Crystal Structure of the FHA Domain of the Chfr Mitotic Checkpoint Protein and Its Complex with Tungstate

Elena S. Stavridi,1 Yentram Huyen,1,2 Ivy R. Loreto,1
Daniel M. Scolnick,1,3 Thanos D. Halazonetis,1,4,7
Nikola P. Pavletich,5,6,7 and Philip D. Jeffrey9
1Molecular Genetics Program
Structural Biology Program
Wistar Institute
Philadelphia, Pennsylvania 19104
2Cell and Molecular Biology Program
3Biochemistry and Biophysics Program
Biomedical Graduate Studies
4Department of Pathology and Laboratory Medicine
University of Pennsylvania
Philadelphia, Pennsylvania 19104
5Cellular Biochemistry and Biophysics Program
6Howard Hughes Medical Institute
Memorial Sloan-Kettering Cancer Center
New York, New York 10021

Summary

The Chfr mitotic checkpoint protein is frequently inactivated in human cancer. We determined the three-dimensional structure of its FHA domain in its native form and in complex with tungstate, an analog of phosphate. The structures revealed a β sandwich fold similar to the previously determined folds of the Rad53 N- and C-terminal FHA domains, except that the Rad53 domains were monomeric, whereas the Chfr FHA domain crystallized as a segment-swapped dimer. The ability of the Chfr FHA domain to recognize tungstate suggests that it shares the ability with other FHA domains to bind phosphoproteins. Nevertheless, differences in the sequence and structure of the Chfr and Rad53 FHA domains suggest that FHA domains can be divided into families with distinct binding properties.

Introduction

The forkhead-associated (FHA) domain was originally described as a small domain in transcription factors with forkhead-type DNA binding domains [1]. More recently, FHA domains have also been identified in protein kinases and other proteins that are not transcription factors such that most of the proteins now known to contain FHA domains function in cell cycle control [2]. Biochemical and structural studies suggest that FHA domains are protein-protein interaction domains with specificity for phosphorylated targets [3–7]. The structural studies have focused on the two FHA domains present within Rad53, a budding yeast DNA damage checkpoint kinase. Three-dimensional structures of the N-terminal FHA domain (N-FHA) bound to a phosphothreonine-containing peptide and of the C-terminal FHA domain (C-FHA) bound to a phosphotyrosine-containing peptide have been determined by crystallography and NMR, respectively, and show overall similar folds for the two FHA domains. However, there are differences in the interactions with the phosphopeptides, and several highly conserved FHA domain residues are not involved in phosphopeptide binding in either structure [4–7].

We recently reported the cloning of chfr, a mitotic checkpoint gene that functions in early mitosis to delay progression through prophase in response to mitotic stress [8]. chfr is frequently inactivated in human cancer cell lines, and its inactivation confers sensitivity to microtubule poisons, such as taxol, which is being used for cancer therapy. The protein encoded by chfr contains an FHA domain at its N terminus. By analogy to Rad53, whose two FHA domains have been shown to bind phosphopeptides, Chfr may utilize its FHA domain to bind to phosphorylated proteins. However, neither physiological substrates of Chfr nor peptides that can bind to its FHA domain in vitro are known. To better understand the function of the FHA domain of Chfr and to explore its ability to recognize phosphoproteins, we determined its three-dimensional structure in its native form and in complex with tungstate, an analog of phosphate. The structures revealed a β sandwich fold similar to the previously determined folds of the Rad53 N- and C-terminal FHA domains. However, the Rad53 domains were monomeric, whereas the Chfr FHA domain crystallized as a segment-swapped dimer. The tungstate ion was bound by a cluster of conserved FHA domain residues; this is similar to the way that phosphate is recognized by the Rad53 N-terminal FHA domain residues; this is similar to the way that phosphate is recognized by the Rad53 C-terminal FHA domain. Thus, the Chfr FHA domain shares the ability with other FHA domains to bind phosphoproteins, although differences in the mechanism of tungstate/phosphate binding suggest that FHA domains can be divided into families with distinct binding properties.

Results and Discussion

Overall Structure

Several fragments of Chfr containing the FHA domain were expressed in E. coli and characterized for solubility and resistance to proteolysis in the presence of subtilisin. The smallest soluble and proteolysis-resistant polypeptides identified by this analysis consisted of an N-terminal methionine and residues 14–124 or 14–128 of Chfr. Both polypeptides eluted as single peaks on gel filtration chromatography, consistent with a monomer form. A selenomethionine derivative of the polypeptide containing residues 14–128 formed crystals with P3(2)121 symmetry that diffracted to 2.7 Å resolution and had one molecule in the asymmetric unit. The structure of this polypeptide was determined using anomalous scattering from the selenium and the isomorphous difference from a selenomethionyl crystal soaked in mercury (Table 1). The polypeptide containing residues 14–124 formed...
<table>
<thead>
<tr>
<th>Molecule</th>
<th>Chfr</th>
<th>Chfr</th>
<th>Chfr-Tungstate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P3(2)21</td>
<td>P2(1)2(1)2(1)</td>
<td>P3(2)21</td>
</tr>
<tr>
<td>SeMet MAD Inflection</td>
<td>0.979</td>
<td>0.979</td>
<td>0.909</td>
</tr>
<tr>
<td>SeMet MAD Peak</td>
<td>0.979</td>
<td>0.979</td>
<td>2.0</td>
</tr>
<tr>
<td>SeMet MAD Remote</td>
<td>0.964</td>
<td>2.7</td>
<td>3.1</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.979</td>
<td>0.979</td>
<td>0.964</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>17.346</td>
<td>17.281</td>
<td>17.359</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>3,642</td>
<td>3,622</td>
<td>3,643</td>
</tr>
<tr>
<td>Data coverage (%)</td>
<td>97.2</td>
<td>97.1</td>
<td>96.9</td>
</tr>
<tr>
<td><strong>Phasing Statistics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>—</td>
<td>18–4.0</td>
<td>18–4.0</td>
</tr>
<tr>
<td>Phasing power (centric)</td>
<td>—</td>
<td>0.46</td>
<td>0.68</td>
</tr>
<tr>
<td>Phasing power (acentric)</td>
<td>—</td>
<td>0.78</td>
<td>1.22</td>
</tr>
<tr>
<td>Rcullis (centric)</td>
<td>—</td>
<td>0.87</td>
<td>0.79</td>
</tr>
<tr>
<td>Rcullis (acentric)</td>
<td>—</td>
<td>0.90</td>
<td>0.78</td>
</tr>
<tr>
<td>Rcullis (anomalous)</td>
<td>—</td>
<td>0.62</td>
<td>0.73</td>
</tr>
<tr>
<td><strong>Refinement Statistics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>18–2.7</td>
<td>18–2.1</td>
<td>18–2.0</td>
</tr>
<tr>
<td>Reflections used</td>
<td>3,448</td>
<td>11,878</td>
<td>6,719</td>
</tr>
<tr>
<td>Protein atoms</td>
<td>904</td>
<td>1,790</td>
<td>904</td>
</tr>
<tr>
<td>Water molecules</td>
<td>25</td>
<td>144</td>
<td>63</td>
</tr>
<tr>
<td>R factor</td>
<td>0.247</td>
<td>0.242</td>
<td>0.232</td>
</tr>
<tr>
<td>R free</td>
<td>0.294</td>
<td>0.288</td>
<td>0.289</td>
</tr>
<tr>
<td>Rms deviations</td>
<td>0.008</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>Bonds (%)</td>
<td>81.4</td>
<td>90.1</td>
<td>84.5</td>
</tr>
<tr>
<td>Angles (%)</td>
<td>1.534</td>
<td>1.298</td>
<td>1.463</td>
</tr>
<tr>
<td>Ramachandran plot Most favored regions (%)</td>
<td>18.6</td>
<td>9.9</td>
<td>15.5</td>
</tr>
</tbody>
</table>

Table 1. Data Collection and Refinement Statistics

Crystals with P2(1)2(1)2(1) symmetry that diffracted to 2.1 Å resolution and had two molecules in the asymmetric unit. The structure of this polypeptide was solved by molecular replacement (Table 1). Both structures were virtually identical and, unless otherwise specified, the higher resolution structure will be described.

The FHA domain of human Chfr crystallized as a dimer (Figure 1A). Each subunit consists of ten β strands and a very short α helix between the two C-terminal strands. The dimer formed by segment swapping (a.k.a. domain swapping) [9, 10], whereby each folded domain consists of two segments, each derived from a different molecule. The first segment contributes the N-terminal six strands and the N-terminal half of the seventh strand of the domain; the second segment contributes the C-terminal half of the seventh strand, the C-terminal three strands, and the α helix. The overall fold of the domain is similar to that previously reported for the Rad53 N-FHA and C-FHA domains. The seventh strand of Chfr corresponds to Rad53 strands 7 and 8 and the loop connecting these two strands (Figure 1B). Thus, to have a consistent numbering scheme when comparing monomeric and segment-swapped FHA structures, we will refer to the seventh strand of Chfr as strand 7/8.

Structure of Conserved Regions

FHA domains have four conserved regions that are referred to as boxes A, B, C, and D (Figure 2A) [2]. Box A consists of a hydrophobic residue followed by a glycine, an arginine, and a polar residue. The center two residues (Gly41 and Arg42 in Chfr) are conserved in all FHA domains and map to the C terminus of strand 3 and the loop that connects strands 3 and 4, respectively (Figure 2B). Box B contains serine and histidine residues (Ser57 and His60 in Chfr) that are conserved in all FHA domains. Ser57 is at the loop that connects strands 4 and 5, and His60 is at the N terminus of strand 5. The hydroxyl oxygen of Ser57 is within 4.0 Å of the Ni of His60, as the peptide backbone in this region adopts a kinked conformation that allows the side chains of these two residues to be close to each other. Box C represents the longest stretch of conserved residues in FHA domains and consists of highly conserved aspartic acid, serine, glycine, threonine, and asparagine residues...
Figure 1. Structure of the Chfr FHA Domain
(A) Structure of the segment-swapped dimer. The two polypeptide chains are colored red and green, respectively. N, N terminus; C, C terminus.
(B) Comparison of the structures of the Chfr FHA and Rad53 N-terminal FHA domains. Only one domain of Chfr is shown, and the strands involved in segment swapping are not shown in their entirety. In both structures, strands 1–7 and 8–11 are colored green and red, respectively, to maintain the same coloring scheme as in (A).

(Asp76, Ser78, Gly81, Thr82, and Asn85 in Chfr). Asp76 at the C-terminal end of strand 6, Ser78 at the loop connecting strands 8 and 7, and Thr82 at the N terminus of strand 7, together with box B conserved residues Ser57 and His60, form a cluster of polar/charged residues that interact with each other through hydrogen bonds. Specifically, the side chain of Asp76, which is completely buried, forms hydrogen bonds with Ser78 and Thr82, whose side chains are also buried. Thr82 also forms a hydrogen bond with His60. In addition, the hydroxyl oxygens of Ser78 and Ser57 and the Nε of His60 are within 3.6–4.3 Å of each other. The C terminus of box C corresponds to the middle of strand 7/8, where the polypeptide chain traverses from one folded domain to another. This region adopts a structure referred to as a polar zipper, an antiparallel β sheet, which, in addition to the main chain hydrogen bonds, is stabilized by hydrogen bonds formed by asparagine or glutamine side chains [11]. In Chfr, the side chain hydrogen bonds are formed by the highly conserved Asn85, which interacts with Asn85 from the other subunit (Figure 2B). A very similar structure is adopted by the linker region of segment-swapped dimers of RNase A and possibly by other protein oligomers [12–14]. Finally, box D contains a glycine followed by an aspartic acid (Gly100 and Asp101 in Chfr). In Chfr, Asp101 is at the N terminus of strand 10 and forms a salt bridge with Lys86 from the other subunit (Figure 2B).

Structure of the Chfr-Tungstate Complex
To explore the ability of the Chfr FHA domain to recognize phosphate, we formed crystals of the polypeptide containing residues 14–128 of Chfr in the presence of tungstate. Tungstate has similar molecular geometry and charge as phosphate and has been used extensively as a phosphate analog in crystallization studies [15–17]. The Chfr-tungstate complex formed crystals with P3(2)21 symmetry that diffracted to 2.0 Å resolution. The structure factor amplitudes derived from such a crystal and the previously refined model (without tungstate) in the P3(2)21 space group were used to calculate a model-phased \( F_o - F_c \) difference map. This map showed a single peak with intensity greater than 4.0 \( r \), indicating the presence of a tungstate ion in the crystal. The structure of the Chfr-tungstate complex was refined and showed interactions of the tungstate with conserved...
Figure 2. Structure of Conserved Regions

(A) Amino acid sequence and secondary structure of the crystallized Chfr FHA domain (residues 14–124) aligned to conserved sequences (boxes A–D) of FHA domains from other proteins. The amino acid sequence numbering refers to the full-length proteins and indicates the position of the conserved boxes or, for Chfr, marks the entire sequence and selected residues. Rad53-N and Rad53-C, N- and C-terminal FHA domains of budding yeast Rad53; Dun1, FHA domain of budding yeast Dun1 checkpoint kinase; NBS1, FHA domain of human DNA repair and cell cycle checkpoint protein NBS1; MNF, FHA domain of mouse MNF transcription factor; Fkh1, FHA domain of budding yeast Fkh1 transcription factor.

(B–C) Comparison of the structures of Chfr FHA and Rad53 N-FHA domains showing the side chains of amino acids conserved in FHA domains. The Rad53 N-FHA domain is bound to a phosphothreonine (pT)-containing peptide. Backbone of the FHA domains corresponding to boxes A–D, orange; phosphopeptide, light blue; phosphate group, red. D/H11001, aspartic acid three residues C-terminal to the phosphothreonine.

residues mostly in boxes A and B (Figure 3B). Specifically, the backbone nitrogen of Gly58 (box B) and the hydroxyl oxygen of Thr79 (box C) are within 2.6–2.9 Å of a tungstate oxygen atom; in addition, the hydroxyl oxygen of Ser57 (box B) and the N$_1$ of Arg42 and Arg43 (box A) are within 3.0–3.2 Å of a tungstate oxygen atom (Figures 3B and 3F). Comparison of the native Chfr domain to the Chfr-tungstate complex reveals that tungstate binding has very minor effects on the conformation of Chfr limited to the side chains of Arg42, Arg43, and Thr79 (Figures 3A and 3B).

Comparison to Rad53 FHA Domains

To identify structural features that are conserved in FHA domains, we compared the structures of the Chfr FHA domain to the structures of the Rad53 N-terminal and C-terminal FHA domains, which have been determined in complex with phosphopeptide ligands [4–7]. We focused our comparison on boxes A, B, C, and D, which show the highest sequence similarity.

Boxes A and B are involved in phosphate/tungstate binding in all three FHA domains, but there are differences in the way the ligand is recognized; the Chfr and Rad53 N-terminal FHA domains turn out to be closely related, whereas the Rad53 C-terminal FHA domain is more divergent (Figure 3). The Chfr-tungstate contacts that are conserved in the Rad53 N-terminal FHA-phosphopeptide structure involve the side chains of Arg42 of box A (Arg70 in Rad53), Ser57 of box B (Ser85 in Rad53), and Thr79 of box C (Thr106 in Rad53) and the backbone nitrogen of Gly58 of box B (Asn86 in Rad53). Chfr makes an additional contact with tungstate through the side chain of Arg43 of box A (Asn71 in Rad53), whereas the Rad53 N-FHA domain makes an additional contact to phosphate through the side chain of Asn86 of box B (Gly58 in Chfr). The C-terminal Rad53 FHA domain contacts phosphate through the side chains of Arg605 of box A (Arg42 in Chfr) and Arg617 and Arg620 of box B (Leu55 and Gly58 in Chfr). Of these three interactions, only the first is conserved in Chfr. The third...
interaction involving Arg620 of box B, although not conserved in Chfr, shares some similarity to the interaction mediated by Arg43 of box A of Chfr (Figures 3B, 3D, and 3E). Consistent with the high number of conserved protein-ligand interactions between the Chfr and Rad53 N-terminal FHA domains, the tungstate and phosphate ligands of the domains superimposed quite well (within 0.7 Å of each other), whereas the phosphate ligand of the Rad53 C-terminal FHA domain was about 5 Å away (Figure 3E).

Despite the similarities in tungstate/phosphate recognition between the N-terminal Rad53 and Chfr FHA domains, the two domains are likely to recognize different peptide sequences. The Rad53 N-FHA domain has a preference for aspartic acid three amino acids C-terminal to the phosphothreonine [4, 7], a preference mediated by an electrostatic interaction between the aspartic acid of the peptide and Arg83 of box B (Figure 3C). The corresponding residue in Chfr, Leu55, would be unable to make a similar interaction.

Box C represents the longest stretch of conserved residues in FHA domains, yet these residues, with the exception of Thr79 (Thr106 in Rad53), do not directly contact tungstate/phosphate. Most of the conserved box C residues form a cluster of amino acids with buried charged and polar side chains (Asp76, Ser78, and Thr82 in Chfr) at the interface of strands 6 and 7 and in close proximity to conserved box B residues (Ser57 and His60 in Chfr). The entire cluster is remarkably similar in the Chfr and Rad53 N-FHA domains (Figures 2B and 2C), even though these buried residues cannot directly bind ligand. The high degree of similarity prompted us to compare the amino acid sequences of multiple FHA domains. Based on the sequences corresponding to boxes C and D [1, 2], we propose that FHA domains can be divided into families. One family includes FHA domains present in cell cycle checkpoint proteins, such as Chfr and Rad53. Another family includes FHA domains present in transcription factors with forkhead-DNA binding domains, such as MNF and Fkh1. The first family is characterized by the presence of charged and polar residues at the ends of the FHA domain. The corresponding residues are mostly hydrophobic (Figure 2A). Because of their close proximity to ligand binding residues, the buried box C residues at the interface of strands 6 and 7 may modulate the binding properties of the FHA domain.

The observation that the FHA domain of Chfr crystallized as a segment-swapped dimer was unanticipated from the structures of the N-FHA and C-FHA domains of Rad53. We therefore sought to obtain additional crystal-
lographic and noncrystallographic evidence for segment swapping in Chfr FHA domain crystals. First, we explored whether it was possible to refine the structure using a chain trace that conferred the same topology as monomeric Rad53 instead of the chain trace that connected two adjacent FHA domains to make a segment-swapped dimer. The differences in the two traces involved very few residues (Asn85, Lys86, and Leu87). However, refinement of the trace with the Rad53 topology led to several problems, including steric hindrance involving the Cα, Cβ, Hα, and Hβ atoms of Lys86 across crystallographic symmetry-related monomers (Figure 4A), disallowed phi/psi angles for residues Asn85 and Lys86, negative Fo – Fc density for the bond we arbitrarily introduced between Asn85 and Lys86 to create the monomer model, and higher R and Rfree values. As a second way to validate the segment-swapped model, we calculated simulated annealing omit maps for both P3(2)21 and P2(1)2(1)2(1) space groups using models in which all atoms within 3 Å of Lys86 were omitted. These maps clearly showed density consistent with a segment-swapped dimer (Figure 4B).

To obtain noncrystallographic evidence for segment swapping, we studied the oligomerization state of the Chfr FHA domain using biophysical approaches. The first step in this analysis was to determine whether, after purification, the Chfr FHA domain was a monomer or dimer. Freshly purified polypeptide containing residues 14–128 was examined by analytical ultracentrifugation and was shown to be a monomer that could, however, form dimers with a dissociation constant of 0.15 mM for the monomer-dimer equilibrium (Figure 4C). Unfortunately, the analytical ultracentrifugation analysis does not identify whether the dimers are formed by segment swapping or by simple juxtaposition of two independently folded domains. Nevertheless, these results indicate that, during purification, when the protein concentration would be below 0.15 mM, the Chfr FHA domain would be mostly a monomer, whereas, after concentration, in preparation for crystallization, the domain would be mostly a dimer.

Analytical ultracentrifugation was not an option for analysis of Chfr FHA domain protein extracted from crystals because of its low sensitivity; we therefore resorted to gel filtration analysis using a protein purification system designed for protein quantities in the micromolar range. About 20 Chfr FHA domain crystals were washed extensively in stabilization buffer and dissolved by dilution in buffer that contained no PEG precipitant. After spinning to remove particulate material, the solubilized protein was run on the gel filtration column. As a control, we examined the entire contents of a drop that had been set up for crystallization and which had a considerable amount of amorphous protein precipitate but no crystals. The amorphous precipitate in the drop was dissolved by dilution in buffer without PEG, and, therefore, this sample was very similar to the sample of solubilized FHA crystals in buffer and protein composition. A further control included purified protein that had not been set up for crystallization. During the gel filtration run, the protein concentration in all three samples was at least ten times below the value of the monomer-dimer dissociation constant calculated by analytical ultracentrifugation. The protein that had not been set up for crystallization and the protein that had formed amorphous precipitate showed identical elution profiles, corresponding to monomers. In contrast, the protein derived from the crystals eluted more slowly, corresponding to the size of a dimer (Figure 4D). This result is consistent with the Chfr FHA domain in the crystals existing as segment-swapped dimers that did not dissociate to monomers during the 30 min time frame of the gel filtration chromatography run. Dissociation of segment-swapped dimers requires at least partial unfolding of the domain and would be predicted to be slow, based on the kinetics of protein unfolding [20]. Indeed, the conversion of the cell cycle yeast protein suc1 from its monomeric to segment-swapped dimeric form is very slow, due to a kinetic barrier imposed by the denatured state intermediate [21].

Biological Implications

The ability of the Chfr FHA domain to bind tungstate suggests that, in vivo, Chfr interacts through its FHA domain with phosphorylated proteins. The identity of these proteins remains elusive. However, since Chfr is a mitotic checkpoint protein, its ligands might be proteins phosphorylated by kinases, such as Bub1, BubR1, and Mps1, that are activated in response to mitotic stress [22–24]. At the atomic level, it is interesting that the Chfr and the N-terminal Rad53 FHA domains recognize tungstate/phosphate mostly through the same conserved residues, whereas the structure of the C-terminal Rad53 FHA domain bound to a phosphotyrosine-containing peptide shows that this domain recognizes phosphate through residues that are mostly not conserved in the former two FHA domains. The distinct architecture may reflect intrinsic differences in the mechanism by which distinct FHA domains recognize phosphate or may be due to the different nature of the ligands used. In relation to the former, a cluster of buried polar and charged residues, adjacent to the residues that directly contact ligand, is conserved in FHA domains present in cell cycle checkpoint proteins but not in FHA domains present in transcription factors. Based on this difference, we propose that FHA domains can be divided into families that will likely have distinct binding properties.

A surprising feature of the Chfr FHA domain structure was the presence of segment swapping involving an exchange of four C-terminal β strands and an α helix between molecules. This phenomenon was first observed in crystals of the diphtheria toxin, and the term domain swapping was introduced to describe it because an entire structural domain of the multidomain diphtheria toxin protein is exchanged [25]. However, in Chfr, as in most examples described since the term domain swapping was introduced, only part of a protein domain is exchanged [9, 10]. We therefore adopted the term segment swapping to describe examples, such as that of Chfr, in which multiple secondary structure elements, but not entire structural domains, are exchanged between subunits.

There are several examples of segment-swapped structures, but they represent a very small fraction of all determined protein structures [9, 10]. The physiological
Figure 4. Segment Swapping
(A) Refined monomer and segment-swapped dimer models in the P3(2)21 space group. The hydrogens of the Lys86 side chain are shown to indicate the presence of steric hindrance in the monomer model. The two polypeptide chains are colored green and red, respectively.
(B) Simulated annealing omit electron density maps contoured at 1.5 \(\sigma\). The maps were generated using the refined monomer model in the P3(2)21 space group (left) or the refined segment-swapped dimer model in the P2(1)2(1)2(1) space group (right), with all atoms within 3 \(\AA\) of Lys86 omitted. For both maps, the refined segment-swapped dimer models are shown to indicate that these models fit theomit electron density maps.
(C) Analytical ultracentrifugation analysis of the polypeptide containing residues 14–128 of Chfr. Lower panel: protein concentrations at equilibrium for the three initial loading concentrations (1.0, 0.5, and 0.25 mg/ml) at 33,800 rpm and 4°C. Circles, raw data; lines, fitted curves for monomer-dimer equilibrium. Upper panels: residuals of the fitted curves to the data points for each concentration from highest to lowest (top to bottom, respectively).
(D) Gel filtration chromatography elution profiles of Chfr FHA domain immediately after purification (Stock) or after extraction from amorphous precipitate (Precip.) or crystals (Crystal). The Stock, Precip., and Crystal elution profiles correspond to molecular sizes of 11.6, 11.5, and 20.6 kDa, respectively.

The significance of segment swapping is unclear, although evidence is accumulating that it is relevant in vivo. For example, amyloid formation in patients with amyloid angiopathy and transmissible encephalopathies may be mediated by segment swapping and subsequent oligomerization of cystatin C and the cellular prion protein PrP(C), respectively [26, 27]. In addition, segment swapping may be required for Flp recombinase to be active and may regulate the enzymatic activity of nitric-oxide synthase and the ligand binding properties of suc1 [28–30]. It is too premature to state whether segment swapping in the Chfr FHA domain structure is relevant in vivo. As yet, there is no evidence that Chfr or other FHA domain-containing proteins form dimers in vivo. Nevertheless, certain features important for the stabilization of the segment-swapped dimer, such as the asparagine at the end of strand 7 (Asn85 in Chfr) and the length of the loop between strands 7 and 8, are conserved in most FHA
domains, raising the possibility of functional significance.

In conclusion, the Chfr FHA domain structure reveals the capacity for phosphate binding, a structurally conserved cluster of buried polar and charged residues next to the phosphate binding site, and segment swapping. Amino acid substitution mutants, whose designs are based on the three-dimensional structure, can now be used to probe the functional significance of these structural features and better understand a protein that is mutated in human cancer.

Experimental Procedures

Protein Expression and Purification

The human Chfr FHA domain (residues 14–124 or 14–128) was expressed in E. coli. Buffer exchange to a final concentration of 25 mM bis-Tris propane (pH 6.8). Proteins were purified by cation exchange (Resource S column; Pharmacia) and gel filtration (Superdex 200 column; Pharmacia) chromatography. The protein eluted from the gel filtration column in buffer consisting of 20 mM bis-Tris (pH 6.5), 150 mM KCl, and 5 mM DTT.

Crystallization and Data Collection

Crystals were grown at 4°C by the hanging drop vapor diffusion method by mixing the protein with an equal volume of reservoir solution containing 14%–18% polyethylene glycol (PEG) 8000, 100 mM bis-Tris (pH 6.5), and 5 mM DTT. Crystallization started at 2–3 days, and crystals grew to a final size over a 2–4 week period. Crystals of the native polypeptide containing residues 14–124 formed in space group P2₁(1)2₁2₁, with a = 54.90 Å, b = 52.90 Å, and c = 77.25 Å, and contained two molecules in the asymmetric unit. Crystals of selenomethionine (SeMet)-labeled polypeptide containing residues 14–128 formed in space group P3(2)2₁, with a = 62.36 Å and c = 54.25 Å, and contained one molecule in the asymmetric unit. Some of the SeMet-labeled crystals were soaked in harvest buffer (21% PEG 8000, 100 mM bis-Tris (pH 6.5), and 75 mM KCl) supplemented with 0.5 mM mercury chloride and 1 mM β-mercaptoethanol for 1 hr. Finally, crystals of the nonlabeled polypeptide containing residues 14–128 were grown in crystallization buffer containing 4 mM sodium tungstate and formed in space group P3(2)2₁, with a = b = 62.21 Å and c = 54.11 Å, and contained one molecule and a symmetric unit.

All data sets were collected using flash-frozen crystals. The native P2₁(1)2₁2₁ data set was collected at the F1 beamline of the Cornell High Energy Synchrotron Source (MACCHESS); the selenomethionine and selenomethionine-mercury chloride derivative data sets were collected at the X9A beamline of the Brookhaven National Laboratory (BNL) Synchrotron Source; the Chfr-tungstate complex data set was collected using a Rigaku R-AXIS IV imaging plate detector mounted on a Rigaku rotating anode X-ray generator. Reflection data were indexed, integrated, and scaled using programs DENOVO and SCALPACK [31].

Structure Determination and Refinement

Phases for the P3(2)2₁ space group were calculated at 3.1 Å using the anomalous signal of the Hg-SeMet derivative and the anomalous MAD data collected from the SeMet crystals using the program MLPHARE [32]. The inflection point data of the SeMet crystal was used as the native for the phasing calculations. The anomalous signal was weak (1 Se in 112 residues) and only useful for phasing at lower resolution. The phases had a mean figure of merit of 0.51 at 3.1 Å resolution (Table 1). After solvent flattening using DM [32], the experimental electron density map showed clear density for nearly all of the backbone and most of the side chains. The model was built using the program O [33] and refined with the program CNS [34]. The structure in the P2₁(1)2₁2₁ space group was solved by molecular replacement using the P3(2)2₁ space group structure and the program AMORE [32] and was refined using the programs O and CNS. The presence of the tungstate ion in the Chfr-tungstate complex structure was determined by refining the P3(2)2₁ space group structure against the structure factor amplitudes obtained from the Chfr-tungstate complex crystal and calculating a model-phased Fo – Fc difference map, which showed the position of the tungstate ion. The Chfr-tungstate complex structure was refined using the programs O and CNS; after multiple cycles of refinement, the occupancy of the tungstate ion was calculated to be 0.47. Figures were prepared using the programs MOLSCRIPT [35], RASTER3D [36], and BOBSCRIPT [37].

Analytical Ultracentrifugation

The oligomerization state of the Chfr FHA domain was analyzed using sedimentation equilibrium in a Beckman XL-1 analytical ultracentrifuge. The polypeptide containing residues 14–128 of Chfr was tested at protein concentrations of 0.25, 0.5, and 1.0 mg/ml in buffer containing 20 mM bis-Tris (pH 6.5), 150 mM KCl, and 5 mM DTT at centrifugation speeds of 33,800 and 48,900 rpm at a temperature of 4°C. Attainment of equilibrium was determined by comparing successive scans acquired using interference optics. The scans acquired at 28 hr were used to calculate the monomer-dimer association constants as previously described [38].

Gel Filtration Analysis for Domain Swapping

Fifty microliters of stabilization buffer without PEG (100 mM bis-Tris [pH 6.5], 150 mM KCl, and 5 mM DTT) were added either to Chfr FHA protein that had not been set up for crystallization, to one drop of protein that had been set up for crystallization and which contained amorphous precipitate but no crystals, or to about 20 harvested crystals that had been extensively washed in stabilization buffer. The samples were spun to remove any particulate material and run on a micropurification gel filtration column (Sephadex 200 PC; column volume, 2.4 ml) on a SMART protein purification system (Pharmacia) at a flow rate of 80 μl/min. Each chromatography run lasted about 30 min. To assign subunit stoichiometry, we compared the elution profiles of the samples to the elution profiles of molecular size standards.

Acknowledgments

The authors thank Steven Smerdon for providing the coordinates of the Rad53 N-FHA domain before publication, the Wistar Institute Protein Chemistry and Nucleic Acid Facilities for protein N-terminal sequencing, mass spectrometry analysis, and DNA sequencing analysis, the staff of the Brookhaven National Laboratory and Cornell High Energy Synchrotron Source for help with data collection, Sandy Harper for help with the analytical ultracentrifugation analysis, and David Speicher, Ronen Marmorstein, and Roger Burnett for support and helpful discussions. This research was funded by grant CA96930 to the National Cancer Institute. D.M.S. was supported by the National Cancer Institute Training Grant CA09171 awarded to the Wistar Institute.

Received: January 25, 2002
Revised: April 17, 2002
Accepted: April 18, 2002

References

5. Liao, H., Byeon, I.J., and Tsai, M.D. (1999). Structure and func-
von a new phosphopeptide-binding domain containing the FHA2 of Rad53. J. Mol. Biol. 294, 1041–1049.

Accession Numbers

The atomic coordinates and structure factors have been deposited with the Protein Data Bank with accession codes 1lgq for the native structure and 1lgp for the tungstate-bound structure.