Iron center, substrate recognition and mechanism of peptide deformylase

Eubacterial proteins are synthesized with a formyl group at the N-terminus which is hydrolytically removed from the nascent chain by the mononuclear iron enzyme peptide deformylase. Catalytic efficiency strongly depends on the identity of the bound metal. We have determined by X-ray crystallography the Fe^{2+}, Ni^{2+} and Zn^{2+} forms of the *Escherichia coli* enzyme and a structure in complex with the reaction product Met-Ala-Ser. The structure of the complex, with the tripeptide bound at the active site, suggests detailed models for the mechanism of substrate recognition and catalysis. Differences of the protein structures due to the identity of the bound metal are extremely small and account only for the observation that Zn^{2+} binds more tightly than Fe^{2+} or Ni^{2+}. The striking loss of catalytic activity of the Zn^{2+} form could be caused by its reluctance to change between tetrahedral and five-fold metal coordination believed to occur during catalysis.

N-terminal formylation and subsequent deformylation appears to be a characteristic feature of eubacterial protein synthesis[1]. Peptide deformylase (PDF, EC 3.5.1.31), the enzyme responsible for cleaving the formyl group[2,3], is essential for bacterial growth[4] but seems to be absent in eucaryotic cells (see references cited in ref. 5) which makes it an attractive target for the design of new antibiotics. PDF is a mononuclear iron enzyme[5–7] of catalytic efficiency ($k_{cat}/K_M = 10^5\ M^{-1}s^{-1}$) comparable to that of other metallopeptidases. Fe^{2+} can be replaced by Ni^{2+} without significant loss of catalytic efficiency, whereas the Zn^{2+} form, prepared from the apoenzyme or by displacement of Fe^{2+} or Ni^{2+} proved virtually inactive[6].

Previous structural work on the Zn^{2+} form by NMR spectroscopy of the core domain (residues 1–147)[8] and X-ray crystallography of the full-length protein (residues 1–168) at 2.9 Å resolution[9] implicates Cys 90, His 132, His 136 and a water molecule as metal ligands. Although the overall fold of PDF differs from other metalloenzymes[8], the finding that the two histidine ligands reside in the sequence motif HEXXH[10], known[11,12] to be involved in zinc binding in metalloproteases like thermolysin[13], has led to the assumption of a similar enzyme mechanism[9]. Recently, we have solved the X-ray structure of the catalytically active Ni^{2+} form at 1.9 Å resolution in complex with the competitive inhibitor PEG identifying the substrate binding site[14]. Now we report the structures of the Fe^{2+}, Ni^{2+} and Zn^{2+} forms of the enzyme along with the structure in complex with the reaction product Met-Ala-Ser, and derive detailed models for the mechanism of substrate recognition and catalysis that will help in the design of new antibacterial agents.

Crystal structures

We have solved the crystal structures of the native Fe^{2+}-enzyme in complex with PEG (PDF–Fe/PEG), the nearly inactive Zn^{2+}-forms in the presence (PDF–Zn/PEG) and absence (PDF–Zn) of PEG, the catalytically active Ni^{2+}-form in complex with the product Met-Ala-Ser (PDF–Ni/MAS), and the Zn^{2+}-form cocrystallized with substrate formyl-Met-Ala-Ser (PDF–Zn/fMAS). The structures were solved by molecular replacement using the PDF–Ni/PEG structure[14] as starting model. Data sets and results of structure determinations are summarized in Table 1 which includes the entries of the previously described PDF–Ni and PDF–Ni/PEG crystals[14] to simplify the discussions.

All crystals possess space group symmetry C2 and contain three molecules in the asymmetric unit[6], named A, B and C. Cell constants are a = 140.8 Å, β = 63.4 Å, c = 86.8 Å, β = 120.6° for PDF/PEG, a = 143.4 Å, β = 64.0 Å, c = 84.6 Å, β = 123.2° for PDF, and a = 143.4 Å, β = 64.1 Å, c = 84.9 Å, β = 123.3° for PDF/(f)MAS crystals; they do not depend on the type of the bound metal. Typical deviations in bond length, bond angles, dihedral and improper angles from their ideal values in the atomic models are 0.011 Å, 1.2°, 25.5°, and 2.1° respectively. Refined[15] metal occupancies in each structure are at least 0.92 and have been reset to 1.0 in the final structures.

Peptide binding site

Analysis of cocrystals of PDF–Ni^{2+} with Met-Ala-Ser (PDF–Ni/MAS), a product of PDF catalysis, reveals clear electron density for the tripeptide in all three crystallographically independent monomers (Fig. 1*a*). Cα-traces of the three monomers together with bound Ni^{2+} and tripeptide are shown superimposed in Fig. 1*b*. Generally, the monomer structures are rather similar except for the more flexible regions 62–68 and 148–168 as observed for the PDF–Ni structures[14]. More subtle structural differences are found in regions far from direct crystal contacts, notably in the position of water W2. In contrast to all other structures, W2 is found in close proximity of the metal in monomers A and C of the PDF/(f)MAS structures which makes it a fifth ligand (Table 2). Ligation of W2 to the metal is accompanied by a shift of ~0.4 Å of the turn Gly 89–Leu 91 which leads to a more open metal center.

The tripeptide binding region of the PDF–Ni/MAS structure (Fig. 1*b*) overlaps well the the PEG binding region of the PDF–Ni/PEG structure (Fig. 1 of ref. 14). Since PEG is a competitive inhibitor of PDF activity[14], we can reasonably assume that Met-Ala-Ser binds to the enzyme in a similar manner as formyl-Met-Ala-Ser which is a substrate of the enzyme. The binding of formyl-Met-Ala-Ser can be modeled by shifting Met-Ala-Ser (Fig. 1*c*) towards the metal by 0.8 Å and replacing W2 by the carbonyl oxygen of the formyl group (Fig. 1*d*). This hypothetical model of the enzyme–substrate complex, generated using X-PLOR[15] with harmonic restraints to the starting structure, obeys stereochemical constraints and requires only minor changes in the active site region of the protein as observed for the PDF–Ni/MAS structure. Moreover, it explains why PDF has a much lower activity for peptides acetylated at the N-terminus[2,16,17]. In the model the additional methyl group (which substitutes for the hydrogen in the formyl group) would overlap with the side chains of Leu 91 and Leu 46 and the carbonyl oxygen of Gly 45, as noted by Chan *et al.*[9]. Also, displacement of Leu 91, which in our model of the enzyme mechanism would be part of the oxyanion hole, is likely to reduce catalytic efficiency. Thus, we speculate that substrate recognition by the enzyme may involve most of the interactions between the peptide and PDF as seen in the PDF–Ni/MAS structure.

Interactions of Met-Ala-Ser with PDF are depicted in Fig. 2. The N-terminal amino group of the peptide forms hydrogen bonds with Gly 45, Glu 133 and the waters W1 and W2, whereas the methionine side chain fits neatly into a hydrophobic pocket formed by residues Ile 44, Ile 86, Glu 88, Leu 125, Ile 128, Cys 129 and His 132 (Fig. 2*b*). Clearly, amino acids like

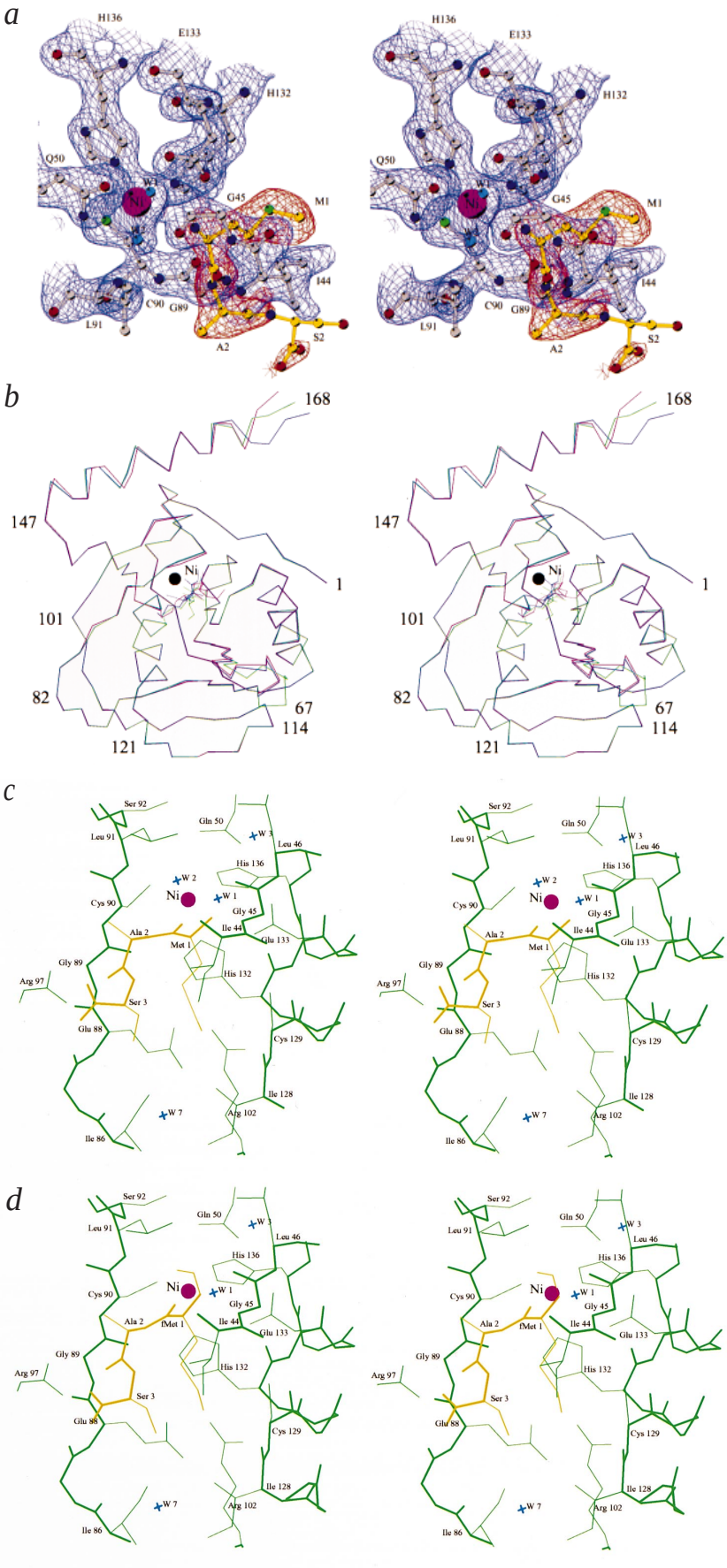

Fig. 1 ***a***, Omit map of Met-Ala-Ser in the PDF–Ni/MAS structure contoured at 1σ . Figure prepared with BobScript[25]. ***b***, Stereo-view of superimposed Cα-traces of the three crystallographically independent copies of peptide deformylase in complex with the reaction product Met-Ala-Ser. The numbers refer to amino acid residues. The transformations for optimal superposition were determined from equivalent Cα-atoms using molecule A as a reference and applied to the Ni^{2+} ion (marked as Ni) and the peptide as well. Molecule complexes A, B, C are shown in green, magenta, blue, respectively. ***c***, Active site of PDF–Ni (monomer A) with bound catalytic product Met-Ala-Ser, ordered water molecules W1, W2 and the Ni^{2+} ion. ***d***, A hypothetical model of PDF with bound substrate formyl-Met-Ala-Ser.

phenylalanine or ethionine instead of methionine also bind to PDF (unpublished results, S.S. and A.F.V.W.) and would also fit into the pocket pointing to significance for inhibitor design. Consistent with the enzyme's low sequence specificity, all hydrogen bonds involve only atoms of the peptide backbone. Moreover, as shown in Figs. 1*c*,*d* and 2*b*, the tripeptide is bound at a position that leaves enough room to accommodate bulky side chains for the second and following residues. Compared with longer substrates or with formyl-Met-amide, the minimal substrate formyl-Met has a much lower catalytic efficiency[7] which points to the importance of the interaction with Gly 89. Perhaps the presence of the hydrogen bond is communicated to the metal ligand Cys 90 thereby reducing the energy of the transition state. Interestingly, Cys 90 is located in the β-turn Gly 89–Leu 91 which together with water W2 is susceptible to weak crystal packing forces.

Motivated by the finding that PDF–Zn^{2+} is nearly catalytically inactive, we have determined its structure using crystals grown in the presence of the substrate formyl-Met-Ala-Ser (data set PDF–Zn/fMAS). Unfortunately, because of residual activity of the Zn^{2+} form in the crystal or because of small contamination by more active metallo forms of the enzyme, no density was found in the map that could by assigned to the formyl group; this is in sharp contrast to the clear density seen for the peptide product Met-Ala-Ser. In fact, both the Ni^{2+}- and Zn^{2+}-forms of PDF with the bound tripeptide are remarkably similar.

Metal center

The various metal forms of the PDF/PEG, PDF and PDF/(f)MAS crystals, each grown under nearly identical chemical conditions, allow us to study the influence of the metal on the structure. In agreement with earlier

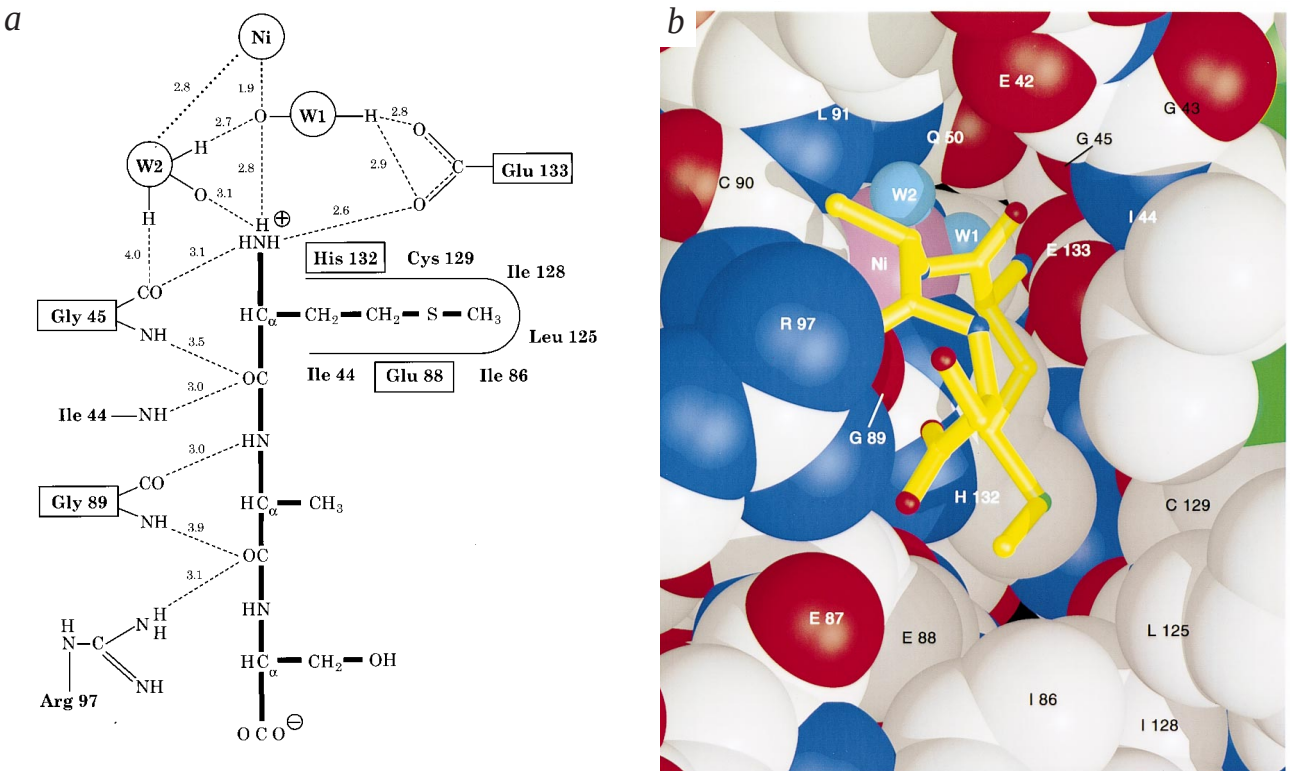

Fig. 2 Peptide deformylase in complex with the reaction product Met-Ala-Ser. ***a***, Peptide binding scheme; Gly 45, Glu 88, Gly 89, His 132 and Glu 133 are conserved residues. Dashed lines indicate hydrogen bonds. Mean distances between donor and acceptor atoms as observed in the three crystallographically independent monomers are given in Å . The distance between the N-terminal amino group of the peptide to the Ni^{2+}-ion is ~3.9 Å. ***b***, Protein atoms (white, carbon; red, oxygen; dark blue, nitrogen; green, sulphur) and catalytic metal (magenta, Ni^{2+}) are depicted as space-filling spheres. Met-Ala-Ser is in ball-and-stick representation with carbon atoms and bonds colored yellow. Water molecules W1, W2 are shown as small light blue spheres. Figure prepared using MOLSCRIPT[26] and Raster3D[27].

NMR[8] and X-ray studies of PDF–Zn[9], the metal ion in the structures presented here (except those of the peptide complexes) is tetrahedrally ligated. Ligands are the Sγ-atom of Cys 90, the Nε2-atoms of His 132 and His 136 in the HEXXH motif, and an oxygen atom of the group W1 modeled as a water molecule. All ligands, occupying well defined positions in the electron density maps, are precisely aligned by an intricate network of hydrogen bonds involving conserved residues of the enzyme family[14]. As shown in Table 2, the arrangement of the ligands, in particular of W1, deviates significantly from ideal tetrahedral symmetry in all structures.

Considering the variability found in the metal ligand positions for the three crystallographically independent molecules in each structure, it appears that crystal packing effects or random errors in the atomic models obscure all structural differences between the Fe^{2+}, Ni^{2+} and Zn^{2+} forms of the enzyme. In fact, the influence of the metal on the structure is extremely small and barely detectable. It is known that Zn^{2+} is bound more tightly by the enzyme than the other metals[6]. This effect becomes visible in the structures if the tetrahedral volumes, spanned by the four metal ligands, are compared as shown in Table 2. In all cases, volumes around the Zn^{2+} ion are smaller than in the presence of Fe^{2+} or Ni^{2+}.

The extreme similarity among the structures in each of the three groups of crystals has also been confirmed by Fourier analysis of differences between observed structure factors. This analysis is known to be very sensitive, yet robust towards errors in the atomic model. Moreover, we have excluded the metal from calculating model phases. In all cases, the dominant maxima in the resulting difference maps varied between 10σ and 36σ and account for the difference of two or four electrons between the compared metal forms. Minima and minor maxima are predominantly found in the vicinity of the ligand Cys 90. These structural differences are remarkably small and of a magnitude well below the contribution expected from an ordered water molecule that may be present, absent or displaced due to the metal type. We conclude that the enormous differences in catalytic activity of the metal forms of the enzyme ultimately result from the tighter binding of the Zn^{2+} ion as compared to Fe^{2+} or Ni^{2+}.

Catalytic model

It has been noted that the active site of PDF is similar to thermolysin although its overall fold differs from other metalloproteases[8]. Both proteins use two histidines from the common HEXXH sequence motif for metal ligation and require the glutamate residue for catalytic activity. In fact, mutations of Glu 133 to Ala, Asp or Gln were found to eliminate catalytic activity of PDF[18,19]. This has led to the proposal of a mechanism of deformylation[9] that resembles the mechanism suggested for thermolysin[20]. Below we propose a model of the catalytic cycle of PDF that is consistent with the available

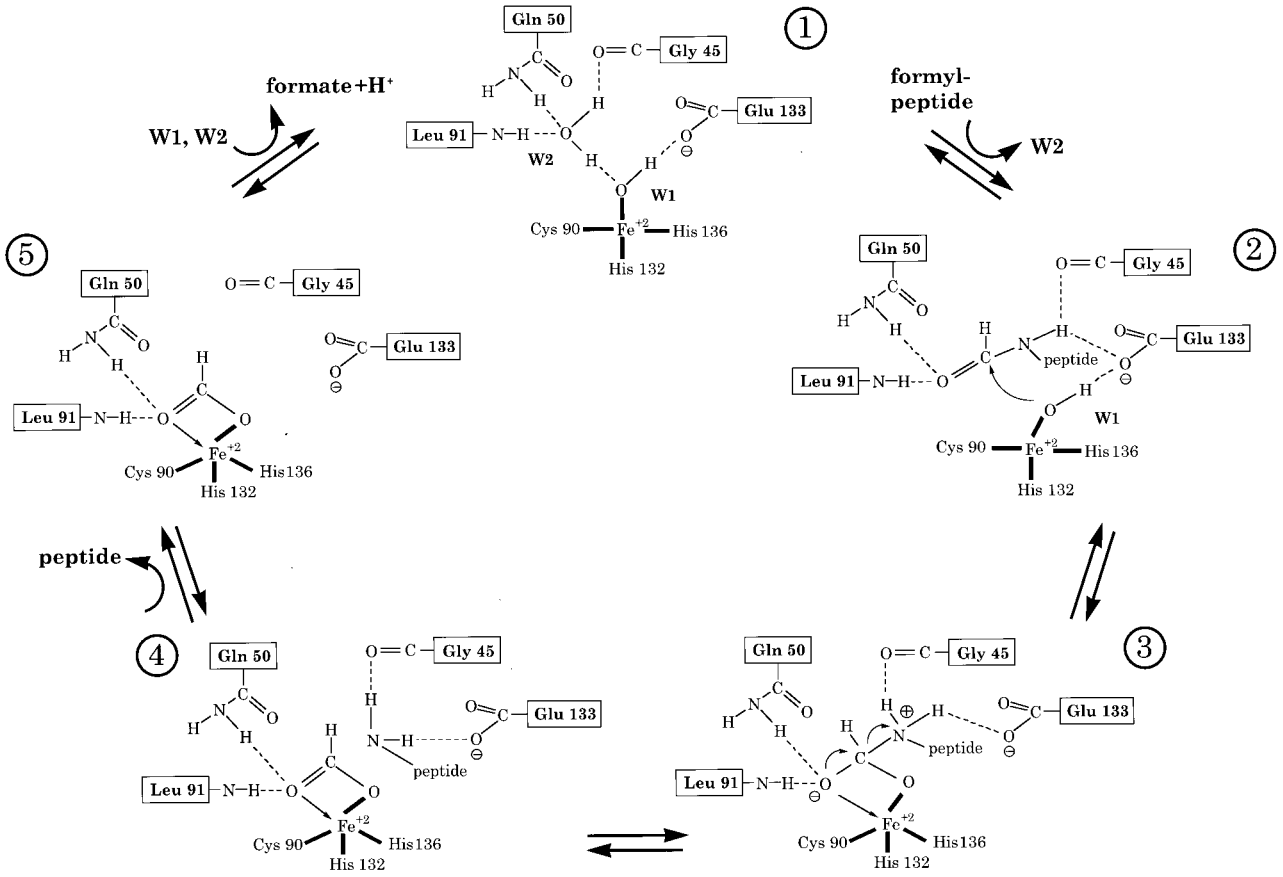

Fig. 3 Proposed reaction cycle of peptide deformylase.

structures and biochemical data. Besides being more detailed, it differs from the model of Chan *et al.*[9] by the inclusion of water W2 and the role of residues Leu 91 and Gln 50 which form an oxyanion hole as described for serine and cysteine proteases[21].

The reaction cycle could proceed in the following steps (Fig. 3). Step 1 represents the initial state of PDF as found in the PDF–Ni and PDF–Zn structures. In the next step the formylated peptide binds as shown in the hypothetical model of the enzyme–substrate complex (Fig. 1*d*) thereby replacing water W2 by the carbonyl oxygen of the formyl group. The carbonyl oxygen is polarized by hydrogen bonds to the amide of Leu 91 and the side chain of Gln 50. This supports the nucleophilic attack of W1 (probably a deprotonated water) on the carbonyl carbon of the formyl group that leads to the transition state of the reaction as depicted in step 3. The carbonyl oxygen is tetrahedrally ligated by the metal, carbonyl carbon, side chain amide of Gln 50, and main chain amide of Leu 91. This arrangement of amide hydrogen bond donors resembles the oxyanion hole in serine and cysteine proteases for stabilizing the transition state[21]. The absence of ionizable groups in the oxyanion hole is well in line with the broad pH-activity profile of PDF–Fe which is nearly constant between pH 6.1 and 11.2 (ref. 7). The electronic state of the attacked carbonyl carbon changes from sp^2 to sp^3 accompanied by a transition from the tetrahedral to a five-coordinate metal center, this transition being strongly inhibited in the case of PDF–Zn due to the tighter binding of this metal ion as compared with Fe^{2+} or Ni^{2+}. Probably the rearrangement of the metal ligands is supported indirectly by the hydrogen bond between Gly 89 and the peptide substrate as mentioned above.

The proton of W1 is transferred with the help of Glu 133 to the amide at the N-terminus of the peptide, the added positive charge making the nitrogen suitable as a leaving group. Subsequent bond cleavage leads to the ternary enzyme–formate–peptide complex shown in step 4. Here, the formate is bound to the five-coordinated metal center and the free N-terminus of the peptide is hydrogen bonded to Glu 133 as found in the PDF–Ni/MAS structure. The reaction proceeds by dissociation of the peptide which leaves an activated enzyme–formate complex as shown in step 5. The existence of such a complex is consistent with biochemical data which show that PDF can transfer the formyl group from one formyl peptide to another peptide by a 'ping-pong' mechanism[7]. The reaction cycle closes with the release of formate and the uptake of two water molecules, W1 and W2. It is very likely that formate release proceeds by attack of a water molecule on the metal rather than on the formate carbon atom. Evidence comes from the lack of ^{17}O exchange during incubation of ^{17}O formate in the presence of PDF[7].

Table1 Data collection and refinement statistics

Compound	PDF–Fe/PEG	PDF–Ni/PEG	PDF–Zn/PEG	PDF–Ni	PDF–Zn	PDF–Ni/MAS	PDF–Zn/fMAS
PDB accession number	1BSZ	1ICJ	1BS4	1BS7	1BS5	1BS6	1BS8
Outer shell (Å)	2.01–1.9	2.01–1.9	2.01–1.9	2.64–2.5	2.64–2.5	2.23–2.1	2.33–2.2
Measured reflections	104,453	281,561	188,577	90,069	130,772	118,387	117,576
Unique reflections	50,232	51,138	51,676	21,929	22,399	37,363	32,674
Completeness (%)	96.1	97.9	99.2	97.2	99.3	98.0	99.2
R_{sym} (%)[1]	5.9	7.5	4.8	8.6	7.6	5.6	8.0
$\langle I/\sigma \rangle$ in outer shell	2.2	2.7	2.8	2.4	1.9	1.8	2.0
R-factor (%)[2]	19.7	19.3	19.3	20.3	20.8	20.6	20.4
R_{free} (%)[3]	24.7	23.8	24.0	27.2	25.8	25.8	26.1
Coordinate error (Å)	0.23	0.22	0.22	0.33	0.33	0.29	0.29
Protein atoms	4,122	4,122	4,122	4,038	4,038	4,107	4,107
Alternate locations	84	84	84	0	0	69	69
Heterogen atoms	97	97	97	13	13	73	73
Solvent molecules	211	211	211	132	132	165	165

[1] $R_{sym} = \Sigma_h \Sigma_i |I_{hi} - I_h| / \Sigma_{hi} I_{hi}$ where h are unique reflection indices and I_{hi} are the intensities of symmetry equivalent reflections giving a mean value of I_h.
[2] $R = \Sigma |F_{obs}-F_{model}| / \Sigma F_{obs}$ where F_{obs} and F_{model} are observed and atomic model structure factor amplitudes respectively.
[3] R-factor calculated for 10% of randomly chosen reflections that were excluded from the refinement

Methods

Preparation of all crystals except PDF–Zn/fMAS, enzymatic activity tests and determinations of metal content were performed as described[6]. PDF–Zn/fMAS crystals were grown under conditions similar to those for PDF–Ni/MAS with 20 mM formyl-Met-Ala-Ser instead of MAS and in addition 500 mM ammonium formate and 100 mM KF.

Diffraction data were collected at room temperature by the rotation method and recorded by an electronic area detector (X-rays: CuKα, focused by Franks double-mirror optics; generator: GX-18, Elliot/Enraf-Nonius, Delft, operated at 35 kV, 50 mA; detector: X100, Siemens/Nicolet; crystal to detector distance: 10 cm; rotation/image: 0.0417° or 0.0833°). Integrated intensities were extracted from the rotation images by the program package XDS[22] which includes routines for scaling data from several crystals as well as for space group determination from the observed diffraction pattern[23].

All structures were solved by molecular replacement using the PDF–Ni/PEG structure[14] as starting model, followed by model correction[24] and refinement[15]. For comparison of two isomorphous crys-

Table 2 Metal coordination[1]

Compound	Bond angles[2] (°)						Bond length[3] (Å)					V[4] (Å^3)
	α_{12}	α_{13}	α_{14}	α_{23}	α_{24}	α_{34}	d_1	d_2	d_3	d_4	d_5[5]	
PDF–Fe/PEG	119	105	128	103	103	93	2.3	2.1	2.1	2.2	3.7	5.0
	108	112	131	101	107	95	2.3	2.3	2.1	2.0	3.7	5.0
	116	111	132	103	104	85	2.2	2.2	2.2	2.3	3.6	5.1
PDF–Ni/PEG	116	102	133	100	98	103	2.2	2.1	2.1	1.9	3.7	4.2
	111	105	134	104	100	98	2.3	2.2	2.0	2.0	3.6	4.6
	115	104	133	102	102	95	2.3	2.1	2.1	2.1	3.7	4.7
PDF–Zn/PEG	118	102	125	104	101	104	2.3	2.1	2.0	1.8	3.9	4.1
	117	105	124	104	100	103	2.3	2.0	2.1	1.9	3.8	4.3
	116	103	126	106	101	102	2.2	2.1	2.1	2.0	3.8	4.6
PDF–Ni	127	107	137	96	95	73	2.1	2.1	2.4	2.5	3.9	4.8
	115	116	137	107	92	84	2.2	2.2	2.0	1.9	3.7	4.0
	114	110	149	99	92	79	2.2	2.2	2.1	2.2	n.o.	4.2
PDF–Zn	118	102	131	107	101	93	2.1	2.0	2.2	2.2	4.1	4.7
	118	107	127	109	96	96	2.1	2.1	2.1	1.8	3.7	4.0
	115	107	143	105	90	91	2.1	2.1	2.1	1.8	n.o.	3.8
PDF–Ni/MAS	121	105	157	98	81	77	2.4	2.2	2.3	2.2	2.3	4.3
	122	111	142	100	88	84	2.3	2.2	2.2	1.9	3.4	4.3
	115	98	155	102	89	82	2.4	2.2	2.2	2.0	2.7	4.2
PDF–Zn/fMAS	117	109	156	101	83	77	2.1	2.3	2.1	1.9	2.3	3.6
	116	104	133	113	93	96	2.2	2.1	2.0	1.9	3.7	4.1
	115	109	154	102	87	77	2.2	2.0	2.0	2.3	2.2	3.8

[1] Bond angles, bond lengths and volumes are given for monomers A, B, C in the asymmetric unit of each compound.
[2] Bond angle α_{ij} is the angle between ligand i, metal and ligand j (i,j = 1,…,4).
[3] Bond lengths d_1, d_2, d_3, d_4 are distances from the central metal to its four ligands, Cys 90 (Sγ), His 132 (Nε2), His 136 (Nε2), and water molecule W1 respectively.
[4] V is the volume of the tetrahedron formed by the four metal ligands.
[5] d_5 is the distance between water molecule W2 and the metal. W2 is not observed in monomer C of the PDF–Ni and PDF–Zn structures.

tal structures difference electron density maps were computed from structure factors calculated as difference between the measured amplitudes of each data set multiplied by a phase factor derived from the atomic model of one of the crystals.

Coordinates. Protein Data Bank accession numbers are given in Table 1.

Acknowledgments
We thank D. Madden, K. Fritz-Wolf and K. Scheffzek for critical discussions and help at various stages of the project, H. Wagner for excellent maintenance of the X-ray facilities at the MPI Heidelberg, I. Dehof for help with the figures, and K. Holmes and J. Knappe for continuous support.

Andreas Becker[1], Ilme Schlichting[2], Wolfgang Kabsch[1], Dieter Groche[3], Sabine Schultz[3] and A. F. Volker Wagner[3]

[1]Max-Planck-Institut für medizinische Forschung, Abteilung Biophysik, Jahnstraße 29, 69120 Heidelberg, Germany. [2]Max-Planck-Institut für molekulare Physiologie, Abteilung Physikalische Biochemie, Rheinlanddamm 201, 44139 `Dortmund, Germany. [3]Biochemie-Zentrum Heidelberg, Ruprecht-Karls Universität, Im Neuenheimer Feld 501, 69120 Heidelberg, Germany.
Correspondence should be addressed to W.K. *email: kabsch@mpimf-heidelberg.mpg.de*

Received 22 June, 1998; accepted 11 September, 1998.

1. Kozak, M. *Microbiol. Rev.* **47**, 1–45 (1983).
2. Adams, J.M. *J. Mol. Biol.* **33**, 571–589 (1968).
3. Ball, L. A. & Kaesberg, P. *J. Mol. Biol.* **79**, 531–537 (1973).
4. Mazel, D., Pochet, S. & Marlière, P. *EMBO J.* **13**, 914–923 (1994).
5. Rajagopalan, P.T.R., Yu, X.C. & Pei, D. *J. Am. Chem. Soc.* **119**, 12418–12419 (1997).
6. Groche, D. *et al. Biochem. Biophys. Res. Commun.* **246**, 342–346 (1998).
7. Groche, D. Ph.D. thesis, Universität Heidelberg. Charakterisierung des Eisenzentrums und des Katalysemechanismus von Peptid-Deformylase aus *Escherichia coli* (1995).
8. Meinnel, T., Blanquet, S. & Dardel, F. *J. Mol. Biol.* **262**, 375–386 (1996).
9. Chan, M.K. *et al. Biochemistry* **36**, 13904–13909 (1997).
10. Meinnel, T. & Blanquet, S. *J. Bacteriol.* **175**, 7737–7740 (1993).
11. Vallee, B.L. & Auld, D.S. *Biochemistry* **29**, 5647–5659 (1990).
12. Jongeneel, C.V., Bouvier, J. & Bairoch, A. *FEBS Lett.* **242**, 211–214 (1989).
13. Holmes, M.A. & Matthews, B.W. *J. Mol. Biol.* **160**, 623–629 (1982).
14. Becker, A., Schlichting, I., Kabsch, W., Schultz, S. & Wagner, A.F.V. *J. Biol. Chem.* **273**, 11413–11416 (1998).
15. Brünger, A.T. *J. Mol. Biol.* **203**, 803–816 (1988).
16. Wei, Y. & Pei, D. *Anal. Biochem.* **250**, 29–34 (1997).
17. Meinnel, T. & Blanquet, S. *J. Bacteriol.* **177**, 1883–1887 (1995).
18. Meinnel, T., Lazennec, C. & Blanquet, S. *J. Mol. Biol.* **254**, 175–183 (1995).
19. Meinnel, T., Lazennec, C., Villoing, S. & Blanquet, S. *J. Mol. Biol.* **267**, 749–761 (1997).
20. Matthews, B.W. *Acc. Chem. Res.* **21**, 333–340 (1988).
21. Ménard, R. & Storer, A.C. *Biol. Chem. Hoppe-Seyler* **373**, 393–400 (1992).
22. Kabsch, W. *J. Appl. Crystallogr.* **21**, 916–924 (1988).
23. Kabsch, W. *J. Appl. Crystallogr.* **26**, 795–800 (1993).
24. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. *Acta crystallogr. A* **47**, 110–119 (1991).
25. Esnouf, R.M. *J. Mol. Graphics* **15**, 132–134 (1997)
26. Kraulis, P.J. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
27. Merritt, E.A. & Murphy, M.E.P. *Acta Crystallogr.* D **50**, 869–873 (1994).

MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily

The DNA glycosylase MutY, which is a member of the Helix-hairpin-Helix (HhH) DNA glycosylase superfamily, excises adenine from mispairs with 8-oxoguanine and guanine. High-resolution crystal structures of the MutY catalytic core (cMutY), the complex with bound adenine, and designed mutants reveal the basis for adenine specificity and glycosyl bond cleavage chemistry. The two cMutY helical domains form a positively-charged groove with the adenine-specific pocket at their interface. The Watson-Crick hydrogen bond partners of the bound adenine are substituted by protein atoms, confirming a nucleotide flipping mechanism, and supporting a specific DNA binding orientation by MutY and structurally related DNA glycosylases.

Reactive oxygen species that are generated during aerobic respiration and immune response cause oxidative lesions in DNA that are repaired by the base excision repair pathway[1]. Oxidative DNA damage is implicated in mutagenesis, carcinogenesis and aging[2]. Hydroxyl radicals rapidly react with the guanine C8 position producing a steady-state level of mutagenic 8-oxoguanines (^{8}OG) in cells[3]. Oxidation also creates 8-oxo-GTP that, like ^{8}OG in template DNA, is generally mispaired with adenine by replicative polymerases, creating A–T to C–G and G–C to T–A transversion mutations[4]. In *Escherichia coli*, three key enzymes, MutT, MutM (FPG), and MutY, function synergistically to protect cells from the deleterious effects of guanine oxidation[5,6]. MutT removes 8-oxo-GTP[7] to reduce its misincorporation by polymerases, and MutM excises ^{8}OG from ^{8}OG–C DNA base pairs[5]. Only the mismatched-adenine specific glycosylase MutY recognizes the mutational intermediate ^{8}OG–A mispair[5], and loss of the *E. coli mutY* gene leads to elevated mutation rates[8]. Mutations in the three analogous human enzymes may increase cancer susceptibility due to elevated tranversion frequencies[6,9]. Importantly, among these three enzymes, only the sequence of MutY is conserved from bacteria to humans suggesting its fundamental importance in DNA repair[10]. Moreover, as only the ^{8}OG in ^{8}OG–A mispairs is correctly coding, restoration of the original G–C base pair requires a two step process. First, MutY excises adenine mispaired with ^{8}OG, creating an abasic site that is processed by the multi-enzyme base excision repair pathway[5,11] to yield an ^{8}OG–C base pair. Second, MutM removes ^{8}OG to initiate a final cycle of base excision repair that restores the original G–C base pair.

Cloning and sequencing of *mutY*[12] and other DNA repair genes revealed a base excision repair glycosylase superfamily[13], which includes both pure DNA glycosylases, like 3-methyladenine glycosylase II (AlkA), and DNA glycosylase/lyases, like endonuclease III (EndoIII). Crystal structures of EndoIII[14] and AlkA[15,16] show that the most conserved superfamily structural element is the Helix-hairpin-Helix (HhH) motif[17]. However, no crystal structures with bound DNA or with bound target bases exist for any member of the HhH glycosylase superfamily. Thus, the structural and chemical basis of DNA binding, substrate specificity and glycosyl bond cleavage by MutY and other members of the HhH DNA glycosylase superfamily is largely uncharacterized.

MutY structure

Here we characterize the structural chemistry of MutY and HhH glycosylase catalysis and recognition of damaged bases using: three designed MutY mutants, MutY inhibition by relevant base analogs, and high-resolution crystal structures of the fully-active, 26,000 M_r catalytic core of MutY (cMutY) that lacks the 13,000 M_r C-terminal region of unknown function[18]. Three cMutY structures were determined: wild type cMutY (1.4 Å resolution), an inactive site-directed