Chiral organocatalysts based on lipopeptide micelles for aldol reactions in water

Article

Accepted Version

Physical Chemistry Chemical Physics

ARTICLE

Chiral Organocatalysts Based on Lipopeptide Micelles for Aldol Reactions in Water

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

B. M. Soares,[a] A. M. Aguilar,[b] E. R. Silva,[c] M. D. Coutinho-Neto,[a] I. W. Hamley,[d] M. Reza,[e] J. Ruokolainen[e] and W. A. Alves[a]†

A comprehensive study of the self-assembly in water of a lipopeptide consisting of a sequence of *L*-proline, *L*-arginine and *L*-tryptophan with a hydrocarbon chain has been performed. Fluorescence assays were used to determine critical aggregation concentration. *In situ* small-angle X-ray scattering (SAXS) and molecular dynamics simulations showed the presence of spherical micelles with diameters around 6 nm. In agreement with these results, *cryo*-TEM images showed globular aggregates with diameters ranging from ≈ 4 nm up to ≈ 9 nm. Furthermore, the lipopeptide catalytic activity has been tested for the direct aldol reaction between cyclohexanone and *p*-nitrobenzaldehyde, and we have observed that the self-association of the organocatalyst played critical role in the enhanced activity. Water affects the selectivity, and poor results are obtained under neat reaction conditions. The location of the catalytic groups at the lipopetide/water solvent interface also endowed unusual selectivity in the catalyzed aldol reactions. Under optimized reaction conditions, high yields (up to >99%), good enantioselectivity (*ee* up to 85%) and high diastereoselectivity (*ds* up to 92:8) were obtained.

Introduction

Aldol reactions are one of the most important and investigated transformations in organic synthesis because a new C-C bond is formed from two carbonyl compounds.[1-3] This reaction generates products that contain a β-hydroxyl carbonyl skeleton, which is largely found in natural products with outstanding pharmacological activities. Furthermore, in the asymmetric version of the reaction, new stereogenic carbons can be generated with high levels of selectivity.[4-8] In biological systems, aldol reactions are promoted reversibly and stereoselectively by a group of enzymes known as fructose 1,6-bisphosphate aldolases, or simply aldolases.[1, 4, 9-11] Since its first use in 1971,[12-14] the amino acid *L*-proline has showed itself as an efficient and versatile organocatalyst for aldol reactions. Thus, the design and synthesis of highly efficient *L*-proline derivatives as organocatalysts is still of considerable interest.[15-17]

Many *L*-proline-catalyzed aldol reactions are usually carried out in polar organic solvents such as DMSO, DMF or chloroform. Therefore, one of the current challenges regarding aldol reactions is replacing organic solvents (often volatile, toxic and flammable) by environmentally friendly compounds that retain the effectiveness of product conversion and selectivity.[18-21] Water is therefore a very attractive solvent for promoting such reactions because, in addition to its obvious environmental benefit, the presence of hydrophobic interactions and mediation of hydrogen bonds could easily promote self-assembly of catalyst superstructures,[22-24] whereas the high surface tension and high dielectric constant of water potentially make reactions faster and more selective.[21, 25-28]

The use of *L*-proline-derived organocatalysts containing long hydrophobic chains has been reported as a strategy for performing aldol reactions in solutions mixing water and organic solvents.[29, 30] A proposition emerging from these studies is that amphiphilic species appear uniformly distributed at the water-oil interfaces of H_2O droplets dispersed in the organic phase[30] increasing the catalytic surface area. Furthermore, organization into two-dimensional chiral surfaces likely induces spatial ordering of reactants and enhances stereoselectivity. Interestingly, when the reaction does not favor emulsion production, satisfactory results are not observed due to the random distribution of the catalyst.[30] Compared to enzymes, supramolecular catalysts based on self-assembled peptides appear as a promising strategy since they are much more stable, cheaper and yet still allow catalyst sites to be mimicked. In addition, they exhibit polymorphic richness at the nanoscale and enable reactions in confined environments with diverse geometrical constraints.[31-34]

[a] *Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André 09210-580, Brazil.*
[b] *Instituto de Ciências Ambientais, Químicas e Farmacêuticas, Universidade Federal de São Paulo, Diadema 09972270, Brazil.*
[c] *Departamento de Biofísica, Universidade Federal de São Paulo, São Paulo 04023-062, Brazil.*
[d] *Department of Chemistry, University of Reading, Whiteknights, Reading RG6 6AD, United Kingdom.*
[e] *Department of Applied Physics, Aalto University School of Science, P. O. Box 15100, FI- 00076, Finland.*
† Corresponding author: wendel.alves@ufabc.edu.br
Electronic Supplementary Information (ESI) available: spectroscopic, scattering and molecular dynamics simulations data, and details on the aldol product characterization. See DOI: 10.1039/x0xx00000x

Recently, a few research groups have proposed the use of peptide-based supramolecular assemblies for mimicking catalytic activity of aldolases.[35, 36] Escuder *et al.*[24, 37] and Miravet *et al.*[38] have reported examples of aldol reactions catalyzed by *L*-proline-derived amphiphilic peptides, where the catalytic performance is activated by self-aggregation of the catalyst within hydrogelated matrices. In this same way, Zhao *et al.*[39] have shown that assembled prolinamide surfactants can enhance activity and stereoselectivity of aldol reactions in nonpolar solvents, like benzene. These results provided new insights into the role of the hydrophobic part of the proline-functionalized organocatalysts on the observed high reactivity and selectivity in water.

Inspired by these previous investigations, we describe here a study where the self-assembly of an *L*-proline-containing lipopeptide (Figure 1) was explored for promoting aldol reactions in excess water. The functional group of the lipopeptide was made up from a tripeptide PRW sequence (P: *L*-proline, R: *L*-arginine, W: *L*-tryptophan), which acts as a catalyst site for reactions. This tripeptide was lipidated to a C16:0 palmitoyl hydrocarbon chain at the carboxyl group of the tryptophan amino acid, leading to a surfactant-like design. Detailed investigations on the self-assembly of the PRW-C_{16} lipopeptide showed spontaneous ordering into spherical micelles with diameters around 6 nm at concentrations as low as 0.19 mmol L^{-1}. In the following, these structures were successfully applied as a supramolecular catalytic system for direct aldol reactions between cyclohexanone and *p*-nitrobenzaldehyde in excess water. Our results show, for the first time, the viability of *L*-proline amphiphilic derivative micelles as a catalyst with excellent conversion and stereoselectivity. In contrast, poor results were obtained in the absence of the lipopeptide assemblies. The excellent results provided by the self-assembled structures demonstrate the potential of lipopeptide assemblies to catalyze aldol reactions in an environmentally-friendly manner.

Experimental

Peptide samples

The lipopeptide was synthesized by CS Bio (Merlo Park, California) using TFA as a counter ion (catalog number CS12677, batch O051). The sequence had a tripeptide as a functional group with sequence PRW (P: *L*-proline, R: *L*-arginine and W: *L*-tryptophan) and it was lipidated by bonding a C16:0 fatty acid chain to the carboxylic acid portion of the tryptophan amino acid (henceforth, the lipopeptide will be referred to as PRW-C_{16}).

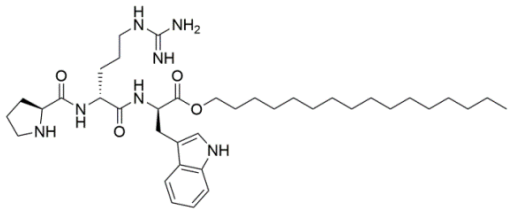


Figure 1. Structure of the lipopeptide PRW-C_{16}.

Purity was determined at 95.43% and the molar mass was measured at 681.60 Da (expected: 681.98). Aqueous solutions were prepared by dissolving PRW-C_{16} into ultrapure water (resistivity > 18 MΩ·cm at 25°C), obtained from a Mili-Q system. Concentrations varied from 4.5×10^{-3} to 15 mM and measurements performed with indicator paper revealed pH values between ~ 6 (higher lipopeptide concentrations) and ~ 7.5 throughout the concentration range. After mixing, samples were submitted to ultrasonication (50°C, 15 minutes) to aid solubilization. In the following, solutions were kept in the dark, at room temperature, for about 24 hours prior to spectroscopic assays. Samples used for scattering and microscopy assays were kept refrigerated for a period of 4 days before analysis.

Critical aggregation concentration

The critical aggregation concentration (CAC) of the lipopeptide in water was determined through fluorescence assays. Two methods were used for obtaining independent estimations on the CAC. The first strategy was based on the intrinsic fluorescence of the tryptophan side-chain.[40] Aqueous PRW-C_{16} solutions were irradiated with excitation wavelength λ_{exc} = 280 nm and emission spectra were probed in the range 300 nm ≤ λ_{em} ≤ 450 nm. In the second strategy, pyrene was used as an extrinsic fluorophore for assessing the formation of peptide aggregates. Peptide mixtures were prepared across the same concentration range described above into solutions containing pyrene at 1×10^{-4} wt%. Samples were then irradiated with λ_{exc} = 338 nm and pyrene emission spectra were registered in the range 350 nm ≤ λ_{em} ≤ 550 nm. The pyrene fluorescence is significantly enhanced in nonpolar environments;[41] thus, a sudden growth on emission intensity is observed upon self-assembly because chromophores remain trapped in the hydrophobic core of peptide nanostructures. Fluorescence data were recorded using a Carian Varian Ellipse spectrometer.

Cryo-TEM imaging

Cryogenic transmission electron microscopy was performed as reported in our previous work.[42-44] Briefly, imaging was conducted on a JEOL JEM-3200FSC microscope, using acceleration voltage of 300 kV. Slit width was kept at 20 eV and micrographs were recorded by a Gatan Ultrascan 4000 CCD camera. Droplets containing about 3 µl solutions were cast onto copper grids and then vitrified at – 180 °C in a 1:1 ethane:propane mixture.

Small-angle scattering assays

High-resolution small-angle X-rays scattering (SAXS) was performed on the SWING beamline of synchrotron SOLEIL (Saint Aubin, France). A HPLC-based bioSAXS setup was used to flow about 50 µl of sample across a 1 mm quartz capillary during data acquisition. 60 frames, 1 s each, were grabbed during the flow. Frames were compared and, if no radiation damage was detected, they were averaged and background subtracted. The beam energy was set at *E* = 12 KeV, the sample-to-detector distance was *SD* = 1526.4 mm and data

were recorded with an *Aviex* 1024 × 1024 detector (pixel size: 41.7 μm). The *q*-range was scanned in the interval 0.1 $nm^{-1} \leq q \leq 5$ nm^{-1}. Complementary small-angle neutron scattering (SANS) data were obtained on the time-of-flight LOQ beamline of ISIS spallation source (Didcot, UK). A 1 wt% solution containing PRW-C_{16} was prepared into D_2O and loaded into a Hellma banjo cell with pathlength *l* = 1 mm. A "white" neutron beam, with cross-section of 12 mm, irradiated the sample with wavelengths in the range 0.22 nm ≤ λ ≤ 1 nm. The sample-to-detector distance was *SD* ~ 4 m and data was registered in the interval 0.1 $nm^{-1} \leq q \leq 2.5$ nm^{-1}. All scattering experiments were conducted at 25 °C and scattering data treatment was carried out by model fitting with SASFit program.[45]

Molecular dynamics simulation setup

The PRW-C_{16} lipopeptide was modeled using the AMBER family of force fields. The amber LIPID14[46] has been used for the palmitoyl chain while the peptide part was described using the ff99SB-ILDN[47, 48] updated parameter set. We used the standard RESP[49] charge scheme to derive charges for the alkyl section using an acetyl cap. The ester bond connecting the peptide and alkyl chain employed the same parameters used in LIPID11[50] and LIPID14. The cyclohexanone molecule was modeled using the GAFF[51] parameter set and RESP charges. We employed the Hartree-Fock method and a 6-31G* basis set in all RESP charge derivations. Tip3P[52] water and Cl- counter ions were added to the system afterwards. Note that Cl- was used instead of TFA as the counter ions in our simulations for the sake of simplicity.

A typical spherical micelle initial condition for the simulation was constructed using packmol[53] from minimized PRW-C_{16} conformations forming a circa 60 Å object from 160 individual molecules. The reasoning for such aggregation number is given in the results section. Ions were added to the ensure charge neutrality using the xleap program from the AmberTools suite of codes. Three configurations were considered: configuration A had a PRW-C_{16} micelle in water while configurations B and C had a micelle in water/cyclohexanone mixtures using 575 cyclohexanone molecules. This number is close to the cyclohexanone saturation concentration in water of 0.86 M, being 0.55 M and 0.66 M respectively. The resulting concentration of PRW-C_{16} in these runs is roughly ten times the CAC value measured in our experiments (see results).

Simulation protocol

Two thousand steps of steepest descent minimization was followed by a short heating step of 140 ps to 300K using the Berendsen[54] thermostat on a NVT ensemble. In this initial run we used a 0.01 fs timestep and restrained all alkyl chains with a force constant of 10 $kcal \cdot Å^{-2}$. In the remainder of the simulations we employed an NPT ensemble using the Berendsen barostat with a 1.0 ps relaxation time and pressure equals to 1 ATM, the Langevin[55] thermostat with a 2 ps^{-1} collision frequency and a timestep of 0.02 fs. Shake[56] was used to constrain the hydrogen atoms to their initial positions. After the initial heating step, we performed another 300 ps of simulation keeping the alkyl chains restrained followed by unrestrained dynamics. All production data was collected after 5ns of equilibration time. All molecular dynamics simulations were performed using the AMBER14 package.[57-59] Figures were rendered using VMD.[60]

Reagents and aldol reactions

All reagents had analytical grade and, unless indicated to the contrary, were used without further treatment. Cyclohexanone was distillated under inert nitrogen atmosphere prior to experiments. Aldol reactions were monitored through thin layer chromatography (TLC), using 200 μm silica gel plates (Supelco) inspected under UV light or revealed trough vanillin method. High resolution mass spectroscopy (HRMS) was performed on a MicrOTOF-QII (Bruker) spectrometer. High performance liquid chromatography (HPLC) was carried using a Shimadzu LC-10AD apparatus equipped with an UV/Vis detector Shimadzu SPD-10AV and an integrator Shimadzu C-R8A, using a Chiralpak AD-H column (hexane/2-propanol = 90/10), 30 °C, 254 nm, 1.0 mL/min. NMR spectra were obtained with a Bruker DPX-300 instrument (Bruker BioSpin GmbH), operating at 300 MHz. Chemical shifts (δ) are shown in parts per million (ppm), whereas coupling constants (*J*) are exhibited in Hz. Data were normalized by the peak of the corresponding solvent used during analysis.

General procedure for aldol reactions in water. Cyclohexanone (1.7 mmol, 178 μL) was added to the catalyst (7.3 μmol, 5.0 mg) in water (178 μL). Then, *p*-nitrobenzaldehyde (147.0 μmol, 22.0 mg) was added to the mixture. After being stirred at room temperature for 2 days, the mixture was extracted with ethyl acetate. The organic layer was concentrated to give crude aldol product.

^{1}H NMR (300 MHz, $CDCl_3$): δ 8.22-8.18 (m, 2H, ArH); 7.51-7.47 (m, 2H, ArH); 5.47 (br s, 1H, C*H*OH-*syn*); 4.89 (dd, *J* = 7.5 Hz, 3.0 Hz, 1H, C*H*OH-*anti*); 4.08 (d, *J* = 3.0 Hz, 1H, CHO*H*-*anti*); 3.21 (d, *J* = 3.0 Hz, 1H, CHO*H*-*syn*); 2.66-2.30 (m, 1H, C*H*CHOH); 2.66-2.30 (m, 2H, $CH_2C(O)$); 2.16-1.24 (m, 6H, chex-H).

^{13}C NMR (50 MHz, $CDCl_3$): δ 214.9 (C=O); 148.5 (ArC); 147.7 (ArC); 128.0 (2 x ArCH); 123.7 (2 x ArCH); 74.1 (CHOH); 57.3 (*C*HCHOH); 42.8 (CH_2); 30.9 (CH_2); 27.8 (CH_2); 24.8 (CH_2).

Results and discussion

Structural determinations

We first analyzed the self-assembly behavior of the PRW-C_{16} lipopeptide by determining its critical aggregation concentration (CAC) in water through fluorescence assays. Our first attempt to estimate the CAC was made by assessing the intrinsic fluorescence of tryptophan available at the peptide head of our lipopeptide. The presence of an indole chromophore in the tryptophan side-chain enabled direct measurements of fluorescence and, since this amino acid is widely-known to be sensitive to local polarity,[40, 61] it was here considered as a natural candidate for use as a probe in aggregation studies. In Figure 2A, we show fluorescence data

from PRW-C_{16} solutions illuminated with wavelength λ_{exc} = 280 nm, corresponding to the absorbance maximum of tryptophan.[62] The emission range exhibits a broad peak centered at $\lambda_{em} \approx$ 357 nm, whose intensity appears to be strongly correlated to peptide concentration (see inset in Fig. 2A). By plotting emission intensity as a function of log [peptide concentration], one clearly observes that fluorescence is characterized by two distinct regimes. In the first domain, associated to dilute solutions, intensity increases regularly upon peptide concentration in the mixture. On the other hand, the concentrated regime is typified by a systematic decay of fluorescence, which can be assigned to the inner-filter effect due to the tryptophan residue absorption or peptide aggregation in solution (or both).[63] Unfortunately, the number of underlying parameters contributing for the emission behavior is large and correction of filter effect across the concentration range is very complex; however, by investigating absorbance through UV-Vis assays (SI file, Fig. S2), one observes saturation for peptide concentrations above ~ 0.1 mM, close to the separation between the emission domains. This finding suggests that, above this concentration, saturation of detector leads to the strong decrease observed in emission intensities.

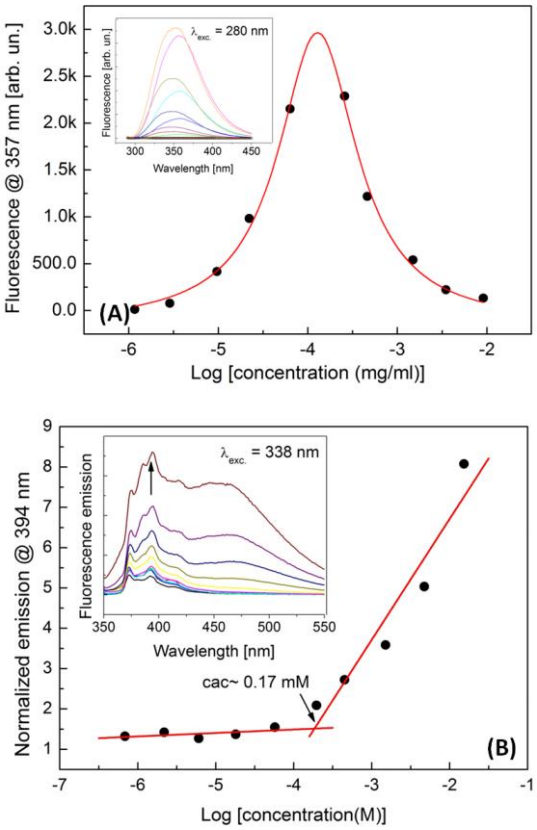


Figure 2. Fluorescence data used for estimating the CAC of the lipopeptide. (A) Intensity of tryptophan fluorescence as a function of log [peptide concentration]. The red line is a Lorentzian fit whose turning point has been ascribed to appearance of aggregates in solution (see text for details). Inset: full-range fluorescence spectra from PRW-C_{16} solutions. (B) Pyrene emission as function of log [peptide concentration]. The intersection between straight lines fitted to data from different fluorescence regimes gives estimation of the CAC. Inset: fluorescence curves showing characteristic peaks associated to pyrene fluorescence. Intensities have been normalized to pyrene emission in a peptide-free solution.

To provide a more reliable estimation on CAC, we performed fluorescence experiments using solutions containing pyrene, a hydrophobic fluorophore whose quantum yield is strongly dependent on polarity of local environment;[41, 42] in this case, when aggregates appear in solution, the hydrophobic cores of these particles provide nonpolar clusters where the fluorophore is trapped and its emission is raised. In Figure 2B the intensity of pyrene emission at λ_{em} = 394 nm is shown as a function of log [peptide concentration]. Again, two regimes are clearly observed: in the dilute one, nearly-constant fluorescence is found, indicating that lipopeptides remain in the monomeric form and no hydrophobic clusters are created in solution. On the other hand, in the concentrated range, one observes sharp fluorescence growth and a linear relationship between emission intensity and log [peptide concentration]. Extrapolations from linear fits performed across the two regimes reveal a crossover at 0.17 mmol L^{-1}, which is associated to the CAC. Also, this value is close to the the separation between the emission domains found in Fig. 2A, suggesting that saturation in tryptophan absorbance is also related to appearance of aggregates in solution.

Once the CAC was estimated, a sophisticated combination of *in-situ* techniques was used to characterize the morphology and structure of self-assemblies in detail using experimental and theoretical techniques. These assays included *cryo*-TEM imaging from vitrified specimens and scattering experiments performed in solutions similar to those where aldol reactions were further investigated. In Figure 3, we show *cryo*-TEM micrographs from mixtures containing the lipopeptide dissolved in water at a concentration of 15 mmol L^{-1}, two orders of magnitude above the CAC and representative of catalyst concentrations used in the reactions below. One observes that samples are mostly populated by nanostructures with globular morphology and diameters between ≈ 4 nm and ≈ 9 nm. A few structures with bigger volumes (sizes ≈ 15 nm) and indefinite shapes – see detail in Figure 3C – are also found in the medium. The hydrodynamic radius of particles in 15 mM PRW-C_{16} solutions was estimated at 3.4 nm, with polydispersity of about 12%, from dynamic light scattering assays (SI file).

Information on the inner structure of the self-assemblies was obtained from high-resolution SAXS. In Figure 4A, SAXS curves from solutions prepared at different PRW-C_{16} concentrations show scattering profiles characterized by oscillations across the intermediate-to-high *q*-range. At moderate *q*-values, the outline is dominated by a broad band centered at $q \sim 1.7$ nm^{-1}, flanked by pronounced minima at $q \sim 1$ nm^{-1} and $q \sim 2.5$ nm^{-1}. This shape of the data is coherent with globular structures and the presence of sharp minima along with diffuse maxima, is also consistent with high-symmetry particles suggesting high homogeneity of sizes and shapes in the solutions.[64, 65] Despite increasing concentration leads to better signal-to-noise ratios, general features of data at higher *q*-values remain unchanged and support globular structures as dominant polymorphs in all solutions investigated.

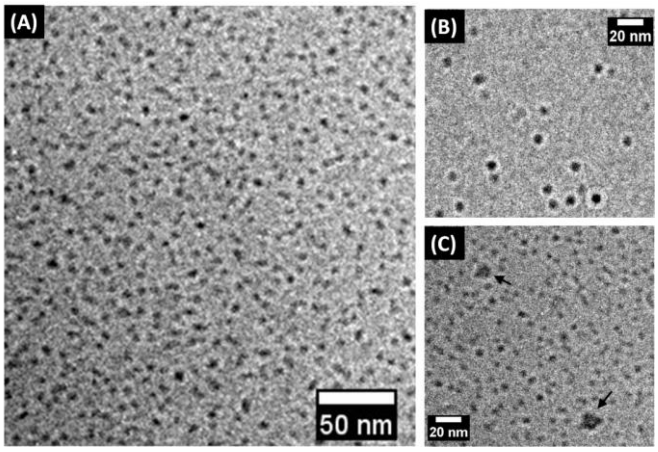


Figure 3. *Cryo*-TEM micrographs from PRW-C_{16} solutions prepared at a concentration of 15 mM into water. (A) Samples appear populated by extensive numbers of globular structures, most of them with diameters between ≈ 4 nm and ≈ 9 nm. (B) Magnified image showing round micelles in detail. (C) Aggregates of micelles are highlighted by arrows.

Conversely, some important differences are observed in the low *q*-range and are evidence for different interacting regimes as a function of peptide concentration. In the sample containing the lowest amount of lipopeptide, 1.5 mM, a low *q* upturn is found which has been ascribed to coexistence of bigger structures in solution.[64] At higher concentrations, 7.3 mM and 15 mM, a downturn of scattered intensities is observed and appearance of a Bragg peak at small angle accounts for repulsive interactions between micelles.[64, 66]

To quantitatively describe these features and extract accurate structural parameters, we performed model fitting of the data using the SASFit program.[45] SAXS intensities in the intermediate-to-high *q*-range have been properly fitted according to a spherical-shell form factor, convoluted with a Gaussian radii distribution, in accordance with the presence of micelles in the mixtures (see SI file for equations and further details). Values arising from this procedure are listed in Table 1 and they reveal that the inner part of micelles have radii between 1.66 and 1.81 nm, with polydispersity of about 8%. These values are in good agreement with the length estimated for C_{16} alkyl chains, $l_c \approx 1.7$ nm, and comply with previous work on micelles formed from self-assembly of palmitoyl lipopeptides.[67, 68] Furthermore, relative negative electronic contrasts of the cores are consistent with that expected for hydrophobic media, being also found in lipid chains.[68, 69]

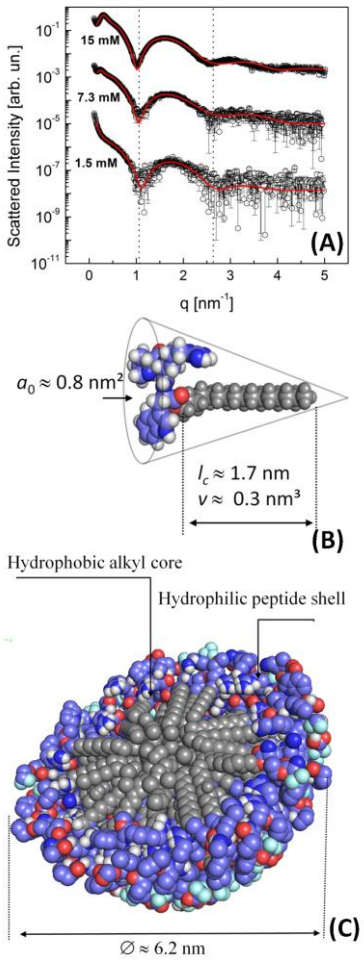


Figure 4. (A) SAXS data from lipopetide solutions at concentrations indicated. The intermediate-to-high *q*-range has been fitted according to a spherical-shell model leading to structural parameters described in Table 1. (B) 3D model of the PRW-C_{16} monomer, after energy minization (see text for details). Estimations of the geometric characteristics of the molecule revealed a packing parameter ≈ 0.2. (C) Micelle model obtained from geometrical constraints and association numbers revealed by SAXS data. The model was built up using Packmol package and representations were rendered using Pymol.

Still looking at Table 1, we see that thicknesses of spherical shells disclosed by SAXS are between 1.30 and 1.38 nm, with positive contrasts agreeing with their polar nature. These values are consistent with the size of a tripeptide group (3 × 0.34 = 1.02 nm) plus a TFA counterion shell.[70] Putting these findings together, we found that micelles have diameters of about 6 nm, in close agreement with DLS data and consistent with *cryo*-TEM imaging (Figure 3).

Table 1. Structural parameters obtained from model fitting of SAXS data (R: inner radius, σ_R: standard deviation of R; ΔS: shell thickness, η_R: core electronic contrast, η_S: shell electronic contrast; R_g: gyration radius; *D*: fractal dimension; RHS: hard sphere repulsion radius; F_p: volume fraction; *V*: micelle volume; *n*: association number).

Conc. (mM)	Spherical shell form factor					Mass fractal form factor		Hard sphere structure factor		Calculated parameters	
	R (nm)	σ_R (nm)	ΔS (nm)	η_R	η_S	R_g (nm)	*D*	RHS (nm)	F_p	V (nm³)	*n*
1.5	1.66	0.12	1.30	-1.67	1.61	13.6	2.54	---	---	109	145
7.3	1.76	0.14	1.33	-3.47	3.63	---	---	11.8	0.18	123	164
15	1.81	0.12	1.38	-4.95	4.95	---	---	9.5	0.21	136	181

Self-assembly into spherical micelles is consistent with the geometry of the PRW-C_{16} monomer, which presents a shape similar to an inverted cone (Figure 4B).

In line with previous description of SAXS data, fitting at very low *q* values required different strategies according to peptide concentration used in the formulation. To fit this range for samples containing the lowest peptide concentration, 1.5 mmol L^{-1}, we introduced a mass fractal form factor with a radius of gyration of R_g = 13.6 nm and fractal dimension *D* = 2.56 (see SI file for details). These measurements of size and dimension are consistent with a few aggregates found in *cryo*-TEM images (Figure 3C), which were around 30 nm in extent and have an irregular shape. In samples prepared at higher peptide concentrations, 7.3 mmol L^{-1} and 15 mmol L^{-1}, the small-angle data exhibits a Bragg peak arising from inter-particle repulsion. To fit this region, a hard sphere structure factor[64, 66, 71] was included in the model revealing repulsion radii of 11.8 and 9.5 nm, respectively. The volume fractions associated to the effective radii of the particles were found around 0.2. Complementary neutron scattering data performed on samples containing 15 mM PRW-C_{16} dissolved in D_2O – and thus exhibiting scattering contrasts very different from those observed in X-ray measurements – also indicated core-shell structures with dimensions compatible with those exhibited in Table 1 and the presence of inter-particle repulsion (see SI file). We ascribe the emergence of this repulsive regime at higher concentrations to increasing charge density at the micelle surface. Indeed, arginine groups in functional group of lipopeptides remain protonated in solution and the appearance of inter-micelle repulsion agrees with the growth of association number. Of course, such an increase in the number of monomers per micelle also means that the density of proline catalyst sites also grows with concentration, with strong repercussions on aldol reaction pathway as discussed below.

Molecular Dynamics Simulations

To investigate the micelle structure in more detail we performed molecular dynamics simulations. An aggregation number range was estimated from molecular dynamics simulations of a single PRW-C_{16} molecule in water and SAXS data. The molecule surface area computed using the Connolly[72] approach gave an average value of 0.8 nm^2 for the whole system and 0.470 nm^2 for the hydrophilic head. The solvent exclusion volume computed from the solvent radial distribution function gave values in the 0.8 to 0.9 nm^3 range for the whole system and 0.4 nm^3 for the hydrophobic tail. Simple geometrical considerations using an alkyl chain length of 1.9 nm (l_c) and these estimated values allowed us to calculate the packing parameter of the molecule at $p = v / (l_c \times a_0) \approx 0.18$, much lower than 0.33 which is the critical *p* expected for self-assembly of surfactant-like moieties into spherical micelles.[73-75] From the solvent exclusion volume it is possible estimate that micelles have association numbers between $n \approx 121$ and $n \approx 170$ across the concentration range investigated. Information derived above for both geometric constraints and association numbers were inputted in the program *Packmol*[53] for obtaining an approximate configuration for PRW-C_{16} micelles that was used as initial condition in our simulations as described in the simulation setup section.

The micelle simulation in water only (simulation A) over 50 ns revealed a considerable variability in its shape. Cluster analysis of the structures using a hierarchical clustering method[76] having the C1-C15 distance as a metric (excluding intra-molecular pairs) showed that only two important topologies with approximately equal weights dominate 68% of all conformations. The C1-C15 (carbon one to carbon 15 from the alkyl chain) distance was chosen as it represents the average radius of the inner hydrophobic part of the micelle when intra-molecular pairs are excluded. Figure 5 shows the average cluster configurations that resulted from this analysis: panel A shows a single more cohesive but irregular structure, while panel B shows two micellar bodies with slightly different sizes. The second configuration is representative of the later stages of the dynamics and formed spontaneously. The larger single structure shown in panel A is flatter in one dimension while the smaller structures are more spherical in shape as seen on *cryo*-TEM images.

Interdigitation of the alkyl chains associated with the monomer PRW-C_{16} charge character resulted in a rather irregular micelle surface with increased surface area, which can be important to catalysis. Figure 6 shows the micelle using the surface representation for the amino acid moieties. As expected, there is a large concentration of counter ions at the micelle surface, with ions frequently being present in pockets closer to arginine residues.

Radial distribution functions (RDFs or g(r)) using the C15 alkyl carbon atom (C15 meaning carbon 15 from the alkyl residue) as a reference revealed several interesting micellar structural features (Figure S6). A value of g(r) greater than one means a probability of finding a pair higher than in bulk (statistically speaking) while a value less than one means a lower than in bulk probability. The RDF between the C15 alkyl carbon and water oxygen atoms revealed a drop in water density starting at about 40 Å from the hydrophobic core.

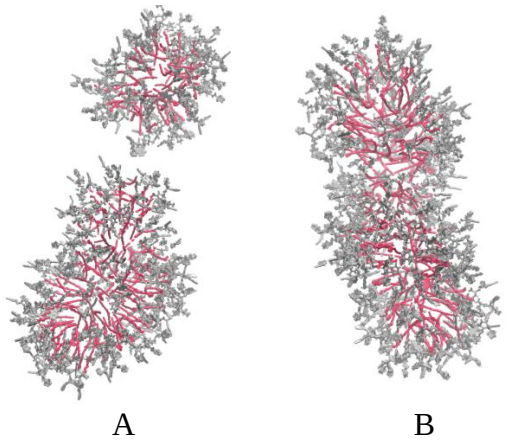


Figure 5. Average representative cluster configurations from 50 ns simulations. The amino acids are represented in gray while the alkyl chains are represented in red.

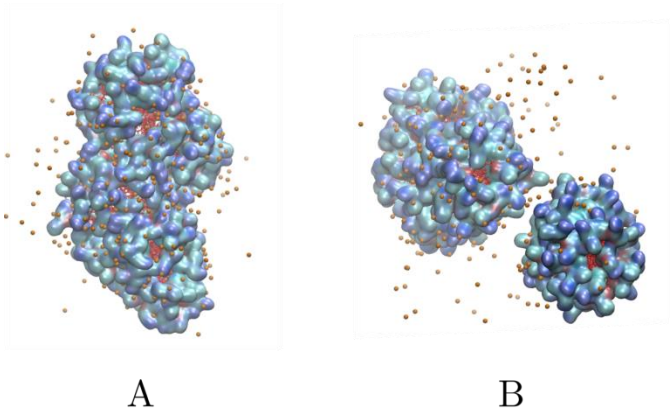


Figure 6. Micelle structures after 15 ns (panel A) and 45 ns (panel B) of simulation time for the water only simulation (simulation A). The amino acids are rendered using a surface representation, the inner alkyl core using a bond representation in red and chlorine ions are rendered in yellow.

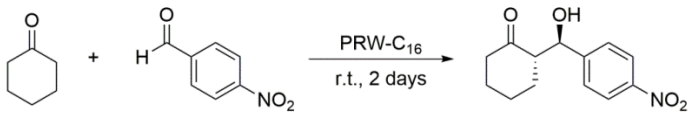
Meanwhile the C1-C15 g(r) peaks at 1.4 nm, defining a hydrophobic core domain with C15-Wat(O) g(r) having only 1/3 of its bulk value at that distance. The RDFs also revealed that the arginine and proline residues remain at about the same distance from the core, with g(r) for the C15-Arg(NE) and C15-Pro(N) peaking at about 2.4 nm. The g(r) for C15-Trp(NE) distribution peaks at 1.8 nm, corroborating a picture that the tryptophan ring interacts with the hydrophobic core of the micelle. These RDFs are consistent with an inverted cone configuration adopted by the PRW-C_{16} lipopeptides during the dynamics. They are also consistent with the presence of a well-defined hydrophobic core composed of alkyl chains and tryptophan with a dimension in the 1.4-1.8 nm range. These results are consistent with SAXS results for the inner radius and with an ordered core shell structure present in the micelle. Results for the water/cyclohexanone mixture simulations, namely simulations B and C, indicate a similar atomistic arrangement in solution as revealed by very similar RDFs to the water only case (Figure S6). The overall structure of the simulated object, however, changed considerably. The addition of cyclohexanone lead to much more irregular shaped micelles, with smaller aggregates separating from the larger bodies. Simulations B and C differ in their initial condition, with simulation B having all cyclohexanone molecules on the outside of the initial micelle while simulation while C had 50 cyclohexanone molecules at the center. Despite the differences in initial conditions, both the final structures and RDFs were nearly identical for the two simulations. Most relevant to the results at hand, the RDFs indicate a similar ordering in solution as in the water only case, with one important difference: there is a strong preference for cyclohexanone to be near or at the hydrophobic core of the micellar structure as shown by Figure S7 and by the g(r) presented in Figure S8. This generates a higher concentration of cyclohexanone close to catalytic *L*-proline residues when compared to bulk values, possibly allowing to a more efficient catalysis. In fact, the RDFs presented in Figure S7 for proline and ciclohexanone that have the same carbon 15 from the alkyl chain as a reference show considerable overlap, suggesting a kinetic enhancement allowed by the micellar structure. In summary, our simulations suggest that the micellar structure provides a unique arrangement for the aldol reaction, with higher local concentration of cyclohexanone close to the catalytic proline in a less hydrophilic environment. All presented analyses were made using the AmberTools 15[57] suite of codes.

Aldol Reactions

The catalytic activity of the PRW-C_{16} lipopeptide for the direct asymmetric aldol reaction was investigated by performing a model reaction using *p*-nitrobenzaldehyde and cyclohexanone at room temperature, in the presence and absence of water. The results are summarized in Table 2.

The reactions performed in aqueous media provided products in excellent yields with good to excellent diastereoselectivities and enantioselectivities for the *anti* aldol product. The best observed catalytic efficiency is shown in entry 1 with full conversion to *anti*-diastereomeric product, *anti*/*syn* diastereomeric ratio of 92:8 and enantiomeric excess (*ee*) of 85%. By decreasing the catalyst loading from 5 to 1 mol%, the reaction exhibited lower diastereoselectivity (entry 4, 87:13 *ds*).

Table 2. Direct asymmetric aldol reaction of *p*-nitrobenzaldehyde and cyclohexanone catalyzed by PRW-C_{16} lipopeptide.[a]



anti

Entry	Loading (mol%)	Cyclohexanone (eq.)	H_2O[b] (eq.)	Conv.[c] (%)	*anti*/*syn*[c]	*ee*[d] (%)
1	5	11.7	0.5	>99	92:8	85
2	5	11.7	1.0	>99	91:9	71
3	5	11.7	2.0	>99	91:9	84
4	1	11.7	1.0	>99	87:13	85
5	1	15.0	-[e]	36	81:19	63
6	5	15.0	-[e]	94	86:14	73

[a] The reactions were performed at room temperature, for 2 days, with vigorous stirring; [b] Water excess towards cyclohexanone (v/v); [c] The diastereomeric ratios and conversion were determined by ^{1}H NMR analysis of the crude mixture; [d] Determined by chiral-phase HPLC analysis of the *anti* isomer; [e] Neat (no water addition).

Different results were obtained when water was removed from the system and the cyclohexanone volume was increased. With 1 mol% loading of catalyst, both the conversion and stereoselectivity were reduced (entry 5, 36% conversion, 81:19 *ds*, 63% *ee*). Using 5 mol% loading of catalyst, there was an improvement in conversion but the stereoselectivity remained lower in comparison with the tests promoted in water (entry 6, 94% conversion, 86:14 *ds*, 73% *ee*).

The amphiphilic peptide showed better catalytic efficiency in aqueous media, by comparison with its performance in organic media (neat cyclohexanone, entries 5 and 6), since the micelle environment seems to provide an environment that enhances catalysis. Therefore, water is responsible for optimizing the catalyst and substrates approaching by forming a self-assembling structure.

The observed differences in results point to the existence of a structural feature directly involved in the reaction mechanism. It seems that the micelles revealed by *cryo*-TEM and SAXS data and molecular dynamics simulations are responsible for the catalysis promotion by the amphiphilic peptide. The effect of phase transfer may contribute to enhancement of stereoselectivity in aldol coupling reactions. Thus, the higher activity of the lipopeptide in water is mainly due to the unique environment around the proline residue that is similar to the one present in a natural enzyme.

Corroborating the structural data previously presented for the lipopeptide with the model aldol reaction results, we propose based on molecular dynamics results that the micellar aggregate formed in water concentrates the substrates in its inner region, keeping the transition states mostly isolated from water, affording, thereby, aldol products in excellent yields with high diastereo- and enantioselectivities. In fact, macroscopically, the reaction presents itself as an emulsion.

When it is performed in the absence of water (i.e., neat cyclohexanone), the aldol reaction shows reduced yields and stereoselectivity due to a lack of an aggregate that provides the catalyst and substrates meeting. So, since the catalytic performance was only activated when the catalyst self-in its assembled form, it was shown that the supramolecular chemistry has a vital effect in aldol reaction catalysis.

Conclusions

Herein we have succeeded in presenting a new amphiphilic peptide whose structural features and catalytic efficiency in the direct asymmetric aldol reaction between cyclohexanone and *p*-nitrobenzaldehyde were evaluated. Critical aggregation concentration was obtained by using a pyrene as a probe, provided a detailed picture of the aggregation process at around CAC = 0.17 ± 0.01 mmol L^{-1}. *Cryo*-TEM micrographs showed nanostructures with globular morphology and diameters between ≈ 4 nm and ≈ 9 nm. High-resolution SAXS data have provided information about the inner structure of the self-assemblies, which indicated spherical micelles with good homogeneity of size and shape. We found that PRW-C_{16} micelles have diameters of about 6 nm, consistently from SAXS data, *cryo*-TEM imaging, and molecular dynamics simulations. Based on its amphiphilic nature, the peptide was found to be an excellent catalyst for aldol reactions performed in water under the optimal reaction conditions, where excellent diastereoselectivities (92:8) and enantioselectivities (85%) with very good conversions (>99%) were obtained by using as little as 5 mol% of catalyst. It is noteworthy that we did not use any additives in the experiments. A suggestive agreement between reaction results and concentration effects computed by molecular dynamics simulations of the cyclohexanone residues is noted.

Acknowledgements

This work was supported by FAPESP (grant no. 2015/24018-1) and CNPq (grant no. 302923/2015-2, 400239/2014-0). INCT in Bioanalytics (FAPESP grant no. 08/57805-2 and CNPq grant no. 573672/2008-3) is kindly acknowledged for the grants. B.M.S. acknowledges CAPES for a doctoral fellowship and E.R.S. is grateful to FAPESP for post-doctoral grants (Proc. 2013/12674-6). I.W.H. is recipient of a Royal Society-Wolfson Research Merit Award and this work was supported by EPSRC Platform Grant "Nanostructured Polymeric Materials for Healthcare" (EP/L020599/1). Synchrotron SOLEIL is kindly acknowledged for awarding beamtime for SAXS experiments (proposal 20140919). E.R.S. and I.W.H. thank Steven Kirkham and Dr. Javier Perez for assistance during SAXS experiments at SWING beamline. SANS measurements were performed at ISIS under proposal RB1510029 and Dr. Stephen King is kindly acknowledged for help during this beamtime. Prof. Alessandro Rodrigues is kindly acknowledged for his help and support in HPLC experiments. The anonymous Referee is also acknowledged for leading us to consider the presence of inner-filter effect and correcting discussion on tryptophan fluorescence assays.

References

1. R. Mahrwald, ed., *Modern Aldol Reactions - Vol. 1*, Wiley-VCH, Weinheim, 2004.
2. R. Mahrwald, ed., *Modern Aldol Reactions - Vol. 2*, Wiley-VCH, Weinheim, 2004.
3. R. Mahrwald, *Aldol Reactions*, Springer, New York, 2009.
4. B. M. Trost and C. S. Brindle, *Chem. Soc. Rev.*, 2010, **39**, 1600-1632.
5. V. Bisai, A. Bisai and V. K. Singh, *Tetrahedron*, 2012, **68**, 4541-4580.
6. L. C. Dias and A. M. Aguilar, *Chem. Soc. Rev.*, 2008, **37**, 451-469.
7. L. C. Dias and A. M. Aguilar, *Quim. Nova*, 2007, **30**, 2007-2015.
8. L. C. Dias, E. C. Lucca Jr, M. A. B. Ferreira and E. C. Polo, *J. Braz. Chem. Soc.*, 2012, **23**, 2137-2158.
9. T. D. Machajewski and C.-H. Wong, *Angew. Chem. Int. Ed.*, 2000, **39**, 1352-1374.

10. P. Clapés and X. Garrabou, *Adv. Synth. Catal.*, 2011, **353**, 2263-2283.
11. S. J. Cooper, G. A. Leonard, S. M. McSweeney, A. W. Thompson, J. H. Naismith, S. Qamar, A. Plater, A. Berry and W. N. Hunter, *Structure*, 1996, **4**, 1303-1315.
12. U. Eder, G. Sauer and R. Wiechert, *Angew. Chem. Int. Ed.*, 1971, **10**, 496-497.
13. Z. G. Hajos and D. R. Parrish, *J. Org. Chem.*, 1974, **39**, 1615-1621.
14. B. List, R. A. Lerner and C. F. Barbas, III, *J. Am. Chem. Soc.*, 2000, **122**, 2395-2396.
15. J. Alemán and S. Cabrera, *Chem. Soc. Rev.*, 2013, **42**, 774-793.
16. J. Mlynarski and S. Bas, *Chem. Soc. Rev.*, 2014, **43**, 577-587.
17. U. Scheffler and R. Mahrwald, *Chem. Eur. J.*, 2013, **19**, 14346-14396.
18. F. M. Kerton and R. Marriott, *Alternative Solvents for Green Chemistry*, RSC Publishing, Cambridge, 2009.
19. R. A. Sheldon, I. W. C. E. Arends and U. Hanefeld, *Green Chemistry and Catalysis*, Wiley-VCH, Weinheim, 2007.
20. U. M. Lindström, *Chem. Rev.*, 2002, **102**, 2751-2772.
21. P. T. Anastas, in *Green Solvents - Vol. 5: Reactions in Water*, ed. C.-J. Li, Wiley-VCH, Weinheim, 2010.
22. M. Raynal, P. Ballester, A. Vidal-Ferran and P. W. N. M. van Leeuwen, *Chem. Soc. Rev.*, 2014, **43**, 1660-1733.
23. A. Sorrenti, O. Illa and R. M. Ortuño, *Chem. Soc. Rev.*, 2013, **42**, 8200-8219.
24. C. Berdugo, J. F. Miravet and B. Escuder, *Chem. Commun.*, 2013, **49**, 10608-10610.
25. J. G. Hernández and E. Juaristi, *Chem. Commun.*, 2012, **48**, 5396.
26. Y. Jung and R. A. Marcus, *J. Am. Chem. Soc.*, 2007, **129**, 5492-5502.
27. S. Bhowmick and K. C. Bhowmick, *Tetrahedron: Asymmetry*, 2011, **22**, 1945-1979.
28. R. Ballini, ed., *Eco-Friendly Synthesis of Fine Chemicals*, RSC Publishing, Cambridge, 2009.
29. Y.-Q. Fu, Y.-J. An, W.-M. Liu, Z.-C. Li, G. Zhang and J.-C. Tao, *Catal. Lett.*, 2008, **124**, 397-404.
30. L. Zhong, Q. Gao, J. Gao, J. Xiao and C. Li, *J. Catal.*, 2007, **250**, 360-364.
31. Z. Huang, S. Guan, Y. Wang, G. Shi, L. Cao, Y. Gao, Z. Dong and J. Xu, *J. Mater. Chem. B*, 2013, **1**, 2297-2304.
32. M. Reches and E. Gazit, *Nano Lett.*, 2004, **4**, 581-585.
33. E. T. Pashuck, H. Cui and S. I. Stupp, *J. Am. Chem. Soc.*, 2010, **132**, 6041-6046.
34. S. Kwon, H. S. Shin, J. Gong, J.-H. Eom, A. Jeon, S. H. Yoo, I. S. Chung, S. J. Cho and H.-S. Lee, *J. Am. Chem. Soc.*, 2011, **133**, 17618-17621.
35. M. Tena-Solsona, J. Nanda, S. Díaz-Oltra, A. Chotera, G. Ashkenasy and B. Escuder, *Chem. Eur. J.*, 2016, **22**, 6687-6694.
36. F. Rodríguez-Llansola, J. F. Miravet and B. Escuder, *Chem. Commun.*, 2009, 7303-7305.
37. S. Díaz-Oltra, C. Berdugo, J. F. Miravet and B. Escuder, *New J. Chem.*, 2015, **39**, 3785-3791.
38. C. Berdugo, B. Escuder and J. F. Miravet, *Org. Biomol. Chem.*, 2015, **13**, 592-600.
39. P. R. Arivalagan and Y. Zhao, *Org. Biomol. Chem.*, 2015, **13**, 770-775.
40. J. T. Vivian and P. R. Callis, *Biophys. J.*, 2001, **80**, 2093-2109.
41. F. M. Winnik, *Chem. Rev.*, 1993, **93**, 587-614.
42. C. C. Decandio, E. R. Silva, I. W. Hamley, V. Castelletto, M. S. Liberato, V. X. Oliveira, C. L. P. Oliveira and W. A. Alves, *Langmuir*, 2015, **31**, 4513-4523.
43. E. R. Silva, W. A. Alves, V. Castelletto, M. Reza, J. Ruokolainen, R. Hussain and I. W. Hamley, *Chem. Commun.*, 2015, **51**, 11634-11637.
44. E. R. Silva, M. N. M. Walter, M. Reza, V. Castelletto, J. Ruokolainen, C. J. Connon, W. A. Alves and I. W. Hamley, *Biomacromolecules*, 2015, **16**, 3180-3190.
45. I. Bressler, J. Kohlbrecher and A. F. Thünemann, *J. Appl. Crystallogr.*, 2015, **48**, 1587-1598.
46. C. J. Dickson, B. D. Madej, Å. A. Skjevik, R. M. Betz, K. Teigen, I. R. Gould and R. C. Walker, *J. Chem. Theory Comput.*, 2014, **10**, 865-879.
47. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins Struct. Funct. Bioinf.*, 2006, **65**, 712-725.
48. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins Struct. Funct. Bioinf.*, 2010, **78**, 1950-1958.
49. C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269-10280.
50. Å. A. Skjevik, B. D. Madej, R. C. Walker and K. Teigen, *J. Phys. Chem. B*, 2012, **116**, 11124-11136.
51. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
52. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
53. L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, *J. Comput. Chem.*, 2009, **30**, 2157-2164.
54. H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684-3690.
55. R. J. Loncharich, B. R. Brooks and R. W. Pastor, *Biopolymers*, 1992, **32**, 523-535.
56. J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
57. D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham, III, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu and P. A. Kollman, *AMBER 14*, University of California, San Francisco, 2014.
58. S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374-380.
59. R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.
60. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33-38.

61. L. A. Munishkina and A. L. Fink, *Biochim. Biophys. Acta, Biomembr.*, 2007, **1768**, 1862-1885.
62. H. Edelhoch, *Biochemistry*, 1967, **6**, 1948-1954.
63. M. Kubista, R. Sjöback, S. Eriksson and Bo Albinsson, *Analyst*, 1994, **119**, 417-419.
64. T. Zemb and P. Lindner, *Neutrons, X-rays and Light: Scattering Methods Applied to Soft Condensed Matter*, Elsevier, Amsterdam, 2002.
65. M. H. Koch, P. Vachette and D. I. Svergun, *Q. Rev. Biophys.*, 2003, **36**, 147-227.
66. J. S. Pedersen, *Adv. Colloid Interface Sci.*, 1997, **70**, 171-210.
67. A. Dehsorkhi, R. M. Gouveia, A. M. Smith, I. W. Hamley, V. Castelletto, C. J. Connon, M. Reza and J. Ruokolainen, *Soft Matter*, 2015, **11**, 3115-3124.
68. J. F. Miravet, B. Escuder, M. D. Segarra-Maset, M. Tena-Solsona, I. W. Hamley, A. Dehsorkhi and V. Castelletto, *Soft Matter*, 2013, **9**, 3558-3564.
69. G. Pabst, M. Rappolt, H. Amenitsch and P. Laggner, *Phys. Rev. E: Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, 2000, **62**, 4000-4009.
70. R. B. Freedman, in *Trends in Biochemical Sciences*, eds. T. E. Creighton and W. H. Freeman, Cell Press, Cambridge, 1983.
71. C. G. De Kruif, W. J. Briels, R. P. May and A. Vrij, *Langmuir*, 1988, **4**, 668-676.
72. M. L. Connolly, *J. Appl. Crystallogr.*, 1983, **16**, 548-558.
73. J. Israelachvilli, *Intermolecular and surface forces*, Academic Press Limited, London, 1998.
74. I. W. Hamley, *Soft Matter*, 2011, **7**, 4122-4138.
75. S. G. Zhang, *Acc. Chem. Res.*, 2012, **45**, 2142-2150.
76. J. Shao, S. W. Tanner, N. Thompson and T. E. Cheatham, *J. Chem. Theory Comput.*, 2007, **3**, 2312-2334.