Oligomeric structure of proclavaminic acid amidino hydrolase: evolution of a hydrolytic enzyme in clavulanic acid biosynthesis

Jonathan M. ELKINS^{*1}, Ian. J. CLIFTON^{*}, Helena HERNÁNDEZ[†], Linh X. DOAN^{*}, Carol V. ROBINSON[†], Christopher J. SCHOFIELD^{*} and Kirsty S. HEWITSON^{*1,2}

^{*}The Dyson Perrins Laboratory, Oxford Centre for Molecular Sciences, South Parks Road, Oxford OX1 3QY, U.K., and [†]Central Chemistry Laboratory, Oxford Centre for Molecular Sciences, South Parks Road, Oxford OX1 3QT, U.K.

During biosynthesis of the clinically used β -lactamase inhibitor clavulanic acid, one of the three steps catalysed by clavaminic acid synthase is separated from the other two by a step catalysed by proclavaminic acid amidino hydrolase (PAH), in which the guanidino group of an intermediate is hydrolysed to give proclavaminic acid and urea. PAH shows considerable sequence homology with the primary metabolic arginases, which hydrolyse arginine to ornithine and urea, but does not accept arginine as a substrate. Like other members of the bacterial sub-family of arginases, PAH is hexameric in solution and requires Mn^{2+} ions for activity. Other metal ions, including Co^{2+} , can substitute for Mn^{2+} . Two new substrates for PAH were identified, *N*-acetyl-(L)-arginine and (3*R*)-hydroxy-*N*-acetyl-(L)-arginine. Crystal structures of PAH from *Streptomyces clavuligerus* (at 1.75 Å and

2.45 Å resolution, where 1 Å = 0.1 nm) imply how it binds β -lactams rather than the amino acid substrate of the arginases from which it evolved. The structures also suggest how PAH selects for a particular alcohol intermediate in the clavam biosynthesis pathway. As observed for the arginases, each PAH monomer consists of a core of β -strands surrounded by α -helices, and its active site contains a di- Mn^{2+} centre with a bridging water molecule responsible for hydrolytic attack on to the guanidino group of the substrate. Comparison of structures obtained under different conditions reveals different conformations of a flexible loop, which must move to allow substrate binding.

Key words: arginase, manganese enzyme, oligomerization.

INTRODUCTION

The most prevalent mechanism of bacterial resistance to β -lactam antibiotics involves β -lactamase enzymes, which catalyse the hydrolysis of the β -lactam ring to give biologically inactive products [1]. Attempts to overcome the action of β -lactamases have centred upon discovering β -lactam antibiotics resistant to β -lactamase-mediated hydrolysis and discovering efficient inhibitors of β -lactamases [1]. The most widely used β -lactamase inhibitor is clavulanic acid, which is administered in combination with the penicillin amoxycillin. Despite its small size, the lability and density of functionalization in clavulanic acid renders its preparation via chemical synthesis difficult, with no asymmetric synthesis having been reported to date. Thus clavulanic acid is isolated directly from fermentation broths of *Streptomyces clavuligerus*.

Since the chemical synthesis of clavulanic acid and its analogues for medicinal use is impractical, clavam biosynthesis has been of considerable interest. The early stages of the pathway have now been elucidated using a combination of genetic and biochemical techniques (Scheme 1) [2]. The first-formed monocyclic β -lactam, compound 2, is biosynthesized in two enzyme-catalysed steps from arginine and glyceraldehyde-3-phosphate [3] via a β -amino acid, compound 1 (Scheme 1). The latter step involves the β -lactam synthetase-mediated cyclization of a β -amino acid to a monocyclic β -lactam in an almost reverse β -lactamase reaction [3a,4]. This intermediate is then converted in four steps

in the central part of the pathway into clavaminic acid, which undergoes a double epimerization to give the highly labile clav-aldehyde intermediate. This intermediate is subsequently reduced in an NADPH-dependent reaction to clavulanic acid [5]. Proclavaminic acid amidino hydrolase (PAH) together with a 2-oxoglutarate oxygenase, clavaminic acid synthase (CAS), catalyse the central part of the pathway in which the labile bicyclic clavam ring system is constructed from the monocyclic intermediate (compound 2) [6]. In the first CAS-catalysed reaction, compound 2, which has a guanidino side chain, is hydroxylated to give a secondary alcohol, guanidinoproclavaminic acid (compound 3). This intermediate is neither an efficient substrate nor inhibitor of CAS. Instead, PAH [7–9] catalyses the hydrolysis of its guanidino side chain to give proclavaminic acid, the amino side chain of which renders it a second substrate for CAS in an oxidative cyclization process. A subsequent CAS-catalysed desaturation yields clavaminic acid. The role of PAH in the pathway may thus be regarded as ‘mutation’ or ‘engineering’ of an intermediate side chain to allow further involvement of a preinvolved enzyme (i.e. CAS). Structures have been reported for two of the enzymes in the clavam pathway, the β -lactam synthetase [10] and CAS [11].

The position of PAH within the clavulanic acid biosynthetic pathway was first established by workers at Beecham [8], where the enzyme was purified directly from *S. clavuligerus*. PAH has also been isolated by Aidoo et al. [9], but as a ‘contaminating’ protein during the purification of δ -(L- α -aminoadipoyl)-(L)-

Abbreviations used: CAS, clavaminic acid synthase; CNS, crystallography and NMR system; ESI, electrospray ionization; *N*-AA, *N*-acetyl-(L)-arginine; *N*-AA-OH, (3*R*)-hydroxy-*N*-acetyl-(L)-arginine; NCS, non-crystallographic symmetry; PAH, proclavaminic acid amidino hydrolase; PEG 400, poly(ethylene glycol) 400.

¹ These authors contributed equally to the work.

² To whom correspondence should be addressed (e-mail kirsty.hewitson@chemistry.ox.ac.uk).

Scheme 1 Biosynthetic pathway leading to clavulanic acid

Abbreviations used: BLS, β -lactam synthetase; CAD, clavuldehyde dehydrogenase; CEAS, carboxyethylarginine synthase.

cysteinyl-(D)-valine synthetase, an enzyme involved in penicillin biosynthesis. Sequencing has revealed that the *pah* gene was adjacent to the *cas* gene and that disruption of *pah* led to loss of clavulanic acid production, reflecting its absolute requirement in clavulanic acid biosynthesis. Recombinant PAH, both as a maltose-binding protein fusion and in a cleaved form, has also been partially characterized [7]. In the latter instance, the PAH produced was modified at its N-terminus by the addition of four amino acids, Ile-Ser-Glu-Phe.

PAH is a member of the arginase family of enzymes [12] for which only two X-ray crystal structures have been solved; that of type I rat liver arginase [13] and the arginase from *Bacillus caldovelox* [14]. The structures revealed the presence of an active site containing two manganese atoms with a bridging water molecule or hydroxide ion that initiated nucleophilic attack at the carbon atom of the guanidino functionality of the arginine substrate.

In the present study, we report the biochemical characterization and crystal structure of recombinant PAH from *S. clavuligerus*. The crystal structure defines the differences between it and the arginases from which it presumably evolved, and implies conformational changes during catalysis. Together with MS analyses, the results also provide new insights into the oligomerization and role of conformational changes in the catalytic mechanism of di-Mn²⁺-dependent hydrolases.

MATERIALS AND METHODS

Materials

All chemicals were of reagent grade and obtained from Sigma-Aldrich Chemical Co. (Poole, Dorset, U.K.) or Fluka (Gillingham, Dorset, U.K.) unless otherwise stated. Restriction enzymes and molecular biology reagents were purchased from Promega (Southampton, U.K.).

Cloning, overexpression and purification

DNA manipulations were carried out by standard protocols [15]. The *pah* gene [9] was PCR amplified and subcloned directly as an *NdeI/BamHI* fragment into the pET24a(+) vector (Novagen, Nottingham, U.K.). Primers used were: forward, 5'-GG GAA TTC CAT ATG GAG CGC ATC GAC CGC ATC GAC TCG C-3'; and reverse, 5'-CGC GGA TCC TCA CAA CTG GGT TCT GTG GGC-3'. Following amplification, the integrity of *pah* was confirmed by DNA sequencing. The *pah*/pET24a(+) construct was transformed into *Escherichia coli* BL21(DE3) and grown at 37 °C in 2TY media (16 g/litre tryptone, 10 g/litre yeast extract and 5 g/litre NaCl) containing 30 μ g/ml kanamycin. When the D_{600} reached 0.8, isopropyl β -D-thiogalactoside was added to 1.0 mM, and the cells were harvested 3 h later by centrifugation at 14300 *g* for 20 min at 4 °C.

PAH/BL21(DE3) cells were thawed and resuspended in 50 mM Tris/HCl (pH 7.5) containing 1 mM dithiothreitol and then lysed by sonication. Cell debris was removed by centrifugation at 14300 *g* for 20 min at 4 °C, and the resulting supernatant was applied directly on to a Q-Sepharose FF column (55 ml) pre-equilibrated with buffer A [50 mM Tris/HCl (pH 7.5)]. Protein was eluted with a 0–1 M gradient of NaCl in buffer A. The purest samples, as judged by SDS/PAGE, were pooled, concentrated to approx. 50 mg/ml and chromatographed on a Superdex-S75 column (700 ml) that had been equilibrated with 100 mM Tris/HCl (pH 7.5). The purest fractions, > 95% pure by SDS/PAGE, were pooled, concentrated to approx. 18 mg/ml and stored at –80 °C until further use. Protein concentrations were determined by the method of Bradford [16].

Molecular mass determination

The molecular mass of PAH was determined using a Superdex 200 HR 10/30 column equilibrated with 100 mM Tris/HCl (pH 7.5) at a flow rate of 0.5 ml/min. The column was calibrated

with ovalbumin (43 kDa), BSA (67 kDa), aldolase (158 kDa), catalase (232 kDa) and ferritin (440 kDa) (Gel Filtration Calibration Kit, Amersham Biosciences, Little Chalfont, Bucks., U.K.). An elution-volume parameter, K_{av} , was calculated for each of the calibration proteins and a calibration curve constructed. By calculating K_{av} for PAH, the native molecular mass was established.

MS

A portion (10 μ l) of the purified protein (32 mg/ml) in 10 mM Tris/HCl (pH 7.5) was buffer-exchanged twice with 100 mM ammonium acetate (pH 7.0) using micro Biospin-6 chromatography columns (BioRad Laboratories, Hercules, CA, U.S.A.). The sample was diluted to a final concentration of 33 μ M in 100 mM ammonium acetate (pH 7.0) for electrospray ionization (ESI)-MS experiments under native conditions. An aliquot was denatured for MS analysis [33 μ M sample in 100 mM ammonium acetate (pH 7.0)/acetonitrile containing 1% acetic acid, 1:1 (v/v)].

Mass spectra were acquired in positive-ion mode using a Q-ToF instrument (Micromass, Altrincham, Cheshire, U.K.) fitted with nanoflow ESI source. Samples were loaded into borosilicate capillaries (1.0 mm outside diameter $\times$ 0.5 mm inside diameter; Clark Electromedical Instruments, Reading, Berks., U.K.), which were drawn down to a fine taper and coated with gold in-house. The capillary tip was cut manually under a stereomicroscope to give the required diameter and flow. A nitrogen-backing gas line was used to initiate and maintain a flow from the capillary. Nitrogen at room temperature was also used as a drying gas and the ESI source was not heated. Caesium iodide in water was used to calibrate the instrument over the acquisition range 100–20000 m/z , with an acquisition step of 5 s. Collisional cooling was achieved through adjustment of the source rotary pump isolation valve to give the required pressure in the source and transfer hexapole region. Argon gas was also introduced into the collision cell, as this was found to increase the signal from the hexamer-charge states. All spectra were smoothed using MassLynx software (Micromass).

CD spectroscopy

CD spectra were recorded using a Jasco 720 spectropolarimeter (Great Dunmow, Essex, U.K.) at 20 °C in a 1-mm pathlength cell over the range 190–250 nm at a protein concentration of 0.10 mg/ml in 10 mM potassium phosphate buffer (pH 7.5).

Activity assays

Product formation was monitored as a function of urea production, which was quantified using the Sigma Diagnostics blood urea nitrogen determination kit (procedure number 535). Each assay mixture contained 30 mM Tris/HCl (pH 9.0), 0.05 mM $MnCl_2$ and a range of substrate concentrations (0–100 mM) in a total volume of 50 μ l. The reaction mixture was preincubated for 5 min at 37 °C, PAH was added, and the reaction was allowed to continue for various lengths of time (0–5 min) at 37 °C. The reaction was stopped by the addition of 950 μ l of blood urea nitrogen acid reagent/blood urea nitrogen colour reagent (1.4:1, v/v). The absorbance at 526 nm was measured and the amount of urea produced was calculated using a preconstructed calibration curve as per the manufacturer's instructions. Data were fitted to the appropriate equations using SigmaPlot (SPSS Inc., Woking, Surrey, U.K.) and rate constants were derived. A unit of PAH activity was defined as the amount of enzyme that catalyses the formation of 1 nmol of product per min.

Product formation was analysed further by HPLC using an HPLC system fitted with an octadecasilane C_{18} column (250 mm $\times$ 4.6 mm; Gilson, Luton, Beds., U.K.). The mobile phase was 10% methanol and the flow rate was 1 ml/min with the detection wavelength set at 218 nm. Approx. 20 μ l of a 0.2 μ M-filtered reaction volume was loaded when using this system. Product verification was confirmed by comparison with standards.

Activity assays with varying metal ions

Standard PAH assays were performed under aerobic conditions using *N*-acetyl-(L)-arginine (*N*-AA) as substrate at 10 mM, but with a variety of metal ions: Mg^{2+} , Ca^{2+} , Mn^{2+} , Fe^{2+} , Co^{2+} , Ni^{2+} , Cu^{2+} and Zn^{2+} at a final concentration of 0.5 mM. All ions were present as their chloride salts. Details of further incubations and concentrations of additives are given in the text.

Crystallization of PAH

Conditions for the crystallization of PAH were initially screened using Hampton Research Crystal Screens 1 and 2 (Laguna Niguel, CA, U.S.A.) employing the vapour-diffusion method. Hanging-drops of 2 μ l of protein (approx. 11 mg/ml) containing a 3-fold excess of $MnCl_2$ and 2 μ l of well solution were suspended at 17 °C over a well solution containing either 2.2 M ammonium formate/0.2 M Hepes (pH 7.5) or 24% (v/v) poly(ethylene glycol) 400 (PEG 400)/0.1 M Hepes (pH 7.5)/0.2 M $CaCl_2$. Under both conditions, crystals of three different shapes, but with identical space group, grew overnight, reaching a maximum size after 3 days. The crystals initially belonged to space group $P2_1$, with unit cell dimensions $a = 95.4$ Å, $b = 81.9$ Å, $c = 118.6$ Å and $\beta = 95.5^\circ$. Crystals grown from ammonium formate and cryo-cooled using 25% (v/v) glycerol as cryoprotectant belonged to this $P2_1$ symmetry and showed diffraction to 2.8 Å in-house, whereas the same crystals cryocooled using 45% (w/v) sucrose as cryoprotectant belonged to space group $C2$, with unit cell dimensions $a = 140.0$ Å, $b = 79.0$ Å, $c = 93.5$ Å and $\beta = 124.0^\circ$, and diffracted to better than 2.0 Å resolution in-house. The crystals grown from PEG400 did not require an additional cryoprotectant for low-temperature data collection and diffracted to 2.45 Å in-house.

Data collection and phasing

Crystals were cryocooled by plunging into liquid nitrogen, and X-ray data were collected at 100 K using a nitrogen stream. Cryoprotection for crystals grown from ammonium formate was accomplished by sequential transfer in steps of 15% (w/v) sucrose to a solution containing well solution with the addition of 45% (w/v) sucrose. Native and selenomethionine data from crystals grown from ammonium formate were collected on beamline BM14 at the European Synchrotron Radiation Facility, Grenoble, France using a MarCCD detector. Other data were collected on a Rigaku rotating anode generator with osmic focusing mirrors (The Woodlands, TX, U.S.A.) and a 345 mm image plate (Mar Research, Hamburg, Germany). The data were processed using MOSFLM [17] and the CCP4 suite of programs [18].

Using Crystallography and NMR System (CNS) [19], 15 selenium positions were determined, corresponding to the expected three molecules per asymmetric unit, excluding the N-terminal methionine, which was present by N-terminal sequencing and ESI-MS. Phasing and density modification were performed using CNS. The phases had a figure of merit of 0.70, which rose to 0.91 following density modification. Data collection, phasing and refinement statistics are shown in Table 1.

Table 1 Summary of data collection (a) and refinement statistics (b)

Se $\lambda 1$, Se $\lambda 2$ and Se $\lambda 3$ refer to the three different wavelengths at which data were collected using the selenomethionine-substituted protein. RMSD, root-mean-square deviation.

(a)

Data Collection	Datasets for PAH				
	Native	Se $\lambda 1$	Se $\lambda 2$	Se $\lambda 3$	PEG 400
Wavelength (Å)	0.8856	0.9787	0.9789	0.8856	1.5418
Spacegroup	<i>C2</i>	<i>C2</i>	<i>C2</i>	<i>C2</i>	<i>P2₁</i>
Unit cell*	139.86, 78.93, 93.20, 123.90	140.76, 79.22, 93.50, 123.28	140.76, 79.22, 93.50, 123.28	140.76, 79.22, 93.50, 123.28	94.92, 81.35, 120.39, 99.55
Resolution range† (Å)	54.76–1.75 (1.84–1.75)	59.76–2.32 (2.44–2.32)	69.01–2.32 (2.44–2.32)	46.13–1.85 (1.95–1.85)	35.58–2.45 (2.58–2.45)
No. of unique observations†	82 386 (11 034)	37 275 (5115)	37 271 (5109)	71 576 (9335)	65 708 (8947)
No. of total observations†	380 778 (27 406)	252 380 (15 807)	251 576 (15 670)	370 411 (23 594)	262 831 (31 969)
Completeness† (%)	97.1 (89.7)	99.1 (94.0)	99.1 (93.9)	98.3 (88.5)	98.5 (98.5)
Multiplicity†	4.6 (2.5)	6.8 (3.1)	6.7 (3.1)	5.2 (2.5)	4.0 (3.6)
$R_{\text{merge}}†$ (%)	6.7 (26.5)	7.1 (19.5)	7.1 (21.1)	8.2 (33.1)	8.9 (32.1)
$\langle I/\sigma(I) \rangle†$	9.5 (2.7)	9.3 (3.5)	9.4 (2.8)	7.1 (2.1)	7.1 (2.3)
R_{cullis}		0.59	0.63	0.80	
Phasing power		2.10	1.85	1.13	

(b)

Refinement	Datasets for PAH				
	Native	Se $\lambda 1$	Se $\lambda 2$	Se $\lambda 3$	PEG 400
<i>R</i> factor (%)	14.2				22.0
$R_{\text{free}}†$ (%)	17.0				23.6
RMSD-bond length§ (Å)	0.014 (1.54)				0.008 (1.40)
<i>B</i> -factors (Å ²)	7.3, 9.0, 25.9				32.2, 32.6, 26.4

* Unit cell values refer to *a*, *b*, *c* (Å) and β (°) respectively.

† Values in parentheses refer to data in the highest resolution bin.

‡ R_{free} is based on 5% of the data.

§ Values in parenthesis refer to the data for the angles (°).

|| Mean *B*-factors for main chain, side chain and solvent respectively.

Model building and refinement

Skeletonized electron density was calculated using MAPMAN [20] and the protein model was built using the program O [21]. Refinement of the initial model against the selenomethionine data (remote wavelength) was performed using CNS. One cycle of simulated annealing followed by grouped *B*-factor refinement brought the R_{free} to 32.5%. Data to 1.85 Å resolution were used in refinement and non-crystallographic symmetry (NCS) restraints were employed; 5% of the reflections were excluded for calculating R_{free} . Following a further cycle of manual rebuilding and refinement, which brought the R_{free} to 26.0%, the model was transferred to the 1.75 Å native data. Nine further rounds of manual rebuilding and refinement using the program REFMAC5 [22], including individual *B*-factor and TLS parameter refinement and the addition of active-site Mn²⁺ and solvent molecules, brought the conventional *R*-factor to 14.2% and R_{free} to 17.0%. The current model contains residues 9–309 for all three molecules in the asymmetric unit, except for residues 187–192, which were not visible in the electron density for chains A and C, and residues 70–72, which were not visible in chain C. According to PROCHECK, there was one Ramachandran outlier, Gln-14, for which the electron density was well defined. The 2.45 Å native data from PEG 400-grown crystals were solved by molecular replacement with the 1.75 Å native structure using CNS. There were six molecules per asymmetric unit. Refinement of these data was performed by simulated annealing and individual *B*-factor refinement using CNS. The initial refinement was performed using coordinates for one protomer with strict NCS constraints being employed, switching to NCS restraints for subsequent refinement steps.

RESULTS AND DISCUSSION

Expression and purification of PAH

The *pah* gene was overexpressed in *E. coli* BL21(DE3) as approx. 15% of the total soluble protein, as judged by SDS/PAGE, and was the most abundant band. A facile two-step purification protocol was developed from which PAH was isolated to > 95% purity and which was suitable for subsequent biochemical characterization and crystallization. The identity of the isolated protein was confirmed as PAH by N-terminal sequencing (Met-Glu-Arg-Ile-Asp-Ser) and its mass was consistent with that expected from its predicted amino acid sequence. The CD spectrum from 190–250 nm of PAH revealed negative bands at 220 nm and 211 nm together with a positive band at 193 nm indicative of significant α -helical secondary structure. However, the flattening of the α -helix region between the two negative bands implied a β -sheet structure was also present. Although PAH has previously been purified directly from *S. clavuligerus* [8,9] and as a recombinant fusion protein [7], this is the first time production of substantial quantities (> 50 mg) of native PAH has been reported.

Effect of metal ions on activity

It has been reported previously [8] that Mn²⁺ has a stimulatory effect on the activity of PAH. Other divalent metal ions, Mg²⁺, Ca²⁺, Fe²⁺, Co²⁺, Ni²⁺, Cu²⁺ and Zn²⁺, were assayed as their chloride salts for their ability to replace Mn²⁺ using *N*-AA as a substrate (see below). As shown in Figure 1, there was little or no stimulation with Ca²⁺, Fe²⁺ and Zn²⁺. Note that analyses were carried out under aerobic conditions and so it is possible

Figure 1 Effects of different divalent metal ions on PAH activity

Arginase activity is expressed as a percentage of that obtained with Mn^{2+} . 0 refers to the activity obtained in the absence of any added metal cofactor.

that Fe^{2+} is oxidized to Fe^{3+} . However, Co^{2+} increased the amount of urea formed relative to Mn^{2+} , whereas Ni^{2+} was of approximate parity with Mn^{2+} (Figure 1). Similar observations have been reported for the arginases from rat liver, axolotl liver and *B. caldovelox* [23–25]. Unlike PAH, these arginases showed no activity in the presence of Mg^{2+} , but it cannot be ruled out that the PAH activity observed in the present study with Mg^{2+} is due to contamination by another metal either directly or by Mg^{2+} -mediated release of a metal from the PAH protein purification. Presumably the ability of PAH to function efficiently with an alternative metal cofactor allows it to function when there is low Mn^{2+} availability. There is also the possibility that Ni^{2+} or Co^{2+} , rather than Mn^{2+} , is the natural PAH cofactor.

To test the ability of each of the stimulatory metals, Mn^{2+} , Co^{2+} , Ni^{2+} and Mg^{2+} , to stabilize PAH, it was preincubated for 15 min at 37 °C with a metal cofactor, followed by the addition of substrate. In the case of Co^{2+} , Ni^{2+} and Mg^{2+} , full activity was retained. However, addition of Mn^{2+} to PAH during the preincubation period resulted in complete loss of activity. Analogous observations, with respect to the preincubation with Mn^{2+} (but not other metal ions), have also been reported with the arginase from *B. caldovelox* [24], where inactivity was proposed to be due to multiple binding of Mn^{2+} . With the *B. caldovelox* arginase, inhibition was abolished by inclusion of either aspartic acid (25 mM) or BSA (1 mg/ml), both of which are known chelators of Mn^{2+} . Inclusion of aspartic acid or BSA during the preincubation of Mn^{2+} with PAH resulted in activity (results not shown), suggesting that a similar phenomenon was occurring with PAH. However, preincubation of PAH with Mn^{2+} , followed by addition of either BSA or aspartic acid, did not restore activity (results not shown). Nor did the addition of fresh Mn^{2+} . This suggested that effectively irreversible inhibition of PAH had occurred, although it cannot be ruled out that a way of restoring activity may be found. SDS/PAGE (15% gel) of PAH following incubation with Mn^{2+} showed no evidence for hydrolytic degradation of PAH.

PAH was preincubated with various concentrations of Mn^{2+} (0–1 mM) for 15 min at 37 °C before addition of substrate to determine the concentration at which inactivation was achieved. Maximum activity was observed with 0.05 mM Mn^{2+} , followed by a rapid decrease in activity at higher concentrations. This particular Mn^{2+} concentration (0.05 mM) corresponds to an approx. 2-fold excess of Mn^{2+} to PAH relative to the enzyme concentration used. In all subsequent assays, 0.05 mM Mn^{2+} was used to prevent complications with inactivity due to excess Mn^{2+} . The observation that greater than a 2-fold excess of Mn^{2+} led to a decrease in PAH activity is consistent with the proposal of a di- Mn^{2+} active site and that another Mn^{2+} -binding site(s) causes inhibition.

Activity assays

During clavam biosynthesis, one of the primary metabolic precursors, arginine, and three intermediates, compounds 1, 2 and 3, possess a guanidino side chain. Previous work [8] has shown that arginine and compound 1 are not accepted as substrates by PAH, whereas compounds 2 and 3 are both hydrolysed, with the latter at a greater rate. Samples of compounds 2 and 3 are prepared by synthesis and their scarcity limits kinetic analyses of PAH. Alternative substrates were thus sought. *N*-AA is commercially available and (3*R*)-hydroxy-*N*-acetyl-(*L*)-arginine (*N*-AA-OH) can be prepared enzymically from *N*-AA by the use of CAS [26].

The effect of variations in the concentrations of *N*-AA and *N*-AA-OH on PAH activity was determined at pH 9.0 and 37 °C. In each case, a peak corresponding to the expected product was observed by HPLC and urea release was detected. Kinetic analyses using this methodology revealed that PAH catalysis with this substrate fits the Michaelis–Menten model. With *N*-AA as substrate, the K_m was 30 ± 2.5 mM (mean $\pm$ calculated error bounds), whereas with *N*-AA-OH this value fell to 12 ± 2.2 mM (mean $\pm$ calculated error bounds). Arginine and compound 1 were also both tested as substrates, but no hydrolysis was observed.

PAH modified at its N-terminus gave a K_m for its natural substrate, compound 3, of 10–15 mM, but with this particular form of PAH no hydrolysis of *N*-AA was observed [7]. Since *N*-AA can be hydrolysed by wild-type PAH, it seems that N-terminal modification affects substrate hydrolysis. The K_m for *N*-AA-OH was measured as 12 mM, and it is expected that the K_m for the natural substrate would be lower. The K_m for the rat liver and *B. caldovelox* arginases with arginine are 3.0 and 3.4 mM respectively [24,25]. No K_m values have been determined for the PAH isolated directly from *S. clavuligerus*, possibly due to low PAH yields. Sequence alignment of PAH with the rat liver and *B. caldovelox* arginases shows that PAH possesses an N-terminal extension relative to these two arginases (Figure 2). In the hexameric form of PAH, the N-terminus lies close to the active site (see below) and the addition of four non-native amino acids may partially disrupt substrate binding. The rate of hydrolysis of *N*-AA-OH by PAH is faster than for *N*-AA, reflecting the result obtained previously [8] in which compound 3 was hydrolysed at a greater rate than compound 2. It thus appears that the presence of a hydroxyl group increases substrate recognition by PAH. Structural explanations for these proposals are discussed below.

Crystallization and the crystal lattice

Crystals of PAH were obtained under a variety of conditions with the most useful crystals obtained using either ammonium formate or PEG 400 as precipitant. The crystals formed under both conditions appeared in three crystal morphologies (tapered rod, hexagonal plate and oblong block), in each case with the same space group ($P2_1$) and similar cell dimensions. The ammonium formate crystals required the presence of a cryoprotective solution for low-temperature data collection. When well solution with 45% (w/v) sucrose was used as a cryoprotectant, a space group change was observed. Thus, as grown, the crystals belonged to space group $P2_1$, with unit cell dimensions $a = 95.4$ Å (where 1 Å = 0.1 nm), $b = 81.9$ Å, $c = 118.6$ Å and $\beta = 95.5^\circ$, but following transfer into the sucrose solution they belonged to $C2$, with unit cell dimensions $a = 140.0$ Å, $b = 79.0$ Å, $c = 93.5$ Å and $\beta = 124.0^\circ$. During the transfer process, the $P2_1$ crystals were observed to form multiple cracks, which subsequently reformed leaving, once again, an apparently flawless crystal. The two cells are related in that $c(C2)$ is equivalent to $a(P2_1)$, whereas $a(C2)$ is equivalent to the short diagonal of the $P2_1$ cell.

Figure 2 Sequence alignment of rat liver arginase, *B. caldovelox* arginase and PAH

Sequences for rat liver arginase (ARGlrat), *B. caldovelox* arginase (ARGlbaccd) and PAH were obtained from the National Center for Biotechnology Information with the accession numbers P07824, P53608 and P37810 respectively. Conserved Mn^{2+} -binding ligands are indicated by black triangles, guanidino ligands with grey squares and the α -amino acid ligands of the arginases with grey circles. The PAH secondary structure is shown above the alignment with α -helices in dark grey and β -sheets in light grey. Conserved amino acids are shown in light-grey shading, identity between two of the aligned proteins is shown in dark-grey lettering, and identical amino acid residues in the aligned proteins are shown with a black background. The alignment was produced using ClustalW [39] and Alscript [40].

Figure 3 Structure of the PAH monomer

(A) Ribbon diagram of the PAH monomer with the Mn^{2+} ions coloured magenta, α helices coloured red and β -strands coloured yellow. (B) The PAH monomer coloured according to crystallographic temperature factors, with blue as the lowest and red as the highest value. Mn^{2+} ions are coloured magenta. Produced using Molscript [41] and Raster3D [42].

Figure 4 Comparison of the PAH monomer (right) with that of the *B. caldovelox* arginase (left)

Produced using Molscript [41] and Raster3D [42].

In both $P2_1$ and $C2$ space groups, a hexameric arrangement of closely associated PAH molecules was observed. The hexamer is formed by association of two trimers by a 2-fold rotation. In the $C2$ case, this is a crystallographic axis, whereas, in the $P2_1$ case, it is a non-crystallographic 2-fold axis. There was no significant difference in the relative arrangement of monomers within the hexamers of the different symmetries.

Structure of the PAH monomer

The structure of the PAH monomer is shown in Figure 3. The overall similarity of the α/β fold of PAH compared with the rat liver and *B. caldovelox* arginases (Figure 4) is reflected in the superposition of their protein backbones of 70% and 72% with root-mean-square deviations of 1.24 Å and 1.27 Å respectively, over C α atoms. The overall fold comprises a central nine-stranded parallel β -sheet core flanked by α -helices that forms a single α/β domain with a globular structure of approximate size $40 \times 50 \times 40$ Å. Eight of the PAH core β -strands correspond to those of the other arginases, but an additional strand is present in the PAH N-terminal region which flanks β -strand 2. For consistency of nomenclature with the arginases, the N-terminal β -strand is numbered β_0 in the secondary structure numbering scheme (Figure 3). As with the arginases, the eight β -strands are arranged in the order 2-1-3-8-7-4-5-6 in the β -sheet topology, with intervening α -helices linking each of the strands. Five of the helices, α_A , α_F , α_G , α_H and α_I , are packed against one face of the β -sheet core with the four remaining helices, α_B , α_C , α_D , and α_E , against the other face. A small piece of 3_{10} helix, which is not observed in the other arginases for which crystal structures are known, was formed by residues 263–268 in the middle of helix α_H .

Metal-binding site

The PAH active site, formed by the combination of loops that link together the secondary structural elements, is situated at the bottom of a deep cleft and contains two asymmetrically arranged Mn^{2+} ions approx. 3.3 Å apart and bridged by a water molecule (Figure 5). No additional Mn^{2+} -binding sites are evident in the structure. One Mn^{2+} ion, designated Mn_A , is positioned deep in the active-site cavity, whereas Mn_B is closer to the surface. Each Mn^{2+} ion is co-ordinated by a distorted octahedral arrangement of ligands. Mn_A is bound by monodentate interactions from the side chains of His-121, Asp-148, Asp-144 and Asp-235, the

bridging water molecule (as hydroxide ion) and a second water molecule at a distance of 2.2 Å. The bridging water molecule also binds Mn_B as does one oxygen from the carboxylate side chain of Asp-235. Asp-237 binds Mn_B as a bidentate ligand through both carboxylate oxygens, with monodentate interactions from the side chains of Asp-144 and His-146 completing the octahedral co-ordination. This active-site arrangement is similar to that observed with the arginases from both rat liver [13] and *B. caldovelox* [14], with all the Mn^{2+} -binding ligands conserved in the three enzymes. The bridging water molecule is symmetrically positioned between the two Mn^{2+} ions with Mn–O separations of 2.2 Å. This is apparently shorter than that observed for rat liver arginase, which has a corresponding distance of 2.4 Å, but the difference may be within experimental error.

A *cis*-peptide bond is observed between Gly-118 and Gly-119 of the Gly-Gly-Asp-His-Ser motif that is conserved in the arginase family. This motif may be important in positioning His-121 to co-ordinate Mn_A . Although *cis*-peptide bonds rarely involve amino acids other than proline residues [27], they sometimes occur to allow shaping of the active-site cavity [28,29].

Oligomerization of PAH

Following the first isolation of PAH, gel-filtration studies suggested that PAH existed in its native state as a homo-octamer [9]. We found that gel-filtration analysis gave the native molecular mass of PAH to be approx. 199.5 kDa. Since the subunit molecular mass of PAH is 33401 Da, this implies that PAH exists as a hexamer under these solution conditions. To investigate further the oligomeric state of PAH in solution, nanoflow ESI-MS analyses were carried out. Using mild conditions, which might be expected to maintain the protein in a native state, a hexamer (Figure 6) was the major species observed (experimental mass, 200925 ± 51 Da; calculated mass, 200405.4 Da), together with a lower level of dodecamer (experimental mass, 402545 ± 47 Da; calculated mass, 400810.8 Da). Collisional cooling and a further increase in pressure in the collision cell region was found to improve the intensity of the hexamer signal. Under these conditions, it was not possible to dissociate the hexamer by further increasing the extractor voltage, suggesting that the hexamer is the most stable *in vivo* entity and not a 'gas phase' species. The conclusions of the ESI-MS were confirmed by the observation that PAH crystallized as a hexamer composed of two trimeric species (Figure 7). In none of these MS analyses was there any evidence for other oligomeric forms of PAH, including trimers. Eukaryotic forms of arginases appear to exist as trimers and prokaryotic forms as hexamers [30]. The ESI-MS results with PAH support these proposals and also emphasize the strength of the interactions forming the hexamers.

The crystal structure of the rat liver arginase indicated that the 19 amino acids at its C-terminus formed an S-shaped oligomerization motif, which mediated approx. 54% of the inter-monomer contacts between monomers in the trimeric units [13]. This C-terminal tail was subsequently proposed to stabilize the oligomeric nature of the arginase by a network of salt bridges initiating from Arg-308. The crystal structure of PAH reveals that, instead of a C-terminal extension, it possesses an additional flexible loop at its N-terminus, including residues 9–30 and presumably residues 1–8 (which are not visible in the crystal structure), relative to the two arginases for which there are crystal structures (Figures 2 and 4). The native PAH crystal structure shows that the side chain of Arg-11 from one monomer is in position to form a hydrogen bond with the alcohol of Ser-44 in the active site of an adjacent monomer. The additional role

Figure 5 Stereo-view of the electron density map of the active site of PAH

A $2mF_o - DF_c$ electron density map (in green) contoured at 1.7σ that defines the di- Mn^{2+} active site of PAH. An $mF_o - DF_c$ difference map (in red) contoured at 6.0σ calculated after omitting the active-site water molecules. Each Mn^{2+} (coloured magenta) is octahedrally co-ordinated by six ligands. Produced using Bobsript [43] and Raster3D [42].

Figure 6 Nanoflow ESI mass spectrum of PAH showing the hexameric and dodecameric species

Hexamer charge states +33 to +27 and dodecamer charge states +47 to +43.

of the side chains of both Arg-11 and Ser-44 may be in catalysis via hydrogen bonding the hydroxyl group of compound 3, the substrate for PAH, and is discussed further below. Thus, within each trimeric unit of PAH, each monomer (A, B and C) is linked by binding of the N-terminus of the preceding monomer to a region close to the active site. These intermonomer associations are, however, stabilized further by apparent salt bridges from Arg-258 (monomer A)–Glu-259 (monomer B) and Asp-69 (A)–Arg-197 (B), and hydrogen bonds from Glu-59 (A)–Arg-46 (B), Gly-287 (A)–Arg-46 (B) and Gly-67 (A)–His-186 (B). In addition, there is a sizeable matching hydrophobic patch at the intermonomer interfaces composed of Ile-289, Ile-292 and Leu-293 of one monomer with Pro-241, Pro-245, Pro-250 and Pro-252 of another. The intermonomer interactions bury a surface area of $2200 \text{ \AA}^2$.

In contrast, the interactions between the two trimeric units, that together compose the hexamer in which each monomer from one trimeric unit interacts with two monomers from the adjacent trimer, reveal fewer direct associations (Figure 7). At one interface, hydrogen bonds are formed around Gln-14, including

Figure 7 Hexameric nature of PAH

Left-hand panel: surface representation of the PAH hexamer. Right-hand panel: ribbon diagram of the PAH hexamer. Surface representation was produced using Swiss PDB-Viewer [44] and Persistence of Vision Raytracer (POV-ray; <http://www.povray.org>). Ribbons were produced using Molscript [41] and Raster3D [42].

Gln-14 NE2–Asn-87 OD1, Gln-14 NE2–Asp-85 O, Gln-14 OE1–Asp-85 N and Gln-14 N–Asp-85 OD2, for a total of eight hydrogen bonds after considering the 2-fold rotational symmetry. This interface buries a surface area of $1940 \text{ \AA}^2$. At the other interface, burying a surface area of $1650 \text{ \AA}^2$, there is only one hydrogen bond formed directly between protein residues (Arg-50 NH1–Tyr-45 O). However, there is an extensive hydrogen-bond network involving at least 15 crystallographic water molecules, which would be expected to provide considerable stabilization. At each interface there are small complementary hydrophobic areas, such as the ones formed by Pro-45 from two monomers or that between Phe-93 and Phe-153. Despite the apparently lower number of interactions between the trimers, no trimeric species was observed during ESI-MS, indicating that the interactions present are sufficient to allow stabilization of the hexamer.

Substrate binding

When the arginase from *B. caldovelox* was crystallized in the presence of arginine, the resultant structure revealed two arginine molecules bound per monomer [14]. As expected, one was located

Figure 8 Comparison of the *B. caldovelox* arginase and PAH active sites with arginine bound

Left-hand panel: overlay of the *B. caldovelox* arginase active site with that of PAH crystallized using ammonium formate as precipitant. Right-hand panel: overlay of the *B. caldovelox* arginase active site with that of PAH crystallized using PEG 400 as precipitant. Mn^{2+} ions are coloured magenta, *B. caldovelox* arginase is coloured red, PAH is coloured grey and the arginine molecule bound at the *B. caldovelox* arginase active site is coloured green. Note Glu-181 and Asp-178 in the *B. caldovelox* arginase are 'replaced' by Pro-190 and Asn-187 in PAH, which may be involved in binding the substrate β -lactam (see text). Ser-44 is likely to be involved in binding the substrate hydroxyl group. Produced using Swiss-PDB Viewer [44] and POV-ray.

at the active site, but the second was bound at the interface between two monomers of a trimer. This arginine molecule was stabilized by interactions with the side chains of Glu-256 from one monomer and Asp-199 of another, residues that are conserved in all arginase sequences excluding those from *Synechocystis* and *Arabidopsis thaliana* [12]. In the rat liver arginase, the side chain of Arg-308 replaces this additionally bound arginine molecule, forming hydrogen bonds with the conserved aspartic acid and glutamic acid residues. Furthermore, the native *B. caldovelox* arginase structure showed that a flexible loop, residues 12–20, blocked the active site for substrate binding in the absence of arginine. This was due to Arg-18 occupying an internal position that did not allow certain substrate-binding residues to adopt their 'active' conformations. However, upon binding of the second arginine molecule, a rearrangement of residues 12–20 occurred that shifted Arg-18 to an external position, opening up the active site for substrate binding. The corresponding loop in PAH, residues 38–46, does not encroach on the active site. Sequence alignments (Figure 2) also indicate that neither Arg-18, Asp-199 and Glu-256 (*B. caldovelox* arginase) nor Arg-308 (rat liver arginase) are conserved in PAH.

The 1.75 Å native PAH structure from ammonium formate, however, revealed that a loop containing residues 184–193 passes 'through' the active site and would block substrate binding (Figure 8). Two hydrogen bonds stabilize this loop; one from the main chain carbonyl oxygen of Pro-188 to His-64 of an adjacent monomer and a second from the main chain carbonyl oxygen of Gly-67 to His-186 of the same monomer. The corresponding residues in both rat liver and *B. caldovelox* arginase form part of an α -helix that is not directly adjacent to the active site. Consideration of the native PAH structure obtained using PEG 400 showed that this flexible loop (residues 184–193) has moved from the active site, consistent with it undergoing a conformational change upon substrate binding (Figure 8).

Despite attempts to co-crystallize PAH with a substrate (*N*-AA) or substrate analogue [*N*-acetyl-(L)-ornithine or *N*-acetyl-(L)-lysine], no structure of a small molecule bound to PAH has been obtained with either the ammonium formate- or PEG 400-crystallization conditions. In the case of the ammonium formate

crystals, this is presumably due to the presence of a loop (residues 184–193) in the active site, which must move to allow substrate binding. Although this loop no longer blocks the active site, under the PEG 400 conditions, it may still not be in its native conformation and further movement may be required for substrate co-ordination. It is also possible that this additional movement may be achieved by the binding of an 'activator', such as observed with the *B. caldovelox* arginase, where a second molecule of arginine substrate is believed to act in this way [14]. Alternatively, activation may be achieved by another protein. Given the intimate association of CAS with PAH in the catalytic pathway, the former is a possibility for the role. Even if CAS is not responsible for activation of PAH, it seems plausible that the two enzymes are spatially orientated *in vivo* to facilitate transfer of substrate between their active sites. In this regard, it is possible that the hexameric nature of PAH may be important in providing a template for spatial oligomerization, possibly as part of a 'metabolon' as proposed for other pathways [31–33]. Spatial organization of proteins during clavam biosynthesis may help to account for high-production titres and rationalize how highly labile and, presumably, toxic intermediates, such as clavaldehyde, are tolerated.

Despite the absence of a PAH–substrate structure, superposition of that of the *B. caldovelox* arginase containing arginine at the active site [14] and that of native PAH, crystallized under both ammonium formate and PEG 400, is informative with respect to substrate binding by PAH (Figure 8). Essentially, binding by arginases of their substrate can be divided into two regions: (1) those residues that bind the guanidino functionality; and (2) those that bind the α -carboxylate and amino groups. Significantly, all the amino acid residues that bind the guanidino functionality, common to both rat liver and *B. caldovelox* arginases, are also maintained in PAH (Figure 2). The side chains of Glu-279, His-160, Asp-148 and Thr-249, on the basis of PAH numbering, are in position to interact with the guanidino group of the PAH substrate compound 3 (Figure 9). As for the arginases [13], the guanidino functionality of compound 3 would then be positioned between the imidazole ring of His-160 and the di- Mn^{2+} site, allowing the carbon atom of the guanidino group to be

Figure 9 Stereo-view of the substrate-binding region of PAH with the arginine substrate of *B. caldovelox* arginase

The arginine substrate of *B. caldovelox* arginase (coloured green; PDB accession number 3CEV) is superimposed on the substrate-binding region of PAH. Mn^{2+} ions are coloured magenta. Potential substrate carboxylate binding interactions to Arg-11 (from a different monomer) and Asn-150 (from the same monomer) are shown. Note roles of Glu-279, Asp-148 and Thr-249 in binding the guanidino group of the substrate. Produced using Swiss-PDB Viewer [44] and POV-ray.

Scheme 2 Proposed mechanism of hydrolysis by PAH

placed directly above the di- Mn^{2+} -bridging hydrolytic water molecule at a distance of 2.25 Å (Scheme 2). The precise role of His-160 in catalysis for either PAH or the arginases is unclear, since modelling studies indicate that, although His-160 is positioned close to the guanidino group of the substrate, it is incorrectly orientated to form a hydrogen bond. It may be involved in substrate binding/product release or acid catalysis via a stacking arrangement with the guanidino group of the substrate. It seems possible that the amino group of the guanidino group of compound 3 would replace the 'terminal' water ligand of Mn_A to become its sixth co-ordinating ligand, as proposed for the *B. caldovelox* arginase [14].

Evolution of PAH from arginases

The evolution of PAH contrasts with that of the β -lactam synthetase involved in clavam biosynthesis. The β -lactam synthetase apparently evolved from an asparagine synthetase via 'gross' structural changes to one of its domains and active-site expansion [10]. The major difference between the active-site residues of *B. caldovelox* arginase and PAH is in the binding pocket at the ' α -amino terminus' of the substrate. Although both substrates, arginine for *B. caldovelox* arginase and compound 3 for PAH, contain an α -carboxylate group, the α -amino group of arginine is replaced by a β -lactam ring in compound 3.

In the *B. caldovelox* structure, the alcohol side chain of Ser-135 and N δ 2 of Asn-128 form hydrogen bonds directly to the

α -carboxylate oxygens of arginine. On the basis of sequence alignment data, Ser-135 of the *B. caldovelox* arginase is replaced by Gly-156 in PAH, whereas Asn-128 is conserved as Asn-150 (Figures 2 and 9). By using the *B. caldovelox* arginase as a template, Asn-150 is suitably positioned to co-ordinate the C-2- α -carboxylate of compound 3. It also appears that Arg-11 is in a position to bind the substrate carboxylate and would also fulfil the function of neutralizing the substrate negative charge (Figure 9). Arg-11 comes from an adjacent PAH monomer in a trimeric unit and therefore may serve two roles: to mediate intermonomer contacts, but also to aid in the binding of substrate. The involvement of residues from one monomer at the active site of an adjacent monomer also suggests that there may be co-operativity during catalysis, although there is no evidence for this as yet.

The side chains of Glu-181 and Asp-178 of the *B. caldovelox* arginase both bind the α -amino group of arginine. These are replaced by Pro-190 and Asn-187 in PAH (Figure 2). It can be envisaged that the side chain of Asn-187 in PAH binds to the carbonyl oxygen and/or the nitrogen of the β -lactam ring of compound 3 through its N δ 2 hydrogens. Pro-190 may help to create a hydrophobic pocket for binding of the methylene carbon atoms of the β -lactam ring in compound 3.

There is a hydroxyl group present at C3 of compound 3, which is absent in compound 1, compound 2 and arginine (intermediates on the clavulanic acid biosynthetic pathway, which also contain a guanidino functionality). The intermediate compound 2 is a poorer substrate for PAH than compound 3, whereas arginine

and compound 1 are not substrates for PAH [7,8]. In the latter two cases, lack of binding to PAH may be readily rationalized by the presence of polar groups rather than a β -lactam ring. Comparison of the crystal structures for arginases PAH and CAS also reveals: (1) why the alcohol compound 3, rather than compound 2, is the preferred substrate for PAH; and (2) an apparent convergence of arginine and serine residues involved in binding the C-3 alcohol of substrates for both PAH and CAS. In the crystal structure of CAS complexed with proclavaminate acid [11], the product of PAH, the hydroxyl group of proclavaminate acid is in position to hydrogen bond with both the side chains of Arg-297 and Ser-134 (CAS numbering). In PAH, the side chains of two residues, Arg-11 and Ser-44, are similarly positioned and Ser-44 is a likely candidate for substrate hydroxyl binding, itself hydrogen bonded to Arg-11 (Figure 9). Neither Ser-44 nor Arg-11 of PAH are conserved in the rat liver or *B. caldovelox* arginases.

The way in which loops block the active sites of PAH (residues 184–193 from the ammonium formate conditions) and the *B. caldovelox* arginase (residues 12–20) is different. In the *B. caldovelox* arginase, Arg-18 of this loop apparently mimics the substrate and hydrogen bonds with two of the substrate-binding ligands, His-139 and Thr-240, in the same monomer through its guanidino functionality. Movement of this loop is brought about by the binding of a second substrate molecule at the monomer interface and frees the active-site residues for substrate co-ordination. However, in PAH, by superposition with the *B. caldovelox* arginase, no disruption of the active-site residues that bind the guanidino functionality of compound 3 has occurred. Instead, the loop would apparently disrupt binding of the β -lactam portion of the substrate and, in the ammonium formate-crystal form, completely blocks this part of the active site. Comparison with the corresponding polypeptide segment in the *B. caldovelox* arginase reveals that an additional two amino acids are present in this loop in PAH (residues 184–193). These additional amino acid residues may be required to accommodate the larger β -lactam ring of compound 3 compared with the α -amino group of arginine.

The active-site chemistry of PAH is related to that of the class B metallo- β -lactamases that catalyse the hydrolysis of almost all β -lactam antibiotics [34–38]. Most metallo- β -lactamases are believed to use a di- Zn^{2+} -binding centre with a bridging water molecule and catalyse hydrolysis via a related mechanism to that of the arginases [34,35]. Conformational changes involving a mobile loop are also important in substrate binding by the metallo- β -lactamases [34–38]. Although Zn^{2+} is a poor substitute for Mn^{2+} in PAH catalysis and there is no apparent overall structural similarity between PAH and the metallo- β -lactamases [37,38], the convergence with a similar di-metal-using active site with a bridging water molecule is notable; however, the question of any evolutionary significance beyond coincidental convergence remains open.

We thank the Biotechnology and Biological Sciences Research Council, the Engineering and Physical Sciences Research Council, the Medical Research Council, the European Union and the Wellcome Trust for funding, and Amura Ltd for a CASE award to L.X.D. We also thank Dr R. T. Applin for mass spectrometric analyses, Dr M. Gross for CD analysis and Dr Martin Walsh at BM14 of the European Synchrotron Radiation Facility for expert assistance.

REFERENCES

- Matagne, A., Dubus, A., Galleni, M. and Frère, J.-M. (1999) The β -lactamase cycle: a tale of selective pressure and bacterial ingenuity. *Nat. Prod. Rep.* **16**, 1–18
- Jensen, S. E. and Paradkar, A. S. (1999) Biosynthesis and molecular genetics of clavulanic acid. *Antonie Leeuwenhoek* **75**, 125–133
- Khaleeli, N., Li, R. and Townsend, C. A. (1999) Origin of the β -lactam carbons in clavulanic acid from an unusual thiamine pyrophosphate-mediated reaction. *J. Am. Chem. Soc.* **121**, 9223–9224
- McNaughton, H. J., Thirkettle, J. E., Zhang, Z. H., Schofield, C. J., Jensen, S. E., Barton, B. and Greaves, P. J. (1998) β -Lactam synthetase: implications for β -lactamase evolution. *J. Chem. Soc. Chem. Commun.* 2325–2326
- Bachmann, B. O., Li, R. and Townsend, C. A. (1998) β -lactam synthetase: a new biosynthetic enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 9082–9086
- Nicholson, N. H., Baggaley, K. H., Cassels, R., Davison, M., Elson, S. W., Fulston, M., Tyler, J. W. and Woroniecki, S. R. (1994) Evidence that the immediate biosynthetic precursor of clavulanic acid is its N-aldehyde analog. *J. Chem. Soc., Chem. Commun.* 1281–1282
- Baldwin, J. E., Adlington, R. M., Bryans, J. S., Brighen, A. O., Coates, J. B., Crouch, N. P., Lloyd, M. D., Schofield, C. J., Elson, S. W., Baggaley, K. H. et al. (1991) Isolation of dihydroclavaminic acid, an intermediate in the biosynthesis of clavulanic acid. *Tetrahedron* **47**, 4089–4100
- Wu, T.-K., Busby, R. W., Houston, T. A., McIlwaine, D. B., Egan, L. A. and Townsend, C. A. (1995) Identification, cloning, sequencing, and overexpression of the gene encoding proclavaminic acid amidino hydrolase and characterisation of protein function in clavulanic acid biosynthesis. *J. Bacteriol.* **177**, 3714–3720
- Elson, S. W., Baggaley, K. H., Davidson, M., Fulston, M., Nicholson, N. H. and Risbridger, G. D. (1993) The identification of three new biosynthetic intermediates and one further enzyme in the clavulanic acid pathway. *J. Chem. Soc., Chem. Commun.* 1212–1214
- Aidoo, K. A., Wong, A., Alexander, D. C., Rittamer, R. A. R. and Jensen, S. (1994) Cloning, sequencing and disruption of a gene from *Streptomyces clavuligerus* involved in clavulanic acid biosynthesis. *Gene* **147**, 41–46
- Miller, M. T., Bachmann, B. O., Townsend, C. A. and Rosenzweig, A. C. (2001) Structure of β -lactam synthetase reveals how to synthesize antibiotics instead of asparagine. *Nat. Struct. Biol.* **8**, 684–689
- Zhang, Z., Ren, J., Stammers, D. K., Baldwin, J. E., Harlos, K. and Schofield, C. J. (2000) Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. *Nat. Struct. Biol.* **7**, 127–133
- Perozich, J., Hempel, J. and Morris, Jr, S. M. (1998) Roles of conserved residues in the arginase family. *Biochim. Biophys. Acta* **1382**, 23–37
- Kanyo, Z. F., Scolnick, L. R., Ash, D. E. and Christianson, D. W. (1996) Structure of a binuclear manganese cluster in arginase. *Nature (London)* **383**, 554–557
- Bewley, M. C., Jeffrey, P. D., Patchett, M. L., Kanyo, Z. F. and Baker, E. N. (1999) Crystal structures of *Bacillus caldovelox* arginase in complex with substrate and inhibitors reveal new insights into activation, inhibition and catalysis in the arginase superfamily. *Structure (London)* **7**, 435–448
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254
- Leslie, A. G. W. (1999) Integration of macromolecular diffraction data. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 1696–1702
- Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760–763
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S. et al. (1998) Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54**, 905–921
- Kleywegt, G. J. and Jones, T. A. (1996) xDIMPAN and xDATAMAN. Programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **52**, 826–828
- Jones, T. A., Zou, J. Y., Cowan, S. W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **47**, 110–119
- Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **53**, 240–255
- Palacios, R., Huitron, C. and Soberon, G. (1969) Studies on the advent of ureotelism. *Biochem. J.* **114**, 449–454
- Patchett, M. L., Daniel, R. M. and Morgan, H. W. (1991) Characterisation of arginase from the extreme thermophile '*Bacillus caldovelox*'. *Biochim. Biophys. Acta* **1077**, 291–298
- Tarrab, R., Rodriguez, J., Huitron, C., Palacios, R. and Soberon, G. (1974) Molecular forms of rat-liver arginase. Isolation and characterisation. *Eur. J. Biochem.* **49**, 457–468
- Baldwin, J. E., Lee, V., Lloyd, M. D., Schofield, C. J., Elson, S. W. and Baggaley, K. H. (1993) Substrate analogue studies on clavaminic acid synthase. *J. Chem. Soc. Chem. Commun.* 1694–1696

- 27 Stewart, D. E., Sarkar, A. and Wampler, J. E. (1990) Occurrence and role of *cis* peptide bonds in protein structures. *J. Mol. Biol.* **214**, 253–260
- 28 Varghese, J. N., Garrett, T. P. J., Colman, P. M., Chen, L. and Hoj, P. B. (1994) Three-dimensional structures of two plant β glucan endohydrolases with distinct substrate specificities. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2785–2789
- 29 Bond, C. S., Clements, P. R., Ashby, S. J., Collyer, C. A., Harrop, S. J., Hopwood, J. J. and Guss, J. M. (1997) Structure of a human lysosomal sulfatase. *Structure (London)* **5**, 277–289
- 30 Jenkinson, C. P., Grody, W. W. and Cederbaum, S. D. (1996) Comparative properties of arginases. *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* **114**, 107–132
- 31 Velot, C., Mixon, M. B., Teige, M. and Srere, P. A. (1997) Model of a quinary structure between Krebs TCA cycle enzymes: a model for the metabolon. *Biochemistry* **36**, 14271–14276
- 32 Elcock, A. H. and McCammon, J. A. (1996) Evidence for electrostatic channeling in a fusion protein of malate dehydrogenase and citrate synthase. *Biochemistry* **35**, 12652–12658
- 33 Burbulis, I. E. and Winkel-Shirley, B. (1999) Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway. *Proc. Nat. Acad. Sci. U.S.A.* **96**, 12929–12934
- 34 Wang, Z. G., Fast, W., Valentine, A. M. and Benkovic, S. J. (1999) Metallo- β -lactamase: structure and mechanism. *Curr. Opin. Chem. Biol.* **3**, 614–622
- 35 Galleni, M., Lamotte-Brasseur, J., Rossolini, G. M., Spencer, J., Dideberg, O. and Frère, J. M. (2001) Standard numbering scheme for class B β -lactamases. *Antimicrob. Agents Chemother.* **45**, 660–663
- 36 Page, M. I. and Laws, A. P. (1998) The mechanism of catalysis and the inhibition of β -lactamases. *Chem. Commun. (Cambridge)* **16**, 1609–1617
- 37 Carfi, A., Pares, S., Duee, E., Galleni, M., Duez, C., Frère, J. M. and Dideberg, O. (1995) The 3-D structure of a zinc metallo- β -lactamase from *Bacillus cereus* reveals a new type of protein fold. *EMBO J.* **14**, 4914–4921
- 38 Concho, N. O., Janson, C. A., Rowling, P., Pearson, S., Cheever, C. A., Clarke, B. P., Lewis, C., Galleni, M., Frère, J. M., Payne, D. J. et al. (2000) Crystal structure of the IMP-1 metallo β -lactamase from *Pseudomonas aeruginosa* and its complex with a mercaptocarboxylate inhibitor: binding determinants of a potent, broad-spectrum inhibitor. *Biochemistry* **39**, 4288–4298
- 39 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680
- 40 Barton, G. J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.* **6**, 37–40
- 41 Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950
- 42 Merritt, E. A. and Bacon, D. J. (1997) Raster 3D: photorealistic molecular graphics. *Methods Enzymol.* **277**, 505–524
- 43 Esnouf, R. M. (1997) An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graphics Modell.* **15**, 132–134
- 44 Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723

Received 21 January 2002/9 May 2002; accepted 21 May 2002

Published as BJ Immediate Publication 21 May 2002, DOI 10.1042/BJ20020125