ACCEPTED VERSION

Belinda J. Boehm, Charles Whidborne, Alexander L. Button, Tara L. Pukala and David M. Huang

DNA triplex structure, thermodynamics, and destabilisation: insight from molecular simulations

Physical Chemistry Chemical Physics, 2018; 20(20):14013-14023

Published at: http://dx.doi.org/10.1039/c8cp02385a

PCCP

ROYAL SOCIETY OF CHEMISTRY

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

DNA triplex structure, thermodynamics, and destabilisation: insight from molecular simulations†

Belinda J. Boehm[a], Charles Whidborne[a], Alexander L. Button[a‡], Tara L. Pukala[a], David M. Huang[a]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

Molecular dynamics simulations are used to elucidate the structure and thermodynamics of DNA triplexes associated with the neurodegenerative disease Friedreich's ataxia (FRDA), as well as complexes of these triplexes with the small molecule netropsin, which is known to destabilise triplexes. The ability of molecular simulations in explicit solvent to accurately capture triplex thermodynamics is verified for the first time, with the free energy to dissociate a 15-base antiparallel purine triplex-forming oligomer (TFO) from the duplex found to be slightly higher than reported experimentally. The presence of netropsin in the minor groove destabilises the triplex as expected, reducing the dissociation free energy by approximately 50%. Netropsin binding is associated with localised narrowing of the minor groove near netropsin, an effect that has previously been under contention. This leads to localised widening of the major groove, weakening hydrogen bonds between the TFO and duplex. Consequently, destabilisation is found to be highly localised, occurring only when netropsin is bound directly opposite the TFO. The simulations also suggest that near saturation of the minor groove with ligand is required for complete triplex dissociation. A structural analysis of the DNA triplexes that can form with the FRDA-related duplex sequence indicates that the triplex with a parallel homopyrimidine TFO is likely to be more stable than the antiparallel homopurine-TFO triplex, which may have implications for disease onset and treatment.

1 Introduction

Although first reported by Watson and Crick as a double helical structure,[1] DNA is able to adopt a variety of non-canonical structures, many of which are able to form under physiological conditions and are generally accepted to have biological relevance, including involvement in certain diseases.[2–5] Notably, the formation of the DNA triple helix, in which a third strand of nucleic acid binds in the major groove of the DNA double helix (duplex) (see Fig. 1 and ESI Fig. S1), has been shown to compete with transcription factor binding[6,7] and is known to be involved in cellular regulatory processes.[2–4] Accordingly, when these structures form outside of normal circumstances, they can affect gene expression and lead to disease.[5] This is typified by the neurodegenerative disorder Friedreich's ataxia (FRDA), which has its roots in a repeat expansion of the GAA triplet within the FXN gene, which is important in the production of the iron transport protein frataxin.[8,9] It is generally accepted that this repeat expansion leads to the formation of an intramolecular triplex that significantly reduces gene expression.[8] Accordingly, the ability to destabilise such structures is therapeutically promising as a means to treat FRDA.

One means to achieve this destabilisation is through the use of ligands that bind in a sequence-specific fashion to DNA. Binding of ligands in the minor groove is generally associated with triplex destabilisation[10] and provides a promising starting point. Netropsin is one such ligand and is typical of a class of compounds, also including molecules such as distamycin, SN-6999, and Hoechst-33258,[10–12] which commonly comprise a series of heterocyclic or aromatic hydrocarbon rings possessing rotational freedom (see ESI Fig. S4 for examples) allowing the molecule to fit into the minor groove, with displacement of water. They generally feature a crescent shape, which complements the shape of the groove, and facilitate binding by promoting van der Waals interactions.[13]

Structurally, netropsin features the typical curved and planar shape of minor-groove binders and is able to bind with good selectivity for AT-rich regions, such as those present in Friedreich's ataxia (Fig. 1). In doing so, it has been reported to decrease the propeller, roll, and inclination of the DNA double helix,[14] and also to affect the minor groove width, although reports on the

[a] Department of Chemistry, School of Physical Sciences, The University of Adelaide, Adelaide, Australia, Tel: +61 8 8313 5580; Email: david.huang@adelaide.edu.au

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/cXCP00000x/

‡ Present address: Institute of Pharmaceutical Sciences, ETH Zürich, Switzerland.

latter point have been contradictory.[15–17] Furthermore, netropsin has been shown, through UV and circular dichroism spectroscopy, to destabilise triplex DNA relative to the corresponding duplex.[18] While this is a promising result, netropsin is toxic[19] and would likely not be suitable for triplex destabilisation *in vivo*. A more thorough understanding of the mechanism of triplex destabilisation by netropsin is therefore important in that it will pave the way for the development of other such destabilising molecules with greater therapeutic utility.

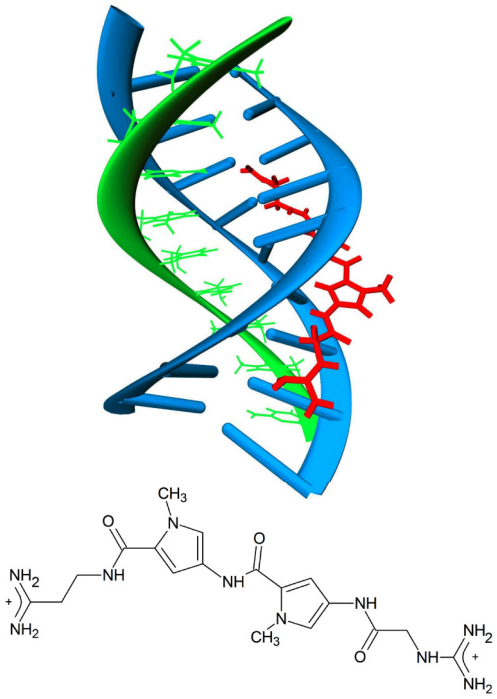


Fig. 1 (Top) Netropsin (red, stick representation) binds in the minor groove of DNA (blue, ribbon representation) while the triplex-forming oligomer (TFO) (green, ribbon + stick representation) binds in the major groove. (Bottom) Chemical structure of netropsin.

A number of different triplex structures are possible. Either Hoogsteen or reverse-Hoogsteen hydrogen bonds can be formed between the triplex-forming oligomer (TFO) and duplex, differing by the atoms of each nucleotide base involved in the pairing (see ESI Fig. S2). Additionally, the directionality of the third strand with respect to the strand of the duplex to which the TFO directly binds can be either parallel, with both strands in the same orientation, or antiparallel, with the strands orientated in the opposite directions (see ESI Fig. S3). In terms of TFO composition, for a given duplex sequence, either a homopurine or homopyrimidine triplex-forming oligomer (TFO), consisting entirely of purine (adenine or guanine) or pyrimidine (thymine or cytosine) nucleotide bases, respectively, can bind to produce a triple helix. These triplexes differ in the directionality of the TFO, with the purine TFO binding antiparallel and the pyrimidine parallel to the homopurine strand of the duplex, and the types of hydrogen-bonding interactions present, being Hoogsteen and reverse-Hoogsteen, respectively.

Both antiparallel-purine and parallel-pyrimidine triplexes are able to form with the FRDA-related duplex sequence and there is contradictory evidence for whether the purine[20] or pyrimidine[21,22] structure is most stable. However, recent experiments have shown that the pyrimidine-type triplex may in fact be significantly more stable.[23] Despite this, it remains unclear which sequence is involved in the onset of FRDA, or even whether the TFO is DNA based, with RNA TFOs also possible and shown to form more stable triplexes.[23–26] The pyrimidine-TFO triplex would be formed by the binding of a strand with C^+TT repeats, either intramolecularly through unbinding and folding back of the original sequence or intermoleculary from free nucleic acids. In this case, for a stable triplex to form, protonation of the cytosine residues is required for Hoogsteen base-pairing with guanine. Although cytosine has a pK_a of approximately 5, precluding it from being protonated at physiological pH, it has previously been shown that its pK_a may significantly increase on binding to the triplex due to the local electrostatic environment.[27,28] Furthermore, the stability of the antiparallel purine-TFO triplex has been shown to depend strongly on the guanine (G) content of the TFO, with a threshold of 40–50 % G content required for triplex formation.[29,30] Given that the antiparallel triplex potentially responsible for FRDA, at 33 % G, does not meet this G content criteria and the aforementioned experimental stability of the parallel pyrimidine-TFO structure, an examination of the relative stabilities of the triplexes that can form with the FRDA-related duplex is of interest and may help to clarify which structure is involved in the disease.

Molecular simulation provides a means of gaining an atom-level understanding of the processes of binding of the TFO in the major groove and binding of ligands such as netropsin in the minor groove. For the theoretical study of DNA structure and thermodynamics, CHARMM36[31] and AMBER parmbsc0[32,33] are the most common force fields currently in use. Both were parametrised using quantum chemical data to accurately reproduce the experimental structure of the B-DNA duplex. The AMBER force field in particular has been shown to accurately represent the structure of Z-DNA, DNA triplexes and quadruplexes, and DNA:RNA hybrids on time scales up to 10 ns.[33] Few studies have looked at the accuracy of thermodynamics calculated using this (or any other) force field, although it has been shown to overestimate base stacking free energies in DNA and RNA duplexes.[34]

The structure of the antiparallel-purine triplex potentially associated FRDA has been well characterised through simulation, including with the AMBER parmbsc0 force field,[35,36] and experiment.[37,38] Furthermore, it has been shown experimentally that triplex stability depends strongly on salt concentration, pH, and sequence.[38,39] Quantitatively, the thermodynamic stability of this triplex has been experimentally determined for a 15-base triplex-forming oligomer (TFO) under physiological conditions.[20] No computational studies, which would give an atom-level understanding of the binding or unbinding processes, have quantified the stability of the FRDA triplex nor compared it with the equivalent pyrimidine-TFO triplex. While previous experimental studies have shown that netropsin binds with sequence specificity to DNA triple helices, destabilising them,[18,19,40] a clear mechanism of how it functions in doing so has not been determined. It is also unclear whether netropsin is a useful starting point for develop-

ment of destabilisers of the FRDA triplex.

Here we examine the ability of all-atom molecular dynamics to accurately predict thermodynamic and structural properties of the DNA triple helix and to provide molecular-scale insight into the mechanism of triplex destabilisation by a model ligand. Through comparison of experimental and theoretical results, we investigate whether the currently available AMBER force fields are applicable to the quantitative analysis of the structure and thermodynamics of triple helices. In addition, the structural and thermodynamic changes on netropsin binding are examined to give insight into its mechanism of action, and the relative stabilities of the purine and pyrimidine triplexes are compared with the aim of determining their potential involvement in the onset of FRDA. Given the similar binding mode of these minor-groove binders, this work provides the first mechanistic understanding of the dissociation free energies of triplexes in the presence of such molecules, as a basis for further development of triplex destabilisers in applications of gene regulation.

2 Methods

All-atom molecular dynamics (MD) simulations were conducted on a 19 base-pair d(GC(GAA)$_5$GC) duplex with a 15-base (GAA)$_5$ homopurine TFO bound antiparallel to the central purine strand. The initial triplex geometry was generated using NAMOT.[41] The resultant DNA triplex represents a short segment of the triplex sequence thought to be responsible for FRDA and previously studied experimentally. For this triplex, two different systems were studied: one with one molecule of netropsin bound directly opposite the TFO and one with no netropsin. Additional simulations were carried out on a shorter 3-base GAA TFO bound to the same 19-base triplex. Three difference cases were studied for this triplex: netropsin bound directly opposite the TFO, underneath it, and absent. To compare the parallel and antiparallel triplexes that can be formed with the aforementioned 19-base duplex sequence, binding of a 15 base (C^+TT)$_5$ homopyrimidine TFO (generated with 3DNA[42,43]) parallel to the central purine strand of the duplex was also considered.

Most simulations were carried out in explicit solvent, with the system solvated using VMD's solvate plugin[44] with approximately randomly placed 15,000 TIP3P water molecules to give an $\approx$(80 Å)^3 cubic solvation box and neutralised with Na^+. 48 Na^+ ions were required to neutralise the 15-base system with netropsin, and 50 without. For the 3-base system, 36 Na^+ ions were required with netropsin present, and 38 when absent. The NaCl concentration was set to 0.15 M, requiring an additional 43 of both Na^+ and Cl^- ions. Long-range interactions were evaluated by means of the particle-mesh Ewald (PME) method. Bonds involving hydrogen were constrained using SHAKE,[45] allowing for a time step of 2 fs. Simulations were conducted in the NPT ensemble, with temperature and pressure maintained at 310 K and 1 atm, respectively, using a Langevin thermostat and barostat. The triplex was initially constrained using NAMD's harmonic restraints with a force constant of 10 kcal/mol/Å^2. The energy of the system was minimised and the ions and water allowed to equilibrate around the restrained triplex for 20 ps. Restraints on the triplex were released and the system equilibrated for 40 ns. The structures appeared stable in terms of the energy and root-mean-square deviation (RMSD) of atomic positions within approximately 15–20 ns (see ESI Fig. S5; for example starting and equilibrium structures see ESI Fig. S7). The final 20 ns of the simulations were used for structural measurements of the triplexes.

Additional simulations were run in implicit solvent using the generalised Born implicit solvent (GBIS) model with a solvent dielectric constant of 78.5 and ion concentration of 0.15 M to match the conditions in the explicit-solvent simulations. All structures were deemed to have reached equilibrium, through analysis of the energy and RMSD, by 10 ns of simulation (see ESI Fig. S6). All other parameters were the same as those used in the explicit-solvent simulations.

All simulations were carried out using NAMD 2.11[46] and the AMBER parm99 force field[32] with the BSC0 modification (parmbsc0)[33] and visualised with VMD.[44]

2.1 Free energy calculations

2.1.1 Free energy perturbation (FEP)

Free energy perturbation (FEP) simulations were carried out in explicit solvent for triplexes with 3- and 15-base homopurine TFOs with netropsin in one of two different positions, or absent. Simulations were set up as outlined above, with the starting configuration taken as the end point of the previously described 40 ns simulations. For the 3-base triplex, the simulation was divided into 50 windows equally spaced in the coupling parameter λ, with 0.2 ns of equilibration in each window prior to 0.3 ns of data collection. This gave a total time of 25 ns per simulation. Longer simulations were required to study the 15-base TFO as it is a large flexible molecule with a slowly fluctuating structure. FEP calculations on this system were divided into 16 windows with 5 ns of equilibration in each window prior to 7 ns of data collection. Windows were unevenly spaced to give more sampling over the first half of the simulation, where greater free energy changes occurred. The values of λ chosen were 0, 0.04, 0.08, 0.12, 0.16, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.60, 0.70, 0.80, 0.90, and 1. This gave a total simulation time of 192 ns per netropsin position. In both cases, electrostatics were decoupled linearly over the λ range 0–0.5, and vdW interactions were decoupled between $\lambda = 0$ and 1.

The thermodynamic cycle used is outlined in Fig. 2. The flexible nature of the TFO means that the coupling transformation in which the TFO's interactions with the environment are gradually introduced to produce the bound triplex are not feasible, since the time required for the TFO to spontaneously achieve the required conformation for binding would be prohibitive. Accordingly, only the forward (decoupling) transformations were considered. Restraints were applied in order to maintain the position of the TFO in the binding site as it was decoupled from the environment. The center-of-mass distance between the TFO and duplex was constrained to its average equilibrium value with a force constant of 10 kcal/mol/Å^2. The effect of these constraints on the binding free energy was accounted for as ΔG_{rest}, calculated according to the method outlined by Wang *et al*. with the coupling parameter taking 11 evenly spaced values between 0 and 1.[47]

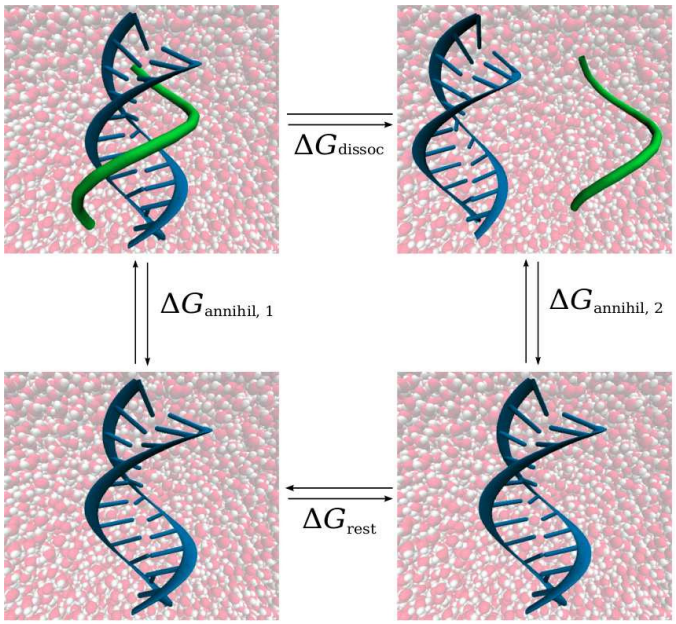


Fig. 2 Thermodynamic cycle for binding free energy calculations of TFO (green, shown in tube representation) to duplex (blue, shown in ribbon representation): $\Delta G_{\text{dissoc}} = \Delta G_{\text{annihil},1} - \Delta G_{\text{annihil},2} + \Delta G_{\text{rest}}$.

2.1.2 Replica exchange molecular dynamics

Replica-exchange molecular-dynamics simulations were conducted in implicit solvent to examine the melting behaviour of the 3-base and 15-base TFOs. For the 3-base TFO, 12 replicas at 10 K intervals between 240 and 350 K were used. Exchange acceptance ratios ranged between 0.21 and 0.32. Prior to attempting exchanges the system was equilibrated for 2.5 ns, after which each replica was at the desired temperature. This was followed by 10 ns of replica exchange. After approximately 5 ns, the number of hydrogen bonds between the TFO and duplex had plateaued at each temperature. The melting curve was constructed from the final 4 ns of data. The 15-base TFO was studied with 8 replicas at 5 K intervals between 315 and 350 K, giving exchange acceptance ratios between 0.2 and 0.5. It was simulated for 28 ns. After approximately 25 ns, the number of hydrogen bonds at each temperature had plateaued. The melting curve was constructed from the final 1.5 ns of data. Melting curves were calculated as the average number of hydrogen bonds over 5 K intervals versus temperature. The melting temperature T_{m} was determined as the temperature where the average hydrogen bond fraction was half of its maximum value.

2.2 Structural calculations

Calculations of structural properties were carried out on the final 20 ns of the 40 ns equilibrium simulations described above for both antiparallel-purine TFO (netropsin and no netropsin) and parallel-pyrimidine TFO (no netropsin) triplexes. The minor groove width was measure as the average interstrand distance between the two phosphorous atoms at residues n on strand A and $n-4$ of strand B (P_n – P_{n-4}). The major groove width was similarly defined as the average interstrand P_n – P_{n+5} distance.

The average number of hydrogen bonds between the TFO and duplex was measured over the same simulation period using the Hbonds plugin for VMD.[44] A donor–acceptor distance cut-off of 3.0 Å and an angle cut-off of 25° were employed. To calculate the average hydrogen-bond length, the distance between each possible pair of hydrogen-bonding atoms was averaged. A similar process was employed for the hydrogen bond angle which was calculated between the hydrogen bonding atoms and one of the atoms to which they are attached. The atoms in each bond length or angle measured are illustrated in Fig. 3.

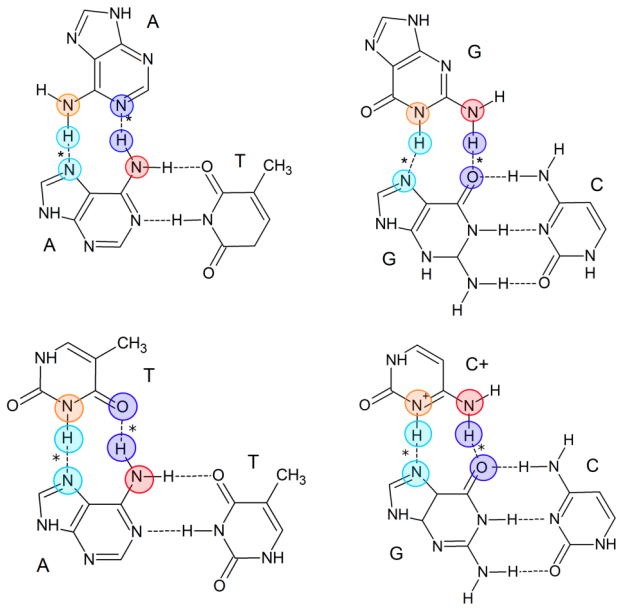


Fig. 3 Atoms between which hydrogen bond distances were calculated (light blue/dark blue) and the third atom added to calculate the hydrogen bond angle (orange/red). The top two structures are for the purine-TFO triplex and the bottom for the pyrimidine-TFO triplex.

Average triplex structures, with and without netropsin, were calculated by aligning the position of the triplex in every frame of the 20 ns of interest and then averaging the coordinates of each atom over this period using VMD's 'measure avpos' function.[44] The structures were positioned such that the long axis of the triplex was aligned with the z axis. The coordinates of the aligned backbone phosphorous atoms were compared to determine whether netropsin binding had any effect on the overall width of the helix.

In order to examine the change in degree of solvation of major groove and of the TFO on netropsin binding, the average number of water molecules in these regions was calculated. So that only the water molecules within the groove were counted (not those located close to the exterior of the triplex, effectively part of the bulk solvent), water molecules were considered to be solvating the groove if they were within 4 Å of the chain(s) of interest, but not within 4 Å of the backbone phosphorus atoms. For the solvation of the major groove, the chains involved were the two chains of the duplex, and for solvation of the TFO, water molecules were counted if within 4 Å of the third strand.

3 Results and discussion

3.1 Triplex stability: simulation vs experiment

The dissociation free energy ΔG_{dissoc} of the antiparallel purine-TFO triplex which can form from the GAA·TTC duplex sequence associated with FRDA has previously been measured experimentally for the same 15-base GAA-repeat TFO studied here.[20] This comparison with the FEP results provides a means of assessing the accuracy of the force field for calculating triplex DNA binding thermodynamics. The experimental study used the same NaCl concentration (0.15 M) as simulated here, but also used 0.01 M $MgCl_2$. Although divalent Mg^{2+} ions are expected to stabilise the triplex,[48,49] their concentration is so low that it would be equivalent to ≈2 ions in the simulations and therefore unlikely to have an appreciable effect on the calculated free energy. The free energy to dissociate the TFO from the duplex was measured in FEP simulations in explicit solvent without netropsin and is compared with the experimental result in Table 1. The simulation result is in reasonable agreement with the literature value, although the simulated triplex does appear to be overstabilised compared with experiment. However, even with long simulations of almost 200 ns in length, which appear sufficient to reach convergence, the error bars are still large, being approximately 40 % of the calculated free energy. (An analysis of the convergence in the calculated free energies is presented in the ESI, section S5.) Further increasing the simulation time may improve the precision of the calculations. Nevertheless, given that the AMBER parmbsc0 force field has been parametrised for structure rather than thermodynamics, the agreement between the computational and experimental results can be considered remarkably good.

Table 1 Free energy of dissociation, ΔG_{dissoc}, at 310 K of a 15-base TFO from the FRDA triplex from FEP simulations and experiment without netropsin and from FEP simulations with netropsin

Method	Netropsin?	ΔG_{dissoc} (kcal/mol)
Experiment[20]	Absent	10.7 ± 4.0^a
FEP	Absent	27.9 ± 9.5
FEP	Present	13.8 ± 8.5

[a] Estimated as $\Delta G_{\text{dissoc}} = \Delta H_{\text{dissoc}} - T\Delta S_{\text{dissoc}}$ from enthalpy and entropy changes, $\Delta H_{\text{dissoc}} = 48.0$ kcal/mol and $\Delta S_{\text{dissoc}} = 120.4$ cal/mol/K, obtained from van't Hoff analysis of UV-melting curves of the triplex. Note that the error bars for ΔG_{dissoc} and ΔH_{dissoc} are contradictorily given by Jain *et al.*[20] as ± 0.7 kcal/mol and ± 4.0 kcal/mol, respectively. Since ΔG_{dissoc} is derived from ΔH_{dissoc}, propagation of error requires the former value to be greater than or equal to the latter, so we expect the experimental error in ΔG_{dissoc} to be at least 4 kcal/mol, as indicated.

DNA base-stacking free energies have been shown previously to be overestimated by simulations using the same parmbsc0 force field and TIP3P water model as this work,[50] which may contribute to the overstabilisation of the triplex observed here. Furthermore, short preliminary simulations that we have carried out using the recent BSC1[51] modifications to the parm99 force field for nucleic acids (parmbsc1) showed a significant decrease in ΔG_{dissoc} relative to that measured in analogous simulations with the parmbsc0 force field. The two force fields differ in the parametrisation of certain dihedrals, which are modified in parmbsc1 to provide better structural agreement with experiment on longer time scales. Although further investigation is needed, it appears that these slight structural modifications may improve estimates of DNA triplex binding thermodynamics.

To further compare with experiment, the melting temperature T_m of the triplex was estimated from the temperature dependence of the number of TFO–duplex hydrogen bonds in the triplex measured in replica-exchange MD simulations in implicit solvent. T_m fot the 15-base TFO was found to be 332 K, slightly higher than the 325.6 K experimental value from the literature.[20] This agrees with the overstabilisation observed in the explicit-solvent FEP simulations, but also suggests that triplex binding thermodynamics may be better captured in implicit solvent with this force field. The melting curve of a 3-base TFO triplex was also measured: it displayed a broader transition and a lower melting temperature than the 15-base TFO as expected (Fig. 4), although experimental data on the melting of this structure is not available.

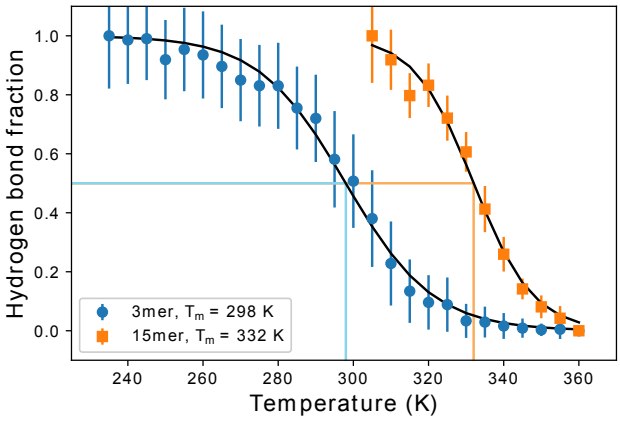


Fig. 4 Melting curves calculated from implicit-solvent replica-exchange MD simulations for the 3-base and 15-base homopurine TFOs. Vertical lines indicate the melting temperature, defined as the temperature where the hydrogen bond fraction is 0.5.

3.2 Triplex destabilisation by netropsin

3.2.1 Free energies

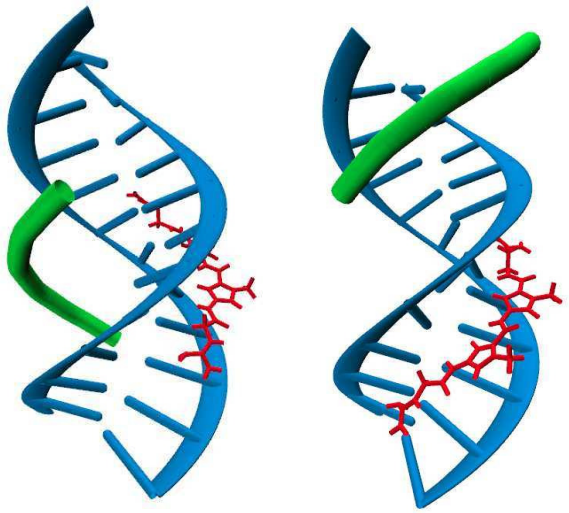
To investigate the effect of netropsin on triplex destabilisation, dissociation free energies for the 15-base homopurine TFO with and without netropsin were compared. The presence of netropsin destabilises the triplex by 14 ± 13 kcal/mol (Table 1), in line with previous qualitative reports of netropsin destabilising DNA triplexes. Nevertheless, this work is the first to quantify the impact of netropsin on triplex destabilisation.

Although clearly acting to destabilise the triplex, a single netropsin molecule is unable to induce triplex dissociation (i.e. reduce ΔG_{dissoc} to 0 or lower) under the physiological conditions simulated, even for a relatively short 15-base TFO. This observation is indicative of a localised mode of action. Binding of at least two netropsin molecules would be required for spontaneous decomposition of the 15-base triplex into its corresponding duplex and single strands, given that netropsin binding approximately halves the stability of the triplex. Given netropsin's reported 7-base binding site in a triple helix[18] and assuming that the binding of multiple netropsin molecules does not have a cooperative

effect (as would be expected for a localised mode of action), near saturation of the minor groove would be required for destabilisation of longer biologically relevant triplexes.

The solution concentration of the destabiliser required to achieve this degree of minor-groove occupation is certainly feasible: using the Langmuir adsorption model for non-cooperative binding on an adsorbate (here, the destabiliser) to a substrate (here, DNA), we estimate that a physically reasonable binding strength of the destabiliser to the minor groove of ≈7 kcal/mol would be required for a solution concentration of ≈1 μM, which is below typical LD50 concentrations of drugs of 10–100 μM (see ESI section S6 for details). Nevertheless, the relatively strong destabiliser binding required to induce TFO unbinding and the necessity of covering the entire triplex would likely lead to adverse side effects, particularly as netropsin is known to stay bound to the duplex after triplex dissociation.[18] Our results highlight the importance in designing novel therapeutics for treating FRDA of the destabiliser having a strong non-local destabilising effect, so that the binding of a single small molecule is able to induce unwinding of large sections of the TFO. This would allow the destabiliser to have a weaker binding constant, or lower solution concentration, leading to fewer potential side effects.

To explore whether netropsin exerts its destabilisation effects only on the local region where it is bound, FEP calculations were carried out on a 3-base homopurine TFO bound to the same 19-base duplex with netropsin in two different positions in the minor groove, either opposite the TFO or underneath it, as shown in Fig. 5. Although a significant decrease in triplex stability (-12.1 ± 3.1 kcal/mol) was observed when netropsin was bound opposite the TFO, a very small change in stability (-3.8 ± 3.2 kcal/mol) was observed when netropsin was bound underneath the TFO, confirming netropsin's localised mode of triplex destabilisation.



(a) Opposite **(b)** Underneath

Fig. 5 Netropsin (red) bound (a) opposite or (b) underneath a 3-base homopurine TFO (green). The dissociation free energy ΔG_{dissoc} in configurations (a) and (b) was, respectively, 12.1 ± 3.1 kcal/mol and 3.8 ± 3.2 kcal/mol lower than that without netropsin present.

3.2.2 Structural effects

As described above, netropsin appears to have a highly localised mode of action, inducing destabilisation only in the region where it is bound. To ascertain the mechanistic details of this localised destabilisation, the effect of netropsin binding on the structure of the triplex was examined.

Position of TFO in major groove The overall helix diameter was compared for the two structures of the 15-base purine TFO, with and without netropsin. On netropsin binding, the TFO was found to be pushed slightly out of the groove in part of the region opposite the netropsin molecule (Fig. 6). While very little change is seen in the structure of the two duplex chains on netropsin binding, the observed change in the position of the TFO is a clear indication of destabilisation. Additionally, the localised nature of this change, occurring solely in a region opposite part of the bound netropsin molecule, is further support for the localised mode of action suggested by the free energy calculations.

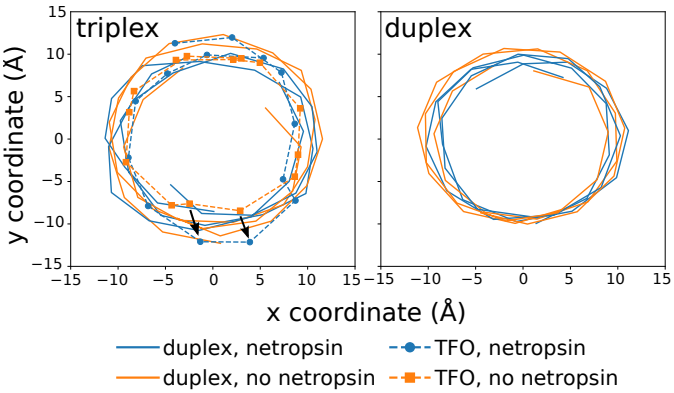


Fig. 6 The backbone structure of a (left) triplex and (right) duplex looking down the z axis. Binding of netropsin locally pushes the TFO out of the major groove. Backbone phosphorus atoms of the duplex moiety are joined by solid lines, while TFO phosphorus atoms are shown as points joined by dashed lines. Arrows indicate the TFO bases where the greatest change in position is observed.

Major and minor groove widths Structural changes to the minor groove upon netropsin binding are potentially responsible for its observed destabilisation of triplexes, yet the published results on netropsin's effects on the width of the DNA minor groove is contradictory. Experimentally, the binding of netropsin to a DNA duplex has been studied by crystallography and has been shown to widen the minor groove by 0.5–2 Å.[15,16,52] On the other hand, computational studies of the netropsin–duplex complex in solution have shown that the binding of netropsin decreases the minor-groove width by 1–2 Å, particularly over the base pairs where the charged groups of netropsin are found.[17] The solvent environment has previously been shown to influence the minor groove width[36,53] and, accordingly, it is possible that the solid-state structure measured in crystallographic studies differs from the arguably more biologically relevant aqueous structure studied by simulation. It should be noted that the aforementioned studies focused on the structure of the netropsin–duplex complex, while the structure of the netropsin–triplex complex has not previously

been examined.

MD simulations of both a 15-base homopurine-TFO triplex and a duplex carried out in this work indicate that binding of netropsin reduces the minor groove width in both cases by approximately 2–4 Å in the region where netropsin is bound (Fig. 7b). This result contrasts with experimental results [15,16,52] but agrees with previous simulations for the duplex, which employed the unmodified AMBER parm99 force field. [17] As similar reductions in groove width were observed for both duplex and triplex, it is unlikely to be the presence of the TFO that results in this deviation from experiment. However, molecular simulations have shown that the interaction of atomic ions or charged molecular species with the minor groove can cause significant narrowing [17,54] which may explain the observed effect for positively charged netropsin. It therefore appears that the structure of DNA triplexes depends strongly on the environment and that solid-state structures may not accurately describe all facets of the solution structure.

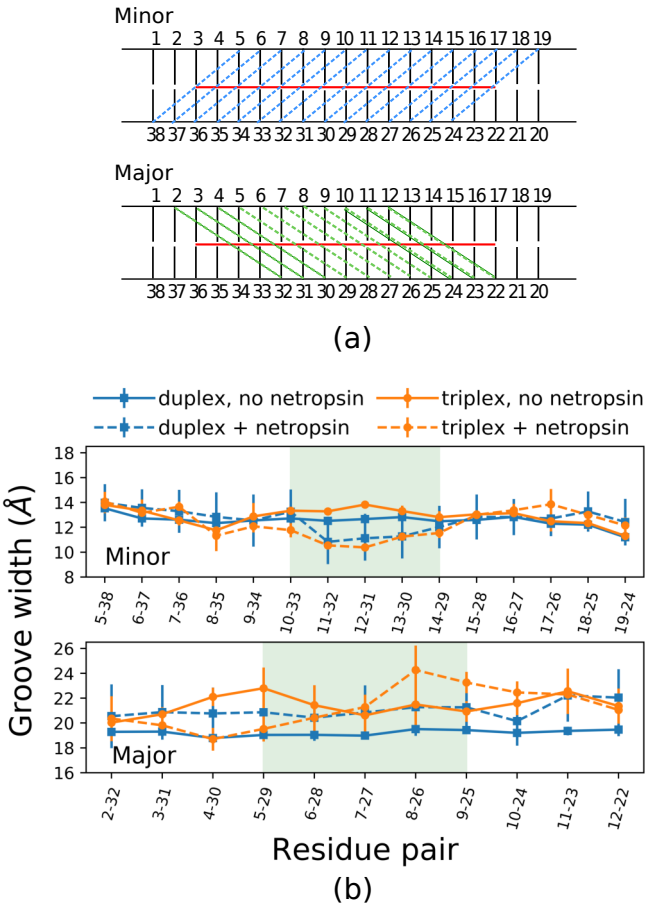


Fig. 7 (a) Definition of the minor (top) and major (bottom) groove widths. The dashed lines connect the bases between which the width is measured. The red line indicates the region where the 15-base TFO binds. (b) Changes in minor (top) and major (bottom) groove widths on binding either a 15-base purine TFO or netropsin. Dashed lines indicate the structures (duplex or triplex) in the absence of netropsin and solid lines the structures with netropsin present. The shaded region indicates the region of the major groove opposite netropsin, or the region of the minor groove where netropsin is bound. The 15-base TFO, when bound, covers the entire plotted region.Residue numbers correspond to (a).

In contrast to the minor groove, the effects of netropsin binding on the major groove are not so simple. On binding to the duplex, netropsin induces a global, uniform, widening by 2–3 Å of the major groove. However, uniform widening is not observed on netropsin binding to the triplex. Instead, relative to the triplex without netropsin, the major groove width is increased in certain regions, but decreased in others. With reference to the position of the TFO in the major groove (Fig. 6), the residues where the greatest change occurs in TFO backbone position occur where the major groove is widened by netropsin, which also correspond to where the minor groove is narrowed. This correlation between the width of the major groove and how well the TFO sits in it indicates that major groove widening may be an important factor in the netropsin's mechanism of triplex destabilisation.

TFO and major groove solvation In considering the slight displacement of the TFO from the major groove seen in Fig. 6, the solvation of both the TFO and this groove may potentially play a role in destabilising the triplex as both should become more solvent-accessible on unbinding. However, little difference was observed between the overall number of solvating water molecules either in the absence or presence of netropsin (Table 2). It is therefore unlikely that changes in the solvation of the major groove are significant contributors to the destabilisation of triplexes by netropsin. However, as the TFO begins to unwind more completely, such as in the presence of more netropsin molecules, solvation may begin to have a larger effect.

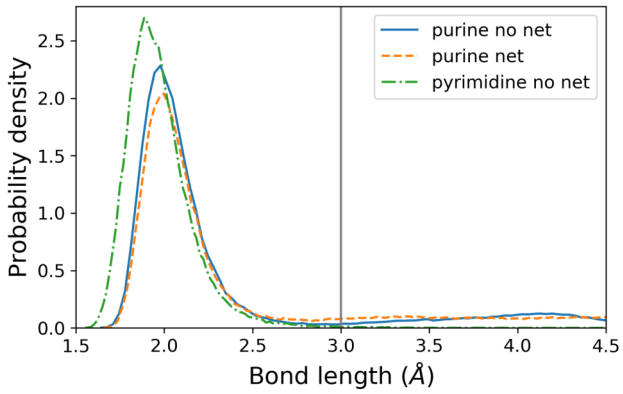
Table 2 Average number of water molecules within 4 Å of either the TFO or the two chains of the duplex, but not within 4 Å of backbone phosphorus atoms, for the triplex with the 15-base homopurine TFO in the presence or absence of netropsin.

	Netropsin?	Solvating TFO	Solvating groove
	Absent	48.2 ± 0.8	157.9 ± 0.7
	Present	50.8 ± 0.8	155.6 ± 0.8
Difference	*Overall: 0.2 ± 1.5*	*2.6 ± 1.1*	*-2.3 ± 1.0*

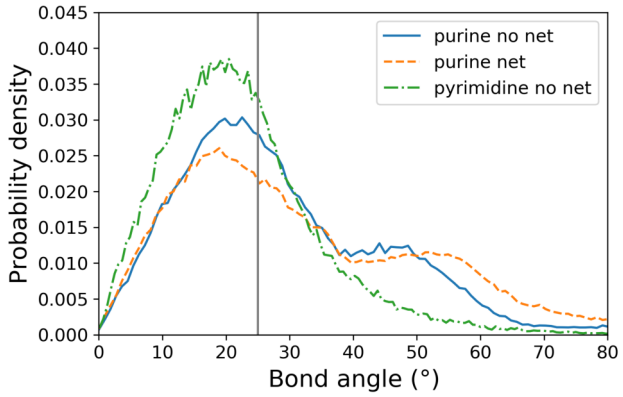
Hydrogen bonding The displacement of the TFO from the major groove, as shown in Fig. 6, could be expected to be associated with a decrease in the number of hydrogen bonds between TFO and duplex. However, the number of hydrogen bonds between the purine TFO and the duplex was found not to change significantly on netropsin binding (Table 3). In an ideal DNA triplex structure, each Hoogsteen base pair should contain two hydrogen bonds, giving a total of 30 hydrogen bonds for the 15-base TFO. The results shown in Table 3 indicate approximately a third of this number. Structurally, this is associated with a misalignment of adenine bases as the TFOs show some distortion in the base stacking, which will be discussed in detail below in comparing antiparallel purine-TFO and parallel pyrimidine-TFO triplexes.

While the addition of netropsin has a negligible effect on the number of hydrogen bonds, it does noticeably change the distribution of hydrogen bond lengths and angles, broadening and shifting the distributions to larger lengths and angles, as shown in Fig. 8 (distributions for different hydrogen bond types and joint length/angle distributions are also given in the ESI in Figs. S11 and S12). We have quantified the impact of the

hydrogen-bond length distribution on the TFO–duplex binding free energy by estimating the average hydrogen-bond strength between the TFO and duplex from the simulated bond-length distributions and distance-dependent hydrogen-bond potentials from quantum-chemical calculations,[55,56] as described in detail in ESI section S8. (We have ignored the effect of the bond-angle distribution, since angle-dependent potentials are not available and because the length dependence is expected to make the dominant contribution, the effect of which would only be enhanced by accounting for the angular dependence.) Summing the contributions of all possible TFO–duplex hydrogen bonds yields the total hydrogen-bonding interactions energies given in Table 3, which shows that the change in the hydrogen-bond length distribution upon netropsin binding reduces the hydrogen-bonding energy by approximately 11 kcal/mol, in line with the 14 ± 13 kcal/mol reduction in the calculated TFO–duplex binding free energy due to netropsin. This result suggests that triplex destabilisation by netropsin can be explained entirely by weakening of TFO–duplex hydrogen bonds.



(a) Bond length distributions



(b) Bond angle distributions

Fig. 8 Hydrogen bond (a) length and (b) angle distributions for the three triplexes studied in this work. Grey lines indicate the cutoffs employed when counting the average number of hydrogen bonds.

Overall structural effects of netropsin binding Based on the above analysis, several factors appear to combine to produce the observed destabilisation of the DNA triplex by netropsin. Local narrowing of the minor groove around netropsin leads to widening of the major groove near netropsin, which weakens but does not completely break TFO–duplex hydrogen bonds and results in the TFO moving 2–3 Å out of the groove in this region. These structural changes are highly localised and explain the localised mode of action of netropsin in triplex destabilisation.

Table 3 Average number of hydrogen bonds and total hydrogen-bonding energy between TFO and duplex in a triplex with a 15-base homopurine or homopyrimidine TFO in the presence or absence of netropsin. The hydrogen-bonding energy was estimated using the method described in ESI section S8.

TFO type	Netropsin?	Number	Total Energy (kcal/mol)
Purine	Absent	8.6 ± 0.2	-139
Purine	Present	9.0 ± 0.5	-128
Pyrimidine	Absent	15.2 ± 0.5	-251

Given that minor groove binders such as berenil and distamycin have similar structures to netropsin, are known to destabilise triplexes, and have been shown to have a similar dependence on triplex:ligand ratio require for unbinding,[37] it is likely that our results are generalisable to other small molecules and they have similar mechanisms of action. Further study of these and other similar ligands would be required to confirm this hypothesis.

3.3 Comparison of parallel and antiparallel triplexes

As previously stated, whether the triplex responsible for FRDA is of the antiparallel-purine or parallel-pyrimidine type is not definitively known.[20–23] It is also possible that the GAA stretch responsible for the disease is a favourable binding site for free nucleic acids in solution, such as mRNA.[23–26] Although the effect of changing the backbone composition from DNA to RNA is beyond the scope of this work, exploring the relative stability of parallel and antiparallel DNA triplex can provide new insight into why one triplex is more likely to form with a given duplex sequence.

The average structures of parallel and antiparallel triplexes comprising the same 19 base-pair d(GC(GAA)$_5$GC) duplex sequence and a 15-base (GAA)$_5$ homopurine TFO and (C$^+$TT)$_5$ homopyrimidine TFO, respectively, were compared in equilibrium simulations. Protonated cytosine was considered in the latter case as it has previously been shown to be stable at physiological pH in non-terminal residues in DNA triplexes.[28] Several factors likely to affect their relative stabilities were examined, namely the number of hydrogen bonds between the TFO and purine strand of the duplex, the structure of the base triplets, overall positioning of the TFO in the major groove, and the sugar pucker preference of each base of the TFO on binding to the duplex.

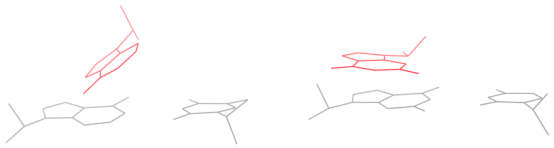
3.3.1 Hydrogen bonds and base-triplet structure

The purine TFO in the antiparallel triplex forms only just over half the number of hydrogen bonds with the purine strand of the duplex as the pyrimidine TFO in the parallel triplex (Table 3). A closer examination of the purine-TFO triplex structure reveals that the A*A pairs, where the * indicates a Hoogsteen or reverse-Hoogsteen pair, do not sit in the plane of the other base with which they pair, but are twisted with respect to the plane of the duplex rungs (Fig. 9). As these distorted bases lie further from

the duplex, this distortion accounts for the lower than expected hydrogen bond counts in this type of triplex discussed earlier. This distortion can be explained by the non-isomorphism of A*A and G*G base pairs, which requires deformation of the backbone if both are to bind simultaneously,[57] and weaker reverse Hoogsteen A*A bonding interactions compared with their G*G counterparts (as indicated by quantum chemical calculations of A*A·T and G*G·C reverse-Hoogsteen trimers[56]), which leads to deviations of the A*A pairs from the ideal (planar) configuration in order to form the stronger G*G interactions.

Significant distortion is observed immediately following every A–G junction in the purine-TFO triplex, whereas all bases in the pyrimidine-TFO triplex pair well, giving an undistorted backbone (Fig. 10) and a higher number of TFO–duplex hydrogen bonds (Table 3). These structural differences suggest, contrary to some measurements[8,20] but consistent with recent experiments,[23] that the parallel pyrimidine-TFO triplex could be more stable than the antiparallel purine-TFO triplex.

Additionally, the estimated total hydrogen-bonding energy of the pyrimidine-TFO triplex given in Table 3 indicates that it is much more stable than the purine-TFO triplex. This is predominantly due to the C^+G Hoogsteen pairs, which are known to form significantly stronger hydrogen bonds than all other pairs, including even the Watson–Crick CG equivalent found in duplex DNA.[56] The presence of these stronger hydrogen bonds therefore also indicate that the pyrimidine-TFO triplex is likely to be more stable.



(a) T.A*A **(b)** C.G*G

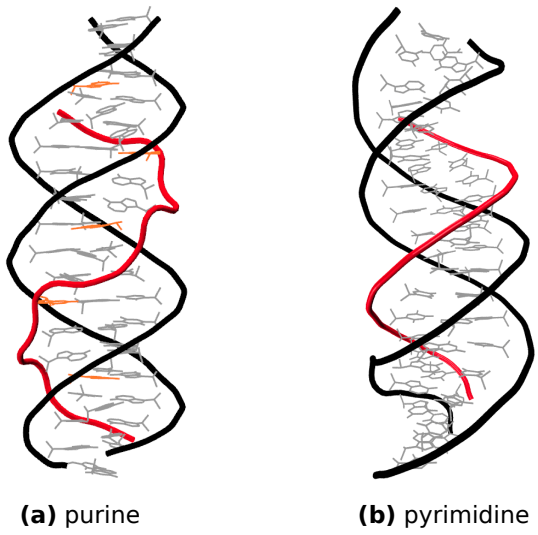
Fig. 9 Average structure of a (a) T.A*A and (b) C.G*G base triplet in a d(GC(GAA)$_5$CG) duplex with (GAA)$_5$ TFO. Bases belonging to the third strand are shown in red. Hydrogens have been removed for clarity.

4 Conclusions

Molecular dynamics simulations were used in this work to provide new molecular-level insight into the structure and thermodynamics of the DNA triplex associated with the neurodegenerative disease Friedreich's ataxia and the mechanism of action of the small molecule netropsin known to destabilise DNA triplexes.

Dissociation free energy calculations were carried out on a 15-base purine DNA triplex representative of that thought to be associated with Friedreich's ataxia. Calculated free energies, obtained using free energy perturbation, slightly overestimate experimental values from the literature, but are in reasonable agreement. This result is the first to validate the ability of molecular dynamics simulations in explicit solvent to accurately capture DNA triplex binding thermodynamics.

Binding of netropsin to the same triplex is associated with a 14 kcal/mol decrease in the dissociation free energy of the triplex, in line with previous reports of it destabilising the triplex. This result is the first quantitative report of this free energy for the antiparallel purine-TFO triplex potentially associated with Friedreich's ataxia. The position of the netropsin molecule relative to the TFO was also shown to significantly affect the stability of the triplex, with netropsin binding below the TFO inducing little destabilisation. These results indicate a localised mode of action and imply that near saturation of the minor groove with ligand is required for destabilisation. As this will require either strong binding or high concentrations of ligand, this requirement of saturation is likely to lead to adverse side effects. In considering the design of future destabilisers, a more global mode of action should be favoured.



(a) purine **(b)** pyrimidine

Fig. 10 Average structures of 15-base (a) purine-TFO and (B) pyrimidine-TFO d(GC(GAA)$_5$CG) triplexes. TFO shown in red. For the purine triplex, guanines are shown in orange.

A structural analysis of netropsin binding to DNA indicated that triplex destabilisation by netropsin results from weakening of TFO–duplex hydrogen bonds, likely due to widening of the major groove induced by narrowing of the minor groove in the immediate vicinity of the netropsin binding site. The observed narrowing of the minor groove upon binding of netropsin contradicts previous crystallographic measurements but agrees with previous simulation work on binding to a DNA duplex. It is therefore likely that the triplex environment has a significant effect on its structure. The effect on DNA structure was highly localised to the 4–5 bases directly surrounding netropsin, explaining the similarly localised effect on the free energy of destabilisation. As minor groove binders have generally been found to destabilise DNA triplexes, the results presented here, namely the local narrowing of the minor groove and weakening of hydrogen bonds, are likely generalisable to other structures of the same class. Although further study of other minor-groove binders is required, this work represents an important step towards understanding the mechanism of DNA triplex destabilisation.

A comparison of the structures of the two different DNA triplexes than can form from the duplex sequence associated with Friedreich's ataxia showed that the structure of the antiparallel

purine TFO in the antiparallel triplex is highly distorted compared with the pyrimidine TFO in the parallel triplex, suggesting that the parallel triplex, consistent with recent experimental results, is more stable than the antiparallel triplex. This finding provides further insight into how Friedreich's ataxia may be caused, and may aid in guiding the design and development of future ligands to target this sequence.

Acknowledgements

B.J.B. thanks the Australian Government for a Research Training Program Stipend Scholarship. This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government, and from The University of Adelaide's Phoenix High Performance Computing Service.

References

1 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737–738.
2 N. Saini, Y. Zhang, K. Usdin and K. S. Lobachev, *Biochimie*, 2013, **95**, 117–123.
3 S. W. Paugh, D. R. Coss, J. Bao, L. T. Laudermilk, C. R. Grace, A. M. Ferreira, M. B. Waddell, G. Ridout, D. Naeve, M. Leuze, P. F. LoCascio, J. C. Panetta, M. R. Wilkinson, C.-H. Pui, C. W. Naeve, E. C. Uberbacher, E. J. Bonten and W. E. Evans, *PLoS Comput. Biol.*, 2016, **12**, 1–20.
4 B. P. Belotserkovskii, R. Liu, S. Tornaletti, M. M. Krasilnikova, S. M. Mirkin and P. C. Hanawalt, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 12816–12821.
5 A. Bacolla and R. D. Wells, *J. Biol. Chem.*, 2004, **279**, 47411–47414.
6 N. Tsuda, A. Matsumoto, A. Ito, T. Uneda, A. Tanabe, S. Obika and T. Imanishi, *Nucleic Acids Symp. Ser.*, 2005, **49**, 335–336.
7 C. Hélène, N. T. Thuong and A. Harel, *Ann. N. Y. Acad. Sci.*, 1992, **660**, 27–36.
8 E. Grabczyk and K. Usdin, *Nucleic Acids Res.*, 2000, **28**, 2815–2822.
9 V. Campuzano, L. Montermini, M. D. Molto, L. Pianese, M. Cossee, F. Cavalcanti, E. Monros, F. Rodius, F. Duclos and et al., *Science*, 1996, **271**, 1423–1427.
10 A. K. Jain and S. Bhattacharya, *Bioconjug. Chem.*, 2010, **21**, 1389–1403.
11 T. J. Stonehouse and K. R. Fox, *Biochim. Biophys. Acta, Gene Struct. Expression*, 1994, **1218**, 322–330.
12 M. Durand, N. Thuong and J. Maurizot, *Biochimie*, 1994, **76**, 181–186.
13 S. Neidle, *Nat. Prod. Rep.*, 2001, **18**, 291–309.
14 J. Dolenc, C. Oostenbrink, J. Koller and W. F. van Gunsteren, *Nucleic Acids Res.*, 2005, **33**, 725–733.
15 C. M. Nunn, E. Garman and S. Neidle, *Biochemistry*, 1997, **36**, 4792–4799.
16 K. Van Hecke, P. C. Nam, M. T. Nguyen and L. Van Meervelt, *FEBS J.*, 2005, **272**, 3531–3541.
17 B. Wellenzohn, W. Flader, R. H. Winger, A. Hallbrucker, E. Mayer and K. R. Liedl, *Biopolymers*, 2002, **61**, 276–286.
18 Y.-W. Park and K. J. Breslauer, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 6653–6657.
19 R. M. Wartell, J. E. Larson and R. D. Wells, *J. Biol. Chem.*, 1974, **249**, 6719–6731.
20 A. Jain, M. R. Rajeswari and F. Ahmed, *J. Biomol. Struct. Dyn.*, 2002, **19**, 691–699.
21 V. N. Potaman, E. A. Oussatcheva, Y. L. Lyubchenko, L. S. Shlyakhtenko, S. I. Bidichandani, T. Ashizawa and R. R. Sinden, *Nucleic Acids Res.*, 2004, **32**, 1224–1231.
22 S. Mariappan, P. Catasti, L. A. Silks, E. Bradbury and G. Gupta, *J. Mol. Biol.*, 1999, **285**, 2035 – 2052.
23 J. Li, A. Begbie, A. Button, C. Whidborne, Y. Pouferis, D. Huang and T. Pukala, *J. Am. Soc. Mass Spectrom.*, submitted.
24 S. Wang and E. T. Kool, *Nucleic Acids Res.*, 1995, **23**, 1157.
25 R. W. Roberts and D. M. Crothers, *Science*, 1992, **258**, 1463–1466.
26 C. H. Gotfredsen, P. Schultze and J. Feigon, *J. Am. Chem. Soc.*, 1998, **120**, 4281–4289.
27 G. R. Pack, L. Wong and G. Lamm, *Int. J. Quantum Chem.*, 1998, **70**, 1177–1184.
28 J. L. Asensio, A. N. Lane, J. Dhesi, S. Bergqvist and T. Brown, *J. Mol. Biol.*, 1998, **275**, 811–822.
29 P. Vekhoff, A. Ceccaldi, D. Polverari, J. Pylouster, C. Pisano and P. B. Arimondo, *Biochemistry*, 2008, **47**, 12277–12289.
30 A. Debin, C. Laboulais, M. Ouali, C. Malvy, M. Le Bret and F. Svinarchuk, *Nucleic Acids Res.*, 1999, **27**, 2699–2707.
31 K. Hart, N. Foloppe, C. M. Baker, E. J. Denning, L. Nilsson and A. D. MacKerell, *J. Chem. Theory Comput.*, 2012, **8**, 348–362.
32 T. E. Cheatham, P. Cieplak and P. A. Kollman, *J. Biomol. Struct. Dyn.*, 1999, **16**, 845–862.
33 A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton and M. Orozco, *Biophys. J.*, 2007, **92**, 3817–3829.
34 R. F. Brown, C. T. Andrews and A. H. Elcock, *J. Chem. Theory Comput.*, 2015, **11**, 2315–2328.
35 C. Laughton and S. Neidle, *Nucleic Acids Res.*, 1992, **20**, 6535–6541.
36 A. Arcella, G. Portella, M. L. Ruiz, R. Eritja, M. Vilaseca, V. Gabelica and M. Orozco, *J. Am. Chem. Soc.*, 2012, **134**, 6596–6606.
37 S. S. Mariappan, X. Cheng, R. B. van Breemen, L. A. Silks and G. Gupta, *Anal. Biochem.*, 2004, **334**, 216–226.
38 R. W. Roberts and D. M. Crothers, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 4320–4325.
39 E. G. Plum and K. J. Breslauer, *J. Mol. Biol.*, 1995, **248**, 679–695.
40 M. Durand, N. T. Thuong and J. C. Maurizot, *J. Biol. Chem.*, 1992, **267**, 24394–24399.
41 C.-S. Tung and E. S. Carter, II, *Bioinformatics*, 1994, **10**, 427.
42 X. J. Lu and W. K. Olson, *Nucleic Acids Res.*, 2003, **31**, 5108–5121.
43 X. J. Lu and W. K. Olson, *Nature Protocols*, 2008, **3**, 1213–1227.

44 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.
45 J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
46 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
47 J. Wang, Y. Deng and B. Roux, *Biophys. J.*, 2006, **91**, 2798–2814.
48 M. Rougee, B. Faucon, J. L. Mergny, F. Barcelo, C. Giovannangeli, T. Garestier and C. Helene, *Biochemistry*, 1992, **31**, 9269–9278.
49 N. Sugimoto, P. Wu, H. Hara and Y. Kawamoto, *Biochemistry*, 2001, **40**, 9396–9405.
50 F. Häse and M. Zacharias, *Nucleic Acids Res.*, 2016, **44**, 7100.
51 I. Ivani, P. D. Dans, A. Noy, A. Perez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goni, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpi, C. Gonzalez, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case and M. Orozco, *Nat. Methods*, 2016, **13**, 55–58.
52 M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura and R. E. Dickerson, *Proc. Natl. Acad. Sci. U. S. A.*, 1985, **82**, 1376–1380.
53 L.-A. Harris, L. D. Williams and G. B. Koudelka, *Nucleic Acids Res.*, 2014, **42**, 14053–14059.
54 D. Hamelberg, L. McFail-Isom, L. D. Williams and W. D. Wilson, *J. Am. Chem. Soc.*, 2000, **122**, 10513–10520.
55 K. Brameld, S. Dasgupta and W. A. Goddard, *J. Phys. Chem. B*, 1997, **101**, 4851–4859.
56 M. Peters, I. Rozas, I. Alkorta and J. Elguero, *J. Phys. Chem. B*, 2003, **107**, 323–330.
57 J.-S. Sun and C. Hélène, *Curr. Opin. Struct. Biol.*, 1993, **3**, 345–356.