The alkene monooxygenase from *Xanthobacter* Py2 is a binuclear non-haem iron protein closely related to toluene 4-monooxygenase

Ning-Yi Zhou, Alistair Jenkins, Chan K.N. Chan Kwo Chion, David J. Leak*

Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AZ, UK

Received 14 April 1998; revised version received 11 May 1998

Abstract The genes encoding the six polypeptide components of the alkene monooxygenase from *Xanthobacter* Py2 have been sequenced. The predicted amino acid sequence of the first ORF shows homology with the iron binding subunits of binuclear non-haem iron containing monooxygenases including benzene monooxygenase, toluene 4-monooxygenase (>60% sequence similarity) and methane monooxygenase (>40% sequence similarity) and that the necessary sequence motifs associated with iron coordination are also present. Secondary structure prediction based on the amino acid sequence showed that the predominantly α-helical structure that surrounds the binuclear iron binding site was conserved allowing the sequence to be modelled on the coordinates of the methane monooxygenase α-subunit. Significant differences in the residues forming the hydrophobic cavity which forms the substrate binding site are discussed with reference to the differences in reaction specificity and stereospecificity of binuclear non-haem iron monooxygenases.

Key words: Alkene monooxygenase; Epoxide; Non-heme iron; Methane monooxygenase; Toluene 4-monooxygenase; *Xanthobacter*

1. Introduction

Xanthobacter Py2 is a Gram-negative bacterial strain which grows on propene as sole carbon and energy source [1]. Metabolism of propene involves an alkene specific monooxygenase which converts propene to epoxypropane (Fig. 1). Further metabolism involves isomerisation and carboxylation of the epoxide, ultimately yielding acetoacetate which feeds into central metabolism [2–4].

The alkene monooxygenase (AMO) will catalyse the epoxidation of a range of alkenes, some with a high degree of stereospecificity, but will not hydroxylate the homologous alkanes [5]. This contrasts with the alkane monooxygenases such as those based on cytochrome P450 [6] or non-haem iron (e.g. ω-hydroxylase [7] and methane monooxygenase [8]) which appear to generate a highly reactive iron-oxygen intermediate which can attack both unactivated C-H bonds and carbon-carbon double bonds. The origin of both the reaction specificity and stereospecificity of alkene monooxygenases is, therefore, of considerable interest. Recently, the AMO from *Xanthobacter* Py2 has been resolved into four components: an NADH dependent reductase; a Rieske-type ferredoxin; an oxygenase; and a small protein which may be a coupling or effector protein [9]. The oxygenase is an $\alpha_2\beta_2\gamma_2$ hexamer which was reported to contain approximately four atoms of non-haem iron per hexamer on the basis of colorimetric iron analysis and the lack of a significant UV/visible chromophore. However, it was not possible to obtain an EPR spectrum which might have confirmed the nature of the iron binding site.

Sequence evidence [10] and, more recently, EPR evidence [11] has demonstrated that an alkene specific monooxygenase derived from the Gram-positive *Rhodococcus rhodochrous* (formerly *Nocardia corallina*) B276 has a binuclear non-haem iron centre of the type found in soluble methane monooxygenase. However, the *Xanthobacter* enzyme is considerably more complex than that from the *Rhodococcus*, having a two component (reductase and ferredoxin) redox system typical of aromatic dioxygenases [12], and a more complex oxygenase structure.

We have previously reported the cloning and expression of the AMO from *Xanthobacter* Py2 [13] and in this paper present sequence and molecular modelling evidence indicating that the *Xanthobacter* AMO is also a binuclear non-haem iron enzyme with a strong resemblance to the recently characterised [14] toluene 4-monooxygenase from *Pseudomonas mendocina* KR1.

2. Materials and methods

2.1. Subcloning and sequencing strategy

An 11.2 kb *Eco*RI fragment from the cosmid clone pNY2 had previously been subcloned as pNY2C and shown to complement mutants of *Xanthobacter* Py2 which were unable to grow on propene but retained the ability to grow on propene oxide [13]. This insert was transferred to plasmid pDELTA2 and used for generating nested deletions using the Deletion Factory system version 2 (Gibco-BRL). Initially, 120 individual deletion clones from each deletion direction were screened to estimate the size of the remaining inserts and subsequently 36 clones from each deletion direction were sequenced. The entire region from the end of the cosmid to the beginning of the isomerase/carboxylase genes [15] has been sequenced in both directions using BigDye Terminator Cycle Sequencing (Perkin Elmer) with Ampli*Taq* FS DNA polymerase and analysed on an ABI 377 automated sequencer (Applied Biosystems).

2.2. Sequence analysis and structural modelling

DNA and deduced amino acid sequences were analysed using MacVector (v. 4.5.3) and AssemblyLIGN (v. 1.07) software packages (Eastman Kodak Co). Homologous protein searches were done with FASTA 3[16] and PSI-BLAST [17]. Pairwise alignments were made with the GAP program in the Wisconsin package [18]. Multiple alignments were run under clustalW [19]. Secondary structure predictions were made by building a consensus from predicted results of nnpredict [20], Predict Protein [21], PREDATOR [22], SOPMA [23], GORIV [24] and 123D prediction which was run on the website http://genomic.sanger.ac.uk/123D/run123D.html.

The model protein 3D structure of the *xamoA* gene product was generated using the Modeller 4 program, an automated approach to comparative protein structure modelling by satisfaction of spatial constraints [25], using the α-subunit of the methane monooxygenase hy-

*Corresponding author. Fax: (44) (171) 5945207.
E-mail: d.leak@bc.ic.ac.uk

PII: S0014-5793(98)00653-X

droxylase (1mmo.pdb) from *Methylococcus capsulatus* [26] as the template structure. The *xamoA* encoded sequence was aligned with the related *memA* encoded sequence (accession number A67077, A.C. Rosenzweig, C.A. Frederick, S.J. Lippard, P. Nordlund). The alignment was then used to calculate an all atom 3D model of the *xamoA* gene product. The model was visualised using Insight II (MSI, San Diego, CA, USA).

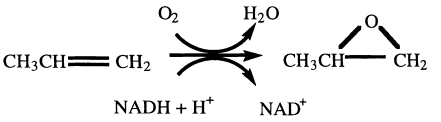

Fig. 1. Epoxidation reaction catalysed by alkene monooxygenase.

3. Results and discussion

The genes encoding AMO from *Xanthobacter* Py2 (XAMO) have previously been cloned in the broad host range cosmid vector pLAFR5 and shown to express propene inducible AMO when transferred to *X. autotrophicus* JW33, although it was not expressed in *Escherichia coli*. [13]. A 2.5 kb fragment from one end of this cosmid was able to complement mutants lacking AMO activity, indicating the probable location of the *amo* genes. Furthermore, a region stretching from the middle of the cosmid to the distal end has been sequenced and shown to contain the genes encoding components of the epoxide isomerase/carboxylase complex [15,4]. We have now sequenced the 2.5 kb complementing region and the bulk of the intervening sequence between this and the isomerase/carboxylase genes and located six closely spaced ORFs, consistent in size and predicted amino acid sequence with the reported components of the AMO [9]. The first ORF in this sequence encodes a polypeptide of 497 amino acids (58 037 Da) which has a significant degree of sequence similarity to the iron containing subunits of a number of monooxygenases including toluene (TMO) [27], benzene (BMO) [28] and methane (MMO) [29] monooxygenases (Fig. 2). Although it also has sequence similarity to the homologous subunit in the alkene monooxygenase from *R. rhodochrous* B276 [10] it is more closely related to the toluene/benzene monooxygenase family (Table 1). Spectroscopic studies of toluene 4-monooxygenase (T4MO) have recently demonstrated that this is a binuclear non-haem iron monooxygenase [14] and, like the XAMO, is a two redox component system involving both a reductase and a ferredoxin.

The binuclear non-haem iron centre in MMO is co-ordinated via two histidine and four glutamate residues [26]. In common with other binuclear non-haem iron proteins, two glutamate and the two histidine residues form a pair of characteristic Glu-X-X-His sequence motifs. Alignment of the predicted amino acid sequence of the first ORF (*xamoA*) in the *xamo* gene cluster clearly demonstrates that the pair of Glu-X-X-His sequence motifs are present with an intervening region of identical size (97 amino acids) to that in T4MO. Additionally, the other co-ordinating ligands equivalent to Glu^{114} and Glu^{209} in MMO are also conserved. This strongly suggests that XAMO is also a binuclear non-haem iron monooxygenase. Furthermore, protein secondary structure prediction demonstrated that the predominantly α-helical structure which surrounds the iron binding site is also present in the predicted amino acid sequence from *xamoA* (Fig. 2). Conservation of these key structural elements indicated that it should be possible to generate a 3D model of the *xamoA* encoded polypeptide on the co-ordinates available for the α-subunit in the crystal structure of the MMO. Although there are some obvious differences at some distance from the active site, the strong degree of secondary structure conservation in the region corresponding to the 4-helix bundle in MMO that co-ordinates the iron binding site meant that it was possible to thread the structure of the *xamoA* encoded polypeptide on the backbone of the MMO α-subunit almost without the need for energy minimisation (Fig. 3). Within the limits of this type of approach this confirms that all of the structural elements are present in the *xamoA* encoded polypeptide to confidently assign the XAMO to the family of binuclear non-haem iron monooxygenases.

Until recently, most of the information on binuclear non-haem iron monooxygenases had originated from studies on MMO. Like cytochrome P450 monooxygenases, MMO must be capable of generating a sufficiently reactive iron bound oxygen intermediate to facilitate C-H bond cleavage. The proposed mechanism(s) for MMO catalysed hydroxylation are analogous to those for P450 involving iron bound O-O bond cleavage leaving the equivalent of an $Fe^{IV}=O$ radical cation species, known as compound Q [33]. However, the observation that aromatic monooxygenases and now alkene monooxygenases are part of the same family but with evidently different reaction specificities means that some subtle constraints must be operating within the framework of the MMO mechanism. While the ancillary components of XAMO may exert some discrimination between alkenes and alkanes it is evident that the mechanism of substrate attack at the active site must also differ as, for instance, with substrates like 2-butene the alkene specific monooxygenases only carry out epoxidation whereas the main product from MMO is the terminal allylic alcohol [34]. Furthermore, when MMO does carry out alkene epoxidation this is not stereospecific, whereas

Table 1
Proteins showing homology to the *Xanthobacter* Py2 alkene monooxygenase *xamoA* encoded polypeptide

Genes with homologous products	Size (aa)	% identity	% similarity	Organism	Accession number	Reference
bmoA	500	47.69	67.0	*Pseudomonas aeruginosa* JI104	D83068	[28]
tmoA	499	45.07	63.98	*Pseudomonas mendocina* KR1	Q00456	[27]
tbuA1	501	44.06	64.39	*Ralstonia pickettii* PKO1	U04052	[30]
phhN	516	27.16	46.91	*Pseudomonas putida* P35X	1093589	[31]
memA	527	22.86	45.16	*Methylococcus capsulatus* (Bath)	78074	[29]
memA	527	22.82	46.47	*Methylosinus trichosporium* OB3b	95287	[32]
amoC	501	22.08	48.33	*Rhodococcus rhodochrous* B-276	D37875	[10]

Percentage identity and similarity values were obtained with the GAP program (see Section 2). GenBank accession numbers are also given.
bmo = benzene monooxygenase; tmo = toluene 4-monooxygenase; tbu = toluene 3-monooxygenase; phh = phenol hydroxylase; mem = methane monooxygenase; amo = alkene monooxygenase.

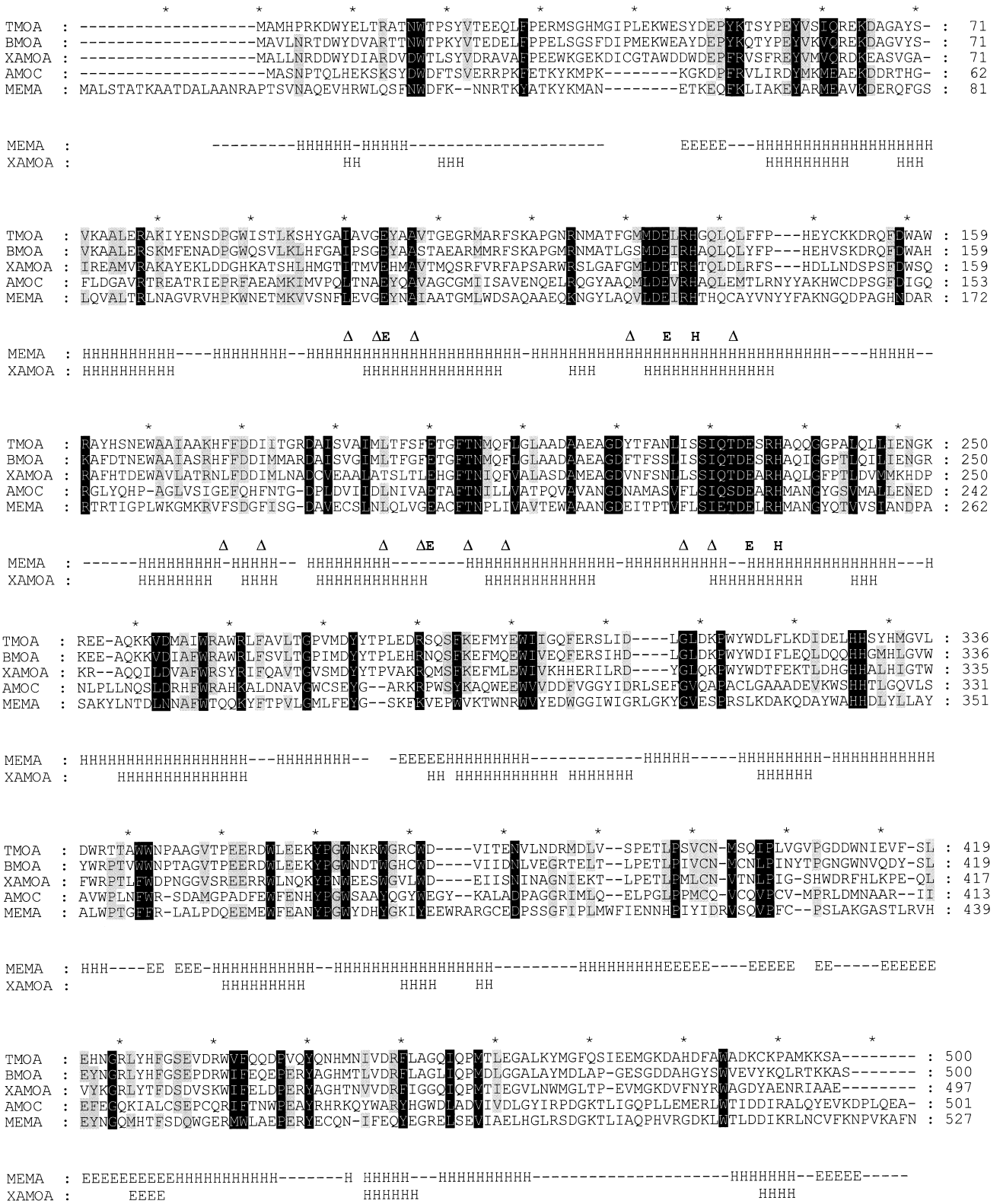

Fig. 2. Conserved features of the *Xanthobacter xamoA* encoded polypeptide sequence. Numbers indicate the positions of the respective amino acids within the polypeptide. Highlighting indicates identical or functionally conserved amino acid residues amongst all five (dark) or four (light) sequences. Triangles indicate the positions of residues equivalent to those surrounding the substrate binding pocket in MMO, which are referred to in the text; the six ligands to the binuclear iron are also indicated (E and H). Underneath the aligned sequences the secondary structural features of the *M. capsulatus* MMO α-subunit (MEMA) obtained from the crystal structure are shown, together with the predicted sheet (E) and helical regions (H) for the *xamoA* encoded polypeptide. Predicted polypeptides are abbreviated on the basis of their respective genes, as in the legend to Table 1.

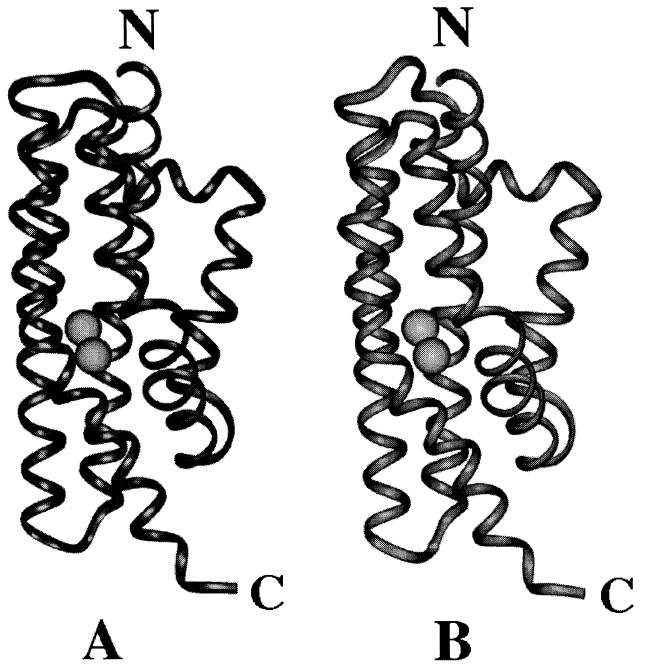

Fig. 3. A ribbon diagram representation of the helical regions surrounding the iron binding site in (A) the model of the *Xanthobacter* Py2 xamoA gene product and (B) the methane monooxygenase α-subunit. The two iron atoms are represented as spheres and the N-termini (N) and C-termini (C) of the regions displayed are indicated.

XAMO has a high degree of stereospecificity with some alkene substrates [35].

The binuclear iron centre in the MMO α-subunit is situated at the base of a hydrophobic cavity formed by amino acid side chains [26]. The cavity is believed to form the substrate binding site and must therefore determine the access and orientation of the substrate with respect to the iron bound oxygen. Given the subtle differences between AMO and MMO, comparison of the residues surrounding the active site might be informative and serve as a guide for site directed mutagenesis. In MMO the residues include Leu^{110}, Ala^{117}, Phe^{188}, Leu^{204}, Gly^{208}, Ile^{217}, Phe^{236} and Ile^{239} [26,36]. A Monte-Carlo simulation of substrate-subunit interaction has also suggested that Gly^{113}, Gln^{140} and Phe^{192} are important in substrate binding [37]. The equivalent residues in the homologous monooxygenases, including XAMO, are highlighted in Fig. 2. With numerical references based on the MMO residues it is evident that Ala^{117} and Ile^{239} are completely conserved in all homologues. Leu^{110} has recently been proposed to be a key component in a substrate/product gating mechanism in MMO [36] and appears to be functionally conserved by replacement with an ile in XAMO and the aromatic monooxygenases, while differences at Gln^{140}, Phe^{192}, Leu^{204}, Ile^{217} and Phe^{236} are not informative, apparently falling into phylogenetic rather than functional groupings. However, there are significant differences in the equivalent positions to Gly^{113}, Phe^{188} and Gly^{208}. At Gly^{113} and Gly^{208} we see a homologous series of increasingly bulky residues going from MMO α-subunit to *amoC* to *xamoA* encoded polypeptides, consistent with their increasing stereospecificity in propene epoxidation. The differences at Phe^{188} in both *amoC* and *xamoA* encoded polypeptides are significant in the sense that they break with the phylogenetic groupings. However, they are less systematic.

Two other residues are worth highlighting. Thr^{213} is completely conserved and is believed to be involved in providing a proton for the initial O-O bond cleavage[8]. Cys^{151} is not conserved but the equivalent residue in the other monooxygenases appears to vary in a functional rather than a phylogenetic manner. Thus, in both of the aromatic monooxygenases this is a Gln residue, while in both *amoC* and *xamoA* encoded polypeptides this is an acidic residue. This is significant because, in one of the mechanisms proposed for MMO, hydrogen is abstracted from Cys^{151} in the α-subunit forming a cysteinyl radical which gives rise to a concerted hydroxylation of the substrate [8]. The equivalent residues in the aromatic and alkene monooxygenases would not be able to fulfil this function which could well explain the inability of these monooxygenases to catalyse C-H bond oxidation.

Acknowledgements: We are grateful to BBSRC for financial support to A.J. (Studentship) and to CKNCKC (Grant 28/T07300).

References

[1] Van Ginkel, C.G. and De Bont, J.A.M. (1986) Arch. Microbiol. 145, 403–407.
[2] Weijers, C.A.G.M., Jongejan, H., Franssen, M.C.R., De Groot, A.E. and De Bont, J.A.M. (1995) Appl. Microbiol. Biotechnol. 43, 775–781.
[3] Chan Kwo Chion, C.K.N. and Leak, D.J. (1996) Biochem. J. 319, 499–506.
[4] Allen, J.R. and Ensign, S.A. (1997) J. Biol. Chem. 272, 32121–32128.
[5] Van Ginkel, C.G., Welten, H.G.J. and De Bont, J.A.M. (1987) Appl. Environ. Microbiol. 53, 2903–2907.
[6] Sligar, S.G. and Murray, R.I. (1986) in: Cytochrome P450 (Ortiz de Montellano, P.R., Ed.), pp. 429–503, Plenum, New York.
[7] Shanklin, J., Whittle, E. and Fox, B.G. (1994) Biochemistry 33, 12787–12794.
[8] Feig, A.L. and Lippard, S.J. (1994) Chem. Rev. 94, 759–805.
[9] Small, F.J. and Ensign, S.A. (1997) J. Biol. Chem. 272, 24913–24920.
[10] Saeki, H. and Furuhashi, K. (1994) J. Ferment. Bioeng. 78, 399–406.
[11] Gallagher, S.C., Cammack, R. and Dalton, H. (1997) Eur. J. Biochem. 247, 635–641.
[12] Butler, C.S. and Mason, J.R. (1997) Adv. Microb. Physiol. 38, 47–84.
[13] Zhou, N.-Y., Chan Kwo Chion, C.K.N. and Leak, D.J. (1996) Appl. Microbiol. Biotechnol. 44, 582–588.
[14] Pikus, J.D., Studts, J.M., Achim, C., Kauffmann, K.E., Munck, E., Steffan, R.J., McClay, K. and Fox, B.G. (1996) Biochemistry 35, 9106–9119.
[15] Swaving, J., Weijers, C.A.G.M., Van Ooyen, J.J. and De Bont, J.A.M. (1995) Microbiology 141, 477–484.
[16] Pearson, W.R. and Lipman, D.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2444–2448.
[17] Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Nucleic Acids Res. 25, 3389–3402.
[18] Devereux, J., Haeberli, P. and Smithies, O. (1984) Nucleic Acids Res. 12, 387–395.
[19] Thompson, J.D., Higgin, D.G. and Gibson, T.J. (1994) Nucleic Acids Res. 22, 4673–4680.
[20] Kneller, D.G., Cohen, F.E. and Langridge, R. (1990) J. Mol. Biol. 214, 171–182.
[21] Rost, B., Sander, C. and Schneider, R. (1994) CABIOS 10, 53–60.
[22] Frishman, D. and Argos, P. (1996) Protein Eng. 9, 133–142.
[23] Geourjon, C. and Deleage, G. (1994) Protein Eng. 7, 157–164.
[24] Garnier, J., Gibrat, J.F. and Robson, B. (1996) Methods Enzymol. 266, 540–553.

[25] Sali, A., Potterton, L., Yuan, F., van Vlijmen, H. and Karplus, M. (1995) Proteins 23, 318–326.
[26] Rosenzweig, A.C., Frederick, C.A., Lippard, S.J. and Nordlund, P. (1993) Nature 366, 537–543.
[27] Yen, K.M., Karl, M.R., Blatt, L.M., Simon, M.J., Winter, R.B., Fausset, P.R., Lu, H.S., Harcourt, A.A. and Chen, K.K. (1991) J. Bacteriol. 173, 5315–5327.
[28] Kitayama, A., Suzuki, E., Kawakami, Y. and Nagamune, T. (1996) J. Ferment. Bioeng. 82, 421–425.
[29] Stainthorpe, A.C., Lees, V., Salmond, G.P., Dalton, H. and Murrell, J.C. (1990) Gene 91, 27–34.
[30] Byrne, A.M., Kukor, J.J. and Olsen, R.H. (1995) Gene 154, 65–70.
[31] Ng, L.C., Shingler, V., Sze, C.C. and Poh, C.L. (1994) Gene 151, 29–36.
[32] Cardy, D.L., Laidler, V., Salmond, G.P. and Murrell, J.C. (1991) Mol. Microbiol. 5, 335–342.
[33] Wallar, B.J. and Lipscomb, J.D. (1996) Chem. Rev. 96, 2625–2657.
[34] Leak, D.J., Aikens, P.J. and Seyed-Mahmoudian, M. (1992) Trends Biotechnol. 10, 256–261.
[35] Weijers, C.A.G.M., Van Ginkel, C.G. and De Bont, J.A.M. (1988) Enzyme Microb. Technol. 10, 214–218.
[36] Rosenzweig, A.C., Brandstetter, H., Whittington, D.A., Nordlund, P., Lippard, S.J. and Frederick, C.A. (1997) Proteins Struct. Funct. Genet. 29, 141–152.
[37] George, A., Wilkins, P. and Dalton, H. (1996) J. Mol. Catal. 2, 103–113.