

The cloning, expression and crystallisation of a thermostable arginase

Maria C. Bewley*, J. Shaun Lott, Edward N. Baker, Mark L. Patchett

Department of Biochemistry, Massey University, Palmerston North, New Zealand

Received 15 March 1996; revised version received 22 April 1996

Abstract The gene for the thermostable arginase from the thermophilic bacterium '*Bacillus caldovelox*' has been cloned and sequenced. Expression of recombinant arginase at high levels has been achieved in *E. coli* using an inducible T7 RNA polymerase-based system. A facile purification procedure incorporating a heat-treatment step yielded 0.2 g of recombinant arginase per litre of induced culture. The kinetic properties of the purified recombinant protein are essentially identical to the native enzyme. The recombinant protein has been crystallised and one crystal form is isomorphous to crystals of the native protein.

Key words: Arginase; Gene cloning; Thermophilic; Thermostable; *Bacillus caldovelox*

1. Introduction

The recent realisation that arginases may play a key regulatory role in nitric oxide metabolism [1,2] and hence affect cytotoxic processes in immunological defence [3] has rekindled interest in this family of enzymes. Arginases (L-arginine amidohydrolases, E.C. 3.5.3.1) are metal ion-activated enzymes which catalyse the hydrolysis of L-arginine to L-ornithine and urea. This reaction is part of the urea cycle in ureotelic animals, and is the initial step of arginine catabolism in certain aerobic bacteria [4]. Despite considerable research on trimeric eukaryotic arginases, notably from rat liver [5–8] and yeast [9–11], these enzymes remain very much a mechanistic and structural mystery. The wide range of transition (VO^{2+} , Fe^{2+} , Co^{2+} , Ni^{2+}) and heavy metal ions (Cd^{2+}) that are able to substitute for the in vivo Mn^{2+} cofactor of arginases also make them an attractive target for the study of metal ion-mediated catalysis and stability [12].

We aim to determine the X-ray crystal structure of the homo-hexameric arginase from the thermophilic bacterium '*B. caldovelox*'. Crystals of the native enzyme have been obtained [13], but the yields of arginase from '*B. caldovelox*' are low. Here we report the cloning, sequencing and expression in *E. coli* of the '*B. caldovelox*' arginase gene, and crystallisation of the recombinant protein, in order to facilitate further structural and physical studies of this enzyme.

2. Materials and methods

'*Bacillus caldovelox*' (*Bacillus* species DSM (Deutsche Sammlung von Mikroorganismen) 411) was grown with strong aeration at 70°C [12]. Enzymes for DNA manipulation were obtained from Gibco BRL and Boehringer Mannheim; Pbluescript KS(–) from Stratagene; radioisotopes from ICN; Sequenase from USB; pGEM-3Zf(–) and *Taq* DNA polymerase from Promega and oligonucleotides from Oligos Etc. Inc.

2.1. Isolation and sequencing of the arginase gene

Genomic DNA from '*B. caldovelox*' was isolated as described for *Thermus aquaticus* [14]. Degenerate oligonucleotides, oligo 1 and oligo 2, were designed, on the basis of N-terminal protein sequence [12] and the conserved internal amino acid sequence -GGDHS-, respectively. Oligo 1: 5'-ATG AAA CCG ATY TCG ATY ATY GGI GTS CCG ATG GA-3'; oligo 2: 5'-RTG RTC SCC SCC MA-3' (M=A or C; R=A or G; S=C or G and Y=C or T).

A 30 cycle PCR was performed using oligos 1 and 2, with a 45°C annealing temperature (first cycle 35°C, second cycle 40°C) and a 15 s extension at 72°C in the presence of 2 mM MgCl_2 with '*B. caldovelox*' genomic DNA as a template. The PCR product was radiolabelled with ^{32}P (Random Primers DNA Labelling System, Gibco BRL) and used to probe Southern blots of genomic DNA digested with a panel of restriction enzymes (Fig. 2). Hybridisation was carried out in 5× Denhardt's reagent, 5× SSC and 0.5% SDS at 68°C, and the filters were washed in 1× SSC, 0.1% SDS at 68°C. A hybridising fragment containing the arginase gene was subcloned into the vector pGEM-3Zf(–), and the coding region of the arginase gene sequenced on both strands using Sequenase. Standard techniques were used for DNA manipulation and analysis [15].

2.2. Subcloning and expression of arginase

The expression vector pKS-rbs is a derivative of pBluescript KS(–) in which the *Xba*I-*Nde*I fragment of pT7-7 [16] containing the ribosome binding sequence has been subcloned downstream of the T7 promoter. The coding region of the arginase gene was amplified from the genomic clone in a 30 cycle PCR using *Pfu* polymerase (Stratagene) at an annealing temperature of 52°C with oligonucleotides 3 and 4, which contain the recognition sites for *Nde*I and *Hind*III, respectively (underlined). Oligo 3: 5'-TGG GAG ACC ATA TGA AGC CA-3'; oligo 4: 5'-CCC ACA AGC TTT ACA TGA AGT-3'.

The PCR product was digested with *Nde*I and *Hind*III and ligated into *Nde*I/*Hind*III digested pKS-rbs to yield the arginase expression plasmid pKS-Arg, in which the initiation codon of the arginase gene is 7 bp downstream of the ribosome binding site. Cells containing pKS-Arg were grown aerobically at 37°C in M1 medium [17], supplemented with 10 mM glucose, 8 μM MnCl_2 and 100 $\mu\text{g}/\text{ml}$ ampicillin, to a culture OD_{600} of 2.0 and induced by adding IPTG to a final concentration of 0.25 mM. Cultures were grown for a further 4 h and cells harvested by centrifugation.

2.3. Protein purification

The cell pellet from a 650 ml culture (3.7 g) was resuspended in 4.5 volumes of lysis buffer (100 mM MOPS/NaOH, 50 mM MnCl_2 , pH 7.5) and lysed by sonication on ice. Cell debris was removed by ultracentrifugation at 220 000×g for 1 h at 5°C. The supernatant was rapidly heated to and incubated at 70°C for 15 min. After rapid cooling on ice, insoluble material was removed by centrifugation at 31 000×g for 15 min at 5°C. Solid ammonium sulphate (enzyme grade, BRL) was added to the supernatant to 30% saturation and the solution was centrifuged at 31 000×g for 10 min at 20°C. The supernatant was made 50% saturated by further addition of ammonium sulphate and the solution was centrifuged as before. The pellet was solubilised in three volumes of buffer (200 mM MOPS/NaOH, pH 7.5) and centrifuged as before. The supernatant was dialysed against 200 volumes of 20 mM MOPS/NaOH, pH 7.5, overnight at 4°C. After dialysis, any insoluble material was removed by further centrifugation at 31 000×g.

2.4. Characterisation of recombinant protein

The purified recombinant protein was subjected to N-terminal sequencing. Arginase activity and protein concentration were determined as previously described [12].

*Corresponding author. Fax: (64) (6) 350 5682.


```

-221 GGGGCCCTAGCTCAGCTGGGAGAGNCCCTGCTTTGACGCGAGGAGTTCATCGGTTTCGATC -162
-161 CCGATAGGCTCCATCAAGAGCAGACTGCTTGACAAACCAAGCAGCGCTCTTTTTTGT -102
-101 TATATCGGTATAAATATTCAAATTTTGTCAAGTTGAGTCATCCCTTTCTTCTCTAC -42
-41 AATAAAGACGTAAACAGGAACCGGAAGGAATGGGAGACAGCATGAAGCAATTTCAATTA 19
1 M K P I S I I 7
20 TCGGGGTTCCGATGGATTAGGGCGAGACACGCCGCGGCTTGATATGGGCGGAGCGCAA 79
8 G V P M D L G Q T R R G V D M G P S A M 27
80 TCGGTTATGACAGCGTCATCGAACGCTGGAACGCTCTCATTACGATATTGAAGATTGG 139
28 R Y A G V I E R L E R L H Y D I E D L G 47
140 GAGATATCCGATTGGAAAGCAGAGCGGTTGCACAGCAAGGAGATTACCGGTTGCGCA 199
48 D I P I G K A E R L H E Q G D S R L R N 67
200 ATTTGAAAGCGGTTGCGGAGCGAACGAGAACTTGGCGGCGGCGTTGACCAAGTCGTT 259
68 L K A V A E A N E K L A A A V D Q V V Q 87
260 AGCGGGGCGGATTTCCGCTTGTGTGGGCGCGACCATGCGCATGCGCATGGCAGCTCG 319
88 R G R F P L V L G G G D H S I A I G G T L A 107
320 CCGGGGTGGCGAAACATTATGAGCGGCTTGGAGTGATCTGGTATGACGCGCATGGCAGC 379
108 G V A K H Y E R L G V I W Y D A H G D V 127
380 TCAACACCGCGAAACGTCGCGCTCTGGAACATTCATGGCATGCGCTGGCGGCGAGCC 439
128 N T A E T S P S G N I H G M P L A A S L 147
440 TCGGGTTGGCCATCCGCGGCTGACGCAATCGGCGGATACAGCCCAATCAAGCGG 499
148 G F G H P A L T G C I I G G Y S P K I K I P E 167
500 AACATGTCGTGTGATCGGCTCCGTTCCCTTGATGAAGGGGAGAAGATTATTCGCG 559
168 H V V L I G V R S L D E G E K K F I R E 187
560 AAAAAGGAATCAAAATTTACACGATGCATGAGGTTGATCGGCTCGGAATGACAGGGTGA 619
188 K G I K I Y T M H E V D R L G M T R V M 207
620 TGAAGAAGCATCGCCTATTAAAGAACGAACGGATGGCGTTCATTGTGCGCTGACT 679
208 E E T I A Y L K E R T D G V H L S L D L 227
680 TGGATGGCCTTGACCAACGACGACCGGAGTCCGGAACGCTGTCTATTGGAGGATTGA 739
228 D G L D P S D A P G V G T P V I G G L T 247
740 CATACCGGAAAGCCATTGGCGATGGAGATGCTGGCGAGGCACAAATCATCACTTCAG 799
248 Y R E S H L A M E M L A E A Q I I T S A 267
800 CGGAATTTGTCGAAGTGAACCGCATCTTGATGAGCGGAACAAACAGCATCAGTGGCTG 859
268 E F V E V N P I L D E R N K T A S V A V 287
860 TAGCGCTGATGGGTCGTTGTTGGTGAAAACTCATGTAATGCATGTGGGCAAGAGGG 919
288 A L M G S L F G E K L M * 299
920 TTGGTTGCGGGATTACCGGATATGATGATTGTTCCCGCTGTTGTGAAAGAGGGGC 979
980 TGCCATTGTTTGGAGTCCCTCTTTTNCGGTTGATGATACAATAAATATGTTGCATGA 1039
1040 AACTTTCTTGATCGAGCGCGTAATATACAGTAAGCCGTGAGCGNGGTATTGTATTTA 1099

```

Fig. 3. DNA sequence of the portion of the genomic arginase clone around the coding region. The deduced arginase protein sequence is shown below the DNA sequence. Potential -35 and -10 promoter regions are highlighted in bold type, and the likely transcriptional start site is underlined. The probable Shine-Dalgarno sequence is marked in bold italics. The double-underlined region contains several potential stem-loop structures, one of which may act as a transcriptional termination signal.

approximately 30–40 mg of purified protein per gram wet weight of induced cells. Approximately 90% of protein contaminants are removed during heat treatment. The purified protein ran as a single band on SDS-PAGE and appeared identical to purified native protein. Fig. 4 illustrates the purification protocol. Amino-terminal amino acid sequencing of the purified recombinant protein was performed and yielded the sequence M-K-P-I-S-I-I-G-V-P-M-D-L-G-Q-, showing that the recombinant protein was correctly expressed in *E. coli*.

The specific activity and K_m of the protein are 4.6 kU/mg (at 60°C) and 3.5 mM respectively. Both values are in close agreement with those reported for the native protein [12].

Crystals were observed under two separate conditions and were subsequently found to be distinct crystal forms: type I and type II. Type I crystals were grown in 14% methoxy-PEG 5000, 200 mM MES/KOH, pH 6.5 and were fully grown with-

in 24 h. A typical crystal grew as a long needle (10 mm×0.1 mm×0.1 mm) and diffracted to 2.2 Å resolution. The space group was $P2_12_12$ with cell dimensions $a=83.4$ Å, $b=146.6$ Å, $c=155.1$ Å. Self-rotation functions calculated for two-fold and three-fold axes showed correlation maxima which were consistent with the arginase hexamer having 32 symmetry. Assuming six molecules in the asymmetric unit, the solvent content is ~53% which is in the normal range found in protein crystals [25]. Type II crystals, grown in 28% PEG 6000, 200 mM Bis-Tris propane/KOH, pH 8.5, appeared after 48 h and were fully grown within 1 week. They grew as square prisms (0.4×0.4×0.5 mm) and diffracted to 3.0 Å, with $\alpha=90^\circ$, $\beta=90^\circ$, $\gamma=120^\circ$ and unit cell dimensions of $a=81.1$ Å, $c=152.0$ Å. The point group symmetry is 622 and thus assuming one monomer in the asymmetric unit, the estimated solvent content is 46% [25]. Arginase undergoes a pH-driven cofactor release, and hence the difference in crystal packing at low and high pH as illustrated by type I and II crystals may reflect a conformational change caused by release of manganese ions.

Comparison of the deduced amino acid sequence of '*B. caldovelox*' arginase with those of other arginases from a diverse range of species (Fig. 1) shows that this is a highly conserved family of homologous proteins. Sequence identity between the '*B. caldovelox*' and *B. subtilis* enzymes is 70% and between '*B. caldovelox*' and human arginase is 43%. The latter figure seems remarkable considering the distinct metabolic roles played by arginase in the two species; clearly the mechanism of action is highly conserved. As well as the sequence -GGDHS-, found in all species, a conserved triplet -DAH- is present 22 residues downstream. The two histidine residues are potential ligands to the cofactor in the active site of arginase.

A number of general rules have been suggested in the literature which attempt to account for the thermostability or

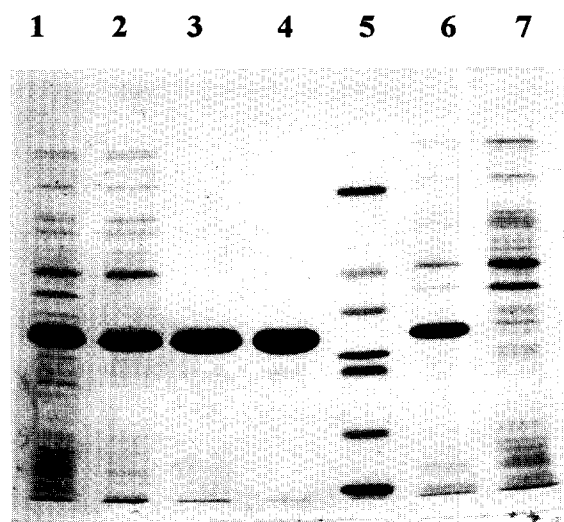


Fig. 4. SDS-PAGE analysis of '*B. caldovelox*' arginase purification and expression. A 15% acrylamide gel running a Laemmli buffer system was stained with Coomassie blue R-250. Lanes: (1) cell lysate; (2) cell-free supernatant; (3) heat treatment supernatant; (4) dialysed $(\text{NH}_4)_2\text{SO}_4$ pellet; (5) SDS-7 molecular weight markers (Sigma) containing a mixture of seven proteins of molecular weights 66, 45, 36, 29, 24, 20 and 14 kDa; (6) IPTG-induced BL21(DE3)/pKS-Arg (whole cell lysate); (7) IPTG-induced BL21(DE3)/pKS-rbs (whole cell lysate).

otherwise of proteins. For instance, comparison of thermostable proteins with non-thermostable homologues seems to indicate that increased thermostability correlates with decreased cysteine, asparagine and glutamine residue content [26], or an increased proline residue content [27]. Cysteine, asparagine and glutamine are thought to be susceptible to covalent damage at high temperatures (e.g. deamidation) and thus tend to be replaced in thermostable proteins [26]. A comparison of the amino acid sequence of the arginase from '*B. caldovelox*' with that of the arginase from the mesophile *B. subtilis* shows that there are fewer cysteines (none compared to one) and asparagines (6 compared to 13) in the more thermostable protein. However, the number of prolines remains unchanged at twelve, and the number of glutamines increases from two to five. Clearly, the phenomenon of protein thermostability cannot be generally explained by amino acid sequence data alone.

However, the '*B. caldovelox*' arginase does contain an increased number of charged amino acid residues compared to the *B. subtilis* protein, which would appear to give weight to the hypothesis that thermostability is achieved by an increase in the number of salt bridges buried within the protein [28]. That said, it is clear that protein thermostability is not a property which can be convincingly explained by general rules, and specific structural information is required to fully understand and analyse the interactions that are responsible for the increased thermostability of a protein relative to non-thermostable homologues.

The previous structural studies on '*B. caldovelox*' arginase [13] were frustrated by the low yield of native protein: even large-scale fermentations produced only 0.1 mg of purified protein per litre of culture. The expression and facile purification of recombinant enzyme presented here allow the structural characterisation to be a viable project. The type I crystals are isomorphous to those previously grown from native protein, and the search for heavy atom derivatives is under way.

Acknowledgements: We gratefully acknowledge financial support from the Foundation for Research, Science and Technology. E.N.B. also receives support as an International Research Scholar of the Howard Hughes Medical Institute.

References

- [1] Bune, A.J., Shergill, J.K., Cammack, R., and Cook, H.T. (1995) FEBS Lett. 366, 127–130.

- [2] Daghighi, F., Fukoto, J.M., and Ash, D.E. (1994) Biochem. Biophys. Res. Commun. 202, 174–180.
- [3] Wang, W.W., Jenkinson, C.P., Griscavage, J.M., Kern, R.M., Arabolos, N.S., Byrns, R.E., Cederbaum, S.D., and Ignarro, L.J. (1995) Biochem. Biophys. Res. Commun. 210, 1009–1016.
- [4] Cunin, R., Glansdorff, N., Piérard, A. and Stalon, V. (1986) Microbiol. Rev. 50, 314–352.
- [5] Kanyo, Z.F., Chen, C.-Y., Daghighi, F., Ash, D.E. and Christianson, D.W. (1992) J. Mol. Biol. 224, 1175–1177.
- [6] Reczkowski, R.S. and Ash, D.E. (1994) Arch. Biochem. Biophys. 312, 31–37.
- [7] Khangulov, S.V., Pessiki, P.J., Barynin, V.V. and Ash, D.E. (1995) Biochemistry 34, 2015–2025.
- [8] Cavalli, R.C., Burke, C.J., Kawamoto, S., Soprano, D.R. and Ash, D.E. (1994) Biochemistry 33, 10652–10657.
- [9] Green, S.M., Eisenstein, E., McPhie, P. and Hensley, P. (1990) J. Biol. Chem. 265, 1601–1607.
- [10] Green, S.M., Ginsburg, A., Lewis, M.S. and Hensley, P. (1991) J. Biol. Chem. 266, 21474–21481.
- [11] Hensley, P. (1988) Curr. Top. Cell. Regul. 29, 35–75.
- [12] Patchett, M.L., Daniel, R.M. and Morgan, H.W. (1991) Biochim. Biophys. Acta 1077, 291–298.
- [13] Smith, C.A., Pratchett, M.L. and Baker, E.N. (1995) Acta Crystallogr. D51, 840–841.
- [14] Lawyer, F.C., Stoffel, S., Saiki, R.K., Myambo, K., Drummond, R. and Gelfand, D.H. (1989) J. Biol. Chem. 264, 6427–6437.
- [15] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- [16] Tabor, S. and Richardson, C.C. (1985) Proc. Natl. Acad. Sci. USA 82, 1074–1078.
- [17] McKie, N., Keep, N.H., Patchett, M.L. and Leadlay, P.F. (1990) Biochem. J. 269, 293–298.
- [18] Kingston, R.L., Baker, H.M. and Baker, E.N. (1994) Acta Crystallogr. D50, 429–440.
- [19] Arndt, U.W. and Wonacott, A.J. (Eds.) (1977) The Rotation Method in Crystallography. North Holland, Amsterdam.
- [20] Otwinowski, Z. (1990) DENZO Data Processing Package. Yale University, New Haven, CT.
- [21] Collaborative Computational Project Number 4 (1994) Acta Crystallogr., D50, 760–763.
- [22] Tong, L. and Rossmann, M.G. (1990) Acta Crystallogr. A46, 783–792.
- [23] Bleasby, A.J. and Wooton, J.C. (1990) Protein Engineering, 3, 153–159.
- [24] Thompson, J.D., Higgins, D.G. and Gison, T.J. (1994) Nucleic Acids Res. 22, 4673–4680.
- [25] Matthews, B.W. (1968) J. Mol. Biol. 33, 491–493.
- [26] Mrabet, N.T. et al (1992) Biochemistry 31, 2239–2253.
- [27] Muir, J. et al. (1995) Protein Engineering 8, 583–592.
- [28] Tomschy, A., Boehm, G. and Jaenicke, R. (1994) Protein Engineering 7, 1471–1478.