Nanostructured assemblies from nucleotide-based amphiphiles†

Nathalie Campins,[ab] Philippe Dieudonné,[c] Mark W. Grinstaff[d] and Philippe Barthélémy*[ab]

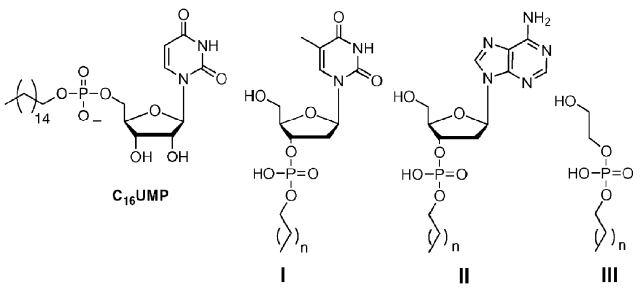
Received (in Montpellier, France) 2nd April 2007, Accepted 4th July 2007
First published as an Advance Article on the web 16th August 2007
DOI: 10.1039/b704884j

A synthetic route to 3′-*O*-alkyl nucleotides is described that provides easy access to a series of nucleotide-based amphiphiles. Physicochemical experiments demonstrate that these nucleotide amphiphiles possess different aggregating properties compared to both their non-nucleotide analogues and simple nucleotides. The data collected from TEM, SEM, FT-IR and WAXS studies indicate that the aqueous dispersion of these nucleotide amphiphiles provides nanostructured assemblies such as ribbon-like structures for the thymidine derivatives and small aggregates for adenine. These results indicate that the nature of the base has an impact on aggregate morphology. A molecular model illustrating the 3D supramolecular arrangements of a thymidine nucleoamphiphile forming ribbons is proposed.

Introduction

Over the past decade, substantial research effort has been focused on the design, synthesis, and characterization of materials structured at the nanometre scale.[1–3] Today, nanometre-scale engineering based on the chemical principles of non-covalent interactions has emerged as a technology, which promises to deliver the next generation of materials for a broad spectrum of applications as diverse as from energy conversion and storage,[4] to medicine.[5–7] However, despite the many supramolecular systems developed, there still exists a need for additional methods to construct supramolecular structures ranging in size from a few nanometres to micrometres. Small molecule and macromolecular amphiphiles offer one of the most efficient methods for producing organic nanostructured materials over a range of size scales. In the presence of water, amphiphiles are known to spontaneously assemble into a large variety of aggregated structures with specific geometries. Of the amphiphiles investigated, nucleotide-based amphiphiles are particularly interesting because the combination of molecular recognition (*e.g.*, H-bonding, π-stacking, electrostatics, *etc.*) and hydrophobic interactions provide an efficient and powerful approach to prepare assemblies.[8–13] A variety of synthetic strategies have been used to prepare nucleotide-based amphiphiles,[14–16] including coupling reactions of a monoalkylphosphate with a nucleoside[17] and enzymatic catalysis.[18–21] A significant amount of the literature devoted to this family of amphiphiles describes mainly long-chain alkyl ester nucleoside-5′-monophosphate derivatives (Scheme 1, *i.e.* C_{16}UMP). These amphipathic nucleotides are known to form micelles spontaneously in water.[21] Interestingly, selective interactions between nucleobases have been reported in the mixed micellar systems and a preferential orientation of the purine and pyrimidine heterocycles perpendicular to the micellar radius has been deduced from NMR studies. The modulation of the polar head structure, in particular the spatial disposition of the base relative to the chain, could influence the aggregation properties of the amphipathic nucleotides. Therefore, we elected to design nucleotide-based amphiphile analogues with a monoalkylphosphate chain attached to the 3′-secondary hydroxyl. Although, to the best of our knowledge, nucleoside-3′-monophosphate bolaamphiphile structures have been reported by Shimizu and co-workers,[22,23] no single-chain nucleoside-3′-monophosphate amphiphile has been described so far.

With the aim of preparing and characterizing new nucleotide-based amphiphiles capable of forming supramolecular structures through concerted weak interactions, we describe here a series of nucleoside-3′-monophosphate and abasic amphiphiles featuring single hydrophobic chains (Scheme 1, compounds **I**, **II** and **III**). Electronic microscopy, X-ray scattering and FT-IR spectroscopy data indicate that the



Scheme 1 Chemical structures of amphipathic nucleotides. Left, C_{16}UMP (hexadecylmonophosphoryluridine) is an example of nucleoside-5′-monophosphate amphiphile.[21] Right, the thymine **I**, adenosine **II**, 3′-monophosphate derivatives and their abasic analogue **III** (n = 14, 18) studied in this paper.

[a] *Université Victor Segalen, Bordeaux, F-33076, France. E-mail: barthelemy@bordeaux.inserm.fr; Fax: +33 5 5757 1015; Tel: +33 5 5757 1014*
[b] *INSERM U869, Bordeaux, F-33076, France*
[c] *Université Montpellier II, UMR 5587, place Eugène Bataillon, 34095 Montpellier Cedex 5 France*
[d] *Departments of Chemistry and Biomedical Engineering, Metcalf Center for Science and Engineering, Boston University, Boston, MA 02215, USA*

† Electronic supplementary information (ESI) available: FT-IR, TEM and SEM. See DOI: 10.1039/b704884j.

nucleotide portion of the molecular structures is essential for the formation of nanostructure assemblies.

Experimental

Materials and instruments

Unless noted otherwise, all starting materials were obtained from commercial suppliers and were used without further purification. The solvents were redistilled over calcium chloride, calcium hydride, potassium hydroxide or sodium according to the solvent used. All compounds were characterized using standard analytical and spectroscopic techniques such as ^{1}H, ^{13}C and ^{31}P NMR spectroscopy (apparatus BRUKER Avance DPX-300, ^{1}H at 300.13 MHz, ^{13}C at 75.46 MHz and ^{31}P at 121.49 MHz) and mass spectroscopy (instrument JEOL SX 102, NBA matrix). The chemical shifts are reported in ppm relative to tetramethylsilane using the deuterium signal of the solvent ($CDCl_3$, CD_3OD) as a heteronuclear reference for ^{1}H and ^{13}C. The ^{1}H NMR coupling constants J are reported in Hz. TEM microscopy experiments were performed on a Philips CM 10 (negative staining with ammonium molybdate 1% in water, Cu/Pd carbon coated grids). The aqueous dispersions of the amphiphiles were dripped onto the grids and the excess water was blotted with paper. SEM secondary electron microscopy experiments were conducted on a JSM-6300F. SEM samples were deposited on carbon coated support, dried under vacuum, and platinum shadowed. Silica gel 60 (particle size 40–60 μm) was employed for flash chromatography. Thin layer chromatograms were performed with aluminium plates coated with silica gel 60 F_{254} (Merck). FT-IR was performed on a Thermo-Nicolet IR200 piloted by EZ OMNIC 7.2a. The technique of ATR was used. The powder was deposited on a diamond crystal. X-Ray diffraction experiments were carried out on solid powders in 1 mm diameter glass capillaries, in the Laboratoire des Colloïdes, Verres et Nanomatériaux (UMR5587), Université de Montpellier II, France. We worked in a transmission configuration. A copper rotating anode X-ray source (functioning at 4 kW) with a multilayer focusing "Osmic" monochromator giving high flux (10^8 photons sec^{-1}) and punctual collimation was employed. An "image plate" 2D detector was used. Diffraction curves were obtained giving diffracted intensity as a function of the wave vector q. Diffracted intensity was corrected by exposition time, transmission and intensity background coming from diffusion by an empty capillary. Indexation of the diffraction pattern corresponding to solid **3c** was performed using WinPlotR, Dicvol and ChekCell software. HyperChem software was used to build the structural model. Geometry optimization was obtained from energy minimum calculations using the molecular mechanic MM+ force field (using default data base parameters and a Polak-Ribiere algorithm), which optimize covalent bonds (distance, angles, torsion) and non-covalent interactions (van der Waals, hydrogen bonding) and which is suited for organic molecular systems. The potential energy of the molecular system is the sum of individual components such as bond, angle, van der Waals potential. HyperChem uses harmonic functions to calculate potentials for bonds, angles and 6–12 or 10–12 Lennard-Jones functions to simulate van der Waals and hydrogen bonding interactions. No specific constraints were imposed for the structure calculation.

Synthesis

5′-*O*-(4,4′-Dimethoxytrityl)deoxyadenosine 3′-eicosylphosphate 2a. 5′-Dimethoxytrityl-*N*-benzoyl-2′-deoxyadenosine 3′-[(2-cyanoethyl)-*N*,*N*-(diisopropyl)]phosphoramidite (500 mg, 1 eq, 0.58 mmol), eicosanol (226 mg, 1.3 eq, 0.76 mmol) and a tetrazole solution (0.45 M in acetonitrile, 1.7 ml, 0.76 mmol) were dissolved in dry acetonitrile under nitrogen. The reaction mixture was stirred for 24 h at room temperature, followed by oxidation with 40 ml of a solution of 0.02 M I_2 in THF–Pyr–H_2O. After 12 h at room temperature, the solvent was evaporated under high vacuum to yield crude product. The contents of the reaction flask were dissolved in 46 ml of NH_4OH 30% in water and heated at 55 °C under stirring in a sealed tube for 24 h. Product **2a** (68 mg) was isolated after purification on silica gel (MeOH–CH_2Cl_2 8 : 2). Yield: 13%. R_f = 0.48 (CH_2Cl_2–MeOH 9 : 1). MS FAB$^+$: m/z = 914 [M + 1]; FAB$^-$ m/z = 912 [M − 1].

Deoxyadenosine 3′-eicosylphosphate 3a. Phosphate **2a** (50 mg, 1 eq, 547 μmol) was dissolved in 3 ml of freshly distilled methylene chloride. An excess of 3% trichloroacetic acid solution in methylene chloride was added. After 2 h under nitrogen, 0.6 ml of MeOH were poured into the solution. The solvent was removed under vacuum and the residual product was isolated on silica gel. Product **3a** (327 mg) was isolated after purification on silica gel (MeOH–CH_2Cl_2 8 : 2). Total yield: 92%. R_f = 0.40 (CH_2Cl_2–MeOH 86 : 14). **^{1}H NMR (300 MHz, CD_3OD)**: δ 0.88 (t, 3H, J = 6.65 Hz, CH_3), 1.32 (dd, 34H, J_1 = 4.37 Hz, J_2 = 7.26 Hz, 17CH_2), 1.62 (dd, 2H, J_1 = 6.75 Hz, J_2 = 13.96 Hz, CH_2), 2.60–2.67 (m, 1H, $H_{2'}$), 2.85–2.90 (m, 1H, $H_{2'}$), 3.34 (t, 1H, $H_{3'}$), 3.84 (dd, 1H, J_1 = 2.93 Hz, J_2 = 4.62 Hz, $H_{4'}$), 3.88 (dd, 2H, J_1 = 6.42 Hz, J_2 = 12.73 Hz, $H_{5'}$), 4.96–5.00 (m, 2H, CH_2), 6.44 (dd, 1H, J_1 = 5.87 Hz, J_2 = 8.30 Hz, $H_{1'}$), 8.16 (s, 1H, $H_{2(base)}$), 8.34 (s, 1H, $H_{8(base)}$). **^{13}C NMR (75 MHz, CD_3OD)**: δ 14.49 (CH_3), 23.81 (CH_2), 27.03 (CH_2), 30.50 (CH_2), 30.79 (CH_2), 33.11 (CH_2), 63.61 ($C_{5'}$), 66.70 (CH_2OP), 77.43 ($C_{3'}$), 87.39 ($C_{1'}$), 88.99 ($C_{4'}$), 141.72 ($C_{2(base)}$), 153.46 ($C_{8(base)}$). **^{31}P NMR (121 MHz, CD_3OD)**: δ 1.08 ppm. High resolution FAB MS: theoretical m/z = 612.3884; observed m/z = 612.3890.

Deoxyadenosine 3′-hexadecylphosphate 3b. A similar procedure was followed as for deoxyadenosine 3′-eicosylphosphate **3a**. 500 mg of 5′-dimethoxytrityl-*N*-benzoyl-2′-deoxyadenosine 3′-[(2-cyanoethyl)-*N*,*N*-(diisopropyl)]phosphoramidite (0.58 mmol, 1 eq), 1-hexadecanol (185 mg, 0.76 mmol, 1.3 eq) and a tetrazole solution (0.45 M in acetonitrile, 1.7 ml, 0.76 mmol) were dissolved in dry acetonitrile under nitrogen. The reaction mixture was stirred for 24 h at room temperature. The resulting mixture was oxidized by adding 40 ml of a solution of 0.02 M I_2 in THF–Pyr–H_2O. After 12 h at room temperature the solvent was evaporated under high vacuum to yield crude product. The contents of the reaction flask were dissolved in 50 ml of NH_4OH 30% in water and heated at 55 °C under stirring in a sealed tube for 24 h. The deprotection

of dimethoxytrityl group occurs *in situ* during the purification. Product **3b** (231 mg) was isolated after purification on silica gel (MeOH–CH_2Cl_2 8 : 2). Total yield: 92%. R_f = 0.40 (CH_2Cl_2–MeOH 86 : 14). **^{1}H NMR (300 MHz, CD_3OD)**: δ 0.88 (t, 3H, J = 6.65 Hz, CH_3), 1.32 (dd, 26H, J_1 = 4.7 Hz, J_2 = 7.26 Hz, 13CH_2), 1.62 (dd, 2H, J_1 = 6.75 Hz, J_2 = 13.96 Hz, CH_2), 2.60–2.67 (m, 1H, $H_{2'}$), 2.85–2.90 (m, 1H, $H_{2'}$), 3.34 (t, 1H, $H_{3'}$), 3.84 (dd, 1H, J_1 = 2.93 Hz, J_2 = 4.62 Hz, $H_{4'}$), 3.88 (dd, 2H, J_1 = 6.42 Hz, J_2 = 12.73 Hz, $H_{5'}$), 4.96–5.00 (m, 2H, CH_2), 6.44 (dd, 1H, J_1 = 5.87 Hz, J_2 = 8.30 Hz, $H_{1'}$), 8.16 (s, 1H, $H_{2(base)}$), 8.34 (s, 1H, $H_{8(base)}$). **^{13}C NMR (75 MHz, CD_3OD)**: δ 14.49 (CH_3), 23.81 (CH_2), 27.03 (CH_2), 30.50 (CH_2), 30.79 (CH_2), 33.11 (CH_2), 63.61 ($C_{5'}$), 66.70 (CH_2OP), 77.43 ($C_{3'}$), 87.39 ($C_{1'}$), 88.99 ($C_{4'}$), 141.72 ($C_{2(base)}$), 153.46 ($C_{8(base)}$). **^{31}P NMR (121 MHz, CD_3OD)**: δ 1.08 ppm. High resolution FAB MS: theoretical m/z = 556.3264; observed m/z = 556.3278.

5′-*O*-(4,4′-Dimethoxytrityl)thymidine 3′-eicosylphosphate 2c. 5′-Dimethoxytrityl-2′-deoxythymidine 3′-[(2-cyanoethyl)-*N*,*N*-(diisopropyl)]phosphoramidite (500 mg, 1 eq, 0.67 mmol), eicosanol (260 mg, 1.3 eq, 0.87 mmol) and a tetrazole solution (0.45 M in acetonitrile, 2 ml, 0.87 mmol) were dissolved in dry acetonitrile under nitrogen. The reaction mixture was stirred for 24 h at room temperature. The resulting mixture was oxidized by adding 40 ml of a solution of 0.02 M I_2 in THF–Pyr–H_2O. After 12 h at room temperature, the solvent was evaporated under high vacuum to yield crude product. The contents of the reaction flask were dissolved in 50 ml of NH_4OH 30% in water and heated at 55 °C under stirring in a sealed tube for 24 h. Product **2c** (396 mg) was isolated after purification on silica gel (MeOH–CH_2Cl_2 8 : 2). Yield: 65%. R_f = 0.58 (CH_2Cl_2–MeOH 86 : 14). MS FAB$^-$ m/z = 904 [M − 1].

Thymidine 3′-eicosylphosphate 3c. Phosphate **2c** (396 mg, 1 eq, 0.47 mmol) was dissolved in 20 ml of freshly distilled methylene chloride. An excess of 3% trichloroacetic acid solution in methylene chloride was added. After 2 h under nitrogen, 2.5 ml of MeOH were poured into the solution. The solvent was removed under vacuum and the residual product **3c** (60 mg) was isolated on silica gel (CH_2Cl_2–MeOH 1 : 0 to 0 : 1). Yield: 67%. R_f: 0.32 (CH_2Cl_2–MeOH 86 : 14). **^{1}H NMR (CD_3OD)**: δ in ppm 0.82 (t, 3H, J = 6.01 Hz, CH_3), 1.20 (m, 34H, 17CH_2), 1.54 (dd, 2H, J_1 = 6.79 Hz, J_2 = 13.73 Hz, CH_2), 1.80 (s, 3H, CH_3), 2.13–2.23 (m, 1H, $H_{2'}$), 2.33–2.43 (m, 1H, $H_{2'}$), 3.68–3.72 (m, 2H, $2H_{5'}$), 3.79 (m, 2H, CH_2O), 4.06 (m, 1H, $H_{4'}$), 4.68–4.75 (m, 1H, $H_{3'}$), 6.22 (t, 1H, J_1 = 6.78 Hz, $H_{1'}$), 7.79 (s, 1H, CH). **^{13}C NMR (CD_3OD)**: δ in ppm 12.33 (CH_3), 14.33 (CH_2), 23.63 (CH_2), 26.61 (CH_2), 30.36–33.66 (CH_2), 39.72 (C5′), 68.36 (CH_2OP), 78.29 (C3′), 86.29 ($C_{1'}$), 87.54 ($C_{4'}$), 137.92 (CH), 152.34 (CO, thymine). **^{31}P NMR (CD_3OD)**: δ in ppm −0.31 ppm. High resolution FAB$^+$: [M − H], theoretical m/z = 603.3774; observed m/z = 603.3766.

Thymidine 3′-hexadecylphosphate 3d. A similar procedure was followed as for deoxythymidine 3′-eicosylphosphate. 5′-Dimethoxytrityl-2′-deoxythymidine 3′-[(2-cyanoethyl)-*N*,*N*-(diisopropyl)]phosphoramidite (500 mg, 1 eq, 0.67 mmol), 1-hexadecanol (213 mg, 1,3 eq, 0.87 mmol) and a tetrazole solution (0.45 M in acetonitrile, 2 ml, 0.87 mmol) were dissolved in dry acetonitrile under nitrogen. The reaction mixture was stirred for 24 h at room temperature. The resulting mixture was oxidized by adding 45 ml of a solution of 0.02 M I_2 in THF–Pyr–H_2O. After 12 h at room temperature the solvent was evaporated under high vacuum to yield crude. The contents of the reaction flask were dissolved in 50 ml of NH_4OH 30% in water and heated at 55 °C under stirring in a sealed tube for 24 h. The deprotection of the dimethoxytrityl group occurs *in situ* during the purification Product **3d** (57 mg) was isolated after purification on silica gel (MeOH–CH_2Cl_2 8 : 2). Yield: 21%. R_f: 0.35 (CH_2Cl_2–MeOH 7 : 3). **^{1}H NMR (300 MHz, CD_3OD)**: δ in ppm 0.92 (t, 3H, J = 6.23 Hz, CH_3), 1.30 (s, 26H, 13CH_2), 1.60–1.71 (m, 2H, CH_2), 1.92 (s, 3H, CH_3), 2.22–2.35 (m, 1H, $H_{2'}$), 2.47–2.57 (m, 1H, $H_{2'}$), 3.38 (s, 1H, $H_{3'}$), 3.84 (m, 2H, $H_{5'}$), 3.91 (dd, 2H, J_1 = 6.46 Hz, J_2 = 12.46 Hz, CH_2O), 4.19 (m, 1H, $H_{4'}$), 6.32 (t, 1H, J = 6.59 Hz, $H_{1'}$), 7.84 (s, 1H, CH). **NMR ^{13}C (75 MHz, CD_3OD)**: δ in ppm 12.54 (CH_3), 14.45 (CH_2), 23.62 (CH_2), 26.88 (CH_2), 30.35 (CH_2), 32.92 (CH_2), 33.53 (CH_2). **NMR ^{31}P (121 MHz, CD_3OD)**: δ in ppm −0.25. High resolution FAB$^-$ MS: theoretical m/z = 545.2992; observed m/z = 545.3013.

Eicosylhydroxyethylphosphate 5a. Freshly distilled THF (16 ml) and dry triethylamine (TEA) (900 µl, 1.8 eq, 6.03 mmol) were added under nitrogen to 1-eicosanol (1 g, 3.35 mmol). The mixture was cooled to 0 °C and 2-chloro-2-oxo-1,3,2-dioxaphospholane (500 µl, 1.6 eq, 5.36 mmol) was added dropwise. The reaction mixture was stirred at room temperature for 3 h. Next, the TEA salts were removed by filtration under vacuum. Most of the solvent was evaporated under reduced pressure and the residual solution was used directly without further purification in the following step. The residual solution was diluted in acetonitrile (130 ml) and water (10 ml) was added dropwise at 0 °C. Next, the solution was warmed to 35 °C under stirring overnight. The product precipitated at 0 °C. 1.215 g of a white solid was obtained after drying under high vacuum. Yield: 86%. R_f: 0.25 (CH_2Cl_2–hexane 7 : 3). **^{1}H NMR (300 MHz, $CDCl_3$)**: δ 0.87 (t, 3H, J = 7.88 Hz, CH_3), 1.24 (s, 26H, 13 CH_2), 1.32 (t, 10H, J = 7.33 Hz, 5 CH_2), 3.89 (dd, 2H, J_1 = 6.51 Hz, J_2 = 13.04 Hz, CH_2OH), 4.03 (m, 2H, $CH_2CH_2OP(O)(OH)CH_2$). **^{13}C NMR (75 MHz, $CDCl_3$)**: δ 14.08 (CH_3), 22.66 (CH_2), 25.78 (C_{18}), 29.34 (C_4 and C_{17}), 29.62 (C_5 to C_{16}), 29.67 (C_{19}), 31.89 (C_3), 63.46 (C_{20}), 67.76 ($C_{1'}$). **^{31}P NMR (121 MHz, $CDCl_3$)**: δ 3.22 ppm. High resolution FAB MS: theoretical m/z = 423.3239 observed m/z = 423.3214.

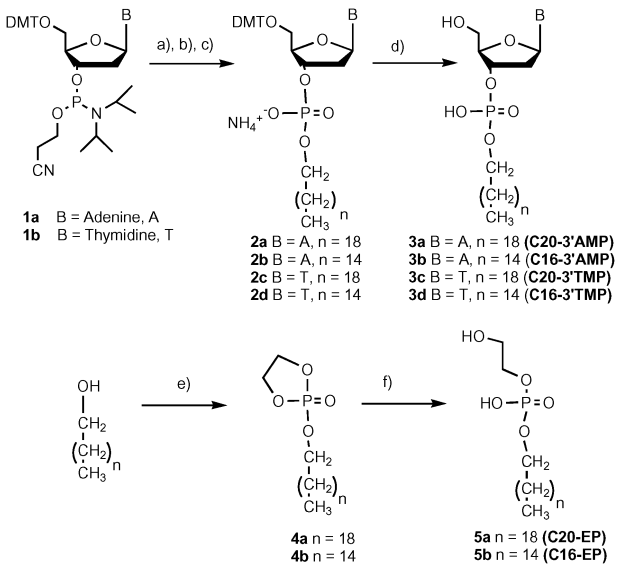
Hexadecylhydroxyethylphosphate 5b. A similar procedure was followed as for eicosylhydroxyethylphosphate **5a**. Freshly distilled THF (15 ml) and dry TEA (1.04 ml, 1.8 eq, 7.42 mmol) were added under nitrogen to 1-hexadecanol (1 g, 4.12 mmol). The mixture was cooled to 0 °C and 2-chloro-2-oxo-1,3,2-dioxaphospholane (600 µl, 1.6 eq, 6.60 mmol) was added dropwise. The reaction mixture was stirred at room temperature for 3 h. Next, the TEA salts were removed by filtration under vacuum. Most of the solvent was evaporated under reduced pressure and the residual solution was used directly

without further purification in the following step. The residual solution was diluted in acetonitrile (160 ml) and water (6 ml) was added dropwise at 0 °C. Next, the solution was warmed to 35 °C under stirring overnight. The product precipitated at 0 °C. Yield: 86%. R_f: 0.26 (CH_2Cl_2–hexane 7 : 3). **^{1}H NMR (300 MHz, $CDCl_3$)**: δ 0.87 (t, 3H, J = 6.59 Hz, CH_3), 1.25–1.30 (m, 26H, $13CH_2$), 1.51–1.70 (m, 2H, CH_2), 3.99 (dd, 2H, J_1 = 6.46 Hz, J_2 = 13.63 Hz, CH_2OH), 4.07 (m, 2H, $CH_2OP(O)$). **^{13}C NMR (75 MHz, $CDCl_3$)**: δ 14.09 (CH_3), 22.67 (CH_2), 25.72 (C_{18}), 29.34 (C_4 and C_{15}), 29.69 (C_5 to C_{14}), 31.91 (C_3), 63.03 ($C_{1'}$). **^{31}P NMR (121 MHz, $CDCl_3$)**: δ 0.26 ppm. High resolution FAB MS: theoretical m/z = 367.2608; observed m/z = 367.2627.

Results and discussion

Synthesis

To assess the effect of the amphiphile structure (base, chain length) on the self-assembly properties, a family of nucleotide derivatives composed of adenosine, thymidine and abasic lipophilic phosphates bearing either eicosyl (C20) or palmityl (C16) hydrophobic chains was synthesized. The single chain nucleoside-3′-monophosphate amphiphiles were prepared by using a convenient phosphoramidite synthetic approach as shown in Scheme 2. After preparation of the lipophilic intermediates in acidic conditions (tetrazole) from 5′-DMT (dimethoxytrityl) protected commercial phosphoramidites, these intermediates were treated with iodine in a THF–pyridine–water mixture at room temperature under nitrogen to afford the cyanoethyl protected phosphates. The cyanoethyl protecting group was then removed *via* a β-elimination reaction under basic conditions to provide the expected products **2**.



Scheme 2 Synthesis of nucleotide-based amphiphiles. (a) $CH_3(CH_2)_nCH_2OH$; tetrazole (0.45 M in MeCN); THF; RT; N_2; 24 h. (b) I_2 0.02 M in THF–Pyr–H_2O 70 : 20 : 10; RT; N_2; 12 h. (c) NH_4OH 30% in water; 55 °C; 24 h. (d) CCl_3COOH 3% in CH_2Cl_2; RT; 2 h. Total yield: (**3a**) 92%, (**3b**) 92%, (**3c**) 67%, (**3d**) 21%. (e) 2-Chloro-2-oxo-1,3,2-dioxaphospholane; THF; TEA; 0 °C; N_2; then RT; N_2; 3 h. (f) H_2O; MeCN; 0 °C; 16 h. Total yield (**5a**): 86%, (**5b**) 86%.

Removal of the dimethoxytrityl protecting group leads to final products **3**. Abasic phosphate analogues featuring a primary hydroxyl mimicking the 5′-hydroxyl of nucleoside were prepared *via* a 2 step synthesis (Scheme 2). In that case, the long-chain alcohols were first reacted with 2-chloro-1,3,2-dioxaphospholane-2-oxide in THF in the presence of triethylamine. Cyclic intermediates **4a** and **4b** were then simply hydrolyzed to afford the expected abasic analogues **5a** and **5b**. Both synthetic strategies developed in this work used commercially available starting materials and are amenable to the preparation of a large number of derivatives.

Microscopy

In order to determine the impact of the nucleoamphiphile structure on the formation of the supramolecular assemblies at the nanometre and the micrometre scales, the single chain nucleoside-3′-monophosphate amphiphiles **3a–d** and **5a** were first investigated by transmission and scanning electronic microscopies (TEM, SEM). A typical experiment involves forming a colloidal dispersion of phosphate derivatives at high concentrations (higher than 1 mM) to afford self-assemblies. The experimental protocol is as follows. First, 1 mg of compound **3a–d** or **5a** is dispersed in 1 ml of an aqueous medium (deionized water pH 5.5, tris buffer pH 7.4, or tris buffer pH 9.4) and the dispersions are heated to 70 °C for 30 s. Then, the samples are sonicated in a Branson 3200 cleaning bath for 30 s at room temperature. This cycle (heating and sonicating) was repeated 5 times to obtain homogeneous dispersions. As shown in Fig. 1a and 1b, this simple procedure afforded ribbon-like aggregates for thymidine derivatives **3d** (C16-3′TMP) and **3c** (C20-3′TMP) whereas, under similar conditions, small heterogeneous objects are observed for the adenine nucleotide amphiphiles **3a** (C20-3′AMP, Fig. 1d) and **3b** (C16-3′AMP, data not shown). Both A and T amphiphiles produce self-assembled solids independently of the aqueous media used. In all cases (deionized water pH 5.5, tris buffers 10 mM, pH 7.4, or 9.4), the pH of the media is higher than the

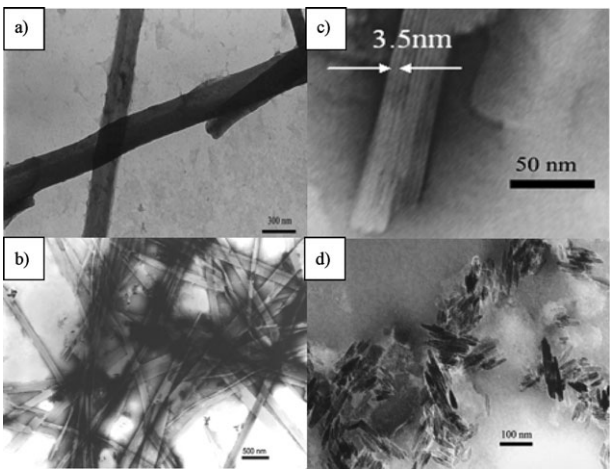


Fig. 1 TEM images of C16-3′TMP (a), C20-3′TMP (b, c) and C16-3′AMP (d) from samples dispersed in deionized water (1 mg ml^{-1}). Negative staining ammonium molybdate 1% in water. For C20-3′TMP, lamellar structure was observed with interlamellar spacing of about 3.5 nm.

pK_a of the phosphoryl group. T Amphiphiles C20-3′TMP and C16-3′TMP exhibited ribbons with 80–150 nm width (Fig. 1a and 1b, and ESI†). In the case of adenosine amphiphiles, C20-3′AMP and C16-3′AMP, micron-sized aggregates composed of several small crystals structures with a thickness of roughly 10 nm varying in length between 20 to 100 nm, were seen. These results show that the nucleotide component of the amphiphile plays an important role in producing supramolecular objects organized at the nano–micro scales. In contrast, the shape and the size of the ribbons and crystal structures do not depend on the aqueous medium used in this work or the chain length of the amphiphilic structures.

Interestingly, at higher magnification, TEM images of ribbons obtained from C20-3′TMP (Fig. 1c) show aligned supramolecular structure well organized at the nanometre scale. The repetitive distance is estimated to be 3.5 nm from the ribbons shown in the TEM image (Fig. 2). To confirm the importance of the base on the self assembly process, the abasic analogues **5a** (C20-EP) and **5b** (C16-EP) were investigated under similar conditions by TEM. No supramolecular assemblies were observed, indicating that the presence of the nucleotide is clearly required to afford the self assemblies.

Scanning electron microscopy (SEM) of A and T amphiphiles C20-3′AMP (Fig. S7, see ESI†) and C20-3′TMP (Fig. S8, see ESI†), respectively, dispersed in deionized water and lyophilized, exhibit different organizations. Fig. S7 and S8 show aggregates of crystals at different magnification, confirming the presence of the aggregates observed by TEM. The formation of ribbons, in the case of the T derivative C20-3′TMP, was also confirmed by SEM. The image shown in Fig. S8 exhibits similar ribbons as the ones previously observed by TEM. Considering the possible complementary molecular recognition phenomena between purine and pyrimidine bases, equimolecular mixtures of A and T amphiphiles C20-3′AMP and C20-3′TMP were dispersed (1 mg ml^{-1}, deionized water, heat at 70 °C and sonication). Surprisingly, the TEM images collected from the resulting dispersions show the cohabitation of both unaffected systems (ribbons and small crystal structures). The absence of a third expected supramolecular system involving A and T nucleotide-based amphiphiles may be explained by the limited water solubility of the supramolecular objects observed at high concentrations.

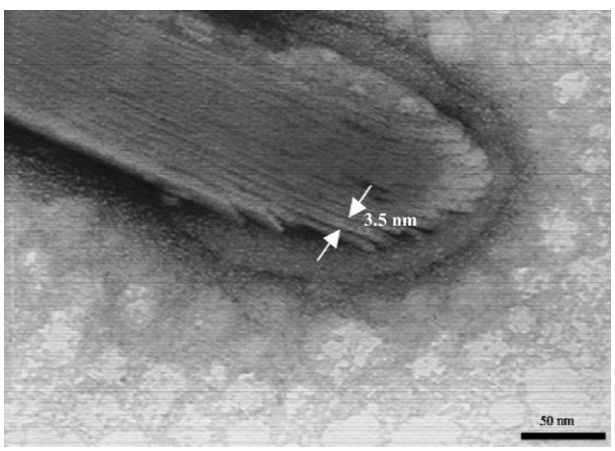


Fig. 2 High resolution TEM image of C20-3′TMP dispersed in deionized water (1 mg ml^{-1}). Negative staining ammonium molybdate 1% in water.

FT-IR studies

To determine whether the formation of the aggregates observed by electron microscopies depends on particular molecular interactions, the dried samples (lyophilized) obtained from the dispersions were studied by infrared (IR) spectroscopy. The FT-IR spectra collected from the supramolecular assemblies can be divided into spectral regions that originate from molecular vibrations of the polar heads (nucleotide moieties) and those of the hydrocarbon chains.

The FT-IR absorption spectrum in the region of 1500–2000 cm^{-1} of the assemblies formed from C20-3′TMP is shown in Fig. 3a.

The infrared spectrum of C20-3′TMP contains a high-wavenumber C_2=O carbonyl stretching vibration band, at 1722 cm^{-1}, that is not present in the thymidine spectrum (Fig. S3, ESI†). A second C_2=O carbonyl stretching vibration band at 1684 cm^{-1} is also observed for C20-3′TMP. The presence of these carbonyl vibration absorptions can be correlated to the existence of C_2=O hydrogen-bonded carbonyl groups in the structure, indicating that T–T base pairing occurs in the thymidine-based amphiphile supramolecular assemblies. Note that an associated carbonyl stretch band at 1654 cm^{-1}, characteristic of the C_4=O associated stretching vibration, is also observed, confirming that the C_4=O is also involved in T–T hydrogen bonding.[24]

For the adenosine-based supramolecular assemblies, the intense bands around 1610 and 1690 cm^{-1}, corresponding to the ring vibration and NH_2 scissoring, respectively, were found to be greatly reduced in the IR spectrum of adenosine derivative C20-3′AMP (Fig. 3b). The new bands, observed at 1652 and 1698 cm^{-1} in the case of C20-3′AMP, strongly suggest the presence of hydrogen-bonded scissoring,[25] which is compatible with A–A hydrogen bonding occurring in the adenosine amphiphile self-assemblies.

The hypochromic effect in the band of adenine in-plane ring vibrations near 1300 cm^{-1}, has been reported as a base stacking marker.[26] This vibration band, observed at 1297 cm^{-1} for adenosine (see ESI Fig. S2†), is reduced dramatically in the spectrum of C20-3′AMP. This observation indicates that base stacking A–A stabilize the supramolecular assemblies obtained from adenosine derivative C20-3′AMP. The weak intensities of the N1–C6–H and C5–C6–H deformations of thymine appearing at 1281 cm^{-1} for compound C20-3′TMP suggest a T–T base stacking for C20-3′TMP supramolecular assemblies (see ESI Fig. S3†).[26]

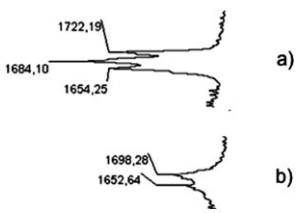


Fig. 3 FT-IR profiles of C20-3′TMP (a) in the region of 1500–2000 cm^{-1}. (b) Show FT-IR spectra in the same region for C20-3′AMP. For FT-IR profiles of adenosine and thymidine see Fig. S2 and S4 (ESI†).

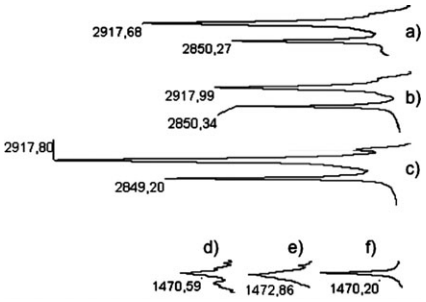


Fig. 4 FT-IR profiles of C20-3′TMP (a), C20-3′AMP (b) and C20-EP (c) in the region of 2700–3000 cm^{-1}. Symmetric and antisymmetric C–H stretching bands are observed at 2917 and 2850 cm^{-1} for all the amphiphiles. δ CH_2 scissoring vibrations observed for derivatives C20-3′TMP (d), C20-3′AMP (e) and C20-EP (f).

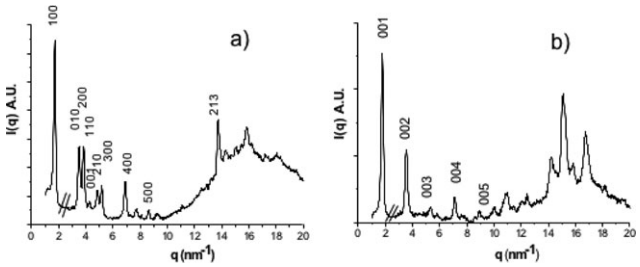


Fig. 5 X-Ray diffraction pattern of C20-3′TMP (a), C20-3′AMP (b). Scattering intensity is plotted as a function of the wave vector q (nm^{-1}).

It was reported previously that the antisymmetric (γ_{asym}) and symmetric (γ_{sym}) methylene stretching vibrations give rise to two strong bands located at 2920 and 2850 cm^{-1} regions, respectively.[27] The exact frequencies of these vibrations are strongly dependent on the chain conformations. Acyl chains with all-*trans* conformation exhibit a γ_{sym} stretching band frequency around 2846–2849 cm^{-1}, whereas an increase of gauche conformers shifts this band to 2853 cm^{-1}. The γ_{asym} and γ_{sym} C–H stretching bands observed at 2917 and 2850 cm^{-1}, respectively for both nucleotide amphiphiles C20-3′AMP, C20-3′TMP and the abasic derivative C20-EP (Fig. 4a, b and c) are compatible with high *trans* conformational population in the alkyl chains.[28,29] The (δ CH_2) scissoring vibrations are very sensitive to the packing of the hydrocarbon chains. A splitting of the (δ CH_2) scissoring mode in two bands, as it is observed for thymidine derivative C20-3′TMP (Fig. 4d), is indicative of alkyl chains arranged in an orthorhombic lattice.[30] The sharp band observed at 1473 cm^{-1} on the FT-IR spectrum of adenosine derivative C20-3′AMP, indicates the presence of a triclinic packing of the alkyl chains in the supramolecular assemblies, whereas a hexagonal packing is marked by a shift to 1470 cm^{-1} in the case of the abasic derivative C20-EP.

X-Ray diffraction

The wide angle X-ray diffraction patterns of the dried self assemblies from nucleoamphiphiles C20-3′TMP and C20-3′AMP are shown in Fig. 5. For abasic compound C20-EP, diffraction from only one family of reticular planes is observed (Fig. S11, ESI†). The most intense first order reflection (001) is followed by subsequent second order (002) to sixth order (006) reflections. No other reflections can be observed. This result is in accordance with a bidimensional smectic (lamellar)-like structure with long range structural ordering in the direction perpendicular to the smectic planes but without periodicity in the smectic planes. In contrast, for A and T derivatives, C20-3′AMP and C20-3′TMP, respectively, different families of reticular planes can be observed, indicating crystal-like 3D structural ordering. A better structural ordering likely for the solids is obtained from the nucleoamphiphiles compared to the abasic derivative C20-EP. A structural model is proposed for C20-3′TMP as shown in Fig. 6, which takes into account the various intermolecular interactions: hydrogen bonding between thymine groups (Fig. 6a), van der Waals interactions between alkyl chains (Fig. 6b) and crystal lattice parameters

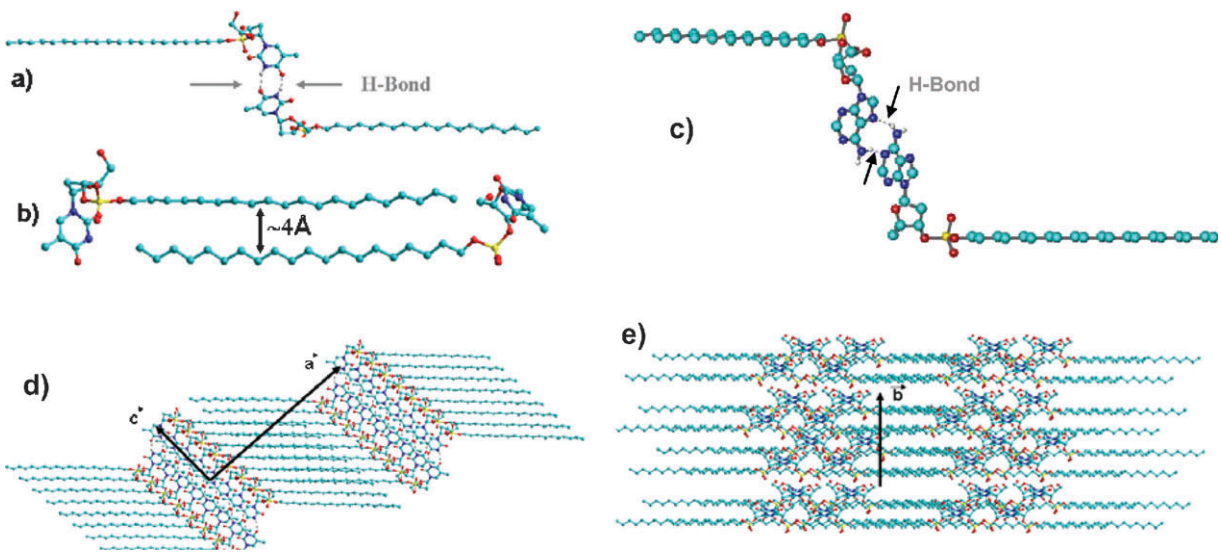


Fig. 6 Structural model proposed for C20-3′TMP. (a) Local hydrogen bonding interaction between two thymine groups. (b) Local hydrophobic (van der Waals) interactions between two alkyl chains. (c) Hydrogen bonding interaction involving two adenosine groups (C20-3′AMP). (d) and (e) 3D structure with monoclinic crystal lattice parameters a (36.485 Å), b (17.962 Å) and c (14.681 Å). (α, β = 90° and γ = 91.91°).

(*a*, *b*, *c*) determined after indexation of the diffraction pattern.[31] The best indexation calculation gave a monoclinic crystal lattice, very close to an orthorhombic symmetry, which is consistent with the FT-IR data obtained from C20-3′TMP supramolecular assemblies. The lamellar-like structure observed in the structural model proposed in Fig. 6d in the (*a*, *b*) plane gives a lamellar periodicity of 36.471 Å in the *a* direction, in accordance with an interlamellar spacing of about 35 Å deduced from the TEM image (Fig. 1). Although the resolution of the X-ray spectrum (Fig. 5b) is fairly poor in the case of the A derivative C20-3′AMP, to determine the all the crystal lattice parameters, a lamellar periodicity of 36.0 Å can be deduced from the diffraction pattern. In summary, T and A amphiphiles self-assemble thanks to hydrogen bonding and van der Waals interactions to provide nanostructured materials. In the case of T derivative C20-3′TMP, the ribbon structures feature an orthorhombic system, whereas the small crystals observed in the case of A amphiphile C20-3′AMP would be due to a triclinic arrangement of molecules.

Conclusion

The present study has demonstrated that the coupling of a 3′-phosphoryl nucleoside unit to the molecular structure of single-chain amphiphiles leads to the preparation of assemblies, which are structured at the nanoscale. As supramolecular assemblies are not fully predictable by molecular design, a variety of nucleotide-based derivatives bearing either a purine or pyrimidine as a hydrophilic moiety have been prepared. When adenosine and thymidine amphiphiles are homogeneously dispersed in aqueous media, small aggregates organized at the nanoscale are observed. Scanning and electron microscopy studies reveal that these nucleotide amphiphiles self-assemble into ribbon-like structures with 80–150 nm width for the thymidine derivatives and small aggregates with a thickness of a few nanometres varying in length between 20 to 100 nm in the case of adenine. The data collected from TEM, SEM, FT-IR and WAXS (wide-angle X-ray scattering) studies indicate that the morphologies of the supramolecular systems observed are driven by the nature of the base inserted in the molecular structure. A molecular model illustrating the supramolecular arrangements of a thymidine nucleoamphiphile forming ribbons is also proposed. The purine- and pyrimidine-based amphiphiles reported in this study extend the current family of small molecules that possess chemical features capable of forming specific non-covalent interactions and these nucleotide-based amphiphiles can be used as tools to construct nanostructured assemblies.

Acknowledgements

This research was supported by the Army Research Office (Grant DAAD 19-02-1-0386). The authors would like to thank Dr Stephen J. Lee for valuable discussions.

References

1 S. Nayak and L. A. Lyon, *Angew. Chem., Int. Ed.*, 2005, **44**(47), 7686–7708.
2 C. J. Hawker and L. J. Wooley, *Science*, 2005, **309**(5738), 1200–1205.
3 P. Podsiadlo, S. Paternel, J. M. Rouillard, Z. Zhang, J. Lee, J. W. Lee, E. Gulari and N. A. Kotov, *Langmuir*, 2005, **21**(25), 11915–11921.
4 A. S. Arico, P. Bruce, B. Scrosati, J. M. Tarascon and W. van Schalkwijk, *Nat. Mater.*, 2005, **4**(5), 366–377.
5 E. H. Lan, B. Dunn and J. I. Zink, *Methods Mol. Biol. (Totowa, N. J.)*, 2005, **300**, 53–79.
6 A. Kikuchi and T. Okano, *J. Controlled Release*, 2005, **101**(1–3), 69–84.
7 D. A. LaVan, T. McGuire and R. Langer, *Nat. Biotechnol.*, 2003, **21**(10), 1184–1191.
8 C. Xu. P. Taylor, P. D. I. Fletcher and V. N. Paunov, *J. Mater. Chem.*, 2005, **15**, 394–402.
9 P. Barthélémy, S. J. Lee and M. W. Grinstaff, *Pure Appl. Chem.*, 2005, **77**(12), 2133–2148.
10 D. Berti, P. Baglioni, S. Bonaccio, G. Barsacchi-Bo and P. L. Luisi, *J. Phys. Chem. B*, 1998, **102**, 303–308.
11 M. Onda, K. Yoshihara, H. Koyano, K. Ariga and T. Kunitake, *J. Am. Chem. Soc.*, 1996, **118**(36), 8524–8530.
12 P. Baglioni and D. Berti, *Curr. Opin. Colloid Interface Sci.*, 2003, **8**, 55–61.
13 L. Moreau, P. Barthelemy, M. El Maataoui and M. W. Grinstaff, *J. Am. Chem. Soc.*, 2004, **126**(24), 7533–7539.
14 T. Tanaka, K. Nakayama, K. Machida and M. Taniguchi, *Microbiology (Reading, U. K.)*, 2000, **146**, 377–384.
15 G. M. van Wijk, K. Y. Hostetler and H. van den Bosch, *Biochim. Biophys. Acta*, 1991, **1084**(3), 307–310.
16 C. I. Hong, A. J. Kirisits, A. Nechaev, D. J. Buchheit and C. R. West, *J. Med. Chem.*, 1985, **28**(2), 171–177.
17 J. Smrt and S. Hynie, *Collect. Czech. Chem. Commun.*, 1980, **45**, 927.
18 G. Zandomeneghi, P. L. Luisi, L. Mannina and A. Segre, *Helv. Chim. Acta*, 2001, **84**(12), 3710–3725.
19 U. Rädler, C. Heiz, P. L. Luisi and R. Tampé, *Langmuir*, 1998, **14**, 6620–6624.
20 D. Berti, P. L. Luisi and P. Baglioni, *Colloids Surf., A*, 2000, **167**, 95–103.
21 C. Heiz, U. Radler and P. L. Luisi, *J. Phys. Chem. B*, 1998, **102**, 8686–8691.
22 R. Iwaura, K. Yoshida, M. Masuda, K. Yase and T. Shimizu, *Chem. Mater.*, 2002, **14**, 3047–3053.
23 R. Iwaura, K. Yoshida, M. Masuda, M. Ohnishi-Kameyama, M. Yoshida and T. Shimizu, *Angew. Chem., Int. Ed.*, 2003, **42**(9), 1009–1012.
24 C. Dagneaux, J. Liquier and E. Taillandier, *Biochemistry*, 1995, **34**, 14815–14818.
25 T. Shimizu, R. Iwaura, M. Masuda, T. Hanada and K. Yase, *J. Am. Chem. Soc.*, 2001, **123**, 5947–5955.
26 M.-C. Lin, P. Eid, P. T. T. Wong and R. B. Macgregor, Jr, *Biophys. Chem.*, 1999, **76**, 87–94.
27 P. Garidel, A. Blume and W. Hubner, *Biochim. Biophys. Acta*, 2000, **1466**(1–2), 245–259.
28 P. Garidel, *Phys. Chem. Chem. Phys.*, 2002, **4**(10), 1934–1942.
29 R. G. Snyder, H. L. Strauss and C. A. Elliger, *J. Phys. Chem.*, 1982, **86**(26), 5145–5150.
30 P. Garidel, W. Richter, G. Rapp and A. Blume, *Phys. Chem. Chem. Phys.*, 2001, **3**(8), 1504–1513.
31 Indexation was performed using WinPlotR, Dicvol and ChekCell software. Relatively large diffraction peaks giving overlap between closed reflections made indexation difficult, especially in the high *q* domain of the diffraction pattern. The best indexation calculation gave a monoclinic crystal lattice with $a = 36.485$ Å, $b = 17.962$ Å and $c = 14.681$ Å, (α, $\beta = 90°$ and $\gamma = 91.91°$). The indexation pattern is shown in Fig. S11 (see ESI†).

