Chain stretching in brushes favors sequence recognition for nucleobase-functionalized flexible precise oligomers

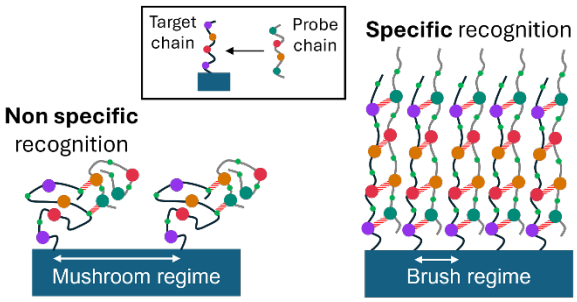
Kseniia Grafskaia,[(1,†)] Qian Qin,[(1,†)] Jie Li,[(1,††)] Delphine Magnin,[(1)] David Dellemme,[(2)] Mathieu Surin,[(2)] Karine Glinel,[(1,*)] and Alain M. Jonas[(1,*)]

(1) Institute of Condensed Matter and Nanosciences, Université catholique de Louvain, Croix du Sud 1 L7.04.02, Louvain-la-Neuve, Belgium
(2) Laboratory for Chemistry of Novel Materials, Université de Mons – UMONS, Avenue Maistriau, 17, B-7000 Mons, Belgium
(*) Corresponding authors
(†) These authors contributed equally to this work
(††) Present address: State Key Laboratory of Chemical Resource Engineering, Beijing Laboratory of Biomedical Materials, Beijing University of Chemical Technology, Beijing 100029, P. R. China

Abstract

Six different flexible stereocontrolled oligo(triazole-urethane)s substituted by precise sequences of nucleobases or analogs are synthesized. Molecular dynamics simulations indicate that the flexibility of the backbone leads to unspecific complexation of pairs of oligomers, irrespective of the complementarity of their sequences. This is ascribed to the existence of other interactions between pairs of oligomers, as well as to the spatial blurring of the sequence order encoded in the chemical structure of the chain due to its flexibility. The same conclusions are drawn when investigating the irreversible adsorption of different probe oligomers onto a layer of target oligomers grafted by click chemistry in a mushroom configuration on a silicon substrate. In contrast, when the target oligomers are grafted in denser brush configurations, irreversible adsorption becomes more specific, with probe chains of complementary sequence being twice more probable to be irreversibly-bound to the layer of target chains than those of non-complementary sequence. This is ascribed to lateral excluded volume interactions between chains in the brush, leading to partial chain stretching and increased spatial preservation of the information contained in the monomer sequence of the chains. At even higher grafting densities, however, the penetration of the probe chains in the brush becomes increasingly difficult, resulting in a loss of binding efficiency. Our work thus demonstrates the adverse role of chain flexibility in the specificity of complexation between nucleobase-functionalized oligomers and provides directions for an improvement of specificity by tuning the grafting density of target chains on a substrate.

TOC graphic and summary statements



□ The high flexibility of stereo-controlled oligomers substituted by precise sequences of nucleobases prevents sequence-specific recognition with surface-grafted target chains at low grafting density.
□ Moderately higher grafting densities of surface-grafted target chains promote sequence-specific recognition thanks to chain stretching.

Introduction

The specificity of DNA duplex formation has been used to develop applications such as DNA origami,[1–3] programmed nanoparticle assembly[4–14] or molecular algorithms.[15] These rely on the formation of double helices, in which hydrogen bonds between nucleobases on complementary strands and base-stacking interactions provide the stabilization mechanism for duplex formation.[16] Importantly, the large electrostatic persistence length of single-stranded DNA[17] compared to the distance between successive nucleobases along the DNA sugar-phosphate backbone results in substantial chain rigidity, which contributes to the local spatial preservation of the chemical information stored in the sequence of nucleobases along the chains.

Synthetic oligomers bearing complementary recognition units can also form sequence-selective duplexes.[18–24] Among those, a number of synthetic sequence-defined oligomers bearing nucleobases or analogous side groups have been demonstrated to form H-bonded duplexes, sometimes including with complementary oligonucleotides;[25–29] peptide nucleic acids are a well-known example.[30] Most nucleobase-functionalized synthetic systems have relatively rigid backbones and/or moderate distances between successive bases, leading to the short-range spatial preservation of the chemically-encoded sequence information contained in the primary structure of the chains.[29] This partial spatial preservation presumably helps in the recognition of a complementary sequence. Here, we want to address the issue of recognition when the distances between bases and the flexibility of the backbone scramble more the information contained in the primary sequence of the chains. Would the blurring resulting from conformational fluctuations completely prevent recognition between complementary chains and if this is the case, what would be possible solutions to this issue?

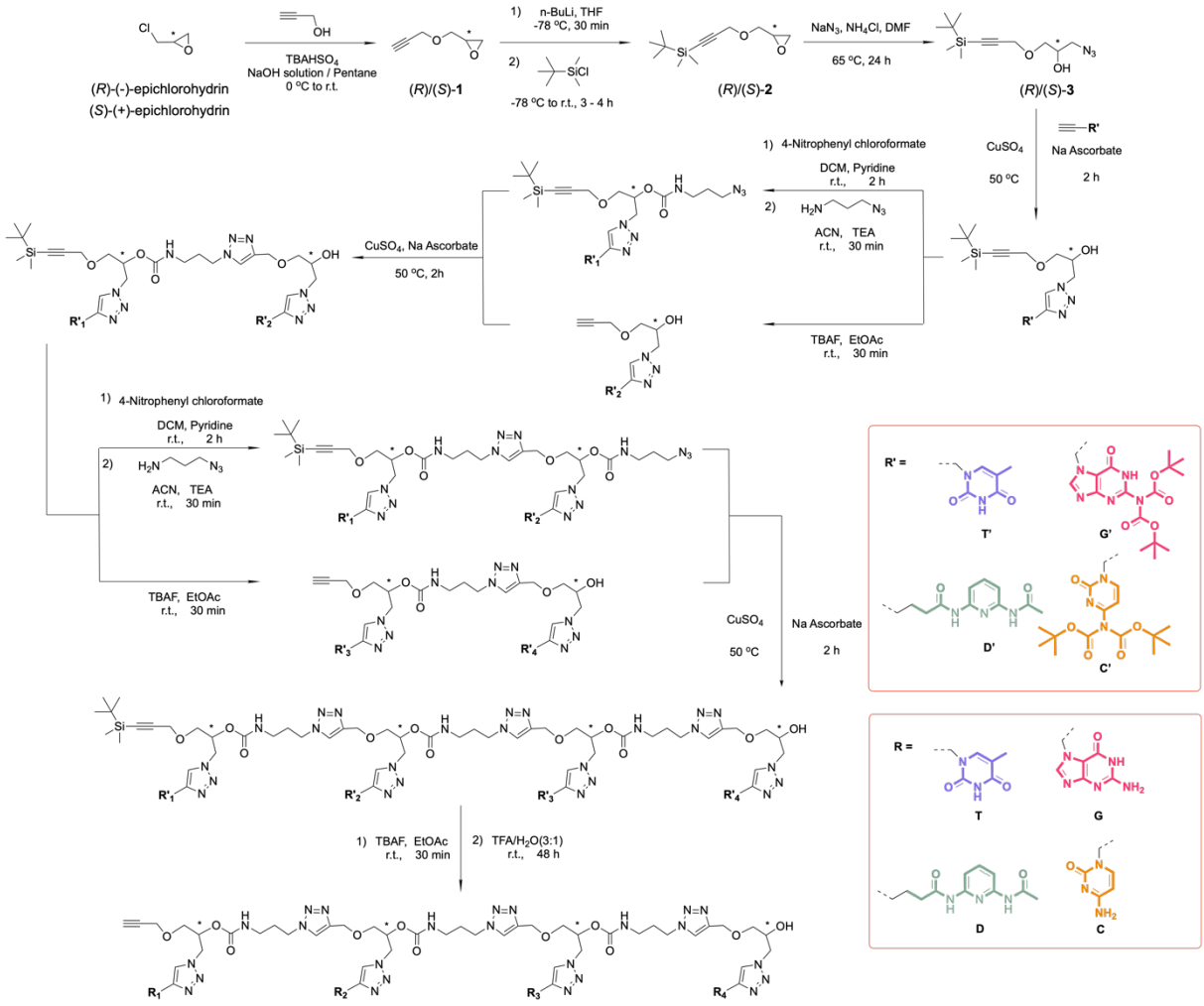
To answer this question, we synthesized flexible precise oligomers with a well-defined sequence of nucleobases. A variety of synthetic routes towards precision oligomers are available, as reviewed in recent articles;[31–35] they differ in their scalability, tolerance towards different functional groups, chemical stability of the products, and ability to control chirality. Considering our aims, we selected a flexible and tolerant synthetic route which offers a complete control over molar mass, monomer sequence, and chirality, as demonstrated in our previous work (Scheme 1).[36,37] Molecular dynamics (MD) simulations of short oligomers prepared by this route have shown that they tend to form random coils of small radius of gyration (in the range of 1 nm and below) owing to the flexibility of their urethane-triazole backbone.[36–39] Therefore, owing to our precise control over sequence and configuration, conformational blurring will be the only scrambling mechanism that contributes to complicate sequence recognition in this system.

Here, we use this route to synthesize a series of complementary and non-complementary nucleobase tetramers, and we show by MD simulations that chain pairing occurs in solution irrespective of the sequence-complementarity of the chains. We confirm this conclusion by studying with *in situ* ellipsometry the binding of complementary and non-complementary probe chains in solution to a layer of target chains grafted in the mushroom regime on a silicon surface. We then show how the specificity (or selectivity) of the recognition process can be substantially improved by playing with excluded volume interactions in the grafted layer using the grafting density of the chains as controlling parameter.

Results and Discussion

Synthesis of precision oligomers decorated with nucleobases or analogues

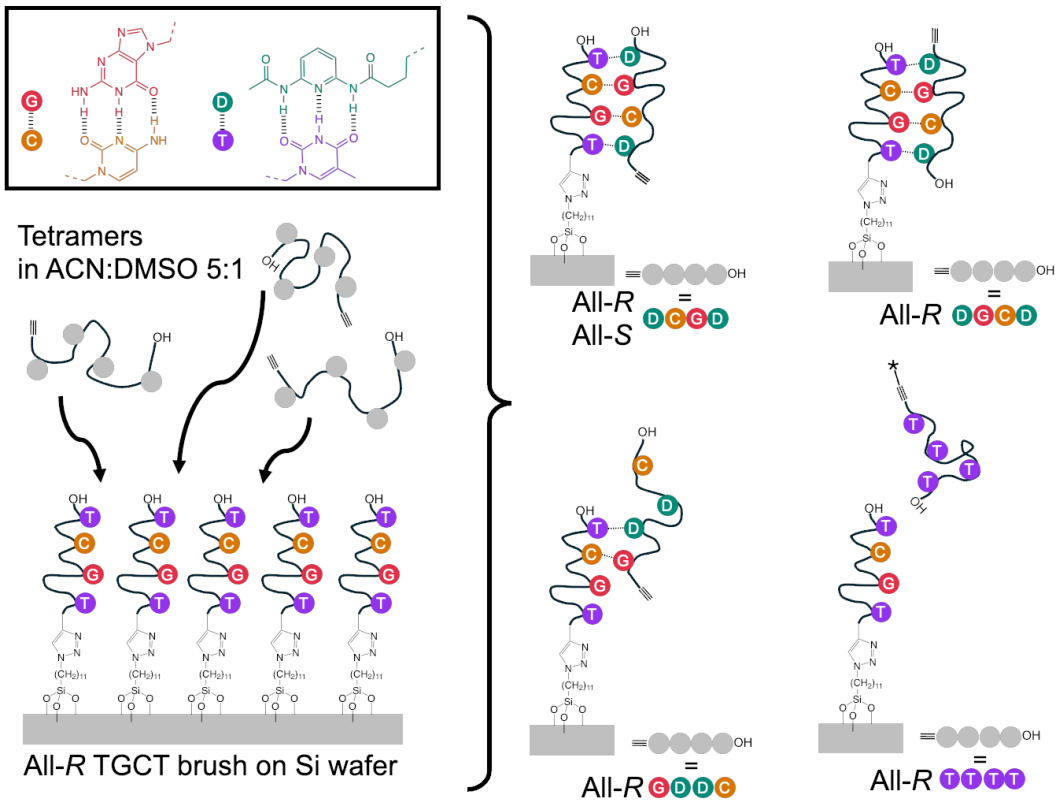
The synthesis of sequence-defined oligo(urethane-triazole) tetramers of controlled chirality was adapted from our previous work (Scheme 1 and ESI section S2).[36,37,39] The monomers are equipped with guanine (G), cytosine (C), a 2,6-diaminopyridine xenobase (D) or thymine (T) side groups (Scheme 1); trivalent H-bonded duplexes can be formed between the complementary base pair analogues C and G, or D and T (Scheme 2).



Scheme 1. General synthetic route of the nucleobase-functionalized oligo(triazole-urethane) tetramers.

Six different tetramers were obtained from these monomer units; they are designed by a succession of letters corresponding to the successive nucleobases, starting from the alkyne end group and ending in the hydroxyl terminal group (which defines the direction of the chain). An all-*R* TGCT tetramer was selected as target chain to be end-grafted on a silicon surface (Scheme 2). Then, we prepared three probe oligomers that have a sequence fully complementary to the target chain when in extended conformation but differ in their chirality or chain direction: all-*R* DCGD, all-*R* DGCD (flipped version of the previous one) and all-*S* DCGD tetramers. Additionally, we prepared a permuted all-*R* GDDC sequence which can only form a partial complementary duplex with the TGCT target chain unless a strongly folded conformation is taken (Scheme 2). Finally, a non-complementary all-*R* TTTT chain which cannot

form a duplex with the target chain was also synthesized as a negative control probe chain. The chain structures were checked by nuclear magnetic resonance (NMR) and tandem mass spectrometry as fully detailed in ESI sections S2 and S3.



Scheme 2. Different sequences of synthesized tetramers and schematics of possible binding patterns between the target grafted chains and the different probe chains. The pattern of hydrogen bonding between the nucleobases or analogues installed on the tetramers is displayed in the top left box. The alkyne end-group of the TTTT oligomer was not deprotected (as indicated by the '*') to improve solubility; it is terminated by a *tert*-butyl-silyl group.

Binding constants of the monomer units

NMR titration and dilution experiments at different temperatures were used to investigate H-binding between the monomers in a 5:1 v:v mixture of acetonitrile (ACN) and dimethyl sulfoxide (DMSO) (ESI, section S4). DMSO was added to ensure complete solubility of all monomer units. The binding constants and standard binding enthalpies and entropies of the different monomer units at 295K are given in Table 1. From these values, an estimation for the binding constant between the probe and target chains when four nucleobases are paired can be obtained if we assume that the pairing of a nucleobase is independent of the pairing of neighboring nucleobases. In this case, the chain binding constant is the product of the binding constants of each pair of nucleobases, leading to *ca.* 11000 L/mol at 295K. This estimation is very imprecise, since it ignores other interactions between chains, the decrease of the number of degrees of freedom resulting from the restriction of the conformational space of the chains upon complexation, the effects of chirality and chain direction, and synergetic binding effects.

Table 1. Binding constants (K_a) and standard binding enthalpies and entropies of the different monomer units, obtained from NMR measurements performed in deuterated ACN:DMSO 5:1 v:v at 295K. "-" indicates that no measurement was performed but that the parameter is likely to be negligeable; “0” indicates that a measurement was performed but that the variations of NMR shift were too small to extract a significant value for the parameter. The selected standard concentration is 1 M.

Complex	Ka (L/mol)	Δ H° (kJ/mol)	Δ S° (J/K·mol)
G-G	1.4±0.2	-	-
C-C	0.1±0.1	-	-
T-T	0	-	-
D-D	0	-	-
D-T	0.99±0.06	-24±4	-43±14
G-C	104±15	-5.7±2.3	-19±8
D-G	0	-	-

Molecular dynamics (MD) simulations of the interaction between probe and target chains in solution

MD simulations of assemblies comprising the target chain in the presence of each of the probe chains were performed over 1 µs time scale with a 1 fs time step, using the AMBER package (details in ESI, section 5).[40] The simulations were carried out starting from the target chain and the probe chain relatively distant (around 25 Å), to avoid any bias in the initial intermolecular contacts. All the simulations were carried implicitly in acetonitrile, using a Generalized Born solvation model considering the acetonitrile dielectric constant (37.5 at room temperature).[41–43] This implicit solvent model is accurate to model aprotic solvents and permits a large conformational sampling at reasonable computational cost. At the early stages of the MD simulations, the probe and target oligomers were found to rapidly interact (in less than 1 ns) to form heteromolecular assemblies. In each assembly, both chains were found to be very flexible and fold into compact globular shapes, with an average radius of gyration R_g ~0.85 nm (ESI Fig.S5.1). The flexibility of the oligomeric chains is also evident from the large variations of the chain end-to-end distances with simulation time, ranging from 0.3 nm to *ca.* 3.9 nm (ESI Fig.S5.2-3). Additionally, the chain-ends that are far apart in the primary sequence remain, on average, close in the 3D structure (average chain end-to-end distances of *ca.* 1.3 and 1.4 nm for target and probe chains, respectively) (ESI Table S5.1). On average, the assembly of the probe and target chains form compact assemblies in which the two oligomers remain close together but keep a high flexibility; a snapshot is shown in Fig.1.

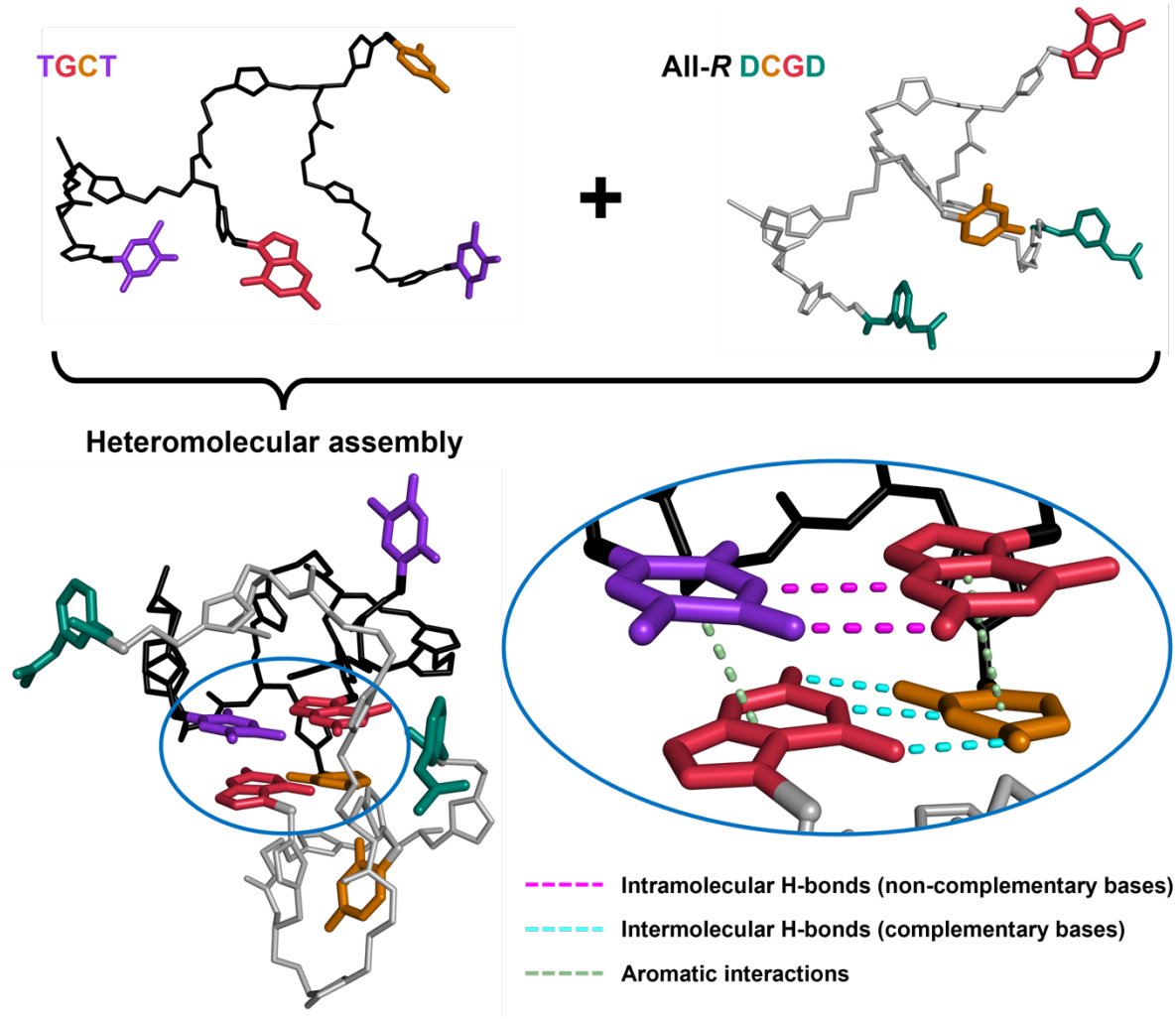


Figure 1. MD snapshots of the heteromolecular assembly of the target chain, TGCT, and one of the probe chains, all-*R* DCGD. The backbones of TGCT and all-R DCGD are coloured in black and grey, respectively. The nucleobases are colored as follows: thymine (T) in purple, guanine (G) in red, cytosine (C) in orange and the xenobase 2,6-diaminopyridine (D) in blue-green. Top: starting structure of the target and probe chains, before their assembly and folding. Bottom: one conformation of the heteromolecular assembly, showing its globular shape. A zoom shows the four bases within the assembly and the coloured dots highlight different types of interactions.

The occurrence of H-bonds and parallel stacking of aromatic cycles was computed for each MD snapshot (total of 4000 conformations for each assembly), with on average 4.5-4.7 aromatic interactions and *ca.* 4.5 H-bonds per configuration. Less than 20 % of these H-bonds occur between complementary base pairs (ESI Fig.5.6). Hence, in acetonitrile, the oligomers do not form well-organized duplexes with a regular pairing, but rather form disordered assemblies, regardless of the complementarity of the monomer sequences. This is due to the high intrinsic flexibility of the chain backbones, the formation of a competitive network of H-bonds (involving H-bonding units such as triazole, carbamate and non-complementary bases), and significant folding promoted by the stacking of aromatic cycles (Figure 1 bottom). This folding also brings numerous intramolecular interactions (around 40 % of the total H-bonds and

around 46 % of the total aromatic interactions). These results indicate that a specific recognition of complementary tetramers is prevented by the flexibility of the backbone of the oligomers which spatially blurs their sequence.

We surmised that a better recognition could be achieved by grafting the target chains on a surface, which should restrict the available conformational space of the chains provided the grafting density be large enough. Indeed, in the brush regime in which the average distance between the chains becomes lower than their unperturbed radius of gyration, the chains should on average stretch vertically, resulting in a vertical spatial distribution of the nucleobases on average closer to the chemical sequence of the chains.

Grafting of target oligomers onto azido-silanized wafers

Therefore, we grafted the all-*R* TGCT target chains onto azido-silanized silicon wafers, using their alkyne terminal groups to click them unto a monolayer of 11-azidoundecyltrimethoxysilane (AzUTMS**)**. A reproducible AzUMTS monolayer could be obtained over the native oxide of Si wafers by overnight gas phase silanation at 100°C, using a well-established silanation protocol.[44–49] XPS confirmed the grafting of the silane onto the silicon wafer (ESI section 7). Measurements of the layer thickness by spectroscopic ellipsometry provided an average thickness of 1.1±0.03 nm (standard deviation 0.32 nm for 124 different samples prepared at different times, each measured in triplicate). Additionally, x-ray reflectivity (XRR) was performed on 35 different samples (ESI Fig.S6.1); from the first minimum arising from destructive interference in the reflectivity curves, an average thickness of 0.9±0.02 nm (standard deviation of 0.12 nm) was found. The XRR data was fit to extract electron density profiles (ESI Fig.S6.1) as described in previous work,[47] from which an average electron density profile was computed (Fig.2). An average grafting density of 2.9±0.1 molecules/nm^2 (0.32 molecules/nm^2 standard deviation) was obtained by integration of the profiles after subtracting the one of a clean Si wafer, considering that each grafted AzUMTS molecule contains 147 electrons. This is a typical grafting density for silanation in the gas phase.[44] As a comparison, the chain areal density of a polyethylene crystal over its *ab* plane (perpendicular to the chain axes) is *ca.* 5.5 molecules/nm^2; hence, the layer is made of chains in a partially coiled conformation. The average water contact angle of 82.8±0.2° (standard deviation 2.1° for five measurements on 46 samples) is in good agreement with what was reported before by Vos *et al*.[50]

The alkyne-ended all-*R* TGCT target chains were then grafted by Copper(I)-catalyzed Azide-Alkyne Cycloaddition (CuAAC), coupling the terminal alkyne group of the all-*R* TGCT oligomers with the azide-terminated AzUTMS layer. The grafting was performed overnight at 80°C in four solvents, DMF, DMF:DMSO 75:25, DMF:DMSO 50:50 v:v and DMSO. XPS again confirmed the successful grafting of the TGCT chains (ESI section 7). Spectroscopic ellipsometry indicated an average increase of the layer thickness upon reaction (Table 2) from 0.59±0.06 nm in DMF to 1.55±0.06 nm in DMSO, revealing a higher number of grafted chains when the click reaction is performed in the presence of increasing DMSO content. Consistently, the water contact angle (Table 2) decreased from 60.5±1.1° in DMF to 40.5±0.4° in DMSO, due to the higher number of grafted chains whose terminal ends comprise a hydrophilic hydroxyl group and a polar triazole-thymine side chain. The differences in the density of the brush depending on the solvent used for the grafting process result from differences in the ratio between intramolecular segmental interactions and segment/solvent interactions depending on solvent

quality. DMSO being a better solvent, it favours segment/solvent interactions, brush swelling and penetration of oligomers during the grafting-to process, resulting in a higher final grafting density.

Table 2. Morphological characterization of the target layer composed of grafted all-*R* TGCT molecules, depending on the solvent used for the click reaction.

DMF:DMSO vol:vol	*d* (nm)[(a)]	error[(b)]	*N*[(c)]	σ (molec/nm^2)[(d)]	error[(b)]	*N*[(c)]	*x* (nm)[(e)]	θ (°)[(f)]	error[(b)]	N[(c)]
100:0	0.62	0.06	15	0.30	0.11	2	1.8	60.5	1.1	25
75:25	1.13	0.12	12	0.60[(g)]	0.06[(g)]	4[(g)]	1.3[(g)]	51.9	2.1	19
50:50	1.26	0.10	15					47.1	1.1	20
0:100	1.55	0.06	63	0.74	0.03	8	1.2	40.5	0.4	70

(a) Average ellipsometric thickness of the grafted all-*R* TGCT layer; (b) standard error on the measurement (same units); (c) number of measurements; (d) XRR-determined average number of grafted all-*R* TGCT molecules per nm^2; (e) average distance between grafted all-*R* TGCT molecules, from XRR; (f) water contact angle; (g) the density profiles of samples clicked in solvents containing from 25 to 50 vol% of DMSO were averaged together.

The electron density profiles of all-*R* TGCT grafted layers prepared in different solvents were obtained by XRR (Fig.2), confirming the increasing grafting density when increasing the proportion of DMSO in the solvent. Owing to the limited precision of XRR, and the proximity of the ellipsometry-derived thickness and contact angles of samples clicked in solvents having from 25 to 50% DMSO, the electron density profiles of these samples were averaged together. For pure DMF and DMSO, the profiles of different samples were also averaged to increase precision.

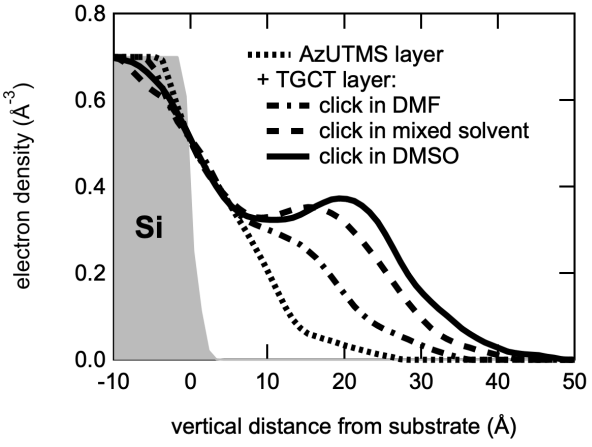


Figure 2. Average electron density profiles of a bare Si wafer (gray), a wafer silanized by AzUTMS, and a silanized wafer reacted with the all-*R* TGCT oligomer by CuAAC in DMF, DMSO, or mixtures of DMF and DMSO containing from 25 to 50% DMSO. All XRR data and fits are shown in Fig.S6.2 of the ESI.

The average grafting densities were then obtained by subtracting the average electron density profile of the AzUTMS-silanized wafer, integrating the subtracted profiles and dividing by the number of electrons per TGCT molecule (874). This provided an average grafted density σ increasing from 0.30±0.11 molecule per nm^2 when the click reaction is performed in pure DMF to 0.74±0.03 molecule per nm^2 in DMSO (Table 2). This corresponds to average distances $x = \sigma^{-1/2}$ between all-*R* TGCT grafted chains of *ca.* 1.2 and 1.8 nm for DMSO and DMF,

respectively, with an intermediate value for the mixed solvents. Given a radius of gyration R_g of 0.85 nm from the MD simulations, the oligomer layers are in the brush regime ($d < 2R_g$) when the click reaction is performed in pure DMSO, and in the mushroom regime ($d > 2R_g$) when performed in pure DMF.[51] Hence, by selecting the click conditions, the TGCT chains can be placed under different lateral constraints, corresponding to more or less stretched chain conformations, with consequently different degrees of spatial memory loss of the primary sequence.

Recognition between sequence-defined oligomers and the grafted all-R TGCT layers

The recognition experiments were performed at 25°C in a temperature-controlled ellipsometry flow cell (Fig.3). Importantly, all experiments were performed in ACN with 16.6 vol% of DMSO added to ensure full solubility of the tetramers, irrespective of the solvent previously used for the grafting of the target layer. The experimental procedure involved flowing pure degassed solvent in the ellipsometer cell, measuring a complete spectroscopic ellipsometry spectrum of the grafted target layer (A in Fig.3), switching to a 0.4 mM solution of a probe tetramer in the solvent while monitoring the evolution of the ellipsometric angle Δ at 400 nm with time until a plateau was reached, switching to the pure solvent while monitoring Δ at 400 nm until a plateau was reached, and finally measuring the complete ellipsometry spectrum of the target layer plus irreversibly-adsorbed probe layer in the pure solvent (B in Fig.3).

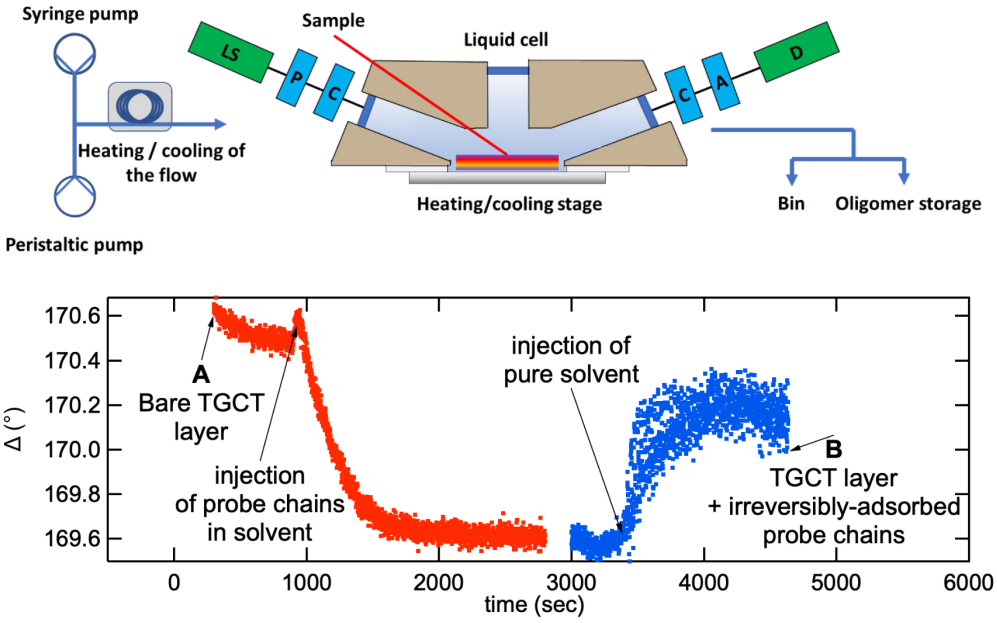


Figure 3. Top. Scheme of the ellipsometry flow cell used for the measurement of adsorption (LS: light source, P: polarizer, C: (rotating) compensator, A: analyzer, D: detector). **Bottom**. Typical variation of the ellipsometric angle Δ measured at 400 nm during adsorption at 25°C of a probe tetramer (here, all-*R* DCGD) on a substrate (here, a grafted target layer of all-*R* TGCT of 0.74 molecule/nm^2 grafting density) from a 0.4 mM solution in ACN:DMSO 5:1. A and B indicate the times at which complete ellipsometric spectra were measured (two zone averaging).

The volume of irreversibly-adsorbed probe tetramer per surface area of target layer (nm^3/nm^2) was computed as $V_{irr} = c(\overline{\Delta_B} - \overline{\Delta_A})$, in which $\overline{\Delta_A}$ and $\overline{\Delta_B}$ are the ellipsometric angles measured at time *A* and *B*, respectively, averaged over the 400-700 nm wavelength range. This corresponds to a Taylor expansion limited to first order of the ellipsometric angle versus surfacic adsorbed volume, which is valid for limited thicknesses of the adsorbed layer as in our case. The calibration constant *c*=-1/1.2615 nm/° was obtained from simulations of ellipsometric angles depending on thickness and solvent content of the adsorbed layer (ESI, section 8).

The results gathered in Fig.4 (left) show that substantial irreversible binding occurs, confirming that the equilibrium is strongly displaced towards complex formation despite the relatively moderate binding constant estimated above. This is due to the locally very high local concentration of nucleobases in the brush, to the existence of other types of H-bonds in the system as shown by MD simulations, as well as to the low probability to have a simultaneous release of the four pairs of nucleobases.

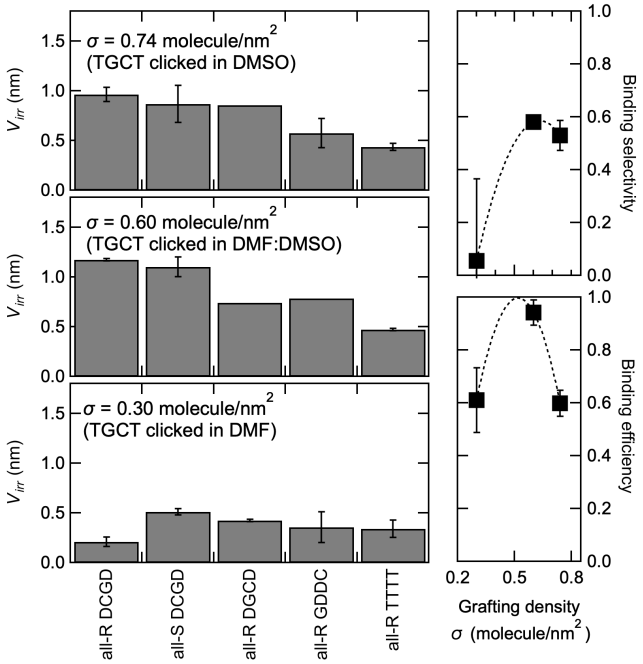


Figure 4. (**left**) Volume V_{irr} of probe chains irreversibly-adsorbed per area of target chain layer, for different grafting densities σ of the grafted all-*R* TGCT target chains, depending on the sequence of the probe chains. The experiments were performed at 25°C in ACN:DMSO 5:1 at a 0.4 mM concentration of probe chains. (**right**) Binding selectivity and binding efficiency of the target layer depending on its grafting density; the dashed lines are parabolic fits drawn to guide the eye.

We computed two indicators that quantify the recognition of the target layer by probe chains. First, we defined the *binding efficiency E* as the number of irreversibly-bound probe chains per target chain in the layer, for fully complementary target and probe chains. It can be estimated by the ratio of the average surfacic volume V_{irr} of irreversibly adsorbed DCGD oligomers irrespective of chirality (hence probe chains having strictly a complementary sequence and same direction as the target chains), to the dry thickness *d* of the target layer:

$$E = \frac{V_{irr(R-DCGD)} + V_{irr(S-DCGD)}}{2d_{R-TGCT}}$$

An efficiency *E*=1 is obtained when each TGCT target chain is complexed by a complementary DGCD probe chain, whatever their chirality; in contrast, when no complementary probe chain irreversibly adsorbs in the target layer, *E*=0.

The second indicator is the *binding selectivity S*, defined as one minus the ratio of the average irreversibly adsorbed surfacic volume of fully non-complementary TTTT oligomers to the one of fully complementary DCGD oligomers (whatever their chirality):

$$S = 1 - \frac{2V_{irr(R-TTTT)}}{V_{irr(R-DCGD)} + V_{irr(S-DGCD)}}$$

A selectivity *S*=0 indicates that fully non-complementary TTTT chains adsorb identically to fully complementary DCGD chains, whereas *S*=1 when non-complementary chains do not adsorb at all. The binding efficiency and selectivity are displayed in Figure 4 (right) versus the grafting density of the target layer.

Regarding the selectivity of the irreversible adsorption, when the all-*R* TGCT target layer is in the mushroom regime after having been clicked in DMF (bottom panel in Fig.4 left), the adsorption is not sequence-specific with, *e.g.*, complementary all-*R* DCGD probe chains being even less irreversibly-adsorbed than non-complementary all-*R* TTTT probe chains. This is in good agreement with expectations from the MD simulations. In contrast, for all-*R* TGCT target chains in the brush regime (grafted in DMSO, top panel in Fig.4 left), complementary probe chains (all-*R* DCGD, all-*S* DCGD, and even all-*R* DGCD) are more bound to the layer, irrespective of their chirality or directionality, compared to all-*R* GDDC probe chains with a permuted sequence or completely non-complementary all-*R* TTTT probe chains. The indicator of binding selectivity in Figure 4 (top right) confirms this conclusion, with almost no selectivity at the lower grafting density whereas values of *ca.* 0.6 are found for denser brushes.

Hence, these results indicate some sequence selectivity in the recognition process when the target chains adopt a more stretched average conformation in denser brushes, although it is not perfect since fully non-complementary all-*R* TTTT probe chains still bind to the all-*R* TGCT target layer (*ca*. 50% less than the complementary probe chains). Clearly, the lateral constraints due to volume exclusion in the brush, which result in thicker layers corresponding to more extended average TGCT chain conformations, favour the sequence-specificity of the recognition process. However, this effect remains moderate, with probe chains having the proper sequence only adsorbing approximately twice as much as chains of improper sequence.

It would thus be tempting to try to increase further the grafting density of the target chains to stretch them even more. This would not be easy to perform in a grafting-to process and might require grafting-from strategies which are known to generally result in higher grafting densities. However, this is not likely to be of much interest, considering the way binding efficiency varies with grafting density (Figure 4, bottom right): when increasing too much the grafting density of the target chains, the binding efficiency decreases due to increased

excluded volume interactions within the brush, which limits the penetration of probe chains in the brush. As a result, even though the binding selectivity might be better, the number of probe chains bound per target chain will decrease, which will impair applications requiring the binding of a substantial number of chains.

Conclusions

In this work, we have shown that flexible precision tetramers substituted by nucleobases or analogs can be synthesized by a versatile route leading to alkyne-ended chains that can be grafted onto azide-modified silicon wafers in configurations ranging from mushroom to brush layers. When in the mushroom regime, the chains do not exhibit any recognition specificity towards probe chains in solution. This is also the conclusion drawn from molecular modelling simulations of the recognition process in solution, the reason being the existence of numerous competing H-bonding interactions with the triazole-urethane backbone of the oligomers as well as the complete scrambling of the spatial positioning of the nucleobases due to the flexibility of the backbone.

In contrast, when the target chains are grafted in the brush regime, the recognition process becomes more specific, with probe chains of complementary sequence being twice more probable to be irreversibly-adsorbed than non-complementary sequences. This undoubtedly results from lateral volume interactions between chains leading to partial chain stretching and a decreased scrambling of the information contained in the chemical sequence of the chains. However, although encouraging, this remains much less specific than DNA, in which the charged chain backbone is of higher rigidity due to electrostatic repulsion, and for which the distance between successive nucleobases is smaller than in our synthetic system by a factor of *ca*. 2. Simulations could be extended to other types of synthetic backbones and recognition units to improve selectivity in the brush regime.

It is thus very improbable that applications similar to DNA origami, which need a high binding selectivity of specific nucleobase sequences in the presence of many competitors, could be transposed to more flexible precision oligomers. Nevertheless, the irreversible binding of probe and target chains we have shown is probably strong enough for simpler applications such as the formation of superlattices of nanoparticles, because such applications do not require to single out a specific sequence among many. Further work in these directions is thus of interest to explore more completely the limits and advantages of flexible precision oligomers as alternatives to DNA strands.

Author Contributions (CRediT)

Investigation and Formal analysis: K.Gr. and A.M.J. (XRR and ellipsometry experiments), Q.Q. and L.J. (synthesis and characterization of the molecules), D.D. and M.S. (molecular modeling simulations); Methodology: K.Gr. and D.M. (*in-situ* ellipsometry); Supervision: K.G., M.S. and A.M.J.; Writing – original draft: K.Gr., D.D., A.M.J.; Writing – review & editing: K.G., M.S., A.M.J.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data supporting this article have been included as part of the Supplementary Information.

Acknowledgments

We acknowledge the Fonds de la Recherche Scientifique - FNRS and the Fonds Wetenschappelijk Onderzoek under EOS project n° 30650939 for financial support. We also acknowledge support from the ASM and SUCH platforms of UCLouvain for MS, NMR and XPS analyses. Computational resources are provided by the Consortium des Equipements de Calcul Intensif (CECI), funded by FNRS (Grant No. U.G.018.18) and Wallonia Region. K.G. and M.S. are Research Directors of the F.R.S.−FNRS. D.D. is supported by a FRIA grant funded by F.R.S.-FNRS.

References

1 G. Tikhomirov, P. Petersen and L. Qian, *Nature*, 2017, **552**, 67–71.

2 K. F. Wagenbauer, C. Sigl and H. Dietz, *Nature*, 2017, **552**, 78–83.

3 P. W. K. Rothemund, *Nature*, 2006, **440**, 297–302.

4 G. Chen, K. J. Gibson, D. Liu, H. C. Rees, J.-H. Lee, W. Xia, R. Lin, H. L. Xin, O. Gang and Y. Weizmann, *Nat. Mater.*, 2018, **18**, 1–8.

5 G. Yao, J. Li, Q. Li, X. Chen, X. Liu, F. Wang, Z. Qu, Z. Ge, R. P. Narayanan, D. Williams, H. Pei, X. Zuo, L. Wang, H. Yan, B. L. Feringa and C. Fan, *Nat. Mater.*, 2020, 1–10.

6 E. Auyeung, T. I. N. G. Li, A. J. Senesi, A. L. Schmucker, B. C. Pals, M. O. de la Cruz and C. A. Mirkin, *Nature*, 2014, **505**, 73–77.

7 M. R. Jones, R. J. Macfarlane, B. Lee, J. Zhang, K. L. Young, A. J. Senesi and C. A. Mirkin, *Nature*, 2010, **9**, 913–917.

8 R. J. Macfarlane, B. Lee, M. R. Jones, N. Harris, G. C. Schatz and C. A. Mirkin, *Science*, 2011, **334**, 204–208.

9 P. J. Santos, P. A. Gabrys, L. Z. Zornberg, M. S. Lee and R. J. Macfarlane, *Nature*, 2021, **591**, 586–591.

10 K. J. Si, Y. Chen, Q. Shi and W. Cheng, *Adv. Sci.*, 2018, **5**, 1700179.

11 M. Xie, J. Jiang and J. Chao, *Sensors*, 2023, **23**, 9229.

12 S. Julin, S. Nummelin, M. A. Kostiainen and V. Linko, *J. Nanoparticle Res.*, 2018, **20**, 119.

13 D. J. Lewis, D. J. D. Carter and R. J. Macfarlane, *J. Am. Chem. Soc.*, 2020, **142**, 19181–19188.

14 S. Wang, X. Xie, Z. Chen, N. Ma, X. Zhang, K. Li, C. Teng, Y. Ke and Y. Tian, *Int. J. Mol. Sci.*, 2021, **22**, 7558.

15 D. Woods, D. Doty, C. Myhrvold, J. Hui, F. Zhou, P. Yin and E. Winfree, *Nature*, 2019, **567**, 366–372.

16 P. Yakovchuk, *Nucleic Acids Res.*, 2006, **34**, 564–574.

17 B. Tinland, A. Pluen, J. Sturm and G. Weill, *Macromolecules*, 1997, **30**, 5763–5765.

18 H. Zeng, X. Yang, A. L. Brown, S. Martinovic, R. D. Smith and B. Gong, *Chem. Commun.*, 2003, 1556–1557.

19 A. Marquis, V. Smith, J. Harrowfield, J. Lehn, H. Herschbach, R. Sanvito, E. Leize-Wagner and A. V. Dorsselaer, *Chem. Eur. J.*, 2006, **12**, 5632–5641.

20 B. Gong, *Acc. Chem. Res.*, 2012, **45**, 2077–2087.

21 D. Núñez-Villanueva, G. Iadevaia, A. E. Stross, M. A. Jinks, J. A. Swain and C. A. Hunter, *J. Am. Chem. Soc.*, 2017, **139**, 6654–6662.

22 Y. Zhang, R. Cao, J. Shen, C. S. F. Detchou, Y. Zhong, H. Wang, S. Zou, Q. Huang, C. Lian, Q. Wang, J. Zhu and B. Gong, *Org. Lett.*, 2018, **20**, 1555–1558.

23 J. A. Swain, G. Iadevaia and C. A. Hunter, *J. Am. Chem. Soc.*, 2018, **140**, 11526–11536.

24 D. Núñez-Villanueva and C. A. Hunter, *J. Am. Chem. Soc.*, 2022, **144**, 17307–17316.

25 D. H. Appella, *Curr. Opin. Chem. Biol.*, 2009, **13**, 687–696.

26 H. V. Nguyen, Z. Zhao, A. Sallustrau, S. L. Horswell, L. Male, A. Mulas and J. H. R. Tucker, *Chem. Commun.*, 2012, **48**, 12165–12167.

27 P. Karri, V. Punna, K. Kim and R. Krishnamurthy, *Angew. Chem. Int. Ed.*, 2013, **52**, 5840–5844.

28 L. Maes, D. M. Roqeuro, L. M. Pitet, P. Adriaensens and T. Junkers, *Polym. Chem.*, 2020, **11**, 2027–2033.

29 J. Li, Z. Wang, Z. Hua and C. Tang, *J. Mater. Chem. B*, 2020, **8**, 1576–1588.

30 P. E. Nielsen, *Chem. Biodivers.*, 2010, **7**, 786–804.

31 M. A. R. Meier and C. Barner-Kowollik, *Adv. Mater.*, 2019, **31**, 1806027.

32 R. Aksakal, C. Mertens, M. Soete, N. Badi and F. D. Prez, *Adv. Sci.*, 2021, **65**, 2004038.

33 B. van Genabeek, B. A. G. Lamers, C. J. Hawker, E. W. Meijer, W. R. Gutekunst and B. V. K. J. Schmidt, *J. Polym. Sci.*, 2021, **59**, 373–403.

34 Z. Deng, Q. Shi, J. Tan, J. Hu and S. Liu, *ACS Mater. Lett.*, 2021, **3**, 1339–1356.

35 J.-F. Lutz, *Eur. Polym. J.*, 2023, **199**, 112465.

36 J. Li, M. Leclercq, M. Fossepré, M. Surin, K. Glinel, A. M. Jonas and A. E. Fernandes, *Polym. Chem.*, 2020, **11**, 4040–4046.

37 Q. Qin, J. Li, D. Dellemme, M. Fossépré, G. Barozzino-Consiglio, I. Nekkaa, A. Boborodea, A. E. Fernandes, K. Glinel, M. Surin and A. Jonas, *Chem. Sci.*, 2023, **14**, 9283–9292.

38 S. Kardas, M. Fossépré, V. Lemaur, A. E. Fernandes, K. Glinel, A. M. Jonas and M. Surin, *J. Chem. Inf. Model.*, 2022, **62**, 2761–2770.

39 J. Li, Q. Qin, S. Kardas, M. Fossépré, M. Surin, A. E. Fernandes, K. Glinel and A. M. Jonas, *ACS Catal.*, 2022, **12**, 2126–2131.

40 D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668–1688.

41 G. D. Hawkins, C. J. Cramer and D. G. Truhlar, *Chem. Phys. Lett.*, 1995, **246**, 122–129.

42 G. D. Hawkins, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem.*, 1996, **100**, 19824–19839.

43 V. Tsui and D. A. Case, *Biopolymers*, 2000, **56**, 275–291.

44 A. Pallandre, K. Glinel, A. M. Jonas and B. Nysten, *Nano Lett.*, 2004, **4**, 365–371.

45 A. M. Jonas, Z. Hu, K. Glinel and W. T. S. Huck, *Macromolecules*, 2008, **41**, 6859–6863.

46 A. M. Jonas, Z. Hu, K. Glinel and W. T. S. Huck, *Nano Lett.*, 2008, **8**, 3819–3824.

47 X. Pei, A. Fernandes, B. Mathy, X. Laloyaux, B. Nysten, O. Riant and A. M. Jonas, *Langmuir*, 2011, **27**, 9403–9412.

48 O. Deschaume, D. Magnin, Z. A. Cheng, C. Douchamps, P. Labbé, S. Yunus, M.-C. Durrieu, B. Nysten, K. Glinel, S. Demoustier-Champagne and A. M. Jonas, *Biomacromolecules*, 2014, **15**, 3706–3716.

49 A. Dirani, A. E. Fernandes, D. R. Wong, P. Lipnik, C. Poleunis, B. Nysten, K. Glinel and A. M. Jonas, *Langmuir*, 2014, **30**, 10057–10065.

50 R. Vos, T. Steylaerts, A. Franquet, A. Moussa, T. Stakenborg and K. Jans, *Sol. St. Phen.*, 2018, **282**, 31–36.

51 R. R. Netz and D. Andelman, *Phys. Rep.*, 2003, **380**, 1–95.

Supporting Information

Chain stretching in brushes favors sequence recognition for nucleobase-functionalized flexible precise oligomers

Kseniia Grafskaia,[(1,†)] Qian Qin,[(1,†)] Jie Li,[(1,††)] Delphine Magnin,[(1)] David Dellemme,[(2)] Mathieu Surin,[(2)] Karine Glinel,[(1,*)] and Alain M. Jonas[(1,*)]

(1) Institute of Condensed Matter and Nanosciences, Université catholique de Louvain, Croix du Sud 1 L7.04.02, Louvain-la-Neuve, Belgium

(2) Laboratory for Chemistry of Novel Materials, Université de Mons – UMONS, Avenue Maistriau, 17, B-7000 Mons, Belgium

(*) Corresponding authors

(†) These authors contributed equally to this work

(††) Present address: State Key Laboratory of Chemical Resource Engineering, Beijing Laboratory of Biomedical Materials, Beijing University of Chemical Technology, Beijing 100029, P. R. China

Table of Contents

S1. General considerations and instrumentation

Reagents were obtained from commercial sources and used without further purification. All reactions were carried out under argon. Flash column chromatography was carried out using Silica gel 230-400 mesh (Sigma-Aldrich) as the stationary phase. Milli-Q water (resistivity 18.2 MΩ.cm) was obtained from a Millipore system.

NMR spectra were recorded on Bruker-300 and Bruker-500 spectrometers. Chemical shifts (δ) are reported in parts per million (ppm) from low to high field and referenced to residual solvent. Coupling constants (J) are reported in hertz (Hz). Standard abbreviations indicating multiplicity are used as follows: br= broad, s= singlet, d= doublet, t= triplet, q= quartet, quint= quintet, m= multiplet.

High-resolution mass spectra (HRMS) were measured on a Q-Exactive (Orbitrap) from ThermoFisher using an atmospheric pressure chemical ionization (APCI) source. Electrospray ionisation mass spectrometry (ESI-MS) and ESI-MS/MS were performed on an SYNAPT G2-Si high definition mass spectrometer (Waters) equipped with a NanoLockSpray dual electrospray ion source (Waters). Precut fused silica PicoTipR Emitters (outer diameters: 360 µm; inner diameter: 20 µm; 10 µm tip; 2.5" length (Waters)) were used to carry samples for nanoelectrospray-injecting the test solution.

LC-QTOF-MS/MS Analysis were performed on an SYNAPT G2-Si high definition mass spectrometer (Waters) equipped with a NanoLockSpray dual electrospray ion source (Waters). Samples were diluted 2 to 10 times in 50% (v/v) acetonitrile, 0.1% formic acid or in 50% (v/v) methanol depending on the quality and the intensity of the signal. Coated fused silica PicoTip™ Econo12 Emitters for nanoelectrospray, outer diameters: 1 µm Tip (New Objective, USA) were filled with 5 µL of samples and placed on the Universal NanoFlow™ Sprayer (Waters). The eluent was sprayed at a spray voltage of 2.8 kV. The source temperature was set to 100°C. The cone gas flow was 20 liter/h with a nano flow gas pressure of 0.3 bar. MS spectra were acquired and processed with MassLynx software (Waters). Full scan MS and MS2 spectra (*m/z* 50 to 2000) were in resolution mode (20,000 resolution FWHM at *m/z* 400) with a scan time of 0.1 sec. Tandem mass spectra of the precursor were generated in the trapping region of the ion mobility cell by using a collision energy ramp from 10 V to 70 V. Charged ions are selected to be submitted to the MS/MS fragmentation over the *m/z* range from 50 to 2000 with a scan time of 0.25 sec. For the post-acquisition lock mass correction of the data in the MS method, the doubly charged monoisotopic ion of [Glu^1]-fibrinopeptide B was used at 100 fmol/µL using the reference sprayer of the nanoESI source with a frequency of 30 s at 0.5 µL/min into the mass spectrometer.

X-ray photoelectron spectroscopy (XPS). XPS measurements were performed on VersaProbe III photoelectron spectrometer from Physical Electronics (USA) equipped with a monochromatized micro focused Al X-ray source (powered at 50 W). The pressure in the analysis chamber was around 10^{-6} Pa. The angle between the surface normal and the axis of the analyser lens was 45°. High-resolution scans of the C 1s, O 1s and N 1s photoelectron peaks were recorded from a spot diameter of 200 µm using a pass energy of 55 eV and a step size of 0.1 eV. Charge stabilization was achieved thanks to a combination of Argon and electron guns. The C-C component of the C1s peak of carbon was fixed to 284.8 eV to set the binding energy scale. Data treatment was performed with the CasaXPS program (Casa Software Ltd, UK). Some spectra were decomposed with the least squares fitting routine provided by the software with a Gaussian/Lorentzian (85/15) product function and after subtraction of a non-linear baseline. Molar fractions were calculated using peak areas normalised based on acquisition parameters and sensitivity factors provided by the manufacturer.

X-ray Reflectometry (XRR). XRR measurements were carried out with a modified Siemens D5000 2-circle goniometer (0.002° positioning accuracy). X-rays of 0.15418 nm wavelength (Cu Kα) were obtained from a Rigaku rotating anode operated at 40 kV and 300 mA, fitted with a collimating mirror (Osmic, Japan) delivering a close-to-parallel beam of ~0.0085° vertical angular divergence. The beam size was defined by a 40-µm-wide slit placed 17.5 cm away from the focal spot. The sample was placed within 2 µm of the center of the goniometer, and the reflected beam was collected through a 200-µm-wide detector slit. Soller slits in the incident and reflected beam limited the axial divergence to 0.02°. The data were corrected for spill over and normalized to unit incident intensity; they are reported as a function of k_{z0}, the vertical component of the photon wavevector in a vacuum.

The XRR data were analysed by procedure described elsewhere.[1] Thus, an estimation of the average thickness of the monolayer, d_x, and grafting densities σ_g from the electron density profiles $\rho(z)$ were obtained.

Water contact angle. The water contact angles were measured at ambient temperature using the sessile drop method and image analysis of the drop profile. A contact angle goniometer from OCA DataPhysics (Germany) was used. The Milli-Q water droplet volume was 0.5 µL, and the contact angle was measured 5 s after the drop was deposited on the sample. For each sample, the reported value is the average of the results obtained on at least five droplets.

In situ ellipsometry. Ellipsometry measurements were performed on Jobin-Yvon UVisel spectroscopic ellipsometer at an incidence angle of *ca*. 65°. Spectroscopic scans were recorded for wavelengths ranging from 400 to 700 nm. Kinetic measurements were performed at *ca*. 65° incidence angle and at a wavelength of 400 nm. All measurements were recorded at room temperature. An Accurion temperature-controlled liquid cell (volume 0.7 mL) was fixed on a homemade multi-axis sample stage attached to the ellipsometer. The solvent flow rate of the solvent was set to *ca*. 0.09 mL/min, with two pumps (peristaltic and syringe) used to inject the solvent and the oligomer solution, respectively. The scheme of the setup is in Fig.3 of the article.

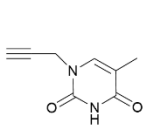
Cleaning of Silicon Wafers. One-side-polished (100) silicon wafers with their *ca*. 1.3 nm thick native oxide layer were purchased from Wacker (Germany) and cut in squares of 1.5 cm side. The wafers were cleaned by dipping into a freshly-prepared piranha solution (H_2SO_4 (98%)/H_2O_2 (30%) v/v 3:1 – *caution: piranha solution is an extremely strong oxidant and should only be handled with proper equipment*). The substrates were then rinsed 3 times with Milli-Q water and dried under a flow of nitrogen. The silicon wafers were immediately used to prevent surface contamination.

Deposition of the Silane self-assembled monolayers (AzUTMS SAMs). The 11-azidoundecyltrimethoxysilane (AzUTMS) layer was formed by gas-phase silanation in a Schlenk tube at 100 °C as described previously.[2] After reaction, the samples were washed twice with acetone and methanol. After a subsequent rinse, these were dried with nitrogen.

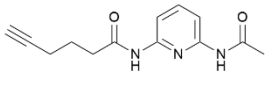
Grafting of the oligomers by CuAAC. Sequence-controlled oligomers were immobilized on the AzUTMS SAMs by Cu-Catalyzed Azide-Alkyne Cycloaddition Reactions (CuAAC), in mixtures of DMF and DMSO. Substrates with the AzUTMS SAM layer were placed in Schlenk tubes containing 13.25 mL of a DMF:DMSO mixture, 0.75 mL of a copper(I) iodide solution in the same solvent (160 mM), 1 mL of oligomer dissolved in pure DMSO (3 mM) and 95 µL of Et_3N. After overnight stirring at 80°C under argon, the substrates were rinsed with DMSO, methanol, 1M HCl, MilliQ water, MeOH, Et3N/DCM (v/v 5%) and dichloromethane before being dried with nitrogen.

S2. Synthesis of the stereo-controlled and sequenced-defined tetramers.

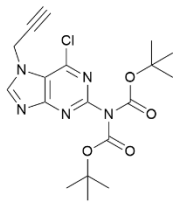
The compounds (*R*)/(*S*)-1, (*R*)/(*S*)-2 and (*R*)/(*S*)-3 (Scheme 1 of the article) were prepared according to the literature and our previous works. [1-4] The synthesis of bases and monomers in the scheme below were reported in our previous work.[3]



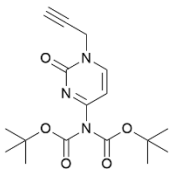
Base T



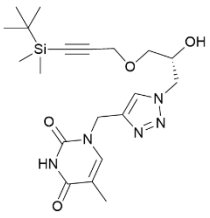
Base D



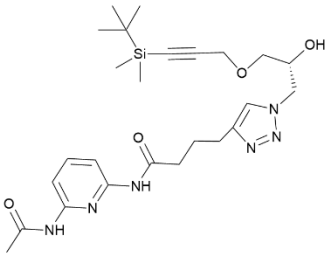
Base G



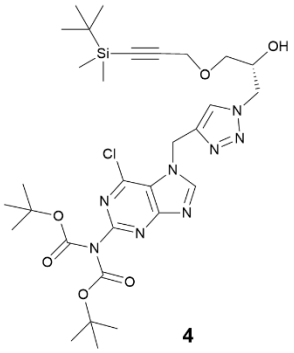
Base C



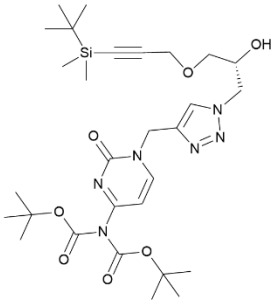
T



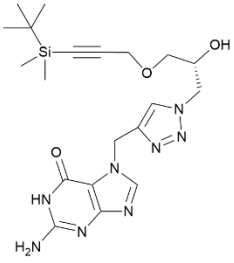
D



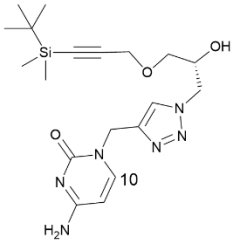
4



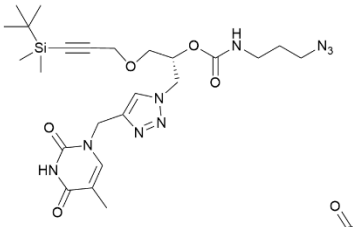
5



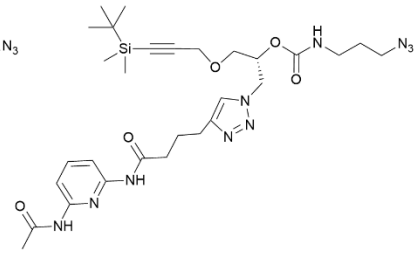
G



C

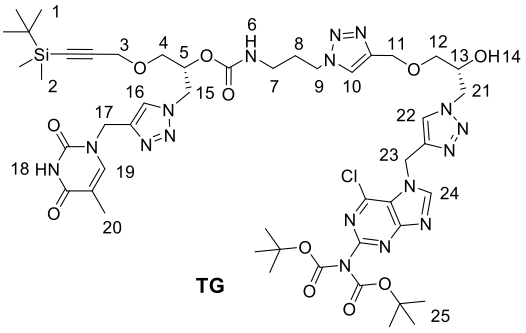


6



7

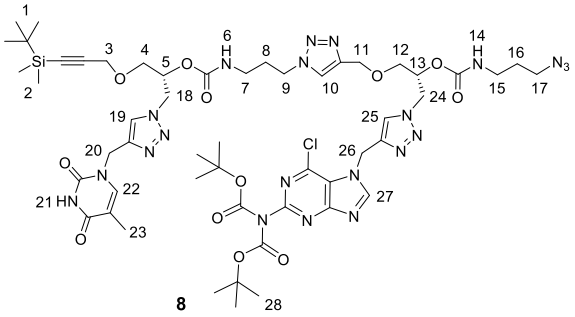
2.1. Synthesis of all-*R* tetramer TGCT



TG

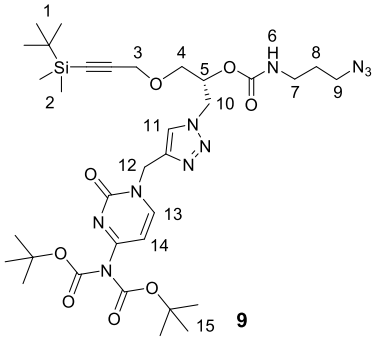
Compound **4** (0.88 g, 1.3 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (1.95 mL, 1.5 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (50 mL) and washed with brine (3 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound **6** (0.728 g, 1 equiv.)) was added into the residue followed by EtOH (9.1 mL) with stirring. Milli-Q water (3.9 mL), sodium ascorbate solution (52 mg in 2 mL water, 0.2 equiv.) and $CuSO_4$ solution (20.8 mg in 1 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The

residue was purified by column chromatography (MeOH /DCM = 2/98 to 6/94) to give the final product **TG** as foamy white solid (1.314 g, 90%). **TG**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.69 (s, 1H, H_{18}), 8.38 (s, 1H, H_{24}), 7.95 (s, 1H, H_{16}), 7.78 (s, 1H, H_{22}), 7.69 (s, 1H, H_{10}), 7.30 (d, J = 1.5 Hz, 1H, H_{19}), 5.62 (t, J = 6.2 Hz, 1H, H_6), 5.55 (s, 2H, H_{23}), 5.18 (dd, J = 7.2, 4.0 Hz, 1H, H_5), 4.94 – 4.82 (m, 2H, H_{17}), 4.69 – 4.59 (m, 4H, H_{11} and H_{15}), 4.51 (dd, J = 14.0, 3.7 Hz, 1H, H_{21}), 4.32– 4.41 (m, 3H, H_9 and H_{21}), 4.25 – 4.16 (m, 2H, H_3 and H_{13}), 3.70 – 3.57 (m, 3H, H_4 and H_{12}), 3.51 – 3.46 (m, 1H, H_{12}), 3.08 (q, J = 6.3 Hz, 2H, H_7), 2.08 – 1.99 (m, 2H, H_8), 1.85 (d, J = 1.2 Hz, 3H, H_{20}), 1.46 (s, 18H, H_{25}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 164.49, 155.41, 152.55, 151.96, 151.31, 151.15, 150.98, 146.32, 144.54, 142.26, 141.09, 140.56, 130.09, 125.26, 125.10, 123.53, 111.28, 101.17, 90.97, 84.03, 71.60, 71.19, 69.15, 68.22, 64.64, 59.64, 53.19, 50.90, 47.39, 43.35, 39.41, 37.70, 28.48, 28.05, 26.16, 16.57, -4.57. **HRMS** m/z= 1122.4815 (calcd. for $C_{48}H_{68}ClN_{17}O_{11}Si$ 1122.4815 $[M+H]^+$).



8

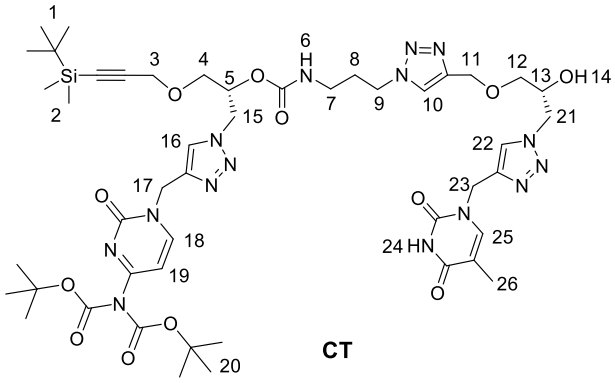
Compound **TC** (1.123 g, 1 mmol) was added in a flask followed by DCM (5 mL), 4-nitrophenyl chloroformate (0.503 g, 2.5 equiv.) and pyridine (202 µL, 2.5 equiv.). The mixture was stirred for 2 h at room temperature followed by solvent evaporation. The residue was dissolved in CH_3CN (5 mL) and added to a flask followed by 3-azido-1-propylamine (0.3 g, 3 eq.) and Et_3N (0.697 mL, 5 eq.). The mixture was stirred for 2 h at room temperature, then DCM (50 mL) and water (50 mL) were added to the mixture. The organic phase was washed with water (2 × 50 mL) and brine (2 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The final product **8** was obtained after passing the residue through a chromatography column (MeOH/DCM= 0/100 to 6/94) as a foamy-yellow solid (0.937 g, 75%). **8**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.55 (s, 1H, H_{21}), 8.40 (s, 1H, H_{27}), 7.86 (s, 1H, H_{19}), 7.78 (s, 1H, H_{25}), 7.72 (s, 1H, H_{10}), 7.35 – 7.28 (m, 1H, H_{22}), 5.68 – 5.42 (m, 4H, H_6, H_{14} and H_{23}), 5.19 (s, 1H, H_5), 5.07 (s, 1H, H_{13}), 4.98 – 4.82 (m, 2H, H_{20}), 4.72 – 4.51 (m, 6H, H_{11}, H_{18} and H_{24}), 4.40 (t, J = 6.7 Hz, 2H, H_9), 4.27 – 4.14 (m, 2H, H_3), 3.73 – 3.42 (m, 4H, H_4 and H_{12}), 3.31 – 3.22 (m, 2H, H_{17}), 3.18 – 2.98 (m, 4H, H_7 and H_{15}), 2.06 – 1.96 (m, 2H, H_8), 1.87 (d, J = 1.1 Hz, 3H, H_{23}), 1.66 – 1.59 (m, 2H, H_{16}), 1.49 (s, 18H, H_{28}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 164.43, 155.38, 155.20, 152.59, 151.87, 151.39, 151.26, 151.14, 146.43, 144.45, 142.30, 140.46, 130.22, 125.21, 125.08, 123.56, 111.28, 101.19, 90.93, 84.28, 71.13, 70.95, 68.43, 68.18, 64.84, 59.62, 50.88, 49.01, 47.39, 43.30, 39.54, 38.49, 37.78, 30.27, 28.94, 28.47, 28.04, 26.15, 16.56, 12.43, -4.58. **HRMS** m/z= 1248.5367 (calcd. for $C_{57}H_{83}ClN_{19}O_{13}Si$ 1248.5356 $[M+Na]^+$).



9

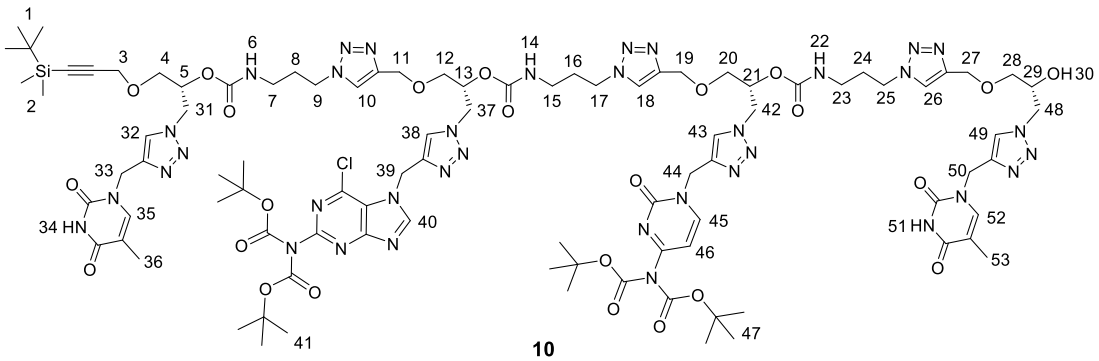
Compound **5** (1 g, 1.62 mmol) was added in a flask followed by DCM (5 mL), 4-nitrophenyl chloroformate (0.49 g, 1.5 equiv.) and pyridine (200 µL, 1.5 equiv.). The mixture was stirred for 1 h at room temperature followed by solvent evaporation. The residue was dissolved in CH_3CN (5 mL) and added to a flask followed by 3-azido-1-propylamine (0.324 g, 2 eq.) and Et_3N (0.68 mL, 3 eq.). The mixture was stirred for 1 h at room temperature, then DCM (50 mL) and Milli-Q water (50 mL) were added to the mixture. The organic phase was washed with Milli-Q water (2 × 50 mL) and brine (2 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The final product **9**

was obtained after passing the residue through a chromatography column (EtOAc /n-hexane = 5/5 to 10/0) as a foamy-yellow solid (1.207 g, 100%). **10**, **^{1}H NMR (300 MHz, $CDCl_3$)** δ 7.84 (t, J = 3.7 Hz, 2H, H_{11} and H_{13}), 7.05 (d, J = 7.4 Hz, 1H, H_{14}), 5.25 – 5.10 (m, 2H, H_5 and H_6), 5.07 (d, J = 5.2 Hz, 2H, H_{12}), 4.71 – 4.55 (m, 2H, H_{10}), 4.20 (d, J = 4.0 Hz, 2H, H_3), 3.60 (d, J = 4.2 Hz, 2H, H_4), 3.43 – 3.32 (m, 4H, H_7 and H_9), 1.83 – 1.73 (m, 2H, H_8), 1.54 (s, 18H, H_{15}), 0.92 (s, 9H, H_1), 0.11 (s, 6H, H_2). **^{13}C NMR (75 MHz, $CDCl_3$)** δ 162.73, 158.39, 155.11, 149.58, 148.01, 141.70, 125.76, 101.23, 96.89, 90.83, 85.15, 71.00, 68.00, 59.59, 50.66, 49.08, 38.65, 37.98, 29.08, 27.80, 26.15, 16.55, -4.59. **HRMS** m/z= 745.3810 (calcd. for $C_{33}H_{52}N_{10}O_8Si$ 745.3812 $[M+H]^+$).



CT

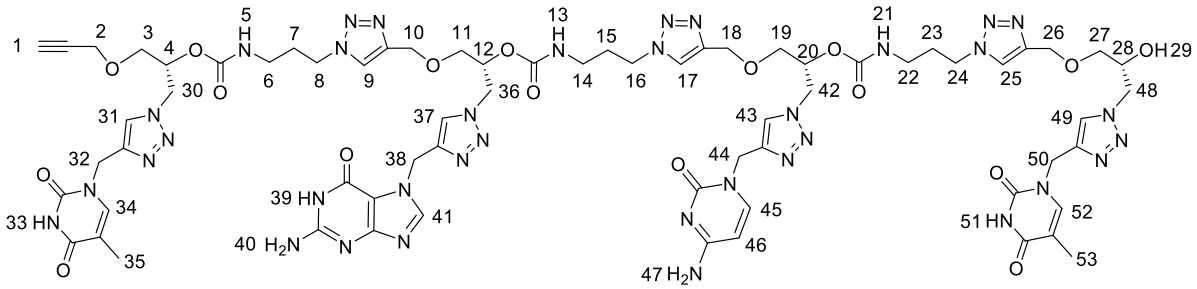
Compound **T** (0.88 g, 1.2 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (1.95 mL, 1.8 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (50 mL) and washed with brine (3 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound **9** (0.728 g, 1 equiv.)) was added into the residue followed by EtOH (9.1 mL) with stirring. Water (3.9 mL), sodium ascorbate solution (52 mg in 2 mL Milli-Q water, 0.2 equiv.) and $CuSO_4$ solution (20.8 mg in 1 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 2/98 to 6/94) to give the final product TG as foamy white solid (1.314 g, 96%).**CT**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.36 (s, 1H, H_{24}), 7.94 – 7.84 (m, 3H, H_{16}, H_{18} and H_{22}), 7.79 (s, 1H, H_{10}), 7.37 – 7.31 (m, 1H, H_{25}), 7.04 (d, J = 7.4 Hz, 1H, H_{19}), 5.59 (t, J = 6.1 Hz, 1H, H_6), 5.20 – 5.01 (m, 3H, H_5 and H_{23}), 4.93 (s, 2H, H_{17}), 4.70 – 4.49 (m, 5H, H_{11}, H_{15} and H_{21}), 4.47 – 4.34 (m, 3H, H_9 and H_{21}), 4.26 – 4.13 (m, 3H, H_3 and H_{13}), 3.71 – 3.60 (m, 2H, H_4), 3.58 – 3.45 (m, 2H, H_{12}), 3.21 – 2.92 (m, 2H, H_7), 2.05 (m, 2H, H_8), 1.88 (d, J = 1.2 Hz, 3H, H_{26}), 1.51 (s, 18H, H_{20}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 164.32, 162.74, 155.51, 155.27, 151.06, 149.60, 148.39, 144.50, 142.04, 141.68, 140.43, 126.10, 125.26, 123.72, 111.28, 101.25, 97.00, 90.88, 85.26, 71.49, 71.12, 69.25, 68.20, 64.78, 59.62, 53.11, 50.76, 47.45, 45.75, 43.15, 37.76, 30.33, 27.80, 26.16, 16.56, -4.57. **HRMS** m/z= 1064.5099 (calcd. for $C_{47}H_{69}N_{15}O_{12}Si$ 1064.5092 $[M+H]^+$).



10

Compound **CT** (0.341 g, 0.32 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.64 mL, 2 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (50 mL) and washed with brine (3 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification.

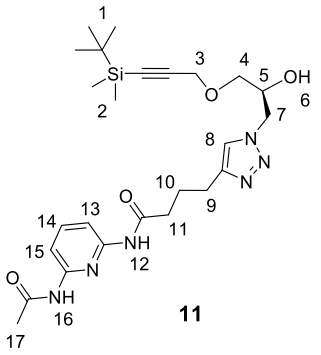
Compound **9** (0.400 g, 1 equiv.) was added into the residue followed by EtOH (2.24 mL) with stirring. Milli-Q water (0.96 mL), sodium ascorbate solution (12.8 mg in 0.5 mL water, 0.2 equiv.) and $CuSO_4$ solution (5.1 mg in 0.3 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 2/98 to 12/88) to give the final product **10** as foamy white solid (0.457 g, 65%). **10**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 10.27 – 9.98 (m, 2H, H_{34} and H_{51}), 8.43 (s, 1H, H_{40}), 7.98 – 7.68 (m, 8H, H_{10}, H_{18}, H_{26}, H_{32}, H_{38}, H_{43}, H_{45} and H_{49}), 7.36 – 7.28 (m, 2H, H_{35} and H_{52}), 7.02 (d, J = 7.3 Hz, 1H, H_{46}), 6.31 – 5.94 (b, 3H, H_6, H_{14}, H_{22}), 5.61 – 5.34 (s, 2H, H_{39}), 5.23 – 4.80 (m, 9H, H_5, H_{13}, H_{21}, H_{33}, H_{44} and H_{50}), 4.58 (m, J = 26.1, 17.4, 7.5 Hz, 13H, H_{11}, H_{19}, H_{27}, H_{31}, H_{37}, H_{42} and H_{48}), 4.44 – 4.27 (m, 7H, H_9, H_{17}, H_{25} and H_{48}), 4.25 – 4.09 (m, 3H, H_3 and H_{29}), 3.74 – 3.41 (m, 8H, H_4, H_{12}, H_{20} and H_{28}), 3.16– 2.90 (m, 5H, H_7, H_{15} and H_{23}), 2.09 – 1.89 (m, 6H, H_8, H_{16} and H_{24}), 1.86 – 1.77 (m, 6H, H_{36} and H_{53}), 1.56 – 1.40 (m, 36H, H_{41} and H_{47}), 0.91 (s, 9H, H_1), 0.09 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 164.68, 162.68, 155.61, 155.49, 155.22, 152.64, 151.27, 151.14, 149.60, 148.60, 146.72, 144.53, 144.27, 144.18, 143.31, 142.30, 142.03, 141.69, 141.38, 140.64, 140.53, 126.03, 125.27, 125.14, 123.85, 111.09, 101.26, 96.90, 90.84, 85.21, 84.17, 71.57, 71.14, 69.17, 68.74, 68.33, 64.73, 59.59, 53.25, 50.97, 50.52, 47.49, 45.60, 43.16, 39.46, 37.83, 30.23, 29.82, 28.47, 28.03, 27.78, 26.15, 16.55, 12.40, -4.57. **TOF MS ES +** m/z= 2197.0074 (calcd. for $C_{93}H_{129}ClN_{36}O_{24}Si$ 2197.9511 $[M+H]^+$).



TGCT

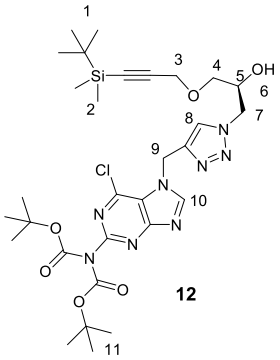
Compound **10** (0.41 g, 0.19 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.57 mL, 3 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (50 mL) and washed with brine (3 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The residue was obtained and followed by adding TFA (3 mL) and Milli-Q water (1 mL) stirring at room temperature for 48 h. The crude product was precipitated from the solution with the neutralization of saturated $NaHCO_3$. Then, the solid product was wash with brine, water, acetone and Et_2O, respectively. The final product was dried under vacuum to give the final product **TGCT** as foamy yellow solid (0.22 g, 71%). **TGCT**, **^{1}H NMR (500 MHz, DMSO)** δ 11.30 (s, 2H, H_{33} and H_{51}), 11.06 (s, 1H, H_{39}), 8.19 – 7.85 (m, 7H, H_9, H_{17}, H_{25}, H_{31}, H_{37}, H_{41} and H_{49}), 7.78 – 7.48 (m, 3H, H_{34}, H_{43} and H_{52}), 7.46 – 7.32 (m, 2H, H_{45} and H_{47}), 7.11 (d, J = 124.8 Hz, 1H, H_{46}), 6.68 (s, 2H, H_{40}), 5.31 (d, J = 5.6 Hz, 1H, H_{21}), 5.21 (s, 2H, H_{38}), 5.09 (s, 3H, H_4, H_{12} and H_{20}), 4.88 (d, J = 4.2 Hz, 6H, H_{32}, H_{44} and H_{50}), 4.64 – 4.46 (m, 12H, H_{10}, H_{18}, H_{26}, H_{30}, H_{36} and H_{42}), 4.42 (dd, J = 13.9, 3.5 Hz, 1H, H_{48}), 4.36 – 4.20 (m, 7H, H_8, H_{16}, H_{24} and H_{48}), 4.17 (d, J = 2.4 Hz, 2H, H_2), 3.96 (s, 1H, H_{28}), 3.62 – 3.39 (m, 8H, H_3, H_{11}, H_{19} and H_{27}), 2.98 – 2.79 (m, 6H, H_6, H_{14} and H_{22}), 2.46 (t, J = 4.0, 2.0 Hz, 1H, H_1), 1.95 – 1.80 (m, 6H, H_7, H_{15} and H_{23}), 1.78 – 1.68 (m, 6H, H_{35} and H_{53}). **^{13}C NMR (126 MHz, DMSO)** δ 166.10, 166.01, 164.33, 155.60, 155.20, 151.23, 150.77, 145.90, 145.74, 143.67, 143.40, 142.96, 142.41, 142.13, 141.15, 141.01, 124.53, 124.45, 124.28, 124.13, 108.84, 93.62, 79.82, 77.68, 71.62, 70.43, 68.67, 68.21, 63.87, 57.92, 54.93, 52.90, 50.06, 48.61, 46.96, 42.15, 37.41, 29.95, 26.24, 11.98 **HRMS** m/z= 1665,6914 (calcd. for $C_{67}H_{84}N_{36}O_{17}$ 1665,6888 $[M+H]^+$).

2.2. Synthesis of all-*R* and all-*S* tetramer DCGD



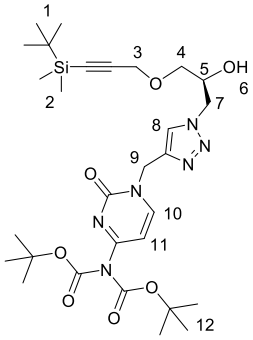
11

Compound **(*S*)-3** (0.808 g, 3 mmol) and **base D** (0.735 g, 1 equiv.) were added into a flask followed by EtOH (21 mL) with stirring. Water (9 mL), sodium ascorbate solution (120 mg in 4 mL water, 0.2 equiv.) and $CuSO_4$ solution (48 mg in 3 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, EtOAc (100 mL) and Na_2EDTA (0.05 M, 100 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with EtOAc (100 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was crystallized from cold EtOAc, the final product **11** was obtained as a yellow solid (1.544 g, 100%) after removing the solvent. **11**, **^{1}H NMR (300 MHz, $CDCl_3$)** δ 8.55 (s, 1H, H_{12}), 8.34 (s, 1H, H_{16}), 7.78 – 7.55 (m, 2H, H_{13} and H_{15}), 7.67 (t, J = 8.0 Hz, 1H, H_{14}), 7.49 (s, 1H, H_8), 4.59 – 4.34 (m, 2H, H_7), 4.30 – 4.18 (m, 3H, H_3 and H_5), 3.67 – 3.52 (m, 2H, H_4), 2.80 (t, J = 6.9 Hz, 2H, H_9), 2.43 (t, J = 7.1 Hz, 2H, H_{11}), 2.21 (s, 3H, H_{17}), 2.08 (t, J = 7.0 Hz, 2H, H_{10}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (75 MHz, $CDCl_3$)** δ 171.95, 169.20, 149.57, 149.41, 146.90, 141.11, 123.09, 109.50, 101.31, 90.97, 70.94, 69.32, 59.69, 53.21, 36.38, 26.14, 25.18, 24.83, 24.22, 16.56, -4.59. **HRMS** *m/z*= 537.2637 (calcd. for $C_{25}H_{38}N_6O_4Si$ 537.2640 $[M+Na]^+$).



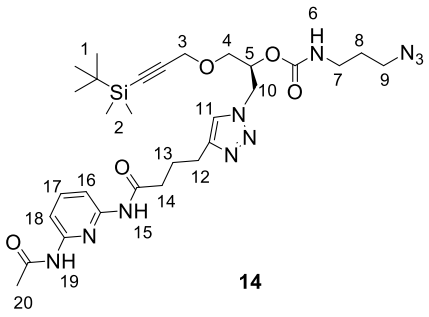
12

Compound **(*S*)-3** (0.404 g, 1.5 mmol) and **base G** (0.612 g, 1 equiv.) were added into a flask followed by EtOH (10.5 mL) with stirring. Water (4.5 mL), sodium ascorbate solution (60 mg in 2 mL water, 0.2 equiv.) and $CuSO_4$ solution (24 mg in 1.5 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, EtOAc (60 mL) and Na_2EDTA (0.05 M, 60 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (60 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (EtOAc /n-hexane = 5/5 to 0/10) to give the final product **12** as foamy white solid (0.944 g, 93%). **12, ^{1}H NMR (300 MHz, $CDCl_3$)** δ 8.32 (s, 1H, H_{10}), 7.92 (s, 1H, H_8), 5.53 (s, 2H, H_9), 4.59 – 4.29 (m, 2H, H_7), 4.23 – 4.14 (m, 3H, H_3 and H_5), 3.61 (dd, J = 9.7, 4.5 Hz, 1H, H_4), 3.49 (dd, J = 9.7, 5.6 Hz, 1H, H_4), 1.48 (s, 18H, H_{11}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (75 MHz, $CDCl_3$)** δ 152.47, 152.01, 151.49, 151.01, 145.94, 141.18, 130.15, 125.11, 101.27, 90.99, 84.03, 70.70, 69.19, 59.64, 53.36, 39.54, 28.07, 26.15, 16.56, -4.58. **HRMS** *m/z*= 677.2990 (calcd. for $C_{30}H_{45}ClN_8O_6Si$ 677.2993 $[M+H]^+$).



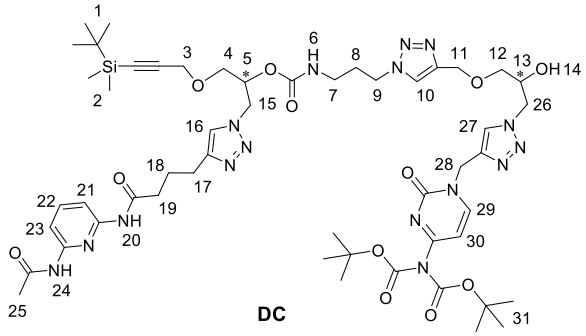
13

Compound **(*S*)-3** (0.539 g, 2 mmol) and **base C** (0.699 g, 1 equiv.) were added into a flask followed by EtOH (14 mL) with stirring. Water (6 mL), sodium ascorbate solution (80 mg in 6 mL water, 0.2 equiv.) and $CuSO_4$ solution (32 mg in 4 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, EtOAc (60 mL) and Na_2EDTA (0.05 M, 60 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with EtOAc (60 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (EtOAc /n-hexane = 5/5 to 0/10) to give the final product **13** as foamy light-yellow solid (0.998 g, 81%). **13**, **^{1}H NMR (300 MHz, $CDCl_3$)** δ 7.93 (s, 1H, H_8), 7.85 (d, J = 7.4 Hz, 1H, H_{10}), 7.03 (d, J = 7.3 Hz, 1H, H_{11}), 5.09 (s, 2H, H_9), 4.30 – 4.57 (m, 2H, H_7), 4.27 – 4.13 (m, 3H, H_3 and H_5), 3.66 – 3.43 (m, 2H, H_4), 1.54 (s, 18H, H_{12}), 0.93 (s, 9H, H_1), 0.11 (s, 6H, H_2). **^{13}C NMR (75 MHz, $CDCl_3$)** δ 162.74, 155.05, 149.64, 147.96, 141.75, 125.84, 101.42, 96.83, 90.83, 85.11, 70.66, 69.24, 59.63, 53.28, 45.51, 27.84, 26.17, 16.57, -4.57. **HRMS** *m/z*= 619.3270 (calcd. for $C_{29}H_{46}N_6O_7Si$ 619.3270 $[M+H]^+$).



14

Compound **11** (0.566 g, 1.1 mmol) was added in a flask followed by DCM (5 mL), 4-nitrophenyl chloroformate (0.333 g, 1.5 equiv.) and pyridine (138 µL, 1.5 equiv.). The mixture was stirred for 1 h at room temperature followed by solvent evaporation. The residue was dissolved in CH_3CN (5 mL) and added to a flask followed by 3-azido-1-propylamine (0.22 g, 2 eq.) and Et_3N (0.46 mL, 3 eq.). The mixture was stirred for 1 h at room temperature, then DCM (50 mL) and Milli-Q water (50 mL) were added to the mixture. The organic phase was washed with water (2 × 50 mL) and brine (2 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The final product **14** was obtained after passing the residue through a chromatography column (MeOH/DCM= 0/100 to 5/95) as a foamy-white solid (0.705 g, 100%). **14**, **^{1}H NMR (300 MHz, $CDCl_3$)** δ 8.60 (s, 1H, H_{15}), 8.17 (s, 1H, H_{19}), 7.91 – 7.77 (m, 2H, H_{16} and H_{18}), 7.68 (t, J = 8.0 Hz, 1H, H_{17}), 7.39 (s, 1H, H_{11}), 5.37 – 5.12 (m, 2H, H_5 and H_6), 4.59 (d, J = 5.6 Hz, 2H, H_{10}), 4.22 (s, 2H, H_3), 3.65 (d, J = 4.9 Hz, 2H, H_4), 3.33 (t, J = 6.5 Hz, 2H, H_9), 3.21 (q, J = 6.5 Hz, 2H, H_7), 2.90 – 2.71 (m, 2H, H_{12}), 2.39 (t, J = 8.0 Hz, 2H, H_{14}), 2.19 (s, 3H, H_{20}), 2.09 (q, J = 7.0 Hz, 2H, H_{13}), 1.73 (p, J = 6.5 Hz, 2H, H_8), 0.92 (s, 9H, H_1), 0.11 (s, 6H, H_2). **^{13}C NMR (75 MHz, $CDCl_3$)** δ 171.96, 169.08, 161.28, 155.34, 149.62, 147.20, 140.93, 122.46, 109.47, 101.07, 91.08, 71.38, 68.35, 59.72, 50.75, 49.09, 38.72, 36.06, 28.98, 26.13, 25.19, 24.79, 24.09, 16.56, -4.59. **HRMS** *m/z*= 641.3338 (calcd. for $C_{29}H_{44}N_{10}O_5Si$ 641.3316 $[M+H]^+$).

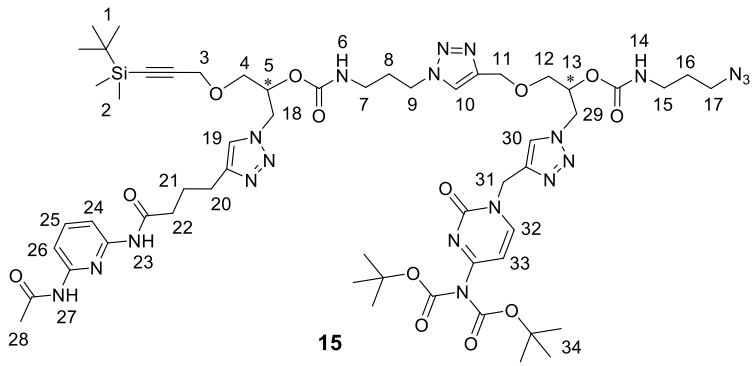


DC

Compound **5** (1.134 g, 1.8 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (2.7 mL, 1.5 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (60 mL) and washed with brine (3 × 60 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound **7** (1.152 g, 1 equiv.) was added into the residue followed by EtOH (12.6 mL) with stirring. Milli-Q water (5.4 mL), sodium ascorbate solution (72 mg in 3 mL water, 0.2 equiv.) and $CuSO_4$ solution (28.8 mg in 2 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (60 mL) and Na_2EDTA (0.05 M, 60 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (60 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 2/98 to 6/94) to give the final product (*R*, *R*)-**DC** as foamy white solid (1.443 g, 70%). The products (*S*, *S*)-**DC** (0.659 g, 72%) was prepared following the same protocol.

(*R*, *R*)-**DC**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 8.81 (s, 1H, H_{20}), 8.55 (s, 1H, H_{24}), 8.00 – 7.58 (m, 6H, H_{10}, H_{21}, H_{22}, H_{23}, H_{27} and H_{29}), 7.46 (d, *J* = 7.0 Hz, 1H, H_{16}), 7.04 (d, *J* = 7.3 Hz, 1H, H_{30}), 5.64 (b, 1H, H_6), 5.26 (d, *J* = 5.9 Hz, 1H, H_5), 5.10 (s, 2H, H_{28}), 4.70 – 4.31 (m, 8H, H_9, H_{11}, H_{15}, H_{26}), 4.29 – 4.08 (m, 3H, H_3 and H_{13}), 3.77 – 3.60 (m, 2H, H_4), 3.57 – 3.42 (m, 2H, H_{12}), 3.22 – 2.94 (m, 2H, H_7), 2.80 – 2.64 (m, 2H, H_{17}), 2.39 (s, 2H, H_{19}), 2.20 (s, 3H, H_{25}), 2.09 – 1.93 (m, 4H, H_8 and H_{18}), 1.51 (s, 18H, H_{31}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 172.08, 169.25, 162.70, 155.59, 155.19, 149.67, 148.28, 147.31, 144.91, 144.67, 144.18, 141.69, 140.75, 125.73, 123.69, 122.63, 109.53, 101.19, 96.85, 90.98, 85.28, 71.43, 71.24, 69.17, 68.50, 64.81, 59.70, 53.10, 50.90, 47.45, 45.68, 37.83, 36.17, 30.27, 27.80, 26.15, 25.15, 24.72, 24.41, 16.57, -4.56. **HRMS** *m/z*= 1145.5678 (calcd. for $C_{52}H_{76}N_{16}O_{12}Si$ 1145.5671 $[M+H]^+$).

(*S*, *S*)-**DC**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.07 – 8.43 (m, 2H, H_{20} and H_{24}), 7.94 – 7.59 (m, 6H, H_{10}, H_{21}, H_{22}, H_{23}, H_{27} and H_{29}), 7.52 – 7.40 (m, 1H, H_{16}), 7.05 (d, *J* = 7.4 Hz, 1H, H_{30}), 5.77 (b, 1H, H_6), 5.37 – 5.21 (m, 1H, H_5), 5.10 (s, 2H, H_{28}), 4.67 – 4.29 (m, 8H, H_9, H_{11}, H_{15}, H_{26}), 4.28 – 4.10 (m, 3H, H_3 and H_{13}), 3.77 – 3.58 (m, 2H, H_4), 3.57 – 3.40 (m, 2H, H_{12}), 3.25 – 2.93 (m, 2H, H_7), 2.85 – 2.66 (m, 2H, H_{17}), 2.40 (s, 2H, H_{19}), 2.20 (s, 3H, H_{25}), 2.10 – 1.89 (m, 4H, H_8 and H_{18}), 1.51 (s, 18H, H_{31}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 171.99, 169.21, 162.71, 155.61, 155.25, 149.66, 148.33, 147.29, 144.82, 144.64, 144.29, 141.64, 140.79, 125.72, 123.69, 122.67, 109.49, 101.21, 96.90, 90.95, 85.29, 71.47, 71.21, 69.15, 68.48, 64.78, 59.68, 53.12, 50.90, 47.47, 45.68, 37.83, 36.17, 30.26, 27.79, 26.15, 25.10, 24.69, 24.43, 16.57, -4.57. **HRMS** *m/z*= 1145.5672 (calcd. for $C_{52}H_{76}N_{16}O_{12}Si$ 1145.5671 $[M+H]^+$).



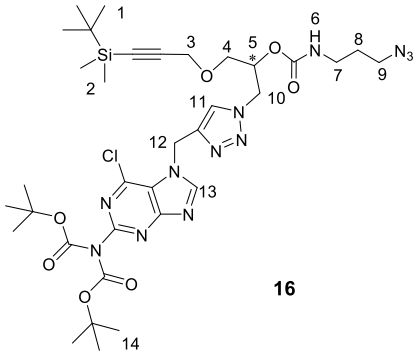
15

Compound (*R*, *R*)-**DC** (0.573 g, 0.5 mmol) was added in a flask followed by DCM (5 mL), 4-nitrophenyl chloroformate (0.252 g, 2.5 equiv.) and pyridine (101 µL, 2.5 equiv.). The mixture was stirred for 2 h at room temperature followed

by solvent evaporation. The residue was dissolved in CH_3CN (5 mL) and added to a flask followed by 3-azido-1-propylamine (0.15 g, 3 eq.) and Et_3N (0.348 mL, 5 eq.). The mixture was stirred for 2 h at room temperature, then DCM (50 mL) and Milli-Q water (50 mL) were added to the mixture. The organic phase was washed with water (2 × 50 mL) and brine (2 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The final product (*R*, *R*)-**15** was obtained after passing the residue through a chromatography column (MeOH/DCM= 3/97 to 8/92) as a foamy-white solid (0.636 g, 100%). The products (*S*, *S*)-**15** (0.425 g, 88%) was prepared following the same protocol.

(*R*, *R*)-**15**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 8.92 – 8.38 (m, 2H, H_{23} and H_{27}), 7.97 – 7.64 (m, 6H, H_{10}, H_{24}, H_{25}, H_{26}, H_{30} and H_{24}), 7.44 (s, 1H, H_{19}), 7.07 (d, *J* = 7.3 Hz, 1H, H_{33}), 5.92 – 5.54 (b, 2H, H_6 and H_{14}), 5.26 (d, *J* = 6.2 Hz, 1H, H_5), 5.17 – 5.05 (m, 3H, H_{13} and H_{31}), 4.68 – 4.51 (m, 6H, H_{11}, H_{18} and H_{29}), 4.39 (t, *J* = 6.6 Hz, 2H, H_9), 4.22 (d, *J* = 2.3 Hz, 2H, H_3), 3.80 – 3.58 (m, 2H, H_4), 3.54 – 3.39 (m, 2H, H_{12}), 3.33 (t, *J* = 6.6 Hz, 2H, H_{17}), 3.25 – 3.10 (m, 3H, H_7 and H_{15}), 3.08 – 2.96 (m, 1H, H_{15}), 2.88 – 2.67 (m, 2H, H_{20}), 2.38 (s, 2H, H_{22}), 2.20 (s, 3H, H_{28}), 2.13 – 1.96 (m, 6H, H_8 and H_{21}), 1.78 – 1.70 (m, 2H, H_{16}), 1.52 (s, 18H, H_{34}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 172.27, 169.45, 162.75, 155.56, 155.33, 155.25, 149.92, 149.66, 148.34, 148.17, 147.28, 144.44, 141.73, 140.72, 125.92, 123.74, 122.61, 109.51, 101.17, 96.93, 91.00, 85.28, 71.23, 70.64, 68.47, 68.13, 64.88, 59.69, 50.86, 49.05, 47.42, 45.77, 38.57, 37.84, 36.08, 30.20, 29.05, 27.79, 26.14, 25.23, 24.72, 24.31, 16.56, -4.58. **HRMS** *m/z*= 1271.6209 (calcd. for $C_{56}H_{82}N_{20}O_{13}Si$ 1271.6212 $[M+H]^+$).

(*S*, *S*)-**15**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.13 – 8.20 (m, 2H, H_{23} and H_{27}), 7.97 – 7.59 (m, 6H, H_{10}, H_{24}, H_{25}, H_{26}, H_{30} and H_{24}), 7.46 (d, *J* = 6.8 Hz, 1H, H_{19}), 7.07 (t, *J* = 8.3 Hz, 1H, H_{33}), 5.93 – 5.55 (b, 2H, H_6 and H_{14}), 5.26 (d, *J* = 5.6 Hz, 1H, H_5), 5.18 – 4.99 (m, 3H, H_{13} and H_{31}), 4.68 – 4.50 (m, 6H, H_{11}, H_{18} and H_{29}), 4.40 (t, *J* = 6.6 Hz, 2H, H_9), 4.30 – 4.16 (m, 2H, H_3), 3.80 – 3.60 (m, 2H, H_4), 3.56 – 3.39 (m, 2H, H_{12}), 3.33 (t, *J* = 6.6 Hz, 2H, H_{17}), 3.25 – 3.11 (m, 3H, H_7 and H_{15}), 3.02 (dd, *J* = 14.0, 6.5 Hz, 1H, H_{15}), 2.85– 2.67 (m, 2H, H_{20}), 2.40 (s, 2H, H_{22}), 2.21 (s, 3H, H_{28}), 2.14 – 1.94 (m,, 4H, H_8 and H_{21}), 1.74 (p, *J* = 6.6 Hz, 2H, H_{16}), 1.52 (s, 18H, H_{34}), 0.92 (s, 9H, H_1), 0.11 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 172.29, 169.78, 162.54, 155.67, 155.63, 155.19, 150.28, 150.11, 149.64, 148.63, 146.48, 144.57, 141.82, 140.20, 126.32, 124.18, 123.55, 109.30, 101.32, 96.85, 90.62, 85.08, 71.42, 70.84, 69.04, 68.08, 64.76, 59.54, 53.00, 50.76, 47.43, 45.01, 37.65, 36.11, 35.97, 30.17, 28.08, 27.74, 26.11, 25.09, 24.67, 24.38, 16.50, -4.62. **HRMS** *m/z*= 1271.6187 (calcd. for $C_{56}H_{82}N_{20}O_{13}Si$ 1271.6212 $[M+H]^+$).



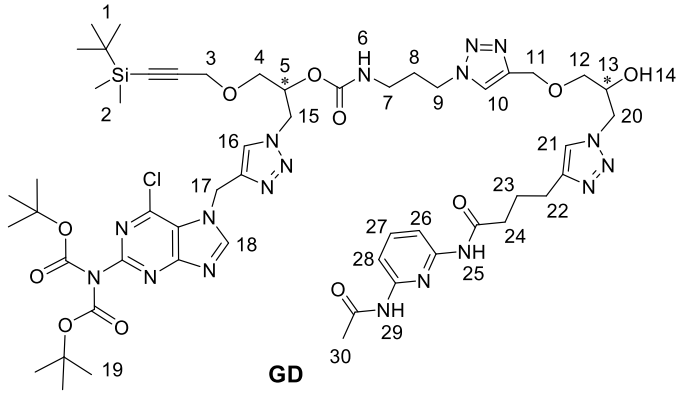
16

Compound **4** (1.3 g, 1.9 mmol) was added in a flask followed by DCM (10 mL), 4-nitrophenyl chloroformate (0.585 g, 1.5 equiv.) and pyridine (235 µL, 1.5 equiv.). The mixture was stirred for 1 h at room temperature followed by solvent evaporation. The residue was dissolved in CH_3CN (10 mL) and added to a flask followed by 3-azido-1-propylamine (0.38 g, 2 eq.) and Et_3N (0.794 mL, 3 eq.). The mixture was stirred for 1 h at room temperature, then DCM (60 mL) and Milli-Q water (60 mL) were added to the mixture. The organic phase was washed with water (2 × 60 mL) and brine (2 × 60 mL), dried with Na_2SO_4 and concentrated under vacuum. The final product (*R*)-**16** was obtained after passing the residue through a chromatography column (MeOH/DCM= 3/97 to 8/92) as a foamy-white solid (1.53 g, 100%). The products (*S*)-**17** (0.521 g, 88%) was prepared following the same protocol.

(*R*)-**16**, **^{1}H NMR (300 MHz, $CDCl_3$)** δ 8.30 (s, 1H, H_{13}), 7.81 (s, 1H, H_{11}), 5.53 (s, 2H, H_{12}), 5.28 (t, *J* = 6.0 Hz, 1H, H_6), 5.09 (dd, *J* = 8.4, 4.2 Hz, 1H, H_5), 4.69 – 4.41 (m, 2H, H_{10}), 4.18 (d, *J* = 3.6 Hz, 2H, H_3), 3.69 – 3.54 (m, 2H, H_4), 3.23 (t, *J* = 6.5 Hz, 2H, H_9), 3.13 – 2.89 (m, 2H, H_7), 1.62 – 1.54 (m, 2H, H_8), 1.51 (s, 18H, H_{14}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (75 MHz, $CDCl_3$)** δ 155.09, 152.57, 151.93, 151.56, 151.31, 146.06, 141.39, 130.31, 125.02, 101.20, 90.84,

84.35, 71.18, 68.22, 59.63, 51.10, 49.04, 39.79, 38.49, 28.95, 28.07, 26.15, 16.55, -4.57. **HRMS** *m/z*= 803.3535 (calcd. for $C_{34}H_{51}ClN_{12}O_7Si$ 803.3534 $[M+H]^+$)

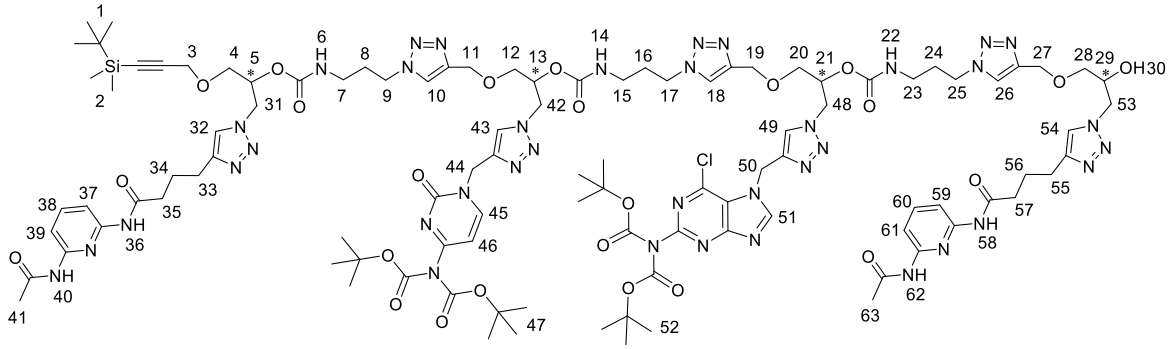
(*S*)-**16**, **^{1}H NMR (300 MHz, CDCl$_3$)** δ 8 8.30 (s, 1H, H_{13}), 7.81 (s, 1H, H_{11}), 5.53 (s, 2H, H_{12}), 5.27 (t, *J* = 6.0 Hz, 1H, H_6), 5.09 (dd, *J* = 8.1, 4.0 Hz, 1H, H_5), 4.72 – 4.45 (m, 2H, H_{10}), 4.18 (d, *J* = 3.6 Hz, 2H, H_3), 3.63 (dd, *J* = 9.0, 4.6 Hz, 2H, H_4), 3.30 – 3.19 (m, 2H, H_9), 3.12 – 2.90 (m, 2H, H_7), 1.83 – 1.68 (m, 2H, H_8), 1.51 (s, 18H, H_{14}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (75 MHz, CDCl$_3$)** δ 155.09, 152.57, 151.92, 151.56, 151.31, 146.06, 141.38, 130.32, 125.02, 101.19, 90.84, 84.35, 71.18, 68.22, 59.63, 51.10, 49.04, 39.79, 38.49, 28.95, 28.07, 26.15, 16.56, -4.57. **HRMS** *m/z*= 803.3534 (calcd. for $C_{34}H_{51}ClN_{12}O_7Si$ 803.3534 $[M+H]^+$)



Compound (*R*)-**D** (0.309 g, 0.6 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.9 mL, 1.5 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (30 mL) and washed with brine (3 × 30 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound (*R*)-**16** (0.48 g, 1 equiv.) was added into the residue followed by EtOH (4.2 mL) with stirring. Milli-Q water (1.8 mL), sodium ascorbate solution (24 mg in 1 mL water, 0.2 equiv.) and $CuSO_4$ solution (9.6 mg in 0.5 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 2/98 to 6/94) to give the final product (*R*, *R*)-**GD** as foamy yellow solid (0.65 g, 90%). The products (*S*, *S*)-**GD** (0.613 g, 79%) was prepared following the same protocol.

(*R*, *R*)-**GD**, **^{1}H NMR (500 MHz, CDCl$_3$)** δ .95 – 8.42 (m, 2H, H_{25} and H_{29}), 8.38 (s, 1H, H_{18}), 7.90 – 7.74 (m, 3H, H_{10}, H_{16} and H_{26}), 7.70 – 7.62 (m, 2H, H_{27} and H_{28}), 7.50 (s, 1H, H_{21}), 5.76 (t, *J* = 6.0 Hz, 1H, H_6), 5.52 (d, *J* = 4.5 Hz, 2H, H_{17}), 5.15 (dd, *J* = 8.3, 4.1 Hz, 1H, H_5), 4.70 – 4.09 (m, 11H, H_3, H_9, H_{11}, H_{13}, H_{15} and H_{20}), 3.72 – 3.41 (m, 4H, H_4 and H_{12}), 2.97 – 2.86 (m, 2H, H_7), 2.74 (t, *J* = 7.0 Hz, 2H, H_{22}), 2.40 (t, *J* = 7.2 Hz, 2H, H_{24}), 2.20 (s, 3H, H_{30}), 2.04 (p, *J* = 7.0 Hz, 2H, H_{23}), 1.97 – 1.82 (m, 2H, H_8), 1.49 (s, 18H, H_{19}), 0.91 (s, 9H, H_1), 0.09 (s, 6H, H_2). **^{13}C NMR (126 MHz, CDCl$_3$)** δ 172.23, 169.35, 155.45, 152.59, 151.90, 151.34, 151.31, 149.77, 149.58, 146.82, 146.55, 144.53, 141.31, 140.90, 130.16, 125.19, 123.44, 123.18, 109.51, 109.48, 101.12, 90.95, 84.45, 71.37, 71.17, 69.27, 68.30, 64.70, 59.63, 52.96, 51.18, 47.45, 39.69, 37.76, 36.30, 30.17, 28.06, 26.14, 25.11, 24.77, 24.28, 16.55, -4.57. **HRMS** *m/z*= 1225.5204 (calcd. for $C_{53}H_{75}ClN_{18}O_{11}Si$ 1225.5213 $[M+Na]^+$)

(*S, S*)-**GD**, **^{1}H NMR (500 MHz, CDCl$_3$)** δ δ 9.13 – 8.47 (m, 2H, H_{25} and H_{29}), 8.39 (s, 1H, H_{18}), 7.94 – 7.73 (m, 3H, H_{26}, H_{27} and H_{14}), 7.69 – 7.62 (m, 2H, H_{10} and H_{16}), 7.51 (s, 1H, H_{21}), 5.85 – 5.74 (t, *J* = 6.0 Hz, 1H, H_6), 5.53 (d, *J* = 3.4 Hz, 2H, H_{17}), 5.15 (dd, *J* = 7.9, 3.9 Hz, 1H, H_5), 4.74 – 4.06 (m, 11H, H_3, H_9, H_{11}, H_{13}, H_{15} and H_{20}), 3.71 – 3.43 (m, 4H, H_4 and H_{12}), 2.99 – 2.86 (m, 2H, H_7), 2.74 (t, *J* = 6.9 Hz, 2H, H_{22}),), 2.40 (t, *J* = 7.2 Hz, 2H, H_{24}), 2.20 (s, 3H, H_{30}), 2.04 (q, *J* = 7.2 Hz, 2H, H_{23}), 1.89 (p, *J* = 27.5, 6.8 Hz, 2H, H_8), 1.49 (s, 18H, H_{19}), 0.91 (s, 9H, H_1), 0.09 (s, 6H, H_2). **^{13}C NMR (126 MHz, CDCl$_3$)** δ 172.14, 169.55, 155.45, 152.58, 151.87, 151.33, 151.29, 149.69, 149.48, 146.79, 146.59, 144.50, 141.30, 141.01, 130.14, 125.21, 123.47, 123.22, 109.48, 109.44, 101.11, 90.93, 84.46, 71.40, 71.15, 69.24, 68.28, 64.67, 59.63, 52.98, 51.17, 47.45, 39.67, 37.74, 36.26, 30.16, 28.05, 26.13, 25.07, 24.75, 24.27, 16.54, -4.58. **HRMS** *m/z*= 1203.5392 (calcd. for $C_{53}H_{75}ClN_{18}O_{11}Si$ 1203.5393 $[M+H]^+$)

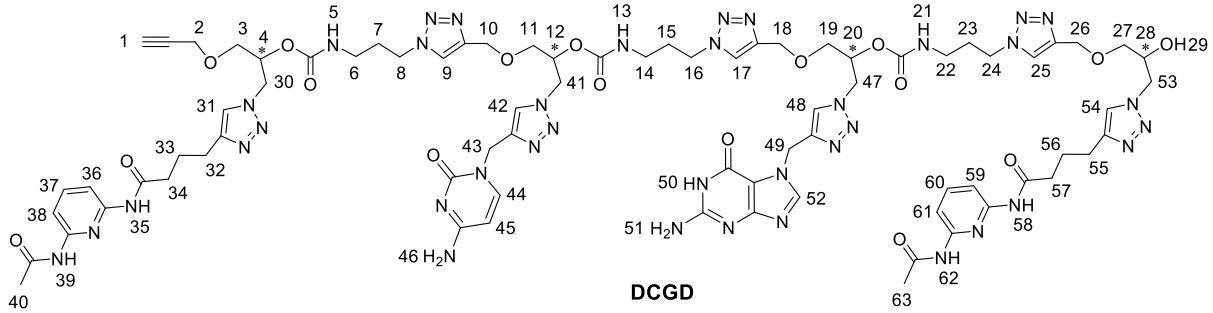


17

Compound (*R*, *R*)-**GD** (0.309 g, 0.6 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.9 mL, 1.5 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (30 mL) and washed with brine (3 × 30 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound (*R*, *R*)-**16** (0.48 g, 1 equiv.) was added into the residue followed by EtOH (4.2 mL) with stirring. Milli-Q water (1.8 mL), sodium ascorbate solution (24 mg in 1 mL water, 0.2 equiv.) and $CuSO_4$ solution (9.6 mg in 0.5 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 2/98 to 8/92) to give the final product all-*R* **17** as foamy yellow solid (0.766 g, 87%). The products all-*S* 19 (0.623 g, 88%) was prepared following the same protocol.

(*R, R, R, R*)-**17**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.48 – 8.48 (m, 4H, H_{36}, H_{40}, H_{58} and H_{62}), 8.41 (s, 1H, H_{51}), 8.01 – 7.57 (m, 12H, H_{10}, H_{18}, H_{26}, H_{37}, H_{38}, H_{39}, H_{43}, H_{45}, H_{49}, H_{59}, H_{60} and H_{61}), 7.51 (s, 1H, H_{54}), 7.44 (s, 1H, H_{32}), 7.08 – 6.99 (m, 1H, H_{46}), 6.38 – 5.91 (b, 3H, H_6, H_{14} and H_{22}), 5.61 – 4.91 (m, 7H, H_5, H_{13}, H_{21}, H_{44} and H_{50}), 4.87 – 4.02 (m, 25H, H_3, H_9, H_{11}, H_{17}, H_{19}, H_{25}, H_{27}, H_{29}, H_{31}, H_{42}, H_{48} and H_{53}), 3.77 – 3.41 (m, 8H, H_4, H_{12}, H_{20} and H_{28}), 3.30 – 2.86 (m, 6H, H_7, H_{15} and H_{23}), 2.76 – 2.64 (m, 4H, H_{33} and H_{55}), 2.43 – 1.87 (m, 20H, H_8, H_{16}, H_{24}, H_{34}, H_{36}, H_{41}, H_{56}, H_{58} and H_{63}), 1.59 – 1.38 (m, 36H, H_{47} and H_{52}), 0.91 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 172.31, 169.26, 161.91, 155.85, 155.76, 155.73, 155.11, 152.56, 152.00, 151.12, 150.94, 149.69, 148.80, 147.28, 147.09, 146.71, 144.54, 144.17, 144.12, 141.90, 141.37, 125.68, 125.19, 123.53, 123.44, 123.32, 122.71, 109.54, 109.49, 101.30, 96.84, 90.86, 84.16, 83.25, 77.00, 71.12, 71.06, 71.01, 70.91, 70.88, 69.11, 68.69, 68.49, 64.85, 64.78, 64.64, 59.65, 53.07, 50.94, 50.89, 50.66, 50.64, 47.59, 45.94, 37.83, 37.79, 36.15, 36.08, 30.20, 30.15, 29.82, 28.13, 28.02, 26.15, 25.17, 25.01, 24.64, 24.51, 16.55, -4.56. **TOF MS ES +** *m/z*= 2360.1295 (calcd. for $C_{103}H_{143}ClN_{38}O_{24}Si$ 2360.0668 $[M+H]^+$).

(*S, S, S, S*)-**17**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.42 – 8.42 (m, 4H, H_{36}, H_{40}, H_{58} and H_{62}), 8.41 (s, 1H, H_{51}), 7.98 – 7.39 (m, 14H, H_{10}, H_{18}, H_{26}, H_{32}, H_{37}, H_{38}, H_{39}, H_{43}, H_{45}, H_{49}, H_{54}, H_{59}, H_{60} and H_{61}), 7.04 (d, *J* = 7.4 Hz, 1H, H_{46}), 6.41– 5.90 (b, 3H, H_6, H_{14} and H_{22}), 5.52 (s, 2H, H_{50}), 5.30 – 5.21 (m, 1H, H_5), 5.16 – 4.93 (m, 4H, H_{13}, H_{21} and H_{44}), 4.77 – 4.11 (m, 25H, H_3, H_9, H_{11}, H_{17}, H_{19}, H_{25}, H_{27}, H_{29}, H_{31}, H_{42}, H_{48} and H_{53}), 3.80 – 3.36 (m, 8H, H_4, H_{12}, H_{20} and H_{28}), 3.22 – 2.87 (m, 6H, H_7, H_{15} and H_{23}), 2.77 – 2.63 (m, 4H, H_{33} and H_{55}), 2.48 – 1.86 (m, 20H, H_8, H_{16}, H_{24}, H_{34}, H_{36}, H_{41}, H_{56}, H_{58} and H_{63}), 1.52 – 1.42 (m, 36H, H_{47} and H_{52}), 0.91 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 172.25, 169.71, 162.73, 155.71, 155.64, 155.59, 155.30, 152.63, 151.86, 151.23, 151.15, 149.67, 148.60, 147.28, 146.82, 146.71, 144.50, 144.21, 144.12, 141.72, 141.28, 125.87, 125.27, 123.89, 123.69, 123.21, 122.66, 109.50, 109.46, 101.19, 96.99, 90.97, 85.37, 84.31, 71.54, 71.44, 71.15, 70.88, 70.86, 69.17, 68.72, 68.49, 64.73, 64.63, 64.53, 59.67, 53.03, 50.96, 50.89, 50.55, 50.52, 47.51, 45.67, 39.52, 37.88, 37.77, 36.25, 36.12, 30.24, 30.10, 29.82, 28.03, 27.77, 26.14, 25.16, 25.06, 24.66, 24.43, 16.55, -4.57. **TOF MS ES +** *m/z*= 2360.1712 (calcd. for $C_{103}H_{143}ClN_{38}O_{24}Si$ 2360.0668 $[M+H]^+$).



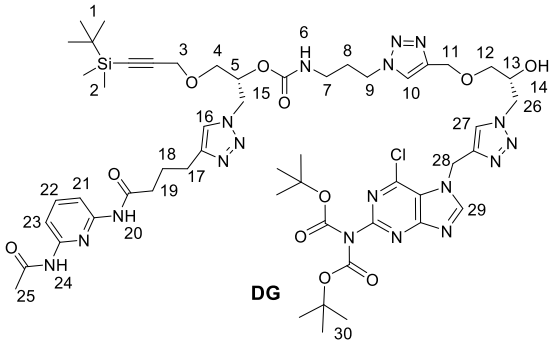
DCGD

Compound **17** (0.731 g, 0.31 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.57 mL, 3 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (50 mL) and washed with brine (3 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The residue was obtained and followed by adding TFA (3 mL) and Milli-Q water (1 mL) stirring at room temperature for 48 h. The crude product was precipitated from the solution with the neutralization of saturated $NaHCO_3$. Then, the solid product was wash with brine, water, acetone and Et_2O, respectively. The final product was dried under vacuum to give the final product all-*R* **DCGD** as foamy yellow solid (0.3 g, 51%). The products all-*S* **DCGD** (0.213 g, 92%) was prepared following the same protocol.

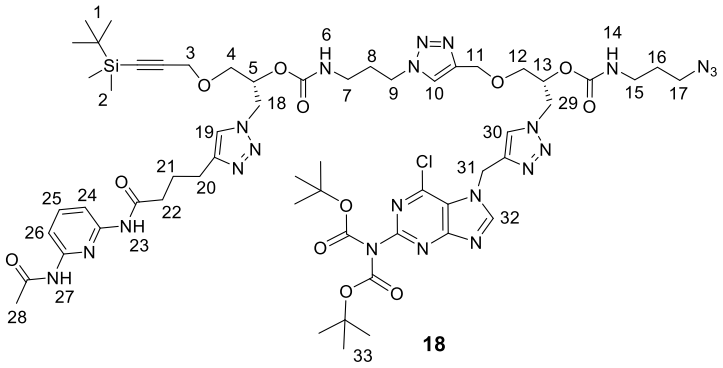
(*R, R, R, R*)-**DCGD**, **^{1}H NMR (500 MHz, DMSO)** δ ^{1}H NMR (500 MHz, DMSO) δ 11.10 (s, 1H, H_{50}), 10.20 – 9.85 (m, 4H, H_{35}, H_{39}, H_{58} and H_{62}), 8.13 – 8.06 (m, 3H, H_9, H_{17} and H_{25}), 7.98 – 7.91 (m, 2H, H_{48} and H_{52}), 7.83 – 7.77 (m, 2H, H_{42} and H_{44}), 7.74 – 7.64 (m, 8H, H_{31}, H_{36}, H_{37}, H_{38}, H_{54}, H_{59}, H_{60} and H_{61}), 7.40 (m, 5.6 Hz, 3H, H_{45} and H_{46}), 7.22 (s, 2H, H_{51}), 6.71 – 6.63 (b, 2H, H_5 and H_{13}), 5.68 (b, 1H, H_{21}), 5.21 (s, 2H, H_{49}), 5.13 – 5.05 (m, 3H, H_4, H_{12} and H_{20}), 4.93 – 4.83 (m, 2H, H_{43}), 4.62 – 4.47 (m, 13H, H_{10}, H_{18}, H_{26}, H_{30}, H_{41}, H_{47} and H_{53}), 4.39 (dd, *J* = 13.9, 3.7 Hz, 1H, H_{53}), 4.35 – 4.26 (m, 6H, H_8, H_{16} and H_{24}), 4.18 (d, *J* = 2.5 Hz, 2H, H_2), 4.10 (d, *J* = 5.5 Hz, 1H, H_{28}), 3.60 – 3.40 (m, 8H, H_3, H_{11}, H_{19} and H_{27}), 2.97 – 2.84 (m, 6H, H_6, H_{14} and H_{22}), 2.68 – 2.57 (m, 4H, H_{32} and H_{55}), 2.48 – 2.41(m, *J* = 7.9 Hz, 5H, H_1, H_{34} and H_{57}), 2.09 (s, 6H, H_{40} and H_{63}), 1.94 – 1.80 (m, 10H, H_7, H_{15}, H_{23}, H_{33} and H_{56}). **^{13}C NMR (126 MHz, DMSO)** δ 172.17, 169.67, 166.31, 157.51, 155.49, 153.92, 151.43, 150.51, 150.44, 146.66, 146.31, 146.15, 143.96, 143.65, 143.12, 142.80, 140.22, 137.53, 124.78, 124.64, 124.38, 124.34, 123.12, 123.07, 116.42, 109.19, 109.08, 94.22, 80.09, 77.80, 71.81, 70.78, 68.84, 68.55, 68.47, 64.01, 58.14, 52.89, 50.28, 50.06, 48.85, 47.21, 43.65, 37.61, 35.83, 35.73, 30.14, 25.04, 24.69, 24.65, 24.20. **TOF MS ES +** *m/z*= 1827,8045 (calcd. for $C_{77}H_{99}N_{38}O_{17}$ 1827,8045 $[M+H]^+$).

(*S, S, S, S*)-**DCGD**, **^{1}H NMR (500 MHz, DMSO)** δ 11.03 (s, 1H, H_{50}), 10.10 – 9.96 (m, 4H, H_{35}, H_{39}, H_{58} and H_{62}), 8.14 – 8.05 (m, 3H, H_9, H_{17}, H_{25}), 8.00 – 7.89 (m, 2H, H_{48} and H_{52}), 7.83 – 7.76 (m, 2H, H_{42} and H_{44}), 7.75 – 7.60 (m, 8H, H_{31}, H_{36}, H_{37}, H_{38}, H_{54}, H_{59}, H_{60} and H_{61}), 7.46 – 7.34 (m, 3H, H_{45} and H_{46}), 7.22 (s, 2H, H_{51}), 6.82 – 6.50 (b, 2H, H_5 and H_{13}), 5.68 (b,, H_{21}), 5.21 (s, 2H, H_{49}), 5.14 – 5.04 (m, 3H, H_4, H_{12} and H_{20}), 4.95 – 4.81 (m, 2H, H_{43}), 4.63 – 4.46 (m, 12H, H_{10}, H_{18}, H_{26}, H_{30}, H_{41} and H_{47}), 4.42 – 4.15 (m, 10H, H_2, H_8, H_{16}, H_{24} and H_{53}), 4.10 (d, *J* = 5.1 Hz, 1H, H_{28}), 3.64 – 3.40 (m, 7H, H_3, H_{11}, H_{19} and H_{27}), 2.91 (m, 6H, H_6, H_{14} and H_{22}), 2.62 – 2.57 (m, 5H, H_1, H_{32} and H_{55}), 2.47 – 2.40 (m, 4H, H_{34} and H_{57}), 2.09 (s, 6H, H_{40} and H_{63}), 1.94 – 1.80 (m, 10H, H_7, H_{15}, H_{23}, H_{33} and H_{56}). **^{13}C NMR (126 MHz, DMSO)** δ 171.87, 169.28, 166.09, 157.19, 155.20, 153.71, 151.19, 150.36, 150.28, 146.38, 146.02, 145.87, 143.70, 143.39, 142.97, 142.66, 139.86, 137.14, 124.47, 124.31, 124.12, 124.09, 122.80, 122.76, 116.32, 109.00, 108.87, 93.82, 79.85, 77.67, 71.66, 70.52, 70.46, 68.64, 68.33, 68.24, 63.85, 57.91, 52.69, 49.99, 49.78, 48.61, 46.95, 37.41, 35.57, 35.48, 29.96, 24.87, 24.54, 24.50, 24.00. **TOF MS ES +** *m/z*= 1827,7686 (calcd. for $C_{77}H_{99}N_{38}O_{17}$ 1827,8045 $[M+H]^+$).

2.3 Synthesis of all-R tetramer DGCD

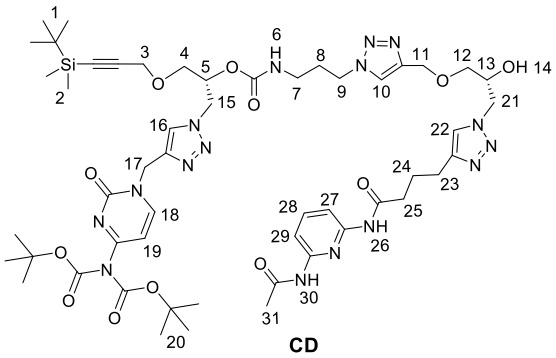


Compound **4** (0.541 g, 0.8 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (1.2 mL, 1.5 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (30 mL) and washed with brine (3 × 30 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound **7** (0.513 g, 1 equiv.) was added into the residue followed by EtOH (5.6 mL) with stirring. Milli-Q water (2.4 mL), sodium ascorbate solution (32 mg in 1 mL water, 0.2 equiv.) and $CuSO_4$ solution (12.8 mg in 0.8 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 2/98 to 8/92) to give the final product **DG** as foamy yellow solid (0.78 g, 81%).**DG**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.41 – 8.89 (m, 2H, H_{20} and H_{24}), 8.43 (s, 1H, H_{29}), 7.99 (s, 1H, H_{10}), 7.86 – 7.72 (m, 3H, H_{21}, H_{23} and H_{27}), 7.65 – 7.51 (m, 2H, H_{16} and H_{22}), 6.14 (t, J = 6.1 Hz, 1H, H_6), 5.51 (s, 2H, H_{28}), 5.27 – 5.14 (m, 1H, H_5), 4.84 (d, J = 5.8 Hz, 1H, H_{15}), 4.66 – 4.07 (m, 10H, H_3, H_9, H_{11}, H_{13}, H_{15} and H_{26}), 3.72 – 3.40 (m, 4H, H_4 and H_{12}), 3.23 – 3.16 (m, 2H, H_7), 2.75 – 2.62 (t, 2H, H_{17}), 2.37 (q, J = 7.1 Hz, 2H, H_{19}), 2.16 (s, 3H, H_{25}), 2.07 – 1.91 (m, 4H, H_8 and H_{18}), 1.43 (s, 18H, H_{30}), 0.90 (s, 9H, H_1), 0.09 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 172.22, 169.58, 155.64, 152.58, 151.93, 151.12, 150.93, 150.07, 149.94, 147.23, 146.57, 144.44, 140.82, 140.50, 129.96, 125.23, 123.82, 122.88, 109.46, 101.20, 90.91, 84.00, 71.53, 71.19, 69.09, 68.45, 64.67, 59.02, 53.26, 50.89, 47.47, 39.32, 37.77, 36.20, 30.26, 28.00, 26.13, 25.12, 24.70, 24.49, 16.54, -4.59. **HRMS** *m/z*= 1203.5397 (calcd. for $C_{53}H_{75}ClN_{18}O_{11}Si$ 1203.5393 $[M+H]^+$)



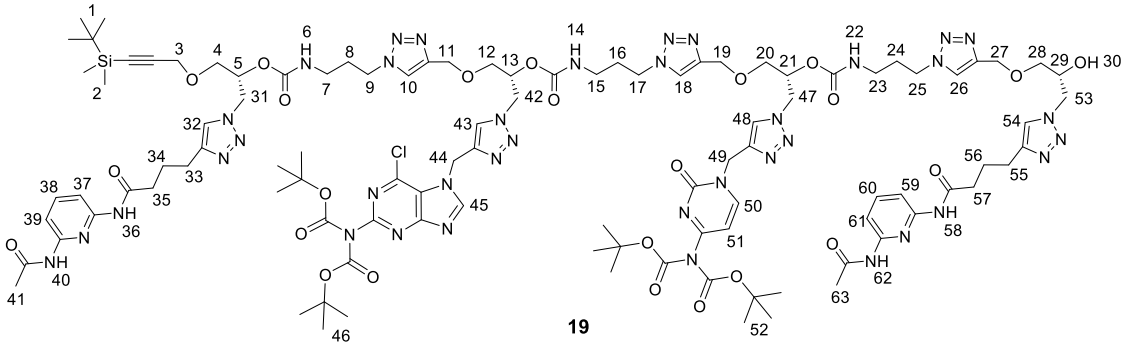
Compound **DC** (0.54 g, 0.45 mmol) was added in a flask followed by DCM (5 mL), 4-nitrophenyl chloroformate (0.227 g, 2.5 equiv.) and pyridine (98 µL, 2.5 equiv.). The mixture was stirred for 2 h at room temperature followed by solvent evaporation. The residue was dissolved in CH_3CN (5 mL) and added to a flask followed by 3-azido-1-propylamine (0.135 g, 3 eq.) and Et_3N (0.314 mL, 5 eq.). The mixture was stirred for 2 h at room temperature, then DCM (50 mL) and Milli-Q water (50 mL) were added to the mixture. The organic phase was washed with water (2 × 50 mL) and brine (2 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The final product **19** was obtained after passing the residue through a chromatography column (MeOH/DCM= 3/97 to 8/92) as a foamy-white solid (0.598 g, 100%). **18**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 8.61 (s, 1H, H_{27}), 8.51 – 8.32 (m, 2H, H_{23} and H_{32}), 7.93 – 7.77 (m, 3H, H_{10}, H_{24} and H_{30}), 7.73 – 7.62 (m, 2H, H_{25} and H_{26}), 7.42 (s, 1H, H_{19}), 5.64 – 5.48 (m, 3H, H_6 and H_{31}), 5.28 (t, J =

5.8 Hz, 1H, H_{14}), 5.10 (s, 1H, H_5), 4.74 – 4.11 (m, 10H, H_3, H_9, H_{11}, H_{13}, H_{18} and H_{29}), 3.86 – 3.37 (m, 4H, H_4 and H_{12}), 3.32 – 2.87 (m, 6H, H_7, H_{15} and H_{17}), 2.74 (d, *J* = 9.2 Hz, 2H, H_{20}), 2.36 (d, *J* = 4.7 Hz, 2H, H_{22}), 2.20 (s, 3H, H_{28}), 2.12 – 1.97 (m, 4H, H_8 and H_{21}), 1.62 (d, *J* = 1.4 Hz, 2H, H_{16}), 1.48 (s, 18H, H_{33}), 0.92 (s, 9H, H_1), 0.10 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** *δ* 171.92, 169.21, 155.52, 155.27, 152.59, 151.91, 151.40, 151.28, 149.96, 149.80, 147.29, 146.41, 144.30, 141.36, 140.67, 130.19, 125.12, 123.62, 122.56, 109.61, 109.57, 101.10, 91.09, 84.33, 71.31, 70.97, 68.45, 68.38, 64.78, 59.74, 50.89, 50.86, 49.01, 47.42, 39.59, 38.50, 37.85, 36.06, 30.30, 28.94, 28.04, 26.15, 25.22, 24.77, 24.26, 16.56, -4.57. **HRMS** *m/z*= 1329.5743 (calcd. for $C_{57}H_{81}ClN_{22}O_{12}Si$ 1329.3935 $[M+H]^+$)



CD

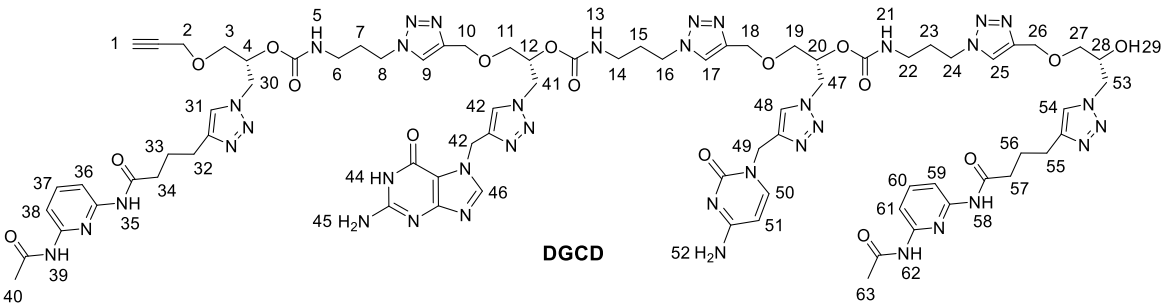
Compound **D** (0.468 g, 0.91 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (1.37 mL, 1.5 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (30 mL) and washed with brine (3 × 30 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound **9** (0.68 g, 1 equiv.) was added into the residue followed by EtOH (6.37 mL) with stirring. Milli-Q water (2.73 mL), sodium ascorbate solution (36.4 mg in 1.5 mL water, 0.2 equiv.) and $CuSO_4$ solution (14.6 mg in 1 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 2/98 to 6/94) to give the final product **CD** as foamy yellow solid (0.624 g, 60%). **CD**, **^{1}H NMR (500 MHz, $CDCl_3$)** *δ* 9.43 – 9.17 (m, 2H, H_{26} and H_{30}), 8.12 (s, 1H, H_{10}), 8.04 – 7.96 (m, 2H, H_{16} and H_{18}), 7.85 – 7.75 (m, 2H, H_{27} and H_{28}), 7.66 (s, 1H, H_{29}), 7.58 (t, *J* = 8.1 Hz, 1H, H_{22}), 6.98 (d, *J* = 7.4 Hz, 1H, H_{19}), 6.57 (b, 1H, H_6), 5.19 – 5.01 (m, 4H, H_5 and H_{17}), 4.72 – 4.32 (m, 8H, H_9, H_{11}, H_{15} and H_{21}), 4.22 – 3.08 (m, 3H, H_3 and H_{13}), 3.68 – 3.45 (m, 4H, H_4 and H_{12}), 3.06 (dd, *J* = 14.4, 6.6 Hz, 2H, H_7), 2.73 (t, *J* = 6.8 Hz, 2H, H_{23}), 2.39 (dt, *J* = 7.2, 3.2 Hz, 2H, H_{25}), 2.18 (s, 3H, H_{31}), 2.10 – 1.98 (m, 4H, H_8 and H_{24}), 1.49 (s, 18H, H_{20}), 0.89 (s, 9H, H_1), 0.08 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** *δ* 172.22, 169.68, 162.72, 155.64, 155.56, 155.24, 149.70, 149.66, 148.36, 147.32, 143.61, 141.74, 140.65, 125.70, 123.74, 122.60, 109.57, 109.54, 101.17, 96.92, 91.02, 85.28, 70.62, 68.48, 64.88, 59.69, 50.86, 50.47, 49.05, 47.64, 45.68, 38.57, 37.84, 36.10, 29.82, 27.79, 26.15, 25.21, 24.64, 24.30, 16.56, -4.57. **HRMS** *m/z*= 1145.5679 (calcd. for $C_{52}H_{76}N_{16}O_{12}Si$ 1145.5671 $[M+H]^+$).



19

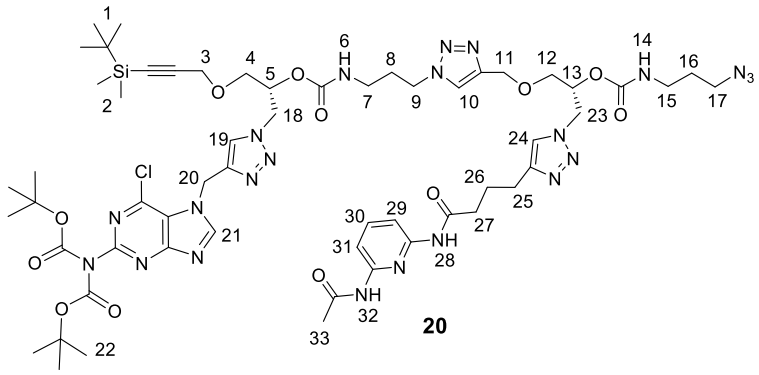
Compound **CD** (0.378 g, 0.33 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.66 mL, 2 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (30 mL) and washed with brine (3 × 30 mL), dried with

Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound **18** (0.44 g, 1 equiv.) was added into the residue followed by EtOH (2.31 mL) with stirring. Milli-Q water (0.99 mL), sodium ascorbate solution (13.2 mg in 0.5 mL water, 0.2 equiv.) and $CuSO_4$ solution (5.3 mg in 0.3 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 3/97 to 12/88) to give the final product **19** as foamy yellow solid (0.584 g, 75%). **19**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.00 – 8.67 (m, 4H, H_{36}, H_{40}, H_{58} and H_{62}), 8.43 (s, 1H, H_{45}), 7.94 – 7.38 (m, 14H, H_{10}, H_{18}, H_{26}, H_{32}, H_{37}, H_{38}, H_{39}, H_{43}, H_{45}, H_{49}, H_{54}, H_{59}, H_{60} and H_{61}), 7.02 (d, J = 7.3 Hz, 1H, H_{51}), 6.31 – 6.06 (b, 2H, H_6 and H_{14}), 5.93 (b, 1H, H_{22}), 5.52 (s, 2H, H_{44}), 5.38 – 4.95 (m, 5H, H_5, H_{13}, H_{21} and H_{49}), 4.77 – 4.08 (m, 24H, H_3, H_9, H_{11}, H_{17}, H_{19}, H_{25}, H_{27}, H_{29}, H_{31}, H_{42}, H_{47} and H_{53}), 3.78 – 3.41 (m, 8H, H_4, H_{12}, H_{20} and H_{28}), 3.25 – 2.87 (m, 6H, H_7, H_{15} and H_{23}), 2.69 (d, J = 7.8 Hz, 4H, H_{33} and H_{55}), 2.44 – 1.83 (m, 20H, H_8, H_{16}, H_{24}, H_{34}, H_{36}, H_{41}, H_{56}, H_{58} and H_{63}), 1.55 – 1.38 (m, 36H, H_{46} and H_{52}), 0.91 (s, 9H, H_1), 0.09 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 172.15, 169.58, 162.69, 155.68, 155.65, 155.57, 155.27, 152.63, 151.87, 151.21, 151.15, 149.68, 148.58, 147.29, 146.83, 146.74, 144.48, 144.24, 144.09, 141.70, 141.29, 125.86, 125.24, 123.86, 123.79, 123.19, 122.65, 109.54, 109.50, 101.17, 96.93, 91.00, 85.36, 84.29, 71.42, 71.16, 70.93, 69.20, 68.62, 68.46, 64.65, 59.68, 53.00, 50.87, 50.56, 47.54, 45.60, 39.47, 37.82, 36.24, 36.11, 30.28, 30.17, 29.81, 28.02, 27.77, 26.13, 25.17, 25.09, 24.65, 24.42, 16.55, -4.58. **TOF MS ES +** *m/z*= 2360.1666 (calcd. for $C_{103}H_{143}ClN_{38}O_{24}Si$ 2360.0668 $[M+H]^+$).



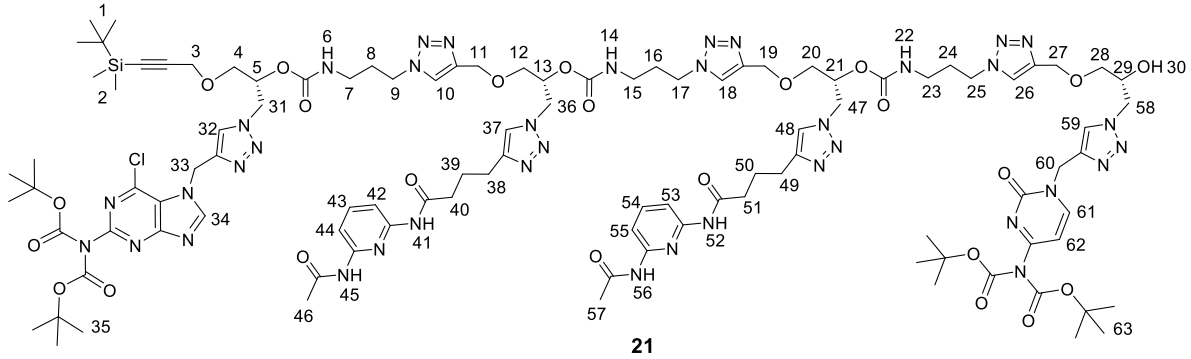
Compound **19** (0.5 g, 0.21 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.63 mL, 3 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (50 mL) and washed with brine (3 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The residue was obtained and followed by adding TFA (3 mL) and Milli-Q water (1 mL) stirring at room temperature for 48 h. The crude product was precipitated from the solution with the neutralization of saturated $NaHCO_3$. Then, the solid product was wash with brine, water, acetone and Et_2O, respectively. The final product was dried under vacuum to give the final product all-*R* **DGCD** as foamy yellow solid (0.32 g, 83%). **DGCD**, **^{1}H NMR (500 MHz, DMSO)** δ 11.17 (s, 1H, H_{44}), 10.03 (dd, J = 24.2, 8.0 Hz, 4H, 4H, H_{35}, H_{39}, H_{58} and H_{62}), 8.18 – 8.02 (m, 3H, H_9, H_{17} and H_{25}), 7.95 (d, J = 16.5 Hz, 2H, H_{44} and H_{46}), 7.86 – 7.59 (m, 10H, H_{31}, H_{36}, H_{37}, H_{38}, H_{42}, H_{44}, H_{54}, H_{59}, H_{60} and H_{61}), 7.48 – 7.32 (m, 3H, H_{51} and H_{52}), 7.22 (s, 1H, H_{45}), 7.05 (b, 1H, H_5), 6.72 (b, 1H, H_{13}), 5.68 (b, 1H, H_{21}), 5.37 – 5.02 (m, 5H, H_4, H_{12}, H_{20} and H_{49}), 4.87 (dd, J = 8.7, 6.1 Hz, 2H, H_{43}), 4.70 – 4.43 (m, 12H, H_{10}, H_{18}, H_{26}, H_{30}, H_{36} and H_{47}), 4.43 – 4.07 (m, 10H, H_2, H_8, H_{16}, H_{24} and H_{53}), 3.95 (d, J = 15.6 Hz, 1H, H_{28}), 3.61 – 3.39 (m, 8H, H_3, H_{11}, H_{19} and H_{27}), 3.03 – 2.78 (m, 6H, H_6, H_{14} and H_{22}), 2.70 – 2.56 (m, 4H, H_{32} and H_{55}), 2.45 (d, J = 8.3 Hz, 4H, H_{34} and H_{57}), 2.09 (s, 6H, H_{40} and H_{63}), 1.97 – 1.75 (m, 10H, H_7, H_{15}, H_{23}, H_{33} and H_{56}). **^{13}C NMR (126 MHz, DMSO)** δ 171.88, 169.29, 166.02, 157.33, 155.22, 153.80, 151.21, 150.37, 150.29, 146.39, 146.02, 145.92, 143.68, 143.39, 143.00, 142.65, 139.88, 137.12, 124.54, 124.26, 124.13, 122.82, 122.77, 116.28, 109.02, 108.88, 93.85, 79.85, 77.69, 71.65, 70.53, 70.48, 68.67, 68.34, 68.24, 63.85, 57.92, 52.70, 50.13, 49.94, 48.62, 46.96, 37.42, 35.58, 35.49, 29.99, 24.88, 24.55, 24.50, 24.01. **TOF MS ES +** *m/z*= 1827, 8105 (calcd. for $C_{77}H_{99}N_{38}O_{17}$ 1827,8045 $[M+H]^+$).

2.4 Synthesis of all-R tetramer GDDC



20

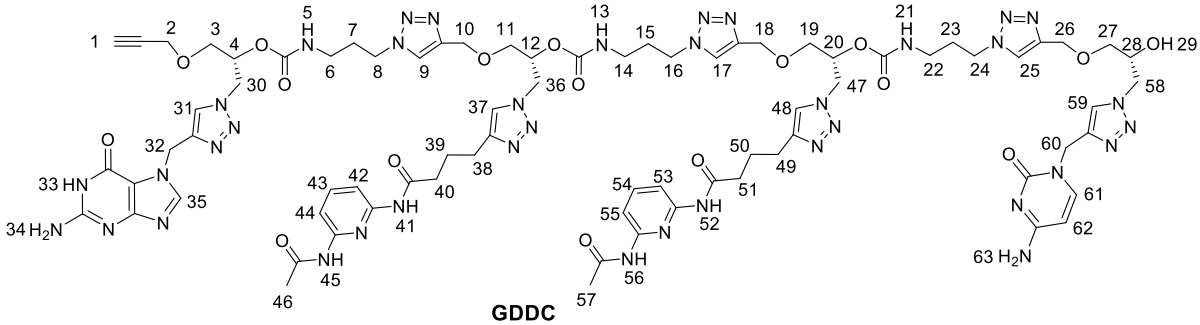
Compound **GD** (0.3 g, 0.25 mmol) was added in a flask followed by DCM (5 mL), 4-nitrophenyl chloroformate (0.126 g, 2.5 equiv.) and pyridine (51 µL, 2.5 equiv.). The mixture was stirred for 2 h at room temperature followed by solvent evaporation. The residue was dissolved in CH_3CN (5 mL) and added to a flask followed by 3-azido-1-propylamine (0.75 g, 3 eq.) and Et_3N (0.139 mL, 5 eq.). The mixture was stirred for 2 h at room temperature, then DCM (30 mL) and Milli-Q water (30 mL) were added to the mixture. The organic phase was washed with water (2 × 30 mL) and brine (2 × 30 mL), dried with Na_2SO_4 and concentrated under vacuum. The final product **20** was obtained after passing the residue through a chromatography column (MeOH/DCM= 2/98 to 6/94) as a foamy-white solid (0.335 g, 100%).**20**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.37 – 8.18 (m, 3H, H_{21}, H_{28} and H_{32}), 7.99 – 7.55 (m, 5H, H_{10}, H_{19}, H_{29}, H_{30} and H_{31}), 7.41 (s, 1H, H_{24}), 5.93 – 5.64 (b, 2H, H_6 and H_{14}), 5.53 (s, 2H, H_{20}), 5.59– 5.46 (m, 2H, H_5 and H_{13}), 4.78 – 4.44 (m, 6H, H_{11}, H_{18} and H_{23}), 4.38 – 4.02 (m, 4H, H_3 and H_9), 3.69 – 3.57 (m, 2H, H_4), 3.33 (t, *J* = 6.5 Hz, 2H, H_{12}), 3.21 (q, *J* = 6.5 Hz, 2H, H_{17}), 3.14 – 3.05 (m, 4H, H_7 and H_{15}), 2.74 (d, *J* = 7.2 Hz, 2H, H_{25}), 2.38 (t, *J* = 7.4 Hz, 2H, H_{27}), 2.19 (s, 3H, H_{33}), 2.08 – 1.89 (m, 4H, H_8 and H_{26}), 1.73 (q, *J* = 6.6 Hz, 2H, H_{16}), 1.48 (s, 18H, H_{22}), 0.90 (s, 9H, H_1), 0.08 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 172.05, 169.39, 155.44, 152.60, 151.87, 151.26, 151.26, 149.88, 149.69, 147.09, 146.58, 144.27, 141.31, 140.72, 130.14, 125.15, 123.65, 122.74, 109.50, 101.11, 90.90, 84.37, 71.06, 68.25, 68.14, 64.69, 59.61, 51.13, 50.22, 49.07, 47.44, 39.63, 38.62, 37.76, 36.13, 30.11, 29.03, 28.03, 26.11, 25.04, 24.71, 24.30, 16.52, -4.60. **HRMS** *m/z*= 1329.5933 (calcd. for $C_{57}H_{81}ClN_{22}O_{12}Si$ 1329.5935 $[M+H]^+$).



21

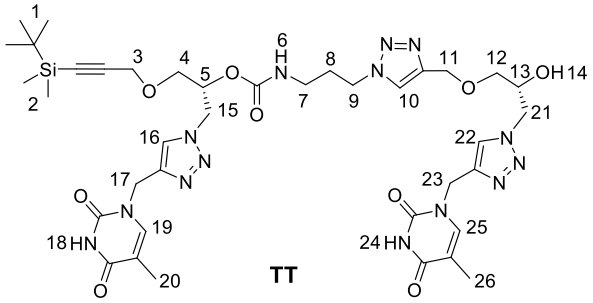
Compound **DC** (0.241 g, 0.21 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.42 mL, 2 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (30 mL) and washed with brine (3 × 30 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound **20** (0.28 g, 1 equiv.) was added into the residue followed by EtOH (1.47 mL) with stirring. Milli-Q water (0.63 mL), sodium ascorbate solution (8.4 mg in 0.3 mL water, 0.2 equiv.) and $CuSO_4$ solution (3.4 mg in 0.2 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 3/97 to 14/86) to give the final product **21** as foamy yellow solid (0.43 g, 87%). **21**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 9.40 – 8.37 (m, 4H, H_{41}, H_{45}, H_{52} and H_{56}), 8.44 (s, 1H, H_{34}), 8.02 – 7.54 (m, 12H, H_{10}, H_{18}, H_{26}, H_{32}, H_{42}, H_{43}, H_{44}, H_{53}, H_{54}, H_{55}, H_{59} and H_{61}), 7.51– 7.39 (m, 2H, H_{37} and

H_{48}), 7.03 (t, *J* = 7.2 Hz, 1H, H_{62}), 6.36 (b, 2H, H_6 and H_{14}), 5.86 (b, 1H, H_{22}), 5.54 (s, 2H, H_{33}), 5.12 (m,, 5H, H_5, H_{13}, H_{21} and H_{60}), 4.70 – 4.09 (m, 24H, H_3, H_9, H_{11}, H_{17}, H_{19}, H_{25}, H_{27}, H_{29}, H_{31}, H_{36}, H_{47} and H_{58}), 3.79 – 3.42 (m, 7H, H_4, H_{12}, H_{20} and H_{28}), 3.17 – 2.83 (m, 6H, H_7, H_{15} and H_{23}), 2.77 – 2.59 (m, 4H, H_{38} and H_{49}), 2.46 – 2.11 (m, 10H, H_{40}, H_{46}, H_{51} and H_{57}), 2.08 – 1.80 (m, 11H, H_8, H_{16}, H_{24}, H_{39} and H_{50}), 1.49 (m, 36H, H_{35} and H_{63}), 0.90 (s, 9H, H_1), 0.08 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** *δ* 172.35, 169.72, 162.67, 155.85, 155.80, 155.46, 155.24, 152.63, 151.86, 151.30, 151.18, 149.69, 148.53, 147.19, 146.91, 146.77, 144.60, 144.17, 144.12, 141.63, 141.36, 125.66, 125.26, 123.94, 123.78, 122.91, 109.65, 109.55, 101.16, 96.88, 90.90, 85.33, 84.41, 71.55, 71.06, 69.08, 68.67, 68.31, 64.71, 64.53, 59.61, 53.19, 51.17, 50.52, 47.58, 45.56, 39.62, 37.85, 37.78, 36.06, 29.82, 28.05, 27.78, 26.14, 25.05, 24.60, 24.46, 16.54, -4.58. **TOF MS ES +** *m/z*= 2360.2353 (calcd. for $C_{103}H_{143}ClN_{38}O_{24}Si$ 2360.0668 $[M+H]^+$).



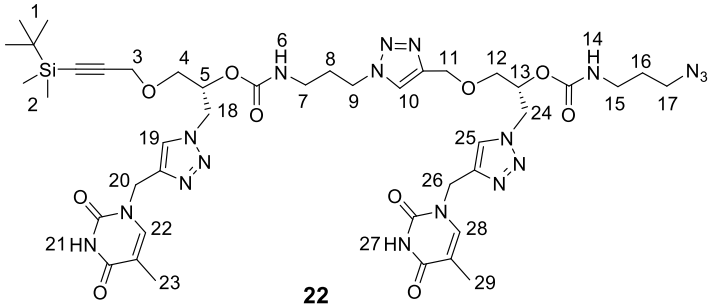
Compound **21** (0.283 g, 0.12 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.36 mL, 3 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (50 mL) and washed with brine (3 × 50 mL), dried with Na_2SO_4 and concentrated under vacuum. The residue was obtained and followed by adding TFA (3 mL) and Milli-Q water (1 mL) stirring at room temperature for 48 h. The crude product was precipitated from the solution with the neutralization of saturated $NaHCO_3$. Then, the solid product was wash with brine, water, acetone and Et_2O, respectively. The final product was dried under vacuum to give the final product all-*R* **GDDC** as foamy yellow solid (0.191 g, 87%). **GDDC**, **^{1}H NMR (500 MHz, DMSO)** *δ* 10.97 (s, 1H, H_{33}), 9.25 – 9.08 (m, 4H, H_{41}, H_{45}, H_{52} and H_{56}), 8.16 – 8.05 (m, 3H, H_9, H_{17} and H_{25}), 8.02 (s, 1H, H_{35}), 7.91 (s, 1H, H_{31}), 7.84 – 7.57 (m, 9H, H_{37}, H_{42}, H_{43}, H_{44}, H_{48}, H_{53}, H_{54}, H_{55} and H_{59}), 7.50 – 7.36 (m, 3H, H_{62} and H_{63}), 7.13 (d, *J* = 79.0 Hz, 2H, H_{34}), 6.87 – 6.66 (b, 2H, H_5 and H_{13}), 5.69 (b, 1H, H_{21}), 5.20 (s, 2H, H_{32}), 5.12 – 5.02 (m, 3H, H_4, H_{12} and H_{20}), 4.88 (d, *J* = 2.5 Hz, 2H, H_{60}), 4.73 – 4.04 (m, 22H, H_2, H_8, H_{10}, H_{16}, H_{18}, H_{24}, H_{26}, H_{30}, H_{36}, H_{47} and H_{58}), 4.00 – 3.89 (m, 1H, H_{28}), 3.71 – 3.41 (m, 8H, H_3, H_{11}, H_{19} and H_{27}), 2.91 (q, *J* = 6.4 Hz, 6H, H_6, H_{14} and H_{22}), 2.67 – 2.53 (m, 5H, H_1, H_{38} and H_{49}), 2.44 (t, *J* = 7.4 Hz, 4H, H_{40} and H_{51}), 2.08 (d, *J* = 3.4 Hz, 6H, H_{46} and H_{57}), 1.95 – 1.77 (m, 10H, H_7, H_{15}, H_{23}, H_{38} and H_{49}). **^{13}C NMR (126 MHz, DMSO)** *δ* 171.92, 169.36, 166.08, 156.95, 155.30, 155.22, 153.93, 150.39, 150.30, 146.40, 145.87, 143.71, 143.42, 142.80, 139.90, 137.06, 124.53, 124.45, 124.21, 124.12, 122.81, 109.03, 108.90, 93.78, 79.84, 79.29, 77.71, 71.62, 70.64, 70.45, 68.66, 68.23, 63.85, 57.94, 52.83, 50.03, 49.98, 48.62, 46.99, 37.43, 35.53, 30.03, 24.89, 24.52, 24.04. **TOF MS ES +** *m/z*= 1827, 7659 (calcd. for $C_{77}H_{99}N_{38}O_{17}$ 1827,8045 $[M+H]^+$).

2.5 Synthesis of all-R tetramer TTTT



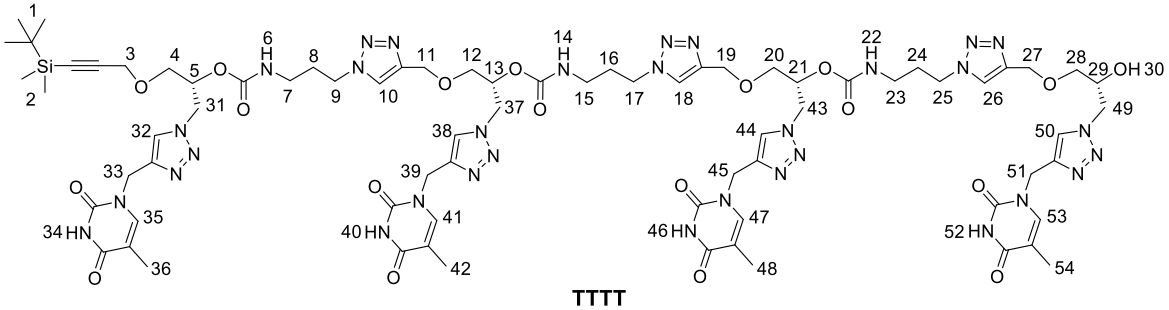
Compound **T** (0.334 g, 0.77 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (1.16 mL, 1.5 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by solvent evaporation. The residue was dissolved in DCM (30 mL) and washed with brine (3 × 30 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification.

Compound **6** (0.431 g, 1 equiv.) was added into the residue followed by EtOH (5.39 mL) with stirring. Milli-Q water (2.31 mL), sodium ascorbate solution (30.8 mg in 1.2 mL water, 0.2 equiv.) and $CuSO_4$ solution (12.3 mg in 1 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 2/98 to 10/90) to give the final product **TT** as foamy white solid (0.59 g, 87%). **TT**, **^{1}H NMR (500 MHz, $CDCl_3$)** δ 10.26 (s, 1H, H_{18}), 10.05 (s, 1H, H_{24}), 8.02 – 7.92 (m, 2H, H_{16} and H_{22}), 7.84 (s, 1H, H_{10}), 7.39 – 7.31 (m, 2H, H_{19} and H_{25}), 6.18 (t, *J* = 6.0 Hz, 1H, H_6), 5.16 (d, *J* = 3.6 Hz, 1H, H_5), 4.91 (d, *J* = 12.8 Hz, 4H, H_{17} and H_{23}), 4.72 – 4.35 (m, 8H, H_9, H_{11}, H_{15} and H_{21}), 4.25 – 4.13 (m, 3H, H_3 and H_{13}), 3.72 – 3.47 (m, 4H, H_4 and H_{12}), 3.15 – 3.00 (m, 2H, H_7), 2.12 – 1.97 (m, 2H, H_8), 1.84 (dd, *J* = 7.8, 1.1 Hz, 6H, H_{20} and H_{26}), 0.91 (s, 9H, H_1), 0.09 (s, 6H, H_2). **^{13}C NMR (126 MHz, $CDCl_3$)** δ 164.82, 164.76, 155.60, 151.57, 151.49, 144.53, 142.26, 141.91, 140.67, 125.56, 123.88, 111.33, 111.27, 101.31, 90.78, 71.59, 71.13, 69.09, 68.27, 64.66, 59.59, 53.15, 50.99, 47.49, 43.19, 37.74, 30.21, 26.15, 16.54, 12.39, -4.58. **HRMS** *m/z*= 879.4043 (calcd. for $C_{38}H_{54}N_{14}O_9Si$ 879.4040 $[M+H]^+$).



22

Compound (*R*)-**TT** (0.282 g, 0.32 mmol) was added in a flask followed by DCM (5 mL), 4-nitrophenyl chloroformate (0.161 g, 2.5 equiv.) and pyridine (65 µL, 2.5 equiv.). The mixture was stirred for 2 h at room temperature followed by solvent evaporation. The residue was dissolved in CH_3CN (5 mL) and added to a flask followed by 3-azido-1-propylamine (0.96 g, 3 eq.) and Et_3N (0.162 mL, 5 eq.). The mixture was stirred for 2 h at room temperature, then DCM (30 mL) and Milli-Q water (30 mL) were added to the mixture. The organic phase was washed with water (2 × 30 mL) and brine (2 × 30 mL), dried with Na_2SO_4 and concentrated under vacuum. The final product **23** was obtained after passing the residue through a chromatography column (MeOH/DCM= 2/98 to 10/90) as a foamy-white solid (0.28 g, 87%). **22, ^{1}H NMR (500 MHz, DMSO)** δ 11.30 (s, 2H, H_{21} and H_{27}), 8.12 – 7.96 (m, 3H, H_{10}, H_{16} and H_{22}), 7.56 (dd, *J* = 2.3, 1.3 Hz, 2H, H_{19} and H_{25}), 7.46 – 7.27 (b, 2H, H_6 and H_{14}), 5.24 – 4.99 (m, 2H, H_5 and H_{13}), 4.88 (t, *J* = 1.7 Hz, 4H, H_{20} and H_{26}), 4.70 – 4.46 (m, 6H, H_{11}, H_{18} and H_{24}), 4.40 – 4.16 (m, 4H, H_3 and H_9), 3.70 – 3.46 (m, 4H, H_4 and H_{12}), 3.29 (t, *J* = 6.7 Hz, 2H, H_{17}), 3.02 – 2.81 (m, 4H, H_7 and H_{15}), 1.93 – 1.83 (m, 2H, H_8), 1.74 (dd, *J* = 5.6, 1.2 Hz, 6H, H_{23} and H_{29}), 1.56 (t, *J* = 6.8 Hz, 2H, H_{16}), 0.89 (s, 9H, H_1), 0.07 (s, 6H, H_2). **^{13}C NMR (126 MHz, DMSO)** δ 164.28, 155.16, 150.72, 143.42, 142.39, 142.34, 141.03, 124.41, 124.12, 108.85, 102.78, 89.27, 70.49, 70.44, 68.70, 68.27, 63.87, 58.61, 50.12, 48.18, 46.97, 42.01, 37.52, 37.42, 29.99, 28.49, 25.85, 16.11, 11.96, -4.79. **HRMS** *m/z*= 1005.4585 (calcd. for $C_{42}H_{60}N_{18}O_{10}Si$ 1005.4582 $[M+H]^+$).



TTTT

Compound **TT** (0.193 g, 0.22 mmol) was added in a flask followed by EtOAc (5 mL) and TBAF (0.44 mL, 2 equiv.). The mixture was stirred for 30 min at room temperature. 5 mL MeOH was added to quench the reaction followed by

solvent evaporation. The residue was dissolved in DCM (30 mL) and washed with brine (3 × 30 mL), dried with Na_2SO_4 and concentrated under vacuum. The crude product was obtained without any further purification. Compound **22** (0.201 g, 1 equiv.) was added into the residue followed by EtOH (1.54 mL) with stirring. Milli-Q water (0.66 mL), sodium ascorbate solution (8.8 mg in 0.3 mL water, 0.2 equiv.) and $CuSO_4$ solution (3.5 mg in 0.2 mL water, 0.1 equiv.) were successively added to the reaction mixture. The reaction was stirred for 2 h at 50 °C followed by cooling to room temperature. Then, DCM (30 mL) and Na_2EDTA (0.05 M, 30 mL) were added into the mixture followed by air bubbling for 30 min. After that, the organic phase was collected, and the water phase was extracted with DCM (30 × 3 mL). The organic phases were combined, dried with Na_2SO_4 and concentrated under vacuum. The residue was purified by column chromatography (MeOH /DCM = 3/97 to 14/86) to give the final product **TTTT** as foamy white solid (0.322 g, 83%). **TTTT**, **^{1}H NMR (500 MHz, DMSO)** δ 11.30 (d, J = 2.1 Hz, 4H, H_{34}, H_{40}, H_{46}, H_{52}), 8.10 (d, J = 3.3 Hz, 3H, H_{32}, H_{38} and H_{44}), 8.05 – 7.94 (m, 4H, H_{10}, H_{18}, H_{26} and H_{50}), 7.63 – 7.53 (m, 4H, H_{35}, H_{41}, H_{47} and H_{53}), 7.36 – 7.44 (b, 3H, H_6, H_{14} and H_{22}), 5.17 – 5.06 (m, 3H, H_5, H_{13} and H_{21}), 4.88 (d, J = 6.1 Hz, 8H, H_{33}, H_{39}, H_{45} and H_{51}), 4.65 – 4.46 (m, 12H, H_{11}, H_{19}, H_{27}, H_{32}, H_{37} and H_{43}), 4.45 – 4.20 (m, 10H, H_3, H_9, H_{17}, H_{25} and H_{49}), 4.00 – 3.92 (m, 1H, H_{29}), 3.56 (t, J = 24.1, 10.6, 4.7 Hz, 6H, H_4, H_{12} and H_{20}), 3.44 – 3.36 (m, 2H, H_{17}), 2.90 (d, J = 7.4 Hz, 6H, H_7, H_{15} and H_{23}), 1.94 – 1.82 (m, 6H, H_8, H_{16} and H_{24}), 1.78 – 1.67 (m, 12H, H_{36}, H_{42}, H_{48} and H_{54}), 0.88 (s, 9H, H_1), 0.07 (s, 6H, H_2). **^{13}C NMR (126 MHz, DMSO)** δ 164.28, 155.21, 150.73, 143.70, 143.41, 142.38, 142.13, 141.18, 141.04, 124.47, 124.43, 124.14, 124.09, 108.86, 102.78, 89.27, 71.64, 70.56, 70.43, 68.69, 68.26, 68.21, 63.88, 63.86, 58.61, 54.94, 52.90, 50.11, 46.96, 42.15, 42.03, 37.41, 29.97, 25.84, 16.10, 11.95, -4.79. **TOF MS ES +** m/z= 1769,7665 (calcd. for $C_{74}H_{100}N_{32}O_{19}Si$ 1769,7685 $[M+H]^+$).

S3. TOF-MS/MS of the oligomers

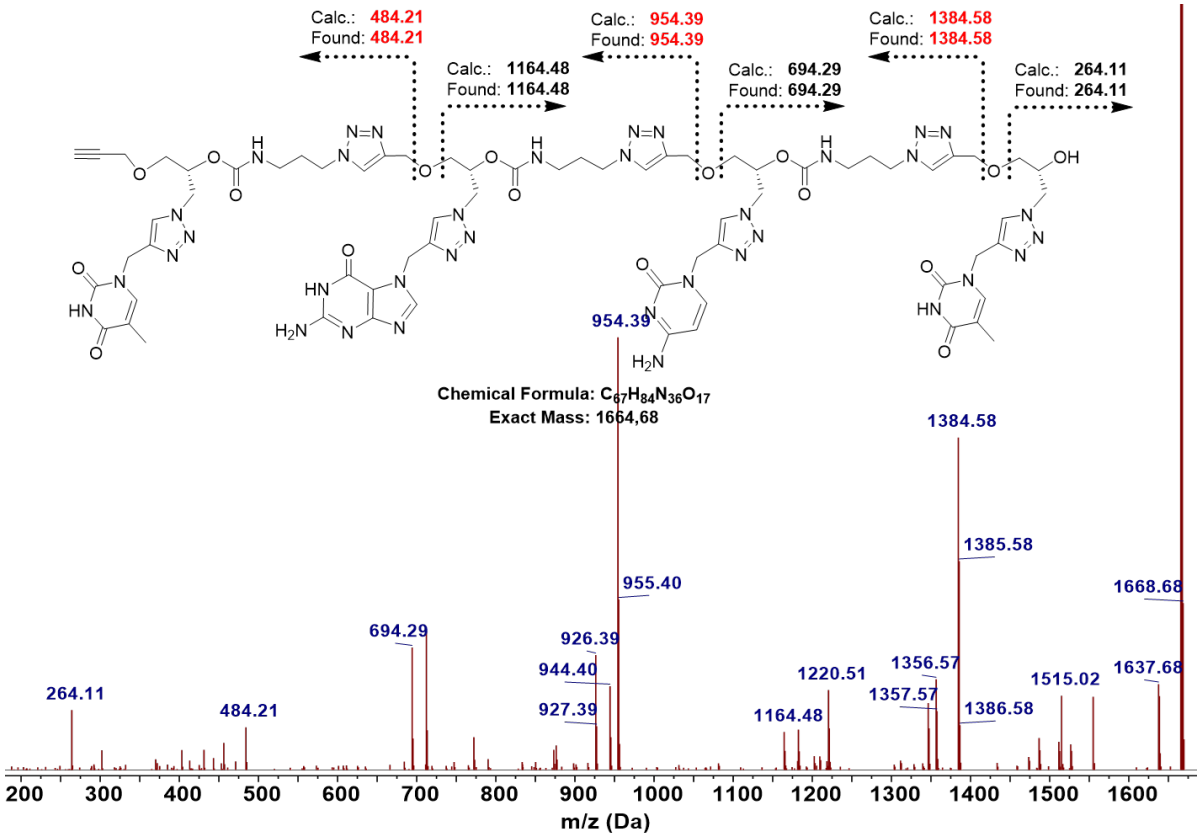


Figure S3.1 - TOF-MS/MS of TGCT.

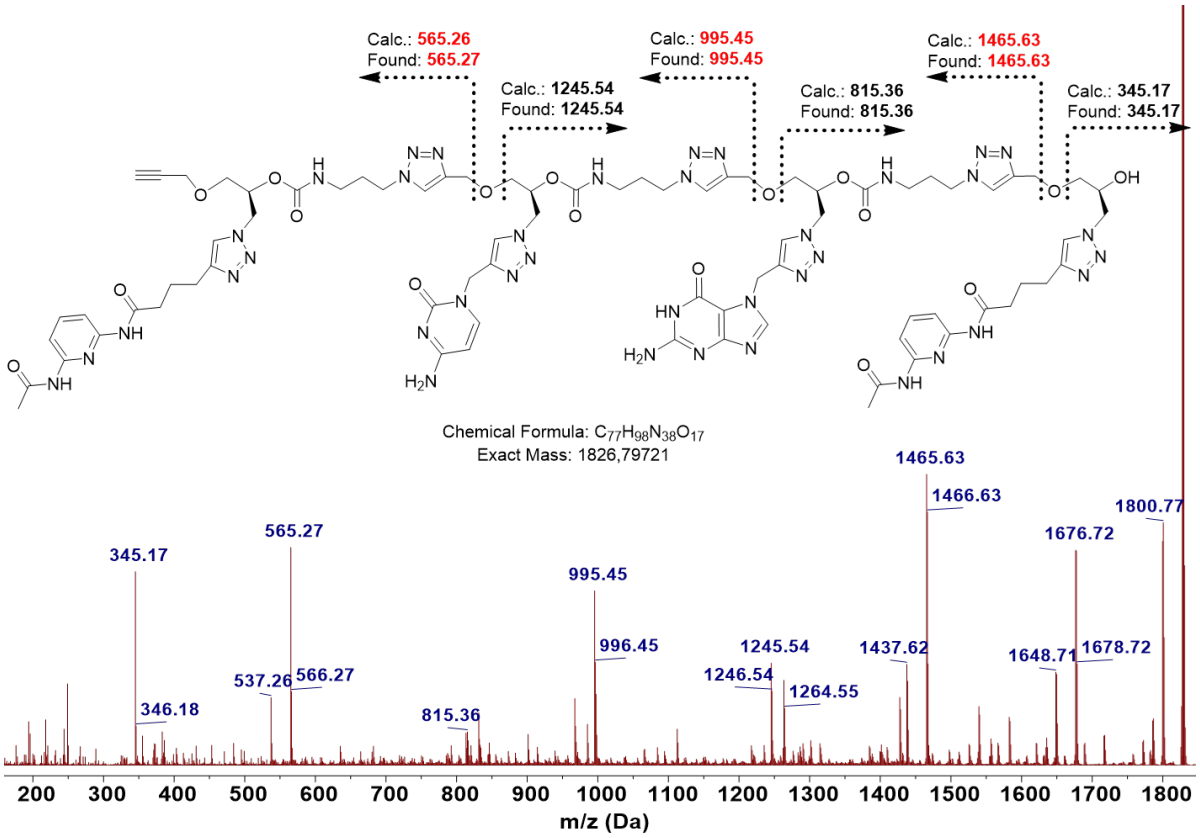


Figure S3.2 - TOF-MS/MS of all-*R* DCGD.

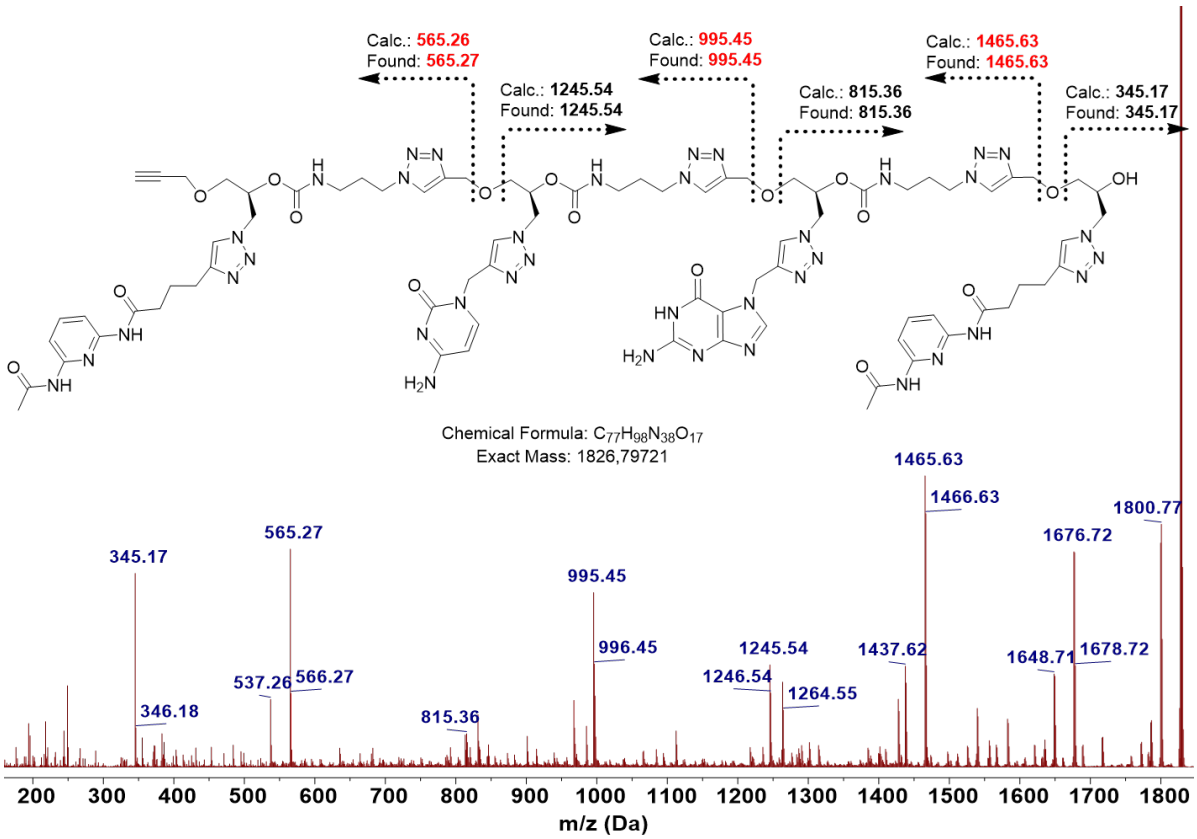


Figure S3.3 - TOF-MS/MS of all-*S* DCGD.

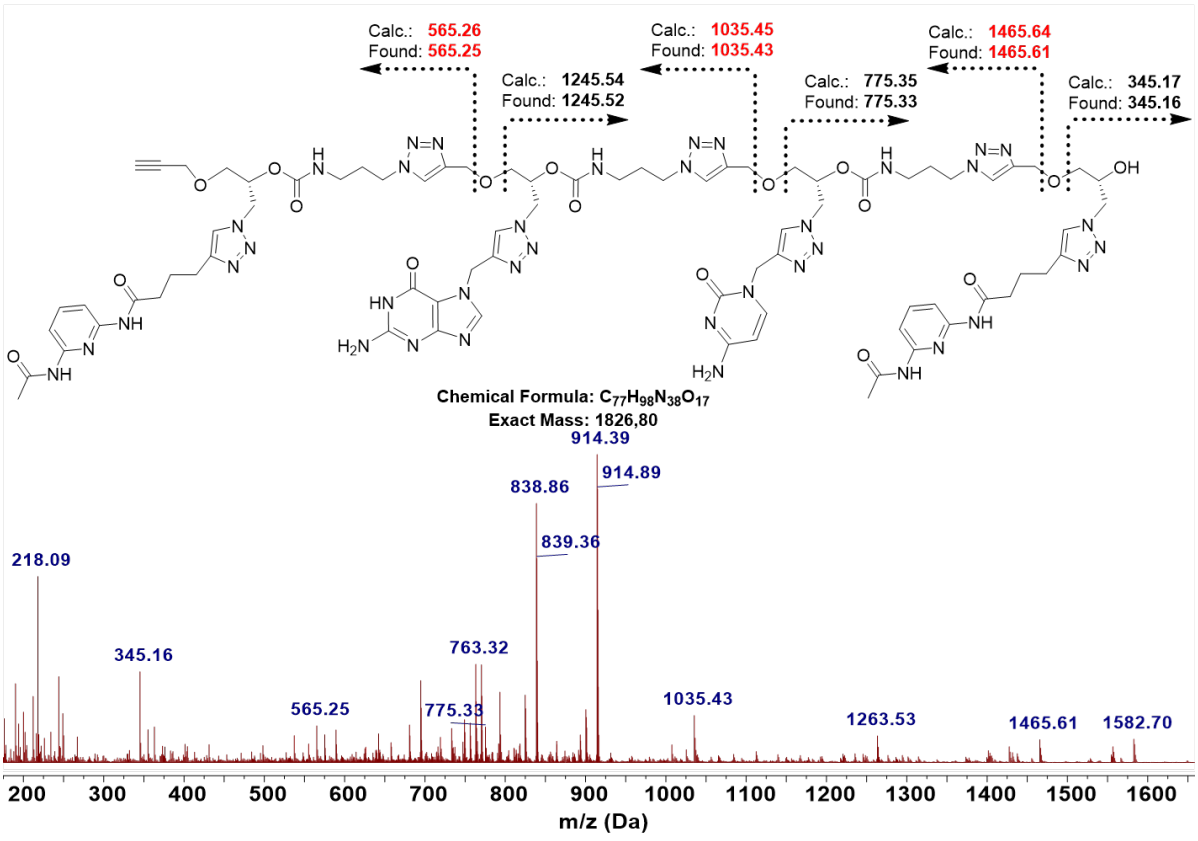


Figure S3.4 - TOF-MS/MS of DGCD.

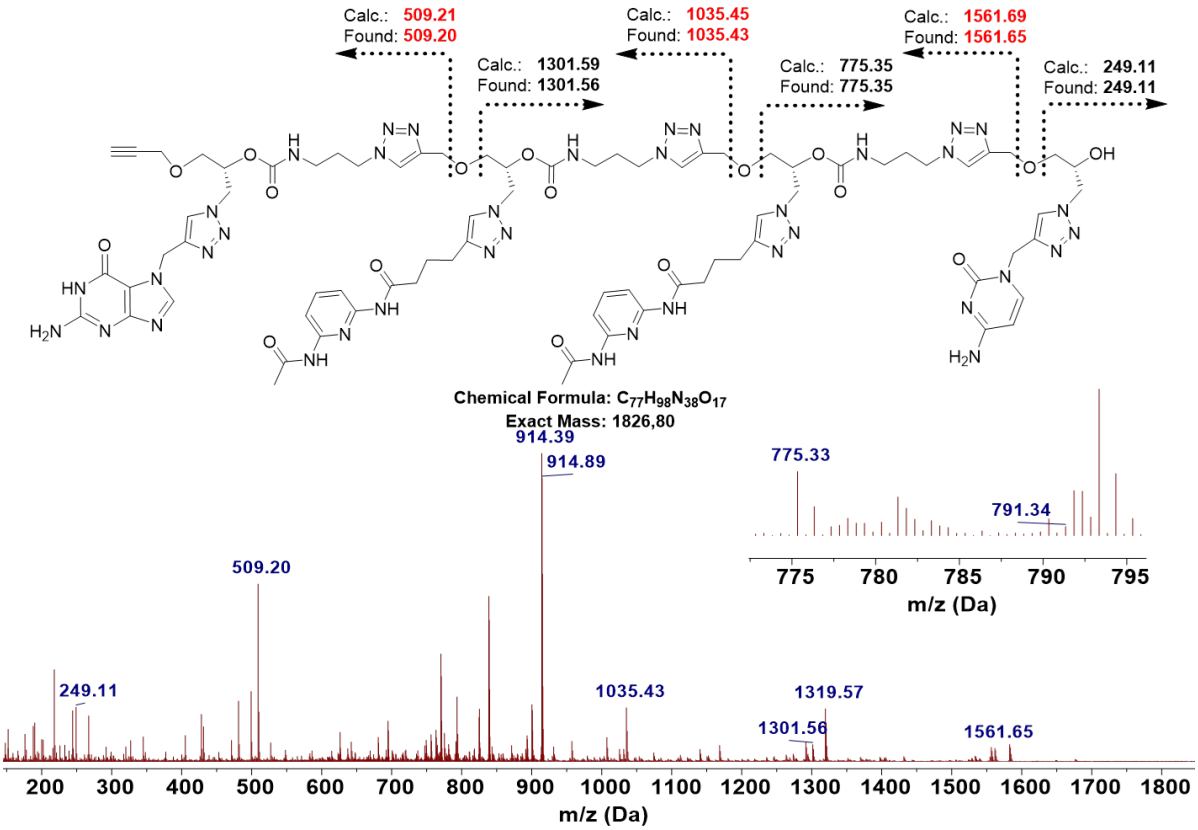


Figure S3.5 - TOF-MS/MS of GDDC.

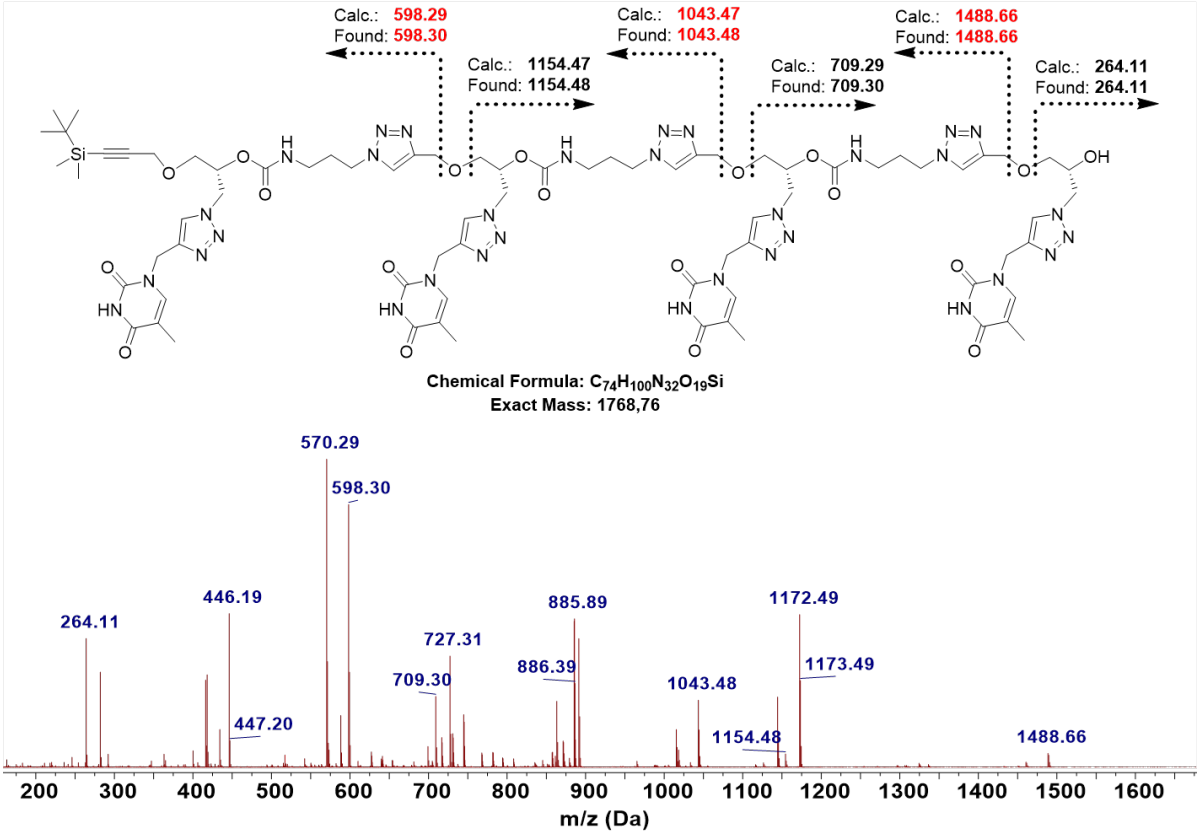


Figure S3.6 - TOF-MS/MS of TTTT.

S4.1 General formalism

S4.1.1 Mass balance equations

Consider two molecules, A and B, which can form a trivalent hydrogen-bonded duplex $A\cdot\overset{3}{\cdot}\cdot B$ in which two donor sites of A (hereafter called Dn1 and Dn3) and one donor site of B (hereafter called Dn2) are involved. The protons associated to these sites are numbered 1 and 3 for A, and 2 for B. Reciprocally, A and B have one and two hydrogen bond-accepting sites, respectively, hereafter called Ac2, Ac1 and Ac3, respectively. This is for instance the case for the pairs of monomers D and T, and G and C, displayed in Figure S4.1.

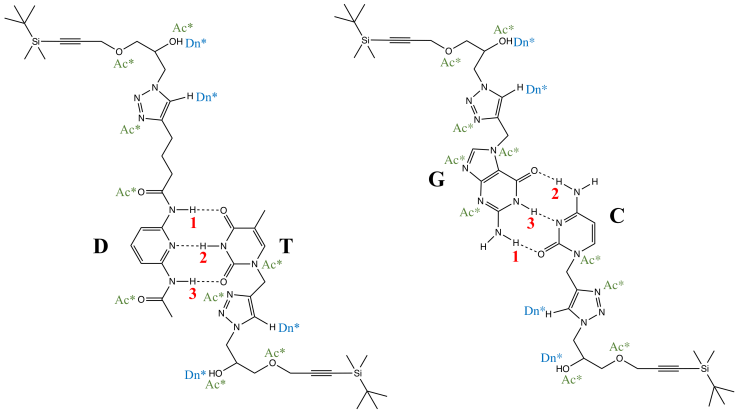


FIGURE S4.1 – Chemical structure of the H-bonding monomers of this study, with the main trivalent duplexes shown.

Additionally, there are a_A and a_B generic accepting sites Ac* not involved in the main binding triad of A and B, respectively (here, we neglect the dual donating/accepting character of Dn1, Dn2 and Dn3). Likewise, there are d_A and d_B generic donor sites Dn* in A and B, respectively. For instance, for the D/T pair, $a_D = 5$, $a_T = 4$, and $d_D = d_T = 2$; for the G/C pair, $a_G = 6$, $a_C = 4$, and $d_G = d_B = 2$.

In addition to the desired trivalent duplex, A and B can form a series of other hydrogen-bonded complexes. Here, we will we restrict ourselves to measurements performed in hydrogen bond-accepting solvents ; therefore, the generic acceptor sites Ac* of A and B are completely negligible compared to the vast amount of accepting sites of the solvent, and may be ignored in the equilibrium equations. The remaining possible hydrogen-bonded complexes involving the sites of the trivalent duplex are :

1. the trivalent duplex of interest, $A\cdot\overset{3}{\cdot}\cdot B$;
2. possible trivalent self-duplexes $A\cdot\overset{3}{\cdot}\cdot A$ or $B\cdot\overset{3}{\cdot}\cdot B$;
3. possible divalent $A\cdot\overset{2}{\cdot}\cdot B$, $A\cdot\overset{2}{\cdot}\cdot A$ and $B\cdot\overset{2}{\cdot}\cdot B$ duplexes ;
4. a collection of simple complexes involving the accepting sites Ac1–3 and generic donating sites Dn*, such as $A\cdot\overset{1}{\cdot}\cdot Dn^*$, $B(Ac2)\cdot\overset{1}{\cdot}\cdot Dn^*$, $B(Ac3)\cdot\overset{1}{\cdot}\cdot Dn^*$ and $B\cdot\overset{2}{\cdot}\cdot Dn^*_2$.

Given the impossibility to obtain separate values for all equilibrium constants associated to all the possible complexes of the problem, we will combine equilibrium constants in the following way, defining :

1. $k_{AB} = K_{AB}/(VC^\circ)$, in which K_{AB} is the sum of all equilibrium constants for the formation of $A\cdots B$ duplexes, be them trivalent or divalent ;
2. $k_{AA} = K_{AA}/(VC^\circ)$, in which K_{AA} is the sum of all equilibrium constants for the formation of $A\cdots A$ duplexes, be them trivalent or divalent ;

3. $k_{BB} = K_{BB}/(VC^\circ)$, in which K_{BB} is the sum of all equilibrium constants for the formation of $\mathrm{B}\cdots\mathrm{B}$ duplexes, be them trivalent or divalent ;
4. $k_A^* = K_A^*/(VC^\circ)$, in which K_A^* is the equilibrium constant for the formation of a simple $\mathrm{A}\cdot\overset{1}{\cdot}\cdot\mathrm{Dn}^*$ complex ;
5. $k_B^* = K_B^*/(VC^\circ)$, in which K_B^* is the equilibrium constant for the formation of a simple $\mathrm{B}\cdot\overset{1}{\cdot}\cdot\mathrm{Dn}^*$ complex ;
6. $k_{B_2}^* = K_{B_2}^*/(VC^\circ)^2$, in which $K_{B_2}^*$ is the equilibrium constant for the formation of a doubly bound $\mathrm{B}\cdot\overset{2}{\cdot}\cdot\mathrm{Dn}_2^*$ complex. For convenience, we define $r_2 = k_{B_2}^*/k_B^*$.

In these equations, V is the volume of solution, and $C^\circ = 1$ mol/L is the standard concentration.

Suppose now that we introduce $n_{0\mathrm{A}}$ and $n_{0\mathrm{B}}$ moles of A and B in the solvent. The introduced number of moles of generic acceptor and donor sites are thus $n_{0\mathrm{Ac}^*} = a_\mathrm{A} n_{0\mathrm{A}} + a_\mathrm{B} n_{0\mathrm{B}}$ and $n_0^* \triangleq n_{0\mathrm{Dn}^*} = d_\mathrm{A} n_{0\mathrm{A}} + d_\mathrm{B} n_{0\mathrm{B}}$, respectively. The mass balance equations are :

$$\frac{n_{0\mathrm{A}}}{n_\mathrm{A}} = 1 + k_{AB} n_\mathrm{B} + 2k_{AA} n_\mathrm{A} + k_A^* n^* \tag{1}$$

$$\frac{n_{0\mathrm{B}}}{n_\mathrm{B}} = 1 + k_{AB} n_\mathrm{A} + 2k_{BB} n_\mathrm{B} + k_B^*\big(1 + r_2 n^*\big) n^* \tag{2}$$

$$\frac{n_0^*}{n^*} = 1 + k_A^* n_\mathrm{A} + k_B^*\big(1 + 2r_2 n^*\big) n_\mathrm{B} \tag{3}$$

in which n_A, n_B and n^* are the number of moles of non-complexed A, B and generic donor sites, respectively.

For a faster numerical solution of eqs.(1)–(3), we first solve them when the number of generating donating sites is vanishingly small ($n_0^* \approx n^* \to 0$), using : [1]

$$2k_{AA} n_\mathrm{A}^2 + (1 + k_{AB} n_\mathrm{B}) n_\mathrm{A} - n_{0\mathrm{A}} = 0 \tag{4}$$

$$\text{with } n_\mathrm{B} = \frac{1 + k_{AB} n_\mathrm{A}}{4k_{BB}}\Big(\sqrt{1 + \frac{8k_{BB} n_{0\mathrm{B}}}{(1 + k_{AB} n_\mathrm{A})^2}} - 1\Big) \tag{5}$$

or, alternatively,

$$2k_{BB} n_\mathrm{B}^2 + (1 + k_{AB} n_\mathrm{A}) n_\mathrm{B} - n_{0\mathrm{B}} = 0 \tag{6}$$

$$\text{with } n_\mathrm{A} = \frac{1 + k_{AB} n_\mathrm{B}}{4k_{AA}}\Big(\sqrt{1 + \frac{8k_{AA} n_{0\mathrm{A}}}{(1 + k_{AB} n_\mathrm{B})^2}} - 1\Big) \tag{7}$$

Once this approximative solution is found, the complete set of eqs.(1)–(2) is solved numerically, starting from the approximate solution obtained for $n_0^* \approx n^* \to 0$, and using eq.(3) to express n^* as a function of n_A and n_B :

$$n^* = \frac{1 + k_A^* n_\mathrm{A} + k_B^* n_\mathrm{B}}{4r_2 k_B^* n_\mathrm{B}}\Big(-1 + \sqrt{1 + \frac{8r_2 k_B^* n_\mathrm{B} n_0^*}{(1 + k_A^* n_\mathrm{A} + k_B^* n_\mathrm{B})^2}}\Big) \tag{8}$$

with the following limit when n_B tends to zero :

$$\lim_{n_\mathrm{B} \to 0} n^* = \frac{n_0^*}{1 + k_A^* n_\mathrm{A}} \tag{9}$$

1. When $n_{0\mathrm{A}} = 0$, $n_\mathrm{A} = 0$ and $n_\mathrm{B} = \frac{\sqrt{1+8k_{BB}n_{0\mathrm{B}}}-1}{4k_{BB}}$. Likewise, when $n_{0\mathrm{B}} = 0$, $n_\mathrm{B} = 0$ and $n_\mathrm{A} = \frac{\sqrt{1+8k_{AA}n_{0\mathrm{A}}}-1}{4k_{AA}}$.

S4.1.2 NMR displacements of protons 1–3

When the rate of exchange between the complexes is much faster than the NMR characteristic time, the displacement δ of a proton involved in hydrogen bonding is given by the molar average of the displacements of this proton in the different coexisting species. The NMR displacements can then be expressed from the knowledge of the composition of the solution (remember that protons 1 and 3 belong to A, while proton 2 belongs to B) :

$$\delta_1 = \delta_{10} + k_{AB}\frac{n_\mathrm{A} n_\mathrm{B}}{n_{0\mathrm{A}}}\alpha_{1AB}\overline{\Delta_{1AB}} + k_{AA}\frac{n_\mathrm{A}^2}{n_{0\mathrm{A}}}\alpha_{1AA}\overline{\Delta_{1AA}} \tag{10}$$

$$\delta_2 = \delta_{20} + k_{AB}\frac{n_\mathrm{A} n_\mathrm{B}}{n_{0\mathrm{B}}}\alpha_{2AB}\overline{\Delta_{2AB}} + k_{BB}\frac{n_\mathrm{B}^2}{n_{0\mathrm{B}}}\alpha_{2BB}\overline{\Delta_{2BB}} \tag{11}$$

$$\delta_3 = \delta_{30} + k_{AB}\frac{n_\mathrm{A} n_\mathrm{B}}{n_{0\mathrm{A}}}\alpha_{3AB}\overline{\Delta_{3AB}} + k_{AA}\frac{n_\mathrm{A}^2}{n_{0\mathrm{A}}}\alpha_{3AA}\overline{\Delta_{3AA}} \tag{12}$$

In these equations, $\overline{\Delta_{iXY}} \triangleq \overline{\delta_{iXY}} - \delta_{i0}$ in which δ_{i0} is the NMR displacement of proton i in the solvent, and $\overline{\delta_{iXY}}$ is an average NMR displacement of proton i in the ensemble of XY duplexes, whose exact expression is complex and not especially useful. The α_{iXY} coefficients appearing in equations (10)–(12) refer to the ratio between the number of times proton i is involved in duplexes of type XY, and the number of such duplexes. These coefficients are introduced to have the $\overline{\Delta_{iXY}}$ shifts normalized per proton, but are not *per se* required. Examples will provided later on.

The previous equations are actually a simplification of a large ensemble of equations. Actually, the average chemical shifts $\overline{\Delta_{iXY}}$ involve a family of duplexes, whose composition might vary slightly when in competition with other species resulting in variations of $\overline{\Delta_{iXY}}$. However, without this simplification, the number of unknowns would be too large to extract reliable numbers from experimental data.

S4.2 Complexation of monomer units D & T

NMR displacements were obtained for a range of concentrations in deuterated acetonitrile (CD_3CN) for temperatures ranging from room temperature to 333 K, in deuterated chloroform/deuterated dimethylsulfoxide (CD_3CN:DMSO-d_6 95:5 and 5:1 v :v) at room temperature, and in $CDCl_3$:DMSO-d_6 5:1 v:v at room temperature, 300 K and 333 K (Figure S4.2). Two types of experiments were performed : dilution experiments of T or D alone, to check the self-complexation of these molecules ; and titration experiments, in which the concentration of one molecule was kept constant while the other's was varied.

The possible bivalent and trivalent duplexes are shown in Figure S4.3. In the formalism of the previous section, A=D and B=T, and $a_\mathrm{D} = 5$, $a_\mathrm{T} = 4$, $d_\mathrm{D} = d_\mathrm{T} = 2$ as mentioned before. Proton 1 is involved twice in hydrogen bonds in the possible two D$\cdots$T duplexes ; hence, $\alpha_{1AB} = 2/2$. Likewise, it is involved four times in the four possible D$\cdots$D duplexes (once for c, once for d, and twice for f) ; hence, $\alpha_{1AB} = 4/4$. Similarly, it can easily be found that $\alpha_{2AB} = 2/2$, $\alpha_{2BB} = 8/4$, $\alpha_{3AB} = 2/2$, and $\alpha_{3AA} = 4/4$.

The number of unknowns being large, simplifications have to be introduced to ensure statistical significance. The symmetry of the H-binding part of monomer D results in very close NMR shifts for protons 1 and 3, with an identical variation with concentration or temperature. These displacements were therefore averaged and the fits performed on the average only. Additionally, K_D^* and K_T^* were invariably found to converge towards very small values, indicating that the generic donor sites of the molecules can be ignored in the equilibrium. Therefore, we set $K_D^* = K_T^* = 10^{-20}$ and $r_2 = 1$, with no dependence on temperature.

Furthermore, an inspection of the pattern of hydrogen bonds in the trivalent D$\cdots$T and divalent D$\cdots$D and T$\cdots$T duplexes (Figure) indicates that $\overline{\Delta_{1DD}} \simeq \overline{\Delta_{2DT}}$ and that $\overline{\Delta_{2TT}} \simeq \overline{\Delta_{1DT}}$. Therefore, these equalities were enforced during the fits.

Finally, a linear dependence of the δ_{i0}'s was assumed with temperature, with no temperature dependence of $\overline{\Delta_{iXY}}$'s ; the temperature dependence of the equilibrium constants was taken into account using :

$$K(T) = K(T_0)\exp\Big(-\frac{\Delta H^\circ}{R}(\frac{1}{T} - \frac{1}{T_0})\Big) \tag{13}$$

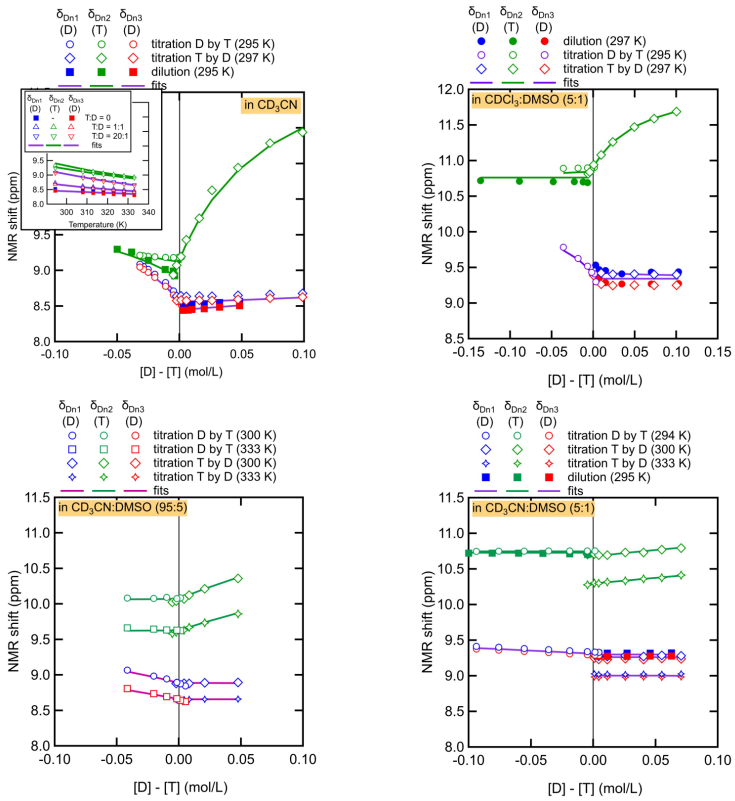


FIGURE S4.2 – NMR shifts of protons 1 (blue symbols), 2 (green symbols) and 3 (red symbols) acquired at different temperatures for different concentrations of D and T in different solvents, including simple dilution experiments of D and T alone, and titration experiments. The shifts are plotted versus the difference of concentration of D and T in mol/L. The continuous lines are fits to the data using the model developed above, with fit parameters collected in Table S4.1. Because protons 1 and 3 have similar NMR shifts, the fits were performed on the average shift of these two protons (purple lines).The inset in the top left panel is the temperature variation of the NMR shifts in deuterated acetonitrile, together with the fits.

in which $T_0 = 295$ K is an arbitrarily-selected reference temperature and ΔH° is the standard enthalpy of binding.

For experiments performed in the more hydrogen-binding CD_3CN:DMSO solvents, we also observed that K_{DD} and K_{TT} converged towards low values, presumably because divalent duplexes were essentially destroyed in this solvent ; they were therefore neglected as well. We also requested that $\delta_{i0} + \overline{\Delta_{iDT}}$ in these mixed solvents be equal at 295K to the value found in pure acetonitrile at this temperature. This corresponds to neglecting solvent effects on the NMR shifts of the D$\cdots$T duplexes, which is reasonable since the protons in duplexes are less prone to interact with solvent molecules.

As for experiments performed in $CDCl_3$:DMSO, for which experiments were only performed close to room temperature, all thermal dependences were ignored.

The fits are displayed in Figure S4.2, and represent properly the experimental data, testifying for the plausibility of the model. The resulting fit parameters are collected in Table S4.1. The equilibrium constant K_{DT} for the formation of the trivalent duplex remains small in the selected solvents, with a decrease of a factor of *ca.* 5 and 16 observed when adding 5 and 16% of DMSO to acetonitrile, respectively. The equilibrium

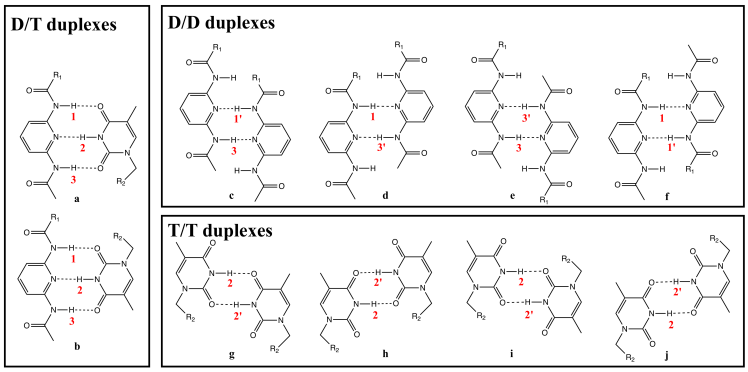


FIGURE S4.3 – Possible hydrogen-bonded duplexes of monomers D and T.

constant in chloroform with 16% added DMSO is *ca.* 20 times higher than in acetonitrile added with the same amount of DMSO, which results from acetonitrile being a much stronger hydrogen bond acceptor than chloroform.

	in CD_3CN		in CD_3CN: DMSO-d_6 95:5		in CD_3CN: DMSO-d_6 5:1		in $CDCl_3$: DMSO-d_6 5:1	
Parameter	at 295 K	Error	at 295 K	Error	at 295 K	Error	at 295 K	Error
K_{DT}	16.0	1.7	3.3	0.2	0.99	0.06	20.1	1.2
K_{DD}	0.4	0.1	10^{-20}	-	10^{-20}	-	0.09	0.06
K_{TT}	3.2	0.6	10^{-20}	-	10^{-20}	-	0.09	0.05
ΔH°_{DT} (kJ)	-21	5.7	-6.1	1.8	-5.7	2.3	-	-
ΔH°_{DD} (kJ)	24	50	0	-	0	-	-	-
ΔH°_{TT} (kJ)	-21	10	0	-	0	-	-	-
ΔS°_{DT} (J/K)	-48	19	-10.8	6	-19	8	-	-
$\frac{\delta_{10}+\delta_{30}}{2}$ (ppm)	8.44	0.01	8.898	0.006	9.297	0.002	9.346	0.004
δ_{20} (ppm)	8.88	0.02	10.097	0.006	10.729	0.002	10.771	0.004
$\frac{\mathrm{d}(\delta_{10}+\delta_{30})/2}{\mathrm{d}T}$ (ppm/K)	-0.0036	0.0022	-0.0068	0.0002	-0.0079	0.0001	0	-
$\frac{\mathrm{d}\delta_{20}}{\mathrm{d}T}$ (ppm/K)	-0.0048	0.002	-0.0132	0.0002	-0.0116	0.0002	0	-
$\overline{\Delta_{1DT}}$ (ppm)	1.91	0.13	1.45	-	1.05	-	0.919	0.04
$\overline{\Delta_{2DT}}$ (ppm)	3.64	0.17	2.42	-	1.79	-	1.375	0.03

TABLE S4.1 – Fit parameters for D$\cdots$T complexation in different solvents (Fig.S4.2). The errors on the fit parameters were estimated from the diagonal terms of the covariance matrix, assuming a standard error on NMR shifts of 0.01 ppm ; '-' in the error columns indicates that the corresponding parameter was not fitted. The standard entropies of complexation were not fitted and are computed from the equilibrium constants and the standard enthalpies of complexation.

S4.3 Complexation of monomer units G & C

NMR displacements were obtained for a range of concentrations in a deuterated chloroform/deuterated dimethylsulfoxide mixture ($CDCl_3$:DMSO-d_6 5:1 v:v) at room temperature, and in CD_3CN:DMSO-d_6 5:1 v:v at room temperature, 300 K and 333 K (monomer G is not soluble in pure acetonitrile, nor with only 5% of added DMSO). Again, two types of experiments were performed : dilution experiments of C or G alone, to check the self-complexation of these molecules ; and titration experiments, in which the concentration of one molecule is kept constant while the other's is varied (Figure S4.4).

The possible bivalent and trivalent duplexes are shown in Figure S4.5. In the formalism of the previous section, A=G and B=C, and $a_G = 6$, $a_C = 4$, and $d_G = d_B = 2$ as mentioned before. Proton 1 is involved four

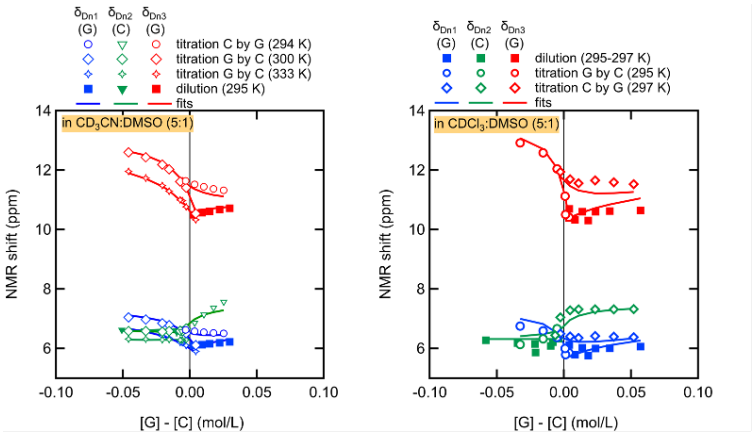


FIGURE S4.4 – NMR shifts of protons 1 (blue symbols), 2 (green symbols) and 3 (red symbols) acquired at different temperatures for different concentrations of G and C in different solvents, including simple dilution experiments of G and C alone, and titration experiments. The shifts are plotted versus the difference of concentration of G and C in mol/L. The continuous lines are fits to the data using the model developed above, with fit parameters collected in Table S4.2. In some experiments, the NMR signal of the primary amine protons 2 of C could be differentiated (triangles), in which case the fit was performed on the average shift.

times in hydrogen bonds in the possible four G···C duplexes; hence, $\alpha_{1AB} = 4/4$. Likewise, it is involved six times in the five possible G···G duplexes; hence, $\alpha_{1AA} = 6/5$. Similarly, it can easily be found that $\alpha_{2AB} = 3/4$, $\alpha_{2BB} = 4/2$, $\alpha_{3AB} = 2/4$, and $\alpha_{3AA} = 4/5$.

As before, simplifications were performed to increase the significance of the fit parameters. Here again, K_G^* and K_C^* always converged towards very small values; hence, we set $K_G^* = K_C^* = 10^{-20}$ and $r_2 = 1$, with no dependence on temperature. Furthermore, we imposed $\overline{\Delta_{2GC}} = \overline{\Delta_{1GC}}$ and $\overline{\Delta_{2CC}} = \overline{\Delta_{1GG}}$, because protons 1 and 2 have similar NMR shifts and binding patterns, and $\overline{\Delta_{3GC}} = \overline{\Delta_{3GG}}$ again because the binding pattern of proton 3 in the C···G and G···G duplexes is similar. Finally, no temperature dependence was considered for experiments performed in $CDCl_3$:DMSO, which were only performed close to room temperature, whereas only the temperature dependence of K_{GC}, δ_{10}, δ_{20} and δ_{30} had to be taken into account in CD_3CN:DMSO.

The fits are displayed in Figure S4.4, and represent properly the experimental data, admittedly less so in $CDCl_3$:DMSO. The resulting fit parameters are collected in Table S4.2. The equilibrium constants K_{GC} for the formation of the trivalent duplex are similar for both solvents, and much larger than for the D···T case, showing the greater stability of G···C complexes compared to D···T complexes. Additionally, the self-complexation of G is significant, with a K_{GG} constant only *ca.* 70 times smaller than K_{GC}. This comes as no surprise, as it is well-known that guanine has a strong tendency to form self-duplexes; this is most probably the cause for the insolubility of G in pure acetonitrile.

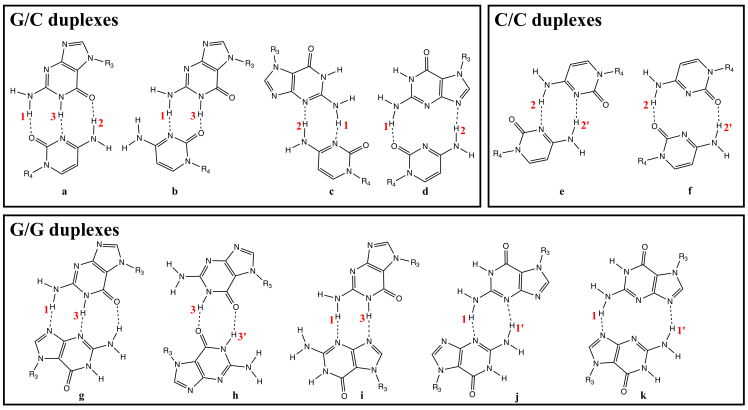


FIGURE S4.5 – Possible hydrogen-bonded duplexes of monomers G and C.

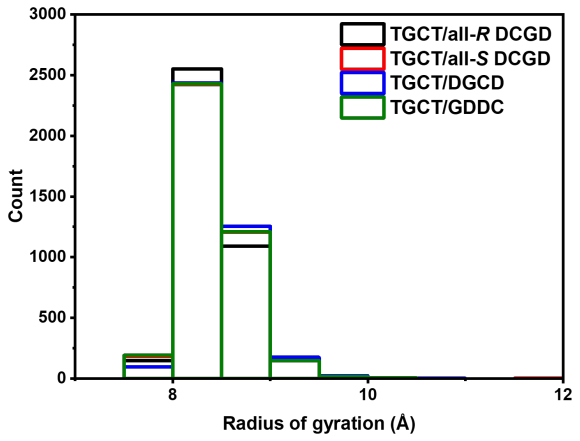
Parameter	in CD_3CN: DMSO-d_6 5:1 at 295 K	Error	in $CDCl_3$: DMSO-d_6 5:1 at 295 K	Error
K_{GC}	104	15	173	3
K_{GG}	1.4	0.2	5.1	0.1
K_{CC}	0.11	0.11	10^{-20}	-
ΔH°_{GC} (kJ)	-24	4	0	-
ΔS°_{GC} (J/K)	-43	14	-	-
δ_{10} (ppm)	6.05	0.05	5.661	0.004
δ_{20} (ppm)	6.53	0.03	6.315	0.003
δ_{30} (ppm)	10.62	0.03	10.293	0.006
$\frac{d\delta_{10}}{dT}$ (ppm/K)	-0.0061	0.0015	0	-
$\frac{d\delta_{20}}{dT}$ (ppm/K)	-0.0092	0.0011	0	-
$\frac{d\delta_{30}}{dT}$ (ppm/K)	-0.0077	0.002	0	-
$\overline{\Delta_{1GC}}$ (ppm)	1.36	0.09	1.572	0.009
$\overline{\Delta_{1GG}}$ (ppm)	5.89	1.91	3.48	0.06
$\overline{\Delta_{3GC}}$ (ppm)	5.10	0.25	6.54	0.03

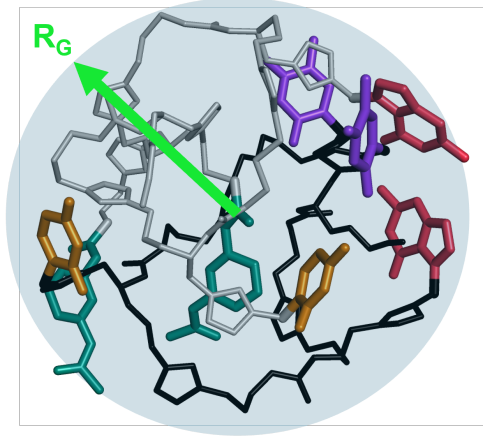
TABLE S4.2 – Fit parameters for G$\cdots$C complexation in different solvents (Fig.S4.4). The errors on the fit parameters were estimated from the diagonal terms of the covariance matrix, assuming a standard error on NMR shifts of 0.01 ppm ; '-' in the error columns indicates that the corresponding parameter was not fitted.The standard entropies of complexation were not fitted and are computed from the equilibrium constants and the standard enthalpies of complexation.

S5. All-atom molecular dynamics (MD) simulations

All-atom molecular dynamics (MD) simulations were performed using the GPU version of AMBER16 package. The five oligomeric strands (all-*R* TGCT, all-*R* DCGD, all-*S* DCGD, all-*R* DGCD and all-*R* GDDC) studied by MD were built by joining small molecular fragments together. The monomers were first constructed within the Avogadro software.[4] The partial atomic charges of each fragment were calculated using the semi-empirical AM1-BCC model[5] as implemented within the antechamber module of AMBER16, whereas other force field parameters are from the 'General AMBER Force Field (GAFF 2.1, version updated in April 2016).[6] The individual molecular fragments were then connected in the desired sequence with the LEaP module of AMBER16 to constitute the complete oligomeric chains. When building the target chain/probe chain assemblies, the latter was translated by 25 Å in the x, y and z directions to avoid intermolecular contacts in the starting structure. A geometry optimization was then performed by molecular mechanics, with a total of 10,000 steps distributed in 1,000 steps of steepest descent and 9,000 steps of conjugated gradient, to get a stable starting point for the subsequent MD simulations. These were carried out with an implicit solvent model, the Generalized Born (GB) model,[7] to ensure a sufficient conformational sampling in a reasonable computational time. A dielectric constant of 37.5 was considered, i.e., the dielectric constant of acetonitrile at room temperature. The temperature in all MD simulations was set to 300 K and controlled by the Langevin thermostat with a coupling constant of 1.0 ps and combined with a pseudo-random seed generator. Frames were collected at 1-ns intervals for a total simulation length of 1 µs, resulting in a set of 1,000 conformations per replica. Four independent replicas were launched for each heteromolecular assembly, leading to a total of 4,000 conformations by duplex. The four duplexes studied by MD are hereafter referred to as **TGCT/all-*R* DCGD**, **TGCT/all-*S* DCGD**, **TGCT/DGCD** and **TGCT/GDDC**. The resulting trajectories were visualized and the MD snapshots were created with PyMOL.[8]

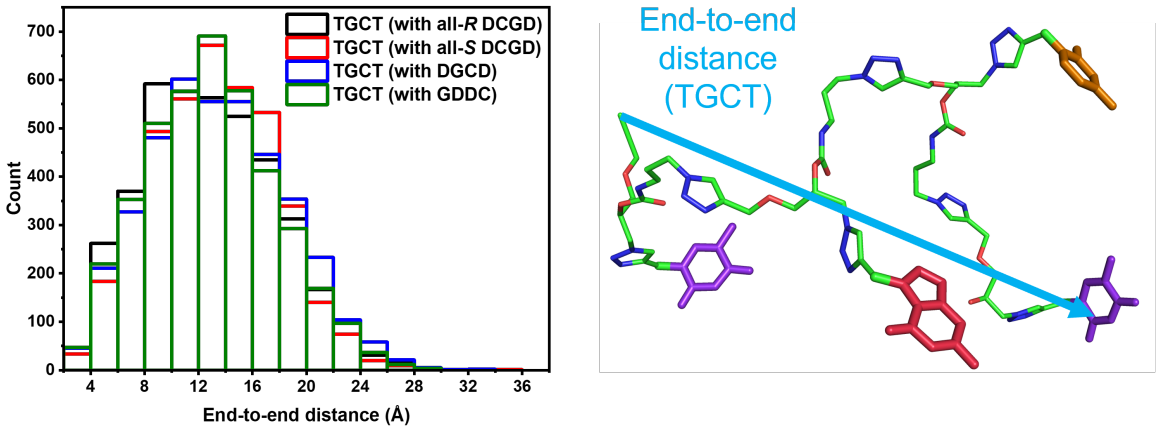
The analyses of the MD trajectories were performed using the cpptraj module available in AmberTools. The Radii of gyration (R_G) were calculated in reference to the heavy atoms, with omission of hydrogen atoms. The end-to-end distances were measured between the first carbon atom of the oligomers and a carbon atom of the last nucleobase analog. A cut-off of 3.0 Å for the donor (D)-acceptor (A) distance and 135° for the A-H-D angle were set as criteria to detect possible hydrogen-bonds. An H-bond is described as "complementary" if it is intermolecular and occurs between a T-D or G-C pair of nucleobase analogs. An aromatic interaction between two cycles was detected if two geometric criteria were satisfied: the distance between the center-of-mass (COM) is less than 5 Å and the angle between the planes of the cycles is less than 45° or more than 135°.





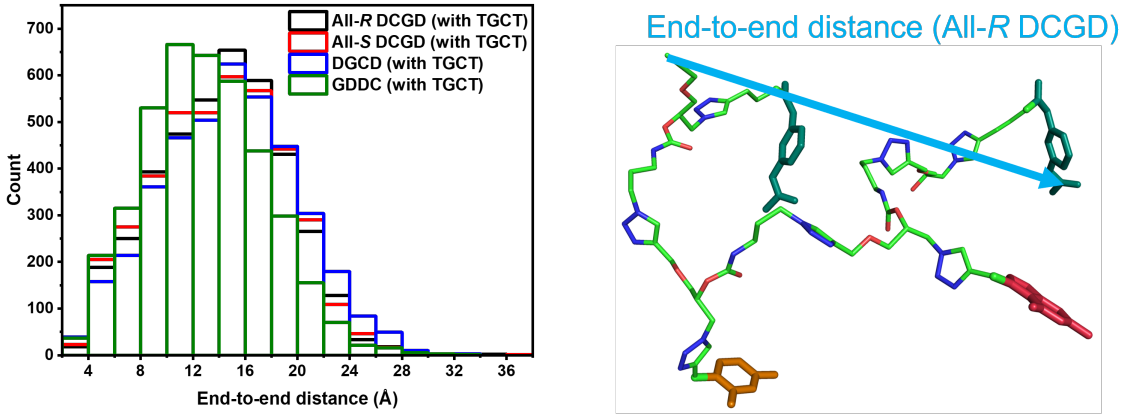
	TGCT/all-*R* DCGD	TGCT/all-*S* DCGD	TGCT/DGCD	TGCT/GDDC
Average ± Std dev (Å)	8.5 ± 1.1	8.5 ± 0.8	8.5 ± 0.9	8.4 ± 0.6
Max (Å)	27.4	30.5	27.4	24
Min (Å)	7.8	7.7	7.8	7.7

Figure S5.1 - Distribution profiles of the radius of gyration (R_G, illustrated on a snapshot on the right) for a total of 4,000 conformations for the pairs of oligomers **TGCT/all-*R* DCGD**, **TGCT/all-*S* DCGD**, **TGCT/DGCD** and **TGCT/GDDC**. Bin sizes were set to 0.5 Å for all distributions. Averages, standard deviations, minimum and maximum values of the radii of gyration for the different sequences are gathered in the table.



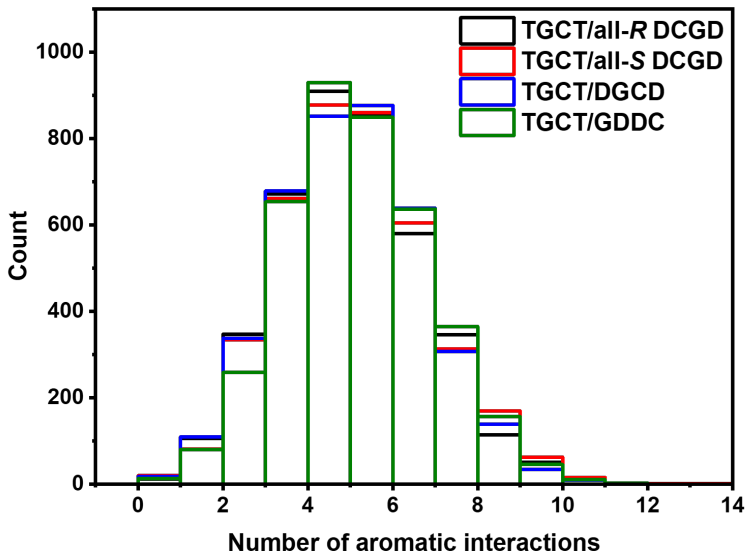
	TGCT (with all-*R* DCGD)	TGCT (with all-*S* DCGD)	TGCT (with DGCD)	TGCT (with GDDC)
Average ± Std dev (Å)	12.8 ± 4.9	13.2 ± 4.5	13.5 ± 5	13 ± 4.7
Max (Å)	34.2	34.2	32	29.7
Min (Å)	3.3	3.2	3.2	3.2

Figure S5.2 - Distribution profiles of the end-to-end distance for a total number of 4,000 conformations for the target chain TGCT, when assembled with the four different probe chains. Bin sizes were set to 2 Å for all distributions. Averages, standard deviations, minimum and maximum values of the end-to-end distances for the different sequences are gathered in the table. The carbon atoms of the target chain, TGCT, between which the distance is measured are shown by the cyan arrow, on the snapshot on the right of the figure.



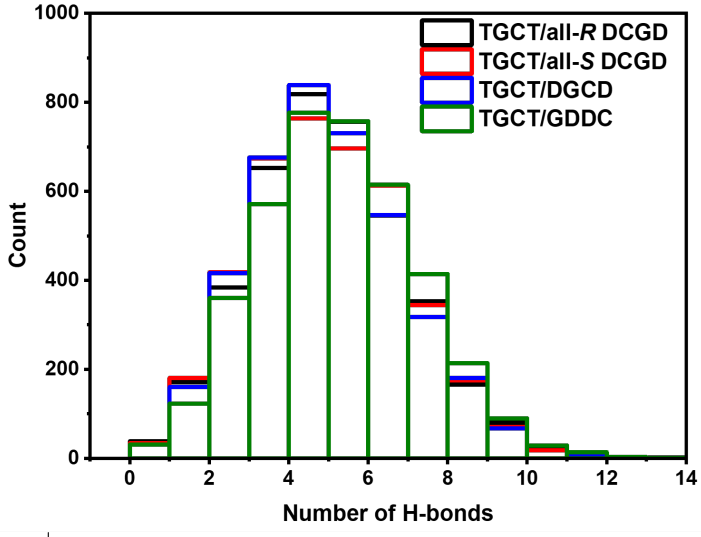
	All-*R* DCGD (with TGCT)	All-*S* DCGD (with TGCT)	DGCD (with TGCT)	GDDC (with TGCT)
Average ± Std dev (Å)	14.2 ± 4.8	14.1 ± 4.9	14.8 ± 5.2	13 ± 4.6
Max (Å)	35.8	37.4	33	39
Min (Å)	3.2	3.4	3.4	3.1

Figure S5.3 - Distribution profiles of the end-to-end distance for a total number of 4,000 conformations for the four different probe chains, when assembled with the target chain, TGCT. Bin sizes were set to 2 Å for all distributions. Averages, standard deviations, minimum and maximum values of the end-to-end distances for the different sequences are gathered in the table. The carbon atoms between which the distance is measured are shown for the All-*R* DCGD probe chain by the cyan arrow, on the snapshot on the right of the figure.



	TGCT/all-*R* DCGD	TGCT/all-*S* DCGD	TGCT/DGCD	TGCT/GDDC
Average ± Std dev	4.5 ± 1.7	4.6 ± 1.8	4.5 ± 1.7	4.7 ± 1.7
Max	11	13	11	11
Min	0	0	0	0

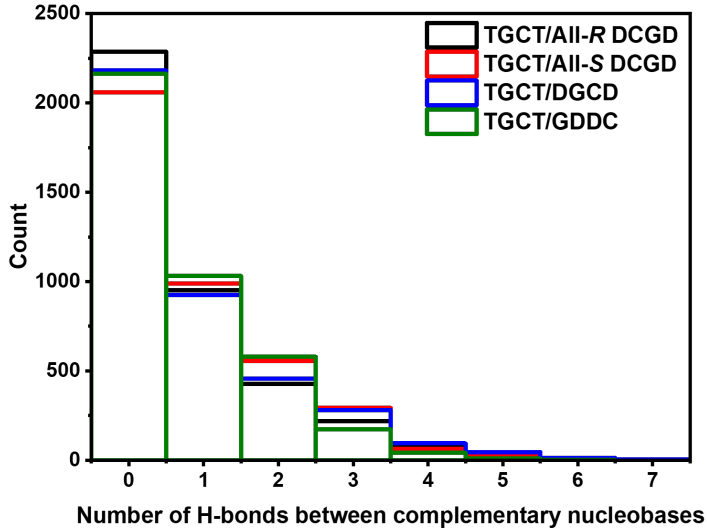
Figure S5.4 - Distribution profiles of the number of aromatic interactions for a total number of 4,000 conformations for the pairs of oligomers **TGCT/all-*R* DCGD**, **TGCT/all-*S* DCGD**, **TGCT/DGCD** and **TGCT/GDDC**. Bin sizes were set to one interaction for all distributions. Averages, standard deviations, minimum and maximum values of the number of aromatic interactions for the different sequences are gathered in the table.



	TGCT/all-*R* DCGD	TGCT/all-*S* DCGD	TGCT/DGCD	TGCT/GDDC
Average ± Std dev	4.5 ± 2	4.5 ± 2	4.5 ± 1.9	4.8 ± 2
Max	13	12	11	13
Min	0	0	0	0

Figure S5.5 - Distribution profiles of the number of hydrogen bonds for a total number of 4,000 conformations for the pairs of oligomers **TGCT/all-*R* DCGD**, **TGCT/all-*S* DCGD**, **TGCT/DGCD** and **TGCT/GDDC**. Bin sizes were set to one

interaction for all distributions. Averages, standard deviations, minimum and maximum values of the number of H-bonds for the different sequences are gathered in the table.



	TGCT/all-*R* DCGD	TGCT/all-*S* DCGD	TGCT/DGCD	TGCT/GDDC
Average ± Std dev	0.7 ± 1.1	0.9 ± 1.1	0.8 ± 1.2	0.7 ± 1
Max	13	12	11	13
Min	0	0	0	0

Figure S5.6 - Distribution profiles of the number of hydrogen bonds between complementary nucleobase pairs (T-D and G-C) for a total number of 4,000 conformations for the pairs of oligomers **TGCT/all-*R* DCGD**, **TGCT/all-*S* DCGD**, **TGCT/DGCD** and **TGCT/GDDC**. Bin sizes were set to one interaction for all distributions. Averages, standard deviations, minimum and maximum values of the number of H-bonds for the different sequences are gathered in the table.

	TGCT/all-*R* DCGD	TGCT/all-*S* DCGD	TGCT/DGCD	TGCT/GDDC
R_G (Å)	8.5 ± 1.1	8.5 ± 0.8	8.5 ± 0.9	8.4 ± 0.6
End-to-end distance: target chain (Å)	12.8 ± 4.9	13.2 ± 4.5	13.5 ± 5	13 ± 4.7
End-to-end distance: probe chain (Å)	14.2 ± 4.8	14.1 ± 4.9	14.8 ± 5.2	13 ± 4.6
Number of aromatic interactions per conformation	4.5 ± 1.7	4.6 ± 1.8	4.5 ± 1.7	4.7 ± 1.7
Number of H-bonds per conformation	4.5 ± 2	4.5 ± 2	4.5 ± 1.9	4.8 ± 2
Number of complementary H-bonds per conformation	0.7 ± 1.1	0.9 ± 1.1	0.8 ± 1.2	0.7 ± 1

Table S5.1 – Summary table of the MD data presented in the figures above.

S6. X-ray reflectometry results

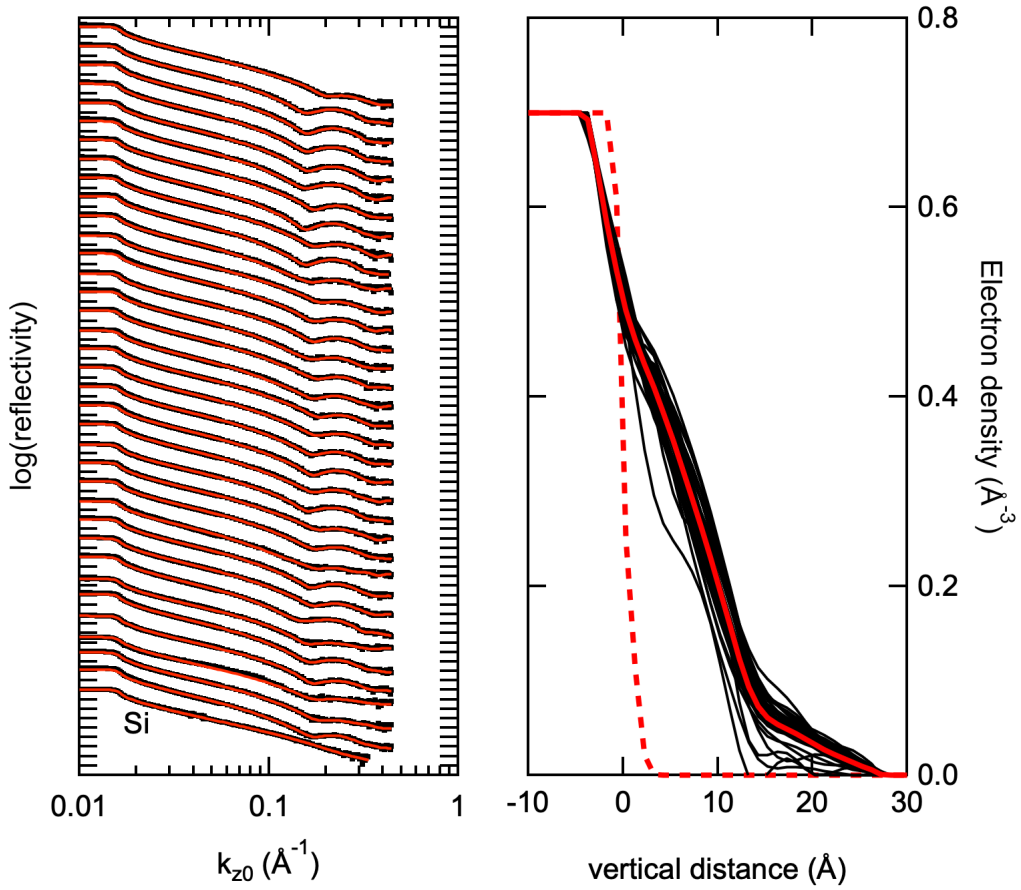


Figure S6.1 – **Left.** X-ray reflectograms (black dots) and fits (red lines) of thirty-fiveSi wafers ((100), native oxide-covered) gas-phase silanized by AzUTMS. The bottom curve is the XRR of a bare Si wafer. The curves are displaced vertically for clarity. **Right.** Electron density profiles corresponding to the fits of the XRR (black lines except for the bare Si wafer which is dislayed as the dashed red line). The thick red line is the average electron density profile of these thirty-five AzUTMS-silanized wafers.

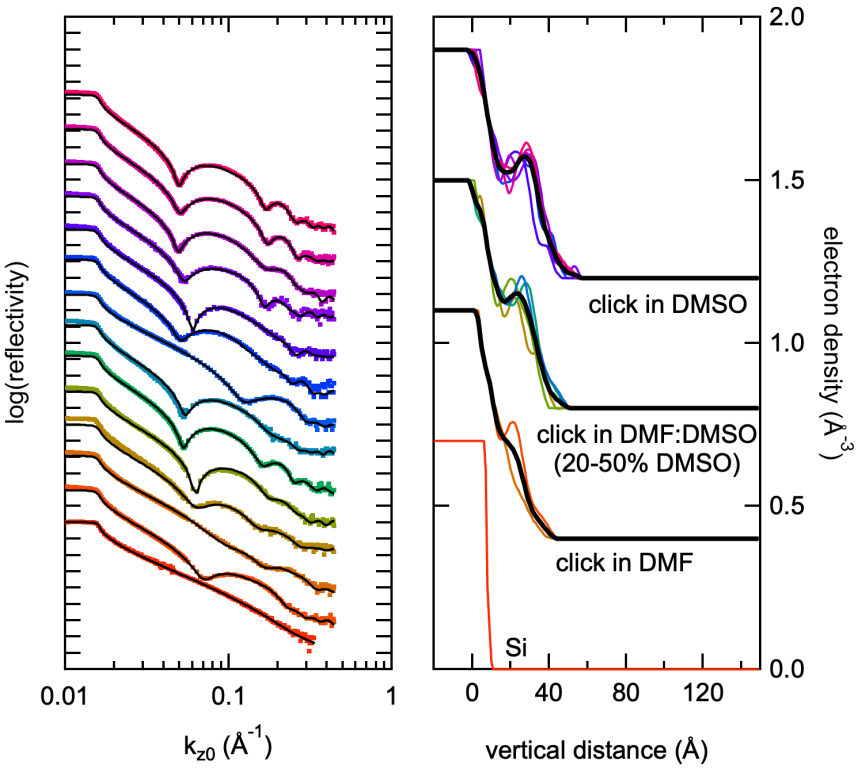


Figure S6.2 – **Left.** X-ray reflectograms (colored dots) and fits (black lines) all-R TGCT layers clicked in different solvents onto AzUTMS-silanized Si wafers. The bottom curve is the XRR of a bare Si wafer. The curves are displaced vertically for clarity. **Right.** Electron density profiles corresponding to the fits of the XRR (colored lines; same color code as in the left panel). The thick black lines are the average electron density profiles of the modified wafers.

S7. XPS of grafted layers

S7.1. AzUTMS layers

The AzUTMS layer grafted on a silicon wafer was characterized by X-ray photoelectron spectroscopy (XPS). The XPS survey spectrum is shown in Figure S7.1, showing the presence of Si, O, C, N and adventitious contaminants such as Na or Zn.

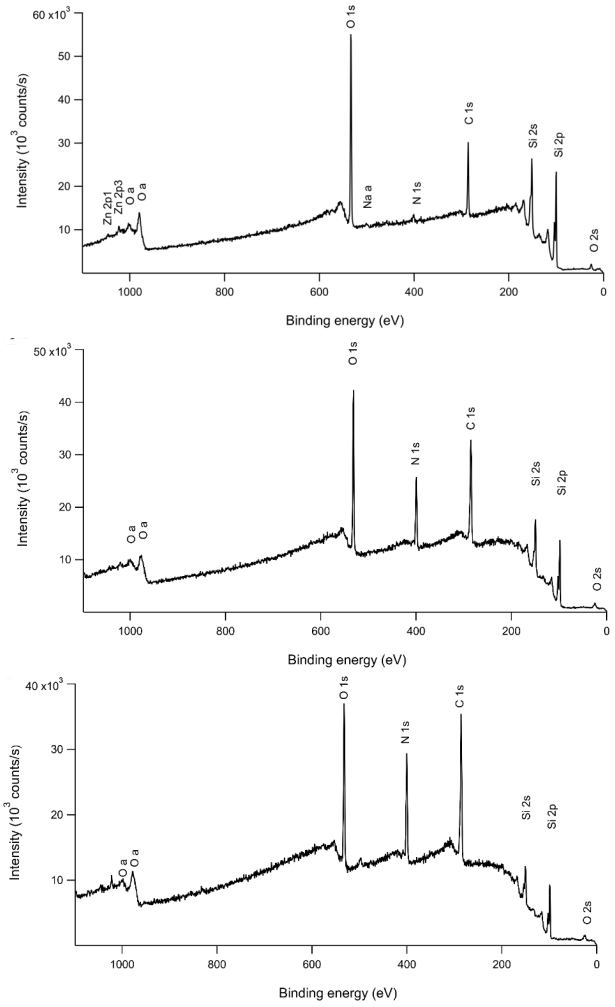


Figure S7.1 - XPS survey scans of (**top**) a monolayer assembled from AzUTMS on a Si (100) wafer with its native oxide layer; (**middle**) the same after clicking the TGCT oligomer in DMSO; (**bottom**) the same after recognition with all-*R* DCGD.

High resolution XPS spectra in the N 1s and C 1s regions are shown in Fig.S7.2 (top). In the N 1s region (Fig.S7.2, top left), the AzUTMS layer shows two azide peaks in a 2:1 ratio, the larger one arising from N and N^- atoms and the smaller one from N^+ atoms. An additional peak (400.3 eV) corresponding to decomposed azide groups upon X-ray exposure is also observed.[9] In the C 1s region (Fig.S7.2 right), the high-resolution spectrum consists of a broad signal at a mean binding energy of 285.0 eV. The large dispersion value (1.4 eV FWHM) of the fitted function is consistent with a peak essentially resulting from contributions of C-C, C-H and C-N bonds.

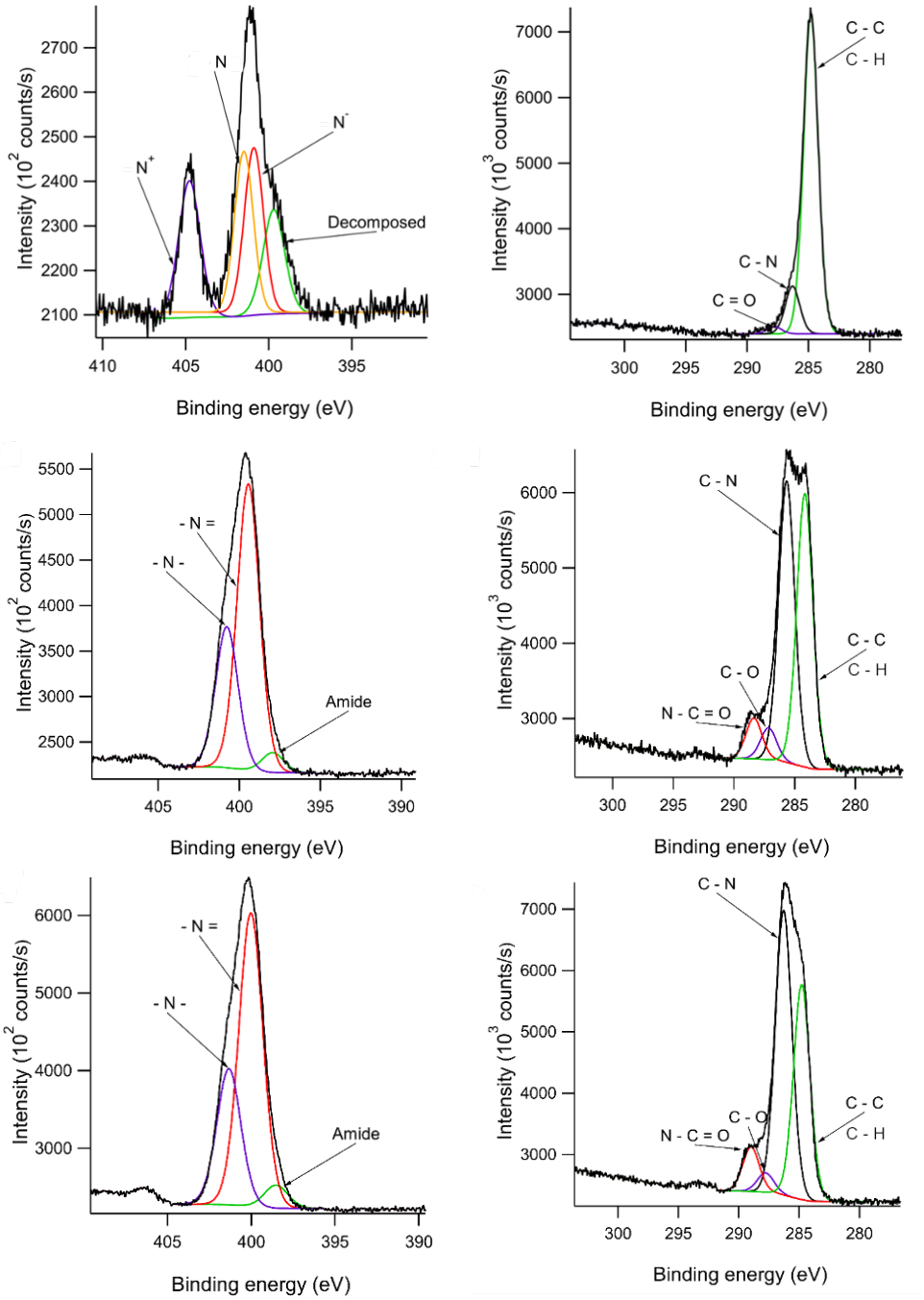


Figure S7.2 – High resolution XPS spectra in the N 1s (**left**) and C 1s (**right**) regions of (**top**) a monolayer assembled from 11-azidoundecyltrimethoxysilane on a Si (100) wafer with its native oxide layer; (**middle**) the same after clicking TGCT in DMSO; (**bottom**) the same after recognition with all-*R* DCGD. Black curves are experimental spectra, colored curves show the components as obtained from spectral decomposition by fitting.

S7.2. TGCT layers

The sample obtained after clicking TGCT target chains on the AzUTMS layer in DMSO was also measured by XPS. The survey scan and high-resolution scans are displayed in the middle panels of Fig.S7.1 and Fig.S7.2, respectively. The decrease of the Si 2p signals together with the increase of N 1s and C 1s signals correspond to the burial of the Si wafer below a layer of TGCT. On the high-resolution scan of the N 1s region, the azide signal vanishes (signal at 405.4 eV), and new peaks corresponding to the nucleobases and triazole rings appear, together with a smaller signal from the urethane and amide groups. Simultaneously, the C 1s region of the XPS spectra exhibits a major component at 287.5 - 287.6 eV which is assigned to C–N. Additionally, N-C=O groups appear at 290.2-290.3 eV. These observations fully confirm the grafting of the TGCT layer on the surface.

S7.3. TGCT layer after recognition by all-R DCGD

The sample grafted by TGCT then exposed to all-*R* DCGD probe chains was also measured by XPS. The comparison of survey scans of both AzUTMS (Fig.S7.1, top panel) and TGCT samples (Fig.S7.1, middle panel) shows a significant decrease of Si 2p signals and an increase of N 1s and C 1s signals, confirming the adsorption of the probe oligomer on the surface. The N 1s spectrum of this sample (Fig.S7.2, bottom left) is similar to the one of the TGCT-grafted layer whereas the C 1s spectrum shows an increase of the C-N signal and a decrease of the C-C/H signal, resulting from a decreased contribution of the AzUTMS layer.

S8. In-situ ellipsometry

S8.1. Determination of the relationship between the change of Δ and the amount of adsorbed probe oligomer.
The amount of oligomer irreversibly-adsorbed over a TGCT-modified wafer was obtained from the variation of the ellipsometric angle Δ before and after adsorption/rinsing of the oligomer over the TGCT layer. Simulations of the (Ψ,Δ) ellipsometric trajectories parametrized by wavelength λ (from 400 to 700 nm) are shown in Fig.S8.1 (left). To draw these trajectories, we considered a model consisting of a Si wafer with 1.3 nm of native oxide, modified by a layer of AzUTMS of 1.05 nm thickness (the average experimental value) covered by a solvated layer of thickness d_{layer} made of a mixture of target/probe oligomers in acetonitrile:DMSO 5:1 at 25°C; this solvated layer contains a volume fraction f_{oligo} of oligomers, the remaining part being the solvent. The indices of refraction used for each layer are detailed in section S8.2. The simulations were performed with a home-made program written in the Igor Pro (Wavemetrics) macro-programming language based on standard ellipsometry equations.[10] The program is available upon request.

In Fig.S8.1, the black dotted line shows the trajectory for an AzUTMS-modified wafer measured in the ACN:DMSO 5:1 solvent. When adding a layer of precision oligomers containing no solvent (red curves), the trajectories shift towards lower values of Δ without significant changes of Ψ, the trajectories staying parallel within experimental precision. If layers containing 50% oligomer and 50% solvent are considered (blue curves), the trajectories behave similarly but the amplitude of the shift of Δ is approximately decreased by a factor of two. A master curve is drawn in Fig.S8.1 (right), in which the average value of Δ computed over the 400-700 nm wavelength range is plotted versus the product d_{layer} x f_{oligo} for different values of f_{oligo}. All data points fall over a single line of slope -1.2615 °/nm, in the range of possible values for d_{layer} x f_{oligo}.

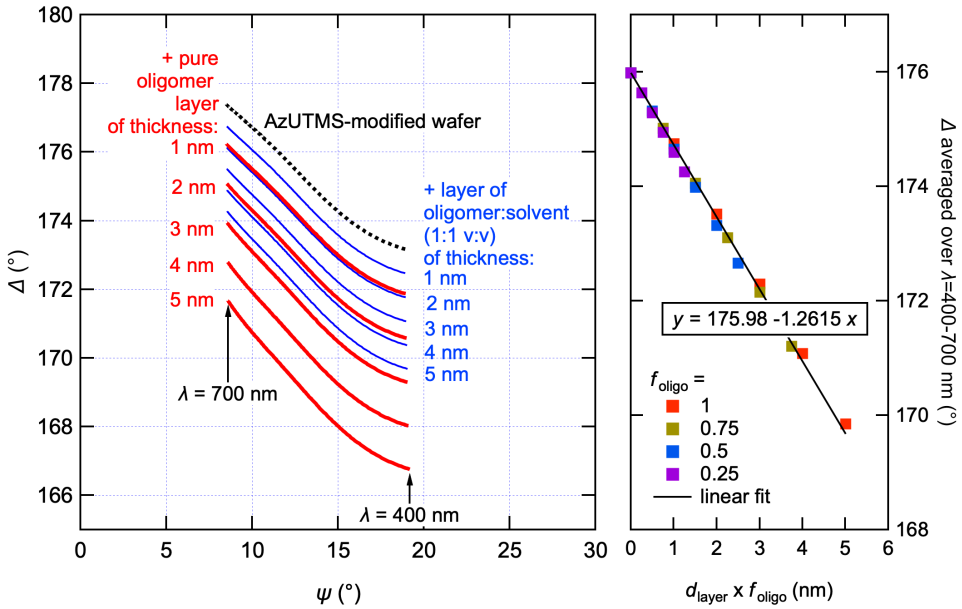


Figure S8.1. (**left**) Simulated ellipsometric trajectories drawn with wavelength as parameter for AzUTMS-modified wafers covered by layers of oligomer:solvent of varying thickness and solvent content, as indicated, measured in ACN:DMSO 5:1 v:v solvent. (**right**) Master line relating the average value of the ellipsometric angle Δ to the product of thickness and oligomer volume fraction in the layer.

The product d_{layer} x f_{oligo} is the volume of oligomer resting over the silanized wafer, per surface area of wafer. Therefore, suppose an average value Δ_A is measured for a sample grafted by a TGCT target layer at equilibrium with the solvent, and an average value Δ_B is measured after adsorption of a probe chain and rinsing; then, the difference (Δ_B - Δ_A)/(-1.2615) [nm] is the adsorbed volume of probe chain per unit area of wafer. The advantage of subtracting these two numbers is that systematic errors due to the presence of cell windows cancel out.

S8.2. Indices of refraction.
The simulations of Fig.S8.1 require the complex indices of refraction of the different materials in the 400-700 nm wavelength range.

- For Si and silicon oxide, the values tabulated in Palik's handbook were used.[11]
- For the AzUTMS layer, a constant value of 1.458 was taken for the real part of the index of refraction, with no absorption, as derived from the ellipsometric analysis of the AzUTMS-silanized wafers discussed in the companion article. Note that the conclusions of section S8.1 are insensitive to the exact values taken for this index of refraction owing to the subtraction of two values of Δ.
- For the ACN:DMSO 5:1 mixture, the refractive index was obtained by the Newton approximation (eq.1)[12] and the empirical dependences on wavelength of DMSO and ACN refractive indices described elsewhere (eqs.2 and 3):[13]

$$n_{solvent}^2 = \frac{n_{DMSO}^2 + 5n_{ACN}^2}{6} \quad (1)$$

$$n_{ACN}(\lambda) = 1.33212 + \frac{3525.78231}{\lambda^2} - \frac{32631699.6}{\lambda^4} + \frac{2.267 \cdot 10^{12}}{\lambda^6} \quad (2)$$

$$n_{DMSO}(\lambda) = \sqrt{1 + \frac{0.04419\,\lambda^2}{\lambda^2 - 46390.67309} + \frac{1.09101\,\lambda^2}{\lambda^2 - 12215.43949}} \quad (3),$$

  with the wavelength λ expressed in nm.
- For the oligomers, we assumed that all of them have very similar refractive indices owing to their close chemical compositions. Therefore, we deposited by spin-coating a TGCT solution in methanol (10 mg/mL, spin rate 500 rpm) on a silicon wafer, and measured the thickness of the layer by XRR. The sample was then measured by spectroscopic ellipsometry. The XRR-determined value of thickness was then used to extract the index of refraction of the TGCT layer from the ellipsometric data, using a Cauchy transparent model to express the dependence of index of refraction with wavelength. The final result at 25°C is:

$$n_{TGCT} = 1.49944 - \frac{1503.02}{\lambda^2} + \frac{1.32138\,10^9}{\lambda^4}$$, with λ the wavelength in vacuum, expressed in nm.

- For the solvent-swollen layer of oligomers with an oligomer volume fraction f_{oligo}, the index of refraction was computed with Bruggeman's effective medium approximation.[14]

S9. References

1 X. Pei, A. Fernandes, B. Mathy, X. Laloyaux, B. Nysten, O. Riant and A. M. Jonas, *Langmuir*, 2011, **27**, 9403–9412.

2 A. M. Jonas, Z. Hu, K. Glinel and W. T. S. Huck, *Nano Lett.*, 2008, **8**, 3819–3824.

3 Q. Qin, J. Li, D. Dellemme, M. Fosséprés, G. Barozzino-Consiglio, I. Nekkaa, A. Boborodea, A. E. Fernandes, K. Glinel, M. Surin and A. Jonas, *Chem. Sci.*, 2023, **14**, 9283–9292.

4 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminf.*, 2012, **4**, 17.

5 A. Jakalian, B. L. Bush, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2000, **21**, 132–146.

6 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

7 G. D. Hawkins, C. J. Cramer and D. G. Truhlar, *Chem. Phys. Lett.*, 1995, **246**, 122–129.

8 *Schrödinger LLC, The PyMOL Molecular Graphics System, Version 2.5.4*.

9 T. Heinrich, C. H.-H. Traulsen, E. Darlatt, S. Richter, J. Poppenberg, N. L. Traulsen, I. Linder, A. Lippitz, P. M. Dietrich, B. Dib, W. E. S. Unger and C. A. Schalley, *RSC Adv.*, 2014, **4**, 17694–17702.

10 R. Azzam and N. M. Bashara, *Ellipsometry and polarized light*, Elsevier, Amsterdam, 1987.

11 E. D. Palik, *Handbook of Optical Constants of Solids*, Academic Press, San Diego, 1985, vol. I.

12 J. C. R. Reis, I. M. S. Lampreia, Â. F. S. Santos, M. L. C. J. Moita and G. Douhéret, *ChemPhysChem*, 2010, **11**, 3722–3733.

13 I. Z. Kozma, P. Krok and E. Riedle, *J. Opt. Soc. Am. B*, 2005, **22**, 1479.

14 H. G. Tompkins, *A User's Guide to Ellipsometry*, Academic Press, San Diego, 1993.