Self-assembling of Fmoc-GC Peptide Nucleic Acid dimers into highly fluorescent aggregates

Concetta Avitabile,[b] Carlo Diaferia, [a] Bartolomeo della Ventura, [c] Flavia Anna Mercurio, [b] Marilisa Leone, [b] Valentina Roviello, [d] Michele Saviano, [e] Raffaele Velotta, [c] Giancarlo Morelli, [a] Antonella Accardo, [a] and Alessandra Romanelli*[a].

[a] : University of Naples "Federico II" via Mezzocannone 16, 80134 Naples (Italy)

[b]: Institute of Biostructure and Bioimaging

National Research Council"

via Mezzocannone 16, 80134 Naples (Italy)

[c] Department of Physics

University of Naples "Federico II"[]

via Cintia Naples (Italy)

[d] Analytical Chemistry for the Environment and CeSMA (Centro Servizi Metrologici Avanzati)

University of Naples "Federico II"

via N. Protopisani 80146 Naples (Italy)

[e] Dr. M. Saviano

Institute of Cristallography

National Research Council

Via Amendola 126/O Bari (Italy)

*. A. Romanelli present address: University of Milan, via Venezian 21-20133 Milan (Italy)

Abstract: Self-assembling of molecules by non-covalent interactions is one of the most attracting topics in supramolecular chemistry. The use of short peptides or modified nucleotides as building blocks of the aggregates is particularly intriguing, as these are very easy to synthesize; moreover subtle changes in the chemical structure of such building blocks may drastically affect the properties of the aggregates. The ability of Peptide Nucleic Acids to aggregate is yet very little explored, despite its practical applications. In this work we investigated the self-assembling properties of a PNA dimer, conjugated at the N-terminus to a fluorenylmethoxycarbonyl group. This PNA dimer forms nano-aggregates at low concentration in $CHCl_3/CH_3OH$ mixtures. The aggregates keep very interesting fluorescent properties (high quantum yield in the visible region

with lifetime in the nanoseconds scale), which make them a promising material for applications in optoelectronic.

Introduction

The self-assembly of nucleobases has been extensively investigated with the aim to produce new materials, electronic nanodevices and biosensors.[1] Interactions between nucleobases occur not only by Watson-Crick or Hoogsteen hydrogen bonding, as observed in the DNA double and triple helices; 28 base pairing motifs between the four standard nucleobases characterized by two hydrogen bonds have been identified and these are at the base of self-assembling processes.[2] Purines, guanine in particular, are prone to aggregation. Guanine with its two hydrogen bond acceptors (N7 and O6) and two hydrogen bond donors (N1 amide and N2 amino) produces different self-assembled structures, depending on the conditions in which it is studied.[3] Self-assembly of cytosine, thymine, adenine and uracil has been explored mostly on solid surfaces, such as on Au or on Cu surfaces, revealing the formation of 1D and 2D supramolecular nanostructures.[4] The structure and the stability of the aggregates depend on the linkers attached to the nucleobases and on the experimental conditions of assembly. Guanine and cytosine connected by a rigid, linear, p-diethynylbenzene self-assemble into stable tetrameric macrocycles.[5] Nucleopeptides, composed of amino-acids connected to the bases through a methylene carbonyl linker at the N-terminus or on the side chain, self-assemble in different ways, depending on the pH or addition of metal ions.[6].

The assembly of nucleobases in the context of Peptide Nucleic Acids (PNAs) is so far very little investigated. A PNA-amphiphile peptide conjugate was reported to self-assemble into nanostructures, which retain the ability to bind with high affinity and specificity complementary oligonucleotides.[7] Recently Gazit described the ability of PNA dimers to self-assemble into organized structures, guided by stacking interactions and Watson-Crick base pairing.[8]. GC dimers exhibit interesting fluorescent properties, being able to strongly emit in the wavelength range between 420 and 490 nm, upon excitation between 330 and 430 nm. GC assembly possesses intriguing optoelectronic properties, showing promise for application as an organic light emitting diode. Fluorenylmethoxycarbonyl/benzhydryloxycarbonyl (Fmoc/Bhoc) protected guanine monomers were demonstrated to assemble into spheres which further organize in a bi-dimensional periodic arrangement.[9]

Encouraged by these interesting results, we initiated our investigation on the self-assembly of the PNA dimer Fmoc-GC. As reported for self-assembling peptides, it is reasonable to hypothesize that subtle chemical changes in the structure of the PNAs affect the structure and the properties of the aggregates[10]. In this context, we explored the effects of the introduction of the fluorenylmethoxy carbonyl (Fmoc) moiety on the self-assembling properties and also on the optical properties of the GC. The Fmoc is an aromatic and hydrophobic group, able to self-assemble by π-π stacking and hydrophobic interactions.[10-11] The Fmoc group is reported to promote aggregation of amino-acids, even non aromatic ones.[10, 12] Fluorene derivatives are also employed as blue or green emitters for optoelectronic applications.[13] The effect of Fmoc group on the association of PNAs has not been investigated yet. In this paper we report the characterization of the structure and

fluorescence properties of Fmoc-GC assemblies. We show that assemblies are formed at low concentration in organic solvents, and exhibit high quantum yield in the visible region with lifetime in the nanoseconds scale.

Results and Discussion

The chemical structure of the PNA Fmoc-GC dimer is shown in Figure 1. As the aggregation of modified guanine is reported to occur in organic solvents, we initially tested the solubility of our PNA in $CHCl_3$.[14] The PNA is soluble in $CHCl_3$ only at a very low concentration, but dissolves quickly in the presence of methanol. We performed experiments in $CHCl_3/CH_3OH$ mixtures 1:1 v/v. Initially we explored the absorption and emission properties of our Fmoc-GC by UV and fluorescence spectroscopy.

The UV spectrum of the PNA dimer at 0.5 mg/mL exhibits a wide and off scale peak between 230 and 300 nm, due to the bases and the fluorenylmethoxycarbonyl, and a peak at 376 nm, which is reasonably attributed to the aggregate (Figure S1A). This peak is clearly distinguishable up to 0.1 mg/mL concentration, suggesting that aggregates are present also at such a low concentration (Figure S1B). Diluted solutions of Fmoc-GC show absorption around 260 and a shoulder at 301 nm, as expected for nucleotidic bases and Fmoc (Figure S1C). Fluorescence spectra were recorded upon excitation in the wavelength range between 260 and 540 nm. No fluorescence emission was observed in the excitation wavelength range 260-300 nm, suggesting lack of formation of fluorene excimer.[15] The most intense emission peak appears around 480 nm upon excitation at 376 nm. Plots of fluorescence vs concentration allowed evaluating the minimum aggregation concentration, which is 0.139 mg/mL. Similar results were obtained when the absorbance was plotted versus concentration (Supplementary Figures S1D, S1E). In Figure 2A we report the excitation and emission spectra of Fmoc-GC recorded setting up emission wavelength and excitation wavelength at 480 and 376 nm respectively.

As the fluorescence quantum yield of nucleobases is extremely low, ranging from 0.3 $x10^{-4}$ to 1.2 $x10^{-4}$, fluorescence emission here is a feature of the Fmoc-GC aggregates.[16] The phenomenon of aggregation induced emission (AIE) is associated to chromophore aggregation and has been observed for a number of molecules containing highly conjugated π systems such as tetraphenylethene and 10,10',11,11'-tetrahydro-5,5'-bidibenzo[a,d]annulenylidene in which AIE effects are believed to arise from restriction of intramolecular rotations (RIR) and restriction of intramolecular vibrations (RIV) and also in electron rich molecules, such as tetrahydropyrimidines lacking extended conjugation.[17] In these compounds aggregation produces clusters of different size where the electron clouds are supposed to be overlapped and shared, originating the intermolecular through space conjugation. Aromatic rings in Fmoc-GC dimers are not conjugated, therefore we speculate that fluorescence emission may be due to aggregation and through space conjugation. Fluorescence quantum yield of the aggregate was also determined, using as a reference coumarin 1. The fluorescence quantum yield was found to be 0.53 in $CHCl_3/CH_3OH$ 1/1 v/v, that argues for an efficient conversion of the absorbed light into emitted light. Figure 2B reports the steady state fluorescence of Fmoc-GC in $CHCl_3/CH_3OH$ at several excitation wavelengths, which suggests the occurrence of the so-called red edge excitation shift (REES)

clearly highlighted in Figure 2C where the emission peak wavelength is reported as a function of the excitation wavelength.[18] This behavior has been observed also for GC in alkaline buffer and takes place when both the solvent and the sample are polar; [8] in this condition the molecules of the solvent are not randomly oriented around the molecule of the sample, but rather they tend to orient themselves for minimizing the potential energy. REES can be understood by considering a two-state system in which the ground state and the excited state get closer by virtue of oriented solvent molecules. [18] Statistically, in a fraction of the sample molecules the solvent relaxation will be more pronounced thereby making such molecules more prone to absorb in red edge part of the absorption spectrum with a consequent red-shifted emission spectrum (Figure 2C). REES manifests itself also in the fluorescence lifetime measurement; in fact the excitation of the molecule at a wavelength far from the edges of the absorption band will populate the excited state in sample molecules around which solvent reorientation has not yet occurred. Upon excitation, the newly created excited-state dipole moment triggers the re-adjustment of the solvent shell (solvent relaxation), a process which may occur on a time scale long enough that the fluorescence emission takes place during the solvent relaxation with the consequence that the energy difference between the excited state and the ground state reduces and the emitted photon is red-shifted.[19] Essentially, the red-edge part of the emission spectrum is produced by transitions from states which appear at later times than the excitation. This gives rise to the interesting fluorescence decays reported in Figure 3A, which were obtained by exciting the sample at 406 nm, a wavelength at which REES is barely appreciable (see Figure 2C). As it is clearly visible, the time resolved fluorescence emission at longer wavelength is delayed, a distinct signature of REES already reported for GC.[8] The results obtained at all the wavelengths the lifetime is measured are reported in Figure 3B, where the delay time in the emission is plotted against the emission wavelength. The decay curves are biexponential and the two exponential were calculated by iterative reconvolution with a least squares analysis [(DecayFit - Fluorescence Decay Analysis Software 1.3, FluorTools, www.fluortools.com]. The shorter lifetime τ_1 resulting from the fit is always approximately 0.9 ns, whereas the longer lifetime τ_2 increases with the wavelength (Figure 3C). We ascribe such a behavior to the presence of a mixture of solvents of different polarity ($CHCl_3$ and CH_3OH), which makes the fluorophores interacting both with the more polar and the less polar solvent. The fluorescence arising from fluorophores interacting with the less polar molecules are essentially unaffected by the relaxation of the more polar molecules and have a lifetime that does not depend on the wavelength emission (τ_1=0.9 ns). On the contrary, the fluorophores surrounded by the more polar molecules are quenched by the reduction of the energy difference between the excited and ground state due to the solvent relaxation. Thus, the blue part of the emission is quenched in favor of the red part of the spectrum, the latter arising from a “steady state” condition and as such is not quenched. Thus, the linear increase of the lifetime τ_2 observed in Figure 3C is another signature of REES.

Structural characterization of the aggregate was performed in solution by nuclear magnetic resonance (NMR) spectroscopy and by dynamic light scattering (DLS). NMR experiments were performed in diverse solvents at several concentrations. Spectra were initially recorded in $CDCl_3/CD_3OD$ 1/1 v/v mixtures. Line broadening in the NMR spectra suggests that in this condition Fmoc-GC undergoes aggregation (Figure 4A). Larger aggregates are formed in $CDCl_3/CD_3OD$ 9/1

(v/v) : the 1D [1H] NMR spectrum shows broad peaks in the 7.0-7.5 ppm region, arising from the Fmoc aromatic unit whereas, extensive enlargement of lines due to the fast relaxation of high molecular weight species brings to complete disappearance of other PNA characteristic signals (Figure 4B). Fluorescence spectra recorded in $CHCl_3/CH_3OH$ 9:1 v/v look very similar to those recorded in 1:1 $CHCl_3/CH_3OH$. The assembly process of Fmoc-GC was monitored by NMR, following the shift of the iminic proton signal upon aggregation. In presence of CD_3OD it is difficult to observe the HN iminic and aminic-protons of G and C due to H/D exchange phenomena.[20] However, soon after dissolving the sample at 6 mg/mL concentration in $CDCl_3/CD_3OD$ 1/1 v/v a broad peak from the G iminic proton resulted evident around 10.3 ppm (Figure S2). This iminic proton chemical shift is characteristic of a solvent exposed G that is not involved in canonical H-bond pairing with the C[21]. Previous studies on supramolecular aggregates formed by modified GC dinucleosides indicate that addition of increasing amounts of $CDCl_3$ to DMSO solutions of the dinucleosides brings to the rise of ordered aggregates characterized by GC base pairing.[5, 22] The aggregation process can be easily followed by looking at the gradual disappearance of the iminic proton signal close to 10.5 ppm and its re-appearance at higher ppm values.[5, 22] For this reason we monitored the evolution of the spectra of Fmoc-GC recorded at increasing concentrations of $CDCl_3$ in DMSO, as $CDCl_3$ drives aggregation. The 1D [^{1}H] spectrum of Fmoc-GC in DMSO is characterized by sharp signals, supporting the lack of aggregation (Figure S3); the iminic proton resonating at 10.5 ppm is indicative of a solvent exposed HN. But co-existence of different conformers in DMSO solution can be envisioned due to the known flexibility of a PNA chain.[23] Addition of $CDCl_3$ causes progressive line broadening and shift of the iminic proton signal to 11.8 ppm (Figure 5). Appearance of this broad peak along with other two iminic proton signals around 11.3 ppm (Figure 5) may indicate an equilibrium between monomeric and aggregated species with unstable G-C base pairing. The more low-field shifted peak (i.e., the one at 11.8 ppm) resonates at a chemical shift value closer to that reported from a coupled G-C base pair (generally ~14 ppm) (Figure 5).[5, 22] Unfortunately, even in presence of only 10% DMSO and 90% $CDCl_3$, signals from disaggregated Fmoc-GC units are dominant in spectra. In fact in the 2D [^{1}H, ^{1}H] TOCSY experiments the presence of 4 characteristic conformations for the monomeric GC PNA chain can be revealed (Figure S4A) and no NOE contacts between G and C could be clearly noticed (Figure S4B) in the 2D [1H, 1H] NOESY spectrum[24] [25]. The presence of DMSO, the flexibility of the PNA backbone and the extensive line broadening of the aggregates with larger molecular weights, is surely influencing this outcome, hindering the observation of ordered aggregates in solution provided with canonical GC H-bonds. Aggregates formed in $DMSO/CH_3OH$ show indeed fluorescence properties, but the absorption and emission peaks appear at slightly different wavelengths suggesting a difference either in the organization of the aggregates or in the aggregation degree (Figure S5). Moreover, in a solution containing $CDCl_3$ and CD_3OD, there is no clear evidence of canonical Watson and Crick GC pairing thus letting us hypothesize that aggregation might be driven by $\pi-\pi$ interactions in between aromatic rings and possible involves the Fmoc unit. To better investigate the structural role that the Fmoc unit could play in the aggregates, we compared NMR spectra of Fmoc-Gly-OH with those of Fmoc-GC (Figure S6). As can be clearly seen, under the different experimental conditions, Fmoc signals from Fmoc-GC undergoes large line broadening with respect to Fmoc-Gly-OH and upfield chemical shift changes (Figure S6). This enlargement of NMR lines underlies

formation of aggregates for the modified PNA chain. Observed intensity and chemical shift changes resemble those observed in self-assembly of Fmoc-Phenylalanine into hydrogel and points to $\pi-\pi$ stacking interactions mediated by Fmoc.[27]
We further investigated possible structural properties of aggregates by in silico molecular modeling. Monomeric Fmoc-GC (Figure 6A) could be present in solution in equilibrium with aggregated forms. Furthermore, monomeric Fmoc-GC chains may interact through canonical Watson and Crick G-C pairing (Figure 6B).
The resultant dimeric units can then self-assemble into different supramolecular aggregates, namely H-type and J-type like aggregates.[28] Based on NMR data highlighting the importance of Fmoc unit in the Fmoc-GC unit PNA self-assembly process through formation of $\pi-\pi$ aromatic interactions and also based on spectroscopic evidences, such as the red shift of the absorption spectra of the aggregates as compared to the monomeric form and the fluorescence of the aggregates, we speculate that aggregates might be J-like aggregates (Figure 6 C,D).

Dimension of the aggregates was determined in solution by DLS. The DLS profile for Fmoc-GC aggregates in $CHCl_3/CH_3OH$ 1:1 v/v is reported in Figure 7A. The distribution profile shows two well-separated modes, due to translational diffusion processes with apparent translational diffusion coefficients D_{fast} and D_{slow}. These modes can be attributed to the presence in solution of nanostructures with different size. According to the Stokes–Einstein equation, at infinite dilution, it is possible to evaluate the hydrodynamic diameter of these nanostructures. The mean diameter values obtained for slow and fast modes, are 167 ± 50 nm and 40 ± 6 nm, respectively. Finally, we characterized the solid obtained upon drying the Fmoc-GC solution. The morphology of the aggregates deposited into an aluminum support was investigated by scanning electron microscopy (SEM). Interestingly SEM (Figures 7B-7D) experiments reveal the formation of network of nano-spheres. The diameter of nanosphere obtained at 1.0 mg/mL is in good agreement with the value calculated by DLS measurement (150-500 nm); whereas sample at 5.0 mg/mL shows larger spheroidal aggregates (400-1300 nm).

Importantly, fluorescence microscopy experiments carried out on the solid revealed the ability of Fmoc-GC to emit in the blue and green, suggesting the existence of supramolecular aggregates in solution as well as in the solid state. (Figures 8A-D)

Conclusions

Fmoc-GC PNA dimers produce aggregates in $CHCl_3/CH_3OH$ at low concentration kept together by $\pi-\pi$ interactions and Watson Crick hydrogen bond. The introduction of a fluorenylmethoxycarbonyl at the N-terminus of the PNA GC affects the aggregation properties as well as the morphology of the aggregates. The fluorenylmethoxycarbonyl group drives the aggregation process, as suggested by NMR, likely yielding to three-dimensional structures of the aggregates different from those observed for GC. Interestingly, spectroscopic properties observed for Fmoc-GC are similar to those observed for GC, confirming that the aggregation of the nucleic acid bases triggers fluorescence. Dimension of the aggregates as well as the optical properties

shown in solution are consistent with those found at the solid state. Solubility of the aggregates in organic solvents renders them compatible with conventional lithography processes. The chemical stability of PNAs, the ease of preparation and the fluorescence properties of such aggregates encourage us to further investigate their optical and electronic properties for technological applications The possibility to derivatize the PNA backbone with a variety of functional groups will help us to explore new molecular architectures and to correlate the chemical structure of the nucleic acid building blocks to 3D structure of the aggregates and ultimately to their chemical and physical properties.

Experimental Section

PNA synthesis: PNA were obtained by solid phase synthesis using standard protocols.[29] Purification was carried out by RP-HPLC on a Shimadzu LC-8A, equipped with a SPD-M10 AV diode array detector using a Phenomenex Jupiter 10 µm Proteo (90 Å 250 × 10.00 mm) column with a flow rate of 5 mLmin^{-1} and with a gradient of CH_3CN (0.1% TFA) in H2O (0.1% TFA) from 5 to 50 % in 20 minutes. Pure compounds were analyzed by LC-MS on a LC-MS Agilent Technologies 6230 ESI-TOF on a Phenomenex Jupiter 3 µ C18 (150 × 2.0 mm) column with a flow rate of 0.2 mL·min^{-1}. Products were liophylized three times, the first to remove the HPLC solvents, the second upon dissolution of the samples in H_2O/CH_3COOH 7/3 v/v, the third following dissolution in H_2O.

Fmoc-GC. Calculated mass (Da): 781.510; Found (Da): 782.338 $[M+H]^+$, 391.676 $[M+2H]^{2+}$.

Sample preparation for absorption and emission experiments: Samples were dissolved in $CHCl_3/CH_3OH$ 1/1 v/v; concentration of the Fmoc-GC was evaluated using diluted solutions based on the absorbance at 301nm of the fluorenylmethoxycarbonyl group (ε=7800 M^{-1}cm^{-1}).

Steady state fluorescence: Spectroscopy experiments were performed on a VARIAN Cary Eclipse Fluorimeter. Samples were dissolved in $CHCl_3/CH_3OH$ 1:1 v/v at 0.5 mg mL-1 concentration. Emission spectra were recorded upon excitation at various wavelength in the range 300-540 nm; excitation and emission slits were set respectively at 5 and 10 nm for measurements up to excitation wavelength of 400 nm, while slits were both set at 10 nm for measurements with an excitation wavelength > 420 nm. Plot of excitation vs emission wavelength was obtained plotting wavelengths obtained by three independent experiments; the software Prism was used. The excitation spectrum was obtained setting the emission wavelength at 480 nm, excitation and emission slits respectively 5 and 10 nm.

Time-Resolved Fluorescence Measurements: Time-resolved fluorescence experiments were carried out using the experimental setup shown in Supporting information Scheme 1. Excitation for the solution was provided by a picosecond diode laser (Picoquant, LDH-P-C-405B) emitting pulses at repetition rate of 40 MHz and a wavelength of 406 nm. The laser beam was focused into a 10 mm sample cell by a simple len. The fluorescence light was detected at 90° to the incident light beam to minimize the amount of transmitted or reflected beam light reaching the detector. An Asahi ZBPA410 bandpass filter blocks the residual laser beam and allows only radiation with a wavelength range above 410 ± 10 nm to reach the detector. The detection apparatus was

composed of a fast compact PMT module (H9305-03, Hamamatsu, Tokyo, Japan) and a time-correlated single-photon counting (TCSPC) electronics (SPC130M, Becker and Hickl GmbH, Berlin, Germany). The instrument response function (IRF) determined by TCSPC was about 320 ps FWHM. Analysis of fluorescence transients was performed as described in the Supporting information.

Quantum yield: Fmoc-GC samples were dissolved in 1:1 v/v $CHCl_3$/CH_3OH and Coumarin 1 (7-Diethylamino-4-methylcoumarin) was dissolved in ethanol. UV spectra were recorded at different concentrations and the absorbance at 376 nm was recorded. Fluorescence spectra were recorded using an excitation wavelength of 376 nm, setting excitation and emission slits at 10 nm. The integrated fluorescence intensity was determined in the range 390-600. Detail on the calculation are reported in the Supporting information.

Fluorescence microscopy: 15 μL of Fmoc-GC solution at 5.0 mg/mL were deposited on a clean coverslip glass, dried and imaged with fluorescence microscopy. Immunofluorescence images were taken with a Leica DFC320 video-camera (Leica, Milan, Italy) connected to a Leica DMRB microscope equipped with a 10 X and 40 X objectives and the Image J Software (National Institutes of Health, Bethesda, MD) was used for analysis.

Dynamic Light Scattering (DLS): Mean diameter and diffusion coefficient (D) of Fmoc-GC based nanostructures in 1:1 $CHCl_3$/CH_3OH (5.0 mg/mL) were measured by DLS measurements using a Zetasizer Nano ZS (Malvern Instruments, Westborough, MA) that employs a 173° backscatter detector. Other instrumental settings were: measurement position (mm) of 4.65; attenuator of 9; temperature of 25 °C; disposable sizing cuvette as cell. DLS measurements in triplicate were carried out also on the sample at 1 mg/mL, after room temperature centrifugation at 13,000 rpm for 5 minutes.

Scanning electron microscopy: Morphological analysis of the self-assembled PNA was carried out using field emission scanning electron microscope (Nova NanoSem 450-FEI). Samples at 1.0 and 5.0 mg/mL were prepared as above described. Solutions were dropped off on an aluminium stub. The air dried samples were covered with a thin coat of gold and palladium sputtered on the stub at a current of 20 mA for 90 s. The images of the sample were acquired with an Everhart Thornley Detector (ETD) and a Through the Lens Detector (TLD) using an accelerating voltage of 2–5 kV.

NMR: NMR spectra were acquired on a Varian (Varian by Agilent Technologies, Milan, Italy) Unity Inova spectrometer operating at a proton frequency of 400 MHz equipped with z-axis pulsed-field gradients and a triple resonance probe or a 600 MHz provided with a cold-probe. Spectra were recorded at 298 K with sample volumes equal to 500 μl.

1D [^{1}H] spectra were recorded for Fmoc-GC dissolved in mixtures $CDCl_3$ (Chloroform-d, 99.8% d, Sigma-Aldrich, Milan, Italy)/CD_3OD (methanol-d4, 99% d, Sigma-Aldrich, Milan, Italy) 9/1 v/v in the concentration range 1-10 mg/mL and 1/1 v/v in the concentration range 1.5-6.0 mg/mL. Experiments were as well conducted in 99.9% DMSO (dimethyl sulfoxide-d_6, 99.9% d, Sigma-Aldrich, Milan, Italy) (8 mg/mL concentration) and in mixtures DMSO/$CDCl_3$ containing increasing amounts of $CDCl_3$ (i.e., 20% v/v; 50% v/v; 80% v/v; 90% v/v) in the concentration range 0.9-1.8

mg/mL. 1D [1H] spectra were generally recorded with a relaxation delay of 1 sec and 0.5-4 k scans. 2D [^{1}H,^{1}H] NOESY[25] and TOCSY[24] spectra were recorded at 600 MHz with mixing times of 300 and 70 ms respectively, with 16-64 scans, 128-256 FIDs in t1, 1024 or 2048 data points in t2. Spectra were processed and analyzed with Varian software (vnmrj_1.1D) and for 2D experiments with the NEASY utility of the CARA software package (Computer Aided Resonance Assignments http://cara.nmr.ch/doku.php). [30]

Conformational analysis and molecular modelling. 3D models for Fmoc-GC in the monomeric and dimeric states were build with the software UCSF Chimera[31] starting from the atomic coordinates of the -GC- fragments in the crystal structure of the PNA p(GCTGCTGC)$_2$ duplex (pdb entry code 5EMG[32]). The starting models were subjected to unrestrained energy minimizazion with UCSF Chimera (including 2000 steps of steepest descent and 200 steps of conjugate gradient).[31] Models of aggregates, made up of ripetition of the obtained dimeric unit, were built as well in UCSF Chimera and energy minimized (100 steps of steepest descent and 10 steps of conjugate gradient).[31]

The conformational space accessible to Fmoc-GC was explored through the software Cyndi[33]. The Cyndi run was conducted setting 200 populations and 200 generations. The probabilities for crossover and mutation operation were kept to 0.85 and 0.1, respectively. The epsilon values for a) VDW energy, b) torsion energy, c) Geometric dissimilarity, and d) gyration radius were 20 kcal/mol, 5 kcal/mol, 0.2 Å and 0.1 Å, respectively. The maximum iteration for post processing of conjugated gradient minimization was set to 100, and the convergence criterion -relying on gradient RMS - was set to 0.1 kcal·mol-1·Å-1. A scale factor equal to 0.3 Å was chosen for RMSD filter.

References

[1] J. T. Davis and G. P. Spada, Chemical Society reviews 2007, 36, 296-313.

[2] S. Sivakova and S. J. Rowan, Chemical Society reviews 2005, 34, 9-21.

[3] a) A. Wong, R. Ida, L. Spindler and G. Wu, Journal of the American Chemical Society 2005, 127, 6990-6998; b) A. Ciesielski, M. El Garah, S. Masiero and P. Samori, Small 2016, 12, 83-95.

[4] L. Liu, D. Xia, L. H. Klausen and M. Dong, International journal of molecular sciences 2014, 15, 1901-1914.

[5] C. Montoro-Garcia, J. Camacho-Garcia, A. M. Lopez-Perez, N. Bilbao, S. Romero-Perez, M. J. Mayoral and D. Gonzalez-Rodriguez, Angewandte Chemie 2015, 54, 6780-6784.

[6] a) G. N. Roviello, D. Musumeci, E. M. Bucci and C. Pedone, Molecular bioSystems 2011, 7, 1073-1080; b) G. N. Roviello, A. Ricci, E. M. Bucci and C. Pedone, Molecular bioSystems 2011, 7, 1773-1778; c) X. Li, Y. Kuang, H. C. Lin, Y. Gao, J. Shi and B. Xu, Angewandte Chemie 2011, 50, 9365-9369.

[7] M. O. Guler, J. K. Pokorski, D. H. Appella and S. I. Stupp, Bioconjugate Chemistry 2005, 16, 501-503.

[8] O. Berger, L. Adler-Abramovich, M. Levy-Sakin, A. Grunwald, Y. Liebes-Peer, M. Bachar, L. Buzhansky, E. Mossou, V. T. Forsyth, T. Schwartz, Y. Ebenstein, F. Frolow, L. J. W. Shimon, F. Patolsky and E. Gazit, Nature Nanotechnology 2015, 10, 353-360.

[9] O. Berger, E. Yoskovitz, L. Adler-Abramovich and E. Gazit, Advanced materials 2016, 28, 2195-2200.

[10] K. Tao, A. Levin, L. Adler-Abramovich and E. Gazit, Chemical Society reviews 2016, 45, 3935-3953.

[11] K. Tao, E. Yoskovitz, L. Adler-Abramovich and E. Gazit, Rsc Advances 2015, 5, 73914-73918.

[12] B. Adhikari, J. Nanda and A. Banerjee, Soft Matter 2011, 7, 8913-8922.

[13] M. R. Zhu and C. L. Yang, Chemical Society reviews 2013, 42, 4963-4976.

[14] G. Gottarelli, S. Masiero, E. Mezzina, S. Pieraccini, J. P. Rabe, P. Samori and G. P. Spada, Chemistry 2000, 6, 3242-3248.

[15] a) F. L. Minn, J. P. Pinion and N. Filipescu, The journal of physical chemistry 1971, 76, 1794-1798; b) B. H. Boo, M. Lee, K. S. Jeon and S. J. Kim, The journal of physical chemistry. A 2014, 118, 2269-2278.

[16] P. R. Callis, Ann. Rev. Phys. Chem. 1983, 34, 329-357.

[17] a) J. Mei, N. L. Leung, R. T. Kwok, J. W. Lam and B. Z. Tang, Chemical reviews 2015, 115, 11718-11940; b) G. F. Zhang, Z. Q. Chen, M. P. Aldred, Z. Hu, T. Chen, Z. Huang, X. Meng and M. Q. Zhu, Chemical communications 2014, 50, 12058-12060; c) N. L. Leung, N. Xie, W. Yuan, Y. Liu, Q. Wu, Q. Peng, Q. Miao, J. W. Lam and B. Z. Tang, Chemistry 2014, 20, 15349-15353.

[18] J. R. Lakowicz, Eds Springer 2006, , 257-258.

[19] a) Y. V. Zvinevich, N. A. Nemkovich and A. N. Rubinov, Journal of Applied Spectroscopy 1993, 58, 171-186; b) Y. V. Zvinevich, N. A. Nemkovich, A. N. Rubinov and V. I. Tomin, Journal of Molecular Liquids 1990, 45, 1-6.

[20] J. L. Sessler, J. Jayawickramarajah, M. Sathiosatham, C. L. Sherman and J. S. Brodbelt, Org Lett 2003, 5, 2627-2630.

[21] C. Montoro-Garcia, J. Camacho-Garcia, A. M. Lopez-Perez, N. Bilbao, S. Romero-Perez, M. J. Mayoral and D. Gonzalez-Rodriguez, Angew Chem Int Ed Engl 2015, 54, 6780-6784.

[22] a) J. L. Sessler, J. Jayawickramarajah, M. Sathiosatham, C. L. Sherman and J. S. Brodbelt, Organic Letters 2003, 5, 2627-2630; b) S. Romero-Perez, J. Camacho-Garcia, C. Montoro-Garcia, A.

M. Lopez-Perez, A. Sanz, M. J. Mayoral and D. Gonzalez-Rodriguez, Organic Letters 2015, 17, 2664-2667.

[23] M. Leijon, A. Graslund, P. E. Nielsen, O. Buchardt, B. Norden, S. M. Kristensen and M. Eriksson, Biochemistry 1994, 33, 9820-9825.

[24] C. Griesinger, G. Otting, K. Wuthrich and R. R. Ernst, Journal of the American Chemical Society 1988, 110, 7870-7872.

[25] A. Kumar, R. R. Ernst and K. Wuthrich, Biochem Biophys Res Commun 1980, 95, 1-6.

[26] H. E. Gottlieb, V. Kotlyar and A. Nudelman, J Org Chem 1997, 62, 7512-7515.

[27] V. Singh, K. Snigdha, C. Singh, N. Sinha and A. K. Thakur, Soft Matter 2015, 11, 5353-5364.

[28] B. Heyne, Photochem Photobiol Sci 2016, 15, 1103-1114.

[29] C. Avitabile, L. Moggio, L. D. D'Andrea, C. Pedone and A. Romanelli, Tetrahedron Letters 2010, 51, 3716-3718.

[30] C. Bartels, Xia, T., Billeter, M., Güntert, P. & Wüthrich, K, J. Biomol. NMR 1995, 1-10.

[31] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, J Comput Chem 2004, 25, 1605-1612.

[32] A. Kiliszek, K. Banaszak, Z. Dauter and W. Rypniewski, Nucleic Acids Res 2016, 44, 1937-1943.

[33] a) X. Liu, F. Bai, S. Ouyang, X. Wang, H. Li and H. Jiang, BMC Bioinformatics 2009, 10, 101; b) M. Gaglione, G. Malgieri, S. Pacifico, V. Severino, B. D'Abrosca, L. Russo, A. Fiorentino and A. Messere, Molecules 2013, 18, 9147-9162.

Figure 1. Chemical structure of Fmoc-GC.

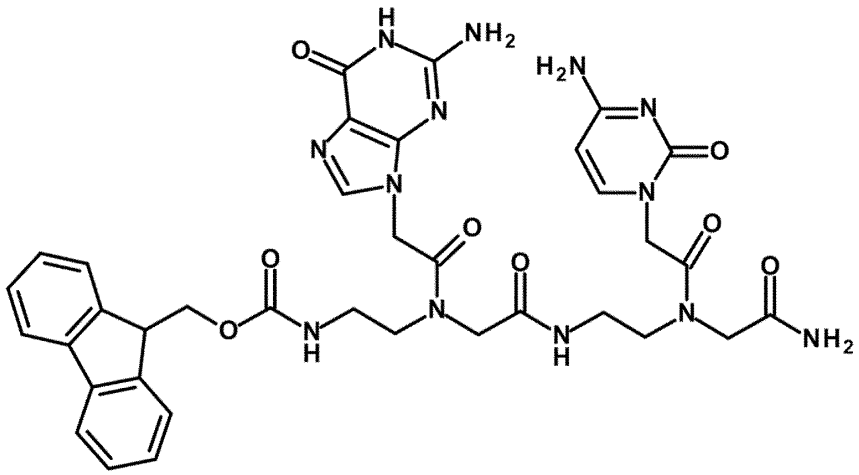


Figure 2. A: excitation (blue) and emission (red) spectra recorded for Fmoc-GC in $CHCl_3/CH_3OH$ 1/1 v/v at 0.5 mg/mL. B: normalized fluorescence spectra recorded at different wavelength for Fmoc-GC 0.5 mg/mL (magenta: λ_{ex}= 400 nm; blue λ_{ex} = 420 nm, light blue λ_{ex} = 440 nm, green λ_{ex} = 460 nm. C: plot of the emission wavelength versus excitation wavelength. .

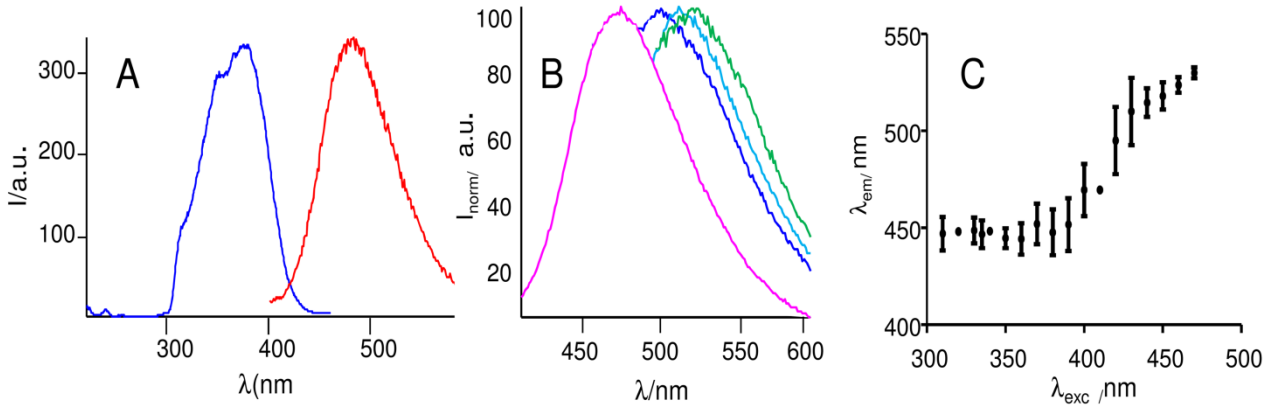


Figure 3: A: time-resolved emission at several wavelengths with λ_{ex} =406 nm. The continuous line is the best fit of the experimental data at 430 nm. The fits at the other wavelength emissions are of the same quality. B: fluorescence delay vs wavelength emission. C: fluorescence lifetime τ_2 vs wavelength emission.

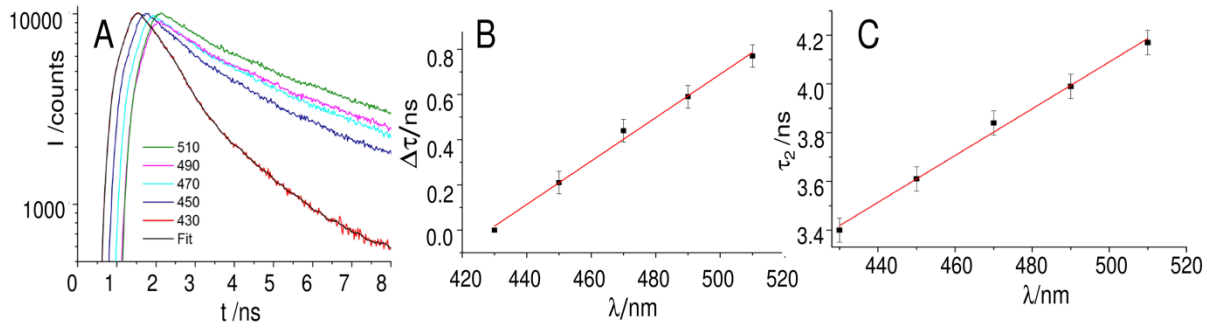


Figure 4: Aromatic region of 1D [^{1}H] spectra of Fmoc-GC at increasing PNA concentrations recorded in: A) $CDCl_3/CD_3OD$ 1/1 v/v B) $CDCl_3/CD_3OD$ 9/1 v/v. Spectra were calibrated on the $CDCl_3$ peak at 7.26 ppm.

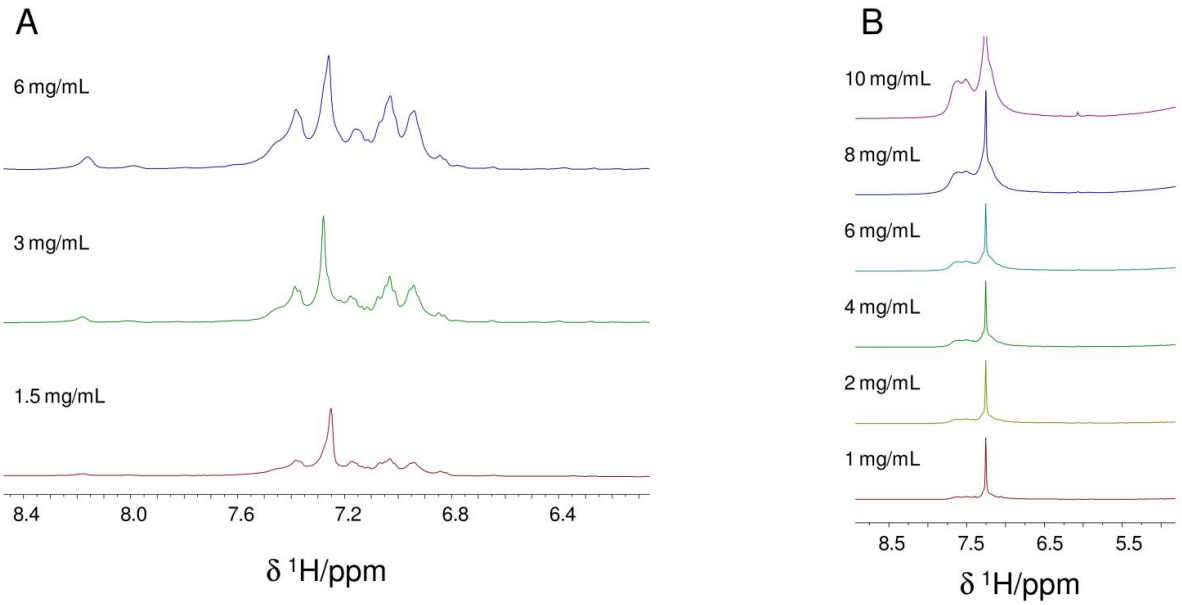


Figure 5. Imino region of the 1D [^{1}H] spectra of Fmoc-GC acquired at different DMSO/$CDCl_3$ ratios. Fmoc-GC concentration is 1.8 mg/mL in the spectra recorded at lowest $CDCl_3$ percentages (till 80%) and 0.9 mg/mL at 90% $CDCl_3$. Spectra were calibrated on the DMSO residual peak at 2.5 ppm until a DMSO/$CDCl_3$ ratio equal to 1/1 v/v, at the highest $CDCl_3$ percentages the DMSO peak was set at 2.62 ppm.[26]

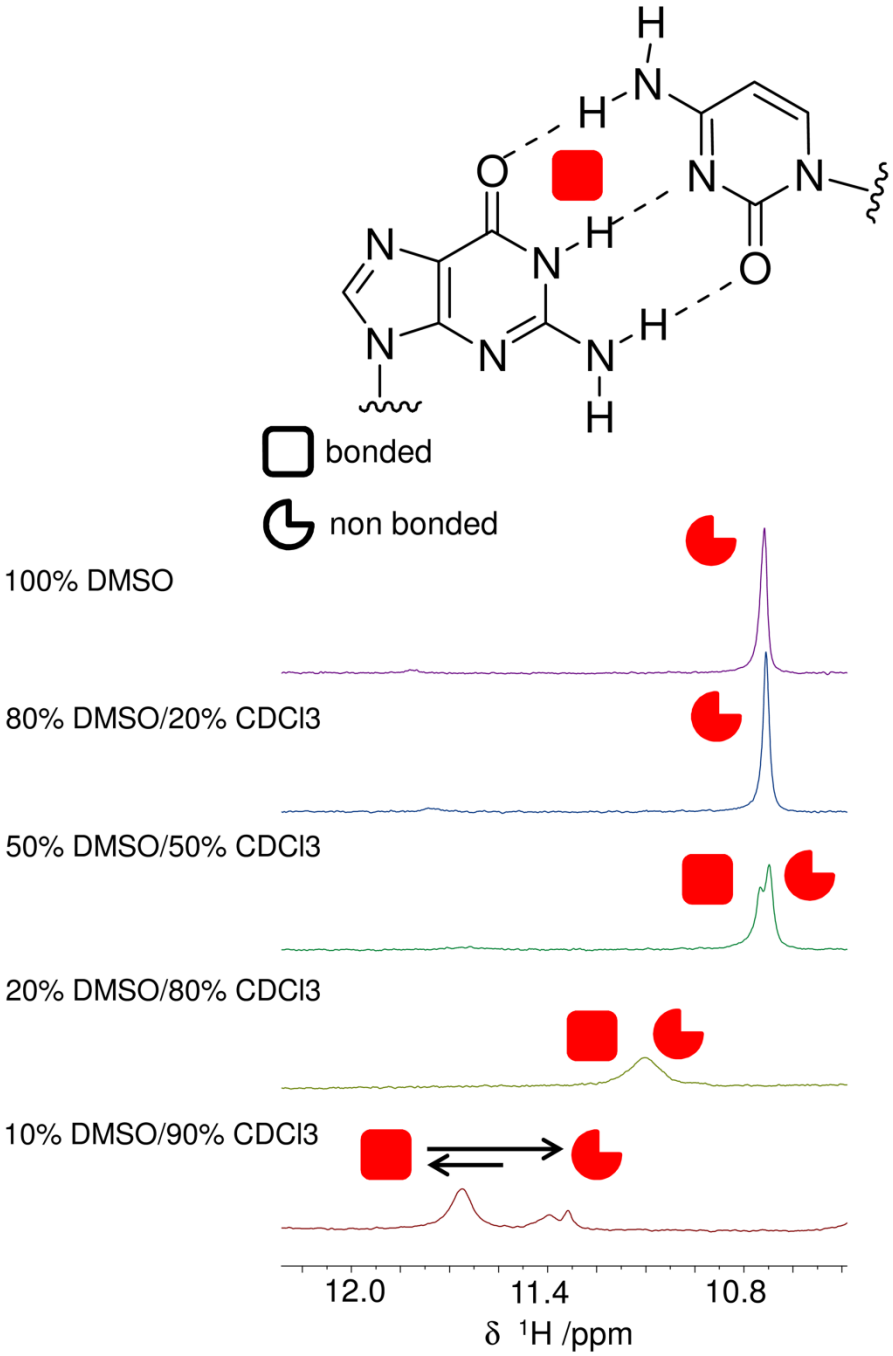


Figure 6. (A) Ensemble of 12 possible Fmoc-GC conformers calculated by the software Cyndi. Theorical speculative 3D models of (B) Fmoc-GC dimer, (C, D) J-aggregates.

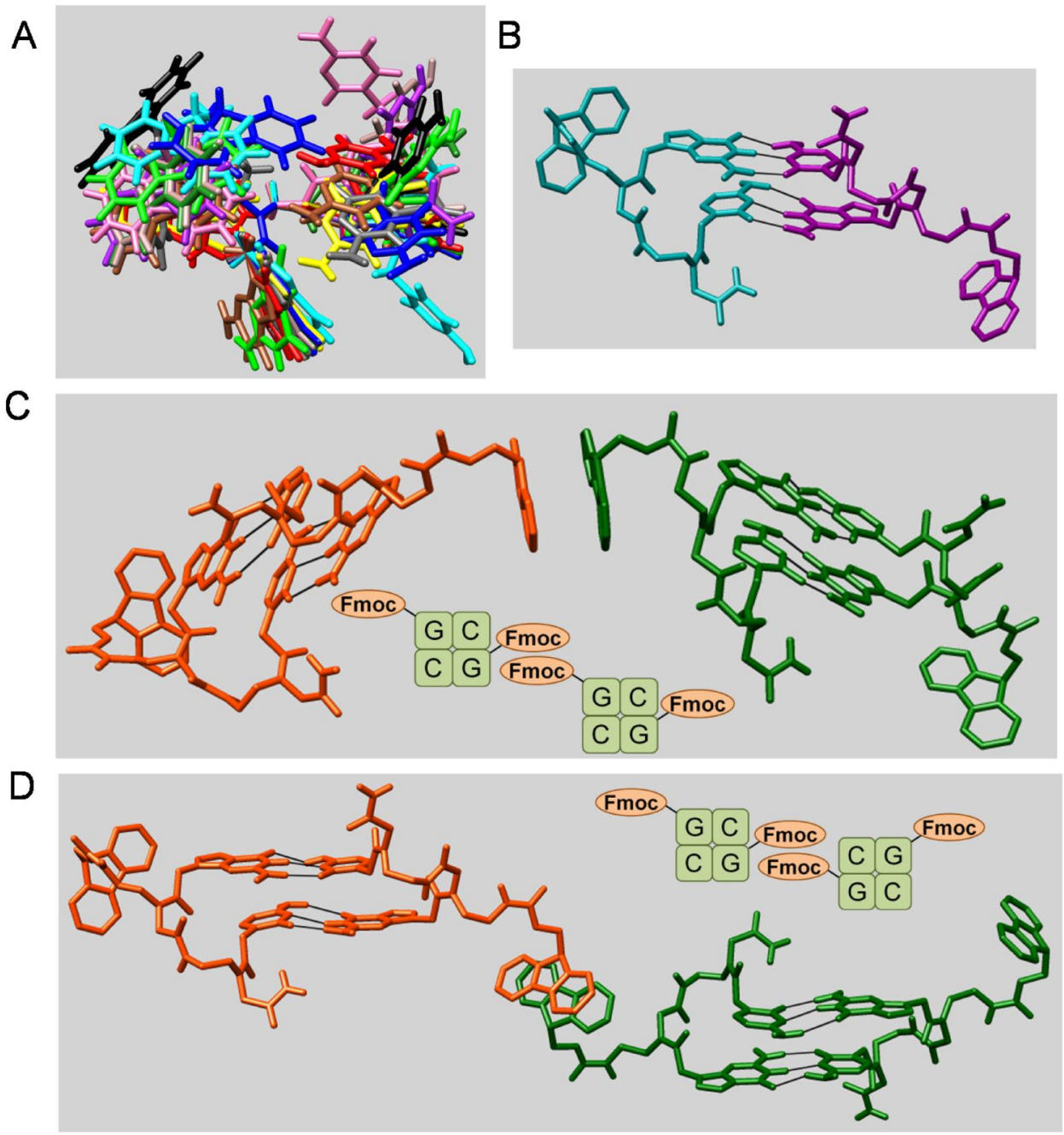


Figure 7. A: DLS profile of Fmoc-GC at 1.0 mg/mL. B: SEM microphotos for Fmoc GC at 1.0 mg/mL, 3 μm and 50000x. C: SEM microphotos for Fmoc-GC at 5.0 mg/mL 10 μm and 12000x. D: SEM microphotos for Fmoc-GC at 5.0 mg/mL 3 μm and 50000x.

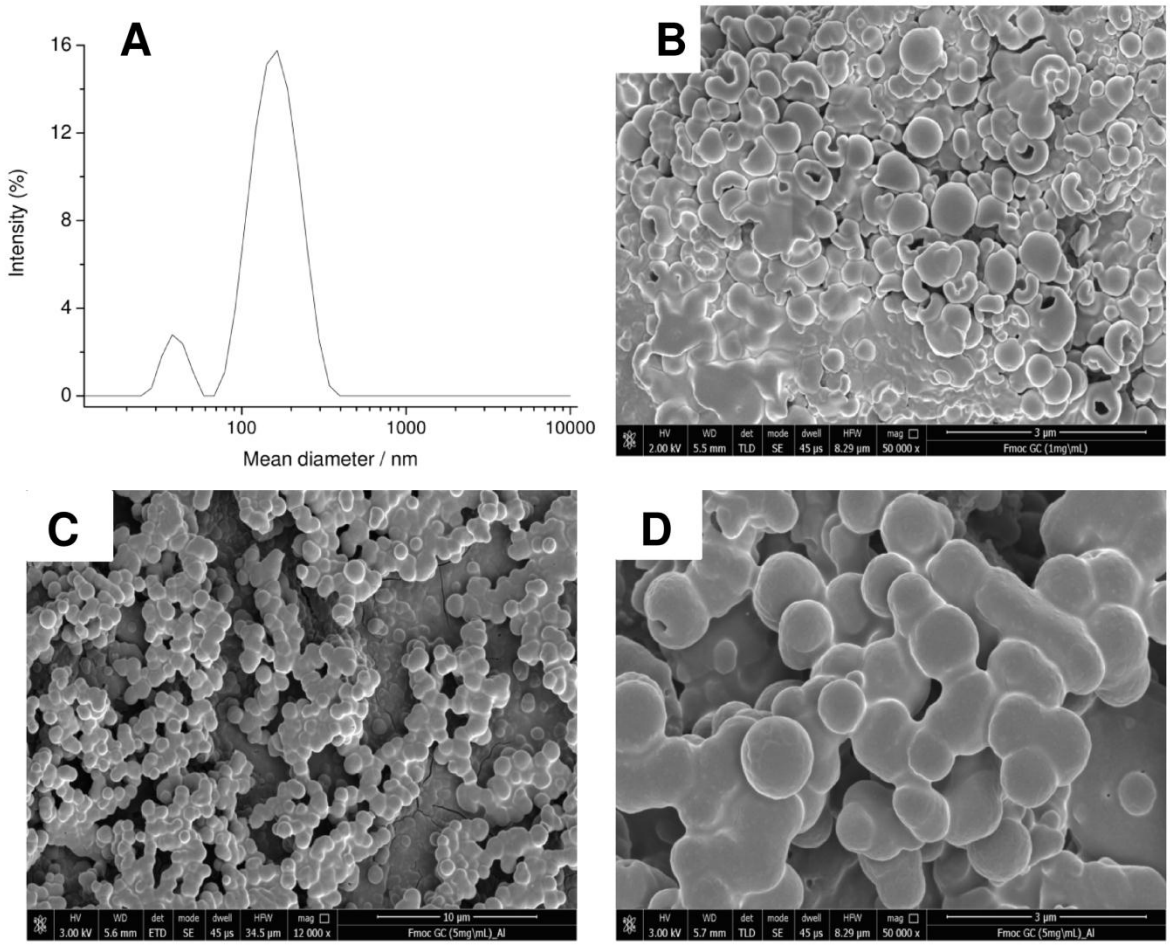


Figure 8. Fluorescence microscopy images of Fmoc-GC at 5.0 mg/ mL deposited on clean coverslip glasses and slowly dried at room temperature. The images were obtained by exciting the air-dried sample in the spectral regions of (A) DAPI (λ_{ex}=359 nm, λ_{em}=461 nm), (B) GFP (λ_{ex}=488 nm, λ_{em}=507 nm); (D) brightfield, E: superimposition of images A to C . The scale bar is 50 µm for all images.

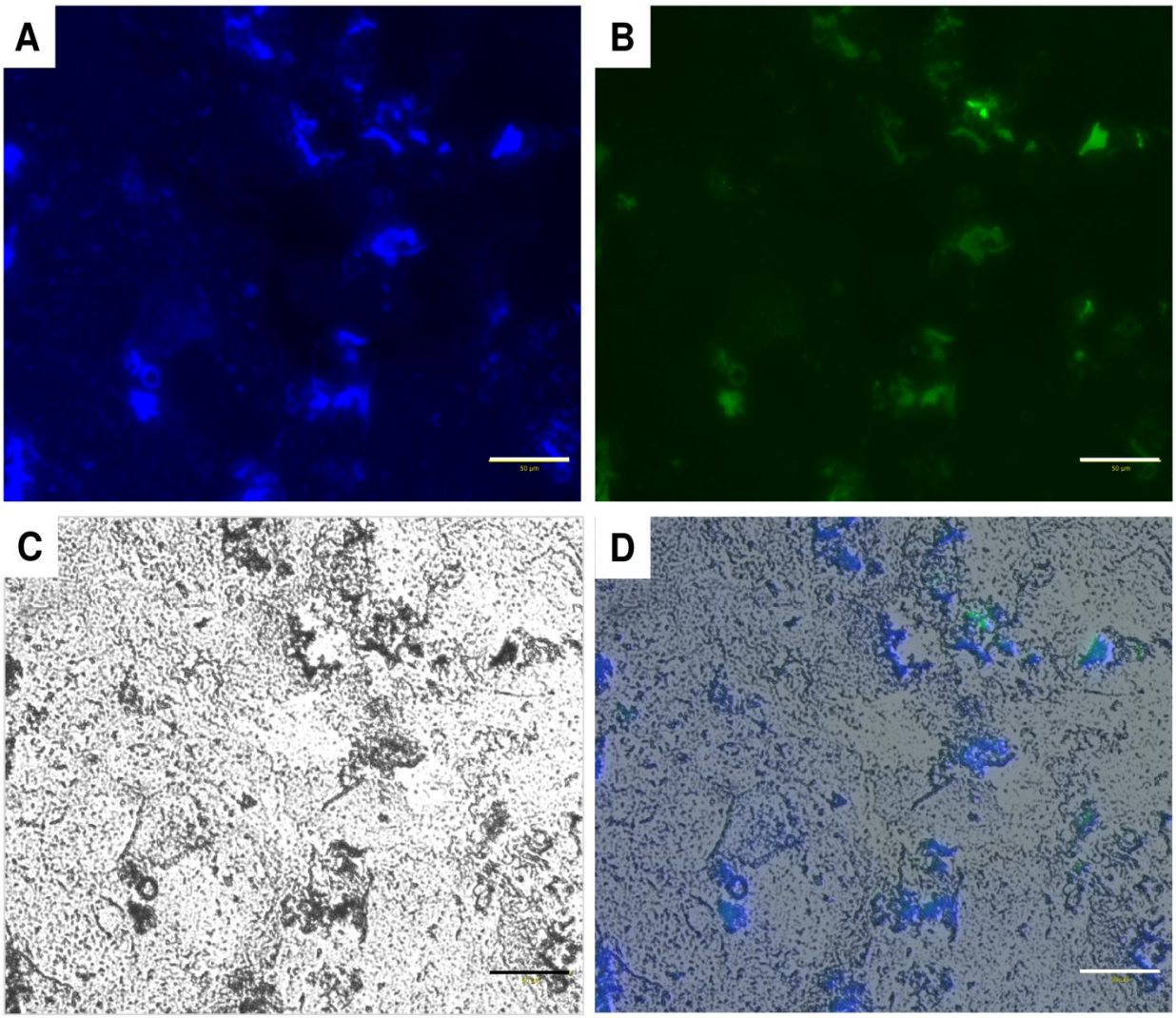
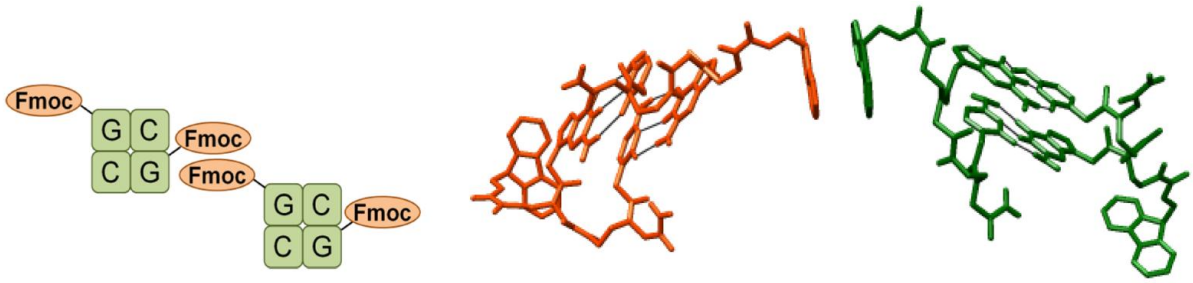


Table of content



Text for the table of content: Fmoc promotes the aggregation of PNA dimers

Keywords: Peptide Nucleic Acids • aggregation • fluorescence • REES • nanosphere