Structure-Based Design on the Way to New Anti-Infectives
Hirsch, Anna K.H.

Published in:
Ideas in Chemistry and Molecular Sciences: Where Chemistry Meets Life

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:
2010

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.

Download date: 31-12-2019
Structure-Based Design on the Way to New Anti-Infectives

Anna Katharina Herta Hirsch

7.1 Introduction

With an estimated 300–500 million new infections and three million deaths annually, malaria and tuberculosis undoubtedly still pose a major health concern [1]. The need for the development of novel therapeutic approaches is ever-growing in light of the emergence of multi-drug-resistant parasites [2]. How can a new drug be identified?

First, a lead compound, that is, a molecule with a promising biological activity that does not yet fulfill all the requirements but represents the starting point on the way to a new drug, has to be identified. A number of different strategies exist to achieve this goal:

• High-throughput screening (HTS) of a large library of small molecules is of particular interest in cases in which no structural information or characterization of the biological target is available. The majority of lead compounds still comes from hits identified by HTS [3].
• Virtual screening has established itself as an alternative or complementary approach to classical HTS. Potentially active and/or drug-like compounds are selected from a library of compounds, using elaborate docking and scoring functions [4].
• Combinatorial chemistry is useful for the formation of large, small-molecule libraries. However, this approach is less effective for generating a great deal of structural diversity.
• Nature can be of help by providing a rich and diverse source of structural inspiration [5]. The scaffold of a natural product, displaying interesting biological properties, could be developed into a new drug.
• Structure-based design, a comparatively new field, has established itself in pharmaceutical research as a valuable alternative to traditional screening; the X-ray crystal structure of a target enzyme is used as a basis for lead compound identification and optimization. The increasing number of leads, identified and/or optimized using this rational approach, used for the development of new drugs illustrates
this fact [6]. Until now, this strategy has been mainly used in the later stages of lead optimization.

The last mentioned strategy will be employed for the purposes of this project. The aim of this approach is to identify synthetically accessible target molecules, with optimal stereoelectronic properties that are complementary to the binding site of the target enzyme and show minimal or no repulsive interactions when complexed to the enzyme. The identification of promising leads is aided by the medicinal chemist’s understanding of molecular recognition. While hydrophobic interactions between a lead and an enzyme are the main driving force for complexation, H-bonding interactions account for selectivity (Figure 7.1). Later stages of drug development are not the subject of this chapter.

To determine the biological activity of a potential lead compound, a new biological target [7], that is, an enzyme or receptor that upon interference by the ligand/drug has an impact on the disease causing pathogen in the desired way, has to be identified. In an ideal case, the target is essential for the pathogen and not present in humans, thereby precluding any selectivity issues. Isoprenoids are an essential class of natural products, requiring the essential precursors isopentenyl diphosphate (IPP, 1) and dimethylallyl diphosphate (DMAPP, 2) for their biosynthesis (Scheme 7.1). Until recently, only one route to the universal isoprenoid precursors was known, the so-called mevalonate pathway, using acetyl-coenzyme A as the only building block [8]. A completely distinct alternative to this well-established biosynthetic route, now known as the nonmevalonate pathway, was discovered in the early 1990s, starting from pyruvate (3) and glyceraldehyde 3-phosphate (4) [9]. Interestingly, this biosynthetic pathway is exclusively used by a number of pathogens such as the malarial parasite *Plasmodium falciparum* and the tuberculosis-causing *Mycobacterium tuberculosis* and not by higher eukaryotes (e.g., humans), which means that inhibition of the constituent enzymes of the nonmevalonate pathway affects and kills only the parasites, leaving the patient untouched (Scheme 7.1). Thus, this pathway has provided a rich source of new, highly attractive drug targets.

To illustrate the use of a structure-based design cycle, the development of the first inhibitors of the kinase IspE of the nonmevalonate pathway is described below, constituting a novel approach toward anti-infectives. First, the biosynthetic pathway

\[ 	ext{Inhibitor} \]

\[ \text{Key requirements:} \]
- Steric complementarity
- Electronic complementarity
- Noncovalent interactions

**Figure 7.1** Schematic representation of structure-based inhibitor design.
7.2 Isoprenoids and the Nonmevalonate Pathway

There are more than 35 000 known isoprenoids, which fulfill a myriad of important biological functions. Despite their striking structural diversity, all isoprenoids are biosynthesized from the two simple five-carbon building blocks 1 and 2 (Scheme 7.1). This concept is also known as the isoprene rule [10]. The nonmevalonate pathway starts with the head-to-tail condensation of the two- and three-carbon precursors 3 and 4. A total of seven enzymes catalyze the conversion of these starting materials into the essential isoprenoid precursors 1 and 2. Fosmidomycin, an inhibitor of the second enzyme of the pathway (IspC), has been shown to cure malaria in rodents, thereby validating the constituent enzymes as drug targets [11]. This finding triggered research efforts aimed at the elucidation of the structures
and mechanisms of the participating enzymes. As a result, detailed structural and mechanistic data exist for most enzymes [12], setting the stage for lead generation by structure-based design. A number of enzymes of the nonmevalonate pathway have been chosen as targets to achieve this goal. The rather hydrophilic nature of the active sites renders the development of low-molecular-weight inhibitors challenging. Thus, it comes as no surprise that the few reported inhibitors either bear phosphate or phosphonate groups or display rather modest inhibition (Table 7.1) [13].

7.2.1
4-Diphosphocytidyl-2C-methyl-D-erythritol Kinase (IspE)

The absence of known inhibitors, the fact that the kinase IspE belongs to the nonmevalonate pathway, and the availability of an X-ray crystal structure make the fourth enzyme of the pathway an ideal target for rational design of potent, drug-like inhibitors without the use of the problematic phosphate or phosphonate moieties. IspE (EC 2.7.1.148) employs adenosine 5′-triphosphate (ATP) and Mg2+ cations for the phosphorylation of the C(2)-hydroxyl group of 4-diphosphocytidyl-2C-methyl-D-erythritol (5) to afford 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate (6, Scheme 7.2) [18].

This central reaction is the only ATP-dependent step of the whole biosynthetic pathway. Sequence comparisons have shown that IspE belongs to the galactose/homoserine/mevalonate/phosphomevalonate (GHMP) kinase superfamily [19].

7.2.2
Structure of IspE

In 2003, the first crystal structures of IspE from *Thermus thermophilus* and *Escherichia coli* were solved, of the apoenzyme and of a ternary complex, respectively [20, 21]. Recently, the crystal structure of *Aquifex aeolicus* IspE was solved as a complex with a number of natural ligands [22], a synthetic substrate mimic [22], and synthetic cytidine [23] as well as cytosine (Section 7.4) [24] derivatives. To date, no crystal structure is available for IspE from a pathogen, for example, *M. tuberculosis* or *P. falciparum*. IspE generally crystallizes as a homodimer with each monomer displaying the characteristic two-domain fold of the GHMP kinase superfamily that consists of an ATP- and a substrate-binding domain. The dimer clasps around a solvent-filled channel, featuring two active sites at either end (Figure 7.2).

7.2.3
Active Site of IspE

The active site of *E. coli* IspE was used for modeling. It can be divided into three main pockets: the ATP-, the cytidine 5′-diphosphate (CDP)-, and the methylerythritol (ME)-binding pockets. Molecular modeling, using the program MOLOC
Table 7.1 Known inhibitors of enzymes of the nonmevalonate pathway.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Inhibitors</th>
<th>Inhibition</th>
</tr>
</thead>
<tbody>
<tr>
<td>DXS</td>
<td><img src="image1" alt="Inhibitor structure" /> <img src="image2" alt="Inhibitor structure" /></td>
<td>IC₅₀ = 0.08 mM [14]</td>
</tr>
<tr>
<td>IspC</td>
<td><img src="image3" alt="Inhibitor structure" /> <img src="image4" alt="Inhibitor structure" /></td>
<td>Kᵢ = 0.9 nM [15]</td>
</tr>
<tr>
<td>IspE</td>
<td><img src="image5" alt="Inhibitor structure" /> <img src="image6" alt="Inhibitor structure" /></td>
<td>IC₅₀ = 0.45 mM [16]</td>
</tr>
<tr>
<td>IspF</td>
<td><img src="image7" alt="Inhibitor structure" /> <img src="image8" alt="Inhibitor structure" /></td>
<td>Kᵢ = 0.011 µM [17]</td>
</tr>
</tbody>
</table>

The inhibition of the best inhibitor is given. CDP = cytidine 5′-diphosphate; DXS and IspF: first and fifth enzyme of the nonmevalonate pathway, respectively; IC₅₀ = concentration of inhibitor at which 50% maximum initial velocity is observed; Kᵢ = inhibition constant.
Scheme 7.2  Simplified version of the nonmevalonate pathway.

Figure 7.2  Schematic representation of the ternary complex of E. coli ispE, 5′-adenyl-β,γ-amidotriphosphate (AMP-PNP) and the substrate 5, cocrystallized as a homodimer (Protein Data Bank (PDB) code: 1OJ4) [20].
[25], revealed that an additional small, hydrophobic pocket lies adjacent to the CDP-binding site (Figure 7.3).

The ATP-binding pocket features a glycine-rich phosphate-binding loop, typically displaying an adjacent positively charged N terminus of an α helix [20]. The adenine moiety is accommodated in a hydrophobic cleft lined by Val57, Val60, Leu66, Ile67, Lys96, and Met100. Numerous hydrophobic contacts offered by this pocket certainly make a large contribution to the binding enthalpy, as predicted by MOLOC. Additional stabilization is derived from a network of H bonds to the nucleobase moiety. The ribose moiety of 5′-adenyl-β, γ-amidotriphosphate (AMP-PNP) is solvent-exposed and does not show any contacts with the protein.

The CDP-binding pocket accommodates the cytosine moiety in a π sandwich, consisting of the side chains of Tyr25 and Phe18 held in place by – stacking interactions. The ribose moiety benefits from stabilization by a pseudo-π sandwich that is composed of the aromatic side chain of Tyr25 and the aliphatic Pro182. A pseudo-π sandwich can be defined as a π sandwich, in which one of the two aromatic rings is replaced by an aliphatic ring. His26 is a key residue for the recognition of cytosine, which involves a total of three H-bonding interactions. The ribose and phosphate groups are stabilized by a number of solvent-mediated interactions and a H bond from the side chain of Tyr25 (Figure 7.4).

IspE was shown to have high substrate specificity [21]. Hence, the cytosine moiety must play a key role in substrate recognition and thus selectivity, providing the starting point for the structure-based design of the first-generation inhibitors (Section 7.3).
7.3 Targeting the CDP-Binding Pocket of IspE

7.3.1 Design

Careful examination of the active site revealed the CDP-binding pocket to be more attractive as a target for inhibitor design than the other binding pockets for a number of reasons. A perfect setup for a double π sandwich and other recognition features should endow potential inhibitors with higher selectivity. Furthermore, given the myriad of proteins that use ATP as a cofactor, inhibitors designed to target the ATP-binding pocket bear a high risk of selectivity issues. Given that neither substrate nor cofactor bind to the adjacent, hydrophobic pocket, presumably, both affinity and selectivity could be gained by occupying it. As this small cavity lies adjacent to the CDP-binding pocket, potential inhibitors should be designed to occupy both the CDP-binding pocket and the newly discovered subpocket, leaving the hydrophilic ME- as well as the ATP-binding pockets unoccupied.

In a first round of design, cytosine was chosen as a central scaffold to position potential inhibitors in the CDP-binding pocket. According to modeling, the nucleobase moiety was predicted to be sandwiched between Tyr25 and Phe185. By analogy to the natural substrate S, the cytosine moiety was postulated to be able to form H bonds to His26. It was concluded that the central platform should be decorated with a suitable ribose analogue at the N(1) position, which should be held in place by the pseudo-π sandwich and a vector designed such that its...
substituent would be placed in the hydrophobic subpocket. If this were to be achieved, numerous hydrophobic interactions would result. By connecting the linker to the C(5) position, it should serve to address the catalytically essential residues Lys10 and Asp141. Because of the modular design, the different components, that is the ribose analogue, the vector, and the central scaffold, should be easy to vary and optimize (Figure 7.5a).

### 7.3.1.1 Possible Ribose Analogues

As ribose analogues, both heteroalicyclic and aromatic rings can be envisaged to fill the space provided by the pseudo-π sandwich (Figure 7.5b, compounds of type 7 and 8). Introduction of a saturated ring featuring a sulfur atom, for example, a tetrahydrothiophenyl ring, presumably would enable an additional sulfur–aromatic interaction with the phenolic ring of Tyr25 [27]. Modeling predicted both enantiomers of a tetrahydrothiophenyl derivative to bind with similar strength due to the conformational flexibility of five-membered rings.

Four questions need to be answered regarding the choice of the ribose analogue:

- Which is the ideal ring size?
- Is an aromatic or an aliphatic system favored?
- Are heteroatoms beneficial?
- Is a connecting methylene linker beneficial or detrimental to affinity?

![Diagram](a) Modular design of first-generation inhibitors. (b) Ribose analogues, (c) different vectors, and (d) sulfone substituents envisaged for first-generation inhibitors. *clogP values (calculated partitioning coefficient) were calculated with the program ACD/LogP [26]. The tetrahydrothiophenyl and the tetrahydrofuranyl derivatives of type 7 and 8, respectively, are chiral.

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>( R^1 )</th>
<th>clogP ( ^a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Saturated cyclic</td>
<td>0.3 to 1.0</td>
</tr>
<tr>
<td>8</td>
<td>Methylene bridged</td>
<td>–0.4 to 1.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>( R )</th>
<th>clogP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(±)-11</td>
<td>Acyclic</td>
<td>0.3 to 2.0</td>
</tr>
<tr>
<td>(±)-12</td>
<td>Cyclic</td>
<td>–0.03 to 2.6</td>
</tr>
</tbody>
</table>

![Diagram](a) Modular design of first-generation inhibitors. (b) Ribose analogues, (c) different vectors, and (d) sulfone substituents envisaged for first-generation inhibitors. *clogP values (calculated partitioning coefficient) were calculated with the program ACD/LogP [26]. The tetrahydrothiophenyl and the tetrahydrofuranyl derivatives of type 7 and 8, respectively, are chiral.
7.3.1.2 Design of the Vector

For the vector, a propargyl sulfonamide substituent at the C(5) position of cytosine was envisaged, which displays three attractive features. First, the alkyne ensures a certain rigidity and linearity. Second, N-substituted sulfonamides are known to prefer a conformation in which the lone pair of the nitrogen atom bisects the O–S–O angle, resulting in a staggered arrangement [28]. In its preferred conformation, the sulfonamide is expected to form ionic H bonds to the side chains of Lys10 and Asp141. Finally, small complementary sulfone substituents ($R^2 = Et$, Figure 7.5a), should orient directly into the subpocket. However, before this part of the vector is optimized (Section 7.3.4), the importance of the sulfonamide’s contribution to affinity should be evaluated (Section 7.3.3).

An exemplary inhibitor of type 7 features a tetrahydrothiophenyl ring as a ribose mimic ($R^1$) and an ethyl group as a sulfone substituent ($R^2$) (Figure 7.5b). According to modeling, the ligand should benefit from H bonds to His26, Lys10, and Asp141 as well as from the postulated sulfur–aromatic interaction (Figure 7.6a). An overlay of the natural substrate 5 and the potential inhibitor showed the sulfone moiety to be almost perfectly superimposed with the C(2)-hydroxyl group that is to be phosphorylated (Figure 7.6b).

The synthesis of the first representative of the first-generation inhibitors was achieved using a convergent strategy based on the Sonogashira cross-coupling reaction [29, 30]. With the synthetic route in place, the optimization of both modules of the inhibitors could be initiated in parallel.

7.3.2 Optimization of the Ribose Analogue

It was envisaged to introduce a number of different ribose analogues (Figure 7.5b). Different ring types had to be tested: saturated, aromatic, or heteroaromatic rings, all of which should be directly attached to the N(1) of the nucleobase (ligands of type 7). Furthermore, the influence of a connecting methylene group between N(1)
of cytosine and the ribose analogue needed to be evaluated both for aliphatic and (hetero-)aromatic substituents (compounds of type 8). In addition, it was deemed important to test an acyclic ribose mimic, featuring an ester group. Replacement of the sulfur atom of the tetrahydrothiophenyl ring of the exemplary inhibitor by a methylene moiety to afford a cyclopentyl ring as a ribose analogue was expected to yield information on the postulated sulfur–aromatic interaction. Synthesis of this set of potential inhibitors was achieved using similar synthetic routes [24, 29].

Using the established photometric assay to determine the activity of potential inhibitors against IspE, the IC\textsubscript{50} and \(K_i\) values for all new ligands were determined (Table 7.2) [24, 29, 31]. A competitive mechanism with respect to substrate binding was assigned to most ligands. In a few cases, however, a mixed competitive (\(K_c\))–uncompetitive (\(K_{iu}\)) mode of inhibition was found. In agreement with the observation that small cytosine derivatives possess remarkably low water solubility [24], none of the inhibitors showed high water solubility, despite the low calculated partitioning coefficient (clogP) values (Figure 7.5b).

The results obtained are highly satisfactory: First, as no inhibitors of IspE had been described at the onset of this design cycle, obtaining the first active compounds with \(K_i\) values in the upper-nanomolar range constitutes an important achievement. Second, inhibition was possible in the absence of a phosphate or phosphonate group. The ligands described constitute the first example of potent, drug-like inhibitors of an enzyme of the nonmevalonate pathway. Furthermore, variation of the ribose mimic has a clear effect on affinity, that is, structure–activity relationships (SARs) could be observed. A methylene-bridged tetrahydrofuranyl ring is the poorest ribose substitute with a double-digit micromolar \(K_i\) value. Introduction of a methylene-bridged aromatic ring (benzyl) or an open alkyl chain bearing an ester moiety as a ribose mimic clearly also has a negative effect on affinity. However, it seems that a methylene-bridged ring is tolerated as long as it is not too bulky: the methylene-bridged cyclobutyl and pyrazolyl derivatives feature similar affinities to that of the inhibitor bearing a cyclopentyl ring directly attached to N(1) of cytosine. On the basis of the affinities, there does not seem to be any

Table 7.2 Inhibitory activities (\textit{E. coli} IspE) of compounds of type 7 and 8.

<table>
<thead>
<tr>
<th>Ribose analogue</th>
<th>(K_i) ((\mu)M)(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tetrahydrothiophenyl</td>
<td>0.29 ± 0.1</td>
</tr>
<tr>
<td>Cyclopentyl</td>
<td>1.5 ± 0.2</td>
</tr>
<tr>
<td>Cyclobutylmethyl</td>
<td>1.5 ± 0.2</td>
</tr>
<tr>
<td>CH(_2)–3-pyrazolyl</td>
<td>1.6 ± 0.1</td>
</tr>
<tr>
<td>Benzy1</td>
<td>3.7 ± 0.5(^b)</td>
</tr>
<tr>
<td>CH(_2)CO(_2)Et</td>
<td>4.2 ± 0.6(^b)</td>
</tr>
<tr>
<td>CH(_2)–2-tetrahydrofuranyl</td>
<td>32.3 ± 2.8</td>
</tr>
</tbody>
</table>

\(^a\)The IC\textsubscript{50} values can be found in 24, 29, 30.

\(^b\)Mixed inhibition: \(K_{iu} = 23.5 ± 7.1\) \(\mu\)M.

\(^c\)Mixed inhibition: \(K_{iu} = 21.6 ± 6.2\) \(\mu\)M.
significant difference between a heteroaromatic ring such as pyrazolyl and the alicyclic cyclopentyl ring. However, if clogP values are taken into account, the latter inhibitor is significantly more lipophilic; hence, the resulting more favorable partitioning from the aqueous buffer into the less polar protein environment could partially account for its affinity. This implies that heteroatoms afford some affinity.

Finally, the tetrahydrothiophenyl ring is the best ribose mimic, affording the lowest $K_i$ value. Presumably, its improved affinity could be ascribed to the favorable sulfur–aromatic interaction. This interaction was quantified when comparing the affinities of the tetrahydrothiophenyl derivative with an inhibitor bearing the same sulfone substituent (cyclopropyl) and a cyclopentyl ring as a ribose analogue. The affinity is lowered by a factor of five, corresponding to a free-enthalpy increase of $\Delta\Delta G_{300 K} \approx 1$ kcal mol$^{-1}$, in agreement with published data [27]. Both rings benefit from stabilization by the pseudo-$\pi$ sandwich. Thus, the higher affinity could be explained by favorable interactions of the phenolic side chain and the sulfur atom of the sandwiched ring.

7.3.3 Importance of the Vector

Before optimizing the propargylic sulfonamide vector, its significance has to be confirmed. According to modeling, the sulfonamide group forms two H bonds to Lys10 and Asp141. To verify this prediction experimentally, four derivatives featuring modified vectors and established ribose analogues were designed and synthesized (Figure 7.5c): (i) The N-methylated derivative of the most potent inhibitor, featuring a cyclopropyl ring as sulfone substituent and a tetrahydrothiophenyl ring as ribose analogue, served to quantify the importance of the ionic H bond between the sulfonamide NH and the side chain of Asp141. (ii) Different propargyl amine derivatives [compounds of type (±)-9] were envisaged to evaluate the effect of substituting the sulfonamide altogether. The set of target molecules were synthesized following similar routes [24].

Using the established enzymatic assay, the IC$_{50}$ and $K_i$ values of the potential inhibitors were determined [24, 29, 31]. N-methylation of the sulfonamide nitrogen atom clearly affected affinity, illustrating the importance of the postulated ionic H bond. Modeling showed the inhibitor to be still accommodated in the active site without any repulsive interactions but with one less H bond. Comparison with the most potent inhibitor enabled the quantification of the postulated ionic H bond, given that both ribose analogue and sulfone substituent remained unchanged. The inhibitory potency was reduced by a factor of nearly 10 upon N-methylation ($K_i = 2.5$ μM); thus, a contribution of up to $\Delta\Delta G_{300 K} = 1.3$ kcal mol$^{-1}$ to the overall binding free enthalpy could be ascribed to this H bond alone (Figure 7.7a).

As expected, substitution of the sulfonamide moiety by heterocyclic amines led to a decrease in inhibitory potency. A piperidyl-substituted derivative maintains the highest affinity ($K_i = 4.7$ μM). A possible explanation could be an ionic H bond between the side chain of Asp141 and the piperidinium residue (Figure 7.7b). According to modeling, the piperidinyl ring should be located at the entrance of the
7.3 Targeting the CDP-Binding Pocket of IspE

small, hydrophobic pocket. In this way, a number of hydrophobic interactions with the residues lining this cavity are still possible. A pyrrolidinyl-substituted derivative was proposed to bind in a similar manner, displaying fewer hydrophobic contacts because of its smaller ring. As a result, a decrease in binding affinity was observed ($K_i = 11.8 \mu M$). The alcohol shows very weak affinity for IspE; consequently, no $K_i$ value could be determined. This observation could be explained by the lack of both a sulfonamide moiety and an alkyl substituent to benefit from potential interactions in the small, hydrophobic cavity.

In summary, important contributions of the sulfonamide moiety to the observed binding affinity were clearly confirmed through this set of derivatives. With the sulfonamide moiety validated as a good vector, the next logical step was therefore the optimization of the sulfone substituent R (Figure 7.5d).

7.3.4 Optimization of the Filling of the Small, Hydrophobic Pocket

Inspired by the finding that optimal volume occupancy can be an important contributor to affinity [32], the small, hydrophobic pocket of IspE was carefully examined. For this purpose, a series of derivatives of type ($\pm$)-11 and ($\pm$)-12 were designed, differing only in the sulfone substituent as shown in Figure 7.5d. By keeping the scaffold, the vector and the ribose analogue constant, the effects of different substituents on affinity should be directly comparable. The target molecules were readily synthesized [24]. The choice of substituents was guided by the predictions made by molecular modeling and a concept from conventional supramolecular chemistry – the “55% rule” – that was recently applied to enzymes for the first time [33].

7.3.4.1 The “55% Rule”

Investigation of the optimal volume occupancy of the cavity space confined by capsular synthetic receptors by Mecozzi and Rebek led to the “55% rule,” stating
that the most stable inclusion complex forms if 55 ± 9% of the apolar space is occupied by the guest [34]. This concept holds true, from synthetic supramolecular chemistry [35]. Recently, Zürcher et al. applied it to the filling of a hydrophobic cavity in the active site of the antimalarial target plasmepsin II [33]. The van-der-Waals interactions in the cavity are not ideal at smaller packing coefficients (PCs). At higher PCs, however, large entropic losses, resulting from a decrease in the mobility of the binding partners, counteract enthalpic gains.

The small, hydrophobic pocket of *E. coli* IspE was estimated to be around 100 Å³ by filling the pocket with a hydrocarbon network [34, 36]. Thus, the PCs were calculated for different target molecules, which guided their design.

### 7.3.4.2 Evaluation of Inhibitors Featuring Different Sulfone Substituents

A series of potential inhibitors were subjected to the enzymatic assay. The inhibitory activities as well as the calculated PCs for selected compounds are summarized in Table 7.3.

A number of derivatives prepared feature *K*ᵢ values in the nanomolar range. In general, small alkyl chains with a maximum length of two carbon atoms or cyclic alkyl substituents up to a ring size of five are well suited to fill the cavity; hence, it is presumably more flexible on the sides than at its bottom. The three-membered ring seems to be ideal to fill this lipophilic pocket. A 2,2,2-trifluoroethyl group almost affords the same affinity as the most potent inhibitor (featuring a cyclopropyl ring).

#### Table 7.3 Inhibitory activities (E. coli IspE) and PCs of compounds of type (±)-11 and (±)-12.

<table>
<thead>
<tr>
<th>Sulfone substituent</th>
<th><em>K</em>ᵢ (µM)</th>
<th>PC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclopropyl</td>
<td>0.29 ± 0.1</td>
<td>56</td>
</tr>
<tr>
<td>2,2,2-Trifluoroethyl</td>
<td>0.36 ± 0.1</td>
<td>n.d.</td>
</tr>
<tr>
<td>Isopropyl</td>
<td>0.52 ± 0.1</td>
<td>62</td>
</tr>
<tr>
<td>Cyclobutyl</td>
<td>0.56 ± 0</td>
<td>69</td>
</tr>
<tr>
<td>Ethyl</td>
<td>0.64 ± 0.1</td>
<td>45</td>
</tr>
<tr>
<td>Cyclopentyl</td>
<td>0.89 ± 0.1</td>
<td>83</td>
</tr>
<tr>
<td>1,1,1-Trifluoromethyl</td>
<td>1.2 ± 0.3</td>
<td>n.d.</td>
</tr>
<tr>
<td>sec-Butyl</td>
<td>1.8 ± 0.3</td>
<td>n.d.</td>
</tr>
<tr>
<td>n-Hexyl</td>
<td>2.0 ± 0.3</td>
<td>107</td>
</tr>
<tr>
<td>Cyclohexyl</td>
<td>2.5 ± 0.4₈</td>
<td>97</td>
</tr>
<tr>
<td>Methyl</td>
<td>2.6 ± 0.1</td>
<td>28</td>
</tr>
<tr>
<td>n-Butyl</td>
<td>8.0 ± 0.1</td>
<td>76</td>
</tr>
<tr>
<td>n-Propyl</td>
<td>8.2 ± 1.7₉</td>
<td>61</td>
</tr>
<tr>
<td>Phenyl</td>
<td>16.3 ± 1.0</td>
<td>n.d.</td>
</tr>
</tbody>
</table>

*₈Mixed inhibition: *K*ᵢ = 19.0 ± 9.2 µM.
*₉Mixed inhibition: *K*ᵢ = 67.7 ± 35 µM.
*₉Mixed inhibition: *K*ᵢ = 27.3 ± 11 µM.

n.d. = not determined.
7.3 Targeting the CDP-Binding Pocket of IspE

and is nearly twice as strong as the corresponding ethyl derivative (Figure 7.6a). The affinity of isopropyl- and cyclobutyl-substituted ligands is very similar to that of the ethyl-substituted derivative; however, the alkyl residues seem to be slightly too large for optimal volume occupancy, leading to a decrease in affinity. As the pocket is not properly filled by ligands with a smaller alkyl substituent (methyl), the binding affinity is reduced.

Evaluation of the set of derivatives in terms of their PCs showed binding affinity to correlate with volume occupancy: a cyclopropyl ring has a PC of 56% and the lowest \( K_i \) value \( (K_i = 0.29 \mu M, \text{ Figure 7.8a}). \)

Lower (28% for methyl and 45% for ethyl substituents) or higher PCs (62% for isopropyl and 69% for cyclobutyl) are mirrored by weaker inhibition. A slight increase in size of the sulfone substituent – that is, extension of the ethyl substituent by one or two carbon atoms – led to a strong decrease in affinity. According to modeling, the \( n \)-propyl substituent might still fit into the cavity at the cost of adopting the energetically less favorable gauche conformation (PC 61%). The \( n \)-butyl substituent, however, cannot be accommodated by the pocket, even when contorted. Thus, the propargylic sulfonamide linker could equally well undergo a conformational change to direct this substituent out of the pocket into solvent-exposed space. Larger substituents (PC > 80%), such as cyclopentyl, cyclohexyl, or \( n \)-hexyl, were predicted to direct their alkyl substituents toward the opposite direction, that is, toward the solvent (Figure 7.8b).

The increased lipophilicity (cf. clogP values) and the resulting more favorable partitioning to the less polar protein environment from the aqueous buffer could explain the increased binding affinity.

In summary, by exploring the small, hydrophobic pocket, the affinity of the inhibitors was improved and another example for the application of the “55% rule” to an enzymatic context was provided.

Figure 7.8 (a) van-der-Waals surfaces of the cyclopropyl ring of the most potent inhibitor of type (±)-11 and the protein in the small, hydrophobic pocket. (b) MOLOC-generated molecular model of the alkyl-substituted inhibitors in the active site of \textit{E. coli} IspE \,(PDB code: 1OJ4) [20]. Color code: C skeleton of the inhibitors: methyl, cyan; ethyl, magenta; \( n \)-propyl, green; \( n \)-hexyl, light pink. Reproduced with permission from the Royal Society of Chemistry.
7.3.5 Summary of the First-Generation Inhibitors

In conclusion, addressing the CDP-binding pocket was rewarded with the discovery of the first, potent, small-molecule inhibitors of IspE. The affinity could be improved through rational modifications of the lead compound, giving access to an extensive set of SARs; for the time being, they provide the only indication that the proposed binding mode is correct. X-ray-crystallographic studies had to be performed to validate this hypothesis (Section 7.4).

7.4 X-ray Cocrystal Structure Analysis

Efforts were undertaken to obtain a cocrystal structure of the first-generation inhibitors and IspE. The kinase is very sensitive to crystallization conditions; even traces of organic cosolvents are sufficient to preclude crystal growth. Presumably, because of the lack of water solubility of the first-generation inhibitors, no cocrystal structures could be obtained. Thus, a water-soluble derivative had to be designed (Section 7.4.1). Provided it resembled the structure of the inhibitors in hand, it could prove the binding mode of this class of compounds.

7.4.1 Design of Water-Soluble Inhibitors

Modeling as well as clogP values were used to guide the selection of promising target molecules (Figure 7.9).

To improve the water solubility of the first-generation inhibitors without a concomitant decrease in affinity, the sulfone substituent or the sulfonamide vector (target molecules of type (±)-13 or the ribose analogue (ligands of type 14) could be modified. Given that modestly potent inhibitors could be cocrystallized (Section 7.2.2), the proposed series of target molecules looked promising as long as water solubility could be achieved. Two derivatives were obtained by varying

![Chemical Structure](image)

Figure 7.9 Target molecules to improve the water solubility of the first-generation inhibitors.

<table>
<thead>
<tr>
<th>Inhibitor</th>
<th>R¹</th>
<th>R²</th>
<th>clogP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(±)-13</td>
<td>2-Tetrahydrothiophenyl</td>
<td>Morpholinyl based</td>
<td>−0.48 to 0.9</td>
</tr>
<tr>
<td>14</td>
<td>Carboxylic-acid-based</td>
<td>Cyclopropyl</td>
<td>−0.6 to 0.4</td>
</tr>
</tbody>
</table>
the vector, whereas the remaining three resulted from a modification of the ribose analogue. Introduction of a carboxylic acid or the corresponding ester functionality, a morpholinyl, or an oxetanyl substituent should afford the desired property. Oxetanes, in particular, were recently described to endow compounds with improved physicochemical properties, namely, affording enhanced solubility and decreased lipophilicity [37].

7.4.2 Enzyme Assays of Inhibitors Designed to be Water Soluble

Among the six compounds specifically prepared to obtain water-soluble inhibitors, two (the morpholinyl-substituted sulfonamide ($K_i = 13.1 \mu M$) and the oxetanyl derivative (compound 15, Figure 7.10b, $K_i = 28.7 \mu M$)) do not require addition of dimethyl sulfoxide as a cosolvent to perform the enzyme assay [24]. While the former still requires ethanol as a cosolvent, the latter is water soluble, setting the stage for X-ray crystallographic studies (Section 7.4.3).

Water solubility was achieved at the expense of potency: the affinity was decreased by nearly two orders of magnitude, when replacing the tetrahydrothiophenyl ring of the most potent inhibitor by the oxetanyl substituent of inhibitor 15. Because of their poor affinities, only the IC$_{50}$ values of the carboxylic acid derivatives were determined (upper micromolar range). The weak inhibitory potency of the morpholinyl derivative lacking the sulfonamide moiety could be explained by the bad match between the electronics of the morpholinyl moiety and the small, hydrophobic subpocket. Thus, introduction of an oxygen atom to the piperidinyl derivative of type ($\pm$)-9 (Section 7.3.3), affording the corresponding morpholinyl-substituted ligand, resulted in a decrease in affinity by almost a factor of nine. A similar decrease in affinity was observed when changing the sulfone substituent from cyclohexyl (compound of type ($\pm$)-12) to morpholinyl. The poor affinity of the carboxylic acid derivatives could be attributed to the less than optimal ribose mimics.

![Figure 7.10](image-url) (a) Binding mode of inhibitor 15 in the CDP-binding pocket of active site B (PDB code: 2VF3) [24]. (b) Structure of compound 15.
7.4.3 Structural Analysis

The structure of the oxetanyl derivative 15 in complex with \textit{A. aeolicus} IspE was determined to 2.2 Å resolution (PDB code: 2VF3) [24]. Two molecules are present in the asymmetric unit, featuring “active sites A and B.” Evidence for the proposed binding mode of inhibitor 15 was provided by the cocrystal structure. The water-soluble ligand indeed binds in the CDP-binding pocket. As a result of the differences in amino-acid sequences of \textit{A. aeolicus} and \textit{E. coli} IspE [30], the sulfonamide and nucleobase moieties form somewhat different H-bonding patterns compared to those predicted by modeling. In particular, two additional H bonds stabilize the cytosine moiety of ligand 15 (Figure 7.10a). The cytosine moiety is engaged in four H bonds to the side chain and backbone amide of His25 in active site B. The amino group forms an additional H bond to the backbone amide C=O of Lys145. The cytosine ring is sandwiched by the side chains of Tyr24 and Tyr175 (\textit{=} Phe185 in \textit{E. coli} IspE). A network of H bonds to the side chains of Asn11, Tyr31, and Asp130 stabilizes the sulfonamide moiety (Figure 7.10a).

Although AMP was present in the crystallization conditions, the electron density observed in the glycine-rich loop of the ATP-binding pocket was incompatible with this compound. Modeling and refinement of the density as diphosphate proved successful. The latter unit forms numerous H bonds to the backbone amide NH groups of the glycine-rich loop, in analogy to the diphosphate moiety of AMP-PNP in the crystal structure of the \textit{E. coli} enzyme complex [20]. With one exception (Gly95), all the backbone amide NH moieties of glycine residues of the loop are involved in H-bonding. The residues of the glycine-rich loop, which are not glycine, however, do not participate in H-bonding interactions with diphosphate. This underlines the importance of the presence of glycine residues in phosphate-binding pockets. Presumably, glycine is ideally suited for the binding of phosphate groups in such loops as it can adopt the required conformation to wrap the loop around the bound phosphate group [38].

For future efforts aimed at the design of inhibitors for IspE from pathogens such as \textit{P. falciparum} or \textit{M. tuberculosis}, the finding that the cyclopropyl substituent of ligand 15 is accommodated in the hydrophobic pocket is of particular interest (Figure 7.10b). Given that the change from Phe185 (in \textit{E. coli} IspE) to Tyr175 (in \textit{A. Aeolicus} IspE) increases the hydrophilicity of the small pocket, the cyclopropyl ring was not expected to be located in this subpocket as it precludes solvation of the hydroxyl group of Tyr175. This is an important finding, given that the enzymes from \textit{P. falciparum} and \textit{M. tuberculosis} possess a conserved tyrosine residue at this position, and that – according to homology modeling – the phenolic hydroxyl group is directed into the small pocket. The propargylic sulfonamide vector of this class of inhibitors is therefore highly suitable for precisely addressing the phenolic hydroxyl group in future design cycles.
7.4.4 Lessons Learnt from the Cocrystal Structure

By designing water-soluble inhibitors and solving the cocrystal structure of one ligand with *A. aeolicus* IspE, the suggested binding mode was validated. Thus, the first design cycle was successfully completed. The inhibitor occupies the CDP-binding pocket and exhibits a $K_i$ value in the lower micromolar activity range. This proof of concept opens the way for further modification and optimization of the inhibitors aimed at the development of ligands with activity against IspE from medically important organisms. Thus, the cocrystal structure solved represents an important step on the way to anti-infectives with a novel mode of action.

7.5 Conclusions and Outlook

7.5.1 Conclusions

The active site of the kinase IspE features both highly polar and hydrophobic subpockets. Because the highly polar subpockets do not lend themselves well to structure-based drug design, the inhibitors were targeted to the lipophilic regions of the active site. Using structure-based design, the first inhibitors of the enzyme were developed, which display drug-like properties and highly satisfactory potency. A number of conclusions can be drawn from the series of inhibitors:

- Inhibition of IspE is possible whilst bypassing the highly polar phosphate-binding pocket in the absence of a phosphate moiety. Therefore, the ligands synthesized constitute the first potent and drug-like inhibitors of an enzyme of the nonmevalonate pathway.
- The inhibitors target only the CDP-binding pocket, potentially affording better selectivity than inhibitors designed to bind to the ATP-binding pocket.
- The majority of the binding free enthalpy is derived from the cytosine scaffold and the propargylic sulfonamide vector.
- A systematic variation of the vector’s substituent provided access to SARs, confirming the proposed binding mode. In addition, these modifications gave rise to the second example of the “55% rule” for the optimal filling of cavities applied to an enzymatic context.
- Preparation of a water-soluble derivative enabled crystallographic studies of an enzyme–inhibitor complex, validating the proposed binding mode. This cocrystal structure provided invaluable structural information, especially in view of fine-tuning the ligands to obtain activity against the pathogenic enzymes.

In conclusion, the new target IspE revealed itself to be ideally suited to structure-based inhibitor design. The inhibitors synthesized in the context of this project represent a successful application of this design strategy.
7.5.2 Outlook

The design and synthesis of the first-generation inhibitors of IspE opened up a number of future research avenues.

First, as a result of the observed difference in amino acid sequences between the model system \textit{E. coli} and the pathogenic enzymes – marked by the replacement of the phenylalanine residue lining the hydrophobic pocket by a tyrosine – the structure of the inhibitors will have to be fine-tuned to address this structural difference. Second, the bisubstrate approach could be tried to obtain even more potent and selective inhibitors of IspE. In principle, it should be possible to decorate the established scaffold to occupy the different pockets of the active site. Third, it would be very rewarding to identify an attractive mimic in the quest for new cytosine analogues. As opposed to the other nucleobases, very few cytosine substitutes have been described to date. Finally, identification of a generally applicable way of rendering promising inhibitors water soluble without having to undertake a time-consuming design cycle and synthesis of the new derivatives would be of great interest. In this way, attractive structures could be designed, synthesized, and assayed without worrying about their physicochemical properties such as solubility.

Acknowledgments

This work was carried out under the supervision of Prof. F. Diederich and in collaboration with the groups of Prof. A. Bacher and Prof. W. N. Hunter. Financial support by the Roche Research Foundation and the ETH Research Council is gratefully acknowledged.

List of Abbreviations

\begin{itemize}
\item AMP-PNP \textit{\textsuperscript{5}}-adenyl-\textit{\textbeta}, \textit{\textgamma}-amidotriphosphate
\item ATP adenosine \textit{\textsuperscript{5}}-triphosphate
\item clogP calculated partitioning coefficient
\item CDP cytidine \textit{\textsuperscript{5}}-diphosphate
\item DMAPP dimethylallyl diphosphate
\item GHMP galactose/homoserine/mevalonate/phosphomevalonate
\item HTS high-throughput screening
\item IC\textsubscript{50} concentration of inhibitor at which 50\% maximum initial velocity is observed
\item IPP isopentenyl diphosphate
\item IspE 4-diphosphocytidyl-2\textit{\textC}-methyl-\textalpha-erythritol kinase
\item \textit{K}_{i} inhibition constant
\item ME methyl-\textalpha-erythritol
\end{itemize}
References


