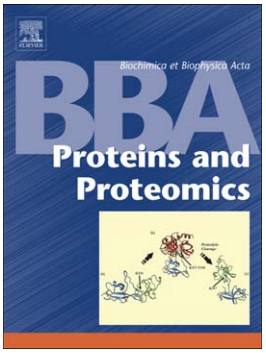


Crystal structures of *T.vivax* nucleoside hydrolase in complex with new potent and specific inhibitors

Wim Versées, Annelies Goeminne, Maya Berg, An Vandemeulebroucke, Achiel Haemers, Koen Augustyns, Jan Steyaert

PII: S1570-9639(09)00033-8
DOI: doi:10.1016/j.bbapap.2009.02.011
Reference: BBAPAP 38138

To appear in: *BBA - Proteins and Proteomics*

Received date: 1 October 2008
Revised date: 2 February 2009
Accepted date: 18 February 2009

Please cite this article as: Wim Versées, Annelies Goeminne, Maya Berg, An Vandemeulebroucke, Achiel Haemers, Koen Augustyns, Jan Steyaert, Crystal structures of *T.vivax* nucleoside hydrolase in complex with new potent and specific inhibitors, *BBA - Proteins and Proteomics* (2009), doi:10.1016/j.bbapap.2009.02.011

Crystal structures of *T.vivax* nucleoside hydrolase in complex with new potent and specific inhibitors

Wim Versées[1,2,§], Annelies Goeminne[3,‡], Maya Berg[3], An Vandemeulebroucke[1,2], Achiel Haemers[3], Koen Augustyns[3], Jan Steyaert[1,2]

[1] Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium

[2] Department of Molecular and Cellular Interactions, VIB, Pleinlaan 2, B-1050 Brussels, Belgium.

[3] Department of Medicinal Chemistry, University of Antwerp, Universiteitsplein 1, 2610 Antwerp, Belgium.

[§]Corresponding author:
Wim Versées
Structural Biology Brussels, Vrije Universiteit Brussel, Oefenplein, Gebouw E, Pleinlaan 2, 1050 Brussel, Belgium
Tel: ++32 2 629 18 49
Fax: ++32 2 629 19 63
Email: wversees@vub.ac.be

[‡] Present address: Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Science, Monash University, 381 Royal Parade, Melbourne VIC 3052, Australia.

Keywords: nucleoside hydrolase, inhibitor design, transition state inhibitors, iminoribitol, leaving group activation, loop movement

Summary

Diseases caused by parasitic protozoa remain a major health problem, mainly due to old toxic drugs and rising drug resistance. Nucleoside hydrolases are key enzymes of the purine salvage pathway of parasites from the *Trypanosomatidae* family and are considered as possible drug targets. *N*-Arylmethyl substituted iminoribitols have been developed as selective nanomolar affinity inhibitors against the purine-specific nucleoside hydrolase of *Trypanosoma vivax*. The current paper describes the crystal structures of the *T. vivax* nucleoside hydrolase in complex with two of these inhibitors, to 1.3 and 1.85 Å resolution. These high resolution structures provide an accurate picture of the mode of binding of these inhibitors and their mechanism of transition-state mimicry, and are valuable tools to guide further inhibitor design. Comparison of the current structures with previously solved structures of the enzyme in complex with ground-state and transition-state-analogue inhibitors also allows for the elucidation of a detailed molecular mechanism of active-site loop opening/closing. These loop movements can be coupled to the complex kinetic mechanism of the *T. vivax* nucleoside hydrolase.

Introduction

Diseases caused by protozoan parasites remain an important health care problem, especially in developing countries. Among these protozoa, different species from the *Trypanosomatidae* family are pathogenic to humans and cause African trypanosomiasis (sleeping sickness), American trypanosomiasis (Chagas' disease) or leishmaniasis. Together these three infections are responsible for over one hundred thousand deaths annually [1]. African trypanosomiasis is endemic in large areas of Sub-Saharan Africa. The human form of the disease is caused by the *Trypanosoma* species *T. brucei gambiense* and *T. brucei rhodesiense* which are transmitted to the human host by the bite of a tsetse fly (*Glossina*). Trypanosomiasis is moreover one of the most important illnesses of domestic livestock in Africa, with gigantic economic losses as a result. The disease in domestic animals and wildlife, also known under the local name nagana, is mainly caused by the species *T. brucei brucei, T. vivax* and *T. congolense*. Drugs that are currently in use to treat the disease, like melarsoprol and eflornithine, are old and suffer from many drawbacks such as host toxicity and spreading resistance [2].

In an effort to develop new and more efficient therapies, research has been focusing in recent years on unique features of parasitic protozoa as targets for drug design. It has been proposed that the purine metabolism of the parasite could provide a good drug target [3]. In contrast to mammals, parasitic protozoa are unable to synthesize purines *de novo*, and rely instead on the purine salvage pathway to obtain purines which are essential for their survival. Nucleoside hydrolases (NHs) are key enzymes of this purine salvage pathway [4]. After uptake of purine nucleosides by specific transporters they catalyze the hydrolysis of the *N*-glycosidic bond with formation of ribose and the nucleic base. The latter is in turn further converted to the corresponding nucleotides through the action of a phosphoribosyltransferase. Although the NHs are wide-spread in nature, they do not occur in the mammalian host organism, a feature which makes the NHs particularly attractive as targets for drug design. On the basis of substrate specificity profiles the NHs have been initially classified into four groups [5], including the non-specific inosine-uridine preferring NHs (IU-NH) [6;7], the pyrimidine specific NHs (cytidine-uridine-preferring, CU-NHs) [8;9], the purine specific NHs (inosine-adenosine-guanosine-

preferring, IAG-NHs) [10;11], and the 6-oxopurine specific isozymes (inosine-guanosine-preferring, IG-NHs) [12]. Later, the availability of many sequences and the structural characterization of several NHs suggested a homology-based subdivision in three groups, where IU-NHs and the CU-NHs belong to group I, IAG-NHs to group II, and a yet uncharacterized family to group III [9]. Until now, three distinct NH activities have been reported in different trypanosome species, namely IU-NH, IAG-NH and IG-NH. Our labs have used the NH from *Trypansoma vivax* as a model system to study the mechanism of the purine specific IAG-NHs and to initiate an inhibitor design program.

Based on the transition state structure of the enzyme-catalyzed reaction, the immucillins (compound **A1**, Figure1) were developed by Vern Schramm and co-workers as nanomolar inhibitors against NHs and purine nucleoside phosphorylases (PNP) [13]. They are characterized by an iminoribitol (1,4-dideoxy-1,4-imino-D-ribitol) moiety with a deazapurine analogue attached to the C1'. Because the iminoribitol N4' atom is partially protonated at physiological pH it mimics the partial positive charge that develops in the oxocarbenium ion-like transition state. One potential drawback of these immucillins in terms of treatment against parasites via NH inhibition, is their cross-reactivity with human PNP (hPNP). hPNP is especially important in T-cells and its inhibition could lead to T-cell mediated immunosuppression and host toxicity [14]. Starting from the crystal structure of *T. vivax* IAG-NH in complex with immucillin H (compound **A1**) [15], molecular modeling studies indicated that attachment of the base moiety to the iminoribitol N4' via a methylene linker could also result in a proper fitting into the active site [16]. Consequently, a library of *N*-arylmethyl substituted iminoribitols was synthesized and screened for IAG-NH inhibition. Three low nanomolar inhibitors were identified in this study (compound **B**, **C**, **D**; see figure 1). Apart from being among the most potent inhibitors of NHs known to date [17,18], two of them (**C** and **D**) showed nearly absolute discrimination against hPNP [19].

The crystal structure of the *T. vivax* IAG-NH has previously been solved unliganded and in complex with the inhibitor 3-deaza-adenosine, the substrate inosine and more recently with the transition state analogue, immucillin H. In unliganded structures of *Tv*NH, two flexible loops (called loops 1 and 2) with low or no associated electron density are present on either side of the active site [11]. Binding of the ground state inhibitor 3-

deaza-adenosine to *Tv*NH, or of the substrate inosine to a slow mutant of this enzyme orders loop 1 [20]. In complex with the transition state analogue immucillin H, loop 1 and loop 2 are ordered and loop 2 folds over the active site, establishing a previously unobserved closed conformation of the active site of *Tv*NH [15]. This latter structure combined with biochemical studies and quantum chemical calculations revealed a remarkable mechanism for activation of the purine leaving group in IAG-NHs [21;22]. In the active site of *T. vivax* IAG-NH the purine leaving group is stacked between the indole side chains of two tryptophans (Trp83 and Trp260). These stacking interactions, in combination with an unusual intra-molecular H-bond between the O5' and C8 atoms, increase the pK_a of the nucleobase N7 allowing its direct protonation by solvent, via a well ordered channel of water molecules.

In this paper we present the crystal structures of *T. vivax* IAG-NH in complex with the nanomolar inhibitors **B** ((2*R*,3*R*,4*S*)-1-[(4-hydroxy-5*H*-pyrrolo[3,2-*d*]pyrimidin-7-yl)methyl]-2-(hydroxymethyl)pyrrolidin-3,4-diol) and **D** ((2*R*,3*R*,4*S*)-2-(hydroxymethyl)-1-(quinolin-8-ylmethyl)pyrrolidin-3,4-diol), to 1.3 Å and 1.85 Å resolution, respectively. The former structure also represents the highest resolution structure ever solved for a nucleoside hydrolase. The crystal structures provide insight into the mode of inhibitor binding, and into the catalytic mechanism and mechanism of active site loop movement of the IAG-NHs.

Materials and Methods

Expression and purification

Expression and purification of the *T. vivax* IAG-NH were performed as described previously [11]. *E. coli* (WK6) cells containing the *iagnh* open reading frame cloned in the pQE-30 expression vector (Qiagen) were used for recombinant protein production. The presence of an N-terminal His_6-tag permitted purification in two chromatographic steps, consisting of a Ni-NTA affinity chromatography step (Qiagen) and gel filtration on a Superdex-200 column (Amersham Bioscience). The concentration of pure protein (expressed per monomer) was determined spectrophotometrically using a ε_{280} of 47752 $M^{-1}cm^{-1}$. Typically, 80 mg of purified protein was obtained from a 1 litre fermentation. SDS-polyacrylamide gel electrophoresis was used to confirm enzyme purity.

Crystallisation and data collection

For co-crystallisation of the *T. vivax* IAG-NH with compounds **B** and **D**, a solution of 8 mg/ml (0.2 mM) wild-type *Tv*NH in 20 mM Hepes pH 7.0, 150 mM NaCl, 1 mM $CaCl_2$ was preincubated with 0.5 mM of **B** and **D** respectively for 1 hour at 4°C. Crystals of the complexes were grown using the hanging drop vapour diffusion method at 20°C. For obtaining the structure in complex with compound **B** (named TvNH-InhB hereafter), equal volumes of protein solution and precipitant, containing 25 % (w/v) PEG 3350 in 100 mM Bis-Tris buffer pH 6.5 and 0.2 M $MgCl_2$, were mixed. For obtaining the structure in complex with compound **D** (named TvNH-InhD hereafter), equal volumes of protein solution and precipitant, containing 2.1 M ammonium sulphate in 100 mM Bis-Tris buffer pH 5.5, were mixed. For TvNH-InhB, a crystal was transferred to a cryo-solution containing 35 % (w/v) PEG 3350 in 100 mM Bis-Tris buffer pH 6.5 and 0.2 M $MgCl_2$ and X-ray diffraction data were collected to a resolution of 1.30 Å at 100 K on beamline X13 (EMBL, DESY, Hamburg) using an X-ray wavelength of 0.8051 Å. For TvNH-InhD, a crystal was transferred to a cryo-solution containing 2.1 M ammonium

sulphate, 10 mM Bis-Tris pH 5.5 and 30% glycerol and X-ray diffraction data were collected to a resolution of 1.85 Å at 100 K on beamline X11 (EMBL, DESY, Hamburg) using an X-ray wavelength of 0.8152 Å.

The diffraction data were indexed and integrated using DENZO and scaled using SCALEPACK [23]. Intensities were converted to structure factor amplitudes using TRUNCATE [24]. Table 1 summarizes the data collection and processing statistics.

Structure determination and refinement

Initial phases for TvNH-InhB and TvNH-InhD were obtained by molecular replacement with the program PHASER [25] using a monomer of the wild type *T. vivax* IAG-NH (pdb 1HOZ), with all flexible parts removed. For both datasets, a solution was found, consisting of the expected two subunits forming a relevant dimer in the asymmetric unit. The molecular replacement solution was subsequently used as a starting model for 10 rounds of automated model building (WarpNTrace) in ARP/WARP [26]. The warp model was checked, corrected and further build in COOT [27]. For both structures, the

($2F_o$-F_c, ϕ_c) and (F_o-F_c, ϕ_c) maps in both active sites in the asymmetric unit showed

unambiguous electron density for the inhibitors **B** and **D**. After several cycles of restrained refinement using REFMAC [28] combined with manual corrections, the inhibitors and solvent molecules were included in the model. Structure refinement was considered complete after crystallographic R-factor and free R-factor had converged, and the difference density was without interpretable features. In TvNH-InhB the whole protein backbone (except for the residues of the N-terminal tag) could be traced. For TvNH-InhD, amino acids 250 to 253 in the A subunit and amino acids 253 and 254 in the B subunit were excluded from the model, due to very weak electron density in this region.

The final models were checked with the Molprobity web server [29]. Refinement statistics are summarized in table 1. The structural superpositions were performed using the program LSQMAN [30]. Figures were prepared with Pymol [31].

The coordinates and structure factors have been deposited in the Protein Data bank with accession codes 3EPW (TvNH-InhB) and 3EPX (TvNH-InhD).

Results and discussion

Structure solution and overall fold of the TvNH-inhibitor complexes

Compound **B** (K_I= 4.4 nM) and **D** (K_I= 10.8 nM) are potent inhibitors of the *T. vivax* NH (TvNH)-catalyzed reaction. In order to obtain insight in the molecular mechanism of this inhibition, crystal structures of the wild-type TvNH, co-crystallized with both compounds were solved using molecular replacement with the uncomplexed TvNH as a search model. The structure in complex with inhibitor **B** (TvNH-InhB) was refined to a resolution of 1.3 Å, yielding an R-factor of 16.5 % (R_{free} of 18.3 %). As such this structure represents the highest resolution structure of a nucleoside hydrolase reported up to this moment, providing a detailed view of the interactions. The structure in complex with inhibitor **D** (TvNH-InhD) was refined to a resolution of 1.85 Å, yielding an R-factor of 16.0 % (R_{free} of 21.0 %).

Both complexes contain a homodimer in the asymmetric unit, with in each case the two active sites fully occupied with the catalytic Ca^{2+}-ion and the inhibitor [5]. Each subunit adopts the characteristic nucleoside hydrolase fold [32] consisting of a central mixed β-sheet, with seven parallel and one anti-parallel strand, and several surrounding α-helices [11]. The TvNH-InhB model shows clear and well defined density for the entire protein backbone (except for the N-terminal his-tag) and the inhibitors. Also in the TvNH-InhD model the inhibitors and most of the amino acid residues were clearly defined. However in the latter, residues 250-253 of the A chain and residues 253-254 of the B chain could not be modeled due to lack of electron density. These residues correspond to the tip of active site loop 2 which is often unordered in non-bound or ground-state-inhibitor-bound complexes (see further).

No large scale conformational changes could be observed upon comparing the two newly solved inhibitor complexes with the previously solved structure of TvNH co-crystallized with the inhibitor Immucillin H (compound **A1**, named hereafter TvNH-ImmH). The root mean square deviation (rmsd) upon superposition of the A subunit of TvNH-ImmH (pdb 2FF2) on TvNH-InhB and TvNH-InhD (for the main chain atoms of residues A2-A327 for TvNH-InhB and A2-A249 and A254-A327 for TvNH-InhD) is 0.386 Å and 0.578 Å,

respectively. Furthermore, no significant asymmetry is observed between the two sub-units of the TvNH-InhB and TvNH-InhD complexes, with rmsd values of 0.318 Å and 0.634 Å for superposition of the main chain atoms of the A and B subunits of TvNH-InhB (residues 2-327) and TvNH-InhD (residues 2-249 and 254-327), respectively. Noticeably however, a disulfide bridge, which has never been observed before, is formed between Cys85 of the A chain and Cys248 of the B chain of the TvNH-InhD complex. While Cys85 is part of what is called loop 1 in TvNH, Cys248 is part of the so-called loop 2 (see further). Previously it has been shown that binding of a transition state analogue to TvNH causes structuring of loop1 and loop2, where part of these flexible regions now fold to α-helices (see further and Figure 2) [15]. Moreover, the α-helical regions of loop 1 and loop 2 of different subunits interact through the dimer interface, thus bringing Cys85 (A) and Cys248 (B) in close proximity. If the tip of loop 2 does not fold as in TvNH-InhD (see further) this could cause the observed disulfide bond to be formed in the crystal. The cysteine bond causes a significant distortion of loop 2 belonging to the B chain of TvNH-InhD, so that chain A of TvNH-InhD where loop 2 is unperturbed will be used for further analysis and discussion. The observed disulfide bond is probably an artifact and is likely formed during crystallization, as no such bond could be observed on SDS-PAGE (without reductant) of the purified protein sample.

TvNH-inhibitor interactions: molecular basis for high affinity binding and transition state mimicry

Both active sites of the TvNH-InhB and TvNH-InhD dimers are fully occupied with inhibitor (Figure 2c and 2d). The iminoribitol moieties of compound **B** and **D** adopt an N4'-*exo* conformation (characterized by a pseudorotation phase angle P of about 278°). This ring pucker conformation differs from the preferred pucker of free iminoribitol (C2'-*endo*) [33], but also differs from the pucker of ribose and iminoribitol previously observed in the active site of NH's (C4'-*endo*) [15;20]. The 5'OH torsion angle (O5'-C5'-C4'-N4') is about 75° for compound B and 85° for compound **D**. Probably the more bulky bicyclic quinoline group attached to the iminoribitol in compound **D** causes a slight distorsion of the 5'OH away from the face of the iminoribitol ring. Although in the strict

sense no glycosidic torsion angle χ (relating the ribose to the aglycone) can be defined due the presence of the methylene bridge, the aromatic base analogue is pointing away from the face of the iminoribitol ring, corresponding to an *anti* conformation. This conformation is also observed for ImmH and inosine binding, although a *syn* conformation has been observed for the inhibitor 3-deaza-adenosine [11].

The iminoribitol moieties of compound **B** and **D** bound to the active site of TvNH nearly perfectly superimpose, as do the residues surrounding them (see Figures 2 and 3). A highly conserved, catalytic Ca^{2+}-ion in the active site pocket coordinates the 2'OH and 3'OH of the iminoribitol [5]. The 2'OH further interacts with the side chain of Asp14, while the 3'OH forms H-bonds with the Asn186 and Asp261 side chains. The 5'OH finally is anchored via interactions with the Glu184 and Asn173 (ND2) side chains. The side chain carbonyl of Asn186 is located at about 4 Å from the iminoribitol nitrogen (N4'), probably to far to form a considerable interaction. In contrast to the iminoribitol moiety, the aglycones of compound **B** and **D** differ quite considerably. Compound **B** possesses a purine-like aglycone group and is bound in an active site with fully defined and closed active site loops. The base is stacked between the indole side chains of Trp83 and Trp260 in a parallel fashion (with Trp260 slightly inclined); while further hydrogen bonds are made between Arg252 (located on loop 2) and the exocyclic O6 keto group and between Asp40 and the N3 atom (see figure 2c). In contrast, compound **D** possesses a non-purine quinoline attached to the iminoribitol. The absence of an aglycone with purine properties prevents the tip of loop 2 from closing (see further). Consequently no interaction with Arg252 can be formed. This is probably compensated by an increased binding energy stemming from the more extensive stacking surface of the quinoline with Trp83 and Trp260. A short hydrogen bond is also formed between Asp40 and the endo-cyclic nitrogen of the quinoline ring (N1, superposing with N3 of the 9-deaza-purine ring of compound **B**). This latter nitrogen was predicted from modeling studies to be crucial for strong binding of compound **D** to the active site of TvNH [16]. Recent studies involving derivatives of compound **D** with substituted quinolines showed that this interaction strongly depends on an optimal pK_a (of around 5) of the quinoline N1 nitrogen [19]. On the other hand, the pK_a of Asp40 in a docked enzyme-substrate complex was

recently calculated to be approximately 7.2 to 7.8, making it protonated at physiological pH and allowing a hydrogen bond with N1 [34].

Comparison of the TvNH-InhB and TvNH-InhD structures with the previously determined TvNH-ImmH structures shows some unexpected results, unanticipated by our previous docking experiments (Figure 3). While all active site residues are virtually perfectly superimposable (at least between TvNH-ImmH and TvNH-InhB), the position of the iminoribitol moiety of the inhibitor is shifted in TvNH-InhB and TvNH-InhD compared to the iminoribitol of ImmH in TvNH-ImmH. This shift is most pronounced by a displacement of the N4' by about 1.3-1.4 Å (see Figure3A). The difference in ring pucker (N4'-*exo* in TvNH-InhB and TvNH-InhD compared to C4'-*endo* in TvNH-ImmH, see above) allows this shift to occur, while keeping the hydroxyls at a fixed position. The methylene linker in compound **B** and **D** moreover allows the aglycone to keep its ideal stacking geometry. This shift of the iminoribitol position has as a result that the N4' of compound **B** and **D** superimposes with the C1' of the iminoribitol of ImmH, with the nucleophilic water located now at about 2.9 Å and 3.1 Å from the N4' for TvNH-InhB and TvNH-InhD respectively (compared to a water – C1' distance of 3.45 Å in TvNH-ImmH). It is tempting to speculate that this geometry in fact allows for a better mimicry of the transition state with concomitant gain in binding energy [13;35]. Indeed, quantum chemical calculations have shown that the positive charge in the oxocarbenium-ion transition state of the NH-catalyzed reaction mainly resides on the C1', rather than on the O4' [22]. Considering the pK_a of the iminoribitol N4', it will be partially protonated at physiological pH. Hence, orienting this nitrogen at the C1' position of the substrate will more closely resemble the electrostatic situation of the transition state. A second feature in which compound **B** might more closely mimic the transition state compared to ImmH, is the rather large spacing between the ribose and base analogue, mediated by the methylene bridge. NHs catalyze a dissociative D_NA_N nucleophilic displacement with low bond order to the leaving group and incoming nucleophile in the transition state [36]. The transition state of NHs is hence considerably expanded with a C1'-N9 distance close to 2 Å. The 1.5 Å C1'-C9 bond length of ImmH is a relatively poor geometric mimic of this expanded transition-state. Due to the methylene linker this distance is increased to about 2.5 Å in compound **B** (and **D**). These two extra transition state features of compound **B**

and **D** make them (unexpectedly) closer structural and electronic analogues of the DADMe immucillins than of the immucillins [37]. In DADMe immucillins the position of the nitrogen in the iminoribitol ring is shifted to the N1' position while a methylene bridge is also used to increase the spacing with the aglycone (Figure 1, compound **A2**). However, due to the absence of a 2'OH group (for stability reasons) the DADMe immucillins have been proven to be poor inhibitors for the IAG-NH's while they are excellent inhibitors for PNPs [15;37]. Compounds **B** and **D** hence provide valid IAG-NH specific alternatives.

Further structural corroboration of the leaving group activation mechanism

Nucleoside hydrolases catalyze the hydrolysis of the *N*-glycosidic bond in ribonucleosides using a combination of three catalytic strategies: (1) activation of the nucleophile water via a general base (Asp10 in TvNH), (2) steric and electrostatic stabilization of the oxocarbenium ion, and (3) activation of the leaving nucleic base by protonation [5]. For activation of the leaving group, the IAG-NHs have been shown to utilize a rather peculiar mechanism, rather than classical general acid catalysis. Using experimental approaches and *ab initio* quantum chemical calculations we showed that a parallel aromatic stacking interaction with the enzyme's Trp260 activates the purine leaving group by raising the pK_a of N7, hence facilitating its protonation [21]. The N7 pK_a is further increased via an unusual intramolecular O5'-HC8 hydrogen bond [22]. Together these two effects raise the pK_a of the purine nucleoside sufficiently to promote direct protonation by solvent. The closed TvNH-ImmH structure gave structural support for this mechanism as it showed no enzyme residues contacting the purine N7. Rather, the N7 is connected to bulk solvent via a network of highly ordered water molecules. Therefore this water channel was proposed to function as a proton relay shuttling protons into the active site to the "activated" N7 atom [15].

This proposed mechanism of leaving group activation is also corroborated by the current structures of TvNH in complex with compound **D** and especially compound **B**. The latter structure shows the analogue of the purine base stacked between Trp260 and Trp83. In comparison to the ground state complex of the slow mutant TvNH-Asp10Ala in complex

with inosine, Trp260 has approached to the purine base by about 0.4 Å. In the closed TvNH-InhB structure there are no enzyme residues in direct contact with the N7 – the site of protonation. Rather the same network of ordered water molecules connecting the N7 to bulk solvent is found as in the TvNH-ImmH structure (Figure 3). This water channel is lined by charged and polar residues of loop 2 and could serve as a proton relay to the N7 [15]. Many of these water molecules are also found in the TvNH-InhD structure. However, since there is no nitrogen corresponding to N7 in the quinoline of compound **D**, no water matching the water directly contacting N7 in TvNH-InhB and TvNH-ImmH is present in TvNH-InhD.

Molecular mechanism of two-step loop opening/closing

In unliganded structures of both IAG-NH and IU-NH, two flexible loops are present on either side of the active site pocket. Loop 1 (residues 75-85 in TvNH) is located at the C-terminal end of β-strand β3, while loop 2 (residues 244-258) is situated on the opposite side of the active site at the C-terminal end of a long α-helix, α10 [5]. Binding of ground state inhibitors or the substrate inosine to TvNH causes ordering of loop 1, bringing Trp83 into the active site pocket in a parallel stacking orientation with the nucleic base of the substrate [11;20]. Binding of the transition state analogue ImmH also allows ordering of loop 2, bringing Arg252 into the active site pocket, within interaction distance of the exocyclic O6 keto group of ImmH [15]. The TvNH-ImmH structure showed that upon ordering of loop 2, the N-terminal part of this loop adopts a α-helical secondary structure which forms a kinked C-terminal extension to helix α10, while the C-terminal part forms a coil connecting α10 to α11. Within the dimer interface, the ordered loop 1 of one subunit interacts with the ordered loop 2 of the other subunit, leading to an increase in dimer surface upon loop ordering (see Table 2). A recent study using a loop deletion mutant showed that loop 2 plays an important role in leaving group activation and governs the rate of the slow isomerizations associated with product release [38;39]. It was moreover suggested that the fast release of the nucleic base product only involves opening of the coil region at the tip of loop 2, while the slow release of ribose requires the complete unfolding of the helical region of loop 2. Comparison of “open” and

"closed" TvNH complexes allowed a mechanism of loop closure to be proposed, where loop closure is initiated by formation of a strong hydrogen bond between the 2'OH of the iminoribitol and the Asp14 side chain [15]. The induced catalytic conformation of Asp14 in turn causes a conformational change of the side chain of Trp242, accompanied by ordering of the entire loop 2. While Trp242 was found in one single conformation ("open") in all ground state complexes, it was found in a conformation with the side chain flipped around by 180° ("closed") in the TvNH-ImmH complex. The current TvNH-InhB and TvNH-InhD complexes now allow corroborating and refining this model (see figure 4) and also account for the observed kinetic behavior. Both complexes contain an iminoribitol analogue of the transition state in their active site and consequently a strong hydrogen bond is formed between the 2'-OH and Asp14 (2.5 to 2.6 Å). In accordance with the previously proposed model, the Trp242 side chain is found in the "closed" conformation, and concomitantly the helical region of loop 2 is ordered in both complexes (partially in TvNH-InhD). However, while in the TvNH-InhB complex the tip of the loop is ordered, folding over the active site pocket, this coil region is unordered in TvNH-InhD (Figure 4). This difference can be attributed to the difference in the properties of the aglycone moiety of the inhibitor. In TvNH-InhB, as in TvNH-ImmH, the presence of the exocyclic 6-keto group allows for an interaction with the Arg252 residue on loop2 with concomitant ordering of the coil region of loop 2. In contrast, the quinoline moiety of InhD does not allow for such an interaction, leaving the tip of loop 2 unstructured. As such, these structures indicate that closing of loop 2 is a process consisting of two uncoupled steps. Transition state interactions formed with the ribose moiety of the substrate (or inhibitor) cause the folding of the helical part of loop 2 (via Trp242). Interactions with the nucleic base are subsequently required to cause ordering of the coil region of this active site loop. This model can also fully account for the kinetically observed ordered release of products. After a fast hydrolysis reaction, release of the base requires opening of the tip of loop 2. This process is energetically not very demanding and occurs relatively fast. Release of the ribose moiety on the other hand requires the reversal of the process described for loop closing. This means that after breaking of the bond between the 2'-OH and Asp14, Trp242 needs to flip to the "open" state and the helical part of loop 2 needs to unwind. Since the helix of loop 2 is partially

buried in the dimer interface (figure 2), this is very likely to be a slow process, accounting for the observed very slow release of ribose, by far the rate limiting step on the reaction pathway. Combining the kinetic studies of the *T. vivax* IAG-NH with the different structures of this enzyme in "open", "half-closed" and "closed" states thus allows the deduction of a detailed and quantitative image of loop closing and opening events on the reaction coordinate.

Acknowledgements

The authors acknowledge the EMBL for use of beamlines X13 and X11 (DESY, Hamburg, Germany). This work was funded by a research project from the IWT-Vlaanderen. W.V. has a postdoctoral grant from the Research Foundation Flanders (FWO-Vlaanderen). A.G. had a Ph.D. Grant from the Institute for the promotion of Innovation through Science and Technology in Flanders. M.B. has a Ph.D. Grant from the Research Foundation Flanders.

Abbreviations

IAG-NH, inosine-adenosine-guanosine preferring nucleoside hydrolase; IU-NH, inosine-uridine preferring nucleoside hydrolase; TvNH, nucleoside hydrolase from *Trypanosoma vivax*; TvNH-InhB, TvNH in complex with compound **B** (*(2R,3R,4S)-1-[(4-hydroxy-5H-pyrrolo[3,2-d]pyrimidin-7-yl)methyl]-2-(hydroxymethyl)pyrrolidin-3,4-diol)*; TvNH-InhD, TvNH in complex with compound **D** (*(2R,3R,4S)-2-(hydroxymethyl)-1-(quinolin-8-ylmethyl)pyrrolidin-3,4-diol*); TvNH-ImmH, TvNH in complex with Immucillin H (compound **A**) ; rmsd, root mean square deviation; ImmH, immucillin H or [(1S)-1-(9-deazahypoxanthin-9-yl)-1,4-dideoxy-1,4-imino-d-ribitol]; DADMe-ImmH, 4′-deaza-1′-aza-2′-deoxy-1′-(9-methylene)-immucillinH

References

[1] (a) World Health Report (2002). (b) www.who.int/tdr (2005).

[2] V. Delespaux and H. P. de Koning, Drugs and drug resistance in African trypanosomiasis, *Drug Resist.Updat.* 10 (2007) 30-50 Scopus.

[3] M. H. El Kouni, Potential chemotherapeutic targets in the purine metabolism of parasites, *Pharmacol.Ther.* 99 (2003) 283-309 Scopus.

[4] D. J. Hammond and W. E. Gutteridge, Purine and pyrimidine metabolism in the *Trypanosomatidae, Mol.Biochem.Parasit.* 13 (1984) 243-261 Scopus.

[5] W. Versées and J. Steyaert, Catalysis by nucleoside hydrolases, *Curr.Opin.Struc.Biol.* 13 (2003) 731-738 Scopus.

[6] M. Degano, D. N. Gopaul, G. Scapin, V. L. Schramm, J. C. Sacchettini, Three-Dimensional Structure of the Inosine-Uridine Nucleoside N-Ribohydrolase from *Crithidia fasciculata, Biochemistry* 35 (1996) 5971-5981 Scopus.

[7] W. Shi, V. L. Schramm, S. C. Almo, Nucleoside Hydrolase from *Leishmania major*. Cloning, expression, catalytic properties, transition state inhibitors, and the 2.5-Å crystal structure, *J.Biol.Chem.* 274 (1999) 21114-21120 Scopus.

[8] R. Mitterbauer, T. Karl, G. Adam, *Saccharomyces cerevisiae* URH1 (encoding uridine-cytidine N-ribohydrolase): functional complementation by a nucleoside hydrolase from a protozoan parasite and by a mammalian uridine phosphorylase, *Appl.Environ.Microb.* 68 (2002) 1336-1343 Scopus.

[9] B. Giabbai and M. Degano, Crystal structure to 1.7 Å of the *Escherichia coli* pyrimidine nucleoside hydrolase YeiK, a novel candidate for cancer gene therapy, *Structure* 12 (2004) 739-749 Scopus.

[10] D. W. Parkin, Purine-specific Nucleoside N-ribohydrolase from *Trypanosoma brucei brucei*. Purification, specificity, and kinetic mechanism, *J.Biol.Chem.* 271 (1996) 21713-21719 Scopus.

[11] W. Versées, K. Decanniere, R. Pellé, J. Depoorter, E. Brosens, D. W. Parkin, J. Steyaert, Structure and function of a novel purine specific nucleoside hydrolase from *Trypanosoma vivax, J.Mol.Biol.* 307 (2001) 1363-1379 Scopus.

[12] B. Estupiñán and V. L. Schramm, Guanosine-Inosine-preferring Nucleoside N-Glycohydrolase from *Crithidia fasciculata, J.Biol.Chem.* 269 (1994) 23068-23073 Scopus.

[13] V. L. Schramm, Enzymatic transition states: thermodynamics, dynamics and analogue design, *Arch.Biochem.Biophys.* 433 (2005) 13-26.

[14] G.A. Kicska, L. Long, H. Hörig, C. Fairchild, P. C. Tyler, R. H. Furneaux, V. L. Schramm, H. L. Kaufman, Immucillin H, a powerfull transition-state analog inhibitor of purine nucleoside phosphorylase, selectively inhibits human T lymphocytes, *Proc.Natl.Acad.Sci.USA* 98 (2001) 4593-4598.

[15] W. Versées, J. Barlow, J. Steyaert, Transition-state complex of the purine-specific nucleoside hydrolase of *T. vivax*: enzyme conformational changes and implications for catalysis, *J.Mol.Biol.* 359 (2006) 331-346 Scopus.

[16] A. Goeminne, M. Berg, M. McNaughton, G. Bal, G. Surpateanu, P. Van Der Veken, S. De Prol, W. Versées, J. Steyaert, A. Haemers, K. Augustyns, N-Arylmethyl substituted iminoribitol derivatives as inhibitors of a purine specific nucleoside hydrolase, *Bioorg Med Chem* 16 (2008) 6752-6763 Scopus.

[17] A. Goeminne, M. McNaughton, G. Bal, G. Surpateanu, P. Van Der Veken, S. De Prol, W. Versées, J. Steyaert, S. Apers, A. Haemers, K. Augustyns, 1,2,3-Triazolylalkylribitol derivatives as nucleoside hydrolase inhibitors, *Bioorg Med Chem Lett.* 17 (2007) 2523-2526 Scopus.

[18] A. Goeminne, M. McNaughton, G. Bal, G. Surpateanu, P. Van Der Veken, S. De Prol, W. Versées, J. Steyaert, A. Haemers, K. Augustyns, Synthesis and biochemical evaluation of guanidino-alkyl-ribitol derivatives as nucleoside hydrolase inhibitors, *Eur.J.Med Chem* 43 (2008) 315-326.

[19] M. Berg, G. Bal, A. Goeminne, P. Van Der Veken, W. Versées, J. Steyaert, A. Haemers, K. Augustyns, Synthesis of bicyclic N-arylmethyl substituted iminoribitol derivatives as selective nucleoside hydrolase inhibitors, *ChemMedChem* (2009) **In Press.**

[20] W. Versées, K. Decanniere, E. Van Holsbeke, N. Devroede, J. Steyaert, Enzyme-substrate interactions in the purine-specific nucleoside hydrolase from *Trypanosoma vivax., J.Biol.Chem.* 277 (2002) 15938-15946.

[21] W. Versées, S. Loverix, A. Vandemeulebroucke, P. Geerlings, J. Steyaert, Leaving group activation by aromatic stacking: an alternative to general acid catalysis, *J.Mol.Biol.* 338 (2004) 1-6 Scopus.

[22] S. Loverix, P. Geerlings, M. McNaughton, K. Augustyns, A. Vandemeulebroucke, J. Steyaert, W. Versées, Substrate-assisted leaving group activation in enzyme-catalyzed N-glycosidic bond cleavage, *J.Biol.Chem.* 280 (2005) 14799-14802.

[23] Z. Otwinowski and W. Minor, in *Methods in Enzymology Volume 276: Macromolecular Crystallography Part A*, C. W. Carter and R. M. Sweet, Eds. (Academic Press, New York, 1997).

[24] S. French and K. Wilson, On the treatment of negative intensity observations, *Acta Crystallogr.* A34 (1978) 517-525.

[25] A. J. McCoy, R. W. Grosse-Kunstleve, L. C. Storoni, R. J. Read, Likelihood-enhanced fast translation functions, *Acta Crystallogr.* D61 (2005) 458-464 Scopus.

[26] A. Perrakis, R. Morris, V. S. Lamzin, Automated protein building combined with iterative structure refinement, *Nat.Struct.Biol.* 6 (1999) 458-463.

[27] P. Emsley and K. Cowtan, Coot: model-building tools for molecular graphics, *Acta Crystallogr.* D60 (2004) 2126-2132 Scopus.

[28] G. N. Murshudov, A. A. Vagin, E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr.D* 53 (1997) 240-255 Scopus.

[29] S. C. Lovell, I. W. Davis, W. B. Arendall, P. I. de Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, D. C. Richardson, Structure validation by C alpha geometry: phi,psi and C beta deviation, *Proteins* 50 (2003) 437-450.

[30] Collaborative Computational Project Number 4, The CCP4 suite: programs for protein crystallography, *Acta Crystallogr.* D50 (1994) 760-763 Scopus.

[31] W. L. DeLano, The PyMOL Molecular Graphics System (DeLano Scientific, San Carlos, CA, USA, 2002).

[32] A. G. Murzin, S. E. Brenner, T. Hubbard, C. Chothia, SCOP: a structural classification of proteins database for the investigation of sequences and structures, *J.Mol.Biol.* 247 (1995) 536-540 Scopus.

[33] A. Fedorov, W. Shi, G.A. Kicska, E. Fedorov, P. C. Tyler, R. H. Furneaux, J. C. Hanson, G. J. Gainsford, J. Z. Larese, V. L. Schramm, S. C. Almo, Transition state structure of purine nucleoside phosphorylase and principles of atomic motion in enzymatic catalysis, *Biochemistry* 40 (2001) 853-860 Scopus.

[34] D. Mazumder-Shivakumar and T. C. Bruice, Computational study of IAG-nucleoside hydrolase: determination of the preferred ground state conformation and the role of active site residues, *Biochemistry* 44 (2005) 7805-7817 Scopus.

[35] R. Wolfenden and M. J. Snider, The depth of chemical time and the power of enzymes as catalysts, *Acc.Chem.Res.* 34 (2001) 938-945 Scopus.

[36] V. L. Schramm, Enzymatic transition state poise and transition state analogues, *Acc.Chem.Res.* 36 (2003) 588-596 Scopus.

[37] A. Lewandowicz, W. Shi, G. B. Evans, P. C. Tyler, R. H. Furneaux, L. A. Basso, D. S. Santos, S. C. Almo, V. L. Schramm, Over-the-barrier transition state analogues and crystal structure with *Mycobacterium tuberculosis* purine nucleoside phosphorylase, *Biochemistry* **42** (2003) 6057-6066 Scopus.

[38] A. Vandemeulebroucke, S. De Vos, E. Van Holsbeke, J. Steyaert, W. Versées, A flexible loop as a functional element in the catalytic mechanism of nucleoside hydrolase from *Trypanosoma vivax, J.Biol.Chem.* (2008).

[39] A. Vandemeulebroucke, W. Versées, J. Steyaert, J. N. Barlow, Multiple transients in the pre-steady-state of nucleoside hydrolase reveal complex substrate binding, product base release, and two apparent rates of chemistry, *Biochemistry* 45 (2006) 9307-9318 Scopus.

Figures

Figure 1: Structures of Immucillin H (**A1**), DADMe-ImmH (**A2**) and the inhibitors **B**, **C** and **D**. The crystal structures of TvNH co-crystallized with **B** and **D** are described in this paper. Inhibition constants for TvNH are given [16]. In (**A1**) the atom numbering for Immucillin H is given.

Figure 2: Overall structure and close-up view of the active sites of TvNH-InhB and TvNH-InhD. (**A**) and (**B**): Comparison of the quaternary structure and subunit interface of TvNH-InhB (A) and TvNH-InhD (B). Loop 1 and 2, which adopt in part a α-helical conformation in the closed active site, are shown in blue and green respectively. The inhibitors bound to the active site are shown in stick representation. Also the disulfide bond between Cys85A and Cys248B in TvNH-InhD is shown. (**C**) and (**D**): Active sites of TvNH-InhB (**C**) and Tvnh-InhD (**D**). The bound inhibitors are shown in yellow. The electron density, contoured at 5σ, of an fo-fc simulated annealed omit map calculated without the inhibitor is shown as a blue mesh. The Ca^{2+}-ion and water molecules are represented as grey and red spheres respectively. Water molecules in the vicinity of the N7 and N1 nitrogen atoms of the inhibitor in TvNH-InhB, and the corresponding water molecules (if present) in TvNH-InhD are shown. The residue Arg252 is provided by the tip of loop 2 and is only observable in the TvNH-InhB structure.

Figure 3: Conformation of bound inhibitors. A "top view" (A) and "lateral view" (B) are shown of a superposition of the active sites of TvNH-ImmH, TvNH-InhB and TvNH-InhD. The active site residues of the three structures superimpose nearly perfectly (except for loop 2 residues) and their carbon atoms are colored in three different shades of grey. The carbon atoms of the inhibitors are colored yellow (ImmH), cyan (compound B) and green (compound D). The nucleophilic water molecules and the first water molecules of the water channel leaving from the N7 atom are also shown using the same color code as the inhibitors. In the top view the rotation of the iminoribitol ring of TvNH-InhB and TvNH-InhD compared to the iminoribitol of TvNH-ImmH is apparent. This conformation places the N4' of compound B and D on the position of the C1' of ImmH. The lateral view shows the increased spacing between the iminoribitol and aglycone moiety caused by the methylene spacer in TvNH-InhB and TvNH-InhD.

Figure 4: Mechanism of loop 2 opening / closing. The backbone of TvNH-InhB is shown with residues of the A chain in light grey and the residues of the B-chain in dark grey. Inhibitors, residues of loop2 and residues important in the loop opening/closing mechanism of TvNH(D10A) in complex with inosine (PDB 1KIC) and of TvNH-InhD are superimposed on this structure and colored cyan for TvNH-InhB, green for TvNH-InhD and yellow for TvNH(D10A)-inosine. In TvNH-InhB and TvNH-InhD a strong hydrogen bond with the 2'OH of the iminoribitol ring pulls Asp14 into one catalytic conformation. This in turn causes a rearrangement of Trp242 and Tyr257 with concomitant ordering of the helical region of loop2 (see text for details). However a hydrogen bond between the exocyclic O6-keto group of the purine analogue and Arg252, as in TvNH-InhB, is further needed to order the tip of loop 2.

Table 1: Data collection and refinement parameters

	TvNH-Inh**B**	TvNH-Inh**D**
	Diffraction data	
Space group	$P2_1$	$P2_1$
a (Å)	51.4	54.4
b (Å)	73.7	74.2
c (Å)	81.1	73.6
β (deg)	75.2	98.4
Resolution range (Å) [a]	50.0 – 1.30 (1.35 – 1.30)	50.0 – 1.85 (1.92 – 1.85)
R_{sym} (%) [a]	8.0 (37.0)	10.7 (51.1)
I / σI [a]	26.2 (2.7)	15.4 (3.1)
Completeness (%) [a]	96.6 (94.0)	99.9 (100.0)
Multiplicity	6.6 (2.9)	5.8 (5.8)
	Structure refinement	
R_{cryst} (%)	16.51	16.05
R_{free} (%)	18.27	20.96
rmsd for bond lengths (Å)	0.009	0.009
rmsd for bond angles (deg)	1.294	1.184
Ramachandran plot (% favoured, allowed, outliers)	98.0, 100, 0	98.3, 100, 0
No. atoms per a.u.	5950	5696

[a] Values in parentheses are for the highest resolution shell

$R_{sym} = \Sigma \mid I_i (hkl) - \langle I_i (hkl) \rangle \mid / \Sigma I_i (hkl)$, where $I_i (hkl)$ are the intensities of multiple measurements and $\langle I_i (hkl) \rangle$ is the average of the measured intensities for the ith reflection.

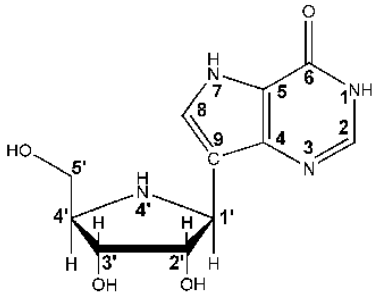
$R_{cryst} = \Sigma_h |F_{obs,\,h} - F_{calc,\,h}| / \Sigma_h F_{obs,\,h}$, where F_{obs} and F_{calc} are observed and calculated structure amplitudes, respectively

$R_{free} = R_{cryst}$ calculated for the test set of reflections not used in refinement

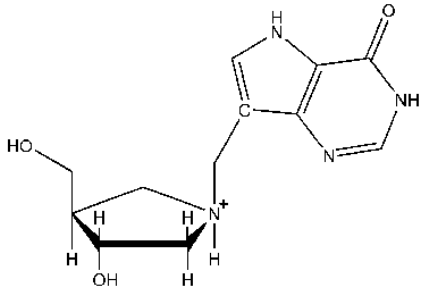
Table 2: Accessible surface area buried in the dimer interface of *T. vivax* nucleoside hydrolase structures

Structure	PDB code	Loop 1	Loop 2	**ASA in dimer interface (Å^2)**
TvNH-native[a]	1HOZ	open	open	876
TvNH-3-deaza-adenosine[a]	1HP0	closed	open	1043
TvNH(D10A)-inosine[a]	1KIC	closed	open	920
TvNH-ImmH (soaked)[a]	2FF1	closed	open	867
TvNH-ImmH (co-crystallized)[a]	2FF2	closed	closed	1232
TvNH-InhB		closed	closed	1297
TvNH-InhD		closed	half-closed	1288

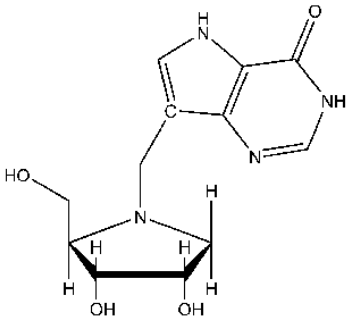
[a] Data from [15]



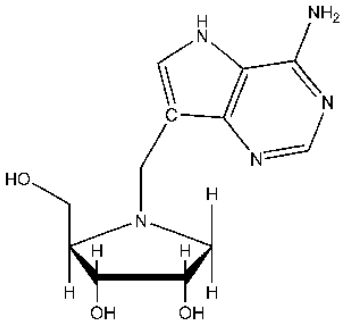
A1 (Immucillin H)
K_I = 6.2 nM



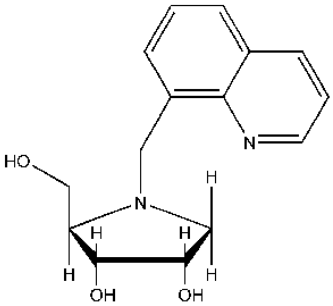
A2 (DADMe-ImmH)



B
K_I = 4.4 nM



C
K_I = 4.1 nM



D
K_I = 10.8 nM