Towards a structural understanding of Friedreich's ataxia: the solution structure of frataxin

Giovanna Musco[1], Gunter Stier[2], Bernhard Kolmerer[2†], Salvatore Adinolfi[1], Stephen Martin[1], Tom Frenkiel[1], Toby Gibson[2] and Annalisa Pastore[1*]

Background: Lesions in the gene for frataxin, a nuclear-encoded mitochondrial protein, cause the recessively inherited condition Friedreich's ataxia. It is thought that the condition arises from disregulation of mitochondrial iron homeostasis, with concomitant oxidative damage leading to neuronal death. Very little is, as yet, known about the biochemical function of frataxin.

Results: Here, we show that the mature form of recombinant frataxin behaves in solution as a monodisperse species that is composed of a 15-residue-long unstructured N terminus and an evolutionarily conserved C-terminal region that is able to fold independently. The structure of the C-terminal domain consists of a stable seven-stranded antiparallel β sheet packing against a pair of parallel helices. The structure is compact with neither grooves nor cavities, features that are typical of iron-binding modules. Exposed evolutionarily conserved residues cover a broad area and all cluster on the β-sheet face of the structure, suggesting that this is a functionally important surface. The effect of two clinically occurring mutations on the fold was checked experimentally. When the mature protein was titrated with iron, no tendency to iron-binding or to aggregation was observed.

Conclusions: Knowledge of the frataxin structure provides important guidelines as to the nature of the frataxin binding partner. The absence of all the features involvement of frataxin in iron metabolism. The effects of point mutations associated with Friedreich's ataxia can be rationalised by knowledge of the structure and suggest possible models for the occurrence of the disease in compound heterozygous patients.

Addresses: [1]NIMR, The Ridgeway, Mill Hill, London NW7 1AA, UK and [2]EMBL, Postfach 102209, 69012 Heidelberg, Germany.

[†]Present address: Schering AG, CNS Research, Muellerstrasse 170–178, 13342 Berlin, Germany.

*Corresponding author.
E-mail: apastor@nimr.mrc.ac.uk

Key words: iron metabolism, mitochondrial disease, neurodegenerative diseases, nuclear magnetic resonance, protein structure

Received: **16 March 2000**
Revisions requested: **18 April 2000**
Revisions received: **4 May 2000**
Accepted: **5 May 2000**

Published: **15 June 2000**

Structure 2000, **8**:695–707

0969-2126/00/$ – see front matter

Introduction

In recent years a new mechanism of genetic disease has been revealed in which repeats of DNA trinucleotide sequences become amplified in the germline, with consequent anomalies in gene expression. To date, more than 12 human genetic diseases have been associated with the expansion of CAG, CGG and GAA repeats. The diseases can be grouped into two classes according to whether or not the triplet repeats occur in coding or in noncoding regions (see [1] and references therein). In the first class, which includes Huntington's disease, the spinocerebellar ataxias and spinobulbar muscular atrophy, expanded CAG repeats encode a polyglutamine segment within an open-reading frame of an expressed sequence. The poly-Gln repeats behave as 'gain-of-function' mutations and the associated diseases, characterised by progressive neuronal loss, are dominantly inherited. In the second class, which includes fragile X syndrome and myotonic dystrophy, the triplet expansion occurs either in the untranslated regions (UTRs) or in introns. These inherited diseases present a multisystem pathology. The most recent addition to this second group is Friedreich's ataxia (FRDA; OMIM code 229300; http://www3.ncbi.nlm.gov/omim) (for reviews see [2]), which is the most common form of hereditary ataxia, affecting 1 out of 50,000 individuals [3,4]. FRDA is a neurodegenerative disorder in which neurons with very long axons die back from the periphery. It usually manifests itself by early adolescence and progresses to severe physical disability, often complicated by hypertrophic cardiomyopathy leading to premature death. The disorder is generally characterized by one or more of the following: an unsteady gait; dysarthria; muscular weakness of the legs; lower limb areflexia; diabetes mellitus; carbohydrate intolerance [5]. It is now well established that FRDA is a genetically homogenous recessive entity, of which the locus has been mapped to chromosome 9 [6,7]. The gene encodes a novel small protein of 210 amino acids called frataxin [8].

About 94% of FRDA patients are homozygous for an intronic GAA repeat expansion in the frataxin gene [9].

The number of repeats ranges from 8–22 in normal individuals and from 120–1700 in affected patients [8]. Homozygous FRDA patients have undetectable or extremely low levels of frataxin mRNA [8], which suggests that the disease is caused by deficiency of frataxin protein. In a few cases, FRDA patients are heterozygous for an expanded repeat in one allele and a deleterious point mutation in the other. Missense, nonsense and splicing point mutations have all been found.

The biochemical function of frataxin is unknown. Frataxin is ubiquitously (though not abundantly) expressed, reaching the highest concentration in the heart, spinal cord and dorsal root ganglia [10], all tissues that are heavily dependent on oxidative respiration. The expression levels correlate well with the pattern of neuronal degeneration, cardiomyopathy and increased risk of diabetes observed in the disease. No sequence homology is detected with proteins of known function. However, mitochondrial localization of frataxin was predicted by the finding of homologues (designated cyaY) in purple bacteria, the closest relatives of the mitochondrial genome, suggesting that chronic peroxidative damage is the cause of neuronal death [11]. A mitochondrial location was rapidly demonstrated experimentally in yeast and human cells [10,12–15]. The yeast studies also showed that frataxin mutants are highly sensitive to oxidants and readily accumulate lethal quantities of iron in the mitochondria.

The evidence that frataxin is closely involved in iron homeostasis has steadily increased. Some years ago, unique iron deposits were reported in cardiac myocytes [16,17]. Mitochondrial iron is significantly higher in fibroblasts from patients with FRDA than in fibroblasts from control patients [18,19]. Additional evidence from the yeast system that implicates mitochondrial dysfunction in FRDA includes the observation that deletion of the yeast frataxin gene (YFH1) results in a decrease of respiration, a depletion of mitochondrial DNA, oxidative stress and an increase in mitochondrial iron concentration [10,12–15,20,21]. Collectively these results have led to the view that frataxin has a role in controlling mitochondrial iron homeostasis and that FRDA is caused by oxidative damage. Consistent with this view, some success in alleviating clinical progression of FRDA has been reported recently; combined treatment with the antioxidants *N*-acetylcysteine, selenium and vitamin E led to a decrease in the rate of clinical decline in an FRDA patient [22]. Treatment with the lipid-soluble antioxidant idebenone (a coenzyme Q analogue) resulted in a dramatic decrease in myocardial hypertrophy [23]. These preliminary findings could be consistent with either a direct or an indirect involvement of frataxin in iron utilisation. Recently, it has been reported that the yeast homologue of frataxin shows a ferritin-like behaviour *in vitro*: in the presence of an iron excess, the protein has been reported to form large aggregates able to segregate iron from solution [24]. Several potential ligands besides iron have been envisaged for frataxin, ranging from vitamin E to mitochondrial DNA (mtDNA) to another protein [11]. Experiments searching for proteic partners have so far identified only the mitochondrial-processing peptidase MPP-β, a protein that is responsible for cleavage of the frataxin mitochondrial targeting sequence [25–27].

Some help in providing a clue to function might come from the increasing number of clinically relevant frataxin mutants and their correlation with phenotype [9]. The possibility of distinguishing residues that are essential for the stability of the three-dimensional fold from those that are directly implicated in the function might help to shed light onto the cellular role of frataxin. Mapping loss-of-function mutations to functional positions would provide useful tools for testing candidate ligands of frataxin. Knowledge of the tertiary structure is essential to reveal these positions.

Our project aims at the structural and functional characterization of frataxin. In the following sections, we describe the structure determination of frataxin in solution. We show that the evolutionary conserved C terminus is necessary and sufficient for forming an independently folded domain. The structure allows us to compare it with other iron-binding proteins and to rationalize the role of each of the clinically observed point mutations. The effect on the structure of two mutations was checked experimentally.

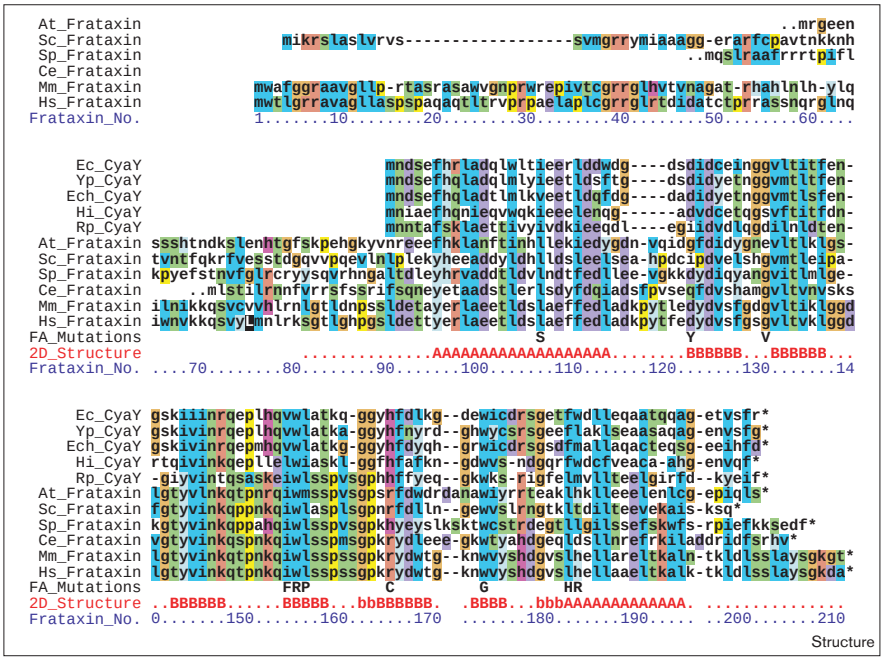
Results

Frataxin is a monomer and contains two structurally distinct regions

Two distinct regions can be identified in the frataxin sequences: a non-conserved N-terminal region, absent from the bacterial cyaY proteins, and a strongly conserved C-terminal block (Figure 1). The N terminus includes the mitochondrial import sequence, which is cleaved off the mature frataxin protein spanning residues 75–210. The sequence present in the mature protein has characteristics of a globular domain with a predicted α–β structure [11]. No classical iron-binding motifs (e.g., conserved cysteines) are identifiable in its sequence. Two constructs of human frataxin were prepared: one construct (hereafter referred to as L_Fra) spanned residues 63–210 of the human protein and a shorter one (hereafter referred to as S_Fra) contained the conserved C-terminal domain (amino acids 90–210) — at the time the project was started, the mature form was not known. Interestingly, the longer construct was degraded naturally and quantitatively during purification to a fragment spanning residues 75–210, which corresponds to the mature form of the protein [26]. Characterization of L_Fra carried out by analytical gel filtration showed that recombinant frataxin behaves as a monodisperse species (data not shown). According to circular dichromism (CD) and nuclear magnetic resonance (NMR) spectroscopies, both constructs are folded, but the

Figure 1

Alignment of currently known frataxin and cyaY sequences, displayed with Clustal X colours to emphasise conserved sequence features [46]. The white-on-black L indicates the N terminus of mature frataxin. Point mutations causing Friedreich's ataxia (Table 2) are listed in the line FA_Mutations. The 2D_structure line summarises the content of secondary structure in terms of 'A' helical and 'B' strand residues. N termini are incomplete for At, Ce and Sp frataxins. See the Materials and methods section for species identifiers.



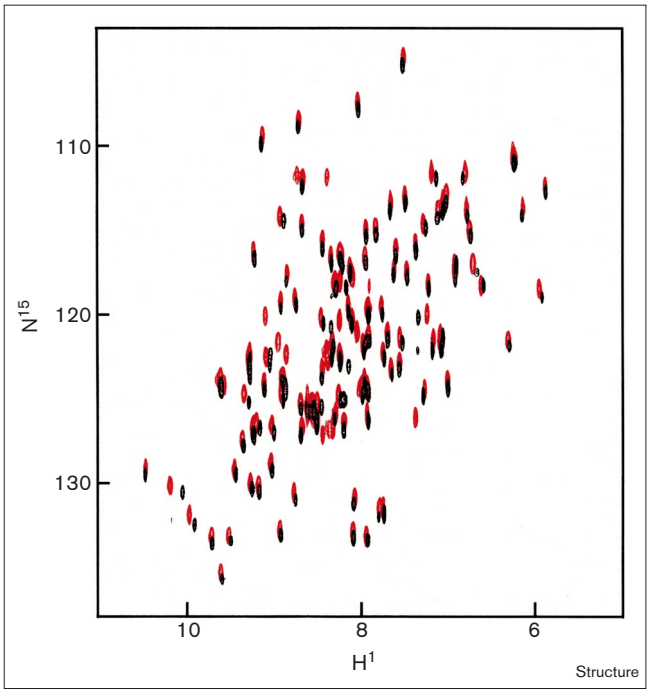
additional residues in L_Fra do not contribute to thermodynamic stability as demonstrated by comparable values of the melting points of the two constructs (62°C). This observation strongly suggests that the additional residues are unfolded. Furthermore the heteronuclear single quantum correlation (HSQC) NMR spectra of the two constructs are superimposable (Figure 2) with the implication that the N terminus does not alter the folded structure. As these N-terminal residues provide no significant information about the tertiary fold, we decided to concentrate our attention on the S_Fra construct for the purpose of structure determination. The long construct L_Fra, which corresponds to the mature frataxin, was used for the binding assays and for production of the mutants.

Frataxin is an α–β protein

Structure determination of S_Fra was carried out using NMR. Complete spectral assignment was achieved as described by Musco *et al.* [28]. Typical NOE effects, such as the secondary chemical shift pattern and $^3J_{HNHA}$ scalar coupling constants, together with amide hydrogen–deuterium exchange experiments indicate that frataxin has an α–β fold (Figure 3). The structure consists of two helical regions (residues 95–113 and 182–194), that flank an antiparallel β sheet, which is formed by six consecutive strands, designated β1 to β6 spanning residues 122–127, 131–136, 142–147, 154–159, 162–169 and 172–175. The presence of a seventh short strand between residues 179 and 181 is suggested by typical β-sheet connectivities, even though it is not supported by the secondary chemical shifts. The overall topology is α1β1β2β3β4β5β6β7α2.

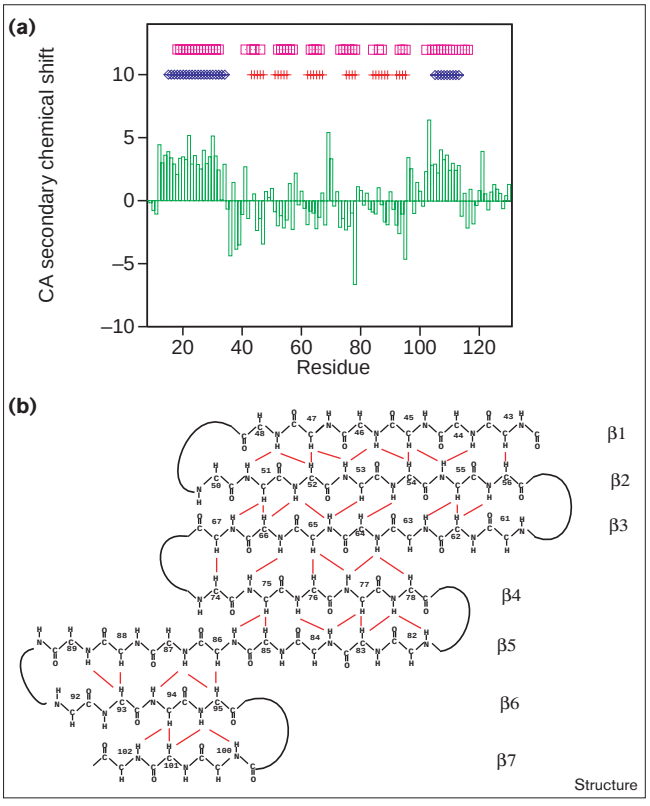
Relevant parameters that define the quality of the calculated structures are reported in Table 1. Figure 4a shows a bundle of the 15 lowest energy structures. Apart from the

Figure 2



Superposition of the HSQC spectra of S_Fra and L_Fra. The spectra were recorded at 600 MHz and 27°C.

Figure 3



NMR parameters related to the secondary structure of S_Fra. **(a)** Plot of the secondary Cα chemical shifts versus the frataxin sequence. In the top line, squares indicate slow-exchanging amide protons in hydrogen-deuterium exchange experiments. In the second line, diamonds and bars refer to $^3J_{HNHA}$ values typical of α-helical ($^3J_{HNHA}$ smaller than 5.5 Hz) and β sheet ($^3J_{HNHA}$ larger than 8 Hz) conformation, respectively. **(b)** The seven-stranded antiparallel β sheet of frataxin. Continuous lines identify observed interstrand NOEs.

Table 1

Relevant parameters that define the quality of the calculated structures of S_Fra.

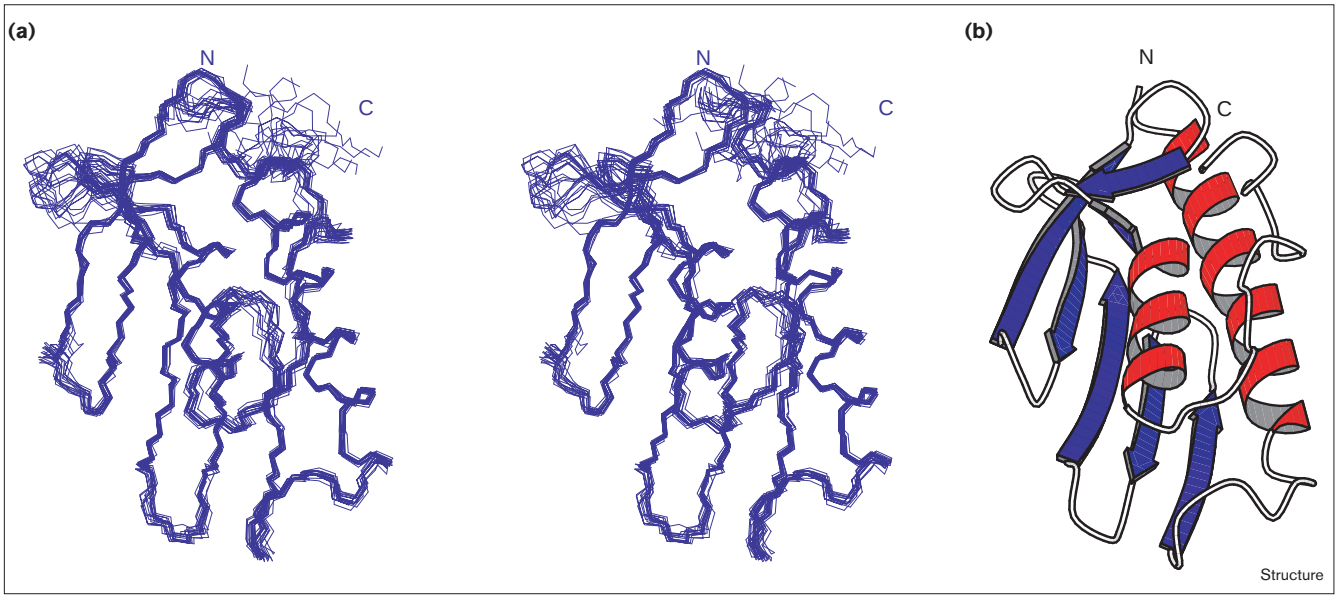
No. of restraints	*	†	‡
Total	951	2114	234
Intraresidue	328	1022	55
Sequential	326	399	25
Medium-range	112	236	43
Long-range	185	457	111
Angle restraints§	0	42	
Hydrogen bonds	50	50	
Rmsd from ideality			
Bonds Å		1.13E-3 ± 1.36E-4	
Angles (°)		0.31 ± 1.06E-2	
Impropers (°)		0.32 ± 0.06	
Energies (kcal/mol)			
E total		143.78 ± 19.410	
E bond		3.393 ± 0.783	
E angle		56.450 ± 0.784	
E vdW		67.228 ± 9.370	
E NOE		12.701 ± 5.398	
E CDIH		0.024 ± 0.017	
Quality control			
PROSA (kT)#		–1.92 ± 0.09	
WHATIF¶		–1.009 ± 0.050	
PROCHECK¥		71.1	
Fit to experimental data**			
Unambiguous NOEs		9.942E-3 ± 2.368E-3	
Ambiguous NOEs		8.714E-3 ± 4.121E-3	
All NOEs		9.909E-3 ± 2.257E-3	
Hydrogen bonds		1.078E-2 ± 3.908E-3	
Dihedrals		8.109E-2 ± 3.587E-2	
Rmsd††			
Backbone‡‡		0.43 ± 0.07	
Heavy atoms‡‡		0.98 ± 0.11	
Backbone (95–205)		0.70 ± 0.12	
Heavy atoms (95–205)		1.20 ± 0.15	

*Distribution of the hard restraints used in DYANA. †Distribution of the unambiguous restraints at the last ARIA iteration. ‡Distribution of the ambiguous restraints at the last ARIA iteration. #Backbone dihedral angle restraints were based on $^3J_{HNHA}$ coupling constants measured in a 3D HNHA spectrum. For α helical regions ($^3J_{HNHA}$ smaller than 5.5 Hz) dihedral angle values of $\phi = -60$ (± 40) were used. For β strands ($^3J_{HNHA}$ larger than 8 Hz) the values were taken as $\phi = -120$ (± 40). #Calculated according to [44]. ¶Calculated according to [45]. ¥Percentage of residues in most favourable regions of the Ramachandran plot [43]. **Given in Å. ††Rmsd in Å are calculated from the first best structure in terms of total energy. ‡‡Rmsd on the elements of secondary structure 95–113, 122–127, 131–136, 142–147, 154–158, 162–169, 172–175 and 182–194.

N-terminal histidine tag and the last three C-terminal residues, most of the fold is well defined in the bundle with a root mean square deviation (rmsd) of 0.70 ± 0.12 Å for the backbone atoms from residue 95–205. The loops are relatively well structured with the exception of the β6β7 loop (residues 176–179) which is less ordered, for example as shown by the absence of nuclear Overhauser enhancements (NOEs). Many loop residues, including small hydrophilic residues (threonines and serines), and glycines and prolines, are well conserved among the frataxin family.

The seven-stranded antiparallel β sheet is solvent-exposed on one side and packs tightly against the two helices on the other (Figure 4b). The two helices lie roughly parallel to each other and their orientation with respect to the β sheet is unambiguously defined by NOE data (e.g., between Tyr118 and Thr196, Thr102 and Leu182, Thr102 and His183). The C-terminal tail spanning residues 194–206 fills a groove formed between the two helices as supported by several NOEs (e.g., Tyr205 forms a dense network of contacts with distal residues). Despite the presence of a number of sequential HN–HN connectivities, all the other NMR observables exclude the presence of a helical conformation in this region.

Figure 4



Tertiary structure of S_Fra. **(a)** NMR bundle (residues 91–210 of the human frataxin sequence). Stereo pair of a backbone bundle of the 15 best calculated structures. The fitting was obtained by superimposing the backbone atoms of residues 95–205 onto the average structure. **(b)** Ribbon representation of S_Fra with helices in red and strands in blue.

Structure similarity searches did not produce a significant match with any other protein except for p32 (with an rmsd of 3.2 Å over 56 residues and a Z score of 3.0), a mitochondrial matrix protein important for the maintenance of oxidative phosphorylation [29]. Despite some similarities in the fold and mitochondrial localization, however, frataxin and p32 have very different contact patterns, which implies no detectable evolutionary relatedness.

Sequence conservation can be mapped onto the structure

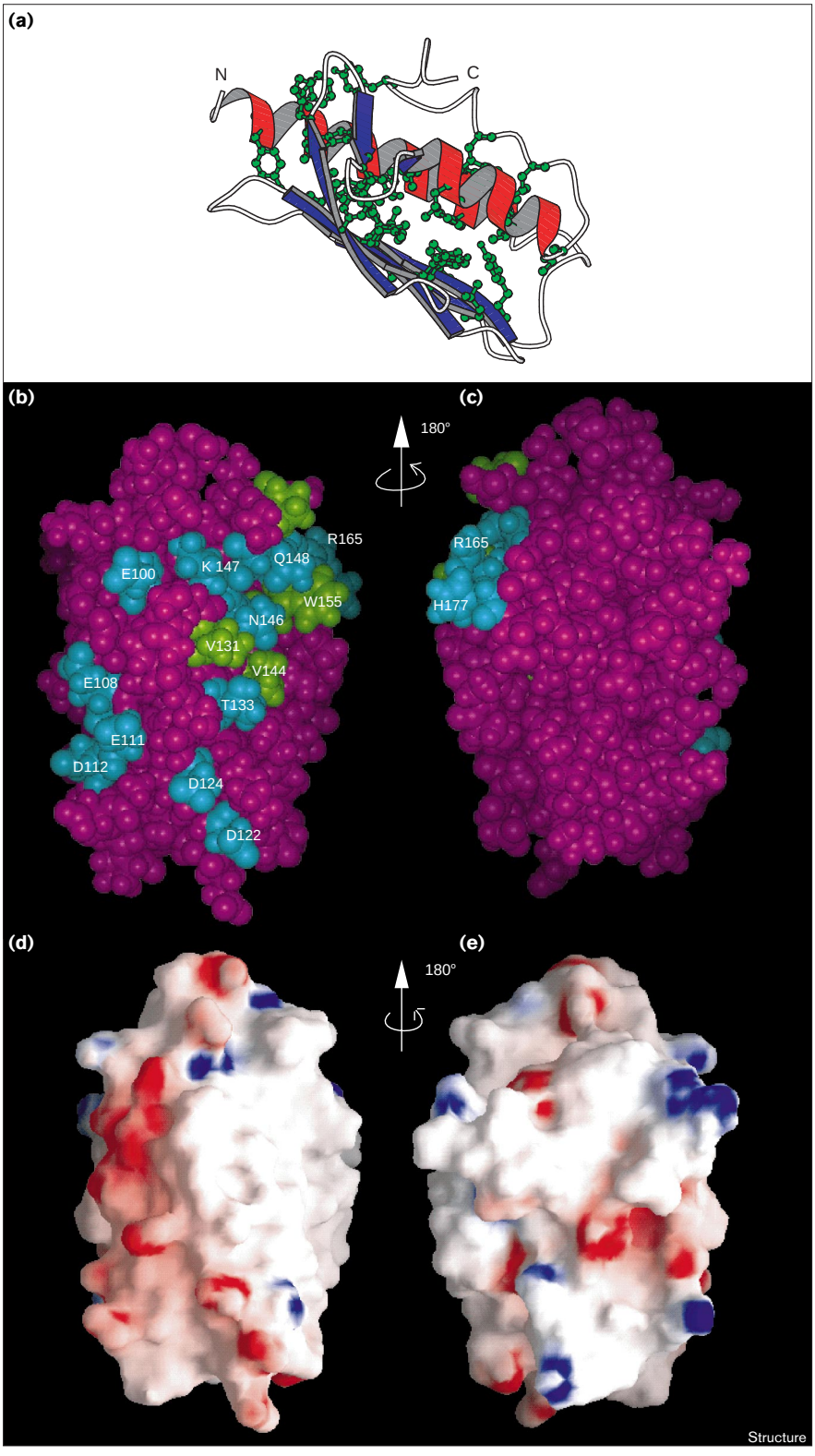
The frataxin C-terminal domain folds as a very compact globular protein with no cavities or grooves. The helices, together with the C-terminal tail, shield the hydrophobic face of the β sheet. This must be crucial for the stability of the fold as it allows formation of extensive interactions between the hydrophobic residues that are conserved in the frataxin family (Figure 5a). The hydrophobic core is dominated by the presence of aromatic sidechains (11 residues). One of the key residues is the conserved Trp173 (accessible surface area of 0.87 $Å^2$) that is surrounded by Tyr95, Trp168, Tyr205 (all well conserved). Together, they build a very compact aromatic cluster around Leu98, of which the sidechain points directly into the aromatic ring of Trp168 as reflected from its δ-protons, which are ring-current shifted to the most high field resonances of the spectrum (–0.18 and –0.6 ppm). A second aromatic cluster involves residues located at the end of the first helix (Phe109, Phe110), which have long-range contacts with residues in the first strand (Tyr123) and the first loop (Tyr118).

The conserved exposed residues all cluster on one side of the protein, encompassing the solvent-exposed face of the first helix and the β sheet (Figures 5b,c). Most of them are, as expected, hydrophilic residues (Glu100, Glu108, Glu111, Asp112, Asp122, Asp124, Thr133, Asn146, Lys147, Gln148, Phe150, Gln153, Ser157, Arg165 and His177) with the exception of Val131, Val144 and Trp155, Val174. The negatively charged residues are aligned along an edge formed by α1 and β2, producing a negatively charged ridge (Figure 5d). Given that the protein is monomeric in solution (see Materials and methods section), conservation of exposed residues must be explained by formation of a functional surface and therefore provide clues to the nature of the ligand.

The frataxin structure has distinct features from iron-binding proteins

In iron transporter proteins, which contain high affinity iron-binding sites, iron is either chelated by a haem or forms an iron–sulphur cluster or, as in transferrin, binds to a cluster of glutamate, histidine and tyrosine residues that sit in a pocket formed between two globular subunits. A strikingly similar di-iron binding centre is found in ribonucleotide reductase, methane monoxygenase hydroxylate, haemerythrin and ferritin (for a review see [30]). The di-iron centres of these proteins are all hidden in a hydrophilic cavity buried in a four-helix bundle. The frataxin structure does not have any pockets or cavities able to accommodate a haem in the C-terminal domain

Figure 5

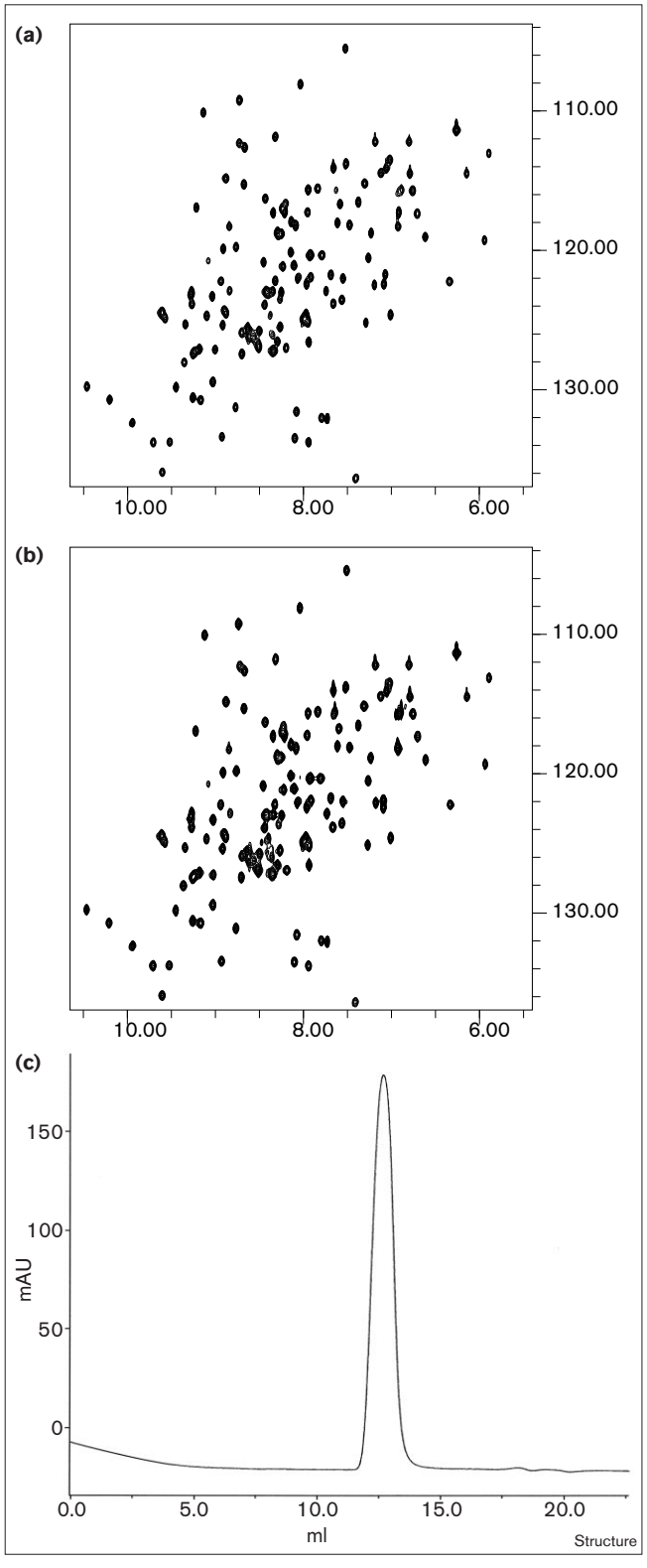


Mapping residue conservation to the structure of S_Fra. **(a)** Ribbon representation of S_Fra showing the conserved and hydrophobic sidechains (in green) implicated in packing of secondary structure elements. **(b)** Surface distribution of phylogenetically conserved and exposed amino acids. CPK representation of the asymmetric distribution of the exposed conserved residues (in cyan hydrophilic sidechains, in green hydrophobic sidechains). **(c)** View of the same representation after 180° rotation around the y axis. **(d)** Charge distribution on the protein surface as viewed from the same direction as in Figure 5b. Positively and negatively charged electrostatic potentials are indicated in blue and red, respectively. The surface was calculated by using a probe radius of 1.4 Å, and the potential displayed at the scale of −14 kbT to +14 kbT, where kb is the Boltzmann constant. A significantly high number of negatively charged residues is present on the first helix. **(e)** Charge distribution on the opposite side of the protein (same orientation as in Figure 5c).

(Figure 4). The additional 15 amino acids present in the mature form of frataxin are not conserved among species and, even assuming species diversity, do not contain residues that could help to fix the haem to the structure.

Figure 6

Iron binding assays. ^{15}N-edited HSQC spectra recorded at 500 MHz and 27°C of L_Fra using a 0.4 mM solution sample at pH 7.2. **(a)** Spectrum of the protein alone and **(b)** spectrum of the protein with an excess of Fe^{2+} ([frataxin]:[Fe^{2+}]1:3). **(c)** Elution profile of L_Fra after the addition of ferrous ammonium sulphate in a protein:iron ratio of 1:40 loaded on a high load 10/30 Superdex column. The scale on the vertical axis is in milli-absorption units (mAU). The same profile was invariably observed for all the different protein:iron ratios tested.



Cysteines are rare in the frataxin family and the few that are present are not conserved.

Although ferritin-like properties have been proposed for frataxin [26], comparison of the structures shows some weak similarities and a number of important differences. The negatively charged ridge along α1 could be imagined to provide an ion-binding site but, in contrast to that observed for most ferritins, it would be highly exposed and would require more than one protein molecule for its formation. An exposed ferroxidase centre has been described recently, however, for *Listeria innocua* ferritin, the only example of a ferritin in which iron is chelated by two different subunits [31]. Apoferritins spontaneously assemble into large aggregates (of 24 or 12 subunits) in the absence of iron [30] whereas frataxin has no tendency to self-aggregate, even at the millimolar concentrations that are used for NMR. Overall, the surface of frataxin is much less charged than that of ferritin. The hollow core of ferritin 24-mers is lined with negative charges that interact with the enclosed inorganic iron precipitate. The frataxin surface conservation is highly asymmetric; one side only of the molecule is conserved, with the negative ridge forming a border between conserved and unconserved residues (Figure 5). This configuration does not seem suitable for the formation of an analogous ferritin-like aggregate with negative charges pointing into the centre. Finally, the presence in frataxin of an exposed surface including absolutely conserved hydrophobic residues such as Trp155 strongly suggests a partner other than iron.

Is frataxin an iron-binding protein?

Different approaches were used to study the potential binding properties of frataxin to iron. In order to check the presence of site-specific conformational changes at high-affinity binding sites such as those observed in monomeric/dimeric iron-transporter proteins (e.g., ferrodoxin, transferrin) NMR was used. L_Fra that corresponds to the mature form of frataxin was titrated with Fe^{2+} and Fe^{3+} up to a protein/iron ratio of 1:3. On addition of iron ions of either oxidative state, no chemical shift changes were observed in the HSQC spectra (Figure 6a). Given that neither the linewidths nor the intensities of the resonances showed any variation, formation of large aggregates and/or precipitation and removal of signals by paramagnetic ion line broadening can be excluded. These *in vitro* data show conclusively that monomeric frataxin does not have a high-affinity iron-binding site.

To further check whether or not an excess of iron could promote aggregation similar to what is observed for ferritin,

L_Fra was titrated with an iron excess (up to a ratio of 1:1000) and the solution applied either on a native gel or on an analytical gel-filtration column. In both cases, however, no aggregation was observed (Figure 6b).

All the clinically important mutations occur at conserved sites of the frataxin structure

Given that the GAA expansion alters protein levels, but the FRDA point mutations directly affect function of the expressed protein, examination of the point mutations in the FRDA gene might be useful to gain new insights into the role of frataxin in the ataxia. The currently known missense mutations are restricted to the C-terminal domain of frataxin. The mutations can thus be mapped on the structure and their effects rationalised according to their location (summarised in Table 2). Seven mutations occur in positions occupied by hydrophobic residues conserved in the frataxin family and are either involved in the protein core or important for the conformation of loop regions (Figure 7a). These mutations might be expected to have deleterious consequences on the fold stability. Three mutations (Asp122→Tyr, Trp155→Arg and Arg165→Cys) occur at surface sites and all affect strongly conserved residues (Figures 7a,b).

Table 2

Summary of the mutations described in FRDA patients.

Mutation	Effect	References
Lys106→Ser, Trp173→Gly	Internal cavities with loss of packing energy	[47] [9]
Ile154→Phe, Leu182→Phe	Steric strain by replacing a buried hydrophobic residue by a larger one	[8] [48]
Leu156→Phe	Disruption in the β-sheet hydrogen bonds by introducing a proline	[9]
Gly130→Val	Steric strain by replacing a glycine which is in a conformation that is not allowed for other residues	[49]
Leu182→His	Electrostatic strain by replacing a buried residue with a hydrophilic one	[9]
His183→Arg	Electrostatic strain by replacing a partially uncharged buried residue with a bulky charged one	[9]
Asp122→Tyr	Change of identity of a conserved negatively charged group on the surface	[9]
Trp155→Arg	Replacement of a bulky highly conserved aromatic sidechain with a positively charged residue	[33]
Arg165→Cys	Replacement of a conserved positively charged group with an exposed cysteine that might form intermolecular disulphide bridges	[50]

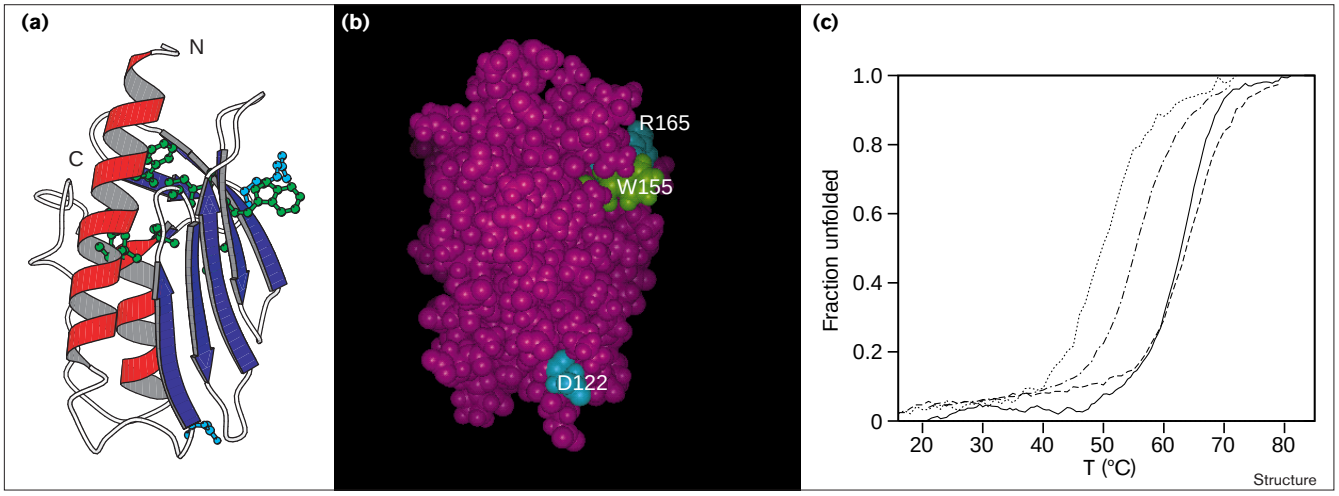
Two representative FRDA point mutations (Ile154→Phe, Trp155→Arg) were chosen for expression in *Escherichia coli*. The first is one of the most common mutations, described in several independent families [9], and leads to a severe phenotype. The second was chosen because the mutation involves an exposed, but absolutely conserved tryptophan [32]. The mutations were introduced into the L_Fra construct. As compared with the wild-type, the expression yields of the two mutants were both low with part of the protein forming inclusion bodies, suggestive of a difficulty in adopting a native conformation. In both cases, however, protein could be purified from the soluble fraction. The fold and the stability of the purified soluble fraction of both mutants were compared with those of the wild-type using CD spectroscopy. At room temperature, the far-UV CD spectra of the three proteins have similar features, indicating that the fold is retained in the mutants (data not shown). The wild-type melting temperature, as followed by CD thermal denaturation curves, was found to be 62°C. The two mutants are clearly less stable than the wild-type and the melting points are 12°C lower for Ile154→Phe and 7°C lower for Trp155→Arg (Figure 7c). Of the two mutants, only the refolding of the Trp155→Arg mutant is reversible under the conditions used *in vitro*.

Discussion

Frataxin seems to be an important protein present in all species and is essential for survival [8,11]. We have shown here that the mature protein forms two structurally distinct regions of which only the evolutionary conserved C terminus is folded and forms a compact globular domain. Even if folded, the remaining 15 residues present in the mature form of frataxin do not influence the stability of the C terminus, therefore suggesting that they do not interact with it. Although the biochemical function of frataxin remains unclear, the structure provides the first guidelines to the nature of the ligand and allows us to rationalise a number of observations that could not be explained by the primary structure of the protein alone.

From the tertiary fold, we can distinguish residues that are buried inside the fold and are important for its stability from those that are exposed and yet conserved through evolution. Residues on the surface are probably conserved for functional reasons and therefore can be suggestive of interaction sites with potential frataxin ligands. Residues that are exposed but conserved in the whole frataxin family cluster only on one face of the structure, the one containing the β sheet and the edge of α1. The conservation covers a span of approximately 30 Å and an area of more than 1400 $Å^2$, a region larger than would be expected for a small molecule as a single ligand. This suggests either that two or more ligands are brought together, or that a single partner, for example another protein or a nucleic acid molecule, is involved. In either case, the

Figure 7



Characterization of clinically occurring mutations. **(a)** Mapping on the S_Fra structure of the known point mutations occurring in Friedreich's ataxia patients: the sidechains of the amino acids corresponding to the point mutations are shown in green (hydrophobic residues) and cyan (hydrophilic residues). The view is the same as in Figures 5b,d. **(b)** Space-filling representation of the surface of frataxin in the same orientation as in 7a. **(c)** Thermal denaturation curves monitored by far-UV circular dichromism. The melting curves were followed by recording the signal of the spectrum of L_Fra at 223 nm (dashed line), the Trp155→Arg mutant (dashed-dotted line) and the Ile154→Phe mutant (dotted line).

ligand(s) will have to match a large composite surface and complement the pattern of conservation formed by the negative ridge composed of Glu108, Glu111 and Asp124, by the central hydrophobic patch and the conserved positive groups of Arg165 and His177 (Figure 5b).

Mapping the pathological mutations onto the structure is the basis for rational design of experiments to test the role of individual residues of frataxin in ligand recognition and for understanding the effect of mutations that cause the disease. Interestingly, all missense mutations observed so far map in the evolutionary conserved frataxin C terminus encoded by exons 4 and 5a (Figures 1 and 7). This observation strongly supports the hypothesis that the C terminus is the functionally most relevant region of the protein. The two mutants that were tested are both folded at room temperature but show distinct features from the wild-type. The Ile154→Phe mutation has greatly reduced thermodynamic stability and poor refolding ability. These effects could lead to much reduced levels of functional frataxin in the cell, either because of refolding difficulty after import into the mitochondria, or because of a more rapid turnover of poorly stable mature frataxin. A lower refolding efficiency would also agree with the slower maturation processing observed by H Koutnikova and colleagues [25] (but not validated by Gordon *et al.* [27]). The Trp155→Arg mutation is even more interesting. Because of the conservation of the tryptophan on the protein surface, the effect of the Trp155→Arg mutation should not only interfere with the refolding pathway (although to a lesser extent than the Ile154→Phe mutant) but also disrupt the ligand-binding site. Such expectation makes this mutation an excellent candidate for the design of *in vivo* knockouts in model organisms and to test ligands *in vitro*.

Molecular research into FRDA has burgeoned since the gene responsible was cloned [8]. Previously, almost nothing was known about the underlying biochemical defect. It is now well-established that the disease results from oxidative damage associated with iron accumulation in mitochondria and experimental antioxidant therapies are being evaluated [23]. The prevailing hypothesis supposes that frataxin is involved in the regulation of mitochondrial iron import/export, although it has been unclear how direct this might be. There has been some expectation that frataxin would itself prove to be an iron-binding protein. The presence of conserved negatively charged residues in the sequence, which cluster together on the same surface of the structure, suggests a binding site for positive ions. However, experimentally we do not observe any high-affinity iron-binding site. Furthermore, whereas all of the known high-affinity iron-binding sites are formed by conserved pockets formed in the interior of the protein monomer, no pocket is observed in the frataxin structure.

On the other hand the recent report that yeast frataxin has ferritin-like iron storage properties and is able to form large aggregates that sequester iron from solution is a very interesting possibility [24]. Ferritin has features very distinct from those of iron-transporters as it forms large

aggregates of 24 subunits. The polymer forms a large hollow cavity in which up to 4500 Fe^{3+} molecules can be stored as an inorganic complex. In frataxin, the conserved exposed hydrophobic residues could be part of a self-aggregation surface whereas the negatively charged ridge could be the equivalent of the ferroxidase ferritin centre reminiscent of that observed for *Listeria innocua* ferritin [31]. The fine details (the monodisperse behaviour of mature frataxin, the lack of a buried hydrophilic site, a much less charged exposed surface, the nature of the conserved hydrophobic residues) suggest that, if frataxin had iron storage properties, its assembly behaviour and iron storage would have a very different mechanism from that of ferritin. Experimentally, we were not able to reproduce the experiments described for yeast frataxin [24], which suggests that the two proteins could have a different *in vitro* behaviour. On the other hand it is true that, because a ferritin-like behaviour has been observed *in vitro* without having relevance *in vivo* [33,34], more work at the cellular level might be necessary to prove iron storage properties for frataxin.

An indirect involvement of frataxin in iron homeostasis cannot, therefore, be ruled out. Iron accumulation as a secondary effect of another misfunction is observed in other diseases: the X-linked childhood sideroblastic anaemia and ataxia (XLSA/A: OMIM 301310) is caused by a partial loss-of-function mutation in the ABC7 gene, which mediates export of iron–sulphur (Fe–S) cluster precursors [35]. An indirect involvement of frataxin in iron metabolism would also agree with the recent report of the knockout of the *E. coli* cyaY, a gene conserved in all the purple bacterial species so far examined [11,36]. Although cyaY is the orthologue of eukaryotic frataxin and shares with it a common conserved exposed pattern (Figures 1 and 5b), the *E. coli* knockout shows neither an effect on the cellular iron content nor sensitivity to oxidants [36]. Although it is tempting to argue that as we could find no clear differences in the pattern of conserved residues, the biochemical functions of cyaY and frataxin have diverged, the sequences provide no grounds for making this assertion; that is despite the dramatic differences in phenotype, the precise biochemical functions of cyaY and frataxin could ultimately prove to be identical.

There are a number of ways in which frataxin might indirectly fit into the iron regulation system. Frataxin could function as a carrier of an antioxidant (such as vitamin E or glutathione). If its partner is a protein (which cannot be excluded because failure of the two-hybrid screen experiments attempted so far could be simply because of various technical reasons), it is tempting to suggest an interaction with one of the Fe–S proteins of a respiratory complex. Fe–S cluster proteins in respiratory complexes I, II and III, together with aconitase have indeed been found to be deficient in FRDA [20]. Atm1 (the yeast ABC7 orthologue) has also been proposed to be regulated by the frataxin homologues [37]. Frataxin could be involved in protein import and processing, perhaps as a chaperone or adaptor.

Biological implications

Friedreich's ataxia (FRDA) is a recessive neurodegenerative disorder that involves progressive physical disability, dysarthria, muscular weakness, and can lead to premature death due to hypertrophic cardiomyopathy [2,3]. The locus for FRDA was recently mapped to a single gene located in chromosome 9 and encodes a small mitochondrial protein of 210 amino acids named frataxin [8–10]. The protein is essential for life and is present in living organisms ranging from purple bacteria to humans [11]. Reduction of the frataxin concentration in mitochondria leads to disregulation of iron homeostasis, with concomitant oxidative damage leading to neuronal death [12,14–20]. An important point that has been questioned is, therefore, whether or not frataxin is an iron-binding protein. Structural characterization of recombinant frataxin shows that the mature protein is composed by an N-terminal tail and an independently folded C-terminal domain with an α–β globular fold. The structure of the protein does not share any of the features typical for iron-binding proteins, strongly suggesting that the involvement of frataxin in iron metabolism is indirect. Iron-binding could also not be supported experimentally. Mapping of sequence conservation onto its surface suggests a large surface of interaction, which selectively clusters on one face. Although the specific partner (or partners) of frataxin remain for the moment only speculative, the speed of progress made in the past few years not only in experimental treatment but also in a deeper comprehension of the metabolic bases of the disease holds a realistic hope that FRDA might soon be a treatable condition. In this perspective, the establishment of the frataxin structure and of an interaction surface discussed in the present work, together with rationalization of the effect of clinically occurring mutations, will provide useful guidelines for future biochemical and clinical experimentation.

Materials and methods

Sequence analysis

New frataxin homologues were retrieved by using an alignment of the C-terminal domain for profile searches on the Bioccellerator (http://shag.embl-heidelberg.de:8000/). Accession numbers of sequence entries in SWISS-PROT or SPTREMBL are: Q16595, Human FRDA (Hs in Figure 1); O35943, *Mus musculus* FRDA (Mm); AAC71159, *Caenorhabditis elegans* FRDA (Ce); O74831, *Schizosaccharomyces pombe* FRDA (Sp); Q07540, *Saccharomyces cerevisiae* FRDA (Sc); Q9ZR07, *Arabidopsis thaliana* FRDA (At); P27838, *E. coli* cyaY (Ec); P71358, *Haemophilus influenzae* cyaY (Hi); P46356, *Yersinia pestis* cyaY (Yp); P40128, *Erwinia chrysanthemi* cyaY (Ech); Q9ZDK5, *Rickettsia prowazekii* cyaY (Rp).

Protein expression and purification

The two C-terminal L_Fra and S_Fra fragments of human frataxin (residues 63–210 and 91–210 respectively; accession number

U43752) were amplified using the polymerase chain reaction (PCR). The S_Fra construct was expressed with a His-tag of ten additional residues at the N terminus (MKHHHHHHPM) included for purification purposes and not removed. (Met10 substitutes for the semiconserved residue Leu90). L_Fra was obtained by cloning it into a pET9-derived plasmid vector behind the encoded His-tag and glutathione-*S*-transferase (GST). The Ile154→Arg and Trp155→Arg mutants were derived, also as GST-fusion proteins, from L_Fra using a site-directed mutagenesis kit (Quiagen). The constructs were expressed in *E. coli* strain BL21 (DE3) transformed with a modified pET vector DNA. For protein expression, the cells were inoculated into 4 l luria broth medium with ampicillin (100 mg/l, for the short construct) or kanamycin (30 mg/l, for the long construct and the mutants), induced for 3–4 h by addition of 0.2 mM isopropyl β-D-thiogalacto-pyranoside (IPTG) after the cultures reached an optical density (OD) between 0.6–0.8 at 600 nm. The cell pellets were harvested and frozen. The frozen cells were thawed in a lysis buffer and subsequently sonicated and centrifuged. Ni-NTA chromatography was applied as a first step of purification both to the His-tagged S_Fra and to the GST-fused constructs (L_Fra and mutants). The soluble over-expressed proteins were passed through a nickel column and eluted with imidazole (pH 8.0).

Fra was further purified over a MonoQ column (Pharmacia). The peptide (with an isoelectric point of 5.82) was eluted at about 0.3 M NaCl and concentrated by ultrafiltration. Final purification was achieved by gel filtration chromatography on a Superose-12 column (Pharmacia) equilibrated in 40 mM potassium phosphate pH 7. The frataxin peptide eluted in the retention volume predicted for a monomeric 14.7 kDa protein. L_Fra and its mutants were dialyzed against 20 mM Tris-HCl buffer with 100 mM NaCl pH 8, concentrated with Amicon concentrator (model 8050) and loaded on a Superdex 16/60 high load gel-filtration column equilibrated with 20 mM Tris-HCl, 500 mM NaCl at pH 8. The fractions containing the pure proteins were pooled together and concentrated. The purity of the recombinant proteins was controlled by SDS–PAGE after each step of the purification and by mass spectroscopy on the final products.

^{15}N-labelled samples were produced by growing the bacteria in minimal medium using $(^{15}NH_4)_2SO_4$ as sole source of nitrogen, and doubly labelled samples were produced using $(^{15}NH_4)_2SO_4$ and $^{13}C_6$-D-glucose.

Analytical gel filtration was used to check whether L_Fra is mono-disperse in solution. Ovalbumin (43 kDa), chymotrypsinogen A (25 kDa) and ribonuclease A (13.7 kDa) were used as molecular standards for the mass calibration curve. The molecular weight of a 25 μM solution of L_Fra is consistent with that expected for a monomeric species.

Structure calculations

Spectral assignment was performed as described in [28]. In order to reduce ambiguities in the assignment of the aromatic sidechains, 3D Constant Time [^{1}H/^{13}C]-TROSY CCH-COSY [38], 2D CT-Cγ(CC)H (mixing time 4.2 ms and 13.6 ms), and 2D CT-Cγ(Cβ)Hβ [39] were performed on a BRUKER 800 MHz DRX spectrometer. Deuterium exchange studies were performed on a lyophilised sample of ^{15}N-labelled protein redissolved in deuterated phosphate buffer (pH 6.8). Six ^{15}N HSQC spectra were collected at 25°C on a BRUKER DMX over a period of 24 h. Amide protons were considered to exchange slowly if they were still visible in the spectrum recorded 24 h after adding D_2O. Interproton distance restraints were derived from three heteronuclear 3D-NOESY spectra at 27°C with mixing time of 100 ms (NOESY-^{15}N-HSQC, NOESY-^{13}C-HSQC and NOESY-^{13}C-HSQC optimised for aromatic residues).

Structure calculations were performed using the programs DYANA (version 1.5) [40] and ARIA interfaced with CNS-0.5 [41] (http://www.NMR.EMBL-Heidelberg.DE/nmr/nilges/). Preliminary DYANA runs were performed to clear the data from possible mistakes and inconsistencies. For these runs, coupling constants, hydrogen bonds and up to 951 manually assigned upper distance constraints, autocalibrated with the DYANA associated module CALIBA, were progressively incorporated. In the final DYANA run, 60 random conformers were annealed in 10,000 torsion angle dynamics steps using the standard annealing protocol. The average residual target function for the ten best conformers was 3.85 ± 0.67 Å^2, the mean global backbone rmsd and heavy atom rmsd (on the secondary structure elements) was 3.60 ± -0.67 and 4.44 ± 0.67 Å, respectively). The structures had a low convergence, as expected from the relative small number of constraints used in the calculation. However, this result was sufficient to get a preliminary picture of the fold of the protein.

To confirm the fold and automatically to increase the number of NOE-derived restraints, the refined structures were calculated by ARIA. For a typical ARIA run, consisting of eight iterations, a total of 20 structures per iteration were calculated. In each ARIA iteration, structures were calculated with a CNS simulated annealing protocol with 10,000 initial MD steps, 2000 MD steps in the high-temperature stage at 2000K, 5000 steps in the first cooling stage from 2000K to 1000K, and 2000 steps in the second cooling stage from 1000K to 50K. The time step was 0.005 ps. Floating assignment for prochiral groups was applied as described in [42]. In the first iteration, initial structures were calculated on the basis of the manually assigned XEASY peak lists, 42 hydrogen bonds and 50 ϕ dihedral angles. In the following iterations, the remaining NOEs were automatically calibrated and assigned by the ARIA subroutines. The seven best structures resulting from each run were selected and used to extract additional NOE restraints from the peak lists. These raw data were converted into distance restraints based on the assigned chemical shifts with a tolerance of 0.5 ppm on the heteroatom dimension, of 0.04 ppm on the direct dimension and of 0.03 ppm in the indirect dimension. The three peak lists were inspected with respect to the seven lowest energy structures from the previous iteration. A restraint was not used if it was violated by more than a target value in maximally 50% of the seven structures. The target value was gradually tightened using a violation tolerance (*violtol*) from 5 to 0.0 Å, the assignment parameter (*ambigcutoff*) was varied from 1.01 to 0.80. With the decreasing level of ambiguity in the ARIA/CNS iterative procedure, more restraints could be assigned unambiguously resulting in a rapid convergence of the structure. The total number of unambiguous and ambiguous NOE restraints after ARIA refinement was 2114 and 234, respectively. In the last iteration, 50 structures were generated by ARIA and the 15 having the lowest energy were used for statistical analysis. None of these structures exhibits distance violations greater than 0.5 Å or dihedral angle violations greater than 5°.

Structure analysis

The Ramachandran plot was analyzed with PROCHECK-NMR (version 3.4) [43]. Additional quality controls were performed with the programs PROSA [44] and WHATIF [45]. Secondary structure elements and Rmsd values were calculated using MOLMOL (version 2.6). Structure and folding similarity searches in the PDB were performed using DALI (http://www2.ebi.ac.uk/dali/).

Iron binding assays

^{15}N-labelled L_Fra was used for iron titration. To avoid any other metal contamination, an excess of EDTA (frataxin:EDTA 1:4) was added to the sample and then removed by gel filtration on a PD-10 column (Pharmacia). In order to solubilize the iron ions, $Fe(NH_4)_2(SO_4)_2{\cdot}6H_2O$ and $FeCl_3$ were dissolved in 20 mM citrate, which was used as iron carrier. For each of the titration experiments, 500 μl of 0.4 mM frataxin, 10 mM Tris (nondeuterated), 10 mM NaCl, pH 7.2 (90% H_2O and 10% D_2O) were prepared in a 5 mm NMR tube. In order to keep dilution and NMR signal loss at minimum, titrations with both Fe^{2+} and Fe^{3+} were carried out by adding microliter amounts of concentrated ion stock solutions of 10 mM (pH 8.2) to the NMR samples, up to a stoichiometric ratio of frataxin:iron of 1:3. Concentrations of the protein were derived by UV using an extinction coefficient of 26030 $M^{-1}cm^{-1}$ at 280 nm. For each titration point, a two-dimensional ^{15}N-edited HSQC spectrum was recorded at 27°C.

A screening to check conditions that could promote aggregation of frataxin in the presence of iron was achieved by native gel and gel filtration. Increasing quantities of ferrous ammonium sulphate (SIGMA) were added to a solution of protein (protein concentrations used were either 8 or 25 μM) to reach protein:iron ratios between 1:20 and 1:1000. Three sets of titrations were performed, one at pH 7 in 100 mM HEPES and incubated overnight at 30°C. Two more titrations were performed at lower pH (pH 6 in 100 mM MES) and incubated for 60 min at room temperature and 4°C, respectively. These conditions were chosen to slow down Fe(II)/Fe(III) oxidation and allow interaction with the protein if this interacted with Fe(II) only. The iron precipitate was centrifuged and the supernatant loaded on a 6% polyacylamide native gel. In all cases, the protein was visible in the gel and ran as a monomer. The same procedure was repeated and the supernatants analyzed by gel filtration using a high load column 10/30 Superdex equilibrated with the corresponding buffer, 300 mM NaCl at a flow rate of 0.5 ml/min.

CD measurements

CD measurements were performed on a Jasco J-715 spectropolarimeter equipped with a thermostatically controlled cell holder stabilized by circulating water from a Neslab RTE-111 water bath. A rectangular quartz cuvette with 0.1 cm pathlength was used (Hellma). All the spectra were baseline corrected by subtracting buffer spectra and the observed ellipticity was then converted to mean residue weight ellipticity $[\theta]/(\text{degree*cm*dmol}^{-1})$. CD spectra were averaged over five scans obtained by collecting data at 0.5 nm intervals from 259–190 nm. A rectangular quartz cuvette of 1 cm pathlength was used. All the experiments were carried out in 20 mM Tris-HCl buffer at pH 8 using a protein concentration of 9 μM. The thermal denaturation of the wild-type and mutant frataxins were followed by monitoring the ellipticity at 223 nm from 16°C to 80°C.

Accession numbers

The coordinates have been deposited in the Protein Data Bank with the accession code 1DLX.

Acknowledgements

We would like to thank M Koenig and F Foury for interesting discussions, J Feeney and DJ Thomas for critical reading of the manuscript, and the NMR Centre of Mill Hill, J Linge and D Nietlisbach for technical support. We are particularly indebted to P Arosio for scientific advice and to Grazia Isaya for sharing her results before publication. The work was supported by a Human Frontier (HFSP) grant.

References

1. Reddy, P.S. & Housman, D.E. (1997). The complex pathology of trinucleotide repeats. *Curr. Opin. Cell Biol.* **3**, 364-372.
2. Koenig, M. & Mandel, J.L. (1997). Deciphering the cause of Friedreich ataxia. *Curr. Opin. Neurobiol.* **7**, 689-694.
3. Pandolfo, M., *et al.* Didonato (1990). Friedreich ataxia in Italian families: genetic homogeneity and linkage disequilibrium with the marker loci D9S5 and D9S15. *Am. J. Hum. Genet.* **47**, 228-235.
4. Pandolfo, M. (1999). Molecular pathogenesis of Friedreich ataxia. *Arch. Neurol.* **56**, 1201-1208.
5. Durr, A., *et al.,* & Koenig, M. (1996). Clinical and genetic abnormalities in patients with Friedreich's ataxia. *N. Engl. J. Med.* **335**, 1169-1175.
6. Chamberlain, S., *et al.,* & Williamson, R. (1988). Mapping of mutation causing Friedreich's ataxia to human chromosome 9. *Nature* **334**, 248-250.
7. Chamberlain, S., *et al.,* & Pandolfo, M. (1993). Genetic recombination events which position the Friedreich ataxia locus proximal to the D9S15/D9S5 linkage group on chromosome 9q. *Am. J. Hum. Genet.* **52**, 99-109.
8. Campuzano, V., *et al.,* & Pandolfo, M. (1996). Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* **271**, 1423-1427.
9. Cossee, M., *et al.,* & Pandolfo, M. (1999). Friedreich's ataxia: point mutations and clinical presentation of compound heterozygotes. *Ann. Neurol.* **45**, 200-206.
10. Koutnikova, H., Campuzano, V., Foury, .F, Dolle, P., Cazzalini, O. & Koenig, M. (1997). Studies of human, mouse and yeast homologues indicate a mitochondrial function for frataxin. *Nat. Genet.* **16**, 345-351.
11. Gibson, T.J., Koonin, E.V., Musco, G., Pastore, A. & Bork, P. (1996). Friedreich's ataxia protein: phylogenetic evidence for mitochondrial dysfunction. *Trends Neurosci.* **19**, 465-468.
12. Babcock, M., *et al.,* & Kaplan, J. (1997). Regulation of mitochondrial iron accumulation by Yfh1p, a putative homolog of frataxin. *Science* **276**, 1709-1712.
13. Campuzano, V., *et al.,* & Koenig, M. (1997). Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes. *Hum. Mol. Genet.* **6**, 1771-1780.
14. Foury, F. & Cazzalini, O. (1997). Deletion of the yeast homologue of the human gene associated with Friedreich's ataxia elicits iron accumulation in mitochondria. *FEBS Lett.* **411**, 373-377.
15. Wilson, R.B. & Roof, D.M. (1997) Respiratory deficiency due to loss of mitochondrial DNA in yeast lacking the frataxin homologue. *Nat. Genet.* **16**, 352-357.
16. Lamarche, J.B., Cote, M. & Lemieux, B. (1980). The cardiomyopathy of Friedreich's ataxia morphological observations in three cases. *Can. J. Neurol. Sci.* **7**, 389-396.
17. Lamarche, J.B., Shapcott, D., Cote, M. & Lemieux, B. (1993). Cardiac iron deposits in Friedreich's ataxia. In *Handbook of Cerebellar Diseases.* (Lechtenberg, R., ed.), pp. 453-548, Marcel Dekker, New York.
18. Wong, A., *et al.* & F. Cortopassi, G. (1999). The Friedreich's ataxia mutation confers cellular sensitivity to oxidant stress which is rescued by chelators of iron and calcium and inhibitors of apoptosis. *Hum. Mol. Genet.* **8**, 425-430.
19. Delatycki, M.B., *et al.,* & Forrest, S.M. (1999). Direct evidence that mitochondrial iron accumulation occurs in Friedreich ataxia. *Annu. Neurol.* **45**, 673-675.
20. Rotig, A., *et al.,* & Rustin, P. (1997). Aconitase and mitochondrial iron-sulphur protein deficiency in Friedreich ataxia. *Nat. Genet.* **17**, 215-217.
21. Radisky, D.C., Babcock, M.C. & Kaplan, J. (1999). The yeast frataxin homologue mediates mitochondrial iron efflux. Evidence for a mitochondrial iron cycle. *J. Biol. Chem.* **274**, 4497-4499.
22. Helveston, W., Cibula, J.E., Hurd, R., Uthman, B.M. & Wilder, B.J. (1996). Abnormalities of antioxidant metabolism in a case of Friedreich's disease. *Clin. Neuropharmacol.* **19**, 271-275.
23. Rustin, P., von Kleist-Retzow, J.C., Chantrel-Groussard, K., Sidi, D., Munnich, A. & Rotig, A. (1999). Effect of idebenone on cardiomyopathy in Friedreich's ataxia: a preliminary study. *Lancet* **354**, 477-479.
24. Adamec, J., Rusnak, F., Owen, W.G., Naylor, S., Benson, L.M. & Isaya, G. (1999) Frataxin is an iron-storage protein. ASHG meeting, abstract.
25. Koutnikova, H., Campuzano, V. & Koenig, M. (1998). Maturation of wild-type and mutated frataxin by the mitochondrial processing peptidase. *Hum. Mol. Genet.* **7**, 1485-1489.
26. Branda, S.S., Yang, Z., Chew, A. & Isaya G. (1999). Mitochondrial intermediate peptidase and the yeast frataxin homolog together maintain mitochondrial iron homeostasis in *Saccharomyces cerevisiae*. *Hum. Mol. Genet.* **8**, 1099-1110.
27. Gordon, D.M., Shi, Q., Dancis, A. & Pain, D. (1999). Maturation of frataxin within mammalian and yeast mitochondria: one-step processing by matrix processing peptidase. *Hum. Mol. Genet.* **8**, 2255-2262.
28. Musco, G., *et al.,* & Pastore A. (1999). Assignment of the ^{1}H, ^{15}N, and ^{13}C resonances of the C-terminal domain of frataxin, the protein responsible for Friedreich ataxia. *J. Biomol. NMR* **15**, 87-88.
29. Jiang, J., Zhang, Y., Krainer, A.R. & Xu, R.M. (1999). Crystal structure of human p32, a doughnut-shaped acidic mitochondrial matrix protein. *Proc. Natl Acad. Sci. USA* **96**, 3572-3577.
30. Harrison, P. & Arosio, P. (1996) The ferritins: molecular properties, iron storage function and cellular regulation. *Biochim. Biophys. Acta* **1275**, 161-203.
31. Ilari, A., Stefanini, S., Chiancone, E. & Tsernoglou, D. (2000) The dodecameric ferritin from *Listeria innocua* contains a novel intersubunit iron-binding site. *Nat. Struct. Biol.* **7**, 38-43.
32. Labuda, M., Poirier, J. & Pandolfo, M. (1999) A missense mutation (W155R) in an American patient with Friedreich ataxia. *Human Mutation* **13**, 506-507.
33. Yariv, J, Kalb, A.J., Helliwell, J.R., Papiz, M.Z., Bauminger, E.R. & Nowick, I. (1988). Chemical and Mossbauer spectroscopic evidence that iron-containing concanavalin A is a ferritin. *J. of Biol. Chem.* **263**, 13508-13510.
34. Briat, J.F., Massenet, O. & Laulhere, J.P. (1989) A jack bean protein similar to pea seed ferritin and unrelated to concanavalin A is the only iron storage protein in Jack bean seeds, although concanavalin A can form ferritin-like iron cores *in vitro*. *J. of Biol. Chem.* **264**, 11550-11553.

35. Allikmets, R., Raskind, W.H., Hutchinson, A., Schueck, N.D., Dean, M. & Koeller, D.M. (1999). Mutation of a putative mitochondrial iron transporter gene (ABC7) in X-linked sideroblastic anemia and ataxia (XLSA/A). *Hum. Mol. Genet.* **8**, 5743-5749.
36. Li, D.S., Ohshima, K., Jiralerspong, S., Bojanowski, M.W. & Pandolfo, M. (1999). Knock-out of the cyaY gene in *Escherichia coli* does not affect cellular iron content and sensitivity to oxidants. *FEBS Lett.* **456**, 13-16.
37. Kispal, G., Csere, P., Prohl, C. & Lill, R. (1999). The mitochondrial proteins Atm1p and Nfs1p are essential for biogenesis of cytosolic Fe/S proteins. *EMBO J.* **18**, 3981-3989.
38. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. (1998). Transverse relaxation optimized spectroscopy (TROSY) for NMR studies of aromatic spin systems in 13C-labelled proteins. *J. Am. Chem. Soc.* **120**, 6394-6400.
39. Prompers, J.J., Groenewegen, J.J., Hilbers, C.W. & Pepermans, H.A.M. (1998). Two-dimensional NMR experiments for the assignment of aromatic sidechains in 13C-labelled proteins. *J. Mag. Res.* **130**,63-67.
40. Guntert, P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283-298.
41. Nilges, M., Macias, M.J., O'Donoghue, S.I. & Oschkinat, H. (1997). Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from β-spectrin. *J. Mol. Biol.* **269**, 408-422.
42. Folmer, R.H., Hilbers, C.W., Konings, R.N. & Nilges, M. (1997). Floating stereospecific assignment revisited: application to an 18 kDa protein and comparison with J-coupling data. *J. Biomol. NMR* **9**, 245-258.
43. Laskowski, R.A., Rullmannn, J.A., 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. Biomol. NMR* **8**, 477-486.
44. Sippl, M.J. (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins* **17**, 355-362.
45. Vriend, G. (1990). WHAT IF: a molecular modelling and drug design program. *J. Mol. Graph.* **8**, 52-56.
46. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D. & D.G (1997). The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876-4882.
47. Bartolo, C., Mendell, J.R. & Prior, T.W. (1998). Identification of a missense mutation in a Friedreich's ataxia patient: implications for diagnosis and carrier studies. *Am. J. Med. Genet.* **79**, 396-399.
48. Filla, A., *et al.* & Cocozza, S. (1996). The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia. *Am. J. Hum. Genet.* **59**, 554-660.
49. Bidichandani, S.I., Ashizawa, T. & Patel, P.I. (1997). Atypical Friedreich ataxia caused by compound heterozygosity for a novel missense mutation and the GAA triplet-repeat expansion. *Am. J. Hum. Genet.* **60**, 1251-1256.
50. Forrest, S.M., *et al.,* & Nicholson, G. (1998). The correlation of clinical phenotype in Friedreich ataxia with the site of point mutations in FRDA gene. *Neurogenetics* **1**, 253-257.

Because *Structure with Folding & Design* operates a 'Continuous Publication System' for Research Papers, this paper has been published on the internet before being printed (accessed from http://biomednet.com/cbiology/str). For further information, see the explanation on the contents page.