Penicillin V acylase crystal structure reveals new Ntn-hydrolase family members

Two enzyme types, penicillin V acylases (PVA) and penicillin G acylases (PGA), with distinct substrate preferences, account for all the enzymic industrial production of 6-aminopenicillanic acid[1,2]. This β-lactam compound is then elaborated into a range of semi-synthetic penicillins. Although their industrial substrates are very similar, representative examples of the two enzyme types differ widely in molecular properties. PVA from *Bacillus sphaericus* is tetrameric with a monomer M_r of 35,000 while PGA from *Escherichia coli* is a heterodimer of M_r 90,000. Furthermore, they have no detectable sequence homology. These differences, which exist in spite of the similarity of their industrial substrates, provoked us to determine the crystal structure of PVA to establish the nature of its catalytic mechanism and to identify any biochemical and structural relationships with PGA and other Ntn (N-terminal nucleophile) hydrolases.

The PVA molecule is a well-defined tetramer with 222 organization made up of two obvious dimers (A and D) and (B and C), which generate a flat disc-like assembly (Fig. 1*a*). The X-ray analysis revealed that the PVA monomer contains two central anti-parallel β-sheets above and below which is a pair of anti-parallel helices (Fig. 1*b*). There are two extensions, one from the upper pair of helices and the other at the C-terminal segment, that interact with other monomers in the tetramer and help stabilize it. The β-sheet and helix organization and connectivity are characteristic of members of the Ntn hydrolase family, which have an N-terminal catalytic residue that is often created by autocatalytic processing[3,4]. In the PVA structure, cysteine was observed as the N-terminal residue, whereas the gene sequence predicts an N-terminal sequence of Met-Leu-Gly-Cys[5]. This finding shows that three amino acids are processed from the precursor N-terminus to unmask a nucleophile with a free α-amino group. Since PVA is an Ntn hydrolase, we can deduce that the N-terminal cysteine in PVA is the catalytic residue.

The PVA and PGA enzymes thus share a distinctive structural core but are otherwise unrelated in primary sequence, including the active site residue. Both PGA and PVA have approximately the same angle (+30°) between the β-strands of the two β-sheets, which are decorated by the active site residues in Ntn hydrolases. Using these β-sheets for structural alignment reveals that the catalytic regions of PVA and PGA overlap (Fig. 1*c*) with a root mean square (r.m.s.) deviation of 0.46 Å for the catalytic atoms. The oxyanion hole in PVA consists of the Nδ2 of Asn 175 and the NH of Tyr 82, corresponding to the Nδ2 of Asn B241 and NH of Ala B69 in PGA. Other common interactions at the active site include those associated with the arginine (residue 228 in PVA and B263 in PGA), and the oxygens of Asp 20 and Asn 175 in PVA and Gln B23 and Asn B241 in PGA, which are critical for positioning the lone pair of the unprotonated N-terminal α-amino group. In spite of these common features, the pH profiles for optimum catalysis by PGA and PVA are quite different[2]. That for PGA is broad (pH 7–9) while it is sharper for PVA, centering on pH 5.5.

Comparison of the catalytic cysteine and serine residues in the two enzymes indicates that their mechanisms may be subtly different. It is notable that in PGA the α-amino nitrogen and the seryl Oγ have a dihedral angle of 60° while in PVA these atoms are eclipsed (Fig. 1*c*,*d*). In PGA the α-amino nitrogen makes a third and well-defined tetrahedral bond to a water molecule that bridges to the nucleophilic seryl oxygen, thus completing its tetrahedral hydrogen-bonding potential. However, the contact from the α-amino group to the seryl OH is unfavorable for a hydrogen bond. It is argued therefore that proton transfer from the seryl Oγ to the neutral α-amino group in PGA will take place via the water molecule, which acts as a virtual base[3]. In PVA a bridging water occurs in a similar position, stabilizing an eclipsed geometry for the N-terminal cysteine. This eclipsed geometry between Sγ and the α-amino group (which presents the hydrogen 2.8 Å from the nitrogen lone pair), favors direct proton transfer. Such a mechanism would also be facilitated by the greater nucleophilic character of sulfur.

The absence of sequence similarity hides the evolutionary relationship between PVA and PGA now demonstrated by X-ray analysis. The connectivity of sequential elements of secondary structure favors divergent evolution of the Ntn hydrolases from a common ancestor[6]. This family contains enzymes of markedly varying size and complexity, with a wide range of substrates and operating in very different biological contexts; it includes for example penicillin G acylase, with a catalytic serine; the proteasome[7,8] and aspartylglucosaminidase[9], with threonine; and *B. subtilis* glutamine PRPP amidotransferase[10], with cysteine. All are processed autocatalytically from inactive precursors[4,11–14]. It is increasingly apparent, however, that in spite of this variety, the simple presence of an N-terminal serine, threonine or cysteine in an enzyme with a nucleophilic mechanism may be sufficient to place it in the Ntn hydrolase family[15]. Thus since the conjugated bile acid hydrolases (CBH) have extensive sequence similarity to PVA[16] and typically have an N-terminal methionine followed directly by cysteine, we can assign these enzymes to the Ntn hydrolase family. We propose by analogy with PVA that Cys 2 is the active residue in CBH—that is, the active residue is revealed by simple processing of the initiation formyl-methionine. This is equivalent to the process seen in the Ntn hydrolase glutaminase domains of *E. coli* glucosamine-6-P-synthase[17] and glutamine PRPP amidotransferase[18].

The relationship of PVA to the choloylglycine hydrolase enzymes sub-family may well help define the real physiological role of penicillin acylases. In addition, PVA and CBH are a fresh test-bed for protein engineering experiments to address the pharmaceutical need to enlarge penicillin acylase specificity to cephalosporins. Indeed, there are natural antibiotics (collectively known as cephalosporin P)[19] which are elaborations of the C_{27} steroid cholesterol nucleus of conjugated bile acids, hinting at a possible distant evolutionary link between antibiotics and cell-signaling molecules.

Methods

Crystallization. Purified PVA from *Bacillus sphaericus*[20] was crystallized by the hanging drop method. The well contained 700 μl 0.2 M sodium phosphate buffer at pH 6.4, 5 mM dithiothreitol (DTT) in 300 μl saturated ammonium sulphate and 100 μl 10% (w/v) sucrose solution. The protein concentration was 15 mg ml^{-1}. Triclinic crystals were obtained when the protein was mixed with phenylacetic acid in 0.2 M Na_2HPO_4. The hexagonal crystals belonged to space group

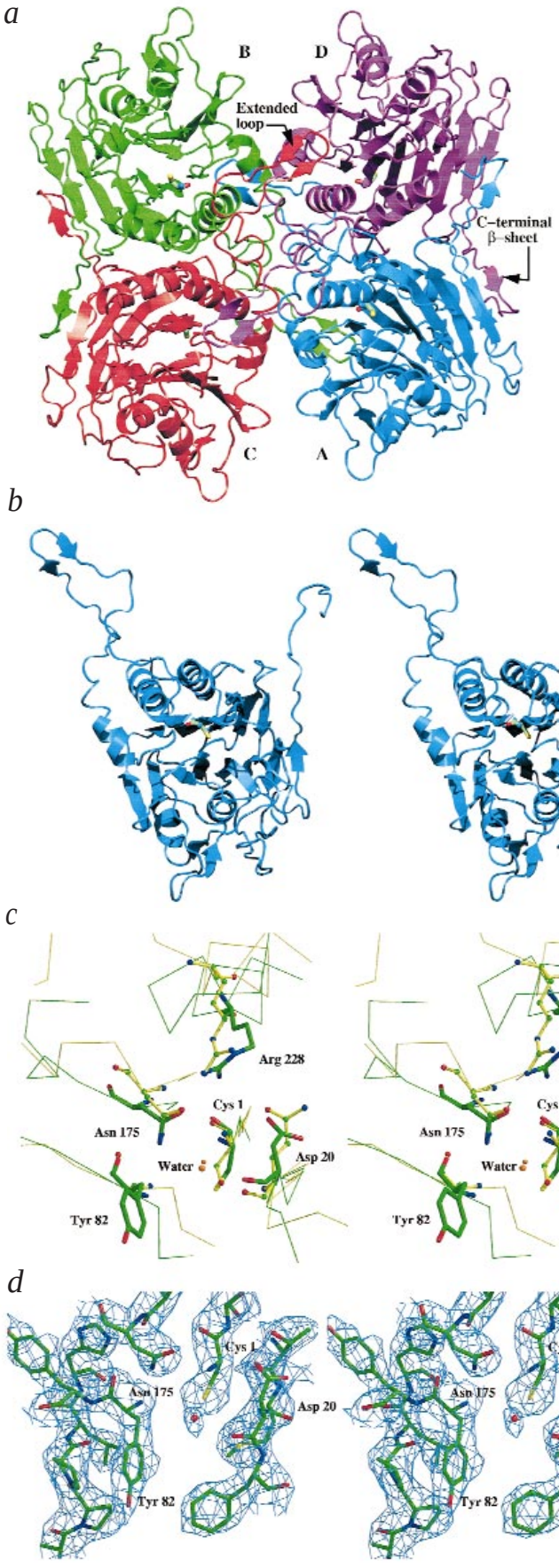

Fig. 1 ***a***,The PVA tetramer showing the main chain as ribbons. The A-D (and B-C) dimer is packed around the central helix with other interactions occurring at the tetramer core and through the cross-over of the C-terminal β-strands. The tetramer is further stabilized by the contacts of the extended loop to the adjacent subunit. The N-terminal cysteine is exposed in all subunits, it is presented towards the reader in A and B. Thus the quaternary organization will not be expected to affect reactivity in any significant way. ***b***,The PVA monomer with its N-terminal catalytic cysteine highlighted in the center of the figure. The αββα sandwich that comprises the central core is characteristic of Ntn hydrolases, and the structural elements share the same topology and connectivity as those in PGA. The two excursions are the extended loop, which stabilizes both the dimer and the tetramer structures, and the C-terminal strand, which interacts with the other monomer of the dimer. ***c***, The overlap of the catalytic environment in PVA and PGA. The PVA structure is represented with a green Cα trace and thick bonds for the amino acids in the catalytic site and the oxyanion hole; the PGA structure has a yellow Cα trace with the highlighted residues drawn in ball and stick figures. The oxygen atoms are colored red, nitrogen blue, sulfur yellow and carbon green. The smaller sphere representing water (orange) belongs to PVA. Note the coincident position of atoms at Asn 175, at the main chain NH 82 and at the NH_2 atom of Arg 228. The more eclipsed conformation of the cysteine is apparent. ***d***, The electron density of PVA calculated with coefficients ($2F_o$ - F_c) and with 2.8 Å data in the region of the catalytic cysteine. The map is contoured at 1σ. In this view the water molecule can be seen near the Sγ. The eclipse of the N and the Sγ in the cysteine completes the tetrahedral interactions at the α-amino group.

P65 with cell dimensions $a = b = 208.4$, $c = 96.3$ Å. The triclinic cell belonged to P1 with unit cell $a = 47.4$, $b = 129.6$, $c = 156.7$ Å, $\alpha = 88.3$, $\beta = 83.4$, $\gamma = 84.6$. 30% glycerol was used as cryoprotectant for freezing triclinic crystals at 120 K. Heavy atom derivatives of the crystals were prepared by soaking overnight in crystallization solution containing 10 mM potassium dicyanoaurate or ethylmercurichloride.

Structure solution. For a summary of structure determination and refinement, see Table 1. The diffraction data were processed using DENZO[21] and the CCP4 programme suite[22]. The Au and Hg derivative sites were located from a combination of difference Patterson maps and Au phased Fourier maps. The parameters for the sites of heavy atoms in both the derivatives were then refined using MLPHARE[23]. The boundaries of the tetramer were selected from the initial electron density map improved by DM[24] and the Cα trace for one monomer defined using QUANTA[25]. Cycles of map improvement using DM and model building located the N-terminal 330 residues. After modelling 203 water molecules in the monomer using ARP[26], followed by several cycles of refinement in REFMAC[27], the R_{cryst} and R_{free} values[28] were 0.23 and 0.25, respectively. Further refinement without non-crystallographic restraints reduced the R_{cryst} and R_{free} to 0.19 and 0.24. The difference map (F_o - F_c) now showed three distinct blobs of density about the S of Cys 1; they were interpreted to be three oxygens forming cysteic acid. Extra density in the vicinity of the N-terminal cysteine fitted an oxidized DTT molecule[29]. Finally, isotropic temperature factors for individual atoms were introduced, refined and anisotropic scaling was used.

The structure in the triclinic crystal form was solved by molecular replacement with AMORE[30], using the refined tetramer coordinates from the hexagonal crystal form. Two tetramers were present in the unit cell. Initial refinement using rigid body and maximum-likelihood calculations on atomic parameters brought down the R_{cryst} and R_{free} to 0.25 and 0.31. No extra density was detected around sulfur at the N-terminal cysteine in the difference density map. Thus it was clear that the cysteine was not oxidized. Water molecules were placed above the 3σ level in the difference density using the program ARP and retaining NCS restraints. Individual atomic B factors were refined in the final stages.

Table 1 Data collection, structure solution and refinement parameters

Data set	**Native 1a**	**Native 1b**	**Hg deriv.**	**Au deriv.**	**Native 2**
Space group	P65	P65	P65	P65	P1
Collection temperature	293 K	293 K	293 K	293 K	120 K
Resolution (Å)	20 - 3.5	20 - 2.5	20 - 3.5	20 - 3.5	18 - 2.8
Reflections measured	72,423	239,296	43,844	52,945	111,412
Unique reflections	29,644	80,217	25,027	24,481	77,056
Completeness (%)	98	97	82	80	85
R_{merge}[1](I)	0.093	0.082	0.167	0.216	0.053
Collection facility	R-axis	Daresbury	R-axis	R-axis	R-axis
R_{iso}[2]			0.23	0.24	
Number of sites			12	8	
R_{cullis}[3] (a/c)			0.69/0.68	0.82/0.85	
Phasing power[4] (a/c)			1.63/1.20	1.10/0.84	
No.reflections phased	24,032				
Mean figure of merit	0.37				
Refinement					
Structure	**Native 1**	**Native 2**			
Resolution (Å)	20.0–2.5	14.0–2.8			
NCS restraint	0	8			
No. protein atoms	10,372	2,605			
Solvent	897	350			
DTT	32				
Reflections (work/free)	75,753/3,963	72,876/3,858			
R_{cryst}/R_{free}[5]	0.193/0.245	0.211/0.243			
R.m.s. deviation[6]					
Bonds (Å)/angles (°)	0.02/2.0	0.02/2.1			
Dihedral angles (°)					
Planar/staggered/ortho	7.0/15.0/20.0	7.0/15.0/20.0			
Average B factor (Å^2)	57.3	56.6			

[1]R_{merge} (I) = $\Sigma_h\Sigma_i|I_{h,i} - \langle I_h\rangle|/\Sigma_h\Sigma_i I_{h,i}$ where I is the observed intensity. <I> is the average intensity of multiple observations of symmetry related reflections.
[2]$R_{iso} = \Sigma||F_{PH}| - |F_P||/\Sigma|F_P|$, where $|F_P|$ is the protein structure factor amplitude and $|F_{PH}|$ is the heavy atom derivative structure factor amplitude.
[3]$R_{Cullis} = \Sigma|E|/\Sigma||F_{PH}|-|F_P||$ for acentric (a) and centric (c) reflections.
[4]Phasing power for acentric (a) and centric (c) reflections = r.m.s ($|F_H|/E$) where F_H is the heavy atom structure factor amplitude and E is the residual lack of closure error.
[5]$R_{cryst} = \Sigma||F_o|-|F_c||/\Sigma|F_o|$; R_{free} is the same as R_{cryst} but was calculated using a separate validation set of reflections that was excluded from the refinement process.
[6]R.m.s. deviation in bond length and angle distances from Engh and Huber.

Coordinates. Coordinates have been deposited in the Protein data bank (accession codes 2pva (oxidized form) and 3pva).

Acknowledgments
We would like to thank G. Murshudov, A. Anston and C. SivaRaman for helpful discussions, the CLRC at Daresbury for synchrotron facilities and M. Vijayan for data collection facilities at the Indian Institute of Science in Bangalore. The British Council Higher Education Link Programme and a travel grant from Glaxo-Wellcome is acknowledged for the support that began the research. The BBSRC provided support through the Rolling Project Mode Time allocation to York. JAB and CSV are funded by the York Structural Biology Centre.

C.G., Suresh[1], A.V. Pundle[1], H. SivaRaman[1], K.N. Rao[1], J.A. Brannigan[2], C.E. McVey[2], C.S. Verma[2], Z. Dauter[2], E.J. Dodson[2] and G.G. Dodson[2]

[1]*Division of Biochemical Sciences, National Chemical Laboratory, Pune-411008, India.* [2]*Chemistry Department, University of York, York, YO10 5DD, UK.*

Correspondence should be addressed to C.G.S email: suresh@ems.ncl.res.in or J.A.B. email: jab@yorvic.york.ac.uk

Received 16 September, 1998; accepted 4 December, 1998.

1. Shewale, J.G. & SivaRaman, H. *Process Biochem.* **24**, 146–154 (1989).
2. Shewale, J.G. & Sudhakaran, V.K. *Enzyme Microb. Technol.* **20**, 402–410 (1997).
3. Duggleby, H.J. *et al. Nature* **373**, 264–268 (1995).
4. Brannigan, J.A. *et al. Nature* **378**, 416–419 (1995).
5. Olsson, A. & Uhlen, M. *Gene* **45**, 175–181 (1986).
6. Artymiuk, P.J. *Nature Struct. Biol.* **2**, 1035–1037 (1995).
7. Lowe, J. *et al. Science* **268**, 533–539 (1995).
8. Groll, M. *et al. Nature* **386**, 463–471 (1997).
9. Oinonen, C., Tikkanen, R., Rouvinen, J. & Peltonen, L. *Nature* Struct. Biol. **2**, 1102–1108 (1995).
10. Smith, J.L. *et al. Science* **264**, 1427–1433 (1994).
11. Tikkanen, R., Riikonen, A., Oinonen, J. & Peltonen, L. *EMBO J.* **15**, 2954–2960 (1996).
12. Guan, C. *et al. J. Biol. Chem.* **271**, 1732–1737 (1996).
13. Brannigan, J.A. & Dodson, G.G. *Nature Struct. Biol.* **4**, 334–337 (1997).
14. Seemuller, E., Lupas, A. & Baumeister, W. *Nature* **382**, 468–470 (1996).
15. Aronson, N.N. *Glycobiology* **6,** 669–675 (1996).
16. Christiaens, H., Leer, R.J., Pouwels, P.H. & Verstraete, W. *Appl. Environ. Microbiol.* **57**, 3792–3798 (1992).
17. Isupov, M.N. *et al. Structure* **4**, 801–810 (1996).
18. Kim, J.H., Krahn, J.M., Tomchick, D.R., Smith, J.L. & Zalkin, H. *J. Biol. Chem.* **271**, 15549–15557 (1996).
19. Burton, H.S. & Abraham, E.P. *Biochem. J.* **50**, 168–174 (1951).
20. Pundale, A. & SivaRaman, H. *Current Microbiol.* **34**, 144–148 (1997).
21. Otwinowski, Z. & Minor, W. *Meth. Enz.* **276**, 307–326 (1997).
22. Collaborative Computational Project No. 4. *Acta Crystallogr. D* **50,** 760–763 (1994).
23. Otwinowski, Z. Proceedings of the CCP4 Study Weekend 80–86 (1991).
24. Cowtan, K. Joint CCP4 and ESF-EACBM Newsletter on protein crystallography **31**, 34–38 (1994).
25. *QUANTA96 X-ray structure analysis user's reference.* (Molecular Simulations, San Diego; 1996).
26. Lamzin, V.S. & Wilson, K.S. *Acta. Crystallogr. D* **49**, 129–147 (1993).
27. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. *Acta Crystallogr. D* **53**, 240–255 (1997).
28. Brunger, A.T. Nature **355**, 472–475 (1992).
29. Stehle, T., Ahmed, S.A., Claiborne, A. & Schulz, G.E. *J. Mol. Biol.* **221**, 1325–1344 (1991).
30. Navaza, J. *Acta Crystallogr. A* **50**, 157–163 (1994).