

Fas Apoptosis Inhibitory Molecule Contains a Novel β -Sandwich in Contact with a Partially Ordered Domain

Michael Hemond¹, Thomas L. Rothstein² and Gerhard Wagner^{1*}

¹Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA

²Center for Oncology and Cell Biology, The Feinstein Institute for Medical Research, 350 Community Drive, Manhasset, NY 11030, USA

Received 14 October 2008;

received in revised form

3 January 2009;

accepted 6 January 2009

Available online

13 January 2009

Fas apoptosis inhibitory molecule (FAIM) is a soluble cytosolic protein inhibitor of programmed cell death and is found in organisms throughout the animal kingdom. A short isoform of FAIM is expressed in all tissue types, while an alternatively spliced long isoform is specifically expressed in the brain. Here, the short isoform is shown to consist of two independently folding domains in contact with each other. The NMR solution structure of the C-terminal domain of murine FAIM is solved in isolation and revealed to be a novel protein fold, a noninterleaved seven-stranded β -sandwich. The structure and sequence reveal several residues that are likely to be involved in functionally significant interactions with the N-terminal domain or other binding partners. Chemical shift perturbation is used to elucidate contacts made between the N-terminal domain and the C-terminal domain.

© 2009 Elsevier Ltd. All rights reserved.

Edited by M. F. Summers

Keywords: FAIM; apoptosis; Fas; NMR structure; β -sandwich

Introduction

Programmed cell death, or apoptosis, plays a vital role in numerous physiological processes. During embryonic development, it is essential for the formation of anatomical structures such as fingers and some tubular organs. It also has numerous regulatory roles in adult life, notably in the regulation of the immune system and deletion of autoreactive lymphocytes. Dysregulation of apoptosis is also thought to play a role in many types of neoplastic disease.

Many stimuli can trigger apoptosis; one such stimulus is engagement of the members of the

transmembrane tumor necrosis factor receptor superfamily with a cognate ligand. Fas/FasL is an example of this type of receptor–ligand pair: binding of FasL can promote the binding of the Fas-associated protein with death domain (FADD) to the intracellular portion of Fas via a “death domain” carried by each protein. FADD also contains a death effector domain that binds to procaspase-8 and promotes its proteolytic activation. This leads to activation of downstream “executioner caspases,” such as caspase-3, caspase-6, and caspase-7, which carry out the apoptotic process and represent the final step common to all apoptotic pathways.

Fas apoptosis inhibitory molecule (FAIM) was first identified by Schneider *et al.* as an inducible inhibitor of apoptosis, using differential mRNA display between primary murine B-lymphocytes stimulated with antibodies directed against the surface immunoglobulin receptor (anti-IgM) in the presence of a soluble dimeric CD40 ligand construct and B-lymphocytes stimulated with CD40 ligand alone.¹ Reverse transcription and PCR resulted in construction of a probe that was used to screen a murine cDNA library, from which was isolated a full-length clone of the short isoform of FAIM (FAIM-S). In the same work, the antiapoptotic activity of FAIM was confirmed by overexpressing FAIM-S in a murine B-lymphoma cell line (BAL-17)

*Corresponding author. E-mail address:

gerhard_wagner@hms.harvard.edu.

Abbreviations used: FAIM, Fas apoptosis inhibitory molecule; FADD, Fas-associated protein with death domain; FAIM-S, short isoform of FAIM; FAIM-L, long isoform of FAIM; NGF, nerve growth factor; CTD, C-terminal domain; NTD, N-terminal domain; HSQC, heteronuclear single-quantum coherence; TROSY, transverse relaxation optimized spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; NOE, nuclear Overhauser enhancement; FABP, fatty acid binding protein; GST, glutathione S-transferase; OD, optical density; TBS, Tris-buffered saline.

and by subjecting transfected and control cells to Fas-mediated apoptotic stimuli, including Fas ligand-bearing AE7 T-cells and Jo-2 anti-Fas antibody; overexpression of FAIM gave partial protection against cell lysis (^{51}Cr release) and the DNA degradation characteristic of apoptosis.

FAIM is a 20-kDa, soluble, evolutionarily conserved protein known to be expressed at the mRNA level in a wide variety of mouse tissues. In murine brain (and, to a lesser extent, in testis and embryonic tissue), alternative splicing is known to yield an mRNA whose protein product, the long isoform of FAIM (FAIM-L), is longer than FAIM-S by 22 residues at the N-terminus.² The *faim* gene localizes to chromosome 9 (9f1) in mice and to chromosome 3 (3q21–25) in humans;² these authors note that abnormalities in this region have been associated with a number of types of human cancer, hinting at a possible role for FAIM in oncogenesis.

Sole *et al.* reported that overexpression of FAIM-S, but not of FAIM-L, in conjunction with nerve growth factor (NGF), promoted the outgrowth of neurites in cultured neurons from the superior cervical ganglion of the mouse and in the neuron-like PC12 cell line.³ Conversely, knockdown of endogenous FAIM with small interfering RNA decreased neurite outgrowth. More recently, Segura *et al.* also found that FAIM-L, but not FAIM-S, is induced in cultured neurons in response to NGF⁺ signaling and that overexpression of FAIM-L antagonizes apoptotic signaling from FasL and tumor necrosis factor α .⁴ Thus, curiously, it would appear that FAIM-S is an inducible inhibitor of Fas-mediated apoptosis in cultured murine lymphocytes, while that function is performed exclusively by FAIM-L in the neuron model despite the presence of FAIM-S.

The molecular mechanisms of FAIM are not understood in detail, but studies have shed some light on the subject. With regard to apoptosis inhibition in lymphocytes, an experiment by Schneider *et al.* showed that the anti-apoptotic activity of FAIM occurs upstream of poly(ADP-ribose) polymerase cleavage by demonstrating a decrease in poly(ADP-ribose) polymerase cleavage in FAIM-transfected cells treated with the Jo-2 anti-Fas antibody.¹ The neurite outgrowth activity of FAIM has also been investigated in molecular detail: it has been found that nuclear factor κB is activated by FAIM-S in the presence of NGF and that transfection of a stabilized form of I κB , an inhibitor of nuclear factor κB , into cultured neurons abolished the ability of FAIM-S to promote neurite outgrowth.³ The same study showed immunoprecipitation evidence of an interaction between FAIM-S and both TrkA and p75^{NTR} neurotrophin receptors when NGF was present; FAIM-L was not studied in this context. The latter study,⁴ using similar techniques, found an interaction between FAIM-L and the Fas receptor that was blocked by overexpression of the known Fas-binding partner FADD. These researchers also showed that the antiapoptotic activity of FAIM-L in neurons is effective upstream of both caspase-8 and caspase-

3, since the activity of these caspases in the presence of apoptotic stimuli is blocked by FAIM-L.

The presence of FAIM nucleotide sequences in several vertebrate and invertebrate animal species has been confirmed by Southern blot analysis¹ and by analysis of publicly available nucleotide databases.^{1,3} Wiens *et al.* and Wiens and Müller even reported a homolog of FAIM that was 41% identical to human FAIM in *Suberites domuncula*, a poriferan lacking lymphocytes and neurons but possessing a receptor, DD2, that contains Fas-like death domains.^{5,6} Although the *faim* gene product is highly conserved among species, there is no known sequence homology to other antiapoptotic proteins¹ or, indeed, to any protein of known structure. It has been noted that these facts point toward a vital physiological function for FAIM.

Although progress has been made toward understanding the activities of FAIM at the cellular level, little is known about its biochemistry or the molecular mechanisms by which it acts. Because FAIM has no significant primary sequence homology to any protein of known structure or function, homology modeling currently can provide no insight. Experimental structure determination will therefore be necessary to interpret mechanistic data and may give clues to the precise function(s) of FAIM as well.

In this article, we (1) describe a construct of FAIM and a set of buffer conditions conducive to NMR spectroscopy, in the process gaining insight into its domain organization; (2) report and analyze the NMR solution structure of the FAIM C-terminal domain (CTD), a novel protein fold, noting residues that are likely to be involved in functionally significant interactions; and (3) identify certain sites of interaction between the N-terminal domain (NTD) and the CTD.

Results

FAIM is composed of two independently folding domains

Wild-type FAIM-S exhibits fair-quality NMR spectra and has limited solubility. Previous work for a similar situation⁷ suggested that introducing rationally chosen point mutations could improve the physical and/or spectral properties of a protein. Point mutations to improve solubility were selected from among the residues seen at the homologous position in the orthologs identified by Schneider *et al.*, namely, the *Caenorhabditis elegans*, mouse, and human forms.¹ Substitutions were chosen subjectively based on criteria including lack of conservation among orthologs, qualitative decrease in hydrophobicity or reduction in *pI* achieved by mutation, and the degree to which the substitution could be considered conservative. The murine FAIM-S triple mutant F55Y/N92D/G148D was identified as a candidate for further study because it was well-behaved at reasonable NMR concentrations and temperatures in low-salt

1010TBS buffer, and its ^{15}N heteronuclear single-quantum coherence (HSQC) spectrum was similar to that of the wild type (data not shown). Backbone assignment of the triple mutant FAIM-S revealed that the majority of the peaks visible in the ^{15}N HSQC spectrum were from the C-terminal half of the protein: residues 91–179 were assigned with 85% completeness (76 of 89), while residues 1–90 were assigned with only 36% completeness (32 of 90). This result suggested that the N-terminal half of the protein might be responsible for the poor overall spectral quality and perhaps the poor physical behavior of the wild type as well. This idea was supported by limited proteolysis experiments in which the triple mutant FAIM-S could be cleaved, resulting in one or several smaller, more stable fragments (Fig. 1a).

Purification of a construct (N Δ 90/G148D) incorporating residues 91–179 of murine FAIM-S and the G148D mutation revealed that the peaks of this construct are well-dispersed in the ^{15}N HSQC spectrum. Thus, this construct is well-folded and is an independent domain, termed FAIM-CTD. FAIM-CTD had no requirement for detergent and could readily be concentrated to more than 1 mM. Limited proteolysis experiments confirmed that it was similar in size to one of the more trypsin-resistant products of triple mutant FAIM-S cleavage (Fig. 1a). It was extremely stable in low-salt buffer as long as reducing conditions were maintained: samples showed no precipitation or changes in the ^{15}N HSQC spectrum when analyzed for several weeks at 25 °C and stored for more than 1 year at 4 °C. Comparison of the ^{15}N HSQC spectra of FAIM-CTD and triple mutant FAIM-S did show some significant differences, as discussed below, suggesting that removal of the N-terminal region could have caused structural perturbation within FAIM-CTD. Structural perturbation associated with the G148D mutation was minimal and generally limited to the region near the site of mutation (Fig. 1b). FAIM-CTD appeared to be monomeric by ultracentrifugation. At 22,500 rpm (34,000g), ultracentrifugation data were well fitted by a straight line (Fig. 1c). The temperature, buffer conditions, and protein concentrations used for ultracentrifugation were very similar to those subsequently used for NMR spectroscopy, except for the unavoidable radial gradient in protein concentration, so the absence of detectable dimerization of FAIM-CTD during ultracentrifugation strongly suggests its absence during NMR spectroscopy.

To ensure that the FAIM-CTD construct contained the complete domain, another construct composed of residues 82–179 (N Δ 81) was synthesized, expressed, and purified. Its ^{15}N HSQC spectrum was very similar to that of the shorter FAIM-CTD construct N Δ 90 (residues 91–179) (data not shown), demonstrating that residues 82–90 do not interact strongly with FAIM-CTD and implying that the N Δ 90 construct encompasses all the residues involved in the folding of FAIM-CTD.

A FAIM-NTD construct (C Δ 97/F55Y), incorporating residues 1–96 of murine FAIM-S and F55Y

mutation, was also expressed and evaluated. Although its ^{15}N HSQC spectrum shows heterogeneous peak intensity, it was stable at NMR concentrations (0.1–0.2 mM) for hours in 1010TBS with 3% Tween 20 at 35 °C. The ^{15}N HSQC spectra of this construct at 0.1–0.2 mM, acquired at 500 MHz in 1010TBS and 3% Tween 20 (without DTT) at three different temperatures, are shown in Fig. 1d. The visible residues are well dispersed, indicating that large regions within this construct are also well folded. However, the number of visible peaks is about one-third less than expected. The intensity and line shape of some of these peaks are highly temperature dependent, suggesting the presence of chemical exchange or temperature-dependent partial unfolding.

A very short truncation of murine FAIM-S, incorporating residues 1–62 (C Δ 63), was also made by introducing a stop codon into the FAIM-S gene at residue 63. This protein was expressed at very low levels under the usual expression conditions and could not otherwise be evaluated. A longer C-terminal protein consisting of residues 52–179 was also synthesized and expressed, but expression yield was very poor, quantitative aggregation was seen by fast protein liquid chromatography, and NMR spectroscopy of the aggregated fraction failed to detect a protein signal. These results indicate that truncation within the NTD yields unstable protein constructs.

FAIM tolerates apolar but not polar substitution of cysteine residues

To assess the structural significance, if any, of the two cysteine residues (Cys122 and Cys149) in FAIM-CTD, sequence alignments and mutation studies were undertaken. Both cysteine residues in murine FAIM-S are conserved among vertebrate orthologs but not among more distant relatives³ (Fig. 1e). Cysteine-to-serine mutations were thought to be conservative and most likely to preserve native structure in a surface-exposed position, while mutations to alanine would be less hydrophilic and also reasonably conservative. Substitution of Cys122 with alanine is observed in a known homolog, indicating that the protein was likely to tolerate this mutation. In one experiment, both cysteines were mutated to serine (C122S/C149S). Another construct in which both residues were mutated to alanine (C122A/C149A) was likewise made. Finally, since there was no known homolog in which Cys149 was replaced by alanine, Cys149 was also mutated to phenylalanine to yield a third cysteine-free construct, C122A/C149F.

Expression levels of various cysteine mutants were compared; results are shown in Fig. 1f. The majority of the C122S/C149S double-mutant protein was found in the cell pellet after lysis. Expression of the cysteine-to-alanine double mutant C122A/C149A, however, was similar to the expression of the wild type. Its behavior was also similar to that of the wild type during purification. The ^{15}N HSQC spectrum at 400 MHz showed some differences from

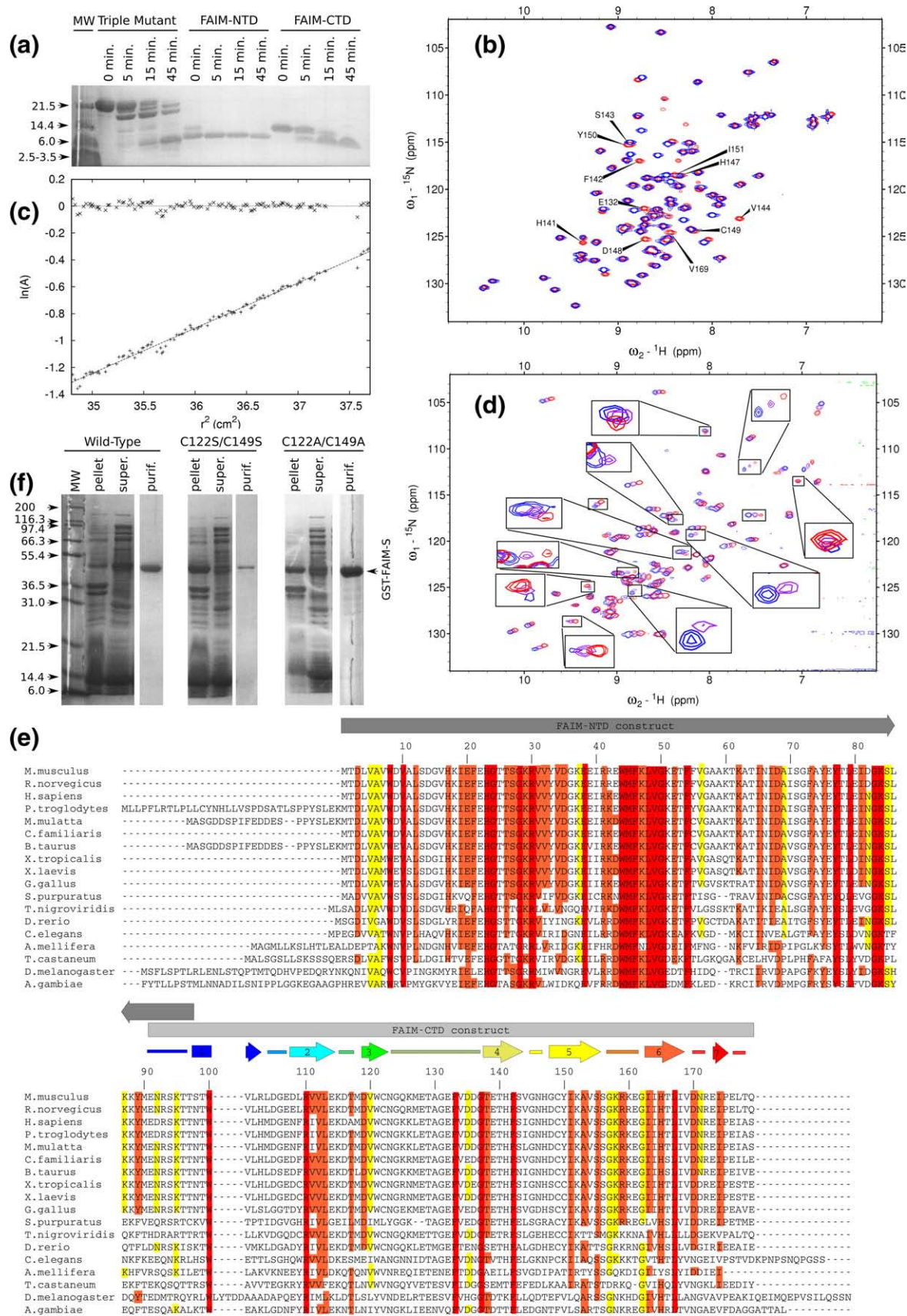


Fig. 1 (legend on next page)

that of wild-type FAIM-S. Many of these differences were from residues near the sites of mutation. Remarkably, the C149F mutation was also well-tolerated (data not shown). Taken together, these results imply that both cysteine residues reside within a hydrophobic environment that cannot tolerate disruption by polar substitutions. However, the sulfhydryl moieties of the cysteine side chains themselves are not vital to the stability of the protein.

FAIM-CTD exhibits a novel β -sandwich fold

Due to the low molecular mass of the FAIM-CTD construct (10 kDa), transverse relaxation optimized spectroscopy (TROSY) experiments were not used, and deuteration was unnecessary during NMR spectroscopy. Triple-resonance NMR spectra were of generally high quality, permitting 100% backbone assignment of nonproline amide protons between residues 96 and 179 inclusive. Because the backbone assignment was done without the benefit of a CO pair and was later checked against the CO pair and found to agree, it is very likely to be error-free. Side-chain atoms were assigned with the use of standard techniques. Stereospecific methyl assignments were obtained for 10 of 15 valine and leucine residues. Nuclear Overhauser enhancement (NOE) spectroscopy (NOESY) peaks were integrated with SPARKY, and CYANA was used to calculate 100 structures, of which the lowest 20% in energy are presented. The structure was then validated by measuring N-HN residual dipolar couplings (RDCs) and by calculating Q_{free} . The final NOE statistics and structural parameters are tabulated in Table 1.

A ribbon diagram showing the tertiary structure is shown in Fig. 2a. The quality of the structure is illustrated in Fig. 2b with a superposition of the

Table 1. Figures of merit for the FAIM-CTD structure

Restrictive NOE distance restraints (unique)	934
Short range	399
Medium range	114
Long range	421
TALOS dihedral angle restraints	106
Side-chain dihedral angle restraints	37
Hydrogen-bonding distance restraints	16
Dihedral angle statistics (%)	
Residues in most favored regions	78.9
Residues in additionally allowed regions	20.9
Residues in generously allowed regions	0.1
Residues in disallowed regions	0.2
Average backbone RMSD to mean structure (residues T96–T178) (Å)	0.51
Average heavy atom RMSD to mean structure (residues T96–T178) (Å)	1.02
Maximum NOE violation (Å)	0.37
Cornilescu Q_{free} from RDC validation (lowest-energy structure; secondary structural elements and turns only)	0.505

Dihedral angles and NOE restraints were analyzed with PROCHECK.⁸ Q_{free} was computed using PALES, as described in the text. Other statistics were calculated using CYANA.

peptide backbones of all accepted structures. FAIM-CTD is a “ β -sandwich” reminiscent of the classical immunoglobulin fold; however, it is somewhat smaller, has a simpler topology, and is stable in the absence of a disulfide bond between the two sheets. The two β -sheets are canted with respect to each other by an angle of approximately 30°.

The topology of FAIM-CTD is shown schematically in Fig. 2c. The topology is simple: there is no interleaving of β -strands, and all strands are anti-parallel. Successive strands are connected by four-residue turns in all but two cases. Between strands 5 and 6 is a longer loop comprising exclusively hydrophilic residues. Residues in the distal region of this loop (from K158 to K160), as well as residues pre-

Fig. 1. Biochemical, physical, and spectral properties of FAIM constructs. (a) Limited proteolysis of FAIM constructs. Truncation mutants FAIM-NTD (CΔ111) and FAIM-CTD reproduce the same stable low-molecular-mass band as FAIM-S itself when subjected to limited tryptic proteolysis, indicating that both constructs are well folded and that truncation has not removed residues necessary for the stability of either domain. A 0.5% bed volume of immobilized trypsin treated with L-1-tosylamido-2-phenylethyl chloromethyl ketone (15–25 U/mL; Sigma) was added to protein in TBS and shaken gently at 4 °C. The reactions were centrifuged, and aliquots were removed from the supernatant at the indicated time points. (b) G148D mutation only minimally perturbs the structure of FAIM-CTD. The ¹⁵N HSQC spectra of native G148 (blue) and mutant G148D (red) NΔ90 FAIM-CTD constructs are shown. The most strongly perturbed residues, near the site of the mutation, are labeled. (c) FAIM-CTD is monomeric by ultracentrifugation at NMR concentration. Equilibrium ultracentrifugation of FAIM-CTD at 1.7 mM was performed for 12 h at 22,500 rpm (34,000g) in freshly prepared degassed 1010TBS identical to that used for NMR spectroscopy. The slope of the line corresponds to 11 kDa, assuming a partial specific volume of 0.73. The residuals are shown at the top of the figure. (d) The NTD of FAIM-S has well-dispersed backbone amide chemical shifts and is largely structured. The ¹⁵N HSQC spectra at 500 MHz of the CΔ97/F55Y FAIM-NTD construct at 15 °C (blue), 25 °C (purple), and 35 °C (red) are overlaid. Contour levels are the same for each spectrum. Only about 60–70% of nonproline residues are visible, and a significant portion of these exhibit temperature-dependent signal broadening or loss (see insets). There is no trend toward improved spectra at any temperature: some peaks exhibit broadening at higher temperatures, while others are broadened at lower temperatures. (e) Alignment of selected FAIM sequences from publicly available databases. Secondary structural elements, where known, are shown above the sequence. Identical residues are boxed and colored according to the degree of substitution at each position: red, strictly conserved; orange, highly similar; yellow, less similar; white, not conserved. Numbering is performed according to murine FAIM-S. (f) Mutation of cysteine residues reveals that they exist within a hydrophobic environment, but are dispensable to folding of FAIM-S. SDS-PAGE of wild-type and cysteine-free full-length FAIM-S at various stages during purification shows that alanine substitutions are well-tolerated, while serine substitutions are not. All constructs have a C-terminal cloning artifact. Results were similar when single cysteine substitutions were expressed (data not shown).

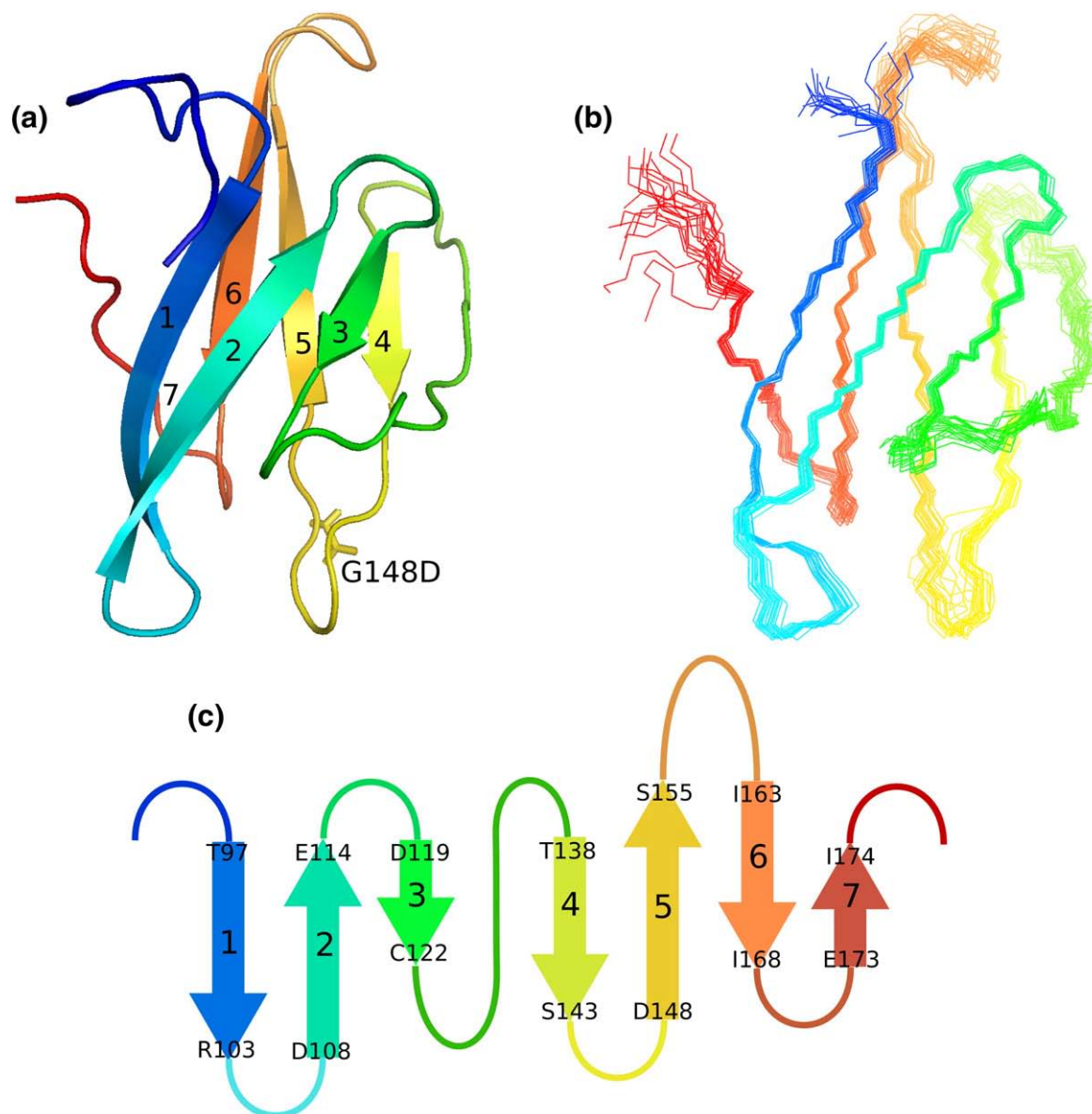


Fig. 2. FAIM-CTD is a noninterleaved β -sandwich. (a) Ribbon diagram of FAIM-CTD, with the N-terminus shown in red. The side chain of the nonnative D148 is also shown. (b) Bundle of 20 structures of FAIM-CTD, colored as in (a). Nine residues preceding K95 are omitted from this figure. (c) Topology diagram of FAIM-CTD. All molecular models were rendered with PyMOL.

ceding T97, are not constrained by any experimental data. Only weak cross-strand NOE cross peaks are seen between the backbone amide protons of residues Ser156 and Gly162, implying that the β -structure becomes more disordered toward the end of the hydrophilic loop under the conditions used in this study. The extended loop connecting the two sheets between strands 3 and 4 is well ordered but does not adopt a regular secondary structure, and its primary sequence, GQKME-TAGEFVDDG, does not have the characteristic pattern of alternating hydrophobic and hydrophilic residues that would be expected of a β -structure.

A search of the DALI protein structure database⁹ in May 2008 yielded 664 matches, with Z-scores ranging from 2.0 to 4.5. However, the vast majority

of these had structural homology only with the C-terminal four-stranded β -sheet of FAIM-CTD and were therefore discarded. All other matches were examined manually and determined to be false positives, indicating that FAIM-CTD represents a novel fold.

The N-terminal face of FAIM-CTD interacts with FAIM-NTD

There are a number of changes seen in the ¹⁵N HSQC spectrum when comparing the peak positions of residues in free FAIM-CTD and the corresponding residues in the CTD of the (full-length) triple mutant FAIM-S, as seen in Fig. 3a. A plot of these chemical shift differences is shown in Fig. 3b.

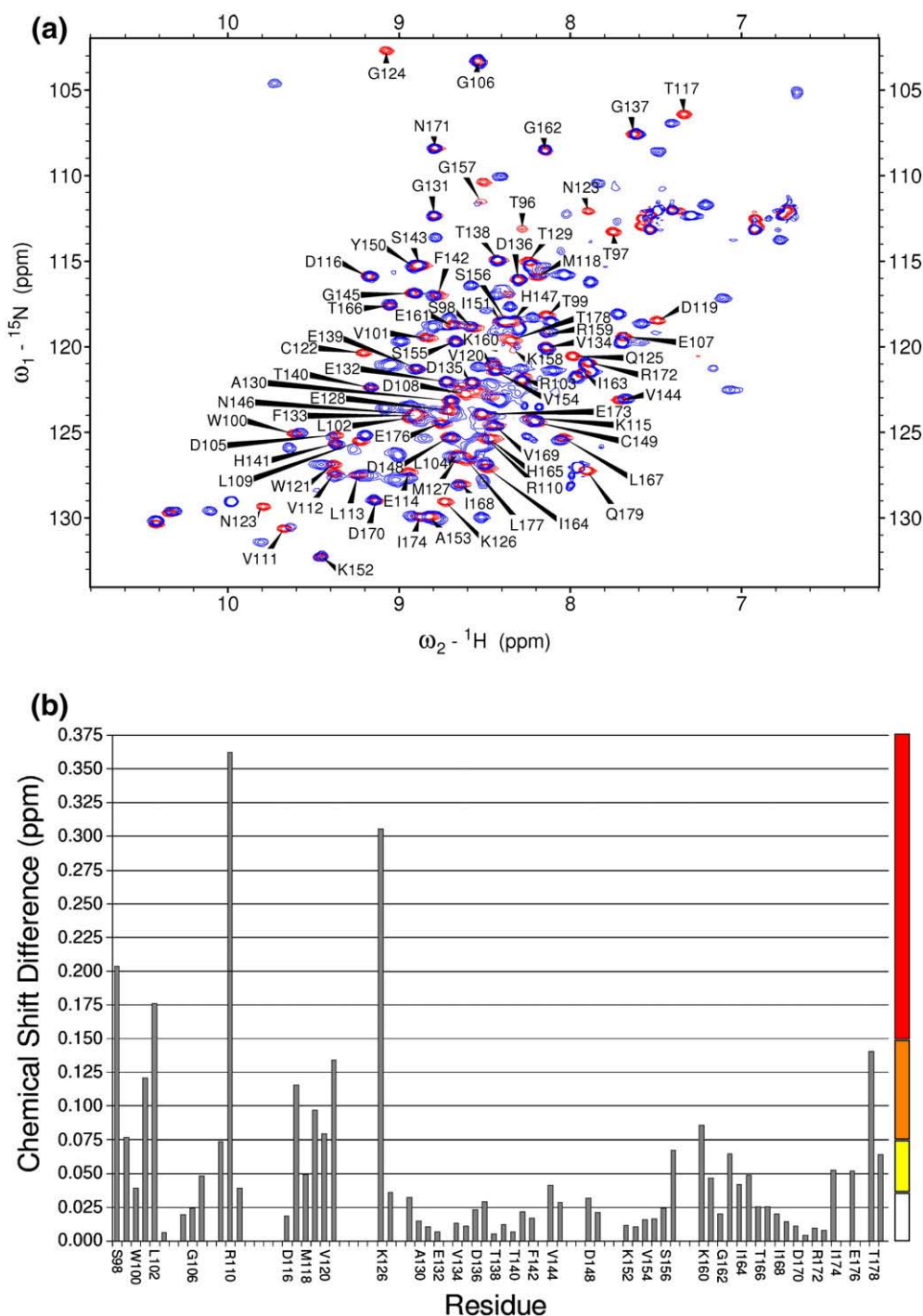


Fig. 3. FAIM-CTD is perturbed by removal of the NTD. (a) Overlaid ${}^{15}\text{N}$ HSQC spectra of isolated FAIM-CTD (red) and full-length triple mutant FAIM-S (blue). Assignments are of FAIM-CTD. The buffer pH is 6.9–7.1. The buffer of full-length FAIM-S has 3% Tween 20. Titration experiments (not shown) indicated that chemical shift perturbation of FAIM-CTD by 3% Tween 20 was minimal and does not explain the observed differences. (b) Chemical shift differences in the ${}^{15}\text{N}$ HSQC spectra in (a) are calculated according to the formula $\Delta\delta = \sqrt{(\Delta\delta_{\text{H}})^2 + (\Delta\delta_{\text{N}}/6)^2}$, plotted as a function of residue, and used to color Fig. 4b according to the scale on the right. No value is shown for residues that could not be assigned in one or both spectra.

There is a trend toward larger differences at the N-terminal end of FAIM-CTD. This is particularly true if one interprets the inability to assign a number of residues in the range V112–Q125 in the full-length FAIM-S construct as additional evidence of perturbation. The chemical shifts of the full-length triple mutant FAIM-S are unaffected by

dilution from 150 to 50 μM and, at lower concentrations, the 20-kDa protein is found to exhibit an apparent molecular mass of approximately 28 kDa on a size-exclusion column (data not shown). Taken together, these data imply that FAIM-S is largely monomeric and that the NTD and CTD of FAIM-S interact in cis.

It is remarkable that the most perturbed residue R110 is one of the few residues that are strictly conserved among all known orthologs of FAIM (Fig. 4). One could therefore hypothesize that this

residue plays a key role in mediating interactions between FAIM-NTD and FAIM-CTD. Notably, Glu38 in FAIM-NTD is also strictly conserved and could interact electrostatically with Arg110. Other

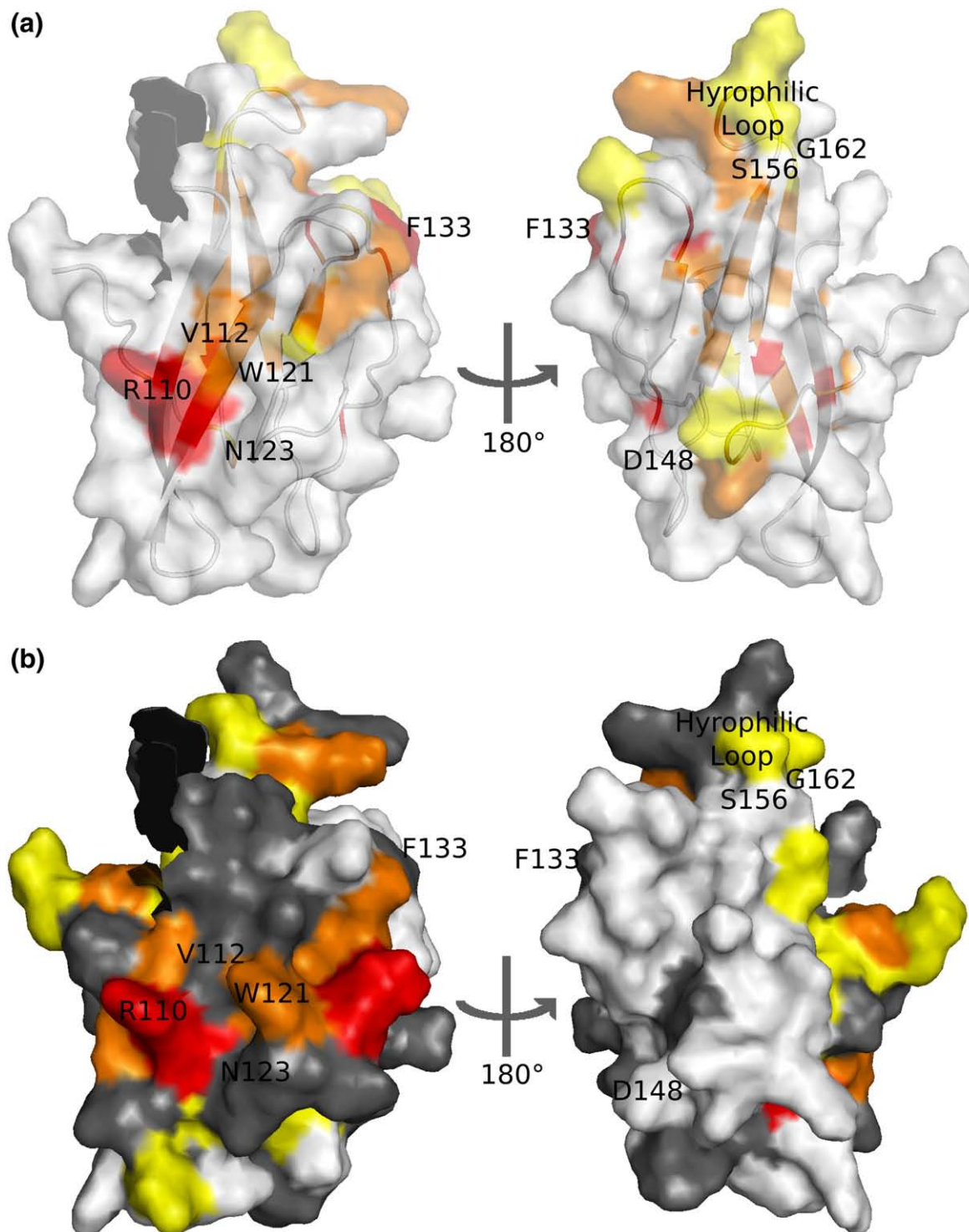


Fig. 4. Structure and sequence analyses highlight the residues of FAIM-CTD involved in interactions with FAIM-NTD and other entities. Surface-exposed residues discussed in the text are labeled in this figure. The position of the nonnative G148D mutation is shown. Nine residues preceding K95 are omitted. (a) Superimposed surface and ribbon models of FAIM-CTD with conserved residues colored as in Fig. 2. (b) Surface model of FAIM-CTD with residues colored according to chemical shift perturbation (Fig. 3b). Residues for which chemical shifts are not known in one or both constructs are shown in gray.

candidates for this role include position 52, which contains a Glu or Asp residue in known homologs, and positions 3, 20, 22, 43, 68, and 76, which have a strong bias toward negatively charged residues as well.

To identify additional surface-exposed sites of high conservation, the FAIM-CTD structure and FAIM-S multiple sequence alignment were submitted to the online ConSurf server.¹⁰ In addition to the residues already discussed, ConSurf identified Val112, Asn123, Gly137, and Gly157 as solvent-exposed and highly conserved. Because the glycine residues have no side chains, they are not highly visible when plotted on the surface of FAIM-CTD. Val112 and Asn123, however, are adjacent to the strictly conserved Arg110 on the N-terminal surface of FAIM-CTD (data not shown). Thus, ConSurf discerns a larger region on this face of the protein that is likely to be involved in interactions with FAIM-NTD or other binding partners.

Discussion

The well-dispersed ¹⁵N HSQC spectra of FAIM-CTD and FAIM-NTD show that each can fold independently of the other and clearly constitute distinct domains, although the structure of each domain may be affected by the loss of interdomain interactions. At 90 residues or less, both domains are relatively small in size.

FAIM-CTD shares some topological features with the “up-and-down barrel” proteins in the fatty acid binding protein (FABP) and P2 families.¹¹ By definition, these proteins are composed of antiparallel β -strands with no interleaving. They differ from FAIM-CTD by their larger molecular mass and the angle at which the strands on opposite sides of the protein appear to cross (approximately 90° in FABP) when viewed in a two-dimensional pro-

jection in the plane of the strands. Unlike FABP, FAIM-CTD appears to be too small to accommodate a comparable fatty acid within a cavity and has no known role in lipid binding. The size and architecture of FAIM-CTD are also quite similar to those of type III fibronectin domains, but the β -strands of the latter have an interleaved topology. Structurally important residues such as Trp100 and Leu167 in FAIM-CTD (see below) are not conserved in type III fibronectin domains.

Both cysteine residues appear to be embedded in the hydrophobic core in the FAIM-CTD structure (Fig. 5a), although Cys122 is near the surface. This observation is consistent with the findings that FAIM-S becomes insoluble during expression if either residue is substituted by serine, but that alanine substitutions are well tolerated (Fig. 1f). Similarly, while neither residue is strictly conserved, substitutions to polar residues are not found in known sequences. Despite their relative positions, several lines of argument support the absence of an intramolecular disulfide bond between these residues. FAIM is physiologically cytosolic⁴ and would not likely have a disulfide bond *in vivo*. Secondly, NMR samples of murine FAIM were kept under reducing buffer conditions and were, in fact, not stable under oxidizing conditions. Thirdly, the C ^{β} chemical shifts of both Cys122 and Cys149 (28.7 and 30.0 ppm) are consistent with the absence of a disulfide bond. Finally, these cysteine residues are not conserved in the more distant homologs of FAIM³ (Fig. 1e) and thus cannot be of vital structural importance.

There exists a network of hydrophobic contacts among Leu167, Trp100, Ile174, Val169, and Leu102 that appears to be responsible for holding the N- and C-terminal edges of the protein in contact and for preventing clamshell-like movements of the two β -sheets (Fig. 5b). In particular, the side chain of Leu167 packs closely against the indole side chain of Trp100, consistent with the observed upfield shift

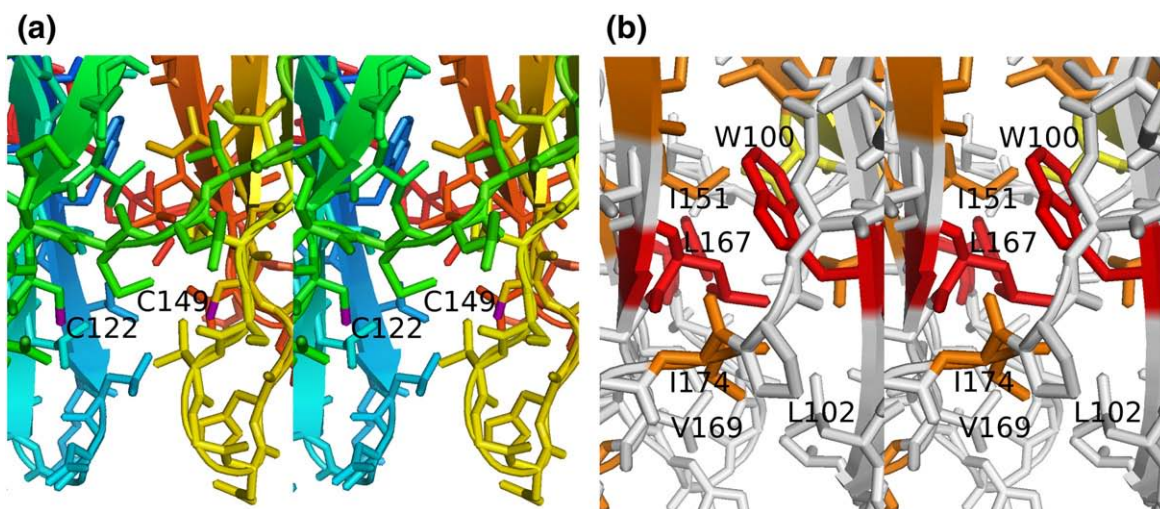


Fig. 5. Structural features of FAIM-CTD. (a) Stick model in stereo of FAIM-CTD, with residues colored as in Fig. 2a, showing cysteine residues 122 and 149. The distance between the two sulfur atoms (purple) in this figure is 8.2 Å. (b) Stick model in stereo of FAIM-CTD, with residues colored according to residue conservation as in Fig. 1e, illustrating the network of hydrophobic contacts involving Trp100, Leu167, Leu102, Val169, and Ile174 that hold together the two β -sheets.

of the side-chain protons of Leu167. These two residues are strictly conserved and may therefore be essential to maintaining the stability of this fold. Ile174 is highly conserved as well, and Val169 and Leu102, like much of the hydrophobic core of FAIM-CTD, suffer only relatively conservative substitutions. The FAIM-CTD structure calculation is also affected by unambiguous NOE restraints between Glu176 and Trp100, although the physical basis for the close contact between these two residues is less clear. One possible explanation for the high degree of conservation at this site is that the lack of interleaving of the β -strands in FAIM-CTD increases the mechanical demands on the interface between the two β -sheets, thereby applying strong selective pressure to maintain an optimized set of hydrophobic contacts throughout evolutionary history.

Conversely, one can predict that conserved non-structural features likely reflect involvement in interactions with FAIM-NTD or other binding partners. These features may be seen in Fig. 4. Notably, the hydrophilic character of the loop between Ser156 and Gly162 is maintained among all the known homologs of FAIM, although no single residue is conserved absolutely. Arg110 is strictly conserved, but serves no structural role in FAIM-CTD. Phe133 also serves no structural purpose and is in an energetically unfavorable solvent-exposed position, yet it is also strictly conserved. While conservation of the hydrophilic loop could simply reflect the deleterious effect of introducing hydrophobic residues into a solvent-exposed position on protein stability, conservation of the other two residues is difficult to explain in the absence of functional significance. Phe133 is therefore likely to be involved in interactions with a binding partner of FAIM-CTD. Arg110 is strongly perturbed in the presence of FAIM-NTD and is surrounded by similarly perturbed residues, as discussed above, and is thus likely to play a vital role in interacting with FAIM-NTD (perhaps via Glu38). It remains possible, however, that the essential function of Arg110 could involve interaction with a binding partner of FAIM-CTD; its interaction with FAIM-NTD could then be incidental.

N-terminal sequencing of the fragments of triple mutant FAIM-S yielded by limited tryptic proteolysis (data not shown) shows that trypsin was able to cleave at residues deep within the first β -sheet (Arg103 and Arg110), while products of cleavage at the unstructured Lys95 were not observed. In contrast, one would expect that the resistance to proteolysis of the folded segments would be much greater than that of the unstructured region. It seems possible that Lys95 could be protected within full-length FAIM or that partial unfolding events are common within FAIM-CTD.

The presence of inward-oriented cysteine and histidine residues raises the question of whether FAIM-CTD might bind metal ions. The solution structure of FAIM-CTD reveals that at least one of the conserved histidine residues (His141) is com-

pletely solvent-exposed and is in a poor position to chelate a metal ion to stabilize the fold. His147 is not well conserved across all phyla, indicating that it likely does not play a role in the chelation of a structurally important metal ion either. His165, in addition to the lack of strict conservation, is also poorly situated to cooperate with other functional groups near the hydrophobic core of the protein. Neither cysteine residue is conserved in insect gene products. Another residue that could potentially assist in the chelation of a metal ion is Thr140, which is buried (although not deeply) within the protein. Conservation of this residue is high but not absolute, with a serine and an alanine being the only known substitutions. Although backbone carbonyl groups have been known to interact directly with cations, as seen, for example, in the EF hand motif of calmodulin¹² and in the selectivity filter of the KcsA potassium channel,¹³ the backbone of FAIM-CTD is primarily involved in a hydrogen-bonding network and does not project these groups into the core of the protein with suitable geometry. Thus, the absence of well-positioned conserved residues that are able to act as electron donors implies that FAIM-CTD does not bind metal ions at a central location within the protein.

With the exception of the Trp121 indole side-chain proton restraints, no NOE-derived distance restraints were discarded during the calculation of the structures presented here. The Trp121 restraints were treated as an exception to the rule only because it was felt that violation of these restraints could readily be rationalized on the grounds that the large side chain may be mobile, possibly flipping, and thereby sampling different environments on the time scale of the NOESY mixing delay.

The lack of structural homology to other proteins means that the FAIM-CTD structure cannot yet play a direct role in efforts to determine the molecular mechanisms of FAIM activity. Interpretation of the structure from a functional point of view must await the identification of, for example, a protein-binding partner or substrate. However, based on the chemical shift differences between FAIM-CTD and full-length FAIM-S, FAIM-CTD clearly interacts with FAIM-NTD. The extent of the structural perturbations induced by this interaction is unknown at this time, but it is now apparent that the interdomain contact surface on FAIM-CTD is on the N-terminal sheet of the protein. This information may be important to properly orient the two domains when modeling the structure of full-length FAIM-S, interpreting functional data, and designing mutation studies. However, preliminary experiments (T.R., unpublished results) indicate that the activity of FAIM-CTD in B cells is as high as, if not higher than, the activity of FAIM-S. This suggests that interaction with FAIM-NTD is not required for activity in this system and underscores the importance of FAIM-CTD in mediating FAIM activity. It is possible that interdomain interaction serves a regulatory purpose or plays a role in the activity of FAIM in another context; this is likely true in cultured neurons, for

example, in which FAIM-S and FAIM-L have different functions but share a common FAIM-CTD.

Because all FAIM-NTD constructs known to date have far fewer peaks than residues in the ^{15}N HSQC spectrum, it is not clear whether it is feasible to solve the solution structure of FAIM-NTD. Limited proteolysis results (Fig. 1a) indicate that the trypsin resistance of FAIM-NTD is comparable to or greater than that of FAIM-CTD despite the presence of at least 10 arginine and lysine residues, implying that there are a few unstructured regions within FAIM-NTD. Taken together, these results suggest that there are structured regions of considerable size within FAIM-NTD that are inaccessible to study by NMR under any currently known conditions.

Distinct functions can often be attributed to individual domains of proteins. It is therefore interesting that a protein-protein (blastp) BLAST search (against the nonredundant database; September 2008) revealed no homology of either domain with any sequence other than the orthologous animal gene products, suggesting that FAIM-CTD only occurs in nature when covalently attached to FAIM-NTD and vice versa. Although the structural basis for FAIM function is not yet fully understood, one could speculate that the small size of the domain could limit the diversity of functions that could have evolved to make use of this particular fold, thereby preventing FAIM-CTD from acquiring functions that would be useful in other contexts. The same could be true of FAIM-NTD if it, too, were found to lack structural homology to other known proteins. However, full-length FAIM-S has a molecular mass of more than 20 kDa and also has no known paralogs, so it would seem that a more parsimonious explanation for the absence of paralogs of either domain would involve some property of full-length FAIM-S that is lost if the protein is truncated, perhaps a property that has not yet been identified. It is also possible that structurally homologous molecules do exist but simply lack sufficient primary sequence homology to be detected by BLAST. Although pairwise identity is high among known orthologs, there are few regions of strict conservation and few structurally irreplaceable residues in FAIM-CTD, so the pairwise primary sequence identity of a paralog could be low.

Materials and Methods

Primers and PCR protocols

Point mutations were introduced using QuikChange PCR (Stratagene). Typically, 100 ng or more of template DNA was used. FAIM-NTD and similar constructs were made by introducing a stop codon in the murine FAIM-S sequence with QuikChange PCR. For the FAIM-CTD constructs, primers containing restriction sites and annealing to the desired start and end sites were designed. The gene fragment was amplified by PCR, digested with BamHI and EcoRI, and then ligated into a pGEX-6P-3 vector. Before use, all constructs were

sequenced by the Dana Farber/Harvard Cancer Center DNA Resource Core.

Protein expression

Murine FAIM constructs were cloned into pGEX-6P-3 glutathione *S*-transferase (GST) fusion expression vectors and transfected into BL21 cells (Novagen; Stratagene). These were grown in M9 medium to an optical density (OD) of 0.6 and induced with 0.1 mM IPTG for 4 h at 37 °C or overnight at 20 °C. Both temperature profiles yielded acceptable expression levels. For expression in LB, cultures were grown to a final OD of about 3.0. To confirm protein identity, a sample of ^{15}N -labeled FAIM-CTD in 1010TBS was analyzed by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry by the Biopolymers Laboratory at Harvard Medical School. As an exception, the 1.8 mM ^{13}C -labeled FAIM-CTD sample used for the ^{13}C NOESY-HSQC spectrum was purified from 1 L of BL21 cell culture and 2 L of XL-1 Blue (Stratagene) cell culture. The protein yield per liter of culture from the latter was 71% of the former, and induction at an OD of 0.7–0.8 with 0.1 M IPTG for 52 h at 20 °C was required.

Deuterated protein expression

A representative protocol is as follows. For triple mutant $^2\text{H}^\alpha, ^{13}\text{C}, ^{15}\text{N}$ -FAIM-S and $^2\text{H}^\alpha, ^{15}\text{N}$ -FAIM-S, aliquots of a starter culture grown in H_2O LB were transferred to 90% D_2O LB to adapt the cultures to growth in D_2O . The culture was then added to 1 L of D_2O M9 with protonated glucose for 24–28 h of growth at 37 °C; at this point, the OD was 0.3–0.4, and the culture was induced to a final OD of 1.5 with 0.1 mM IPTG for 16–19 h at 20 °C. The notation $^2\text{H}^\alpha$ indicates that side chains beyond the H^α position were not fully deuterated because protonated glucose was used during expression.

Residue-specific labeling

For residue-specific labeling with ^{15}N , 500 mL of a modified M9 medium was used for protein expression. The medium was supplemented with 150 mg/L of the ^{15}N -labeled amino acid and 200 mg/L of all 19 other amino acids, as well as 500 mg/L adenine, 650 mg/L guanine, 200 mg/L thiamine, 500 mg/L uracil, and 200 mg/L cytosine. The medium was then filtered to remove undissolved solids. Induction was performed for 4 h at 37 °C. Scrambling appeared to be minimal, since the ^{15}N HSQC spectra of labeled FAIM-S showed no more strong peaks than expected based on the primary sequence (12 of 17 valine, and 11 of 11 isoleucine residues).

Reagents

Tris-buffered saline (TBS) [150 mM NaCl and 50 mM Tris-HCl (pH 7.5)] was used for routine handling of FAIM constructs. An optimized low-salt buffer [1010TBS: 10 mM NaCl and 10 mM Tris-HCl (pH 7.5)] was also used. Tween 20 (0.1%; Calbiochem) was added when handling FAIM-NTD and full-length FAIM-S constructs. DTT (5 mM) was also added to all TBS buffers, except where noted. In preliminary experimental work, Tris-HCl (pH 7.0) was used instead of Tris-HCl (pH 7.5), and 0.1% 2-mercaptoethanol was substituted for DTT. Deuterated 1010TBS, using Tris- d_{11} (D, 98%; Cambridge Isotope Labs) and DTT- d_{10} (D, 98%;

CIL), was prepared in small batches prior to use. pH was adjusted with DCl and NaOD; the nominal pH values given for D₂O buffers are presented as pH*.¹⁴

Protein purification

All steps were carried out on ice or at 4 °C unless otherwise noted. A representative protocol is as follows. Fresh or frozen cell pellets were resuspended in 30–100 mL of ice-cold lysis buffer per liter of cell culture. Lysis was accomplished using either sonication on ice or disruption in an Avestin Emulsiflex C5 homogenizer in TBS with 0.1% Tween 20, 200 μ M 4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride, and trace amounts of DNase I; 1 mg/mL lysozyme was also added during early investigations. Lysate was centrifuged at 10,000 rpm for 25 min. The supernatant was filtered at 0.22 or 0.45 μ m, then incubated with glutathione-conjugated agarose beads (Glutathione Sepharose 4 Fast Flow; Amersham; 6-mL bed volume per liter of culture). Beads were washed with 150 mL of TBS or 1010TBS (without Tween 20, regardless of construct). GST-fused rhinoviral 3C protease was then added to the column to remove the GST fusion protein. Tween 20 (0.1%) was also added before the cleavage of full-length or NTD FAIM-S constructs. Cleaved protein was eluted with 1010TBS or TBS. Eluted protein was typically more than 90% pure by SDS-PAGE, with the major visible contaminants (if any) being uncleaved GST fusion protein and free GST. Eluted protein was purified by gel filtration in TBS or 1010TBS. Full-length FAIM-S constructs typically eluted at 66–73 mL, FAIM-CTD constructs eluted at 73–77 mL, and FAIM-NTD constructs eluted at 84–94 mL.

Limited proteolysis

For N-terminal sequencing, limited proteolysis was carried out at 0 °C using 350 μ g/mL wild-type FAIM-S and 4 μ g/mL trypsin (Sigma) in 110 mM ammonium bicarbonate in a volume of 80 μ L. Forty-microliter aliquots were removed at 10 and 90 min. The reaction was stopped by addition of SDS loading buffer and heating to 70 °C for 5 min. SDS-PAGE and electroblotting were performed essentially in accordance with Matsudaira.¹⁵ N-terminal sequencing using Edman chemistry was performed with the Molecular Biology Core Facilities at the Dana Farber Cancer Institute.

NMR spectroscopy

The pH of the buffer used for NMR spectroscopy (1010TBS with 5 mM DTT) was 7.3 at room temperature. Tween 20 (0.1%) was added for full-length FAIM-S and FAIM-NTD constructs. Because detergent was retained during the concentration of the protein sample, the concentration of Tween 20 was approximately 3–5% for the most concentrated samples during NMR spectroscopy. For experiments in deuterated 1010TBS, however, the buffer was prepared in small batches, and the buffer pH was titrated to 7.0. In titration experiments (not shown), it was found that in the ¹⁵N HSQC spectrum of FAIM-CTD, only the peak corresponding to V144 had a highly pH dependent chemical shift between pH 6.8 and 8.0. NMR samples were filtered, and the sample tube was flushed with argon. All spectra were acquired at 25 °C unless stated otherwise. Linear prediction was used to double the number of points in one indirect dimension when pro-

cessing most three-dimensional experiments. Spectra were processed with NMRPipe.¹⁶ Assignment was done manually using SPARKY version 3.110.¹⁷

Backbone assignment of triple mutant FAIM-S

Standard non-TROSY triple-resonance experiments [HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, and HN(CO)CACB] were acquired on a 250 μ M ²H α , ¹³C, ¹⁵N-labeled sample using a Varian Inova 500 spectrometer except where noted. A ¹⁵N NOESY-HSQC spectrum with a mixing time of 120 ms, acquired with 600 MHz on a Bruker Avance 600, was also used to assist in backbone assignment. For this experiment, a 0.7 mM triple mutant ²H, ¹⁵N-labeled FAIM-S sample was used. Residue-specific labeling with isoleucine and valine was done to help confirm the assignment and to match sequences of successive unidentified residues to the known primary sequence. In this case, ¹⁵N HSQC spectra were acquired on a Varian UnityPlus 400-MHz spectrometer.

Backbone assignment of FAIM-CTD

Standard non-TROSY triple-resonance experiments, excluding HN(CA)CO [¹⁵N HSQC, HNCA, HNCO, HN(CO)CA, HNCACB, CBCA(CO)NH, ¹⁵N NOESY-HSQC], were acquired on a Varian Inova 500-MHz spectrometer. Knowledge of the backbone assignment of the CTD of the triple mutant FAIM-S was incorporated into the assignment procedure. An HN(CA)CO experiment was then acquired after completion of the FAIM-CTD backbone assignment and used to confirm the assignment.

Side-chain assignment of FAIM-CTD

Most aliphatic side-chain protons were assigned based on ¹³C HSQC/D₂O, C(CO)NH, H(CCO)NH, HCCH-COSY/D₂O, HCCH-TOCSY (total correlated spectroscopy)/D₂O, and two-dimensional NOESY/D₂O spectra. A ¹³C NOESY-HSQC/D₂O spectrum acquired with a 900-MHz spectrometer was also used. Stereospecific assignments were obtained for 10 of 15 valine and leucine residues in accordance with Neri *et al.*¹⁸ Aromatic protons were assigned using the ¹⁵N NOESY-HSQC, two-dimensional NOESY/D₂O, aromatic ¹³C HSQC/D₂O, and CCH-COSY¹⁹ spectra. Histidine imidazole side-chain protons were assigned using the ¹⁵N NOESY-HSQC spectrum and a modified version of the procedure used by Pelton *et al.* employing a ¹⁵N HSQC spectrum in lieu of a ¹⁵N HMQC spectrum.¹⁹ The H ϵ protons of M127 were assigned, after calculation of preliminary structures, on the basis of numerous cross peaks in the ¹³C NOESY-HSQC/D₂O spectrum.

NOESY of FAIM-CTD

A ¹⁵N NOESY-HSQC spectrum was acquired in H₂O at 500 MHz with a 100-ms mixing time on a 0.9 mM ¹⁵N-labeled FAIM-CTD sample over a period of 3.5 days. A ¹³C NOESY-HSQC/D₂O spectrum was acquired at 900 MHz with a 120-ms mixing time on a 1.8 mM ¹³C-labeled FAIM-CTD sample in 1010TBS/D₂O in a 5-day acquisition. To minimize the effects of spin diffusion, a two-dimensional NOESY/D₂O spectrum at 900 MHz on a 1.4 mM unlabeled FAIM-CTD sample in D₂O 1010TBS (pH 7.0) was also acquired with 70-ms mixing time; any

distance restraint arising from a peak in the ^{13}C NOESY-HSQC/D $_2$ O spectrum for which a corresponding peak was not seen in the two-dimensional NOESY/D $_2$ O spectrum was removed.

Other restraints

TALOS²⁰ was used to obtain backbone torsion angle constraints. These reduced the backbone RMSD by approximately 0.5 Å when added to the force field of the simulated annealing calculations. Proton chemical shifts were referenced to an external reference sample of 2,2-dimethyl-2-silapentane-5-sulfonate dissolved in 1010TBS; nitrogen and carbon chemical shifts were referenced using virtual referencing.²¹ Side-chain rotamer constraints were added based on the work presented by Matsuo *et al.*²² and on data compiled by Dunbrack and Karplus,²³ but individual constraints were removed if they were consistently violated. Hydrogen-bonding restraints were introduced between adjacent β -strands when there was clear spectroscopic evidence that this structure existed. The secondary structural elements formed in the absence of hydrogen-bonding restraints; the restraints were used only to decrease the RMSD of the structure bundle.

Structure calculation

CYANA version 2.1 was used to calculate structures. For the final structure bundles, 100 structures were calculated using 100,000 steps of simulated annealing from a random starting structure. The lowest 20% in energy were accepted. In early investigations, it was found that the residue Trp121 was involved in large numbers of constraint violations. NOE assignments were double-checked and found to be correct. As Trp121 is solvent-exposed in the calculated structures (see Discussion) and appears to make a NOE cross peak with water (not shown), its large side chain could be highly mobile, resulting in inaccurate distance restraints. Therefore, all NOE restraints involving a side-chain proton of Trp121 were removed.

Validation with RDCs

Pf1 Magnetic Resonance Cosolvent (6 mg/mL; ASLA Biotech, Latvia) was added to one 250- μg /mL sample of ^{15}N , ^{13}C -labeled FAIM-CTD in 1010TBS. Buffer was added to an identical sample to ensure that the buffer conditions were identical in the isotropic sample. The splitting of the HDO peak in the aligned sample was 7 Hz. The one-bond scalar coupling constant $^1J_{\text{NH}}$ was measured by two redundant means. One method used the J -modulated HSQC pulse sequence described by Tjandra *et al.*²⁴ The second method was based on the chemical shift difference between TROSY and non-TROSY HNCO spectra. Values from each method were compared to ensure consistency and to estimate error. The alignment tensor, predicted RDCs, and Q_{free} factor²⁵ were calculated using PALES.²⁶ RDCs for residues not known to be involved in the formation of secondary structural elements were excluded; this modestly improved the overall correlation between observed and predicted RDCs.

Accession numbers

Atomic coordinates for FAIM-CTD have been deposited in the Protein Data Bank with accession number 2KD2.

Chemical shift assignments have been deposited in the Biological Magnetic Resonance Data Bank with accession number 16103.

Acknowledgements

We would like to acknowledge Dmitri Ivanov and Alexey Lugovskoy for preliminary work related to the expression and purification of FAIM. This research was supported by a fellowship from The Whitaker Foundation and by National Institutes of Health grants GM47467, EB002026, and GM066360.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jmb.2009.01.004](https://doi.org/10.1016/j.jmb.2009.01.004)

References

- Schneider, T. J., Fischer, G. M., Donohoe, T. J., Colarusso, T. P. & Rothstein, T. L. (1999). A novel gene coding for a Fas apoptosis inhibitory molecule (FAIM) isolated from inducibly Fas-resistant B lymphocytes. *J. Exp. Med.* **189**, 949–955.
- Zhong, X., Schneider, T. J., Cabral, D. S., Donohoe, T. J. & Rothstein, T. L. (2001). An alternatively spliced long form of Fas apoptosis inhibitory molecule (FAIM) with tissue-specific expression in the brain. *Mol. Immunol.* **38**, 62–76.
- Sole, C., Dulcet, X., Segura, M., Gutierrez, H., Diaz-Meco, M.-T., Gozzelino, R. *et al.* (2004). The death receptor agonist FAIM promotes neurite outgrowth by a mechanism that depends on ERK and NF- κ B signaling. *J. Cell Biol.* **167**, 479–492.
- Segura, M. F., Sole, C., Pascual, M., Mubarak, R. S., Perez-Garcia, M. J., Gozzelino, R. *et al.* (2007). The long form of Fas apoptosis inhibitory molecule is expressed specifically in neurons and protects them against death receptor-triggered apoptosis. *J. Neurosci.* **28**, 11228–11242.
- Wiens, M., Perović-Ottstadt, S., Müller, I. M. & Müller, W. E. G. (2004). Allograft rejection in the mixed cell reaction system of the desmosponge *Suberites domuncula* is controlled by differential expression of apoptotic genes. *Immunogenetics*, **56**, 597–610.
- Wiens, M. & Müller, W. E. G. (2006). Cell death in Porifera: molecular players in the game of cell death in living fossils. *Can. J. Zool.* **84**, 307–321.
- Ito, T. & Wagner, G. (2004). Using codon optimization, chaperone co-expression, and rational mutagenesis for production and NMR assignments of human eIF2 α . *J. Biol. NMR*, **28**, 357–367.
- Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biol. NMR*, **8**, 477–486.
- Holm, L. & Sanders, C. (1995). Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478–480.

10. Armon, A., Grauer, D. & Ben-Tal, N. (2001). ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. *J. Mol. Biol.* **307**, 447–463.
11. Jones, T. A., Bergfors, T., Sedzik, J. & Unge, T. (1988). The three-dimensional structure of P2 myelin protein. *EMBO J.* **7**, 1597–1604.
12. Babu, Y. S., Bugg, C. E. & Cook, W. J. (1988). Structure of calmodulin refined at 2.2 Å resolution. *J. Mol. Biol.* **204**, 191–204.
13. Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J., Cohen, S. L. *et al.* (1998). The structure of the potassium channel: molecular basis of K⁺ conduction and selectivity. *Science*, **280**(5360), 69–77.
14. Krężel, A. & Bal, W. (2003). A formula for correlating pK_a values determined in D₂O and H₂O. *J. Inorg. Chem.* **98**, 161–166.
15. Matsudaira, P. (1987). Sequence from picomolar quantities of protein electroblotted onto polyvinylidene fluoride membrane. *J. Biol. Chem.* **262**, 10035–10038.
16. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
17. Goddard, T. G. & Kneller, D. G. (undated). SPARKY 3. University of California, San Francisco.
18. Neri, D., Szyperski, T., Otting, G., Senna, H. & Wuthrich, K. (1989). Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ¹³C labeling. *Biochemistry*, **28**, 7510–7516.
19. Pelton, J. G., Torchia, D. A., Meadow, N. D. & Roseman, S. (1993). Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated III^{Glc}, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques. *Protein Sci.* **2**, 543–558.
20. Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.
21. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E. *et al.* (1995). ¹H, ¹³C, and ¹⁵N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR*, **6**, 135–140.
22. Matsuo, H., Li, H., McGuire, A., Fletcher, C. M., Gingras, A.-C., Sonenberg, N. & Wagner, G. (1997). Structure of translation factor eIF4E bound to m⁷GDP and interaction with 4E-binding protein. *Nat. Struct. Biol.* **4**, 717–724.
23. Dunbrack, R. L. & Karplus, M. (1993). Backbone-dependent rotamer library for proteins: application to side-chain prediction. *J. Mol. Biol.* **230**, 543–574.
24. Tjandra, N., Grzesiek, S. & Bax, A. (1996). Magnetic field dependence of nitrogen–proton *J* splittings in ¹⁵N-enriched human ubiquitin resulting from relaxation interference and residual dipolar coupling. *J. Am. Chem. Soc.* **118**, 6264–6272.
25. Cornilescu, G., Marquadt, J. L., Ottinger, M. & Bax, A. (1998). Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **120**, 6836–6837.
26. Zwecksetter, M. & Bax, A. (2000). Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR. *J. Am. Chem. Soc.* **122**, 3791–3792.