International Journal of
Molecular Sciences

MDPI

Article

Supramolecular Polymorphism of $(G_4C_2)_n$ Repeats Associated with ALS and FTD

Melani Potrč [1,2,†], Nerea Sebastián [2,†], Miha Škarabot [3], Irena Drevenšek-Olenik [2,4] and Lea Spindler [2,5,*]

1 Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška 160, 2000 Maribor, Slovenia; melani.potrc@student.um.si
2 Department of Complex Matter, Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia; nerea.sebastian@ijs.si (N.S.); irena.drevensek@ijs.si (I.D.-O.)
3 Department of Condensed Matter Physics, Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia; miha.skarabot@ijs.si
4 Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia
5 Faculty of Mechanical Engineering, University of Maribor, Smetanova 17, 2000 Maribor, Slovenia
* Correspondence: lea.spindler@um.si
† These authors contributed equally to the work.

Abstract: Guanine-rich DNA sequences self-assemble into highly stable fourfold structures known as DNA-quadruplexes (or G-quadruplexes). G-quadruplexes have furthermore the tendency to associate into one-dimensional supramolecular aggregates termed G-wires. We studied the formation of G-wires in solutions of the sequences $d(G_4C_2)_n$ with n = 1, 2, and 4. The $d(G_4C_2)_n$ repeats, which are associated with some fatal neurological disorders, especially amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), represent a challenging research topic due to their extensive structural polymorphism. We used dynamic light scattering (DLS) to measure translational diffusion coefficients and consequently resolve the length of the larger aggregates formed in solution. We found that all three sequences assemble into longer structures than previously reported. The $d(G_4C_2)$ formed extremely long G-wires with lengths beyond 80 nm. The $d(G_4C_2)_2$ formed a relatively short stacked dimeric quadruplex, while $d(G_4C_2)_4$ formed multimers corresponding to seven stacked intramolecular quadruplexes. Profound differences between the multimerization properties of the investigated sequences were also confirmed by the AFM imaging of surface films. We propose that π-π stacking of the basic G-quadruplex units plays a vital role in the multimerization mechanism, which might be relevant for transformation from the regular medium-length to disease-related long $d(G_4C_2)_n$ repeats.

Keywords: DNA-quadruplex; G-wires; d(GGGGCC) repeats; self-assembly; dynamic light scattering; AFM

Citation: Potrč, M.; Sebastián, N.; Škarabot, M.; Drevenšek-Olenik, I.; Spindler, L. Supramolecular Polymorphism of $(G_4C_2)_n$ Repeats Associated with ALS and FTD. *Int. J. Mol. Sci.* **2021**, *22*, 4532. https://doi.org/10.3390/ijms22094532

Academic Editor: Aldo Galeone

Received: 6 April 2021
Accepted: 23 April 2021
Published: 26 April 2021

1. Introduction

Although DNA sequences forming guanosine quadruplexes (G-quadruplexes) have been extensively studied for several decades, new folding patterns in widely investigated strands are still being discovered [1,2]. Therefore, the conformational polymorphism of G-quadruplex structures remains an important research challenge. Quite frequently, in standard conditions, several quadruplex folds simultaneously coexist in a solution, which makes a precise determination of different folding topologies very difficult. Such behavior is also typical for GGGGCC (G_4C_2) repeat sequences associated with neurological disorders amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) [3–6]. The most common mutation related to these diseases is an increased number of $d(G_4C_2)$ repeats within the *C9orf72* gene. Individuals with ALS and FTD generally have hundreds to thousands of G_4C_2 repeat units, while healthy individuals typically have less than 25 of them [7,8]. The molecular mechanisms through which these repeat expansions and their

quadruplex structures cause the neurodegenerative modifications are not yet known. Nevertheless, some potential therapeutics targeting them are already being investigated [9,10].

A recent study indicated that, in addition to their conformational diversity, G-quadruplex structures from $d(G_4C_2)$ repeats are prone to supramolecular assembly beyond the simple G-quadruplex formation [4]. This feature brings additional difficulties to the experimental determination of the exact G-quadruplex folding topologies. To tackle this problem, different approaches were developed, like 8Br-dG substitution [11] or anion exchange chromatography [5]. Only a few of the possible G-quadruplex folds could be pinned down. For $d(G_4C_2)$, the basic building block is a parallel tetrameric G-quadruplex (Figure 1a) [4]. For $d(G_4C_2)_2$, several folding topologies are possible, but most of them are dimeric quadruplexes (Figure 1b) [5]. For $d(G_4C_2)_4$, a monomolecular antiparallel quadruplex was identified as the predominant species (Figure 1c) [11,12]. A possible role of $d(G_4C_2)$ supramolecular assembly, in particular multimerization of the associated G-quadruplexes, in the development of ALS and FTD is, however, still a fully open problem.

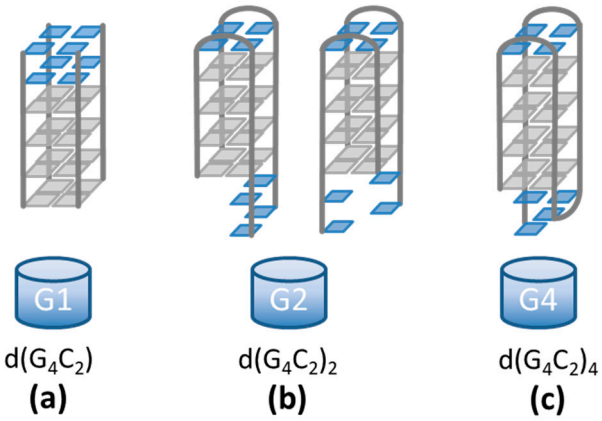


Figure 1. $d(G_4C_2)_n$ sequences form quadruplex structures in which four guanines (gray squares) associate via Hoogsteen-type hydrogen bonding. (**a**) The $d(G_4C_2)$ sequence forms a tetrameric symmetric quadruplex with 3′-CC overhangs (blue squares) [4]. (**b**) The $d(G_4C_2)_2$ sequence shows a high polymorphism with the exact structures still being unknown [5]. Two possible dimeric folds are depicted. (**c**) The $d(G_4C_2)_4$ sequence is believed to form a unimolecular antiparallel quadruplex with edgewise loops [11–13].

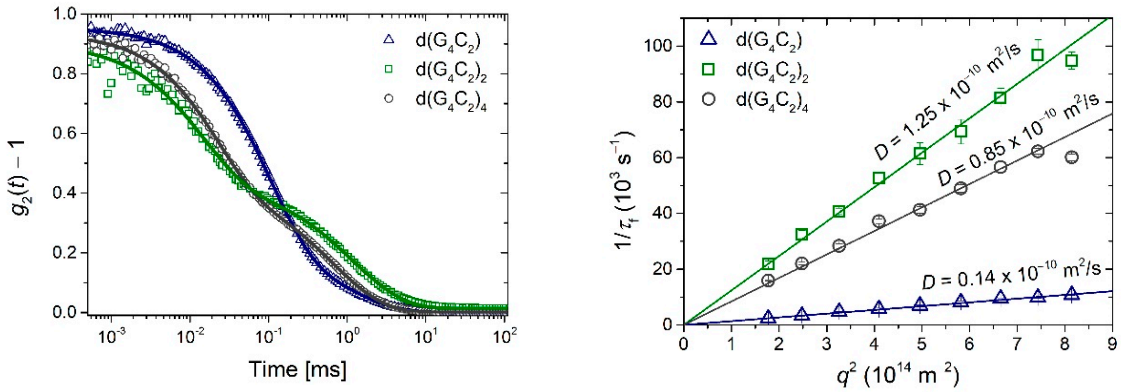
This work is focused on the investigation of supramolecular structures formed in solutions of $d(G_4C_2)$, $d(G_4C_2)_2$, and $d(G_4C_2)_4$. For this purpose, we used two methods that are not routinely used in DNA-quadruplex research, namely dynamic light scattering (DLS) and atomic force microscopy (AFM). Both of them give information on the size and shape of the objects present in the system, while they cannot resolve atomistic details of their internal structure. With DLS measurements, we determined the length of the larger $d(G_4C_2)_n$ aggregates in the solution, while with AFM, we tested the presence of such macromolecular assemblies in the drop cast films deposited from the solution onto a solid substrate. The obtained results revealed a surprising ability of the shortest of the three investigated sequences, i.e., $d(G_4C_2)$, to form remarkably long G-quadruplex-based assemblies associated with the so-called guanine wires (G-wires) [14]. In contrast, the other two sequences, namely $d(G_4C_2)_2$ and $d(G_4C_2)_4$, formed much shorter assemblies.

2. Results

2.1. Dynamic Light Scattering

Although the three investigated sequences seem to be very similar, DLS measurements revealed their profoundly different self-assembly processes. Before the measurements, freshly prepared solutions were left to equilibrate for a week to make sure that they reached thermodynamically stable structures. After measuring the autocorrelation functions $g_2(t)$

(Figure 2a) at several scattering angles and fitting them to Equation (1) (see Materials and Methods section), the obtained inverse relaxation times of the fast mode $1/\tau_f$ were plotted versus q^2 (Figure 2b). The slope of the corresponding linear fits was used to deduce the diffusion coefficients in accordance with the Equation (2). Surprisingly, the shortest sequence, namely d(G_4C_2), exhibited the lowest value of $D = (0.14 \pm 0.02) \cdot 10^{-10}$ m^2/s, indicating extremely long rod-like assemblies. The sequence d(G_4C_2)$_2$ exhibited the highest value of $D = (1.25 \pm 0.02) \cdot 10^{-10}$ m^2/s, while the sequence d(G_4C_2)$_4$ exhibited an intermediate value of $D = (0.85 \pm 0.02) \cdot 10^{-10}$ m^2/s. The measurements were also repeated a week later, and, within the experimental error, the same values of the diffusion coefficients were obtained. To estimate the length distribution of aggregates an average stretch factor s for all scattering angles was calculated. For completely monodisperse scatterers the expected value is $s = 1$, while lower s values are typical for more polydisperse objects. We obtained $s = 0.86 \pm 0.01$ for d(G_4C_2), $s = 0.98 \pm 0.02$ for d(G_4C_2)$_2$ and $s = 0.90 \pm 0.02$ for d(G_4C_2)$_4$ indicating quite narrow size distributions for all three sequences.



(**a**) (**b**)

Figure 2. DLS measurements: (**a**) Autocorrelation curves taken at $\vartheta = 90°$ clearly reveal very slow dynamics of the fast mode in solution of the d(G_4C_2) sequence. (**b**) The measured dependencies of the inverse relaxation time of the fast mode $1/\tau_f$ versus q^2 were used to calculate the translational diffusion coefficients D (Equation (2)). Quite notably, there is no evident correlation between the obtained values of D and the number of G_4C_2 repeats in the investigated sequence, which indicates that these sequences exhibit very diverse self-assembly patterns.

Additional information on the assembly process was obtained by denaturation of the solutions and subsequent monitoring of the DLS signal. Before DLS measurements, the denatured samples were left for one hour to cool down to the ambient temperature. The results are shown in Figure 3. In solution of d(G_4C_2), long assemblies started to form already during the cooling process, and their growth continued for another 4 h, with the final value of D getting below $0.2 \cdot 10^{-10}$ m^2/s. In contrast, in solution of d(G_4C_2)$_4$, the observed value of $D \approx 0.9 \cdot 10^{-10}$ m^2/s was practically constant during the first 6 h. Afterward, it very slowly decreased with the increasing time, gradually reaching the final value of $D \approx 0.8 \cdot 10^{-10}$ m^2/s. In solution of d(G_4C_2)$_2$, the self-assembly process was rather slow, and it took more than 2 h before enough scattering objects were formed to produce a reasonable DLS signal. The initially detected very large value of $D \approx 1.6 \cdot 10^{-10}$ m^2/s indicates that at first very small folds were formed, which in the next 6 h developed into the final structures with $D \approx 1.2 \cdot 10^{-10}$ m^2/s.

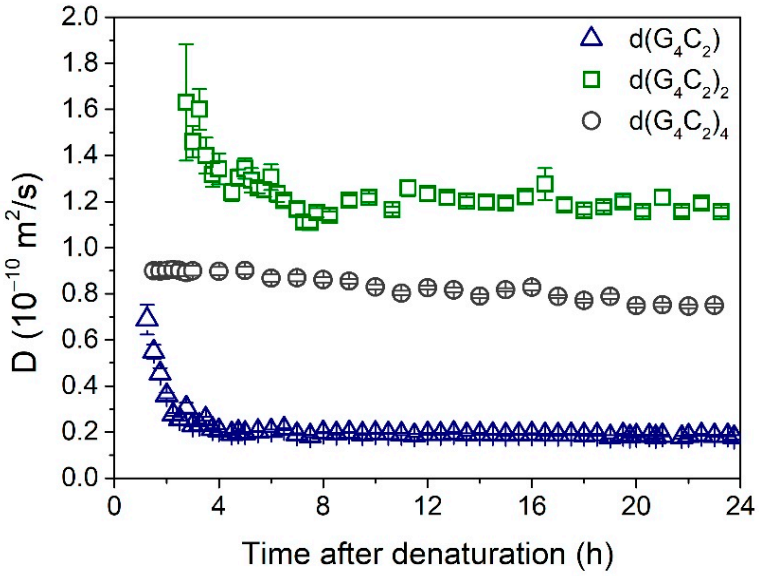


Figure 3. Time dependences of the diffusion coefficient after thermal denaturation (performed at $t = 0$ h). The highest values of D observed in solution of $d(G_4C_2)_2$ indicate the presence of small structures, most probably two stacked bimolecular quadruplexes. The intermediate values of D observed in solution of $d(G_4C_2)_4$ are probably associated with higher-order quadruplexes. The lowest values of D observed in solution of $d(G_4C_2)$ signify a formation of extremely long G-wires created by stacking of individual parallel quadruplex units.

2.2. Atomic Force Microscopy

AFM was used to additionally test the extent of the self-assembly of the three investigated sequences and especially to confirm the ability of the shortest hexanucleotide $d(G_4C_2)$ to form extremely long G-wires. AFM images of the films formed from the solutions of $d(G_4C_2)_2$ and $d(G_4C_2)_4$ reveal numerous small globular islands with variable sizes (Figure 4). All islands have a very similar height of $h = (2.4 \pm 0.4)$ nm for $d(G_4C_2)_2$ and $h = (2.3 \pm 0.6)$ nm for $d(G_4C_2)_4$. Although this height is smaller than the actual G-quadruplex diameter in solution and also smaller than the G-wire height detected in solution AFM imaging [15], it is the usual value found in AFM imaging of air-dried G-quadruplexes on mica substrates [16,17]. While in AFM imaging, the height of the structures can be measured with a $\pm$ 0.1 nm precision, the lateral spatial resolution is much worse due to the tip curvature. Similar surface structures, however, were observed in the films formed by parallel tetrameric G-quadruplexes from the *Tetrahymena* telomeric repeat sequence $d(TG_4T)$ [18] and the G-rich oligonucleotides $d(TG_8T)$ and $d(TG_9)$ [19].

In addition to spherical islands, surface structures from $d(G_4C_2)$ also revealed the formation of long wire-like assemblies with lengths spanning from few nanometers up to around one hundred nanometers (Figure 5). The actual G-wire lengths were around 15 nm smaller than seen by AFM imaging due to the finite tip size effect [20]. Despite large differences in their lengths, all wires showed the same relatively homogeneous height of $h = (2.3 \pm 0.6)$ nm. This height is typical for G-wires lying flat on the mica surface and being buttressed by magnesium ions [21–23]. The fact that $d(G_4C_2)_2$ and $d(G_4C_2)_4$ did not form such G-wires, is in agreement with their limited stacking ability in solution as witnessed by the DLS measurements.

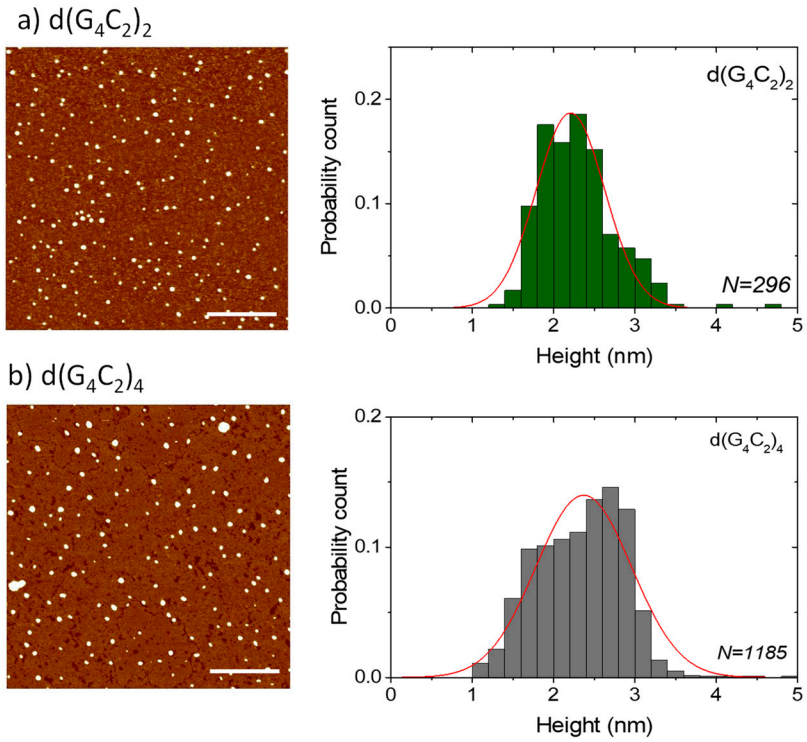


Figure 4. AFM images of small globular islands formed in drop-cast surface films of (**a**) $d(G_4C_2)_2$ and (**b**) $d(G_4C_2)_4$. The height distribution of the structures is given in the right panels and shows an average height of around 2.4 nm. The lines are guides to the eye. The scale bars in the AFM images correspond to 500 nm.

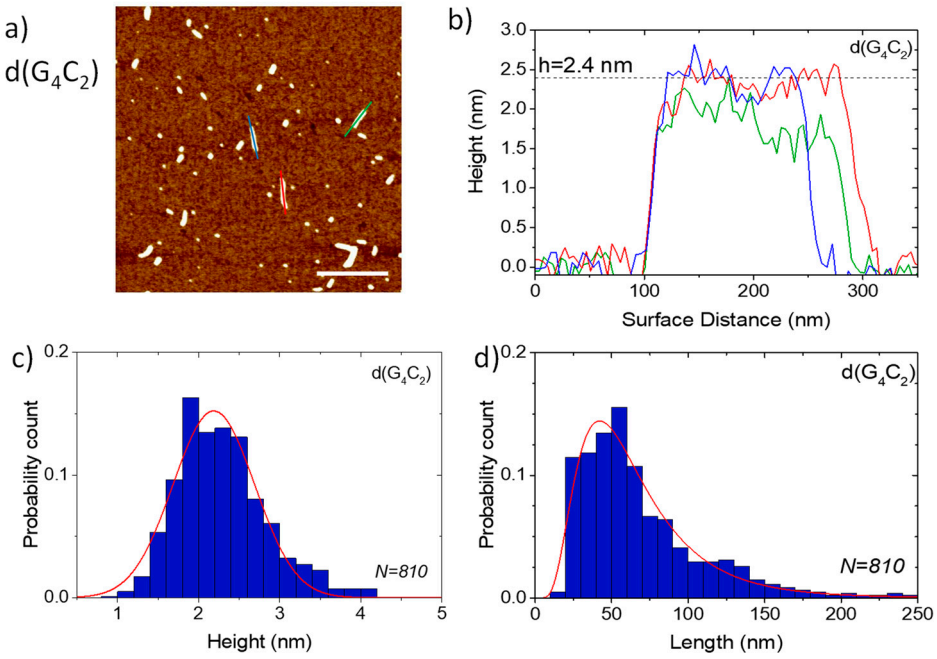


Figure 5. (**a**) AFM image of a drop cast film of $d(G_4C_2)$. The scale bar corresponds to 500 nm. (**b**) The height profiles along three single G-wires (see colored lines in (**a**)). (**c**) The length and (**d**) the height distribution of G-wires. The lines are guides to the eye. Length analysis was limited to objects above 1 nm and elongated in shape.

3. Discussion

Combining the obtained DLS and AFM results with the previously published work on $d(G_4C_2)_n$ repeats [4,5,11,12,24–26], we propose the following models for their self-assembly.

3.1. Self-Assembly of d(G4C2)

For the $d(G_4C_2)$ hexanucleotide the basic structural unit was proposed to be a tetramolecular parallel G-quadruplex [4] (Figure 1a). The NMR spectra for this sequence already indicated formation of larger assemblies, but with NMR it was not possible to comprehend the extent of the stacking process. As it turns out, the $d(G_4C_2)$ hexanucleotide forms exceptionally long G-wires, that have never been reported for this sequence before. The diffusion coefficient $D = (0.14 \pm 0.02) \cdot 10^{-10}$ m^2/s is outside the limits of the Tirado and Garcia de la Torre theory ($L/d > 30$). The length of the aggregates in solution is believed to be larger than ~80 nm, which corresponds to more than ~50 stacked basic G-quadruplex units (Figure 6a). Such results are supported by the AFM findings showing long wire-like structures up to one hundred nanometers. Although stacking of G-quadruplexes was reported for a number of G-rich DNA sequences [27–30], such extremely long aggregates have been until now observed only for 5′-dGMP in the form of diammonium salt [31,32], for the *Tetrahymena thermophila* telomeric sequence $d(G_4T_2)_4$, and Human telomeric sequence $d(TTAG_3)_4$ in the presence of Sr^{2+} ions [33,34]. From our measurements, the exact mechanism of the stacking cannot be resolved, but there exist a few possible scenarios. Šket et al. [4] proposed stacking of two basic units through 5′-5′ interfaces. Such dimerization through cofacial stacking was indeed proven to be energetically favorable to 5′-3′ or 3′-3′ stacking [35]. This dimerized structure, however, is flanked by four CC overhangs on each end. In order to continue the stacking process, these CC overhangs need to re-arrange themselves. One possibility is the exclusion of the 3′ flanking overhangs [36]. Another possibility is the formation of a CCCC tetrad that is formed above the last G-tetrad and stabilized by π-π stacking [26]. The stacking of two such loosely formed end surfaces, however, does not seem very probable. An alternative possibility is the energetically less favorable 5′-3′ stacking (head-to-tail) [35] of single G-quadruplexes that could give more stable long aggregates beyond the dimer formation. Whatever the stacking mechanism is, our measurements show that the final G-wires resulting from it are extremely long.

3.2. Self-Assembly of d(G4C2)2

DLS data show that $d(G_4C_2)_2$ repeats have the highest diffusion coefficient of the three studied repeats (Figure 2b) and thus form the shortest aggregates. The first structures detected after denaturation (Figure 3) have a diffusion coefficient of $D \approx 1.6 \cdot 10^{-10}$ m^2/s. This value is outside the limits of the TGT theory, but, based on previous work [37] it is associated with quadruplex folds having four stacked G-quartets. The structures forming soon after the denaturation are therefore most likely dimeric G-quadruplexes (Figure 1b). Their exact structure is up to date still unknown since G-quadruplexes from $d(G_4C_2)_2$ exhibit a number of different topologies and are a mix of parallel and antiparallel conformations that are simultaneously existing in the solution. The equilibrium assembly of $d(G_4C_2)_2$ is reached 6 h after denaturation and has a value of $D = (1.25 \pm 0.02) \cdot 10^{-10}$ m^2/s. There are two possible structures that could correspond to such diffusion coefficients: either two stacked dimeric quadruplexes (higher-order structure) or a symmetric tetramer. Both were proven to exist together with the dimeric form in $d(G_4C_2)_2$ repeat solutions [5]. However, for the parallel tetrameric form it is expected that is should easily form G-wires, which were not observed by the AFM imaging. Thus, the stacked dimeric quadruplex remains the most likely option (Figure 6b) for being the largest aggregate of self-assembled $d(G_4C_2)_2$ repeats, while the presence of other, smaller, topologies cannot be excluded based on our DLS data.

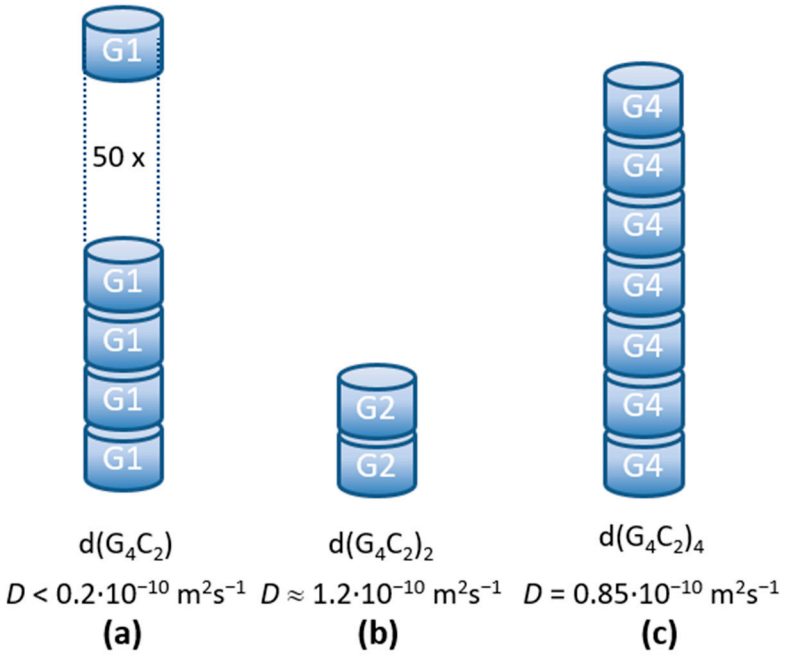


Figure 6. The model for the largest diffusing structures in the solutions [38] as deduced from the DLS data. The quadruplexes from $d(G_4C_2)$, $d(G_4C_2)_2$, and $d(G_4C_2)_4$ show very different stacking ability with $d(G_4C_2)$ forming the longest (**a**), $d(G_4C_2)_2$ the shortest (**b**) and $d(G_4C_2)_4$ intermediate sized G-wires (**c**). All aggregates determined by DLS are longer than previously reported with other methods.

3.3. Self-Assembly of $d(G_4C_2)_4$

Finally, the aggregates from $d(G_4C_2)_4$ repeats exhibit translational dynamic features in between the two smaller sequences. Their size is the only one that fits into the TGT theory limits with the measured value of $D \approx 0.9 \cdot 10^{-10}$ m^2/s. This result is rather surprising, since several papers reported that $d(G_4C_2)_4$, as well as $r(G_4C_2)_4$, repeats formed a small unimolecular G-quadruplex [4,11–13,24,39]. Only Reddy et al. [24] reported for $r(G_4C_2)_4$ repeats that they could also form electrophoretically slow migrating structures and confirmed their multimolecular origin. These multimers are most likely stacked monomolecular quadruplexes. Based on the NMR studies, the $d(G_4C_2)_4$ sequence in K^+ solutions forms an antiparallel quadruplex with 4 G-quartet planes (Figure 1c). From the TGT theory we can estimate the number of stacked units using Equation (3) with $D = 0.85 \cdot 10^{-10}$ m^2/s (Figure 2b) and consequently obtain the length of the structures to be $L = 10$ nm. With an average stacking height of 0.34 nm this yields 29 G-quartet planes, which corresponds to around seven stacked monomolecular quadruplexes (Figure 6c). Although this finding is a bit surprising, similar higher-order structures were previously reported also for a number of other G-quadruplex forming sequences investigated by DLS measurements and other methods including theoretical modeling [33,34,40–42]. We should also note that the diffusion coefficient $D = 0.85 \cdot 10^{-10}$ m^2/s and the corresponding dimensions of the aggregates could be associated with a parallel tetrameric G-quadruplex [34]. All previous work on this sequence, however, indicates that such a tetrameric G-quadruplex in an unlikely fold.

In general, DLS results for $d(G_4C_2)_2$ and $d(G_4C_2)_4$ do not rule out the simultaneous existence of smaller, monomolecular, or bimolecular, G-quadruplexes in solution. Unfortunately, AFM imaging for these two sequences did not give more insight into the polymorphism of these smaller structures due to the limitations of AFM imaging in the air [15].

4. Materials and Methods

4.1. Materials

All three oligonucleotides (d(G_4C_2), d$(G_4C_2)_2$, and d$(G_4C_2)_4$) were purchased in lyophilized form from Eurogentec (Seraing, Belgium), purified by HPLC-RP, and desalted on a Sephadex G25 column. The material was dissolved in 100 mM KCl to a final strand concentration of 1 mM, and the aliquots of about 300 µL were put into 5 mm diameter glass capillaries. The denaturation was performed by heating the samples to 95 °C and holding the temperature constant for 5 min. Afterward, the samples were left to slowly cool down to room temperature.

4.2. Dynamic Light Scattering

Dynamic light scattering (DLS) measurements were performed using either a He-Ne laser operating at λ = 633 nm or a frequency-doubled Nd:YAG laser operating at 532 nm as the light source. For detection of scattered light, a digital correlator (ALV-7002 Multiple Tau) in combination with an avalanche photodiode was used. The scattered light was detected at scattering angles θ ranging from 30 to 130°.

The theoretical background and the experimental setup were described in detail in our previous work [43]. In the majority of the measurements, two diffusive dynamic modes were observed that are referred to as the fast and the slow mode. The measured intensity autocorrelation functions were fitted to the following relation [44]:

$$g_2(t) - 1 = \left(1 + j_d\left(a_f \exp\left(-\left(\frac{t}{\tau_f}\right)^{s_f}\right) + \left(1 - a_f\right)\exp\left(-\left(\frac{t}{\tau_s}\right)^{s_s}\right) - 1\right)\right)^2 + y_0, \quad (1)$$

where y_0 is the baseline correction, j_d is the ratio between the intensity of the inelastically scattered light and the intensity of total scattered light, a_f is the amplitude of the fast mode, and τ_f and τ_s are the relaxation times of the fast and the slow mode, respectively. The stretch exponent parameters, s_f and s_s, are associated with the relaxation times distribution and can attain values between 0 and 1.

The translational diffusion coefficient D was obtained from the relation:

$$D = \frac{1}{\tau q^2}, \quad (2)$$

where q is the scattering wave vector defined as $q = (4\pi n/\lambda)\sin(\theta/2)$, n = 1.33 is the solution refractive index, and λ the laser wavelength. From the two observed dynamic modes, the fast mode was associated with the diffusion of G-quadruplexes and their multimerized assemblies, while the slow mode was assigned to the diffusion of large, unspecific "clusters". This mode is generally known as the »slow mode« of polyelectrolyte solutions and was observed in solutions of various polyelectrolytes, including quadruplex-forming G-rich DNA sequences [45–47]. However, the associated loose clusters have dimensions in the micrometer range and are believed not to be directly connected to the G-quadruplex formation.

To obtain the size of the diffusing objects from the measured value of D, an appropriate theoretical model is needed. For diluted solutions of rod-like particles, the hydrodynamic theory of Tirado and Garcia de la Torre (TGT theory) can be applied [48]. For the length to diameter ratio $p = L/d$ in the range $2 \leq p \leq 30$, the expression for D in this theory is given by the interpolation equation

$$D = \frac{k_B T}{3\pi\eta L}(\ln p + \nu), \quad (3)$$

where ν is the end-effect correction given as $\nu = 0.312 + 0.565/p + 0.100/p^2$, k_B the Boltzmann constant, T the solution temperature, and η the solvent viscosity. In the experiments, T was 296 K, and η was taken to be 0.932 mPa·s, which is solvent viscosity at this temperature. Previous NMR studies [4] have shown that the basic quadruplex unit of d(G_4C_2) does not

have any side loops, so the value of d = 2.6 nm that takes into account also the hydration sphere around such G-quadruplexes was used [43]. For this case, Equation (3) can be reasonably applied for scattering objects with the lengths in the range 5 nm $\leq L \leq$ 80 nm corresponding to the diffusion coefficients in the range $1.2 \cdot 10^{-10}$ m^2/s $\geq D \geq 0.22 \cdot 10^{-10}$ m^2/s.

If a solution contains scattering objects of different sizes, the DLS signal "favors" larger ones since the intensity of scattered light is proportional to the square of the scattering volume [38]. For rod-like particles with the same diameter d this gives the $\propto L^2$ dependence of the scattered intensity. In case of polymorphic structures in solution, DLS cannot resolve only one of possible coexisting topologies. However, because light scattering cross section increases with the square of the scattering volume, the contribution from larger scatterers predominates in the DLS signal. In addition, one also needs to take into consideration the fact that G-quadruplexes and their assemblies behave as strong polyelectrolytes, and consequently, electrostatic interactions influence their solution dynamics. However, based on previous work [33], we can assume that in 1 mM oligonucleotide solutions with the addition of 100 mM KCl the electrostatic interactions are strongly screened. Therefore, the values of L calculated by the use of the Equation (3) are very close to the actual length of the scattering objects.

4.3. Atomic Force Microscopy Imaging

Atomic force microscopy (AFM) imaging was performed in the tapping mode using the Nanoscope IIIa-MultiMode AFM device (Digital Instruments, Santa Barbara, CA, USA) and silicon cantilevers (Bruker OTESPA-R3) with the nominal resonance frequency of 300 kHz and a nominal tip radius of 7 nm. Oligonucleotide solutions for AFM imaging were prepared by diluting the solutions used for DLS measurements with an imaging buffer (10 mM KCl + 10 mM $MgCl_2$) to the concentrations of either 5 µM or 20 µM [21]. The diluted solution was drop-cast onto the substrate (freshly cleaved V-1 grade muscovite mica) and left to adsorb for 15 min. After this, the excess material was removed by washing the substrate with distilled water. Then, the sample was left to dry for at least one day at ambient conditions. AFM images were recorded on several different parts of the substrate. The obtained images were reproducible in the large areas. Particle height analysis was performed with the particle analysis option of the NanoScope Analysis software. To obtain the length distribution, the images were analyzed with the ImageJ program [49,50] and the Ridge Detection plugin [51]. The latter is based on the algorithm for detecting ridges and lines as described by Steger [52]. Both analysis methods were applied on a number of images and the cumulative statistics is presented. In the case of d(G_4C_2), length statistics were obtained by limiting the detection to elongated objects with a height equal or greater than 1 nm. To identify the artefacts from residual buffer salts in AFM images a set of measurements with only the buffering salt on mica was made. Drop cast films without water rinsing had circular structures of varying diameter and height (many of them much higher than 2 nm). After rinsing with distilled water, nearly all surface structures disappeared. We can therefore assume that circular structures observed in Figure 4 are not made of salt. Nevertheless, a small amount of salt could still be attached to the G-quadruplex structures and affect the AFM measurements.

5. Conclusions

Based on our DLS and AFM measurements no generalization of the expected self-assembly of d$(G_4C_2)_n$ repeats could be made. We found, however, that all the studied repeats form much larger aggregates than previously reported due to a strong tendency for π-π stacking of the basic quadruplex units. The assembly of numerous d$(G_4C_2)_n$ repeats, which are found in the defected gene of patients with ALS and FTD, could therefore be governed by the same assembling pattern: first the formation of basic, smaller G-quadruplexes, followed by stacking of these quadruplexes in a "beads on a string" fashion. Thus, DLS together with AFM was able to provide new and important insights into the folding scheme and supramolecular assembly of the highly polymorphic d$(G_4C_2)_n$ sequences.

Author Contributions: Conceptualization and Methodology, M.P., N.S., and L.S.; Investigation, M.P. and N.S.; Writing—original draft preparation, M.P., N.S., and L.S.; Writing—review and editing, N.S., M.Š., I.D.-O., and L.S.; Supervision, M.Š., I.D.-O. and L.S.; M.P. and N.S. contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Slovenian Research Agency within the Project Grant J7-9399 and research core funding Grant No. P1-0192.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, C.; Geng, Y.; Miao, H.; Shi, X.; You, Y.; Xu, N.; Zhou, B.; Zhu, G. G-quadruplex structures formed by human telomeric DNA and C9orf72 hexanucleotide repeats. *Biophys. Rev.* **2019**, *11*, 389–393. [CrossRef]
2. Le, H.T.; Dean, W.L.; Buscaglia, R.; Chaires, J.B.; Trent, J.O. An investigation of g-quadruplex structural polymorphism in the human telomere using a combined approach of hydrodynamic bead modeling and molecular dynamics simulation. *J. Phys. Chem. B* **2014**, *118*, 5390–5405. [CrossRef]
3. Kovanda, A.; Zalar, M.; Sket, P.; Plavec, J.; Rogelj, B. Anti-sense DNA d(GGCCCC)(n) expansions in C9ORF72 form i-motifs and protonated hairpins. *Sci. Rep.* **2015**, *5*, 17944. [CrossRef]
4. Sket, P.; Pohleven, J.; Kovanda, A.; Stalekar, M.; Zupunski, V.; Zalar, M.; Plavec, J.; Rogelj, B. Characterization of DNA G-quadruplex species forming from C9ORF72 G(4)C(2)-expanded repeats associated with amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* **2015**, *36*, 1091–1096. [CrossRef] [PubMed]
5. Zhou, B.; Geng, Y.; Liu, C.; Miao, H.; Ren, Y.; Xu, N.; Shi, X.; You, Y.; Lee, T.; Zhu, G. Characterizations of distinct parallel and antiparallel G-quadruplexes formed by two-repeat ALS and FTD related GGGGCC sequence. *Sci. Rep.* **2018**, *8*, 2366. [CrossRef]
6. Bozic, T.; Zalar, M.; Rogelj, B.; Plavec, J.; Sket, P. Structural Diversity of Sense and Antisense RNA Hexanucleotide Repeats Associated with ALS and FTLD. *Molecules* **2020**, *25*, 525. [CrossRef]
7. DeJesus-Hernandez, M.; Mackenzie, I.R.; Boeve, B.F.; Boxer, A.L.; Baker, M.; Rutherford, N.J.; Nicholson, A.M.; Finch, N.A.; Flynn, H.; Adamson, J.; et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **2011**, *72*, 245–256. [CrossRef]
8. Leko, M.B.; Župunski, V.; Kirincich, J.; Smilović, D.; Hortobágyi, T.; Hof, P.R.; Šimić, G. Molecular mechanisms of neurodegeneration related to C9orf72 hexanucleotide repeat expansion. *Behav. Neurol.* **2019**. [CrossRef]
9. Simone, R.; Balendra, R.; Moens, T.G.; Preza, E.; Wilson, K.M.; Heslegrave, A.; Woodling, N.S.; Niccoli, T.; Gilbert-Jaramillo, J.; Abdelkarim, S.; et al. G-quadruplex-binding small molecules ameliorate C9orf72 FTD/ALS pathology in vitro and in vivo. *Embo Mol. Med.* **2018**, *10*, 22–31. [CrossRef] [PubMed]
10. Ursu, A.; Wang, K.W.; Bush, J.A.; Choudhary, S.; Chen, J.L.; Baisden, J.T.; Zhang, Y.-J.; Gendron, T.F.; Petrucelli, L.; Yildirim, I.; et al. Structural features of small molecules targeting the RNA repeat expansion that causes genetically defined ALS/FTD. *ACS Chem. Biol.* **2020**, *15*, 3112–3123. [CrossRef] [PubMed]
11. Brcic, J.; Plavec, J. NMR structure of a G-quadruplex formed by four d(G(4)C(2)) repeats: Insights into structural polymorphism. *Nucleic Acids Res.* **2018**, *46*, 11605–11617.
12. Zhou, B.; Liu, C.; Geng, Y.; Zhu, G. Topology of a G-quadruplex DNA formed by C9orf72 hexanucleotide repeats associated with ALS and FTD. *Sci. Rep.* **2015**, 5. [CrossRef]
13. Brčić, J.; Plavec, J. ALS and FTD linked GGGGCC-repeat containing DNA oligonucleotide folds into two distinct G-quadruplexes. *Biochim. Biophys. Acta Gen. Subj.* **2017**, *1861*, 1237–1245. [CrossRef]
14. Marsch, T.C.; Henderson, E. G-wires: Self-assembly of a telomeric oligonucleotide, d(GGGGTTGGGG), into large superstructures. *Biochemistry* **1994**, *33*, 10718–10724. [CrossRef]
15. Bose, K.; Lech, C.J.; Heddi, B.; Phan, A.T. High-resolution AFM structure of DNA G-wires in aqueous solution. *Nat. Commun.* **2018**, *9*, 1959. [CrossRef]
16. Varizhuk, A.M.; Protopopova, A.D.; Tsvetkov, V.B.; Barinov, N.A.; Podgorsky, V.V.; Tankevich, M.V.; Vlasenok, M.A.; Severov, V.V.; Smirnov, I.P.; Dubrovin, E.V.; et al. Polymorphism of G4 associates: From stacks to wires via interlocks. *Nucleic Acids Res.* **2018**, *46*, 8978–8992. [CrossRef] [PubMed]
17. Mergny, J.-L.; Sen, D. DNA quadruple helices in nanotechnology. *Chem. Rev.* **2019**, *119*, 6290–6325. [CrossRef] [PubMed]
18. Rodrigues Pontinha, A.D.; Chiorcea-Paquim, A.-M.; Eritja, R.; Oliveira-Brett, A.M. Quadruplex nanostructures of d(TGGGGT): Influence of sodium and potassium ions. *Anal. Chem.* **2014**, *86*, 5851–5857. [CrossRef]
19. Chiorcea-Paquim, A.-M.; Santos, P.V.; Eritja, R.; Oliveira-Brett, A.M. Self-assembled G-quadruplex nanostructures: AFM and voltammetric characterization. *Phys. Chem. Chem. Phys.* **2013**, *15*, 9117–9124. [CrossRef]

20. Keller, D. Reconstruction of STM and AFM images distorted by finite-size tips. *Surf. Sci.* **1991**, *253*, 353–364. [CrossRef]
21. Vesenka, J. Preparation and atomic force microscopy of quadruplex DNA. In *DNA Nanotechnology: Methods and Protocols*; Zuccheri, G., Samori, B., Eds.; Methods in Molecular Biology; Springer: Cham, Switzerland, 2011; Volume 749, pp. 105–113.
22. Troha, T.; Drevensek-Olenik, I.; da Silva, M.W.; Spindler, L. Surface-adsorbed long g-quadruplex nanowires formed by G:C linkages. *Langmuir* **2016**, *32*, 7056–7063. [CrossRef] [PubMed]
23. Shankaraswamy, J.; Tyagi, S.; Singh, A.; Miyoshi, D.; Saxena, S. Metal sensitive and DNA concentration dependent structural rearrangement of short oligonucleotide into large suprastructures. *J. Biomol. Struct. Dyn.* **2019**, *37*, 2211–2218. [CrossRef]
24. Reddy, K.; Zamiri, B.; Stanley, S.Y.R.; Macgregor, R.B., Jr.; Pearson, C.E. The Disease-associated r(GGGGCC)(n) Repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J. Biol. Chem.* **2013**, *288*, 9860–9866. [CrossRef] [PubMed]
25. Haeusler, A.R.; Donnelly, C.J.; Periz, G.; Simko, E.A.J.; Shaw, P.G.; Kim, M.-S.; Maragakis, N.J.; Troncoso, J.C.; Pandey, A.; Sattler, R.; et al. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **2014**, *507*, 195. [CrossRef]
26. Zhang, Y.; Roland, C.; Sagui, C. Structural and dynamical characterization of DNA and RNA quadruplexes obtained from the GGGGCC and GGGCCT hexanucleotide repeats associated with C9FTD/ALS and SCA36 Diseases. *ACS Chem. Neurosci.* **2018**, *9*, 1104–1117. [CrossRef]
27. Parkinson, G.N.; Lee, M.P.H.; Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **2002**, *417*, 876–880. [CrossRef]
28. Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E.S.; Bowers, M.T.; De Pauw, E.; Gabelica, V. G-quadruplex DNA assemblies: Loop length, cation identity, and multimer formation. *J. Am. Chem. Soc.* **2008**, *130*, 10208–10216. [CrossRef]
29. Adrian, M.; Ang, D.J.; Lech, C.J.; Heddi, B.; Nicolas, A.; Phan, A.T. Structure and conformational dynamics of a stacked dimeric G-quadruplex formed by the human CEB1 minisatellite. *J. Am. Chem. Soc.* **2014**, *136*, 6297–6305. [CrossRef] [PubMed]
30. Do, N.Q.; Lim, K.W.; Teo, M.H.; Heddi, B.; Phan, A.T. Stacking of G-quadruplexes: NMR structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity. *Nucleic Acids Res.* **2011**, *39*, 9448–9457. [CrossRef] [PubMed]
31. Spindler, L.; Olenik, I.D.; Copic, M.; Romih, R.; Cerar, J.; Skerjanc, J.; Mariani, P. Dynamic light scattering and (31)P NMR spectroscopy study of the self-assembly of deoxyguanosine 5′-monophosphate. *Eur. Phys. J. E* **2002**, *7*, 95–102. [CrossRef]
32. Spindler, L.; Drevensek-Olenik, I.; Copic, M.; Cerar, J.; Skerjanc, J.; Mariani, P. Dynamic light scattering and P-31 NMR study of the self-assembly of deoxyguanosine 5′-monophosphate: The effect of added salt. *Eur. Phys. J. E* **2004**, *13*, 27–33. [CrossRef] [PubMed]
33. Wlodarczyk, A.; Grzybowski, P.; Patkowski, A.; Dobek, A. Effect of ions on the polymorphism, effective charge, and stability of human telomeric DNA. Photon correlation spectroscopy and circular dichroism studies. *J. Phys. Chem. B* **2005**, *109*, 3594–3605. [CrossRef]
34. Wlodarczyk, A.; Patkowski, A.; Grzybowski, P.; Dobek, A. The effect of mono- and divalent cations on Tetrahymena thermophila telomeric repeat fragment. A photon correlation spectroscopy study. *Acta Biochim. Pol.* **2004**, *51*, 971–981. [PubMed]
35. Kogut, M.; Kleist, C.; Czub, J. Why do G-quadruplexes dimerize through the 5′-ends? Driving forces for G4 DNA dimerization examined in atomic detail. *PLoS Comput. Biol.* **2019**, *15*, e1007383. [CrossRef] [PubMed]
36. Marzano, M.; Falanga, A.P.; Dardano, P.; D'Errico, S.; Rea, I.; Terracciano, M.; De Stefano, L.; Piccialli, G.; Borbone, N.; Oliviero, G. pi-pi stacked DNA G-wire nanostructures formed by a short G-rich oligonucleotide containing a 3′-3′ inversion of polarity site. *Org. Chem. Front.* **2020**, *7*, 2187–2195. [CrossRef]
37. Pavc, D.; Wang, B.; Spindler, L.; Drevensek-Olenik, I.; Plavec, J.; Sket, P. GC ends control topology of DNA G-quadruplexes and their cation-dependent assembly. *Nucleic Acids Res.* **2020**, *48*, 2749–2761. [CrossRef]
38. Berne, B.J.; Pecora, R. *Dynamic Light Scattering*; Wiley: New York, NY, USA, 1976.
39. Fratta, P.; Mizielinska, S.; Nicoll, A.J.; Zloh, M.; Fisher, E.M.C.; Parkinson, G.; Isaacs, A.M. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.* **2012**, *2*, 1016. [CrossRef]
40. Petraccone, L.; Spink, C.; Trent, J.O.; Garbett, N.C.; Mekmaysy, C.S.; Giancola, C.; Chaires, J.B. Structure and stability of higher-order human telomeric quadruplexes. *J. Am. Chem. Soc.* **2011**, *133*, 20951–20961. [CrossRef]
41. Rebic, M.; Mocci, F.; Laaksonen, A.; Ulicny, J. Multiscale simulations of human telomeric G-quadruplex DNA. *J. Phys. Chem. B* **2015**, *119*, 105–113. [CrossRef]
42. Tothova, P.; Krafcikova, P.; Viglasky, V. Formation of highly ordered multimers in G-quadruplexes. *Biochemistry* **2014**, *53*, 7013–7027. [CrossRef]
43. Ilc, T.; Sket, P.; Plavec, J.; da Silva, M.W.; Drevensek-Olenik, I.; Spindler, L. Formation of G-wires: The role of G:C-base pairing and G-quartet stacking. *J. Phys. Chem. C* **2013**, *117*, 23208–23215. [CrossRef]
44. Mertelj, A.; Cmok, L.; Copic, M. Anomalous diffusion in ferrofluids. *Phys. Rev. E* **2009**, *79*, 1402. [CrossRef] [PubMed]
45. Protozanova, E.; Macgregor, R.B. Frayed wires: A thermally stable form of DNA with two distinct structural domains. *Biochemistry* **1996**, *35*, 16638–16645. [CrossRef]
46. Zimbone, M.; Bonaventura, G.; Baeri, P.; Barcellona, M.L. Unusual salt-induced behaviour of guanine-rich natural DNA evidenced by dynamic light scattering. *Eur. Biophys. J. Biophys. Lett.* **2012**, *41*, 425–436. [CrossRef]
47. Prislan, I.; Jamnik, A.; Tomsic, M. Kinetically governed formation of d(G(4)T(2)G(4)) assemblies. *Acta Chim. Slov.* **2012**, *59*, 590–600.

48. Tirado, M.M.; García De La Torre, J. Translational friction coefficients of rigid, symmetric top macromolecules. Application to circular cylinders. *J. Chem. Phys.* **1979**, *71*, 2581–2587. [CrossRef]
49. Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; et al. Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **2012**, *9*, 676–682. [CrossRef] [PubMed]
50. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671–675. [CrossRef]
51. Wagner, T.; Hiner, M. Thorstenwagner/ij-Ridgedetection: Ridge Detection 1.4.0. Available online: 10.5281/zenodo.845874 (accessed on 5 March 2021).
52. Steger, G. An unbiased detector of curvilinear structures. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20*, 113–125. [CrossRef]