) approaches, enabling testing large libraries of compounds in an automated fashion, have become popular over the past two decades mostly in the pharmaceutical industry but increasingly also in academia. Since its introduction in the 1990s, fast advances in automation and miniaturization have rapidly made HTS feasible, rendering it ever faster and efficient: 96-well plates have been replaced by higher density microtiter plates with up to 1586 wells per plate, low-volume assays (most commonly \( \leq 50 \mu L \)) and novel fluorescence-based detection systems have been developed and screening robots have been fully adapted to desktop environments. The main criticism of HTS campaigns, is associated with the size and quality of the libraries that are being screened. Nevertheless, a lot of progress has been made in this field since the 1990s, when the compound collections were assembled without giving much consideration to their suitability for drug discovery. Nowadays, libraries, which are tested in HTS, only include highly pure, drug-like compounds, taking into account parameters such as molecular weight (M\(_w\)) and cLogP. Another notorious problem associated with HTS hits are false positives, arising mainly from promiscuous binders and solubility issues. The fact that assays are typically performed at a single concentration point, might lead to both false positives and false negatives. Despite being the main source of hits entering the drug-discovery pipeline, HTS requires very expensive equipment, which is usually not affordable for academic research. In general, target validation and assay development continue to present major bottlenecks: the availability or stability of the protein over the time requested for the HTS analysis or the kind of assay can make a target incompatible with HTS. For example, if the assay is based on kinetic measurements, the screening would require an enormous amount of instrument time.

Virtual screening (VS), where large databases of compounds are screened \textit{in silico} against a certain target, constitutes a complementary approach to classical HTS. The two main strategies used nowadays are ligand-based virtual screening (LBVS) and structure-based virtual screening (SBVS). Although structures of target proteins, which can be used as VS templates, are becoming increasingly available, LBVS continues to be prevalent, owing to the fact that hit (or lead) information is still the predominant type of knowledge used in numerous projects. Nevertheless, the two strategies are perfectly complementary. When the crystal structure of the protein of interest has been reported, SBVS campaigns can be carried out: a library of compounds is docked into the binding site of the protein and each compound is scored according to how well it fits the pocket in terms of calculated binding energy. On the one hand, SBVS strictly relies on the resolution of the crystal structure of the protein and suffers from the same limitations related to its static nature, namely protein
Ligand-based virtual screening for the discovery of inhibitors of DXS

Flexibility in solution or definition of a certain pocket only upon ligand binding. One the other hand, the only prerequisite for a LBVS campaign is the knowledge of a ligand for the protein of interest. The structure of this reference molecule is compared with a database of compounds according to several criteria, as discussed in Section 6.2. The database can be composed of commercially available molecules or can be an in-house database of compounds, as well as a combinatorial library of compounds. Each molecule of the database is ranked according to its similarity to the reference molecule and the top-scoring compounds are selected and evaluated in vitro for their inhibitory potency. The main advantage of LBVS over SBVS, besides not requiring the structure of the protein to be solved, is that the calculation power required is remarkably lower.

6.2 Ligand-based virtual screening

LBVS relies on the use of descriptors of molecular structures and properties to compare various molecules. As shown schematically in Figure 1, the molecular descriptors can be classified as one- two- or three-dimensional (1D, 2D or 3D) ones.

![Diagram of molecular descriptors](image)

**Figure 1.** Meaning of the 1D, 2D and 3D molecular descriptors used in virtual screening.

The most straightforward and efficient similarity-based calculations are based on 2D molecular descriptors using fingerprints. They consist in binary bit string representations of many aspects and properties of the template molecule, encoding for the presence or absence of given properties. The fingerprint of the template molecule is then compared to those of the molecules present in the database, and the similarity of two different molecules according to their fingerprints is determined using similarity coefficients such as the Tanimoto coefficient. Going a step beyond 2D structures, similarity between molecules can also be analyzed using 3D shape similarities or pharmacophore models.
In this project, two methods developed in the group of Prof. Reymond, at the University of Bern, were applied for the virtual screening of (parts of) the ZINC database.\textsuperscript{14, 15} The first method, called Xfp,\textsuperscript{16} is based on the concept of atom pairs originally proposed by Carhart\textsuperscript{17} and co-workers and relies on atom pair fingerprints counting atom pairs at increasing topological distances in 2D structures, without any atom-property assignment. The hydrophobicity, the number of hydrogen-bond donors and acceptors, and the planarity of the molecule are parameters, which are taken into account during the count of atom pairs. The method was developed to overcome the computational demand associated with precise 3D-shape screening of large databases and it was shown to correlate very well with various representations of molecular shape extracted from 3D structures.

The second method, called translational Ligand Overlap Score (tLOS), compares the 3D shape of two molecules using an algorithm to optimize the spatial overlap between atoms. The fitting function takes into account the hydrophobicity and the hydrogen-bond donors and acceptors of the molecules, which are being compared. tLOS is computationally more demanding than the previously described Xfp, thus being limited to databases with restricted size.\textsuperscript{18, 19}

In the context of this project, we will also apply clustering methods, allowing for the classification of compounds obtained from the screening in groups of molecules according to their similarity.\textsuperscript{20} Each cluster will include a set of molecules with a high degree of diversity (however defined, depending on the algorithm) between the different clusters. This method is very useful when one has the possibility to purchase only a limited number of compounds for testing, but still would like to explore the inhibitory potency of very different chemical scaffolds in the first stage of a project.

### 6.3 First round of LBVS based on known inhibitors for DXS

Fluoropyruvate (1),\textsuperscript{21} ketoclozamone (2),\textsuperscript{22} gibberellic acid (3)\textsuperscript{23} and compound 4\textsuperscript{24} are known inhibitors of various DXS orthologues (Table 1). 2 binds to an unknown binding site of DXS, which differs from the binding site of the two substrates. The binding modes of 3 and 4 are unknown, the latter being identified starting from the scaffold of a known inhibitor of transketolase (TK).\textsuperscript{25} The fact that 4 targets \textit{M. tuberculosis} DXS and its close analogue had shown activity against human TK – both thiamine diphosphate-dependent enzymes – could suggest that 4 targets the cofactor-binding pocket of the two enzymes. 1 is supposed to be a covalent, irreversible inhibitor of DXS, presumably acting in a similar manner to the one shown for the E1 component of the pyruvate dehydrogenase complex (PDH).\textsuperscript{26}
We performed the docking of compounds 1–4 with the FlexX\textsuperscript{27} docking module of the LeadIT suite using the complete structure of \textit{D. radiodurans} DXS (Protein Data Bank code: 2O1X),\textsuperscript{28} trying to identify trends, which could suggest a preferential binding pocket for each scaffold. Careful analysis of the distribution and ranking of the docked poses – according to the scoring function HYDE\textsuperscript{29,30} in the LeadIT suite –, however, did not enable us to draw any conclusions. When testing 2 and 3 for their inhibitory potency against both \textit{D. radiodurans} and \textit{M. tuberculosis} DXS, we observed no activity at 2000 μM. Although there is a high degree of sequence homology between different orthologues of DXS, one often observes tremendous differences in inhibitory potency for the same compound against different orthologues (\textit{e.g.}, IC\textsubscript{50} of 2 with respect to \textit{Arabidopsis thaliana} and \textit{Chlamydomonas} DXS, Table 1).

\textbf{Table 1.} Known inhibitors of DXS with their reported inhibitory potency. We tested 2 and 3 for their inhibitory potency against \textit{D. radiodurans} and \textit{M. tuberculosis} DXS.

<table>
<thead>
<tr>
<th>Known inhibitor</th>
<th>Known inhibition properties</th>
<th>IC\textsubscript{50} (μM)[a]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{[a]} IC\textsubscript{50} values were determined using a photometric assay using the program Dynafit. Full details of the biochemical assay conditions are provided in the Experimental part of Chapter 2. Every IC\textsubscript{50} value reported in this chapter has been determined with this assay, if not stated otherwise.

\textsuperscript{[b]} Whenever no assay was performed, we report not determined (nd) throughout the chapter.
Given that these compounds constituted the only known inhibitors of DXS at the onset of the project, we decided to initiate a LBVS campaign by using them as reference molecules. First, the 3D structures of 1–4 were generated using the software CORINA,31 and their 3D shape compared with those of the commercially available compounds from the Princeton database. The compounds present in this database were ranked according to their similarity towards each reference resulting in four different series. Compounds with undesired functional groups such as ester and hydrazine moieties, were immediately excluded. The ten top-scoring compounds of each series were selected (40 compounds in total). Moreover, the 1000 top-scoring compounds of each series were clustered independently using the k-mean clustering algorithm. 20 to 30 compounds were manually selected in each series by visual inspection of the clusters, and were combined with the previously selected compounds. The derivatives of 1 resulting from the screening did not give sufficiently good scores when compared to the other derivatives. As a result, we decided to focus on the other three scaffolds. The selected compounds were docked within D. radiodurans DXS and the top-ranked poses were selected so as to reduce the number of candidates. Finally, eleven compounds were purchased, dissolved in DMSO at the highest concentration possible (ideally 80 mM) and tested for their inhibitory activity against DXS using the spectrophotometric assay as described in Chapter 2.

Unfortunately, none of them inhibited D. radiodurans nor M. tuberculosis DXS at the highest concentration tested (depending on the solubility of each compound in DMSO and in the assay buffer, concentrations vary from 12 to 2000 μM). One explanation for these disappointing results might be that the subset of compounds tested was too small. Indeed, since a hit rate of 1–3% might be reasonably expected for this kind of approach, biochemical testing of 100–200 compounds might have led to the identification of some hits. Moreover, as discussed above, 2 and 3 used as references for our virtual screening campaign were active on some DXS orthologues (Table 1) but we found them to be inactive against both D. radiodurans and M. tuberculosis DXS.

6.4 First round of LBVS based on three novel scaffolds

In Chapter 2, we introduced deazathiamine (5) as a moderate, thiamine diphosphate (TDP)-competitive inhibitor of D. radiodurans DXS. Moreover, we showed how we designed de novo two fragments, 6 and 7, which inhibit D. radiodurans DXS in the low millimolar and upper micromolar range, respectively. We validated the binding mode of 5 and 6 by NMR studies and we proposed that 7 inhibits DXS with a mixed mode of inhibition, suggesting a binding mode similar to that proposed in Figure 5 (Chapter 2). We decided
to use 5–7 as reference molecules for performing a LBVS campaign aimed at obtaining more chemical diversity while enhancing the *in vitro* inhibitory potencies. Although we found the reference scaffolds 5–7 to be inactive against *M. tuberculosis* DXS, we decided to assess the inhibitory capacity of the hits found by LBVS also against the pathogenic enzyme.

**Table 2.** Inhibitory potency of compounds 5–7, used as reference molecules for our ligand-based virtual screening project, against *D. radiodurans* and *M. tuberculosis* DXS.

<table>
<thead>
<tr>
<th>Compound</th>
<th><em>D. rad. DXS</em></th>
<th><em>M. tub. DXS</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>430 ± 68</td>
<td>&gt;2000</td>
</tr>
<tr>
<td>6</td>
<td>1810 ± 480</td>
<td>&gt;2000</td>
</tr>
<tr>
<td>7</td>
<td>762 ± 199</td>
<td>&gt;2000</td>
</tr>
</tbody>
</table>

First, the 3D shape of 5–7 generated with CORINA was screened against the Princeton database. Clustering of the 1000 compounds with the highest similarity to each reference molecule was performed to support the manual selection. We selected the 20 compounds with the highest similarity score and combined them with the analogues selected by clustering. Visual analysis of the selection was performed so as to explore as much chemical diversity as possible given that this constitutes the first round of screening. This selection helped us in reducing the number of hits to purchase: 21, 25 and 27 analogues of 5, 6 and 7, respectively, were finally purchased in 1 mg batches. Stock solutions in DMSO were prepared for each compound based on the 1 mg batch provided by the company and a preliminary biochemical assay was carried out at one concentration point (in triplicate) or three concentration points. Compounds, which resulted in some inhibitory activity, were then tested with a complete dilution series of eight concentration points and an IC$_{50}$ value was determined. Among the first series of compounds we tested, we could identify nine hits. Given the uncertainty regarding the exact weight of the purchased batches, we repurchased the hits in higher amounts to determine accurate IC$_{50}$ values and confirm their inhibitory potency (Table 3, 8–11). We used this protocol (preliminary biochemical evaluation using the 1 mg batch based on one or three concentration points and repurchasing of promising hits) throughout the whole project. Nevertheless, the values reported in Table 3 will have to be confirmed, given that the biochemical assay was carried out at concentrations, which are very close to the solubility limits of the compounds.
Studies are currently ongoing to solve this problem, in which the biochemical assays are being carried out at higher percentage of DMSO in the assay mixture, namely 10%.

Table 3. Most potent hits from the first round of ligand-based virtual screening starting with 5–7 as reference molecules.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Reference</th>
<th>IC$_{50}$ (μM)</th>
<th>IC$_{50}$ (μM)</th>
<th>IC$_{50}$ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Molecule 8" /></td>
<td>5</td>
<td>17 ± 4</td>
<td>60 ± 5</td>
<td>3 ± 0.4</td>
</tr>
<tr>
<td><img src="image" alt="Molecule 9" /></td>
<td>5</td>
<td>252 ± 72</td>
<td>145 ± 17</td>
<td>&gt;150</td>
</tr>
<tr>
<td><img src="image" alt="Molecule 10" /></td>
<td>6</td>
<td>398 ± 30</td>
<td>317 ± 29</td>
<td>&gt;500</td>
</tr>
<tr>
<td><img src="image" alt="Molecule 11" /></td>
<td>7</td>
<td>1395 ± 150</td>
<td>528 ± 43</td>
<td>&gt;1000</td>
</tr>
<tr>
<td><img src="image" alt="Molecule 12" /></td>
<td>7</td>
<td>&gt;500</td>
<td>455 ± 81</td>
<td>nd$^{[b]}$</td>
</tr>
<tr>
<td><img src="image" alt="Molecule 13" /></td>
<td>6</td>
<td>&gt;120</td>
<td>&gt;120</td>
<td>nd</td>
</tr>
</tbody>
</table>

$^{[a]}$ Preliminary results that have to be validated, given that the highest concentrations used are close to the solubility limit.
Preliminary results on an exemplary compound of the LBVS series, provided the same IC\textsubscript{50} value when the compound was tested in the presence of 5% and 10% DMSO in the assay mixture, suggesting that the data set is reliable and presumably was acquired sufficiently far from the solubility limit. Further assays of this type with the best hits from our LBVS campaign should be carried out to corroborate this assumption. By performing measurements at long wavelength (e. g., 800 nm), one could also monitor whether precipitates or aggregates are being formed. Given the fact that increasing percentage of DMSO could have severe effects on the protein, a dilution series should be carried out to check the effect of multiple DMSO concentrations on the activity of DXS. A fluorescent dye could be used to check whether any aggregates are present in the assay mixture, which could affect the IC\textsubscript{50} determination.

Structure-activity relationships (SARs) of analogues resulting from the whole LBVS campaign, will be discussed in Section 6.7.

We obtained very interesting hits from all the reference molecules, with moderate up to striking enhancements of the inhibitory potency against \textit{D. radiodurans} DXS. Moreover, the best hits reported in Table 3, are also active against \textit{M. tuberculosis} DXS, with the deazathiamine derivative \textit{8} being the most active one, having an IC\textsubscript{50} of 60 ± 5 \textmu M. Given that the spectrophotometric assay that we use to determine IC\textsubscript{50} values uses IspC as an auxiliary enzyme, we also assessed the inhibitory potency of our hits against \textit{Escherichia coli} IspC. \textit{8} was the only compound, which we found to be active against \textit{E. coli} IspC, with an IC\textsubscript{50} of 3 ± 0.4 \textmu M. In this case, a direct assay with DXS should be carried out so as to confirm the inhibitory activity of \textit{8} against DXS, in the absence of the auxiliary enzyme. To do so, an NMR-based assay in which the DXS-catalyzed conversion of \textsuperscript{13}C-sodium pyruvate to \textsuperscript{13}C-1-deoxy-D-xylulose 5-phosphate is monitored, could be used. Alternatively, a fluorometric end-point assay based on the reaction of 1-deoxy-D-xylulose 5-phosphate with 3, 5-diaminobenzoic acid in an acidic medium has been proposed, based on the formation of a highly fluorescent quinaldine derivative under these conditions.\textsuperscript{32}

Compounds \textit{8–11} were used as references for performing another round of VS. Compound \textit{(E)-12}, relatively close to \textit{(E)-11}, was also used as reference, given its moderate activity against \textit{M. tuberculosis} DXS combined with the fact that it is still a fragment-like molecule, with a \textit{M\textsubscript{w}} of 230 Dalton. Compound \textit{13} was included as a reference molecule because its preliminary biochemical evaluation against \textit{M. tuberculosis} DXS based on the 1 mg batch resulted in a moderate inhibitory activity, which we did not confirm at a later stage.
Chapter 6

6.5 Second round of LBVS

For the second round of LBVS, we used two different similarity-search methods. In addition to the algorithm comparing the 3D shape previously used, the similarity between the reference molecules and the compounds present in the database was measured with the Xfp method, described in Section 6.2, comparing atom pair 2D-fingerprints. Besides the Princeton database, we also explored the libraries of Vitas-M and Specs. We selected 126 compounds by visual inspection of the 1000 top-scoring compounds. We tried to include as many direct analogues as possible, to get a first insight into the SARs of these scaffolds and to focus our efforts, we decided to perform the biochemical assay only against M. tuberculosis DXS, our target enzyme. We found many commercially available derivatives of 8 and 9 and, although not leading to an improvement in the IC₅₀ values with respect to 8 and 9, interesting SARs emerged (Section 6.7), which could further guide semi-rational chemical modification.

We only found a few derivatives of (E)-12, which did not display enhanced inhibitory potency, and no SARs could be obtained.

The best hits found within the second screening, are shown in Table 4, where we report the confirmed IC₅₀ values after reordering them in higher amount. 14 was identified from 12, which we eventually found to be inactive at 120 μM against both orthologues of DXS. Interestingly, 14, the most promising compound identified during this round of VS with an IC₅₀ of 37 ± 5 μM against M. tuberculosis DXS, was retrieved using the Xfp algorithm as an analogue of 12 but did not appear in the 1000 top-scoring molecules when the 3D shape of 12 was compared with the compound library. This fact illustrates the complementarities of different similarity-search methods.

15 and 16 proved to be eight to ten times more active than the corresponding reference molecules (10 and 11, respectively), although being structurally very similar to them. As one can observe from the cLogP values reported in Table 4, hits 14–16 are slightly too hydrophobic considering their Mₘ. This feature renders their in vitro biochemical evaluation challenging and should be optimized on the way to larger and more potent inhibitors, which can effectively proceed through the drug-discovery pipeline.
**Table 4.** Most potent hits from the second round of virtual screening starting with 10, 11 and 13 as reference molecules.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Reference</th>
<th>$IC_{50}$ (μM)[a]</th>
<th>$IC_{50}$ (μM)[a]</th>
<th>$M_w$ (cLogP)[b]</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image13" alt="Molecule 13" /></td>
<td>13</td>
<td>37 ± 5</td>
<td>&gt;40</td>
<td>335.87 (4.82)</td>
</tr>
<tr>
<td><img src="image10" alt="Molecule 10" /></td>
<td>10</td>
<td>56 ± 12</td>
<td>&gt;60</td>
<td>306.19 (4.46)</td>
</tr>
<tr>
<td><img src="image11" alt="Molecule 11" /></td>
<td>11</td>
<td>56 ± 11</td>
<td>&gt;250</td>
<td>277.70 (3.37)</td>
</tr>
</tbody>
</table>

[a] Preliminary results that have to be validated, given that the highest concentrations used are close to the solubility limit.

[b] cLogP values were calculated using ChemAxon.33

6.6 Third and fourth round of LBVS

We used the molecules depicted in Table 4 as references to run a third round of VS, using 3D-shape similarity and substructure similarity. We decided to restrict our selection to direct analogues of the reference molecules to gain additional insight into the SARs and not to purchase any additional new scaffold. It should be noted that the activities of the derivatives tested within this round of VS are based on the preliminary biochemical evaluation of the purchased 1 mg batch and should be interpreted carefully. As a result, any difference in activity less than a factor of two, has not been considered as significant.

We purchased in total 34 compounds. Only four analogues of 15 were available (compounds 17–20, Table 5), thus we purchased them and tested them against *M. tuberculosis* DXS. Unfortunately, no direct analogues of 16 were available neither at Princeton nor at Vitas-M or Specs. Therefore, to get clearer insights into the SARs for 15 and 16 and at the same time enhance their hydrophilicity, derivatives of these scaffolds should be obtained. Few synthetic protocols will be proposed in Section 6.12.1. Numerous derivatives of 14 were commercially available, and we present here a selection, collected in the multiple rounds of screening. These include also derivatives of 14 that we identified...
during a fourth round of LBVS focused specifically on 14, during which we purchased 35 more derivatives.

**Table 5.** Analogues of 15 purchased and tested in the third round of LBVS.

<table>
<thead>
<tr>
<th>IC_{50} (µM)</th>
<th>Mtub DXS</th>
</tr>
</thead>
<tbody>
<tr>
<td>142 ± 16</td>
<td>279 ± 20</td>
</tr>
<tr>
<td>284 ± 16</td>
<td>&gt; 150</td>
</tr>
</tbody>
</table>

Several derivatives include modifications of part A of 14. The presence of a fluorine atom at the *para* position of ring A (20) led to a small loss in inhibitory potency (IC_{50} = 89 ± 14 µM). Unfortunately, a direct comparison of 20 with 21, bearing the fluorine atom at the *ortho* position, is not possible given that 21 resulted in an IC_{50} > 60 µM. The presence of a trifluoromethyl group at the *meta* position (22) led approximately to a three-fold loss in inhibitory potency (IC_{50} = 124 ± 6 µM). When trying to introduce more polar substituents on ring A, such as hydroxyl and carboxyl groups, which also help enhance the hydrophilicity of these scaffolds, the inhibitory potency clearly decreases (23, 24). The presence of a methyl, methoxy or acetyl group at the *para* position of ring A (25, 26 and 27, respectively), also did not lead to any improvement in the inhibitory potency. The fact that 25–27 give essentially the same inhibitory potency (111–149 µM), shows that several types of substituents are largely tolerated at this position. However, no inhibitory potency is observed at 150 µM when a diethylamino group is present (28). Replacing ring A with a *para*-chloro substituted or unsubstituted pyridine rings (29–31) caused at least four-fold loss in inhibitory potency with respect to 14.

An interesting series of compounds resulted from the substitution of the methyl group of ring C with an amino group. 32 is three times more active than its methylated analogue 22. Interestingly, changing the position of the trifluoromethyl group (33, 34), hardly affects the inhibitory potency. When replacing ring A with a pyridinyl ring (35), the activity drops, consistent with our observation for the previously discussed series of compounds. Nevertheless, as shown by the biochemical activity of 36 and 37, the presence of certain substituents on this ring can help to recover the inhibitory potency. On the one hand, 36 resulted in a double-digit micromolar IC_{50} value, while its methylated analogue 30 did not show any inhibition at 250 µM.
As discussed above, 32, bearing the amino group, was also approximately three-fold more active than the corresponding methylated compound 22. One the other hand, 35 and 31, which differ only in the presence of the amino group on ring C, resulted in comparable IC_{50} values. The same holds true for 38, closely related to 23. It is therefore rather difficult to draw any conclusion about how important the amino group on these scaffolds is for the inhibitory potency.

If the amino-substituted thiazolium ring in C is replaced by a six-membered aromatic ring, ideally substituted with polar groups to decrease the lipophilicity (39–42), the inhibitory
activity decreases. It might therefore be better to keep the amino-substituted thiazolium ring in C and exploit different substituents on part A. However, compounds 43 and 44 (Figure 2), might be worthwhile exploring further, given that they display double-digit micromolar IC$_{50}$ values while, being more polar than most of the compounds screened, including their direct analogues (22, 32 (Table 6) and 45 (Table 7), analogues of 43).

Figure 2. Structures of 43 and 44.

An interesting series of compounds resulted from the replacement of the thiazole ring C with a pyridine ring, as shown in Table 7 (45–50). However, the inhibitory potencies are overall worse than those observed for compounds 32–38, except for 49 and 50, which have IC$_{50}$ values of 25 ± 6 μM and 55 ± 5 μM, respectively. The presence of the pyridinyl nitrogen atom seems essential given that 51 did not result in any inhibitory activity at 60 μM.

Table 7. Derivatives of 14 with a pyridinyl ring in part C

<table>
<thead>
<tr>
<th>Comp.</th>
<th>X</th>
<th>R</th>
<th>IC$_{50}$ M. tub DXS (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>46</td>
<td>N</td>
<td>2-CF$_3$</td>
<td>86 ± 5</td>
</tr>
<tr>
<td>45</td>
<td>N</td>
<td>3-CF$_3$</td>
<td>168 ± 18</td>
</tr>
<tr>
<td>47</td>
<td>N</td>
<td>2-F</td>
<td>100 ± 27</td>
</tr>
<tr>
<td>48</td>
<td>N</td>
<td>4-F</td>
<td>101 ± 35</td>
</tr>
<tr>
<td>49</td>
<td>N</td>
<td>2, 5-CH$_3$</td>
<td>25 ± 6</td>
</tr>
<tr>
<td>50[c]</td>
<td>N</td>
<td>3, 4-CH$_3$</td>
<td>55 ± 5[a, b]</td>
</tr>
<tr>
<td>51</td>
<td>C</td>
<td>3, 4-CH$_3$</td>
<td>&gt;60</td>
</tr>
</tbody>
</table>

[a] IC$_{50}$ values confirmed with exact amount of material;  
[b] The value has been extrapolated from an incomplete dataset, where only with one point is below the second plateau.  
[c] This compound was tested also against E. coli IspC, resulting in an IC$_{50}$ value > 60 μM
Another series of compounds we found during the LBVS of 14 has the general scaffold depicted in Figure 3. Although some derivatives gave double-digit micromolar IC\textsubscript{50} values, we believe it is not worthwhile investigating this scaffold further, given that the inhibitory potency is in the same range as the one displayed by much smaller molecules as discussed before.

![Figure 3. General scaffold of the tetracyclic scaffolds coming from the LBVS of 14.](image)

6.7 Discussion of structure-activity relationships (SARs)

In this Section, we will present the most active scaffolds we identified within our LBVS campaign and we will briefly discuss the SARs. Although we will report also the inhibitory potencies against the model enzyme \textit{D. radiodurans} DXS (when available), we will focus on the results we obtained against \textit{M. tuberculosis} DXS, our target enzyme.

8, derived from 5 as reference molecule, is a double-digit micromolar inhibitor of \textit{M. tuberculosis} DXS. As one can observe by comparing the IC\textsubscript{50} value of 8 with those of 52, 53 and 54, the presence of the ethylene hydroxyl group as R\textsuperscript{2} seems fundamental for the inhibitory potency (Table 8). As for R\textsuperscript{1}, a correct interpretation of the SARs is not possible, given that we could only test 55–59 at relatively high concentrations because of solubility issues (Table 8). Methyl substituents on the fused pyridine ring (structures not shown) did not lead to any improvement in the inhibitory potency. 8 turned out to be a potent inhibitor of \textit{E. coli} IspC, with an IC\textsubscript{50} in the single-digit micromolar range. As a result, it could be interesting to test it against pathogenic orthologues of IspC as well as to exploit it as a multi-target inhibitor of both DXS and IspC. In the context of our project, we decided not to follow up on this compound given that, despite displaying a double-digit micromolar IC\textsubscript{50} value against \textit{M. tuberculosis} DXS, it has a higher M\textsubscript{w} compared to 14, 15 and (E)-16. Moreover, as we will discuss in Section 6.10, it did not show any activity in the cell-based assay against multiple strains of \textit{M. tuberculosis}. 
Table 8. SARs of compound 8.

![Chemical Structure](image)

<table>
<thead>
<tr>
<th>Comp.</th>
<th>R¹</th>
<th>R²</th>
<th>R²</th>
<th>D. rad DXS</th>
<th>M. tub DXS</th>
</tr>
</thead>
<tbody>
<tr>
<td>8[b]</td>
<td>cyclohexyl</td>
<td>CH₂CH₂OH</td>
<td>C-</td>
<td>2-CH₃, 3-Cl</td>
<td>17 ± 4[a]</td>
</tr>
<tr>
<td>52</td>
<td>cyclohexyl</td>
<td>CH₃</td>
<td>CH</td>
<td>2-F</td>
<td>12 ± 4</td>
</tr>
<tr>
<td>53</td>
<td>cyclohexyl</td>
<td>CH₂CH₃</td>
<td>CH</td>
<td>4-F</td>
<td>&gt;30</td>
</tr>
<tr>
<td>54</td>
<td>cyclohexyl</td>
<td>CH₂CH₂CH₃</td>
<td>C-</td>
<td>3-CF₃</td>
<td>&gt;30</td>
</tr>
<tr>
<td>55</td>
<td>CH₃</td>
<td>CH₂CH₂OH</td>
<td>C-</td>
<td>3-OH, 4-COOH</td>
<td>nd</td>
</tr>
<tr>
<td>56</td>
<td>CH₂CH₂CH₃</td>
<td>CH₂CH₂OH</td>
<td>C-</td>
<td>3-COOH, 4-OH</td>
<td>nd</td>
</tr>
<tr>
<td>57</td>
<td>CH₂CH₂CH₃</td>
<td>CH₂CH₂OH</td>
<td>CH</td>
<td>4-CH₃</td>
<td>42 ± 10</td>
</tr>
<tr>
<td>58</td>
<td>CH₂CH₂OH</td>
<td>CH</td>
<td>4-OCH₃</td>
<td>nd</td>
<td>136 ± 32</td>
</tr>
<tr>
<td>59</td>
<td>CH₂CH₂OH</td>
<td>CH</td>
<td>4-N(CH₂CH₃)₂</td>
<td>&gt;40</td>
<td>&gt;60</td>
</tr>
</tbody>
</table>

[a] IC₅₀ values confirmed with exact amount of material.
[b] This compound was also tested against E. coli Ispc, resulting in an IC₅₀ = 3 ± 0.4 μM

All the derivatives of compound 9 (obtained from 5 as the reference compound for the LBVS) that we purchased, bearing different substituents on the phenyl ring (Figure 4; e.g., R¹, R², R³ = halogens, hydroxyl groups) resulted in complete loss of the inhibitory potency. One could still explore multiple substitution patterns on the fused aromatic ring. However, given the fact that the scaffold is particularly hydrophobic and planar, we do not believe 9 is the ideal hit to be followed up.

![Chemical Structure](image)

IC₅₀ = 257 ± 72 μM (D. radiodurans DXS)
IC₅₀ = 145 ± 17 μM (M. tuberculosis DXS)

Figure 4. Compound 9, one hit identified during the LBVS campaign.

10 is closely related to 15, discussed in Section 6.6. As one can see from Table 9, substitution on the benzyl ring is essential. In fact, the completely unsubstituted derivative
Ligand-based virtual screening for the discovery of inhibitors of DXS (60), is ten times less active than 10 against M. tuberculosis DXS. The only substitution pattern among the ones shown in Table 9 (compounds 61–66), which retains the inhibitory potency to a certain extent, is the 2,6-dichlorinated compound 65. Although 15 is more promising than 10 in terms of its inhibitory potency ($IC_{50} = 56 \pm 12 \mu M$, M. tuberculosis DXS, Table 4), the data reported in Table 9 might be useful for the optimization of 15.

**Table 9. SARs of compound 10.**

![Structure of compound 10](image)

<table>
<thead>
<tr>
<th>Comp.</th>
<th>R1</th>
<th>IC$_{50}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>D. rad. DXS</td>
</tr>
<tr>
<td>10</td>
<td>4-iPr</td>
<td>298 ± 30[a]</td>
</tr>
<tr>
<td>60</td>
<td>H</td>
<td>nd</td>
</tr>
<tr>
<td>61</td>
<td>3-CH$_3$</td>
<td>&gt;1000</td>
</tr>
<tr>
<td>62</td>
<td>4-tBu</td>
<td>243 ± 89</td>
</tr>
<tr>
<td>63</td>
<td>4-Cl</td>
<td>1308 ± 138</td>
</tr>
<tr>
<td>64</td>
<td>4-F</td>
<td>&gt;1000</td>
</tr>
<tr>
<td>65</td>
<td>2,6-Cl</td>
<td>nd</td>
</tr>
<tr>
<td>66</td>
<td>2-F, 6-Cl</td>
<td>nd</td>
</tr>
</tbody>
</table>

[a] IC$_{50}$ values confirmed with exact amount of material.

(E)-11 is closely related to (E)-16, featuring two phenyl substituents on the oxime carbon rather than one phenyl and one benzyl substituent. As for the derivatives of 10, derivatives of (E)-11 shown in Table 10 might be useful to improve the inhibitory activities of (E)-16, which resulted in a much better inhibitory potency ($IC_{50} = 56 \pm 11$, M. tuberculosis DXS, Table 4) than (E)-11 ($IC_{50} = 528 \pm 43$, M. tuberculosis DXS, Table 10). As expected, the absence of any substituent on both phenyl rings, causes loss of activity (67). Moving the fluorine atom from the para to the ortho position ((E)-68), retains the inhibitory potency, while substituting the para-fluorine atom with a methyl group causes loss of the inhibitory activity (69). The presence of a 2-amino-5-chloro pattern ((E)-70) replacing the dihydroxyl pattern, enhances the inhibitory potency almost two fold with respect to (E)-11, although a direct comparison among the two compounds is not possible, due to the different substitution pattern on R$^2$. However, by comparing the inhibitory potencies of 67 and (E)-70, one could conclude that the 2-amino-5-chloro pattern might be worthwhile exploring more. Interestingly, when replacing the oxime moiety with a hydrazone moiety on (E)-70 (structure not shown), the biochemical activity is retained ($IC_{50} = 287 \pm 140 \mu M$).
Table 10. SARs of compound (E)-11.

<table>
<thead>
<tr>
<th>Comp.</th>
<th>X</th>
<th>R¹</th>
<th>R²</th>
<th>IC₅₀</th>
<th>D. rad. DXS</th>
<th>M. tub. DXS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(E)-11[b]</td>
<td>CH</td>
<td>4-F</td>
<td>2, 4-OH</td>
<td>1385 ± 150[a]</td>
<td>528 ± 43[a]</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>CH</td>
<td>H</td>
<td>H</td>
<td>nd</td>
<td>&gt;600</td>
<td></td>
</tr>
<tr>
<td>(E)-68</td>
<td>CH</td>
<td>2-F</td>
<td>2, 4-OH</td>
<td>914 ± 167</td>
<td>520 ± 45</td>
<td></td>
</tr>
<tr>
<td>69[b]</td>
<td>CH</td>
<td>4-CH₃</td>
<td>2, 4-OH</td>
<td>&gt;1000</td>
<td>&gt;1000</td>
<td></td>
</tr>
<tr>
<td>(E)-70</td>
<td>CH</td>
<td>H</td>
<td>2-NH₂, 5-Cl</td>
<td>nd</td>
<td>282 ± 128</td>
<td></td>
</tr>
</tbody>
</table>

[a] IC₅₀ values confirmed with exact amount of material. [b] The geometry of the double bond is unknown.

Other scaffolds resulting in interesting IC₅₀ values against *M. tuberculosis* DXS are reported in Table 11, which one could also explore as hits.

Table 11. Two unrelated scaffolds identified during the multiple cycles of our LBVS campaign.

<table>
<thead>
<tr>
<th>IC₅₀ (μM)</th>
<th>M. tub. DXS</th>
</tr>
</thead>
<tbody>
<tr>
<td>165 ± 21</td>
<td>473 ± 188</td>
</tr>
</tbody>
</table>

6.8 Distribution of molecular weight and cLogP among all hits of our LBVS project

According to Lipinski’s rule of five, drug-like molecules with the potentiality to have good absorption and permeation properties, therefore being likely orally active drugs, should contain no more than five H-bond donors and ten H-bond acceptors, should have a Mᵦ below 500 Dalton and a cLogP value (where c stands for calculated and P is the partition coefficient between octanol and water) not greater than five.³⁴ In analogy to the rule of five, a rule of three has been introduced for fragment-based projects. Typically, fragments should have Mw < 300 Dalton, cLogP < 3, and fewer than three H-bond acceptors and donors.³⁵
We performed a detailed analysis of the whole series of compounds screened in our LBVS in terms of their Mw and cLogP (calculated with Chemaxon). As shown in Figure 5, all compounds tested have Mw < 500, with a mean value of 307 Dalton. In fact, a considerable number of them still belongs to the fragment world (Mw < 300), which is not surprising considering that the initial reference molecules were also fragments. Our most promising hits mainly range from Mw of 290 to 380 Dalton, showing that there is still room for improvement by fragment growing.

![Figure 5](image)

**Figure 5.** Distribution of molecular weights (Mw) among the compounds tested during the three rounds of LBVS.

As shown in Figure 6a, compounds with Mw < 300 have a mean cLogP value of 2.57, but most of the hits tend to be too hydrophobic according to the rule of three. Figure 6b represents the same type of distribution for compounds with Mw > 300. We can observe that most of the compounds have cLogP < 5 and, particularly, our hits all have cLogP < 5.
Figure 6. cLogP distribution of the compounds tested during the three rounds of VS. (a) Compounds with $M_w < 300$; (b) Compounds with $M_w > 300$. The cLogP values were calculated with Chemaxon.

6.9 Biochemical evaluation of the most potent hits against a mammalian thiamine diphosphate-dependent enzyme: pyruvate dehydrogenase

As we discussed in Chapter 1 of this thesis, DXS has weak homology with other TDP-dependent enzymes. The homology (identity) is particularly high within the TDP-binding pocket. Even though the binding mode of our most potent hits has not yet been established, we wanted to evaluate their activity against a representative mammalian TDP-dependent enzyme. Given its commercial availability, we chose porcine pyruvate dehydrogenase (PDH). The PDH complex plays a vital role in cellular metabolism, catalyzing the oxidative decarboxylation of pyruvate and the subsequent acetylation of coenzyme A to acetyl-coenzyme A, which can enter the citric acid (Krebs) cycle. The E1 subunit of the PDH complex is a TDP-dependent enzyme.

We ran the PDH assay in the optimal conditions for this enzyme, which differs from those adopted for the DXS assay, as described in detail in the experimental part. Nevertheless, the relative concentrations of pyruvate and TDP in both assays, with respect to their corresponding $K_m$ values, was kept the same (6 $x$ $K_m$ and 6.5 $x$ $K_m$ for pyruvate and TDP, respectively) for better comparison of the activity of our inhibitors against both enzymes. By dissolving each compound in 100% DMSO and slowly adding water to each solution, we could reach higher concentrations of the inhibitors in the assay media than the highest ones achieved in the DXS assay. We could in fact test most of the compounds at 100 or 200 μM. (E)-16 did not result in any inhibitory potency of porcine PDH at 200 μM, and 15,
36, 32 and 50 did not inhibit porcine PDH at 100 μM. Although every measurement was previously blanked by measuring the absorbance of each solution of inhibitor in DMSO at every concentration, we did observe slightly different initial absorbances for different batches of the same inhibitor at the zero point of the measurement for 32, 36 and 50. This fact could arise from the inhibitors reacting with one or more components of the PDH storage buffer and does not affect our qualitative evaluation of the data, although the actual concentration of every inhibitor in solution might be lower than the one we report.

These results are very promising and suggest that selective inhibition of the pathogenic DXS over mammalian TDP-dependent enzymes might be possible.

6.10 Cell-based assays against multiple strains of *M. tuberculosis*

Thanks to a collaboration with the Rijksinstituut voor Volksgezondheid en Milieu, in Bilthoven (The Netherlands), we could obtain information about the potency of our best hits – based on the results of the in vitro biochemical assay – also in cell-based assays against multiple strains of tuberculosis (TB), including drug-resistant (DR), multi-drug resistant (MDR) and extensively-drug resistant (XDR) strains, as defined in Table 12.

**Table 12.** List of all the strains of *Mycobacterium tuberculosis* used in our cell-based assays with the corresponding commercially available drugs to which they are resistant. DR-, MDR- and XDR-TB indicate drug-resistant, multidrug-resistant and extensively-drug resistant strains of tuberculosis.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>H37Rv</td>
<td>All-susceptible (control strain)</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>1010900234[a]</td>
<td>All-susceptible</td>
<td>-</td>
</tr>
<tr>
<td>C</td>
<td>1011000848[a]</td>
<td>DR-TB</td>
<td>Isoniazide</td>
</tr>
<tr>
<td>D</td>
<td>1011200081[a]</td>
<td>DR-TB</td>
<td>Isoniazide</td>
</tr>
<tr>
<td>E</td>
<td>1011200189[a]</td>
<td>MDR-TB</td>
<td>Isoniazide, rifampicine, ethambutol, streptomycin, pyrazinamide</td>
</tr>
<tr>
<td>F</td>
<td>1011200345[a]</td>
<td>MDR-TB</td>
<td>Isoniazide, rifampicine, ethambutol, streptomycin, pyrazinamide, ciprofloxacin, clarithromycin</td>
</tr>
<tr>
<td>G</td>
<td>1010900268[a]</td>
<td>XDR-TB</td>
<td>Isoniazide, rifampicine, ethambutol, streptomycin, pyrazinamide, kanamycin, capreomycin</td>
</tr>
<tr>
<td>H</td>
<td>1010901542[a]</td>
<td>XDR-TB</td>
<td>Isoniazide, rifampicine, ethambutol, streptomycin, kanamycin.</td>
</tr>
</tbody>
</table>

[a] The numbers correspond to patient isolate strains of *M. tuberculosis*, as identified with the reverse line blot from the company Hain Lifescience.
As mentioned before in the contest of this thesis, the discrepancy between in vitro inhibitory potency of a potential antituberculotic drug and its cell-based activity can be striking due to the remarkably thick and hydrophobic cell wall of \textit{M. tuberculosis}. Therefore, it is particularly important to obtain and compare in vitro and cell-based activities when it comes to the early stages of development of an antituberculotic drug. We first tested the best hits obtained from the first round of virtual screening, namely 8–(E)-11 (Table 13). All compounds, except for 8, showed moderate inhibition of bacterial growth of four different strains of TB, having MIC values between 10 and 100 $\mu$g/mL. When comparing the cLogP values of 8–(E)-11, we realized that 8 has the lowest cLogP value of this series (2.37); the particularly polar character of 8 might explain why it is the only compound not displaying cell-based activity.

\begin{table}[h]
\centering
\caption{Minimum inhibitory concentration (MIC) values of 8–(E)-11, 5 and 71 against multiple strains of \textit{M. tuberculosis}}
\label{tab:table13}
\begin{tabular}{c|c|c|c|c}
\hline
 & Strain A & Strain B & Strain E & Strain F \\
\hline
8 & >100 & >100 & >100 & >100 \\
9 & 10–100 & 10–100 & 10–100 & 10–100 \\
10 & 10–100 & 10–100 & 10–100 & 10–100 \\
(E)-11 & 10–100 & 10–100 & 10–100 & 10–100 \\
5 & >100 & >100 & >100 & >100 \\
71 & >100 & >100 & >100 & >100 \\
\hline
\end{tabular}
\end{table}

Despite being only moderate inhibitors of \textit{D. radiodurans} \textit{DXS} in vitro (5: IC$_{50}$ = 430 ± 68 $\mu$M; 71: IC$_{50}$ = 109 ± 16 $\mu$M) and not showing any activity against \textit{M. tuberculosis} \textit{DXS}, thiamine derivatives 5 and 71 were also tested to investigate their potential as pro-drugs. In fact, whereas polar, diphosphorylated compounds cannot permeate the bacterial cell wall – even in humans, external forms of thiamin such as thiamin phosphates are hydrolyzed to thiamin prior to uptake –, their dephosphorylated analogues could enter the cells, where they would be “activated” by being (di)phosphorylated.

To better understand the role of thiamine analogues as potential pro-drugs, we studied in detail how bacteria synthesize thiamine and thiamine diphosphate, the essential cofactor for many metabolic reactions. In fact, while many eukaryotes have to take up thiamine (vitamin B1) from nutrition, most bacteria can produce thiamine and its diphosphorylated form \textit{de novo}. Moreover, salvage and transport pathways, complement the necessity of
thiamine for most bacteria (Scheme 1).\textsuperscript{38} On the one hand, the synthesis of TDP in bacteria, goes through the corresponding thiamine monophosphate, which is phosphorylated by thiamine phosphate kinase (ThiL). On the other hand, many enzymes involved in the bacterial salvage pathways are responsible for the mono- or diphosphorylation of thiamine or its intermediates, such as ThiD (pyrimidine salvage, not shown), ThiK and ThiN (thiamine salvage, Scheme 1). \textit{M. tuberculosis} lacks these salvage pathways and only relies on the \textit{de novo} thiamine synthesis.\textsuperscript{39} This could be the reason why no cellular activity was observed for 5 and 71, which cannot be mono- or diphosphorylated within mycobacterial cells and are too weak to inhibit mycobacterial growth. It must be underlined that, in general, the strategy of taking advantage of kinases within the cells to activate pro-drugs, might be quite challenging, mainly due to the critical step of the substrate-recognition process.\textsuperscript{40, 41}

\textbf{Scheme 1.} Last three steps of the thiamine diphosphate biosynthesis in bacteria and two steps of thiamine salvage pathway. Enzymes involved in thiamine diphosphate biosynthesis are shown in bold and underlined. Enzymes involved in the thiamine salvage pathway are shown in italics.

We obtained the results of the cell-based assay described in Table 13 at the same time as the \textit{in vitro} inhibitory potencies of compounds 14, 15, 32, (\textit{E})-16 and 36, which turned out to be much more active. Therefore, we decided to focus on the last series, and accurately establish their cell-based activities.

Preliminary tests to check the cell-based activity of 14, 15, 32, 36, (\textit{E})-16 and BAT\textunderscore 225 were performed, where all compounds were tested for their ability to inhibit the growth of multiple strains of \textit{M. tuberculosis} at a concentration of 20 \textmu M. While for 14, 15, 32, 36 and (\textit{E})-16 a MIC value < 20 \textmu M was measured, 50 did not inhibit the growth of most of the strains. The compounds resulting in no growth of \textit{M. tuberculosis} at 20 \textmu M (14, 15, 32, 36 and (\textit{E})-16) were subsequently tested against multiple strains at different concentrations, namely 10, 5, 2.5 and 1.25 \textmu M in a Middlebrook 7H10 agar medium. The results are summarized in Table 14. Although the results for 14 have to be finalized, one can conclude
that 14 has the most potent cellular activity of this series. Its close analogue 32 also displays good cellular activity, while 36 is slightly less potent. The other two scaffolds, 15 and (E)-16 are overall less potent.

Table 14. Minimum inhibitory concentration (MIC) values for compounds 14, 15, 32, (E)-16, 36 against multiple strains of M. tuberculosis, including multi- and extensively-drug resistant strains.

<table>
<thead>
<tr>
<th>Strain</th>
<th>B</th>
<th>A</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>2.5–5[a]</td>
<td>1.25–2.5[a]</td>
<td>nd</td>
<td>1.25–2.5[a]</td>
<td>nd</td>
<td>&lt;1.25</td>
<td>1.25–2.5[a]</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>5</td>
<td>5–10[a]</td>
<td>5–10[a]</td>
<td>5–10[a]</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>5</td>
<td>2.5–5[a]</td>
<td>5</td>
<td>2.5–5[a]</td>
<td>2.5–5[a]</td>
</tr>
<tr>
<td>(E)-16</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>5–10[a]</td>
<td>5–10[a]</td>
</tr>
<tr>
<td>36</td>
<td>5–10[a]</td>
<td>&gt;10</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5–10[a]</td>
<td>5–10[a]</td>
<td></td>
</tr>
</tbody>
</table>

[a] The lower value of the range corresponds to the concentration (μM) at which the inhibitor causes less growth than 1% of the control. The higher value of the range corresponds to the first well of the series, where no growth is observed.

6.11 Cytotoxicity assays and preliminary cell-based assays against Plasmodium falciparum 3D7

A selection of compounds from our LBVS campaign, namely 5, 68, 78, 88 and 14, as DMSO solutions, has also been tested for their potential antimalarial activities. 5, 10 and (E)-11 were incubated in a P. falciparum 3D7 culture at 1 mg/mL for 48 hours, with an initial parasitemia of 2%. 9 was incubated at 8.6 mg/mL. Additionally, a P. falciparum culture control with 1% v/v DMSO was analyzed. The parasitaemia was measured by flow cytometry and standardized to the control. As shown in Figure 7, 5 has almost no antimalarial effect compared to the control, while 14 reduces the parasitemia of about 20%.

![Figure 7](image_url)

**Figure 7.** Evaluation of the antimalarial effect in P. falciparum 3D7 of 5, 9, 10, (E)-11 and 14. As control, P. falciparum 3D7 with 1% v/v DMSO was used.
10 and (E)-11 caused complete lysis of erythrocytes, which is why no data could be collected. A similar result was observed for 9, which caused partial lysis of the red blood cells.

6.12 Conclusions and follow up

LBVS proved to be a successful strategy for the identification of hits that are able to inhibit M. tuberculosis DXS in the low micromolar range. LBVS is a particularly useful hit-identification strategy when no or little structural information is available, which is the case for DXS. LBVS allowed us to explore a wide spectrum of chemical diversity, starting from the structures of three DXS inhibitors that we identified as described in Chapter 2. By purchasing closely related derivatives of the hit or lead compounds, we obtained a first series of SARs; a systematic evaluation of SARs will, however, require the chemical synthesis of more derivatives. A few, straightforward chemical routes for the synthesis of derivatives of the most interesting hits are proposed in Section 6.12.1. Many of the compounds we tested tend to be too hydrophobic (Figure 6), rendering their biochemical evaluation challenging, leading to often incomplete datasets. Moreover, having too hydrophobic scaffolds is not ideal for a potential drug. Therefore, especially for fragment-like hits, the optimization stage should be aimed at enhancing both potency and polarity.

We confirmed that selective inhibition of DXS over mammalian TDP-dependent enzymes is possible, by testing a few compounds against porcine PDH, a representative mammalian TDP-dependent enzyme. Our hits are active against multiple strains of M. tuberculosis, including MDR- and XDR-strains. In particular, 14 was identified as the most potent compound in the cell-based assays. Interestingly, 14 was also found to be the only compound among the LBVS hits we tested, having a weak antimalarial effect, reducing parasitemia by about 20% in a cell-based assay with P. falciparum 3D7. Building on this promising initial result, 15, (E)-16, 32 and 36 are going to be subjected to the same assay.

The fact that the mycobacterial activity of 14 and the other compounds is due to inhibition of DXS should be also confirmed, by performing experiments to rescue the bacterial growth, for instance by supplementing the growth medium with 1-deoxy-D-xylulose (DX) – which is capable of penetrating the cell and undergoing intracellular phosphorylation to yield the product of the DXS-catalyzed reaction – or by overexpressing DXS. The fact that DXS is a branch point in bacterial metabolism, further complicates the validation of the intracellular target. Additional biosynthetic pathways including TDP and pyridoxal
phosphate biosynthesis should be affected by DXS inhibition, therefore rescue studies should also take these pathways into account (e.g., external supplementation of thiamine).

Understanding the binding mode of the most potent representative scaffolds 14, 15 and (E)-16 by establishing their mode of inhibition or by using the STI-NMR methodology described in Chapter 2, is, without any doubt, the next step to rationally improve the in vitro inhibitory potencies of these scaffolds. Once the mode of inhibition has been established, one could perform docking studies focused on the specific subpocket where the molecule binds (substrate- or cofactor-binding pocket). These studies, together with a more complete set of SARs, would certainly guide a rational optimization of the best inhibitors. However, we would like to underline once more that the discrepancy between in vitro and cell-based data, especially for the development of antituberculotic drugs, can be dramatic. Therefore, it might be a good idea to perform the two assays in parallel, as much as is practically possible.

**6.12.1 General approaches for the synthesis of derivatives of the best hits**

As reported in the literature, 15 and many derivatives of this scaffold can be conveniently synthesized according to the retrosynthetic scheme shown in Scheme 2a. This route allows for exploring various substituents on the indole and phenyl rings, according to the commercial availability of the corresponding starting materials (or the possibility to conveniently synthesize them). The importance of the hydroxymethylene moiety on 15 could also be explored by testing all corresponding aldehyde derivatives as well as by replacing it with other functional groups (e.g., -CH₂CH₂OH, CH₂CH₂CH₂OH, H, CN, COCH₃ etc.).

Oxime (E)-16 can be synthesized from the corresponding ketone, which can be obtained by a Friedel-Crafts reaction of resorcinol and para-chloro phenylacetic acid. Many derivatives of (E)-16 have been reported in the literature, but they all contain one phenyl ring (ring A) with two or three hydroxyl or methoxy groups. By using this methodology, one can exploit multiple substitution patterns on ring B (halogens, -OCH₃, -NO₂, Scheme 1b), while the groups on ring A are limited to electron-donating groups. If the ring is sufficiently activated, some alkyl substituents might also be introduced. Derivatives with electron-withdrawing groups on A ring could be accessed by reacting a cyanide with a Grignard reagent (Scheme 2b).

We tested many derivatives of 14, which are all commercially available. However, to explore even more chemical diversity, one could follow the synthetic approach as shown...
in Scheme 2c, where the second thiazolyl ring is formed in a ring-closing reaction in the last step. This synthetic approach allows access to a great variety of scaffolds, as already reported in the literature.47

Scheme 2. Proposed retrosynthetic approaches for synthesizing more derivatives of 15, (E)-16 and 14.

6.13 Experimental

Characterization of compounds purchased in higher amount

The compounds were purchased in 1 mg batches and directly dissolved in DMSO without further purification, so as to obtain 80 mM solutions, when possible. Their solubility in the assay buffer was tested prior to the inhibitory assay so as to reach 1000 μM as the highest concentration tested when possible. The most active compounds were re-ordered in higher amount (typically 25 mg) from Princeton Biomolecular Research, except otherwise mentioned, and analyzed by NMR and MS prior to IC50 determination. If the purity of the
sample as received from the supplier was below 95% (determined by Ultra-High Performance Liquid Chromatography (UHPLC) at 214 nm, they were further purified by reversed phase (RP)-HPLC, as described below. Low-resolution mass spectra were obtained by electron-spray ionization (ESI) in the positive mode on a Thermo Scientific LCQ Fleet. \(^1\)H-NMR spectra were obtained at 300 MHz with a Bruker AV300 spectrometer. Compound purity was determined by analytical RP-UHPLC. Analytical RP-UPLC was performed on a Dionex Ultimate 3000 RSLC System (DAD-3000 RS Photodiode Array Detector) and a Dionex Acclaim RSLC 120 column (C18, 3.0 x 50 mm, particle size 2.2 μm, 120 Å pore size) at a flow rate of 1.2 mL min\(^{-1}\). Compounds were detected by UV-absorption at 214 nm. Data recording and processing was done with Dionex Chromeleon Management System Version 6.8. Preparative RP-HPLC was performed with Waters Prep LC4000 Chromatography System using a Reprospher 100 (C18-DE, 100 mm x 30 mm, particle size 5 μm, 100 Å pore size) column from Dr. Maisch GmbH and a Waters 489 Tunable Absorbance Detector operating at 214 nm.

Eluents for analytical RP-UHPLC were as follows: A: miliQ-deionized water with 0.05% TFA and D: HPLC-grade acetonitrile/miliQ-deionized water (9:1) with 0.05% TFA. The gradients for the analytical RP-UHPLC were: Method A: 100% A to 100% D in 2.2 minutes, Method D: 90%A/10%D to 100%D in 2.2 minutes and Method E: 80% A to 20% D to 100%D in 2.2 minutes.

Eluents for preparative RP-HPLC were as follows: A: miliQ-deionized water with 0.05% TFA, C: HPLC-grade acetonitrile/miliQ-deionized water (6:4) with 0.1% TFA and D: HPLC-grade acetonitrile/miliQ-deionized water (9:1) with 0.1% TFA.

2-(2-Hydroxyethylamino)-3-((Z)-(3-cyclohexyl-4-oxo-2-thioxothiazolidin-5-ylidene)methyl)-4H-pyrido[1,2-a]pyrimidin-4-one (8, supplier ID: OSSK_026083)

\[
\begin{align*}
\text{H-NMR (300 MHz, DMSO-}d_6\text{): } & \delta=8.77 \text{ (dd, } J = 1.6, 6.7, 1H), 8.07 \text{ (br s, } 1H), 7.89 \text{ (ddd, } J = 1.6, 6.7, 8.7, 1H), 7.64 \text{ (s, } 1H), 7.34 \text{ (d, } J = 8.7, 1H), 7.13 \text{ (td, } J = 1.3, 6.7, 1H), 4.93 \text{ (br t, } J = 12.1, 1H), 4.77 \text{ (t, } J = 5.5, 1H), 3.68-3.50 \text{ (m, } 4H), 2.45-2.22 \text{ (m, } 2H), 1.93-1.75 \text{ (m, } 2H), 1.75-1.54 \text{ (m, } 3H), 1.44-1.06 \text{ (m, } 3H) \text{. LC-MS (ESI\textsuperscript{+}): } m/z 431.12 \text{ ([M+H\textsuperscript{+}])}. \text{ Purity of the sample by RP-UHPLC could not be measured because it was insoluble in the eluents required for the measurement.}
\end{align*}
\]

4-(2,4-Dichlorophenyl)-4,10-dihydro-1H-[1,3,5]triazino[1,2-a]benzimidazole-2-amine (9, supplier ID: OSSK_676817)

\[
\begin{align*}
\text{H-NMR (300 MHz, DMSO-}d_6\text{): } & \delta=8.04 \text{ (br s, } 1H), 7.73 \text{ (d, } J = 2.0, 1H), 7.45 \text{ (dd, } J = 2.0, 7.8, 1H), 7.26 \text{ (d, } J = 7.3, 1H), 7.08 \text{ (m, } 2H), 6.96 \text{ (t, } J = 7.3, 1H), 6.80 \text{ (t, } J = 7.3, 1H), 6.62 \text{ (d, } J = 7.8, 1H), 6.40 \text{ (br s, } 2H) \text{. LC-MS (ESI\textsuperscript{+}): } m/z 332.14 \text{ ([M+H, Cl\textsuperscript{35}])}, 334.07 \text{ ([M+H, Cl\textsuperscript{37}])}. \text{ Purity of the purchased sample by RP-UHPLC (method A, } t_{Rx} = 1.8 \text{ min) = 99%.}
\end{align*}
\]

2-(1-(4-Isopropylbenzyl)-1H-benzo[d]imidazol-2-yl)ethanol
Ligand-based virtual screening for the discovery of inhibitors of DXS

(10, supplier ID: OSSK_927261)

\[
\text{\textsuperscript{1}H-NMR (300 MHz, DMSO-}d_6\text{): } \delta = 7.61 - 7.53 (m, 1H), 7.49 - 7.41 (m, 1H), 7.22 - 7.11 (m, 4H), 7.03 (d, J = 8.2, 2H), 5.46 (s, 2H), 4.85 (t, J = 5.5, 1H), 3.85 (q, J = 6.8, 2H), 3.00 (t, J = 6.8, 2H), 2.82 (d, J = 6.9, 1H), 1.15 (d, J = 6.9, 6H). LC-MS (ESI\textsuperscript{+}): m/z 295.12 ([M+H\textsuperscript{+}]). Purity of the purchased sample by RP-UHPLC (method A, t\textsubscript{R} = 1.8 min) = 99%.
\]

4-[(4-Fluorophenyl)(hydroxyimino)methyl]benzene-1,3-diol (11, supplier ID: OSSL_151911)

This compound was purified by RP-HPLC (Gradient 85% A 15% D to 100% D in 40 minutes) prior to testing it. \textsuperscript{1}H-NMR (300 MHz, DMSO-\textit{d}\textsubscript{6}): \delta = 11.37 (s, 1H), 11.26 (s, 1H), 9.77 (s, 1H), 7.39 - 7.25 (m, 4H), 6.53 (d, J = 8.6, 1H), 6.31 (d, J = 2.4, 1H), 6.21 (dd, J = 2.4, 8.6, 1H). LC-MS (ESI\textsuperscript{+}): m/z 248.02 ([M+H\textsuperscript{+}]). Purity of the sample after RP-HPLC purification, assessed by RP-UHPLC (method A, t\textsubscript{R} = 2.0 min) = 98%.

2-Amino-6,7-dihydro-6-(1-(4-isopropylphenyl)ethyl)pyrrolo[3,4-d]pyrimidin-5-one (13, supplier ID: OSSL_332078)

\textsuperscript{1}H-NMR (300 MHz, DMSO-\textit{d}\textsubscript{6}): \delta = 8.51 (s, 1H), 7.38 (br s, 2H), 7.29 - 7.16 (m, 4H), 5.41 (q, J = 7.2, 1H), 4.37 (d, J = 18.4, 1H), 3.91 (d, J = 18.4, 1H), 2.85 (hept, J = 6.9, 1H), 1.57 (d, J = 7.2, 3H), 1.17 (d, J = 6.9, 6H). LC-MS (ESI\textsuperscript{+}): m/z 297.11 ([M+H\textsuperscript{+}]). Purity of the purchased sample by RP-UHPLC (method D, t\textsubscript{R} = 1.9 min) = 98%.

N-(3-Chloro-2-methylphenyl)-4-(2,4-dimethylthiazol-5-yl)thiazol-2-amine (14, supplier ID: OSSK_523229)

\textsuperscript{1}H-NMR (300 MHz, DMSO-\textit{d}\textsubscript{6}): \delta = 9.63 (s, 1H), 7.92 - 7.80 (m, 1H), 7.28 - 7.16 (m, 2H), 6.93 (s, 1H), 2.57 (s, 3H), 2.48 (s, 3H), 2.32 (s, 3H). LC-MS (ESI\textsuperscript{+}): m/z 336.12 ([M+H, Cl\textsubscript{35}]\textsuperscript{+}), 338.01 ([M+H, Cl\textsubscript{37}]\textsuperscript{+}). Purity of the purchased sample by RP-UHPLC (method A, t\textsubscript{R} = 2.2 min) = 99%.

(1-(3,4-Dichlorobenzyl)-1H-indol-3-yl)methanol (15, supplier ID: AH-487/41951505)

\textsuperscript{1}H-NMR (300 MHz, DMSO-\textit{d}\textsubscript{6}): \delta = 7.61 (d, J = 7.8, 1H), 7.57 (d, J = 8.3, 1H), 7.50 (d, J = 2.0, 1H), 7.46 - 7.39 (m, 2H), 7.15 (dd, J = 2.0, 8.3, 1H), 7.12 - 7.07 (m, 1H), 7.06 - 6.99 (m, 1H), 5.39 (s, 2H), 4.82 (t, J = 5.3, 1H), 4.64 (d, J = 5.3, 2H). MS (ESI\textsuperscript{+}): m/z 288.03 ([M-OH, Cl\textsubscript{35}]\textsuperscript{+}), 290.03 ([M-OH, Cl\textsubscript{37}]\textsuperscript{+}). Purity of the sample by RP-UHPLC could not be measured because it was insoluble in the eluents required for the measurement.
Chapter 6

4-[(4-Chlorobenzyl)(hydroxyimino)methyl]benzene-1,3-diol (16, supplier ID: ID STK858385)

(E)-16 was purchased from Vitas-M. $^1$H-NMR (300 MHz, DMSO-$d_6$): $\delta$=11.65 (s, 1H), 11.50 (s, 1H), 9.72 (s, 1H), 7.34 (d, $J$ = 8.7, 2H), 7.30–7.22 (m, 3H), 6.28–6.22 (m, 2H), 4.16 (s, 2H). LC-MS (ESI$^+$): m/z 278.12 ([M+H, Cl$^{35}$]+), 280.11 ([M+H, Cl$^{37}$]+). Purity of the purchased sample by RP-UHPLC (method E, $t_R$ = 1.9 min) > 99%.

4-(2-Amino-4-methylthiazol-5-yl)-N-(4-(trifluoromethyl)phenyl)thiazol-2-amine (27, supplier ID: OSSK_523240)

$^1$H-NMR (300 MHz, DMSO-$d_6$): $\delta$=10.65 (s, 1H), 7.82 (d, $J$ = 8.46, 2H), 7.65 (d, $J$ = 8.46, 2H), 7.01 (s, 2H), 6.72 (s, 1H), 2.33 (s, 3H). LC-MS (ESI$^+$): m/z 357.14 ([M+H]$^+$). Purity of the purchased sample by RP-UHPLC (method E, $t_R$ = 1.6 min) = 97%.

4-(2-Amino-4-methylthiazol-5-yl)-N-(3-(trifluoromethyl)phenyl)thiazol-2-amine (32, supplier ID: STK729380)

32 was purchased from Vitas-M. $^1$H-NMR (300 MHz, DMSO-$d_6$): $\delta$=10.61 (s, 1H), 8.41 (s, 1H), 7.64 (d, $J$ = 7.9, 1H), 7.53 (t, $J$ = 7.9, 1H), 7.28 (d, $J$ = 7.9, 1H), 7.12 (br s, 2H), 6.69 (s, 1H), 2.35 (s, 3H). LC-MS (ESI$^+$): m/z 357.14 ([M+H]$^+$). Purity of the purchased sample by RP-UHPLC (method E, $t_R$ = 1.6 min) = 95%.

N-(4-(2-Amino-4-methylthiazol-5-yl)thiazol-2-yl)-5-chloropyridin-2-amine (36, supplier ID: STK858385)

36 was purchased from Vitas-M. $^1$H-NMR (300 MHz, DMSO-$d_6$): $\delta$=11.50 (s, 1H), 8.34 (d, $J$ = 2.29, 1H), 7.81 (dd, $J$ = 2.62, 8.89, 1H), 7.11 (d, $J$ = 8.48, 1H), 7.01 (br s, 2H), 6.80 (s, 1H), 2.31 (s, 3H). LC-MS (ESI$^+$): m/z 324.14 ([M+H]$^+$). Purity of the purchased sample by RP-UHPLC (method A, $t_R$ = 1.9 min) = 97%.

Cell-based assays with *Mycobacterium tuberculosis* strains. A series of drug-susceptible, DR, MDR and XDR *M. tuberculosis* isolates, and one control strain (H37Rv) were selected and subcultured on a Middlebrook 7H10 agar medium until use. Susceptibility testing using the absolute concentration method was carried out by preparing 25-well plates with solid 7H10 medium containing different concentrations (20, 10, 5, 2.5, 1.25 $\mu$M) of each inhibitor. The plates were subsequently inoculated by adding 10 $\mu$L *Mycobacterium* suspension to each well. Then the plates were incubated at 35.5 °C in a CO$_2$ incubator. After appropriate incubation, the MICs were assessed. The reading of the plates was carried out when the bacterial growth on the two control wells without inhibitor was sufficient, *i.e.*, when colonies were clearly visible.

Preliminary cell-based assays against *Plasmodium falciparum* 3D7. These assays were conducted according to literature procedures.

Spectrophotometric assay for determining the biochemical activity against pyruvate dehydrogenase complex

Pyruvate dehydrogenase from porcine heart was purchased from Sigma Aldrich as a buffered aqueous 50% glycerol solution containing approximately 9 mg/mL bovine serum albumin, 30% sucrose, 1.5 mm EDTA, 1.5 mm ethylene glycol tetraacetic acid (EGTA),
1.5 mM 2-mercaptoethanol, 0.3 TRITON® X-100, 0.003% sodium azide, and 15 mM potassium phosphate, pH 6.8. Each inhibitor was diluted in 100% DMSO, and bidistilled water was slowly added. The solutions were vortexed, and then the rest of the assay components was added in the following order: (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH = 8 (100 mM), bovine serum albumin (1 mg/mL), cysteine (2 mM), TDP (0.325 μM), coenzyme-A (150 μM), (tris(2-carboxyethyl)phosphine (TCEP, 300 μM), pyruvate (150 μM) and MgCl₂ (1 mM). The final % of DMSO in the assay media was 10%. The assay was carried out as previously described.⁴⁸ The UV/Vis spectrophotometer (Beckman Coulter DU800) was first blanked with each solution, and then NAD⁺ was added to a final concentration of 500 μM. The solutions were then preincubated for 5 minutes at 30 °C. Porcine PDH (final concentration of 0.01 U/mL) was then added to initiate the reactions which were monitored spectrophotometrically by measuring the appearance of NADH at 340 nm over time at 30 °C. Initial rates of product formation (graphed as [NADH] versus time) were determined using Microsoft Excel.

6.14 References

Ligand-based virtual screening for the discovery of inhibitors of DXS


