DNA triplex structures in neurodegenerative disorder, Friedreich's ataxia

MOGANTY R RAJESWARI

Department of Biochemistry, All India Institute of Medical Sciences, New Delhi 110 029, India

(Fax, +91-11-26588663; Email, rajeswari3011@hotmail.com)

It is now established that a small fraction of genomic DNA does adopt the non-canonical B-DNA structure or 'unusual' DNA structure. The unusual DNA structures like DNA-hairpin, cruciform, Z-DNA, triplex and tetraplex are represented as hotspots of chromosomal breaks, homologous recombination and gross chromosomal rearrangements since they are prone to the structural alterations. Friedreich's ataxia (FRDA), the autosomal recessive degenerative disorder of nervous and muscles tissue, is caused by the massive expansion of (GAA) repeats that occur in the first intron of Frataxin gene X25 on chromosome 9q13-q21.1. The purine strand of the DNA in the expanded (GAA) repeat region folds back to form the (R·R*Y) type of triplex, which further inhibits the frataxin gene expression, and this clearly suggests that the shape of DNA is the determining factor in the cellular function. FRDA is the only disease known so far to be associated with DNA triplex. Structural characterization of GAA-containing DNA triplexes using some simple biophysical methods like UV melting, UV absorption, circular dichroic spectroscopy and electrophoretic mobility shift assay are discussed. Further, the clinical aspects and genetic analysis of FRDA patients who carry (GAA) repeat expansions are presented. The potential of some small molecules that do not favour the DNA triplex formation as therapeutics for FRDA are also briefly discussed.

[Rajeswari MR CM 2012 DNA triplex structures in neurodegenerative disorder, Friedreich's ataxia. *J. Biosci.* **37** 519–532] **DOI** 10.1007/s12038-012-9219-1

1. Introduction

In 1953, James Watson and Francis Crick, with the help of X-ray diffraction, demonstrated that DNA, the genetic material, exists as right-handed helical duplex structure. This structure of Watson and Crick model of DNA is conventionally referred as B-DNA. Watson and Crick published their work in *Nature*, where they humbly say that their DNA structure has novel features which are of 'considerable biological interest' (Watson and Crick 1953). Today, we know what an earth-shaking discovery it was! Although the B-DNA structure has been widely accepted by all biologists including biochemists and cell biologists, the classical DNA structure made everyone, particularly biophysicists, think of the alternate structures that DNA can perhaps adopt (Mirkin 2008). Indeed, this led to the discovery of other polymorphic forms: A-DNA, Z-DNA, C-DNA, etc. (see recent review by Choi and Majima 2011). It is not surprising that more than 10 types of non-B-DNA structures, including DNA triple helix, quadruplex, cruciform, motifs, -A and I-, have been identified (Zhao *et al.* 2010). The presence of such non-B-DNA structures were initially established in synthetic oligo-/polynucleotides. However, occurrences of unusual DNA structures are now proved in natural systems also. These are particularly seen in the human genome at the repeat DNA sequences, which account for more than 50% of the total genomic DNA while simple sequence repeats account for of approximately 3% of the total DNA. Poly (purine·pyrimidine)-rich regions in the human genome are prone to adopting non-canonical DNA structures such as intra-molecular triplexes or hinge-DNA or H-DNA (Vasquez and Glazer 2002). Such mirror repeats of triplex-structure-forming sequences are abundant and can regulate the expression of several disease-linked genes. Application of DNA

Keywords. DNA triple helix; Friedreich's ataxia; (GAA) repeats; triplet repeat expansion (TRE) neurodegenerative disorder

Published online: 25 June 2012

triplex using antigene strategy has been an active research area for the control of gene expression in several human cancers (Chan and Glazer 1997; Fox 2000).

2. DNA triple helix: General structural aspects

Soon after the double helical structure was declared in 1956, Felsenfeld's group reported triple helical structures in ribonucleic acids (Felsenfeld and Rich 1957) and in deoxyribonucleic acids (Felsenfeld *et al.* 1957). DNA triplex helix, or the triplex as the name indicates, contains three strands, two of the DNA double helix and a triplex forming third strand which usually contains polypurine/polypyrimidine (Kamenetskii-Frank and Mirkin 1995; Duca *et al.* 2008). The third strand winds around the major groove of DNA double helix to form triplex (see the textbook by Soyfer and Potaman 1996). Triplexes can be of two types – Inter-molecular and Intra-molecular – depending on the source of the third strand. Inter-molecular triplexes are formed when the triplex-forming strand is a separate single strand containing deoxyribonucleotide or one of the strands from a different DNA molecule (figure 1A). While the 'intra-molecular triplex' is formed when DNA with polypyrimidine/polypurine sequences having mirror symmetry undergo conformational rearrangement and fold back on to the duplex itself (figure 1B). The *in vivo* intra-molecular triplex also knows as Hinge-DNA (H-DNA). This type of polypyrimidine/polypurine mirror repeats are overrepresented in the human genome and are generally found near promoter regions and recombination hotspots. The binding of the third strand in the triplex is through 'Hoogsteen' or 'reverse-Hoogsteen type' of hydrogen bonding, which is different from the classical Watson–Crick base pairing of B-DNA. The arrangement in which the third strand with a stretch of pyrimidines remains parallel (p) to the central purine strand is termed as 'pyrimidine motif' (Y*R·Y). while in the 'purine motif' (R*R·Y), the third strand comprises of a polypurine stretch that runs antiparallel (ap) to the central polypurine strand. In either motif, the purine strand of the DNA duplex occupies the central position and provides sites for hydrogen bond formation with the complementary pyrimidine strand of the duplex and the third strand containing purine or pyrimidine. Therefore, triplexes are confined to targeting polypurine stretches. This further gives rise to the possible base triplet with purine motif being G*G· and A*A·T and pyrimidine motif being C*G·C and T*A·T. The purine motif can be formed at physiological conditions; however, for the formation of a pyrimidine motif, cytosine, if present, in the third strand needs to be either $5'^{mc}$C methylated or in protonated (C^+) form (pH <6).

The base triplets A*A·T can be found in (ap) triplexes with reverse-Hoogsteen (r-Hg), and C*G·C can only give

Figure 1. Scematic presentation of (**A**) 'inter-molecular' and (**B**) 'intra- molecular' DNA triple helix.

rise to (p) triplexes with Hoogsteen base pair (Hg), whereas G*G·C and T*A·T can occur both in (p) and (ap) triplexes (figure 2). The (ap) R*R·Y triplex containing G*G·C and A*A·T triads is not isomorphic compared to the (p) Y*R·Y type of triplex, which involves isomorphic C^{+}*G·C and T*A·T triads. The base triplets formed in the triplexes indicating Hoogsteen (in red colour), reverse-Hoogsteen (in green) and Watson–Crick base pairing are shown in figure 2. However, G*G·C and A*A·T triads are stabilized by two reversed-Hoogsteen-type hydrogen bonds between bases in the third (purine) strand of the Watson–Crick duplex and the glycosidic torsion angles restricted to the anti-domain (11). The adenine residue can be readily accommodated within the third strand of ap R*R·Y triplex and shows N^6-H—N reversed Hoogsteen bond with adenine in A·T base pairs, while the G*G·C triplets next to A*A·T remain unperturbed.

3. DNA triplet repeat expansions in human genome

In the beginning of the last decade, expanded DNA-trinucleotide repeats in genes were identified as unstable and responsible for a large number of neurological disorders like FRDA, fragile X syndrome, spinocerebellar ataxia and muscular dystrophy. This discovery brought a paradigm shift in genetics. Trinucleotide repeat expansion (TRE) is established as an important mutational mechanism associated with a series of human neurological disorders with the etiology of at least 12 diseases and a few fragile sites associated with TREs (Wilmot and Warren 1998). The inheritance pattern of several more diseases make them candidates for mutation through TRE. Clinically, the triplet repeat diseases are associated with a phenomenon called anticipation, implying progressively earlier onset and worsening severity of the disease

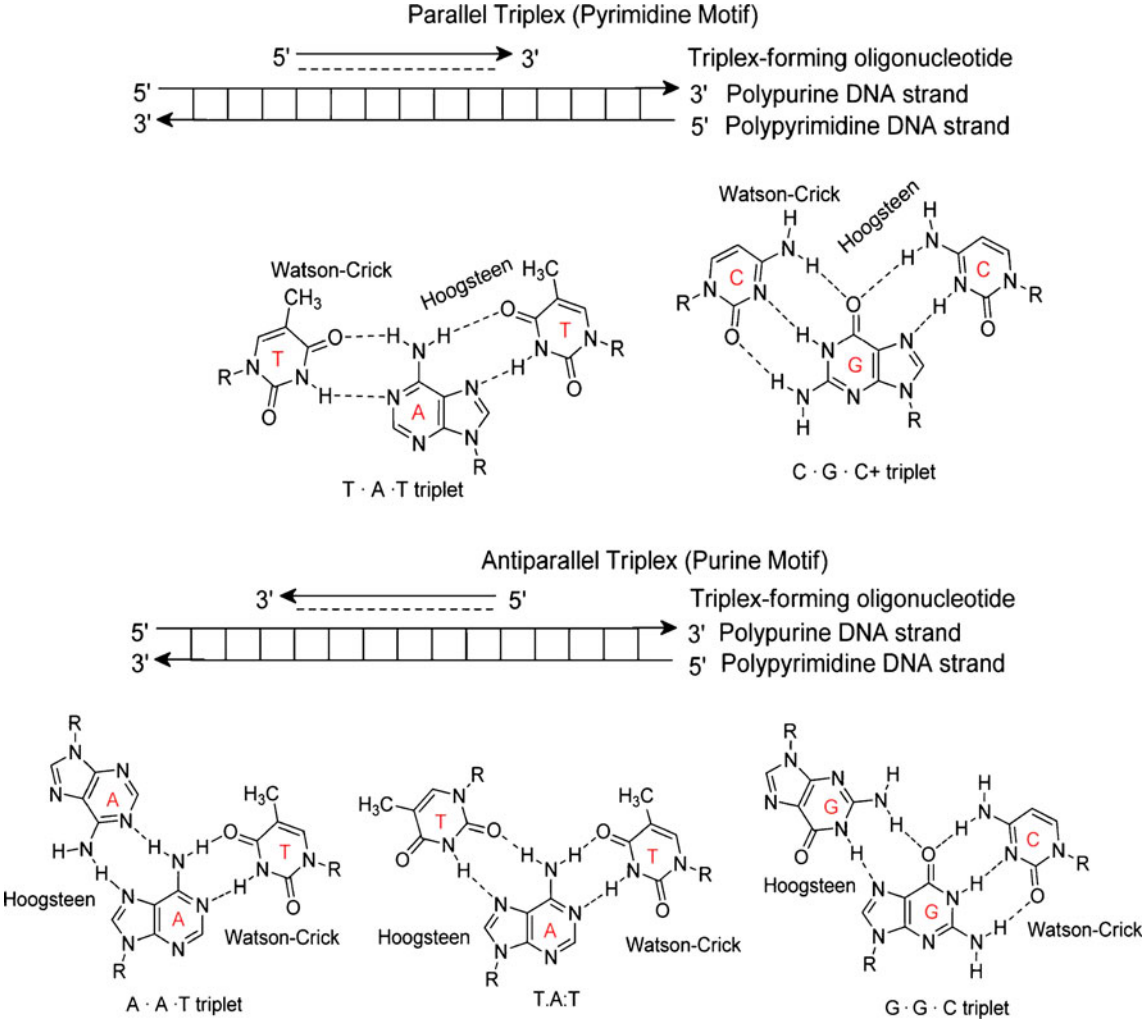

Figure 2. Hydrogen bonding scheme in DNA base triplets: (Top) parallel, pyrimidine motif and (Bottom) antiparallel, purine motif DNA triple helices. These are defined with respect to the orientation of the triplex forming oligonucleotide (TFO) and homopurine Watson–Crick (W-C) strand (Reprinted, with permission, from *Acc. Chem. Res.* **44** 134–146, 2011. Copyright (2011) American Chemical Society).

with successive generations. Of the 10 to 12 TREs known till date, most are rich in G and C content (Warren 1996) and belong to the type (CXG)n, where X=is A, G, C or T . Further, triplet repeats are found in every region of the gene like introns, exons and the 3′ or 5′ untranslated regions (UTR) (Wilmot and Warren 1998). Genes of healthy humans also do contain DNA triplet repeats of short length and their number is distinctly small. However, when the number of repeats reaches a certain value, the symptoms of the disease start manifesting. For full-blown diseases the repeat number is quite high. A glance at table 1 shows that the repeat number *n* in the disease state differs from the repeat type, and the gene on which it appears, however, has no relation between any two different disease conditions. In some cases, permutation alleles with $n \sim 60$ can convert Friedreich's ataxia (FRDA) alleles in just one generation (Cossée *et al*. 1997). The offspring from such parents with intermediate number of DNA triplet repeats are prone to manifest the disease. The repeat number of these triplets found in normal individuals and the patients with intermediate and fully developed diseases are summarized in table 1.

The 24 autosomal dominant ataxias – SCA 1–8, 10–19, 21–23 and 25, dentatorubral-pallidoluysian atrophy (DRPLA) and ataxia caused by mutations in the gene that encodes fibroblast growth factor 14 (FGF14) – have been identified. In 12 of these disorders the genes involved and the underlying mutations are known. Six SCA subtypes (SCA1, SCA2, SCA3, SCA6, SCA7 and SCA17) and DRPLA are caused by (CAG) trinucleotide repeat expansions in the respective genes (table 1).

Table 1. Some of the ataxias associated with DNA triplet repeat expansions and their repeat number in normal and various neurological disorders

		Repeat range		
Disease	Repeat	Normal	Intermediate	Affected
Fragile XA	CGG	6–50	50–200	> 200
Fragile XE	GCC	7–35	130–150	230–750
FRDA	**GAA**	**10 –21**	**40–60**	**> 100**
Dentato rrubral-pallido luysian atrophy (DRPLA)	CAG	6 –39	–	49–75
Spinocrebellar ataxia SCA I	CAG	6–39	–	40–81
SCA II	CAG	15–24	–	35–59
SCA III	CAG	12–40	–	50–84
SCA VI	CAG	4–16	–	21–27
SCA VII	CAG	4–35	28–35	37 –200
SBMA	CAG	15–31	–	40–62
Huntington disease HD	CAG	10–35	36–39?	36 +
Mitonic distrophy (DM)	CTG	5–37	50–100	> 100

These expansions encode polyglutamine repeats, as in Huntington's disease (HD); and these diseases are also known as polyglutamine expansion disorders. Besides CAG repeats in the coding regions, a CAG repeat expansion of more than 66 repeats has been found in the 5′-UTR region of the *PPP2R2B* gene in SCA12.

3.1 *Normal and expansion of (GAA) repeats in FXN gene*

The GAA/TTC is a unique family of triplet repeats that is not GC rich and yet undergoes dynamic expansion in the X25 gene, ultimately leading to the disease FRDA (Gacy *et al.* 1998). The FXN gene responsible for FRDA was mapped in 1988 by Chamberlain *et al.* on chromosome 9 (Chamberlain *et al.* 1988). This was followed by cloning experiments in 1996 by Campuzano *et al.* Moreover, a very rare locus FRDA2 has been found in Spanish family by Smeyers *et al*. (1996) and Christodoulou *et al*. (2001) separately. Christodoulou *et al*. (2001) specified this new locus at 9p23–9p11. The FRDA gene FXN spans 80 kb in human genome and consists of 7 exons 1, 2, 3, 4, 5a, 5b and 6, as shown in figure 3.

The most common transcript of FXN gene arises from exons 1–5a, which is translated into a 210-amino-acid mitochondrial protein called 'frataxin'. By alternative splicing, sometimes exon 5b can be transcribed instead of 5a and translated to a 177-amino-acid protein of unknown function, while exon 6 is non-coding. Normally, healthy individuals have short (GAA) repeats (8–33) in the FXN gene but FRDA patients have expanded (66–1700) repeats. Normal chromosomes have 5–30 (GAA) repeats in which <12 repeats are called short normal (SN) alleles and ≥12 repeats are called long normal (LN) alleles. In FRDA patients alleles having >100 (GAA) repeats are called disease-causing alleles (E) and may vary up to 1700 repeats (Filla *et al*. 1996; Delatycki *et al.* 1999). Repeat number from 20 to 100 is termed as premutation (intermediate) and does not cause disease phenotype. These pre-mutated alleles can be expanded into a fully expanded allele in the next generation by a process called anticipation. Heterozygotes possess one normal (SN/LN) allele and the expanded allele and may not have the disease symptoms. If mutation occurs in normal allele, then heterozygotes can also manifest the FRDA. These carriers have more tendencies to develop disease than pre-mutated or normal alleles.

3.2 *Proposed mechanisms for frataxin gene silencing in FRDA*

There are three different modes of mechanisms that have been proposed in order to explain the gene silencing in FRDA:

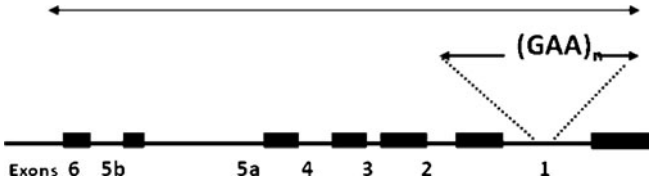

Figure 3. Location of the FRDA gene (fxn) on chromosome 9q13-q21.1; The (GAA) repeats is highlighted between exons 1 and 2.

(i) The first being formation of DNA triplex. In this mode the DNA triplex is very stable and stalls RNA polymerase. This is discussed below in detail. However, it is important to mention the other two mechanisms.

(ii) Formation of RNA·DNA hybrid. As transcription proceeds, the (GAA) non-template strand folds back and intra-molecular DNA triplex and RNA polymerase stall, and further, the nascent RNA re-anneal with CTT template strand form a stable RNA·DNA hybrid (Hebert 2008).

(iii) Formation of heterochromatin. Epigenetic studies in the fxn promoter and intron regions flanking the (GAA) repeat expansions have revealed modifications of condensed heterochromatin. These modifications include increased methylation of specific CpG sites in FRDA lymphoblasts, peripheral blood, brain and heart tissues and reduction of histone H3 and H4 acetylation levels and increased histone H3 lysine 9 (H3K9) trimethylation (Greene *et al*. 2007). Histone hypoacetylation was not observed in the promoter region. The initiating form of RNA polymerase II and histone H3K4 trimethylation, a chromatin mark tightly linked to transcription initiation, were found to be reduced on both FRDA alleles. In addition, a mark of transcription elongation, trimethylated H3K36, shows a reduced rate of accumulation downstream of the repeat (Kumari *et al*. 2011). These data suggest that repeat expansion reduces both transcription initiation and elongation in FRDA cells. The presence of (GAA) repeats might nucleate heterochromatin formation. HDAC along with HMTases makes gene silence, HDAC removes acetyl groups, and HMTases add methyl group, which are hallmarks of heterochromatin.

4. Triple helical DNA containing (GAA) repeats

The TRE stretch of the FRDA gene containing purine strand with (GAA) repeats as non-template strand shall be referred to as ‘ purine strand’ and the pyrimidine-rich strand with CTT repeats as the template strand shall be referred to as ‘pyrimidine strand’. The triplex DNA formed in FRDA with GAA/TTC can be of two types, first, by folding back of the purine strand (with GAA repeats) on to the same purine strand in a antiparallel fashion (figure 4A) and, second, by folding back of the pyrimidine stretch (with CTT repeats) on to the purine strand in a parallel fashion (figure 4B). Formation of such triplexes is in general initiated at a small denaturation bubble in the interior of the co-polymer, which allows the duplexes on either side to rotate slightly and to fold back, in order to make the first base triplet. The levels of DNA supercoiling and sequence of DNA determine which half strand is to become the donated third strand in the triple helix formation.

Grabczyk and Usdin (2000) have proposed a model for the formation of intra-molecular triplex in the course of transcription of GAA/TTC repeat sequences. According to their model, when RNA polymerase is reading the C*T·T strand, the non-template (GAA) strand folds back on to the duplex to form R*R·Y triplex. Similarly, in the course of transcription in the opposite direction, a Y*R·Y triplex is formed by folding back of C*T·T strand. The negative supercoiling behind the advancing polymerase seems to be the driving force behind it. Further, according to the model, the formation of transcription-driven triplex presents an obstacle to subsequent RNA polymerase at the promoter proximal end, thereby causes an increase in the frequency of promoter proximal pause (Bidchandani *et al.* 1998). The curtailed transcription of GAA/TTC repeat sequences leads to decreased production of Frataxin (Babcock *et al.* 1997).

Long GAA/TTC repeats from FRDA patients are reported to have highly tangled triple helical structure, called ‘sticky DNA’ structure. RD wells postulated that $(GAA)_n$ sequences could self-associate to form highly compact DNA structure, which is novel, and he named it ‘sticky DNA’ (Sakamoto *et al.* 1999). Sakamoto *et al.* hypothesized that two triplex structures of the type R*R·Y exchange their pyrimidine strands, and further correlated the diseases phenotype with the extent of formation of sticky DNA, which is further dependent on the GAA/TTC repeat length.

5. How are the DNA triplexes formed *in vivo*?

It is known that DNA double helix gets partially unwound during replication and transcription. Some of the regions that

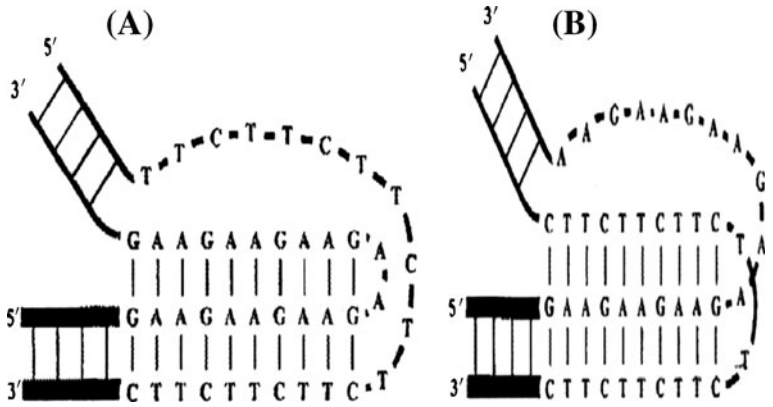

Figure 4. Schematic representation of (**A**) purine (R*R·Y) and (**B**) pyrimidine (Y*R·Y) motifs in intra-molecular DNA triple helices.

are unwound single-strand sequences with repetitive purine-rich sequences can lead to triplexes. The DNA triplexes may be formed at negative supercoil density regions that are perhaps formed transiently or generated after binding to certain proteins or during transcription induced extrusion of non-template strand at GC-rich sites. It has been shown in yeast that the regions in chromosome with (GAA) repeats are fragile and the orientation and mismatch repair play important roles in the stability of the triplex (Kim *et al.* 2008). It is now very clear that DNA structure mediates the instability by creating strong polymerase pause sites at or within the repeats by facilitating slippage or sister chromatid exchange (Rohs *et al.* 2009). Interestingly, Nikolova *et al*. have shown that the Hoogsteen base pairs exist as transient entries (in thermal equilibrium with Watson–Crick base pairs) in some DNA sequences, particularly in the CA and TA dinucleotides (Nikolova *et al.* 2011). There are reports on proteins such as DHX9 that have marked preference for triplex DNA structures containing a 3′-single-stranded overhang over other triplex and duplex DNA substrates with or without 3′-tails. The protein unwinds the Hoogsteen-bound bases by translocating with a 3′→5′ polarity (Jain *et al.* 2010). STM1 protein of *Saccharomyces cerevisiae* recognizes and specifically binds to purine-motif triplex DNA (Katayama *et al.* 2007), while C-terminal of CDP1 protein of yeast also shows high binding efficiency for purine-motif triplex DNA (Musso *et al.* 2000).

6. Characterization of triplex

Various methods can be used to get the to get insights into the various aspects of triplex structures while some of them confirm the formation of triple-stranded structures (Rajeswari 1996; Mills *et al.* 1999; Jain *et al.* 2002). Simple methods like UV absorption and calorimetric melting have been conventionally used for quantitative thermodynamic characterization of ‘inter-molecular triplexes’. More recently, the filter-binding assay was used find the thermodynamic parameters for triplexes at temperatures far from melting intervals. 2-D electrophoresis has proved to be the method of choice for thermodynamic description of triplex formation in ‘intra-molecular’ triplexes. Sedimentation, UV, NMR and IR spectroscopy, gel co-migration, 2-D electrophoresis and affinity chromatography have been used to determine specific triads and the consequences of imperfect triads for triplex stability. NMR and IR spectroscopy and X-ray analysis can also provide more detailed information on triplex formation, such as sugar pucker type or base orientation relative to the backbone, etc. However, there are no X-ray crystallographic data available for the GAA-containing DNA triplexes. The data available on AT-containing triplexes is also of powder diffraction data.

The physical methods described above differ with respect to their requirements for the quantity of DNA or oligonucleotides samples (Cantor and Schimmel 1980). For many of them (UV and CD spectrometry, fluorimetry, equilibrium sedimentation, electrophoresis and electron microscopy), a single experiment requires 1 to 10 μg DNA. Calorimetry, affinity chromatography and IR spectroscopy require an amount on order of 100 times higher. NMR and X-ray techniques demand milligram quantities of poly- or oligonucleotides. In the majority of methods, the preliminary incubations to form the triplexes are relatively long compared to the measurement period of only several minutes. In electrophoretic experiments, time requirements for triplex formation and separation are comparable (up to several hours). Equilibrium sedimentation is the longest procedure, requiring about 20 h. Generally, to avoid misinterpretation of the data, more than one method should be used for the physical characterization of triple-stranded structures. UV, circular dichroic and gel retardation assays are discussed below in detail. Characterization of triplex of

the type R*R·Y related to GAA triplet repeats found in FRDA, using various physicochemical techniques are discussed in sequel.

The three strands 5′-TCGC $(GAA)_5$ CGCT-3′, (23 R); 5′-AGCG (CTT)5GCGA-3′, (23Y) and triplex forming oligonucleotide (TFO) 5′-(AAG)5 –3′, (15R) form:

```
                3′-GAAGAAGAAGAAGAA-5′   15R(TFO)
                  * * * * * * * * * * * * * * *      Hoogsteen base pairing
5′-TCGC  GAAGAAGAAGAAGAA CGCT- 3′       23R
   ●●●●  ●●●●●●●●●●●●●●●●  ●●●●        Watson–Crick base pairing
3′-AGCG  CTT CTT CTT  CTT CTT  GCGA-5′  23Y
```

The melting profile of the duplex 23RY alone is a monophasic with a Tm of 73.10±10°C (curve a) (figure 5). However, the melting curves of the mixture of 23RY:15R (1:1) shows biphasic nature with two distinctly separate melting temperatures Tm1 at 52.6±10°C and Tm2 at 72.20±10°C. The biphasic melting nature of the mixture of 23RY and 15R clearly suggests the presence of a complex between 23RY and 15R which melts at a temperature lower than its respective duplex. This feature can be understood easily as duplex is a 23-mer formed by Watson and Crick base pairing while the triplex is only a 15 mer and is formed by reverse-Hoogsteen base pairs. Therefore, Tm of the triplex is much lower than that of corresponding duplex and the difference is about 17.6°C, and the thermodynamic parameters are also different (Jain *et al.* 2002)**.** The formation and dissociation of triplexes can be seen as an equilibrium as shown in figure 5C. Thermodynamic parameter ΔG, ΔH and ΔS can be evaluated by the shape analysis of the UV melting curves

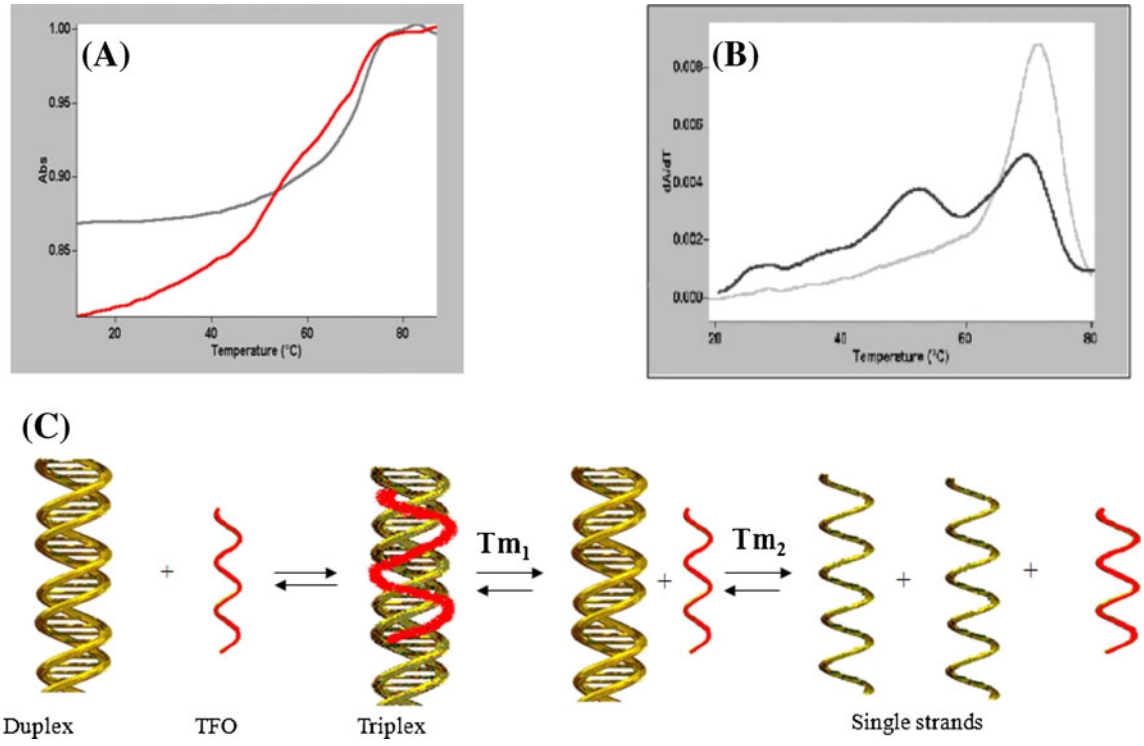

Figure 5. (**A**) **S**chematic representation of dissociation of DNA triplex formed by 5′-TCGC GAAGAAGAAGAAGAA CGTC-3′ (23R), 3′-AGCG CTT CTTCTT CTT CTT GCGA-5′ (23Y) and TFO, 3′-GAAGAAGAAGAAGAA-5′ (15R). The normalized UV thermal melting curves (**B**) and the first derivative profiles (**C**) of duplex, 23RY (curve a) and triplex, 23RY:15R (curve b) in 10 mM Sodium-cacodylate, 150 mM NaCl and 10 mM $MgCl_2$ at pH 7.4 and 20°C. The absorbance was monitored at 260 nm – each strand concentration was kept constant at 1.5 μM. The Tm_1 and Tm_2 from curve b of triplex and duplex are 52.60°C and 72.20°C respectively (Reproduced with permission from *Journal of Biomolecular Structure and Dynamics*).

at different salt concentration using the following bimolecular methods equation:

$$\Delta H(HG) = 2(n+1).\ RT2[\delta\alpha/dT]Tm1 \quad (1)$$

$$\Delta H(WC) = 2(n+1).\ RT2[\delta\alpha/dT]Tm2 \quad (2)$$

Tm1 and Tm2 are the melting temperatures of DNA triplex and duplex respectively. α is the fraction of dissociation at a given temperature, of triplex in equation 1 and duplex in equation 2. While n represents the number of molecular species, is considered to be 1 for the monomolecular process and to be 2 for the bimolecular process (Marky and Breslauer 1987). The estimated values of ΔH using the Tm data were found to be generally in good agreement with the vant Hoff analysis within 10% error. The ΔH_{WC} corresponds to Watson–Crick base pairing in duplex DNA and ΔH_{HG} to that of Hoogsteen base pairing in triplex. Obviously, the triplex is thermodynamically less stable than its host duplex. While the purine–motif triplex shows greater stability, as can be seen from the table 2, the free energy of triplex and duplex are 7.9 and 16.36 kcal/mol respectively. The enthalpy changes ΔH for duplex and its complex with drug at different ratios of D/N ('D' and 'N' represent the concentration of drug and duplex, respectively) were evaluated by the shape analysis of the UV melting curves by using the equation 1.

Circular dichroic spectra also reveal changes in the triplex and duplex. The calculated CD spectrum of the triplex (weighted sum total of 23 RY and 15R) is significantly different from that of the experimentally measured triplex. Further, the spectrum of 23RY duplex corresponds to the usual B-DNA and has the characteristic broad positive band at ~279 nm and negative band at 245 nm (figure 6A).The mathematical addition of 23RY with 15R showed a spectrum similar to the duplex 23RY with positive maxima at 279 nm and 220 nm and minimum at 248 nm. However, an experimentally generated CD spectrum on addition of 15R to the 23RY duplex showed strong changes; the positive band at 220 nm had disappeared while an intense negative band appeared at 210 nm. The negative band ~210 nm is characteristic of the triplex and generally considered as a 'hall mark' for triplex formation in oligonucleotides containing GA or GT or CT repeats (Roberts and Crothers 1992; Kandimalla *et al.* 1996; He *et al.* 1997; Jain *et al.* 2002).

2-D electrophoresis has proved to be the method of choice for thermodynamic description of triplex formation in 'intramolecular' triplexes, while gel retardation assay can be performed using [γ-32P]-labelled DNA. Triplexes exhibit much slower electrophoretic mobility than their corresponding duplexes due to the larger mass. Figure 6B shows the autoradiogram of the gel retardation assay (GRA) of 50 nM duplex, 23RY (lane 1) (with hot 23Y) and mixtures of duplex 23RYand 15R in different mole ratios 2:1 (lane 2); 1:1 (lane 3) and 1:2 (lane 4).

7. Friedreich's ataxia

7.1 *Clinical aspects*

Freidreich's ataxia (FRDA) (Romeo *et al.* 1983), named after the German doctor Nikolaus Friedreich, who first described the disease in 1863, is an autosomal recessive disease, caused by mutations in the FRDA gene, located on chromosome 9 (Campuzano *et al.* 1996). It is the most common inherited ataxia although the incidence is low. The neurodegenerative disorder, affecting both males and females, usually manifests before adolescence and is generally characterized by progressive gait ataxia and ataxia of all four limbs, hypertrophic cardiomyopathy and increased incidence of diabetes mellitus/impaired glucose tolerance. There is a progressive loss of voluntary muscular coordination and most of the patients are wheelchair bound by their late twenties, with myocardial failure being the most common cause of the death.

Table 2. Thermodynamic parameters of structural transitions of the mixtures of 23RY and 15R at different sodium ion concentrations

	Triplex-to-duplex transition			Duplex-to-open strand transition		
NaCl (mM)	ΔH (kcal mol^{-1})	ΔS (cal mol^{-1} K^{-1})	ΔG (kcal mol^{-1})	ΔH (kcal mol^{-1})	ΔS (cal mol^{-1} K^{-1})	ΔG (kcal mol^{-1})
80	46.7	130.0	7.9	73.7	192.4	16.36
150	48.0	120.4	12.3	71.5	185.0	16.40
500	49.2	112.8	15.6	66.9	168.4	16.70

All experiments were performed in cacodylate buffer containing 10 mM $MgCl_2$, pH 7.4. The strands (23R, 23Y and 15R) were mixed in equimolar ratio with each strand of 1.5 μM. The parameters were calculated using the equations 1 and 2 and from the melting curves using a two-state model (Marky and Breslauer 1987). ΔG values calculated at 25°C. Errors are ± 4°C kcal mol^{-1} for ΔH; ± 6°C cal mol^{-1} K^{-1} for ΔS and ± 0.7 kcal mol^{-1} for ΔG.

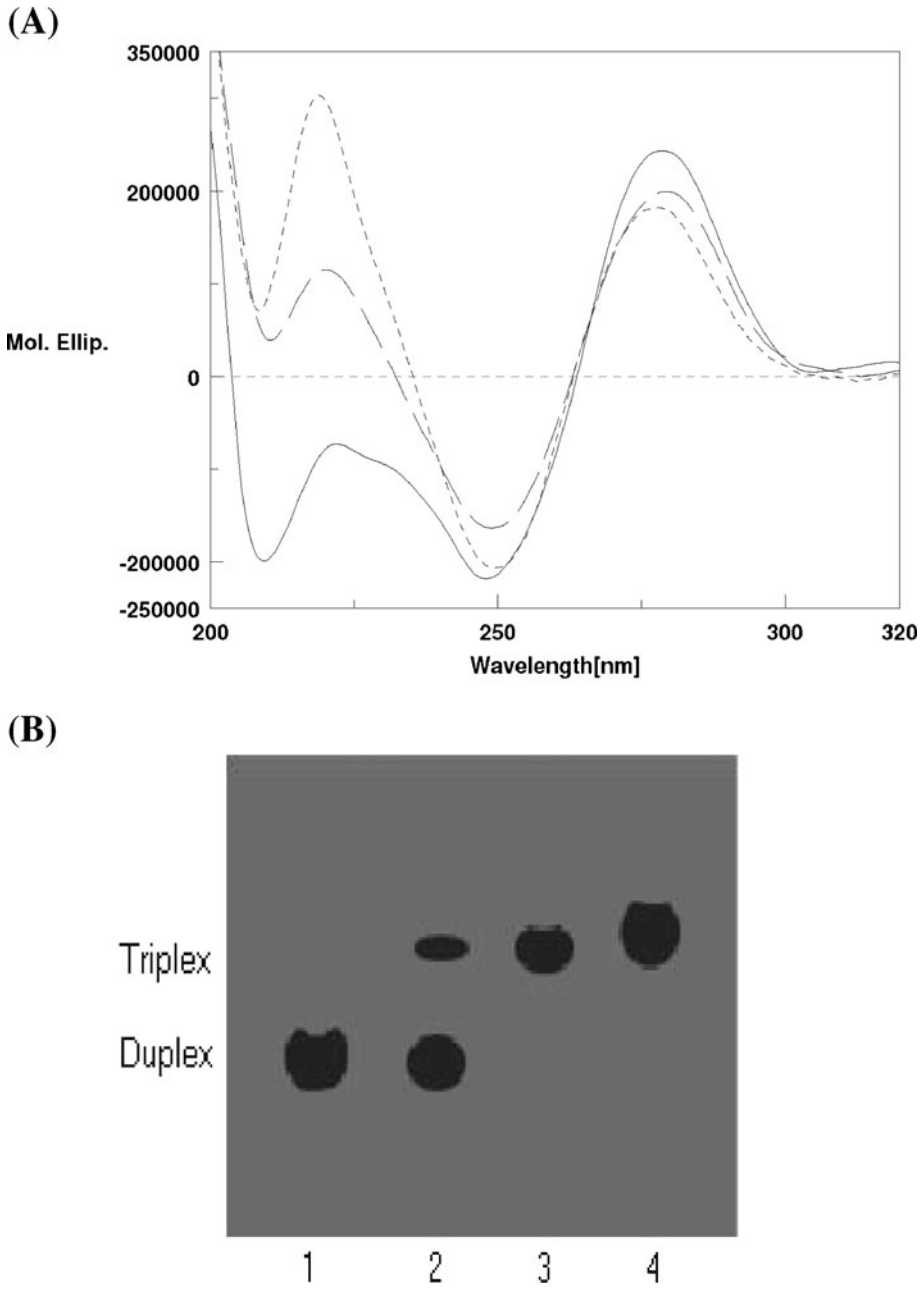

Figure 6. (**A**) Circular dichroic spectra for the duplex, 23RY (….), triplex, (23RY: 15R) experimental (—), and triplex calculated from the weighted sum of duplex and TFO (— · —). The molar ellipticity was calculated per DNA strand concentration. (**B**) Gel retardation assay of 23RY alone and the mixture of 23RY and TFO, 15R on a 15% polyacrylamide gel in Tris-borate buffer containing 10 mM $MgCl_2$, pH 7.4 at 4°C. The pyrimidine strand, 23Y of the duplex 23RY was ^{32}P labelled. The total concentration of oligonucleotide was kept constant at 50 nM. Lane 1, 23RY alone; lanes 2–4 represent mixtures 23RY and 15R in mole ratios, 2:1, 1:1 and 1:2 respectively (Reproduced from *Journal of Biomolecular Structure and Dynamics*).

Harding proposed restricted clinical criteria for FRDA diagnosis (Harding 1981). The first symptoms, as described by Harding, are noticed around the time of puberty. On average, after 10 to 15 years of disease, progressive gait and limb ataxia eventually results in the need for a wheelchair and for help with all activities of daily living. Some patients have a very severe cardiomyopathy that can cause premature death due to cardiac insufficiency or arrhythmia. The disease progresses rapidly in young adults, and patients are confined to wheelchair approximately 20 years after first appearance of

symptoms and die mainly due to heart failure. It has been proved that these neurological manifestations result from primary degeneration of dorsal root ganglion associated with axonal degeneration in posterior columns, spinocerebellar tracts and corticospinal tracts in spinal cord. Neuro-imaging reveals thinning of the cervical spinal cord. In more advanced stages of the disease, cerebellar atrophy and cerebellar vermis atrophy are also seen. However, there is no accurate measurement of clinical progression of disease. International Cooperative Ataxia Rating Scale (Trouillas *et al.*, 1997) has been widely used to assess clinical manifestations of FRDA, but none of the scales is very specific. Some of the ICARS parameters and radiological MR imaging data used in data of Indian patients from our hospital is given in table 3.

7.2 *Genetic analysis*

Identifying two expanded (GAA) repeats in intron-I of FXN gene of suspected FRDA patients confirm the disease; whereas a heterozygous genotype of one expanded and one non-expanded alleles is highly suggestive of FRDA (Montermini *et al.* 1997). Polymerase chain reaction (PCR) is the simplest and most frequently used method. PCR is based on the amplification of the gene of interest by using a polymerase enzyme that can synthesize a complementary strand to a given DNA strand in a mixture containing the 4 DNA bases and 2 DNA fragments (primers, each about 20 bases long) flanking the target sequence. The mixture is heated to separate the strands of double-stranded DNA containing the target sequence and then cooled to allow the primers to bind to their complementary sequences on the separated strands and the polymerase extends the primers into new complementary strands. Repeated heating and cooling cycles multiply the target DNA exponentially, as each new double strand separates to become two templates for further synthesis. In about 1 h, 20 PCR cycles can amplify the target by a million fold.

Table 3. Frequency of clinical signs in all the 42 suspected FRDA patients included in the study

Symptoms	Frequency (%)
Ataxia	100
Areflexia	85
Dysarthria	78
Extensor plantar response	80
Position and vibratory sense	71
Foot deformity	77
Scoliosis	60
Abnormality in ECG	11
Diabetes mellitus	0
Nystagmus	88
Radio diagnosis (CT scan / MRI of brain)	18

Representative gel photographs of 1.2% agarose gel electrophoresis of DNA samples from Friedreich's ataxia are shown in figure 7. The healthy control shows DNA band with 1.3 kb. With clinically suspected patients, genetic tests proved them to be normal for (GAA) repeats. FRDA patients, who are heterozygous in nature, show two bands at 1.4 and 4 kb bands. The FRDA homozygous patients show two bands each from two alleles at 4 and 3 kb bands respectively. Out of 120 patients, only 20 patients were found to be homozygous and 4 patients heterozygous for (GAA) repeats. The GAA repeat number of allele1 and allele2 of each are given below the corresponding lanes. The patients referred by the Neurology or Ataxia-Special clinic are 'suspected for FRDA' as the clinical findings do not confirm the disease. However, results of the genetic analysis of some of these patients reveal normal length of (GAA) repeats, which indicates that those ataxia patients are not FRDA. Such patients are referred to as 'normal' (lane 4, figure 7).

Highly advanced techniques like small pool PCR (SP-PCR) and real-time PCR, triplet repeat primed (TP) PCR (Ciotti *et al.* 2004), etc., are also used to confirm the number of (GAA) repeats. SP-PCR (Gomes-Pereira *et al.* 2004) requires very small amount of DNA and Southern blot to calculate (GAA) repeat number. And, TP PCR can only detect homozygous or heterozygous state of patients.

7.2.1 *Chromosomal sequencing by contig analysis:* In a chromosomal map, genes or other identifiable DNA fragments are assigned to their respective chromosomes, with distances measured in base pairs. These markers can be physically associated with particular bands primarily by *in situ* hybridization, a technique that involves tagging the DNA marker with an observable label, such as fluorescent and radioactive markers. The location of the labelled probe can be detected after it binds to its complementary DNA strand in an intact chromosome. The highest-resolution physical map is the complete elucidation of the DNA base pair sequence of each chromosome in the human genome. Till date, the best chromosomal maps could be used to locate a DNA fragment only to a region of about 10 Mb, the size of a typical band seen on a chromosome. Improvements in fluorescence *in situ* hybridization (FISH) methods allow orientation of DNA sequences that lie as close as 2 to 5 Mb. Modifications to *in situ* hybridization methods, using chromosomes at a stage in cell division (interphase) when they are less compact, increase map resolution to around 100,000 bp. Contig analysis involves cutting the chromosome into small pieces and each piece is cloned. The ordered fragments form contiguous DNA blocks (contigs). Currently, the resulting library of clones varies in size from

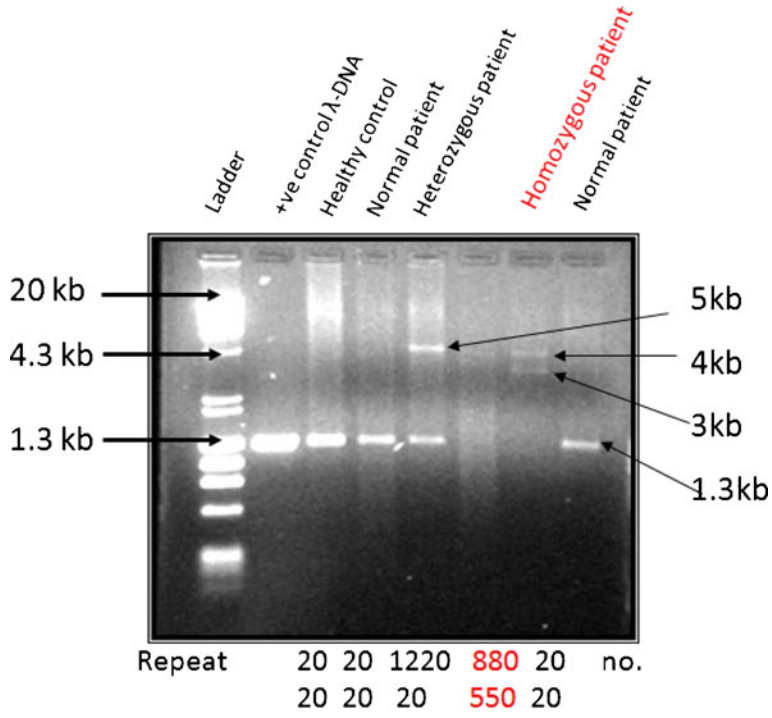

Figure 7. Representative gel photographs of 1.2% agarose gel electrophoresis of DNA samples from Friedreich's ataxia: lane M, 1 kb DNA ladder; lane 1, positive control l-DNA with 1.3 kb primer set; lane 2, healthy control; lanes 3 and 6, clinically suspected patients but whose genetic test proved them normal for (GAA) repeats; lane 4, heterozygous patient with 1.4 and 4 kb bands; lane 5, homozygous patient with 4 and 3 kb bands. Out of 120 patients, only 20 patients were found to be homozygous and 4 patients heterozygous for (GAA) repeats. The (GAA) repeat number of allele1 and allele2 of each are given below the corresponding lanes (Reproduced with permission from *DNA and Cell Biology*).

10,000 bp to 1 Mb. An advantage of this approach is the accessibility of these stable clones to other researchers. DNA sequencing using PCR is laborious as it demands cloning or post-PCR maneuvering of amplified product, while contig analysis is a high-throughput technology and therefore can sequence DNA at a large scale.

The repeat number of the DNA triplets found in normal individuals and the patients with intermediate and fully developed diseases and the plasma DNA levels are summarized in table 4 (Swarup *et al.* 2011). In some cases permutation alleles with $n \sim 60$ can convert FRDA alleles in just one generation.

The incidence of FRDA frequency is rather low: the reported cases being 1–2 per 50,000 in the UK and in Italy 1–2 per 100,000 (Babady *et al.* 2007). The incidence data in India is not available. The carrier frequency varies depending

Table 4. Clinical parameters, genetical analysis and plasma DNA of FRDA (n=15), and healthy controls (n=20), where n is the number of subjects in each group

Clinical findings						Molecular findings			
Age of patients at onset (years)		Duration of disease (years)		ICARS score		(GAA) triplet repeats Number of repeats		Plasma DNA (ng/mL)	
Mean±SD	Range	Mean±SD	Range	Mean±SD	Range	Mean±SD	Range	Mean±SD	Range
13±5	4– 2	5±3	0.5–12	48±11	32– 56	1181±117 (Controls, 21±2)	870–1220 (20 – 25)	167±43 (59±15)	64 –703 (40 – 94)

SD represents standard deviation (based on AIIMS data). Reprinted with permission from *DNA and Cell Biology* (Swarup *et al.* 2011).

on the ethnic group from 1/60 to 1/100 (Bidchandani *et al.* 1998). The majority (>95%) of patients with FRDA are homozygous for large expansions of a GAA triplet repeat sequence (66–1800 triplets). Indian patients clinically diagnosed for FRDA were also found to be homozygous for (GAA) repeat expansion (Mukerji *et al.* 2000). The (GAA) repeat in the normal Indian population shows a bimodal distribution with 94% of alleles ranging from 7 to 16 repeats and a low frequency (6%) of large normal alleles, indicating low prevalence of FRDA in the Indian population. Despite the fact that FRDA incidence is low, because if its severity of the diseases and very low survival chances of the patients, Frataxia challenges the treatment using gene therapy.

7.3 *Circulating plasma DNA*

Cell-free circulating plasma DNA is found in blood and other body fluids like urine, amniotic fluid, etc. Plasma DNA is present in very low quantities in healthy controls. However, elevated levels are being reported in a number of diseases and infections (Swarup and Rajeswari 2007). Therefore, cell-free DNA in plasma has emerged as an attractive tool in early prognosis of several human diseases. Significantly high levels ($p<0.001$) of plasma DNA of 167± 43 ng/mL were found in FRDA patients ($n=15$), while those of healthy controls ($n=20$) was only 59±15 ng/mL (Swarup and Rajeswari 2007; Swarup *et al.* 2011) (table 4). Therefore, we were able to distinguish between ataxia patients and healthy controls using plasma DNA. Although the precise mechanism by which plasma DNA enters into circulation is not known, significantly high concentrations of plasma DNA appears to be due to neuronal and muscular degeneration in these patients. Identification of genes in plasma DNA, which are overexpressed/novel, can be promising tool for prognosis of these diseases.

8. Therapeutic approaches

8.1 *Small molecules that bind to triplex*

Now it is well evident that (GAA) repeats adopt a non-B-DNA triple helical structure that stalls RNA polymerase and finally decrease frataxin protein levels in patients. Therefore, molecules/compounds that disfavour these non-B-DNA structures could be one of the possible ways to restore FXN gene transcription (Hebert 2008). Short oligonucleotides like $(GAA)_7$ are suggested to increase *in vitro* transcription of $(GAA{\cdot}TTC)_{88}$ when used at concentrations higher than 1 μM; however, the permeability of these oligonucleotides was very poor in cells *in vivo* and did not look very promising (Grabczyk and Usdin 2000), while pentamidine was found to be a potent drug for FRDA as it increased the levels of frataxin by twofolds in patient cells (Grant *et al.* 2006). As polyamides are known to preferentially bind to a stretch of purines (adenines/guanines) in DNA, Burnett *et al.* (2006) designed linear β-alanine-linked polyamides FA1 (ImPyβImPyβImβDp) (where Py is pyrrole, Im is imidazole, β is β-alanine and Dp is dimethylaminopropylamine). This β-alanine-linked polyamide increased FXN transcription by approximately threefolds at mRNA and protein levels in FRDA patient lymphoid cell line.

8.2 *Inhibition of triplex formation*

Control of formation of the abnormal DNA structure (triplex) by the GAA expansions represents an alternative strategy that could potentially ameliorate frataxin deficiency. Grabczyk and Usdin have recently designed certain oligodeoxyribonucleotides to block a particular type of triplex formation. As the (GAA·GGA/TCC·TTC) repeats in the frataxin gene locus are non-pathogenic and do not form the triplex or sticky DNA, Sakamoto's group recently tried to explore the possibility of alleviating the transcriptional deficiency of frataxin gene by using GGA·TTC-interrupted GAA triplets (Sakamoto *et al.* 1999). In order to explain the chemical and biological properties of (GAA) repeats, they proposed that the GGA·TCC interruptions introduce base mismatches and phenotypic variability into the R*R·Y triplex. Our laboratory is engaged in studying the binding of certain ligands including anti-cancer drugs to GAA/TTC sequences to find conditions that inhibit DNA triplex formation (unpublished results). Also, functional and structural studies on frataxin are likely to be facilitated by its evolutionary conservation and will hopefully generate new therapeutic possibilities

Acknowledgements

I wish to express my sincere thanks to Dr B Jayaram, Professor, Department of Chemistry and Kusuma School of Biosciences, Indian Institute of Technology, New Delhi, for giving me the opportunity to present this paper at the International Conference on Nucleic Acids in Disease and Disorder. I also want to thank my students Vishnu Swarup and Himanshu N Singh for their support in preparing this review.

References

Babady NE, Carelle N, Wells RD, Rouault TA, Hirano M, Lynch DR, Delatycki MB, Wilson RB, Isaya G, and Puccio H 2007 Advancements in the pathophysiology of Friedreich's Ataxia and new prospects for treatments. *Mol. Genet. Metabol.* **92** 23–35

Babcock M, de Silva D, Oaks R, Davis-Kaplan S, Jiralerspong S, Montermini L, Pandolfo M and Kaplan J 1997 Regulation of

mitochondrial iron accumulation by Yfh1, a putative homolog of frataxin. *Science* **276** 1709–1712

Bidchandani SI, Ashizawa T and Patel PI 1998 The GAA triplet-repeat expansion in Friedreich's ataxia interferes with transcription and may be associated with an unusual DNA structure. *Am. J. Human Genet.* **62** 111–121

Burnett R, Melander C, Puckett JW, Son LS, Wells RD, Dervan PB and Gottesfeld JM 2006 DNA sequence specific polyamides alleviate transcription inhibition associated with long GAATTC repeats in Friedreich's ataxia. *Proc. Natl. Acad. Sci. USA* **103** 11497–11502

Campuzano V, Montermini L, Moltò MD, Pianese L, Cossée M, Cavalcanti F, Monros E, Rodius F, *et al.* 1996 Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* **271** 1423–1427

Cantor CR and Schimmel PR 1980 *Biophysical chemistry* (San Francisco: WH Freeman and Company)

Chamberlain S, Shaw J, Rowland A, Wallis J, South S, Nakamura Y, von Gabain A, Farrall M and Williamson R 1988. Mapping of mutation causing Friedreich's ataxia to human chromosome 9. *Nature* **334** M248–M250

Chan PP and Glazer PM 1997 Triplex DNA: fundamentals, advances, and potential applications for gene therapy. *J. Mol. Med.* **75** 267–282

Choi J and Majima T 2011 Conformational changes of non-B DNA. *Chem. Soc.* **40** 5893–909

Christodoulou K, Deymeer F, Serdaroglu P, Ozdemir C, Poda M, Georgiou DM, Ioannou P, *et al.* 2001 Mapping of the second Friedreich's ataxia (FRDA2) locus to chromosome 9p23-p11: evidence for further locus heterogeneity. *Neurogenetics* **3** 127–132

Ciotti P, Emilio Di M, Bellone E, Ajmar F and Mandich P 2004 Triplet repeat primed PCR (TP PCR) in molecular diagnostic testing for Friedreich ataxia. *J. Mol. Diag.* **6** 285–289

Cossée M, Schmitt M, Campuzano V, Reutenauer L, Moutou C, Mandel JL and Koenig M 1997 Evolution of the Freidreich ataxia trinucleotide repeat expansion: founder effect and permutations. *Proc. Natl. Acad. Sci. USA* **94** 7452–7457

Delatycki MB, Camakaris J, Brooks H, Evans-Whipp T, Thorburn DR, Williamson R and Forrest SM 1999 Direct evidence that mitochondrial iron accumulation occurs in Friedreich ataxia. *Ann. Neurol.* **45** 673–675

Duca M, Vekhoff P, Oussedik K, Halby L and Arimondo PB 2008 The triple helix: 50 years later, the outcome. *Nucleic Acids Res.* **36** 5123–5138

Felsenfeld G and Rich A 1957 Studies on the formation of two- and three-stranded polyribonucleotides. *Biochim. Biophys. Acta* **26** 457–468

Felsenfeld G, Davies DR and Rich A 1957 Formation of three-stranded polynucleotide molecule. *J. Am. Chem. Soc.* **79** 2023

Filla A, De Michele G, Cavalcanti F, Pianese L, Monticelli A, Campanella G and Cocozza S 1996 The relationship between trinucleotide (GAA) repeat length and clinical features in Friedreich ataxia. *Am. J. Human Genet.* **59** 554–560

Fox KR 2000 Targeting DNA with triplexes. *Curr. Med. Chem.* **7** 17–37

Gacy AM, Goellner GM, Spiro C, Chen X, Gupta G, Bradbury EM, Dyer RB, Mikesell MJ, *et al.* 1998 GAA instability in Friedreich's ataxia shares a common, DNA-directed and intraallelic mechanism with other tri-nucleotide diseases. *Mol. Cell* **1** 583–593

Gomes-Pereira M, Bidchandani SI and Monckton DG 2004 Analysis of unstable triplet repeats using small-pool polymerase chain reaction. *Methods Mol. Biol.* **277** 61–76

Grabczyk E and Usdin K 2000 Alleviating transcript insufficiency caused by Friedreich's ataxia triplet repeats. *Nucleic Acids Res.* **28** 4930–4937

Grant L, Sun J, Xu H, Subramony SH, Chaires JB and Herbert MD 2006 Rational selection of small molecules that increase transcription through the GAA repeats found in Friedreich's ataxia. *FEBS Lett.* **580** 5399–5405

Greene E, Mahishi L, Entezam A, Kumari D and Usdin K 2007 Repeat-induced epigenetic changes in intron 1 of the frataxin gene and its consequences in Friedreich ataxia *Nucleic Acids Res.* **35** 3383–3390.

Harding AE 1981 Friedreich's ataxia: a clinical and genetic study of 990 families with an analysis of early diagnostic criteria and intrafamilial clustering of clinical features. *Brain* **104** 589–620

He Y, Scaria PV and Shafer RH 1997 Studies on formation and stability of the d[G(AG)5]*[d(G(AG)5].d[C(TC)5] and d[G(TG)5]*d(G(AG)5].d[C(TC)5] triple helices. *Biopolymer* **41** 431–441

Hebert MD 2008 Targeting the gene in Friedreich ataxia. *Biochimie* **90** 1131–1139

Jain A, Rajeswari MR and Ahmed F 2002 Formation and thermodynamic stability (RRY) DNA triplex in GAA/TTC repeats associated with Freidreich's Ataxia. *J. Biomol. Stuct.. Dyn*. **19** 691–699

Jain A, Bacolla A, Chakraborty P, Grosse F and Vasquez KM 2010 Human DHX9 helicase unwinds triple-helical DNA structures. *Biochemistry* **49** 6992–6999

Kamenetskii-Frank MD and Mirkin SM 1995 Triplex DNA structures. *Annu. Rev. Biochem.* **64** 65–95

Kandimalla ER Manning A and Agarwal S 1996 Single strand targeted triplex formation physicochemical and Biochemical properties of foldback triplexes. *J. Biomol. Struct. Dyn.* **14** 79–90

Katayama T, Inoue N and Torigoe H 2007 Location of the triplex DNA-binding domain of Saccharomyces cerevisiae Stm1 protein. *Nucleic Acids Symp. Ser.* **51** 123–124

Kim HM, Narayanan V, Mieczkowski PA, Petes TD, Krasilnikova MM, Mirkin SM and Lobachev KS 2008 Chromosome fragility at GAA tracts in yeast depends on repeat orientation and requires mismatch repair. *EMBO J.* **27** 2896–2906

Kumari D, Biacsi ER and Usdin K 2011 Repeat expansion affects both transcription initiation and elongation in Friedreich's ataxia cells. *J. Biol. Chem.* **286** 4209–4215

Marky LA and Breslauer KJ 1987 Calculating thermodynamic data for transition of any molecularity from equilibrium melting curves. *Biopolymers* **26** 1601–1620

Mills M, Arimondo PB, Lacroix L, Garestier T, Hélène C, Klump H and Mergny JL 1999 Energetics of strand –displacement reaction in triple helices: a spectroscopic study. *J. Mol. Biol.* **291** 1035–1105

Mirkin SM 2008 Discovery of alternative DNA structures: a heroic decade (1979–1989). *Front. Biosci.* **13** 1064–1071

Montermini L, Andermann E, Labuda M, Richter A, Pandolfo M, Cavalcanti F, Pianese L, Iodice L, *et al.* 1997 The freidreich ataxia GAA triplet repeats: permutation and normal alleles. *Human Mol. Genet.* **6** 1261–1266

Mukerji M, Choudhry S, Saleem Q, Padma MV, Maheshwari MC and Jain S 2000 Molecular analysis of Friedreich's ataxia locus in the Indian population. *Acta Neurol. Scand.* **102** 227–229

Musso M, Bianchi-Scarrà G and Van Dyke MW 2000 The yeast CDP1 gene encodes a triple-helical DNA-binding protein. *Nucleic Acids Res.* **28** 4090–4096

Nikolova EN, Kim E, Wise AA, O'Brien PJ, Andricioaei I and Al-Hashimi HM 2011 Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **470** 498–502

Rajeswari MR 1996 Tryptophan intercalation in G, C containing polynucleotides: Z to B conversion of poly [d(G-5 MC)] in low salt induced by a tetra peptide. *J. Biomol. Struct. Dyn* **14** 25–30

Roberts RW and Crothers, DM 1992 Satbility and properties of double and triple helices: Dramatic effect of RNA or DNA backbone composition. *Science* **258** 1463–1466

Rohs R, West SM, Sosinsky A, Liu P, Mann RS and Honig B 2009 The role of DNA shape in protein-DNA recognition. *Nature* **461** 1248–1253

Romeo G, Menozzi P, Ferlini A, Fadda S, Di Donato S, Uziel G, Lucci B, Capodaglio L, Filla A and Campanella G 1983 Incidence of Freidreich ataxia in Italy estimated from consanguineous marriages. *Am. J. Human Genet.* **35** 523–529

Sakamoto N, Chastain PD, Parniewski P, Ohshima K, Pandolfo M, Griffith JD and Wells RD 1999 Sticky DNA:self-association properties of long GAATTC repeats in RRY triplex structures from Friedreich's ataxia. *Mol. Cell* **3** 465–475

Smeyers P, Monros E, Vilchez J, Lopez-Arlandis J, Prieto F and Palau F 1996 A family segregating a Friedreich ataxia phenotype that is not linked to the FRDA locus. *Human Genet.* **97** 824–828

Soyfer VN and Potaman VN 1996 *Triple-helical nucleic acids* (New York: Springer-Verlag)

Swarup V, Srivastava AK, Padma MV and Rajeswari MR 2011 Quantification of cell-free circulating nucleic acid in neurological disorders, Friedreich's ataxia, spinocerebellar ataxia type 2 and 12. *DNA Cell Biol.* **30** 389–394

Swarup V and Rajeswari MR 2007 Circulating (cell-free) nucleic acids - A promising, non-invasive tool for early detection of several human diseases. *FEBS Lett.* **581** 795–709

Trouillas P, Takayanagi T, Hallett M, Currier RD, Subramony SH, Wessel K, Bryer A, Diener HC, *et al.* 1997 International Cooperative Ataxia Rating Scale for pharmacological assessment of the cerebellar syndrome. *J. Neurol. Sci.* **145** 205–211

Vasquez KM amd Glazer PM 2002 Triplex-forming oligonucleotides: principles and applications. *Q. Rev. Biophys.* **35** 89–107

Warren ST 1996 The expanding world of trinucleotide repeats. *Science* **271** 1374–1375

Watson JD and Crick FH 1953 Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171** 737–738

Wilmot GR and Warren ST 1998 A new mutational basis for disease; in *Genetic instabilities and hereditary neurological diseases* (eds) RD Wells and ST Warren (New York: Academic Press) pp 3–12

Zhao J, Bacolla A, Wang G and Vasquez KM 2010 Non-B DNA structure-induced genetic instability and evolution. *Cell Mol. Life Sci.* **67** 43–62