Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Thiol-reactive lanthanide chelates for phasing protein X-ray diffraction data

Michael D. Purdy,[a] Pinghua Ge,[b] Jiyan Chen,[b] Paul R. Selvin[b] and Michael C. Wiener[c]*

[a]Interdisciplinary Program in Biophysics, University of Virginia, 1300 Jefferson Park Avenue, Charlottesville, VA 22908, USA, [b]Department of Physics and Biophysics Center, University of Illinois Urbana-Champaign, 1110 West Green Street, Urbana, IL 61801, USA, and [c]Department of Molecular Physiology and Biological Physics, University of Virginia, 1300 Jefferson Park Avenue, Charlottesville, VA 22908, USA

Correspondence e-mail: mwiener@virginia.edu

Received 1 January 2002
Accepted 11 April 2002

Lanthanides can contribute a large anomalous component to X-ray scattering when present and ordered in a target crystal. This large anomalous signal is a useful source of phase information in X-ray crystallographic studies of biological macromolecules. Thiol-reactive lanthanide chelates were tested as a means of incorporation of lanthanides into protein crystals. Two compounds, each capable of being loaded with a lanthanide of choice, were synthesized: diethylenetriaminepentaacetic 3-(2-pyridyldithio)propionyl hydrazide (DTPA-PDPH) and 1,4,7,10-tetraazacyclododecane-*N*,*N*′,*N*″,*N*‴-tetraacetic 3-(2-pyridyldithio)propionyl hydrazide (DOTA-PDPH). A cysteine mutant of the 34 kDa phosphate-binding protein (PBP-A197C) from *Escherichia coli* was used as a test case. PBP-A197C was labeled with DTPA-PDPH loaded with dysprosium. Characteristics of DTPA-PDPH enabled spectroscopic monitoring of the labeling reaction. Complete labeling of PBP-A197C was confirmed by mass spectrometry and SDS–PAGE analysis. Labeled PBP-A197C (PBP-A197C–DTPA-Dy) crystallized identically to unlabeled protein. X-ray diffraction data were collected from PBP-A197C–DTPA-Dy crystals in-house with a Cu $K\alpha$ rotating-anode source and with a tuneable synchrotron source (ALS 5.0.2). Synchrotron data were collected at energies corresponding to the Dy $L_{\rm III}$ edge f'' peak and a high-energy remote. Each data set was treated as an independent SAD experiment. A large anomalous signal was present in the data collected in-house and at the synchrotron. The Dy site was easily located in anomalous difference Patterson maps calculated from each of the data sets. In each case, SAD phasing resulted in high-quality electron-density maps, as evidenced by the success of automated model building. The generality of the method was analyzed with several other test proteins. Labeling of some of these proteins with thiol-reactive lanthanide chelates was deleterious to protein solubility or crystallization. In two of the cases the lanthanide chelate was disordered in the crystals. These results suggest that this method may not be well suited for high-throughput crystallography. However, for difficult cases requiring a large anomalous signal, thiol-reactive lanthanide chelates may prove to be a valuable tool.

1. Introduction

Thanks to a rapid increase in the number of tuneable synchrotron beamlines and the attendant development of methods for incorporation of anomalous scatterers into macromolecular crystals, MAD and SAD experiments are now routine. Biosynthetic incorporation of selenium is invaluable (Doublié, 1997), and exhibiting signs of an immi-

nent popular coup is Dauter's quick cryosoaking method (Dauter *et al.*, 2000). These and other methods are and will continue to be excellent tools for phasing in a wide variety of cases. There are, however, crystallographic experiments that necessitate, or at least benefit from, a large number of anomalous scatterers or the inclusion of atoms that contribute a large anomalous component to X-ray scattering. Examples include large complexes, weakly diffracting crystals (*i.e.* many membrane proteins) and SAD experiments conducted in-house with a rotating-anode source.

Long recognized is the utility of the large anomalous signal from the lanthanides in phasing X-ray diffraction data from macromolecular crystals (Lye *et al.*, 1980). Sharp white lines at the lanthanide $L_{\rm III}$ and $L_{\rm II}$ absorption edges are a potent source of phase information when the energy of the incident beam is proximate to these electronic transitions. Illustrative of the benefit of lanthanide incorporation, for anomalous diffraction experiments, is a comparison with two commonly used anomalous scatterers, selenium and mercury. An estimate of the Bijvoet difference ratio at zero scattering angle is given by

$$\text{r.m.s.}(\Delta F_{\pm h})/\text{r.m.s.}(|F|) \simeq (N_A/2N_T)^{1/2}(2f_A''/Z_{\rm eff}),$$

where N_A is the number of anomalous scatterers per protein molecule, N_T is the number of non-H protein atoms, f_A'' is the imaginary component of anomalous scattering and $Z_{\rm eff} = 6.7$ e is the effective normal scattering at zero scattering angle (Hendrickson & Ogata, 1997). Considering a 34 kDa protein associated with one Se ($f_A'' \simeq 7$; Hendrickson & Ogata, 1997), Hg ($f_A'' = 11.6$; Krishna *et al.*, 1994) or Dy ($f_A'' = 28$; this study), the maximum Bijvoet difference ratios are 0.03, 0.05 and 0.13, respectively. Independent of protein size, the signal of one lanthanide is equivalent to that of six mercury or 16 Se atoms. Despite the large anomalous signal obtainable from the lanthanides, their use in phasing has been restricted primarily to metalloproteins, which allow substitution of the transition metals for an endogenous ion (*e.g.* Weis *et al.*, 1991). In some cases, conventional heavy-atom soaks with lanthanide compounds yielded useful derivatives (*e.g.* Gaudet *et al.*, 1996). Methods for lanthanide derivatization recently published include a variation of the quick cryosoaking method (Nagem *et al.*, 2001) and use of an NMR contrast reagent, Gd-HPDO3A, in co-crystallization (Girard *et al.*, 2002). A general means of incorporation of lanthanides into protein crystals would be a useful addition to the crystallographer's toolbox. To this end, we are investigating the use of thiol-reactive lanthanide chelates for phasing protein X-ray diffraction data. These compounds bind to free cysteine residues and can be loaded with the lanthanide ion that best suits the experiment. Because of the advantageous optical properties of lanthanides, particularly terbium and europium, thiol-reactive lanthanide chelates have been used as optical probes (Selvin, 2002). In this study, we labeled a 34 kDa test protein, phosphate-binding protein (PBP) from *E. coli*, with dysprosium loaded DTPA-PDPH. The anomalous signal present in data collected at a synchrotron beamline, ALS 5.0.2, as well as from a copper rotating-anode source, yielded high-quality phases. The same compound was tested unsuccessfully with several other proteins. Results with PBP suggest that thiol-reactive lanthanide chelates may be a powerful tool for phasing in some cases.

2. Materials and methods

2.1. Chemicals

DTPA dianhydride and $DyBr_3$ were purchased from Aldrich; 3-[2-pyridyldithio]propionyl hydrazide (PDPH) was purchased from Pierce; DOTA-Na (1,4,7,10-tetraazacyclododecane-*N*,*N*′,*N*″,*N*‴-tetraacetic acid sodium salt) was purchased from Parish. All these chemicals were used directly as purchased. Anhydrous DMSO and triethylamine was purchased from Aldrich and dried over molecular sieves before use. Triethylammonium acetate (TEAA) was prepared by standard methods.

2.2. Synthesis of DTPA-PDPH (Fig. 1)

DTPA dianhydride (31 mg, 87 µmol) was dissolved in a mixture of 0.5 ml anhydrous DMSO and 0.1 ml dry triethylamine. PDPH (10 mg, 44 µmol) in 0.5 ml DMSO was added to the above solution slowly under N_2. The reaction was stirred for 2 h. 9 ml of 0.1% TFA was added to quench the reaction. The product was purified by reverse-phase HPLC with a 15–40% acetonitrile/TFA linear gradient, monitored at $\lambda = 282$ nm ($\varepsilon_{280} = 4000\ M^{-1}\ {\rm cm}^{-1}$). DTPA-PDPH eluted at 14 min [∼50% yield based on HPLC profile; $M^+ + 1$ (FAB): 605] and concentrated under vacuum. The product was then dissolved in 0.1 *M* TEAA pH 6.5 and a stoichiometric amount of aqueous solution of $DyBr_3$ was added. The reaction was stirred at room temperature for 2 h. The mass spectrum of this reaction showed a DTPA-PDPH-Dy *m*/*z* peak at 765 ($M - 1$, ESI). No free chelate mass peak was found. The solution was used directly for labeling purposes without further purification.

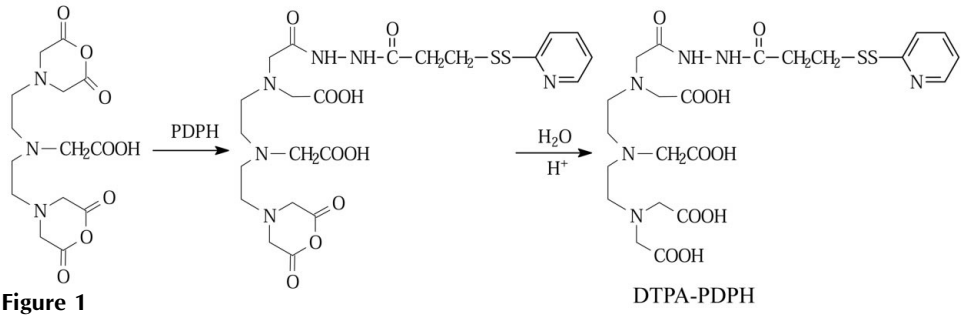

Figure 1
Synthesis of DTPA-PDPH.

2.3. Synthesis of DOTA-PDPH (Fig. 2)

DOTA(Na) (15 mg, 32 µmol) was dissolved in a mixture of DMF (2 ml) and triethylamine (0.2 ml). Isobutyl chloroformate (4.7 µl, 32 mol) was added slowly to the solution at 273 K with vigorous stirring. The cold bath was removed and the solution was allowed to warm up to room temperature. After an additional 30 min stirring, PDPH (8.07 mg, 35 µmol) in 1 ml DMF was added to the solution *via* a syringe. The reaction was stirred at room temperature overnight. The solvent was then evaporated under vacuum. The residue was dissolved in 5 ml 0.1 *M* TEAA. The product was purified by reverse-phase HPLC with a 10–30% acetonitrile/TFA linear gradient, monitored at λ = 282 nm (ε_{280} = 4000 M^{-1} cm^{-1}). DOTA-PDPH eluted at 30.7 min [~40% yield based on HPLC profile; M^{+} + 1 (FAB): 647]. DOTA-PDPH solution (in 0.1 *M* TEAA pH 6.5) was added to 1.2 equivalents of $DyBr_3$ aqueous solution. The reaction was stirred at room temperature for 2 h. The reaction mixture was loaded onto an anion-exchange DE-52 column. The column was then washed with a few millilitres of deionized water to remove excess metal ions. The product was then eluted by washing the column with 1:1 MeOH/H_2O and concentrated under vacuum. The mass spectrum of this concentrated solution shows only a DOTA-PDPH-Dy *m*/*z* peak at 807 (*M* − 1, ESI). This solution was used for labeling purposes.

2.4. Selection of the lanthanide

DTPA and DOTA can be loaded with a variety of ions from the lanthanide series. The primary consideration in the selection of the lanthanide is the source of X-rays that will be used for the diffraction experiment. Generally, crystallographic experiments conducted at tuneable synchrotron sources will benefit from the use of lanthanides with relatively higher energy L_{III} and L_{II} absorption edges (*e.g.* lutetium). Use of these ions will minimize X-ray absorption and result in fewer problems with detector size. Experiments performed in-house with Cu $K\alpha$ radiation require the use of ions that have *L* edges below the energy of Cu $K\alpha$. The energy of the L_{III} absorption edge of Dy is 7790 eV. This edge is close to but below the energy of Cu $K\alpha$ radiation. Dysprosium was chosen for this study because of the large anomalous signal present in-house combined with the fact that the Dy L_{III} edge is accessible at many tuneable synchrotron beamlines.

2.5. Test proteins

Proteins used as test cases for the lanthanide-chelate method were selected based upon the presence of a free cysteine in the wild-type protein or availability of cysteine mutants, as well as ease of expression, purification and crystallization. Inspection of packing in these cases indicated sufficient space for the chelate within the lattice and in none of the test cases were the chelates predicted to disrupt crystal contacts.

Phosphate-binding protein (PBP) from *E. coli* is a 34 kDa periplasmic protein. The crystal structure of the PBP mutant A197C labeled with a probe for inorganic phosphate, MDCC, has been solved previously (Hirshberg *et al.*, 1998). The PBP-A197C mutant was used as the first test protein for the lanthanide-chelate method. Expression, purification and crystallization were performed as described previously (Brune *et al.*, 1998).

The other test proteins used in this study were *Bacillus subtilis* YabJ (Sinha *et al.*, 1999), *Haemophilus influenzae* proteins HI0319 and HI1434 (Fleischmann *et al.*, 1995) and a cysteine mutant of *E. coli* thioesterase II (TEII-S62C; Li *et al.*, 2000). In addition to these proteins, five single-site cysteine mutants of HI1434 (S44C, S78C, S95C, S104C, S136C) were used in this study. HI0319 (Lim *et al.*, 2001) and HI1434 (Zhang *et al.*, 2000) were expressed, purified and crystallized as described previously. The five HI1434 cysteine mutants were generated with Stratagene's Quickchange kit and expressed and purified following the protocol for the wild-type protein.

2.6. Labeling

Following purification, proteins were labeled with either DTPA-PDPH-Dy or DOTA-PDPH-Dy in solution. A twofold to fourfold molar excess of compound was used in all labeling experiments. Release of pyridyl upon reaction of the compounds with cysteine sulfhydryls enabled spectroscopic monitoring of labeling. The extinction coefficient at 343 nm for free pyridyl is 8100 M^{-1} cm $^{-1}$. The expected absorption at 343 nm (A_{343}) for complete labeling of the proteins was calculated based on the fact that the concentration of free pyridyl after complete labeling equals the concentration of protein in the reaction. The A_{343} was monitored until there was no further increase. Labeling of the proteins was

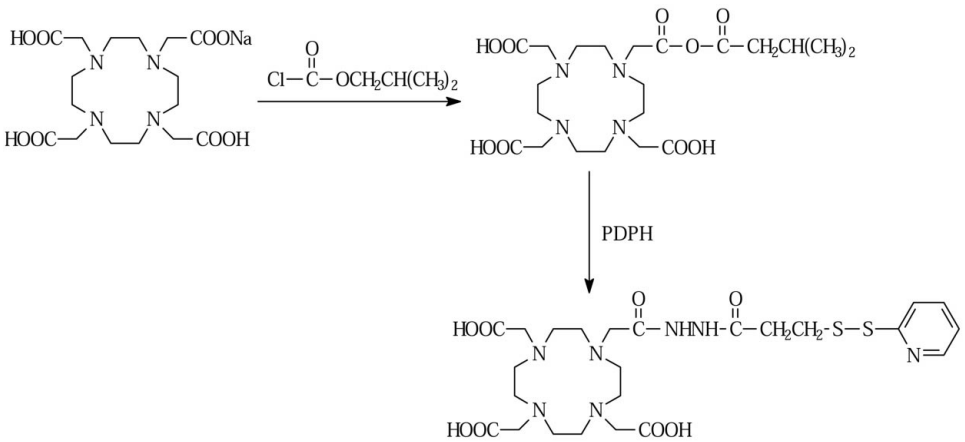

Figure 2
Synthesis of DOTA-PDPH

confirmed by SDS–PAGE analysis of labeled and unlabeled protein run in adjacent lanes of the gels. In some cases mass spectrometry was used as an additional confirmation of labeling.

2.7. Data collection

In-house data sets were collected from crystals of labeled PBP-A197C and TEII-S62C using a Rigaku RU-H2R generator (50 kV, 100 mA) and a Bruker SMART 6000 CCD detector equipped with Goebel mirrors. Data collection was controlled with *HKL*2000 (Otwinowski & Minor, 1997). Data were collected at two wavelengths from a crystal of PBP-A197C–DTPA-Dy at ALS beamline 5.0.2 with an ADSC Quantum 4 detector after performing a fluorescence scan. Wavelengths for data collection corresponded to the Dy L_{III} edge f'' peak (1.5907 Å, 7.794 keV) and a high-energy remote wavelength (1.1800 Å, 10.506 keV). Data were collected from a crystal of HI1434–DOTA-Dy at NSLS beamline X12B at a single wavelength (1.1000 Å) corresponding to a Dy L_{III} edge high-energy remote. All data were collected at ~100 K. All data were reduced with *HKL*2000.

2.8. SAD phasing

Crystallographic calculations were performed with *CNS* v1.0 (Brünger *et al.*, 1998), *CCP*4 v4.1 (Collaborative Computational Project, Number 4, 1994) and *ARP*/*wARP* v5.1 (Perrakis *et al.*, 1999). Specifically for PBP-A197C–DTPA-Dy, the dysprosium sites were located automatically with *CNS* from the Bijvoet differences. SAD phasing and density modification (DM) were performed with *CNS* to yield the resultant PBP-A197C–DTPA-Dy SAD/DM electron-density map. In the cases of the test proteins other than PBP-A197C (TEII-S62C–DTPA-Dy and HI1434–DOTA-Dy), molecular replacement *via AMoRe* (Navaza, 1994) of the known structures (Li *et al.*, 2000; Zhang *et al.*, 2000) was used to calculate Fourier difference maps.

2.9. PBP-A197C–DTPA-Dy automated model building

Success of automated model building with the program *ARP*/*wARP* was used as an objective measure of the quality of the PBP-A197C–DTPA-Dy SAD/DM electron-density maps. A single run of *ARP*/*wARP* was sufficient for model building into density obtained from the synchrotron data. A two-step process was used in the case of in-house data. First, *ARP*/*wARP* was used to build a partial model. Then, phases from the partial model were combined with SAD phases, followed by density modification in *CNS*. The resultant phases and structure-factor amplitudes were used to calculate the initial map in a second round of automated building.

3. Results

PBP-A197C was completely labeled with DTPA-PDPH-Dy in solution as revealed by absorbance measurements (Fig. 3), mass spectrometry and SDS–PAGE. Mass spectra revealed a 485 Da difference between labeled and unlabeled protein. This mass shift corresponds to the mass of DTPA without the Dy, which was lost during ionization. No reduction in the solubility of the protein was observed after labeling.

Table 1
Data and phasing statistics for PBP-A197C–DTPA-Dy.

Data set	Dy f'' peak, ALS 5.0.2	Dy f'' remote, ALS 5.0.2	Cu $K\alpha$, in-house
Wavelength (Å)	1.5907	1.1800	1.5418
Detector 2θ (°)	0.0	0.0	10.0
Resolution (Å) (radial)†	20.00–2.07 (2.45)	20.00–1.52 (1.80)	30.00–1.70
Completeness (%) (radial)†	82.3 (99.6)	79.1 (99.4)	93.0
Unique reflections‡	30826	74913	62528
Redundancy‡	5.0	4.9	6.0
$I/\sigma(I)$§	18.0 (3.1)	24.7 (2.6)	26.6 (3.8)
R_{sym}‡§ (%)	7.5 (24.0)	5.1 (20.3)	4.8 (20.9)
Anomalous difference¶ (%)	12	4	5
f'' (e)††	28	10	10
Phasing power	2.0	1.1	0.7
FOM before/after DM (acentrics)	0.38/0.88	0.39/0.83	0.20/0.84

† Data in the corners of the square detector were used to facilitate automated model building. ‡ Each Friedel mate is treated as a unique reflection. § Values in parentheses are for the highest resolution shells (peak, 2.14–2.07 Å; remote 1.57–1.52 Å; in-house, 1.76–1.70 Å). ¶ $\left(\left\langle\left|\,|f_{obs}+|-|f_{obs}-|\,|\right|^2\right\rangle\right)^{1/2}/\frac{1}{2}\left(\langle|f_{obs}+|\rangle^2+\langle|f_{obs}-|\rangle^2\right)^{1/2}$. †† The f'' peak value was determined from the fluorescence scan; the values for the high energy remote wavelength and in-house Cu $K\alpha$ were obtained from a plot of the theoretical f'' values (http://www.bmsc.washington.edu/scatter/).

Labeled PBP-A197C crystallized identically to unlabeled protein (space group $P2_12_12_1$, unit-cell parameters a = 41.3, b = 62.4, c = 122.2 Å, one molecule in the asymmetric unit). Three data sets were collected from crystals of labeled protein and used independently for location of the Dy site, subsequent SAD phasing and density modification (Table 1). The fluorescence scan collected at ALS 5.0.2 from a crystal of PBP-A197C–DTPA-Dy shows the large f'' peak at the Dy L_{III} edge as well as the source of the significant Bijvoet difference present in data collected at energies above the L_{III} edge (Fig. 4).

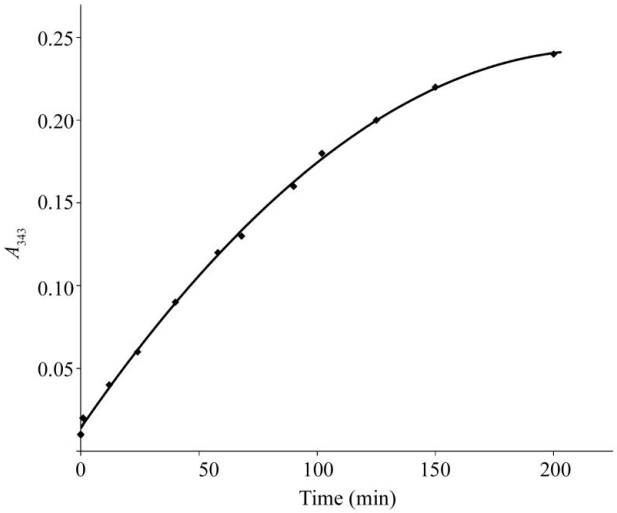

Figure 3
Labeling of PBP-A197C with DTPA-PDPH-Dy monitored spectroscopically at the absorption maximum for free pyridyl, 343 nm. The calculated value for complete labeling of the protein is A_{343} = 0.20.

Anomalous difference Patterson maps calculated from data collected in-house as well as with the tuneable source clearly revealed the position of the Dy (Fig. 5). High-quality SAD phases were obtained for each data set using the Dy anomalous signal (Table 1 and Fig. 6). Electron-density maps calculated following density modification were of excellent quality in each case (Fig. 7).

The success of automated model building with *ARP*/*wARP* confirmed the high quality of the electron-density maps. *ARP*/*wARP* built 312 of 321 residues in two chains (connectivity = 0.99) into the SAD map derived from the ALS 5.0.2 high-energy remote data set. The first round of automated building with maps calculated from the in-house data yielded a partial model. After combination of SAD phases and phases calculated from the partial model and density modification, *ARP*/*wARP* built 309 of the residues in two chains (connectivity = 0.98). Although the maps calculated from the ALS 5.0.2 f'' peak data were of the highest quality based upon visual inspection, the high-resolution limit of the data was not sufficient for automated building with *ARP*/*wARP*. These maps were easily interpretable and if this were a new structure, conventional manual or semi-automated methods would readily yield a structure. However, as we sought the objective criterion of automated map interpretation and building, we did not further pursue interpretation or analysis of this superb, albeit lower resolution, experimentally phased electron-density map.

Several obstacles prevented successful use of the lanthanide-chelate method with test proteins other than PBP-A197C (Table 2). Some of the proteins labeled with DTPA-Dy did not crystallize under the conditions used to crystallize the native proteins. YabJ and HI0319 exhibited poor solubility after labeling with the lanthanide chelate. Absorbance measurements and SDS–PAGE analysis indicated that HI1434 and TEII-S62C were labeled. However, there was no anomalous signal present in data collected from the crystals. Molecular replacement was used to generate difference maps to look for evidence of the lanthanide-containing compounds. These σ_A-weighted $F_o - F_c$ difference maps revealed no electron density for the DTPA-Dy. Mass spectrometry of solubilized crystals of labeled TEII-S62C confirmed that the protein in the crystal used for data collection was labeled with DTPA-Dy. Thus, the only explanation for the absence of an anomalous signal in the data and electron density for the compound in the difference maps is disorder of the compound in the crystals of TEII-S62C and, by inference, HI1434.

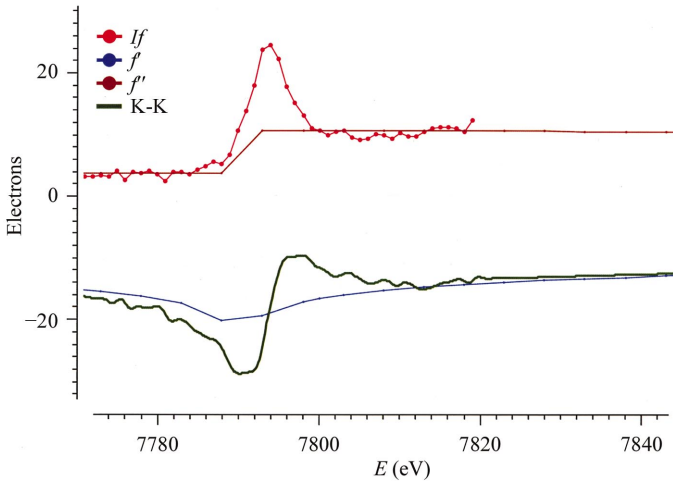

Figure 4
Fluorescence scan collected from a crystal of PBP-A197C–DTPA-Dy at ALS 5.0.2. Theoretical values of the imaginary and real components of Dy anomalous scattering are labeled f'' and f', respectively. The imaginary (Bijvoet) component f'', obtained directly from the measured fluorescence intensity (Hendrickson & Ogata, 1997), is labeled *If*. The real (dispersive) component f', obtained by the Kramers–Kronig transformation (Hendrickson & Ogata, 1997), is labeled K-K.

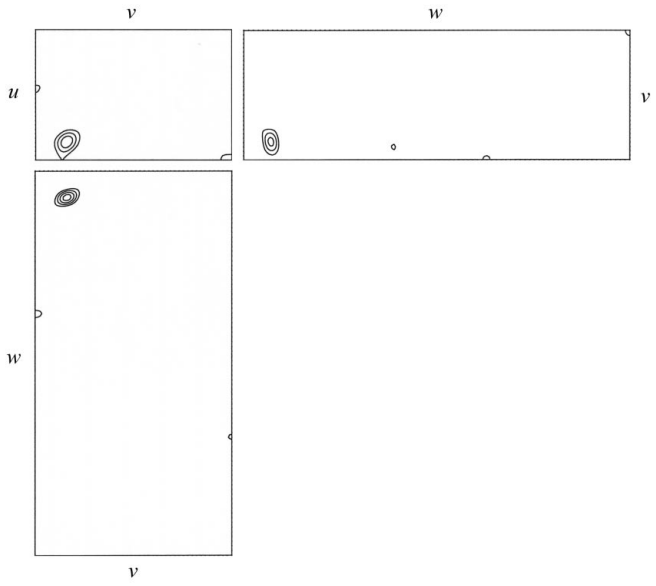

Figure 5
Harker sections of the anomalous difference Patterson map calculated from PBP-A197C–DTPA-Dy data collected in-house, showing the Dy peak. The first contour is at 4σ and the contour increment is 4σ.

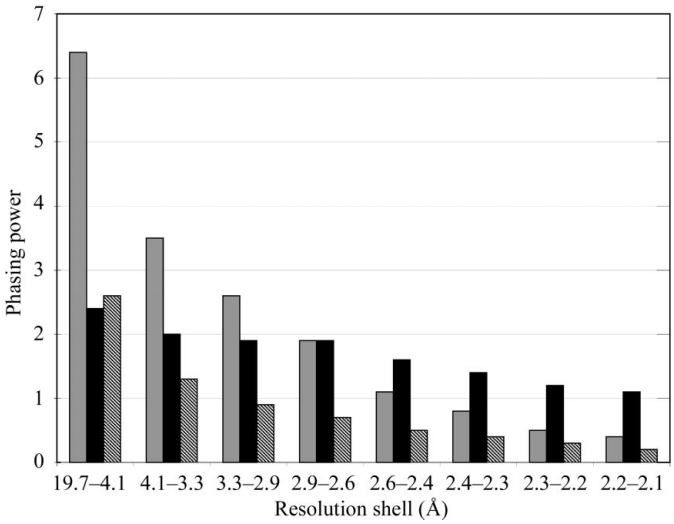

Figure 6
SAD phasing power, from Dy anomalous signal, for the two synchrotron and one in-house data sets. Grey, ALS f'' peak; black, ALS remote; hatched, Cu $K\alpha$.

Table 2
Summary of results of the lanthanide-chelate method with all of the test proteins used in this study.

Protein	Compound	Result
PBP-A197C	DTPA-Dy	Successful with in-house and ALS 5.0.2 data
YabJ	DTPA-Dy	Poor solubility
HI0319	DTPA-Dy	Poor solubility
HI1434	DTPA-Dy	Crystals, but compound disordered
HI1434	DOTA-Dy	Crystals, but compound disordered
HI1434-S44C	DTPA-Dy	Did not crystallize
HI1434-S78C	DTPA-Dy	Did not crystallize
HI1434-S95C	DTPA-Dy	Poor crystals
HI1434-S104C	DTPA-Dy	Did not crystallize
HI1434-S136C	DTPA-Dy	Poor crystals
TEII-S62C	DTPA-Dy	Crystals, but compound disordered

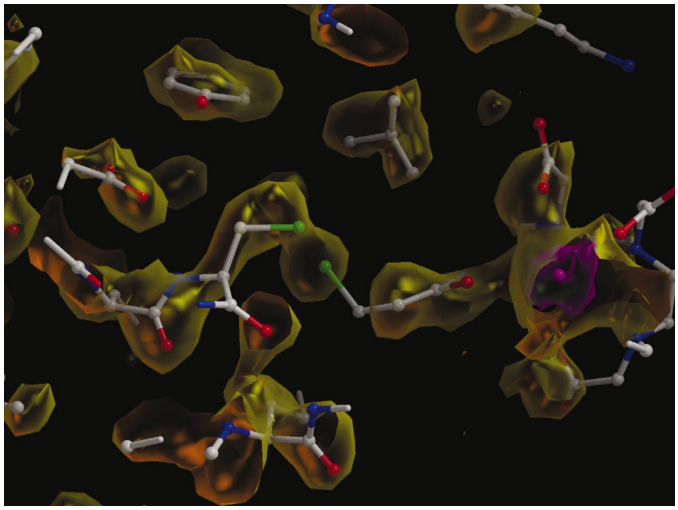

Figure 7
PBP-A197C–DTPA-Dy SAD/DM map calculated from data collected at the Dy f'' peak. The map clearly show the density of the lanthanide chelate and the Dy. The density in yellow/orange is at 1.2σ; the density in magenta is at 12σ. Figure prepared with *Bobscript* v.2.4 (Esnouf, 1997).

4. Discussion

There are obvious similarities between the thiol-reactive lanthanide-chelate method and labeling of free cysteines with mercurials (Dao-Pin *et al.*, 1987; Sanderson *et al.*, 1990; Wiener *et al.*, 1997) and there are notable advantages to each method. In single-crystal experiments, the lanthanides are more powerful for phasing. When non-isomorphism between native and derivative crystals is not a problem, isomorphous replacement with Hg may yield better results in subsequent SIRAS or MIRAS phasing. There is little difference in the anomalous signal from Hg and Dy when SAD experiments are conducted in-house with Cu $K\alpha$ X-rays. Parenthetically, there are a number of other elements useful for in-house phasing; Pb, Pt, Au, U, Xe and I all have f'' between 7 and 13 e at 1.5418 Å (8.041 keV, Cu $K\alpha$). The strength of the thiol-reactive lanthanide-chelate method is the combined benefits that the method offers. These lanthanide compounds are modular, enabling variation of the reactive group, linker, chelate and lanthanide. For example, while both DOTA and DTPA chelates have very high binding constants (much greater than $10^{14}\ M^{-1}$), DOTA is less sensitive to competition from other chelates (such as EDTA) or high concentrations of phosphates. The pyridyl-containing compounds offer the ability to monitor labeling spectroscopically. Loading of the chelate with particular lanthanides (Sm, Eu, Gd, Tb, Dy) provides a significant anomalous signal when data is collected in-house. Bijvoet differences present in the in-house data can be used at a minimum to determine whether the lanthanide is ordered prior to a synchrotron trip. If the lanthanide is ordered, SAD phasing of the in-house data may be possible, as was achieved with PBP-A197C tested in this study. In contrast to the use of weaker anomalous scatterers with Cu $K\alpha$ radiation, unusually high redundancy and $I/\sigma(I)$ are not prerequisites for successful in-house phasing. When data is to be collected at a synchrotron, the chelate can be loaded with the lanthanide that best suits the characteristics of the beamline. At the synchrotron, exploitation of the lanthanide $L_{\rm III}$ or $L_{\rm II}$ edge will result in a large anomalous component of X-ray scattering, facilitating phasing in single-crystal experiments.

We chose to label proteins prior to crystallization to facilitate spectroscopic monitoring of the reactions, through measurement of the free pyridyl absorbance at 343 nm. Derivatizating crystals 'classically' by soaking may also be a viable option. Though technically more challenging, it may even be possible to optically monitor labeling in soak-derivatization experiments. However, the large size, approximately three times that of a tryptophan side chain, of the lanthanide chelate compounds may preclude derivatization by soaking, as is often the case with other large heavy-atom compounds. Based only on the size of the compounds, we would not expect a higher success rate for derivatization of crystals with lanthanide chelates than for that of derivatization with mercurials (Dao-Pin *et al.*, 1987).

In this study, we only attempted to crystallize the labeled test proteins in the conditions established for the unlabeled proteins and in all of the test cases except one, only a single cysteine site was tested with the lanthanide chelates. However, for new structures requiring a large anomalous signal, generation and testing of multiple cysteine mutants and additional screening of crystallization conditions may be warranted for use with lanthanide chelates. Also, modification of reaction or buffer conditions (as well as additional mutagenesis of the protein) may be considered to alleviate any solubility problems that may occur with labeled protein. Several technical aspects may also further justify such an experimental strategy. Firstly, modern mutagenesis methods (*i.e.* kits such as Stratagene's Quickchange) make preparation of multiple single-site cysteine mutants straightforward and fast. Secondly, scanning cysteine mutagenesis is often used to investigate the function and structure of proteins, especially membrane proteins such as channels and transporters (Hubbell *et al.*, 1996; Perozo *et al.*, 1998; Cadieux *et al.*, 2000). Therefore, an ensemble of cysteine mutants can serve multiple roles. As seen in the case of the 34 kDa protein PBP-A197C, labeling the protein with a single DTPA-Dy molecule resulted in excellent phases when the anomalous signal was maximized at a synchrotron source and even when data were collected in-house.

Plasmid containing the PBP-A197C gene and the *E. coli* expression host (JM83) were a gift from Martin R. Webb, National Institute for Medical Research, Mill Hill, England. Janet Smith, Purdue University, donated purified YabJ. Zygmunt Derewenda, University of Virginia, provided purified TEII-S62C. Plasmids containing the HI0319 and wild-type HI1434 genes were from Osnat Herzberg, University of Maryland Biotechnology Institute. Mass-spectrometry analysis was performed by the W. M. Keck Biomedical Mass Spectrometry Laboratory, which is funded by a grant from the Univeristy of Virginia Pratt Committee. We thank Gerry McDermott, Keith Henderson and Thomas Earnest of ALS 5.0.2, and the staff of NSLS X12B for use of their beamlines and assistance during data collection. We appreciate the numerous useful discussions with Wladek Minor during the course of this study. Funding provided by NIH GM56251 and NASA NAG8-1567 (MCW and MDP) and NIH AR44420 and NSF 9984841 (PRS). MDP is a NASA GSRP predoctoral fellow.

References

Brune, M., Hunter, J. L., Howell, S. A., Martin, S. R., Hazlett, T. L., Corrie, J. E. T. & Webb, M. R. (1998). *Biochemistry*, **37**, 10370–10380.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Cadieux, N., Bradbeer, C. & Kadner, R. J. (2000). *J. Bacteriol.* **182**, 5954–5961.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Dao-Pin, S., Alber, T., Bell, J. A., Weaver, L. H. & Matthews, B. W. (1987). *Protein Eng.* **1**, 115–123.

Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.

Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.

Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.

Fleischmann, R. D. *et al.* (1995). *Science*, **269**, 496–512.

Gaudet, R., Bohm, A. & Sigler, P. B. (1996). *Cell*, **87**, 577–588.

Girard, E., Chantalat, L., Vicat, J. & Kahn, R. (2002). *Acta Cryst.* D**58**, 1–9.

Hendrickson, W. A. & Ogata, C. M. (1997). *Methods Enzymol.* **276**, 494–520.

Hirshberg, M., Henrick, K., Haire, L. L., Vasisht, N., Brune, M., Corrie, J. E. T. & Webb, M. R. (1998). *Biochemistry*, **37**, 10381–10385.

Hubbell, W. L., Mchaourab, H. S., Altenbach, C. & Lietzow, M. A. (1996). *Structure*, **4**, 779–783.

Krishna, T. S. R., Kong, X. P., Gary, S., Burgers, P. M. & Kuriyan, J. (1994). *Cell*, **79**, 1233–1243.

Li, J., Derewenda, U., Dauter, Z., Smith, S. & Derewenda, Z. S. (2000). *Nature Struct. Biol.* **7**, 555–559.

Lim, K., Zhang, H., Tempczyk, A., Bonander, N., Toedt, J., Howard, A., Eisenstein, E. & Herzberg, O. (2001). *Proteins Struct. Funct. Genet.* **45**, 397–407.

Lye, R. C., Phillips, J. C., Kaplan, D., Doniach, S. & Hodgson, K. O. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 5884–5888.

Nagem, R. A. P., Dauter, Z. & Polikarpov, I. (2001). *Acta Cryst.* D**57**, 996–1002.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Perozo, E., Cortes, D. M. & Cuello, L. G. (1998). *Nature Struct. Biol.* **5**, 459–469.

Perrakis, A., Morris, R. M. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Sanderson, M. R., Freemont, P. S., Rice, P. A., Goldman, A., Hatfull, G. F., Grindley, N. D. F. & Steitz, T. A. (1990). *Cell*, **63**, 1323–1329.

Selvin, P. R. (2002). *Ann. Rev. Biophys. Biol. Struct.* **31**, 275–302.

Sinha, S., Rappu, P., Lange, S. C., Mantsala, P., Zalkin, H. & Smith, J. L. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 13074–13079.

Wiener, M., Freymann, D., Ghosh, P. & Stroud, R. M. (1997). *Nature (London)*, **385**, 461–464.

Weis, W. I., Kahn, R., Fourme, R., Drickamer, K. & Hendrickson, W. A. (1991). *Science*, **254**, 1608–1615.

Zhang, H., Huang, K., Li, Z., Banerjei, L., Fisher, K. E., Grishin, N. V., Eisenstein, E. & Herzberg, O. (2000). *Proteins*, **40**, 86–97.