



# Computational Methods to Study G-Quadruplex–Ligand Complexes

Shozeb Haider\*

**Abstract** | Quadruplexes are four-stranded DNA or RNA structures formed from tandem repeats of guanine-rich sequences. They are present throughout the genome, namely at telomeres, in promoter and untranslated regions of important genes. Quadruplexes are thermodynamically stable structures and once formed can be a hindrance to cellular processes like transcription, translation, replication among others. Based on these findings, it has been proposed that small molecules can be used to selectively bind and stabilise quadruplexes. In this review, we explore computational methods that can be used to study quadruplex–ligand interactions. These methods not only provide interpretation of experimental data but also generate information that cannot be accessed by experimental methods.

## 1 Introduction

Guanine has hydrogen-bond donors and acceptors present on both the Watson–Crick and Hoogsteen faces. Four guanines can readily self-associate together via Hoogsteen hydrogen bonding in a co-planar array to form structures termed as G-quartets<sup>1</sup>. A total of 8 hydrogen bonds stabilise a quartet (Fig. 1). G-quartets, in the presence of cations, can stack on top of one another to make a four-stranded G-quadruplex (G4) DNA or RNA structure. All G4 structures comprise two distinct structural features, namely a stem formed from a set of G-quartets stacked together and unpaired bases which link the quartets and form the loops. If the two strands are antiparallel, then the loop between them is positioned above the plane of the stacked G-quartets. However, if the two strands are parallel, then the connecting loop is positioned within the groove formed between two sugar-phosphate backbones of the guanines.

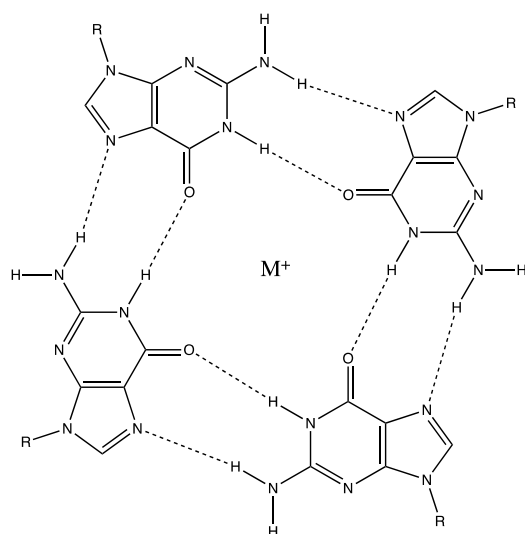
G-quadruplexes are highly polymorphic in nature, with regard to its sequence, nature of cation stabilisation, strand orientation and the glycosidic conformation of the bases. They can be formed from a single strand (monomeric), two strands (dimeric) and four strands (tetrameric). When monomeric, at least four-tandem guanine repeats are needed with the general sequence:  $G_{2-5} L G_{2-5} L G_{2-5} L G_{2-5}$ , where G represents

guanine in the range of 2–5 nucleotides and L represents the intervening loop sequence. These sequences were first identified at the ends of eukaryotic chromosomes. In case of intermolecular dimeric quadruplexes, two strands come together having a consensus sequence of  $G_{2-5} L G_{2-5}$ . Finally, a tetramolecular quadruplex is formed when four separate strands come together. The guanine tracts need not be of identical length and the quadruplexes can be formed within single-stranded or double-stranded DNA or RNA. Of these, the monomeric intramolecular quadruplex is the most important form in terms of occurrence and biological importance.

The human genome contains between 350,000 and 700,000<sup>2–4</sup> distinct putative quadruplex forming sequences, of which the telomeric and those present in the oncogenic promoter sequences have been extensively studied. However, experimental trapping of actual quadruplex structures in human chromatin has found a much lower prevalence, of ca 10,000<sup>5</sup>. The functional relevance of quadruplex formation at different sites has been supported by the isolation of proteins that bind and promote the formation of quadruplex structures, such as transcription factors, nucleases, topoisomerases, resolvases and helicases. Unambiguous evidence to their existence in vivo has been demonstrated recently,

UCL School of Pharmacy,  
University College  
London, 29-39 Brunswick  
Square, London WC1N  
1AX, UK.

\*shozeb.haider@ucl.ac.uk



**Figure 1:** A quartet formed by a co-planar association of four guanine residues.

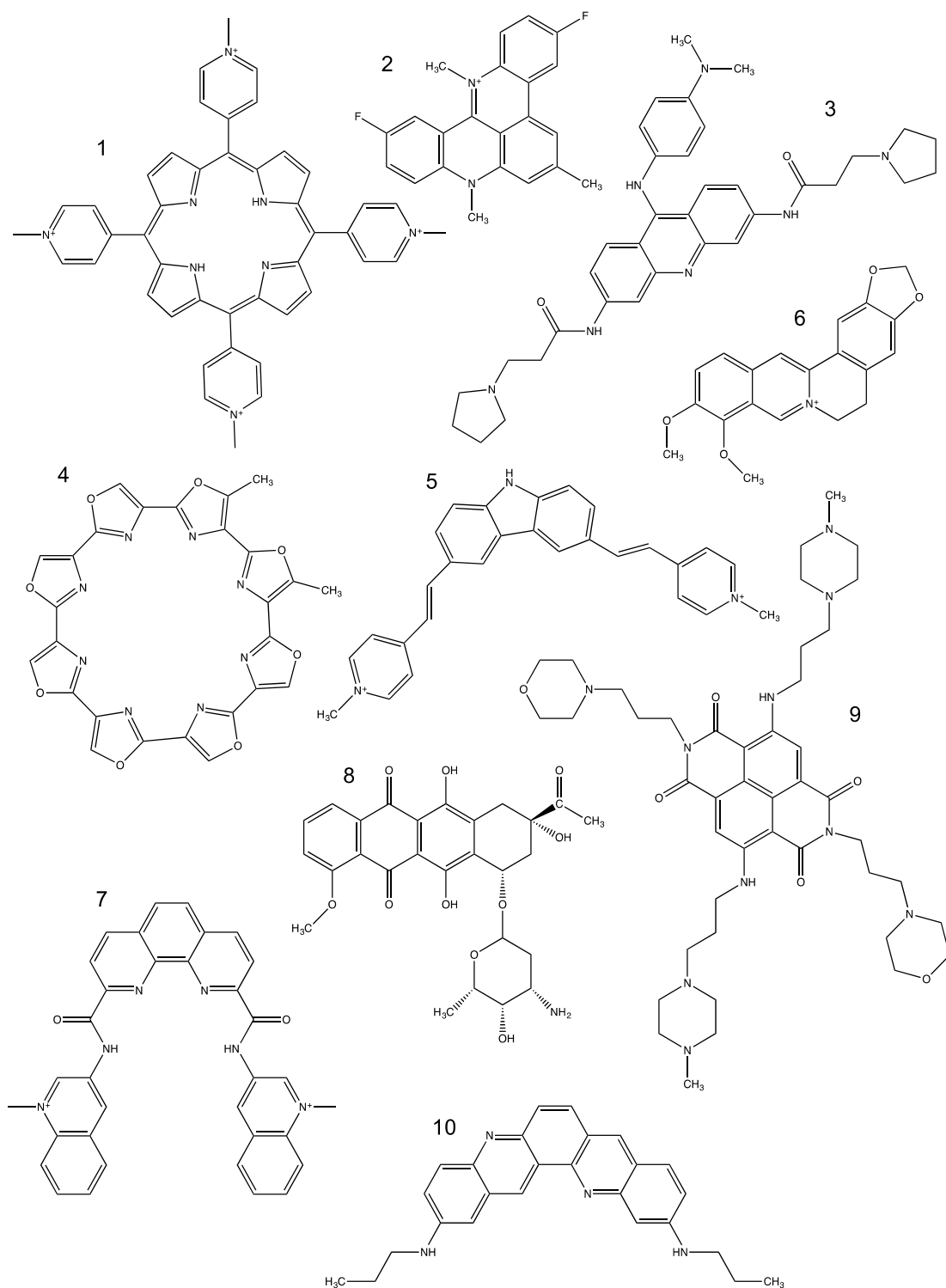
employing monoclonal antibodies as probes<sup>6, 7</sup>. Currently, there are over 200 crystal and NMR-derived structures of quadruplexes present in the protein data bank. These G-quadruplex structures and their ligand complexes have been reviewed elsewhere in detail<sup>8–18</sup>.

Positively charged alkali metal ions are essentially required for the stabilisation of the quadruplexes<sup>19</sup>. The cations are coordinated to carbonyl oxygen (O6) atoms of the guanine, at the centre of the quartet, thereby forming a central ion channel within the quadruplex stem<sup>20–22</sup>. The ions are considered to be an integral part of all quadruplexes<sup>19</sup>. The role of mono- and divalent cations in stabilising quadruplexes has already been reported<sup>19, 23</sup>.  $K^+$  is the strongest stabilising cation and binds in a square antiprismatic bipyramidal geometry<sup>20, 21</sup>. The order of stabilisation by cations is  $K^+ > Rb^+ > Na^+ > Li^+ = Cs^+ > Sr^+ > Ba^{2+} > Ca^{2+} > Mg^{2+}$ <sup>19</sup>. Ions have also been implicated in determining the structure of quadruplexes and multiple cation coordination geometries have also been reported<sup>24</sup>. G-quartets with  $K^+$  ion coordination are larger than their  $Na^+$  counterparts, and their loops are more flexible<sup>25</sup>. The interconversion between antiparallel and parallel quadruplexes can be controlled by  $K^+/Na^+$  balance, even though  $K^+$  favours formation of the antiparallel form<sup>26</sup>. Studies on thrombin-binding aptamer (TBA) have demonstrated that ions can influence structural conformation in addition to stabilisation<sup>27</sup>. Molecular simulations have been crucial in explaining the rapid exchange of cations between the quadruplex and

the solvent. The spontaneous exchange prevents any destabilisation to the quadruplex structure<sup>27</sup>.

A wide range of small molecules have been reported that bind to quadruplexes and a large number of structure–activity relationships were derived (Fig. 2)<sup>13, 14, 28</sup>. The majority of these ligands share a common structural feature of a planar aromatic chromophore. The design is based on the hypothesis that large aromatic chromophores can enhance quadruplex stabilisation and, therefore, interrupt cellular processes such as transcription, translation and replication as well as DNA damage repair, in the absence of mechanisms to unwind them<sup>15</sup>. A major focus is thus to design small molecules with enhanced affinity and selectivity to bind to quadruplexes and inhibit progression of these cellular processes<sup>15</sup>.

BRACO-19 is a 3,6,9-trisubstituted acridine derivative, designed by Neidle et al. to stabilize the telomeric quadruplex DNA structure<sup>29</sup>. It is the first rationally designed telomerase inhibitor<sup>30</sup> that exhibited strong antitumor activity.<sup>31</sup> MMQ1 quinacridine-based ligand was described as an anti-cancer agent in 2007 by Hounsou et al.<sup>32</sup> The G-quadruplex stabilising properties of the MMQ1 series were confirmed by Hou et al.<sup>33</sup> and Gabelica et al.<sup>34</sup> MM41 is a tetra-substituted naphthalene-diimide derivative described first by Micco et al.<sup>35</sup> as a potent stabilizer of human telomeric and gene promoter DNA quadruplexes and can inhibit the growth of human cancer cells in vitro and in vivo. Recently, Ohnmacht et al. reported MM41's significant in vivo anti-tumour activity against the MIA PaCa-2 pancreatic cancer xenograft model.<sup>36</sup> TmPyP4 is a tetracationic porphyrin, which has been extensively studied by Hurley and co-workers.<sup>37</sup> Investigation of this ligand gives interesting results describing high affinity for G-quadruplex binding and anti-telomerase activity, and down-regulating the expression of the oncogenes on promoter region of c-myc.<sup>38</sup> It has been demonstrated that TmPyP4 favours parallel-stranded G-quadruplex and can convert antiparallel to parallel forms of G-quadruplexes.<sup>39</sup> However, this ligand does not bind selectively to G-quadruplex structures. 3,6-bis(1-methyl-4-vinylpyridinium) carbazole diiodide (BMVC) was one of the first selective ligand described as a G-quadruplex binder by Chang et al.<sup>40</sup> It is a diphenylcarbazole derivative, which stabilizes telomeric G-quadruplex and exhibited remarkable inhibition of telomerase ( $IC_{50} = 0.05 \mu M$ ).<sup>41</sup> Phen-DC3 is a phenanthroline analogue described by De Cian et al.<sup>42</sup> and has a perfect geometrical match with a G-quartet resulting in high affinity and selectivity



**Figure 2:** Some ligands that have been studied using computational methods (1. TmPyP4, 2. RHPS4, 3. BRACO-19, 4. Telomestatin, 5. BMVC, 6. Berberine, 7. Phen-DC3, 8. Daunomycin, 9. MM41, 10. MMQ1).

to human telomeric G-quadruplex.<sup>43, 44</sup> Daunomycin has a good affinity to B-DNA structure and has also been crystallised as a trimer with G-quadruplex.<sup>45</sup> Isoquinoline alkaloids including Berberine, which exhibit anti-bacterial activity<sup>46</sup>, have

also been shown to bind to G-quadruplex DNA and inhibit telomere elongation.<sup>47</sup> One of the most interesting G-quadruplex ligands is telomestatin, a natural compound isolated from *Streptomyces annulatus* in 2001 by Shin-ya's et al.<sup>48</sup>.

Telomestatin is a ring-shaped polyheteroaromatic compound, which selectively binds to quadruplex without any affinity towards B-DNA. This compound showed significant activity as a telomerase inhibitor ( $IC_{50}$ -TRAP=5 nM). Telomestatin induced and stabilised G-quadruplex structure even in salt-deficient environment.<sup>49–51</sup> RHPS4 is an N-methylated pentacyclic acridinium, which has shown preference for binding and stabilizing quadruplex isoforms over DNA duplexes.<sup>52</sup> This ligand has shown to decrease telomere length and acts in synergy with the classical anti-cancer agent Taxol.<sup>53, 54</sup>

The ever-increasing use of structural methods for quadruplexes and their ligand complexes provides a strong foundation to carry out molecular modelling and computational approaches. These approaches have been established to predict quadruplex–ligand interactions and rationalisation of biological and biophysical data. This review focuses on the molecular modelling and computational methods to study G-quadruplex–ligand complexes.

## 2 Automated Molecular Docking

Automated molecular docking (AMD) is a computational method that is used to predict ligand binding to a target receptor<sup>55</sup>. It is a widely used approach to screen chemical libraries to identify potential selective ligands that bind to quadruplexes.<sup>14</sup> The results define the putative binding site and highlight atomistic interactions of the docked ligand to its receptor. This has been proven useful in understanding several biochemical processes. Employing such an approach has two-fold advantages. Firstly, it is a rapid method for ligand optimisation and secondly it is relatively cheap when compared with experimental approaches.<sup>56</sup> As a result, AMD is the first step in the study of quadruplex–ligand interactions. Automated docking is carried out in two stages.<sup>57</sup> In the first stage, sampling algorithms are used to predict the position and orientation of the ligand within a defined binding site. The second stage takes into account the chemical interactions between the ligand and the receptor and estimates the binding affinities of the ligand for the receptor. Ideally, the results from AMD should reproduce the experimentally determined binding modes and affinities among the highest generated conformations. Even if this is not the case, AMD has been invaluable in identifying new hit compounds for subsequent experimental validation and lead compound generation.

Virtual screening has become an integral component for hit identification and lead

compound optimisation. It is primarily divided into two main categories, namely structure-based and ligand-based virtual screening. The structure-based approach employs a 3D structure of a target quadruplex, while in the ligand-based screening the structure activity data from a set of known active compounds are employed. A novel form of ‘dynamic docking’ has been developed by Neidle and co-workers<sup>58</sup>. They extracted multiple conformations of the quadruplex (from molecular dynamics simulations). Diverse conformations of the ligand were placed in a grid-like manner and in multiple orientations in the binding site. Multiple starting positions for subsequent MD simulations were generated, allowing large conformational space for both ligand and the quadruplex to be explored.

Interactions of cationic tetraolporphyrin derivatives with antiparallel G-quadruplex (PDB id 143D) were studied by AMD using DOCK6.<sup>59</sup> The first virtual screening of a library of 3000 FDA-approved drugs as *myc* quadruplex stabilising ligands was done using ICM software.<sup>60</sup> The difference in selectivity of a group of naphthalene diimides for telomeric RNA vs DNA quadruplexes was explained by docking using AFFINITY program.<sup>61</sup> Natural products like alkaloids have also been identified by screening a Chinese herbal database of 10,000 compounds with previously generated pharmacophores using Catalyst, Accelrys.<sup>62</sup> A ligand-based screening approach was reported by Chen et al., which centred on a pharmacophore model generated from acridine derivatives.<sup>63</sup> A synthetic substituted indole was identified by screening ~100,000 drug-like compounds.<sup>64</sup> Fonseca B, a natural naphthopyrone compound, was identified by screening 20,000 compound natural product database.<sup>65</sup> AMD, when used in conjunction with NMR experiments, provide in depth description of interactions with the quadruplex.<sup>66</sup> Randazzo et al. screened a small but structurally diverse set of 6000 compound library against parallel  $[d(TG_4T)_4]$  using Autodock. Subsequent NMR screening and ITC measurements of the top hits identified six groove binding ligands that displayed strong quadruplex groove binding properties.<sup>67</sup> A large-scale integrated in silico and in vitro screening platform has been developed by Chaires group and is being used to discover new quadruplex binding scaffolds.<sup>68</sup>

Despite significant advances in method development, a number of pit-falls still exist. The most important being the scoring function to rank various binding poses is still inaccurate. Additionally, the polymorphic nature of quadruplex structures,

resulting from different experimental methods and conditions, can lead to conflicting results. Therefore, the choice of the quadruplex for modelling and structure-based screening should be done with care. Another challenge is the selection of the most favourable pose of the quadruplex for AMD or virtual screening. This is due to the presence of the charged backbone, the presence of the stabilising cations and the conformational flexibility of the loops, all of which can influence the binding of the ligand to the quadruplex.

### 3 Molecular Dynamics Simulations

Molecular dynamics (MD) simulations, in conjunction with automated docking, have been the preferred method to study the dynamic quadruplex–ligand interactions.<sup>69–71</sup> It has been used to supplement other experimental approaches. In MD simulations, the classical Newtonian equation of motion is solved. The force on a simple atomistic system is represented by  $F=ma$ . Covalently bonded parameters like bond length, bond angle, out-of-plane distortions, and torsional terms and non-bonded terms like van der Waals interactions, electrostatic interactions and hydrogen bonds describe contributions towards the force on any individual atom. The position of the atoms is a distribution of velocities and the acceleration is determined by the gradient of the potential energy function. The position vectors are then integrated and their course is traced as a function of time, resulting in a trajectory. The determination of a suitable potential function is extremely important for an MD simulation.<sup>69</sup> The interatomic behaviour of atoms is represented by Lennard–Jones potentials, which describe the attractive and repulsive forces on atoms. The systems are immersed in an explicit solvent, contained in a box that satisfies periodic boundary conditions. MD simulations can be carried out in various ensembles, e.g. canonical (NVT) or isobaric–isothermal (NPT) where the number of atoms, volume, pressure and temperature are kept constant during the simulation.

### 4 Current Force Fields

In molecular mechanics and dynamics, force fields are a set of parameters that describe the molecular behaviour of the system. These parameters can be derived using quantum–mechanical, empirical or experimental methods. All atom contributions like Columbic, polarisation, dispersion and repulsive energies are used to calculate the total energy of a molecule.<sup>72</sup> While there are multiple force fields available, very few are available to accurately simulate quadruplexes. Orozco et al. presented the

parmbsc0 force field.<sup>73</sup> This was a refinement of the AMBER parm99 force field, which focussed on the  $\alpha/\gamma$  torsional terms to resolve various inaccuracies of previous parm94/94 and parm99 force fields.<sup>74</sup> They identified very large  $\alpha/\gamma$  transitions to the *gauche*<sup>+</sup>, *trans* geometry resulting in severe distorted DNA geometry in 50 ns MD trajectories.<sup>75</sup> The new parmbsc0 force field improved the representation of the  $\alpha/\beta$  conformational space, and can recognise and repair larger structural errors while still preserving the flexibility of duplex DNA. Parmbsc0 force field has been successfully used for telomeric quadruplex simulations,<sup>76</sup> reproducing the overall features of the both parallel and antiparallel quadruplexes. A newer and more updated version of the parmbsc0 force field has been introduced. In parmbsc1 force field, modifications to the sugar puckers, the  $\chi$  glyosidic torsion angles, and the  $\epsilon/\zeta$  torsion angles have been made.<sup>77</sup> While the force field can adequately represent guanine stems, the agreement of loop conformations with that observed in experimentally determined structures has still not been achieved due to multiple force field related issues. The Czech group led by Jurecka introduced the OL15 force field.<sup>78</sup> They have made similar improvements to the parmbsc0 force field with refinements in the  $\chi$ <sup>79</sup>,  $\epsilon/\zeta$ <sup>80</sup> and  $\beta$  torsion angles.<sup>78</sup> A subsequent report by Sponer group to study loop conformations in parallel-stranded quadruplexes highlighted the limitations of these force fields to accurately reproduce experimental structure in all these force field versions.<sup>81</sup> However, it is worth mentioning that modifications of the AMBER force field to bsc0, bsc1 and OL15 were developed to improve the accuracy of simulating double-stranded DNA.<sup>74</sup> In spite of their limitations, these force fields have helped in improving the reliability to simulate quadruplexes. Their robustness will be further tested when extended simulations are run, which are closer to biophysical and biological timescales.

The parameters for ligands in the ligand–quadruplex complex can be assigned using the Generalised AMBER force field (GAFF1/2).<sup>82</sup> These force fields can be combined with the bsc1 or OL15 force fields. Several long timescale simulations of ligand–quadruplex complexes have been reported recently. The results from MD simulations corroborated with the experimental data implying bsc0, bsc1 or OL15 force fields are suitable for MD simulations of ligand–quadruplex complexes along with GAFF1/2 method of ligand parameterisation.<sup>83</sup> It is worth mentioning that free energy calculations should be avoided using ligand force fields. The ligand–quadruplex

complexes can be best used to interpret experimental data or make rationalised predictions.

## 5 Long-Range Electrostatic Interactions

In MD simulations, the impact of all contributing factors on energies has to be taken into account. However, it would be ineffective to sum all the non-bonded interactions in an MD simulation setup. Therefore, spherical cutoff distances have been employed to ignore long-range electrostatic interactions beyond a specified distance, without affecting the quality of the simulations.<sup>84</sup> Long-range electrostatics calculations have a significant impact on the stability of the charged quadruplex backbone, cation–quadruplex interactions and the interactions of a charged ligand with the quadruplex. When inaccurately treated, the trajectories of these simulations are unstable; cations are dislodged from the quadruplex stem, resulting in the collapse of the structure.<sup>69</sup> The introduction of the particle-mesh Ewald (PME) truncation method eliminates this fundamental problem and produces stable DNA simulations.<sup>85,86</sup> In spite of being more time consuming, the PME, with a cutoff of 10 Å, is the preferred method of choice to accurately simulate quadruplex–ligand complexes.<sup>70</sup>

## 6 Base stacking interactions and backbone descriptions

The force field parameters also provide descriptions for inter-base interactions. Since there is no quantum mechanical operator or direct experiment to determine atomic charges on the bases, arbitrary charges are assigned by fitting molecular electronic potential charges.<sup>87</sup> These charges are essential to reproduce the electrostatic potential around the molecules.<sup>88</sup> The amino groups of nucleic acid bases are non-planar due to their partial  $sp^3$  pyramidal hybridisation.<sup>87</sup> This can result in the stabilisation of bifurcated H-bonds, close amino group contacts, non-planar G/A base pairs and some other specific interactions.<sup>87,89,90</sup> The force fields assume purely  $sp^2$  amino hybridisation for ring nitrogen atoms. This should be sufficient for most interactions, as primary H-bonds stabilize the  $sp^2$  electronic structure. However, the current force fields do not support out-of-plane H-bonds or amino acceptor interactions. As a result, the force fields are unable to accurately reproduce bifurcated H-bonds or amino group contacts.

Backbone descriptors are very important in simulating nucleic acids for two main reasons. (a) The backbone is highly flexible and negatively charged

and (b) the charge on the backbone changes with the conformation and solvation dynamics. This is not handled by the non-polarizable atom–atom pair-additive force field. The current force fields like parmbsc0<sup>73</sup>, parmbsc1<sup>77</sup> and OL15<sup>78,80</sup> have introduced several correction terms for the DNA backbone profile, which have been able to keep the fluctuations within acceptable limits over the entire course of nano- or microscale simulations.

## 7 Automated Molecular Docking and Dynamic Simulations of Quadruplex–Ligand Complexes: Case Studies

The impetus for automated molecular docking on quadruplex targets first came from the BRACO series of compounds.<sup>29</sup> This series was the first to be rationally designed as quadruplex DNA stabilisers and telomerase inhibitors.<sup>29</sup> Their selectivity and potent telomerase inhibitory activity were subsequently demonstrated experimentally. The NMR structure of the intramolecular human telomeric DNA quadruplex (PDB id 143D) in sodium solution was used to generate a pseudo-ligand-binding site between the T<sub>2</sub>A loop and the top G-quartet. The ligands were manually docked in the generated binding site. The simulations of quadruplex-BRACO-19 complex were then compared to four other ligands, which differed in their functional groups on the three side-chain arms. Later, the crystal structure of quadruplex-BRACO-19 complex<sup>91</sup> validated these predictions and the subsequent MD simulation that followed detailed the stacking interactions between BRACO-19 and the terminal G-quartet.<sup>33,71,83</sup> Molecular structures of other quadruplex–ligand complexes are present and have been simulated including those with daunomycin (PDB id 1L1H),<sup>45</sup> RHPS4 (PDB id: 1NZM),<sup>92</sup> MMQ1 (PDB id: 2JWQ),<sup>32</sup> and TMPyP4 (PDB id: 2A5R).<sup>93</sup> Short MD simulations (of up to 200 ns) are sufficient to stabilise the ligand in the binding site and explore the interactions it makes with the quadruplex.<sup>61,70,71</sup> In order to obtain meaningful binding affinity calculations to study ligand selectivity, it is essential to fully equilibrate the system. Long timescale MD simulations of quadruplexes have shown that quadruplexes are transitionally stabilised on a shorter timescales and are far from equilibrium states.<sup>76,81</sup> A fully converged system requires extended sampling of the quadruplex–ligand systems.<sup>76,81</sup> A good example is the MD simulation used to study the interactions of tetra-substituted naphthalene diimide compound MM41 and quadruplexes formed in the gene

promoter sequences, employing parmbsc0 force field supplemented with  $\chi_{OL4}$  modifications.<sup>36</sup> The results from the MD simulations were able to demonstrate the favourable binding mode of MM41 to the BCL-2 quadruplexes. The interactions that MM41 made were analogous to that made by the ligand with human telomeric quadruplex.<sup>94</sup> This demonstrated that MD simulations could provide useful insights on overall shape, size and nature of ligand skeletons. The substituents can then be optimised and then fine tuned to increase ligand selectivity.

The binding of TMPyP4 to different human telomeric quadruplex topologies has been reported.<sup>95</sup> Automated blind docking was employed to identify the most potent binding sites and MD simulation was performed using parmbsc0 force field. The results showed that the systems were stable over 50 ns simulation time and that TMPyP4 binds most efficiently with a parallel propeller quadruplex topology due to the pronounced effect of the favourable  $\pi$ – $\pi$  stacking interaction of the G-quartets with the core aromatic moieties of the ligand.

MD simulations evaluating the interactions of Telomestatin with human Tel-22 quadruplex have also been reported, using ff99 force field in  $K^+$  solution.<sup>96</sup> Telomestatin was manually docked over the terminal G-quartet, on both 5' and 3' ends of the quadruplex core.<sup>97, 98</sup> A short simulation of just 5 ns highlighted the instability of  $K^+$  ions within the quadruplex core.<sup>96</sup> This indicated that  $K^+$  ions have a significant impact on stability via interactions between telomestatin and several conformations of telomeric quadruplexes. Subsequently, Mulholland and Wu employed the GAFF and parmbsc0 force fields to systematically explore the binding of telomestatin to human telomeric DNA.<sup>99</sup>

More recently, combined experimental and computational methods<sup>100</sup> have been used to investigate the interactions between the isoquinoline alkaloid ligands and a 3+1 hybrid-type human telomeric quadruplex structure (PDB id 2MB3).<sup>47</sup> In this study, Nouredini et al. docked the ligands using the ICM-Pro MolSoft molecular modelling suite. MD simulations were then performed on the quadruplex–ligand complexes selected on the basis of their interaction energy ranking. The results from the simulations indicated that the isoquinolines are stacked on the terminal quartet with the aromatic chromophores maximising the  $\pi$ – $\pi$  stacking interactions. The nitrogen atom in the ligand molecules prefer to position on top of the central electronegative channel that runs along the central axis of

the quadruplex. This behaviour was observed in all the five isoquinolines ligands, indicating that it might be a physical reality.

Distamycin-A, a classic duplex DNA groove binding compound, binds to a tetra-molecular parallel DNA quadruplex (PDB id: 1S45) in the grooves formed between the sugar-phosphate backbone.<sup>101–103</sup> Different models have been proposed for Distamycin-A binding. It can either (a) bind in the groove or (b) stack on a terminal G-quartet or (c) possibly adopting a mixed groove/G-quartet binding mode.<sup>104</sup> The specificity, affinity and binding modes of potential quadruplex groove binders' ligands have been explored by virtual screening docking calculations.<sup>66</sup> The docking results were in sync with the NMR experimental screening and suggested that appropriate groove binders can interact with high affinity to at least this quadruplex.<sup>67</sup>

In addition to DNA quadruplexes, RNA (TERRA) sequences that exhibit quadruplex forming ability have also been explored for ligand binding using computational methods.<sup>61</sup> A comparison of BRACO-19 and naphthalene diimide interactions between DNA and RNA human telomeric quadruplexes has been systematically made by molecular modelling.<sup>61</sup> The 22-mer DNA parallel-stranded human telomeric quadruplex was used (PDB id: 1KF1) in this study and an RNA 22-mer telomeric quadruplex was generated from the DNA structure by direct addition of 2'-OH groups. The differences between DNA and RNA quadruplexes revealed that RNA complexes are more stable and rigid than their DNA counterparts. The stability is due to the interactions made the by 2'-OH groups with the O5' groups and water molecules in the grooves, making the nucleotide backbone more rigid. Additionally, the presence of 2'-OH group in the RNA quadruplex constricts the space that is available to the ligand side chains to interact with the loops, thereby reducing the depth and the width of the UUA loops. There was just one naphthalene diimide ligand that forms a strong complex with a telomeric RNA quadruplex. MD simulations showed that the substitution of one side-chain group ( $-NMe_2$  for an  $-OH$ ) increased RNA quadruplex–ligand affinity by 15-fold. The study suggested that the RNA telomeric quadruplex is less likely to bind ligands with side chains terminating in bulky and/or inflexible functional groups, as are present in the two ligands with reduced affinity to the RNA quadruplex. These differences highlighted that it is possible to rationally design small molecule ligands that have the ability to discriminate between RNA and DNA quadruplexes.<sup>61</sup>

## 8 Enhanced Sampling Methods

### 8.1 Simulated Annealing

Simulated annealing (SA) is a local search algorithm for solving unconstrained and bound-constrained optimisation problems.<sup>105</sup> This method carries out random optimisation based on a Monte Carlo iterative strategy. The name of the algorithm is derived from the physical process of heating and then slowly lowers the temperature to decrease the defects. During the changes of temperature, an SA algorithm searches randomly for the solutions using the metropolis criterion. The process is repetitive while the temperature descends gradually until the global optimisation solution is found. Due to the effectiveness of non-linear combinatorial optimisation methods, simulated annealing algorithms are used in many application fields.<sup>106</sup>

Simulated annealing has been used to determine the optimum ligand orientation and key molecular interactions<sup>107</sup> of telomestatin with a human telomeric intramolecular antiparallel topology quadruplex. The results revealed that two bound telomestatin ligands are most favourable and it can bind via loops or terminal G-quartets to the quadruplex. Furthermore, the quadruplex–telomestatin complex undergoes a conformational rearrangement that leads to significant changes in the relative position, orientation and potential energy of both the telomestatin and the quadruplex.<sup>97</sup>

A simulated annealing docking approach has been used to find the most probable conformation of a novel triazatruxene derivative azatrux binding to the parallel form of human telomeric quadruplex.<sup>108</sup> The method identified ten possible conformations of the ligand bound on the terminal G-quartet face, following which the orientations were evaluated to identify the most appropriate conformation. This protocol was repeated until a stable conformation with optimal G-quartet stacking was identified.<sup>108</sup> A simulated annealing approach has also been used to find reasonable positions for the research compound GQR bound to a quadruplex.<sup>109</sup>

### 8.2 Principal Component Analysis

The principal component analysis (PCA) method is a statistical method, which is used in data reduction. It transforms a number of possibly correlated variables into a smaller number of uncorrelated variables termed principal components. In a simulated system, not all motions are important. PCA separates large amplitude motions from irrelevant fluctuations. The translation and rotation of the structure through the trajectory are

removed by changes in the simulated structure, which is translated to the geometrical centre of the molecule by least-square fit superimposition onto a reference structure.<sup>110</sup> The appropriate configurational space is then constructed using a simple linear transformation in Cartesian coordinate space to generate a  $3N \times 3N$  covariance matrix. The matrices are summed and averaged over the entire trajectory. The resulting matrix is then diagonalised, generating a set of eigenvectors that gives a vectorial description of each component of the motion by indicating the direction of the motion. Each eigenvector describing the motion has a corresponding eigenvalue that represents the energetic contribution of that particular component to the motion. The eigenvalue is the average square displacement of the structure in the direction of the eigenvector. Projection of a trajectory on a particular eigenvector highlights the time-dependent motions that the component performs in the particular vibrational mode. The time average of the projections shows the contribution of components of the atomic vibrations to this mode of concerted motions.<sup>111</sup> The eigenvalues are placed in descending order where the first eigenvectors and eigenvalue describe the largest internal motion of the structure. The eigenvalues decline sharply, highlighting the possibility of separating the dynamics into a small essential space and a relatively large space containing only small atomic fluctuations. On average, only about 5% of eigenvectors are necessary to describe 90% of the total dynamics. Although PCA is a convenient method to visualise trajectories, its limitations should also be taken into consideration when interpreting results.<sup>112</sup> PCA is most suited to analyse trajectories of system that undergo transitional changes instead of trajectories that highlight thermal fluctuations to flexible molecules.<sup>112</sup>

PCA has been used to study the interactions of dimeric quadruplex formed by the sequence d(GGGTGGGTGGGTGGGT)<sup>113</sup> with planar perylene derivative Tel03. The analysis revealed that the flexibility of the loops was reduced by the ligand, leading to more stable Tel03-quadruplex complex. A general conclusion is that using PCA reveals that the dominant motions in the free dimer occur mostly in the loop regions and that the presence of ligand reduces loop motion.<sup>114</sup>

### 8.3 Free Energy Calculations

Accurately calculating free energy of a ligand binding to a quadruplex remains a challenge. There are several methods available, some of which have been used to study quadruplex–ligand interactions.



## 9 Implicit Solvent Methods (MM/PBSA)

A continuum solvent method has been used to assess the structure and free energies from an MD simulation. The electrostatic contribution to the solvation energies is calculated by Poisson–Boltzmann methods.<sup>115</sup> This method combines molecular mechanics, implicit solvation and a hydrophobic contribution, which is represented by a solvent accessible surface area term for energy calculations. The PB model is used to calculate polar contribution while the SASA solvent accessible surface area is used to estimate the non-polar energy component. The molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) relies on calculating the difference in free energy of two states, namely bound (complex) and unbound (ligand and quadruplex).<sup>115</sup> This approach extracts estimates of the free energies from the MD trajectories based on the averages of molecular mechanical energy of the solute with an estimate of solvation energies from a PB continuum model. This method has been used to calculate the free energies of quadruplex–ligand complexes.<sup>116</sup>

Li et al. used MM/PBSA method to study the free energy of binding between Tel03 ligand and an intramolecular parallel-stranded quadruplex.<sup>113</sup> The  $K^+$  ions in the central channel of the quadruplex are treated as an integral part of the quadruplex structure and included in all calculations. The results suggested that Tel03 binds efficiently via end stacking and the interactions formed with the 3' terminal thymine was suggested to be the most favourable binding mode.<sup>113</sup>

MM/PBSA calculations of the binding free energy of tri- and tetra-substituted naphthalene diimides are in excellent accord with the experimental results.<sup>117</sup> They have shown to have greater binding affinity than BRACO-19 for a human telomeric quadruplex.<sup>117</sup> MM-PBSA has also been used to compare the energies for various topologies of the human telomeric

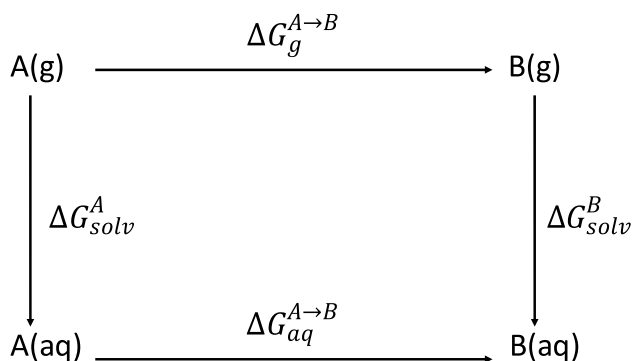
quadruplex.<sup>118</sup> However, variations in energy within a simulation have been a major limitation hindering one from making any more general conclusions from such a study.<sup>118</sup>

## 10 Thermodynamic Integration

In thermodynamic integration (TI), when the two discrete states (e.g. initial state A and final state B) with potential energy  $U_A$  and  $U_B$  are dependent on a different way from the spatial coordinates, then one can compare the difference in free energy between them. The paths between the two states are defined as a thermodynamic cycle, which is used to estimate the energy difference (Fig. 3). Even though TI is computationally expensive, it is commonly used to calculate binding energies and gives accurate results. The binding of TMPyP4 and Phen-DC3 to a human telomeric DNA quadruplex has been characterised using TI.<sup>119</sup> The results revealed that the thermodynamics of ligand–quadruplex binding is complicated and requires a large number of adjustable parameters to fully define the complex.

## 11 Umbrella Sampling

Free energy along a reaction coordinate can be estimated using Umbrella sampling. The method applies a bias potential along a reaction coordinate, which drives the system from state A to B. Any intermediate steps need to be treated as separate systems (bins) and a molecular dynamics simulation run for each of them. This method prefers sampling of many bins with shorter times than fewer bins with longer.<sup>120</sup> The bias potential in the bins depends only on the reaction coordinates and its ability to connect energy minima gives the method its name of umbrella sampling. A weighted histogram analysis (WHAM) can then be used to estimate the free energy of the system.<sup>121</sup>



**Figure 3:** A thermodynamics cycle.

Umbrella sampling has also been effectively used with other modified molecular dynamics methods like Steered molecular dynamics (SMD). A force is applied along reaction coordinates to remove the ligand from its binding site, following which the initial coordinates for Umbrella sampling calculations were generated. These combined methods explained the molecular mechanism and kinetics of how BMVC ligand enters and exits the terminal G-quartet binding site and forms a stable ligand–quadruplex complex.<sup>122</sup>

## 12 Metadynamics

Ligand binding to quadruplexes is a long time-scale event that occurs between micro- to milliseconds, which is extremely difficult to sample using standard molecular dynamics simulations. Metadynamics is an enhanced sampling method that explores long timescale events in a reasonable computational time.<sup>123</sup> The sampling permits the calculations of the free energy landscape of the system. The sampling is accelerated by introducing a history-dependent bias potential on a few degrees of freedom of the system, called collective variables. Parinello et al. applied a modification of this method (funnel metadynamics) to study the binding of berberine to quadruplex and calculating the absolute binding free energy.<sup>124</sup> It took 800 ns to fully explore the free energy landscape of berberine binding to quadruplex DNA and accurately computing the free energy estimate. A comparison of the binding free energy calculated from funnel metadynamics and that from experiments was within 1 kcal/mol.

In another work, Di Leva et al. studied the binding of 3-(benzo[d]thiazol-2-yl)-7-hydroxy-8-((4-(2-hydroxyethyl)piperazin-1-yl)methyl)-2H-chromen-2-one to parallel-stranded tetramolecular TG<sub>4</sub>T quadruplex.<sup>125</sup> The authors propose a hopping binding mechanism, which explores the grooves and the 3' terminal end of the quadruplex. The results are fully able to explain the experimental data at a relatively short timescale.

## 13 Conclusions

There are over 200 quadruplex structures present in the protein data bank. These structures are static time-averaged ensembles that need to be simulated before any meaningful information can be extracted from them. Molecular simulations are invaluable tools to study dynamics of systems such as interactions with solute, solvent and ions, exploring the free energy landscape to identify intermediate, metastable and lowest energy states

that are almost inaccessible to study using any other experimental methods. There is no other method that can provide time-coursed dynamic details of interactions and, therefore, molecular dynamics simulations are the method of choice to study ligands in complex with quadruplexes. The potential of computational methods is realised at best, when they provide explanation and interpretation of experimental data and supplement it with information, which cannot be accessed by experimental methods.

## 14 A Note on Software

Several software is available to carry out automated molecular docking. The most commonly used and freely available are Autodock (<http://autodock.scripps.edu>), Autodock Vina (<http://vina.scripps.edu>) and Dock6 (<http://dock.compbio.ucsf.edu>). While Autodock comes with a graphical user interface (GUI), Dock6 files can be prepared using Chimera molecular visualisation tool (<http://cgl.ucsf.edu/chimera>). The compatibility of Chimera and Dock6 with ZINC database (<http://zinc.dockng.org>) makes it an excellent streamlined suite for virtual screening. In addition to these, several proprietary softwares such as MolSoft-Pro (<http://www.molsoft.com>), Schrodinger suite (<http://www.schrodinger.com>) and Discovery studio (<http://www.accelrys.com>) have also been used for docking ligands to quadruplexes.

AMBER software (<http://www.ambermd.org>) has been the engine of choice to run molecular dynamics simulations. This might be due to (a) the internal implementation of the Barcelona (parmBSc) and its modified variant force fields used to parameterise quadruplex systems, (b) the ease of parameterising ligands using ANTECHAMBER code employing the GAFF/GAFF2 force fields and (c) the fast and efficient PMEMD code that runs on the graphical processing units (GPUs), having demonstrated that long timescale simulations provide meaningful results for quadruplex systems. Recently, latest force fields have been made available and can also be effectively ported for use in GROMACS (<http://www.gromacs.org>), DESMOND (<http://www.schrodinger.com/desmond>) and ACEMD (<http://www.acellera.com>) MD software. CHARMM force fields and molecular simulation program (<http://www.charmm.org>) have also been used to simulate quadruplexes. For scientific groups that do not have access to computing infrastructure, Acellera Inc. (<http://www.acellera.com>) are now providing cloud-based services to run molecular dynamics simulations.

## Open Access

This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Received: 1 May 2018 Accepted: 28 June 2018  
Published online: 21 July 2018

## References

- Gellert M, Lipsett MN, Davies DR (1962) Helix formation by guanylic acid. *Proc Natl Acad Sci USA* 48:2013–2018
- Chambers VS, Marsico G, Boutell JM, Di Antonio M, Smith GP, Balasubramanian S (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat Biotechnol* 33:877–881
- Huppert JL, Balasubramanian S (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res* 33:2908–2916
- Todd AK, Johnston M, Neidle S (2005) Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res* 33:2901–2907
- Hansel-Hertsch R, Beraldi D, Lensing SV, Marsico G, Zyner K, Parry A, Di Antonio M, Pike J, Kimura H, Narita M, Tannahill D, Balasubramanian S (2016) G-quadruplex structures mark human regulatory chromatin. *Nat Genet* 48:1267–1272
- Biffi G, Tannahill D, McCafferty J, Balasubramanian S (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem* 5:182–186
- Henderson A, Wu Y, Huang YC, Chavez EA, Platt J, Johnson FB, Brosh RM Jr., Sen D, Lansdorp PM (2014) Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res* 42:860–869
- Adrian M, Heddi B, Phan AT (2012) NMR spectroscopy of G-quadruplexes. *Methods* 57:11–24
- Cree SL, Kennedy MA (2014) Relevance of G-quadruplex structures to pharmacogenetics. *Front Pharmacol* 5:160
- Islam MK, Jackson PJ, Rahman KM, Thurston DE (2016) Recent advances in targeting the telomeric G-quadruplex DNA sequence with small molecules as a strategy for anticancer therapies. *Future Med Chem* 8:1259–1290
- Monchaud D, Teulade-Fichou MP (2008) A hitchhiker's guide to G-quadruplex ligands. *Org Biomol Chem* 6:627–636
- Muller S, Rodriguez R (2014) G-quadruplex interacting small molecules and drugs: from bench toward bedside. *Expert Rev Clin Pharmacol* 7:663–679
- Neidle S (2009) The structures of quadruplex nucleic acids and their drug complexes. *Curr Opin Struct Biol* 19:239–250
- Neidle S (2016) Quadruplex nucleic acids as novel therapeutic targets. *J Med Chem* 59:5987–6011
- Neidle S (2017) Quadruplex nucleic acids as targets for anticancer therapeutics. *Nat Rev Chem* 1:0041
- Ohnmacht SA, Neidle S (2014) Small-molecule quadruplex-targeted drug discovery. *Bioorg Med Chem Lett* 24:2602–2612
- Sekaran V, Soares J, Jarstfer MB (2014) Telomere maintenance as a target for drug discovery. *J Med Chem* 57:521–538
- Sissi C, Palumbo M (2014) Telomeric G-quadruplex architecture and interactions with potential drugs. *Curr Pharm Des* 20:6489–6509
- Hud NV, Plavec J (2006) The role of cations in determining quadruplex structure and stability. In: Neidle S, Balarubramanian S (eds) *Quadruplex nucleic acids*, chap 4. Royal Society of Chemistry, UK, pp 100–130. ISBN 978-0-85404-374-3
- Haider S, Parkinson GN, Neidle S (2002) Crystal structure of the potassium form of an *Oxytricha nova* G-quadruplex. *J Mol Biol* 320:189–200
- Haider SM, Parkinson GN, Neidle S (2003) Structure of a G-quadruplex–ligand complex. *J Mol Biol* 326:117–125
- Parkinson GN, Lee MP, Neidle S (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* 417:876–880
- Largy E, Marchand A, Amrane S, Gabelica V, Mergny JL (2016) Quadruplex turncoats: cation-dependent folding and stability of quadruplex–DNA double switches. *J Am Chem Soc* 138:2780–2792
- Hud NV, Smith FW, Anet FA, Feigon J (1996) The selectivity for  $K^+$  versus  $Na^+$  in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by  $^1H$  NMR. *Biochemistry* 35:15383–15390
- Schultze P, Macaya RF, Feigon J (1994) Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GTTGGTGTGGTTGG). *J Mol Biol* 235:1532–1547
- Strahan GD, Keniry MA, Shafer RH (1998) NMR structure refinement and dynamics of the  $K^+$ -[d(G3T4G3)]<sub>2</sub> quadruplex via particle mesh Ewald molecular dynamics simulations. *Biophys J* 75:968–981
- Reshetnikov RV, Sponer J, Rassokhina OI, Kopylov AM, Tsvetkov PO, Makarov AA, Golovin AV (2011) Cation binding to 15-TBA quadruplex DNA is a multiple-pathway cation-dependent process. *Nucleic Acids Res* 39:9789–9802

28. Neidle S (2015) A Personal history of quadruplex-small molecule targeting. *Chem Rec* (New York, N.Y.) 15:691–710
29. Read M, Harrison RJ, Romagnoli B, Taniou FA, Gowan SH, Reszka AP, Wilson WD, Kelland LR, Neidle S (2001) Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proc Natl Acad Sci* 98:4844–4849
30. Harrison RJ, Cuesta J, Chessari G, Read MA, Basra SK, Reszka AP, Morrell J, Gowan SM, Incles CM, Taniou FA, Wilson WD, Kelland LR, Neidle S (2003) Trisubstituted acridine derivatives as potent and selective telomerase inhibitors. *J Med Chem* 46:4463–4476
31. Gowan SM, Harrison JR, Patterson L, Valenti M, Read MA, Neidle S, Kelland LR (2002) A G-quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity. *Mol Pharmacol* 61:1154–1162
32. Hounsou C, Guittat L, Monchaud D, Jourdan M, Saettel N, Mergny J-L, Teulade-Fichou M-P (2007) G-quadruplex recognition by quinacridines: a SAR, NMR, and biological study. *ChemMedChem* 2:655–666
33. Hou J-Q, Chen S-B, Tan J-H, Ou T-M, Luo H-B, Li D, Xu J, Gu L-Q, Huang Z-S (2010) New insights into the structures of ligand-quadruplex complexes from molecular dynamics simulations. *J Phys Chem B* 114:15301–15310
34. Gabelica V, Baker ES, Teulade-Fichou M-P, De Pauw E, Bowers MT (2007) Stabilization and structure of telomeric and c-myc region intramolecular G-quadruplexes: the role of central cations and small planar ligands. *J Am Chem Soc* 129:895–904
35. Micco M, Collie GW, Dale AG, Ohnmacht SA, Pazitna I, Gunaratnam M, Reszka AP, Neidle S (2013) Structure-based design and evaluation of naphthalene diimide G-quadruplex ligands as telomere targeting agents in pancreatic cancer cells. *J Med Chem* 56:2959–2974
36. Ohnmacht SA, Marchetti C, Gunaratnam M, Besser RJ, Haider SM, Di Vita G, Lowe HL, Mellinas-Gomez M, Diocou S, Robson M, Šponer J, Islam B, Barbara Pedley R, Hartley JA, Neidle S (2015) A G-quadruplex-binding compound showing anti-tumour activity in an in vivo model for pancreatic cancer. *Sci Rep* 5:11385
37. Han H, Langley DR, Rangan A, Hurley LH (2001) Selective interactions of cationic porphyrins with G-quadruplex structures. *J Am Chem Soc* 123:8902–8913
38. Nasiri HR, Bell NM, McLuckie KIE, Husby J, Abell C, Neidle S, Balasubramanian S (2014) Targeting a c-MYC G-quadruplex DNA with a fragment library. *Chem Commun* (Cambridge, England) 50:1704–1707
39. Zhang H, Xiao X, Wang P, Pang S, Qu F, Ai X, Zhang J (2009) Conformational conversion of DNA G-quadruplex induced by a cationic porphyrin. *Spectrochim Acta Part A Mol Biomol Spectrosc* 74:243–247
40. Chang C-C, Wu J-Y, Chien C-W, Wu W-S, Liu H, Kang C-C, Yu L-J, Chang T-C (2003) A fluorescent carbazole derivative: high sensitivity for quadruplex DNA. *Anal Chem* 75:6177–6183
41. Chang C-C, Kuo IC, Ling IF, Chen C-T, Chen H-C, Lou P-J, Lin J-J, Chang T-C (2004) Detection of quadruplex DNA structures in human telomeres by a fluorescent carbazole derivative. *Anal Chem* 76:4490–4494
42. De Cian A, DeLemos E, Mergny J-L, Teulade-Fichou M-P, Monchaud D (2007) Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J Am Chem Soc* 129:1856–1857
43. Chung WJ, Heddi B, Hamon F, Teulade-Fichou M-P, Phan AT (2014) Solution structure of a G-quadruplex bound to the bisquinolinium compound phen-DC3. *Angew Chem Int Ed* 53:999–1002
44. Piazza A, Boulé J-B, Lopes J, Mingo K, Largy E, Teulade-Fichou M-P, Nicolas A (2010) Genetic instability triggered by G-quadruplex interacting Phen-DC compounds in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 38:4337–4348
45. Clark GR, Pytel PD, Squire CJ, Neidle S (2003) Structure of the first parallel DNA quadruplex-drug complex. *J Am Chem Soc* 125:4066–4067
46. Ghosh AK, Bhattacharyya FK, Ghosh DK (1985) *Leishmania donovani*: amastigote inhibition and mode of action of berberine. *Exp Parasitol* 60:404–413
47. Noureini SK, Esmaili H, Abachi F, Khiali S, Islam B, Kuta M, Saboury AA, Hoffmann M, Sponer J, Parkinson G, Haider S (2017) Selectivity of major isoquinoline alkaloids from *Chelidonium majus* towards telomeric G-quadruplex: a study using a transition-FRET (t-FRET) assay. *Biochim Biophys Acta* 1861:2020–2030
48. Shin-ya K, Wierzbka K, Matsuo K, Ohtani T, Yamada Y, Furihata K, Hayakawa Y, Seto H (2001) Telomestatin, a novel telomerase inhibitor from *Streptomyces anulatus*. *J Am Chem Soc* 123:1262–1263
49. Chung WJ, Heddi B, Tera M, Iida K, Nagasawa K, Phan AT (2013) Solution structure of an intramolecular (3 + 1) human telomeric G-quadruplex bound to a telomestatin derivative. *J Am Chem Soc* 135:13495–13501
50. Shirude PS, Gillies ER, Ladame S, Godde F, Shin-Ya K, Huc I, Balasubramanian S (2007) Macrocyclic and helical oligoamides as a new class of G-quadruplex ligands. *J Am Chem Soc* 129:11890–11891
51. Tera M, Ishizuka H, Takagi M, Suganuma M, Shin-ya K, Nagasawa K (2008) Macrocyclic hexaoxazoles as sequence- and mode-selective G-quadruplex binders. *Angew Chem Int Ed* 47:5557–5560
52. Stanslas J, Hagan DJ, Ellis MJ, Turner C, Carmichael J, Ward W, Hammonds TR, Stevens MFG (2000) Antitumor polycyclic acridines. 7. Synthesis and biological properties of dna affinic tetra- and pentacyclic acridines. *J Med Chem* 43:1563–1572
53. Cookson JC, Dai F, Smith V, Heald RA, Laughton CA, Stevens MFG, Burger AM (2005) Pharmacodynamics of the G-quadruplex-stabilizing telomerase inhibitor 3,11-difluoro-6,8,13-trimethyl-8H-quinolo[4,3,2-kl]

- acridinium methosulfate (RHPS4) in vitro: activity in human tumor cells correlates with telomere length and can be enhanced, or antagonized, with cytotoxic agents. *Mol Pharmacol* 68:1551–1558
54. Salvati E, Leonetti C, Rizzo A, Scarsella M, Mottolese M, Galati R, Sperduti I, Stevens MFG, D'Incalci M, Blasco M, Chiorino G, Bauwens S, Horard B, Gilson E, Stoppacciaro A, Zupi G, Biroccio A (2007) Telomere damage induced by the G-quadruplex ligand RHPS4 has an antitumor effect. *J Clin Invest* 117:3236–3247
  55. Magalhães CSD, Barbosa HJC, Dardenne LE (2004) A genetic algorithm for the ligand-protein docking problem. *Genet Mol Biol* 27:605–610
  56. Hughes J, Rees S, Kalindjian S, Philpott K (2011) Principles of early drug discovery: principles of early drug discovery. *Br J Pharmacol* 162:1239–1249
  57. Meng X-Y, Zhang H-X, Mezei M, Cui M (2011) Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des* 7:146–157
  58. Husby J, Todd AK, Platts JA, Neidle S (2013) Small-molecule G-quadruplex interactions: systematic exploration of conformational space using multiple molecular dynamics. *Biopolymers* 99:989–1005
  59. Ishikawa Y, Higashi E, Morioka H (2007) Molecular docking of porphyrins with cationic limbs on intramolecular G-quadruplex. *Nucleic Acids Symp Ser (Oxf)* 51:247–248
  60. Chan DS, Yang H, Kwan MH, Cheng Z, Lee P, Bai LP, Jiang ZH, Wong CY, Fong WF, Leung CH, Ma DL (2011) Structure-based optimization of FDA-approved drug methylene blue as a c-myc G-quadruplex DNA stabilizer. *Biochimie* 93:1055–1064
  61. Collie GW, Haider SM, Neidle S, Parkinson GN (2010) A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Res* 38:5569–5580
  62. Li Q, Xiang J, Li X, Chen L, Xu X, Tang Y, Zhou Q, Li L, Zhang H, Sun H, Guan A, Yang Q, Yang S, Xu G (2009) Stabilizing parallel G-quadruplex DNA by a new class of ligands: two non-planar alkaloids through interaction in lateral grooves. *Biochimie* 91:811–819
  63. Chen SB, Tan JH, Ou TM, Huang SL, An LK, Luo HB, Li D, Gu LQ, Huang ZS (2011) Pharmacophore-based discovery of triaryl-substituted imidazole as new telomeric G-quadruplex ligand. *Bioorg Med Chem Lett* 21:1004–1009
  64. Ma DL, Lai TS, Chan FY, Chung WH, Abagyan R, Leung YC, Wong KY (2008) Discovery of a drug-like G-quadruplex binding ligand by high-throughput docking. *ChemMedChem* 3:881–884
  65. Lee HM, Chan DS, Yang F, Lam HY, Yan SC, Che CM, Ma DL, Leung CH (2010) Identification of natural product fonsecin B as a stabilizing ligand of c-myc G-quadruplex DNA by high-throughput virtual screening. *Chem Commun (Camb)* 46:4680–4682
  66. Cosconati S, Marinelli L, Trotta R, Virno A, Mayol L, Novellino E, Olson AJ, Randazzo A (2009) Tandem application of virtual screening and NMR experiments in the discovery of brand new DNA quadruplex groove binders. *J Am Chem Soc* 131:16336–16337
  67. Trotta R, De Tito S, Lauri I, La Pietra V, Marinelli L, Cosconati S, Martino L, Conte MR, Mayol L, Novellino E, Randazzo A (2011) A more detailed picture of the interactions between virtual screening-derived hits and the DNA G-quadruplex: NMR, molecular modelling and ITC studies. *Biochimie* 93:1280–1287
  68. Holt PA, Buscaglia R, Trent JO, Chaires JB (2011) A discovery funnel for nucleic acid binding drug candidates. *Drug Dev Res* 72:178–186
  69. Haider S, Neidle S (2010) Molecular modeling and simulation of G-quadruplexes and quadruplex-ligand complexes. *Methods Mol Biol* 608:17–37
  70. Haider S, Parkinson GN, Neidle S (2008) Molecular dynamics and principal components analysis of human telomeric quadruplex multimers. *Biophys J* 95:296–311
  71. Haider SM, Neidle S (2009) A molecular model for drug binding to tandem repeats of telomeric G-quadruplexes. *Biochem Soc Trans* 37:583–588
  72. Maurer P, Laio A, Hugosson HW, Colombo MC, Rothlisberger U (2007) Automated parametrization of biomolecular force fields from quantum mechanics/molecular mechanics (QM/MM) simulations through force matching. *J Chem Theory Comput* 3:628–639
  73. Perez A, Luque FJ, Orozco M (2007) Dynamics of B-DNA on the microsecond time scale. *J Am Chem Soc* 129:14739–14745
  74. Galindo-Murillo R, Robertson JC, Zgarbova M, Sponer J, Otyepka M, Jurecka P, Cheatham TE 3rd (2016) Assessing the current state of amber force field modifications for DNA. *J Chem Theory Comput* 12:4114–4127
  75. Várnai P, Zakrzewska K (2004) DNA and its counterions: a molecular dynamics study. *Nucleic Acids Res* 32:4269–4280
  76. Islam B, Sgobba M, Laughton C, Orozco M, Sponer J, Neidle S, Haider S (2013) Conformational dynamics of the human propeller telomeric DNA quadruplex on a microsecond time scale. *Nucleic Acids Res* 41:2723–2735
  77. Ivani I, Dans PD, Noy A, Perez A, Faustino I, Hospital A, Walther J, Andrio P, Goni R, Balaceanu A, Portella G, Battistini F, Gelpi JL, Gonzalez C, Vendruscolo M, Laughton CA, Harris SA, Case DA, Orozco M (2016) Parmbsc1: a refined force field for DNA simulations. *Nat Methods* 13:55–58
  78. Zgarbova M, Sponer J, Otyepka M, Cheatham TE 3rd, Galindo-Murillo R, Jurecka P (2015) Refinement of the sugar-phosphate backbone torsion beta for AMBER force fields improves the description of Z- and B-DNA. *J Chem Theory Comput* 11:5723–5736

79. Krepl M, Zgarbova M, Stadlbauer P, Otyepka M, Banas P, Koca J, Cheatham TE 3rd, Jurecka P, Sponer J (2012) Reference simulations of noncanonical nucleic acids with different chi variants of the AMBER force field: quadruplex DNA, quadruplex RNA and Z-DNA. *J Chem Theory Comput* 8:2506–2520
80. Zgarbova M, Luque FJ, Sponer J, Cheatham TE 3rd, Otyepka M, Jurecka P (2013) Toward improved description of DNA backbone: revisiting epsilon and zeta torsion force field parameters. *J Chem Theory Comput* 9:2339–2354
81. Islam B, Stadlbauer P, Gil-Ley A, Perez-Hernandez G, Haider S, Neidle S, Bussi G, Banas P, Otyepka M, Sponer J (2017) Exploring the dynamics of propeller loops in human telomeric DNA quadruplexes using atomistic simulations. *J Chem Theory Comput* 13:2458–2480
82. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25:1157–1174
83. Machireddy B, Kalra G, Jonnalagadda S, Ramanujachary K, Wu C (2017) Probing the binding pathway of BRACO19 to a parallel-stranded human telomeric G-quadruplex using molecular dynamics binding simulation with AMBER DNA OL15 and ligand GAFF2 force fields. *J Chem Inf Model* 57:2846–2864
84. Norberg J, Nilsson L (2000) On the truncation of long-range electrostatic interactions in DNA. *Biophys J* 79:1537–1553
85. Cheatham TEI, Miller JL, Fox T, Darden TA, Kollman PA (1995) Molecular dynamics simulations on solvated biomolecular systems: the particle mesh ewald method leads to stable trajectories of DNA, RNA, and proteins. *J Am Chem Soc* 117:4193–4194
86. York DM, Darden TA, Pedersen LG (1993) The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods. *J Chem Phys* 99:8345
87. Šponer J, Leszczynski J, Hobza P (2001) Electronic properties, hydrogen bonding, stacking, and cation binding of DNA and RNA bases. *Biopolymers* 61:3–31
88. Bayly CI, Cieplak P, Cornell W, Kollman PA (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* 97:10269–10280
89. Islam B, Stadlbauer P, Krepl M, Koca J, Neidle S, Haider S, Sponer J (2015) Extended molecular dynamics of a c-kit promoter quadruplex. *Nucleic Acids Res* 43:8673–8693
90. Šponer J, Bussi G, Stadlbauer P, Kúhrová P, Banáš P, Islam B, Haider S, Neidle S, Otyepka M (2017) Folding of guanine quadruplex molecules—funnel-like mechanism or kinetic partitioning? An overview from MD simulation studies. *Biochimica et Biophysica Acta (BBA) Gen Subj* 1861:1246–1263
91. Campbell NH, Parkinson GN, Reszka AP, Neidle S (2008) Structural basis of DNA quadruplex recognition by an acridine drug. *J Am Chem Soc* 130:6722–6724
92. Gavathiotis E, Heald RA, Stevens MFG, Searle MS (2003) Drug recognition and stabilisation of the parallel-stranded DNA quadruplex d(TTAGGGT)<sub>4</sub> containing the human telomeric repeat. *J Mol Biol* 334:25–36
93. Phan AT, Kuryavyy V, Gaw HY, Patel DJ (2005) Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. *Nat Chem Biol* 1:167–173
94. Prato G, Silvent S, Saka S, Lamberto M, Kosenkov D (2015) Thermodynamics of binding of di- and tetrasubstituted naphthalene diimide ligands to DNA G-quadruplex. *J Phys Chem B* 119:3335–3347
95. Ali A, Bansal M, Bhattacharya S (2015) Ligand 5,10,15,20-tetra(*N*-methyl-4-pyridyl)porphine (TMPyP4) prefers the parallel propeller-type human telomeric G-quadruplex DNA over its other polymorphs. *J Phys Chem B* 119:5–14
96. Agrawal S, Ojha RP, Maiti S (2008) Energetics of the human Tel-22 quadruplex-telomestatin interaction: a molecular dynamics study. *J Phys Chem B* 112:6828–6836
97. Kim M-Y, Vankayalapati H, Shin-ya K, Wierzbka K, Hurley LH (2002) Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. *J Am Chem Soc* 124:2098–2099
98. Rezler EM, Seenisamy J, Bashyam S, Kim M-Y, White E, Wilson WD, Hurley LH (2005) Telomestatin and diseleno saphyrin bind selectively to two different forms of the human telomeric G-quadruplex structure. *J Am Chem Soc* 127:9439–9447
99. Mulholland K, Wu C (2016) Binding of telomestatin to a telomeric G-quadruplex DNA probed by all-atom molecular dynamics simulations with explicit solvent. *J Chem Inf Model* 56:2093–2102
100. Arora A, Balasubramanian C, Kumar N, Agrawal S, Ojha RP, Maiti S (2008) Binding of berberine to human telomeric quadruplex—spectroscopic, calorimetric and molecular modeling studies. *FEBS J* 275:3971–3983
101. Cocco MJ, Hanakahi LA, Huber MD, Maizels N (2003) Specific interactions of distamycin with G-quadruplex DNA. *Nucleic Acids Res* 31:2944–2951
102. Pagano B, Virno A, Mattia CA, Mayol L, Randazzo A, Giancola C (2008) Targeting DNA quadruplexes with distamycin A and its derivatives: an ITC and NMR study. *Biochimie* 90:1224–1232
103. Randazzo A, Galeone A, Mayol L (2001) <sup>1</sup>H-NMR study of the interaction of distamycin A and netropsin with the parallel stranded tetraplex [d(TGGGGT)]<sub>4</sub>. *Chem Commun* 11:1030–1031
104. Martino L, Virno A, Pagano B, Virgilio A, Di Micco S, Galeone A, Giancola C, Bifulco G, Mayol L, Randazzo A (2007) Structural and thermodynamic studies of

- the interaction of distamycin A with the parallel quadruplex structure [d(TGGGGT)]<sub>4</sub>. *J Am Chem Soc* 129:16048–16056
105. van Laarhoven PJ (1987) In: Laarhoven PJV, Aarts EH (eds), vol 37. Springer, Netherlands, pp 7–15
  106. Aspbury AS, Gibson RM (2004) Long-range visibility of greater sage grouse leks: a GIS-based analysis. *Anim Behav* 67:1127–1132
  107. Duan W, Rangan A, Vankayalapati H, Kim M-Y, Zeng Q, Sun D, Han H, Fedoroff OY, Nishioka D, Rha SY, Izbicka E, Hoff DDV, Hurley LH (2001) Design and synthesis of fluoroquinophenoxazines that interact with human telomeric G-quadruplexes and their biological effects. *J Am Assoc Cancer Res* 1:103–120
  108. Petraccone L, Fotticchia I, Cummaro A, Pagano B, Ginnari-Satriani L, Haider S, Randazzo A, Novellino E, Neidle S, Giancola C (2011) The triazatruxene derivative azatrux binds to the parallel form of the human telomeric G-quadruplex under molecular crowding conditions: biophysical and molecular modeling studies. *Biochimie* 93:1318–1327
  109. Zhang L, Er JC, Ghosh KK, Chung WJ, Yoo J, Xu W, Zhao W, Phan AT, Chang Y-T (2014) Discovery of a structural-element specific g-quadruplex “light-up” probe. *Sci Rep* 4:3776
  110. Amadei A, Linssen AB, de Groot BL, van Aalten DM, Berendsen HJ (1996) An efficient method for sampling the essential subspace of proteins. *J Biomol Struct Dyn* 13:615–625
  111. van Aalten DM, Amadei A, Linssen AB, Eijssink VG, Vriend G, Berendsen HJ (1995) The essential dynamics of thermolysin: confirmation of the hinge-bending motion and comparison of simulations in vacuum and water. *Proteins* 22:45–54
  112. Réblová K, Štřelcová Z, Kulhánek P, Bešševová I, Mathews DH, Van Nostrand K, Yildirim I, Turner DH, Šponer J (2010) An RNA molecular switch: intrinsic flexibility of 23S rRNA helices 40 and 68 5'-UAA/5'-GAN internal loops studied by molecular dynamics methods. *J Chem Theory Comput* 6:910–929
  113. Li M-H, Luo Q, Xue X-G, Li Z-S (2011) Molecular dynamics studies of the 3D structure and planar ligand binding of a quadruplex dimer. *J Mol Model* 17:515–526
  114. Bhat J, Chatterjee S (2016) Skeleton selectivity in complexation of chelerythrine and chelerythrine-like natural plant alkaloids with the G-quadruplex formed at the promoter of c-MYC oncogene: in silico exploration. *RSC Adv* 6:36667–36680
  115. Fogolari F, Brigo A, Molinari H (2003) Protocol for MM/PBSA molecular dynamics simulations of proteins. *Biophys J* 85:159–166
  116. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model* 51:69–82
  117. Cuenca F, Greciano O, Gunaratnam M, Haider S, Munnur D, Nanjunda R, Wilson WD, Neidle S (2008) Tri- and tetra-substituted naphthalene diimides as potent G-quadruplex ligands. *Bioorg Med Chem Lett* 18:1668–1673
  118. Islam B, Stadlbauer P, Neidle S, Haider S, Sponer J (2016) Can we execute reliable MM-PBSA free energy computations of relative stabilities of different guanine quadruplex folds? *J Phys Chem B* 120:2899–2912
  119. Bončina M, Podlipnik Č, Piantanida I, Eilmes J, Teulade-Fichou M-P, Vesnaver G, Lah J (2015) Thermodynamic fingerprints of ligand binding to human telomeric G-quadruplexes. *Nucleic Acids Res* 43:10376–10386
  120. Beutler TC, Gunsteren WFV (1994) The computation of a potential of mean force: choice of the biasing potential in the umbrella sampling technique. *J Chem Phys* 100:1492–1497
  121. Wang Y-K, Wei D-Q, Gu R-X, Fan H-M, Ulmschneider J (2013) Applications of rare event dynamics on the free energy calculations for membrane protein systems. *Can J Chem* 91:769–774
  122. Zhou J-K, Yang D-Y, Sheu S-Y (2015) The molecular mechanism of ligand unbinding from the human telomeric G-quadruplex by steered molecular dynamics and umbrella sampling simulations. *Phys Chem Chem Phys* 17:12857–12869
  123. Laio A, Parrinello M (2002) Escaping free-energy minima. *Proc Natl Acad Sci USA* 99:12562–12566
  124. Moraca F, Amato J, Ortuso F, Artese A, Pagano B, Novellino E, Alcaro S, Parrinello M, Limongelli V (2017) Ligand binding to telomeric G-quadruplex DNA investigated by funnel-metadynamics simulations. *Proc Natl Acad Sci USA* 114:E2136–E2145
  125. Di Leva FS, Novellino E, Cavalli A, Parrinello M, Limongelli V (2014) Mechanistic insight into ligand binding to G-quadruplex DNA. *Nucleic Acids Res* 42:5447–5455



**Dr Shozeb Haider** is a Senior Lecturer in Computational Chemistry and the Drug Discovery lead at UCL School of Pharmacy, London. His research focuses on using Computational methods to study

G-quadruplex complexes, protein-nucleic acid and protein-protein interactions in Cancer pathways. He is also a Fellow of the Royal Society of Chemistry.