Crystal structure of the worm NitFhit Rosetta Stone protein reveals a Nit tetramer binding two Fhit dimers

H.C. Pace*, S.C. Hodawadekar*, A. Draganescu*†, J. Huang*‡, P. Bieganowski*, Y. Pekarsky§, C.M. Croce§ and C. Brenner*

Background: The nucleotide-binding protein Fhit, among the earliest and most frequently inactivated proteins in lung cancer, suppresses tumor formation by inducing apoptosis. In invertebrates, Fhit is encoded as a fusion protein with Nit, a member of the nitrilase superfamily. In mice, the *Nit1* and *Fhit* genes have nearly identical expression profiles. According to the Rosetta Stone hypothesis, if the separate Nit and Fhit genes could be shown to occur in the same subset of genomes (that is, to share a phylogenetic profile), then the existence of a fusion protein in invertebrates and the coordinated expression of separate mRNAs in mouse suggest that Nit and Fhit function in the same pathway and that the structure of invertebrate NitFhit may reflect the nature of Nit–Fhit interactions.

Results: To satisfy the phylogenetic profile criterion for functional significance of protein fusion events, we cloned additional Nit homologs from organisms with Fhit homologs. We used fluorescent nucleotide analogs of AppppA to follow the purification and to characterize the nucleotide specificity of NitFhit from *Caenorhabditis elegans,* crystallized the 200 kDa tetrameric complex, and solved the structure of NitFhit from a single mercury derivative phased by two-wavelength anomalous diffraction.

Conclusions: Nit monomers possess a new α–β–β–α sandwich fold with a presumptive Cys–Glu–Lys catalytic triad. Nit assembles into a tetrameric, 52-stranded beta box that binds Fhit dimers at opposite poles and displays Nit active sites around the middle of the complex. The most carboxy-terminal β strand of each Nit monomer exits the core of the Nit tetramer and interacts with Fhit. Residence in the NitFhit complex does not alter the nucleotide specificity of Fhit dimers, which are oriented with ApppA-binding surfaces away from Nit.

Addresses: *Structural Biology and Bioinformatics Program, and §Genetics and Molecular Biology Program, Kimmel Cancer Center, Thomas Jefferson University, 233 S Tenth Street, Room 826, Philadelphia, Pennsylvania 19107, USA.

Present address: †Unilever Research, Edgewater, New Jersey 07020, USA. ‡Harvard University, Cambridge, Massachusetts 02138, USA.

Correspondence: C. Brenner
E-mail: Charles.Brenner@tju.edu

Background

Loss of Fhit protein is a frequent and early event in the development of lung cancer, the leading cause of cancer deaths worldwide [1–4]. The *FHIT* gene is located at chromosomal position 3p14.*2* [5] and spans FRA3B [6], the most fragile site in the human genome. Instability of the 1.5 Mb *FHIT* locus [7], coupled with the small size of the transcript and coding region [5], account for inactivation of the gene principally by deletion [7] and, less frequently, by methylation [8], rather than by point mutation. Mutations in the first *FHIT* allele are either inherited as a t(3;8) translocation [5,9] or are acquired by exposure to tobacco carcinogens [3], papilloma virus insertion [10], or other mechanisms. In families carrying a t(3;8) translocation, affected young adults suffer bilateral, multifocal renal carcinomas [11]. Somatic loss of Fhit in humans is associated with cancers in a wide variety of sites, including lung [1–4], kidney [12,13], stomach [14], pancreas [15,16], cervix [10], ovary [17], head and neck [18,19], breast [20,21], and hematopoetic cells [22–24]. In addition, loss of Fhit is found in murine cancer cell lines [25], and targeted disruption of murine *Fhit* results in predisposition to stomach and sebaceous tumors that resemble human Muir–Torre syndrome [26].

Fhit is a member of the histidine triad (HIT) superfamily of nucleotide-binding proteins [27,28]. Members of the Fhit branch of the HIT superfamily bind and cleave diadenosine polyphosphates (Ap_nA) such as ApppppA and ApppA to generate AMP plus ATP and ADP, respectively [29,30]. The tumor-suppressing function of Fhit does not depend on cleavage of Ap_nA [31]. The H96N allele of Fhit, which maintains micromolar binding to ApppA at the expense of a million-fold loss in catalytic activity [32], is functional in tumor suppression [31,33]. Fhit binds two Ap_nA substrates per dimer, presenting all of the phosphates and two adenosines on a surface of the protein that is spatially and electrostatically altered in the substrate-bound form [32]. Thus, by analogy with G proteins, Fhit has been proposed to function as a nucleotide-substrate-dependent molecular switch [28,32].

Re-expression of Fhit in cancer cell lines with *FHIT* deletions induces apoptosis [34,35] through an unknown mechanism. Thus, identification of molecules that interact with Fhit and/or participate in Fhit-dependent pathways is of great interest. Recently, a general method was proposed to identify interacting proteins by identifying a 'Rosetta Stone' protein consisting of two unrelated proteins fused in one organism but expressed as separate polypeptides in other organisms [36]. With few exceptions, experimental evidence and bioinformatic inference suggest that the existence of a fusion protein in one genome powerfully predicts that the separate polypeptides function in the same cellular or biochemical pathway in other organisms [36,37]. The strongest case that Rosetta Stone proteins decode real interactions can be made when the separate genes have similar gene expression patterns [38] and are found in the same subset of genomes (that is, share a phylogenetic profile) [39].

In mammals [5,25,40] and fungi [27,29], Fhit homologs are encoded as single polypeptides that are at least 42% identical within a core of 113 residues. In flies and worms, Fhit-homologous domains are encoded at the carboxyl termini of 460 and 440 amino acid polypeptides in which the ~300 amino acid amino-terminal domains are 22% identical to plant and bacterial nitrilases [41], enzymes that hydrolyze compounds such as indoleacetonitrile to indoleacetic acid plus ammonia [42]. Using Nit domains of fly and worm NitFhit as search molecules, we identified murine and human orthologs encoded as separate polypeptides that are 48% identical to invertebrate Nit domains [41]. We refer to this branch of the nitrilase superfamily as Nit proteins.

Satisfying the first criterion for the likely functional significance of natural fusion proteins [38], we showed that levels of *Nit1* and *Fhit* mRNA are highly correlated in seven of eight tissues examined in mouse, the exception being brain, which has a high level of *Fhit* and a low level of *Nit1* message [41]. In the present work, to address the second criterion [38], we cloned an additional Nit protein from human and mouse, a Nit homolog from frog, two homologs from budding yeast, and identified two Nit sequences from fission yeast. Thus, Nit homologs, having been identified in vertebrate and invertebrate animals and fungi (Figure 1), cover the same phylogenetic space as Fhit homologs [27,28].

Further, we report the purification of worm NitFhit by following the GpppBODIPY (Gppp-S-(4-4-fluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacine-3-yl)methylamino-acetyl) hydrolysis activity [43] of its Fhit active site. We characterized the nucleotide specificity of the Fhit active site, establishing that NitFhit prefers ApppppA > ApppA > ApppppA > pyrophosphate > other compounds. Finally, we determined the crystal structure of NitFhit, defining a

new α–β–β–α sandwich protein fold for the nitrilase superfamily. Nit possesses a novel tetrameric superstructure, termed a beta box, that recognizes a pair of Fhit dimers at opposite poles. Nit and Fhit domains are not merely tethered together in the fusion protein. In contrast, the most carboxy-terminal β strand encoded by the Nit portion of the NitFhit sequence extends out of the Nit globular domain and binds Fhit.

Results

Nit homologs are found in the same organisms as Fhit homologs

In the course of cloning Fhit-homologous cDNAs from *Drosophila melanogaster* and *Caenorhabditis elegans*, we identified NitFhit sequences [41]. The Nit domain of the invertebrate NitFhit proteins was classified as a distinct member of the nitrilase superfamily and used to clone the single most homologous sequence from a human and from a mouse cDNA library [41]. It has been pointed out that events that fuse unrelated proteins [36,37] are most likely to be functionally significant [38] if the separate proteins have similar gene expression patterns and have similar phylogenetic profiles [39]. At the level of tissue specificity, murine *Fhit* and *Nit1* have nearly identical mRNA accumulation patterns [41]. We therefore sought to identify Nit-related genes from divergent organisms known to contain Fhit-homologous genes, namely *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* [27] and *Xenopus laevis* (our unpublished data). In each yeast, two Nit-related sequences were identified (Figure 1) as well as sequences related to plant nitrilases (our unpublished data). The frog also yielded a Nit sequence and further examination of human and murine expressed sequence tag databases allowed us to identify a second Nit coding sequence from human and mouse (Figure 1). Nit sequences have a low level of identity with nitrilases and a substantial level of identity with each other. Nit homologs, having been found fused or coordinately expressed with Fhit homologs [25] and in the same organisms as Fhit homologs (Figure 1), are reasonable candidates for proteins that interact with Fhit homologs.

Characterization of the Fhit active site of NitFhit

Because of the insolubility of *D. melanogaster* NitFhit expressed in *Escherichia coli* [41], we chose to express worm NitFhit, the product of the *nft-1* gene of *C. elegans* [41]. The enzyme was followed fluorimetrically with GpppBODIPY [43], a quenched fluorescent nucleotide substrate developed for use with Fhit. Worm NitFhit was ~20% soluble when expressed in *E. coli* at 20°C. Our conventional purification of NitFhit, based on maximizing GpppBODIPY-hydrolase specific activity, was performed exclusively from the soluble fraction (Table 1).

The function of Fhit is thought to depend on formation of substrate complexes with Ap_nA [32] in the presence of

Figure 1

Sequence alignment of Nit domains with a plant nitrilase. The Nit domains of *C. elegans* (Ce) and *D. melanogaster* (Dm) NitFhit proteins are aligned with Nit homologs from *Homo sapiens* (Hs), *Mus musculus* (Mm), *S. pombe* (Sp), *S. cerevisiae* (Sc), *X. laevis* (Xl), and nitrilase1 from *Arabidopsis thaliana* (At). Human Nit2, murine Nit2, frog Nit1, and budding yeast Nit2 and Nit3 are newly cloned and have been deposited in GenBank with accession numbers AF284574, AF284573, AF284575, AF284571 and AF284572, respectively. Secondary structural elements and sequence numbers correspond to worm NitFhit. Residues shaded purple are identical in all aligned sequences. Residues shaded yellow are identical to those of the worm Nit domain. Triangles, positions of insertions found in some of the sequences (the numbers indicate the size of the insertions); filled circles, residues found in the vicinity of Cys169. The figure was prepared using ALSCRIPT [58].

Current Biology

higher cellular concentrations of purine mononucleotides and other competitors [43]. To assess the nucleotide specificity of the worm Fhit homolog, we performed a series of assays with ApppBODIPY and GpppBODIPY. Titration of non-labeled nucleotides and related compounds into fluorescent nucleotide-hydrolysis assays allows determination of the binding constant of each compound for the Fhit active site [43]. As shown in Table 1, the nucleotide specificity of the Fhit domain of NitFhit was similar to that of Fhit. Whereas human Fhit has a slight binding preference for ApppA over ApppppA [43], the worm enzyme, like the homolog from fission yeast [44], preferred ApppppA. As is the case for the human enzyme, pyrophosphate competed for the Fhit active site of NitFhit more effectively than did purine mononucleotides. After ApppA and ApppppA, both enzymes preferred ApppppppA > pyrophosphate > AMP and ATP–αS > GTP–αS > monophosphate. The single most important measure of an enzyme's activity on a substrate,

k_{cat}/K_m, was measured for ApppBODIPY and GpppBOD-IPY in substrate-decay assays and initial-rate assays, respectively [43]. Although worm NitFhit displayed only 22% of the activity of human Fhit on ApppBODIPY, it displayed 109% of the activity of human Fhit on Gppp-BODIPY. Thus, the Nit domain of NitFhit does not inhibit the nucleotide-binding or hydrolysis activity of the associated Fhit domain.

Nit is a novel α–β–β–α sandwich protein

To determine the structure of Nit and the nature of Nit–Fhit interactions, worm NitFhit was crystallized and its crystal structure was determined. The 440 amino acid polypeptide (molecular weight 49,936 Da) had a molecular weight of 200,000 Da in solution (see Materials and methods) and crystallized with a monomer in the asymmetric unit in space group I222. The symmetry of these crystals suggested that NitFhit tetramers were located at the origin and center of the unit cell and that their

Table 1

Purification and characterization of the Fhit active site of NitFhit.

Purification of NitFhit via the Fhit active site

Fraction	Total protein (mg)	Units (pmol min⁻¹)	Specific activity (pmol min⁻¹ mg⁻¹)	Cumulative yield (%)	Purification (cumulative fold)
Cleared lysate	764	8.00 E7	1.05 E5	100	–
Ammonium sulfate	451	5.60 E7	1.24 E5	70	1.2
HQ	13.5	1.61 E7	1.19 E6	20	11.3
CM	8.9	2.12 E7	2.38 E6	27	22.7

Nucleotide specificity of the Fhit active site of NitFhit

	k_{cat}/K_M (10^5 s⁻¹ M⁻¹)	k_{cat} (s⁻¹)	K_M (µM)	K_I (µM)
ApppBODIPY	5.0 ± 0.1			
GpppBODIPY	6.3	2.4	3.7	
ApppA			4.2 ± 0.4	
ApppppA			2.7 ± 0.3	
ApppppA			4.5 ± 0.7	
Pyrophosphate				78 ± 12
AMP				153 ± 13
ATP-αS			287 ± 73	
GTP-αS			481 ± 27	
Monophosphate				2660 ± 490

Table 2

X-ray data collection, phasing and refinement statistics.

Data collection and phasing of $P2_12_12$-thimerosal crystals

	λ_{edge}	λ_{remote}
Resolution (Å)	30.0–2.8	30.0–2.8
Completeness (%) (outer shell)	97.9 (91.4)	97.6 (90.3)
Multiplicity	7.5	7.0
I/σ	24.6	23.2
R_{sym} (%) (outer shell)	4.5 (8.8)	5.1 (9.9)
R_{anom} (%) (outer shell)	3.8 (6.2)	5.0 (7.5)
Hg sites	5	
Phasing power	1.95	
Figure of merit	0.47	
Figure of merit after density modification	0.96	

Refinement statistics

Non-hydrogen atoms (water molecules)	6708 (161)
Unique reflections (free)	49,386 (3315)
R_{work} (%) (R_{free})	19.0 (23.1)
Rmsd bond lengths (Å)	0.007
Rmsd bond angles (°)	1.4
Average B value (Å²)	28.5

The unit cell measured a = 68.74 Å, b = 100.44 Å, c = 158.65 Å. The λ_{edge} was 1.008989 and λ_{remote} was 0.992782. The f′ and f″ scattering factors refined from reference values [57] of −17.79 e− and 7.24 e− to −13.49 e− and 9.41 e− for λ_{edge} and −10.86 e− and 9.99 e− to −10.86 e− and 11.18 e− for λ_{remote}. $R_{sym} = \Sigma |I - <I>| / \Sigma <I>$ in which I is a measured intensity and <I> is the average intensity from multiple measurements of symmetry-related reflections. $R_{anom} = \Sigma |<F+> - <F->/ \Sigma <F>$ in which <F+> and <F−> are the average structure factors of Friedel pairs. Phasing power, figure of merit, and figure of merit after density modification were as defined in CNS [54] for anomalous difference phasing.

oligomeric symmetry consisted of three mutually perpendicular twofold rotation axes. These crystals were large and single but diffracted weakly to no better than 3.5 Å resolution at synchrotron sources. Crystals soaked with thimerosal or ethylmercuric phosphate exhibited a reduction in crystallographic symmetry to $P2_12_12$, a concomitant increase in the size of the asymmetric unit to two monomers, a 6% reduction in the length of one unit-cell length, and a corresponding 2% reduction in solvent content. These derivatized crystals showed a striking increase in reflection intensities and resolution. Diffraction data from a single thimerosal-soaked crystal, collected at the mercury absorption edge and at one remote wavelength, were used to solve the structure of mercurated NitFhit by two-wavelength anomalous diffraction (TAD) phasing to 2.8 Å resolution (Table 2). The two non-identical NitFhit monomers were built from a density-modified TAD electron density map (Figure 2a) and refined independently.

By sequence alignment, the Nit domain of NitFhit spans from residue 1 through 296 (see reference [41] and Figure 1) and the Fhit domain spans from residue 297–440 [27,28,41]. The Nit domain, defined by continuous electron density from residue 10 to its carboxyl terminus, is a novel protein fold (Figure 2b) consisting of five α helices designated NH1 to NH5 and 13 β strands designated NS1 to NS13. In CATH (class, architecture, topology, homologous superfamily) nomenclature [25], Nit can be assigned to the α–β class and the four-layer sandwich architecture and is the first of its kind in topology and

Figure 2

Structure determination of NitFhit.
(a) A portion of the 2.8 Å experimental
electron density map in stereo. The map was
contoured at 1.5 σ and superimposed on the
refined atomic model. **(b)** Stereo ribbon view
of the Nit domain of a NitFhit monomer
(northern conformation). Secondary structural
elements are indicated. The Fhit domain,
carboxy terminal to the Nit domain, and three
additional subunits of the NitFhit tetramer are
not shown. N, amino terminus; C, carboxyl
terminus of the Nit domain.
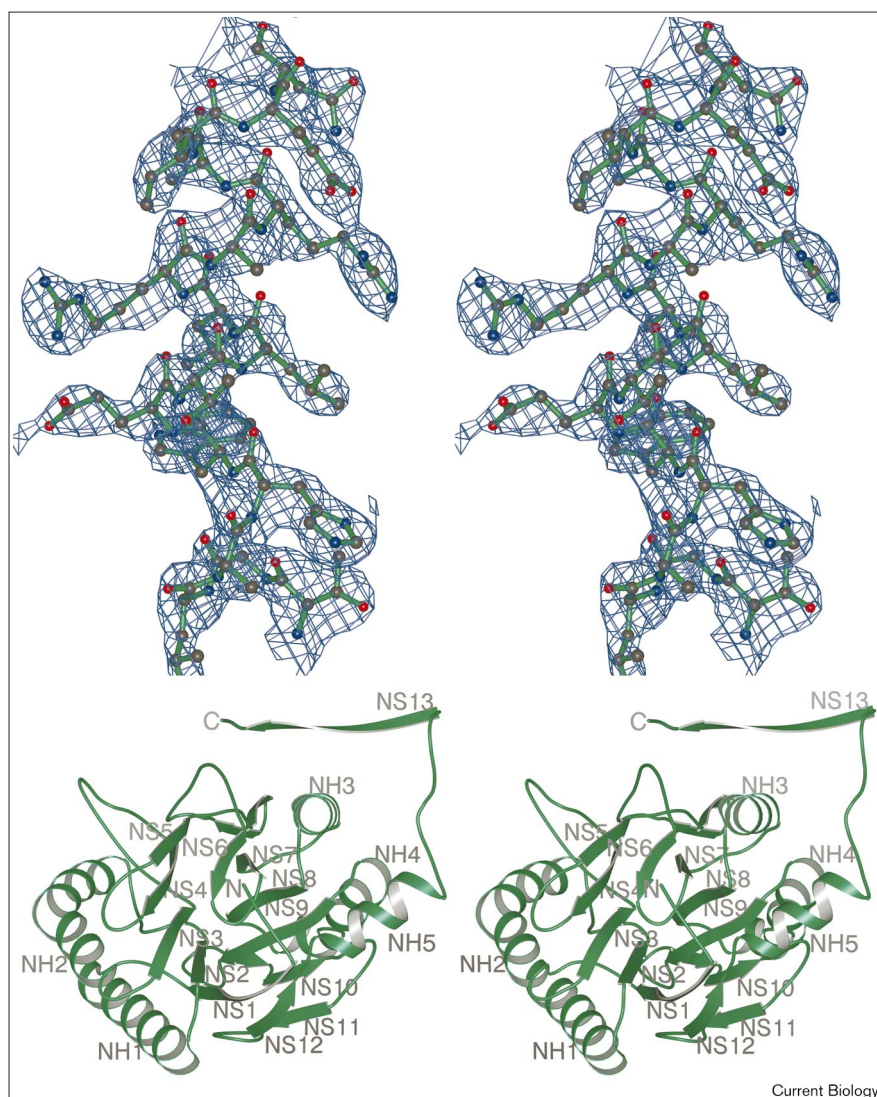
Current Biology

superfamily. The core of Nit is a highly regular α–β–β–α
sandwich structure containing helices NH1 through NH4
and strands NS1 through NS12 (Figures 2b and 3a). A
cross section of the Nit core reveals a layer containing two
α helices, followed by two layers of six β strands, followed
by a layer of two α helices. The most similar of the 12-
stranded α–β–β–α sandwich folds is that of DNase I [45]
and related nucleases. The α–β–β–α sandwich of DNase I
is topologically distinct from that of Nit and is unlikely to
be related, however. The pattern and direction of the first
eight elements of the Nit core (NS1, NH1, NS2, NH2,
NS3, NS4, NS5, NS6 with amino termini of NS1, NS2,
NS3, NS5 and carboxyl termini of NH1, NH2, NS4 and
NS6 facing the viewer in Figure 2b) are repeated by the
second eight elements (NS7 through NS12) by an internal
pseudo-twofold rotation axis. Carboxy terminal to NS12,
Nit contains helix NH5 and strand NS13 orthogonal to

the core. NS13, extended away from the globular Nit
core, makes extensive interactions with the Fhit domain
as discussed below.

At residue 297, the NitFhit polypeptide aligns with
residue 1 of human Fhit. Human Fhit structures are
defined for residues 2–106 and 128 to their carboxyl
termini at residue 147 [32,46]. The Fhit domain of worm
Fhit contains the seven β strands, FS1 through FS7, and
the two α helices, FH1 and FH2, of Fhit and is nearly
identical to human Fhit in all respects. Refined NitFhit
models contain a 20-residue gap in the same location as
the 21-residue gap within human Fhit models. The largest
root mean square differences between superimposed
human Fhit and worm Fhit domain are in the loop
between FH1 and FS6. Even there, aligned Cα positions
differ by 2 Å or less.

Figure 3
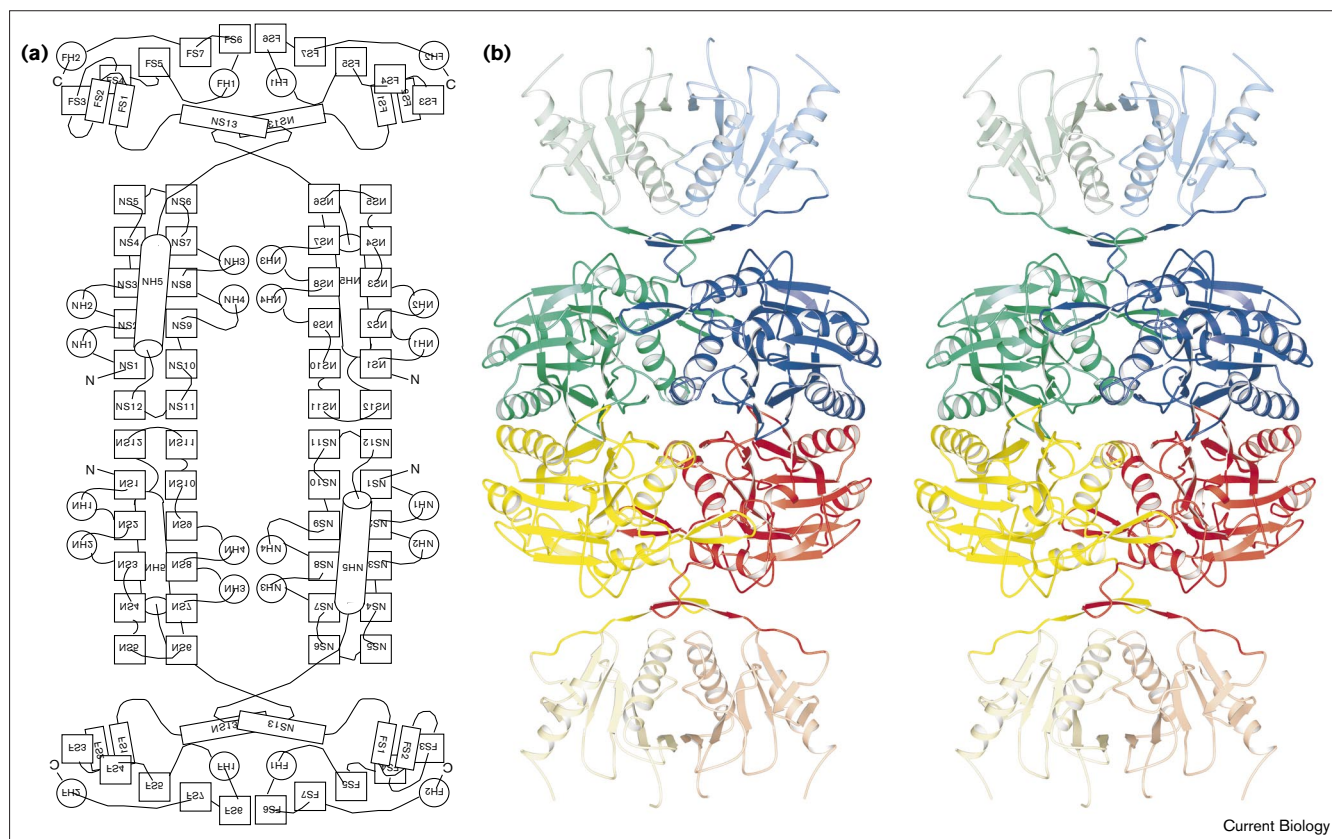
Structure of the NitFhit tetramer. **(a)** Levitt and Chothia [59] representation of the NitFhit tetramer. The 222 point symmetry of the tetramer is indicated by the manner in which labels are flipped across symmetry axes. N, amino terminus; C, carboxyl terminus. **(b)** Stereo ribbon representation of the NitFhit tetramer. NitFhit monomers are colored green and blue in the northern hemisphere, red and yellow in the southern hemisphere. Nit domains (residues 10–296) are in bold colors and Fhit domains (residues 297–440) are in pastel colors. The boldly colored elements in the northern and southern domains correspond to the carboxy-terminal Nit β strand NS13, which is physically located in Fhit domains.

Nit tetramers form a 52-stranded beta box

Nit monomer domains assemble into a tetramer that contains two types of homotypic Nit–Nit interfaces termed north–south and east–west (Figure 3). Heavy atom binding locked the northern and southern hemispheres into different conformations but did not disturb their perfect east–west symmetry. Fhit dimers are located at the north and south poles of the Nit tetramer.

The east–west dimer interface of Nit, parallel to an already extensive Fhit dimer interface formed by FH1 and FS6, is formed by a four-helix bundle (NH3 and NH4 with their symmetry mates) and an anti-parallel β-interaction mediated by NS13 (Figure 3). Thus, the east–west Nit dimer turns a four-layered α–β–β–α sandwich into an eight-layered α–β–β–α–α–β–β–α sandwich. Helices NH1 and NH2 are solvent-exposed on the external layers of the sandwich. Furthermore, one edge of the β-sheets is enclosed by NH5 while the other edge of the β-sheets is exposed to solvent. According to amino-acid conservation detected in our alignment (Figure 1), we would expect all Nit proteins to form α–β–β–α–α–β–β–α dimers.

The north–south Nit interface is formed by anti-parallel, homotypic β interactions involving strands NS11 and NS12 (Figure 3). These interactions double the width of the four β sheets in the α–β–β–α–α–β–β–α sandwich from six strands north to south to 12 strands north to south. The tetrameric Nit assembly can be termed a 52-stranded beta box. The east and west sides of the beta box each consist of two 12-stranded β sheets. Between the east and west sides are two four-helix bundles. The north and south poles of the beta box are capped by the final β strand (NS13) of each monomer as an anti-parallel pair of strands on each pole. By alignment, some of the salt bridges that stabilize the north–south dimer interface appear to be absent in homologous Nit sequences. Thus, it remains to be seen whether vertebrate and fungal Nit proteins will be dimers or tetramers.

Figure 4
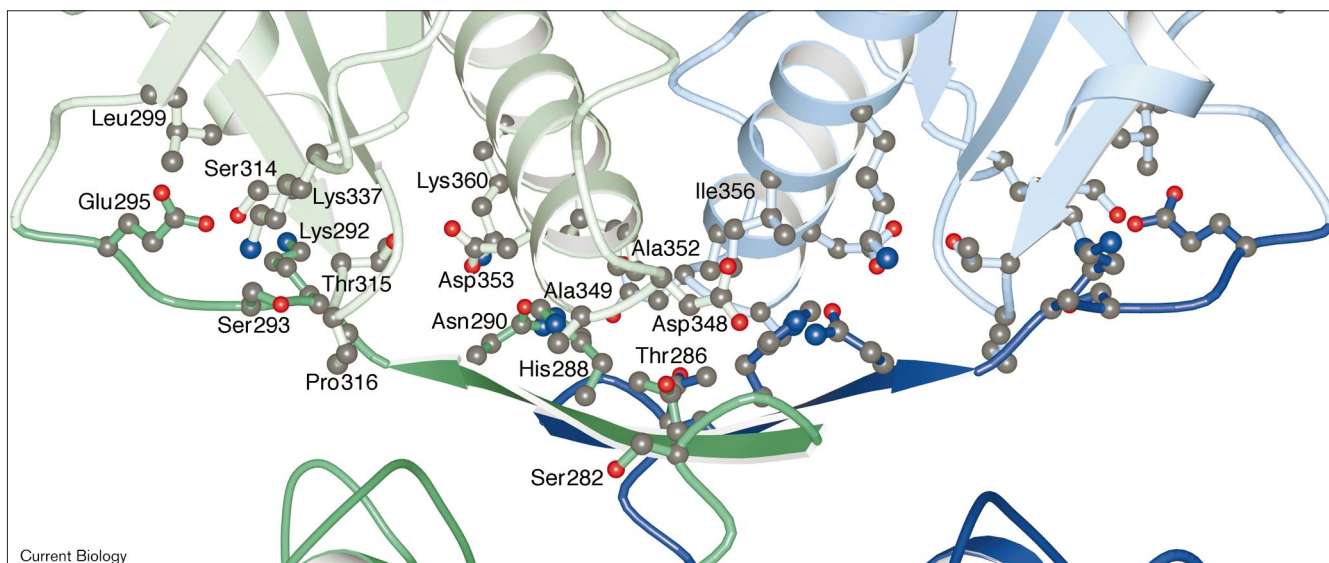
Nit fits. Molecular features of the physical interaction between Nit and Fhit. At the north and south poles of the Nit tetramer, pairs of anti-parallel NS13 Nit strands (bold colors) interact with Fhit domains (pastel colors) beneath pairs of anti-parallel FH1 helices. Side chains are shown at the interface.

Nit sequences bind Fhit

The carboxy-terminal β strand encoded by Nit sequences, NS13, exits the Nit core domain (Figure 2b). Formation of the east–west Nit dimer allows NS13 strands to pair in an anti-parallel fashion, and formation of the north–south Nit tetramer allows these strands to form the top and bottom sides of the beta box (Figure 3). Moreover, in worm NitFhit, the NS13 elements have extensive interactions with Fhit dimers and appear to be physically part of Fhit dimer domains rather than the Nit tetramer (Figures 3 and 4). Nearly all of the interactions between Nit and Fhit are mediated by binding of anti-parallel NS13 strands to the anti-parallel FH1 helices at the bottom of Fhit dimers. The two Nit–Fhit interaction surfaces are extensive, at 1080 Å each, but appear to be more reversible than the east–west interface (7300 Å^2, including 4900 Å^2 of Nit–Nit interactions and two 1200 Å^2 patches of Fhit–Fhit interactions) or the north–south interface (2350 Å^2) within the NitFhit tetramer. Consistent with biochemical data (Table 1), the Nit tetramer does not interact with the nucleotide-binding surface of Fhit dimers. In contrast, the Nit tetramer binds Fhit in a manner that presents nucleotide-binding surfaces of Fhit [32] at the two extreme poles of the complex, potentially for interaction with Fhit effectors.

## Plasticity in Nit and a candidate Nit active site

The differences in protein conformation between the northern and southern molecules are localized to NH2 (Figure 5). Ethylmercury bound to Cys75 in the northern but not the southern chains. Without ethylmercury bound, residues 70 through 76 are not helical. Upon binding of ethylmercury, NH2 becomes helical but bent at residue 75. Though it is not obvious what sequence feature in worm NH2 disrupts helicity, NH2 is unique among the Nit helices in that homologs have insertions, deletions and

Figure 5
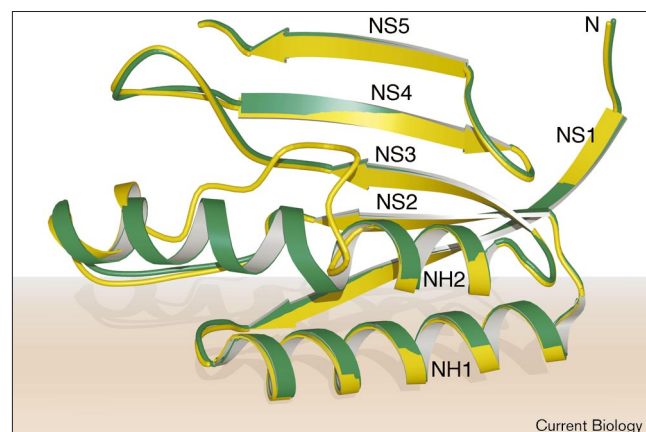
Structural plasticity in Nit tetramers. Two conformations of the second Nit helix, NH2. Portions of two non-identical Nit domains were superimposed to show two conformations at a solvent-exposed surface of the NitFhit tetramer. Without mercury binding, the segment is not helical. With mercury bound, the helix is bent. Both conformations distort helix NH2 at a site in which other Nit sequences, but not worm NitFhit, contain Gly and/or Pro residues.

Figure 6
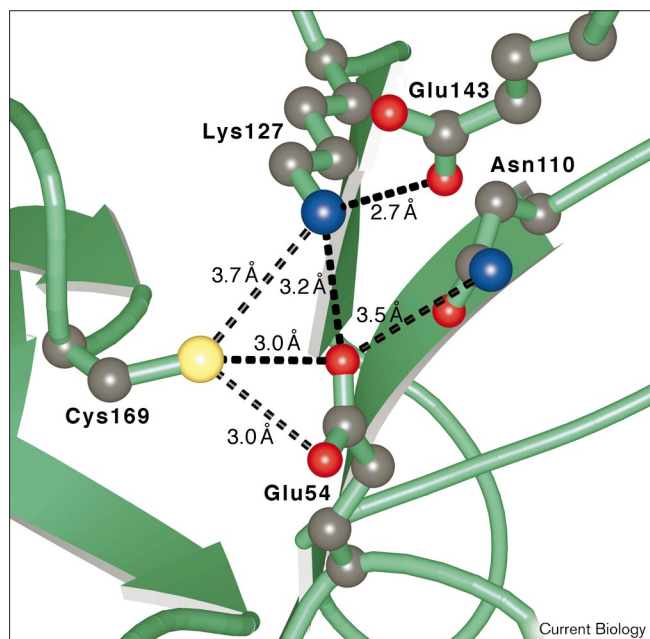
A putative Nit active site. The region around Cys169, a residue conserved in nitrilases, is conserved in Nit proteins. Residues aligning with Cys169, Glu54 and Lys127 are predicted to form a catalytic triad in the nitrilase superfamily.

Gly and Pro substitutions in this region (Figure 1). Thus, the disrupted nature of the NH2 helix appears to be a conserved feature that may be functionally important.

Nitrilases are thiol enzymes that attack the cyano carbon of nitriles (R–C≡N) to form a covalent thioimidate complex [47]. Addition of one water molecule is accompanied by release of ammonia and transformation of the planar thioimidate to a planar thiol acylenzyme through a tetrahedral intermediate. Addition of a second water molecule would allow the acid product to leave and regenerate the enzyme [47]. A related family of aliphatic acid amidases also uses the conserved cysteine to acylate and release ammonia from acid amides (R-CONH$_2$) [48]. Cys169, which aligns with the conserved Cys of nitrilases and amidases, is located on the solvent exposed face of the Nit β sheet and was modified by ethylmercury (Figure 6). Only 3.0 Å and 3.7 Å from Cys169, we located Glu54 and Lys127, both conserved in nitrilases. In nitrilases, the corresponding residues may function as a catalytic triad with Glu acting as the general base for the thiol.

Discussion

According to the theory of Rosetta Stone proteins, proteins that engage in fusion events are expected to jointly participate in a biochemical or cellular pathway and/or to physically interact [36]. Nonetheless, the simplest expectation about NitFhit would have been that homotypic

Fhit interactions would drive dimerization of NitFhit and that Nit would neither be multimerized nor bound to Fhit. Our results prove, however, that NitFhit is a stable tetramer that displays Fhit dimers on opposite poles and Nit active sites around the equator, and is held together by extensive homo- and hetero-oligomeric interactions. Strikingly, the most carboxy-terminal β strand of Nit polypeptide sequences exits the tetrameric Nit domain and binds Fhit dimer domains. Fhit dimer domains are bound with their nucleotide-binding surfaces [32] facing away from Nit. Thus, the Rosetta Stone hypothesis implicates Nit as a candidate Fhit-interacting protein and the crystal structure of NitFhit proves that Fhit can assemble into a 200 kDa complex with Nit in a manner that potentially allows two Fhit dimers to interact with effectors.

It was hypothesized that proteins with low-affinity heterotypic interactions that function in the same process might provide a selection for fusion events [36]. Upon fusion, the local concentration of the binding partner is greatly increased from that of separate polypeptides. The structure of NitFhit supports the view that Nit and Fhit function in the same pathway and that a reversible interaction between mammalian Nit and Fhit would depend on the carboxy-terminal β strand of Nit, NS13.

Though substrates for animal Nit proteins have not been identified, the structure of NitFhit is the basis for prediction of a Cys–Glu–Lys catalytic triad in the nitrilase superfamily. The nitrilase from *Rhodococcus rhodocrous* J1 has been purified and tested for pH dependence of benzonitrile hydrolysis. Consistent with Glu functioning as a general base, recombinant *R. rhodocrous* J1 nitrilase showed no pH dependence between pH 5.5 and 10.0, which were the limits of the enzyme's physical stability (S.K. Milano, M. Schimerlik and C.B., unpublished data).

Cancer cells that are deficient in Fhit are defective in programmed cell death [34,35], but the point of action of Fhit in apoptosis is unclear. Three well-known signals for cell-cycle arrest and programmed cell death, namely contact inhibition of growth [49], interferons [50] and etoposide [51], induce synthesis of diadenosine polyphosphates, the likely positive regulators of the cellular activity of Fhit [32]. Although Fhit is likely to function in an animal cell-death pathway, identification of Fhit and Nit proteins in fungi suggests that these proteins have a fundamental role in maintaining the differentiated states of single cells. Because indole-3-acetonitrile is not a substrate of animal Nit proteins (data not shown), animal Nit proteins may function by producing novel regulatory compounds that are not auxin. By mutating residues in the putative active site of yeast or animal Nit proteins, it may be possible to trap Nit substrates and use genetic approaches to discover the cellular consequences of Nit activity and inactivity, Nit–Fhit heteromultimerization, and the cellular targets of this pathway.

Materials and methods

Protein expression and purification

The *C. elegans* NitFhit cDNA was amplified using primers that generated an *Nde*I site at the initiator codon and an *Xho*I site 3′ of the stop codon. After restriction with *Nde*I and *Xho*I, the fragment was ligated to plasmid pSGA02 [52] that had been digested with the same enzymes. Protein was expressed in *E. coli* strain ER2566 (New England Biolabs) and grown in LB with 150 µg ml⁻¹ ampicillin; 1 L cultures, shaken at 24°C in 2800 ml Fernbach flasks, were induced with 0.4 mM IPTG at an optical density ($\lambda = 600$ nm) of 0.4 and aerated for 9 h. Steps before and including ammonium sulfate precipitation were performed at 0–4°C. Frozen cell pellets were resuspended in 50 mM NaHEPES pH 7.0, 5 mM DTT, 10% glycerol, 0.5 mM PMSF, 2 µg ml⁻¹ leupeptin and 3.4 µg ml⁻¹ pepstatin, and lysed by sonication. Cleared lysate was subjected to protamine sulfate precipitation followed by centrifugation to remove nucleic acids. A 20–60% ammonium sulfate fraction was obtained, and resuspended and dialyzed into 50 mM NaHEPES pH 7.2, 5 mM DTT, 10% glycerol. The dialysate was loaded onto a 58 ml POROS 20 HQ column (PE Biosystems) with 50 mM NaHEPES pH 7.2, 5 mM DTT, 2% glycerol as running buffer and NaCl as eluant. GpppBODIPY hydrolysis activity eluted at 0.2 M NaCl. Peak fractions were pooled and buffer was exchanged to 25 mM NaHEPES pH 7.0, 5 mM DTT, 2% glycerol using 10,000 Da-retaining Ultrafree filters (Millipore). Concentrated and desalted sample was loaded onto a 1.7 ml POROS 20 CM column (PE Biosystems) and chromatographed as before, with the peak of total protein eluting with enzymatic activity at 0.25 M NaCl. Purified NitFhit was concentrated to 7 mg ml⁻¹ in 10 mM NaHEPES pH 7.0, 50 mM NaCl, 5 mM DTT, microaliquoted, and stored at −80°C. Data from Edman degradation, mass spectrometry and analytical ultracentrifugation indicated that purified NitFhit contains an intact amino-terminal Met, displays mass/charge ratios consistent with the predicted monomer size of 49,936 Da, and exists as a stable 200,000 Da tetramer in solution, independent of nucleotide occupancy.

Characterization of the Fhit active site of NitFhit

Standard activity assays of the Fhit active site of NitFhit used GpppBODIPY in an initial rate assay as developed for Fhit [43] except that reactions were initiated by addition of 10–50 ng total protein diluted into 20 mM NaHEPES pH 7.0, 10% glycerol, 5 mM DTT, 0.2 mg ml⁻¹ BSA and incubated at 21°C. Determination of k_{cat}/K_M with ApppBODIPY was performed as a substrate-decay assay [43] with 1.5 µM ApppBODIPY, initiated by 0.15 pmol NitFhit. For GpppBODIPY, substrate concentration was titrated from 40 µM to 2.5 µM, and k_{cat} and K_M were determined from initial rates using 0.5 pmol enzyme [43]. Assays of non-labeled compounds as competitive inhibitors of ApppBODIPY hydrolysis were also performed as developed for human Fhit. K_M values for ApppA, ApppppA, ApppppA, ATP-αS, GTP-αS and K_I values for pyrophosphate, monophosphate and AMP were derived from titration of each non-labeled compound into 1.5 µM ApppBODIPY assays at five concentrations of the competitors [43].

Protein X-ray crystallography

NitFhit crystals were grown by hanging drop vapor diffusion by mixing 2 µl protein (7 mg ml⁻¹) in 10 mM NaHEPES pH 7.0, 50 mM NaCl, 5 mM DTT with 2 µl of 38% 2-methyl-2,4-pentanediol (MPD) and equilibrating against 1 ml 38% MPD. After one week at room temperature, individual microcrystals were seeded into similar drops, equilibrated against 35% MPD, and grown for one month. Crystals (~100 µm × 200 µm × 300 µm) were flash-frozen in liquid nitrogen. On the basis of data collected at Cornell High Energy Synchrotron Source beamline F-1, native crystals had the symmetry of space group I222 or I2₁2₁2₁, contained a NitFhit monomer in the asymmetric unit and 58% solvent, and were ordered to 3.5 Å resolution. A single frozen crystal was thawed into 10 µl 5 mM NaHEPES, 25 mM NaCl, 2.5 mM DTT, 20% MPD with 1 mM thimerosal, and refrozen after 8 h.

At the National Synchrotron Light Source, beamline X8-C was tuned to the measured absorption edge for the mercurated crystal ($\lambda_{edge} = 1.008989$ Å) and to a remote wavelength ($\lambda_{remote} = 0.992782$ Å), and

X-ray diffraction data were measured in 1° oscillations with an ADSC Quantum-4 CCD camera 250 mm from the crystal. Data, indexed and scaled with the HKL package [53], indicated the presence of mercury atoms and a reduction in crystallographic symmetry to P2₁2₁2 (from I222). The change in space group was accompanied by a 6% reduction in the b cell length, a 2% reduction in solvent content, and an increase in resolution to 2.8 Å. Using the CNS package [54], a four atom mercury solution was obtained for the λ_{edge} anomalous difference Patterson map. Using diffraction data from both wavelengths, the heavy atom positions and the scattering factors f′ and f″ of the mercurated protein were refined [54] and used to generate TAD phases (Table 2). The density-modified TAD electron-density map of the refined, enantiomorphic heavy atom solution was interpretable. A fifth mercury position, located in this map, was used to generate the final density-modified TAD electron density map (hereafter, the experimental map) that was used for model building (Figure 3a) with O (http://imsb.au.dk/~mok/o/).

Electron density corresponding to two non-identical Fhit dimers was located. Each Fhit dimer occurs across a crystallographic twofold rotation axis such that the east and west subunits of Fhit are identical. The two non-identical Fhit dimers occur as an imperfect 222 tetramer with an origin of noncrystallographic symmetry (NCS) at 0.5, 0.5, 0.25. The experimental map was of sufficient quality to build from residues 13 to the carboxyl terminus of each non-identical, 440 amino acid polypeptide with 23 residues missing from the Fhit domain of one molecule and 30 residues missing from the other. Non-identical NitFhit molecules were built independently and refined without NCS restraints. The protein atomic model, containing additionally five mercury atoms and a bulk solvent model, was refined by simulated annealing [54] against λ_{edge} data with a maximal-likelihood target function based on experimental phases [55]. All reflections from 30.0 Å to 2.8 Å were included in refinement except 7% reserved for free R factor analysis [56]. The five mercury atoms appear to be ethylmercuric adducts to cysteine residues 55, 75 and 169 of the (northern) A chain and residues 55 and 169 of the (southern) B chain. These adducts were built and refined as ethylmercury with occupancies of 40% to 57%. Nineteen additional amino acids, 161 water molecules, four sodium ions, and one ordered MPD molecule were built from sigma A weighted, cross-validated, phase combined, 2Fo–Fc and Fo–Fc maps [54]. The final atomic model has an R_{work} of 19.0% and an R_{free} of 23.1% with geometry that is neither under-restrained nor over-restrained with respect to ten recently released protein structures refined to 2.8 Å. Molecular graphics methods were as described [32]. The coordinates (1EMS) and structure factors (1EMSsf) have been deposited into the Protein Data Bank.

References

1. Sozzi G, Veronese ML, Negrini M, Baffa R, Cotticelli MG, Inoue H, *et al.*: **The *FHIT* gene at 3p14.2 is abnormal in lung cancer.** *Cell* 1996, **85:**17-26.
2. Mao L, Lee JS, Kurie JM, Fan YH, Lippman SM, Lee JJ, *et al.*: **Clonal genetic alterations in the lungs of current and former smokers.** *J Natl Cancer Inst* 1997, **89:**857-862.
3. Sozzi G, Pastorino U, Moiraghi L, Tagliabue E, Pezzella F, Ghirelli C, *et al.*: **Loss of *FHIT* function in lung cancer and preinvasive bronchial lesions.** *Cancer Res* 1998, **58:**5032-5037.

4. Huebner K, Sozzi G, Brenner C, Pierotti MA, Croce CM: **Fhit loss in lung cancer: diagnostic and therapeutic implications.** *Adv Oncol* 1999, **15**:3-10.
5. Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, *et al.*: **The *FHIT* gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers.** *Cell* 1996, **84**:587-597.
6. Zimonjic DB, Druck T, Ohta M, Kastury K, Croce CM, Popescu NC, *et al.*: **Positions of chromosome 3p14.2 fragile sites (FRA3B) within the *FHIT* gene.** *Cancer Res* 1997, **57**:1166-1170.
7. Mimori K, Druck T, Inoue H, Alder H, Berk L, Mori M, *et al.*: **Cancer-specific chromosome alterations in the constitutive fragile region FRA3B.** *Proc Natl Acad Sci USA* 1999, **96**:7456-7461.
8. Tanaka H, Shimada Y, Harada Y, Shinoda M, Hatooka S, Imamura M, *et al.*: **Methylation of the 5′ CpG island of the *FHIT* gene is closely associated with transcriptional inactivation in esophageal squamous cell carcinomas.** *Cancer Res* 1998, **58**:3429-3434.
9. Huebner K, Garrison PN, Barnes LD, Croce CM: **The role of the *FHIT/FRA3B* locus in cancer.** *Annu Rev Genet* 1998, **32**:7-31.
10. Greenspan DL, Connolly DC, Wu R, Lei RY, Vogelstein JTC, Kim YT, *et al.*: **Loss of *FHIT* expression in cervical carcinoma cell lines and primary tumors.** *Cancer Res* 1997, **57**:4692-4698.
11. Cohen AJ, Li FP, Berg S, Marchetto DJ, Tsai S, Jacobs SC, *et al.*: **Hereditary renal-cell carcinoma associated with a chromosomal translocation.** *New Engl J Med* 1979, **301**:592-595.
12. Xiao GH, Jin F, Kleinszanto AJP, Goodrow TL, Linehan MW, Yeung RS: **The *FHIT* gene product is highly expressed in the cytoplasm of renal tubular epithelium and is down-regulated in kidney cancers.** *Amer J Pathol* 1997, **151**:1541-1547.
13. Hadaczek P, Siprashvili Z, Markiewski M, Domagala W, Druck T, McCue PA *et al.*: **Absence or reduction of Fhit expression in most clear cell renal carcinomas.** *Cancer Res* 1998, **58**:2946-2951.
14. Baffa R, Veronese ML, Santoro R, Mandes B, Palazzo JP, Rugge M *et al.*: **Loss of *FHIT* expression in gastric carcinoma.** *Cancer Res* 1998, **58**:4708-4714.
15. Simon B, Bartsch D, Barth P, Prasnikar N, Münch K, Blum A, *et al.*: **Frequent abnormalities of the putative tumor suppressor gene *FHIT* at 3p14.2 in pancreatic carcinoma cell lines.** *Cancer Res* 1998, **58**:1583-1587.
16. Sorio C, Baron A, Orlandini S, Zamboni G, Pederzoli P, Huebner K, *et al.*: **The *FHIT* gene is expressed in pancreatic ductular cells and is altered in pancreatic cancers.** *Cancer Res* 1999, **59**:1308-1314.
17. Mandai M, Konishi I, Kuroda H, Nanbu K, Matsushita K, Yura Y, *et al.*: **Expression of abnormal transcripts of the *FHIT* (fragile histidine triad) gene in ovarian carcinoma.** *Eur J Cancer* 1998, **34**:745-749.
18. Virgilio L, Schuster M, Gollin SM, Veronese ML, Ohta M, Huebner K, *et al.*: **FHIT gene alterations in head and neck squamous cell carcinomas.** *Proc Natl Acad Sci USA* 1996, **93**:9770-9775.
19. Kisielewski AE, Xiao GH, Liu SC, Klein-Szanto AJ, Novara M, Sina J, *et al.*: **Analysis of the *FHIT* gene and its product in squamous cell carcinomas of the head and neck.** *Oncogene* 1998, **17**:83-91.
20. Ingvarsson S, Agnarsson BA, Sigbjornsdottir BI, Kononen J, Kallioniemi OP, Barkardottir RB, *et al.*: **Reduced Fhit expression in sporadic and BRCA2-linked breast carcinomas.** *Cancer Res* 1999, **59**:2682-2689.
21. Campiglio M, Pekarsky Y, Menard S, Tagliabue E, Pilotti S, Croce CM: **FHIT loss of function in human primary breast cancer correlates with advanced stage of the disease.** *Cancer Res* 1999, **59**:3866-3869.
22. Iwai T, Yokota S, Nakao M, Nakazawa N, Taniwaki M, Kimura T, *et al.*: **Frequent aberration of *FHIT* gene expression in acute leukemias.** *Cancer Res* 1998, **58**:5182-5187.
23. Peters UR, Hasse U, Oppliger E, Tschan M, Ong ST, Rassool FV *et al.*: **Aberrant *FHIT* mRNA transcripts are present in malignant and normal haematopoiesis, but absence of Fhit protein is restricted to leukaemia.** *Oncogene* 1999, **18**:79-85.
24. Hallas C, Albitar M, Letofsky J, Keating MJ, Huebner K, Croce CM: **Loss of *FHIT* expression in acute lymphoblastic leukemia.** *Clin Cancer Res* 1999, **5**:2409-2414.
25. Pekarsky Y, Druck T, Cotticelli MG, Ohta M, Shou J, Mendrola J, *et al.*: **The murine *Fhit* locus: isolation, characterization, and expression in normal and tumor cells.** *Cancer Res* 1998, **58**:3401-3408.
26. Fong LYY, Fidanza V, Zanesi N, Lock LF, Siracusa LD, Mancini R, *et al.*: **Muir–Torre-like syndrome in *FHIT* deficient mice.** *Proc Natl Acad Sci USA* 2000, **97**:4742-4747.
27. Brenner C, Garrison P, Gilmour J, Peisach D, Ringe D, Petsko GA, *et al.*: **Crystal structures of Hint demonstrate that histidine triad proteins are GalT-related nucleotide-binding proteins.** *Nat Struct Biol* 1997, **4**:231-238.
28. Brenner C, Bieganowski P, Pace HC, Huebner K: **The histidine triad superfamily of nucleotide-binding proteins.** *J Cell Physiol* 1999, **181**:179-187.
29. Huang Y, Garrison PN, Barnes LD: **Cloning of the *Schizosaccharomyces pombe* gene encoding diadenosine 5′,5-P1,P4-tetraphosphate (Ap4A) asymmetrical hydrolase: sequence similarity with the histidine triad (HIT) protein family.** *Biochem J* 1995, **312**:925-932.
30. Barnes LD, Garrison PN, Siprashvili Z, Guranowski A, Robinson AK, Ingram SW, *et al.*: **Fhit, a putative tumor suppressor in humans, is a dinucleoside 5′,5‴-P-1,P-3-triphosphate hydrolase.** *Biochem* 1996, **35**:11529-11535.
31. Siprashvili Z, Sozzi G, Barnes LD, McCue P, Robinson AK, Eryomin V, *et al.*: **Replacement of Fhit in cancer cells suppresses tumorigenicity.** *Proc Natl Acad Sci USA* 1997, **94**:13771-13776.
32. Pace HC, Garrison PN, Robinson AK, Barnes LD, Draganescu A, Rosler A, *et al.*: **Genetic, biochemical, and crystallographic characterization of Fhit-substrate complexes as the active signaling form of Fhit.** *Proc Natl Acad Sci USA* 1998, **95**:5484-5489.
33. Werner NS, Siprashvili Z, Fong LYY, Marquitan G, Schroder JK, Bardenheuer W, *et al.*: **Constitutive expression of exogenous Fhit in renal carcinoma cell lines confers differential tumor suppression.** *Cancer Res* 2000, **60**:2780-2785.
34. Ji L, Fang B, Yeh N, Fong K, Minna JD, Roth JA: **Induction of apoptosis and inhibition of tumorigenicity and tumor growth by adenovirus vector-mediated fragile histidine triad (*FHIT*) gene overexpression.** *Cancer Res* 1999, **59**:3333-3339.
35. Sard L, Accornero P, Tornielli S, Delia D, Bunone G, Campiglio M, *et al.*: **The tumor-suppressor gene *FHIT* is involved in the regulation of apoptosis and in cell cycle control.** *Proc Natl Acad Sci USA* 1999, **96**:8489-8492.
36. Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO, Eisenberg D: **Detecting protein function and protein-protein interactions from genome sequences.** *Science* 1999, **285**:751-753.
37. Enright A, Iliopoulos I, Kyrpides N, Ouzounis C: **Protein interaction maps for complete genomes based on gene fusion events.** *Nature* 1999, **402**:86-90.
38. Marcotte E, Pellegrini M, Thompson M, Yeates T, Eisenberg D: **A combined algorithm for genome-wide prediction of protein function.** *Nature* 1999, **402**:83-86.
39. Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D, Yeates TO: **Assigning protein functions by comparative genome analysis: protein phylogenetic profiles.** *Proc Natl Acad Sci USA* 1999, **96**:4285-4288.
40. Glover TW, Hoge AW, Miller DE, Ascara-Wilke JE, Adam AN, Dagenais SL, *et al.*: **The murine *Fhit* gene is highly similar to its human orthologue and maps to a common fragile site region.** *Cancer Res* 1998, **58**:3409-3414.
41. Pekarsky Y, Campiglio M, Siprashvili Z, Druck T, Sedkov Y, Tillib S, *et al.*: **Nitrilase and Fhit homologs are encoded as fusion proteins in *Drosophila melanogaster* and *Caenorhabditis elegans*.** *Proc Natl Acad Sci USA* 1998, **95**:8744-8749.
42. Normanly J, Grisafi P, Fink GR, Bartel B: ***Arabidopsis* mutants resistant to the auxin effects of indole-3-acetonitrile are defective in the nitrilase encoded by the *NIT1* gene.** *Plant Cell* 1997, **9**:1781-1790.
43. Draganescu A, Hodawadekar SC, Gee KR, Brenner C: **Fhit-nucleotide specificity probed with novel fluorescent and fluorigenic substrates.** *J Biol Chem* 2000, **275**:4555-4560.
44. Robinson AK, de la Pena CE, Barnes LD: **Isolation and characterization of diadenosine tetraphosphate (Ap4A) hydrolase from *Schizosaccharomyces pombe*.** *Biochem Biophys Acta* 1993, **1161**:139-148.
45. Lahm A, Suck D: **DNase I-induced DNA conformation. 2 Å structure of a DNase I-octamer complex.** *J Mol Biol* 1991, **222**:645-667.
46. Lima CD, Damico KL, Naday I, Rosenbaum G, Westbrook EM, Hendrickson WA: **MAD analysis of Fhit, a putative human tumor suppressor from the HIT protein family.** *Structure* 1997, **5**:763-774.
47. Stevenson DE, Feng R, Dumas F, Groleau D, Mihoc A, Storer AC: **Mechanistic and structural studies on *Rhodococcus* ATCC 39484 nitrilase.** *Biotechnol Appl Biochem* 1992, **15**:283-302.
48. Novo C, Tata R, Clemente A, Brown PR: ***Pseudomonas aeruginosa* aliphatic amidase is related to the nitrilase/cyanide hydratase enzyme family and Cys166 is predicted to be the active site nucleophile of the catalytic mechanism.** *FEBS Lett* 1995, **367**:275-279.

49. Segal E, Le Pecq JB: **Relationship between cellular diadenosine 5′,5‴-P1,P4-tetraphosphate level, cell density, cell growth stimulation and toxic stresses.** *Exp Cell Res* 1986, **167:**119-126.

50. Vartanian A, Narovlyansky A, Amchenkova A, Turpaev K, Kisselev L: **Interferons induce accumulation of diadenosine triphosphate (Ap3A) in human cultured cells.** *FEBS Lett* 1996, **381:**32-34.

51. Vartanian A, Prudovsky I, Suzuki H, Dal Pra I, Kisselev L: **Opposite effects of cell differentiation and apoptosis on Ap3A/Ap4A ratio in human cell cultures.** *FEBS Lett* 1997, **415:**160-162.

52. Ghosh S, Lowenstein JM: **A multifunctional vector system for heterologous expression of proteins in *Escherichia coli* — expression of native and hexahistidyl fusion proteins, rapid purification of the fusion proteins, and removal of fusion peptide by Kex2 protease.** *Gene* 1997, **176:**249-255.

53. Otwinowski Z, Minor W: **Processing of X-ray diffraction data collected in oscillation mode.** *Methods Enzymol* 1997, **276:**307-326.

54. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse Kunstleve RW, *et al.*: **Crystallography and NMR system: a new software suite for macromolecular structure determination.** *Acta Crystallogr D* 1998, **54:**905-921.

55. Pannu NS, Murshudov GN, Dodson EJ, Read RJ: **Incorporation of prior phase information strengthens maximum-likelihood structure refinement.** *Acta Crystallogr D* 1998, **54:**1285-1294.

56. Brünger AT: **Free R value: cross-validation in crystallography.** *Methods Enzymol* 1997, **277:**366-396.

57. Tesmer JJ, Stemmler TL, Penner-Hahn JE, Davisson VJ, Smith JL: **Preliminary X-ray analysis of *Escherichia coli* GMP synthetase: determination of anomalous scattering factors for a cysteinyl mercury derivative.** *Proteins* 1994, **18:**394-403.

58. Barton GJ: **ALSCRIPT, a tool to format multiple sequence alignments.** *Protein Eng* 1993, **6:**37-40.

59. Levitt M, Chothia C: **Structural patterns in globular proteins.** *Nature* 1976, **261:**552-558.