Calculation of Hydrodynamic Properties for G-Quadruplex Nucleic Acid Structures from in silico Bead Models

ARTICLE *in* TOPICS IN CURRENT CHEMISTRY · AUGUST 2012

Impact Factor: 4.46 · DOI: 10.1007/128_2012_351 · Source: PubMed

5 AUTHORS, INCLUDING:

Huy T Le
University of Louisville
5 PUBLICATIONS **27** CITATIONS
SEE PROFILE

William Dean
University of Louisville
122 PUBLICATIONS **2,486** CITATIONS
SEE PROFILE

Jonathan B Chaires
University of Louisville
195 PUBLICATIONS **11,307** CITATIONS
SEE PROFILE

John O Trent
University of Louisville
131 PUBLICATIONS **5,374** CITATIONS
SEE PROFILE

Published in final edited form as:
Top Curr Chem. 2013 ; 330: 179–210. doi:10.1007/128_2012_351.

Calculation of Hydrodynamic Properties for G-Quadruplex Nucleic Acid Structures from *in silico* Bead Models

Huy T. Le, **Robert Buscaglia**, **William L. Dean**, **Jonathan B. Chaires**, and **John O. Trent**

Abstract

Nucleic acids enriched in guanine bases can adopt unique quadruple helical tertiary structures known as G-quadruplexes. G-quadruplexes have emerged as attractive drug targets as many G-quadruplex-forming sequences have been discovered in functionally critical sites within the human genome, including the telomere, oncogene promoters, and mRNA processing sites. A single G-quadruplex-forming sequence can adopt one of multiple folding topologies often resulting in a lack of a single definitive atomic-level resolution structure for many of these sequences and a major challenge to the discovery of G-quadruplex-selective small molecule drugs. Low-resolution techniques employed to study G-quadruplex structures (e.g. CD spectroscopy) are often unable to discern between G-quadruplex structural ensembles while high-resolution techniques (e.g. NMR spectroscopy) can be overwhelmed by a highly polymorphic system. Hydrodynamic bead modeling is an approach to studying G-quadruplex structures that could bridge the gap between low-resolution techniques and high-resolution molecular models. Here, we present a discussion of hydrodynamic bead modeling in the context of studying G-quadruplex structures, highlighting recent successes and limitations to this approach, as well as an example featuring a G-quadruplex structure formed from the human telomere. This example can easily be adapted to the investigation of any other G-quadruplex-forming sequences.

Keywords

G-quadruplex; Nucleic acids; Drug discovery; Hydrodynamic; Bead Models; Sedimentation

1 Introduction

Hydrodynamic bead modeling has emerged as a useful tool for structural biologists studying biological macromolecules and their complexes for which high-resolution structural coordinates are unavailable [1]. The comparison of directly measurable properties with theoretical calculations on structures and molecular models can provide fundamental insight, albeit at low resolution, into structures in solution. One particular macromolecular family that we and others have applied this technique toward is G-quadruplex, a unique type of tertiary structure formed from guanine-rich DNA and RNA sequences. We will present a brief introduction to G-quadruplex nucleic acids focusing on some of the underlying challenges experienced when studying these structures at atomic-level resolution, followed by a discussion of hydrodynamic bead modeling in the study of G-quadruplex structures, and then conclude with some of the limitations and possible future directions of the application of this technique. Here, we focus on aspects of hydrodynamics as it is related to G-quadruplex structures. For a general overview of hydrodynamics history, theories, and modeling techniques, the readers will benefit from several excellent reviews available by Byron [1], Carrasco and García de la Torre[2], and García de la Torre and Bloomfield [3].

2 G-Quadruplex Nucleic Acids

2.1 G-Quadruplex Structures as Drug Targets

The field of drug discovery is largely protein-centric but there are efforts to find and target non-protein systems. A recent assessment of potential drug targets concluded that approximately only 10–15% of the human proteome is “druggable” [4], which is defined as the intersection of sets of proteins that are capable of binding “drug-like” molecules and proteins that are products of disease modifying genes. The total number of potentially viable protein targets, thus, may be very small and so it is critical to consider options for drug discovery that involve other biomolecules. DNA and RNA are fundamentally attractive, yet an underrepresented and underutilized, alternative to proteins as drug targets [4–6]. In the progression through the central dogma; transcription of DNA produces RNA and translation of RNA produces proteins. At each stage of the process, the absolute number of targets that may be hit by a drug molecule increases. Aiming at a particular drug target at the level of gene expression by targeting DNA and RNA rather than the numerous resultant proteins could allow for more efficient drug action. However, targeting genetic DNA sequence selectively via small molecules [7–9] is extremely challenging, as is using oligonucleotide-based approaches such as the antigene, antisense, and iRNA which have their own set of challenges [10–15].

DNA and RNA are polymorphic and can adopt a variety of other secondary and tertiary arrangements in addition to the standard Watson-Crick double helix. Often these unique structures are localized to functionally important sites known to play a role in regulating gene expression. One of these particular tertiary structures is the G-quadruplex, a unique quadruple helix formed from the self-assembly of guanine-rich DNA and RNA sequences containing four or more runs of at least two guanines. A G-quadruplex usually consists of two or three stacked G-quartets or tetrads, which involve four guanines in a square planar arrangement that are stabilized by Hoogsteen hydrogen bonds. The presence of a monovalent cation in the center of the G-quartet can greatly increase the stability of the structure (stability: $Li^+ < Na^+ < K^+$) [16]. G-quadruplexes can form from the unimolecular folding of a single-strand, or form by the bi- or tetramolecular association of strands.

The human genome is believed to contain more than 370,000 putative G-quadruplex forming sequences [17,18]. These sequences have been discovered throughout the genome [17] and are found to be localized to functionally important sites. Amongst these sites is the human telomere [19], a 5000–8000 base pair long sequence of DNA consisting of the repeats $d(T_2A\ G_3)$ at the end of the human chromosome. The telomere, in addition to serving as the biological clock for the cell, also contributes to genetic stability by capping and protecting the chromosomes from damage during replication [20]. The formation of G-quadruplex is believed to occur at the extreme 3′ end of the telomere that contains a single-stranded overhang of 100 to 200 bases [21]. Small molecules that stabilize G-quadruplexes formed from telomeric oligonucleotide sequences *in vitro* have been shown to decrease the activity of telomerase, an enzyme which is responsible for maintaining the length of telomeric DNA, in addition to leading to a DNA double-stranded break response by the cell that results in cellular senescence and cell death [20,22–25]. Since telomerase activation has been found to be involved in greater than 90% of all cancers [26], G-quadruplex formation in the human telomere is an attractive anti-cancer drug target. In addition to the human telomere, G-quadruplex-forming sequences is overrepresented in the promoter regions of many oncogenes, such as *c-Myc*[27], *c-Kit*[28,29], *Bcl-2*[30], *VEGF*[31], and *HIF-1α*[32], whereas the occurrence of G-quadruplex-forming sequences in tumor suppressor genes tends to be much lower [33]. These sequences were found to be evolutionary conserved between humans and related species [34]. The formation of G-quadruplex structures in these promoter regions is believed to influence gene transcription [35,36]. In the case of the

putative G-quadruplex-forming sequence located upstream of the P1 promoter for the *c-Myc* proto-oncogene, G-quadruplex formation was demonstrated to be the primary regulator for *c-Myc* expression [37]. Last, but not least, putative G-quadruplex-forming sequences have also been found in the 5′- or 3′-untranslated regions of about 20% of genes [38] and were discovered to regulate gene expression by repressing translation [39,40] or promoting alternative splicing [41]. These findings seem to suggest that G-quadruplexes could play a key role in fundamental biological processes essential for normal cellular functions as well as cell growth and differentiation.

2.2 Elucidating the Structures of G-Quadruplex-Forming Sequences

Over four decades before the discovery of the double helix, high concentrations of guanine were observed to form gels in aqueous solution [42]. The structure involved was later determined in 1962, nine years after Watson and Crick presented their findings on the double helix, to be the G-quartet or tetrad (Figure 1A) [43]. The double helical structure discovered by Watson and Crick, known as B-DNA, is one of several forms that DNA can adopt. In addition to the canonical B-DNA, two other major forms, A-DNA and Z-DNA have been studied extensively. A recent review revealed that many other forms of DNA have been reported in the literature [44]. However, most of these DNA structures were observed *in vitro* and it is unclear what biological relevance, if any, they possess. While the structures of double helical nucleic acids appear to be highly polymorphic, this polymorphism pales in comparison to the possible structural diversity that can be assumed by guanine quadruple helices. G-quadruplexes can be classified primarily based on molecularity with a decrease in molecularity associated with an increase in the structural polymorphism (Figure 1B–D). Tetramolecular G-quadruplexes, formed from four guanine-rich DNA or RNA sequences, are the least polymorphic, adopting only a parallel conformation with all four strands in the same direction (Figure 1B) [45,46]. Bimolecular G-quadruplexes, formed from the dimerization of two guanine-rich DNA or RNA sequences folded into double-stranded helical duplex structures, are more diverse with three possible conformations, one parallel and two antiparallels (Figure 1C). In the parallel conformation the two strands of the duplex are connected by a double chain-reversal, or propeller, loop while in the antiparallel conformation the two strands are connected by a hairpin loop – a diagonal loop if the interconnecting strands oppose one another within the G-quadruplex structure [47–50] or a lateral, or edgewise, loop if the strands are adjacent to one another [51]. Lastly, unimolecular G-quadruplexes, formed from the folding of a single guanine-rich DNA or RNA sequence into a four-stranded quadruple helix structure with three connecting loops, are the most polymorphic capable of adopting multiple topologies (Figure 1D) depending on the condition.

Much of our understanding of the polymorphic nature of G-quadruplex structures comes from structural studies involving G-quadruplex-forming sequences derived from the human telomere. In 1993, the first structure of a unimolecular G-quadruplex formed from a human genomic sequence was solved by Wang and Patel (PDB: 143D) [52]. It was discovered by NMR spectroscopy that the human telomeric sequence $dAG_3(T_2AG_3)_3$ in Na^+ solution folds into an anti-parallel G-quadruplex topology consisted of three stacked G-quartets with two lateral and a diagonal connecting loops. The structure in K^+ solution, however, proved to be more difficult. In 2002, a structure for the human telomeric sequence in K^+ was determined by X-ray crystallography (PDB: 1KF1) [53]. Parkinson *et al.* reported that in K^+, the human telomeric sequence adopted a parallel topology consisting of three stacked G-quartets and three double chain-reversal loops. However, the reported crystal structure did not agree with biophysical measurements, including sedimentation velocity and equilibrium experiments, CD spectroscopy measurements, and 2-aminopurine fluorescent quenching studies, in solution and in 2005, Li *et al.*, using a combined approach of hydrodynamic bead modeling

and experimental measurement, demonstrated conclusively that the parallel structure was not the predominant form adopted by the human telomeric sequence in K^+ solution [54]. It was suggested in the same study that perhaps the crowding or dehydrated conditions in the crystals selected for the parallel conformation. This hypothesis was confirmed when in 2010, Miller *et al.*, using CD and NMR spectroscopy, determined that dehydration not molecular crowding determines the conformation of G-quadruplex structure [55]. Heddi and Phan, shortly after in 2011, solved the NMR solution structure (PDB: 2LD8) for the same sequence in 40% polyethylene glycol (PEG 200) and discovered that the structure formed was close to the previously reported crystal structure [56]. After the 2005 Li *et al.* report that the solution structure of the human telomere was not predominantly the crystal form, alternative folding topologies were discovered. In addition to the antiparallel and the parallel conformations, two mixed topologies (i.e. three strands in one direction, one strand in the opposite direction) were reported for the human telomeric sequence in K^+ solution by NMR spectroscopy. The first conformation, hybrid-1, was reported by Ambrus *et al.* in April 2006 but did not publish the full structure[57]. Shortly after, Luu *et al.* in July 2006 reported the full NMR structure for the hybrid-1 topology consisting of a three stack G-quadruplex with a chain reversal loop followed by two lateral loops (PDB: 2GKU) [58]. The second mixed conformation, hybrid-2, reported by *Dai et al.* later in 2007 differed from the hybrid-1 conformation by the order of the connecting loops, two lateral loops followed by a chain reversal loop (PDB: 2JPZ) [59]. In addition, Lim *et al.* in 2009 reported an antiparallel structure of the human telomeric sequence in K^+ solution (PDB: 2KF8) that was very similar to the structure in Na^+ solution [60], but contained only two stacked G-quartets in K^+ as opposed to three stacked G-quartets in Na^+ solution. To further expand the observed folding topologies, another antiparallel folding topology, this one consisted of three lateral loops, have also been reported [61]. However, this topology has only been seen with the thrombin binding aptamer and has not been observed for the human telomere or any other human genomic sequences.

As we and others have reported, the folding of a single DNA or RNA strand into a unimolecular G-quadruplex can be highly polymorphic [62,63]. The human telomeric sequence, while it is a relatively simple G-quadruplex-forming sequence, can potentially fold into more than 200 intramolecular conformations [62] with variations in connecting loop types (i.e. lateral, diagonal, or double chain-reversal), strand/segment orientations (i.e. parallel, antiparallel, or hybrid 3+1), numbers of G-quartets (i.e. two or three), and glycosyl torsion angles (i.e. *syn* or *anti*) depending on the conditions (e.g. buffer composition, the presence of organic solvents such as acetonitrile or ethanol, DNA concentration, ion concentration, annealing profile, and the presence of various biological molecules). Consequently, when a G-quadruplex-forming sequence is studied, certain steps are usually taken to reduce this inherent polymorphism. The most common approach for reducing the structural polymorphism of G-quadruplex-forming sequences is sequence modification [64,65]. A sequence is usually modified through a series of base subtraction, addition, and/or substitution to yield a new modified sequence with reduced polymorphism with the ultimate goal of sufficient enrichment of one species for NMR structure elucidation. Sequence modifications often include the incorporation non-canonical bases, such as 8-methylguanine or 8-bromoguanine, which are known to produce G-quadruplex structures with a *syn* glycosidic configuration [66–68], while use of O6-methylguanine, inosine, or 6-thioguanine have been shown to destabilize G-quadruplex formation [69–72]. Substitution of 8-aminoguanine promotes formation of tetramolecular parallel quadruplexes such as those formed by TG_4T [73]. Modifications of the sugar-phosphate backbone by insertion of 5′-5′ or 3′-3′ polarity inversion have also been shown to have a dramatic effect on G-quadruplex formation and stability [74] and use of RNA or LNA force adoption of a *syn* glycosidic guanosine conformation [75–78]. In some cases, sequence modification can result in new sequences that differ only slightly from the parent sequence, e.g. the human telomere

sequence [58–60,79–81], while often the resulting sequences differ significantly from the parent sequence, e.g. the *c-Myc* promoter sequence [82–86]. In addition to sequence modifications, another common approach to reduce polymorphism is by changing the solution conditions. The addition of biological molecules [87] (e.g. sugar, proteins), presence of co-solvents (e.g. acetonitrile, PEG) [55,88], choice of divalent versus monovalent cations [89,90], and cation concentration [91,92] all play a major role in directing G-quadruplex formation and determining stability. Due to the significant influence sequence and conditions have on G-quadruplex population distribution, it is very important that a complete description of the experimental methods/conditions used in sample preparation be fully reported for reproducibility.

One potential consequence in using any polymorphism selection method described above for altering the distribution of the polymorphic states of the G-quadruplex structures is that such an approach can result in an unpredictable perturbation of the system which may or may not be present in the original population of states (this has a significant effect on what can be claimed as "biologically relevant"). In the case of the human telomeric sequence, it was demonstrated that small changes in the flanking base of the G-quadruplex-forming segment $dG_3(T_2AG_3)_3$, (-core-), resulted in dramatically different topologies, such as hybrid-1 (2GKU – 5′-T_2-core-A-3′, 2HY9 – 5′-A_3-core-A_2-3′, and 2JSM – 5′-TA-core-3′) [58,79,80], hybrid-2 (2JPZ – 5′-T_2A–core-T_2-3′, 2JSL – 5′-TA-core-T_2-3′) [59,80], and the antiparallel conformation (2KF8 – 5′-core-T-3′, and 2KKA – 5′-A-core-T-3′) [60,81]. The untested assumption is that these means of reducing the polymorphic distribution selects a member of the ensemble of species originally formed by the parent sequence. In fact, such perturbation can shift the equilibrium to favor species that might not be indicative of, or be completely dissimilar to, any species actually formed *in vitro* or *in vivo* [62]. Further, G-quadruplex polymorphism has severely hindered investigation of quadruplex structure, biophysical properties, small molecule interaction, and the thermodynamics of quadruplex formation. Techniques commonly used to study G-quadruplex DNA are typically too low in resolution to distinguish between species in a mixture, e.g. CD spectroscopy, UV-Vis spectroscopy, ultracentrifugation, and gel electrophoresis, while high resolution techniques, e.g. NMR spectroscopy, are of limited usefulness for mixtures containing multiple G-quadruplex topologies [63].

3 Hydrodynamic Bead Modeling of G-Quadruplex Structures

3.1 History and Overview

Several approaches have been proposed for studying G-quadruplex structures without perturbing the equilibrium of the G-quadruplex species in solution by sequence modification. In one approach, G-quadruplex species in solution are separated using size-exclusion chromatography and individual species are studied using a combination of biophysical techniques (e.g. analytical ultracentrifugation, CD spectroscopy, UV-Vis spectroscopy, NMR spectroscopy) [63,93]. This approach discovered that the *c-Kit* promoter G-quadruplex-forming sequence exists in solution as a mixture of monomers, dimers, and higher-order G-wire structures [93]. The disadvantage of this approach is that it relies on chromatographic separation and is of limited usefulness for some sequences, such as the human telomeric sequence, that elute and sediment essentially as one species. In such instances, another approach employing hydrodynamics bead modeling has emerged as a useful tool for studying G-quadruplex structures at atomic-level resolution. The formulation and development of hydrodynamic bead modeling came about from the extensive theoretical work of Kirkwood, Riseman, Bloomfield, García de la Torre, and Allison. Their research is discussed in several extensive reviews and the readers are referred to these articles for a more in depth discussion of the topic [1–3]. Briefly, the structure of a macromolecule is represented as a collection of solid spheres (i.e. beads on a string model), for which known

hydrodynamic properties can be calculated. A sphere can represent an individual atom, a particular residue, or even a whole chain/segment of the macromolecule. For each bead in the model, the frictional force being exerted on the bead by the bulk solvent and perturbed by other beads in the system as it travels through the solvent is calculated. The set of N calculated frictional forces (where N is the number of beads in the model) is arranged into a matrix from which the frictional tensors can be derived and ultimately, the intrinsic viscosity and translational and rotational properties can be calculated. The calculation time is usually in the order of N^3 which can be a challenge for large protein systems but is generally not a problem for smaller system, such as G-quadruplexes.

The first study of G-quadruplex structures in solution using hydrodynamic bead modeling was reported in 1999 by Niermann *et al.* [94]. In that study, the rotational and translational diffusion coefficients were calculated for the Watson-Crick double helical B-DNA, the single-stranded duplex "hairpin," and the four-stranded tetramolecular G-quadruplex structures. The findings demonstrated good agreement between the calculated and experimental hydrodynamic properties when proper hydration of the molecules was considered. In 2005, Li *et al.* employed hydrodynamic bead modeling to study unimolecular G-quadruplex structures [54]. In that study, the program HYDROPRO was used to calculate the sedimentation co-efficient ($S_{20,w}$) and the translational diffusion coefficient (D_t) for the human telomeric sequence $dAG_3(T_2AG_3)_3$ from the reported high-resolution NMR structure in Na^+ (143D) [52] and the crystal structure in K^+ (1KF1) [53]. HYDROPRO is a program developed by Garcia de la Torre and colleagues for the calculation of hydrodynamic properties of a macromolecule of arbitrary shape. Contrary to its predecessor HYDRO [95], which requires the input of a premade bead model, HYDROPRO can read in an atomic coordinate file (e.g. PDB formatted file) and generate a primary hydrodynamic bead model for calculations. To generate this bead model, HYDROPRO replaces the non-hydrogen atoms with a collection of overlapping spheres (Figure 2). HYDROPRO then calculates the hydrodynamic properties of this primary model using the shell-model methodology [96] proposed by Bloomfield *et al.*, in which the molecular surface of the model is replaced by a shell of smaller overlapping spheres. The theory behind the shell-model method is that only the surface exposed to the bulk solvent contributes to the frictional properties and thus the hydrodynamic behavior of the molecules.

Using HYDROPRO, it was found that using the Na^+ or "basket" structure experimentally determined by NMR yielded calculated $S_{20,w}$ and D_t values that agreed with the experimentally obtained hydrodynamic measurements via analytical untracentrifugation (3.2% lower and 6.0% lower, respectively). However, using the K^+ or "propeller" structure experimentally determined by crystallography yielded calculated $S_{20,w}$ and D_t values that were significantly lower than and inconsistent with the experimentally obtained hydrodynamic measurements (24.9% lower and 18.5% lower, respectively). HYDROPRO was also employed to calculate the $S_{20,w}$-values for a series of *in silico* models to determine possible consistent topologies for the human telomeric sequence in K^+ solution. However, none of the topologies examined (7 antiparallel, 1 parallel, and 1 mixed) yielded a calculated $S_{20,w}$-value sufficiently close enough to the experimentally determined $S_{20,w}$-value measurement. It should be noted that the hybrid topologies were unknown at that time and with their discovery later, the calculated $S_{20,w}$-values of these structures agreed with the experimentally obtained $S_{20,w}$-value for the human telomeric sequence in K^+ solution.

Between 2008 and 2011, hydrodynamic calculations by HYDROPRO was used by Petraccone *et al.* in a novel strategy to study higher-order G-quadruplex structures formed by the human telomeric sequence $(T_2AG_3)_nT_2$ (n = 4, 8, 12) [97–99]. Since the 3′-single-stranded overhang of the telomere can be up to 200 bases long, it follows that multiple G-quadruplexes must form on one telomeric overhang. High-resolution monomeric G-

quadruplex hybrid topology structures (2HY9, 2JPZ, and 1KF1) were employed as building blocks to construct five different higher-order G-quadruplex models. The five models were hybrid-11 (5′-hybrid-1-hybrid-1–3′), hybrid-12 (5′-hybrid-1-hybrid-2–3′), hybrid-21 (5′-hybrid-2-hybrid-1–3′), hybrid-22 (5′-hybrid-2-hybrid-2–3′), and all-propeller (5′-parallel-parallel-3′). Following molecular dynamics simulations, $S_{20,w}$-value and solvent accessible surface areas (SASA) were calculated using HYDROPRO and NACCESS, respectively, for 400 snapshots over the last 4 ns of simulation. The calculated $S_{20,w}$-values were compared with sedimentation velocity experiments while the SASA-values were compared with 2-aminopurine fluorescent experiments [92]. Only one model (hybrid-12) emerged where both the experimentally determined $S_{20,w}$ (2.3% lower) and SASA-value matched the predicted values and the authors concluded that this structure is the most probable structure in solution [97]. Two models (hybrid-21 and all-propeller) yielded calculated $S_{20,w}$ that were as close or even closer in agreement with the experimentally obtained $S_{20,w}$. However, the predicted SASA-values for these two models did not match the experimentally measured SASA-values. Conversely, one model (hybrid-22), yielded a predicted SASA-value that appeared to match the experimentally determined values, however, the $S_{20,w}$ was not as close as the hybrid-12 model. The findings reported by Petraccone et al. illustrate one of the limitations of hydrodynamic calculations in that it is a low resolution technique that should not be employed solely. However, when used in conjunction with other techniques, it could be a powerful tool for predicting possible high-resolution structures. In a followup study looking at trimeric G-quadruplex models, it was determined that the structure for the trimeric G-quadruplex is consistent with the hydrodynamic calculations and experimentally derived properties and follows the hybrid-121 model [98]. Higher-order G-quadruplex structures are discussed in more detail in another chapter in this volume

3.2 Calibration of HYDROPRO Parameters

Two major releases of HYDROPRO have been made publicly available for download (releases 5 and 10). The first versions (5 through 7) only supported one calculation mode consisting of an atomic-level primary model and shell-model calculation [100]. The latest release (version 10) offers several new enhanced capabilities and features three different calculation modes – atomic-level primary model with shell-model calculation, residue-level primary model with shell-model calculation, and residue-level primary model with bead-model calculation – as well as some support for parallel computing, which offers a great improvement in computing time [101].

Following the procedures outlined by Garcia de la Torre [101], we calibrated the HYDROPRO parameters for use in calculating the hydrodynamic properties of G-quadruplex structures. D_t and $S_{20,w}$ were calculated for nine different structures (Table 1) and compared to the values obtained using sedimentation velocity experiments (Table 2). From these nine structures, seven were used to optimize the HYDROPRO parameters. For the G-quadruplex structures used in this study, the NMR solutions were reported as ensembles of six to twelve structures. HYDROPRO calculations were carried out for all structures in the ensembles and the values from the different reported structures were averaged to obtain the calculated value. To determine the experimental value for a particular G-quadruplex structure, the raw sedimentation data were analyzed by two different software packages, DCDT+ (version 2.3.2, John Philo, Thousand Oaks CA) and Sedfit (www.analyticalultracentrifugation.com). The experimental value is the average of the two values determined by the programs from the analytical ultracentrifugation data. For both experimental and calculated values, the 95% confidence limit is reported as the error (Figure 3). In order to determine the optimum atomic element radius (AER) for the beads used in the primary hydrodynamic model for each calculation mode, we fitted the calculated values to the experimental values using a global-fit approach (Equation 1) described in the latest

HYDROPRO calibration report [101] where the differences between the calculated and experimental values are averaged, first, over all the hydrodynamic properties available for each G-quadruplex structure and then over the entire set of all G-quadruplexes.

$$\Delta^2 = \frac{1}{N_{G-Quadruplex}} \sum^{N_{G-Quadruplex}} \frac{1}{N_{Property}} \sum_{X}^{N_{Property}} \left[\frac{calculated_X - experimental_X}{experimental_X} \right]^2 \quad (1)$$

The Δ value is the root mean-square relative difference between the calculated values and the experimental values with 100Δ representing the percent difference typically used to characterize the goodness of prediction. To evaluate the goodness of prediction for a particular hydrodynamic property, a reduced equation (Equation 2) can be used and the $100\Delta_X$ is the percent difference between calculated and experimental values for a particular hydrodynamic property *X*.

$$\Delta_X^2 = \frac{1}{N_{G-Quadruplex}} \sum_{X}^{N_{G-Quadruplex}} \left[\frac{calculated_X - experimental_X}{experimental_X} \right]^2 \quad (2)$$

From the original set of nine structures, the calculated hydrodynamic values for seven NMR solution structures were selected in the global fit analysis (Figure 4). The PDB IDs for the seven structures were 143D, 2GKU, 2HY9, 2JPZ, 2JSL, 2JSM, and 2KF8. These seven structures were selected because they are solution structures and the sequences contained only canonical bases. For the two structures that were excluded from the global fit analysis (i.e. 1KF1, the crystal structure, and 2KKA, which contains a non-canonical inosine substituted for guanine), the hydrodynamic values were still calculated by HYDROPRO and compared with the experimental values. The results for these two structures are discussed in a subsequent section. For the atomic-level primary hydrodynamic model with HYDROPRO shell-model calculations, the best-fit AER was determined to be 2.53 Å with 100Δ to be 2.90 (Table 3). This can be interpreted to be that hydrodynamic values predicted at this AER using this particular calculation method have an error of about 2.90% compared to the experimental values. For a non-hydrogen atom, the atomic radius is approximately 1.8 Å [101]. The determined best fit AER suggests a hydration sphere of about 0.7Å which is slightly less than the 1.1Å typically used in hydrodynamics calculations to estimate hydration [102]. This smaller hydration sphere represents a hydration of about 0.27g/g (Table 4). For comparison, the typical hydration value for nucleic acids is 0.35g/g [103]. The best fit AER determined for G-quadruplex was in close agreement with the standard 2.84 Å AER that was the default setting. When the default AER was used instead of the best fit AER, the error increased from 2.90 to 3.77% (Table 1) of which the major contribution was due to the increased error in predicting the sedimentation coefficient. In the current release of HYDROPRO (version 10), there is an option for calculating hydrodynamics properties at the residue level rather than the atomic level with the bead representation centered on the Cα carbon for each residue. We performed the global fit analysis for the residue-level primary hydrodynamic model with a HYDROPRO shell-model calculation and found that the best fit AER at this level was 4.35 Å with an error of 3.12%. For the residue-level primary hydrodynamic model with a HYDROPRO bead-model calculation, the best fit AER was 5.54 Å with an error of 3.22%. For both of these calculation modes, the increased error compared to the atomic-level calculation was mainly due to the worse prediction of the diffusion translational coefficient. The best fit AER reported for these two modes were similar to the default AERs, 4.84 Å for the residue-level shell-model calculation (5.04% error) and 6.11 Å for the residue-level bead-model calculation (5.29% error). We should

note that in all three calculations modes, the best fit AER for each individual hydrodynamic property was in very close agreement with the best fit AER for the entire model (Figure 4).

All three modes of calculations were able to accurately predict the hydrodynamic properties of the seven G-quadruplex structures fitted (Table 2). The errors of the prediction was well within the range of experimental error for determining the values of hydrodynamic properties, which can be about 5% or more (Figure 3) [103]. Of the nine original structures considered in this study, two were excluded from the global fit analysis. For one of these, 1KF1, was not used because this structure has been previously demonstrated to not accurately predict the hydrodynamic behavior of the structure in solution [54]. Our current findings agree with the previously reported data (Table 2). The other structure, 2KKA, was excluded because the reported structure contained a non-canonical inosine substituted for a guanine. Inosine disrupts G-quadruplex formation [70] and is used reduce the number of possible folding topologies for a polymorphic G-quadruplex sequence. Instead of using this structure to build our global fit model, we selected it to study how base substitution affects the structure in solution. We found that the calculated hydrodynamic values for 2KKA differs from the experimental values with an error of 7.16, 6.66, and 7.47% depending on the calculation method used. Compared with the training set of the seven structures, the error associated with the 2KKA structure is nearly doubled. The findings indicate that, similar to 1KF1, the structure of 2KKA that is selected by the inosine substitution might not be representative of the original structure of the sequence in solution.

4 Limitations of Hydrodynamic Bead Modeling

4.1 Hydrodynamic Bead Modeling is a Low Resolution Technique

Among the chief limitations to hydrodynamic bead modeling is that hydrodynamic measurement itself is a low-resolution technique. The model that accurately predicts the hydrodynamic parameters measured in solution is not necessarily the definitive structure for that molecule but rather one of the possible conformations amongst many others [1]. For example in the previous study by Petraccone *et al.* [97] where hydrodynamic modeling was employed to predict the structures for higher-order G-quadruplexes, three out of the five models examined agreed with the measured hydrodynamic values. However, only one of those three models agreed with the prediction for solvent accessibility for the adenine bases compared to steady-state fluorescent measurements suggesting that it was the consistent model. Hydrodynamic measurements can help rule out a particular conformation if the predicted values differ greatly with the experimental values. However, when that difference is smaller, hydrodynamic predictions should be confirmed by non-hydrodynamic experiments (e.g. 2-aminopurine fluorescence spectroscopy to probe the solvent accessibility of adenine bases or DMS footprinting to probe guanine base-paring interactions). Such is the case with the antiparallel human telomeric G-quadruplex structure, 2KKA. To obtain the enrichment needed for NMR structure determination, an inosine was substituted for the guanine 14 in the original sequence (G14I). The reported NMR structure, thus contained this substitution and from our calculations and measurements on the parent unsubstituted sequence, we can conclude that the structure selected for by inosine substitution appears to differ from the structure or ensemble of structures in solution (without inosine substitution) because of the relatively large error associated with the predicted hydrodynamic properties for this structure (~7%), almost double the error seen with the predictions from the global fit analysis (Table 3). However, such an error of prediction still falls within the experimental error associated with hydrodynamic measurement, which can be 5% or greater [103]. Therefore we cannot conclude as to the reason for this apparent discrepancy if it is due to the enrichment of a small subset of topologies that lead to altered hydrodynamic behavior or whether it is because of a new structure was populated. Perhaps in this case, a similar

approach to probe the loop structure by 2-aminopurine fluorescent measurement is warranted to provide additional insight into the structures of the of 2KKA G-quadruplex.

4.2 Treatment of Hydration in Hydrodynamics Bead Modeling

The structure of a macromolecule in solution, its size, shape, and hydrodynamic properties, can be influenced by the bulk solvent surrounding it through the process of hydration. The accepted interpretation for hydration is that any water whose residence time on the surface of a macromolecule is longer than the hydrodynamic relaxation time (~10 ns) is considered to be a part of the macromolecule and contributes to it hydrodynamic values. There are two current hypotheses to the shape water molecules assume around the macromolecule. In the first hypothesis, uniform expansion, the size of the macromolecule is assumed to expand by a constant factor in all directions. This uniform expansion hypothesis works well for globular structures but breaks down with elongated structures and can lead to abnormal changes in the predicted properties. The inadequacy of the uniform expansion hypothesis has led to an alternative view of hydration being proposed. Instead of assuming a uniform expansion effect, this second hypothesis assumes that hydration forms a layer of uniform thickness around the macromolecule. However, this uniform thickness can result in a relatively thicker hydration layer in the directions where the macromolecule is longer. In HYDROPRO, by replacing the atoms with a representation of beads with a slightly larger size, the effect of hydration on the hydrodynamics properties of the macromolecule can be implicitly accounted for by this uniform hydration hypothesis. The degree of hydration, δ, in units of g/g (grams of water bound per gram of dry solute), can be described as a ratio between the hydrated volume, V_{hyd}, and the anhydrous volume, V_{anh}. The anhydrous volume for a DNA or RNA molecule can be calculated using equation 3:

$$Vanhydrous = \frac{M\bar{\upsilon}}{N_A} \quad (3)$$

M is the molecular weight, $\bar{\upsilon}$ is the partial specific volume, and N_A is Avogadro's number. For G-quadruplex DNA structures, the partial specific volume is assumed to be 0.55 ml/g although there can be small variations from structure to structure [104]. The degree of hydration, δ, is then calculated using equation 4:

$$\delta = \left(\frac{V_{hydrated}}{V_{anhydrous}} - 1 \right) \bar{\upsilon}\rho \quad (4)$$

Using equations 3 and 4 and the hydrated volume of the G-quadruplexes determined by HYDROPRO, we determined the degree of hydration for the seven G-quadruplex structures employed previously in the global fit analysis (i.e. 143D, 2GKU, 2HY9, 2JPZ, 2JSL, 2JSM, 2KF8). The results are shown in Table 4. We did not calculate the degree of hydration for 1KF1 and 2KKA since these reported structures might not be representative of the structures that exist in solution. Using the best-fit AER for each calculation mode, we obtained a degree of hydration of 0.27 g/g for calculations using the atomic-level model with the shell-model calculation, 0.04 g/g for calculations using the residue-level model with the shell-model calculation, and 0.36 g/g for calculations using the residue-level model with the bead-model calculation. The 0.36 g/g degree of hydration predicted by the residue-level model with bead model calculation is the closest to the value of 0.35 g/g reported for nucleic acids by Fernades *et al.* [103]. However, it should be noted that the nucleic acid sequences examined in that particular study were short sequences whose shapes can be more elongated than the G-quadruplex structures. Thus, it is not inconceivable that the hydration for G-quadruplexes would fall closer to the typical 0.3 g/g for globular proteins [102]. In that case,

the 0.27 g/g estimate by the atomic-level model with shell-model calculations is more appropriate. We also performed calculations using the default AER for HYDROPRO and found that the degree of hydrations were 0.38, 0.21, and 0.56 g/g, respectively, for the three calculation modes. The atomic-level model with shell-model calculation method is the only one with a reasonable estimate of hydration (0.38 g/g). The hydration is overestimated by the residue-level model with bead-model calculation (0.56 g/g) while the residue-level model with shell-model calculation still underestimates it (0.21 g/g). It should be noted that the volume calculated by HYDROPRO is based entirely on the size of the beads in the primary hydrodynamic model and is not dependent on the shell-model/bead-model calculation of frictional properties. Thus, the reason that the residue-level model with shell-model calculation underestimate the size of the G-quadruplexes compared to the residue-level model with bead-model calculation is because the former employs a smaller AER value. In fact, when we used 4.35 Å instead of 5.54 Å for the residue-level model with bead-model calculation, we obtained the same degree of hydration of 0.04 g/g. We can conclude then that the residue-level with shell-model calculation is not an appropriate method for hydrodynamic calculations of G-quadruplex structures.

4.3 Potassium Binding by G-Quadruplex and its Implication on Hydrodynamics

If water binding can potentially influence the hydrodynamic properties of a G-quadruplex structure, it follows that the same could be said about the binding of potassium and other ions. Traditionally, the association of cations with a G-quadruplex structure is believed to occur primarily within the central ion channel formed by the G-quartets. The coordination of potassium (K^+), sodium (Na^+), thallium (Tl^+), rubidium (Rb^+), cesium (Cs^+), calcium (Ca^{2+}), lead (Pb^{2+}), barium (Ba^{2+}), and strontium (Sr^{3+}) to the G-quadruplex central ion channel has been observed in several crystal structures [105]. However, it was recently reported that in addition to the canonical internal binding sites, cations can also bind to a G-quadruplex structure externally, particularly within the loop regions [106]. In a study using KCl titration to the human telomere sequence (Tel22 – $d[A(G_3T_2A)_3]$), it was determined that K^+ binds to Tel22 at a stoichiometry of $3K^+$/Tel22 at $[K^+]_{free}$ = 5mM and that stoichiometry increases to 8–10K^+/Tel22 at $[K^+]_{free}$ = 20 mM [107]. In addition, it was reported that the folding of Tel22 into the G-quadruplex structure is driven almost entirely by cation binding with K^+ binding contributing ~-4.9 $kcal_{/mol}$ (at $[K^+]_{free}$ = 5mM) to the overall folding free energy of -2.4 $kcal_{/mol}$. Together, these findings suggest that in addition to stabilizing the G-quartets within the G-quadruplex core, potassium ions can also play other roles in promoting G-quadruplex formation and stability. This hypothesis is further supported by two structural studies, Phan *et al.* in 2011 that reported a structure of RNA G-quadruplex/duplex junction determined by NMR spectroscopy and was the first structure to demonstrate a possible cation binding site on the surface of a G-quadruplex structure [108] and also by Wei *et al.* in 2012, which reported a structure of the human *c-Kit* DNA promoter sequence determined by X-ray crystallography with two K^+ and one Mg^{2+} ions binding in the G-quadruplex loops and grooves in addition to the K^+ ions within canonical central ion chanel [109]. In the crystal structure, all three external cations are believed to play a role in maintaining the *c-Kit G*-quadruplex structure. It should be noted that one of the external K^+ ions appeared to be transient and capable of adopting one of several distinct positions while the other external K^+ ion appeared to be more static suggesting the existence of a high affinity binding site. While the K^+ ions are believed to have a primary role in stabilizing the G-quadruplex structure by direct electrostatic interaction with the DNA, the Mg^{2+} ion is thought to assume a secondary role of shielding the anionic charge of the phosphate groups. In fact, it is well-known that polyanions, e.g. DNA, can attract a shell of cations to help partially neutralizes the negatively-charged backbone phosphates [110,111].

To examine how K^+ ions can affect the structure of a G-quadruplex-forming sequence hydrodynamically, we conducted sedimentation velocity experiments for two G-quadruplex-forming sequences in the presence of different concentrations of K^+ ions (Figure 5). For both sequences, an increase in sedimentation coefficient was observed with the corresponding increase in the amount of potassium present. In low potassium solutions (25 mM) the low sedimentation coefficients can probably be attributed to hydrodynamic non-ideality. In an ideal solution of uncharged particles, the equilibrium distribution of each individual species subjected to a centrifugation field traveling through an incompressible solvent is independent of one another. However, for charged macromolecules such as DNA and some proteins, their equilibrium distributions can be coupled to the distribution of other charged particles (i.e. counter ions such as potassium in our case) by the Donnan effect resulting in non-ideal behavior [112]. Non-ideality can cause a macromolecule to appear much smaller than it actually is and, in effect, results in a small molecular weight and sedimentation coefficient from sedimentation experiments. The Donnan effect can be corrected by one of two ways. The first option is to reduce the charge of the macromolecule by adjusting the pH but this option is only applicable to zwitterionic species. The second option is to “swamp” out the Donnan effect by adding a large quantity of a charged electrolytes (i.e. potassium) to the solution. However, at the higher concentration of potassium, it is unclear how much of the increase in the sedimentation coefficient can be attributed to a shift toward an ideal solution and how much can be attributed to possible direct potassium binding. The binding of a potassium ion to a G-quadruplex structure can have one of two possible consequences. In the first scenario, the potassium ions are bound to the G-quadruplex on the surface, possibly along the phosphate backbones. In this case, these ions not only will change the apparent molecular weight of the G-quadruplex structures but also its shape. In the second scenario, the potassium ion is trapped within the G-quartet core or is buried deep within one of the grooves. Because these potassium ions are not at the surface of the G-quadruplex structure, they do not change the shape of the hydrodynamic model, however, they would cause an underestimation of the apparent molecular weight leading to an under-prediction of the sedimentation coefficient. To further explore this scenario, we performed a series of HYDROPRO calculations and added potassium or sodium to the G-quadruplex structure by making the appropriate increase in the molecular weight value in the HYDROPRO parameter file. The calculations were done using all three modes of HYDROPRO calculation and the default AER instead of the best-fit AER previously determined. In order to display the data for the sequences on the same scale, we normalized the number of potassium ions bound by equation 5 as follow:

$$\%charge\ saturation = \frac{\#\ potassium\ bound}{sequence\ length\ (bp) - 1} \quad (5)$$

The results of the HYDROPRO calculations are shown in Figure 6. For all three calculation modes, the smallest errors for many of the sequences were observed at 40% charge saturation which is equal to approximately eight to ten bound potassium ions. It should be noted that the correction for molecular weight is as good or even better than the previous correction for size of the beads (Figure 4). These findings would suggest that potassium ions not only play a role in G-quadruplex formation and stability but can also influence how the G-quadruplex structure appears in hydrodynamic observations. However, these results require further investigation and validation as there is yet no structural study to back up this hypothesis.

5 Future Directions for G-Quadruplex Hydrodynamic Bead Modeling

G-quadruplex structures have emerged as attractive drug targets because of the overrepresentation of putative quadruplex-forming sequences in functionally important sites within the human genome [17]. In addition, because of their highly polymorphic nature, G-quadruplex-targeted therapies have the potential to be highly selective compared to other DNA targets. However, because of this polymorphism, it can be sometimes difficult to elucidate the relevant structure for a putative G-quadruplex-forming sequence. Modification of the sequence by base subtraction, addition, and substitution or changing of the experimental conditions has been shown to reduce the polymorphism in some cases. However such perturbation can have an unwanted effect by shifting the equilibrium to favor species that might not be indicative of, or be completely dissimilar to, any species actually formed *in vitro* or *in vivo* [62]. The rational drug design of new small molecule G-quadruplex binding agents can be limited by the lack of a definitive target structure. An example of this is the cationic porphyrin, TMPyP4, a G-quadruplex binding agent whose mode of binding and mechanism of action remains a current topic of investigation [82,113–116]. Hydrodynamic bead modeling presents a potentially useful tool to help address this problem. As seen in the reports by Petraccone et al., measurement of hydrodynamic properties, combined with *in silico* simulations (e.g molecular dynamics) and comprehensive biophysical characterization could be used to predict new structures that might serve as a basis for rational drug design.

We present a case study on how this can be conducted. The sequence for the G-quadruplex hybrid-1 structure 2HY9 was obtained from Integrated DNA Technologies (Coralsville, IA) and consists of the sequence:

5′ – dAAAGGGTTAGGGTTAGGGTTAGGGAA – 3′

A stock solution of the oligonucleotide was dissolved in folding buffer, which is composed of monobasic tetrabutylammoniumphosphate (10 mM), EDTA disodium salt dihydrate (1 mM), pH 7.0. Samples were made to an optical density of 0.80 A_{260} by dilution with folding buffer. KCl was added to bring the final concentration to 400 mM. The oligonucleotide samples were annealed in a water bath by heating to 100°C, holding the samples at that temperature for 10 minutes, and gradually cooling to room temperature overnight. Sedimentation velocity experiments was carried out in a Beckman Coulter ProteomeLab XL-A analytical ultracentrifuge at 20°C overnight at 50,000 rpm in standard 2 sector cells. Data were analyzed using two different software packages, DCDT+ (version 2.3.2, John Philo, Thousand Oaks CA) and Sedfit (www.analyticalultracentrifugation.com). Excellent agreement between the two software packages for $S_{20,w}$ values and species distributions was achieved. Buffer density was determined on a Mettler/Paar Calculating Density Meter DMA 55A at 20.0°C and buffer viscosity was calculated using Sednterp software (www.jphilo.mailway.com).

The model for a hybrid-1 G-quadruplex structure was downloaded from the Protein Bank Database (PDB ID: 2HY9) and to the first structure in the file, two potassium ions were added between the G-quartets to complete the model. In a different study with a novel G-quadruplex-forming sequence, this structure could be generated using a fragment-based approach [117]. A solvation model with explicit water and counter ions representations were generated using the parm99SB.dat Amber force field was used with the following protocol: (i) 75 potassium ions were added randomly around the G-quadruplex structures using Amber 11 leap rules for counter ions, (ii) charges were neutralized by the addition of chloride ions, and (iii) the system was solvated by the addition of a rectangular box of TIP3P water (51Å × 52Å × 46Å). The final model represents a system consisting of 13mM G-quadruplex in 1M KCl. The system was heated slowly and equilibrated for 600 ps using the

following protocol: (i) minimize water, (ii) 50 ps molecular dynamics (heating to 300 K) holding the DNA fixed (2 kcal/mol $Å^{-1}$), (iii) 50 ps molecular dynamics (T = 300K) for density equilibration holding the DNA fixed (2 kcal/mol $Å^{-1}$), and (iv) 500 ps molecular dynamics (T = 300K). Production runs of 200 ns after the final equilibration step to obtain snapshots of 10 ps throughout the trajectory. Simulations were performed in the isothermal isobaric ensemble (P = 1atm, T = 300K). Periodic boundary conditions and the Particle-Mesh-Ewald algorithm were used. A 2.0 fs time step was used with bonds involving in hydrogen atoms frozen using SHAKE. For the equilibration steps and the production steps, molecular dynamics calculations were carried out using AMBER 11 program sander and the cuda version of pmemd, respectively. The trajectories were analyzed using the PTRAJ module in AMBER and visualized using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco. The hydrodynamic properties of the simulated models were computed using HYDROPRO [101] and the best-fit AERs, 2.53 Å, 4.35 Å, 5.54 Å, previously determined by global analysis fitting. The distributions of the sedimentation coefficients were computed over the last 20 ns of the trajectory (2,000 snapshots in the last 20 ns of simulation).

The results are shown in Figure 7 with the calculated sedimentation coefficient distribution (colored lines) overlaying the experimentally derived sedimentation coefficient distribution (black line). The results from all three calculation modes, atomic-level model with shell-model calculation, residue-level model with shell-model calculation, and residue-level model with bead-model calculation, were in close agreement with each other. The average of all the sedimentation coefficients from the 2,000 snapshots was considered as the calculated sedimentation coefficient. The sedimentation coefficient was 2.040 for the atomic-level model with the shell-model calculation, 2.044 for the residue-level with the shell-model calculation, and 2.043 for the residue-level model with the bead-model calculation This corresponds $100\Delta_X$ values of error of 2.65, 2.41, and 2.48% for the three calculation modes respectively (Table 5). These results agreed with previous calculations performed for the global fit analysis (Table 3). It should be noted that in all cases the experimentally derived sedimentation coefficient distribution was wider than the calculated distribution, possibly due to the larger extent of the experimental error (Figure 7). In addition, in all three calculation modes, the distribution of sedimentation coefficient remained unchanged between 180–190 ns and 190–200 ns compared to 180–200 ns indicating that the simulation has reached equilibrium.

6 Conclusion

We have demonstrate that hydrodynamic bead modeling is a powerful tool for studying G-quadruplex structures when experimentally determined high-resolution structural representations are not available or cannot be feasibly obtained. In such cases, the use of low-resolution techniques such as hydrodynamics, combined with readily accessible biophysical measurements (e.g. CD spectroscopy, fluorescent spectroscopy), can be used to provide a high-resolution glimpse into these structures. One of the limitations of hydrodynamic calculations, such as HYDROPRO, is that the calculation is performed on one structure giving a static look at an otherwise dynamic system. Recent advancements in computation technology, particularly the development of molecular dynamic calculations to take advantage of the speed and calculation power of graphical processor units, allows for the researcher to conduct longer and more complex simulations. These can generate a more dynamic representation of the macromolecule structure that can be validated by hydrodynamic bead modeling and be used as a basis for drug design. From our work, we recommend for G-quadruplexes using either the atomic-level model with shell-model calculation mode (AER = 2.53 Å) or the residue-level model with bead-model calculation mode (AER = 5.54 Å) for HYDROPRO calculations, as both modes can predict the

hydrodynamic properties accurately with a reasonable estimate of the size of the macromolecule. The residue-level model with bead-model calculation mode has the advantage of being significantly faster but with a slightly higher error in predicting hydrodynamic properties. The parameters for HYDROPRO presented here can be used as is for hydrodynamic calculation of G-quadruplexes or can be further optimized against a more extensive set of hydrodynamic properties.

Acknowledgments

We would like to acknowledge Dr. Robert D. Gray for useful discussions in the writing of this chapter. Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR001081). This work was supported by NIH Grants CA35635 (J.B.C), GM077422 (J.B.C. & J.O.T) and University of Louisville grant, CTSPGP 20058 Award (J.B.C. & J.O.T).

References

1. Byron, O. Hydrodynamic Modeling: The Solution Conformation of Macromolecules and Their Complexes. In: John, JC.; William Detrich, H., III, editors. Methods in Cell Biology. Vol. 84. Academic Press; Massachusetts: 2008.
2. Carrasco B, García de la Torre J. Hydrodynamic Properties of Rigid Particles: Comparison of Different Modeling and Computational Procedures. Biophysical Journal. 1999; 76 (6):3044–3057. [PubMed: 10354430]
3. García de la Torre J, Bloomfield VA. Hydrodynamic properties of complex, rigid, biological macromolecules: theory and applications. Quarterly Reviews of Biophysics. 1981; 14 (01):81–139. [PubMed: 7025081]
4. Hopkins AL, Groom CR. The druggable genome. Nature Reviews Drug Discovery. 2002; 1 (9): 727–730.
5. Drews J. Drug Discovery: A Historical Perspective. Science. 2000; 287 (5460):1960–1964. [PubMed: 10720314]
6. Imming P, Sinning C, Meyer A. Drugs, their targets and the nature and number of drug targets. Nature Reviews Drug Discovery. 2006; 5 (10):821–834.
7. Chaires JB. Allosteric conversion of Z DNA to an intercalated right-handed conformation by daunomycin. Journal of Biological Chemistry. 1986; 261 (19):8899–8907. [PubMed: 3722181]
8. Chaires JB. Inhibition of the thermally driven B to Z transition by intercalating drugs. Biochemistry. 1986; 25 (26):8436–8439. [PubMed: 3828287]
9. Qu X, Trent JO, Fokt I, Priebe W, Chaires JB. Allosteric, chiral-selective drug binding to DNA. Proceedings of the National Academy of Sciences. 2000; 97 (22):12032–12037.
10. Scanlon K. Anti-genes: siRNA, ribozymes and antisense. Current Pharmaceutical Biotechnology. 2004; 5 (5):415–420. [PubMed: 15544489]
11. Sepp-Lorenzino L, Ruddy MK. Challenges and Opportunities for Local and Systemic Delivery of siRNA and Antisense Oligonucleotides. Clinical Pharmacology and Therapeutics. 2008; 84 (5): 628–632. [PubMed: 18800034]
12. de Fougerolles A, Vornlocher H-P, Maraganore J, Lieberman J. Interfering with disease: a progress report on siRNA-based therapeutics. Nature Reviews Drug Discovery. 2007; 6 (6):443–453.
13. Tiemann K, Rossi JJ. RNAi-based therapeutics–current status, challenges and prospects. EMBO Molecular Medicine. 2009; 1 (3):142–151. [PubMed: 20049714]
14. Garzon R, Marcucci G, Croce CM. Targeting microRNAs in cancer: rationale, strategies and challenges. Nature Reviews Drug Discovery. 2010; 9 (10):775–789.
15. Castanotto D, Rossi JJ. The promises and pitfalls of RNA-interference-based therapeutics. Nature. 2009; 457 (7228):426–433. [PubMed: 19158789]
16. Williamson JR, Raghuraman MK, Cech TR. Monovalent cation-induced structure of telomeric DNA: The G-quartet model. Cell. 1989; 59 (5):871–880. [PubMed: 2590943]

17. Huppert JL, Balasubramanian S. Prevalence of quadruplexes in the human genome. Nucleic Acids Research. 2005; 33 (9):2908–2916. [PubMed: 15914667]
18. Huppert JL, Balasubramanian S. G-quadruplexes in promoters throughout the human genome. Nucleic Acids Research. 2007; 35 (2):406–413. [PubMed: 17169996]
19. Neidle S. Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. FEBS Journal. 2010; 277 (5):1118–1125. [PubMed: 19951354]
20. De Cian A, Lacroix L, Douarre C, Temime-Smaali N, Trentesaux C, Riou JF, Mergny JL. Targeting telomeres and telomerase. Biochimie. 2008; 90 (1):131–155. [PubMed: 17822826]
21. Wright WE, Tesmer VM, Huffman KE, Levene SD, Shay JW. Normal human chromosomes have long G-rich telomeric overhangs at one end. Genes and Development. 1997; 11 (21):2801–2809. [PubMed: 9353250]
22. Riou JF, Guittat L, Mailliet P, Laoui A, Renou E, Petitgenet O, Mégnin-Chanet F, Hélène C, Mergny JL. Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. Proceedings of the National Academy of Sciences. 2002; 99 (5):2672–2677.
23. Cuesta J, Read MA, Neidle S. The Design of G-quadruplex Ligands as Telomerase Inhibitors. Mini Reviews in Medicinal Chemistry. 2003; 3 (1):11. [PubMed: 12570851]
24. Lopes J, Piazza A, Bermejo R, Kriegsman B, Colosio A, Teulade-Fichou M-P, Foiani M, Nicolas A. G-quadruplex-induced instability during leading-strand replication. EMBO Journal. 2011; 30 (19):4033–4046. [PubMed: 21873979]
25. Rodriguez R, Miller KM, Forment JV, Bradshaw CR, Nikan M, Britton S, Oelschlaegel T, Xhemalce B, Balasubramanian S, Jackson SP. Small-molecule–induced DNA damage identifies alternative DNA structures in human genes. Nature Chemical Biology. 2012; 8 (3):301–310.
26. Shay JW, Bacchetti S. A survey of telomerase activity in human cancer. European Journal of Cancer. 1997; 33 (5):787–791. [PubMed: 9282118]
27. Ambrus A, Chen D, Dai JX, Jones RA, Yang DZ. Solution structure of the biologically relevant g-quadruplex element in the human *c-MYC* promoter. implications for g-quadruplex stabilization. Biochemistry. 2005; 44 (6):2048–2058. [PubMed: 15697230]
28. Hsu STD, Varnai P, Bugaut A, Reszka AP, Neidle S, Balasubramanian S. A G-Rich Sequence within the c-kit Oncogene Promoter Forms a Parallel G-Quadruplex Having Asymmetric G-Tetrad Dynamics. Journal of the American Chemical Society. 2009; 131 (37):13399–13409. [PubMed: 19705869]
29. Phan AT, Kuryavyi V, Burge S, Neidle S, Patel DJ. Structure of an unprecedented G-quadruplex scaffold in the human *c-Kit* promoter. J Am Chem Soc. 2007; 129 (14):4386–4392. [PubMed: 17362008]
30. Dai J, Chen D, Jones RA, Hurley LH, Yang D. NMR solution structure of the major G-quadruplex structure formed in the human *BCL2* promoter region. Nucleic Acids Research. 2006; 34 (18): 5133–5144. [PubMed: 16998187]
31. Sun DY, Guo KX, Rusche JJ, Hurley LH. Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human *VEGF* gene by the presence of potassium and G-quadruplex-interactive agents. Nucleic Acids Research. 2005; 33 (18):6070–6080. [PubMed: 16239639]
32. De Armond R, Wood S, Sun DY, Hurley LH, Ebbinghaus SW. Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1 alpha promoter. Biochemistry. 2005; 44 (49):16341–16350. [PubMed: 16331995]
33. Eddy J, Maizels N. Gene function correlates with potential for G4 DNA formation in the human genome. Nucleic Acids Research. 2006; 34 (14):3887–3896. [PubMed: 16914419]
34. Verma A, Halder K, Halder R, Yadav VK, Rawal P, Thakur RK, Mohd F, Sharma A, Chowdhury S. Genome-Wide Computational and Expression Analyses Reveal G-Quadruplex DNA Motifs as Conserved *cis*-Regulatory Elements in Human and Related Species. Journal of Medicinal Chemistry. 2008; 51 (18):5641–5649. [PubMed: 18767830]
35. Brooks TA, Kendrick S, Hurley L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. FEBS Journal. 2010; 277 (17):3459–3469. [PubMed: 20670278]

36. Balasubramanian S, Hurley LH, Neidle S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? Nature Reviews Drug Discovery. 2011; 10 (4):261–275.
37. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress *c-MYC* transcription. Proceedings of the National Academy of Sciences. 2002; 99 (18):11593–11598.
38. Huppert JL, Bugaut A, Kumari S, Balasubramanian S. G-quadruplexes: the beginning and end of UTRs. Nucleic Acids Research. 2008; 36 (19):6260–6268. [PubMed: 18832370]
39. Halder K, Wieland M, Hartig JS. Predictable suppression of gene expression by 5′-UTR-based RNA quadruplexes. Nucleic Acids Research. 2009; 37 (20):6811–6817. [PubMed: 19740765]
40. Arora A, Suess B. An RNA G-quadruplex in the 3′ UTR of the proto-oncogene *PIM1* represses translation. RNA Biology. 2011; 8 (5):802–805. [PubMed: 21734463]
41. Marcel V, Tran PLT, Sagne C, Martel-Planche G, Vaslin L, Teulade-Fichou M-P, Hall J, Mergny J-L, Hainaut P, Van Dyck E. G-quadruplex structures in *TP53* intron 3: role in alternative splicing and in production of p53 mRNA isoforms. Carcinogenesis. 2011; 32 (3):271–278. [PubMed: 21112961]
42. Bang I. Untersuchungen über die Guanylsäure. Biochemische Zeitschrift. 1910; 26:293–311.
43. Gellert M, Lipsett MN, Davies DR. HELIX FORMATION BY GUANYLIC ACID. Proceedings of the National Academy of Sciences. 1962; 48 (12):2013–2018.
44. Ghosh A, Bansal M. A glossary of DNA structures from A to Z. Acta Crystallographica Section D. 2003; 59 (4):620–626.
45. Laughlan G, Murchie A, Norman D, Moore M, Moody P, Lilley D, Luisi B. The high-resolution crystal structure of a parallel-stranded guanine tetraplex. Science. 1994; 265 (5171):520–524. [PubMed: 8036494]
46. Phillips K, Dauter Z, Murchie AIH, Lilley DMJ, Luisi B. The crystal structure of a parallel-stranded guanine tetraplex at 0.95Å resolution. Journal of Molecular Biology. 1997; 273 (1):171–182. [PubMed: 9367755]
47. Smith FFWJ. Quadruplex structure of Oxytricha telomeric DNA oligonucleotides. Nature. 1992; 356 (6365):164–168. [PubMed: 1545871]
48. Smith FW, Feigon J. Strand orientation in the DNA quadruplex formed from the Oxytricha telomere repeat oligonucleotide d($G_4T_4G_4$) in solution. Biochemistry. 1993; 32 (33):8682–8692. [PubMed: 8357810]
49. Haider S, Parkinson GN, Neidle S. Crystal Structure of the Potassium Form of an *Oxytricha nova* G-quadruplex. Journal of Molecular Biology. 2002; 320 (2):189–200. [PubMed: 12079378]
50. Horvath MP, Schultz SC. DNA G-quartets in a 1.86 Å resolution structure of an *Oxytricha nova* telomeric protein-DNA complex. Journal of Molecular Biology. 2001; 310 (2):367–377. [PubMed: 11428895]
51. Kang C, Zhang X, Ratliff R, Moyzis R, Rich A. Crystal structure of four-stranded Oxytricha telomeric DNA. Nature. 1992; 356 (6365):126–131. [PubMed: 1545863]
52. Wang Y, Patel DJ. Solution structure of the human telomeric repeat d[$AG_3(T_2AG_3)_3$] G-tetraplex. Structure. 1993; 1 (4):263–282. [PubMed: 8081740]
53. Parkinson GN, Lee MPH, Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. Nature. 2002; 417 (6891):876–880. [PubMed: 12050675]
54. Li J, Correia JJ, Wang L, Trent JO, Chaires JB. Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. Nucleic Acids Research. 2005; 33 (14):4649–4659. [PubMed: 16106044]
55. Miller MC, Buscaglia R, Chaires JB, Lane AN, Trent JO. Hydration Is a Major Determinant of the G-Quadruplex Stability and Conformation of the Human Telomere 3′ Sequence of d($AG_3(TTAG_3)_3$). Journal of the American Chemical Society. 2010; 132 (48):17105–17107.
56. Heddi B, Phan AT. Structure of Human Telomeric DNA in Crowded Solution. Journal of the American Chemical Society. 2011; 133 (25):9824–9833. [PubMed: 21548653]
57. Ambrus A, Chen D, Dai J, Bialis T, Jones RA, Yang D. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. Nucleic Acids Research. 2006; 34 (9):2723–2735. [PubMed: 16714449]

58. Luu KN, Phan AT, Kuryavyi V, Lacroix L, Patel DJ. Structure of the Human Telomere in K+ Solution: An Intramolecular (3 + 1) G-Quadruplex Scaffold. Journal of the American Chemical Society. 2006; 128 (30):9963–9970. [PubMed: 16866556]
59. Dai J, Carver M, Punchihewa C, Jones RA, Yang D. Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K^+ solution: insights into structure polymorphism of the human telomeric sequence. Nucleic Acids Research. 2007; 35 (15):4927–4940. [PubMed: 17626043]
60. Lim KW, Amrane S, Bouaziz S, Xu W, Mu Y, Patel DJ, Luu KN, Phan AT. Structure of the Human Telomere in K^+ Solution: A Stable Basket-Type G-Quadruplex with Only Two G-Tetrad Layers. Journal of the American Chemical Society. 2009; 131 (12):4301–4309. [PubMed: 19271707]
61. Schultze P, Macaya RF, Feigon J. Three-dimensional Solution Structure of the Thrombin-binding DNA Aptamer d(GGTTGGTGTGGTTGG). Journal of Molecular Biology. 1994; 235 (5):1532–1547. [PubMed: 8107090]
62. Lane AN, Chaires JB, Gray RD, Trent JO. Stability and kinetics of G-quadruplex structures. Nucleic Acids Research. 2008; 36 (17):5482–5515. [PubMed: 18718931]
63. Dailey MM, Miller MC, Bates PJ, Lane AN, Trent JO. Resolution and characterization of the structural polymorphism of a single quadruplex-forming sequence. Nucleic Acids Research. 2010; 38 (14):4877–4888. [PubMed: 20348136]
64. Dai J, Carver M, Yang D. Polymorphism of human telomeric quadruplex structures. Biochimie. 2008; 90 (8):1172–1183. [PubMed: 18373984]
65. Yang D, Okamoto K. Structural insights into G-quadruplexes: towards new anticancer drugs. Future Medicinal Chemistry. 2010; 2 (4):619–646. [PubMed: 20563318]
66. Virgilio A, Esposito V, Randazzo A, Mayol L, Galeone A. 8-Methyl-2′-deoxyguanosine incorporation into parallel DNA quadruplex structures. Nucleic Acids Research. 2005; 33 (19): 6188–6195. [PubMed: 16257981]
67. Virgilio A, Esposito V, Randazzo A, Mayol L, Galeone A. Effects of 8-methyl-2 ′-deoxyadenosine incorporation into quadruplex forming oligodeoxyribonucleotides. Bioorganic and Medicinal Chemistry. 2005; 13 (4):1037–1044. [PubMed: 15670911]
68. Esposito V, Randazzo A, Piccialli G, Petraccone L, Giancola C, Mayol L. Effects of an 8-bromodeoxyguanosine incorporation on the parallel quadruplex structure [d(TGGGT)](4). Organic & Biomolecular Chemistry. 2004; 2 (3):313–318. [PubMed: 14747859]
69. Mekmaysy CS, Petraccone L, Garbett NC, Ragazzon PA, Gray RD, Trent JO, Chaires JB. Effect of O6-methylguanine on the stability of G-quadruplex DNA. Journal of the American Chemical Society. 2008; 130 (21):6710–6711. [PubMed: 18447358]
70. Petrovic AG, Polavarapu PL. Quadruplex structure of polyriboinosinic acid: Dependence on alkali metal ion concentration, pH and temperature. Journal of Physical Chemistry B. 2008; 112 (7): 2255–2260.
71. Marathias VM, Sawicki MJ, Bolton PH. 6-Thioguanine alters the structure and stability of duplex DNA and inhibits quadruplex DNA formation. Nucleic Acids Research. 1999; 27 (14):2860–2867. [PubMed: 10390526]
72. Spackova N, Cubero E, Sponer J, Orozco M. Theoretical study of the guanine -> 6-thioguanine substitution in duplexes, triplexes, and tetraplexes. Journal of the American Chemical Society. 2004; 126 (44):14642–14650. [PubMed: 15521784]
73. Gros J, Avino A, de la Osa JL, Gonzalez C, Lacroix L, Perez A, Orozco M, Eritja R, Mergny JL. 8-amino guanine accelerates tetramolecular G-quadruplex formation. Chemical Communications. 2008; (25):2926–2928. [PubMed: 18566727]
74. Esposito V, Virgilio A, Randazzo A, Galeone A, Mayol L. A new class of DNA quadruplexes formed by oligodeoxyribonucleotides containing a 3′-3′ or 5′-5′ inversion of polarity site. Chemical Communications. 2005; (31):3953–3955. [PubMed: 16075083]
75. Bonifacio L, Church FC, Jarstfer MB. Effect of locked-nucleic acid on a biologically active G-quadruplex. A structure-activity relationship of the thrombin aptamer. International Journal of Molecular Sciences. 2008; 9 (3):422–433. [PubMed: 19325759]
76. Kumar N, Maiti S. Role of locked nucleic acid modified complementary strand in quadruplex/Watson-Crick duplex equilibrium. Journal of Physical Chemistry B. 2007; 111 (42):12328–12337.

77. Tang CF, Shafer RH. Engineering the quadruplex fold: nucleoside conformation determines both folding topology and molecularity in guanine quadruplexes. Journal of the American Chemical Society. 2006; 128 (17):5966–5973. [PubMed: 16637665]

78. Qi J, Shafer RH. Human telomere quadruplex: refolding and selection of individual conformers via RNA/DNA chimeric editing. Biochemistry. 2007; 46 (25):7599–7606. [PubMed: 17539606]

79. Dai J, Punchihewa C, Ambrus A, Chen D, Jones RA, Yang D. Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation. Nucleic Acids Research. 2007; 35 (7):2440–2450. [PubMed: 17395643]

80. Phan AT, Kuryavyi V, Luu KN, Patel DJ. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K^+ solution. Nucleic Acids Research. 2007; 35 (19): 6517–6525. [PubMed: 17895279]

81. Zhang Z, Dai J, Veliath E, Jones RA, Yang D. Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K^+ solution: insights into the interconversion of human telomeric G-quadruplex structures. Nucleic Acids Research. 2010; 38 (3):1009–1021. [PubMed: 19946019]

82. Seenisamy J, Rezler EM, Powell TJ, Tye D, Gokhale V, Joshi CS, Siddiqui-Jain A, Hurley LH. The Dynamic Character of the G-Quadruplex Element in the *c-MYC* Promoter and Modification by TMPyP4. Journal of the American Chemical Society. 2004; 126 (28):8702–8709. [PubMed: 15250722]

83. Phan AT, Modi YS, Patel DJ. Propeller-Type Parallel-Stranded G-Quadruplexes in the Human c-myc Promoter. Journal of the American Chemical Society. 2004; 126 (28):8710–8716. [PubMed: 15250723]

84. Ambrus A, Chen D, Dai J, Jones RA, Yang D. Solution Structure of the Biologically Relevant G-Quadruplex Element in the Human *c-MYC* Promoter. Implications for G-Quadruplex Stabilization. Biochemistry. 2005; 44 (6):2048–2058. [PubMed: 15697230]

85. Phan AT, Kuryavyi V, Gaw HY, Patel DJ. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human *MYC* promoter. Nature Chemical Biology. 2005; 1 (3):167–173.

86. Mathad RI, Hatzakis E, Dai J, Yang D. *c-MYC* promoter G-quadruplex formed at the 5′-end of NHE III1 element: insights into biological relevance and parallel-stranded G-quadruplex stability. Nucleic Acids Research. 2011

87. Sannohe, Y.; Sugiyama, H. Current Protocols in Nucleic Acid Chemistry. John Wiley & Sons, Inc; New York: 2001. Overview of Formation of G-Quadruplex Structures.

88. Xue Y, Kan Z-y, Wang Q, Yao Y, Liu J, Hao Y-h, Tan Z. Human Telomeric DNA Forms Parallel-Stranded Intramolecular G-Quadruplex in K^+ Solution under Molecular Crowding Condition. Journal of the American Chemical Society. 2007; 129 (36):11185–11191. [PubMed: 17705383]

89. Blume SW, Guarcello V, Zacharias W, Miller DM. Divalent Transition Metal Cations Counteract Potassium-Induced Quadruplex Assembly of Oligo(dG) Sequences. Nucleic Acids Research. 1997; 25 (3):617–625. [PubMed: 9016604]

90. Miyoshi D, Nakao A, Sugimoto N. Structural transition of $d(G_4T_4G_4)$ from antiparallel to parallel G-quartet induced by divalent cations. Nucleic Acids Symposium Series. 2001; 1 (1):259–260.

91. Gray RD, Li J, Chaires JB. Energetics and Kinetics of a Conformational Switch in G-Quadruplex DNA. The Journal of Physical Chemistry B. 2009; 113 (9):2676–2683. [PubMed: 19708205]

92. Gray RD, Petraccone L, Trent JO, Chaires JB. Characterization of a K^+-Induced Conformational Switch in a Human Telomeric DNA Oligonucleotide Using 2-Aminopurine Fluorescence. Biochemistry. 2009; 49 (1):179–194. [PubMed: 19961180]

93. Miller MC, Le HT, Dean WL, Holt PA, Chaires JB, Trent JO. Polymorphism and resolution of oncogene promoter quadruplex-forming sequences. Organic & Biomolecular Chemistry. 2011; 9 (22):7633–7637. [PubMed: 21938285]

94. Niermann M, Bolten M, Eimer W. Optimization of the Hydrodynamic Bead Model for the Analysis of DNA Conformations in Solution. The Journal of Physical Chemistry B. 1999; 103 (45):10065–10074.

95. Garcia de la Torre J, Navarro S, Lopez Martinez MC, Diaz FG, Lopez Cascales JJ. HYDRO: a computer program for the prediction of hydrodynamic properties of macromolecules. Biophysical Journal. 1994; 67 (2):530–531. [PubMed: 7948671]

96. Bloomfield V, Dalton WO, Van Holde KE. Frictional coefficients of multisubunit structures. I. Theory. Biopolymers. 1967; 5 (2):135–148. [PubMed: 6040712]

97. Petraccone L, Garbett NC, Chaires JB, Trent JO. An integrated molecular dynamics (MD) and experimental study of higher order human telomeric quadruplexes. Biopolymers. 2010; 93 (6): 533–548. [PubMed: 20095044]

98. Petraccone L, Spink C, Trent JO, Garbett NC, Mekmaysy CS, Giancola C, Chaires JB. Structure and Stability of Higher-Order Human Telomeric Quadruplexes. Journal of the American Chemical Society. 2011; 133 (51):20951–20961. [PubMed: 22082001]

99. Petraccone L, Trent JO, Chaires JB. The tail of the telomere. Journal of the American Chemical Society. 2008; 130 (49):16530–16532. [PubMed: 19049455]

100. García de la Torre J, Huertas ML, Carrasco B. Calculation of Hydrodynamic Properties of Globular Proteins from Their Atomic-Level Structure. Biophysical Journal. 2000; 78 (2):719–730. [PubMed: 10653785]

101. Ortega A, Amorós D, García de la Torre J. Prediction of Hydrodynamic and Other Solution Properties of Rigid Proteins from Atomic- and Residue-Level Models. Biophysical Journal. 2011; 101 (4):892–898. [PubMed: 21843480]

102. Garc a de la Torre J. Hydration from hydrodynamics. General considerations and applications of bead modelling to globular proteins. Biophysical Chemistry. 2001; 93 (2–3):159–170. [PubMed: 11804723]

103. Fernandes MX, Ortega A, López Martínez MC, García de la Torre J. Calculation of hydrodynamic properties of small nucleic acids from their atomic structure. Nucleic Acids Research. 2002; 30 (8):1782–1788. [PubMed: 11937632]

104. Hellman L, Rodgers D, Fried M. Phenomenological partial-specific volumes for G-quadruplex DNAs. European Biophysics Journal. 2010; 39 (3):389–396. [PubMed: 19238377]

105. Campbell N, Neidle S. G-quadruplexes and metal ions. Metal Ions in Life Sciences. 2012:119–134. [PubMed: 22210337]

106. Ida R, Wu G. Direct NMR Detection of Alkali Metal Ions Bound to G-Quadruplex DNA. Journal of the American Chemical Society. 2008; 130 (11):3590–3602. [PubMed: 18293981]

107. Gray RD, Chaires JB. Linkage of cation binding and folding in human telomeric quadruplex DNA. Biophysical Chemistry. 2011; 159 (1):205–209. [PubMed: 21764207]

108. Phan AT, Kuryavyi V, Darnell JC, Serganov A, Majumdar A, Ilin S, Raslin T, Polonskaia A, Chen C, Clain D, Darnell RB, Patel DJ. Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. Nature Structural and Molecular Biology. 2011; 18 (7):796–804.

109. Wei D, Parkinson GN, Reszka AP, Neidle S. Crystal structure of a *c-kit* promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation. Nucleic Acids Research. 2012; 40 (10):4691–4700. [PubMed: 22287624]

110. Manning GS. The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. Quarterly Reviews of Biophysics. 1978; 11 (02):179–246. [PubMed: 353876]

111. Record MT, Anderson CF, Lohman TM. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. Quarterly Reviews of Biophysics. 1978; 11 (02):103–178. [PubMed: 353875]

112. Yphantis DA, Roark DE. Equilibrium centrifugation of nonideal systems. Donnan effect in self-associating systems. Biochemistry. 1971; 10 (17):3241–3249. [PubMed: 5119248]

113. Anantha NV, Azam M, Sheardy RD. Porphyrin Binding to Quadruplexed T_4G_4. Biochemistry. 1998; 37 (9):2709–2714. [PubMed: 9537740]

114. Haq I, Trent JO, Chowdhry BZ, Jenkins TC. Intercalative G-Tetraplex Stabilization of Telomeric DNA by a Cationic Porphyrin1. Journal of the American Chemical Society. 1999; 121 (9):1768–1779.

115. Freyer MW, Buscaglia R, Kaplan K, Cashman D, Hurley LH, Lewis EA. Biophysical Studies of the *c-MYC* NHE III1 Promoter: Model Quadruplex Interactions with a Cationic Porphyrin. Biophysical Journal. 2007; 92 (6):2007–2015. [PubMed: 17172304]

116. Wei C, Jia G, Zhou J, Han G, Li C. Evidence for the binding mode of porphyrins to G-quadruplex DNA. Physical Chemistry Chemical Physics. 2009; 11 (20):4025–4032. [PubMed: 19440632]

117. Fogolari F, Haridas H, Corazza A, Viglino P, Cora D, Caselle M, Esposito G, Xodo L. Molecular models for intrastrand DNA G-quadruplexes. BMC Structural Biology. 2009; 9 (1):64. [PubMed: 19811654]

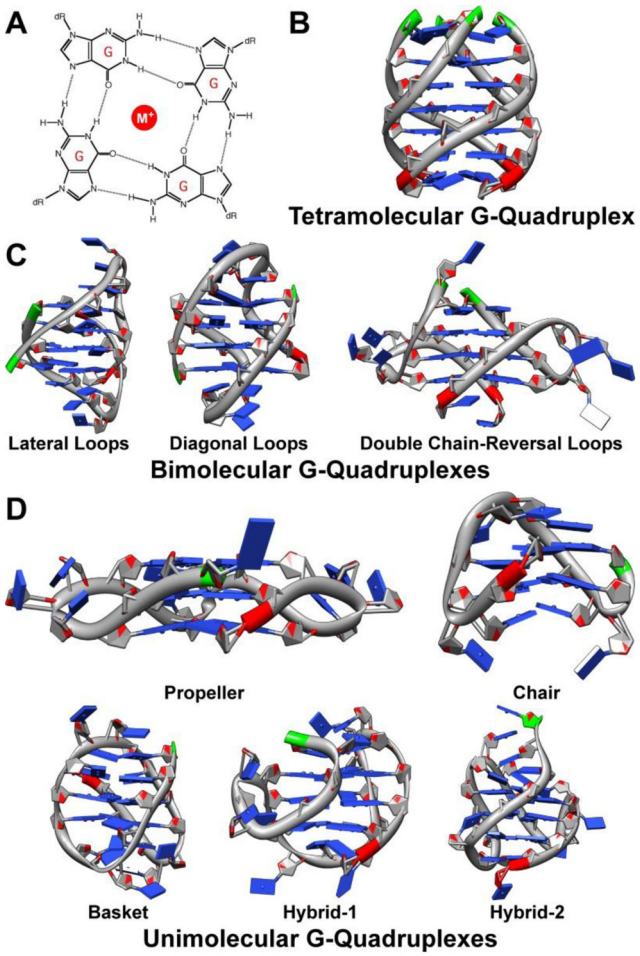

Fig. 1. G-Quadruplex Structures and Polymorphism

The formation of a G-quartet (A), a square planar arrangement of four guanine bases stabilized by centrally coordinating monovalent cation with Hoogsteen hydrogen bonds between guanine bases, is the first step in the polymorphic formation of G-quadruplex structures (B–C). In (B), a tetramolecular G-quadruplex, from four guanine-rich DNA or RNA strands, assumes a parallel topology (PDB ID: 139D). In (C), bimolecular G-quadruplexes, from two guanine-rich sequences, can form either a parallel structure with two double chain reversal loops (3CE5) or one of two possible anti-parallel topologies with either two lateral loops (1D59) or two diagonal loops (1JRN). In (D), G-quadruplex structures are formed from the folding of a single guanine-rich sequence into a quadruple helical structure that can assume one of several topologies including but not limited to propeller (1KF1), chair (148D), basket (143D), hybrid-1 (2GKU), or hybrid-2 (2JPZ). In all structures, the 5′ end is colored green and the 3′ end is colored red

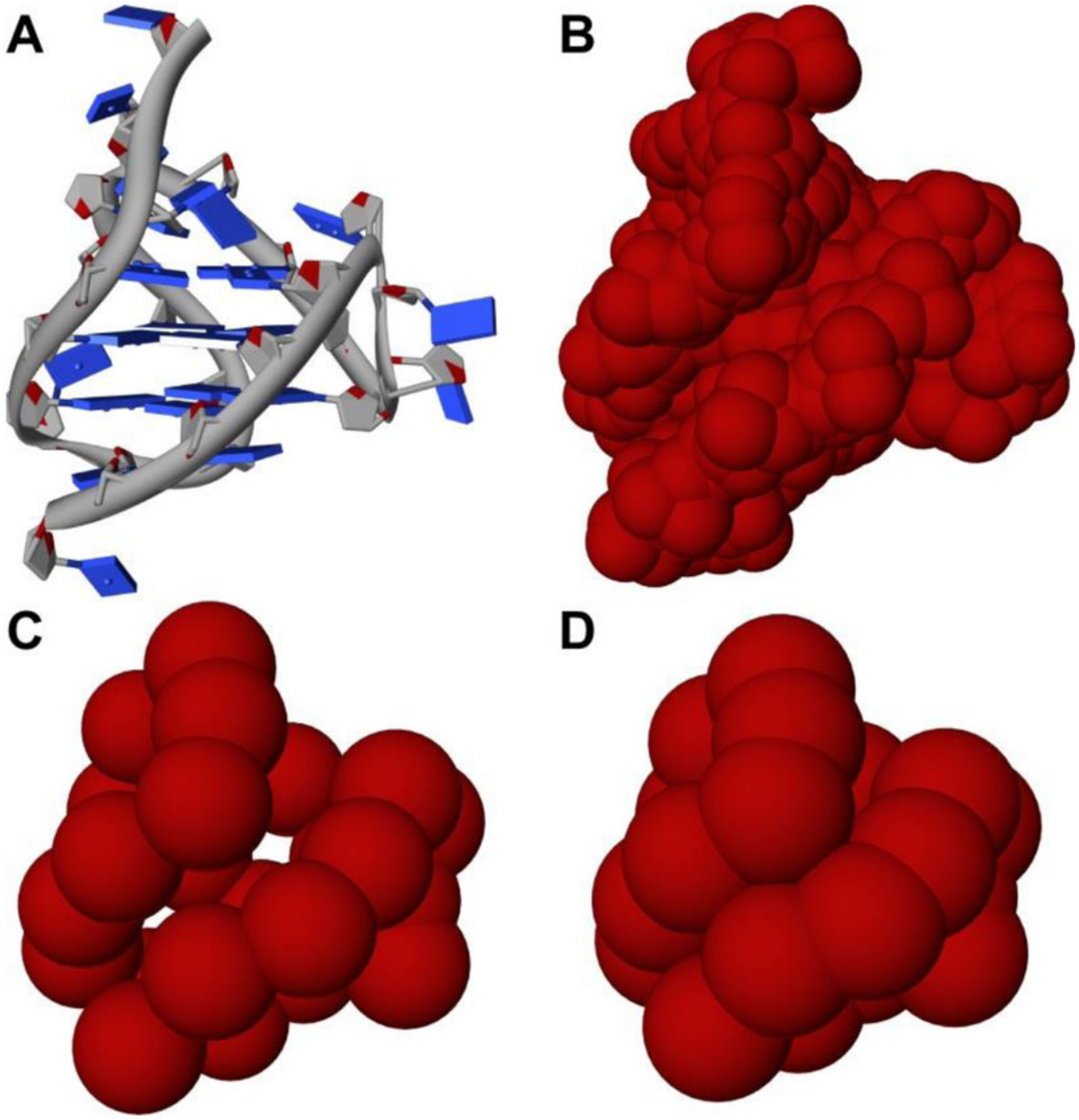

Fig. 2. HYDROPRO Generated Bead Models

The G-quadruplex structure (PDB: 2JPZ) formed from the human telomere sequence, d(T2AG3)T2, is shown as a sample of the structures employed for the HYDROPRO calculations. The G-quadruplex is represented (A) as a ribbon model with rectangular blocks representing the bases. The 5′ end is marked green while the 3′ end is marked red. The primary hydrodynamic model is displayed for (B) atomic-level shell-model calculation mode, (C) residue-level shell-model calculation mode, and (D) residue-level bead-model calculation mode.

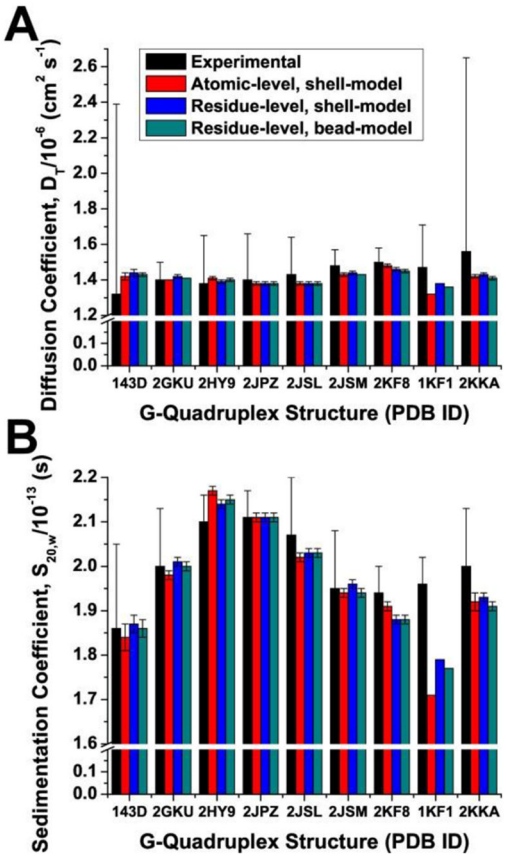

Fig. 3. Summary of HYDROPRO Calculations

The translational diffusion coefficients (A) and sedimentation coefficients (B) for the G-quadruplex structures formed from the human telomere sequence calculated by HYDROPRO using the indicated calculation method and compared to the experimental values. The error bars represent the 95% confidence interval. For experimental values, the errors represent the difference between experimental values determined by the DCDT+ software package compared to the Sedfit software package. For calculated values, the errors represent the difference between values calculated from the different solution structures reported. For 1KF1, there was no error for the calculated value since only one crystal structure was reported.

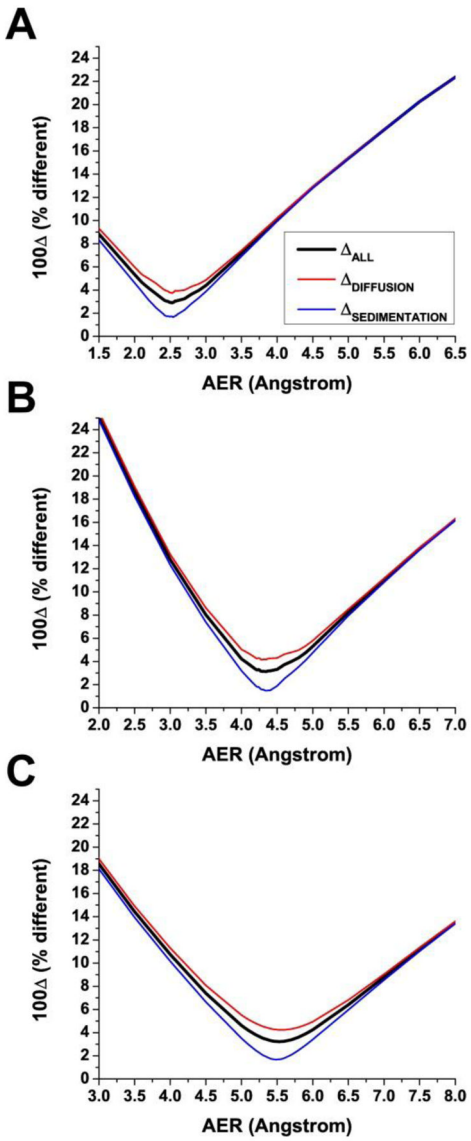

Fig. 4. Global Fitting of HYDROPRO Parameters

The values for 100ΔX and 100Δ as a function of the atomic element radius (AER) for the primary hydrodynamic model as calculated using the seven G-quadruplex structures formed from the human telomere sequence. Hydrodynamic properties of G-quadruplexes were calculated using the atomic-level hydrodynamic shell-model calculation (A), residue-level shell-model calculation (B), and residue-level bead-model calculation (C).

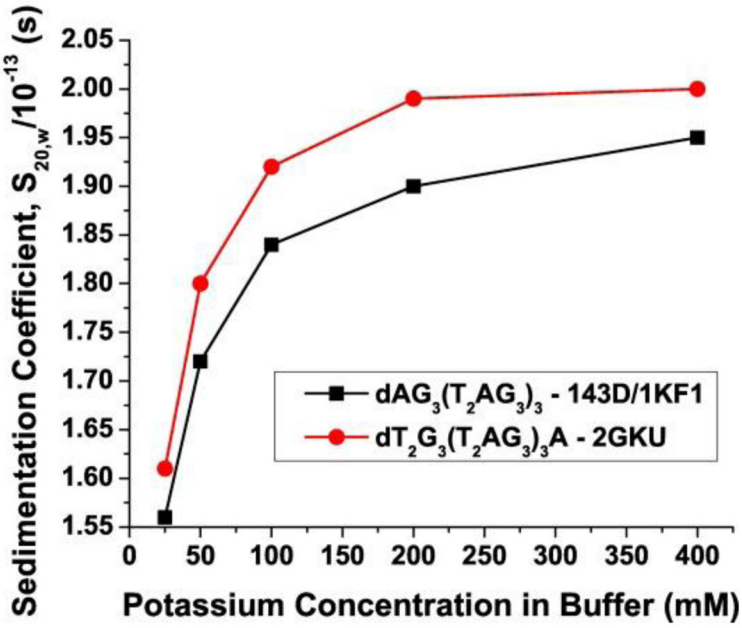

Figure 5.

The experimentally determined value for sedimentation coefficient is dependent on amount of potassium in solution. This dependency was demonstrated for two different sequence $dAG_3(T_2AG_3)_3$ – 22-mer, the same sequence used to derive an X-ray crystal structure in K^+ (1KF1) and an NMR solution structure in Na^+ (143D), and $dT_2G_3(T_2AG_3)A$ – the sequence used to derive an NMR solution structure in K^+ (2GKU).

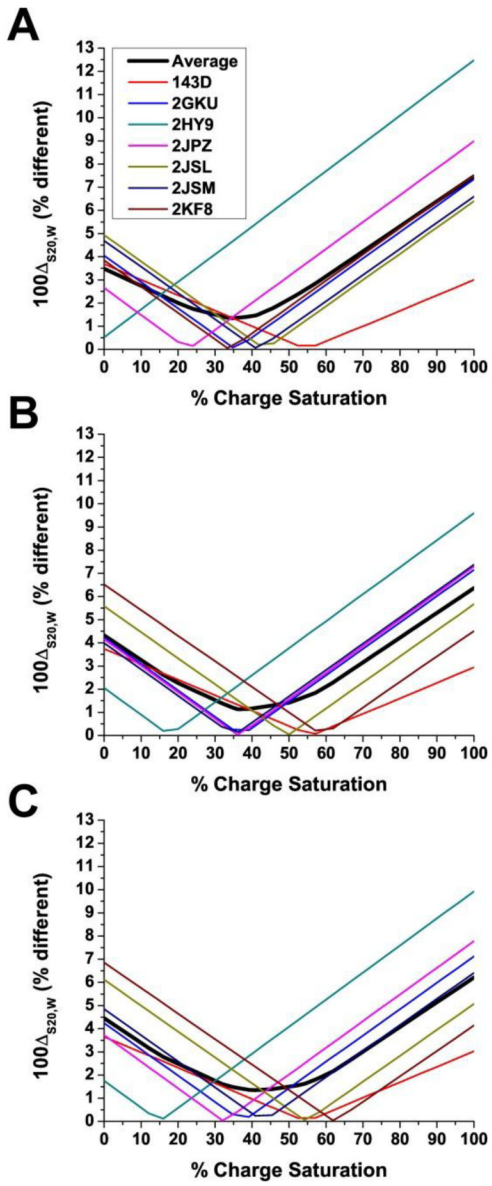

Fig. 6. Results of HYDROPRO Parameters Calibration by Number of Potassium Bound

The values for $100\Delta S_{20,W}$ as a function of percent charge saturation for the primary hydrodynamic model calculated using the seven G-quadruplex structures formed from the human telomere sequence. Hydrodynamic properties of G-quadruplexes were calculated using atomic-level shell-model calculation (A), residue-level shell-model calculation (B), and residue-level bead-model calculation (C).

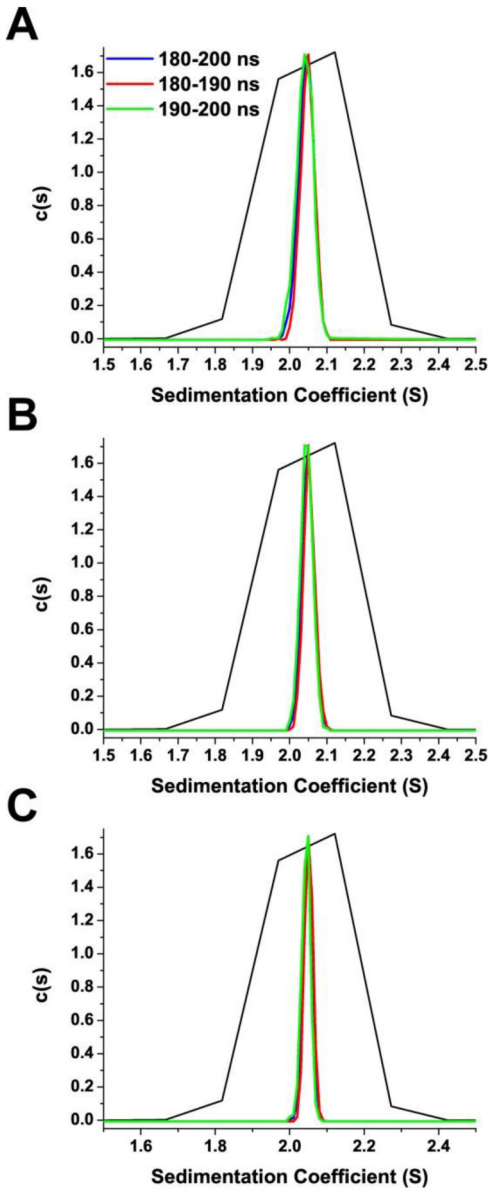

Fig. 7. Results of Molecular Dynamics and HYDROPRO Calculations

The distribution of the sedimentation coefficient, c(s), for the hybrid-1 G-quadruplex structure 2HY9 determined experimentally by sedimentation velocity (black) as compared to the distribution of sedimentation coefficient computed by HYDROPRO following molecular dynamics simulation. The sedimentation coefficient was computed for each snapshot taken every 10 ps between the 180–200 ns (blue), 180–190 ns (red), and 190–200 ns (green) interval of the trajectory. Hydrodynamic properties of G-quadruplexes were calculated using atomic-level shell-model calculation (A), residue-level shell-model calculation (B), and residue-level bead-model calculation (C).

Table 1

G-Quadruplex-Forming Sequences for HYDROPRO Calculations

Sequence	Buffer Condition	Corresponding Structure (PDB ID)
1. $dAG_3(T_2AG_3)_3$	10 mM tetrabutylammoniumphosphate monobasic, 400 mM sodium chloride, pH 7.0	143D
2. $dAG_3(T_2AG_3)_3$	10 mM tetrabutylammoniumphosphate monobasic, 400 mM potassium chloride, pH 7.0	1KF1
3. $dT_2G_3(T_2AG_3)_3A$	10 mM tetrabutylammoniumphosphate monobasic, 400 mM potassium chloride, pH 7.0	2GKU
4. $dA_3G_3(T_2AG_3)_3A_2$	10 mM tetrabutylammoniumphosphate monobasic, 400 mM potassium chloride, pH 7.0	2HY9
5. $dT_2AG_3(T_2AG_3)_3T_2$	10 mM tetrabutylammoniumphosphate monobasic, 400 mM potassium chloride, pH 7.0	2JPZ
6. $dTAG_3(T2AG3)_3T_2$	10 mM tetrabutylammoniumphosphate monobasic, 400 mM potassium chloride, pH 7.0	2JSL
7. $dTAG_3(T_2AG_3)_3$	10 mM tetrabutylammoniumphosphate monobasic, 400 mM potassium chloride, pH 7.0	2JSM
8. $dG_3(T_2AG_3)_3T$	10 mM tetrabutylammoniumphosphate monobasic, 400 mM potassium chloride, pH 7.0	2KF8
9. $dAG_3(T_2AG_3)_3T$	10 mM tetrabutylammoniumphosphate monobasic, 400 mM potassium chloride, pH 7.0	2KKA

Table 2

Summary of HYDROPRO Calculated Hydrodynamics Values

G-Quadruplex (PDB ID)	Experimental*	Atomic, shell model	Residue/Cα, shell model	Residue/Cα, bead model
143D	D_t = 1.32	1.42	1.44	1.43
	$S_{20,w}$ = 1.86	1.84	1.87	1.86
2GKU	D_t = 1.40	1.40	1.42	1.41
	$S_{20,w}$ = 2.00	1.98	2.01	2.00
2HY9	D_t = 1.38	1.41	1.39	1.40
	$S_{20,w}$ = 2.10	2.17	2.14	2.15
2JPZ	D_t = 1.40	1.38	1.38	1.38
	$S_{20,w}$ = 2.11	2.11	2.11	2.11
2JSL	D_t = 1.43	1.38	1.38	1.38
	$S_{20,w}$ = 2.07	2.02	2.03	2.03
2JSM	D_t = 1.48	1.43	1.44	1.43
	$S_{20,w}$ = 1.95	1.94	1.96	1.94
2KF8	D_t = 1.50	1.48	1.46	1.45
	$S_{20,w}$ = 1.94	1.91	1.88	1.88
1KF1	D_t = 1.47	1.32	1.38	1.36
	$S_{20,w}$ = 1.96	1.71	1.79	1.77
2KKA	D_t = 1.56	1.42	1.43	1.41
	$S_{20,w}$ = 2.00	1.92	1.93	1.91

*Dt is expressed in units of 10^{-6} cm^2 s^{-1}; S20,w in units of 10^{-13} s; 1KF1 and 2KKA were not used in global analysis of best fit AER for primary hydrodynamic model.

Table 3

Results of HYDROPRO Global Fit Analysis

Calculation Mode	AER (Å)	% diff.	% diff.$_{S20,w}$	% diff.$_{D}$
Atomic, shell model	2.53	2.90	1.71	3.73
Atomic, shell model	2.84	3.77	3.02	4.39
Residue/Cα, shell model	4.35	3.12	1.47	4.16
Residue/Cα, shell model	4.84	4.44	3.74	5.04
Residue/Cα, bead mode	5.54	3.22	1.70	4.23
Residue/Cα, bead mode	6.11	4.67	3.96	5.29

Table 4

Results of Hydration Analysis

Calculation Mode	AER (Å)	Hydration (g/g)
Atomic, shell model	2.53	0.27
Atomic, shell model	2.84	0.38
Residue/Cα, shell model	4.35	0.04
Residue/Cα, shell model	4.84	0.21
Residue/Cα, bead mode	5.54	0.36
Residue/Cα, bead mode	6.11	0.56

Table 5

Results of 2HY9 Molecular Dynamics Calculations

Calculation Mode	$S_{20,w}$ = 2.10	% diff.
Atomic, shell model	2.040	2.65
Residue/Cα, shell model	2.044	2.41
Residue/Cα, bead mode	2.043	2.48