Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Probing the oxygen-binding site of the human formylglycine-generating enzyme using halide ions

Dirk Roeser,[a] Bernhard Schmidt,[b] Andrea Preusser-Kunze[b] and Markus G. Rudolph[a]*

[a]Department of Molecular Structural Biology, University of Göttingen, D-37077 Göttingen, Germany, and [b]Department of Biochemistry II, University of Göttingen, D-37077 Göttingen, Germany

Correspondence e-mail: markus.rudolph@bio.uni-goettingen.de

Received 27 February 2007
Accepted 1 March 2007

PDB References: FGE–Br, 2hi8, r2hi8sf; FGE–I, 2hib, r2hibsf.

The catalytic residue in sulfatases is a unique formylglycine that is post-translationally generated by oxidation of a cysteine or serine precursor. Molecular oxygen oxidizes the cysteine precursor in eukaryotic sulfatases, a reaction that is catalysed by the formylglycine-generating enzyme FGE. Previously, FGE was crystallized in complex with a chloride ion which, based on its similar polarizability and hydrophobicity, indicates the site of molecular oxygen binding. Here, two structures of FGE in complex with bromide and iodide were determined in order to further delineate the volume and stereochemical restraints of the oxygen-binding site for potential reaction intermediates. Anomalous difference density maps unambiguously assigned the nature of the halide ions. Unexpectedly, data collected at a wavelength of 1.54 Å from the iodide-containing crystal and data collected at a wavelength of 0.8 Å from a bromide-containing crystal were sufficient for SIRAS phasing.

1. Introduction

Eukaryotes employ sulfatases to hydrolyse sulfate esters such as glycosaminoglycans, sulfolipids and steroid sulfates (Hopwood & Ballabio, 2001). All eukaryotic sulfatases described to date contain a unique formylglycine (FGly) as the catalytic residue that is generated by oxidation of a cysteine precursor (Fig. 1). This unusual post-translational modification is catalysed in the endoplasmatic reticulum by the formylglycine-generating enzyme (FGE; Cosma *et al.*, 2003; Dierks *et al.*, 2003). FGE forms the bottleneck in the activation of all sulfatases and inactivity of FGE leads to a severe reduction or lack of all sulfatase activity, termed multiple sulfatase deficiency. What is striking about FGE is that this oxidase lacks any cofactors known to activate molecular oxygen such as a haem group, FAD or an Fe–S cluster. Thus, FGE belongs to the family of cofactor-less oxidases, which also includes quinone-forming monooxygenases involved in polyketide antibiotic synthesis and bacterial dioxygenases such as the carbon monoxide-forming 2,4-dioxygenases involved in

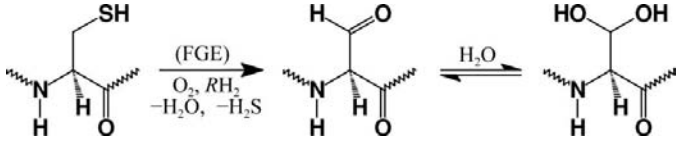

Figure 1
Post-translational modification of eukaryotic sulfatases. The scheme outlines the FGE-catalysed Cys oxidation to FGly by molecular oxygen and a reducing agent (RH_2). The identity of RH_2 *in vivo* has yet to be determined. Spontaneous hydration of FGly yields a geminal diol or 'hydroxylated serine', which acts as the nucleophile in sulfatases.

the degradation of *N*-heteroaromatic compounds (Fetzner, 2002). It is currently unknown how cofactor-less oxidases can activate molecular oxygen and at the same time prevent side reactions originating from reactive oxygen species.

FGE functions as a monooxygenase, using molecular oxygen as a substrate to oxidize cysteine to formylglycine by an intricate mechanism that may involve a cysteine-bound hydroperoxide or a sulfenic acid as an intermediate. The first insights into the FGE catalytic mechanism came from several crystal structures of the human enzyme (Dierks *et al.*, 2005; Roeser *et al.*, 2005, 2006). Structures of the apo enzyme revealed a furrow on the molecular surface that is bordered by a redox-active pair of cysteine residues (Cys336 and Cys341). Depending on the redox potential of the environment, these cysteine residues exist in the thiol or in the disulfide form. Mutagenesis of Cys336 to serine not only inactivated the enzyme but also enabled the trapping of an FGE–substrate complex in which Cys341 formed an intermolecular disulfide bond with a minimal peptide derived from a substrate sulfatase. Furthermore, mutation of Ser333 to alanine also abolished enzyme activity, establishing an important role for Ser333 during catalysis.

The close stereochemical proximity of Ser333, Cys336 and Cys341 allowed the assignment of an oxygen-binding site next to these residues and beneath the disulfide-bound substrate. In line with this assignment, a chloride ion is present in the FGE–substrate complex structure (Roeser *et al.*, 2006). Although negatively charged, owing to their polarizability and roughly similar volume compared with molecular oxygen, halide ions seem to be adequate surrogates for probing the oxygen-binding site in FGE. Here, we describe two crystal structures of FGE in complex with bromide and iodide ions that help in further delineating the mode of oxygen binding. When compared with the chloride-containing FGE–substrate complex, the bromide and iodide structures define the maximum space available for a bound oxygen molecule and for possible oxidation intermediates in FGE.

Table 1
Crystallographic data collection and analysis.

Values in square brackets and parentheses are for the lowest and highest resolution shells, respectively.

	In-house, iodide (2hib)	Synchrotron, bromide (2hi8)
Wavelength (Å)	1.54	0.8015
Source/detector	MicroMax 007/MAR 345	X13/MAR CCD 165
Temperature (K)	100	100
Resolution range (Å)	[50–4.29] 50–2.0 (2.06–2.00)	[50–3.51] 50–1.64 (1.69–1.64)
Space group	$P2_12_12$	$P2_12_12$
Unit-cell parameters (Å)	a = 61.9, b = 109.1, c = 43.5	a = 61.2, b = 108.8, c = 43.1
Measured reflections	[8134] 72271 (3687)	[12951] 124316 (8373)
Unique reflections	[2257] 20270 (1570)	[3896] 35818 (2823)
Anomalous multiplicity	[2.1] 2.0 (1.6)	[1.9] 1.9 (1.6)
Anomalous completeness (%)	[98.5] 97.2 (81.4)	[99.3] 96.6 (74.4)
Mosaicity (°)	0.6	0.3
$R_{\rm sym}$† (%)	[7.0] 13.2 (41.3)	[3.3] 7.5 (39.6)
$R_{\rm p.i.m.}$‡ (%)	[4.4] 8.0 (29.8)	[2.0] 4.7 (27.2)
$R_{\rm anom}$§ (%)	[5.2] 8.6 (33.2)	[2.2] 5.1 (27.8)
Average $I/\sigma(I)$	[15.1] 8.9 (1.9)	[29.8] 15.0 (3.0)

† $R_{\rm sym} = 100 \times \sum_h \sum_i |I_i(h) - \langle I(h)\rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ is the ith measurement of reflection h and $\langle I(h)\rangle$ is the average reflection intensity. Friedel pairs were merged for calculation of $R_{\rm sym}$. ‡ $R_{\rm p.i.m.} = 100 \times \sum_h (N-1)^{-0.5} \sum_i |I_i(h) - \langle I(h)\rangle| / \sum_h \sum_i I_i(h)$, where N is the multiplicity. § $R_{\rm anom} = 100 \times \sum_h |I_i(h) - I_i(-h)| / \sum_h \langle I_i(h)\rangle$, where $I_i(h)$ and $I_i(-h)$ are Friedel pairs.

2. Experimental

2.1. Protein purification, crystallization, data collection and analysis

FGE was cloned and produced as described by Preusser-Kunze *et al.* (2005). For all studies the FGE active-site mutant Cys336Ser was used, which has been shown to allow trapping of a covalent enzyme–substrate complex (Roeser *et al.*, 2005). The pentameric peptide CTPSR was used as a substrate for cocrystallization with FGE. *In situ* limited proteolysis using elastase [1:1000(*w*:*w*)] was essential to generate crystals from a 1:1 mixture of 10 mg ml^{-1} FGE–peptide complex and a solution containing 20–25% PEG 4000, 0.1 *M* Tris–HCl pH 8.0–9.0, 0.2–0.3 *M* $CaCl_2$ (Roeser *et al.*, 2005). Bromide and iodide complexes of FGE were generated by substituting $CaBr_2$ and CaI_2 for $CaCl_2$ and using Tris–acetate as buffer. The crystals were directly vitrified in a 100 K nitrogen-gas stream without additional cryoprotectant. Data for the bromide complex were collected on EMBL beamline X13 at DESY, while data for the iodide complex were collected in-house (Table 1). The *HKL* suite of programs (Otwinowski & Minor, 1997) was used for data reduction. No significant radiation damage was observed in the bromide and iodide data as judged by the comparison of the overall $R_{\rm sym}$ value with the R values for each image and by the constant B value and scale factors for each image. The program *RMERGE* (Weiss, 2001) was used for calculation of the precision-indicating merging R factor $R_{\rm p.i.m.}$ The anomalous R factor $R_{\rm anom}$ was calculated in *SCALEPACK* by comparing I^+ and I^- reflections after merging symmetry-equivalent reflections. For calculation of anomalous ΔF values in *XPREP*, symmetry-related reflections were kept separate during data reduction in *SCALEPACK* to allow for local scaling. The data were analyzed for anomalous signal content by comparing the anomalous signal-to-noise ratio based on the variances of F^+ and F^- [$\Delta F/\sigma(\Delta F)$]. No significant anomalous signal beyond about 4 Å resolution was detectable in the bromide and iodide data (Figs. 2*a* and 2*b*). In addition, the correlation coefficient (CC) between the signed anomalous differences as calculated by *SHELXC* (Pape & Schneider, 2004) for subsets of the nonmerged derivative data set falls below the meaningful value of 30% at about 4 Å resolution for the bromide data and 5 Å resolution for the iodide data (Figs. 2*a* and 2*b*). Another statistic to address the anomalous signal content in the diffraction data is the ratio of the precision-indicating merging R factor $R_{\rm p.i.m.}$ and the

anomalous R factor $R_{\rm anom}$. As $R_{\rm p.i.m.}$ is a measure of the noise in the data, $R_{\rm p.i.m.}/R_{\rm anom}$ can serve as an estimate of the anomalous signal-to-noise ratio (Weiss *et al.*, 2001). In line with the indicators applied above, the ratios of 1.1 and 1.2 for the iodide and bromide data, respectively (Table 1), indicate a poor anomalous signal in the data. In addition, Patterson plots using the anomalous differences calculated at high-resolution limits of 4–2 Å for the iodide data and high-resolution limits of 4–1.6 Å for the bromide data were essentially featureless (not shown). In a *post structuram* analysis the few small Patterson peaks that could represent halide sites were shown not to correlate with the calculated peaks for the correct halide substructures, again indicating no strong anomalous signal in the halide data.

2.2. Structure determination by molecular replacement

Since a high-resolution FGE crystal structure was available that was isomorphous with the halide data, molecular replacement was trivial and the bromide and iodide FGE structures could be determined by direct refinement of the chloride-containing coordinates (PDB code 2aij) against the new data sets. The presence of halide ions was ascertained by large peaks in $F_{\rm o} - F_{\rm c}$ and anomalous difference maps using the phases of the protein and the ΔF values from the halide data sets as Fourier coefficients. The presence of clear density for the halide ions motivated an approach of independent phasing of the bromide and iodide FGE structures. Indeed, despite the disappointing anomalous signal indicators discussed above, both halide structures could be phased by SIRAS and these unbiased electron-density maps were then used for model building (see below).

2.3. Iodide and bromide SIRAS phasing of FGE

The presence of 0.5–0.6 *M* halide ions in the crystallization setups rendered their binding to the protein surface likely, as the concentrations are in the same range as those reported for halide soaks used for SAD or SIRAS phasing (Dauter *et al.*, 2000). Data collection was not designed for anomalous phasing and no care was taken to optimize the anomalous signal by choice of wavelength, collection of Friedel pairs close in time or reciprocal space or maximization of the multiplicity. It nevertheless seemed an interesting question whether the data could be phased by a halide-based anomalous signal, however weak, which would deliver independent and complementary information to the results from molecular replacement such as unbiased protein phases and the location and relative occupancy of the anomalous scatterers.

The calculated f'' value for iodine using Cu $K\alpha$ radiation is 6.91 e and f'' of bromine using 0.8 Å radiation is 2.95 e. Both data sets have a low overall anomalous multiplicity of ∼2 (Table 1). Not surprisingly, SAD and SIR phasing attempts using various cutoffs for the high resolution (6 Å to the respective high-resolution limit in 0.5 Å steps) and E values (values between 1.2 and 1.8 were tried) were unsuccessful in both the bromide and iodide cases. Also, a pseudo-MAD experiment (Evans & Bricogne, 2002) using the bromide and iodide data alone (MIRAS without a native data set) was ineffective. However, SIRAS phasing was successful in both cases when a high-resolution (1.54 Å) chloride-containing data set (Roeser *et al.*, 2005) was used as the native. This native data set was collected in-house and Friedel pairs were merged to eliminate any anomalous signal in this data set and to base all substructure calculations exclusively on the anomalous signal of the halide data. SIRAS substructure calculations were performed with *SHELXD* (Schneider & Sheldrick, 2002) using data truncated at an optimistically high resolution limit of 2.3 Å. However, the halide substructures were also found with high-resolution cutoffs as low as 3.5 Å. *SHELXE* (Sheldrick, 2002) was used to resolve the phase ambiguity during density modification and phase extension to the high-resolution limit of 1.54 Å of the chloride-containing data set (Fig. 3).

Both SIRAS experiments returned a similar substructure with several conserved halide sites. The resulting maps after phase extension were easily interpreted by *ARP*/*wARP* (Collaborative Computational Project, Number 4, 1994) to yield virtually complete models, which were finalized in *Coot*

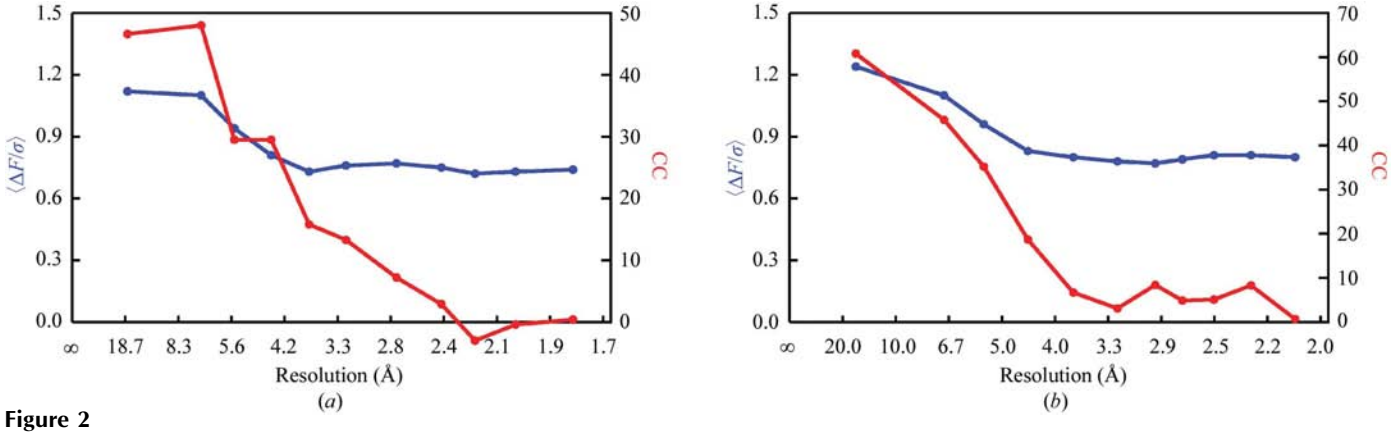

Figure 2
Anomalous signal of the bromide and iodide FGE data sets. (*a*) Anomalous signal and correlation coefficient (CC) for the signed anomalous differences in partial bromide data sets. A value of 0.8 for $\Delta F/\sigma(\Delta F)$ denotes the noise level and CC values below 30% are deemed insignificant for improving the phasing success. (*b*) Anomalous signal and CC for the iodide data set.

Table 2
Refinement statistics and model analysis.

Values in parentheses are for the highest resolution shell.

	Iodide (2hib)	Bromide (2hi8)
Resolution range (Å)	33.8–2.00 (2.05–2.00)	27.2–1.64 (1.68–1.64)
No. of reflections/No. in test set	19211/963	33959/1726
R_{cryst}† (%)	17.2 (23.4)	15.3 (24.0)
R_{free}† (%)	23.8 (29.4)	19.1 (31.9)
No. of residues	276	274
No. of protein atoms	2221	2310
No. of water O atoms	259	436
No. of sugar residues	2	2
No. of Ca^{2+} ions	2	2
No. of halide ions	11	16
Coordinate error‡ (Å)	0.118	0.060
R.m.s.d. bonds (Å)	0.011	0.010
R.m.s.d. angles (°)	1.26	1.31
R.m.s.d. dihedrals (°)	6.7	6.5
Ramachandran plot§		
Core (%)	87.5	88.1
Allowed	11.6	11.0
Generously allowed	0	0
Disallowed	0.9	0.9
Average B values (Å^2)		
Protein	31.6	14.6
Water	38.9	29.7
Ions	36.7	21.6
Sugars¶	61.1	44.1

† $R_{cryst} = \sum |F_o| - |F_c| / \sum |F_o|$, where F_o and F_c are the structure-factor amplitudes from the data and the model, respectively. R_{free} is R_{cryst} with a 5% test set of structure factors. ‡ Based on maximum likelihood after refinement of the final model with *REFMAC*5. § Calculated using *PROCHECK* (Laskowski *et al.*, 1993). Residues Phe284 and Asn297 have disfavoured main-chain dihedral angles, but their conformations are unambiguously defined by electron density. Phe284 is close to an additionally allowed region and Asn297 is involved in Ca^{2+} coordination. ¶ Asn141 is *N*-glcosylated. Two *N*-acetylglucosamine moietes are visible in the electron density.

(Emsley & Cowtan, 2004) and refined with *REFMAC*5 (Collaborative Computational Project, Number 4, 1994) against the respective halide data to arrive at the model statistics collected in Table 2. The coordinates and structure factors of the bromide and iodide FGE structures have been deposited with the PDB (codes 2hi8 and 2hib). Taken together, FGE could be SIRAS phased by halide cocrystallization despite suboptimal data quality. The strongest halide-binding sites are the same for both bromide and iodide (Figs. 4*b* and 4*c*), as has been observed previously with the IGF II binding repeat in the Man6P/IGF II receptor (Usón *et al.*, 2003).

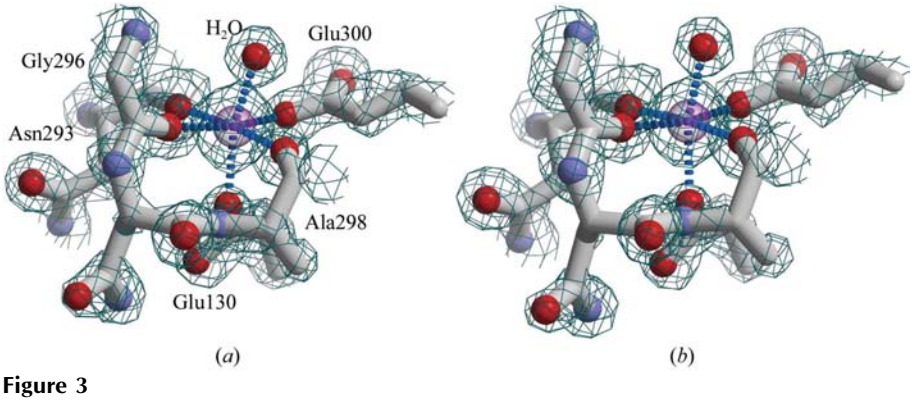

Figure 3
Experimental F_o maps with SIRAS phases of the bromide and iodide cocrystallization experiments after density modification and phase extension in *SHELXE*. All maps are drawn around a Ca^{2+} site and are contoured at 2σ. (*a*) Bromide-phased map after density modification and phase extension to the resolution limit of the native data set of 1.54 Å. (*b*) Results for the iodide-phased map. All maps allowed virtually complete automatic model building at 1.54 Å resolution. At least 257 out of 274 residues were traced and the side chains were docked, resulting in an overall connectivity index of 0.95.

3. Results and discussion

3.1. Probing the oxygen-binding site using halide ions

FGE uses molecular oxygen to convert a cysteine residue in its eukaryotic substrate sulfatases to FGly, which in its hydrated form acts as the active-site nucleophile for sulfate-ester hydrolysis (von Figura *et al.*, 1998). The substrate-binding site has been delineated by a series of apo-FGE structures as a groove on the surface of FGE (Fig. 4*a*) that is bordered by two catalytically important cysteine residues, Cys336 and Cys341 (Dierks *et al.*, 2005). An important intermediate in the FGE catalytic cycle is a mixed disulfide between Cys341 of FGE and a conserved sulfatase cysteine residue that is the precursor of FGly (Fig. 1). Cys341 is close to Cys336, which was speculated to activate molecular oxygen, putting the oxygen-binding site into the immediate vicinity of these residues. The crucial role of Cys336 in catalysis was proven by site-directed mutagenesis to a serine residue, which resulted in an inactive enzyme (Dierks *et al.*, 2005). The precise mode of oxygen binding and mechanism of its activation by FGE in the absence of any cofactors is still elusive. A first glimpse of where exactly molecular oxygen must bind to FGE to be effective in catalysis was provided by crystal structures of the inactive Cys336Ser FGE mutant. This mutant enabled the trapping of mixed-disulfide intermediates of FGE with short peptides that mimic the authentic substrate sulfatase (Roeser *et al.*, 2006). In two independent FGE–peptide complexes (peptide sequences CTPSR and LCTPSR), an oval-shaped cavity of 32 Å^3 volume was formed directly beneath the disulfide bond. The cavity is close to Cys336 and Cys341 and is large enough to harbour a particle of radius 1.97 Å, thus serving as the molecular oxygen-binding site. Based on an anomalous difference map and the chemical environment, a chloride ion was confidently assigned to this site (Roeser *et al.*, 2006; Figs. 4*a* and 4*d*). The ionic radius of chloride is 1.81 Å (corresponding to a volume of 25 Å^3) and is easily accommodated in this cavity formed by covalent substrate binding. The presence of a chloride ion in the oxygen-binding site is a fortuitous outcome of the crystallization mother-liquor composition, but in hindsight any polarisable negatively charged ion or electron-rich entity should be able to exert the same effect. This hypothesis

was tested and verified by redetermination of the FGE structures in the presence of substrate peptide and $CaBr_2$ and CaI_2 (see above). The increased sizes of bromide and iodide allowed determination of the maximum particle size that could be accommodated in the oxygen-binding site during catalysis while still allowing covalent substrate binding.

3.2. FGE structure in complex with chloride and bromide

In the presence of $CaCl_2$ and the pentamer peptide CTPSR, a minimal sulfatase substitute, there is full occupancy of both the buried chloride ion and the peptide as judged by comparison of the chloride and peptide *B* values with those of the surrounding FGE atoms (Roeser *et al.*, 2006). In contrast, in the bromide-containing structure the peptide density was much weaker and refinement of a bromide ion with full occupancy at the former position of the chloride resulted in a $>10\sigma$ negative $F_o - F_c$ difference density peak, clearly suggesting that both the peptide and the bromide had occupancies of <1. A mixture of $CaBr_2$, peptide and FGE was cocrystallized, suggesting that disulfide formation of FGE Cys341 with the CTPSR peptide and bromide binding could compete with each other and may be mutually exclusive. The occupancy of the peptide was estimated as 0.5 by comparing its *B* values after refinement with those of the surrounding FGE atoms. However, adjusting the occupancy of the buried bromide ion to 0.5 in order to reflect a mixture of two structures in the crystal, one with bromide but without peptide and another containing the peptide, left a $>5\sigma$ positive $F_o - F_c$ peak, indicating that either the occupancy chosen was too low

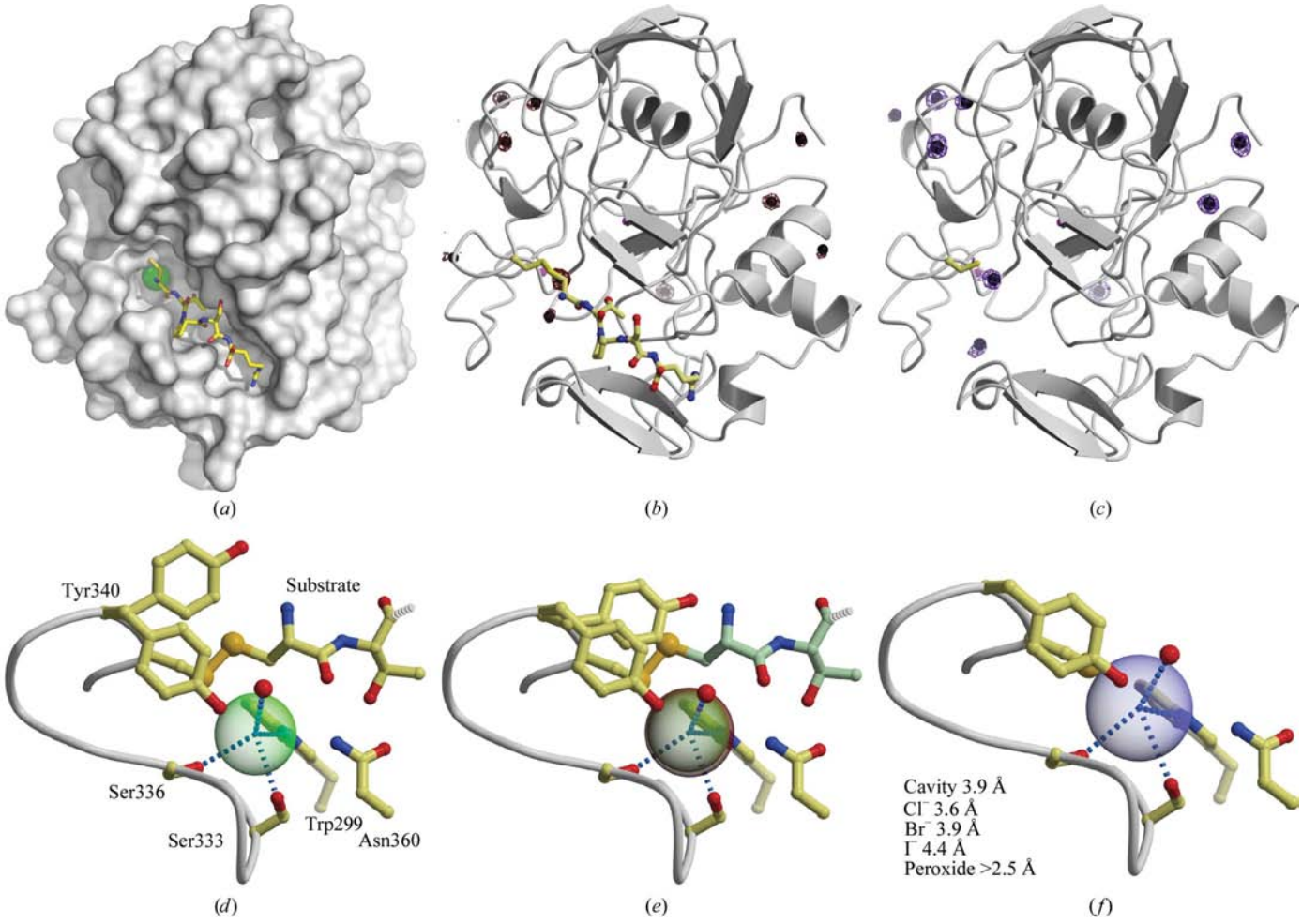

Figure 4
Anomalous difference maps with SIRAS phases and close-up of the bromide and iodide ions in the peptide-binding sites of the three different structures. (*a*) Surface representation of the chloride-containing FGE–peptide complex (Roeser *et al.*, 2006). The chloride ion is buried under the peptide and is represented as a green sphere. The pentamer peptide of sequence CTPSR that represents a minimal substrate sulfatase is shown as a stick model with atoms coloured according to atom type. (*b*) FGE is traced in grey and the location of the ten strongest bromide sites is indicated by black spheres. The anomalous map (brown) is contoured at 4σ. The two Ca^{2+} ions are shown as magenta spheres. (*c*) As in (*b*) for the nine strongest iodide sites. The anomalous map (violet) is contoured at 4σ. (*d*) Close-up of the peptide-binding site for the chloride-containing structure (Roeser *et al.*, 2006). The peptide has full occupancy, burying the chloride ion, which is shown as a green transparent sphere with its ionic radius. The side chain of Tyr340 adopts two alternate conformations. Dashed lines indicate nearest neighbour interactions of <3.4 Å distance. (*e*) In the bromide cocrystallization, the peptide (C atoms in pale green) has partial occupancy and the oxygen-binding site harbours a mixture of bromide (brown sphere) and chloride (green sphere). (*f*) The presence of iodide (violet sphere) prohibits simultaneous binding of the peptide owing to its large ionic radius. Inset: approximate diameters of the cavity and of the discussed particles.

or that this site was occupied by a mixture of bromide and another ion. Finally, the density was satisfactorily explained by a 1:1 mixture of bromide and chloride, each with occupancy of 0.5 (Fig. 4*e*). As chloride was not present in the crystallization mother liquor, the origin of this ion is most likely the protein buffer itself (10 m*M* Tris–HCl pH 7.5). The centres of mass of the chloride and bromide ions in this structure are the same and both ions are shielded from solvent. The complete burial of halide ions in the absence of a neutralizing positive charge has been found in only two other examples: *Vibrio cholerae* extracellular endonuclease I (Altermark *et al.*, 2006) and the hormone-binding domain of the natriuretic peptide receptor (van den Akker *et al.*, 2000). Similar to FGE, the chloride ions in these structures are within 4 Å distance of a disulfide bridge and make hydrogen bonds to serine side chains and NH groups. This situation indicates that polarisable electron-rich environments suffice for buried halide binding even without charge neutralization. While in the endonuclease I and the natriuretic peptide receptor the chloride ion is addressed as a structural component (Altermark *et al.*, 2006), chloride and bromide are located at the enzyme active site in FGE. This site becomes shielded from solvent only upon substrate binding. In the apo-FGE structure without a bound peptide, the oxygen-binding site is solvent-accessible. In addition, the larger bromide ion does not allow simultaneous peptide binding but occupies the same position as the chloride ion in the FGE–peptide complex (Fig. 4*e*), indicating that the oxygen-binding site is pre-formed in the absence of peptide and completed upon covalent substrate binding.

3.3. FGE structure in complex with iodide

The picture is simplified in the iodide-containing structure, where no peptide is bound to FGE at all and the iodide ion has an estimated occupancy of 0.5 based on comparison of its *B* value with those of nearby protein atoms. The assignment of iodide, but not chloride, to this site was based on the strong peak in the anomalous difference map with SIRAS phases (Fig. 4*c*). Apparently, the volume of iodide is too large to allow simultaneous binding of substrate peptide, as has also been inferred for bromide (see above and Fig. 4*f*).

There are only minor (<0.3 Å) conformational changes in FGE upon mixed-disulfide formation (Roeser *et al.*, 2006). Similarly, the FGE structure appears to be rigid when the environments of the bromide and iodide ions in the oxygen-binding pocket is compared. The iodide ion shifts by 0.4 Å away from Cys341, possibly owing to the larger ionic radius of iodide compared with bromide. All residues interacting with iodide essentially remain in place. These are the side chains of Trp299, Ser333 and Ser336 (Cys in wild-type FGE) as well as a water molecule. In addition, the iodide ion is contacted by Cys341 (3.4 Å distance), completing an irregular pyramidal coordination.

Xenon has been used as a probe for molecular oxygen in a number of structures, mostly haem-containing proteins such as myoglobin (Cohen *et al.*, 2006) and cytochrome P450 (Wade *et al.*, 2004), but also FAD-dependent 3-hydroxybenzoate hydroxylase (Hiromoto *et al.*, 2006) and copper amine oxidases (Duff *et al.*, 2004). A survey of 15 unique xenon-bound structures *via* the Macromolecular Structure Database (http://www.ebi.ac.uk/msd-srv/msdsite) shows Xe to prefer hydrophobic cavities. The van der Waals radius of Xe is 2.2 Å, which is the same as the ionic radius of iodide. Thus, xenon could also serve as a molecular oxygen surrogate for FGE, albeit with likely the same outcome as iodide, namely being too large to allow simultaneous peptide binding. Unfortunately, this hypothesis could not be verified as generation of a FGE–Xe derivative was unsuccessful, despite several attempts during which the incubation times and xenon pressures applied to the FGE crystals were varied. All derivatized crystals were unusable as they diffracted to unsatisfactory resolutions lower than 3 Å and displayed excessive mosaicity. However, halide derivatization by cocrystallization yielded the necessary information for probing the oxygen-binding site. Thus, halides could prove useful to act as molecular oxygen surrogates when obtaining a xenon derivative is impossible. The advantage of halides in comparison to Xe is the gradual change in ionic radii, which enables a more precise stereochemical probing of the oxygen-binding site.

In summary, bromide and iodide ions are able to bind to a site in FGE that would normally harbour molecular oxygen but do not fulfil the geometric criteria to serve as perfect oxygen surrogates. Instead, FGE structures containing these ions revealed that the oxygen-binding pocket must be rigid and therefore stereochemically demanding, which is discussed further below.

3.4. Mechanistic implications

FGE is a thiol- and oxygen-dependent monooxygenase. The 32 Å^3 cavity beneath the peptide in the FGE–peptide complexes is more than large enough to harbour an oxygen molecule. The volume of molecular oxygen can be approximated to 7 Å^3 using a bond length of 1.21 Å and an atomic radius of 0.60 Å and on the assumption that O_2 behaves isotropically. While the cavity seems far too large for molecular oxygen at first sight, potential oxygen intermediates must also fit into this volume. Most negatively charged oxygen intermediates have larger volumes compared with O_2; for instance, the bond length of peroxides is in the range 1.30–1.47 Å. Finally, once the oxygen molecule is split, the van der Waals radii of the products impose even higher volume requirements on the cavity. Substrate peptides bind with exquisite surface complementarity and serine-containing peptides that do not bind covalently to FGE display high apparent affinities, with K_d values in the nanomolar range. These observations render a drastic change in volume of the cavity very unlikely. Thus, the 32 Å^3 volume of the cavity imposes an upper limit to any guest particle, which readily explains why binding of iodide (ionic radius of 2.20 Å, volume of 45 Å^3) is incompatible with concomitant binding of substrate peptide (Fig. 4*f*). The bromide ion, on the other hand, has an ionic radius of 1.96 Å (volume of 32 Å^3), which would in principle allow formation of a ternary FGE–

bromide–peptide complex but represents a borderline case: the spherical bromide may not fit into the nonspherical cavity provided by FGE (Fig. 3*c* in Roeser *et al.*, 2006). In addition, binding of the peptide to FGE may be inhibited when iodide and bromide are bound to FGE owing to altered chemical reactivity of Cys336 and, *vice versa*, when peptide is bound exchange of the buried ligand may no longer occur. It could thus be envisaged that covalent binding of the substrate to FGE *via* a disulfide bond provides a mechanism for not only trapping but also fixing an oxygen molecule. Compartmentalization by the electron-rich disulfide bond and also the nearby Trp299 may raise the chemical reactivity of O_2 to a level that does not need further assistance by a cofactor.

We would like to thank Matthew Groves at EMBL beamline X13 at DESY for guidance during synchrotron data collection and Malayalam Mariappan, Kurt von Figura and Dagmar Klostermeier for a critique of the manuscript. This work was supported by grants from the Deutsche Forschungsgemeinschaft to MGR.

References

Akker, F. van den, Zhang, X., Miyagi, M., Huo, X., Misono, K. S. & Yee, V. C. (2000). *Nature (London)*, **406**, 101–104.
Altermark, B., Smalås, A. O., Willassen, N. P. & Helland, R. (2006). *Acta Cryst.* D**62**, 1387–1391.
Cohen, J., Arkhipov, A., Braun, R. & Schulten, K. (2006). *Biophys. J.* **91**, 1844–1857.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cosma, M. P., Pepe, S., Annunziata, I., Newbold, R. F., Grompe, M., Parenti, G. & Ballabio, A. (2003). *Cell*, **113**, 445–456.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
Dierks, T., Dickmanns, A., Preusser-Kunze, A., Schmidt, B., Mariappan, M., von Figura, K., Ficner, R. & Rudolph, M. G. (2005). *Cell*, **121**, 541–552.
Dierks, T., Schmidt, B., Borissenko, L. V., Peng, J., Preusser, A., Mariappan, M. & von Figura, K. (2003). *Cell*, **113**, 435–444.
Duff, A. P., Trambaiolo, D. M., Cohen, A. E., Ellis, P. J., Juda, G. A., Shepard, E. M., Langley, D. B., Dooley, D. M., Freeman, H. C. & Guss, J. M. (2004). *J. Mol. Biol.* **344**, 599–607.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Evans, G. & Bricogne, G. (2002). *Acta Cryst.* D**58**, 976–991.
Fetzner, S. (2002). *Appl. Microbiol. Biotechnol.* **60**, 243–257.
Figura, K. von, Schmidt, B., Selmer, T. & Dierks, T. (1998). *Bioessays*, **20**, 505–510.
Hiromoto, T., Fujiwara, S., Hosokawa, K. & Yamaguchi, H. (2006). *J. Mol. Biol.* **364**, 878–896.
Hopwood, J. J. & Ballabio, A. (2001). *The Metabolic and Molecular Bases of Inherited Diseases*, edited by C. R. Scriver, A. L. Beaudet, W. S. Sly, D. Valle, B. Childs, K. W. Kinzler & B. Vogelstein, pp. 3725–3732. New York: McGraw–Hill.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
Preusser-Kunze, A., Mariappan, M., Schmidt, B., Gande, S. L., Mutenda, K., Wenzel, D., von Figura, K. & Dierks, T. (2005). *J. Biol. Chem.* **280**, 14900–14910.
Roeser, D., Dickmanns, A., Gasow, K. & Rudolph, M. G. (2005). *Acta Cryst.* D**61**, 1057–1066.
Roeser, D., Preusser-Kunze, A., Schmidt, B., Gasow, K., Wittmann, J. G., Dierks, T., von Figura, K. & Rudolph, M. G. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 81–86.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.
Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.
Usón, I., Schmidt, B., von Bulow, R., Grimme, S., von Figura, K., Dauter, M., Rajashankar, K. R., Dauter, Z. & Sheldrick, G. M. (2003). *Acta Cryst.* D**59**, 57–66.
Wade, R. C., Winn, P. J., Schlichting, I. & Sudarko (2004). *J. Inorg. Biochem.* **98**, 1175–82.
Weiss, M. (2001). *J. Appl. Cryst.* **34**, 130–135.
Weiss, M. S., Sicker, T. & Hilgenfeld, R. (2001). *Structure*, **9**, 771–777.