

Hypothesis

The molecular peculiarities of catalase-peroxidases

Marcel Zámocký^{a,b}, Günther Regelsberger^c, Christa Jakopitsch^c, Christian Obinger^{c,*}

^aInstitute of Biochemistry and Molecular Cell Biology, University of Vienna, Dr. Bohrgasse 9, A-1030 Vienna, Austria

^bInstitute of Molecular Biology, Slovak Academy of Sciences, Dubravská cesta 21, SK-84251 Bratislava, Slovak Republic

^cInstitute of Chemistry, University of Agricultural Sciences, Muthgasse 18, A-1190 Vienna, Austria

Received 21 December 2000; revised 5 February 2001; accepted 5 February 2001

First published online 21 February 2001

Edited by Matti Saraste

Abstract In developing ideas of how protein structure modifies haem reactivity, the activity of Class I of the plant peroxidase superfamily (including cytochrome *c* peroxidase, ascorbate peroxidase and catalase-peroxidases (KatGs)) is an exciting field of research. Despite striking sequence homologies, there are dramatic differences in catalytic activity and substrate specificity with KatGs being the only member with substantial catalase activity. Based on multiple sequence alignment performed for Class I peroxidases, we present a hypothesis for the pronounced catalase activity of KatGs. In their catalytic domains KatGs are shown to possess three large insertions, two of them are typical for KatGs showing highly conserved sequence patterns. Besides an extra C-terminal copy of the ancestral hydroperoxidase gene resulting from gene duplication, these two large loops are likely to control the orientation of both the haem group and of essential residues in the active site. They seem to modulate the access of substrates to the prosthetic group at the distal side as well as the flexibility and character of the bond between the proximal histidine and the ferric iron. The hypothesis presented opens new possibilities in the rational engineering of peroxidases. © 2001 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Catalase-peroxidase; Ascorbate peroxidase; Cytochrome *c* peroxidase; Multiple sequence alignment; Site-directed mutagenesis

1. Introduction

Catalases and peroxidases are oxidoreductases involved in the molecular mechanisms of defence against reactive oxygen species. These two abundant groups of enzymes possess striking similarities in their reaction mechanism but have different residues in the haem cavity. The initial step in the catalytic mechanism of a peroxidase or a catalase is heterolysis of the oxygen–oxygen bond of hydrogen peroxide (Fig. 1, reaction 1). This reaction causes the release of one water molecule and coordination of the second oxygen atom to the iron centre [1]. The resulting intermediate, compound I, is two oxidising equivalents above the resting state, two electrons have been

transferred from the enzyme to the coordinated oxygen atom, one from the iron and one from either the porphyrin or an amino acid [1]. The currently accepted structure for a catalase compound I is a ferryl ($\text{Fe}^{\text{IV}}=\text{O}$) porphyrin π -cation radical, whereas with different peroxidases both this form and a ferryl ($\text{Fe}^{\text{IV}}=\text{O}$) protein radical species have been reported [1]. Generally, rate constants for the formation of compound I from peroxidases and catalases were calculated to be in the range from 10^6 to $10^8 \text{ M}^{-1} \text{ s}^{-1}$ [1]. In classical catalases a distal His–Asn and in peroxidases a distal His–Arg pair have been shown to be essential for compound I formation [1,2]. The main difference in the enzymatic mechanism between catalases and peroxidases is compound I reduction. In a catalase cycle, a second peroxide molecule is used as a reducing agent for compound I. This two-electron reduction completes the cycle forming ferric catalase (Fig. 1, reaction 2) and molecular oxygen [3]. In a peroxidase cycle, with most substrates compound I is reduced in two consecutive one-electron steps via compound II back to the ferric enzyme (Fig. 1, reactions 3 and 4).

Catalase-peroxidases (KatGs) can be viewed as a molecular fossil revealing the common phylogeny of catalytic and peroxidative activity during evolution [4]. Up to now the complete genes (*katGs*) were characterised only from prokaryotes

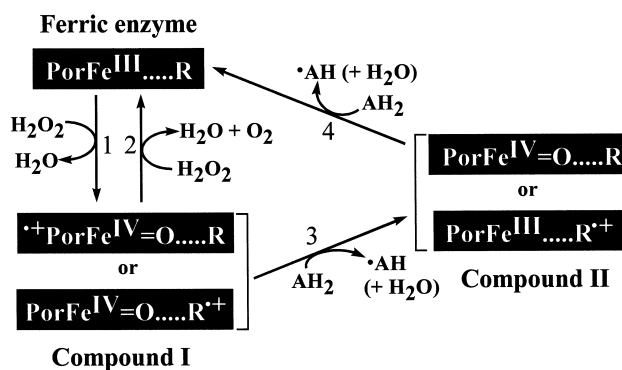


Fig. 1. Putative reaction scheme. In the first step H₂O₂ is used for compound I formation (reaction 1). Compound I is two oxidising equivalents above that of the native enzyme with a porphyrin π -cation radical in combination with an iron(IV) centre or an amino acid radical in combination with iron(IV). Compound I can react with a second H₂O₂ reducing the enzyme back to the ferric state (reaction 2, catalase reaction). In the peroxidase reaction compound I is transformed in the first one-electron reduction to compound II containing either an iron(IV) centre or an amino acid radical (R[•]) in combination with Fe(III) (reaction 3). Compound II is finally reduced back to ferric peroxidase in a second one-electron reduction (reaction 4).

*Corresponding author. Fax: (43)-1-36006 6059.
E-mail: cobinger@edv2.boku.ac.at

Abbreviations: APX, ascorbate peroxidase; CCP, yeast cytochrome *c* peroxidase; KatG, catalase-peroxidase; *katG*, gene encoding catalase-peroxidase; HPI, hydroperoxidase I

W105F, which were characterised as peroxidases with negligible catalase activity. Point mutants of *Synechocystis* PCC 6803 KatG were characterised by steady state and transient state kinetics. The results demonstrated the essential role of amino acid residues R119 and H123 for the catalytic reactivity. After their mutagenesis (mutations R119A, R119N, H123Q, H123E) a drastic loss of the catalase activity was observed ranging from 14.6% down to only 0.02% remaining activity. Furthermore, it was shown that the distal tryptophan (W122) is essential for H₂O₂ oxidation but not for the peroxidative activity. Mutants W122F and W122A completely lost their catalase activity but retained their peroxidase activity [8,9]. Mutations of corresponding residues in *M. tuberculosis* KatG were only partially characterised, but exhibited the same behaviour as the two above mentioned hydroperoxidases [10]. If this highly conserved triad, especially the distal Trp, is so important for the catalytic process in KatGs, why do CCP and APX (which also contain the same invariant catalytic triad) not exhibit catalytic activity? The possible answer could be obtained from comparison of the 3D-structures of the respective proteins. Currently the 3D-structures are known for APX [11] and CCP [12] but, unfortunately, for several reasons [7] no crystal structure is available for a KatG yet. Therefore, in this situation we have to rely on analysis of the sequence of Class I peroxidases which reveals possible structural homologies and/or differences.

3. The presence of three large loops in the catalytic part of KatGs

From the multiple sequence alignment performed for all three branches of Class I we see that there are three large insertions and one short stretch in the sequences of KatGs which APXs and CCP do not possess (Fig. 2). The short stretch occurs at position 102–113 (numbering corresponds to *Synechocystis* KatG) and is positioned in the region homologous to the connection between helices A and B (known from 3D-structures of APX and CCP, Fig. 3). This sequence pattern 'SQXWWPADXGXY' on the distal side of the haem is highly conserved (>95%) among all known KatGs. The first large insertion occurs between positions 210 and 251 and is of variable length among KatGs. This loop, with a length of 36–43 residues, occurs between helices D and E situated between the distal and proximal catalytic domain. There is a conserved sequence on the C-terminus of this loop comprising residues 'AXXMGLIYVN'. The second large insertion, occurring at positions 300–334, has the same length for all known KatGs (35 or 36 residues) and is relatively conserved among all KatGs. There exists a conserved sequence 'GXXPXXAXXEXQGLGW' in this insertion. It is located between helices F and G in the corresponding sequence of peroxidases with known structure. The third large insertion between residues 357 and 392 (between helices G and

[species] AA.pos.↓

```

Arath-APX      117 GREDKQPPP-----PEGR--LPDATKGDHLDVFAKQMGLSDK
Capan-APX     118 GREDKPEPP-----VEGR--LPDATKGDHLDVFAKQMGLSDQ
Cucsa-APX     118 GREDKPEPP-----PEGR--LPDATKGDHLDVFAKQMGLSDQ
Fraan-APX     118 GREDKPEPP-----PEGR--LPDATKGDHLDVFAKQMGLSDQ
Glyma-APX     118 GREDKPEPP-----PEGR--LPDATKGDHLDVFAKQMGLSDQ
Goshi-APX     116 GRKDSNICP-----REG--LPDAKRGAPHLRDIFF-YRMGLSDK
Mescr-APX     112 GRKDTDVAD-----T-L-N--IPNPNGGADHLRTVF-HQMGLVDK
Nicta-APX     118 GREDKTEPP-----VEGR--LPDATKGDHLDVFAKQMGLSDK
Orysa-APX     118 GREDKPAPP-----PEGR--LPDATKGDHLDVFAKQMGLSDQ
Pissa-APX     118 GREDKPEPP-----PEGR--LPDATKGDHLDVFAKQMGLSDQ
Vigun-APX     118 GREDKPEPP-----PEGR--LPDATKGDHLDVFAKQMGLSDQ
Zeama-APX     118 GRQDKPEPP-----PEGR--LPDATKGDHLDVFAKQMGLSDQ
Sacce-CCP     129 GRVDTPEDT-----TPD--NGRLPDADKDADYVRTFF-QRLNMNDR
Arcfu-KatG    165 GREDIFEPDESPDWGPEEEMLT----AKRGEKE-----ELERPFAATEMGLIYVNP
Bacst-KatG    179 GRVDVWHPEEDVYWGSEKEWLA----SERYSG----DRE--LENPLAAVQMGLIYVNP
Ecoli-KatG    184 GREDVWEPDLVDNNGDEKAWLT----HRHPEA-----LAKAPLGATEMGLIYVNP
EcoliP-KatG   174 GREDAFEEDKAVNNGPEDEFETQ----ERFDEP-----GEIQEGLGASVMGLIYVNP
Legpn-KatG    161 GRKDDYTPDEAVDWGPEDEWETTS---GDRFDAD----GSLKWPLGNTVMGLIYVNP
Mycbo-KatG    177 GREDIWHPEKDVYWGSEQEWLG----AKRYDGK----SRESLENPLAAVQMGLIYVNP
Mycfo-KatG    186 GRVDQWEPDE-VYWGKEATWLG----DERYSKG----RDLENPLAAVQMGLIYVNP
Mycin-KatG    195 GRQDIWEPEE-ILWGQEDTWLGT----DKRYSGE----RELAQPYGATMGLIYVNP
Mycin-KatG    192 GREDVWEPEE-ILWGEEEEWLGT----DKRYSGE----RELAQPYGATMGLIYVNP
Myctu-KatG    186 GRVDQWEPDE-VYWGKEATWLG----DERYSKG----RDLENPLAAVQMGLIYVNP
Salty-KatG    185 GREDVWEPDLVDNNGDEKAWLT----HRHPEA-----LAKAPLGATEMGLIYVNP
Scosp-KatG    172 GREDIWHPEKDIYWGPEKEWFPPSTNPNSRYTG---DRE--LENPLAAVQMGLIYVNP
Scysp-KatG    201 GREDIWHPEKDIYWGAEKEWLASS---DHRYGSE---DRESLENPLAAVQMGLIYVNP
Strre-KatG    187 GRADVWEAEDVYWGPETTWLD----DRRYTGD----RELENPLAAVQMGLIYVNP

```

: . : * : : *

phd output for Scysp-KatG

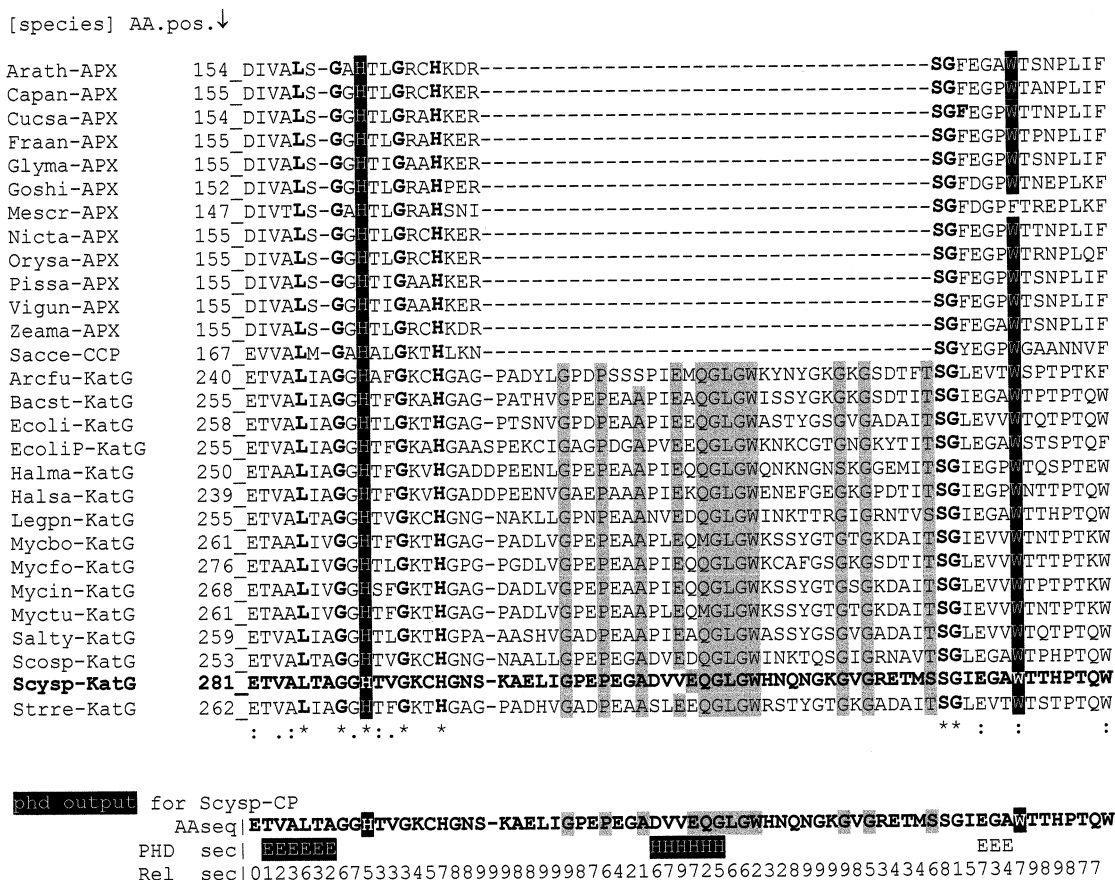
```

AAseq|GREDIWHPEKDIYWGAEKEWLASS---DHRYGSE---DRESLENPLAAVQMGLIYVNP
PHD sec|          HHHHHH          EEEE          HHHHHHHHHHHH-HH
Rel sec|978899987775761446651455  4778988  7743147542112442783699999999833667999999 97578931

```

B

Fig. 2. (continued).



C

Fig. 2. (continued).

H) seems not to be highly conserved among KatGs, on the other hand there exists a similar although shorter insertion in CCP. Therefore, one could suggest that the first and the second large insertion as well as the short distal stretch are of functional importance for KatGs.

4. Structural consequences of the large insertions in KatGs

The first large loop typical for KatGs is positioned between conserved helices D and E known from the crystal structure of APX and CCP. According to the secondary structure prediction [13] there may be a short α -helix at the beginning of this loop and a short extended β -structure at the end which can stabilise this relatively long unique insertion (Fig. 2B). As is obvious from the inspection of the corresponding region in the structure of APX and CCP [11,12] this pattern may possibly control the access to the cavity on the distal side of the haem group and also reorient the essential residues relative to the haem iron. It also can serve as a molecular sieve supporting the diffusion of small peroxide molecules in the distal pocket. It should also be noted that the short stretch on the distal side preceding the essential helix B is of functional importance for KatGs since it can (i) control access to the essential catalytic triad Arg-Trp-His for the potential substrates and (ii) hinder typical peroxidase substrates in their diffusion

to their binding and oxidation sites (e.g. at the periphery of the porphyrin ring) with the consequence that the reduction of compound I is accomplished mainly by the excess hydrogen peroxide molecules present in the distal pocket.

The second large loop is situated on the proximal side of the haem. It is known for peroxidases with known structure that the iron ion is strongly coordinated with the imidazole ring of the essential proximal His [1]. Interestingly this 35-residues long insertion is situated in the region of the short β -structure of APX and CCP and probably this sequence stretch is differently organised in KatGs. It was reported [14] that in CCP after single-residue mutagenesis in this region (W191G) an extra cavity was created and increased flexibility of the neighbouring region resulted. Thus it is logical to postulate that a 35-amino acid-residue-long loop can totally reconstruct the proximal side of the haem and consequently modulate the strength of the Fe-N coordination as mentioned above. Consequently the secondary structure prediction in the region of the proximal His produces a different output for KatGs in comparison to APX and CCP. Whereas the essential proximal His in APX and CCP is known to be located in a truncated α -helix, in KatGs the corresponding His is located on the edge of an extended β -structure (Fig. 2C).

In APX and CCP the proximal histidine has an imidazolate character because its δ -nitrogen seems to be H-bonded to a

Table 1
Sources of analysed hydroperoxidases

Abbreviation	Accession number	Enzyme	Organism (strain)
Arath-APX	spQ05431	APX 1	<i>Arabidopsis thaliana</i>
Capan-APX	gbX81376	APX	<i>Capsicum annuum</i>
Cucsa-APX	gbD88649	APX	<i>Cucumis sativus</i>
Fraan-APX	gbAF022213	APX	<i>Fragaria × ananassa</i>
Glyma-APX	gbL10292	APX	<i>Glycine max</i> (soybean)
Goshi-APX	gbU37060	APX	<i>Gossypium hirsutum</i>
Mescl-APX	gbU43561	APX	<i>M. crystallinum</i>
Nicta-APX	gbU15933	APX	<i>Nicotiana tabacum</i>
Orysa-APX	gbD45423	APX	<i>Oryza sativa</i>
Pissa-APX	spP48534	APX	<i>P. sativum</i>
Vigun-APX	gbU61379	APX	<i>Vigna unguiculata</i>
Zeama-APX	gbZ34934	APX	<i>Zea mays</i>
Sacce-CCP	spP00431	cytochrome <i>c</i> peroxidase	<i>Saccharomyces cerevisiae</i>
Arcfu-KatG	gbAE000951	KatG	<i>Archaeoglobus fulgidus</i>
Bacst-KatG	gbM29876	catalase I	<i>Bacillus stearothermophilus</i>
Ecoli-KatG	spP13029	catalase HPI	<i>E. coli</i>
EcoliP-KatG	gbX89017	EHEC-KatG	<i>E. coli</i> (0157: H7)
Halma-KatG	gbY16851	KatG	<i>Haloarcula marismortui</i>
Halsa-KatG	gbAF069761	KatG	<i>Halobacterium salinarum</i>
Legpn-KatG	gbAF078110	KatG	<i>Legionella pneumophila</i>
Mycbo-KatG	spP46817	KatG	<i>Mycobacterium bovis</i>
Mycfo-KatG	gbY07865	KatG	<i>Mycobacterium fortuitum</i>
Mycin-KatG	spQ04657	KatG	<i>Mycobacterium intracellulare</i>
Myctu-KatG	spQ08129	KatG	<i>M. tuberculosis</i>
Salty-KatG	spP17750	catalase HPI	<i>Salmonella typhimurium</i>
Scosp-KatG	gbD61378	KatG	<i>Synechococcus</i> sp. (PCC7942)
Scysp-KatG	gbD90910	catalase HPI	<i>Synechocystis</i> sp. (PCC6803)
Strre-KatG	gbY14317	KatG	<i>Streptomyces reticuli</i>

Abbreviations included in text, with their accession numbers from GenBank (gb) or Swiss-Prot (sp) and organisms from which they originate.

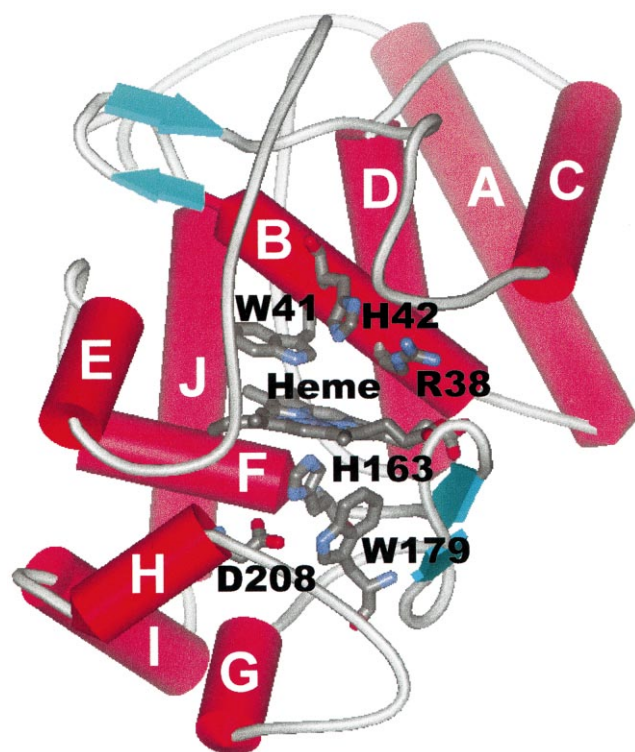


Fig. 3. Ribbon and helical structure of APX from *P. sativum* [11]. Important residues at the haem site are shown. The figure was constructed using WebLab Viewer Pro from Molecular Simulations Inc.

conserved aspartate residue. In APX and CCP, the latter is H-bonded to a tryptophan residue. In KatG proteins both residues also seem to be conserved (D402 and W341 in *Synechocystis* KatG). In APX a conserved Asp (D208 APX from *Pisum sativum*) is located in helix H, whereas in KatGs secondary structure prediction suggests its location is not in the corresponding helix but rather between the third large insertion and helix H (not shown). In APX, a conserved Trp (W179 in APX from *P. sativum*) is located between helix F and G (Fig. 3), whereas in KatGs it is located between the second large insertion and helix G (Fig. 2C). Since both the bond strength of the iron–imidazole link as well as the reduction potential of the protein [15] is determined by the relative position, distance and the basic character of the proximal His, it is reasonable to assume that these KatG-specific insertions contribute to the catalytic features typical for KatGs.

5. The occurrence of an extra copy of the ancestral hydroperoxidase gene and its influence on the overall properties of KatGs

It has been shown that all KatGs possess a second copy of the ancestral hydroperoxidase gene at the C-terminus with no obvious catalytic function ([4] and references within). Even the essential catalytic triad is not conserved in this part of the sequence and it remains as a relic of ancestral gene duplication. However, random mutagenesis experiments revealed the importance of this part for the proper function of KatG, since also mutations in this C-terminal part lost the catalase and peroxidase activities [16]. Unfortunately, the authors monitored only crude lysates after transposition mutagenesis and this negative evidence cannot exclude the formation of misfolded mutants of mycobacterial KatGs. Because

the main insertions are partially conserved in this C-terminal copy of the ancestral hydroperoxidase gene (data not shown) we can expect that the typical fold of the catalytic domains also exists here and promote interdomain and intersubunit interactions (KatGs are known to exist as dimers or tetramers, no monomeric KatG has yet been reported). It is highly probable that for reasonable catalase activity the typical oligomeric assembly must be present, similar to monofunctional haem catalases [17], and for stable folding and assembly the 'extra' C-terminal sequence of KatG is essential.

6. Conclusions: predictions for site-directed mutagenesis of KatG

To study the structure–function relationships in KatGs in the situation where no crystal structure is available one can perform site-directed mutagenesis to remove the extra loops in KatGs and monitor the changed profiles in catalase and peroxidase activity. Of special interest is the mutagenesis of the highly conserved proximal insertion (Fig. 2C). Interestingly an insertion of comparable length was reported in horseradish peroxidase [18] which belongs to Class III of the plant peroxidase superfamily. But this insertion has no homology with those conserved in KatG and it is stabilised by several glycosylation sites. Mutagenesis in the proximal loop of KatG may be of technological importance similar to examples mentioned by Fitzgerald et al. [15] for CCP. By varying the length of this sequence stretch one can modulate the flexibility of the chain around the proximal His and through it also the Fe–N bond, and consequently modulate the reactivity of the haem iron.

References

- [1] Dunford, H.B. (1999) Heme Peroxidases, Wiley-VCH, New York.
- [2] Fita, I. and Rossmann, M.G. (1985) *J. Mol. Biol.* 185, 21–37.
- [3] Nicholls, P. and Schonbaum, G.R. (1963) in: *The Enzymes* (Boyer, P.D., Lardy, H. and Myrbäck, K., Eds.), 2nd edn., pp. 147–225, Academic Press, New York.
- [4] Zámocký, M., Janecek, Š. and Koller, F. (2000) *Gene* 256, 169–182.
- [5] Fraaije, M.W., Roubroeks, H.P., Hagen, W.R. and Van Berkel, W.J.H. (1996) *Eur. J. Biochem.* 235, 192–198.
- [6] Welinder, K.G. (1992) *Curr. Opin. Struct. Biol.* 2, 388–393.
- [7] Hillar, A., Peters, B., Pauls, R., Loboda, A., Zhang, H., Mauk, A.G. and Loewen, P.C. (2000) *Biochemistry* 39, 5868–5875.
- [8] Regelsberger, G., Jakopitsch, C., Rüker, F., Krois, D., Peschek, G.A. and Obinger, C. (2000) *J. Biol. Chem.* 275, 22854–22861.
- [9] Regelsberger, G., Jakopitsch, C., Rüker, F., Switala, J., Loewen, P. and Obinger, C. (2000) *Biochem. Soc. Trans.*, in press.
- [10] Rouse, D.A., DeVito, J.A., Li, Z., Byer, H. and Morris, S.L. (1996) *Mol. Microbiol.* 22, 583–592.
- [11] Patterson, W.R. and Poulos, T.L. (1995) *Biochemistry* 34, 4331–4341.
- [12] Finzel, B.C., Poulos, T.L. and Kraut, J. (1984) *J. Biol. Chem.* 259, 13027–13036.
- [13] Rost, B. and Sander, C. (1993) *J. Mol. Biol.* 232, 584–599.
- [14] Fitzgerald, M.M., Musah, R.A., McRee, D.E. and Goodin, D.B. (1996) *Nat. Struct. Biol.* 3, 626–631.
- [15] Millis, C.D., Cai, D., Stankovich, M.T. and Tien, M. (1989) *Biochemistry* 28, 8484–8489.
- [16] Billman-Jacobe, H., Sloan, J. and Coppel, R.L. (1996) *FEMS Microbiol. Lett.* 144, 47–52.
- [17] Mate, M.J., Zámocký, M., Nykyri, L.M., Herzog, C., Alzari, P.M., Betzel, C., Koller, F. and Fita, I. (1999) *J. Mol. Biol.* 286, 135–149.
- [18] Gajhede, M., Schuller, D.J., Henriksen, A., Smith, A.T. and Poulos, T.L. (1997) *Nat. Struct. Biol.* 4, 1032–1038.