J|A|C|S
ARTICLES
Published on Web 05/12/2007

Double-Lanthanide-Binding Tags for Macromolecular Crystallographic Structure Determination

Nicholas R. Silvaggi,[†] Langdon J. Martin,[‡] Harald Schwalbe,[§] Barbara Imperiali,[‡] and Karen N. Allen*,[†]

Contribution from the Department of Physiology and Biophysics, Boston University School of Medicine, 715 Albany Street, Boston, Massachusetts 02118, Departments of Chemistry and Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, and Institut für Organische Chemie und Chemische Biologie, Center for Biomolecular Magnetic Resonance, Johann Wolfgang Goethe-Universität Frankfurt, Max von Laue-Strasse 7, 60438 Frankfurt/M, Germany

Received January 22, 2007; E-mail: drkallen@bu.edu

Abstract: A double-lanthanide-binding tag (dLBT), a small peptide sequence engineered to bind two lanthanide ions (e.g., Tb^{3+}) with high affinity, was used to solve the phase problem for the structure determination of ubiquitin by the single-wavelength anomalous diffraction (SAD) method. Since the dLBT is comprised exclusively of encoded amino acids, the necessity for the incorporation of unnatural amino acids or chemical modification of the protein as a prerequisite for X-ray structure determination is eliminated. A construct encoding the dLBT as an N-terminal fusion with ubiquitin provides for facile expression and purification using standard methods. Phasing of the single-wavelength X-ray data (at 2.6 Å resolution) using only the anomalous signal from the two tightly bound Tb^{3+} ions in the dLBT led to clear electron-density maps. Nearly 75% of the ubiquitin structure was built using automated model-building software without user intervention. It is anticipated that this technique will be broadly applicable, complementing existing macromolecular phasing methodologies. The dLBT should be particularly useful in cases where protein derivatization with heavy atoms proves to be problematic or synchrotron facilities are unavailable.

Introduction

Discretely bound metal ions can provide essential information for structural studies using NMR[1] spectroscopy and X-ray crystallography.[2] Existing methods for the incorporation of non-native metal-binding sites into proteins rely on introduction of unnatural amino acids,[3] chemical modification of a reactive natural amino acid side chain,[4,5] or native chemical ligation,[6] all of which generally require considerable manipulation and optimization. Recently, the power of combinatorial screening has been harnessed to create lanthanide-binding tags (LBTs), short polypeptides (15−25 amino acids) derived from calcium-binding motifs that have been iteratively selected to bind trivalent lanthanide ions tightly and selectively.[7−9] The size of the tags is minimal, thereby preventing impact on the structure and function of the proteins to which they are fused. Furthermore, because the LBTs are composed exclusively of encoded amino acids, these tags can be fused to proteins of interest using standard molecular biology techniques. The unique physical properties of LBTs make them promising tools for applications in biology, biochemistry, and biophysics. For example, the paramagnetic properties and anisotropic magnetic susceptibility of the bound lanthanide ion facilitate measurement of residual dipolar couplings, which provide distance and orientational information in NMR structure determination as described in the accompanying article.[10] Of the numerous potential applications of LBT technology, one of the most attractive is the use of LBTs as tools for crystallographic phasing of macromolecular structures.

In order to determine the atomic structure of a macromolecule by X-ray diffraction, it is necessary to determine both the amplitudes and the phases of the diffracted X-rays. The amplitude data are collected directly as intensities in the diffraction experiment; however, to obtain phases, techniques such as multiple isomorphous replacement (MIR),[2,11] multi-wavelength anomalous diffraction (MAD),[12,13] or single-

[†] Boston University School of Medicine.
[‡] Massachusetts Institute of Technology.
[§] Johann Wolfgang Goethe-Universität Frankfurt.

(1) Banci, L.; Bertini, I.; Huber, J. G.; Luchinat, C.; Rosato, A. *J. Am. Chem. Soc.* **1998**, *120*, 12903−12909.
(2) Harker, D. *Acta. Crystallogr.* **1956**, *9*, 1−9.
(3) Hendrickson, W. A.; Horton, J. R.; LeMaster, D. M. *EMBO J.* **1990**, *9*, 1665−1672.
(4) Bunni, M. A.; Baindur, S. R.; Douglas, K. T.; Drake, A. F. *Biochim. Biophys. Acta* **1987**, *923*, 421−430.
(5) Yamasaki, R. B.; Osuga, D. T.; Feeney, R. E. *Anal. Biochem.* **1982**, *126*, 183−189.
(6) Cotton, G. J.; Muir, T. W. *Chem. Biol.* **1999**, *6*, R247−R256.
(7) Nitz, M.; Franz, K. J.; Maglathlin, R. L.; Imperiali, B. *ChemBioChem* **2003**, *4*, 272−276.
(8) Franz, K. J.; Nitz, M.; Imperiali, B. *ChemBioChem* **2003**, *4*, 265−271.
(9) Martin, L. J.; Sculimbrene, B. R.; Nitz, M.; Imperiali, B. *QSAR Comb. Sci.* **2005**, *24*, 1149−1157.
(10) Martin, L. J.; Hähnke, M. J.; Nitz, M.; Wöhnert, J.; Silvaggi, N. R.; Allen, K. N.; Schwalbe, H.; Imperiali, B. *J. Am. Chem. Soc.* **2007**, *129*, 7106−7113.
(11) Green, D.; Ingram, V.; Perutz, M. *Proc. R. Soc. London Ser. A, Math. Phys. Sci.* **1954**, *225*, 287−307.
(12) Hendrickson, W. A.; Smith, J. L.; Sheriff, S. *Methods Enzymol.* **1985**, *115*, 41−55.

10.1021/ja070481n CCC: $37.00

wavelength anomalous diffraction (SAD)[14] must be used. With the exception of SAD phasing using the anomalous signal from sulfur atoms and cases where the protein contains a native anomalously scattering atom (e.g., heme-binding proteins), all three phasing methods require the incorporation of heavy atoms (e.g., Hg, U, or Se) into the protein structure.

Incorporation of heavy atoms into the crystal lattice can be accomplished in a number of ways; the simplest method is to soak protein crystals in solutions containing heavy atoms or organometallic compounds. However, the empirical nature of the method can slow the process of obtaining heavy-atom derivatives that do not disrupt the crystal lattice (non-isomorphism). An attractive alternative approach is the incorporation of selenium into recombinant proteins during expression in the presence of selenomethionine (SeMet) with *Escherichia coli* methionine auxotrophs.[3] This technique circumvents the search for derivatives and solves the problem of non-isomorphism. For these reasons, SeMet MAD is currently the most rapid and efficient phasing method. However, there can be significant obstacles preventing the use of SeMet protein. First, a minimal number of methionine residues must occur in the protein sequence to provide a measurable anomalous signal (one Met per ~80 residues). Also, the incorporation of SeMet can dramatically reduce the yield of soluble, correctly folded protein, increasing the time and expense required to obtain sufficient material. Finally, since methionine is incorporated preferentially over the selenium-containing analogue, SeMet may not efficiently replace every methionine.

Given these challenges, complementary methods, including phasing using lanthanide ions, may be advantageous if they can be incorporated routinely. At wavelengths where diffraction data are routinely collected, most of the lanthanides have larger anomalous signals than selenium. In fact, the expected anomalous signal at the peak wavelength is calculated to be as much as 4 times greater for Tb than that for Se (Table 1), giving Tb approximately 4 times the phasing power. An additional advantage is that the anomalous signal is still large far from the absorption edges, so that phasing at a single wavelength (SAD phasing) can be used with Cu Kα radiation.

The utility of lanthanide ions in X-ray structure determination has been demonstrated by exploiting native calcium-binding proteins.[16,17] Lanthanide ions have also been incorporated using a chelating agent that is introduced via chemical modification,[18] or by co-crystallizing proteins with neutral gadolinium complexes, such as Gd−HP DO3A.[19−23] However, currently there is no general method for incorporating lanthanide ions directly and specifically into recombinant proteins. A simple protein expression tag that binds lanthanide ions would be an extremely useful tool to complement existing macromolecular phasing techniques.

Table 1. Comparison of the Phasing Power of Se and Tb

	edge	f' (e^-)[a]	f'' (e^-)[a]	energy (keV)	wavelength (Å)
Se	K	−8.3	3.8	12.6578	0.9795
Tb	L-I	−9.0	12.8	8.7080	1.4238
	L-II	−13.3	12.2	8.2516	1.5025
	L-III	−19.5	10.6	7.5140	1.6500

	wavelength (Å)	expected anomalous signal[b] (e^-) 1000 non-H atoms	3500 non-H atoms
Se	0.9795	0.0363	0.0194
Tb	1.4238	0.1208	0.0646
	1.5418 (Cu Kα)	0.0860	0.0460
	1.6500	0.1000	0.0535
	1.1000	0.0807	0.0432
S	1.5418 (Cu Kα)	0.0052	0.0028

[a] These are theoretical values calculated using the Cromer−Liberman approximation.[15] Such theoretical values do not take into account the effects of neighboring atoms, and so they do not accurately predict the behavior of atoms near absorption edges. The experimental values of f' and f'', as well as the energy of the absorption edge, are likely to differ somewhat from the values listed here. [b] Calculated using the formula[14] $\langle \Delta F_{ANOM} \rangle / \langle F \rangle = (2N_A/N_T)^{1/2}(f''_A/Z_{eff})$, where N_A is the number of anomalously scattering atoms per asymmetric unit (assumed to be 2 in all cases), N_T is the total number of non-hydrogen atoms per asymmetric unit, f''_A is the anomalous scattering contribution, and Z_{eff} is the average number of electrons (estimated to be 6.7) for non-hydrogen protein atoms.

Herein we describe the first use of a lanthanide-binding coexpression tag for phasing in macromolecular structure determination. In order for this approach to succeed, the lanthanide tag must be well ordered with respect to the macromolecule. Indeed, initial attempts to use single-LBTs for phasing failed, presumably because the tag was mobile with respect to the protein. However, successful NMR experiments (see accompanying article)[10] utilizing a "double-LBT" (dLBT), wherein two Tb^{3+}-binding modules are concatenated in a single 32-residue peptide, prompted us to investigate the use of this tag in macromolecular structure determination.

We demonstrate that a dLBT fusion with ubiquitin, **dLBT3**-ubiquitin, can be generated using standard molecular biology techniques, expressed with high efficiency, loaded with Tb^{3+}, and crystallized from commercially available screens.[24] The X-ray diffraction data from the resultant crystals were used to solve the structure via the SAD phasing method in an automated fashion. This work provides the first proof-of-concept of the applicability of dLBTs in X-ray crystallography and underscores the potential for implementing these tools in automated structure solution technologies.

Experimental Section

Protein Expression and Purification. The **dLBT3**-ubiquitin was cloned, expressed, and purified as described[10] with a modification of the cell lysis procedure. The amino acid sequence of the **dLBT3** tag was GPGYIDTNNDGWIEGDELYIDTNNDGWIEGDELLA. For lysis, the cells were resuspended in a buffer consisting of 50 mM Tris (pH 8.0), 300 mM NaCl, and 10 mM imidazole at a volume of 5 mL per gram of cells. Cells were then lysed in one freeze−thaw cycle (−20 and 25 °C) in the presence of 0.5 mg/mL hen egg white lysozyme

(13) Kahn, R.; Fourme, R.; Bosshard, R.; Chiadmi, M.; Risler, J. L.; Dideberg, O.; Wery, J. P. *FEBS Lett.* **1985**, *179*, 133−137.
(14) Hendrickson, W. A.; Teeter, M. M. *Nature* **1981**, *290*, 107−113.
(15) Cromer, D. T.; Liberman, D. *J. Chem. Phys.* **1970**, *53*, 1891−1898.
(16) Weis, W. I.; Kahn, R.; Fourme, R.; Drickamer, K.; Hendrickson, W. A. *Science* **1991**, *254*, 1608−1615.
(17) Weis, W. I.; Crichlow, G. V.; Murthy, H. M.; Hendrickson, W. A.; Drickamer, K. *J. Biol. Chem.* **1991**, *266*, 20678−20686.
(18) Purdy, M. D.; Ge, P.; Chen, J.; Selvin, P. R.; Wiener, M. C. *Acta Crystallogr D* **2002**, *58*, 1111−1117.
(19) Girard, E.; Chantalat, L.; Vicat, J.; Kahn, R. *Acta Crystallogr. D* **2002**, *58*, 1−9.
(20) Girard, E.; Anelli, P. L.; Vicat, J.; Kahn, R. *Acta Crystallogr. D* **2003**, *59*, 1877−1880.
(21) Girard, E.; Stelter, M.; Anelli, P. L.; Vicat, J.; Kahn, R. *Acta Crystallogr. D* **2003**, *59*, 118−126.
(22) Girard, E.; Stelter, M.; Vicat, J.; Kahn, R. *Acta Crystallogr. D* **2003**, *59*, 1914−1922.
(23) Girard, E.; Pebay-Peyroula, E.; Vicat, J.; Kahn, R. *Acta Crystallogr. D* **2004**, *60*, 1506−1508.
(24) Jancarik, J.; Kim, S.-H. *J. Appl. Crystallogr.* **1991**, *24*, 409−411.

Table 2. Data Collection, Structure Determination, and Refinement Statistics

Data Collection	
space group and unit cell	$P3_221$; $a = b = 57.7$ Å, $c = 115.0$ Å
wavelength (Å)	1.1000
resolution limits (Å) (highest resolution shell)	50−2.60 (2.69−2.60)
no. of reflections	
measured	61078
unique	7214
completeness (%)	
all data (highest resolution shell)	98.4 (87.7)
R_{sym}[a] (on I) (highest resolution shell)	0.127 (0.378)
$[I/\sigma(I)]$	
all data (highest resolution shell)	15.8 (2.8)
Structure Determination	
$\langle m \rangle_{BnP}$ (%)	0.252
$\langle m \rangle_{RESOLVE}$ (%)	0.610
Refinement	
resolution (Å)	30.5−2.6
R factor (all data)	0.218
R free	0.254
reflections in test set	285
non-hydrogen atoms	891
rms deviations	
bond lengths (Å)	0.013
angles (°)	1.601
average B factor (Å^2) (all atoms)	60.0

[a] $R_{sym} = \Sigma|I_{obs} - \langle I \rangle|/\Sigma I_{obs}$.

(Sigma-Aldrich). Cell debris was removed from the lysate by centrifugation at 39000*g*. After purification, the protein was dialyzed against 4 L of 20 mM HEPES (pH 7.5) and 100 mM NaCl. A small amount of white, flocculent precipitate appeared after dialysis, which was removed by centrifugation. The >95% pure **dLBT3**-ubiquitin was diluted to 1 mg/mL in the dialysis buffer, and $TbCl_3$ (1 mM in 25 mM ammonium acetate, 1 mM HCl) was added in a 2:1 molar ratio (Tb^{3+}:protein) in 10 aliquots with 2 min of incubation at room temperature after each addition. Finally, the Tb^{3+}-loaded **dLBT3**-ubiquitin was concentrated to 17 mg/mL using an Amicon centrifugal concentrator having a nominal MWCO of 3500 Da.

Crystallization and Data Collection. Crystals of the **dLBT3**-ubiquitin were grown by the vapor diffusion method with hanging-drop geometry. Initial crystallization conditions were identified from the Index Screen (Hampton Research, Aliso Viejo, CA). The most promising condition, 0.1 M HEPES (pH 7.5) and 3.0 M NaCl, yielded thin, hairlike crystals within seconds. The crystal morphology was improved by optimization to 0.1 M HEPES (pH 7.5) and 3.7 M NaCl, and further improvement in size was achieved by decreasing the rate of crystal growth by adding 100% 2-methyl-2,4-pentanediol (MPD) directly to the crystallization drops. Diffraction quality crystals (hexagonal rods 0.5−1.0 mm long and 0.05 mm in cross section) were obtained by mixing 1 μL of 17 mg/mL protein solution (20 mM HEPES, pH 7.5, 100 mM NaCl) with 1 μL of well solution and 1 μL of 100% MPD. The crystal was moved directly from the mother liquor and flash-cooled in a gaseous nitrogen stream at 100 K. X-ray diffraction data were collected from a single crystal using 1.1000 Å wavelength radiation at Beamline X29A at the National Synchrotron Light Source of the Brookhaven National Laboratory. Data were integrated and scaled using the programs DENZO and SCALEPACK.[25] Data collection and model refinement statistics are reported in Table 2.

Structure Determination and Model Refinement. The structure of **dLBT3**-ubiquitin was phased via the SAD method, using the single-wavelength 2.6 Å data set. The Tb^{3+} substructure was solved using BnP,[26] which located the two Tb^{3+} sites. Examination of the anomalous difference Patterson map confirmed the sites identified by BnP. The iterative model building and refinement script in RESOLVE[27] was used to build an initial model, which comprised 75% of the total residues in the final structure. This starting model was subsequently improved by cycles of manual rebuilding in the graphics program COOT,[28] alternating with refinement in REFMAC5[29,30] using the Hendrickson−Lattman coefficients from BnP. Iterative cycles of restrained refinement and rebuilding ultimately yielded a model containing 110 of 111 residues, 18 water molecules, and 2 Tb^{3+} ions. Assessment of the final model by MOLPROBITY[31] indicated that 96.3% of the residues were in the favored regions of the Ramachandran plot, 3.7% were in the additionally allowed regions, and 0.0% were in the disallowed region. The relatively low resolution and anisotropy of the diffraction data account, at least in part, for the slight abnormalities in the Ramachandran analysis.

Results and Discussion

Structure Determination. Recombinant **dLBT3**-ubiquitin was expressed in standard BL21(DE3) cells and purified using conventional immobilized metal-affinity chromatography (IMAC) in a single-column purification protocol followed by cleavage of the His_6 tag. Crystals of purified **dLBT3**-ubiquitin were obtained directly from the Index Screen (Hampton Research). Optimization of the crystallization conditions resulted in hexagonal rods measuring approximately 0.5−1.0 mm long and 0.05 mm wide. When illuminated with a hand-held ultraviolet lamp (254 nm), the crystals were luminescent, indicating that the Tb^{3+} was bound to the dLBT (Figure 1A,B).

X-ray diffraction data were collected at Beamline X29A of the National Synchrotron Light Source at Brookhaven National Laboratory using 1.1000 Å radiation. This wavelength was selected because it is remote from all of the Tb^{3+} absorption edges and therefore more closely resembles data collection conditions at a Cu Kα home source (Table 1). The crystals are in the $P3_221$ space group ($a = b = 57.7$ Å, $c = 115.0$ Å) and contain one monomer in the asymmetric unit. The low resolution (2.6 Å) of the diffraction was unexpected, given the relatively large crystal size and small protein molecular weight. In addition, the diffraction data were somewhat weak and anisotropic (1.9 Å diffraction limit along the X coordinate of the detector and <3.0 Å along Y; Figure 1C). The crystals had an unusually high solvent content (68%), which explains, at least in part, the poor quality of the diffraction data (Table 2). Experiments with other LBT-tagged proteins (unpublished data) indicate that high solvent content is not a general feature of LBT fusion-protein crystals. Since the quality of the data in this case was less than ideal, this data set was not a "best case scenario", but rather it is representative of typical X-ray diffraction data. Therefore, the data thus far suggest that LBT technology may be applicable to a number of real-world crystallographic problems.

The structure of **dLBT3**-ubiquitin was determined using the SAD method (structure factors and coordinates were deposited

(25) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307−326.

(26) Weeks, C. M.; Blessing, R. H.; Miller, R.; Mungee, R.; Potter, S. A.; Rappleye, J.; Smith, G. D.; Xu, H.; Furey, W. *Z. Kristallogr.* **2002**, *217*, 686−693.

(27) Terwilliger, T. C.; Berendzen, J. *Acta Crystallogr. D* **1999**, *55* (Pt. 4), 849−61.

(28) Emsley, P.; Cowtan, K. *Acta Crystallogr. D* **2004**, *60*, 2126−2132.

(29) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr. D* **1997**, *53*, 240−255.

(30) Collaborative Computational Project. *Acta Crystallogr. D* **1994**, *50*, 760−763.

(31) Davis, I. W.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Nucleic Acids Res.* **2004**, *32*, W615−W619.

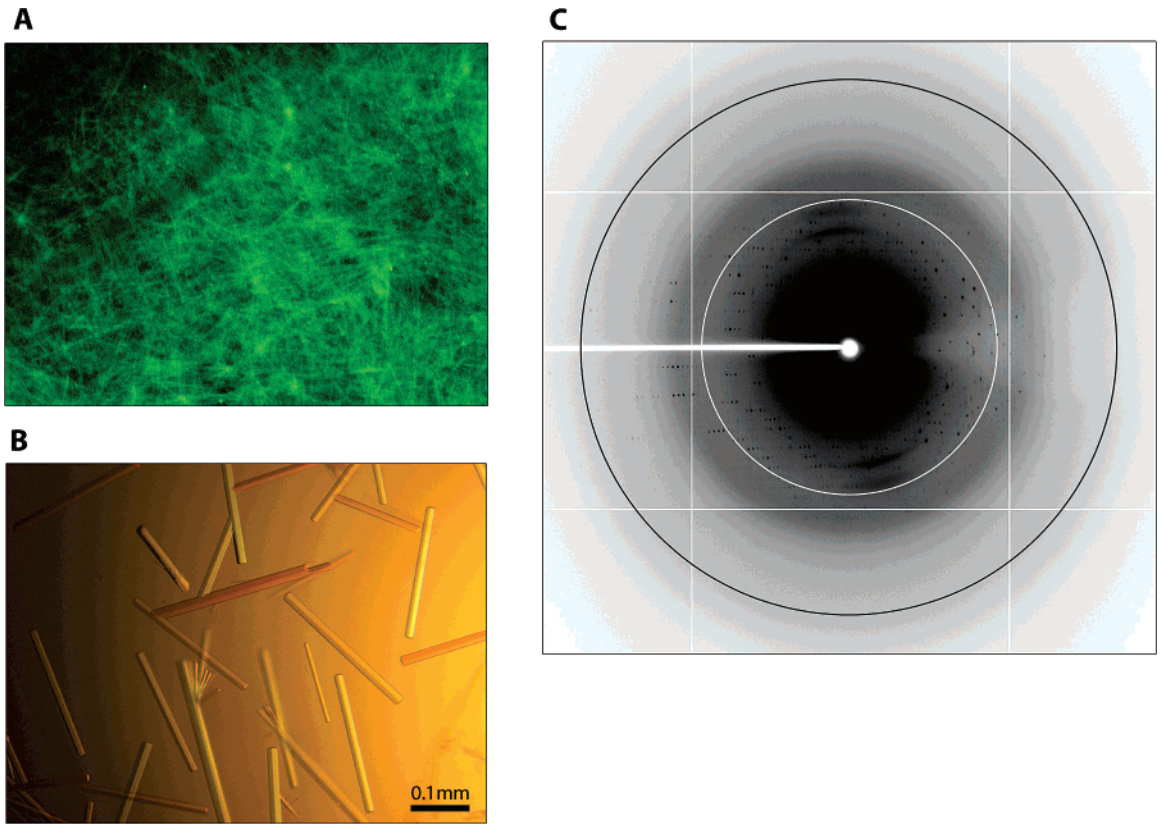

Figure 1. (A) Photomicrograph of the initial **dLBT3**-ubiquitin crystals under ultraviolet illumination. (B) After optimization of the crystallization conditions, the fusion protein crystallized as narrow, hexagonal rods, pictured here under conventional illumination. (C) Detail of a representative diffraction image from the data set collected at NSLS Beamline X29A. Note the low intensity of the diffracted X-rays above ~3.5 Å and the anisotropic nature of the diffraction. The white and black rings represent approximately 3.5 and 1.9 Å resolution, respectively.

in the Protein Data Bank with PDB ID 2OJR), taking advantage of the large and robust anomalous signal from the bound Tb^{3+} ions. The positions of the two anomalously scattering Tb^{3+} atoms were located using the Shake-n-Bake algorithm[32] as implemented in the BnP suite.[26] Initial phases were calculated by the PHASES component of BnP (FoM = 0.252) and improved by solvent flattening in RESOLVE[33] (FoM = 0.610). This compares very favorably with the original ubiquitin structure, which was solved at 1.8 Å resolution by isomorphous replacement using a single mercuric acetate derivative (FoM = 0.68).[34] Iterative cycles of automated model building and refinement in RESOLVE and REFMAC5[29,30] produced a model containing 55 of the 76 residues of ubiquitin. Thus, while the data were collected at 1.1000 Å, 2.56 keV, which is remote from the nearest absorption edge of Tb^{3+} (Table 1), structure determination proceeded in a fully automated fashion. This result indicates that LBT technology can be implemented for structure determination via SAD phasing in situations where synchrotron radiation is not available. Currently, the lack of tunable X-ray sources in most laboratories limits users without access to synchrotron radiation to structure determination by molecular replacement, MIR, or SAD using the anomalous signal from sulfur, a technique pioneered by Hendrickson and Teeter.[14] However, sulfur has a small anomalous signal at 1.54 Å, which is the single wavelength commonly available from in-house Cu Kα X-ray sources (Table 1), and therefore requires high-quality, high-redundancy data in order to detect it. In contrast, the anomalous signal of Tb^{3+} at this wavelength is 16 times larger (Table 1). Given these facts and the performance of the LBT in these experiments, LBT technology has the potential to be a valuable tool for structure determination in the home laboratory.

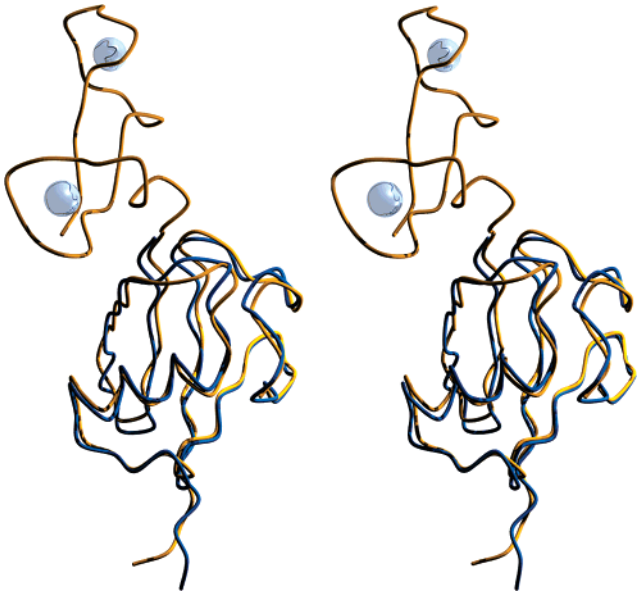

Figure 2. Stereoview showing the Cα trace of **dLBT3**-ubiquitin (orange) overlaid by least-squares fitting to the published structure of ubiquitin (blue, PDB accession code 1UBQ). The rms deviation for Cα atoms in residues 2−72 is 0.80 Å. The Tb^{3+} ions bound by **dLBT3** are represented by metallic spheres.

(32) Hauptman, H. A. *Methods Enzymol.* **1997**, *277*, 3−13.
(33) Terwilliger, T. C. *Acta Crystallogr. D* **1999**, *55*, 1863−1871.
(34) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1987**, *194*, 531−544.

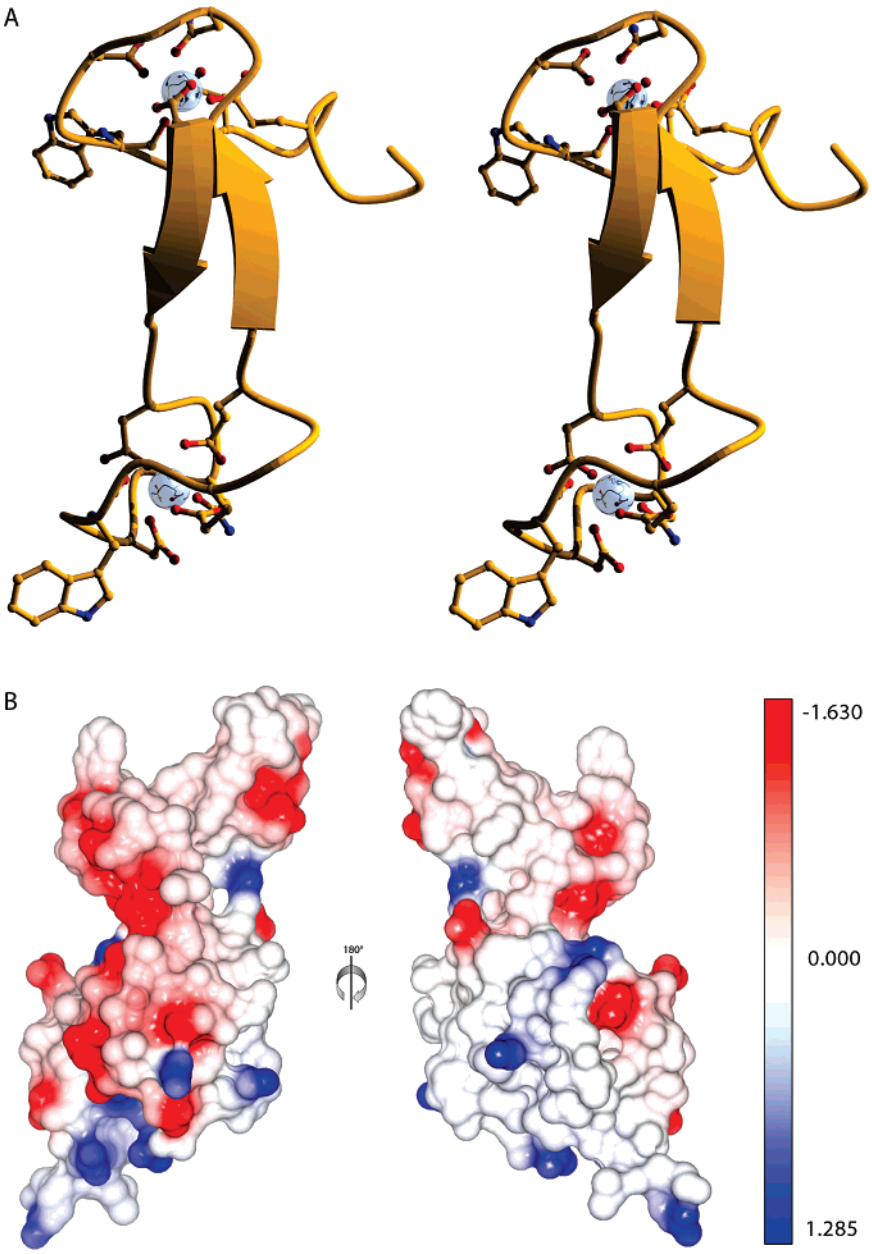

Figure 3. (A) Stereoview showing the overall structure of the **dLBT3** lanthanide-binding tag portion of the **dLBT3**-ubiquitin fusion protein. Residues forming ligands to the Tb^{3+} ions are shown as orange cylinders, and the Tb^{3+} ions are depicted as metallic spheres. (B) Surface representation of the **dLBT3**-ubiquitin fusion protein colored on the basis of electrostatic potential.[37] Note the absence of complementary charges between the **dLBT3** (top) and ubiquitin (bottom) components of the structure.

Overall Structure. The structure of ubiquitin is well characterized (PDB accession code 1UBQ).[34] The ubiquitin structure determined using dLBT-derived phases does not differ significantly from the published ubiquitin structure (0.80 Å root-mean-square deviation for ubiquitin Cα atoms in residues 2−72; Figure 2). Thus, the presence of the dLBT does not affect the final fold of the protein to which it is fused. The **dLBT3** portion of this structure is completely ordered, including, as expected from the robust phasing, both Tb^{3+}-binding sites. In addition, both Tb^{3+}-binding sites of **dLBT3** are very similar to the structure of the free, single-LBT peptide determined previously.[35] As described for the single-LBT peptide, each Tb^{3+}-binding site comprises an EF-hand fold with eight ligands to the metal ion, seven through the Asp, Asn, and Glu side chains, and one through a backbone carbonyl oxygen. One somewhat unexpected but intriguing feature of the **dLBT3** structure was the formation of a short antiparallel β-sheet between the Gly-Pro-Gly sequence of the DAPase stop site and the sequence linking the two LBT modules of **dLBT3** (Figure 3A). This aspect of the **dLBT3** structure is discussed below.

Notably, the ∼430 Å^2 interface between the dLBT and ubiquitin is nearly half the size expected for the interfaces of dimeric proteins of similar size.[36] Additionally, there are no complementary charges between the **dLBT3** and ubiquitin (Figure 3B). Taken together, the limited buried surface area and the lack of electrostatic interactions suggest that there are no fortuitous protein−protein contacts stabilizing the dLBT with

(35) Nitz, M.; Sherawat, M.; Franz, K. J.; Peisach, E.; Allen, K. N.; Imperiali, B. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 3682−3685.

(36) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204*, 155−164.

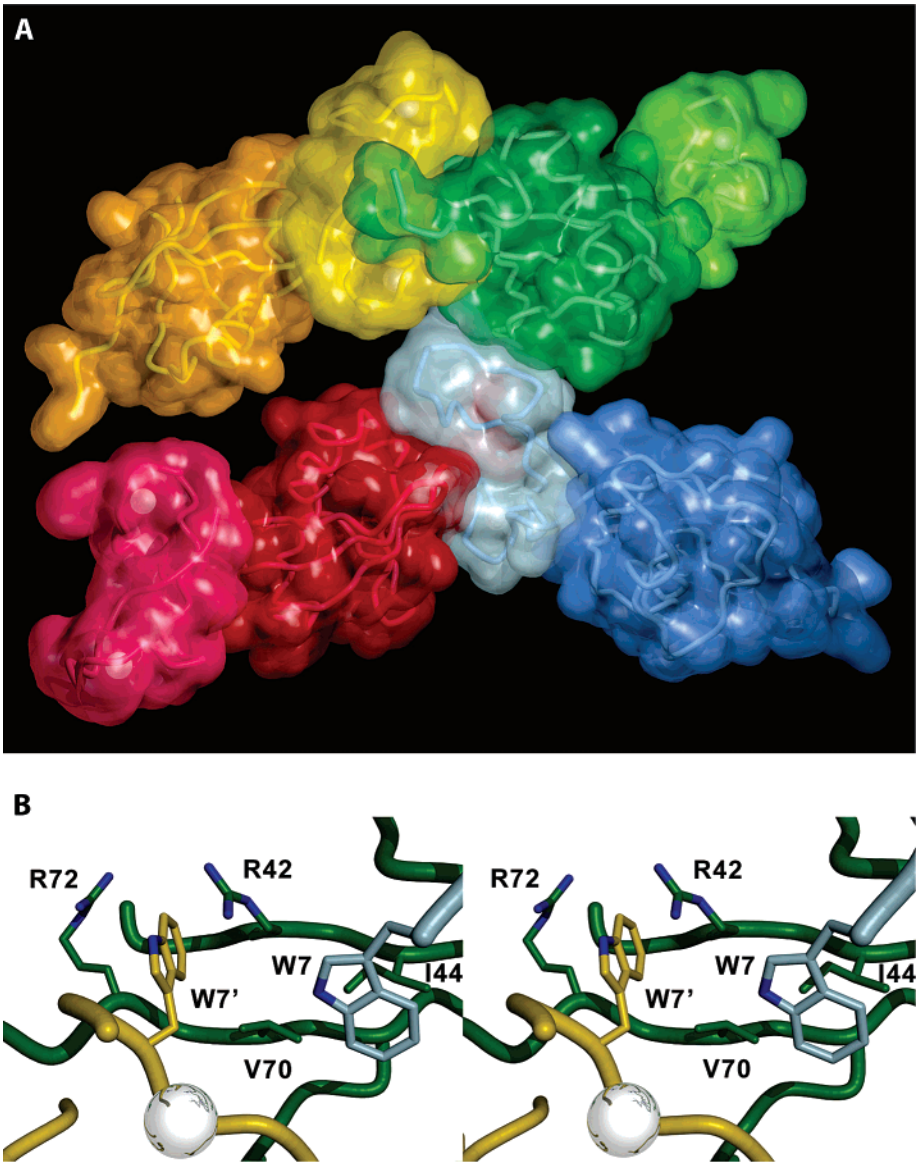

Figure 4. (A) Surface representation of symmetry-related **dLBT3**-ubiquitin molecules in the crystal lattice. The dLBT portion of each monomer is depicted in a lighter shade. Note that all intermolecular contacts are formed between the **dLBT3** of one monomer and the ubiquitin of a neighboring monomer; there are no direct contacts between neighboring ubiquitin domains. (B) Stereoview showing the details of the crystal contacts between the **dLBT3** and ubiquitin in the crystal. Monomers are colored as in (A). The LBT residue numbering is based on the literature.[39,40]

respect to the protein in this particular fusion. This provides further evidence of the general applicability of the LBTs.

It is instructive to contemplate the structural features of the dLBT that contribute to its successful application in structure determination. Preliminary studies (unpublished results) with single-LBT−protein fusions failed to provide phasing or reveal the structure of the attached LBT, presumably due to disorder of the tag relative to the protein. Clearly, the LBT, and hence the Tb^{3+} ion(s), must be well ordered with respect to the protein to be useful for structure determination. What factors provided the necessary immobilization of the successful **dLBT3**-ubiquitin fusion protein that distinguished it from its progenitors? Critically, the dLBT contains two EF-hand motifs instead of one, and is thus nearly twice as large as the single-LBTs. The increased size results in a less mobile tag (as shown by NMR analysis in the accompanying article[10]) that is sufficiently ordered with respect to ubiquitin to form crystal contacts. An additional aspect of the dLBT is the presence of the Gly-Pro-Gly DAPase stop site, which is involved in a short β-sheet between the two LBT modules of the **dLBT3** tag, resulting in a rigid overall structure. This is evidenced by the similar average *B* factors of the two halves of the dLBT (61.7 and 53.3 $Å^2$ for the N- and C-terminal halves, respectively, compared to the ubiquitin portion at 61.6 $Å^2$). The slightly higher average *B* factor of the N-terminal half of the LBT could be explained by its position at the end of the polypeptide chain.

Although all of the crystal contacts are mediated by the dLBT (Figure 4A), there do not appear to be any specialized contacts between symmetry-related ubiquitin molecules and the dLBT. The Trp7 side chain, in the N-terminal Tb^{3+}-binding site of the dLBT, is interacting with Ile44 and Val70 of an adjacent **dLBT3**-ubiquitin monomer. Ile44 and Val70 comprise the so-called hydrophobic patch of ubiquitin involved in recognition of a number of ubiquitin-binding proteins[38] (Figure 4B). The

(37) Nicholls, A.; Sharp, K. A.; Honig, B. *Proteins* **1991**, *11*, 281−296.
(38) Hurley, J. H.; Lee, S.; Prag, G. *Biochem. J.* **2006**, *399*, 361−372.

other tryptophan side chain, Trp7′, slips between the side chains of Arg42 and Arg72 of a second symmetry-related molecule (Figure 4B). Although Trp7 occupies a known hydrophobic protein−protein interaction site on ubiquitin, the interface consists of only two small hydrophobic side chains. Such small hydrophobic surface patches are not unusual features of proteins. Likewise, the occurrence of arginine side chains in close proximity is not an uncommon occurrence on protein surfaces. Thus, similar crystal packing arrangements should be possible with other LBT fusion proteins.

Conclusions

The structure of **dLBT3**-ubiquitin is the first X-ray crystal structure to be determined from a coexpressed lanthanide-binding tag. Production of the tagged protein required only encoded amino acids and standard molecular biology methods, obviating the need for unnatural amino acids or chemical modification of the protein structure. The LBT performed well, providing good-quality phases despite the relatively poor quality of the diffraction data, and the remote wavelength used for data collection. We anticipate that the LBTs will be particularly useful in situations where synchrotron radiation is unavailable or where SeMet fails to provide adequate phasing power. This study indicates that the dLBT technology will complement currently available techniques and augment the suite of tools available to structural biologists.

Acknowledgment. The authors thank the staff of Beamline X29A of the National Synchrotron Light Source, Brookhaven National Laboratory, for their aid during data collection. This research was supported by NSF-CRC program grant CHE-0304832 from the National Science Foundation (to K.N.A., B.I., and H.S.) and by National Institutes of Health training grant HL07291 (to N.R.S.).

JA070481N

(39) Marsden, B. J.; Hodges, R. S.; Sykes, B. D. *Biochemistry* **1988**, *27*, 4198−4206.
(40) MacManus, J. P.; Hogue, C. W.; Marsden, B. J.; Sikorska, M.; Szabo, A. G. *J. Biol. Chem.* **1990**, *265*, 10358−10366.