Soft Matter

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/softmatter

Multi-orthogonal folding of single polymer chains into soft nanoparticles.

Federica Lo Verso,[*a,b] **José A. Pomposo,**[b,c,d,e] **J. Colmenero,**[a,b,c,d] **and Angel J. Moreno** [b,c]

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

First published on the web Xth XXXXXXXXXX 200X

DOI: 10.1039/b000000x

Efficient folding of single polymer chains is a topic of great interest due, mainly, to the challenging possibility of mimicking and controlling the structure and functionality of natural biomacromolecules (e.g., enzymes, drug delivery vehicles, catalysts) by means of artificial single chain nano-objects. By performing extensive molecular dynamics simulations we investigate the formation of soft nanoparticles by irreversible intramolecular cross-linking of polymer precursors of different length. In order to optimize the folding process and to obtain more compact structures we vary the number of chemical species among the linker groups (orthogonal chemistry) which selectively form the bonds. The use of orthogonal chemistry protocols, by increasing the number of different chemical species of the linkers, lead to nanoparticles that are systematically smaller and more spherical than the homofunctional counterparts. We characterize the conformational properties of the resulting nanoparticles. These are intrinsically polydisperse in size, with a significant fraction of sparse topologies. We discuss the relevance of our results for synthesis protocols in real systems.

1 Introdution

Protein folding to its native, functional state is determined by multiple factors including the exact amino acid sequence, or primary protein structure, the interactions between amino acids in solution as well as the interactions of protein residues with water. Through millions of years of nature evolution, the precise conformation and related molecular function of folded proteins (e.g., enzymes), results from a variety of different molecular interactions which determine noncovalent (hydrogen bonding, hydrophobic interactions, π-π stacking, metal coordination) as well as covalent bonds (e.g. disulfide bonds)[1]. As a consequence of the concerted and multi-orthogonal interactions involved in protein folding, the most compact native state of globular proteins is solid-like[2]. Thus, its squared radius of gyration scales as $R_g^2 \sim N^{2/3}$, where N is the number of residues in the polypeptide chain, similarly to the scaling expected for an isolated polymer chain in poor solvent. Conversely, in the unfolded state the polypeptide chain is best described as a self-avoiding random coil[3] showing a scaling ($N^{1.2}$) similar to that of a polymer chain in good solvent.

The controlled folding of single polymer chains into soft nanoparticles has attracted significant interest due to its (partial) resemblance to the polypeptide chain collapse process into the compact native state[4]. In recent years, several intrachain cross-linking chemistries have been introduced for single-chain nanoparticle (SCNP) synthesis involving covalent[5–28], noncovalent[29–35] and dynamic covalent bonds[36–39], most of them comprising a single cross-linking type. In addition, techniques like intrachain homocoupling, intrachain heterocoupling or cross-linking induced collapse are currently available to promote the intramolecular folding of individual polymer chains to SCNPs[40]. Inspired by protein functions, enzyme-mimetic catalytic properties[25,41,42], drug transport/delivery ability[24,43] sensing characteristics[44] or even polymerase activity[25], have been envisioned and established for artificial single-chain nano-objects.

Nevertheless, current SCNPs are still far from being globular objects in solution even by using highly efficient 'click' chemistry techniques[40] or combined hydrogen bonding/chiral folding processes[45]. Small-angle neutron scattering (SANS) and small-angle X-ray scattering (SAXS) measurements of SCNPs in solution have shown form factors[24,25,43,45,46] more closely related to those of crumpled coils (disordered proteins) than those of compact globular objects (native proteins). Indeed scattering form factors consistent with $R_g^2 \sim N^{\nu}$ were observed, with $0.85 < \nu < 1.1$. Though some investigations by molecular dynamics (MD) simulations revealed the existence of SCNPs sparse configurations[47–49], a precise determination of their

[0a] *Donostia International Physics Center, Paseo Manuel de Lardizabal 4, 20018 San Sebastián, Spain. E-mail:federica_loverso001@ehu.es*

[0b] *Materials Physics Center MPC, Paseo Manuel de Lardizabal 5, E-20018 San Sebastián, Spain*

[0c] *Centro de Física de Materiales (CSIC, UPV/EHU), Paseo Manuel de Lardizabal 5, E-20018 San Sebastián, Spain*

[0d] *Departamento de Física de Materiales, Universidad del País Vasco (UPV/EHU), Apartado 1072, E-20080 San Sebastián, Spain*

[0e] *IKERBASQUE Basque Foundation for Science, Alameda de Urquijo 36, E-48011 Bilbao, Spain*

scaling properties was lacking. Recently, we confirmed by means of extensive MD simulations[24,25,43,46] the scaling behaviour suggested by SANS and SAXS experiments. The significant departure from the globular state of SCNPs in solution was attributed to the inherent self-avoiding character of the precursor polymer chain in good solvent conditions. This feature strongly impedes (apart from unfrequent events) the formation of long-range intramolecular loops during the cross-linking process, which is the efficient mechanism for the formation of globular nanoparticles[46]. We observed that, indeed, scaling properties of SCNPs were dominated by these unfrequent long-range loops: the scaling behaviour was essentially *independent* of the fraction of linkers, since most of the cross-linking events involved inefficient short-range loops[46].

In addition, we demonstrated through both MD simulations and experimental results the significant advantages of intramolecular cross-linking of heterofunctional precursors decorated with two species of orthogonal linkers[46]. We showed that long-range looping is enhanced by orthogonal synthesis and the resulting nanoparticles are, on average, smaller and more spherical than their homofunctional counterparts (i.e., those with only one type of linker). Motivated by these results, in this work we investigate the limits of multi-orthogonal folding protocols by supplying the precursor with several linkers of different chemical species, which selectively react with each other, i.e., a couple of linkers create a bond only if they belong to the same chemical species. We analyse the folding and conformational properties of several model systems with a number of chemical species per molecule, x, ranging from three to six. We find that multi-orthogonal folding leads to a systematic compactation of the obtained nanoparticles by increasing the number of chemical species. Still, significant intrinsic polydispersity and non-globular character persist even for $x = 6$. We discuss the implications of these protocols for synthesis in real systems.

The article is organized as follows: In Section 2 we describe the simulated model and give MD details. In Section 3 we discuss our simulation results. In particular we investigate the scaling properties, shape and internal structure of the obtained soft nanoparticles starting from different precursor systems. Finally, Section 4 summarizes the main conclusions of our work.

2 Model and Method

The experimental precursor usually consists of a polymer chain containing bulky side groups[6,7,9,16,24,47–49] (e.g., phenyl groups), some of them being functionalized with short branches ending in reactive groups. We described qualitatively the precursor by a bead-spring model. A bead in this qualitative model represents the center-of-mass of typically 2-4 carbons[50]. In Fig. 1 we show a schematic representation of the architecture

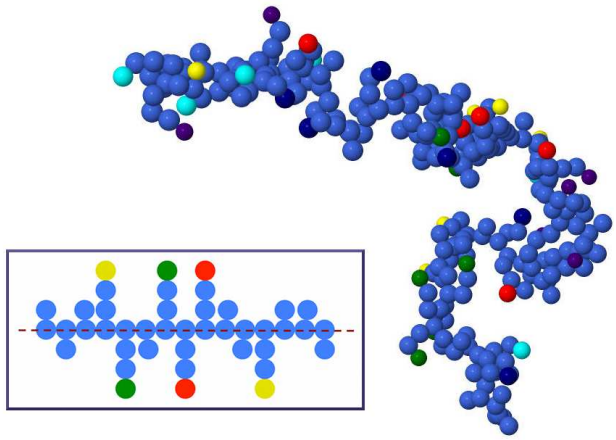


Figure 1 Typical equilibrium configuration of a polymer precursor (main panel) and scheme of the precursor architecture (inset). Dark blue beads form the backbone (indicated by the dashed line in the inset) and the inactive parts of the side groups. Beads of other colours (different for each chemical species) correspond to the linkers (e.g., yellow, green and red beads in the inset).

of the precursor (inset) and a typical equilibrated configuration. The precursor consists of a linear backbone of N_b beads, each of them being attached to a side group. There are two kinds of side groups, which are randomly distributed along the backbone. A number N_l of the side groups (the 'linker side groups') contain three beads. The free end beads are the cross-linkers (e.g., yellow, green and red monomers in the inset of Fig. 1). The remaining $N_b - N_l$ side groups contain one single bead. The total number of beads per molecule is $N = 2(N_b + N_l)$. The fraction of linker side groups is defined as $f = N_l/N_b$. We considered several systems, by changing the chemical species of the reactive linkers. The different systems are coded as SPx where x is the number of the different chemical species of the linkers. In our analysis we considered the cases $1 \leq x \leq 6$. The case $x = 1$ corresponds to standard homofunctional precursors. We fixed the total fraction of cross-linkers, $f = 0.4$, and for each system we varied the number of backbone monomers ($50 \leq N_b \leq 400$). For each system SPx the linkers of each species were randomly distributed along the backbone, and in an identical proportion (f/x) to the other $x - 1$ species.

We performed implicit-solvent simulations, i.e, the solvent molecules were not included. All pairs of beads ('monomers') in the polymer, irrespective whether they were bonded to each other or not, interacted through a purely repulsive Lennard-Jones potential (LJ),

$$V_{LJ}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} + \frac{1}{4} \right], \qquad (1)$$

for $r < r_c$, and $V(r) = 0$ for $r \geq r_c$. By using a cut-off distance $r_c = 2^{1/6}\sigma$, the potential and the forces were continuous at $r = r_c$, and were purely repulsive. In this way we mimicked cross-linking in good solvent conditions. Bonded monomers interacted through the finitely extensible nonlinear elastic (FENE) potential[50]:

$$V_{\mathrm{FENE}}(r) = -\varepsilon K_{\mathrm{F}} R_0^2 \ln\left[1 - \left(\frac{r}{R_0\sigma}\right)^2\right]. \qquad (2)$$

We used $K_{\mathrm{F}} = 15$ and $R_0 = 1.5$ so that the total spring potential between two connected monomers had its minimum at $r_{\mathrm{min}} = 0.96\sigma$. This guaranteed chain uncrossability[50]. We used $\varepsilon = 1$, $\sigma = 1$ for *all* pairs of beads, setting the scales of energy and length, respectively. Thus, for simplicity we used identical bonded and non-bonded interactions for all the chemical species in the molecule.

Standard MD methods were employed. We simulated *isolated* chains coupled to the same thermal bath. In this way we investigated purely intramolecular cross-linking of chains at high dilution conditions. In order to stabilize the selected temperature, $T = \varepsilon/k_{\mathrm{B}}$ (with k_{B} the Boltzmann constant), we implemented a Langevin thermostat (see Ref.[46] for further details). For simplicity we used identical masses $m = 1$ and friction coefficients $\gamma = 0.05$ for all the beads. The simulation time unit is $\tau = (m\sigma^2/\varepsilon)^{1/2}$. The Langevin equations of motion were discretized with a time step $\Delta t = 0.01\tau$ and integrated in the velocity-Verlet scheme, following the impulse approach proposed in Refs.[51,52].

Equilibration runs were performed over several millions steps, in order to sample equilibrium configurations of the unlinked precursors. After equilibration, cross-linking was activated. The detailed implementation of the cross-linking runs is described in Ref.[46]. Briefly, multiple bonding between linkers was not permitted: each linker only formed *one* bond with another linker. Moreover, only bonding between linkers of the *same* species was permitted. Bonds were irreversible: once they were formed, they remained for the rest of the simulation (with the bonding interaction given by the FENE potential of Eq. 2). Besides fulfilling the former conditions, two linkers formed a bond only if they were at a mutual distance smaller than a 'capture distance'[53] of 1.3σ. In case of multiple possibilities within the capture distance, the new bonds were selected randomly (see Ref.[46]).

Where not stated otherwise, we considered simultaneous cross-linking of the reactive groups, where all the linkers were simultaneously activated from the beginning of the cross-linking process. A second possibility is a sequential cross-linking. In this case we activated a single species of linkers at a time. Once the cross-linking of this species was completed, an equilibration run of a few million steps was performed and then cross-linking of another species was started. This process was repeated over the different species until full cross-linking was completed. Sequential cross-linking is computationally more demanding than the simultaneous route, differences quickly increasing with the number x of different chemical species in the polymer. Therefore we investigated sequential cross-linking for a few selected cases with low values of $x = 2$ and 3. The resulting cross-linked nanoparticles did not show significant differences with their counterparts obtained by simultaneous cross-linking (see below).

After completing cross-linking, simulations were further extended over several millions steps to accumulate configurations for statistical averages. For each system we simulated a minimum of $N_{\mathrm{t}} = 200$ polymers. Thus, statistical averages included average over both the N_{t} *different* polymers (*ensemble-average*), and the different configurations sampled by the propagation of the individual polymers (*time-average*). Time-average was performed over several hundreds of conformations for each individual polymer. These conformations were selected at equispaced times. The corresponding intervals were sufficiently long to allow for full decorrelation of the molecular conformations. For the largest investigated systems and the highest number of species, there was a small fraction of polymers that did not complete cross-linking within the simulated time scales. These were excluded in the calculation of the conformational properties presented in next sections, which strictly correspond to fully cross-linked nanoparticles.

3 Results and discussion

By fixing $f = 0.4$ and increasing the number x of different chemical species up to 6, we can investigate quite folded objects SPx, and compare results with trends previously found for the cases $x = 1$ and $x = 2$. Irreversible cross-linking results into different topologies of the obtained soft nanoparticles. In Fig. 2 we show typical topologies for nanoparticles with $N_{\mathrm{b}} = 200$, for the systems SP4 and SP6. By increasing the number of chemical species a clear trend arises: the conformations are on average more compact and sparse nanoparticles are less frequent. A similar behaviour is observed for all the backbone lengths investigated (see e.g., $N_{\mathrm{b}} = 400$, SP6 in Fig. 3). For the purpose of characterizing systematically the size of the nanoparticles we calculate their radius of gyration. Fig. 4 shows the averaged squared radius of gyration $\langle R_{\mathrm{g}}^2 \rangle$ (joint ensemble and time average) as a function of the backbone length, for all the investigated values of x. All data sets are consistent with the power-law scaling $\langle R_{\mathrm{g}}^2 \rangle = b^2 N_{\mathrm{b}}^{\nu}$, with b the effective segment length and ν the effective exponent. The values of ν, as obtained from data fits to the former power law, are indicated in the main panel of Fig. 4. The inset shows the x-dependence of ν. By increasing the number of chemical species of the cross-linkers, the nanoparticles exhibit lower exponents. In particular, there is a

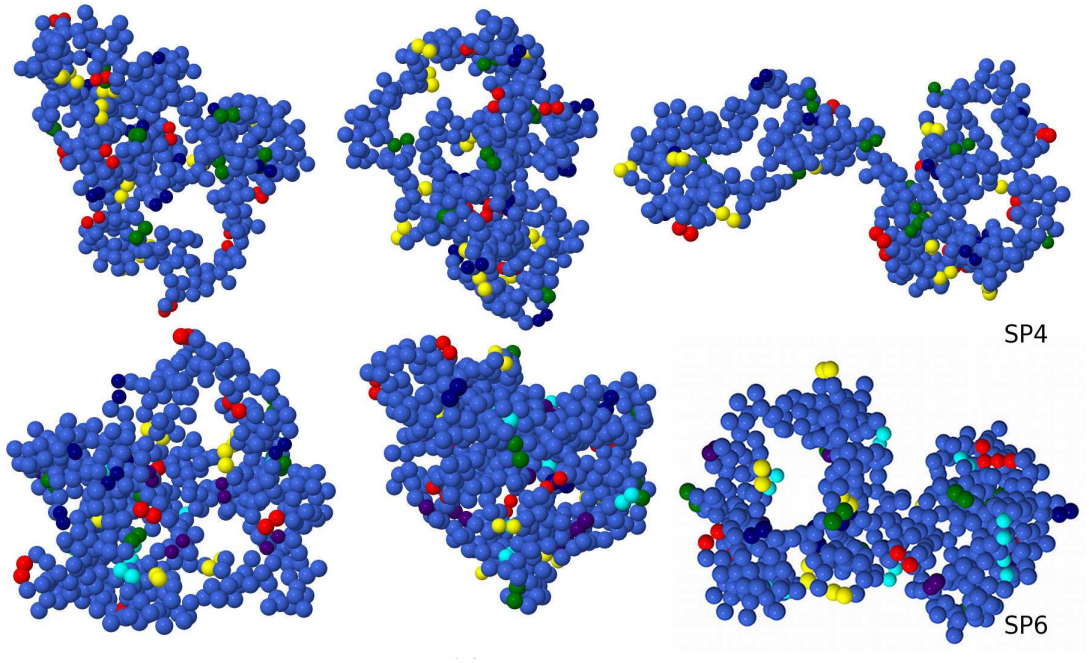


Figure 2 Typical topologies of cross-linked nanoparticles obtained by orthogonal folding of polymer precursors with $N_b = 200$ and $f = 0.4$. Top line: SP4, bottom line: SP6. Dark blue beads correspond to the inactive monomers. Beads of other colours correspond to the linkers (a different colour for each chemical species, note the pairs of bonded linkers).

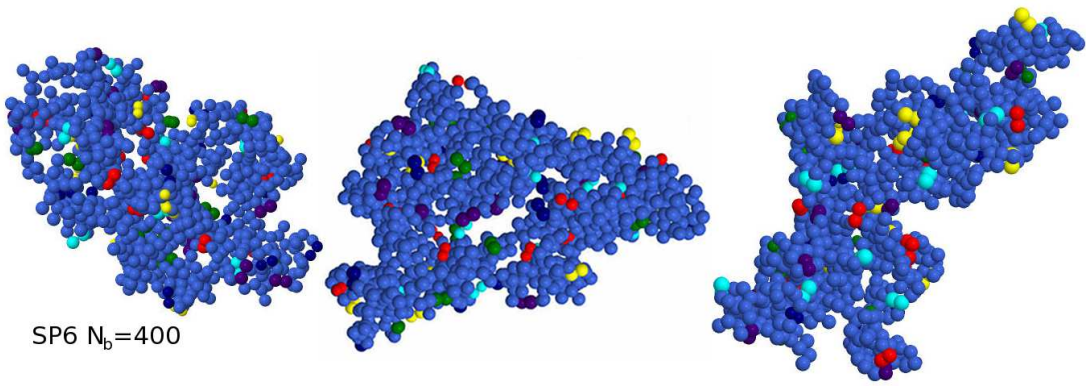


Figure 3 As Fig. 2 for SP6 with $N_b = 400$ and $f = 0.4$.

significant change between the homofunctional case (SP1) and the SP3-nanoparticles. As the number of species is further increased, ν keeps decreasing although at a slower rate.

As already discussed[46] the exponents for SP1 and SP2 resemble that of random walks ($\nu_{RW} \equiv 1$), which describes the (Gaussian) statistics of linear chains in polymer melts or in θ-solvents[54,55]. This is essentially due to the *local*, short-range globulation of the nanoparticles resulting from the cross-linking process. Because of the self-avoiding character of the precursors in the simulated good solvent conditions, long-range looping is unfrequent, preventing strong compactation of the resulting nanoparticles. As already revealed from the analysis of SP2 nanoparticles[46], long-range looping can be favoured by increasing the number of different chemical species, yielding more compact nanoparticles. This is confirmed by representing the distribution $P(s)$ of contour distances between bonded linkers. The contour distance is defined as $s = 1 + n$, with n the number of backbone beads between the two backbone beads to

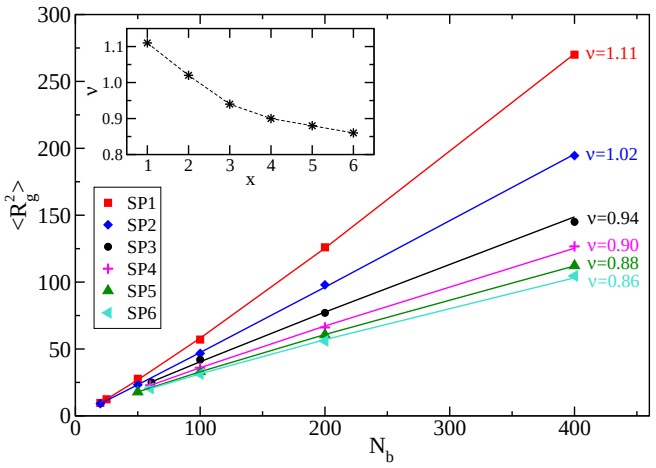


Figure 4 Main panel: average squared radius of gyration of the cross-linked nanoparticles. Symbols are simulation values vs. the backbone length N_b. Different data sets (see legend) correspond to the cases of cross-linked homofunctional (SP1) and heterofunctional (SP2, SP3, SP4, SP5, SP6) nanoparticles. The ν-exponents, as obtained from fits (solid lines) to power laws $\langle R_g^2 \rangle = b^2 N_b^{\nu}$ are indicated in the main panel, and represented in the inset vs. the number x of chemical species of the linkers.

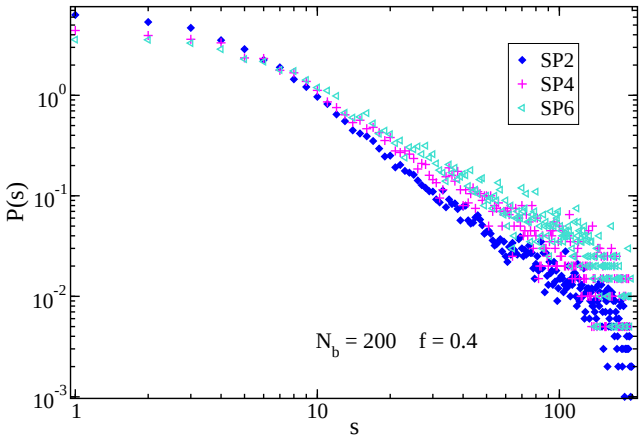


Figure 5 Histogram of contour distances s between bonded linkers for fully cross-linked nanoparticles with identical values of $N_b = 200$ and $f = 0.4$. Different data sets (see legend) correspond to different numbers of chemical species: SP2, SP4, and SP6.

which the corresponding linker side groups are attached. Results for $P(s)$ are represented in Fig. 5 for the cases SP2, SP4, SP6. Increasing the number of different chemical species in the precursor, at fixed N_b and f, increases the average contour distance between linkers that can form mutual bonds. This is due to the fact that the species are randomly distributed (in identical fractions f/x) along the chain contour and cross-linking between different species is not permitted. These features lead to a higher population of long-range loops for larger x, which results in a more efficient compactation of the nanoparticle.

For $x \geq 3$ the scaling exponent ν decreases below the value $\nu_{RW} \equiv 1$ for random walks. However, the lowest exponent observed, $\nu = 0.86$ for SP6, is still clearly above the ideal value for spherical objects $\nu_S = 2/3$. This finding suggests that a significant fraction of sparse topologies still exists even for SP6-nanoparticles, as confirmed by visual inspection (see representative snapshots in Figs. 2 and 3), and by determining the distributions $P(\bar{R}_g)$ of the time-averaged radius of gyration, $\bar{R}_g$. Fig. 6 shows representative results of $P(\bar{R}_g)$ for fixed $N_b = 200$ and $f = 0.4$. For comparison we include the value of $\bar{R}_g$ for the corresponding precursor. Though, because of the random distribution of the side groups along the backbone, the individual precursors are not strictly equivalent, this has a minor effect in the time-averages of their intramolecular fluctuations. As a consequence, all precursors with same N_b, f and x show, within statistics, the same value of $\bar{R}_g$. However, even by starting from the same precursors (identical N_b, f and x) irreversible cross-linking produce rather different topologies for the nanoparticles, which are intrinsically polydisperse in size. Having noted this, consistently with the scaling properties of $\langle R_g^2 \rangle$, a clear shift of $P(\bar{R}_g)$ to lower $\bar{R}_g$ values is found whilst moving from SP1 to SP6. This corroborates the efficiency of the cross-linking protocol in terms of more compact structures. In particular increasing x does not only result, on average, in smaller nanoparticles, but also reduces the asymmetry of the distribution $P(\bar{R}_g)$. Thus, the long tail observed for the SP2-nanoparticles (Fig. 6) almost vanishes for the SP6-nanoparticles.

Further insight on the effect of x on the topology of the nanoparticles can be obtained by analyzing shape parameters as the asphericity and prolateness. These can be obtained from the radius of gyration tensor[56,57], defined as:

$$T_{\alpha\beta} = \frac{1}{N^2} \sum_{i=1}^{N} (r_{i\alpha} - r_{\alpha}^{cm})(r_{i\beta} - r_{\beta}^{cm}), \qquad (3)$$

where α, β denote cartesian components of the position vectors, $\mathbf{r}_i$ and $\mathbf{r}^{cm}$ for the ith-monomer and center-of-mass of the polymer, respectively. Shape parameters can be obtained from the three eigenvalues, $\lambda_1, \lambda_2, \lambda_3$, of the gyration tensor. The asphericity parameter, $0 \leq a \leq 1$, is defined as[56,57]:

$$a = \left\langle \frac{(\lambda_2 - \lambda_1)^2 + (\lambda_3 - \lambda_1)^2 + (\lambda_3 - \lambda_2)^2}{2(\lambda_1 + \lambda_2 + \lambda_3)^2} \right\rangle. \qquad (4)$$

For a perfectly spherical object $a = 0$. The prolateness parameter, $-1 \leq p \leq 1$, is defined as[56,57]:

$$p = \left\langle \frac{(3\lambda_1 - R_g^2)(3\lambda_2 - R_g^2)(3\lambda_3 - R_g^2)}{2(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 - \lambda_1\lambda_2 - \lambda_1\lambda_3 - \lambda_2\lambda_3)^{3/2}} \right\rangle. \qquad (5)$$

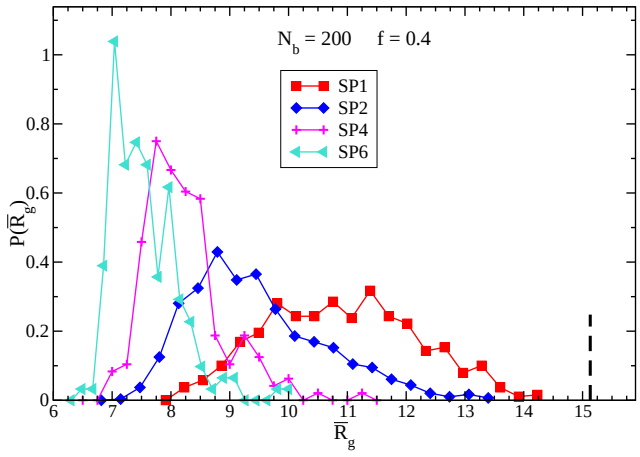


Figure 6 Symbols: for systems with $N_b = 200$ and $f = 0.4$, distributions of time-averaged radii of gyration $\bar{R}_g$ of the cross-linked nanoparticles. Solid lines are guides for the eyes. The vertical dashed line is the (single) value of $\bar{R}_g$ for the corresponding unlinked precursor.

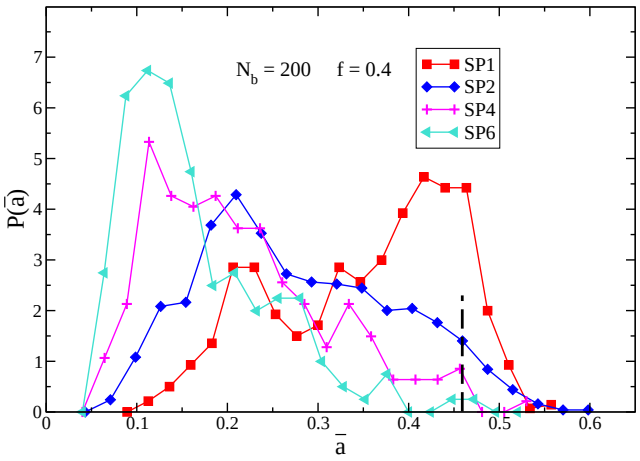


Figure 7 Same as Fig. 6 for the asphericity parameter $\bar{a}$.

For perfectly oblate objects ($\lambda_1 < \lambda_2 = \lambda_3$) the prolateness is $p = -1$. For perfectly prolate objects ($\lambda_1 = \lambda_2 < \lambda_3$), $p = 1$.

Fig. 7 shows results for the distributions $P(\bar{a})$ of the time-averaged asphericity parameter, corresponding to $N_b = 200$, $f = 0.4$ and $x = 1, 2, 4, 6$ used in Fig. 6 for $P(\bar{R}_g)$. Consistently with the observed reduction of the nanoparticle size (Fig. 6), increasing the number of different chemical species at fixed N_b and f results into more spherical objects. However, a significant fraction of quite non-spherical objects is still present even for the SP6-nanoparticles. This is reflected by the broad tail in the distribution $P(\bar{a})$, extending beyond the main peak (around $\bar{a} \approx 0.1$) of quasi-spherical nanoparticles.

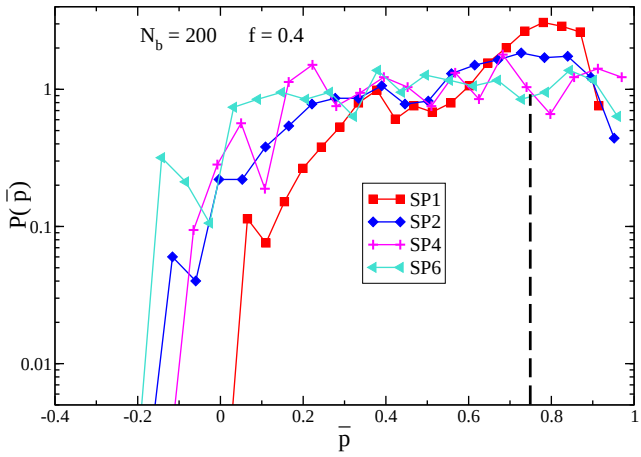


Figure 8 Same as Fig. 7 for the prolateness parameter $\bar{p}$.

Fig. 8 shows the corresponding results for the distributions $P(\bar{p})$ of the time-averaged prolateness parameters. For small x these are dominated by prolate nanoparticles ($\bar{p} \to 1$) — note the logarithmic scale in the ordinate axis of Fig. 8. The distribution tends to flatten in the range $\bar{p} > 0$ for the largest investigated x-values, though in all cases intrinsically oblate nanoparticles ($\bar{p} \to -1$) are extremely rare.

Further details on the structure of the nanoparticles can be obtained by representing the radial monomer density profiles, $\rho(r)$, around the molecular center-of-mass. Results are presented in Fig. 9 for $N_b = 200$ and $f = 0.4$, both for the precursor and the nanoparticles. We describe the density profiles by fits (solid lines in Fig. 9) to generalized exponential functions, $\rho(r) \propto \exp(-\alpha r^m)$. Whilst a nearly Gaussian profile ($m = 2.4$) is found for the precursor, $\rho(r)$ for the nanoparticles becomes strongly non-Gaussian by increasing the number of chemical species of the linkers. Thus, we find $m = 4.5$ for the SP6-nanoparticles, reflecting a sharper decrease of the density from the molecular center (and therefore a more compact structure) than in the low-x counterparts. In all cases density profiles are smooth. No evidence of a core-shell structure, typical of e.g., folded proteins, is found.

Now we analyze the spatial distribution of the linkers in the cross-linked nanoparticles. Obviously, there is a well-defined characteristic (bonding) distance, $r \approx \sigma$, between two bonded linkers. In order to remove this feature, we treat each pair of bonded linkers as a 'single particle' with position at the center of the corresponding bond. Then we characterize correlations between such bond centers through the corresponding radial distribution function $g(r)$. Fig. 10 shows results of $4\pi r^2 g(r)$ for $N_b = 200$, $f = 0.4$ and different x-values. For comparison we include results for the precursor. In the case of the precursor, we show the radial distribution function of the linker posi-

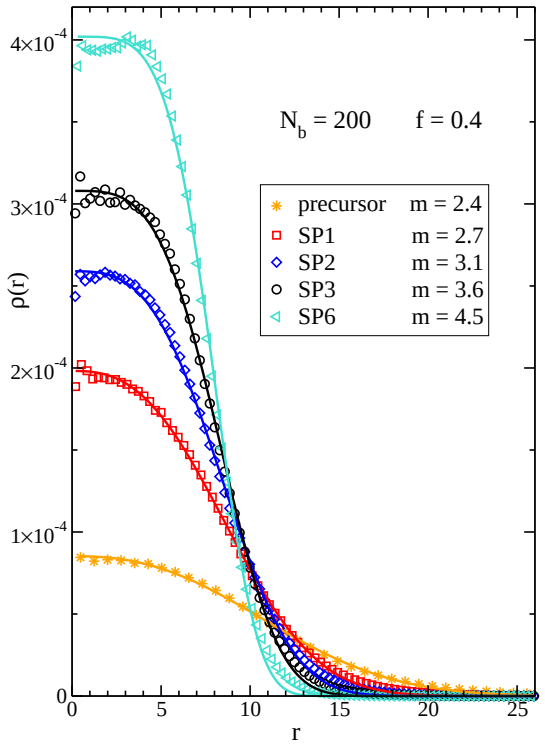


Figure 9 Symbols: for $N_b = 200$ and $f = 0.4$, radial monomer density profiles around the center-of-mass. For comparison we include the corresponding results for the precursor. Lines: fits to generalized exponentials, $\rho(r) \propto \exp(-\alpha r^m)$. The m-exponents are indicated.

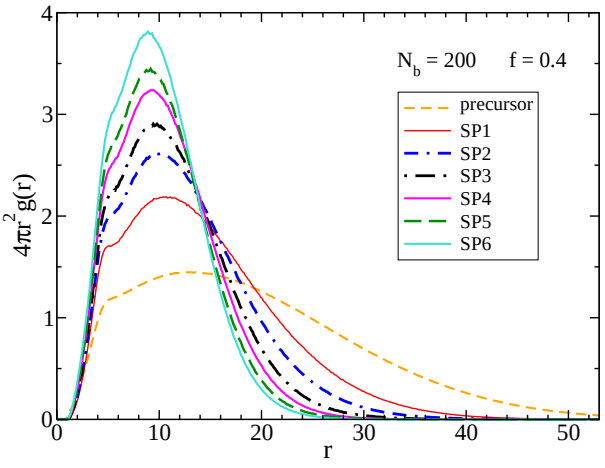


Figure 10 For $N_b = 200$ and $f = 0.4$, radial distribution function (multiplied by the phase factor $4\pi r^2$) for the bonds connecting linkers in the nanoparticles SPx. For comparison we include the corresponding results for the linker positions in the precursor.

tions, but we use the same normalization as for the curves of the nanoparticles. Apart from the expected excluded volume hole at $r < \sigma$, the only significant feature in $g(r)$ is a kink at $r \approx 5\sigma$. This is related to short-range correlations between linkers that are already present in the precursor (see orange dashed line in Fig. 10). The histogram of distances between bonded linkers at longer r does not show relevant features and is essentially Gaussian. These results indicate that the spatial arrangement of the linkers in the nanoparticles is essentially random in all the investigated cases. These results are confirmed by representing (not shown) the radial density profiles $\rho(r)$ of the linkers. Within statistics, they are identical to the corresponding all-monomer profiles of Fig. 9.

Results for other backbone lengths (not shown) exhibit the same qualitative trends presented in Figs. 5 to 10 for $N_b = 200$. All the results presented above correspond to the case of simultaneous cross-linking of all the species. To conclude this section, we discuss the effect of sequential cross-linking, in which bonding of the ith-species is only initiated after full cross-linking of the $(i-1)$th species. Unfortunately, sequential cross-linking becomes computationally more and more demanding with increasing x. Here we limit the comparison between simultaneous and sequential cross-linking to the case $x = 3$, for fixed $N_b = 200$ and $f = 0.4$. Fig. 11 shows results for the corresponding distributions of time-averaged radii of gyration, asphericities and prolateness. As previously observed[46] for the simplest case $x = 2$, nanoparticles with $x = 3$ obtained from the same precursors by simultaneous or sequential cross-linking do not show significant differences in their conformational properties. We expect that negligible differences will persist at higher x, at least up to $x = 6$. From the simulation data we can estimate the effective density of monomers in the nanoparticle as $\rho \sim N(4\pi R_g^3/3)^{-1}$. For the highest investigated value $x = 6$ we obtain $\rho < 0.3$, which is much smaller than the density, $\rho \sim 1$, of the corresponding melt of bead-spring chains[50]. This suggests that the progressive bonding of the linkers does no result in significant crowding effects ('glassy dynamics') hindering the formation of new bonds, and therefore simultaneous and sequential cross-linking produce nanoparticles with very similar statistical properties.

4 Conclusions and outlook

We have presented MD simulations of the intramolecular cross-linking of polymer chains into soft nanoparticles. We have performed a systematic investigation of the size and shape of the nanoparticles, exploring the effect of the number x of different chemical species of the linkers. At fixed backbone length and fraction of linkers, increasing x lead to nanoparticles that are, on average, smaller and more compact. However both the scaling exponent and the distributions of shape parameters reveal

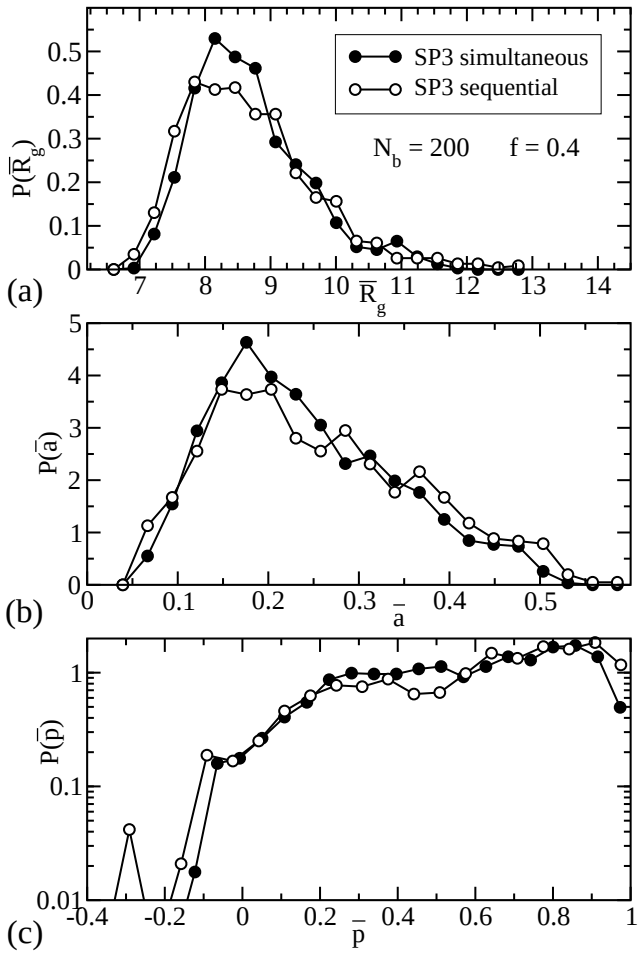


Figure 11 Distributions of time-averaged values for size and shape parameters. (a): $P(\bar{R}_g)$; (b): $P(\bar{a})$; (c): $P(\bar{p})$. Data in all panels correspond to heterofunctional nanoparticles SP3, with identical values of $N_b = 200$ and $f = 0.4$, but obtained through different cross-linking routes. Filled and empty symbols correspond to the simultaneous and sequential route, respectively.

that cross-linking produces, even for $x = 6$, a significant fraction of highly non-spherical, sparse nanoparticles. The density profiles of the nanoparticles are smooth. No evidence of, e.g., core-shell structures is found, neither particular spatial arrangements of the linkers, which are randomly distributed in the cross-linked nanoparticle. The same conclusions are valid for all the backbone lenghts we considered.

Our results are relevant for guiding the implementation of *multi-orthogonal* protocols for the synthesis of real soft nanoparticles, a complex problem that is still scarcely explored. To the best of our knowledge, most of the experimental investigations have been limited to the homofunctional case ($x = 1$) and to the most simple heterofunctional case of $x = 2$. Very recently Berda and co-workers[58] have reported the synthesis of SP3-nanoparticles, based on poly(oxanorbornene anhydride-*co*-cyclooctadiene), formed by supramolecular folding with aniline tetramer, reaction with diamine to covalently fix the folded structure, and second covalent folding reaction via thiol-ene chemistry. The observed reduction of the hydrodynamic radius[58], in comparison with the SP1 and SP2 counterparts, is consistent with the trends observed in our simulations. It will be instructive, although experimentally very demanding, to investigate the folding of precursors containing a higher number of different chemical species. Still, our simulation results suggest that multi-orthogonal protocols have a fundamental limitation for producing globular (i.e., scaling as $N^{2/3}$) single-chain nanoparticles. This limitation is intimately connected to the inherent self-avoiding character of the polymer precursors in good solvent conditions. As a consequence of the self-avoiding equilibrium conformations, long-range intramolecular looping, which is the efficient mechanism for global compactation, is unfrequent, and most of the cross-links involve short contour distances that just produce *local* globulation. This is the case even by using multi-orthogonal folding which, by keeping a high density of linker side groups along the chain backbone, increases the average contour distance between linkers than can potentially form mutual bonds. Even for $x = 6$, long-range looping is not sufficiently promoted to fully prevent the formation of sparse nanoparticles.

Still, there may be alternative routes for the formation of single-chain globular soft nanoparticles. One protocol might consist in using polymers with much longer side arms than in standard precursors, or very long bridging groups connecting the linkers, when the cross-linked induced collapse technique is used[40]. This would largely increase the probability of long-range looping, even for self-avoiding conformations of precursors in good solvent. Another route might be performing cross-linking in bad solvent, and swelling the obtained nanoparticles by recovering good solvent conditions after completing cross-linking. The dense globular conformations of polymers in bad solvent should strongly favour cross-linking at long contour distances, eventually leading to globular conformations of the *swollen* nanoparticles. However, intermolecular aggregation in bad solvent conditions occurs even at very high dilution, preventing the formation of single-chain nanoparticles. On this basis, an alternative strategy to prevent aggregation might be anchoring the precursor to a surface, with low grafting density, during the cross-linking in bad solvent. Another possibility for free precursors would be to use properly taylored *amphiphilic* polymers, in which the solvophobic reactive groups would be fully isolated by a shell formed by the solvophilic parts[59,60], preventing intermolecular aggregation during the cross-linking process. Work in these directions is in progress.

5 Acknowledgements

We acknowledge financial support from Projects No. MAT2012-31088 (MEC, Spain) and No. IT654-13 (GV, Spain), and CSUC for generous allocation of CPU time.

References

[1] C. B. Anfinsen, *Science*, 1973, **181**, 223–230.

[2] D. Thirumalai, E. P. O'Brien, G. Morrison and C. Hyeon, *Annu. Rev. Biophys.*, 2010, **39**, 159–183.

[3] J. B. Udgaonkar, *Arch. Biochem. Biophys.*, 2013, **531**, 24–33.

[4] O. Altintas and C. Barner-Kowollik, *Macromol. Rapid Commun.*, 2012, **33**, 958–971.

[5] D. Mecerreyes, V. Lee, C. J. Hawker, J. L. Hedrick, A. Wursch, W. Volksen, T. Magbitang, E. Huang and R. D. Miller, *Adv. Mater.*, 2001, **13**, 204–208.

[6] E. Harth, B. V. Horn, V. Y. Lee, D. S. Germack, C. P. Gonzales, R. D. Miller and C. J. Hawker, *J. Am. Chem. Soc.*, 2002, **124**, 8653–8660.

[7] J. Jiang and S. Thayumanavan, *Macromolecules*, 2005, **38**, 5886–5891.

[8] A. E. Cherian, F. C. Sun, S. S. Sheiko and G. W. Coates, *J. Am. Chem. Soc.*, 2007, **129**, 11350–11351.

[9] T. A. Croce, S. K. Hamilton, M. L. Chen, H. Muchalski and E. Harth, *Macromolecules*, 2007, **40**, 6028–6031.

[10] C. T. Adkins, H. Muchalski and E. Harth, *Macromolecules*, 2009, **42**, 5786–5792.

[11] A. Ruiz de Luzuriaga, N. Ormategui, H. J. Grande, I. Odriozola, J. A. Pomposo and I. Loinaz, *Macromol. Rapid Commun.*, 2008, **29**, 1156–1160.

[12] N. Ormategui, I. Garcia, D. Padro, G. Cabanero, H. J. Grande and I. Loinaz, *Soft Matter*, 2012, **8**, 734–740.

[13] J. E. F. Radu, L. Novak, J. F. Hartmann, N. Beheshti, A.-L. Kjoniksen, B. Nyström and J. Borbély, *Colloid Poly. Sci.*, 2008, **286**, 365–376.

[14] J. B. Beck, K. L. Killops, T. Kang, K. Sivanandan, A. Bayles, M. E. Mackay, K. L. Wooley and C. J. Hawker, *Macromolecules*, 2009, **42**, 5629–5635.

[15] L. Oria, R. Aguado, J. A. Pomposo and J. Colmenero, *Adv. Mater.*, 2010, **22**, 3038–3041.

[16] A. Ruiz de Luzuriaga, I. Perez-Baena, S. Montes, I. Loinaz, I. Odriozola, I. García and J. A. Pomposo, *Macromol. Symp.*, 2010, **296**, 303–310.

[17] B. Zhu, J. Ma, Z. Li, J. Hou, X. Cheng, G. Qian, P. Liu and A. Hu, *J. Mater. Chem.*, 2011, **21**, 2679–2683.

[18] B. Zhu, G. Qian, Y. Xiao, S. Deng, M. Wang and A. Hu, *J. Polym. Sci., Part A: Polym. Chem.*, 2011, **49**, 5330–5338.

[19] X. Jiang, H. Pu and P. Wang, *Polymer*, 2011, **52**, 3597 – 3602.

[20] P. Wang, H. Pu and M. Jin, *J. Polym. Sci., Part A: Polym. Chem.*, 2011, **49**, 5133–5141.

[21] A. Sanchez-Sanchez, I. Asenjo-Sanz, L. Buruaga and J. A. Pomposo, *Macromol. Rapid Commun.*, 2012, **33**, 1262–1267.

[22] P. Khanjani, I. Perez-Baena, L. Buruaga and J. A. Pomposo, *Macromol. Symp.*, 2012, **321-322**, 145–149.

[23] P. T. Dirlam, H. J. Kim, K. J. Arrington, W. J. Chung, R. Sahoo, L. J. Hill, P. J. Costanzo, P. Theato, K. Char and J. Pyun, *Polym. Chem.*, 2013, **4**, 3765–3773.

[24] A. Sanchez-Sanchez, S. Akbari, A. Etxeberria, A. Arbe, U. Gasser, A. J. Moreno, J. Colmenero and J. A. Pomposo, *ACS Macro Lett.*, 2013, **2**, 491–495.

[25] I. Perez-Baena, F. Barroso-Bujans, U. Gasser, A. Arbe, A. J. Moreno, J. Colmenero and J. A. Pomposo, *ACS Macro Lett.*, 2013, **2**, 775–779.

[26] O. Altintas, J. Willenbacher, K. N. R. Wuest, K. K. Oehlenschlaeger, P. Krolla-Sidenstein, H. Gliemann and C. Barner-Kowollik, *Macromolecules*, 2013, **46**, 8092–8101.

[27] S. Mavila, C. E. Diesendruck, S. Linde, L. Amir, R. Shikler and N. G. Lemcoff, *Angew. Chem. Int. Ed.*, 2013, **52**, 5767–5770.

[28] P. G. Frank, B. T. Tuten, A. Prasher, D. Chao and E. B. Berda, *Macromolecular Rapid Communications*, 2014, **35**, 249–253.

[29] M. Seo, B. J. Beck, J. M. J. Paulusse, C. J. Hawker and S. Y. Kim, *Macromolecules*, 2008, **41**, 6413–6418.

[30] E. J. Foster, E. B. Berda and E. W. Meijer, *J. Am. Chem. Soc.*, 2009, **131**, 6964–6966.

[31] J. He, L. Tremblay, S. Lacelle and Y. Zhao, *Soft Matter*, 2011, **7**, 2380–2386.

[32] E. A. Appel, J. Dyson, J. del Barrio, Z. Walsh and O. A. Scherman, *Angew. Chem. Int. Ed.*, 2012, **51**, 4185–4189.

[33] T. Akagi, P. Piyapakorn and M. Akashi, *Langmuir*, 2012, **28**, 5249–5256.

[34] T. Mes, R. van der Weegen, A. R. A. Palmans and E. W. Meijer, *Angew. Chem. Int. Ed.*, 2011, **50**, 5085–5089.

[35] P. Piyapakorn, T. Akagi, M. Hachisuka, T. Onishi, H. Matsuoka and M. Akashi, *Macromolecules*, 2013, **46**, 6187–6194.

[36] B. S. Murray and D. A. Fulton, *Macromolecules*, 2011, **44**, 7242–7252.

[37] B. T. Tuten, D. Chao, C. K. Lyon and E. B. Berda, *Polym. Chem.*, 2012, **3**, 3068–3071.

[38] A. Sanchez-Sanchez and J. A. Pomposo, *Part. Part. Syst. Charact.*, 2014, **31**, 11–23.

[39] A. Sanchez-Sanchez, D. A. Fulton and J. A. Pomposo, *Chem. Commun.*, 2014, **50**, 1871–1874.

[40] A. Sanchez-Sanchez, I. Perez-Baena and J. A. Pomposo, *Molecules*, 2013, **18**, 3339–3355.

[41] T. Terashima, T. Mes, T. F. A. De Greef, M. A. J. Gillissen, P. Besenius, A. R. A. Palmans and E. W. Meijer, *Journal of the American Chemical Society*, 2011, **133**, 4742–4745.

[42] E. Huerta, P. J. M. Stals, E. W. Meijer and A. R. A. Palmans, *Angewandte Chemie International Edition*, 2013, **52**, 2906–2910.

[43] A. Sanchez-Sanchez, S. Akbari, A. J. Moreno, F. Lo Verso, A. Arbe, J. Colmenero and J. A. Pomposo, *Macromol. Rapid Commun.*, 2013, **34**, 1681–1686.

[44] M. A. J. Gillissen, I. K. Voets, E. W. Meijer and A. R. A. Palmans, *Polym. Chem.*, 2012, **3**, 3166–3174.

[45] N. Hosono, M. A. J. Gillissen, Y. Li, S. S. Sheiko, A. R. A. Palmans and E. W. Meijer, *J. Am. Chem. Soc.*, 2012, **135**, 501–510.

[46] A. J. Moreno, F. Lo Verso, A. Sanchez-Sanchez, A. Arbe, J. Colmenero and J. A. Pomposo, *Macromolecules*, 2013, **46**, 9748–9759.

[47] J. W. Liu, M. E. Mackay and P. M. Duxbury, *EPL*, 2008, **84**, 46001.

[48] J. W. Liu, M. E. Mackay and P. M. Duxbury, *Macromolecules*, 2009, **42**, 8534–8542.

[49] F. Ferrante, F. Lo Celso and D. Duca, *Colloid Polym. Sci.*, 2012, **290**, 1443–1450.

[50] K. Kremer and G. S. Grest, *J. Chem. Phys.*, 1990, **92**, 5057.

[51] J. A. Izaguirre, D. P. Catarello, J. M. Wozniak and R. D. Skeel, *J. Chem. Phys.*, 2001, **114**, 2090.

[52] W. Smith, T. R. Forester and I. T. Todorov, *The DL_POLY_2 User Manual, Version 2.19*, STFC Daresbury Laboratory Daresbury, UK, 2009.

[53] D. R. Rottach, J. G. Curro, J. Budzien, G. S. Grest, C. Svaneborg and R. Everaers, *Macromolecules*, 2006, **39**, 5521–5530.

[54] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*, Oxford University Press, Oxford, UK, 1986.

[55] M. Rubinstein and R. H. Colby, *Polymer Physics*, Oxford University Press, Oxford, UK, 2003.

[56] J. Rudnick and G. Gaspari, *Science*, 1987, **237**, 384–389.

[57] E. J. Rawdon, J. C. Kern, M. Piatek, P. Plunkett, A. Stasiak and K. C. Millett, *Macromolecules*, 2008, **41**, 8281–8287.

[58] D. Chao, X. Jia, B. Tuten, C. Wang and E. B. Berda, *Chem. Commun.*, 2013, **49**, 4178–4180.

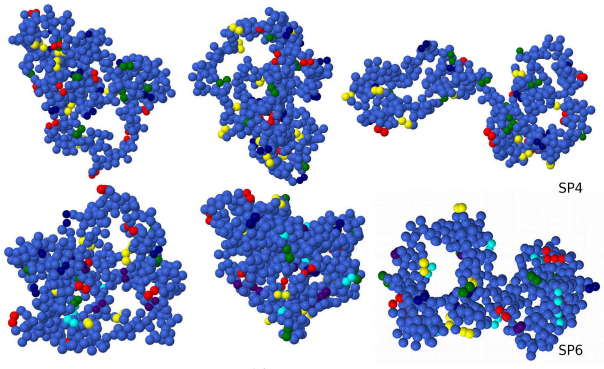
[59] P. G. Khalatur and A. R. Khokhlov, *Adv. Polym. Sci.*, 2006, **195**, 1–100.

[60] G. Zhang and C. Wu, *Adv. Polym. Sci.*, 2006, **195**, 101–176.

For Table of Contents Use Only

Title: 'Typical topologies of cross-linked nanoparticles obtained by orthogonal folding of single chain polymer precursors. The precursor consists of a linear backbone of beads, each of them being attached to a side group. A fixed fraction of the side groups contain a linker in the free end bead. Here the number of different chemical species of the linkers is 4 (top) and 6 (bottom). Dark blue beads correspond to inactive monomers. Beads of other colours correspond to the reactive linkers (a different colour for each chemical species, note the pairs of bonded linkers).'

Authors: Federica Lo Verso, José A. Pomposo, J. Colmenero, and Angel J. Moreno



TOC graph