View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

MAD analysis of FHIT, a putative human tumor suppressor from the HIT protein family

Christopher D Lima[1], Kevin L D'Amico[2], Istvan Naday[3], Gerold Rosenbaum[2], Edwin M Westbrook[2] and Wayne A Hendrickson[1,4]*

Background: The fragile histidine triad (FHIT) protein is a member of the large and ubiquitous histidine triad (HIT) family of proteins. It is expressed from a gene located at a fragile site on human chromosome 3, which is commonly disrupted in association with certain cancers. On the basis of the genetic evidence, it has been postulated that the FHIT protein may function as a tumor suppressor, implying a role for the FHIT protein in carcinogenesis. The FHIT protein has dinucleoside polyphosphate hydrolase activity *in vitro*, thus suggesting that its role *in vivo* may involve the hydrolysis of a phosphoanhydride bond. The structural analysis of FHIT will identify critical residues involved in substrate binding and catalysis, and will provide insights into the *in vivo* function of HIT proteins.

Results: The three-dimensional crystal structures of free and nucleoside complexed FHIT have been determined from multiwavelength anomalous diffraction (MAD) data, and they represent some of the first successful structures to be measured with undulator radiation at the Advanced Photon Source. The structures of FHIT reveal that this protein exists as an intimate homodimer, which is based on a core structure observed previously in another human HIT homolog, protein kinase C interacting protein (PKCI), but has distinctive elaborations at both the N and C termini. Conserved residues within the HIT family, which are involved in the interactions of the proteins with nucleoside and phosphate groups, appear to be relevant for the catalytic activity of this protein.

Conclusions: The structure of FHIT, a divergent HIT protein family member, in complex with a nucleotide analog suggests a metal-independent catalytic mechanism for the HIT family of proteins. A structural comparison of FHIT with PKCI and galactose-1-phosphate uridylyltransferase (GalT) reveals additional implications for the structural and functional evolution of the ubiquitous HIT family of proteins.

Addresses: [1]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA, [2]Structural Biology Center, Argonne National Laboratory, Argonne, IL 60439, USA, [3]Electronics and Computing Technologies Division, Argonne National Laboratory, Argonne, IL 60439, USA and [4]Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA.

*Corresponding author.
E-mail: hendw@convex.hhmi.columbia.edu

Key words: advanced photon source, HIT protein family, PKCI, tumor suppressor, X-ray crystallography

Received: 24 March 1997
Revisions requested: 14 April 1997
Revisions received: 17 April 1997
Accepted: 21 April 1997

Structure 15 June 1997, 5:763–774
http://biomednet.com/elecref/0969212600500763

© Current Biology Ltd ISSN 0969-2126

Introduction

The FHIT (fragile histidine triad) protein is a member of a large and ubiquitous family of proteins termed the HIT family due to a conserved histidine triad (His-X-His-X-His) amino acid sequence motif found at the C-terminal end of their polypeptide sequences [1,2]. Sequences corresponding to HIT proteins have been found in diverse organisms, including *Haemophilus influenzae*, *Methanococcus jannaschii*, *Zea maize* and *Homo sapiens*. Although amino acid sequences from putative *HIT* genes can be aligned to reveal at least two subfamilies within the superfamily, a more correct functional classification of this family is not complete. One subfamily, characterized by a highly conserved C-terminal sequence, has greater than 94% sequence identity over the entire protein length between all known mammalian homologs. In addition, over 66% amino acid sequence identity is observed in this subfamily between organisms as diverse as humans and *Caenorhabditis elegans*, and over 58% identity is observed between humans and the cyanobacterium *Synechococcus sp*. We have previously described the identification, cloning and structure determination of human protein kinase C interacting protein (PKCI), a member of the highly conserved HIT subfamily, in its free form [3]. Recently, the structure of the rabbit HIT protein homolog, renamed as HINT, has been reported [4]. The rabbit HIT protein is 96% identical to human PKCI with five amino acid substitutions. Although the rabbit HIT protein and human PKCI structures were not compared in detail, the sequence conservation suggests that they are virtually identical.

Members of a more divergent HIT subfamily, to which FHIT belongs, share approximately 50% sequence identity between one another, but diverge significantly from

the more conserved subfamily, of which PKCI is a member, by the addition of a C-terminal amino acid sequence extension. A core domain of approximately 100 amino acids has amino acid sequence similarity among all HIT family members. PKCI and FHIT share only 20.7% amino acid sequence identity over an 110 amino acid overlap. We are now able to report the three-dimensional MAD structures and subsequent analysis of a very divergent HIT family member, FHIT, in both apo (free) and partial ligand bound forms.

The *FHIT* gene was recently identified by an exon-trapping strategy from cosmids covering a region of human chromosome 3p14.2 involved in homozygous deletions in epithelial cancer cell lines [5]. Subsequent genetic studies of the *FHIT* locus show abnormalities and homozygous deletions in several cell lines and primary tumors. From these studies, it has been postulated that the FHIT protein may function as a tumor suppressor when expressed from this gene locus, thus implying a possible role for FHIT in carcinogenesis [6–15]. A subsequent genetic study focusing on *FHIT* in colorectal cancers suggests that FHIT does not play a primary role in this type of cancer [16]. It was recently reported that reciprocal hybrid transcripts of *HMGIC* and *FHIT* are expressed in a pleomorphic adenoma of the parotid gland, suggesting that two tumor-associated genes are disrupted in a benign tumor, again demonstrating the high level of recombination activity at this locus [17]. Definitive evidence supporting the putative role for FHIT as a tumor suppressor has yet to be elucidated.

FHIT has been shown to have sequence similarity (52% identity over 110 amino acids) to a *Schizosaccharomyces pombe* HIT family member, which has been previously shown to have diadenosine 5′,5‴-P^1,P^4-tetraphosphate (Ap_4A) asymmetrical hydrolase activity [18]. Dinucleoside polyphosphates, which are primarily the by-products of some tRNA synthetases, have been postulated to function *in vivo* as messengers in cellular responses to growth conditions and stress response [19]. It has recently been demonstrated that the FHIT protein is capable of asymmetrical cleavage of a broad range of dinucleoside polyphosphates, which always yields a nucleotide monophosphate [20]. FHIT shows a preference for diadenosine 5′,5‴-P^1,P^3-triphosphate (Ap_3A) as a substrate, in contrast to the *S. pombe* enzyme which prefers Ap_4A. In this reaction, FHIT reacts with Ap_3A to yield ADP and AMP products. It should be noted that both enzymes cleave a broad range of substrates, but with varying efficiencies depending on the substrate. These previous studies report that catalysis is metal-dependent, as the reactions between FHIT and its substrates can be halted by the addition of EDTA. The present structural study demonstrates that FHIT interacts with a nucleotide analog in the absence of a divalent metal cation, thus, enabling a detailed analysis of the critical residues involved in binding and catalysis of nucleotide polyphosphates which are conserved among divergent HIT family members. The structural analysis of FHIT, a diverse human HIT family member, and PKCI will provide insights into the *in vivo* function of HIT proteins.

Results and discussion

Undulator MAD experiment and structure determination

The Structural Biology Center (SBC) beamline 19-ID at the Advanced Photon Source (APS) utilizes an undulator insertion device to generate very bright X-rays in narrow tunable harmonic peaks. The APS operates at suitably high particle energy to provide the first undulator harmonic in the X-ray range for multiwavelength anomalous diffraction (MAD) experiments at the K edge of selenium. The structure of FHIT has been determined by MAD phasing [21] using the selenomethionyl protein [22], in both its free and adenosine/sulfate-bound forms. The data for free and ligand-bound forms of FHIT were measured with the SBC APS1 3×3 CCD array detector (CCD) [23] and a FUJI image plate (IP) system, respectively. Both four-wavelength experiments took advantage of the ability to tune the peak of the harmonic by altering the gap width of the undulator during data collection, as has been done previously [24]. Three gap widths were chosen to center the peak of the first harmonic at the wavelength of choice. A single setting sufficed for peak and inflection points of the selenium K edge, which are very close together. Both experiments utilized crystals of the selenomethionyl-FHIT protein, although one of the crystals was also soaked in an adenosine/vanadate mixture in order to obtain a complex between protein and nucleoside. The nucleoside soak severely limited the diffraction quality of this sample in comparison with the free form of the protein. Although this enabled two independent structure determinations, it hampered the ability to compare results from the IP system and the CCD detector directly.

An almost complete set of phases was obtained by MADSYS [21] analysis to 2.8 Å spacings for the FHIT–adenosine/sulfate complex IP data set. This gave a fully interpretable Fourier map, whereupon the chain was traced and all ordered amino acid residues were positioned. A partially complete phase set was calculated to 2.0 Å spacings for the free FHIT CCD-data set, which resulted in an incomplete and disconnected electron density map. The limited extent of this data set was due to a time constraint, which prevented the completion of the experiment, and to an abundance of saturated low-angle reflections ($d > 5.0$ Å). The addition of phased reflections out to 5.0 Å spacings from the IP data set and density modification [25] at 2.0 Å yielded readily interpretable maps at 2.0 Å resolution, which allowed the peptide chain to be retraced and sidechains to be positioned [26]. Although these differences make it difficult to compare

Table 1

MAD diffraction data.

Data set	Wavelength (Å)	Reflections (unique/total)*	Completeness (%)	(I/σ)	R_{sym} (%)†
IP data set	0.9951	19709/81943	97.3	8.0	9.7
(10–2.2 Å)	0.9795	19775/83506	97.6	8.0	7.2
	0.9793	19649/81873	96.9	8.0	7.2
	0.9639	19907/84307	98.2	8.0	9.2
CCD data set	0.9951	21007/39596	55.5	44.4	2.5
(10–1.67 Å)	0.9795	41754/82995	65.2	51.3	3.5
	0.9793	41857/84710	66.3	62.3	3.6
	0.9639	43275/87495	67.1	48.6	3.9

*Unique reflections distinguish Bijvoet mates. †$R_{sym} = \Sigma|I - <I>|/\Sigma I$, where I = observed intensity and $<I>$ = average intensity.

the strategies and detectors used in the two experiments, the merging and phasing statistics from both experiments demonstrate that undulator beamline 19-ID and the CCD detector are capable of producing data of sufficient quality for accurate MAD phasing (Tables 1–3; Fig. 1). Ultimately, the native structure was refined against a more complete and higher resolution data set that was measured at beamline X4A at the National Synchrotron Light Source.

FHIT structure description

The electron density for native FHIT was well-enough defined to complete the model for residues 2–106 and 127–147, but not for the intervening segment or the first residue. The Cα backbone from this model is shown in Figure 2a for the FHIT dimer. Elution of FHIT upon size exclusion chromatography suggests that it is a dimer in solution (data not shown) and, although there is only one chain per asymmetric unit in the crystal, a dimer interaction similar to that previously reported for the PKCI dimer [3] is observed across a crystallographic twofold axis. The overall structure of the FHIT protomer can be described as a general α + β type and further subclassed as an α + β meander fold [27]. The FHIT protomer contains two helices, A and B, and seven β strands. These elements are shown in the ribbon diagram of Figures 2b and c. Strands three to seven form a five-stranded antiparallel sheet and antiparallel strands one and two form a β hairpin across from and at an angle to the other sheet. In the protomer, helix A packs on one side of the five-stranded antiparallel sheet. Helix B packs on the same side of the molecule as helix A and primarily interacts with strands three and four and the loop connecting strands two and three. Two protomers are brought together in the dimer by interactions between helices A and A′ and by the formation

Table 2

MAD phasing analysis.

		Anomalous diffraction ratios*					
		pre-edge	edge	peak	above edge	f′†	f″†
IP data set (10–2.8 Å)	pre-edge	0.042 (0.039)	0.077	0.078	0.037	–3.65	1.14
	edge		0.057 (0.034)	0.046	0.077	–10.73	3.98
	peak			0.073 (0.039)	0.077	–9.64	5.78
	above edge				0.054 (0.035)	–3.62	3.83
CCD data set (8–2.0 Å)	pre-edge	0.039 (0.032)	0.066	0.090	0.038	–3.60	1.10
	edge		0.058 (0.025)	0.065	0.057	–10.06	4.26
	peak			0.079 (0.038)	0.077	–10.21	6.10
	above edge				0.051 (0.029)	–4.30	3.59

*Anomalous diffraction ratios = $<\Delta|F|^2>^{1/2} / <|F|^2>^{1/2}$, where $\Delta|F|$ is the absolute value of the Bijvoet (diagonal elements) or dispersive difference (off diagonal elements), respectively. †f′ and f″ are the refined values obtained from MADSYS [21].

Table 3

Anomalous phasing statistics.

	$R(^{\circ}\|F_T\|)$*	$R(^{\circ}\|F_A\|)^{\dagger}$	$<\Delta(\Delta\phi)>^{\ddagger\S}$	$<\sigma(\Delta\phi)>$
IP data set	0.082	0.464	46.26°	19.31°
CCD data set	0.049	0.308	30.93°	18.93°

*$^{\circ}F_T$ is the structure factor due to normal scattering from all the atoms; $^{\dagger}{}^{\circ}F_A$ is the structure factor due to normal scattering from the anomalous scatterers only. $^{\ddagger}\Delta\phi$ is the phase difference between $^{\circ}F_T$ and $^{\circ}F_A$. $^{\S}\Delta(\Delta\phi)$ is the difference between two independent determinations of $\Delta\phi$.

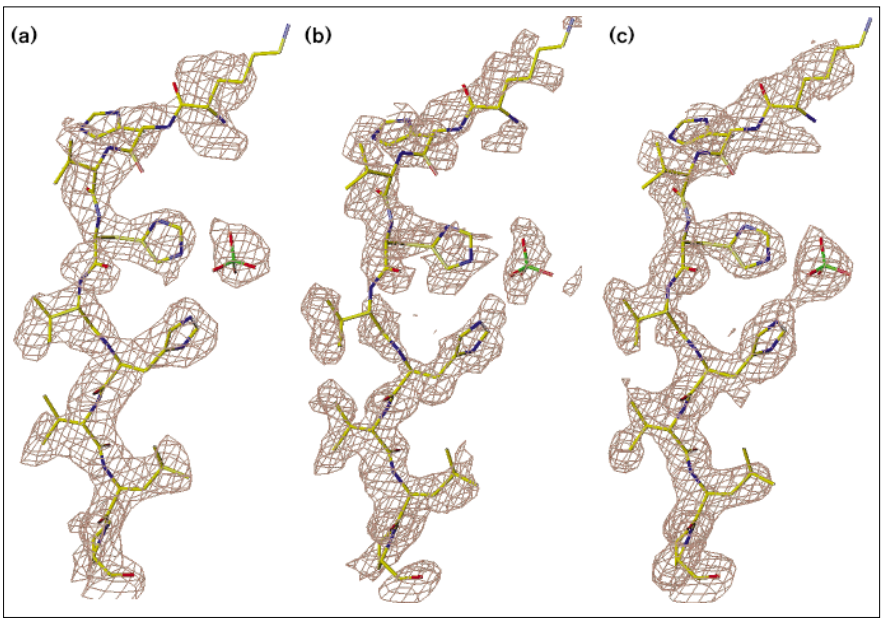
of a ten-stranded antiparallel sheet comprising the respective five-stranded antiparallel sheet of each protomer. A common hydrophobic core is formed from the protomer interactions within the context of the dimer. The extensive dimer interface occupies 2374 Å^2 per dimer compared to the total dimer surface area of 11239 Å^2 as computed with a probe radius of 1.4 Å [28].

With the structure of FHIT at hand, the deletion of specific exons from the *FHIT* locus, which in turn codes for the aberrant *FHIT* transcripts observed in many cancer cell lines and primary tumors can be analyzed in terms of protein structure [5–9]. Several of these aberrant transcripts have been sequenced and details of the specific deleterious effects on *FHIT* transcripts have been described. Most of the observed deletions in the *FHIT* gene either truncate or eliminate the coding regions. Specific exons are deleted in some cases leading to the possibility of a fused protein. Several possible and observed exonic deletions have been mapped onto the FHIT protein structure, including a deletion of exons 8, 8 to 9, 5 to part of 6, 5 to 6, part of 6 to 9, 5 to 9 and 4 to 7. Analysis of any of these deletions mapped onto the three-dimensional structure lead us to believe that none of these mutant proteins would fold and dimerize with the full length FHIT protein; thus, it is unlikely if not impossible to have a dominant negative phenotype when one copy of the gene is found to be aberrant in length or sequence. This does not exclude the possibility of dose-dependent effects of *FHIT* expression when one copy of the gene has been disrupted.

Comparison with PKCI, another human HIT family member

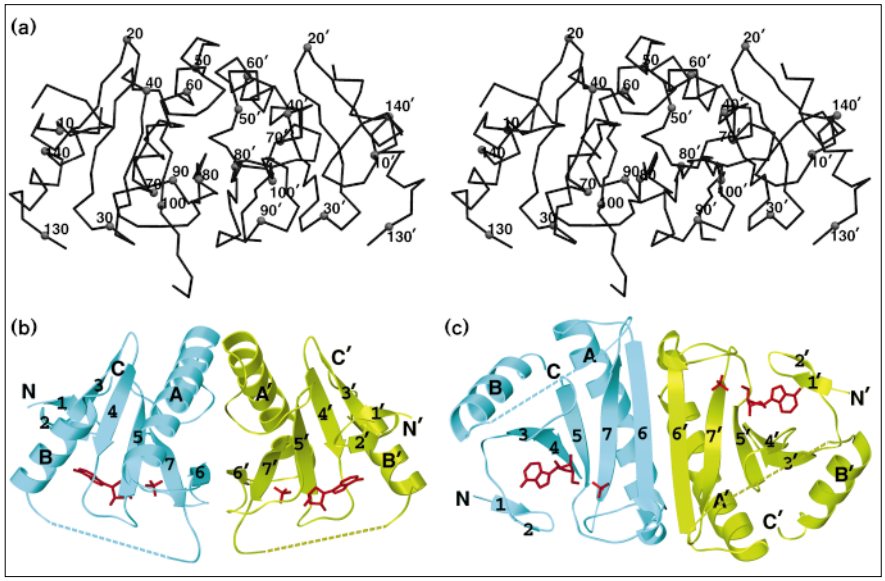
We recently reported the structure of free human PKCI . Human PKCI was identified in a yeast two-hybrid screen using protein kinase C β (PKC-β) as a probe and subsequently cloned from a HL60 cDNA library [3]. Although the role for PKCI in cells is not clear at present, evidence for the physiological role of the protein has been suggested by interactions obtained in the yeast two-hybrid assay between PKCI and PKC-β or ataxia-telangiectasia group D complementing (ATDC) protein, a gene product

Figure 1



Electron density for the active site histidine region of FHIT. **(a)** Experimental electron density generated with MAD phases from the IP data set. The map is at 2.8 Å resolution and is contoured at 1.0σ. **(b)** Experimental electron density generated with MAD phases from the CCD data set. The map is at 2.0 Å resolution and is contoured at 1.0σ. **(c)** Experimental electron density generated with MAD phases from the CCD data set to 2.0 Å spacing augmented with MAD phases and reflections from the IP data set to 5.0 Å spacing after density modification. The map is contoured at 2.0σ. (Figure generated with O [26].)

Figure 2



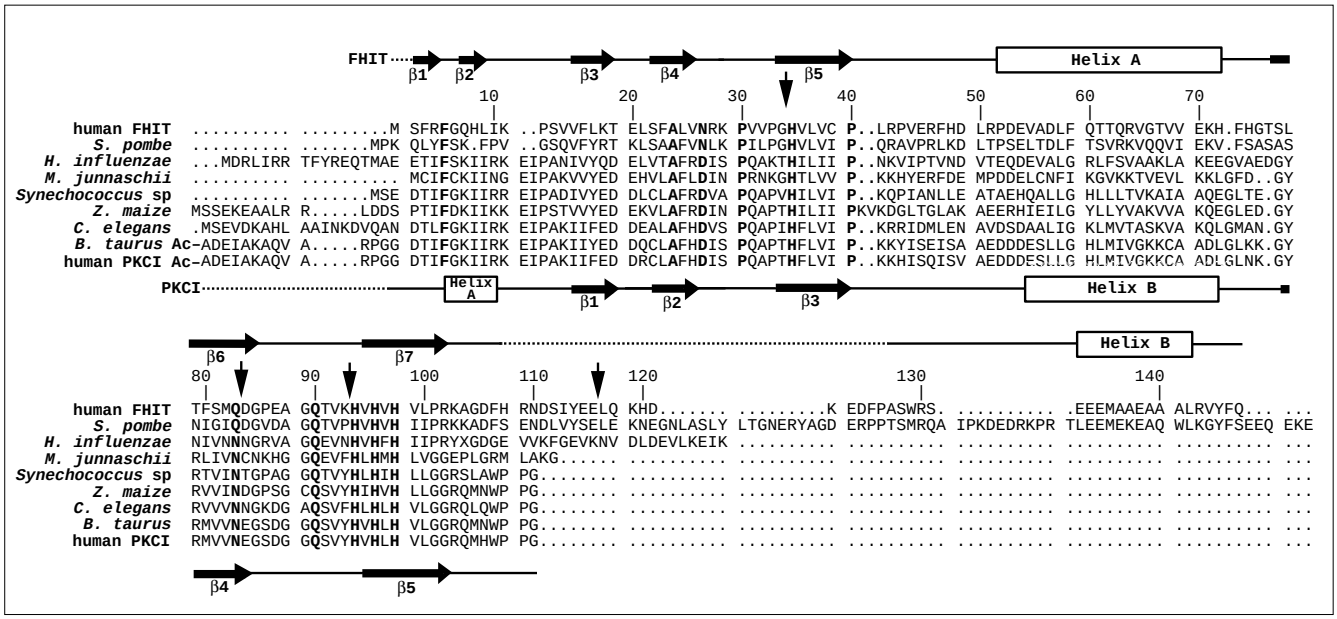
The FHIT dimer. (a) Stereo Cα backbone trace of FHIT with Cα atoms numbered every ten amino acids. Numbering is differentiated between protomers by use of a prime symbol. (b,c) Orthogonal views of a Cα spline Richardson style ribbon diagram of the FHIT dimer with helical elements lettered A (residues 52–72) and B (residues 132–141) and beta elements numbered 1 (residues 2–4), 2 (residues 7–9), 3 (residues 15–19), 4 (residues 22–26), 5 (residues 34–40), 6 (residues 77–85), and 7 (residues 95–102). Protomers are distinguished by use of a prime symbol in the numbering scheme. The disordered residues between amino acids 108 and 125 are indicated by a dashed line. (These figures were generated using SETOR [53].)

isolated for its ability to complement the ionizing radiation sensitivity of AT group D fibroblasts [29]. PKCI has recently been mapped to human chromosome 5q31.2 by fluorescence *in situ* hybridization [30]. In our PKC-β screen, PKCI was shown to interact specifically with the regulatory domain of this PKC isoform. In the ATDC screen, ATDC interacted with both vimentin and PKCI. The first PKCI HIT homolog to be characterized biochemically was bovine PKCI (94.4% identity over 126 amino acids with human PKCI) (Fig. 3). Bovine PKCI was purified directly from brain, reported to be a potent inhibitor of protein kinase C and was subsequently shown to interact with zinc [31–36]. A HIT homolog from maize (53% identity over 115 amino acids with human PKCI) was identified, cloned and the protein biochemically characterized. The maize HIT protein was shown to interact with zinc and, albeit weakly, synergistically inhibit PKC in the presence of 14-3-3 protein [31]. Since these experiments were carried out, doubt has been cast on the ability of these proteins to inhibit protein kinase C *in vitro* [3,31,35]. Although previous studies demonstrate zinc binding by some HIT proteins, we have shown by the crystallographic analysis of PKCI at high resolution in the presence and absence of zinc that zinc does not bind specifically to this protein in the crystal lattice [3].

The amino acid sequences of PKCI and FHIT overlap with a single gap and 20.7% identity spanning residues 1–110 of FHIT (Fig. 3); this value of sequence identity is close to random. Nevertheless, the PKCI and FHIT three-dimensional structures are quite similar for a core corresponding to the five-stranded β sheet and helix A (FHIT), but they diverge markedly at both ends. A least-squares superposition of the 88-residue core (FHIT residues 14–101 and PKCI residues 30–117) gives a root mean square deviation (rmsd) of 2.15 Å between FHIT and PKCI Cα atoms. The related helices (A in FHIT and B in PKCI) have coincident axes, but they superimpose poorly because of a divergence where PKCI is distorted at Gly75, at a protomer–protomer contact (Fig. 4). Consequently, if the helices are excluded from the least-squares calculations (FHIT residues 14–46 and 72–101 against PKCI residues 30–62 and 92–117), the rmsd drops to 1.1 Å. The core structural similarity between FHIT and PKCI extends to the entire dimer. The rmsd values for the dimer superpositions of the entire core and for the β-sheet substructure are 2.2 Å and 1.7 Å, respectively. It is interesting to note that the sequence similarity observed in the N-terminal residues of PKCI and FHIT does not correspond to a similar secondary structure in the two proteins (Figs 3,4). The corresponding residues in FHIT that make up strands one and two are helical in PKCI.

A major structural difference between FHIT and PKCI is that the PKCI HIT subfamily has a highly conserved C-terminal sequence involved in extensive protomer–protomer and PKCI–ligand interactions (Figs 3,4), which apparently serves to segregate two large clefts, one in each protomer, by interactions between the C-terminal amino acids of each protomer. This segment in human PKCI has the sequence Gly117-Gly-Arg-Gln-Met-His-Trp-Pro-Pro-Gly126. These residues wrap around and interact with the respective residues from the other protomer, ending in a salt bridge between the C-terminal Gly126 of one PKCI protomer and Arg119 of the other. FHIT lacks the conserved sequence and subsequent interactions between

Figure 3



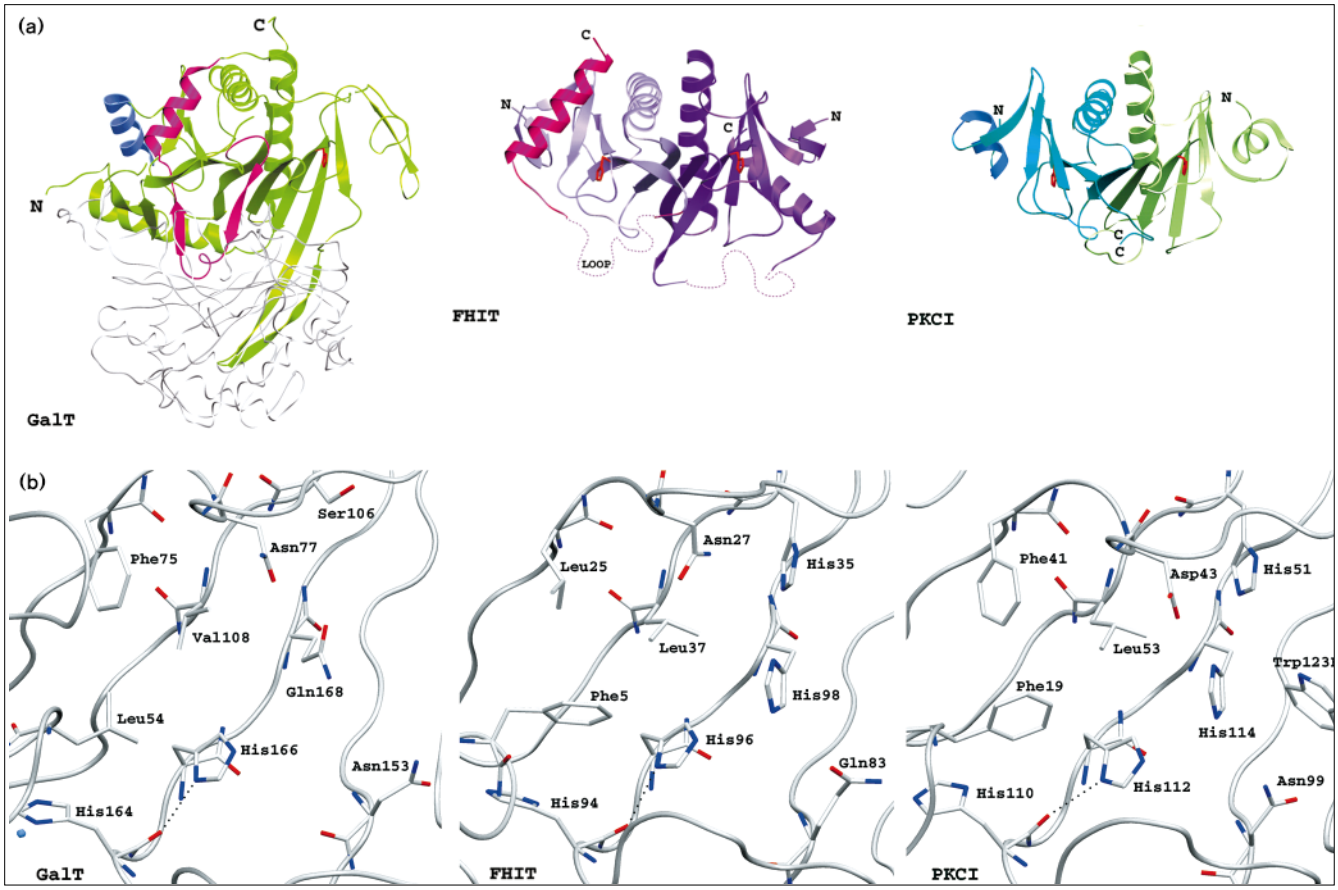
Alignment of several HIT family members (only a subset of HIT family members are depicted in the figure for clarity). Those sequences included are *Homo sapiens* FHIT [6] (GenBank accession (GB) U46922), *Schizosaccharomyces pombe* HIT [54] (GB U32615), *Haemophilus influenzae* HIT [55] (GB U32777), *Methanococcus junnaschii* HIT [56] (GB U67530), *Synechococcus sp.* HIT [57] (GB M34833), *Zea maize* HIT [31] (GB Z29643), *Caenorhabditis elegans* HIT (GB Z71261), *Bos taurus* HIT [32] (GB U09405), and *Homo sapiens* PKCI (HIT) [3] (GB U51004). It was previously determined that PKCI contains an N-terminal acetyl group [3] in place of the N-terminal methionine similar to that identified for endogenous bovine HIT protein. Vertical arrows demarcate exons which includes the region of exons 5–9 in the coding region of *FHIT*. Secondary structural elements for FHIT are indicated above the FHIT polypeptide sequence whereas secondary structural elements for PKCI are indicated below the PKCI polypeptide sequence. Helices are denoted by boxes, β strands by black arrows, and polypeptide chain by a bold black line. Disordered regions in the respective structures are represented by dashed lines. Residues highlighted by bold characters are either totally conserved within the entire HIT family or are involved in ligand binding. All numbering relates to the FHIT sequence (top line). Phe5 of FHIT is highly conserved and located below the nucleotide base in the ligand-binding pocket. Ala24 of FHIT is totally conserved in the HIT family. Residue 27 of FHIT directly contacts the nucleotide ribose. Prolines at FHIT positions 30 and 40 are highly conserved in the HIT family. Histidines at FHIT positions 35, 94, 96 and 98 are totally conserved in the HIT family and comprise a major portion of the active site; they are involved in catalysis and phosphate binding as is residue 83. The glutamine at FHIT position 90 is involved in stabilizing strands six and seven (FHIT notation) in a splayed conformation.

C-terminal amino acids, thus, creating a single large cleft running from one ligand-binding site to the other. It should be noted that FHIT residues 107–126 were not sufficiently ordered for placement into electron density, but the direction of the polypeptide chains preceding these disordered residues and some observable electron density within this region of disorder indicates that at least some of these residues are located above the cleft where the partial substrate is bound (see later discussions).

Ligand binding by HIT family members

General aspects of ligand binding for the HIT family can be described based on the nucleotide-analog–FHIT complex. Due to the very high concentrations of ammonium sulfate in the crystallization and stabilization solutions, it appears vanadate was unable to replace the bound sulfate in the FHIT active site. The adenosine and sulfate groups bound in the FHIT active site do, however, elucidate important interactions between the ligand and protein. The ligand-binding site is located in a deep cleft within each FHIT protomer, the floor of which is formed from the five-stranded β sheet. The sides of the cleft are formed from strands one and two, together with the loop connecting strands six and seven on one side and helix B and the loop between strands four and five on the other side (Figs 2,5). Interactions between ligand and protein include hydrophobic interactions, direct and water-mediated hydrogen-bonding interactions, and the interactions between active-site histidines 96 and 98 and the α phosphate group of the nucleotide analog. The nucleotide analog–FHIT complex represents the FHIT product complex in the asymmetric hydrolysis of Ap_3A to ADP and AMP. Although this study clearly shows how FHIT binds the AMP product of its respective substrates, the components of FHIT responsible for binding the remaining portion of the substrate during catalysis remain elusive. As stated previously in the text, the disordered loop in FHIT, located above the active-site region, could close down on

Figure 4



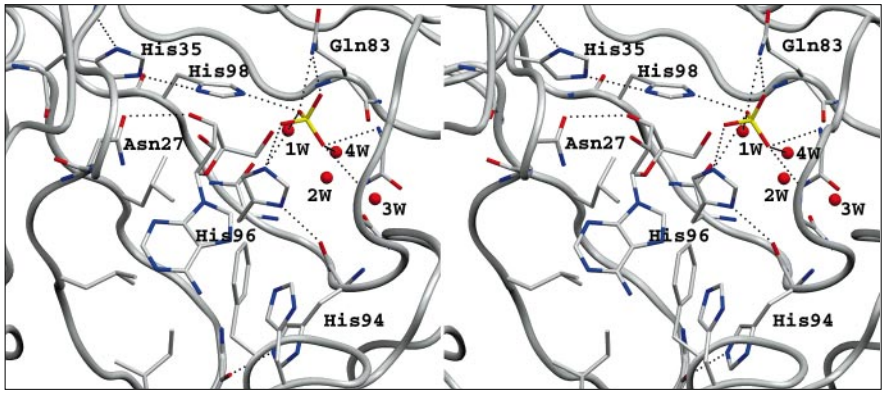
Comparison of the structures of GalT, PKCI and FHIT. (a) Ribbon diagrams of GalT, PKCI and FHIT. Note the similar overall fold conserved between these proteins. Histidine residues implicated in catalysis in the mechanism are shown in red (ball-and-stick) and include: GalT, His166; PKCI, His112; and FHIT, His96. The GalT monomer has only one active-site histidine located on the right side of GalT in the diagram, whereas the HIT dimers have two equivalent active sites. Comparison of PKCI to GalT reveals an additional structural analog to the N-terminal helix (color-coded dark blue) of PKCI, which is found in the dimerization half of the GalT monomer (located on the left side of the GalT monomer in the diagram). This helix is missing in the active-site half (right side) of the GalT monomer. Similar features can be observed in comparison of the FHIT and GalT structures. The C-terminal helix of FHIT (color-coded magenta) is absent from PKCI, but is present in the dimerization half (left side) of the GalT monomer. Again, this helix is absent from the active-site half of the GalT monomer. In addition, the disordered loop in FHIT between strands 7 and helix B corresponds well in position and number of residues with a loop in GalT which is involved in extensive interactions in the GalT dimer and active site of the dimer partner (labeled 'loop' in the FHIT structure and color-coded magenta in the GalT structure). (b) Close-up of the active-site regions for GalT, PKCI and FHIT in their apo or free forms. Only a subset of residues in the active site are depicted on a Cα backbone spline for clarity. Residues depicted most importantly include the active-site histidines of each protein (GalT, His166; FHIT, His96; PKCI, His112). Interestingly, several hydrophobic residues that make contact to the nucleoside base are somewhat conserved (FHIT, Phe5 Leu37 and Leu25, and the homologous residues in GalT and PKCI). The residue responsible for binding the nucleoside ribose is somewhat conserved among FHIT, PKCI and GalT (Asn27, Asp43 and Asn77, respectively). Although these similarities are striking, two major differences exist between GalT and the two HIT family members: firstly, GalT is missing two of the highly conserved histidine residues in the active site; and secondly, FHIT His98 or PKCI His114 is replaced by GalT Gln68 and FHIT His35 or PKCI His51 is replaced by GalT Ser106. These substitutions are likely to play some role in dictating and differentiating substrate specificity between the GalT-related enzymes and the HIT protein family.

the active site, possibly providing specificity for the binding of the non-AMP portion of the respective ligands.

The highest degree of amino acid conservation between PKCI and FHIT is observed in the ligand-binding pocket and notably includes FHIT residues His35, His94, His96 and His98. Three water molecules, 1W–3W, located between strands six and seven, and behind the phosphate, are conserved in all FHIT and PKCI structures, including other complexed and free forms of both protein molecules (CDL, unpublished results). An additional water, 4W, which is within hydrogen-bonding distance to a phosphate oxygen and the other conserved waters, is found only in the product (AMP) complex of FHIT (Fig. 5). Residues

Figure 5

Schematic stereo diagram of FHIT in adenosine/sulfate product complexed form. Hydrogen-bonding interactions thought to be important in catalysis and ligand recognition are depicted by dotted lines. To simplify the depiction, only a subset of residues have been shown superimposed on the Cα backbone spline. Although most residues involved in interactions between protein and adenine base are not conserved between FHIT and PKCI (not shown here), both proteins utilize a deep hydrophobic cleft and water-mediated contacts to the backbone of N-terminal residues in binding the adenosine base.



surrounding the nucleotide base and ribose are generally less well-conserved between PKCI and FHIT, suggesting that alternative ligands/bases could interact with binding sites on either FHIT or PKCI. Especially notable among the predicted common features in ligand binding between FHIT and PKCI are the many direct contacts and interactions between the α phosphate group in respective ligands and the active-site residues in respective proteins, thus, implying that although PKCI and FHIT probably interact with different substrates, catalysis at the α phosphate group proceeds through a similar pathway. Two complexes between a nucleotide monophosphate and the rabbit HIT protein have been reported [4]. The FHIT–nucleotide complex is in agreement with those complexes of the rabbit HIT protein in regards to the general interactions with the nucleotide base, ribose and sugar.

Structural and functional similarity to galactose-1-phosphate uridylyltransferase

A structural analog to the HIT family of proteins has recently been uncovered by analysis of the Protein Data Bank [37,38]. In this analysis, it was found that although no overall sequence similarity is observed, the human PKCI dimer [3] was structurally similar to a protomer of galactose-1-phosphate uridylyltransferase [39,40], also known as GalT. A subsequent comparison between the rabbit HIT protein and GalT was also recently described, although the methods used to determine similarity were not described [4]. GalT catalyzes the exchange of a UMP moiety between the hexose-1-phosphates of glucose and galactose, thus exchanging UDP-glucose and galactose-1-phosphate with UDP-galactose and glucose-1-phosphate, a nucleotidylyl phosphotransferase reaction. The GalT phosphotransferase reaction is catalyzed by a histidyl residue [39], in a manner which we believe to be similar to that observed in the reactions of FHIT and PKCI (CDL, unpublished results), members of the HIT family of proteins. When a detailed structural comparison of these proteins is made, similarities in the overall fold are seen between the GalT protomer and the PKCI and FHIT homodimers (Fig. 4). It should be noted that although two active sites exist per HIT dimer, only one active site exists per GalT protomer while the other half of the GalT protomer is involved in dimerization interactions in the GalT dimer. A structure-based sequence alignment establishes weak sequence similarity between GalT and the HIT proteins. The sequence identity of aligned structural elements is 17.4% between 155 amino acids of the GalT protomer with the two PKCI protomers and 14.9% between 168 amino acid residues of the GalT protomer and the two FHIT protomers. When the HIT protomers are aligned with the active-site half of the GalT protomer they have weaker sequence identity than with the other half, which has 18.8% identity with PKCI and 19.8% identity with FHIT. No available sequence analysis software identified any similarity between PKCI and GalT or FHIT and GalT.

The structural alignment of all corresponding mainchain atoms of the respective protomers in the HIT dimers and the GalT protomer revealed that PKCI aligns with an rmsd of 2.1 Å to the active site half of the GalT protomer and with 2.6 Å to the dimerization half of the GalT protomer. FHIT aligns with an rmsd of 1.9 Å and 2.0 Å with the respective halves of the GalT protomer. A structural comparison of the respective halves of the GalT protomer revealed that they align with one another to an rmsd of 1.8 Å. It is interesting to note that although the active-site half of the GalT protomer does not share common structural elements with either the N-terminal helix of PKCI or the C-terminal helix and loop between strand 7 and the C-terminal helix of FHIT, the dimerization half of the GalT protomer does (Fig. 4a). Thus, aspects of the active site were apparently conserved in one half of the GalT protomer, whereas several common structural elements were conserved in the other. Similarity in the region of the active site is especially notable with respect to the catalytic histidine residues implicated in catalysis in GalT, FHIT

and PKCI (Fig. 4b) (CDL, unpublished results). When the three structures are aligned, the catalytic histidines (PCKI, His112; FHIT, His96; and GalT, His166) superimpose almost perfectly. All these histidines are coordinated through their Nδ atom to the respective carbonyl oxygen of the other conserved histidines (PKCI, His110; FHIT, His94; and GalT, His164) in the active site, leaving the Nε of each catalytic histidine available to attack the α phosphorus atom of the respective substrates.

Although the similarities between these protein families are striking, it should be noted that major differences exist between the GalT and HIT protein families, not least that the HIT protein homodimers correspond to a single GalT protomer. The structural alignment would apparently argue that FHIT and GalT are more structurally similar than are GalT and PKCI, but the differences are so small, and the sequence similarities weak enough that no definitive evolutionary relationship can be constructed, even between the two halves of the GalT protomer. Identical residues observed among all known HIT family members include four histidines in the vicinity of the active site, of which only two are found in the GalT active site, and only one is common to both halves of the GalT monomer and to all other HIT protein sequences. This implies that although the GalT and HIT family share an overall fold capable of binding and catalyzing reactions with a nucleotide polyphosphate, the specificity for respective substrates and products may differ substantially. GalT sequences from human and *Escherichia coli* have 56% identity, whereas the sequences among mammalian GalT family members are approximately 90% identical. In both cases, only weak sequence identity is found between HIT family members and GalT family members. As was stated before, the sequence identity among the PKCI HIT family members apparently corresponds well with the level of conservation within the GalT family, with 47% identity between *E. coli* and human, and greater than 94% identity between mammalian homologs of PKCI. If the GalT and HIT proteins share some very distant evolutionary link, it appears that the two families would have diverged long ago and conserved their independent functions, which subsequently led to two independent well-conserved protein families that share some chemical characteristics in their active sites, but no longer share common substrate specificity.

Mechanistic implications

Several aspects of the HIT family catalytic mechanism can be inferred from the data provided by the structures of HIT proteins in their apo (ligand-free) forms and the structure of the FHIT–adenosine/sulfate complex. If the reaction intermediate proceeds through a nucleotidylyl– histidyl covalent intermediate, as it does in the GalT reaction with its respective substrates, several catalytic requirements must be met. These include the activation of a histidine to initially attack the α phosphate and the subsequent activation of a second nucleophile to attack the phosphoramidate bond, thus releasing the nucleotide monophosphate product. This reaction pathway has been previously observed using two histidines [40–44]. In the mechanism for 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase, one histidine acts as the nucleophile to attack the phosphate, whereas the other acts as a base to deprotonate a water to produce the hydroxyl ion capable of attacking and releasing the covalent phosphohistidyl intermediate.

Structural analogy with this enzyme and GalT indicate that His96 of FHIT and His112 of PKCI act as the nucleophile in the inversion of the α phosphate in these reactions. Evidence for this mechanism includes the fact that respective histidine residues appear highly stabilized by a conserved structural motif, whereby the carbonyl oxygen atom of FHIT His94 or PKCI His110 is hydrogen bonded in a close interaction with the Nδ atom of FHIT His96 or PKCI His112 (with distances between the two histidine residues of 2.5 Å for FHIT and 2.6 Å for PKCI). This indicates indirectly that the Nε atom of FHIT His96 or PKCI His112 is deprotonated in the apo form of the enzyme at neutral pH and properly oriented to attack the phosphate group as it enters the binding pocket.

Kinetic studies on FHIT subjected to the site-directed mutagenesis of its four conserved histidines are consistent with our model for catalysis. Barnes *et al.* have found that mutation of His35 or His94 to asparagine reduces catalytic efficiency of FHIT by 36% and 69%, respectively [20]. Mutation of His98 to asparagine reduces activity to less than 2.5%, and when His96 is mutated to asparagine all detectable activity is ablated. We predict, on the basis of the structural evidence in both PKCI and FHIT and the mutagenesis data, that His96 is required for activity because it is directly involved in the nucleophilic attack and inversion of the α phosphate and in the covalent intermediate with the nucleotide monophosphate. His98 may likely be involved in the deprotonation of the water responsible for attacking and releasing the covalent intermediate. An analogous mutagenic study in 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase revealed similar behavior when the two active-site histidines were mutated to other residues—mutations of His258, the residue shown to be a phosphohistidyl residue during catalysis, ablated all enzymatic activity, and those of His392 diminished activity to less than 5% [42]. In the reaction catalyzed by this enzyme, His392 has been implicated as a proton donor and acceptor. Because of the similarity of residues in the vicinity of the active site among PKCI, FHIT and GalT, we propose the catalytic mechanism for PKCI and FHIT be extended to all HIT family members, especially with respect to the initial attack of the α phosphate in the respective substrates. Although the identity of the *in vivo* substrate is yet to be shown for

either enzyme or any other HIT family member for that matter, it appears that these molecules catalyze reactions that will proceed through a covalent nucleotidyl- phosphoprotein intermediate. Members of the HIT family of enzymes are capable of catalyzing the hydrolysis of a phosphoanhydride bond under certain conditions, however, a covalent enzyme intermediate suggests these enzymes may be capable of catalyzing a transferase reaction of the covalent intermediate instead of just a hydrolysis reaction (a short-sighted and limited aspect of the *in vitro* hydrolase assay). Clearly, this is something that should not be overlooked in the search for *in vivo* HIT protein function.

Biological implications

Here we report the successful MAD structure determination of the fragile histidine triad (FHIT) protein, a member of a large and highly conserved family of proteins known as the histidine triad (HIT) family of proteins. Although their *in vivo* function is still unknown, HIT proteins have been implicated in the binding and hydrolysis of nucleotide polyphosphates. It is likely that such a highly conserved family of enzymes, shown by the high sequence similarity among organisms ranging from mycoplasma, archae, bacteria, plants and humans, are involved in some critical and ubiquitous biochemical pathway that remains undiscovered.

FHIT, a gene that resides within a fragile site of human chromosome 3, has been postulated to be a human tumor suppressor gene. Unlike the p53 gene, no point mutations have been discovered in the coding regions of the *FHIT* gene which would alter the ability of FHIT to catalyze the hydrolysis of nucleotide polyphosphates. Instead, disruptions and translocations seem to be the major factors that alter the *FHIT* translation product. These alterations usually affect only one copy of the *FHIT* gene in the genome, leaving one full length copy available in the cell. On the basis of the structure of the FHIT protein, it is unlikely that truncated forms of the protein resulting from altered *FHIT* transcripts would be capable of folding and interacting with the full length gene product, thus, eliminating the possibility of a dominant negative phenotype when one altered *FHIT* transcript exists in the cell with one full length copy of the gene.

This study provides the first structure of FHIT, a diverse member of the HIT family of proteins, and suggests that distant members within the HIT superfamily share a similar overall fold and catalytic mechanism, some features of which can be extended to and inferred from the GalT family of proteins. As the catalytic activity observed *in vitro* demonstrates the ability of these enzymes to hydrolyze their substrates, it is likely that these enzymes could behave as nucleotidyl phosphotransferases through a pathway mediated via a covalent enzyme intermediate—several features of which can be inferred from the FHIT–nucleotide-analog complex. These studies support a mechanism in which the nucleophilic attack by respective histidines in various HIT family members at the **α**-phosphorus atom would result in a transient covalently attached nucleotidyl–protein intermediate before hydrolysis of the phosphoramidate bond.

Materials and methods

Cloning, protein expression and purification

The cDNA for FHIT was obtained [5] and subcloned into the pGEX-2T vector for over-expression in *E. coli* [45]. The protein was purified initially by passage over a glutathione-sepharose and MonoQ (Pharmacia) columns. The fusion protein was cleaved with thrombin and passed over these two columns again in the same order. A final purification step was applied using a Superdex 200 size exclusion column (Pharmacia). The protein was concentrated to ~3mg/ml and stored at –80°C in 5.0 mM Tris-HCL pH7.4, 1.0 mM DTT.

Crystallization

FHIT was crystallized using a partial factorial screen [46]. Final conditions were obtained whereupon crystals of both native and selenomethionyl FHIT could be grown reproducibly. These conditions are ~10 mg/ml protein, 1.25–1.35 M $AmSO_4$, 100 mM sodium cacodylate pH 6.5. Crystals grown under either condition were in space group $P6_122$ with unit cell dimensions of a = 50.6Å, b = 50.6Å, c = 268.0Å, a = 90°, b = 90° and g = 120°.

Data collection and processing

All data was collected on cryo-preserved crystals at 110K. Crystals were cryo-protected in a sucrose–$AmSO_4$ solution prior to cooling. The two four wavelength MAD experiments were collected at the APS SBC undulator beamline 19-ID on an IP system and the CCD detector [23]. The single wavelength experiment on Native2 was collected at the Howard Hughes beamline X4A at Brookhaven National Laboratory using an IP system. All IP and RaxisII data was process using DENZO and SCALEPACK [47]. CCD data was initially processed at the APS [48]. Subsequent data reduction was done using the CCP4 suite of programs [49] (Tables 2 and 3).

MAD data collection

Crystals were oriented with the c axis on the camera spindle in order to collect Bijvoet pairs on the same or adjacent images. A four wavelength IP data set was collected on a 0.1 mM adenosine/vanadate-soaked selenomethionyl crystal. Eight 30° segments (with the long axis parallel to the spindle) and four 20° segments (orthogonal to the previous orientation) of data were taken on the IP system. Only seven 30° segments of data were taken with the CCD detector. Rotations of 2.0° per 30 seconds with 0.5° overlap were taken using the IP system. Only fully recorded reflections were used in the IP-data processing. Rotations of 0.2° per 10 seconds were taken using the CCD detector by continuously summing partials. The wavelength and gap, if necessary, would be changed between each segment during data collection. Although this was not optimal for MAD phasing due to sample decay and ice build-up, it was a compromise between manual gap width changes and the speed of data collection.

Phasing

A selenium site was identified in a Bijvoet difference Patterson. This site was used to locate a second selenium position in a Bijvoet difference Fourier with phases generated to 2.8 Å from all the IP data using MADSYS [21]. The third potential selenium site corresponding to the N-terminal methionine was not revealed and was later found to be disordered. MAD phases were determined to 2.8 Å using the two selenium positions. MAD phases were determined for the CCD data using only the first four 30° segments of data. Combining further segments of data completely degraded the quality of the phases due to unknown problems with the last two segments of data. Because

Table 4

Crystallographic data and refinement statistics.

	FHIT		
	Native1	Adenosine soak	Native2
Diffraction data			
spacegroup	$P6_122$		
cell (Å)	50.6 50.6 268.0	50.6 50.6 268.0	50.6 50.6 268.0
d_{min} (Å)	1.7	2.2	1.8
λ (Å)	0.9639	0.9551	1.0036
detector	CCD	Fuji image plates	Fuji image plates
unique*/total reflections	30590 / 89229	19709 / 81943	35339 / 197773
coverage (%)†	68.2 (31.0)	97.3 (92.6)	94.7 (69.3)
R_{sym} (%)†‡	3.9 (18.8)	9.7 (45.7)	5.6 (33.3)
source	BL-19-ID	BL-19-ID	BL-X4A
Refinement			
resolution (Å)	8–1.9	8–2.4	8–1.85
R‡	0.222	0.201	0.202
R_{free}‡	0.249	0.249	0.225
no. reflections	23987	14428	29538
no. protein atoms	1049	1030	1021
no. water atoms	112	101	180
rms bonds/angles (°)§	0.012 / 1.768	0.018 / 2.106	0.016 / 1.823
average B (Å^2)§	27.5	34.1	24.6
rms B mc/sc§	1.8 / 3.5	2.3 / 4.6	1.2 / 2.4

*Unique reflections distinguish Bijvoet mates. †Numbers in parentheses in the coverage and R_{sym} columns indicate the values for these quantities in the last resolution bin. Although the R_{sym} is quite high for the last resolution bin for the adenosine complex, the data was apparent upon inspection of the image plates and it is believed to be inflated due to the high degree of anisotropy with this crystal. ‡$R_{sym} = \Sigma|I - <I>|/\Sigma I$, where I = observed intensity and $<I>$ = average intensity; ‡R, R based on 95% of the data used in refinement; ‡R_{free}, R based on 5% of the data withheld for the cross-validation test. §B is the crystallographic B factor and rms B mc/sc is the root mean square deviation of B for mainchain and sidechain atoms; rms bonds/angles is the root mean square deviation from ideal bonds (Å) and angles (°).

these data were very incomplete, phased reflections to 5.0 Å were used from the IP data set to augment the phase set obtained from the CCD data to 2.0 Å which underwent several rounds of density modification [25] at 2.0 Å. Correlation coefficients have been calculated using O between the final F_c maps, MAD phased maps, and solvent flattened maps. The IP-MAD map has a correlation of 0.579 between the 2.8 Å MAD map and refined F_c map for the adenosine/sulfate structure. A correlation coefficient of 0.372 was calculated between the final 2.0Å F_c map of the native structure and the 2.0 Å MAD data augmented with up to 5.0 Å data from the IP data set. This map was not readily interpretable at this stage. When solvent flattening was applied to these data, the calculated correlation coefficient rose to 0.673 between the solvent flattened 2.0 Å map and the final F_c map.

Refinement

Two independent polypeptide chains were traced into the 2.8 Å adenosine-soak MAD map and apo-CCD map, respectively, using O [26]. All subsequent structures were solved using partial models in molecular replacement with AMORE [50]. All models were refined with X-PLOR using the cross-validation test [51,52]. Each model was put through one simulated annealing round with all subsequent phase extension and refinement accomplished with positional least-squares refinement. Each FHIT model roughly includes residues 2–108 and 125–147. Anisotropic B-factor refinement and scaled F_{obs} were generated (as implemented by X-PLOR) on FHIT data sets before the addition of waters to the model. Subsequent refinement and reported R_{values} utilized the scaled data for FHIT (Table 4).

Accession numbers

Coordinates have been deposited in the PDB with accession codes 1fit, 2fit and 3fit.

Acknowledgements

We thank the staff of beamline X4A at the National Synchrotron Light Source (NSLS) and beamline 19-ID at the Advance Photon Source (APS). We particularly thank Daved Fremont, Gwyndaf Evans, Mary Westbrook, Tom Coleman, Bob Daly, Frank Rotella, Carol Zimmer, John Gonczy and Randy Alkire for their help during data collection at beamline 19-ID and all the members of the Hendrickson lab for helpful discussion, especially Hao Wu for her aid in utilizing the newest MADSYS implementation, Craig Bingman for discussions about catalysis, and Larry Shapiro for his critical reading of the manuscript. We thank Michael Klein for continual discussions involving the HIT protein family. Beamline X4A at the NSLS, a Department of Energy facility, is supported by the Howard Hughes Medical Institute. Use of the APS was supported by the US Department of Energy, Basic Energy Sciences, Office of Energy Research, under Contract No. W-31-109-ENG-38. The Structural Biology Center is supported by the US Department of Energy, Office of Health and Environmental Research, Office of Energy Research, under Contract No. W-31-109-ENG-38. This work was supported in part by a Helen Hay Whitney Foundation Fellowship to CDL and by NIH grant GM34102 to WAH.

References

1. Seraphin, B. (1992). The HIT protein family: a new family of proteins present in prokaryotes, yeast and mammals. *J. DNA Seq. Map.* 3, 177–179.
2. Robinson, K. & Aitken, A. (1994). Identification of a new protein family which includes bovine protein kinase C inhibitor-1. *Biochem. J.* 304, 662–664.
3. Lima, C.D., Klein, M.G., Weinstein, I.B. & Hendrickson, W.A. (1996). Three-dimensional structure of human protein kinase C interacting protein 1, a member of the HIT family of proteins. *Proc. Natl. Acad. Sci. USA* 93, 5357–5362.
4. Brenner, C., *et al.*, & Lowenstein, J.M. (1997). Crystal structures of HINT demonstrate that histidine triad proteins are GalT-related nucleotide-binding proteins. *Nat. Struct. Biol.* 4, 231–238.

5. Ohta, M., *et al*., & Huebner, K. (1996). The FHIT gene, spanning chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. *Cell* **84**, 587–597.
6. Sozzi, G., *et al.*, & Croce, C.M. (1996). The *FHIT* gene at 3p14.2 is abnormal in lung cancer. *Cell* **85**, 17–26.
7. Sozzi, G., *et al.*, & Croce, C.M. (1996). Aberrant *FHIT* transcripts in Merkel cell carcinoma. *Cancer Res*. **56**, 2472–2474.
8. Negrini, M., *et al.*, & Croce, C.M. (1996). The *FHIT* gene at 3p14.2 is abnormal in breast carcinoma. *Cancer Res*. **56**, 3173–3179.
9. Virgilio, L., *et al.*, & Croce, C.M. (1996). *FHIT* gene alterations in head and neck squamous cell carcinomas. *Proc. Natl. Acad. Sci. USA* **93**, 9770–9775.
10. Panagopoulos, I., *et al.*, & Åman, P. (1996). The *FHIT* and *PTPRG* genes are detected in benign proliferative breast disease associated with familial breast cancer and cytogenetic rearrangements of chromosome band 3p14. *Cancer Res.* **56**, 4871–4875.
11. Mao, L., Fan, Y.-H., Lotan, R. & Hong, W.K. (1996). Frequent abnormalities of *FHIT*, a candidate tumor suppressor gene, in head and neck cancer cell lines. *Cancer Res.* **56**, 5128–5131.
12. Shridhar, R., *et al*., & Smith, D.I. (1996). Frequent breakpoints in the 3p14.2 fragile site, *FRA3B*, in pancreatic tumors. *Cancer Res*. **56**, 4347–4350.
13. Yanagisawa, K., *et al.*, & Takahashi, T. (1996). Molecular analysis of the *FHIT* gene at 3p14.2 in lung cancer cell lines. *Cancer Res*. **56**, 5579–5582.
14. Druck, T., *et al.*, & Huebner, K. (1997). Structure and expression of the human FHIT gene in normal and tumor cells. *Cancer Res.* **57**, 504–507.
15. Gartenhaus, R.B. (1997). Allelic loss determination in chronic lymphocytic leukemia by immunomagnetic bead sorting and microsatellite marker analysis. *Oncogene* **14**, 375–378.
16. Thiagalingam, S., *et al.*, & Vogelstein, B. (1996). Evaluation of the *FHIT* gene in colorectal cancers. *Cancer Res*. **56**, 2936–2939.
17. Geurts, J.M.W., Schoenmakers, E.F.P.M., Roijer, E., Stenman, G. & Van de Ven, J.M. (1997). Expression of reciprocal hybrid transcripts of *HMGIC* and *FHIT* in a pleomorphic adenoma of the parotid gland. *Cancer Res*. **57**, 13–17.
18. Robinson, A.K., Pena, C.E. & Barnes, L.D. (1993). Isolation and characterization of diadensosine tetraphosphate (Ap_4A) hydrolase from *Schizosaccharomyces pombe*. *Biochem. Biophys. Acta* **1161**, 139–148.
19. McLennan, A.G. (1992). *Ap4A and Other Dinucleoside Polyphosphates*. *CRC Press*.
20. Barnes, L.D., *et al.*, & Huebner, K. (1996). FHIT, a putative tumor suppressor in humans, is a dinucleoside 5′,5‴-P1,P3-triphosphate hydrolase. *Biochemistry* **36**, 11529–11535.
21. Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**, 51–58.
22. Hendrickson, W.A., Horton, J.R. & LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **9**, 1665–1672.
23. Westbrook, E. (1996). CCD-based area detectors. *Methods Enzymol.* **276**, in press.
24. Shapiro, L., *et al.*, & Hendrickson, W.A. (1995). Structural basis of cell-cell adhesion by cadherins. *Nature* **374**, 327–337.
25. Abrahams, J.P. & Leslie, A.G.W. (1996). Methods used in the structure determination of bovine mitochondrial F_1-ATPase. *Acta Cryst. D* **52**, 30–42.
26. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta. Cryst. A* **47**, 110–119.
27. Orengo, C.A. & Thornton, J.M. (1993). Alpha plus beta folds revisited: some favoured motifs. *Structure* **1**, 105–120.
28. Nicholls, A., Sharp, K.A. & Honig, B.H. (1991). Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Prot. Struct. Funct. Genet*. **11**, 281–296.
29. Brzoska, P.M., *et al.*, & Christman, M.F. (1995). The product of the ataxia-telangiectasia group D complementing gene, *ATDC*, interacts with a protein kinase C substrate and inhibitor. *Proc. Natl. Acad. Sci. USA* **92**, 7824–7828.
30. Brzoska, P.M., *et al.*, & Christman, M.F. (1996). Cloning, mapping, and *in vivo* localization of a human member of the PKCI-1 protein family (PRKCNH1). *Genomics* **36**, 151–156.
31. Robinson, K., Jones, D., Howell, S., Soneji, Y., Martin, S. & Aitken, A. (1995). Expression and characterization of maize ZBP14, a member of a new family of zinc-binding proteins. *Biochem. J.* **307**, 267–272.
32. MacDonald, J.R. & Walsh, M.P. (1985). Ca^{2+}-binding proteins from bovine brain including a potent inhibitor of protein kinase C. *Biochem. J.* **232**, 559–567.
33. MacDonald, J.R., Groschel-Stewart, U. & Walsh, M.P. (1987). Properties and distribution of the protein inhibitor (Mr 17,000) of protein kinase C. *Biochem. J.* **242**, 695–705.
34. Pearson, J.D., *et al*., & Walsh, M.P. (1990). Amino acid sequence and characterization of a protein inhibitor of protein kinase C. *J. Biol. Chem*. **265**, 4583–4591.
35. Fraser, E.D. & Walsh, M.P. (1991). The major endogenous bovine brain protein kinase C inhibitor is a heat-labile protein. *FEBS Lett.* **294**, 285–289.
36. Mozier, N.M., Walsh, M.P. & Pearson, J.D. (1991). Characterization of a novel zinc binding site of protein kinase C inhibitor-1. *FEBS Lett.* **279**, 14–18.
37. Holm, L. & Sander, C. (1997). New structure-novel fold? *Structure* **15**, 165–171.
38. Holm, L. & Sander, C. (1997). Enzyme HIT? *Trends Biochem. Sci.* **22**, 116–117.
39. Wedekind, JE., Frey, P.A. & Rayment, I. (1995). Three-dimensional structure of galactose-1-phosphate uridylyltransferase from Escherichia coli at 1.8 Å resolution. *Biochemistry* **34**, 11049–11061.
40. Wedekind, J.E., Frey, P.A. & Rayment, I. (1996). The structure of nucleotidylated histidine-166 of galactose-1-phosphate uridylyltransferase provides insight into phosphoryl group transfer. *Biochemistry* **35**, 11560–11569.
41. Pilkis, S.M., Walderhaug, M., Murray, K., Beth, A., Venkataramu, S.D., Pilkis, J. & El-Maghrabi, M.R. (1983). 6-phosphofructo-2-kinase/fructose 2,6-bisphosphatase from rat liver. *J. Biol. Chem*. **258**, 6135–6141.
42. Tauler, A., Lin, K. & Pilkis, S.J. (1990). Hepatic 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. *J. Biol. Chem*. **265**, 15617–15622.
43. Lin, K., Li, L., Correia, J.J. & Pilkis, S.J. (1992). Glu327 is part of a catalytic triad in rat liver fructose-2,6-bisphosphatase. *J. Biol. Chem.* **267**, 6556–6562.
44. Hasemann, C.A., Istvan, E.S., Uyeda, K. & Deisenhofer, J. (1996). The crystal structure of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase reveals distinct domain homologies. *Structure* **4**, 1017–1029.
45. Smith, R.H. & Johnson, K.S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**, 31–40.
46. Jankarik, J. & Kim, S.H. (1991). Sparse matrix sampling: a screening method for crystallization. *J. Appl. Cryst.* **24**, 409–411.
47. Otwinowski, Z. (1993). Data collection and processing. In *Proceedings of the CCP4 study weekend.* (Sawyer, L., Isaccs, N. & Bailey, S. eds), pp. 56–62, SERC Daresbury Laboratory, Warrington, UK.
48. Westbrook, M.L, Coleman, T.A., Daley, R.T. & Pflugrath, J.W. (1996). Data acquisition and analysis at the Structural Biology Center. In *Proceedings of IUCr Computing School*, in press.
49. CCP4 SERC (UK). (1979). Collaborative Computing Project 4 Daresbury Laboratory, Warrington, UK.
50. Navaza, J. (1984). AMoRe: an automated package for molecular replacement. *Acta. Cryst. A* **50**, 157–163.
51. Brunger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* **235**, 458–460.
52. Brunger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–474.
53. Evans, S.V. (1993). SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graphics* **11**, 134–138.
54. Huang, Y., Garrison, P.N. & Barnes, L.D. (1995). Cloning of the *Schizosaccharomyces pombe* gene encoding diadenosine 5′,5‴-P^1,P^4-tetraphosphate (Ap_4A) asymmetrical hydrolase: sequence similarity with the histidine triad (HIT) protein family. *Biochem. J.* **312**, 925–932.
55. Fleischmann, R.D., *et al*., & Venter, J.C. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512.
56. Bult, C.J., *et al*., & Venter, J.C. (1996). Complete genome sequence of the Methanococcus archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058–1073.
57. Bustos, S., Schaefer, M.R. & Golden, S.S. (1990). Different and rapid response of four cyanobacterial psbA transcripts to changes in light intensity. *J. Bacteriol*. **172**, 1998–2004.