Published in final edited form as:

J Mol Biol. 2010 March 26; 397(2): 481–494. doi:10.1016/j.jmb.2010.01.051.

Crystal structures of trypanosomal histidyl-tRNA synthetase illuminate differences between eukaryotic and prokaryotic homologs

Ethan A Merritt^{a,b,*}, Tracy L Arakaki^{a,b}, J Robert Gillespie^{a,c}, Eric T Larson^{a,b}, Angela Kelley^{a,c}, Natascha Mueller^{a,c}, Alberto J Napuli^{a,c}, Jessica Kim^{a,b}, Li Zhang^{a,b}, Christophe L M J Verlinde^{a,b}, Erkang Fan^{a,b}, Frank Zucker^{a,b}, Frederick S Buckner^{a,c}, Wesley C Van Voorhis^{a,c}, and Wim G J Hol^{a,b}

^a Medical Structural Genomics of Pathogenic Protozoa <http://www.msgpp.org>

^b Department of Biochemistry, University of Washington, Seattle, WA 98195 USA

^c Department of Medicine, University of Washington, Seattle, WA 98195 USA

Abstract

Crystal structures of histidyl-tRNA synthetase from the eukaryotic parasites *Trypanosoma brucei* and *Trypanosoma cruzi* provide a first structural view of a eukaryotic form of this enzyme, and reveal differences from bacterial homologs. Histidyl-tRNA synthetases in general contain an extra domain inserted between conserved motifs 2 and 3 of the Class II aminoacyl-tRNA synthetase catalytic core. The current structures show that the three dimensional topology of this domain is very different in bacterial and archaeal/eukaryotic forms of the enzyme. Comparison of apo and histidine-bound trypanosomal structures indicates substantial active site rearrangement upon histidine binding, but relatively little subsequent rearrangement after reaction of histidine with ATP to form the enzyme's first reaction product, histidyladenylate. The specific residues involved in forming the binding pocket for the adenine moiety differ substantially both from the previously characterized binding site in bacterial structures and from the homologous residues in human histidyl-tRNA synthetases. The essentiality of the single histidyl-tRNA synthetase gene in *T. brucei* is shown by a severe depression of parasite growth rate that results from even partial suppression of expression by RNA interference.

Keywords

aminoacyl-tRNA synthetase; protozoa; drug target; tropical disease; RNAi

Introduction

Aminoacyl-tRNA synthetases (aaRS) mediate the key biological function of translating an RNA message into a corresponding protein sequence. Each specific aminoacyl-tRNA synthetase is responsible for attaching a particular amino acid to its proper cognate tRNA. This

*Corresponding author: Phone: 206-543-1421 Fax: 206-685-7002, merritt@u.washington.edu.

Accession Numbers

Coordinates and structure factors have been deposited in the Protein Data Bank with accession numbers 3HRI, 3HRJ, and 3HRK.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

proceeds in two separate reactions, both catalyzed at the same active site. First, the amino acid is adenylated by reaction with ATP to yield aminoacyladenylate + pyrophosphate. In a second reaction, the adenylate of the activated amino acid is replaced by the terminal adenosine residue of the tRNA acceptor stem. In general this results in a charged tRNA that uniquely associates the anticodon sequence it carries on its anticodon loop with the amino acid it carries on its acceptor stem. A few exceptional cases require subsequent additional chemical modification of the attached amino acid¹. As expected for a molecule with such a key role, the aminoacyl-tRNA synthetases are found in all living organisms. They group into two classes defined by sequence phylogeny and by the consequent structural homology of their core domains. The Class I catalytic domain contains a Rossmann fold and two highly conserved sequence motifs. The Class II catalytic domain contains a core antiparallel β -sheet surrounded by α -helices, and is uniquely identified by three conserved sequence motifs. Further subdivision of Class II into subclasses a, b, and c is based on the nature of the anticodon binding domain. The aaRS for each amino acid except lysine belongs consistently to either Class I or Class II across all forms of life. Despite this broad conservation of core structural and functional features, many individual aaRS have evolved idiosyncratic structural features, additional inserted domains, and even whole new biological functions to supplement their already essential role. This combination of essential biological function and cross-species variation suggests that individual aaRS may be useful targets for drugs targeting infectious disease [review: Hurdle *et al.*2].

We chose trypanosomal Histidyl-tRNA synthetase (HisRS) for study as part of a broader effort by the Medical Structural Genomics of Pathogenic Protozoa collaborative to investigate aminoacyl-tRNA synthetases as possible targets for drug discovery³. Among the target pathogens are *Trypanosoma brucei*, the causative agent of African sleeping sickness, and *Trypanosoma cruzi*, the causative agent of Chagas disease. The chronic forms of these diseases are fatal, and existing therapeutics are unsatisfactory due to poor effectiveness and to toxicity. Many higher eukaryotes, including mammals, have separate cytosolic and mitochondrial forms for each aminoacyl-tRNA synthetase. In contrast, the genomes of many lower eukaryotes, including trypanosomatids, contain only aaRS sequences from the cytosolic grouping. As a further complication, some trypanosomatids have evolved separate mitochondrial forms for specific aaRS, apparently through gene duplication of the cytosolic sequence^{4,5}. However, HisRS is not one of these; the *T. brucei* and *T. cruzi* genomes contain only a single HisRS, of the typical eukaryotic cytosolic type, that is used for both cytosolic and mitochondrial translation.

HisRS is a Class IIa aminoacyl-tRNA synthetase containing three conserved domains. The first of these is the catalytic domain, whose core fold is shared by all Class II aaRS. The second is the C-terminal domain, whose α/β fold is shared by the three Class IIa enzymes HisRS, ThrRS, and ProRS, and by the eukaryotic/archaeal GlyRS⁶. This C-terminal domain holds the binding site for the tRNA anticodon loop, although the anticodon itself is not a primary specificity determinant for recognizing tRNA^{His}. Cognate tRNA recognition is instead believed to be mediated in part by residues of the Class II motif 2 loop in the catalytic domain⁷. No structure has been reported for a HisRS:tRNA^{His} complex, although tRNA binding can be modeled approximately based on structures of bacterial threonyl- and prolyl-tRNA synthetase:tRNA complexes^{8,9}. The third conserved domain is unique to HisRS, and consists of a 50–75 residue α -helical insertion between conserved motifs 2 and 3 of the catalytic domain. Prokaryotic sequences for this insertion domain contain several highly conserved residues on the surface near the active site that have been proposed to interact with the 3' strand backbone of the tRNA acceptor stem¹⁰. Guth *et al.*¹¹ have found that the insertion domain is also important to the histidine adenylation reaction in the absence of tRNA, as fluorescent resonance transfer experiments on the *Escherichia coli* HisRS suggest that the forward rate of adenylation is limited by concomitant rigid body rotation of the insertion domain. Furthermore, this

conformational change introduces an asymmetry in the reaction rate of the two monomers making up the HisRS dimer. No equivalent characterization of the insertion domain has been reported for a eukaryotic HisRS.

In addition to these conserved features, HisRS from many eukaryotes including humans contains one or more copies of a short, N-terminal WHEP domain⁶. WHEP domains are named for their presence in several otherwise non-homologous aaRS (tryptophanyl, histidyl, glutamyl, prolyl), where they mediate various non-canonical biological activities by promoting association of multiple aaRS molecules into a larger assembly^{12,13}. The HisRS WHEP domain has also been implicated as a major epitope in human autoimmune disease¹⁴. NMR studies have shown that the single WHEP domain found in human HisRS forms an antiparallel 2-helix bundle [unpublished; PDB accession 1x59]. This same structure is observed for the individual modules within repeated WHEP domain sequences from other aaRS¹⁵. The N-terminal sequence of trypanosomal HisRS is of equivalent length to the human WHEP domain but has low sequence similarity. It is not known whether the N-terminus of trypanosomal HisRS mediates a secondary biological function.

There are two mammalian HisRS genes, HARS and HARS2 (also known as HARSL), which arose by inverted gene duplication^{16,17}. The human HisRS gene products are 79% sequence identical to each other outside of the WHEP domain. Although the trypanosomal and human HisRS sequences belong to the same, eukaryotic branch of the HisRS phylogenetic tree¹⁸, the maximum pairwise sequence identity between either of the human genes and either trypanosomal homolog discussed here is less than 30%. The pairwise sequence identity between these eukaryotic sequences and previously studied bacterial HisRS homologs is also less than 30%.

Crystal structures of several bacterial HisRS have been published, offering insight into the specific residues and binding interactions involved in binding the substrates histidine and ATP^{10,19,20,9}, and the intermediate product histidyladenylate (HAM)^{10,21}. The apo structure of one archaeal HisRS, that of *Thermococcus acidophilum*, has also been determined crystallographically [unpublished; PDB accession 1wu7]. No structure of a eukaryotic HisRS has been reported to date.

We report here a crystal structure of the apo HisRS from *T. brucei* at 2.85 Å resolution, and crystal structures of the HisRS from *T. cruzi* in complex with the amino acid substrate, histidine, and with the product of the activation reaction, histidyladenylate, at 1.8 Å and 3.05 Å resolution respectively. These first structures of a eukaryotic HisRS show that the insertion domain in eukaryotic HisRS adopts a different three-dimensional structure than the bacterial insertion domain. They also reveal significant differences at the binding site of histidyladenylate when compared to previously reported crystal structures of bacterial homologs^{19,10}. The residues contributing to these differences at the active site also distinguish the trypanosomal HisRS from both of the human HisRS homologs, suggesting that it may be possible to design selective inhibitors for the trypanosomal enzymes.

Results and Discussion

Comparison of the overall structure to HisRS from other organisms

The three-dimensional structures of the *T. cruzi* and *T. brucei* HisRS homologs are very similar, reflecting the fact that they are 84% identical in sequence (Figures 1, 2). The sequence identity between either trypanosomal HisRS and previously studied bacterial homologs is on the order of 25% in the catalytic domain and lower elsewhere. The entire trypanosomal HisRS protein is visible in the crystal structures with the exceptions of the N-terminus, which was truncated in the expressed proteins to begin at residue 45 (*T. cruzi* numbering), and residues 456–467

near the C-terminus of the anticodon binding domain. Residues 158–164 of the conserved Class II aaRS motif 2 are well-ordered in the apo structure and in the histidyladenylate complex, but are not visible in the higher resolution histidine complex. Conversely, residues 220–241, comprising the first half of the insertion domain, are poorly ordered in the apo structure but are well-defined in the ligand-bound structures. The overall folds of the catalytic and anticodon binding domains are as previously seen in the well-studied bacterial homologs, although there are differences among the residues forming the active site that will be discussed separately below. The catalytic and anticodon binding domains are connected by an extended stretch of residues, 376–383, following the Class II motif 3.

All HisRS sequences are characterized by an additional domain inserted between conserved motifs 2 and 3 of the Class II catalytic core. The insertion tends to be larger in eukaryotes than in prokaryotes. Sequence conservation across eukaryotic, archaeal, and prokaryotic insertions is low, but within eukaryotes it is higher. The trypanosomal and human HisRS sequences are roughly 30% identical in this region (Figure 2). The insertion domain has been proposed to play a role in tRNA recognition^{21,10}, but no specific model has been developed. The present work reveals that the eukaryotic and prokaryotic insertion domains have very different three dimensional structures (Figure 3). After superposition of the structurally similar eukaryotic and prokaryotic core domains, there is no evident correspondence in the position of any of the secondary structural elements belonging to their respective insertion domains. Given the divergence of both sequence and structure in this region, it seems unlikely that a single model for interaction between the insertion domain and tRNA would apply to both prokaryotic and eukaryotic HisRS.

The N-terminus of some eukaryotic HisRS folds into a 2-helix bundle (WHEP domain)^{15,22}. Truncation of the 60 N-terminal residues from human HisRS results in an inactive protein as measured by production of His-tRNA^{His}²³, but it is not clear whether this is due to loss of the 45-residue WHEP domain or to the absence of the following 15 residues belonging to the start of the catalytic domain. Both of the trypanosomal proteins in the present work were N-terminal truncations equivalent to removal of only the WHEP domain. The 44-residue N-terminal truncation of the *T. cruzi* HisRS studied here retains the ability to activate histidine, but was not assayed for its ability to charge tRNA^{His}. Although it is obviously not present in the crystal structures, we were interested to see if there is evidence for a WHEP-equivalent domain at the N-terminus of the trypanosomal HisRS. There is less than 10% sequence identity between the trypanosomal and human sequences in the first 46 residues, corresponding to the two helices of the human WHEP domain (Figure 2). To evaluate the likelihood that the trypanosomal HisRS N-terminus also forms a helical bundle, we examined the trypanosomal sequences for helix propensity. Two different secondary structure predictors indicate a strong propensity for helix formation from residues 10–21 (Jpred24) or 5–28 (Porter25) but do not predict a second helical region that would pair with the first to form a WHEP-like 2-helix bundle. Equivalent predictions were obtained for the N-terminal sequence of the HisRS from *Leishmania major*, a related pathogenic protozoan.

Conformational changes associated with substrate binding

The base of the active site in all Class II aaRS is formed by the central strands of the conserved β -sheet. These are $\beta 6$, $\beta 10$, and $\beta 11$ in the present structures (Figure 4). The binding surface for the amino acid and ATP substrates is largely contributed by residues from the conserved Class II sequence motifs 2 and 3. In the case of HisRS, histidine recognition and binding is mediated by two additional motifs, His A and His B, that are not shared with other aaRS²¹. While other Class II aaRS bind ATP in coordination with a divalent metal ion that helps to position the β - and γ - phosphates and acts as an electrophile for the adenylation reaction, HisRS

is atypical in binding ATP and catalyzing adenylation without the involvement of a divalent cation.

HisRS is also atypical in the extent to which substrate binding proceeds via an induced-fit mechanism. The catalytic domain undergoes a substantial conformational change upon histidine binding (Figure 4). Helix $\alpha 3$ shifts up to 10 Å towards the top of the active site. Helices $\alpha 8$, $\alpha 9$, $\alpha 10$, and $\alpha 11$ of the insertion domain shift in concert by approximately 3 Å towards the active site. The HisRS-specific histidine recognition sequence (His A motif), residues 314–318 in the *T. cruzi* enzyme, shifts by more than 3 Å to interact directly with the histidine. This has the effect of repositioning Arg 314 so that it is placed to interact with the α -phosphate of ATP, and probably contributes to the increased affinity for ATP after histidine is bound (Figure 5). As discussed later, we attribute positioning of the β - and γ -phosphates to coordination by three additional arginine residues, one of which is not conserved in other HisRS sequences. Somewhat unexpectedly, the portion of the conserved Class II motif 2 that interacts with the adenosine ring system, residues 158–165 in the *T. cruzi* sequence, is well ordered in the apo structure but not in the histidine-bound structure. We have observed this same lack of order in a different crystal form of the same complex (not detailed here), making it more likely that this is a true consequence of binding histidine rather than an artifact such as crystal lattice packing.

There is further conformational change between the histidine substrate-bound enzyme and the histidyladenylate product-bound enzyme. Helix $\alpha 3$ shifts slightly farther toward the active site. In addition, motif 2 residues 158–165 become fully ordered and interact directly with the ring system of the adenylyl moiety. This constitutes a shift of approximately 7 Å in the position of C^α atoms of the residues at the tip of this loop when compared to their position in the apo crystal structure. As discussed below, the specific residues making up the overall binding pocket for the adenine ring system are significantly different in the current structures than their equivalents in the bacterial or archaeal homologs that were previously studied. There is little if any shift of the insertion domain between the trypanosomal histidine- and histidyladenylate-bound structures. Of course this does not preclude movement of the insertion domain during the course of the reaction; domain flexibility may be necessary to accommodate either the recognition and binding of ATP or the subsequent release of the β - and γ - phosphates by after catalytic displacement by the previously bound histidine.

Order and asymmetry of substrate binding

The structural details of tRNA binding to HisRS have not been directly observed, but by analogy to the binding mode seen for other Class IIa aaRS, the cognate tRNA is believed to interact with only one monomer of the HisRS dimer^{8,9}. The geometry of binding is such that two tRNA molecules can bind simultaneously, one to each monomer of the dimer. The two monomers do not act independently, however. Francklyn and coworkers have extensively characterized the reaction ordering and kinetics of the *E. coli* HisRS dimer. The *E. coli* enzyme interacts weakly, if at all, with ATP alone¹¹. In the presence of tRNA, activation of histidine in the second monomer is dependent on prior completion of aminoacyl transfer to tRNA at the active site of the first monomer²⁶. Even in the absence of tRNA, the rate of formation of histidyladenylate is different at the two active sites¹¹. If the HisRS dimer is repurified by size exclusion chromatography after the reaction, the recovered complex has a molar ratio of one bound histidyladenylate per dimer²⁷. This led to a model in which the activation of histidine by ATP at one site induces a rate-limiting conformational change, possibly rigid-body rotation of the insertion domain, that interferes with ATP binding and/or reactivity at the second site¹¹.

The conformational difference between the apo and histidine-bound trypanosomal HisRS crystal structures is substantial, and provides a structural basis for sequential binding of the substrates, histidine first and ATP second. Thermal melt data for three trypanosomal HisRS

homologs are in agreement with this being the order of binding in solution (Figure 5). In each case addition of AMP, ADP, or ATP alone does not stabilize the HisRS dimer in solution, whereas addition of histidine alone shifts the T_M of the protein by 4 to 6° C. In the presence of histidine, addition of AMP or ADP causes a small but significant increase in stability. Addition of both histidine and ATP leads to substantial further stabilization of the protein corresponding to a >7° C increase in T_M relative to the apo protein. As the enzymes are active, this probably represents the stability of the complex of the protein with the reaction product, histidyladenylate. This ordering of ΔT_M is observed consistently for the *T. cruzi* and *T. brucei* protein constructs used for the current structural analysis, and also for a full length (no N-terminal truncation) HisRS from the related trypanosomatid *Leishmania naiffi*.

However, we found only weak evidence for structural asymmetry of the two active sites in trypanosomal HisRS before or after the activation of histidine by ATP. The monomers making up the crystallographically observed apo and His-bound dimers are structurally the same. Indeed, crystals of the *T. cruzi* HisRS:His complex contain only a single monomer in the asymmetric unit; the two monomers in the biological dimer are related by perfect crystallographic 2-fold symmetry. The two crystallographically independent monomers making up the dimer of the *T. cruzi* HisRS:HAM complex are also identical to within the precision of model; after superposition of the two monomers, the RMSD for all 414 paired C α atoms in the model is 0.17 Å. This does not preclude the existence of slight differences, but their effect on the atomic geometry of the active site in the two monomers is not evident in the crystal structure.

The dimers seen in the previously characterized complexes of *T. thermophilus* with histidine (PDB accession 1adj) and with HAM (PDB accession 1ady) were less symmetric. In these structures, the two monomers within the dimer superimpose with an rmsd of 0.50 Å for all C α atoms not in the insertion domain, but after this superposition the insertion domains of the paired monomers differ by a rotation that leads to a displacement of the distal helices by up to 4 Å. The insertion domain of the single archaeal homolog studied crystallographically, *T. acidophilum*, is structurally much more like the current eukaryotic structures than to those of the bacterial homologs. Only the apo form of the archaeal structure has been reported (PDB 1wu7), but here also the insertion domains of the two monomers within the dimer differ by a rotation that leads to a displacement of roughly 4 Å at the distal end.

Notwithstanding the symmetric geometry of the two monomers in the *T. cruzi* HisRS:HAM complex, there is nevertheless some evidence for a differential binding of histidyladenylate at the two active sites. Electron density corresponding to a complete molecule of histidyladenylate is present in both sites (Figure 6). However, the electron density in monomer A is greater than in monomer B. Furthermore, while in monomer A the density for the adenine moiety is better resolved than the density for the histidine, in monomer B the density for the histidine moiety is stronger than for the AMP moiety. This may indicate that the reaction went to completion in monomer A, while in some fraction of the monomer B copies present in the crystal the histidine remains unactivated.

The trypanosomal HisRS active site

As noted above, the ATP binding site in HisRS only becomes fully formed after a conformational rearrangement induced by prior binding of histidine. In the *E. coli* HisRS complex with histidinol and ATP¹⁹, the three ATP phosphates are positioned by interaction with the guanidinium groups of five arginine residues. The α -phosphate is positioned by association with bacterial residues Arg 113 and Arg 259. The homologous *T. cruzi* residues are Arg 156 and Arg 314, and the analogous interactions with the α -phosphate are seen also in the present complex with histidyladenylate. Arg 314 is believed to act also as the electrophile during catalysis, substituting for the divalent metal cation present in other aaRS. In the absence

of an ATP-bound structure of the eukaryotic enzyme, we used the *E. coli* complex to model interactions of the β - and γ -phosphates. The β -phosphate interacts with bacterial residue Arg 113. The γ -phosphate is positioned by bacterial residues Arg 121 and Arg 311. We infer from sequence and structure alignment that the role of bacterial Arg 121 is played by the homologous trypanosomal Arg 164. However, there is no arginine or similar residue in the trypanosomal sequence alignment equivalent to the bacterial Arg 311. The spatial proximity of the sidechain of trypanosomal residue Arg 334 in the current structures suggests that the role of positioning the ATP γ -phosphate prior to catalysis is played instead by this residue from a non-conserved region of the sequence preceding the His B motif (Figure 2). The *E. coli* and human homologs have a glycine at this position, so an equivalent interaction is not possible.

The hydrogen-bonding network that secures the sugar moiety of the histidyladenylate is conserved in the bacterial and trypanosomal complexes. The sidechain of the strongly conserved Gln 170 bridges ribose O4' and the histidyl O. The ribose O2' donates a hydrogen bond to the backbone carbonyl oxygen of Leu 336 and accepts a hydrogen bond from the backbone N of the strongly conserved residue Gly 363. The ribose O3' donates a hydrogen bond to the backbone carbonyl oxygen of Ala 335. There is also a conserved hydrogen bond from terminal O^c of the strongly conserved residue Glu 158 to N6 of the adenine ring.

However, the binding pocket occupied by the adenine ring system in the current *T. cruzi* complex with histidyladenylate is otherwise formed by an intriguingly different set of residues than previously observed for ATP or histidyladenylate complexes with bacterial homologs. The adenine ring as seen in the *E. coli* and *T. thermophilus* HAM complexes is sandwiched between Phe 125 and Arg 311 (*E. coli* numbering; Figure 7a). In the trypanosomal homologs, the residue equivalent to Phe 125 is His 168 (*T. cruzi* numbering), and the histidine ring does indeed perform the equivalent role in stacking against one side of the adenine ring (Figure 7b). But there is no sequence or structural trypanosomal equivalent for the bacterial Arg 311. On the side of the binding pocket from which the bacterial Arg 311 extends, the trypanosomal enzymes have instead a pair of residues, Cys 365 and Val 366, whose sidechains provide a hydrophobic surface adjacent to the adenine. The role of stacking against the other face of the adenine ring system is instead played by the sidechain of Arg 164, which extends from the opposite side of the binding pocket (Figure 7).

Unlike the structural difference between the bacterial and trypanosomal insertion domains, the differences in ATP binding mode demonstrated by the *T. cruzi* and *E. coli* HisRS do not represent a generic difference between eukaryotic and bacterial forms of the HisRS active site. The corresponding set of key residues found in other eukaryotic HisRS sequences, including in particular the two human homologs HARS and HARS2, are directly equivalent to those of the *E. coli* enzyme rather than to those of the *T. cruzi* enzyme (Table 1, Figure 2). That is, the details of ATP coordination deduced from studies of bacterial HisRS^{21,10,7} are likely to be valid for many HisRS homologs including those from higher eukaryotes, but the trypanosomal homologs deviate quite remarkably from this.

RNA interference of HisRS gene in *T. brucei*

As trypanosomes possess only a single gene coding for a histidyl-tRNA synthetase, it is to be expected that its gene product is an essential protein. We were able to confirm by RNAi silencing that even partial suppression of HisRS expression has a profound effect on the growth of bloodstream-form *T. brucei*. Targeted knockdown of the single HisRS gene of *T. brucei* was performed using established RNA interference methods for *T. brucei*²⁸. Cells induced to express the RNAi construct showed profound suppression of growth, by a factor of $>10^3$ at day 4 and beyond, compared to the uninduced cultures (Figure 8a). Interestingly, the degree of mRNA knockdown was far from complete as indicated by quantitative PCR of the mRNA (Figure 8b). It seems likely that more complete gene knockdown would lead to even more

complete growth arrest and possibly death. The profound growth phenotype with only partial knockdown of mRNA supports the selection of HisRS as a drug target, since pharmacological inhibition of the enzyme may not have to be complete to effectively stop growth and possibly kill the cells.

Potential of HisRS as an anti-protozoan drug target

Humans have two HisRS genes. Both are homologous to the trypanosomal HisRS, but the overall sequence identity between the trypanosomal enzymes and either human homolog is less than 30%. As we see in the current structures, this lack of sequence conservation extends to the residues making up the active site. The distinctive differences found in the adenine binding pocket of the trypanosomal HisRS as compared to the human homologs may well offer an opportunity for the identification of selective inhibitors that could provide a basis for anti-trypanosomal drug design.

In order to determine whether this opportunity extends to other pathogens, we reexamined HisRS sequences from other protozoa and representative higher eukaryotes. In particular we looked for the presence of a cluster of residues equivalent to *T. cruzi* [Arg 165, Arg 334, Cys 365, Val 366] rather than to the human or *E. coli* [Phe/Tyr, Gly, Glu, Arg] at the corresponding positions. We found this cluster in the single HisRS from the closely related trypanosomatid *Leishmania major* and in apicoplast-specific HisRS sequences from various species of the apicomplexan parasites *Plasmodium* and *Toxoplasma*. However, the cytosolic HisRS homologs identified in these same apicomplexan genomes instead matched the cluster of residues from the human HisRS sequences, as did the sequences identified from *Entamoeba*, *Giardia*, and *Cryptosporidia*. It remains an open question whether a hypothetical drug specific for the *T. cruzi*-like HisRS homologs in the apicoplast would be effective against apicomplexans possessing a second, cytosolic HisRS.

Summary

The trypanosomal histidyl-tRNA synthetase structures presented here illuminate differences between eukaryotic and prokaryotic HisRS homologs, and also highlight differences between the trypanosomal and human enzymes. The insertion domain of histidyl-tRNA synthetase in eukaryotes is seen to adopt a very different three-dimensional structure from that reported for bacterial homologs. As the insertion domain of bacterial HisRS has been implicated in the mechanisms of tRNA binding and aminoacylation^{10,11}, this leaves open the question of whether the equivalent eukaryotic mechanisms may differ to some extent. It would be informative to have a structure of a eukaryotic HisRS:tRNA^{His} complex, but to date this is lacking. We now see also that details of substrate binding differ even among eukaryotic HisRS homologs. The substrate binding pocket that accepts the adenine ring of ATP differs significantly in character between the trypanosomal structures reported here and the previously characterized bacterial and archaeal structures. The surface of the binding pocket adjacent to the adenine ring is formed by a pair of hydrophobic residues in trypanosomes, Cys 365 and Val 366, that replace an arginine found in the sequence of previously studied homologs, both bacterial and human. The previously assigned functions of this missing arginine are taken over by Arg 164, which we observe to stack against the adenine ring, and by Arg 311, which we infer to coordinate the ATP γ -phosphate. This may be a peculiarity of trypanosomes and related protozoa, as both of the human HisRS sequences are in these key regions more like those of bacteria than those of trypanosomes. Whatever the evolutionary basis for this difference, it constitutes a possibly exploitable target for the identification of selective anti-protozoan drugs targeting the essential enzyme histidyl-tRNA synthetase.

Materials and Methods

Choice of expression constructs and crystallization strategy

Following standard MSGPP procedure, both full length and truncated variants of the target *T. brucei* and *T. cruzi* sequences were selected for expression. Based on multiple sequence alignment of the HisRS sequence family, first priority was given to a truncation at the boundary between the extra N-terminal (WHEP) domain and the conserved HisRS core domains. Both variants of both proteins were cloned into the *E. coli* expression vector AVA0421, derived from vector pET14b²⁹. The AVA0421 vector contains an N-terminal His-tag followed by a protease 3C cleavage site, giving us the option of purifying cleaved or uncleaved protein for crystallization trials. In the event, no expression was observed for the full length *T. cruzi* variant. Both the cleaved and uncleaved forms of the full length *T. brucei* variant were subject to crystallization trials with and without addition of histidine as a co-crystallant, but no crystals were obtained that diffracted to better than 20 Å resolution. Crystallization trials of the uncleaved truncation variants were screened for crystallization with or without histidine, and with or without ATP. When crystals diffracting to better than 3 Å were obtained for the apo *T. brucei* and histidine-bound *T. cruzi* proteins, we focused on optimizing these rather than exploring additional truncation variants or tag cleavage.

Protein production and crystallography

***T. brucei* HisRS**—The sequence corresponding to GenBank accession no. [AAZ11768.1](#), coding for residues 44–477 of the 477 residue *T. brucei* HisRS, was PCR amplified from genomic DNA of *T. brucei* strain TREU927 GUTat 10.1, cloned into the AVA0421 vector, and expressed in *E. coli*. Protein was purified by a Ni-NTA affinity column followed by a XK 26/60 Superdex 75 column (Amersham Pharmacia Biotech) using the SGPP standard buffer (20mM HEPES, 0.5 M NaCl, 2 mM β-mercaptoethanol, 5% glycerol, 0.025% NaN₃ at pH 7.5)³⁰. Purified protein retained the 22 residue expression tag. Crystals used for data collection were grown by vapor diffusion from sitting drops equilibrated against a reservoir containing 2 M lithium sulfate, 0.1 M HEPES buffer at pH 7.3, and 1 mM TCEP. The initial crystallization drops consisted of 0.2 μl protein at 22 mg/ml in standard buffer plus 0.2 μl of the reservoir solution. Crystals were soaked in 20% ethylene glycol prior to being frozen in liquid nitrogen.

Crystallographic diffraction images were collected from a single crystal at beamline 9–2 of the Stanford Synchrotron Radiation Lightsource and processed using Mosflm³¹. An initial structural model was available from a previous molecular replacement experiment using lower resolution data collected from crystals in a different crystal form (data not shown). This model had been generated by the program Phaser using a monomer of the 35% sequence identical HisRS from the archaeote *Thermoplasma acidophilum* (PDB 1wu7) as a probe. However, this relatively poor model did not refine well and major portions of it were later rebuilt using the higher resolution structure of the *T. cruzi* His:HisRS complex as a guide.

***T. cruzi* HisRS**—The sequence corresponding to residues 45–478 of GenBank sequence accession no. [EAN89395.1](#), coding for the 478 residue *T. cruzi* HisRS, was cloned from genomic DNA of strain TCCL Brener and expressed in *E. coli*. Protein was purified as above. Purified protein retained the 22 residue expression tag. Crystals were grown by vapor diffusion from sitting drops equilibrated against a reservoir containing 0.2 M lithium sulfate, 0.1 M BisTris buffer at pH 5.5, 26–28% w/v PEG 3350, 1 mM TCEP, and 10 mM histidine. The initial crystallization drops consisted of 0.15 μl protein at 24 mg/ml in standard buffer plus 0.15 μl of the reservoir solution. Crystals with differing unit cells and space groups were found to grow from these same conditions. Diffraction data to 1.8 Å resolution from a single crystal of the His:HisRS complex belonging to space group C2 were collected at SSRL beamline 9–2. Integration and scaling were performed in Mosflm³¹ and Scala³². An initial molecular

replacement model was generated by BALBES³³, which used multiple fragments of the *T. acidophilum* homolog (1wu7) as probes.

The histidyladenylate:HisRS complex (HAM:HisRS) was generated by allowing the enzyme to catalyze the aminoacyl activation reaction

in crystals of the His:HisRS complex. The reaction buffer was prepared by adding 0.2 μl 250 mM ATP in standard buffer and 0.2 μl 250 mM MgCl_2 in standard buffer to 5 μl of the crystallization reservoir solution from a neighboring well containing slightly higher PEG concentration (30% w/v). Then 2 μl of the reaction buffer were added directly to the crystallization drop containing already grown crystals. Most crystals crumbled immediately, but a few survived or re-formed. These were left to re-equilibrate against the original reservoir solution for 85 minutes before being frozen directly in liquid nitrogen. Diffraction data from a single crystal were collected at SSRL beamline 9-2, then integrated and scaled using HKL2000³⁴. The space group of this HAM:HisRS crystal was found to be different from those determined for unsoaked His:HisRS crystals (Table 2). The treated crystal diffracted only to 3 Å resolution, a substantial degradation from the diffraction quality of untreated crystals. It is unclear whether the crystal represents a minority crystal form that better tolerated exposure to the reaction buffer, or whether it morphed into a new crystal lattice that reformed rapidly after reaction-induced disruption of the original lattice.

Crystallographic data handling and project management used the CCP4 program suite^{32,35}. Structural superpositions were performed using Coot and SSM³⁶. Refinement of all three structural models proceeded using alternating cycles of refinement in Refmac5³⁷ and model building in Coot³⁸. Moderate non-crystallographic symmetry (NCS) restraints were applied to the apo (6 NCS-related chains) and HAM:HisRS complex (2 NCS-related chains) models throughout refinement. Model validation was performed using Molprobity³⁹. The refined crystallographic model for each structure included a multi-segment model for atomic displacements generated by TLSMD40. Crystallographic statistics and Protein Data Bank accession codes are given in Table 2.

Thermal melt assays

Protein stability was measured in solution as a function of temperature by adding a hydrophobic dye, Sypro Orange (SigmaAldrich). The dye emits low fluorescence in aqueous solution surrounding a well-folded protein but gives increased fluorescent signal as a consequence of binding to hydrophobic patches exposed as the protein releases a bound ligand and eventually denatures. In the presence of a high-affinity ligand, the protein will in general be more resistant to thermal denaturation, resulting in a shift ΔT_M in the inflection point of the melting curve⁴¹. The effect of substrate binding on three trypanosomal histidyl-tRNA synthetases was assayed by adding substrate or other potential ligands to 0.5 mg/ml protein in standard buffer and monitoring fluorescence over a temperature range of 20–90° C. The assay was carried out in 96-well trays in a DNA Engine Opticon 2 RT-PCR machine (BioRad). Melting curve shifts were measured for the *T. cruzi* and *T. brucei* N-terminally truncated HisRS proteins used for structural studies, and also for the full length HisRS from another trypanosomatid, *Leishmania naiffi*.

RNA interference of *T. brucei* HisRS

To confirm that the HisRS gene product is essential, the *T. brucei* gene was subjected to RNAi knockdown. The region within the gene coding sequence to be used for RNAi was selected

using the RNA target selection program RNAit⁴² to ensure that there was no significant sequence homology with other genes within the genome. The sequence of the *T. brucei* genomic DNA used for the construction of the RNAi plasmids was taken from GeneDB accession number Tb927.6.2060. Bases 467 to 1097 were amplified from *T. brucei* 927 genomic DNA using primers 5'-ATGAGGCAATTACTCGTGGG-3' and 5'-ATAACGCAATCACCAAAGCC-3'. The resulting amplicon was ligated by TA cloning into the vector p2T7TABBlue (a gift of D. Horn, London School of Hygiene and Tropical Medicine)²⁸, and the sequence was confirmed by nucleotide sequence analysis of the insert. The construct was linearized with the NotI restriction enzyme (New England Biolabs, Ipswich, MA). *T. brucei* bloodstream-form parasites (provided by G. Cross, Rockefeller University) expressing the T7 RNA polymerase and Tet repressor under a single selection marker (SM), G418 resistance, were cultured in HMI-9 medium with 10% heat-inactivated fetal bovine serum and G418 at 2.5 μ g/ml at 37°C in a 5% CO₂ atmosphere⁴³. Ten micrograms of linearized p2T7TABBlue containing the HisRS RNAi fragment was electroporated into mid-log-phase *T. brucei* (2.5×10^7 cells) suspended in 500 μ l of cytomix (120 mM KCl, 0.15 mM CaCl₂, 10 mM K₂HPO₄-KH₂PO₄, pH 7.6, 25 mM HEPES, 2 mM disodium EDTA, 5 mM MgCl₂). Electroporation was done in 4-mm gap cuvettes at 1.6 kV and 24 Ω resistance, as described previously⁴⁴. The cells were resuspended in HMI-9 medium supplemented with 2.5 μ g/ml hygromycin and 2.5 μ g/ml G418. Stable individual clones were selected 5 to 7 days after transfection. Selected cultures diluted to 1×10^5 cells/ml were induced to express double-stranded RNA by addition of 1 μ g/ml tetracycline. Parasite growth was measured over a period of 8 days following induction of RNAi in cultures passed daily at a 1:2 to 1:20 dilution as needed, and was monitored with ATPLite luminescence ATP detection assay system (catalog no. 6016941; Perkin-Elmer, Waltham, MA). For quantification of gene knockdown, cDNA was prepared from messenger RNA collected at 0h, 24h, and 48h after RNAi induction with tetracycline. mRNA signal knock-down was analyzed by quantitative PCR of the cDNA using primers 5'-TCAATTAAGGCGTCTGATAGCG-3' and 5'-TGTCGTGCAAGTGATGGTGTC-3', which amplified a separate region of the gene than that which was used for the RNAi construct. The amplified products after 23 PCR cycles were quantified by densitometry (normalized to β -tubulin).

Acknowledgments

This work was funded by NIAID award P01AI067921 (Medical Structural Genomics of Pathogenic Protozoa). We thank Angela Criswell of Rigaku Americas Corporation for collecting X-ray data used for initial characterization of crystals of the *T. cruzi* HisRS. Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences.

Abbreviations

HisRS	histidyl-tRNA synthetase
HAM	histidyl-adenosinemonophosphate (histidyladenylate)
aaRS	aminoacyl-tRNA synthetase

References

1. Sheppard K, Yuan J, Hohn MJ, Jester B, Devine KM, Soll D. From one amino acid to another: tRNA-dependent amino acid biosynthesis. *Nucleic Acids Res* 2008;36:1813–25. [PubMed: 18252769]
2. Hurdle JG, O'Neill AJ, Chopra I. Prospects for aminoacyl-tRNA synthetase inhibitors as new antimicrobial agents. *Antimicrob Agents Chemother* 2005;49:4821–33. [PubMed: 16304142]
3. Fan E, Baker D, Gelb MH, Buckner FS, Van Voorhis WC, Phizicky E, Dumont M, Mehlin C, Grayhack EJ, Sullivan M, Verlinde CL, DeTitta G, Meldrum D, Merritt EA, Earnest TN, Soltis M, Zucker F,

- Myler P, Schoenfeld L, Kim D, Worthey EA, LaCount D, Vignali M, Li J, Mondal S, Massey A, Carroll B, Gulde S, Luft JR, DeSoto L, Holl M, Caruthers JM, Bosch J, Robien MA, Arakaki T, Holmes MA, LeTrong I, Hol WG. Structural genomics of pathogenic protozoa: An overview. *Methods in Molecular Biology* 2008;426:497–513. [PubMed: 18542886]
4. Charriere F, O'Donoghue P, Helgadottir S, Marechal-Drouard L, Cristodero M, Horn EK, Soll D, Schneider A. Dual Targeting of a tRNA^{Asp} Requires Two Different Aspartyl-tRNA Synthetases in *Trypanosoma brucei*. *J Biol Chem* 2009;284:16210–7. [PubMed: 19386587]
 5. Charriere F, Helgadottir S, Horn EK, Soll D, Schneider A. Dual targeting of a single tRNA(Trp) requires two different tryptophanyl-tRNA synthetases in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* 2006;103:6847–52. [PubMed: 16636268]
 6. Wolf YI, Aravind L, Grishin NV, Koonin EV. Evolution of aminoacyl-tRNA synthetases—analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res* 1999;9:689–710. [PubMed: 10447505]
 7. Hawko SA, Francklyn CS. Covariation of a specificity-determining structural motif in an aminoacyl-tRNA synthetase and a tRNA identity element. *Biochemistry* 2001;40:1930–6. [PubMed: 11329259]
 8. Sankaranarayanan R, Dock-Bregeon AC, Romby P, Caillet J, Springer M, Rees B, Ehresmann C, Ehresmann B, Moras D. The structure of threonyl-tRNA synthetase-tRNA(Thr) complex enlightens its repressor activity and reveals an essential zinc ion in the active site. *Cell* 1999;97:371–81. [PubMed: 10319817]
 9. Yaremchuk A, Tukalo M, Grotli M, Cusack S. A succession of substrate induced conformational changes ensures the amino acid specificity of *Thermus thermophilus* prolyl-tRNA synthetase: comparison with histidyl-tRNA synthetase. *J Mol Biol* 2001;309:989–1002. [PubMed: 11399074]
 10. Aberg A, Yaremchuk A, Tukalo M, Rasmussen B, Cusack S. Crystal structure analysis of the activation of histidine by *Thermus thermophilus* histidyl-tRNA synthetase. *Biochemistry* 1997;36:3084–94. [PubMed: 9115984]
 11. Guth E, Farris M, Bovee M, Francklyn CS. Asymmetric Amino Acid Activation by Class II Histidyl-tRNA Synthetase from *Escherichia coli*. *J Biol Chem* 2009;284:20753–62. [PubMed: 19487703]
 12. Rho SB, Lee KH, Kim JW, Shiba K, Jo YJ, Kim S. Interaction between human tRNA synthetases involves repeated sequence elements. *Proc Natl Acad Sci U S A* 1996;93:10128–33. [PubMed: 8816763]
 13. Jia J, Arif A, Ray PS, Fox PL. WHEP domains direct noncanonical function of glutamyl-Prolyl tRNA synthetase in translational control of gene expression. *Mol Cell* 2008;29:679–90. [PubMed: 18374644]
 14. Mathews MB, Bernstein RM. Myositis autoantibody inhibits histidyl-tRNA synthetase: a model for autoimmunity. *Nature* 1983;304:177–9. [PubMed: 6866113]
 15. Cahuzac B, Berthonneau E, Birlirakis N, Guittet E, Mirande M. A recurrent RNA-binding domain is appended to eukaryotic aminoacyl-tRNA synthetases. *EMBO J* 2000;19:445–52. [PubMed: 10654942]
 16. O'Hanlon TP, Raben N, Miller FW. A novel gene oriented in a head-to-head configuration with the human histidyl-tRNA synthetase (HRS) gene encodes an mRNA that predicts a polypeptide homologous to HRS. *Biochem Biophys Res Commun* 1995;210:556–66. [PubMed: 7755634]
 17. O'Hanlon TP, Miller FW. Genomic organization, transcriptional mapping, and evolutionary implications of the human bi-directional histidyl-tRNA synthetase locus (HARS/HARSL). *Biochem Biophys Res Commun* 2002;294:609–14. [PubMed: 12056811]
 18. Brindefalk B, Viklund J, Larsson D, Thollessen M, Andersson SGE. Origin and evolution of the mitochondrial aminoacyl-tRNA synthetases. *Mol Biol Evol* 2007;24:743–56. [PubMed: 17182897]
 19. Arnez JG, Augustine JG, Moras D, Francklyn CS. The first step of aminoacylation at the atomic level in histidyl-tRNA synthetase. *Proc Natl Acad Sci U S A* 1997;94:7144–9. [PubMed: 9207058]
 20. Qiu X, Janson CA, Blackburn MN, Chhohan IK, Hibbs M, Abdel-Meguid SS. Cooperative structural dynamics and a novel fidelity mechanism in histidyl-tRNA synthetases. *Biochemistry* 1999;38:12296–304. [PubMed: 10493797]
 21. Arnez JG, Harris DC, Mitschler A, Rees B, Francklyn CS, Moras D. Crystal structure of histidyl-tRNA synthetase from *Escherichia coli* complexed with histidyl-adenylate. *EMBO J* 1995;14:4143–55. [PubMed: 7556055]

22. Shiba K. Intron positions delineate the evolutionary path of a pervasively appended peptide in five human aminoacyl-tRNA synthetases. *J Mol Evol* 2002;55:727–33. [PubMed: 12486531]
23. Raben N, Nichols R, Dohlman J, McPhie P, Sridhar V, Hyde C, Le3 R, Plotz P. A motif in human histidyl-tRNA synthetase which is shared among several aminoacyl-tRNA synthetases is a coiled-coil that is essential for enzymatic activity and contains the major autoantigenic epitope. *J Biol Chem* 1994;269:24277–83. [PubMed: 7523371]
24. Cole C, Barber JD, Barton GJ. The jpred 3 secondary structure prediction server. *Nucleic acids research* 2008;36:W197–201. [PubMed: 18463136]
25. Pollastri G, McLysaght A. Porter: a new, accurate server for protein secondary structure prediction. *Bioinformatics* 2005;21:1719–20. [PubMed: 15585524]
26. Guth EC, Francklyn CS. Kinetic discrimination of tRNA identity by the conserved motif 2 loop of a class II aminoacyl-tRNA synthetase. *Mol Cell* 2007;25:531–42. [PubMed: 17317626]
27. Guth E, Connolly SH, Bovee M, Francklyn CS. A substrate-assisted concerted mechanism for aminoacylation by a class II aminoacyl-tRNA synthetase. *Biochemistry* 2005;44:3785–94. [PubMed: 15751955]
28. Alibu VP, Storm L, Haile S, Clayton C, Horn D. A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. *Mol Biochem Parasitol* 2005;139:75–82. [PubMed: 15610821]
29. Alexandrov A, Vignali M, LaCount DJ, Quartley E, deVries C, Rosa DD, Babulski J, Mitchell SF, Schoenfeld LW, Fields S, Hol WG, Dumont ME, Phizicky EM, Grayhack EJ. A facile method for high-throughput co-expression of protein pairs. *Mol Cell Proteomics* 2004;3:934–938. [PubMed: 15240823]
30. Mehlin C, Boni E, Buckner FS, Engel L, Feist T, Gelb M, Haji L, Kim D, Liu C, Mueller N, Myler PJ, Reddy JT, Sampson JN, Subramanian E, Van Voorhis WC, Worthey E, Zucker F, Hol WGJ. Heterologous expression of proteins from *plasmodium falciparum*: results from 1000 genes. *Molecular and Biochemical Parasitology* 2006;148:144–160. [PubMed: 16644028]
31. Leslie AGW. Recent changes to the mosflm package for processing film and image plate data. *Joint CCP4 and ESF-EACBM Newsletters on Protein Crystallography* 1992;26:x.
32. Collaborative Computational Project No. 4. The ccp4 suite: programs for protein crystallography. *Acta Crystallographica Section D* 1994;50:760–763.
33. Long F, Vagin AA, Young P, Murshudov GN. BALBES: a molecular-replacement pipeline. *Acta Crystallographica Section D* 2008;64:125–132.
34. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. In: Carter, CW., Jr; Sweet, RM., editors. *Methods Enzymol. Part A. Vol. 276*. Academic Press; New York: 1997. p. 307–326. *Macromolecular Crystallography*
35. Potterton L, McNicholas S, Krissinel E, Gruber J, Cowtan K, Emsley P, Murshudov GN, Cohen S, Perrakis A, Noble M. Developments in the ccp4 molecular-graphics project. *Acta Crystallographica Section D* 2004;60:2288–94.
36. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica Section D* 2004;60:2256–2268.
37. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D* 1997;53:240–255.
38. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D* 2004;60:2126–32.
39. Lovell S, Davis I, Arendall WB III, deBakker P, Word J, Prisant M, Richardson J, Richardson D. Structure validation by $C\alpha$ geometry: ϕ , ψ and $C\beta$ deviation. *Proteins: Structure, Function, and Genetics* 2003;50:437–450.
40. Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing tls motion. *Acta Crystallographica Section D* 2006;62:439–450.
41. Lo MC, Aulabaugh A, Jin G, Cowling R, Bard J, Malamas M, Ellestad G. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal Biochem* 2004;332:153–9. [PubMed: 15301960]
42. Redmond S, Vadivelu J, Field MC. RNAit: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Mol Biochem Parasitol* 2003;128:115–8. [PubMed: 12706807]

43. Hirumi H, Hirumi K. Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J Parasitol* 1989;75:985–9. [PubMed: 2614608]
44. Wirtz E, Leal S, Ochatt C, Cross GA. A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol* 1999;99:89–101. [PubMed: 10215027]
45. Eriani G, Delarue M, Poch O, Ganglo3 J, Moras D. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* 1990;347:203–6. [PubMed: 2203971]

Figure 1. The dimer of *T. cruzi* HisRS with bound histidyladenylate

Monomer A is shown in dark brown (catalytic domain), yellow (anticodon binding domain), and light brown (insertion domain). Monomer B is similarly shown in blue. The active site of each monomer is occupied by histidyladenylate, the product of the first half-reaction catalyzed by the enzyme.

Figure 2. Sequence alignment of HisRS from *T. brucei* and *T. cruzi* with the human cytosolic (HARS) and mitochondrial (HARS2) homologs, and with the *E. coli* homolog

The three conserved characteristic Class II aaRS catalytic domain motifs as defined by Eriani *et al.*⁴⁵ are indicated above the sequence, as are two conserved histidine recognition motifs specific to HisRS. Secondary structure elements as observed in the *T. cruzi* His:HisRS structure are shown above the sequences. The α -helices of the N-terminal WHEP domain as seen in NMR studies of human and other eukaryotic homologs are indicated below the human HARS2 sequence. This domain was not included in the expression constructs for the trypanosomal proteins studied here. The region of the *E. coli* sequence highlighted in orange indicates a portion of the insertion domain that is not structurally equivalent in the bacterial and archaeal/eukaryotic structures. Residues interacting directly with histidyladenylate in the *T. cruzi* structure are indicated by blue letters below the sequence: H/P/S/A signify interaction with the histidyl/phosphate/sugar/adenine moieties respectively. The red P below residue 334 indicates inferred interaction with the γ -phosphate of ATP. Points of relevant sequence difference between the trypanosomal and human sequences are highlighted in cyan. These are residues that contribute to a difference in the shape or hydrophobic character of the binding surface in the active site.

Figure 3. The insertion domain of eukaryotic and bacterial HisRS structures

All HisRS contain an additional domain that is inserted between motifs 2 and 3 of the conserved Class II catalytic domains. However, the insertion domains of the eukaryotic and bacterial HisRS homologs are not similar in either sequence or structure. The figures were generated by superimposing the *T. cruzi* (3hrk) and *T. thermophilus* (1ady) dimers and then separating their respective insertion domains by pure translation in the horizontal. The grey helix to the lower left of each structure is the first structurally equivalent element following the insertion domain. (a) The *T. cruzi* insertion domain, residues 210–285, followed by the helix α 13 of the conserved catalytic domain. (b) The *T. thermophilus* insertion domain, residues 164–235, followed by the homologous catalytic domain helix.

Figure 4. Superposition of the apo HisRS onto the histidyladenylate:HisRS complex

The protein undergoes a substantial conformational change upon binding histidine. The apo form is shown in light colors; the histidyladenylate:HisRS complex is shown in dark colors. In this superposition the apo form is from *T. brucei* whereas the histidine-bound form is from *T. cruzi*; however, the sequence identity of the two trypanosomal HisRS is 84% and we believe that the highlighted conformational change can be attributed to histidine binding rather than to sequence differences. The apo form has little affinity for ATP. The rearrangement of residues 314–317, the His A motif, to bind histidine also acts to reposition Arg314 such that it is oriented to interact with the α -phosphate of ATP. Other structural elements also undergo a substantial conformational shift.

Condition	<i>T. cruzi</i>	<i>T. brucei</i>	<i>L. naiiffi</i>
	ΔT_M	ΔT_M	ΔT_M
control (Buffer only)	0 ± 0.1	0 ± 0.1	0 ± 0.1
10 mM ATP/MgCl ₂	-0.7 ± 0.1	-0.7 ± 0.1	-0.8 ± 0.1
10 mM ADP/MgCl ₂	-0.6 ± 0.2	-0.2 ± 0.1	-0.5 ± 0.1
10 mM AMP/MgCl ₂	0.0 ± 0.1	0.1 ± 0.1	-0.3 ± 0.1
10 mM L-Histidine	4.0 ± 0.2	6.0 ± 0.1	4.5 ± 0.3
10 mM L-Histidine + 10 mM ADP/MgCl ₂	4.6 ± 0.2	6.7 ± 0.1	7.0 ± 0.1
10 mM L-Histidine + 10 mM AMP/MgCl ₂	4.9 ± 0.3	8.0 ± 0.9	8.9 ± 0.1
10 mM L-Histidine + 10 mM ATP/MgCl ₂	7.2 ± 0.1	11.2 ± 0.2	11.9 ± 0.1

Figure 5. Effect of substrate binding on thermal melt curves

(a) Fluorescence as a function of temperature. Each curve has been normalized to the maximum peak height. Two replicate curves are shown for each experimental condition. (b) Shift in T_M , defined as the change in the inflection point of the melting curve in the presence of a ligand, relative to the inflection point of the melting curve for the protein without ligand.

Figure 6. Difference electron density at the two active sites of the *T. cruzi* HAM:HisRS complex
(a) Site A. Difference density is contoured at 2.4σ (grey cages) and 4.5σ (green cages) in a sigma-A weighted omit map. **(b)** Site B. Difference density is contoured at 2.0σ and 3.6σ in a sigma-A weighted omit map. The view angle is slightly different from (a) in order to more clearly show the better density for the histidyl moiety.

Figure 7. Adenine binding pocket

(a) The binding mode of histidyladenylate as seen previously in structures of the *T. thermophilus* (1ady) and *E. coli* HisRS (1kmm, not shown). One face of the adenine ring is stacked against the sidechain of Arg 311. (b) The binding mode of histidyladenylate as seen in the structure of *T. cruzi* HisRS (3hrk). The dot surface surrounds the hydrophobic residues Cys 365 and Val 366, which occur at the sequence position equivalent to the bacterial residue Arg 311. In the trypanosomal HisRS, the adenine ring is stacked instead against the sidechain of Arg 164, which extends from the opposite face of the binding pocket. Residue 334, visible at the left of the figure, is too far away to interact directly with the adenine ring, but is hypothesized to coordinate the γ -phosphate of ATP.

Time from TCN Induction

0H

24H

48H

His RS Gene

B-Tubulin

Figure 8. RNAi silencing of *T. brucei* HisRS

(a) Growth of bloodstream *T. brucei* cells in culture at 1 day intervals following RNAi induction. RNA interference of HisRS was induced with tetracycline (TCN +) and compared to cells without RNAi induction (TCN -). The growth rate was also compared with the 'single marker' (SM) strain of *T. brucei* that was not transfected with the RNAi expression vector. Cumulative cell densities are shown on a log-scale as the product of the cell number and the total dilution. Another clone was observed to have a nearly identical phenotype (data not shown). (b) Knockdown of *T. brucei* HisRS mRNA at 0, 24, and 48 hours after induction of RNAi with tetracycline. PCR products at 23 cycles were imaged by staining an agarose gel

with ethidium bromide. Even loading is demonstrated by PCR of the β -tubulin gene. The intensity of the bands at 24h and 48h was 42% and 71% of the 0h control, respectively, as measured by densitometry.

Table 1

Residues involved in ATP binding

	<i>T. cruzi</i> (3HRK)	<i>E. coli</i> (1KMN)	human HARS(2)
H-bond to adenine N6	Glu 158	Glu 114	Asp(Glu) 159
sandwiches Adenine ring	His 168 Arg 164 (not structurally homologous)	Phe 125 Arg 311	Phe 171 Arg 388
coordinates α phosphate	Arg 156 Arg 314	Arg 113 Arg 259	Arg 157 Arg 326
coordinates β phosphate	Arg 156 (inferred)	Arg 113	Arg 157
coordinates γ phosphate	Arg 164 (inferred) Arg 334 (not structurally homologous)	Arg 121 Arg 311	Arg 167 Arg 388
provides hydrophobic surface	Cys 365 Val 366	no equivalent	no equivalent

The residues tabulated for *T. cruzi* are those observed in the HisRS:HAM complex, except that the coordination of the β - and γ -phosphates is inferred. The residue numbering for bacterial HisRS is that of the *E. coli* sequence. No structure is available for human HisRS, but for the set of residues shown, the human HARS and HARS2 sequences align exactly with the *E. coli* sequence (Figure 2). Residue numbering is that of human HARS.

Table 2

Crystallographic data collection and refinement statistics

PDB ID	<i>T. brucei</i> HisRS 3HRI	<i>T. cruzi</i> HisRS 3HRJ	<i>T. cruzi</i> HAM:HisRS 3HRK
Data collection			
Wavelength (Å)	0.97923	0.97946	0.97923
Resolution (Å)	64-2.7 (2.85-2.7)	40-1.8 (1.9-1.8)	50-3.05 (3.16-3.05)
Space group	I2	C2	P2 ₁ 2 ₁ 2
Unit cell (Å)	a=148.6 b=98.4 c=252.0 $\alpha=90^\circ \beta=90.3^\circ \gamma=90^\circ$	a=90.3 b=118.8 c=66.0 $\alpha=90^\circ \beta=132.6^\circ \gamma=90^\circ$	a=112.4 b=167.0 c=66.0 $\alpha=90^\circ \beta=90^\circ \gamma=90^\circ$
Monomers per asymmetric unit	6	1	2
Unique reflections	99984 (14507)	47130 (6847)	19599 (1641)
R_{merge}	0.166 (0.728)	0.042 (0.637)	0.120 (0.905)
Completeness (%)	100 (100)	100 (100)	80 (69)
Mean I/ σ	6.0 (1.8)	14.0 (1.8)	7.7 (1.05)
Redundancy	3.8 (3.7)	3.8 (3.8)	3.0 (2.7)
Wilson B factor (Å ²)	61	33	63
Refinement			
Resolution (Å)	35-2.85	34-1.80	43-3.05
R/R_{free}	0.271/0.313	0.193/0.211	0.232/0.308
Number of protein atoms	18087	3296	6514
Number of other atoms	0	189	67
Mean B_{eq} protein (Å ²)			
rmsd deviation from ideal bond lengths (Å)	0.010	0.011	0.010
rmsd deviation from ideal bond angles (°)	1.173	1.264	1.119
TLS groups per chain	8	8	8
Residues with ϕ, ψ in favored regions (%)	93.8	97.8	90.9
Residues with ϕ, ψ in allowed regions (%)	5.9	1.7	8.3
Number of ϕ, ψ outliers (residues)	7	2	7