The EMBO Journal Vol.16 No.12 pp.3373–3385, 1997

Crystal structure and mechanism of human L-arginine:glycine amidinotransferase: a mitochondrial enzyme involved in creatine biosynthesis

Andreas Humm[1], Erich Fritsche, Stefan Steinbacher and Robert Huber

Max-Planck-Institut für Biochemie, Abt. Strukturforschung, Am Klopferspitz 18a, D-82152 Martinsried, Germany

[1]Corresponding author

L-arginine:glycine amidinotransferase (AT) catalyses the committed step in creatine biosynthesis by formation of guanidinoacetic acid, the immediate precursor of creatine. We have determined the crystal structure of the recombinant human enzyme by multiple isomorphous replacement at 1.9 Å resolution. A telluromethionine derivative was used in sequence assignment. The structure of AT reveals a new fold with 5-fold pseudosymmetry of circularly arranged ββαβ-modules. These enclose the active site compartment, which is accessible only through a narrow channel. The overall structure resembles a basket with handles that are formed from insertions into the ββαβ-modules. Binding of L-ornithine, a product inhibitor, reveals a marked induced-fit mechanism, with a loop at the active site entrance changing its conformation accompanied by a shift of an α-helix by ~4 Å. Binding of the arginine educt to the inactive mutant C407A shows a similar mode of binding. A reaction mechanism with a catalytic triad Cys–His–Asp is proposed on the basis of substrate and product bound states.

Keywords: creatine/induced fit/pseudosymmetry/reaction mechanism/telluromethionine

Introduction

Creatine and its phosphorylated form play an essential role in the energy metabolism of muscle and nerve tissues, acting as a dynamic reservoir of high-energy phosphate which buffers the rapid fluctuations of the ATP/ADP ratio during muscle and nerve action (Walker, 1979). The highest levels of creatine and phosphocreatine are found in skeletal muscle, heart, spermatozoa and photoreceptor cells of the retina. Here, the major part of creatine is phosphorylated in the mitochondrial intermembrane space by creatine kinase (Bessman and Geiger, 1981; Bessman, 1985; Wyss and Wallimann, 1994). The structure of the octameric mitochondrial creatine kinase isoform from chicken cardiac muscle (EC 2.7.3.2) has been solved recently to 3.0 Å resolution (Fritz-Wolf *et al.*, 1996). However, the major sites of creatine biosynthesis are kidney, pancreas and liver. Creatine synthesized in these organs is transported via the blood stream to the organs with a high requirement for creatine and taken up from the blood by Na^+-dependent creatine transporters (Ku and Passow, 1980; Syllm-Rapoport *et al.*, 1980; Guimbal and Kilimann, 1993; Sora *et al.*, 1994). The biosynthesis of creatine comprises the steps shown in Figure 1: formation of guanidinoacetic acid from L-arginine and glycine by L-arginine:glycine amidinotransferase (AT, EC 2.1.4.1) and transfer of a methyl group to guanidinoacetate by *S*-adenosylmethionine:guanidinoacetate methyltransferase (GMT, EC 2.1.1.2) (Walker, 1973). While AT is located in the cytosol and in the intermembrane space of mitochondria, GMT was only found in the cytosol (Magri *et al.*, 1975; Walker and Hannan, 1976; Walker, 1979). The activity of AT, the enzyme catalysing the committed step in creatine biosynthesis, is regulated tightly. The enzyme is inhibited by ornithine, one of the reaction products (Sipila, 1980). The expression of AT was reported to be induced by growth hormone and thyroxine (McGuire *et al.*, 1980) and repressed by creatine (McGuire *et al.*, 1984; Guthmiller *et al.*, 1994) and in Wilms' tumour (Austruy *et al.*, 1993). Inhibition of AT by ornithine was suggested to be a major cause of muscle weakness in gyrate atrophy of the choroid and retina with hyperornithinaemia (Sipila *et al.*, 1979, 1980; Sipila, 1980). The sequences of the human, porcine (Humm *et al.*, 1994) and rat (Guthmiller *et al.*, 1994) enzymes were obtained recently by protein and cDNA sequencing. Human mitochondrial AT is synthesized as a precursor of 423 amino acids from which the N-terminal 37 residues are supposed to be cleaved off upon transport of the protein to the mitochondrial intermembrane space, yielding a mature protein of 386 amino acid residues (Humm *et al.*, 1997). A second cDNA sequence for a human AT was submitted recently to the DDBJ/EMBL/GenBank database (accession No. X86401) by Austruy *et al.* The sequence codes for a 391 amino acid protein, differing from the mature mitochondrial enzyme by only five additional N-terminal residues. This protein is supposed to represent the cytosolic form of AT. Presumably, the mitochondrial and cytosolic forms of the enzyme are derived from the same gene by alternative splicing of the mRNA precursor (Humm *et al.*, 1997).

Active AT was obtained by recombinant expression of the human mature mitochondrial protein with an N-terminal six-histidine tag (AT38H) in *Escherichia coli*. Cys407 was shown to be the active site residue by radioactive labelling of the active site cysteine with L-[guanidino-^{14}C]arginine and by site-directed mutagenesis (Humm *et al.*, 1997).

Here we present the three-dimensional structure of recombinant human AT, the first enzyme in creatine biosynthesis to be structurally characterized. The crystal structure was resolved by multiple isomorphous replacement (MIR). Crystals of AT expressed with *N*-acetyltelluromethionine (Karnbrock *et al.*, 1996) were especially useful to assign the amino acid sequence in the electron density map. In addition, the structures of AT

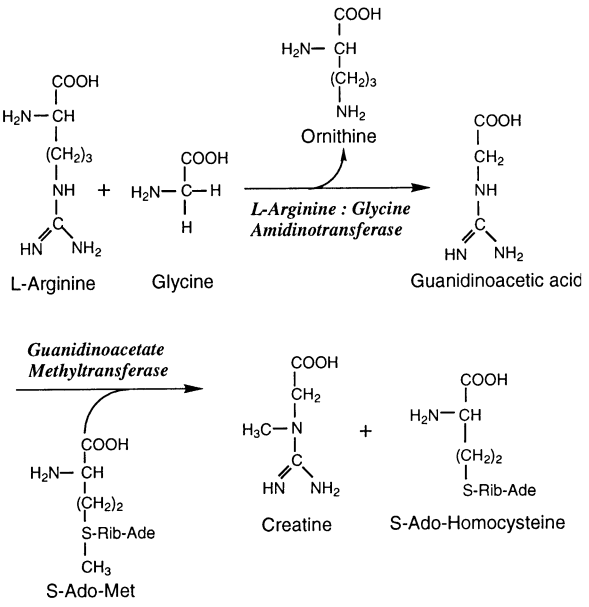

Fig. 1. Enzymatic reactions leading to the formation of creatine.

soaked with L-ornithine and of the active site mutant C407A soaked with L-arginine were determined. This allowed a reaction mechanism for AT to be proposed which involves a covalent amidino–enzyme intermediate as proposed by Grazi and Rossi (1968).

Results

Secondary and tertiary structure

The crystal structure of the mature form of mitochondrial human kidney L-arginine:glycine amidinotransferase (AT38H, residues 38–423 encoded by the cDNA, Humm *et al.*, 1997) was solved by MIR. The model comprises residues 64–423. Residues 38–63 as well as the 14 amino acid fusion peptide that had been cloned to the N-terminus of the protein for affinity purification are not visible in the electron density map and are mobile or disordered.

The polypeptide chain folds into a compact single-domain structure. Viewed from the side, the structure looks like a basket with handles (Figure 2A). When inspected from the bottom, the structure reveals 5-fold pseudosymmetry, built up from five modules consisting of ββαβ-motifs ordered circularly around the 5-fold symmetry axis like ribs of a basket (Figure 2B). The first two strands of the β-sheet of each module are antiparallel, the third strand is parallel to the second strand in each module (Figure 3). Insertions into these building blocks form connectivities between the modules. These insertions contain a further 10 helices and a two-stranded antiparallel β-sheet (strands S3 and S4 in module I). The short helices H5, H8, H10 and H13 are in a 3_{10} conformation. The helices H2 and H9, formed from insertions into the modules I and II (purple and blue in Figures 2 and 3), form the handles of the basket. A very narrow channel leads to the active site residues. It is formed by coil regions of modules I, III and IV. Module I shows only the β-strands two and three of the ββαβ-motif. The first β-strand is delivered to the β-sheet of module I from β-strand S17 of module V. Although the three dimensional structures of the modules are similar, the lengths of helices and strands are different. The primary structures reveal no significant internal repeats. A sequence alignment including the secondary structure elements is depicted in Figure 4.

In the following the individual modules are described (Figure 3). Module I comprises the N-terminus and residues 64–171. After a short coil segment β-strand S1, the helices H1 and the handle helix H2 become part of module I which is formed mainly by the long α-helix H3, β-strand S2 and an insertion containing the antiparallel β-sheet of strands S3 and S4 as well as the short H4 helix. The strands S1 and S2, together with β-strand S17 of module V, form the first rib of the basket. All secondary structure elements added to the (β)βαβ-motif of the first module are arranged perpendicular to the basic rib motif and the 5-fold axis of the protein structure.

Module II is formed from residues 172–256. The antiparallel β-strands S5 and S6 are followed by a small insertion containing the short helices H5 and H6. Helix H6 is joined by a short coil to helix H7 and strand S7, which together with the strands S5 and S6 form the rib of this module. The residual amino acid residues of this module form a further insertion with the short 3_{10}-helices H8 and H10 and the helix H9.

Helix H10 is connected to strand S8 of module III, which comprises residues 257–307. This module contains only the central ββαβ-motif formed from S8, S9, H11 and S10 and a C-terminal coil, from which residues M302, H303 and D305 contribute to the active site of the enzyme.

Module IV consists of residues 308–363 and the secondary structure elements S11, S12, H12, S13 and H13. The first four of these elements form the ββαβ-core motif. The 3_{10} helix H13 is part of the insertion of this module.

Module V consists of residues 364–423. The secondary structure elements S14, S15, H14 and S16 build up the fifth rib of the basket and are followed by an insertion leading to the C-terminus of the polypeptide chain. This insertion contains helix H15 and the active site cysteine residue C407 which is placed just in front of the C-terminus with β-strand S17, forming the inner strand of the β-sheet of module I.

Dimer assembly

L-arginine:glycine amidinotransferase isolated from human kidney was reported to form dimers by sedimentation-equilibrium experiments (Gross *et al.*, 1986). When we performed size exclusion chromatography with purified amidinotransferase from porcine kidney, we found peaks corresponding to the molecular weight of monomeric and dimeric enzyme, suggesting an equilibrium between the monomeric and dimeric forms, shifted towards the side of the monomer. In the bipyramidal, tetragonal crystals of recombinant human AT, which belonged to space group $P4_32_12$, we found one monomer of AT in the asymmetric unit. A dimer of amidinotransferase molecules was built up from two monomers of AT by the crystallographic 2-fold axis (Figure 5).

The interface of the crystallographic AT dimer is formed by amino acid residues 151 and 153–157 of the first module around β-strand S3 and between residues 190, 191, 193, 194, 197, 201, 204, 205, 209–212, 214 and 217 of the second module (Figure 5). This segment contains β-strand S7 and flanking coil regions on both sides of this

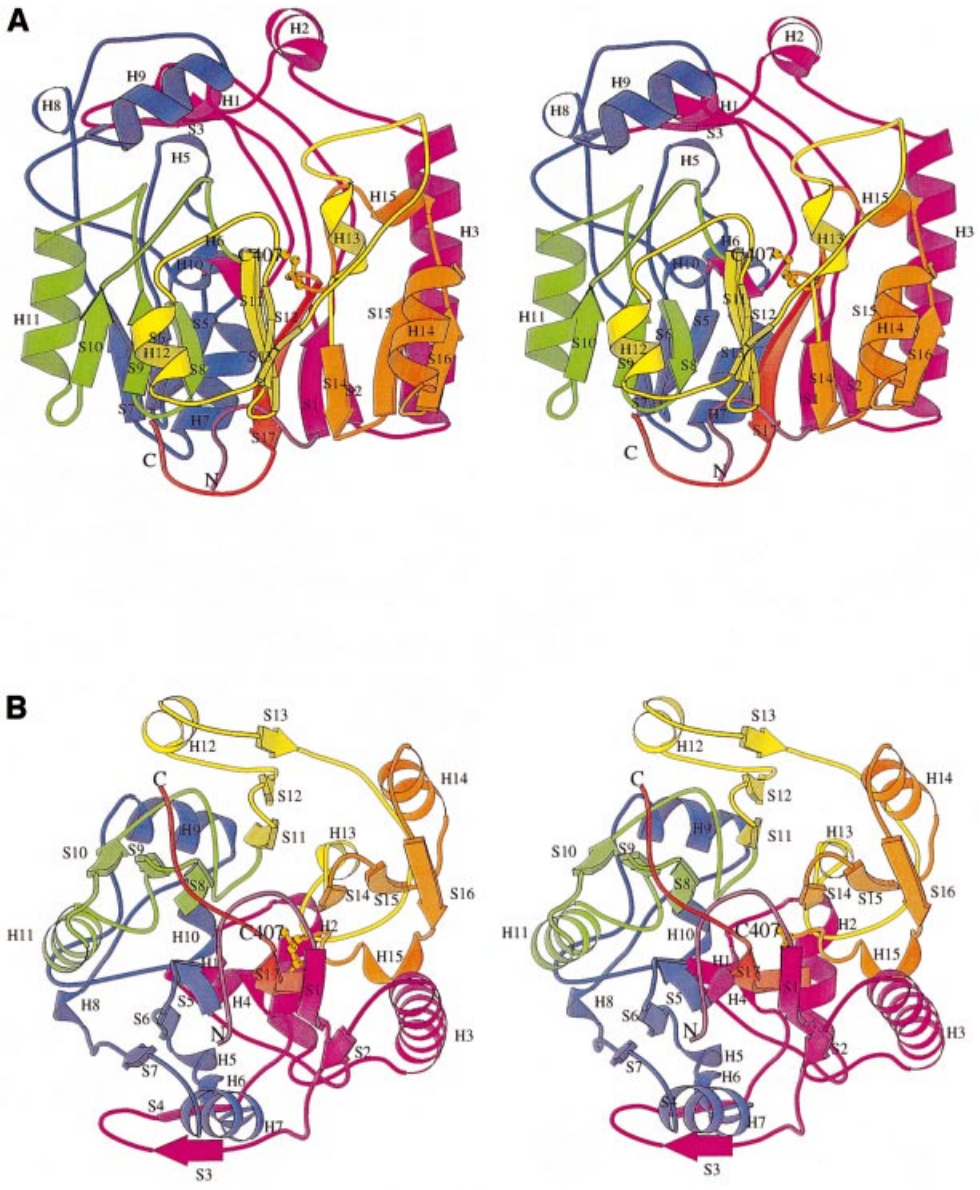

Fig. 2. (**A**) Stereo view of AT in standard orientation with the handles of the basket at the top of the model. The different modules are depicted in different colours (modules I, II, III, IV and V are presented in purple, blue, green, yellow and orange, respectively; the C-terminal β-strand completing module I is shown in red) and the active site cysteine residue is shown in a ball-and-stick representation. Helices are numbered with H_n, strands with S_n. The secondary structure elements were determined with the program DSSP (Kabsch and Sander, 1983) and the figure was produced with MOLSCRIPT (Kraulis, 1991). (**B**) Stereo view of the AT model from the bottom of the basket, revealing the 5-fold pseudosymmetry of the model. The figure was designed with MOLSCRIPT (Kraulis, 1991), the colours are the same as in (A).

strand. The residues W280, R283 and H284 from helix H11 of module III complete the dimer interface. Interactions by aromatic residues provide a major contribution for the stabilization of the dimer; 9 of the 23 residues of the dimerization interface have aromatic side chains. The 2-fold crystallographic axis, building up this dimer, is approximately parallel to the 5-fold pseudosymmetry axis in the centre of the AT basket. The interface between the two AT monomers is rather planar when viewed along the 2-fold symmetry axis.

A calculation of the solvent-accessible surface area (A_s) with XPLOR (Brunger *et al.*, 1987) yielded 25 365 Å^2 for two AT monomers and 21 820 Å^2 for the crystallographic AT dimer. The interface area of 1773 Å^2 is in the range for dimeric proteins of the size of AT according to Janin *et al.* (1988).

Enzyme–substrate interactions

The structure of AT was solved and refined (Tables I and II) with and without β-mercaptoethanol (β-ME) (Figure 6A), which originated from the buffer solution and was found covalently attached to the active site cysteine residue C407 in the initial structure. Subsequently, the structure was determined with L-ornithine bound to the active site channel of the native enzyme (Figure 6B) and with L-arginine bound to the active site channel of the inactive mutant C407A (Figure 6C).

The enzyme exists in a closed conformation when no

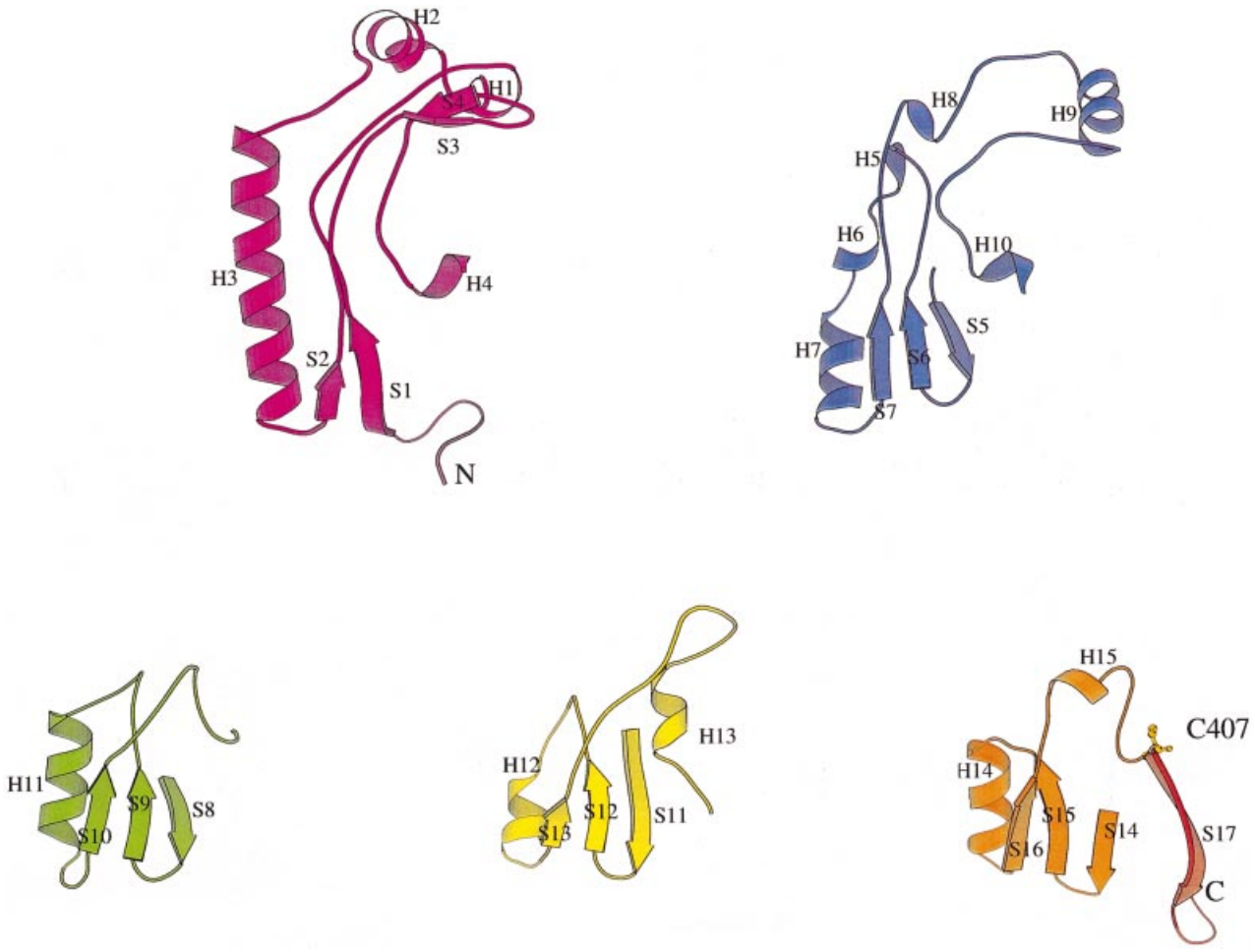

Fig. 3. Representation of the individual modules of the AT structure. The figure was made with MOLSCRIPT (Kraulis, 1991), the colours are the same as in Figure 2A.

substrate is bound. Upon binding of L-ornithine and L-arginine, the enzyme adopts an open conformation. In the open conformation the H9 helix of module II turns back ~3.5 Å and the loop containing residues 298–302 of module III moves aside by 5.0 Å for the α-carbon atom of P299 (Figure 7A and B). Both movements open the channel entrance and may ease the dissociation of reaction products from the channel. The ligands L-ornithine and L-arginine become bound tightly to the residues R322, S354 and S355 with their carboxylate groups forming hydrogen bonds to R322, S354 and S355. The α-amino groups of arginine and ornithine form hydrogen bonds to the main chain carbonyl of M302 and to a water molecule which is further hydrogen-bonded by N98. The ε-amino group of ornithine is hydrogen-bonded to C407, H303 and a water molecule, which is bound to the main-chain carbonyl of G402. The ε-imino group of the arginine substrate is hydrogen-bonded to D305 and weakly to H303. The amino nitrogen atoms of the amidino group are hydrogen-bonded to D170, H303 and D305 (Figure 8). In summary, L-arginine appears to form 11 hydrogen bonds (shorter than 3.2 Å) to amino acid residues of the protein molecule and further hydrogen bonds to water molecules in the active site channel, which are hydrogen-bonded to other residues as well. Alongside the bound arginine ligand there is a chain of nine well-defined water molecules leading out of the channel. This chain of water molecules presumably provides an exit mechanism for water molecules which are displaced upon binding of the ligand.

Active site structure and proposed reaction mechanism

The active site of AT is composed of residues C407, H303 and D254 (Figure 8). The amidino-carbon atom of the arginine substrate is placed between the thiol group of the cysteine residue and the δ1 imino group of the histidine residue. D254 is hydrogen-bonded to the ε2 nitrogen atom of the H303. The arginine is fixed tightly to the active site by formation of hydrogen bonds between the amidino nitrogen atoms and the carboxyl groups of aspartate residues 170 and 305. These orient the substrate's guanidino group such that the sulfur atom of C407, the guanidino-carbon atom and the δ1 nitrogen atom of H303 are in a line orthogonal to the plane of the guanidino group (Figure 8).

On the basis of the structural data a reaction scheme is proposed, in which, during a concerted reaction, the thiol group of C407 adds to the carbon atom of the guanidino group and donates its proton to the ε-imino group of the arginine substrate. A tetrahedral reaction intermediate is formed at this stage (Figure 9). Hydrogen bonds and the

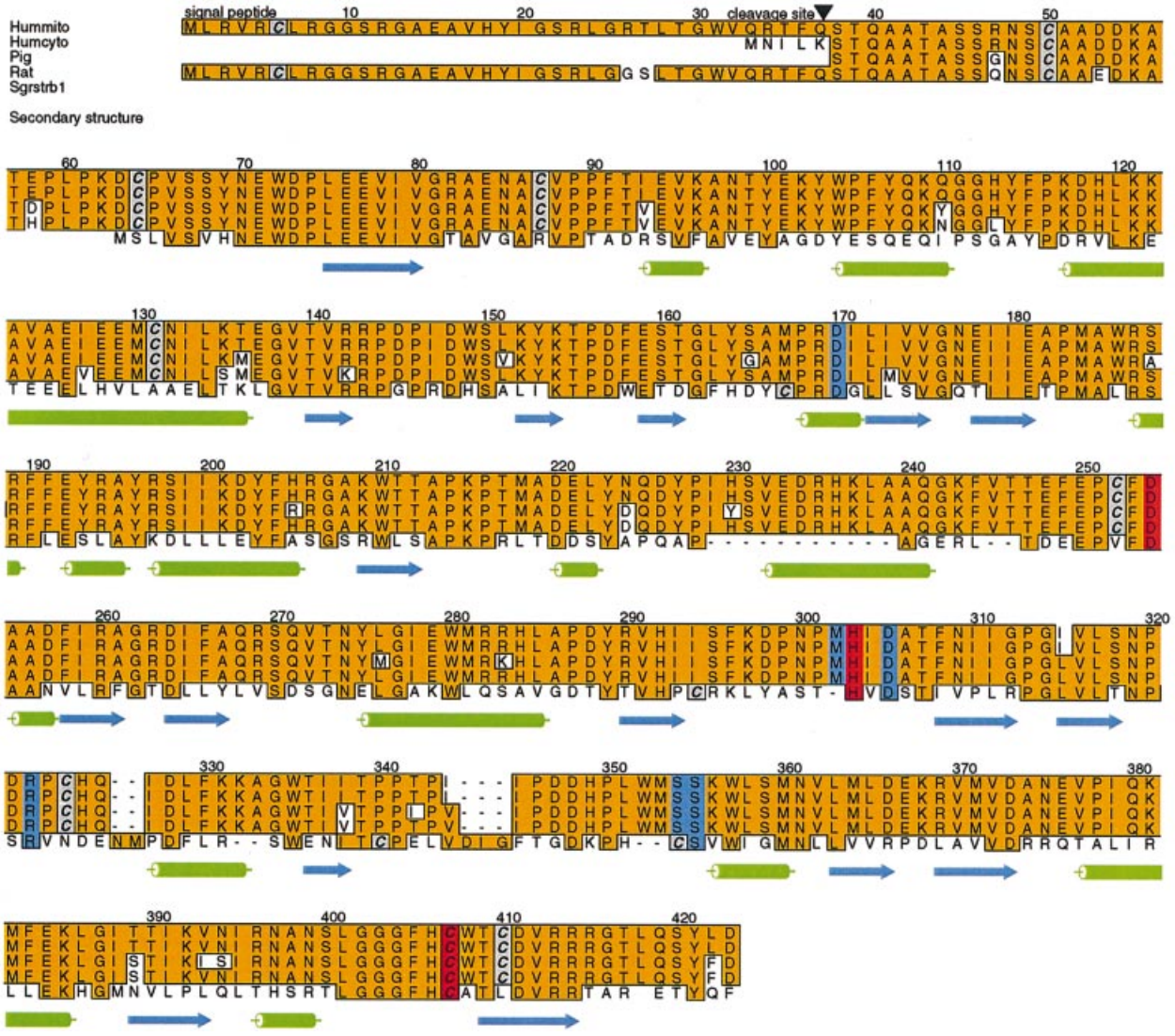

Fig. 4. Multiple sequence alignment of ATs and StrB1 protein. Abbreviations are Hummito (human mitochondrial AT, DDBJ/EMBL/GenBank accession No. S68805), Humcyto (human cytosolic AT, accession No. X86401), Pig (pig AT), Rat (rat mitochondrial AT, accession No. U07971) and Sgrstrb1 (*Streptomyces griseus* StrB1, accession No. S55493/Y00459), an inosaminphosphate amidinotransferase involved in streptomycin biosynthesis. The most abundant amino acid for each position is shown in orange, residues of the proposed catalytic triad are shown in red, residues forming hydrogen bonds to the arginine substrate are depicted in blue and cysteine residues are shown in grey. The numbering refers to the human mitochondrial sequence. The cleavage site for the signal peptide is indicated by a black triangle. The secondary structure elements determined with DSSP (Kabsch and Sander, 1983) are shown below the alignment with blue arrows for β-sheets and green tubes for α-helices.

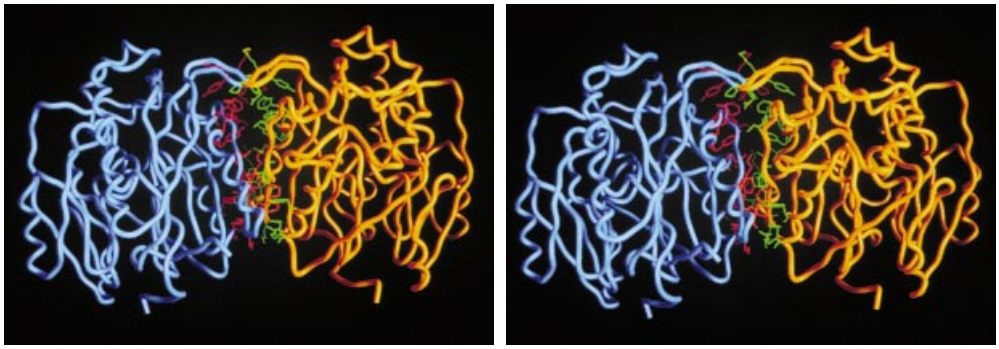

Fig. 5. Stereo image, showing the AT dimer from the side. The residues at the dimer interface are depicted in green and red for the different monomers. The worm plot was produced with GRASP (Nicholls, 1993).

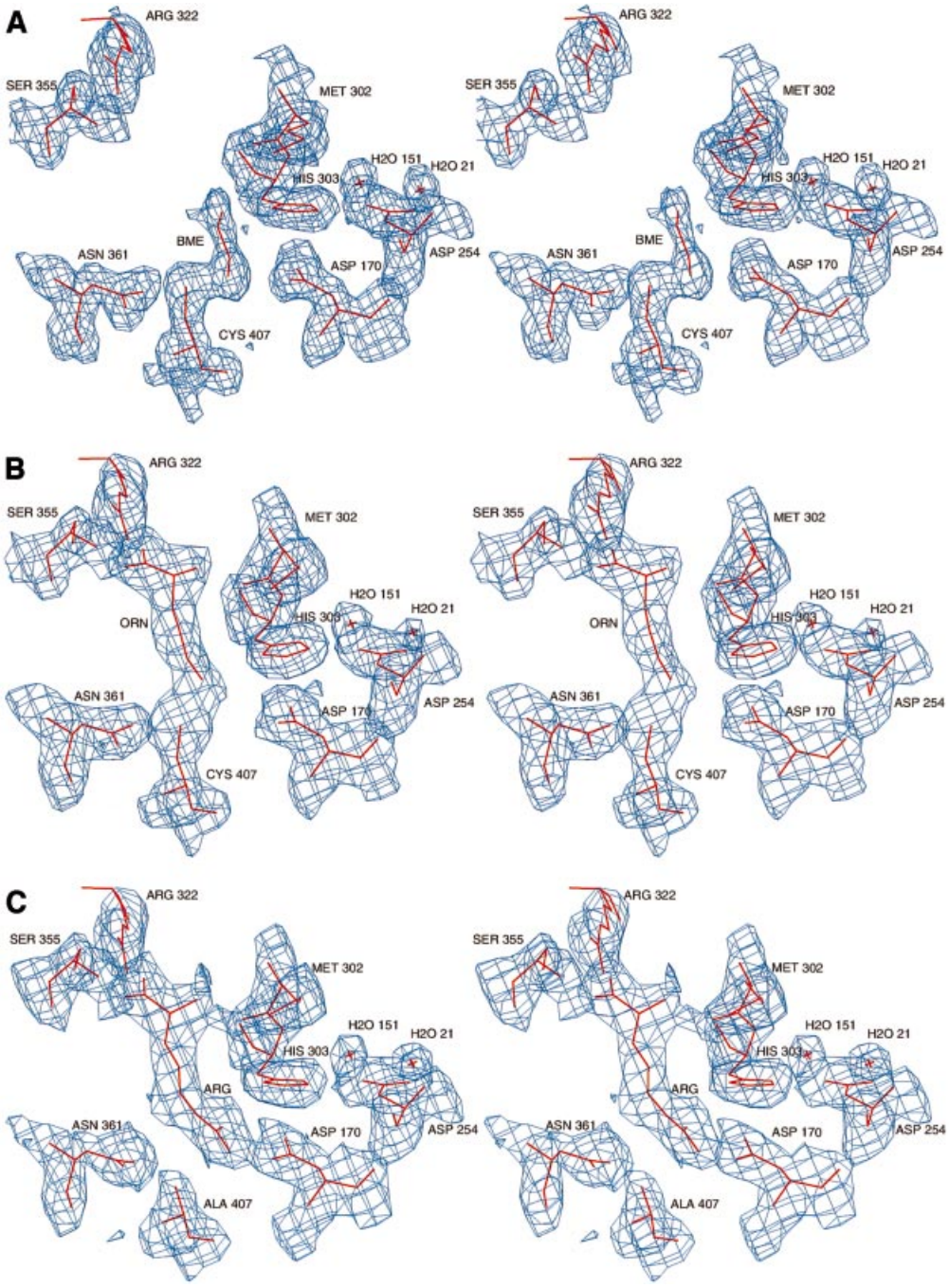

Fig. 6. (**A**) Stereo view of a $2F_o–F_c$ electron density contoured at 1 σ, showing the active site of the enzyme with the β-ME molecule attached to Cys407 at 1.9 Å resolution. (**B**) Stereo view of a $2F_o–F_c$ electron density contoured at 1 σ, showing the active site of the enzyme with an L-ornithine molecule located in the active site channel at 2.4 Å resolution. (**C**) Stereo view of a $2F_o–F_c$ electron density contoured at 1 σ, showing the active site of the enzyme with an L-arginine molecule located in the active site channel of the active site mutant C407A at 2.5 Å resolution. The figure was produced using FRODO (Jones, 1978).

salt linkage of the guanidino nitrogens to D170 and D305 enhance the electrophilicity of the guanidino-carbon atom and a potential hydrogen bond of the ε-imino nitrogen to H303 supports proton transfer to this atom. Subsequently the bond between the ε-imino group and the amidino-carbon atom is broken. The covalent amidino–cysteine product of this reaction is rather stable since the amidino–enzyme complex can be isolated under acidic conditions (Grazi and Rossi, 1968). L-ornithine, the other product of this reaction, is stabilized in the active site channel by hydrogen bonding to H303. L-ornithine binds rather stably to the active site..The K_I for L-ornithine was determined to 253 μM (Sipila, 1980) whereas the K_M for L-arginine lies in the 2.5 mM range (Gross *et al.*, 1986). When ornithine leaves the narrow active site channel, glycine can enter the channel and, in a second half-reaction (Figure 9), bind to the amidino-carbon atom of the amidino–cysteine.

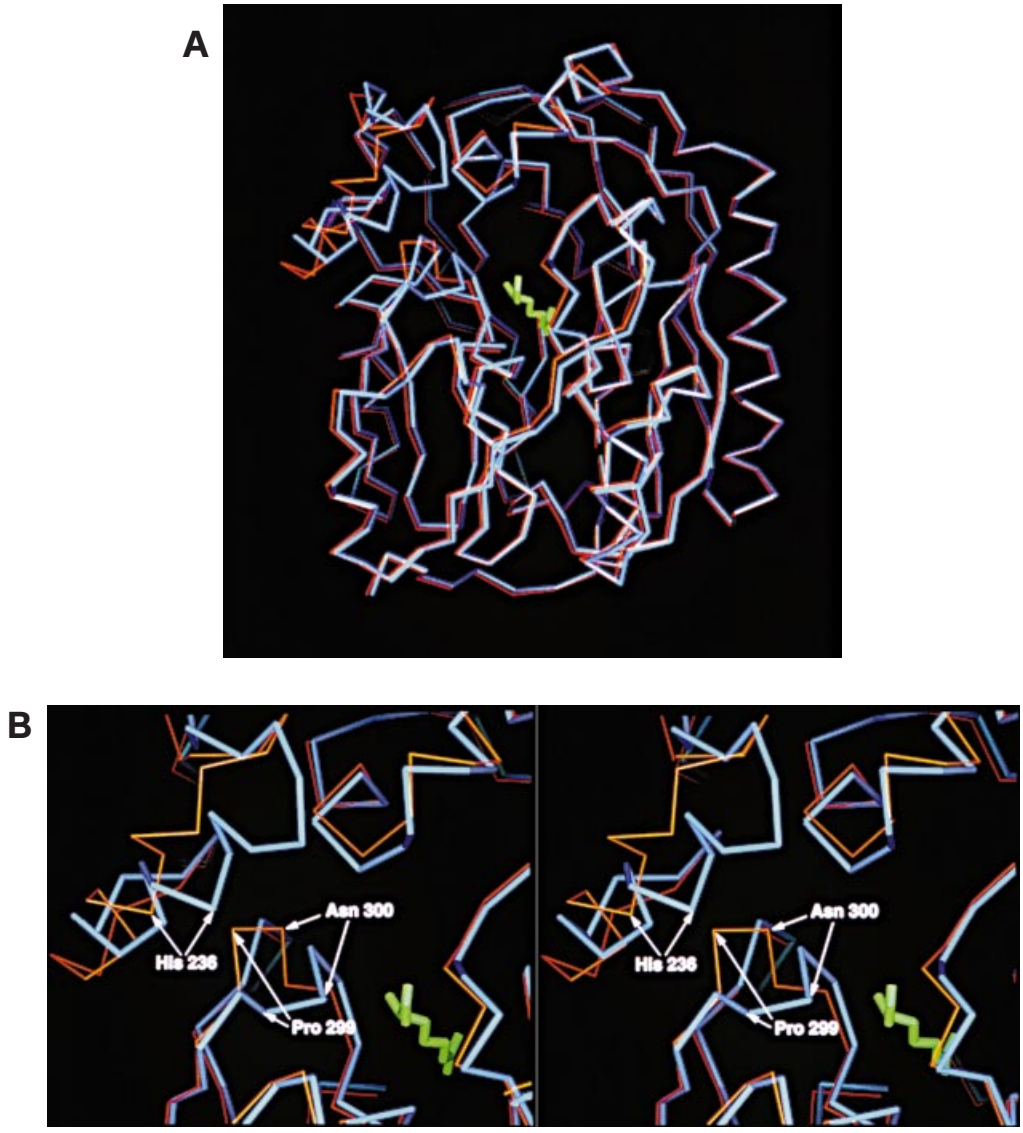

Fig. 7. (**A**) Superposition of Cα-models of AT with β-ME bound to the active site cysteine (blue) and the mutant C407A (red) with L-arginine (green) located in the active site channel. (**B**) Zoom out showing the rearrangements upon binding of L-arginine to the active site. The figure was produced using MAIN (Turk, 1992).

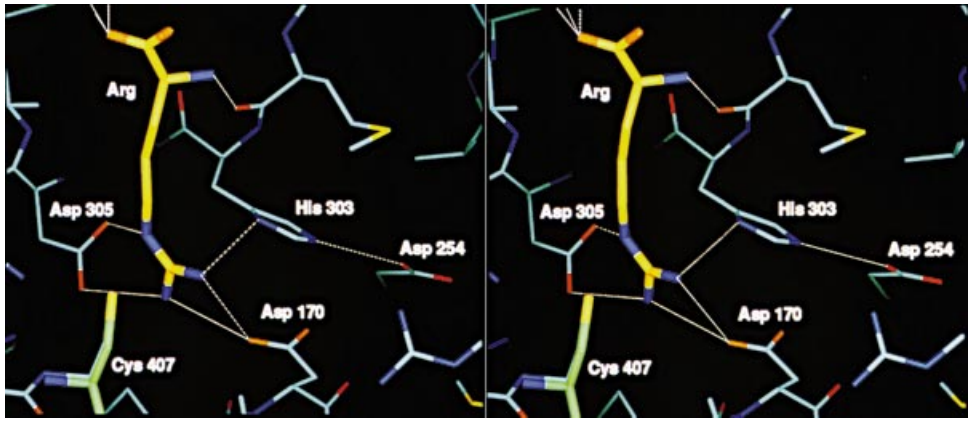

Fig. 8. Stereo view of L-arginine in the active site channel of the mutant C407A. The catalytically active cysteine residue 407 (green) is modelled into the active site model of the mutant C407A by adding the coordinates to the coordinates of C407A. Some hydrogen bonds (bond lengths <3.2 Å) are depicted in white dotted lines. The residues C407, H303 and D254 form the proposed catalytic triad of the enzyme. This image was produced with MAIN (Turk, 1992).

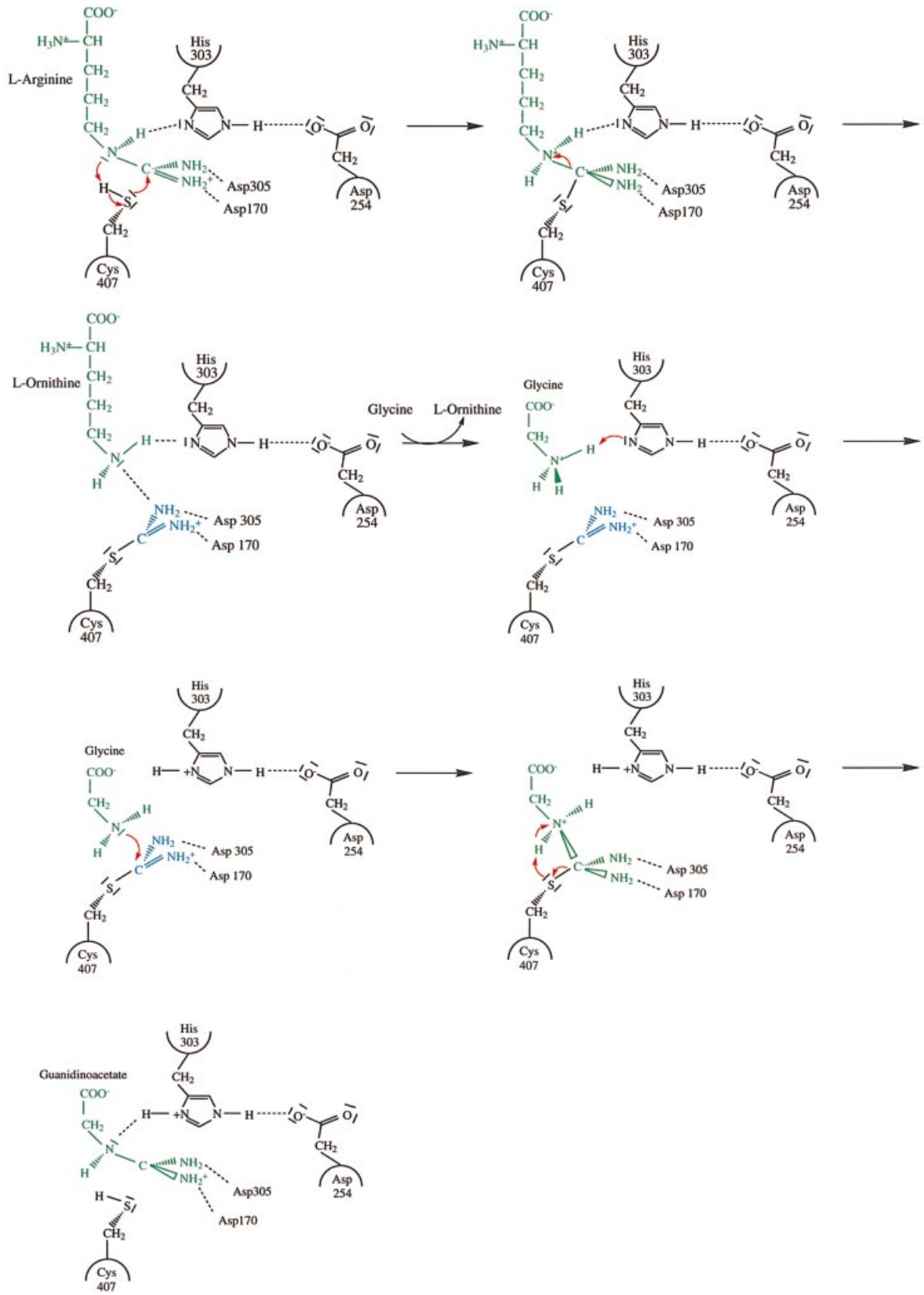

Fig. 9. Transamidination reaction as proposed in the text. Amino acid residues of the protein molecule are shown in black, the substrate in green and the amidino group after its transfer to the Cys407 residue in blue.

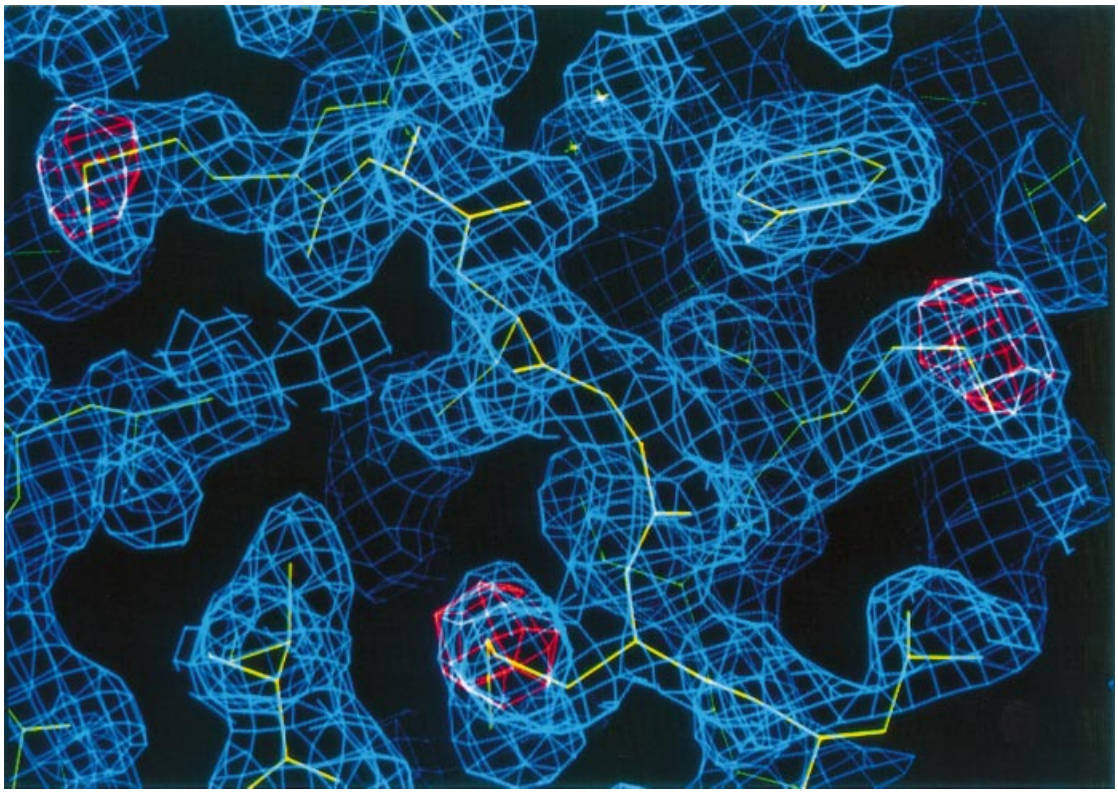

Fig. 10. Superposition of a $2F_o$–F_c electron density from AT at a resolution of 1.9 Å, contoured at 1 σ (blue), with an F_o–F_c electron difference density of AT, expressed with *N*-acetyltelluromethionine, at a resolution of 3.0 Å, contoured at 5 σ (red). The image displays the positions of three tellurium atoms revealing the methionine residues. The figure was produced using FRODO (Jones, 1978).

The reaction may start with a proton transfer from the positively charged glycine substrate to H303 and a nucleophilic attack of the lone electron pair of the glycine nitrogen atom at the carbon atom of the bound amidino group. In the next steps formation of a tetrahedral adduct and its collapse by cleavage of the amidino–carbon–sulfur bond occurs. The thiol group of C407 is regenerated and the reaction product, guanidinoacetic acid, leaves the active site channel to complete the reaction cycle. The transamidination step of the second half-reaction is speculative since no structure of an AT glycine or AT guanidinoacetic acid complex is available.

For the small amidino-acceptor molecule glycine, a direct transfer reaction from L-arginine to the acceptor molecule may be possible since the acceptor and donor molecules could be present in the active site channel at the same time, as simple modelling shows. However, AT also catalyses the transfer of an amidino group from L-canavanine to L-ornithine, from guanidinoacetic acid to L-ornithine and from L-arginine to L-canaline (Ratner and Rochovansky, 1956; Walker, 1956) for which there is insufficient space in the active site channel to be bound at the same time.

Discussion

The crystal structure of recombinant mitochondrial AT from human kidney was solved by MIR. AT containing telluromethionine (TeMet) instead of methionine turned out to be a valuable derivative for assigning the amino acid sequence to the initial polyalanine model (Figure 10 and Table I).

The crystallographic model contains the amino acid residues 64–423 and thus omits the first 26 residues of the mature form of the enzyme (AT38) as well as the 14 amino acids of the fusion peptide. Besides this form of the enzyme, a second, shorter form of mitochondrial AT, starting at the amino acid corresponding to codon 56 of the cDNA of the mitochondrial human enzyme, was found in pig kidney (Humm *et al.*, 1994). The eight amino acid residues (56–63) missing in the model of the enzyme structure compared with the shortest form of the enzyme, which proved to be fully active (AT56) (Humm *et al.*, 1997), point out from the bottom surface of the basket-like structure into the solvent (Figure 2A). We believe that these residues do not play a critical role in the enzymatic activity of AT because they are located far away from the entrance to the active site channel.

The polypeptide chain of AT folds into a compact single domain composed of repeating structural units, the five ββαβ-modules (Figure 3), which surround the central active site compartment in a 5-fold pseudosymmetric manner (Figure 2B). In addition to the characteristic secondary structure elements, the modules contain insertions of variable length which connect different modules and form the handles of the basket-like structure (Figure 2A). Human AT is the only representative of this structural fold. Although the module structures are very similar, no significant homology between the amino acid sequences of different modules could be detected, leaving issues of divergent versus convergent evolution and gene multiplication open.

In the AT crystals which belonged to space group $P4_32_12$ one protein molecule was found in the asymmetric

Table I. Data collection statistics

Derivative	No. of measurements	No. of unique reflections	Resolution (Å)	Completeness overall/ last shell (%/%)	R_m (%)	R_f (%)	No. of sites	Phasing power (25.0–3.0 Å)
NATI 1	26 959	15 648	2.9	91.4/60.2	7.9	3.86		
NATI 2	195 790	51 924	1.9	91.7/67.4	6.6	3.64		
OS6C	24 401	14 291	3.0	89.4/36.2	7.9	4.89	2	0.44
TPPT	16 363	10 139	3.1	77.8/22.6	9.7	6.30	1	0.31
THIO	46 955	15 876	3.0	80.2/31.8	12.4	6.36	1	0.50
HGAC	53 954	19 534	2.5	80.6/76.8	9.9	5.46	1	0.22
TEME	11 660	6 546	3.7	77.4/32.2	17.2	10.53	10	0.76

Overall figure of merit (25.0–3 Å) = 0.28.
Derivative soaking conditions: HGAC, Hg(II) acetate 0.25 mM, 15 h and afterwards 5 mM, 1 day; OS6C, ammoniumhexachloroosmate (IV) 2 mM, 2 days; THIO, thimerosal/cysteamine (1.2 mM/4.8 mM), 5.5 h; TPPT, platinum (II)-(2,2′-6,2″-terpyridinium) chloride cocrystallized in a molar ratio of two terpyridiniumplatinum molecules per protein molecule. TEME, dataset of TeMet containing AT crystals.

$R_m = \sum|I_h - \langle I_h\rangle| / \sum\langle I_h\rangle$; R_F:R_m after independent averaging of Friedel pairs.

Phasing power: $\langle F_H\rangle$/E, where $\langle F_H\rangle = \sum(f_h^2/n)^{1/2}$ is the r.m.s. heavy-atom structure factor amplitude.

$E = \sum[(F_{PHC} - F_{PH})^2/n]^{1/2}$ is the residual lack of closure error with F_{PH} being the structure factor amplitude and $F_{PHC} = |\mathbf{F}_P + \mathbf{f}_H|$ the calculated structure factor amplitude of the derivative.

Table II. Refinement parameters

Structure	NATI2	NATI3	Ornithine	C407AARG	TEM3
Resolution range (Å)	8.0–1.9	8.0–2.1	8.0–2.4	8.0–2.5	8.0–3.0
No. of unique reflections	51 924	37 316	24 010	22 282	12 651
Protein atoms (excluding H)	2 940	2 940	2 940	2 940	2 940
Ligand atoms (excluding H)	4	–	8	11	4
Solvent atoms (excluding H)	195	195	193	194	195
R-factor[a]	0.196	0.199	0.181	0.183	0.251[b]
Free *R*-factor	0.231				
R.m.s. deviations from target values:					
bonds (Å)	0.010	0.011	0.011	0.012	
angles (°)	1.519	1.601	1.620	1.684	
Average B-factor (Å^2) of:					
main chain atoms	23.01	24.68	26.30	22.33	
side chain atoms	26.83	27.70	28.88	24.85	
solvent atoms	36.43	37.87	43.11	40.36	

NATI2, data set of AT with β-ME bound to the active site cysteine; NATI3, data set with free C407 residue; ornithine, data set with L-ornithine bound to the active site, C407AARG, data set with L-arginine bound to the active site of the mutant C407A and TEM3, data set to 3 Å resolution of AT containing telluromethionine.

[a] $R = (\sum|F_o - F_c|)/\sum F_o$.

[b] Model NATI2 without refinement.

unit and an AT-dimer was built up from the monomer by the 2-fold crystallographic axis (Figure 5). The crystallographic dimer is stabilized largely by aromatic interactions at the dimer interface and is thought to represent the naturally occurring AT-dimer, as recorded by sedimentation-equilibrium experiments (Gross *et al.*, 1986), for two reasons. On the one hand, the dimer interface lies at the backside of the monomer, thus dimerization would not affect the accessibility of the active site channel for the substrates. It is formed by interactions between amino acid residues of modules I, II and III. All residue movements during binding of the substrate L-arginine or the product inhibitor L-ornithine occur either at helix H9 of module II, which is one of the handle helices, or at residues 298–302 of module III, which coat the upper part of the active site channel. Thus, the dimerization-interface is not affected by substrate binding to the active site channel. On the other hand, the bottom side of the AT basket, where the N-terminus of the protein protrudes, is unaffected by dimerization and thus the N-terminal residues of the polypeptide chain would possibly be free to anchor the mitochondrial AT to the inner mitochondrial membrane. Mitochondrial AT was localized on the outer surface of the inner mitochondrial membrane (Magri *et al.*, 1975) and the only difference in the sequences of cytoplasmic and mature mitochondrial amidinotransferases was found to be an additional five amino acid residues at the N-terminus of the cytosolic enzyme (Humm *et al.*, 1997).

Upon binding of large ligands (L-arginine or L-ornithine) to the active site channel, conformational changes occur at helix H9 and at residues 298–302, opening the channel entrance slightly with respect to the structure of AT (Figure 7). These ligands form five hydrogen bonds between the

carboxy group of the ligand molecule and the residues R322, S354 and S355 and an additional hydrogen bond between the ligand amino group and the carbonyl oxygen atom of M302. All these hydrogen bonds, which might trigger the conformational changes at the channel entrance to ease the release of reaction products, can only be formed when the ligand is bound. The active site channel is deep enough to allow discrimination of substrates which fit exactly into the channel and reach the active site cysteine residue from e.g. arginine side chains at the surface of protein molecules, which cannot enter deep enough into the channel to reach the active site and to be recognized as substrates. For correct orientation of the substrates, binding of the carboxylate group is necessary.

During the transfer of the amidino group, a covalent intermediate with a reactive cysteine was demonstrated (Grazi and Rossi, 1968) that can be stabilized under acidic conditions. Recently, this active site cysteine residue was identified biochemically as C407 (Humm *et al.*, 1997). The proposed reaction mechanism (Figure 9) begins with a nucleophilic attack of the sulfur atom of C407 onto the carbon atom of the guanidino group and proton transfer from C407 to the ε-imino group to form a tetrahedral reaction intermediate. It continues by cleavage of the bond between the ε-imino group and the amidino-carbon atom. The amidino group at C407 collapses to a planar isothiourea group and the reaction product L-ornithine diffuses out of the active site channel.

The second part of the reaction might be a reversal of the arginine cleavage reaction as described in Figure 9. There is space for a glycine acceptor when the arginine donor is present and direct transamidination might be envisaged, but could not explain the formation of amidinocysteine and would be incompatible with the features of arrangement of donor and active site residues. In addition, AT catalyses the transfer of amidino groups to large acceptors (Ratner and Rochovansky, 1956; Walker, 1956) which cannot be bound simultaneously. Therefore our findings support a ping-pong mechanism for this enzyme as proposed earlier (Walker, 1973).

Different mechanisms for the enzymatic activation and conversion of the stable guanidino and carbamoylamino groups have been found in different systems. The three-dimensional structures of human AT, creatinase from *Pseudomonas putida* (Hoeffken *et al.*, 1988), *N*-carbamoylsarcosine amidohydrolase from *Arthrobacter sp*. (Romao *et al.*, 1992), urease from *Klebsiella aerogenes* (Jabri *et al.*, 1995) and rat liver arginase (Kanyo *et al.*, 1996) show no similarity. The features of the active sites are different with amidinotransferase containing a catalytic triad with an active site cysteine residue, *N*-carbamoylsarcosine amidohydrolase containing an active site cysteine residue but lacking the catalytic triad (Zajc *et al.*, 1996) and creatinase operating with a histidine-activated nucleophilic water molecule (Coll *et al.*, 1990). The catalytic centre of urease contains two nickel ions each bound by two histidine residues (Jabri *et al.*, 1995) and the active site of arginase contains two manganese ions and an activated solvent molecule which bridges the two Mn^{2+} ions. Thus, mechanisms for cleaving of guanidino and carbamoylamino groups and transfer to different acceptor molecules have been developed several times in the course of evolution.

Materials and methods

Recombinant human mitochondrial kidney amidinotransferase was prepared as described in Humm *et al.* (1997). Briefly, AT was expressed as a fusion protein with an N-terminal histidine-tag in *E.coli* BL21(DE3)-pLysS. After harvesting and lysis of the cells, the crude extract was centrifuged and applied to an Ni^{2+}-NTA–agarose column (Qiagen GmbH, Hilden, Germany). The fusion protein was eluted with a gradient of 180–700 mM imidazole in 300 mM NaCl, 2 mM β-ME, 30 mM HEPES, pH 6.2. It was dialysed against 30 mM HEPES, 2 mM β-ME, pH 8.0, concentrated by ultrafiltration and stored at 4°C. Presumably due to toxicity of active amidinotransferase to *E.coli* by consumption of L-arginine and glycine, only 2–5 mg of AT per litre of culture volume were obtained. Enzyme activity and protein concentration were determined according to Van Pilsum *et al.* (1970) and Lowry *et al.* (1951).

AT containing TeMet was expressed in the methionine-auxotrophic *E.coli* strain B834(DE3)pLysS (Budisa *et al.*, 1995). The *E.coli* cells were grown at 37°C in minimal medium with a limiting amount of methionine. Cell growth stopped at an OD_{600} of 0.7, presumably because methionine was exhausted. The cells of 11 l *E.coli* culture were harvested by centrifugation and resuspended in 1 l of fresh minimal medium. Expression of AT was induced by addition of 1 mM IPTG and 50 mg/l *N*-acetyltelluromethionine (Karnbrock *et al.*, 1996). The cells were cultivated for 4 h at 25°C. Telluromethionine–amidinotransferase was isolated by affinity chromatography as the native fusion protein with a yield of 0.4 mg protein per litre of initial culture.

Crystallization experiments were performed at 22°C employing the vapour diffusion technique. Hanging or sitting droplets were made by mixing 2–10 µl of protein solution (13–15 mg/ml) and twice the volume of precipitating buffer (3% PEG 6000, 40 mM HEPES, 2 mM β-ME, pH 7.0). Crystals appeared after 3–10 days but usually remained too small for X-ray analysis. Crystals with dimensions up to 0.8×0.4×0.4 mm^3 were obtained by macroseeding. The crystals showed a tetragonal, bipyramidal shape and belonged to space group $P4_32_12$ with lattice constants $a = b = 83.6$ Å, $c = 200.4$ Å and $\alpha = \beta = \gamma = 90°$.

Determination of the crystal density by volume measurement and amino acid analysis (Kiefersauer *et al.*, 1996) yielded in two measurements 7.24 and 9.35 molecules per unit cell, corresponding to one amidinotransferase molecule per asymmetric unit and a solvent content of 65% ($V_m = 3.8$ Å^3/Da) (Matthews, 1968). Crystals were well ordered and diffracted to 1.9 Å resolution using a synchrotron X-ray source (beamline BW6 at the DESY, Hamburg, Germany). All data, apart from one high-resolution data set collected at the synchrotron in Hamburg with an X-ray beam of $\lambda = 0.90$ Å wavelength on a MAR Research imaging plate detector, were collected on a detector of the same kind with graphite-monochromatized Cu-Kα radiation from a Rigaku RU200 rotating anode generator (Rigaku, Tokyo, Japan). The generator was operated at 5.4 kW with an apparent focal spot size of 0.3×0.3 mm^2. Data were collected in frames of 1.0°, processed with MOSFLM (version 5.23) (Leslie, 1994) and scaled and reduced using ROTAVATA/AGROVATA of the CCP4 package (Collaborative Computational Project, 1994).

Heavy-atom derivatives were prepared by soaking the crystals in heavy-atom solutions for 2 h to 2 days or by cocrystallization of the protein with the heavy-atom compound (Table I). Interpretation of heavy-atom derivatives OS6C, TPPT, THIO and HGAC was done by vector verification with the PROTEIN program package (Steigemann, 1974). Cross-phased-difference Fourier maps were used to confirm heavy-atom positions and to establish a common origin for the derivatives. Anomalous data were included to identify the enantiomorphous space group with the program SIGAT of the PROTEIN package. Phases were calculated with MLPHARE and improved by solvent flattening and histogram matching with the program DM of the CCP4 package. An initial electron density map calculated at 3 Å resolution showed some secondary structure elements, which were fitted as poly(A). The electron density was improved by combining model and MIR phases in iterative cycles, finally yielding a model of poly(A) fragments comprising ~280 out of 360 residues. Hardly any connectivities and side chains electron density was visible. At this stage the derivative TEME became available. Ten Te positions in the derivative TEME with data to 3.7 Å resolution were identified by visual inspection of a difference Fourier at 5 Å resolution phased with model combined MIR phases. Phasing by MIR with the heavy-atom derivatives including the derivative TEME (Table I) yielded a significantly improved electron density map. The tellurium positions were used to assign the amino acid sequence and to identify connectivities. Model building was performed with FRODO on an ESV-30 Graphic system workstation (Evans & Sutherland, Salt Lake City, UT, USA) and

refinement was done with XPLOR (Brunger *et al.*, 1987) using the parameter set of Engh and Huber (1991). After six rounds of model building and refinement, residues 64–423 had been placed into the model (the numbering refers to the cDNA-derived sequence of mitochondrial AT). The current model comprises 2940 protein atoms (excluding H-atoms) and 195 water molecules. Additional electron density at the sulfur atom of the active site Cys407 was interpreted as a covalently bound β-ME molecule originating from the buffer solution. The crystallographic *R*-factor was 19.6% for 51 924 reflections from 8.0–1.9 Å resolution. The r.m.s. deviations from ideal stereochemistry were 0.010 Å for bond lengths and 1.52° for bond angles (Table II). All residues apart from Met184, Asp298 and His303, which are well defined in the electron density, are in allowed regions of the Ramachandran diagram.

A second data set (TEM3) from a TeMet–AT crystal which diffracted up to 3.0 Å was collected (Table II and Figure 10). Soaking of crystals with substrates and inhibitors was hampered by β-ME bound to the active site cysteine residue. Therefore β-ME was replaced by glutathione as reducing agent, necessary for crystallization. From crystals prepared in this way a data set was collected to 2.1 Å resolution. In the structure of AT crystallized in this way β-ME was missing in the active site and the active site cysteine residue was free. All the rest of the structure was identical to the 1.9 Å AT-structure with β-ME. After soaking of such crystals with L-ornithine, an inhibitor of AT, a data set to 2.4 Å resolution could be measured and the model was refined to a crystallographic *R*-factor of 18.1%. Since soaking of such crystals with L-arginine destroyed the crystals, crystals of the active site mutant C407A (Humm *et al.*, 1997) were soaked with L-arginine. L-arginine was found in the active site and the model was refined to an *R*-factor of 18.3% at 2.5 Å resolution (Table II).

References

Austruy,E., Cohen-Salmon,M., Antignac,C., Beroud,C., Henry,I., Van Cong,N., Brugieres,L., Junien,C. and Jeanpierre,C. (1993) Isolation of kidney complementary DNAs down-expressed in Wilms' tumor by a subtractive hybridization approach. *Cancer Res.*, **53**, 2888–2894.

Bessman,S.P. (1985) The creatine-creatine phosphate energy shuttle. *Annu. Rev. Biochem.*, **54**, 831–862.

Bessman,S.P. and Geiger,P.J. (1981) Transport of energy in muscle: the phosphoryl-creatine shuttle. *Science*, **211**, 448–452.

Brunger,A.T., Kuriyan,J. and Karplus,M. (1987) Crystallographic R factor refinement by molecular dynamics. *Science*, **35**, 458–460.

Budisa,N., Steipe,B., Demange,P., Eckerskorn,C., Kellermann,J. and Huber,R. (1995) High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eur. J. Biochem.*, **230**, 788–796.

Coll,M., Knof,S.H., Ohga,Y., Messerschmidt,A., Huber,R., Moellering,H., Russmann,L. and Schumacher,G. (1990) Enzymatic mechanism of creatine amidinohydrolase as deduced from crystal structures. *J. Mol. Biol.*, **214**, 597–610.

Collaborative computational project, Number 4. (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.

Engh,R.A. and Huber,R. (1991) Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr*, **A47**, 392–400.

Fritz-Wolf,K., Schnyder,T., Wallimann,T. and Kabsch,W. (1996) Structure of mitochondrial creatine kinase. *Nature*, **381**, 341–345.

Grazi,E. and Rossi,N. (1968) Transamidinase of hog kidney. *J. Biol. Chem.*, **243**, 538–542.

Gross,M.D., Eggen,M.A., Simon,A.M. and Van Pilsum,J.F. (1986) The purification and characterization of human kidney L-arginine:glycine amidinotransferase. *Arch. Biochem. Biophys.*, **251**, 747–755.

Guimbal,C. and Kilimann,M.W. (1993) A Na^+-dependent creatine transporter in rabbit brain, muscle, heart, and kidney. *J. Biol. Chem.*, **268**, 8418–8421.

Guthmiller,P., Van Pilsum,J.F., Boen,J.R. and McGuire,D.M. (1994) Cloning and sequencing of rat kidney L-arginine:glycine amidinotransferase. *J. Biol. Chem.*, **269**, 17556–17560.

Hoeffken,H.W., Knof,S.H., Bartlett,P.A., Huber,R., Moellering,H. and Schumacher,G. (1988) Crystal structure determination, refinement and molecular model of creatine amidinohydrolase from *Pseudomonas putida*. *J. Mol. Biol.*, **204**, 417–433.

Humm,A., Huber,R. and Mann,K. (1994) The amino acid sequences of human and pig L-arginine:glycine amidinotransferase. *FEBS Lett.*, **339**, 101–107.

Humm,A., Fritsche,E., Mann,K., Gohl,U. and Huber,R. (1997) Recombinant expression and isolation of human L-arginine:glycine amidinotransferase and identification of its active site cysteine residue. *Biochem. J.*, **322**, 771–776.

Jabri,E., Carr,M.B., Hausinger,R.P. and Karplus,P.A. (1995) The crystal structure of urease from *Klebsiella aerogenes*. *Science*, **268**, 998–1004.

Janin,J., Miller,S. and Chothia,C. (1988) Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.*, **204**, 155–164.

Jones,T.A. (1978) A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.*, **11**, 268–272.

Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.

Kanyo,Z.F., Scolnick,L.R., Ash,D.E. and Christianson,D.W. (1996) Structure of a unique binuclear manganese cluster in arginase. *Nature*, **383**, 554–557.

Karnbrock,W., Weyher,E., Budisa,N., Huber,R. and Moroder,L. (1996) A new efficient synthesis of acetyltelluro- and acetylselenomethionine and their use in the biosynthesis of heavy-atom protein analogs. *J. Am. Chem. Soc.*, **118**, 913–914.

Kiefersauer,R., Stetefeld,J., Gomis-Ruth,F.X., Romao,M.J., Lottspeich,F. and Huber,R. (1996) Protein-crystal density by volume measurement and amino-acid analysis. *J. Appl. Crystallogr.*, **29**, 311–317.

Kraulis,P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Ku,C.-P. and Passow,H. (1980) Creatine and creatinine transport in old and young human red blood cells. *Biochim. Biophys. Acta*, **600**, 212–227.

Leslie,A.G.W. (1994) Mosflm user guide, mosflm version 5.20. MRC Laboratory of Molecular Biology, Cambridge, UK.

Lowry,O.H., Rosebrough,N.J., Farr,A.L. and Randall,R.J. (1951) Protein measurement with the folin phenol reagent. *J. Biol. Chem.*, **193**, 265–275.

Magri,E., Balboni,G. and Grazi,E. (1975) On the biosynthesis of creatine. Intramitochondrial localization of transamidinase from rat kidney. *FEBS Lett.*, **55**, 91–93.

Matthews,B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.*, **33**, 491–497.

McGuire,D.M., Tormanen,C.D., Segal,I.S. and Van Pilsum,J.F. (1980) The effect of growth hormone and thyroxine on the amount of L-arginine:glycine amidinotransferase in kidneys of hypophysectomized rats. *J. Biol. Chem.*, **255**, 1152–1159.

McGuire,D.M., Gross,M.D., Van Pilsum,J.F. and Towle,H.C. (1984) Repression of rat kidney L-arginine:glycine amidinotransferase synthesis by creatine at a pretranslational level. *J. Biol. Chem.*, **259**, 12034–12038.

Nicholls,A., Bharadwaj,R. and Honig,B. (1993) GRASP—graphical representation and analysis of surface properties. *Biophys. J.*, **64**, A166.

Ratner,S. and Rochovansky,O. (1956) Biosynthesis of guanidinoacetic acid. II. Mechanism of amidine group transfer. *Arch. Biochem. Biophys.*, **63**, 296–315.

Romao,M.J., Turk,D., Gomis-Ruth,F.-X., Huber,R., Schumacher,G., Mollering,H. and Russmann,L. (1992) Crystal structure analysis, refinement and enzymatic reaction mechanism of N-carbamoylsarcosine amidohydrolase from *Arthrobacter sp.* at 2.0 Å resolution. *J. Mol. Biol.*, **226**, 1111–1130.

Sipila,I. (1980) Inhibition of arginine-glycine amidinotransferase by ornithine. *Biochim. Biophys. Acta*, **613**, 79–84.

Sipila,I., Simell,O., Rapola,J., Sainio,K. and Tuuteri,L. (1979) Gyrate atrophy of the choroid and retina with hyperornithinemia:tubular aggregates and type 2 fiber atrophy in muscle. *Neurology*, **29**, 996–1005.

Sipila,I., Simell,O. and Arjomaa,P. (1980) Gyrate atrophy of the choroid and retina with hyperornithinemia. *J. Clin. Invest.*, **66**, 684–687.

Sora,I., Richman,J., Santoro,G., Wei,H., Wang,Y., Vanderah,T., Horvath,R., Nguyen,M., Waite,S., Roeske,W.R. and Yamamura,H.I. (1994) The cloning and expression of a human creatine transporter. *Biochem. Biophys. Res. Comun.*, **204**, 419–427.

Steigemann,W. (1974) Ph.D. thesis. Technische Universitat Munchen, Germany.

Syllm-Rapoport,I., Daniel,A. and Rapoport,S. (1980) Creatine transport into red blood cells. *Acta Biol. Med. Germ.*, **39**, 771–779.

Turk,D. (1992) Weiterentwicklung eines Programms fur Molekulgraphik und Elektronendichte-Manipulation und seine Anwendung auf verschiedene Proteinstrukturaufklarungen. Ph.D. thesis, Technische Universitat Munchen, Germany.

Van Pilsum,J.F., Taylor,D., Zakis,B. and McCormick,P. (1970) Simplified assay for transamidinase activities of rat kidney homogenates. *Anal. Biochem.*, **35**, 277–286.

Walker,J.B. (1956) Biosynthesis of arginine from canavanine and ornithine in kidney. *J. Biol. Chem.*, **218**, 549–556.

Walker,J.B. (1973) Amidinotransferases. In Boyer,P.D. (ed.), *The Enzymes*, Volume 9, part B, 3rd edn. Academic Press, New York, pp. 497–509.

Walker,J.B. and Hannan,J.K. (1976) Creatine biosynthesis during embryonic development. False feedback suppression of liver amidinotransferase by N-acetimidoylsarcosine and 1-carboxymethyl-2-iminoimidazolidine (cyclocreatine). *Biochemistry,* **15**, 2519–2522.

Walker,J.B. (1979) Creatine: biosynthesis, regulation, and function. *Adv. Enzymol.*, **50**, 177–242.

Wyss,M. and Wallimann,T. (1994) I-4 creatine metabolism and the consequences of creatine depletion in muscle. *Mol. Cell. Biochem.*, **133/134**, 51–66.

Zajc,A., Romao,M.J., Turk,B. and Huber,R. (1996) Crystallographic and fluorescence studies of ligand binding to N-carbamoylsarcosine amidohydrolase from *Arthrobacter sp*. *J. Mol. Biol*. **263**, 269–283.

Received on November 4, 1996; revised on January 31, 1997