Double trouble—Buffer selection and His-tag presence may be responsible for nonreproducibility of biomedical experiments

Karolina A. Majorek,1,2,3,4 Misty L. Kuhn,4,5 Maksymilian Chruszcz,1,3,4,6 Wayne F. Anderson,3,4,5 and Wladek Minor1,3,4*

1Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia 22908
2Bioinformatics Laboratory, Institute of Molecular Biology and Biotechnology, Faculty of Biology, Adam Mickiewicz University, 61–614 Poznan, Poland
3Midwest Center for Structural Genomics, USA
4Center for Structural Genomics of Infectious Diseases (CSGID), USA
5Department of Molecular Pharmacology and Biological Chemistry, Northwestern University Feinberg School of Medicine, Chicago, Illinois 60611
6Department of Chemistry and Biochemistry, University of South Carolina, Columbia, South Carolina 29208

Received 29 May 2014; Revised 6 July 2014; Accepted 11 July 2014
DOI: 10.1002/pro.2520
Published online 16 July 2014 proteinscience.org

Abstract: The availability of purified and active protein is the starting point for the majority of in vitro biomedical, biochemical, and drug discovery experiments. The use of polyhistidine affinity tags has resulted in great increases of the efficiency of the protein purification process, but can negatively affect structure and/or activity measurements. Similarly, buffer molecules may perturb the conformational stability of a protein or its activity. During the determination of the structure of a Gcn5-related N-acetyltransferase (GNAT) from Pseudomonas aeruginosa (PA4794), we found that both HEPES and the polyhistidine affinity tag bind (separately) in the substrate-binding site. In the case of HEPES, the molecule induces conformational changes in the active site, but does not significantly affect enzyme activity. In contrast, the uncleaved His-tag does not induce major conformational changes but acts as a weak competitive inhibitor of peptide substrate. In two other GNAT enzymes, we observed that the presence of the His-tag had a strong influence on the activity of these proteins. The influence of protein preparation on functional studies may affect the reproducibility of experiments in other laboratories, even when changes between protocols seem at first glance to be insignificant. Moreover, the results presented here show how critical it is to adjust the experimental conditions for each protein or family of proteins, and investigate the influence of these factors on protein activity and structure, as they may significantly alter the effectiveness of functional characterization and screening methods. Thus, we show that a polyhistidine tag and the buffer molecule HEPES bind in the substrate-binding site and influence the conformation of the active site and the activity of GNAT acetyltransferases. We believe that such discrepancies can influence the reproducibility of some experiments and therefore could have a significant “ripple effect” on subsequent studies.

Grant sponsor: National Institute of Allergy and Infectious Diseases; Grant numbers: GM053163; GM094585. Grant sponsor: National Institutes of Health, Department of Health and Human Services; Grant number: HHSN272201200026C.

*Correspondence to: Wladek Minor, Department of Molecular Physiology and Biological Physics, University of Virginia, 1340 Jefferson Park Ave., Charlottesville, VA 22908. E-mail: wladek@iwonka.med.virginia.edu.

Published by Wiley-Blackwell. © 2014 The Protein Society

PROTEIN SCIENCE 2014 | VOL 23:1359–1368 1359
Keywords: reproducibility of biomedical experiments; acetyltransferase; polyhistidine affinity tag; His-tag; HEPES; GNAT

Introduction
For nearly 3 decades, polypeptide fusion affinity tags have been used to easily and efficiently produce large quantities of highly purified recombinant proteins. Affinity tags are produced by altering the sequence of a gene to add specific polypeptide sequences at the N- and/or C-terminus of an expressed protein that provides a “handle” for a specific purification.1 Among the most widely used types of affinity tags is the polyhistidine tag (His-tag), which allows the expressed protein to be purified through binding to immobilized metal cations such as Ni2+ or Co2+.

Typically, a His-tag contains six or more consecutive histidine residues at either the N- or C-terminus of the protein. The polyhistidine sequence of the tag may be removed via proteolytic cleavage4 if a specific endoprotease recognition sequence is also introduced between the tag and the main polypeptide sequence. This also permits additional “subtractive” purification by passing cleaved protein through a metal affinity column and collecting the flowthrough. However, tag removal is sometimes impossible or inefficient as cleavage of the tag requires additional purification steps, and thus the tag is sometimes left uncut (particularly in high-throughput protein purification). Additionally, leaving a His-tag is sometimes desirable because it improves protein solubility or alters crystal packing.

Since His-tags are significantly smaller than most other types of affinity tags, they are often thought to only minimally perturb protein structure and function. However, there is evidence that suggests this is not always true.5 For instance, the tag may interfere with the protein’s ability to bind ligands, cause the protein to aggregate, or change its solubility.4 In particular, purification of metalloproteins using polyhistidine is often problematic, as both the adsorption matrix and the affinity tag itself may potentially strip metal ions from the protein.1

Many examples of how a His-tag can affect protein structure, activity, or binding affinity have been described in the literature. For instance, the rat corticotropin-releasing factor receptor type 2a adopts two different disulfide bond patterns at the N-terminus when the His-tag is at either the N- or C-terminus. However, the tag did not significantly influence the binding affinities of the ligands used in the study.6 On the other hand, the His-tag placement on the tumor-associated single-chain Fv protein significantly affects its ability to bind to the epitope.7 Finally, the substrate specificity, kinetic constants, or protein dynamics may also be affected by His-tag placement in certain proteins. Some of these include the enantioselectivity of some lipases from Staphylococcus8 and the protein dynamics of myoglobin.9

These changes can be also observed in crystal structures. Over 16,000 protein structures determined by X-ray diffraction have a polyhistidine tag reported in their SEQRES records, but only a small fraction (over 1000, or about 6%) have all or part of the His-tag in the final refined model. In many cases, the His-tag is indeed too flexible to be modeled, but in some cases it is not modeled, even though clear electron density for it is present (see Results). Sometimes large conformational changes are observed at sites distant from the His-tag, as at the DNA-binding site of the gene regulatory protein AreA from Aspergillus nidulans.10 The His-tag can also mimic interactions between proteins and/or affect the oligomeric state. For example, the structure of the Archaeoglobus fulgidus AF0173 redox maturation protein shows that the His-tag mimics a portion of the twin-arginine motif that binds to this protein.11

In addition to polypeptide fusion affinity tags, another factor that may perturb protein stability is buffer composition. The choice of buffer is crucial, not only for hydrogen ion buffering capacity at the desired pH, but also because of other unwanted chemical effects.12 Buffers have been shown to interfere with proteins by reacting with protein substrates or acting as inhibitors of enzymatic reactions.13,14 Buffers are also frequently observed in crystal structures. Ordered HEPES buffer molecules, for example, are found in hundreds of protein structures and have yielded a variety of information about substrate binding and protein conformational changes that occur during catalysis. For instance, HEPES bound in the structure of an endoglucanase mimics the substrate, which changes the conformation of the protein to the active state.15 Other structures that have HEPES bound in the substrate-binding site and mimics the native substrate include an l-serine dehydrogenase PA0743 from Pseudomonas aeruginosa16 and an aminoglycoside 6’-N-acetyltransferase from Salmonella typhimurium, where HEPES acts as a transition state analog of the reaction.17 While nonphysiological buffers are typically needed to maintain protein stability for in vitro studies, these artificial components can sometimes interact with the system being studied.

In this work, we describe the impact of both polyhistidine affinity tags and buffer molecules on...
GNAT activity and structure. The very large GNAT superfamily comprises proteins that acetylate a variety of small molecule substrates like glucosamine-6-phosphate and aminoglycoside antibiotics, as well as large substrates like proteins and histones.\textsuperscript{18,19} GNAT homologs are found in all kingdoms of life, and many organisms have multiple paralogs. Here we show and analyze the structures of the GNAT protein PA4794 from \textit{Pseudomonas aeruginosa}\textsuperscript{20} in complex with either HEPES or the His-tag in the substrate-binding site. We discuss the influence of those factors on PA4794 conformation and activity, and the influence of His-tags on the activity of two other GNATs—spermidine/spermine acetyltransferase SpeG from \textit{Vibrio cholerae}, and SACOL1063, an N-terminal protein acetyltransferase from \textit{Staphylococcus aureus}.\textsuperscript{21}

**Results**

**His-tag and HEPES binding**

Currently, there are several hundred three-dimensional structures of GNATs in the Protein Data Bank (PDB). Nearly one quarter of these were crystallized in the presence of a poly-histidine tag and many have ligands bound, including HEPES. To determine the influence of the His-tag or HEPES on the structure of a GNAT from \textit{Pseudomonas aeruginosa} (PA4794) we produced and analyzed two crystal structures—one of His-tagged apo-protein and one of the protein:HEPES complex. We previously established that PA4794 has acetylation activity, produced complexes of the protein with products of the reaction (CoA and acetylated peptide), and identified three residues important for PA4794 activity: Tyr128, Cys117, and Cys29.\textsuperscript{20}

The apo-PA4794 protein with and without the His-tag, as well as PA4794 complexed with HEPES, crystallized in the orthorhombic crystal system (space group \textit{P}2\textsubscript{1}2\textsubscript{1}2\textsubscript{1}) with a monomer in the asymmetric unit. The PA4794:HEPES complex structure was the first structure obtained, which contained the HEPES buffer molecule bound in the substrate-binding site [Fig. 1(B)]. The complex was produced using protein purified and crystallized in HEPES buffer in the absence of substrates. The residues that interact with the HEPES molecule include the following: (1) the sulfonate group of HEPES is coordinated through several interactions, including the Ne and Nx atoms of Arg141, the O\textsubscript{x} atom of Tyr68 through a water molecule, and several additional water molecules, (2) the oxygen of the hydroxyethyl group of HEPES interacts with the main chain oxygen of Tyr28, Met81, and (3) the oxygen of the hydroxyethyl group and the nitrogen atom of the piperazine ring of HEPES both interact with main chain oxygen atoms of Cys29 and Gly79 through the same water molecule.

The structure of PA4794 bound to HEPES also shows that HEPES binding induces a conformational change between the apo-form of PA4794 and the HEPES-bound form (Fig. 2). The differences are

**Figure 1.** Structures of PA4794 with either His-tag (A) or HEPES (B) bound to the substrate-binding site.

**Figure 2.** Conformational change of PA4794 upon binding HEPES. The apo structure is shown in gray, and the HEPES-bound structure is shown in green. Upon binding of HEPES we observe a flip of the main chain of Gly79 and Asn80, a shift in the main chain of residues 28 and 29, and a conformational change of the Cys29 side chain.
located in the active site area, involving both the main-chain and the side chains. Interestingly, the most distinct change is a flip of the main chain of Gly79 and Asn80. We have not observed this conformational change in any of the structures of PA4794 previously reported, and thus its biological relevance is unclear. On the opposite side of the active site there is a shift of main chain residues 28 and 29 and a conformational change of the Cys29 side chain. This shows that the buffer-bound structure cannot be treated as a truly apo-form due to its changed conformation, and it may also be of limited use for interpreting function.

The protein that was used for the second structure was purified in Tris-HCl buffer instead of HEPES, and crystallization trials were set in conditions that did not contain HEPES. However, the second structure revealed additional electron density for the His-tag. The analysis of the symmetry-related molecules in the crystal shows that the tag of one molecule is bound in the substrate-binding cleft of the protein molecule in the adjacent asymmetric unit, presumably blocking substrate binding [Fig. 1(A)]. The binding of the tag is driven by several interactions: (1) polar interactions through Lys32, (2) hydrophobic interactions including Phe118, Leu151, Phe27, Pro31, Tyr68, and Ile34, and (3) numerous hydrogen-bonding interactions with the protein backbone as well as the side chains. These include interactions of the His-tag backbone and side chains with the main chain oxygen atoms of Pro31 and Ala33, and the main chain and side chain atoms of Lys32. The interaction with Lys32 seems to be most relevant in the context of peptide binding, as it includes hydrogen bonds with three main chain oxygen atoms of the His-tag. This would give strong, but sequence independent, interaction with the peptide, while other residues may be responsible for the substrate specificity.

To determine whether PA4794 structures in the presence of HEPES or the His-tag provide relevant information regarding substrate binding, we compared them with PA4794 structures containing either the acetylated N-phenylacetyl-Gly-Lys (NPAcGK) substrate or AcCoA (Fig. 3). The comparison of the HEPES structure with the acetylated NPAcGK structure shows that the position of both NPAcGK peptide and HEPES are very similar. When the structures of PA4794 that are in complex with either HEPES or AcCoA are superimposed, the acetyl group of AcCoA is very close to the hydroxyethyl group of HEPES. This confirms that HEPES is positioned similarly to the potential substrate, and thus may be mimicking true substrate binding [Fig. 3(B)]. A comparison of the His-tag and acetylated NPAcGK structures shows how the His-tag may compete with NPAcGK during catalysis by blocking the peptide-binding site if the tag is not removed from the protein before kinetic characterization. Although a portion of the His-tag mimics how a peptide may bind, it is not inserted as deeply into the active site as NPAcGK [Fig. 3(A)].

**Influence of the polyhistidine tag on kinetic activity**

We analyzed the kinetics of the activity of PA4794 in the presence of AcCoA and the NPAcGK peptide to determine if the presence of the His-tag affected the function of PA4794. As expected, the predominant effect of the His-tag, which binds in the acetyl group acceptor binding site, was to decrease the enzyme's affinity for the acceptor—NPAcGK peptide (Fig. 4). Alternatively, the tag had relatively little effect on the enzyme's maximal velocity or affinity for AcCoA (which binds in a different binding site that is
inaccessible to the His-tag; Fig. 4). Contrary to the effects of the His-tag, HEPES did not significantly affect the activity of the enzyme even at very high millimolar (>200) concentrations.

To investigate the possible effect of the His-tag on other types of GNATs, we also analyzed the kinetic parameters of two other proteins: a spermidine/spermine acetyltransferase SpeG from *Vibrio cholerae* and an N-terminal protein acetyltransferase SACOL1063 from *Staphylococcus aureus*. Both tagged SpeG and SACOL1063 enzymes exhibit a decreased activity or affinity for AcCoA compared to the untagged enzyme (Figs. 5 and 6). For SpeG, the activity of the untagged enzyme was significantly higher (8-fold) than that of the His-tagged enzyme, and both $V_{\text{max}}$ and $S_{0.5}$ were significantly affected (Fig. 5). On the other hand, when the substrate saturation curves for AcCoA were compared for both
His-tagged and untagged SACOL1063 enzyme, we observed that the presence of the His-tag mainly decreased the enzyme’s affinity for substrate (Fig. 6). Compared to PA4794, where the His-tag primarily affected the enzyme’s affinity for the NPAcGK peptide, SpeG and SACOL1063 showed that the His-tag is also capable of affecting the enzymes’ affinities for AcCoA and/or their maximal velocities. We were not able to obtain crystals of SACOL1063 in the presence of the His-tag. For SpeG, the tag was not visible in the electron density, and although it has a strong influence on enzyme activity, it may still be relatively flexible in the crystal. Our results show that the kinetic parameters for members of the GNAT superfamily can be affected by the presence of the His-tag in a number of ways; therefore, the His-tag should be removed from all GNATs before functional studies.

**Effects of buffers and His-tags are underestimated**

We found that in more than 18% of structures containing HEPES, the buffer molecule is bound in close proximity to the active site of proteins, but often is not mentioned in accompanying publications. Similarly, many structures of proteins with His-tags have density for the affinity tag in the proximity of active sites, but are either not discussed or are omitted altogether from the model, regardless of crystallographic evidence of its presence. For example, in structure 2QHA, despite the fact that there was clear density for the His-tag in the active site, this motif was not only not modeled, but in the accompanying publication it was also claimed that the motif was not visible in the density map. With manual building of the six histidine residues (Fig. 7) and a few cycles of refinement with Refmac, we were able to decrease the $R/R_{free}$ of this structure from 16.0%/19.0% to 12.2%/16.4%, respectively.

Since many proteins bind the tag with relatively high affinity in their active sites or between oligomers, difficulties may arise once tag cleavage is attempted and may lead to further unexpected behaviors during experiments. For example, a putative redox-enzyme maturation protein (REMP) from *Archaeoglobus fulgidus* (AF0173) crystal structure shows that the tag allows the protein to form dimers. In doing so, it causes the TEV protease recognition site to become inaccessible and therefore unable to be cleaved proteolytically. The presence and position of the tag has also been shown to affect protein stability and can change the substrate specificity of enzymes. For example, the stability of the Msed_1072 carboxylesterase from *Metallosphaera sedula* varies depending upon tag placement and presence or absence of detergents or organic solvents. Additionally, the *Escherichia coli* thioesterase I enzyme substrate specificity shifted its preference towards more hydrophilic substrates when a C-terminal His-tag was added. At times, the presence of a His-tag may be deleterious to protein activity, as was shown for *Rhodobacter sphaeroides* YedY reductase. Many buffers used in biological research have a strong influence in thermodynamics studies due to ionization reactions.

**Discussion**

The results demonstrated that both the polyhistidine affinity tag and the buffer HEPES bind to the active site of PA4794, and the binding of HEPES induced structural changes not observed in the structure of apo-protein. However, the His-tag also inhibited the kinetic activity of the enzyme. It is reasonable to conclude from the structural and kinetic data that the His-tag binds in the substrate site and acts as a weak competitive inhibitor for the peptide. Competitive inhibition of AcCoA by the His-tag was also observed for two other GNAT enzymes, SpeG and SACOL1063, indicating that the effects of the His-tag may be widespread for members of the GNAT superfamily. Most certainly, this will also be true for many other classes of proteins, especially enzymes. Moreover, PA4794 is a very promiscuous enzyme. It is becoming clear that the number of promiscuous enzymes is higher than previously acknowledged, and it is tempting to speculate that the capability of promiscuous enzymes to interact with a range of molecules might make these enzymes more prone to unwanted tag or buffer influence.
Often the presence of an affinity tag can actually aid in protein crystallization, especially when the tag is well ordered and interacts with the protein (for example, see structures with PDB codes 1ZKP, 1ZSW, and 2AUA). The presence of the tag has been reported to either improve or deteriorate the resolution of protein structures. The effect might depend on the localization of the tag and the way it interacts with the protein. For example, binding of the tag in the ligand binding site might stabilize the whole protein, while flexible tag binding to the surface of the protein in disparate ways might increase the overall flexibility and thus have a deteriorating effect on the quality and resolution of the crystals and thereby resolution. Because affinity tags are commonplace for protein production, their potential effects on the structure and function of protein are sometimes overlooked. In the case of PA4794, the His-tag does not influence crystallization of the protein, but it does bind to the substrate-binding site of the protein, with an affinity sufficient to act as a weak competitive inhibitor of substrate. Thus, the presence of a His-tag may influence both the activity and binding affinity for ligands because it blocks the substrate binding site in the crystal and competes for the substrate-binding site in solution.

The presence of nonphysiological small molecules in buffers used in purification and crystallization screening methods can yield important clues as to the types of ligands (e.g., substrates, inhibitors, cofactors, etc.) that may bind to the protein, and analysis of residues involved in these interactions can help with understanding protein function. The crystallization screens often contain molecules that represent a larger class of compounds, but might not be the strongest binder in a particular class (for example, the ceftazidime included in the crystallization ligand screen was the first cephalosporin discovered to bind to PA4794, but further experiments revealed that other cephalosporins bind and inhibit more strongly). Since many of these molecules may compete with binding of biologically relevant ligands, unambiguous identification of physiologically relevant ligands may not be obvious. The conformational changes in protein structure may also give clues for determining the catalytic or regulatory mechanisms of a given protein. However, these components can also interfere with physiological mechanisms. Clearly it is critical to select buffer conditions optimized for a particular protein to ensure and verify that the nonphysiological components of the buffer do not influence the results of biochemical studies in artificial ways.

While the potential effects of both affinity tags and sulfonyl buffers, especially HEPES, on biological activity have been known and observed for many years, these effects are rarely discussed in the published literature. The conditions of protein expression and purification can play a significant role in the outcome of protein functional and structure-function studies. In particular, if these nonphysiological compounds are observed in structures that are used for in silico screening, this can negatively affect in silico binding and screening analyses by occluding binding sites and/or inducing artificial conformational changes in the proteins. This may cause a “ripple effect” as subsequent studies are based on questionable data. To properly interpret results of structural and functional studies, it is essential to pay attention to these types of details. Our results have shown the importance of taking these factors into account for members of the GNAT superfamily, but they may also lend a cautionary word of advice for interpreting results obtained for other proteins in different families. In our opinion, recognition, identification, and clarification of artifacts due to buffers or affinity tags should be described and evaluated in order to truly understand a protein’s function and accurately characterize the enzyme of interest.

Materials and Methods

Cloning, expression, and purification

The PA4794 gene was cloned into the p11 pET expression vector (Structural Genomics Consortium), and the SpeG and SACOL1063 genes were cloned into the pMCSG7 expression vector (Midwest Center for Structural Genomics). The proteins that are produced using these vectors contain an N-terminal poly-His-affinity tag followed by a spacer and a tobacco etch virus (TEV) protease cleavage site. The sequence Gly-His (in the case of the p11 pET vector) or Ser-Asn-Ala (in the case of the pMCSG7 vector) remains on the N-terminus of each protein after tag cleavage with recombinant TEV protease. The PA4794 protein used for the 4M3S structure was overexpressed and purified as described. This protein was stored in a buffer containing 500 mM NaCl and 10 mM HEPES at pH 7.5. The PA4794 fusion protein used for the 3KKW structure and the SpeG and SACOL1063 fusion proteins used for the kinetic assays were overexpressed and purified as described. Both His-tagged and cleaved proteins were purified and stored in a buffer containing 500 mM NaCl and 10 mM Tris-HCl at pH 7.5. The PA4794 proteins were concentrated to 9 (PDB ID: 3KKW) and 10 (PDB ID: 4M3S) mg/mL using an Ultracel-10K concentrator (Millipore) and used for crystallization trials.

Enzyme kinetic assays. Assays were performed using previously described procedures. To determine the effect of the polyhistidine tag on enzyme activity, we measured the kinetic parameters of the enzymes toward different substrates in a 50 μL reaction volume. Substrate saturation curves were produced by holding one substrate at a constant concentration while the other one was varied.
kinetic parameters for PA4794 were determined for AcCoA (0–1 mM) at 10 mM NPAcGK or for NPAcGK (0–20 mM) at 1 mM AcCoA in Bicine buffer pH 9.0 with 2.6 μM enzyme (His-tagged or untagged). Kinetic parameters for AcCoA (0–2 mM) for SpeG and SACOL1063 were determined by holding the concentration of spermine at 3 mM (SpeG) or the concentration of L-threonine at 20 mM (SACOL1063). SpeG was assayed in Bicine buffer pH 9.0 and SACOL1063 was assayed in Tris-HCl buffer pH 8.0. The activity of SpeG was determined using 0.6 μM His-tagged enzyme and 0.044 μM untagged enzyme, whereas the activity of SACOL1063 was determined using 0.32 and 0.12 μM His-tagged and untagged enzyme, respectively. All measurements were performed in duplicate and the data shown is the average of the two individual trials. One unit of activity is defined as the amount of enzyme that produces 1 mol of CoA per min under the described reaction conditions. All kinetic curves were fitted using a modified Hill equation in the program Origin v.8.1 as described previously.

Crystallography. Tracking and analysis of the crystallization experiments were performed with the Xtaldb system. The crystals for 4M3S were grown using vapor diffusion and sitting drop setup. The crystallization drops were composed of a 1:1 μL mixture of the protein solution to precipitant solution (100 mM HEPES-Na pH 7.5, 2% PEG400, 2 M (NH₄)₂SO₄). The crystals for 3KKW were grown using a vapor diffusion and hanging drop setup. The crystallization drops were composed of a 1:1 μL mixture of the protein solution to precipitant solution (2 M (NH₄)₂SO₄, 100 mM Bis-Tris pH 6.5). Crystals were grown overnight at 16°C. Prior to data collection, the crystal was transferred to a solution containing a 2:1 mixture of well solution and ethylene glycol and was immediately cryo-cooled in liquid nitrogen.

Data collection, structure determination and refinement. Data collection for PA4794 was performed at the 19-BM beamline of the Structural Biology Center at the Advanced Photon Source (APS, Argonne National Laboratory). Data were collected at 100 K and processed with HKL-2000. Structure determination for 4M3S was performed using single-wavelength anomalous diffraction and HKL-3000, which is integrated with SHELXD/E, MLPHARE, DM, ARP/wARP, CCP4, and SOLVE and RESOLVE. That structure was then used as a model to solve the structure of PA4794 with the His-tag by molecular replacement (MR), using HKL-3000 coupled with MOLREP. Refinement was performed using HKL-3000 coupled with REFMAC5, COOT, and selected programs from the CCP4 package. The B-factors were refined in each case using translation/libration/screw (TLS) groups assigned using the TLSMD server.

Validation of the structures was performed using MOLPROBITY and ADIT. The coordinates, structure factors, and intensities were deposited in the PDB (PDB IDs: 4M3S, 3KKW). Statistics describing crystallographic data collection and refinement are summarized in Table I.

Acknowledgments
We would like to thank Matthew D. Zimmerman, Alex Wlodawer, Thomas Terwilliger, and Zbigniew Dauter for numerous discussion and suggestions.

Table I. Crystallographic Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th>Complex</th>
<th>His-tag</th>
<th>HEPES</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB code</td>
<td>3KKW</td>
<td>4M3S</td>
</tr>
<tr>
<td>Data collection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2₁2₁2₁</td>
<td>P2₁2₁2₁</td>
</tr>
<tr>
<td>Unit cell</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>57.53, 76.14, 39.45</td>
<td>56.95, 75.51, 39.48</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.0, 90.0, 90.0</td>
<td>90.0, 90.0, 90.0</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>1.41 (1.41–1.43)</td>
<td>1.30 (1.30–1.32)</td>
</tr>
<tr>
<td>No. reflections</td>
<td>32398</td>
<td>42543</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.047 (0.391)</td>
<td>0.103 (0.779)</td>
</tr>
<tr>
<td>Df(Δ)</td>
<td>4.72 (4.2)</td>
<td>37.2 (2.1)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.9 (100.0)</td>
<td>99.9 (100.0)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>6.5 (5.7)</td>
<td>8.8 (8.1)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rwork (%)</td>
<td>15.7/17.9</td>
<td>13.5/16.9</td>
</tr>
<tr>
<td>Rfree (%)</td>
<td>15.7/17.9</td>
<td>13.5/16.9</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>1299</td>
<td>1245</td>
</tr>
<tr>
<td>Ligand/Ion</td>
<td>78</td>
<td>29</td>
</tr>
<tr>
<td>Water</td>
<td>210</td>
<td>318</td>
</tr>
<tr>
<td>B factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wilson B factor</td>
<td>14.6</td>
<td>12.9</td>
</tr>
<tr>
<td>Protein</td>
<td>21</td>
<td>15</td>
</tr>
<tr>
<td>Ligand/Ion</td>
<td>29</td>
<td>19</td>
</tr>
<tr>
<td>Water</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td>Structure quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bond length (Å)</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>bond angles (°)</td>
<td>1.8</td>
<td>1.7</td>
</tr>
<tr>
<td>Ramachandran Statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favored (%)</td>
<td>99.38</td>
<td>98.73</td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>0.62</td>
<td>1.27</td>
</tr>
</tbody>
</table>

a Highest resolution shell values are shown in parentheses. b About 5% of reflections were randomly assigned to the Rfree set.


