University of Groningen

De novo fragment-based design of inhibitors of DXS guided by spin-diffusion-based NMR spectroscopy

Masini, T.; Pilger, J.; Kroezen, B. S.; Illarionov, B.; Lottmann, P.; Fischer, M.; Griesinger, C.; Hirsch, A. K. H.

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

Link to publication in University of Groningen/UMCG research database

EDGE ARTICLE

CrossMark
← click for updates

De novo fragment-based design of inhibitors of DXS guided by spin-diffusion-based NMR spectroscopy†

T. Masini,‡[a] J. Pilger,‡[b] B. S. Kroezen,[a] B. Illarionov,[c] P. Lottmann,[b] M. Fischer,*[c] C. Griesinger*[b] and A. K. H. Hirsch*[a]

We applied for the first time an innovative ligand-based NMR methodology (STI) to a medicinal-chemistry project aimed at the development of inhibitors for the enzyme 1-deoxy-ᴅ-xylulose-5-phosphate synthase (DXS). DXS is the first enzyme of the 2*C*-methyl-ᴅ-erythritol-4-phosphate (MEP) pathway, present in most bacteria (and not in humans) and responsible for the synthesis of the essential isoprenoid precursors. We designed *de novo* a first generation of fragments, using *Deinococcus radiodurans* DXS as a model enzyme, targeting the thiamine diphosphate (TDP) pocket of DXS whilst also exploring the putative substrate-binding pocket, where selectivity over other human TDP-dependent enzymes could be gained. The STI methodology – suitable for weak binders – was essential to determine the binding mode in solution of one of the fragments, circumventing the requirement for an X-ray co-crystal structure, which is known to be particularly challenging for this specific enzyme and in general for weak binders. Based on this finding, we carried out fragment growing and optimisation, which led to a three-fold more potent fragment, about as potent as the well-established thiamine analogue deazathiamine. The STI methodology proved therefore its strong potential as a tool to support medicinal-chemistry projects in their early stages, especially when dealing with weak binders.

Introduction

Fragment-based design (FBD) has emerged as a powerful tool for the discovery and optimisation of drug-like compounds.[1] Whereas the most common approach for the identification of fragments as hits still remains the screening of fragment libraries,[2] *de novo* design of fragments is attracting more and more attention,[3] in particular for targets for which an initial screen of a fragment library is impossible owing to the type of assay used or insufficient quantities of protein available. The use of FBD is particularly attractive due to its potential to cover a greater fraction of chemical space and to afford fragments with high ligand efficiencies (LEs, defined as $LE = -\Delta G/N$ where ΔG corresponds to the Gibbs free energy of binding (kcal mol^{-1}) and N is the number of non-hydrogen atoms) that should be maintained during the optimisation stages.[4] Despite these advantages, several hurdles persist: the fragment hit identified may not be amenable to follow-up synthesis, which is often a time-consuming and resource-intensive undertaking. Furthermore, the fragment-to-lead process consists in optimising (*e.g.*, fragment growing, fragment linking) the initial fragment hit towards a more potent lead-like compound.[5] During this process, the binding mode of the initial fragment must be maintained, requiring ideally a co-crystal structure of the (modified) fragment to be solved at every optimisation step.[6]

Unfortunately, for approximately 40% of pharmaceutically relevant protein targets, crystal structures of sufficient quality cannot be obtained and for those proteins, which are crystallisable, this is often a time- and resource-consuming process, as reliable soakable crystallographic systems are not always available.[7] Moreover, it can be very difficult to co-crystallise fragments obtained in the early stage of the fragment-to-lead process because of their weak affinity for the target, requiring prohibitively high ligand concentrations.

NMR spectroscopy can be particularly helpful in these cases,[8] as it is suited for a wide range of binding affinities and can outperform other methods such as screening assays, surface plasmon resonance and fluorescence anisotropy in terms of the detection of weak binders.[9] A major advantage of NMR spectroscopy with respect to co-crystal structures is that information about the binding mode in solution is obtained. In addition, the

[a]Strathingh Institute for Chemistry, University of Groningen, Nijenborgh 7, NL-9747 AG Groningen, The Netherlands. E-mail: a.k.h.hirsch@rug.nl

[b]Max-Planck-Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. E-mail: cigr@nmr.mpibpc.mpg.de

[c]Hamburg School of Food Science, Institute of Food Chemistry, Grindelallee 117, D-20146, Hamburg, Germany. E-mail: markus.fischer@uni-hamburg.de

† Electronic supplementary Information (ESI) available: Selected docking poses, procedures for modelling and docking, synthesis of the fragments, NMR analysis, MD simulations, kinetic studies as well as NMR spectra. See DOI: 10.1039/c4sc00588k

‡ These authors contributed equally to this work.

rigid nature of co-crystal structures can lead to misinterpretations of the interactions between the designed ligand and the protein.[10]

Whereas protein-observed NMR experiments are limited to small protein targets (M_w < 50 kDa), which can be labelled and purified in large amounts,[5b] ligand-observed NMR experiments require less, unlabelled protein and enable direct identification of the ligands as long as their chemical shifts in solution are known.[11]

We recently developed and validated retrospectively an innovative, fast and reliable ligand-based NMR methodology[12] extending the INPHARMA methodology[13] to decipher the binding mode of a ligand in solution using a very low concentration of unlabelled protein (e.g., 30 μM). The latter aspect makes this technique available for all targets with a certain size (M_w > 30 kDa), including less established ones for which protein expression is often the bottleneck. The method covers a wide range of binding affinities (K_d from 1.0 μM to 2.0 mM), typically observed for weakly to moderately binding fragments.
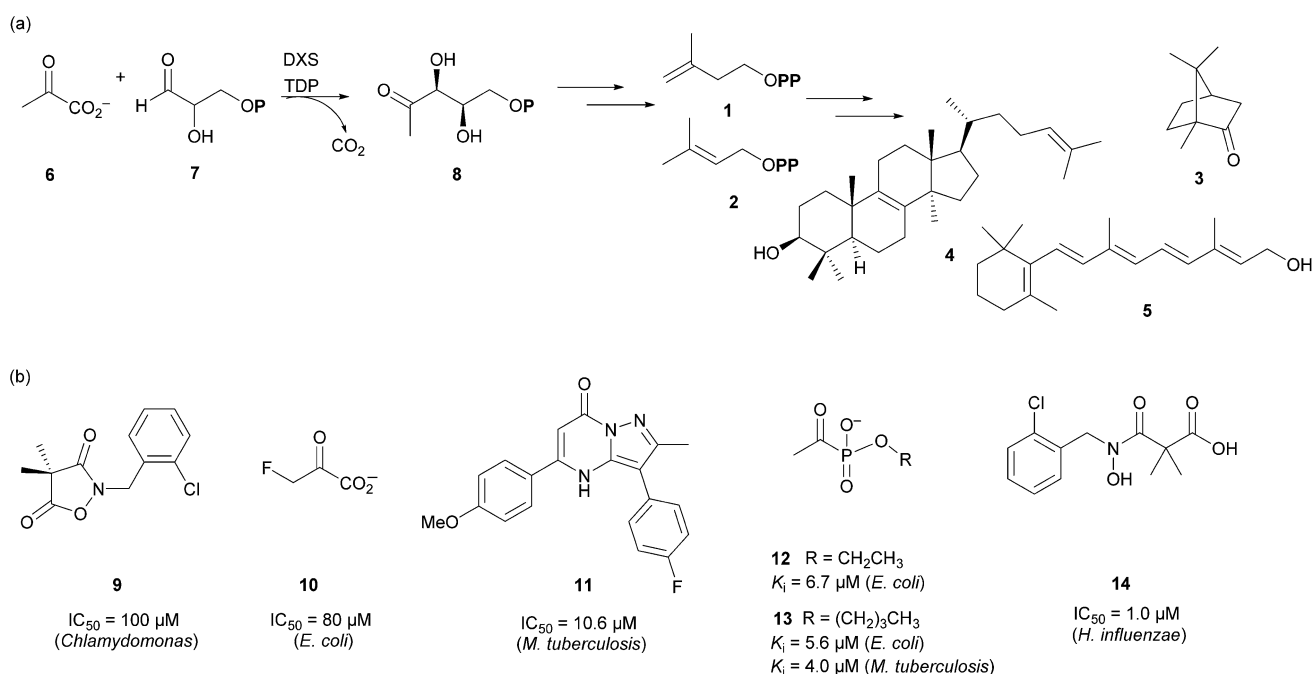
A combination of three NMR techniques is applied to elucidate the binding mode of a ligand or fragment: (i) Saturation-Transfer Difference (STD) yields information on the protein-buried and water-exposed parts of the ligand,[14] (ii) Transfer-NOE (trNOE) reveals the bound ligand conformation and (iii) INPHARMA derives the relative binding mode of two ligands interacting with the same binding site.[13] The methodology is termed STI, referring to its three NMR parameters STD, trNOE and INPHARMA and holds the potential to represent an attractive alternative to protein crystallography in both FBD and structure-based design (SBD) projects. An INPHARMA-based identification of the binding site of fragment-like molecules for

glycogen phosphorylase has been reported recently, which illustrates the potential of this method.[15]

In the present work, we apply the STI methodology for the first time to the *de novo* FBD of inhibitors of the enzyme 1-deoxy-D-xylulose-5-phosphate synthase (DXS), illustrating its potential when applied in the early stages of medicinal-chemistry projects.

DXS catalyses the first and rate-limiting step of the 2*C*-methyl-D-erythritol-4-phosphate (MEP) pathway (also known as non-mevalonate pathway) for the biosynthesis of isopentenyl diphosphate (IPP, **1**) and dimethylallyl diphosphate (DMAPP, **2**).[16] IPP and DMAPP are the universal precursors of isoprenoids, a large and important class of natural products (*e.g.*, camphor (**3**) lanosterol (**4**) and vitamin A (**5**), Scheme 1a) featuring very diverse structures and essential biological functions.[17] The first step of the MEP pathway catalysed by DXS consists in the thiamine-diphosphate (TDP)-dependent decarboxylative condensation of pyruvate (**6**) and D-glyceraldehyde 3-phosphate (**7**, GAP) to afford 1-deoxy-D-xylulose-5-phosphate (**8**, Scheme 1a). Whereas pathogens such as *Mycobacterium tuberculosis* and *Plasmodium falciparum*, causative agents of tuberculosis and malaria, respectively, use the MEP pathway for the biosynthesis of **1** and **2**, humans exclusively utilise the alternative mevalonate pathway, making the enzymes of the MEP pathway attractive drug targets for the development of new drugs against tuberculosis and malaria.[18] Due to the emergence of several multi- and extensively-drug-resistant strains of *M. tuberculosis* and *P. falciparum*, there is a growing need for the development of new anti-infective agents with new mechanisms of action.[19]

The involvement of DXS both in microbial thiamine (vitamin-B_1) biosynthesis[20] as well as in vitamin-B_6 biosynthesis[21] renders it a particularly interesting target, with the benefit of interfering

Scheme 1 (a) First step of the MEP pathway catalysed by DXS to afford the final products of the MEP pathway, **1** and **2**, leading to three representative isoprenoids, namely camphor (**3**) lanosterol (**4**) and vitamin A (**5**). (b) Known inhibitors of DXS.

with three crucial bacterial biosynthetic pathways at once. On the one hand, the similarity between DXS and mammalian TDP-dependent enzymes such as transketolase (TK)[22] and pyruvate dehydrogenase (PDH)[23] could give rise to selectivity problems in the development of inhibitors against DXS, especially when targeting the TDP-binding pocket (Fig. S1,† 20% identity overall, 47% identity in the TDP-binding pocket between human TK and *M. tuberculosis* DXS). On the other hand, DXS has distinctive features compared to other TDP-dependent enzymes suggesting that selective inhibition of DXS over other TDP-dependent enzymes could be possible: the peculiar domain arrangement of *Deinococcus radiodurans* DXS (RCSB Protein Data Bank (PDB) code: 2O1X) where the TDP-binding site is located within the same monomer[24] and not at the dimer interface like in its mammalian homologues,[26] the particularly large active site[27] and its unique catalytic mechanism, which requires the formation of a ternary complex with pyruvate and GAP[28] (in contrast, all the other TDP-dependent enzymes follow classical ping-pong kinetics in which the binding of acceptor substrate is preceded by the activation of pyruvate and release of CO_2)[29] Therefore, ideally, a DXS-inhibitor should be designed so as to target both the TDP-binding pocket to enhance the affinity and the substrate-binding pocket to afford selectivity.

The two crystal structures of DXS in complex with TDP deposited in the PDB[24] are not of the pathogenic organisms, making the development of inhibitors for this enzyme a challenge. As a result, very few inhibitors of DXS have been reported in the literature (**9–14**, Scheme 1b)[30] and for none of them the binding mode has been validated. Only phosphonates **12** and **13** have been shown to compete with the binding of pyruvate by kinetic studies.[30d,f]

Results and discussion

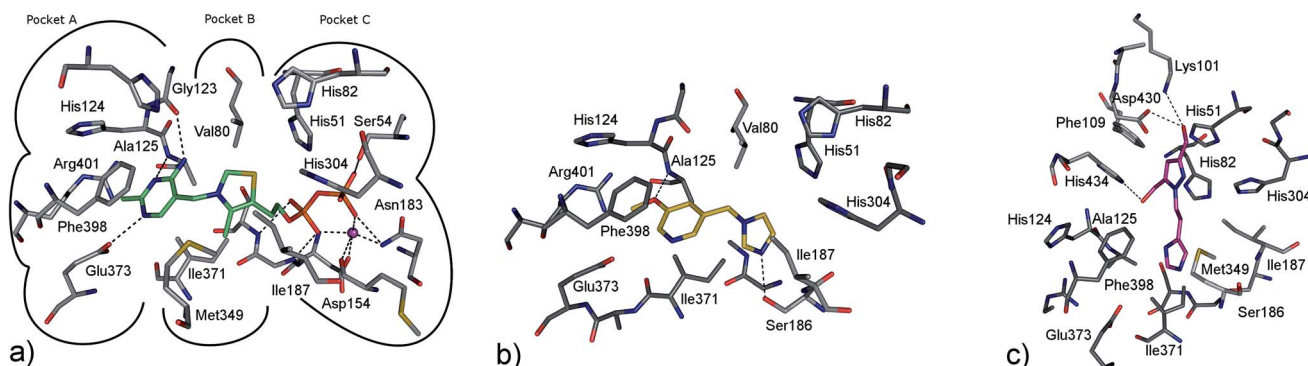### *De novo* design and biochemical evaluation of first-generation fragments

As a starting point for our FBD, we decided to target the TDP-binding pocket rather than the ill-defined substrate-binding

pocket. Despite the low lipophilic character of the TDP pocket with respect to the average for druggable pockets,[32] it obtained a rather high Dscore when analysed with the programme DogSiteScore.[33] The fragments were designed so that they could be further grown either towards the diphosphate pocket or ideally towards the substrate-binding pocket, where selectivity could be gained.

We used *D. radiodurans* DXS as a model enzyme for our studies given the high degree of sequence identity between *D. radiodurans* DXS and *M. tuberculosis* DXS (38% overall, 68% in the TDP-binding pocket (Fig. S1†).[24] As a first step towards the development of inhibitors for DXS, we focused on designing TDP-competitive fragments and therefore first analysed the binding mode of TDP, which is deeply buried inside *D. radiodurans* DXS. As highlighted in Fig. 1a, the aminopyrimidine moiety of TDP is involved in three hydrogen bonds with the protein ($d(\text{N3}\cdots\text{HN–Ala125}) = 3.1$ Å, $d(\text{H}_2\text{N}\cdots\text{O}=\text{C–Gly123}) = 2.9$ Å, $d(\text{N1}\cdots\text{HO–C–Glu373}) = 2.7$ Å) and a π-stacking interaction with the side chain of Phe398. The catalytically active part of TDP, the thiazolium ring, is surface-exposed and only involved in a few hydrophobic interactions (with Ile187, Ile371 and Val80), while the diphosphate moiety engages in numerous polar interactions both with polar amino-acid side chains and with the protein backbone. The catalytic Mg^{2+} ion is complexed by both phosphate groups.
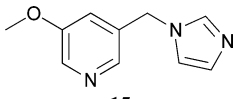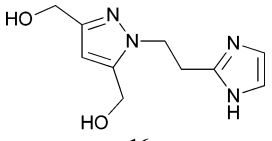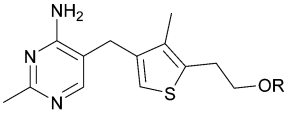
The binding of TDP with DXS is rather tight (*D. radiodurans* DXS, $K_d = 0.114 \pm 0.01$ μM; *M. tuberculosis* DXS, $K_d = 3.1 \pm 0.30$ μM, Fig. S2†). Moreover, by comparing the inhibitory potency of deazathiamine **17** (IC$_{50}$ = 430 ± 68 μM) and its diphosphorylated form, deazathiamine diphosphate **18** (IC$_{50}$ = 0.034 ± 0.006 μM),[34] it is evident that the diphosphate moiety accounts for a good part of its binding energy. Being aware of the challenges in targeting very polar diphosphate-binding pockets, we decided to occupy the more druggable aminopyrimidine and thiazolium pockets.

Using the modelling software MOLOC,[35] we designed fragment **15** *de novo* (Fig 1b, Table 1). Docking with the FlexX docking module in the LeadIT suite further guided our

Fig. 1 (a) TDP in the TDP–binding pocket of *D. radiodurans* DXS (PDB code: 2O1X). The three subpockets are highlighted and labelled as pocket A (aminopyrimidine pocket), B (thiazolium pocket) and C (diphosphate pocket), respectively. (b) Modelled binding mode of fragment **15** in *D. radiodurans* DXS. (c) Modelled binding mode of fragment **16** in *D. radiodurans* DXS. Colour code: protein skeleton: C: grey; O: red; N: blue; S: yellow. Mg^{2+} ion is shown as a purple sphere. Inhibitor **15** skeleton: C: yellow. Inhibitor **16** skeleton: C: pink. Hydrogen bonds below 3.3 Å are shown as dashed lines. The units for the distances indicated and the colour code are maintained throughout, if not stated otherwise. Figures were generated using the software PyMOL.[25]

Table 1 Experimental IC$_{50}$, experimental or calculated K_i values and ligand efficiencies (LEs) of two modelled fragments (**15**, **16**) and deazathiamine (**17**); LEs are derived from experimentally determined IC$_{50}$ values unless stated otherwise

Inhibitors	IC$_{50}^a$ (μM)	K_i^a (μM)	LE^b
15	1810 ± 480	448 ± 129	0.33
16	762 ± 199	183 ± 52	0.32
17: R = H	430 ± 68	151 ± 34^c, competitive mode of inhibition	0.29^c
18, R = P$_2$O$_6^{3-}$	0.034 ± 0.006^d	0.007 ± 0.001	0.44

^{a} IC$_{50}$ and calculated K_i values were determined using a photometric assay using the programme Dynafit.[31] Full details of the biochemical assay conditions and the calculation of the K_i values are provided in the ESI. ^{b} Values indicate the ligand efficiency (LE) calculated as $-\Delta G/N$, where ΔG corresponds to the Gibbs free binding energy (kcal mol^{-1}) calculated as $\Delta G = -RT \ln K_0/K_i$, K_i being the calculated inhibition constant, unless stated otherwise. N corresponds to the number of non-hydrogen atoms of that specific molecule. ^{c} The K_i value and mode of inhibition of **17** were determined experimentally. LE was derived from the experimental K_i. Full details of the biochemical-assay conditions are provided in the ESI. ^{d} **18** was isolated and tested as its ammonium tosylate salt.

modelling studies.[36] The favourable π-stacking interaction with Phe398 is maintained in the same way as for TDP, but the pyridine ring of **15** is slightly displaced with respect to the amino-pyrimidine ring of TDP leading to a more favourable parallel-displaced π-stacking interaction. In the modelled pose, the imidazole ring of **15** is accommodated in the thiazolium pocket and is displaced with respect to the thiazolium ring of TDP. As a result, the van-der-Waals interactions with Ile187 and Ile371 – which help to keep the thiazolium ring in place in the case of TDP – should be strengthened.

According to the modelling studies, the imidazolic N3 is involved in a favourable hydrogen bond with the side chain of Ser186 (d(N···HO–Ser186) = 3.0 Å). The 3-methoxy group of the pyridine ring forms a hydrogen bond with the backbone N–H of Ala125 (d(O···HN–Ala125 = 2.8 Å), mimicking the hydrogen bond of the N3 of the aminopyrimidine ring of TDP. Not surprisingly, the N1 of the pyridine ring of **15** is not involved in a hydrogen bond with Glu373 anymore, due to the displacement of the ring with respect to the aminopyrimidine ring of TDP. We carried out the synthesis of **15** in four steps using an S$_N$2 reaction as a key step to link the two aromatic rings (Scheme S1†). Biochemical evaluation of **15** using a coupled DXS-IspC (1-deoxy-D-xylulose-5-phosphate reductoisomerase) photometric assay showed an IC$_{50}$ = 1.8 ± 0.5 mM against *D. radiodurans* DXS

(experimentally derived LE = 0.33 kcal mol^{-1}). The imidazolyl ring is a suitable platform to address both the diphosphate-binding pocket and the newly identified pocket lined by residues His304, His82, His51, Phe109 and His43. This pocket, from now on referred to as "histidine-rich pocket", is presumably involved in substrate-binding.

Fragment **16** (Fig. 1c, Table 1) was designed *de novo* so as to explore this pocket, presumably involved also in the substrate-recognition process. According to the modelling studies, we introduced a pyrazole moiety bearing a diol functionality to hydrogen bond to Lys101 (d(OH···$_2$HN–Lys101 = 3.3 Å), Asp430 (d(OH···O=C–Asp430 = 2.8 Å) and His434(d(OH···N–His434 = 2.8 Å). Although the pyrazolyl ring is not directly interacting with His51, it is believed to interfere with the catalytic role of His51, supposedly in the substrate-recognition process.[37] An ethylene linker (torsional angle τ(N$_{pyrazole}$–C–C–C$_{imidazole}$): $-168°$) connects the pyrazolyl ring to an imidazolyl ring accommodated in the thiazolium pocket. In addition to the van-der-Waals interactions with the side chains of Ile187 and Ile371, the imidazolyl ring is involved in a favourable S-aromatic interaction with the side-chain sulfur atom of Met349. The linker itself is also involved in numerous lipophilic contacts with the side chain of Ile187 and aromatic hydrogen atoms of Phe398. The synthesis of **16** relied on a regioselective alkylation of an appropriately protected imidazole followed by introduction of the pyrazole in an S$_N$2 reaction (Scheme S2†). Biochemical evaluation of **16** resulted in an IC$_{50}$ of 762 ± 199 μM, with an experimentally derived LE of 0.32 kcal mol^{-1}. Just like for fragment **15**, the imidazolyl ring constitutes an equally versatile platform for further fragment growing and optimisation of **16**. Based on the LEs, both **15** and **16** can be considered suitable starting points for further optimisation.

Determination of the binding mode in solution using the STI methodology

Before starting fragment optimisation, we validated the binding mode of the two fragment hits identified. To do so, we applied the novel STI NMR methodology (Fig. S3†) for the first designed fragments given the lack of a reliable crystallographic protocol and the opportunity to obtain information about the binding mode of small molecules in solution. In the first stage, we conducted STD experiments with ligands **15** and **16** (Fig. S4 and S5†). Both molecules showed STD enhancement in the presence of DXS, indicating that they bind to DXS. As expected, we observed no STD enhancement for the fragments in a buffer solution without DXS. The same observation holds true for trNOE peaks in a NOESY spectrum. For the third step of the methodology (INPHARMA), we chose deazathiamine (**17**) as a reference ligand for our studies (STD experiments shown in Fig. S6†).[34] **17** is structurally related to TDP and docking studies supported our assumption that it binds in the TDP-binding pocket (Fig. S7†). Moreover, the experimental determination of the K_i value for **17** (K_i = 151 ± 34 μM) confirmed that its mode of inhibition is competitive with respect to TDP, making it a suitable ligand for the INPHARMA methodology (Fig. S8†).

In order to observe INPHARMA peaks in a spectrum, the ligands must have an off-rate from the target, which is large enough to ensure that several exchange events can take place during the NOESY mixing time of several 100 ms. When both molecules target the same binding pocket, magnetisation of the protons of the first molecule will be transferred to the protein protons and back to the protons of the second molecule, when it is bound. **17** is a suitable choice in our case, given its IC_{50} of 430 ± 68 μM against *D. radiodurans* DXS.

Whereas we recorded a NOESY spectrum of the ligand combination **17** and **15** (Fig. S9 and S10†) and were able to clearly observe INPHARMA peaks, we were unable to observe INPHARMA peaks for ligand combinations **17** with **16** and **15** with **16**. We used the docking module FlexX to dock **15** and **16** into the crystal structure of *D. radiodurans* DXS. We generated 30 docking poses for each ligand and clustered them with an RMSD difference of 3 Å using GROMACS to yield a set of representative and diverse binding modes. From every cluster, we used the structure with the best ΔG according to the scoring function HYDE in the LeadIT suite for the back-calculation of NMR data (Fig. S11 and S12, Table S1†).[38] We also docked **17** in the TDP-binding site of *D. radiodurans* DXS and, as expected, found the top-ranked poses to overlap with TDP, which we used as a reference for the INPHARMA calculations (Fig. S7, Table S1†). We back-calculated the NMR peak volumes of all three data sets (STD, trNOE and INPHARMA) for every docked pose, using the software SpINPHARMA (http://www.inpharma.de) and scored the correlation between back-calculated and experimental data using the Pearson correlation coefficient to yield R_{STD}, R_{trNOE} and $R_{INPHARMA}$. We then combined these to R_{STI} by averaging and selected the structure pair with the highest score as the most accurate binding mode. The R_{STI} values for every docked pose of each inhibitor are reported in Table S1.† For **17**, we identified three different and representative binding modes by docking (Fig. S7†), all of them preserving the π–π stacking interaction with Phe398 observed for TDP. Nevertheless, the hydroxyethylene tail seems to be rather flexible and adopts different orientations in the three poses, being involved in hydrogen bonds with various amino acids (His82, His51 or His 304). We had observed before for another protein system that different binding modes of one ligand directly contribute to the spin diffusion and the NMR signals represent an average of them.[12] In order to better understand the binding behaviour of **17**, we used all three representative binding poses as a starting point of an MD simulation for 1 ns (Fig. S13†). It is remarkable that only one pose maintains its binding mode during the MD simulation (Fig. S7,† docked pose 2), confirming that this pose corresponds to the binding mode of **17** to *D. radiodurans* DXS in solution (Fig. 2b). The best-scoring binding modes for **17** and **15** based on their R_{STI} are shown in Fig. 2b. It is noteworthy that for ligand **15** the best scoring R_{STI} (0.52) binding mode (Fig. S11,† docked pose 1) is very similar to the modelled one (Fig. 1b). In the other plausible binding mode proposed by the docking studies, **15** binds in the diphosphate pocket of TDP (Fig. S7,† docked pose 2) but displays negative R_{STD} and $R_{INPHARMA}$ values. The STI methodology was therefore able to clearly differentiate between both binding modes and allowed us to validate the
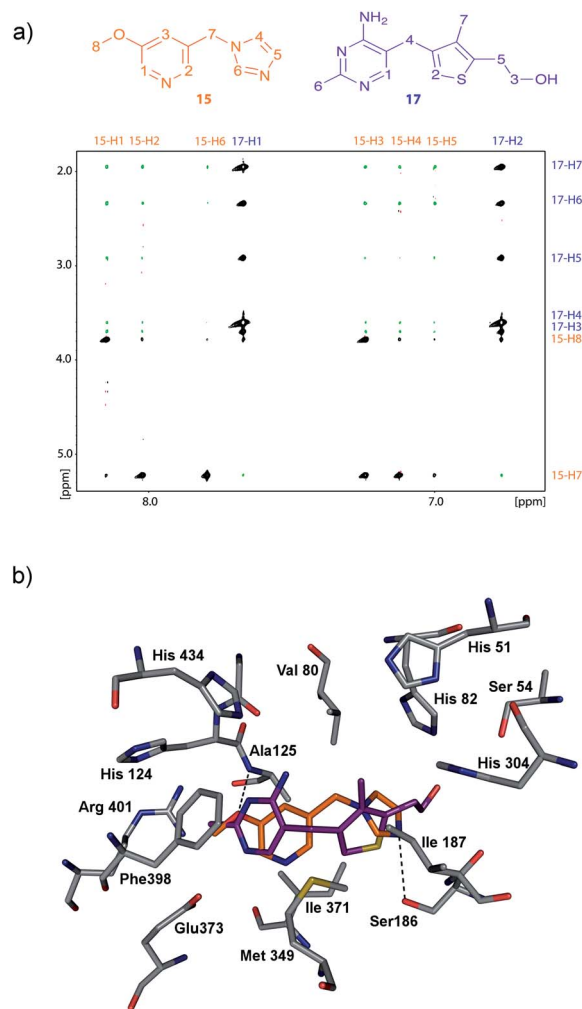

Fig. 2 (a) NOESY spectrum of inhibitors **15** and **17** in the presence of DXS clearly shows trNOE (black) and INPHARMA (green) peaks. Peak assignment is indicated on the axes for **15** (black) and **17** (blue). (b) Binding mode for **15** and **17** in the TDP-binding pocket of *D. radiodurans* DXS validated by the STI methodology. Deazathiamine (**17**) skeleton: C: purple. Inhibitor **15** skeleton: C: orange.

modelled binding mode of **15**. In the case of **16**, we could only record STD and trNOE data. Docking identified two representative binding modes of **16** in the TDP-binding pocket of *D. radiodurans* DXS, which are shown in Fig. 3 (Fig. S12, Table S1†). The absence of INPHARMA peaks could be explained by the small overlap between **16** and **17** in both poses. Owing to the lack of hydrogen-bonding interactions and the inherently flexible ethylene linker, the imidazolyl moiety presumably adopts various conformational states. MD simulations for 1 ns of the two selected representative poses further confirm this assumption. In fact, although the two poses are maintained during the simulation period, we observed several rotations of the imidazolyl moiety. The high degree of flexibility combined with the small overlapping area with **17** are putatively the reasons why we only observed weak trNOE and no INPHARMA peaks. Regarding the R_{ST} values, one of the two poses (Fig. S11,† docked pose 2, R_{ST} = 0.45) seems to be slightly preferred over
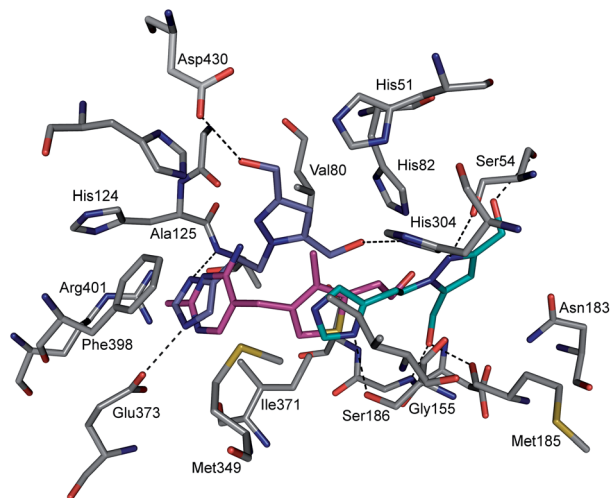
Fig. 3 The two representative docked poses of **16** in the TDP-binding pocket of *D. radiodurans* DXS. The validated binding mode of deaza-thiamine (**17**) is also shown. Deazathiamine (**17**) skeleton: C: magenta. Inhibitor **16** skeleton: C: blue or C: light blue.
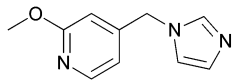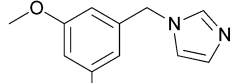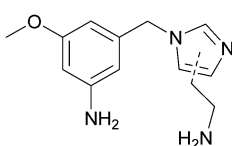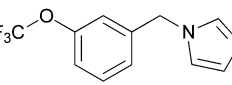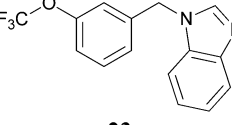
the other (Fig. S12,† docked pose 1, $R_{ST} = 0.42$). Given that the validation of the binding mode within *D. radiodurans* DXS was possible just for **15** and given that **15** and **16** are equally amenable to follow-up synthesis, we focused our attention on the growing and optimisation of fragment **15**.

Fragment growing and optimisation

To do so, we modelled and synthesised the compounds shown in Table 2 (Fig. S14, Scheme S1–S6†). The synthesis involves in every case a late-stage introduction of the moiety targeting the thiazolium and diphosphate pockets in an S_N2 reaction. Fragment **19** constitutes a regioisomer of **15** and we designed it so as to gain a hydrogen-bonding interaction with the side chain of Glu373, provided that the carboxylic acid is present in its protonated form. To circumvent the uncertainty related to the protonation state of the side chain of Glu373, we designed compound **20**. According to the modelled pose, the amino group of **20** should form two strong hydrogen bonds with the side chain of Glu373 – whether protonated or not – as well as with the backbone C=O of Pro347. For both fragments, all the other interactions of the original fragment **15** are preserved. As shown in Table 2, we observed only 20% and 0% inhibition at 2 mM for **19** and **20**, respectively. Addressing the side chain of Glu373 seems therefore unreliable in terms of molecular recognition and addressing other amino-acid residues in the TDP-binding site of DXS would appear to be a better strategy for SBD.

To grow the fragment towards the diphosphate-binding pocket, we designed and tested **21** as a mixture of two regioisomers (**21a** and **21b**), which displays an IC$_{50}$ value of 1.2 ± 0.2 mM. A possible explanation of the higher inhibitory potency of **21** with respect to **20** could be the capability of the aminoethylene linker to anchor **21** in a slightly different orientation with respect to **20**, which might help the binding

Table 2 Experimental IC$_{50}$, calculated K_i values and ligand efficiencies (LEs) of synthesised derivatives of fragment **15**. LEs are derived from experimentally determined IC$_{50}$ values

Inhibitors	IC$_{50}$[a] (μM)	K_i[a] (μM)	LE[b]
19	>2000[c]	—	—
20	>2000	—	—
21a + 21b	1200 ± 200[d]	310 ± 65	0.27
22	>2000[e]	—	—
23	595 ± 81	151 ± 22	0.24

[a] IC$_{50}$ and calculated K_i values were determined using a photometric assay using the programme Dynafit.[31] Full details of the biochemical assay conditions and the calculation of the K_i values are provided in the ESI. [b] Values indicate the experimentally derived ligand efficiency (LE) calculated as $-\Delta G/N$, where ΔG corresponds to the Gibbs free binding energy (kcal mol^{-1}) calculated as $\Delta G = -RT \ln K_0/K_i$, K_i being the calculated inhibition constant. N corresponds to the number of non-hydrogen atoms of that specific molecule. [c] 20% inhibition at 2000 μM. [d] Compound **21** has been tested as a mixture of 4- and 5-substituted regioisomers (**21a** + **21b**). [e] 40% inhibition at 2000 μM.

event to occur. This favourable anchoring both on the left- and right-hand sides of the molecule is probably responsible for the gain in potency. We accessed **22** conveniently in one step, starting from imidazole and the corresponding commercially available bromide. In the modelled pose, the π-stacking inter-action with Phe398 is preserved as well as the two hydrogen bonds of **15**. Moreover, the trifluoromethyl group is involved in numerous lipophilic interactions with the side chains of Phe398, His124 and Ile371. It is also engaged in an electrostatic interaction with the positively charged guanidinium side chain of Arg401. The inhibitory potency of **22** (40% inhibition at 2 mM) can be considered comparable to that of **15** but its optimisation is more straightforward owing to the reduced number of synthetic steps. In addition, a benzene ring has more free substitution points compared to a pyridine ring. To increase the lipophilic contacts of **22**, we expanded the imidazolyl ring into a benzimidazolyl moiety, leading to compound **23**, which
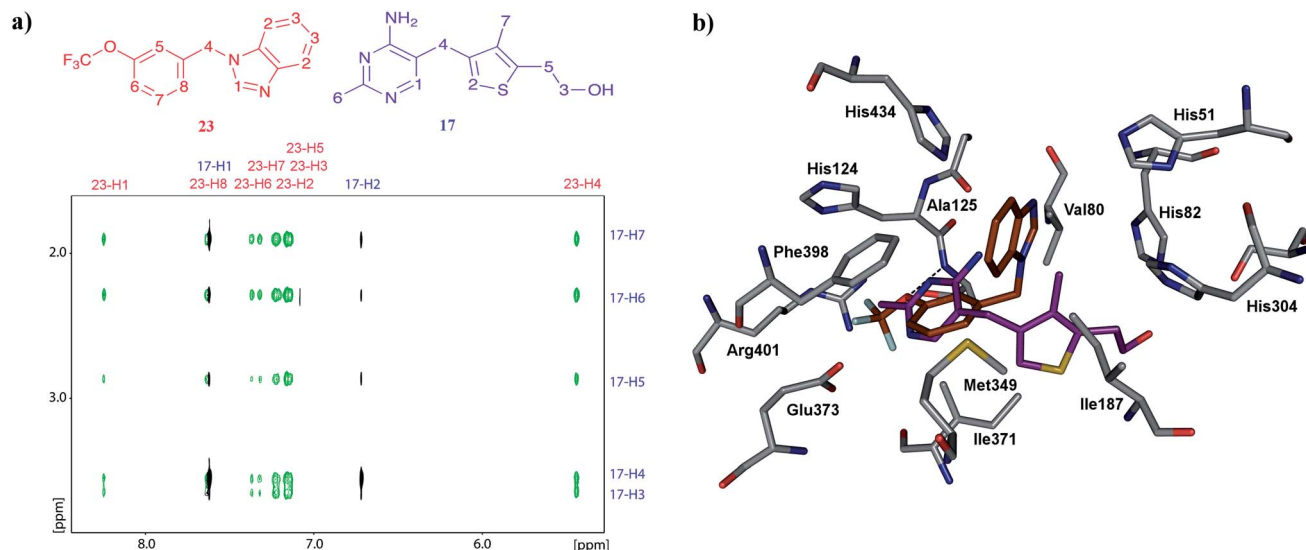
Fig. 4 (a) NOESY spectrum of inhibitors **23** and **17** in the presence of DXS clearly shows trNOE (black) and INPHARMA (green) peaks. Peak assignment is indicated on the axes for **23** (red) and **17** (purple). (b) Validated binding mode for **23** in the TDP-binding pocket of *D. radiodurans* DXS. Colour code: inhibitor **23** skeleton: C: orange.

interacts, according to our modelling, with Val80, His82 and Ile187. As a result, **23** ($IC_{50} = 595 \pm 81\ \mu M$) is three times more potent than the original fragment **15**. The inhibitory potency of **23** is in the same range as that of the non-selective thiamine analogue deazathiamine (**17**; $IC_{50} = 430 \pm 68\ \mu M$), making **23** an attractive and novel thiamine analogue. The fact that this fragment was designed *de novo* could facilitate and enable subsequent tuning of the selectivity over human thiamine- or TDP-binding targets.

We validated the binding mode of **23**, using the STI method. Firstly, we recorded an STD spectrum of **23**, which showed STD enhancement only in the presence of DXS, confirming the binding event (Fig. S15†). We observed the same in a trNOE spectrum (Fig. 4a). In the subsequent NOESY spectrum of **23** and deazathiamine (**17**) we clearly observed INPHARMA peaks (Fig. 4a), confirming that both compounds bind in the same pocket. As described above for **15**, we generated three representative poses for **23** (Fig. S16, Table S1†). The docked pose of **23** displaying the best R_{STI} score ($R_{STI} = 0.57$) is shown in Fig. 4b (Fig. S16,† docked pose 3). Two other plausible docking modes show binding of **23** in the diphosphate pocket (Fig. S16,† docked pose 1, $R_{STI} = 0.4$) and the histidine-rich pocket (Fig. S16,† docked pose 2, $R_{STI} = 0.46$), both having lower ΔG values than the pose shown in Fig. 4b, according to HYDE scoring (Table S1†).

In the best pose according to the NMR studies (Fig. 4b), the left-hand side of **23** maintains the same interactions with DXS with respect to the originally modelled pose, while the benzimidazolyl ring points into the histidine-rich pocket rather than towards the diphosphate pocket as is the case for the modelled pose of **23**. We did not observe the modelled pose of **23** amongst the 30 docked poses but it was included in the evaluation and, not surprisingly, received a lower R_{STI} score ($R_{STI} = 0.29$): this fact further validates the pose shown in Fig. 4b for **23**.

Taking advantage of the NMR studies, applied here for the first time, we were able to investigate the binding mode of **23**, setting the stage for further fragment-growing and optimisation to afford lead-like compounds with improved inhibitory potency, even in the case of a very complicate binding behaviour. It is worthwhile to underline that the binding pose validated with the STI approach taken here, differs from the pose obtained purely by modelling and that the structure-based fragment optimisation of **23** based on the modelled pose alone would have led to fragment growing in the wrong direction. This observation confirms that the validation of the binding mode of the fragments in the first steps of the fragment-to-lead process is fundamental to render the whole process more efficient and less time- and material-consuming.

Conclusions

In conclusion, we have successfully applied the innovative NMR-based STI methodology for the first time to a SBD project. In our *de novo* FBD of inhibitors of *D. radiodurans* DXS, the STI methodology was essential to validate the binding mode of the first promising fragment (**15**, $IC_{50} = 1.8 \pm 0.5\ mM$) and enabled us to focus our efforts on just one fragment and perform further fragment-growing and optimisation leading to **23**. Compound **23** is three times more potent than **15** ($IC_{50} = 595 \pm 81\ \mu M$) and about as potent as deazathiamine (**17**), making **23** an attractive thiamine analogue with the potential for better selectivity over other thiamine- or TDP-dependent enzymes, which will have to be realised in future cycles of optimisation. We investigated the binding mode of **23** using STI and found it to be different from the one modelled in the absence of experimental data. Our findings show that the STI approach represents an attractive alternative to protein X-ray crystallography enabling SBD and FBD projects on targets for which a reliable crystallographic

protocol is not available. A further advantage is that the binding mode in solution is derived. The STI methodology therefore holds the potential to find application in a wide range of medicinal-chemistry or chemical-biology projects.

Acknowledgements

Notes and references

1 (a) P. J. Hajduk and J. Greer, *Nat. Rev. Drug Discovery*, 2007, **6**, 211–219; (b) M. Congreve, G. Chessari, D. Tisi and A. J. Woodhead, *J. Med. Chem.*, 2008, **51**, 3661–3680.

2 D. A. Erlanson, *Top. Curr. Chem.*, 2012, **317**, 1–32.

3 F. X. Talamas, G. Ao-Ieong, K. A. Brameld, E. Chin, J. de Vicente, J. P. Dunn, M. Ghate, A. M. Giannetti, S. F. Harris, S. S. Labadie, V. Leveque, J. Li, A. S.-T. Lui, K. L. McCaleb, I. Nájera, R. C. Schoenfeld, B. Wang and A. Wong, *J. Med. Chem.*, 2013, **56**, 3115–3119.

4 (a) C. Abad-Zapatero, *Expert Opin. Drug Discovery*, 2007, **2**, 469–488; (b) S. D. Bembenek, B. A. Tounge and C. H. Reynolds, *Drug Discovery Today*, 2009, **14**, 278–283.

5 (a) D. C. Rees, M. Congreve, C. W. Murray and R. Carr, *Nat. Rev. Drug Discovery*, 2004, **3**, 660–672; (b) S. B. Shuker, P. J. Hajduk, R. P. Meadows and S. W. Fesik, *Science*, 1996, **274**, 1531–1534.

6 C. W. Murray, M. L. Verdonk and D. C. Rees, *Trends Pharmacol. Sci.*, 2012, **33**, 224–232.

7 A. L. Hopkins and C. R. Groom, *Nat. Rev. Drug Discovery*, 2002, **1**, 727–730.

8 M. Pellecchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giralt, W. Jahnke, T. L. James, S. W. Homans, H. Kessler, C. Luchinat, B. Meyer, H. Oschkinat, J. Peng, H. Schwalbe and G. Siegal, *Nat. Rev. Drug Discovery*, 2008, **7**, 738–745.

9 C. Dalvit, *Drug Discovery Today*, 2009, **14**, 1051–1057.

10 (a) V. S. Honndorf, N. Coudevylle, S. Laufer, S. Becker and C. Griesinger, *Angew. Chem., Int. Ed.*, 2008, **47**, 3548–3551; (b) R. S. Armen, J. Chen and C. L. Brooks III, *J. Chem. Theory Comput.*, 2009, **5**, 2909–2923; (c) M. Vogtherr, K. Saxena, S. Hoelder, S. Grimme, M. Betz, U. Schieborr, B. Pescatore, M. Robin, L. Delarbre, T. Langer, K. U. Wendt and H. Schwalbe, *Angew. Chem., Int. Ed.*, 2006, **45**, 993–997; (d) A. M. Davis, S. J. Teague and G. J. Kleywegt, *Angew. Chem., Int. Ed.*, 2003, **42**, 2718–2736.

11 P. Śledź, C. Abell and A. Ciulli, in NMR of Biomolecules: Towards Mechanistic Systems Biology, ed. I. Bertini, K. S. McGreevy and G. Parigi, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2012, ch. 15, pp. 265–280.

12 J. Pilger, A. Mazur, P. Monecke, H. Schreuder, B. Elshorst, S. Bartoschek, T. Langer, A. Schiffer, I. Krimm, M. Wegstroth, D. Lee, G. Hessler, K.-U. Wendt, S. Becker and C. Griesinger, *Angew. Chem., Int. Ed.*, in review.

13 (a) V. M. Sánchez-Pedregal, M. Reese, J. Meiler, M. J. J. Blommers, C. Griesinger and T. Carlomagno, *Angew. Chem., Int. Ed.*, 2005, **44**, 4172–4175; (b) J. Orts, J. Tuma, M. Reese, S. K. Grimm, P. Monecke, S. Bartoschek, A. Schiffer, K. U. Wendt, C. Griesinger and T. Carlomagno, *Angew. Chem., Int. Ed.*, 2008, **47**, 7736–7740; (c) M. Reese, V. M. Sánchez-Pedregal, K. Kubicek, J. Meiler, M. J. J. Blommers, C. Griesinger and T. Carlomagno, *Angew. Chem., Int. Ed.*, 2007, **46**, 1864–1868; (d) J. Orts, S. K. Grimm, C. Griesinger, K. U. Wendt, S. Bartoschek and T. Carlomagno, *Chem.–Eur. J.*, 2008, **14**, 7517–7520; (e) L. Skjaerven, L. Codutti, A. Angelini, M. Grimaldi, D. Latek, P. Monecke, M. J. Dreyer and T. Carlomagno, *J. Am. Chem. Soc.*, 2013, **135**, 5819–5827; (f) K. Ono, K. Takeuchi, H. Ueda, Y. Morita, R. Tanimura, I. Shimada and H. Takahashi, *Angew. Chem., Int. Ed.*, 2014, **53**, 2597–2601.

14 M. Mayer and B. Meyer, *J. Am. Chem. Soc.*, 2001, **123**, 6108–6117.

15 I. Krimm, *Med. Chem. Commun.*, 2012, **3**, 605–610.

16 (a) A. K. H. Hirsch and F. Diederich, *Chimia*, 2008, **62**, 226–230; (b) I. Hale, P. M. O'Neill, N. G. Berry, A. Odom and R. Sharma, *Med. Chem. Commun.*, 2012, **3**, 418–433; (c) T. Gräwert, M. Groll, F. Rohdich, A. Bacher and W. Eisenreich, *Cell. Mol. Life Sci.*, 2011, **68**, 3797–3814.

17 (a) L. Ruzicka, *Experientia*, 1953, **9**, 357–367; (b) A. Eschenmoser and D. Arigoni, *Helv. Chim. Acta*, 2005, **88**, 3011–3050.

18 J.-Y. van der Meer and A. K. H. Hirsch, *Nat. Prod. Rep.*, 2012, **29**, 721–728.

19 (a) A. Koul, E. Arnoult, N. Lounis, J. Guillemont and K. Andries, *Nature*, 2011, **469**, 483–490; (b) M. Schlitzer and R. Ortmann, *ChemMedChem*, 2010, **5**, 1837–1840.

20 Q. Du, H. Wang and J. Xie, *Int. J. Biol. Sci.*, 2011, 41–52.

21 R. E. Hill, K. Himmeldirk, I. A. Kennedy, R. M. Pauloski, B. G. Sayer, E. Wolf and I. D. Spenser, *J. Biol. Chem.*, 1996, **271**, 30426–30435.

22 E. Fiedler, S. Thorell, T. Sandalova, R. Golbik, S. Konig and G. Schneider, *J. Mol. Biol.*, 1995, **238**, 387–404.

23 P. Arjunan, N. Nemeria, A. Brunskill, K. Chandrasekhar, M. Sax, Y. Yan, F. Jordan, J. R. Guest and W. Furey, *Biochemistry*, 2002, **41**, 5213–5221.

24 S. Xiang, G. Usunow, G. Lange, M. Busch and L. Tong, *J. Biol. Chem.*, 2007, **282**, 2676–2682.

25 *The PyMOL Molecular Graphics System, Version 1.4*, Schrödinger, LLC.

26 R. A. Frank, F. J. Leeper and B. F. Luisi, *Cell. Mol. Life Sci.*, 2007, **64**, 892–905.

27 F. Morris, R. Vierling, L. Boucher, J. Bosch and C. L. Freel Meyers, *ChemBioChem*, 2013, **14**, 1309–1315.

28 (a) L. M. Eubanks and C. D. R. Poulter, *Biochemistry*, 2003, **42**, 1140–1149; (b) N. S. Patel, L. A. Nemeria, C. L. Brammer, C. L. Freel Meyers and F. Jordan, *J. Am. Chem. Soc.*, 2012, **134**, 18374–18379; (c) L. A. Brammer,

J. M. Smith, H. Wade and C. L. Freel Meyers, *J. Biol. Chem.*, 2011, **286**, 36522–36531.

29 R. Kluger and K. Tittmann, *Chem. Rev.*, 2008, **108**, 1797–1833.

30 (*a*) C. Mueller, J. Schwender, J. Zeidler and H. K. Lichtenthaler, *Biochem. Soc. Trans.*, 2000, **28**, 792–793; (*b*) B. Altincicek, M. Hintz, S. Sanderbrand, J. Wiesner, E. Beck and H. Jomaa, *FEMS Microbiol. Lett.*, 2000, **190**, 329–333; (*c*) J. Mao, H. Eoh, R. He, Y. Wang, B. Wan, S. G. Franzblau, D. C. Crick and A. P. Kozikowski, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 5320–5323; (*d*) J. M. Smith, R. J. Vierling and C. L. Freel Meyers, *Med. Chem. Commun.*, 2012, **3**, 65–67; (*e*) D. Hayashi, N. Kato, T. Kuzuyama, Y. Sato and J. Ohkanda, *Chem. Commun.*, 2013, **49**, 5535–5537; (*f*) J. M. Smith, N. V. Warrington, R. J. Vierling, M. L. Kuhn, W. F. Anderson, A. T. Koppisch and C. L. Freel Meyers, *J. Antibiot.*, 2014, **67**, 77–83.

31 P. Kuzmic, *Anal. Biochem.*, 1996, **237**, 260–273.

32 A. Volkamer, D. Kuhn, T. Grombacher, F. Rippmann and M. Rarey, *J. Chem. Inf. Model.*, 2012, **52**, 360–372.

33 T. Masini, B. S. Kroezen and A. K. H. Hirsch, *Drug Discovery Today*, 2013, **18**, 1256–1262.

34 H. Zhao, L. P. S. de Carvalho, C. Nathan and O. Ouerfelli, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 6472–6474.

35 P. R. Gerber and K. Müller, *J. Comput.-Aided Mol. Des.*, 1995, **9**, 251–268.

36 BioSolveIT GmbH, Sankt Augustin, http://www.biosolveit.de, LeadIT, 2.1.2.

37 J. Querol, M. Rodríguez-Concepción, A. Boronat and S. Imperial, *Biochem. Biophys. Res. Commun.*, 2001, **289**, 155–160.

38 (*a*) N. Schneider, S. Hindle, G. Lange, R. Klein, J. Albrecht, H. Briem, K. Beyer, H. Claussen, M. Gastreich, C. Lemmen and M. Rarey, *J. Comput.-Aided Mol. Des.*, 2012, **26**, 701–723; (*b*) I. Reulecke, G. Lange, J. Albrecht, R. Klein and M. Rarey, *ChemMedChem*, 2008, **3**, 885–897.