Impacts of fullerene derivatives on regulating the structure and assembly of collagen molecules†

Xiaohui Yin,‡[ab] Lina Zhao,‡[a] Seung-gu Kang,[c] Jun Pan,[ad] Yan Song,[a] Mingyi Zhang,[a] Gengmei Xing,[a] Fei Wang,[b] Jingyuan Li,*[a] Ruhong Zhou*[ce] and Yuliang Zhao*[a]

During cancer development, the fibrous layers surrounding the tumor surface get thin and stiff which facilitates the tumor metastasis. After the treatment of metallofullerene derivatives $Gd@C_{82}(OH)_{22}$, the fibrous layers become thicker and softer, the metastasis of tumor is then largely suppressed. The effect of $Gd@C_{82}(OH)_{22}$ was found to be related to their direct interaction with collagen and the resulting impact on the structure of collagen fibrils, the major component of extracellular matrices. In this work we study the interaction of $Gd@C_{82}(OH)_{22}$ with collagen by molecular dynamics simulations. We find that $Gd@C_{82}(OH)_{22}$ can enhance the rigidity of the native structure of collagen molecules and promote the formation of an oligomer or a microfibril. The interaction with $Gd@C_{82}(OH)_{22}$ may regulate further the assembly of collagen fibrils and change the biophysical properties of collagen. The control run with fullerene derivatives $C_{60}(OH)_{24}$ also indicates that $C_{60}(OH)_{24}$ can influence the structure and assembly of collagen molecules as well, but to a lesser degree. Both fullerene derivatives can form hydrogen bonds with multiple collagen molecules acting as a "fullerenol-mediated bridge" that enhance the interaction within or among collagen molecules. Compared to $C_{60}(OH)_{24}$, the interaction of $Gd@C_{82}(OH)_{22}$ with collagen is stronger, resulting in particular biomedical effects for regulating the biophysical properties of collagen fibrils.

Introduction

The applications of fullerene derivatives in the biomedical field, such as bioimaging, drug delivery and antitumor therapy, have obtained a lot of attention in recent years.[1,2] The metallofullerene derivative, $Gd@C_{82}(OH)_{22}$, shows great potential in cancer therapy with several advantages over conventional antitumor medicine. It can not only inhibit tumor growth more effectively, but also possesses low toxicity *in vivo* and *in vitro*.[2–7] It is known that cancer cells can induce the degradation of extracellular matrices (ECM), which facilitates the invasiveness of tumor.[8–12] Normally, the content of collagen, an important ECM component,[13–16] rapidly decreases as the tumor weight increases.[17] The anti-tumor capability of $Gd@C_{82}(OH)_{22}$ is

thought to be related to the inhibition of production and activity of matrix metalloproteinases (MMPs) which degrade various ECM proteins and participate in neovascularization and angiogenesis.[8,9] It may enhance the drug selectivity in targeting MMPs as well.[8] After the treatment of $Gd@C_{82}(OH)_{22}$ the thickness of the fibrous layer surrounding the tumor surface significantly increases, resulting in the formation of a fibrous cage to imprison the tumor tissue and prevent the metastasis of the tumor. Meanwhile, recent experimental studies have shown that $Gd@C_{82}(OH)_{22}$ can also directly interact with collagen, the major component of extracellular matrices, and affect the structure and biophysical properties of collagen fibers: the density and stiffness of the collagen matrix remarkably decrease.[9,18] However, the exact molecular mechanism underlying this important process is still largely unknown.

In this work, we perform detailed molecular dynamic simulations to investigate the interaction of $Gd@C_{82}(OH)_{22}$ with molecular collagen and its impact on the structural stability and assembly of the protein. Tropocollagen molecules, the basic unit of collagen with a triple helical structure, are composed of three polypeptide chains. Tropocollagens associate into a microfibril formation, and further assemble to collagen fibrils. Another fullerene derivative $C_{60}(OH)_{24}$ with a similar structure and functional groups is also studied for comparison. We find that both fullerene derivatives can stably bind to the collagen molecule and promote an ordered

[a] CAS Key Laboratory for Biomedical Effects of Nanomaterials & Nanosafety, Institute of High Energy Physics, Chinese Academy of Sciences (CAS), Beijing 100049, China. E-mail: lijingyuan@ihep.ac.cn; zhaoyuliang@ihep.ac.cn

[b] Department of Physics, Beijing Institute of Technology, Beijing 100081, China

[c] IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598, USA. E-mail: ruhongz@us.ibm.com

[d] College of Chemistry and Chemical Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

[e] Department of Chemistry, Columbia University, New York, New York 10027, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3nr01469j

‡ These authors contributed equally to this work.

triple-helix structure. The adhered fullerene derivatives can facilitate the assembly process of molecular collagen which can also enhance the structural stability of the collagen monomer. Interestingly, it has often been reported that the interaction of nanoparticles with protein can disturb the protein structure or induce protein unfolding and/or misfolding,[19–22] while, as indicated in this work, the nanoparticle may also enhance the native structure and assembly of protein. As revealed by our simulation result, both fullerene derivatives can simultaneously interact with multiple polypeptide chains or even with multiple collagen molecules, and enhance the intra- or inter-molecule interaction by means of a "fullerenol-mediated bridge" between the collagen peptides. Moreover, the interference of fullerene derivatives on the the interactions among collagen molecules may affect further the assembly of collagen fibrils, resulting in reduction of the mechanical stability (*i.e.* stiffness) of the collagen matrix. Our simulation results are qualitatively consistent with the experimental observations on the binding affinity of fullerenol to collagen (biolayer interferometry, BLI), secondary structure changes of the collagen molecules (circular dichroism, CD), promoted initial assembly process (turbidity assay), and disturbed collagen fiber structure (atomic force microscopy, AFM).[18]

Compared to the $C_{60}(OH)_{24}$, the impact of $Gd@C_{82}(OH)_{22}$ on the structural stability and assembly of collagen molecules is more significant. Our simulation observations are consistent with the experimental observations. This is found to be largely attributed to different distributions of the hydroxyl groups and the surface charge on the carbon cage. Indeed, $Gd@C_{82}(OH)_{22}$ can have an efficient hydrophobic, electrostatic and hydrogen bond interaction with the collagen molecule, implying the importance of physicochemistry in the *de novo* design of nanoparticles for specific bioactivity.

System and method

The crystal structure of the collagen molecule (PDB code: 1WZB) is used for this study. It is composed of three polypeptide chains (chain 1, 2 and 3, respectively) with a repeating sequence "X-Y-G". There are 29 amino acids in chain 1 and chain 2, 28 amino acids in chain 3, respectively, resulting in a 8.6 nm long triplex. In our simulation, the collagen triplex is constructed with Pro at position X instead of Hyp in the crystal structure. The corresponding sequences of three chains are shown in Table 1.

Two related fullerene derivatives, $Gd@C_{82}(OH)_{22}$ and $C_{60}(OH)_{24}$, are studied in this work, with $C_{60}(OH)_{24}$ for comparison. The optimized structures and atomic partial charges of the fullerene derivatives $Gd@C_{82}(OH)_{22}$ and $C_{60}(OH)_{24}$ (Fig. 1b) are taken from the results of density functional theory (DFT) calculations.[23,24] Due to an embedded Gd^{3+} ion, there is a 3e negative charge on the $[C_{82}(OH)_{22}]$ cage. The morphology of $Gd@C_{82}(OH)_{22}$ is ellipsoidal with hydroxyl groups distributed evenly on the surface of the carbon cage, while $C_{60}(OH)_{24}$ exhibits a discus-shaped structure with hydroxyl groups located around the equator of the carbon cage.[23] The different distributions of hydroxyl groups and surface charge considerably affect the interaction of fullerene derivatives with the collagen triplex, which will be discussed below.

We configure two different systems for each fullerene derivative to study their impact on the structural stability and assembly of collagen triplexes. The structural stability of the collagen triplex has been studied with a system of one collagen triplex surrounded by 14 $Gd@C_{82}(OH)_{22}$ or $C_{60}(OH)_{24}$, where the fullerene derivatives are set to be at least 15 Å away from the central collagen (denoted as Monomer-Gd and Monomer-C60, respectively). For the collagen assembly process, we used four collagen triplexes with 20 $Gd@C_{82}(OH)_{22}$ or $C_{60}(OH)_{24}$ located at least 15 Å away from any collagen (denoted as Tetramer-Gd and Tetramer-C60, respectively). Initially, the collagen triplexes are set to be parallely separated from each other with at least one water layer, while the quasi-hexagonal packing of the collagen assembly remains.[25] In addition, we also ran two control systems with one collagen triplex and four collagen triplexes only, respectively denoted by Monomer-0 and
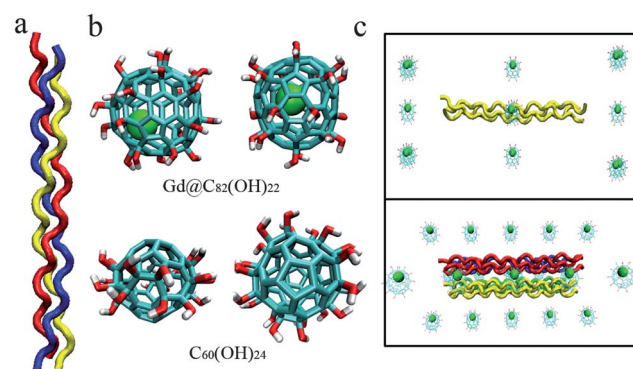

Fig. 1 (a) Structure of the collagen triplex. Three peptide chains are depicted in blue, red and yellow, separately. (b) The fullerene derivatives $Gd@C_{82}(OH)_{22}$ (upper panel) and $C_{60}(OH)_{24}$ (lower panel), the Gd atom is presented as a green ball inside the fullerenol cage. (c) Initial configurations of the simulation (upper panel: Monomer-Gd; lower panel: Tetramer-Gd): $Gd@C_{82}(OH)_{22}$ molecules are placed at least 15 Å away from the collagen triplex.

Table 1 Residue sequence of each collagen polypeptide chain[a]

Peptide chain	N-region	Mid-region	C-region	Residue number
Chain 1	P O G P O	G P O G P O G P O G P O G P O G P O G	P O G P O	29
Chain 2	P O G P O	G P O G P O G P O G P O G P O G P O G	P O G P O	29
Chain 3	O G P O	G P O G P O G P O G P O G P O G P O G	P O G P O	28

[a] P: Pro, O: Hyp, G: Gly.

Tetramer-0, in order to get information of their intrinsic stability and assembly process. All systems are then solvated using the TIP3P water model. The size of the solvation boxes is 78 Å × 72 Å × 146 Å for monomer systems with ∼81 000 atoms, and 100 Å × 80 Å × 126 Å for tetramer systems with ∼100 000 atoms.

The simulations are performed with NAMD2 molecular-dynamics package.[26] CHARMM22 all-atom force field with the addition of parameters specific for Hyp is used.[27] The electrostatic interactions are calculated by the particle-mesh Ewald (PME) method with a cut-off distance of 12 Å. The cut-off distance for the calculation of van der Waals interactions is 12 Å. Periodic boundary conditions are applied in all directions. All simulations are carried out at constant temperature (300 K) and pressure (1 bar).

In order to study possible adsorption of fullerene derivatives, we restrained Cα atoms of 6th and 24th of each chain at their initial positions for the first 20 ns for both Monomer and Tetramer systems, and ran the simulations with no restraints for the subsequent 60 ns. For the control systems, we obtained 60 ns long trajectories without any restraints.

The structural stability of the collagen triplex is characterized by means of the time evolution of backbone root-mean-square deviation (rmsd) and root-mean-square fluctuation (rmsf) with respect to the initial configuration, together with the number of intra-protein backbone hydrogen bonds (H-bonds). The interaction of the fullerene derivative with collagen is analyzed by the number of adsorbed fullerene derivatives, and the hydrogen bonds between the fullerene derivatives and the collagen molecule. The adhesion is defined when the fullerene derivative is in direct contact with collagen as measured with the minimum distance between any heavy atoms less than 3.5 Å. The hydrogen bonds are measured with widely used geometric definition (i.e. donor–acceptor distance ≤ 3.5 Å and acceptor–donor–hydrogen angle ≤ 30°). Because of the linear structure of the collagen triplex, the position of the residue is related to the residue index. For convenience of the discussion, the collagen triplex can be divided into three regions according to the residue index: N-terminal region (Nter-region; residue index = 1–5), C-terminal region (Cter-region; residue index = 25–29) and mid-region (residue index = 6–24) (shown in Table 1). All snapshots are rendered with VMD.[28]

Results and analysis

The impact of adsorbed fullerene derivatives to collagen triplex

Our simulation shows that metallofullerenol $Gd@C_{82}(OH)_{22}$ has strong preference to adsorbing on the surface of the collagen triplex. Fig. 2b shows a time profile of the direct contact of $Gd@C_{82}(OH)_{22}$ on the collagen triplex in the three regions (i.e. Nter-, mid- and Cter-regions). There is at least one $Gd@C_{82}(OH)_{22}$ nanoparticle adsorbed on any region of the triplex in the initial configuration (taken from the result of the first 20 ns of simulation) and the adsorbed $Gd@C_{82}(OH)_{22}$ nanoparticles strongly bind to the triplex during the subsequent 60 ns of simulation. The numbers of adhering $Gd@C_{82}(OH)_{22}$
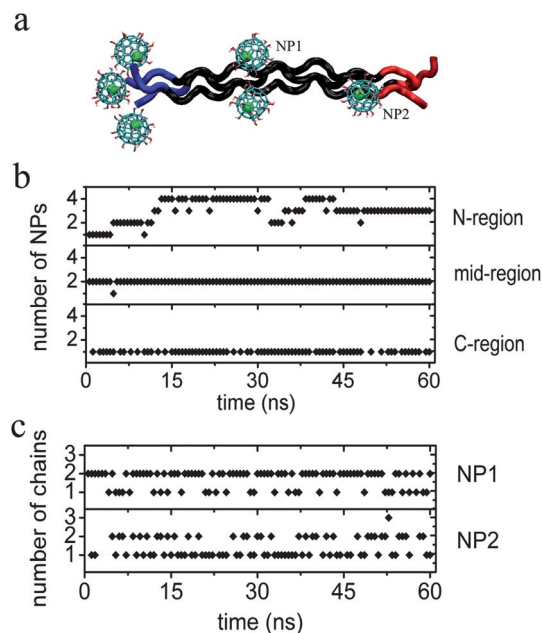

Fig. 2 (a) Representative snapshot of the collagen triplex bound by $Gd@C_{82}(OH)_{22}$ ($t = 50$ ns). The Nter-, mid- and Cter-region of the collagen triplex is colored in blue, black and red, respectively. (b) Number of adsorbed $Gd@C_{82}(OH)_{22}$ molecules in each region. (c) Number of collagen peptide chains with which the representative $Gd@C_{82}(OH)_{22}$ molecules interact. The adsorbed $Gd@C_{82}(OH)_{22}$ can simultaneously interact with two peptide chains, acting as the "fullerenol-mediated bridge" between two chains.

can be up to 4 and 2 in the Nter-region and mid-region, respectively, while a relatively low one $Gd@C_{82}(OH)_{22}$ molecule is found in Cter-region (i.e., see Fig. 2a). It is interesting to note that the number of adhering nanoparticles per residue in Nter-region is 0.6, much more than other regions, with values of 0.1 and 0.2. As discussed above, the fullerenol cage is highly negatively charged (−3e), and there is a strong electrostatic attraction between $Gd@C_{82}(OH)_{22}$ and the positively charged N-terminus of the collagen peptide chains, while it is worth noting that the $Gd@C_{82}(OH)_{22}$ can still strongly bind to Cter-region. The stable adsorption to the collagen triplex can be largely attributed to the strong hydrophobic and hydrogen bond interaction of $Gd@C_{82}(OH)_{22}$ with the collagen triplex, which will be discussed in detail below. Similarly, the stable adsorption of $C_{60}(OH)_{24}$ on the surface of cthe ollagen triplex is also observed in system Monomer-C60, and there are six $C_{60}(OH)_{24}$ molecules adsorbed to the collagen triplex.

The structural impact of adherence of fullerene derivatives is assessed with backbone rmsd and rmsf. The rmsd over all collagen residues is shown more or less as similar regardless of the existence of $Gd@C_{82}(OH)_{22}$, albeit a noticeable decrease from 3.1 Å (Monomer-0) to 2.77 Å (Monomer-Gd) (Fig. 3a). A more detailed analysis with the residue-specific rmsd, however, shows that $Gd@C_{82}(OH)_{22}$ have a different effect on the collagen triplex, depending on the residue position. Even with, in general, more deviation in the terminal regions from the crystal structure (i.e., Monomer-0), Fig. 3b shows that $Gd@C_{82}(OH)_{22}$ makes the terminal regions, especially in the
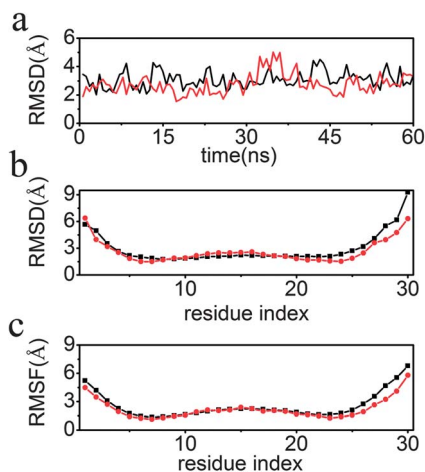
Fig. 3 (a) Time evolution of backbone rmsd of the collagen triplex with reference to the initial conformation in systems Monomer-0 (black) and Monomer-Gd (red). The corresponding averaged rmsd (b) and rmsf (c) of each residue in the collagen triplex. Conformational fluctuation of the collagen triplex is relatively suppressed after an addition of $Gd@C_{82}(OH)_{22}$, especially in the terminal regions.

C-terminus, closer to the crystal structure, while slightly disturbing the middle one.

In addition, we compute the averaged backbone rmsf of each residue to characterize the flexibility of the collagen triplex. Similar to the rmsd patterns for the termini, the averaged rmsf in both terminal regions are inherently larger than that in the *mid-region*, regardless of the existence of nanoparticles. The addition of $Gd@C_{82}(OH)_{22}$, however, tends to stabilize the collagen peptide more effectively in both terminal regions. Taken together, the adhering $Gd@C_{82}(OH)_{22}$ can suppress the thermal fluctuation and gently enhance the triple helical structure of molecular collagen, largely attributed to stabilization in the terminal residues.

As for the case of Monomer-C60, we also find that the averaged values of rmsd and rmsf of the residues are smaller than those of the control system. The averaged rmsd of the collagen triplex in Monomer-C60 is 2.89 Å, comparable to the value shown in Monomer-Gd. Similar to the case of $Gd@C_{82}(OH)_{22}$, the decrement is more obvious in both terminal regions than that in the *mid-region*.

It has been well accepted that the hydrogen bonds between peptide chains play a crucial role for the stability of the collagen triplex structure.[27,29–33] In this regard, we analyzed residue-specific H-bonds among the three chains in the collagen triplex, and compared their distributions along the three separated regions. In the control Monomer-0, the average H-bond numbers in *Nter*- and *Cter-region* are 0.8 and 1.4, respectively. The corresponding numbers increase to 0.9 (*Nter-region*) and 1.6 (*Cter-region*) in system Monomer-Gd. On the other hand, the average H-bond number in *mid-region* decreases from 6.0 (Monomer-0) to 5.0 (Monomer-Gd), whereas the average number of inter-chain H-bonds in system Monomer-C60 is 0.60 (*Nter-region*), 5.0 (*mid-region*) and 1.83 (*Cter-region*), respectively, where the hydrogen bonding is only enhanced in *Cter-region*. In all cases, the average number of hydrogen bonds per residue in

terminal regions is sizably smaller than that in the *mid-region*, largely due to the pronounced thermal fluctuation in the terminal regions. It is not surprising that the average H-bond number of the whole collagen triplex is slightly reduced after adsorption of both $C_{60}(OH)_{24}$ and $Gd@C_{82}(OH)_{22}$ molecules. The fullerene derivatives, having hydrogen bonds with the collagen triplex, might slightly disturb the inter-peptide hydrogen bond network, yet the conformational integrity still remains stable. It is interesting to note that the number of hydrogen bonds in the terminal regions still increased with the fullerene derivatives, which can be largely attributed to the suppression of thermal fluctuation by the nanoparticle adsorption.

The impact of fullerene derivatives on reducing the thermal fluctuation of the collagen triplex can be highly related to their stable adsorption on collagen. Compared to $C_{60}(OH)_{24}$, $Gd@C_{82}(OH)_{22}$ is relatively more hydrophobic: there are only one-fourth of carbon atoms functionalized by hydroxyl groups. Meanwhile, the negatively charged fullerenol cage facilitates the electrostatic interaction with collagen, especially with *Nter*-region. The stronger hydrophobic and electrostatic interactions facilitate a more stable adsorption of $Gd@C_{82}(OH)_{22}$. Besides, even the distribution of hydroxyl groups of $Gd@C_{82}(OH)_{22}$ favors an efficient formation of hydrogen bond interaction with collagen.

Hence, we compute the number of hydrogen bonds (n_{HB}) as well as the number of contacting hydroxyl groups (n_{OH}) between $Gd@C_{82}(OH)_{22}$ and the collagen molecule, where the contact is counted when the smallest distance between the hydroxyl oxygen of $Gd@C_{82}(OH)_{22}$ and any collagen heavy atom is less than 3.5 Å. Since the exchange of nanoparticle still occurs especially in the terminal region, in our analysis we only considered nanoparticles adsorbed on the surface of the collagen triplex for at least 20 ns. All results are summarized in Table 2. Most $Gd@C_{82}(OH)_{22}$ molecules show about 40% of n_{HB}/n_{OH} ratio with average values of $\langle n_{HB} \rangle = 1.8$ and $\langle n_{OH} \rangle = 4.6$. The maximum ratio of n_{HB}/n_{OH} can be up to 62.7%. It is important to note that about one-fifth of the hydroxyl groups get in contact with collagen, and half among the contacts are participating in forming hydrogen bonds with collagen. In contrast, $C_{60}(OH)_{24}$ involves fewer hydroxyl groups in contact with collagen with $\langle n_{OH} \rangle = 3.0$. Moreover, the hydrogen bond participation ratio n_{HB}/n_{OH} is only 14%, much smaller than that of $Gd@C_{82}(OH)_{22}$.

Table 2 Number of H-bonds and hydroxyl groups of adhering fullerene derivatives

Monomer-Gd	NP1	NP2	NP3	NP4	NP5
n_{HB}	1.34	1.80	1.54	2.47	1.74
n_{OH}	5.49	4.67	4.31	3.94	4.50
n_{HB}/n_{OH}	24.4%	38.5%	35.7%	62.7%	38.7%
Monomer-C60	NP1	NP2	NP3	NP4	NP5
n_{HB}	0.65	0.47	0.28	0.20	0.44
n_{OH}	2.49	4.24	2.29	2.65	3.26
n_{HB}/n_{OH}	26.1%	11.1%	12.2%	7.5%	13.4%

Thus, the number of hydrogen bonds formed with collagen is also basically low, $\langle n_{HB} \rangle = 0.4$ which is less than one-forth of the corresponding value for $Gd@C_{82}(OH)_{22}$. The maximum value of n_{HB} is 0.65, still much smaller than all adsorbed $Gd@C_{82}(OH)_{22}$ molecules. In contrast to $Gd@C_{82}(OH)_{22}$, the distribution of hydroxyl groups of $C_{60}(OH)_{24}$ is inhomogeneous. And the nonpolar region of the carbon cage shows stronger preference to contact with protein, resulting in fewer number of a contacted hydroxyl group and much less hydrogen bonds formed with collagen.

More interestingly, we find that $Gd@C_{82}(OH)_{22}$ interact with more than one peptide chain resulting from their comparable size to a triplex radius. The interaction of the fullerene derivative with multiple chains raises a "fullerenol-mediated bridge" mechanism to explain how $Gd@C_{82}(OH)_{22}$ stabilizes the collagen structure. Fig. 2c shows the numbers of peptide chains that simultaneously form hydrogen bonds with one $Gd@C_{82}(OH)_{22}$ for two chosen representative $Gd@C_{82}(OH)_{22}$. Such a bridge occurs more frequently in the terminal regions. In both cases, the $Gd@C_{82}(OH)_{22}$ molecules form stable hydrogen bonds with two peptide chains. The "bridging" $Gd@C_{82}(OH)_{22}$ can enhance the connection between the peptide chains, thus suppressing conformational fluctuation. Meanwhile, the interaction of $Gd@C_{82}(OH)_{22}$ with collagen is still modest with the amplitude of around two hydrogen bonds. In this way, such "fullerenol-mediated bridge" can improve the ordered triple helical structure, without considerably disturbing the original protein structure. As for the $C_{60}(OH)_{24}$ molecules, such "bridging" roles are more modest due to their relatively less effective hydrogen bonds with protein.

Impact of the fullerene derivatives on the assembly of collagen triplexes

Besides impact on the structural stability of the collagen triplex, we also study the effect of fullerene derivatives on the formation and stability of the collagen oligomer. For this purpose, the assembly process of four collagen triplexes is studied. Four collagen triplexes (labeled as A, B, C, and D) are initially placed parallel to each other with separation of at least one water layer. The cross-section of the four triplexes forms a parallelogram, corresponding to the quasihexagonal packing of the collagen molecule within the collagen fibrils.

Similar to the case of the monomer system, $Gd@C_{82}(OH)_{22}$ molecules can stably bind to collagen triplexes, where 19 out of 20 nanoparticles are eventually adsorbed on the collagen triplexes. Fig. 4a shows the time evolution of the backbone rmsd calculated for the four-triplex bundle with respect to the initial configuration. In the case of Tetramer-0, the rmsd significantly increases to 8 Å within 10 ns and as high as ~12 Å in the 60 ns simulation. On the other hand, the bundle rmsd is only 4 Å in Tetramer-Gd, much smaller than the control. The structural inspection reveals that the large rmsd in Tetramer-0 is largely due to a less restricted motion among the constituent collagen triplexes, while the motion is more likely restricted by intermolecular mediation of $Gd@C_{82}(OH)_{22}$, hence facilitating the ordered arrangement among the collagen triplexes. Similarly,

we find that $C_{60}(OH)_{24}$ stabilize the rmsd of a four-triplex bundle with rmsd ~4 Å, meaning the addition of $C_{60}(OH)_{24}$ also can enhance the ordered structure of the four-triplex bundle, albeit it is not as effective as $Gd@C_{82}(OH)_{22}$. The relative transversal motion of the collagen triplex can still be observed in this system (more discussion below).

In another perspective, we measure the crossing angle among the constituent collagen triplexes. The angle is defined by two principal vectors using heavy atoms (residues index 5 to 24) for any chosen pair of collagen triplexes (Fig. 5a). The angles of all six triplex pairs (A–B, A–C, A–D, B–C, B–D, and C–D) are shown in Fig. 5. In Tetramer-0, the crossing angles seem more likely to fluctuate and diverge from the original parallel configurations (>25°). Although the divergences among the three triplexes B, C and D become reduced at the end of the simulation, the relative orientation with triplex A is not improved as reaching to ~40° from others (Fig. 5a). On the other hand, all angles of the triplex pairs in Tetramer-Gd slightly increase to ~8° with much smaller fluctuation, largely retaining the initial parallel configuration of the collagen triplexes. This is also similar in Tetramer-C60 with an average crossing angle of ~10°, implying that $C_{60}(OH)_{24}$ also contribute to the parallel packing of the collagen triplexes.

Moreover, the assembly process is studied by calculating the separation between collagen triplexes, since the distance of the center-of-masses over the whole triplexes may not correctly reflect the association process due to their relative rotation. Given by this, the triplex is equally divided into three segments, *i.e.* residue index 1 to 10, 11 to 20, 21 to 29. The distances between the corresponding segments are calculated and the averaged value is used to characterize the separation of a triplex pair. In the initial configuration the cross-section of the four-triplex bundle forms a parallelogram (Fig. 6a), and the separations of the triplex pairs are 20 Å (A–B), 16 Å (B–C), 19 Å (C–D), 16 Å (D–A), and 21 Å (A–C), 29 Å (B–D). The triplex pairs A–C and B–D correspond to the diagonals of the parallelogram. In the case of Tetramer-0, distances of the triplex pairs change significantly. At the end of the simulation, the separations are 23 Å (A–B), 13 Å (B–C), 13 Å (C–D), 19 Å (D–A), and 21 Å (A–C), 24 Å (B–D), respectively. This shows a clear deviation from the parallelogram arrangement of collagen triplexes, requiring further adjustment of the quasihexagonal packing structure.[25] On the contrary, the parallelogram pattern of the cross-section of the triplex bundle is largely retained in system Tetramer-Gd. The relative separations of the triplex pairs A–B, B–C and A–C, B–D cooperatively decrease to 16, 16 and 14, 27 Å correspondingly, while the distances of triplex pairs C–D, D–A do not change. As for Tetramer-C60, the corresponding separations change to 18 Å (A–B), 13 Å (B–C), 15 Å (C–D), 19 Å (D–A) and 21 Å (A–C), 26 Å (B–D). Although similar to $Gd@C_{82}(OH)_{22}$, the addition of $C_{60}(OH)_{24}$ enhances the longitudinal parallel arrangement for collagen bundles, and it is not as effective as $Gd@C_{82}(OH)_{22}$ in controlling the transversal motion that retains the parallelogram arrangement. That is, the adsorbed $Gd@C_{82}(OH)_{22}$ molecules more effectively suppress the relative transversal motion of the triplex during assembly and facilitate the assembly process of the collagen triplex.
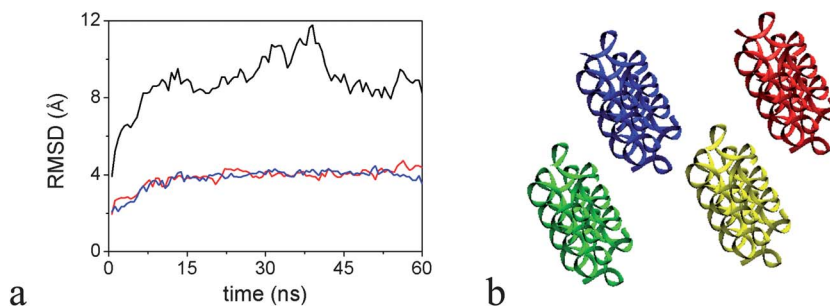
Fig. 4 (a) Backbone rmsd of the whole four-triplex bundle with reference to the initial conformation in system Tetramer-0 (black), Tetramer-Gd (red) and Tetramer-C60 (blue). The addition of both $Gd@C_{82}(OH)_{22}$ and $C_{60}(OH)_{24}$ significantly enhances the ordered packing structure of the collagen triplexes tetramer. (b) Initial parallelogram structural arrangement of the four collagen triplexes.
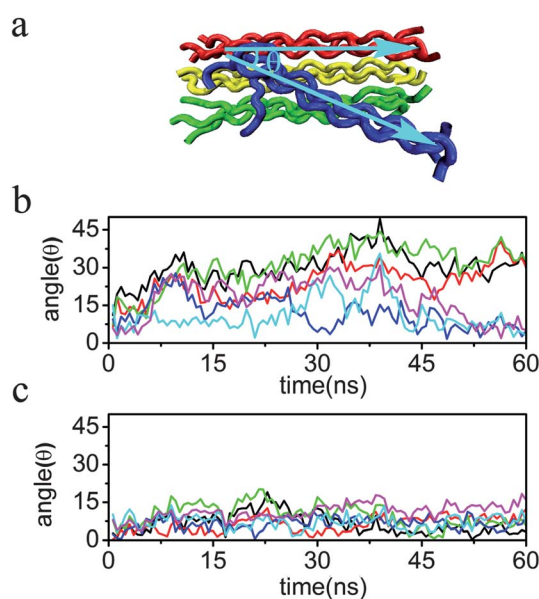

Fig. 5 (a) Snapshot of the collagen triplex tetramer at $t = 60$ ns, collagen triplexes A, B, C and D are colored in blue, red, yellow, and green respectively. Linear collagen triplexes can be fitted to the vectors (cyan). And the angle θ between vectors represents the relative orientation of the collagen triplexes. Time evolution of the angle between all six triplex pairs in system Tetramer-0 (b) and system Tetramer-Gd (c). The angles between triplex pairs A–B, A–C, A–D, B–C, B–D, and C–D are colored to black, red, green, blue, cyan and magenta, respectively. Considerable rotation of the collagen triplexes within a triplex tetramer is observed in system Tetramer-0, while the collagen triplexes keep parallel to each other in system Tetramer-Gd.
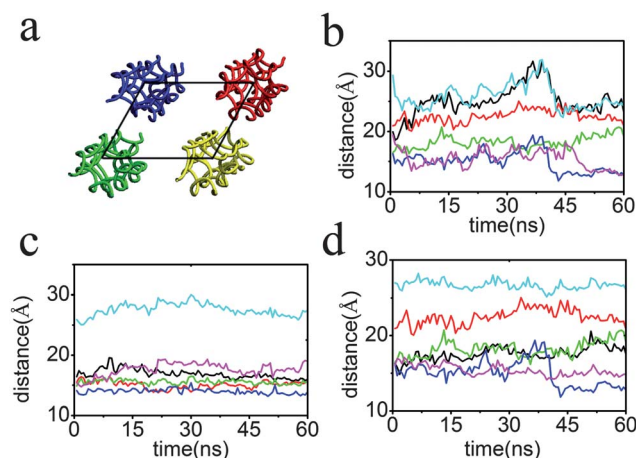

Fig. 6 (a) Cross-section of the initial configuration of the triplex tetramer. Collagen triplexes A, B, C and D are colored in blue, red, yellow, green, respectively. Separation of all six triplex pairs (A–B, black; A–C, red; A–D, green; B–C, blue; B–D, cyan; C–D, magenta) in system Tetramer-0 (b), Tetramer-Gd (c) and Tetramer-C60 (d).

Our analysis shows that $Gd@C_{82}(OH)_{22}$ and $C_{60}(OH)_{24}$ stabilize both a single collagen triplex and a four-triplex bundle, albeit $Gd@C_{82}(OH)_{22}$ is more effective. Interestingly, our findings are supported by our experimental results. The triple helical structure of tropocollagen possesses characteristic circular dichroism (CD) spectra with a positive peak around
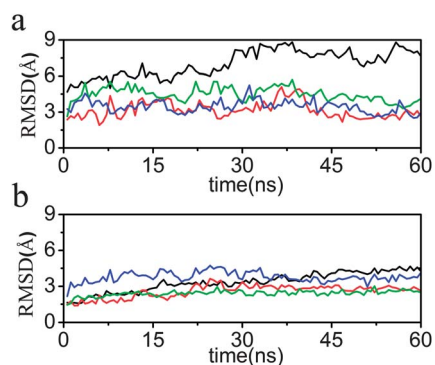
The conformational fluctuation of each collagen triplex within the bundle is analyzed by the backbone rmsd of each triplex with respect to the crystal structure (Fig. 7). In system Tetramer-0, the triplexes generally show higher rmsd from the reference structure. Especially, the triplex A showed the largest deviation of ∼7 Å, possibly due to a lowered inter-chain stability caused by its large rotational and transversal motion. On the other hand, the rmsd of the triplexes in system Tetramer-Gd are smaller than that in Tetramer-0. The averaged rmsd of the triplexes is about 3.0 Å, implying that the individual triplexes are stabilized by an interaction with $Gd@C_{82}(OH)_{22}$ as well. The averaged rmsd of the collagen triplex similarly decreases after the addition of $C_{60}(OH)_{24}$.

Fig. 7 Backbone rmsd of each collagen triplex (A, black; B, red; C, green; D, blue.) in system Tetramer-0 (a) and Tetramer-Gd (b). The structural fluctuation of each triplex in system Tetramer-0 is more significant than in system Tetramer-Gd.

220 nm, a negative peak around 197 nm, and a crossover around 213 nm.[34] We observe that $Gd@C_{82}(OH)_{22}$ and $C_{60}(OH)_{24}$ intensify both peaks of the collagen solution, with more significant enhancement in $Gd@C_{82}(OH)_{22}$, meaning an increase of the order of the triple-helix structure.[18] As discussed above, even though the simulation timescale of the sub-hundred nanoscale is still limited, as compared to the full fibrogenesis process, the adsorption process of fullerenol onto collagen and the subsequent conformation change of a single molecule can be successfully captured in MD simulation. Moreover, the impact on the collagen oligomer formation can also be investigated to some degree with the current computational resources. Taken together, the interaction of fullerenol with collagen and the tendency for interference with the collagen assembly can still be observed during the current MD simulation lengths.

As mentioned above, $Gd@C_{82}(OH)_{22}$ can form hydrogen bonds with multiple polypeptide chains of the collagen triplex, acting as a "fullerenol-mediated bridge". Similarly, the $Gd@C_{82}(OH)_{22}$ adsorbed on the surface of the four-triplex bundle can simultaneously interact with different triplexes. Since the separation between triplexes is larger than the distance between polypeptide chains, $Gd@C_{82}(OH)_{22}$ molecules rarely form direct hydrogen bonds with multiple triplexes at the same time. Such bridging role requires the assistance of water molecules, for example $Gd@C_{82}(OH)_{22}$ can have a direct hydrogen bond with one triplex and the water-mediated hydrogen bond with the other. Hence, the $Gd@C_{82}(OH)_{22}$ is considered to interact with the collagen triplex if the smallest distance between the heavy atoms of $Gd@C_{82}(OH)_{22}$ and triplex is below 4 Å. The numbers of $Gd@C_{82}(OH)_{22}$ nanoparticles simultaneously interacting with any triplex pairs are calculated and shown in Fig. 8. For all collagen triplex pairs except for pair B–D in the vertices of the longer diagonal of the parallelogram, at least one $Gd@C_{82}(OH)_{22}$ is found in simultaneous mediation of two triplexes. Fig. 8 shows representative configurations of the "fullerenol-mediated bridge" at the end of the simulation. The bridges can be observed in both the terminal and *mid-regions*. Over the simulation, $Gd@C_{82}(OH)_{22}$ molecules very stably bind between two adjacent collagen triplexes, thus affecting the overall stability of the collagen bundle.

Similar to the impact on the connection between peptide chains of the collagen triplex, the "fullerenol-mediated bridge" of $Gd@C_{82}(OH)_{22}$ can also enhance the interaction between collagen triplexes and hence suppress the relative rotational and transversal motion of the triplex during the association process. So that the collagen triplexes can assemble in a more efficient manner, the ordered triplex structure of molecular collagen is enhanced. It is important to note that strong adsorption of $Gd@C_{82}(OH)_{22}$ to the collagen molecule and homogenous distribution of the hydroxyl group facilitates the formation of a "fullerenol-mediated bridge", resulting in a more significant impact on the assembly and conformational stability of the collagen triplex than $C_{60}(OH)_{24}$. Interestingly, the addition of $Gd@C_{82}(OH)_{22}$ can indeed promote the nucleation of a fibril, as revealed by an apparent increase in the turbidity of collagen solution in the lag period.[18] Our simulation results are
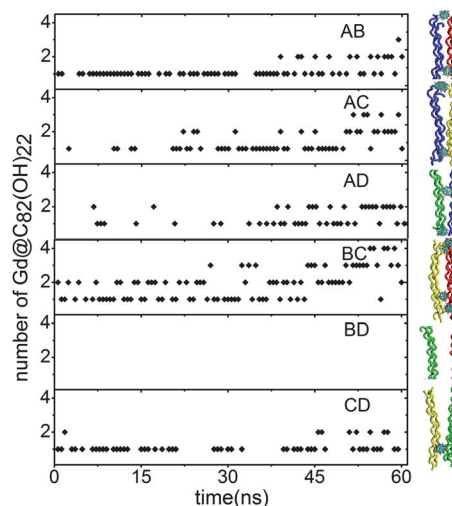

Fig. 8 Number of $Gd@C_{82}(OH)_{22}$ simultaneously contacted with collagen triplex pairs: A–B, A–C, A–D, B–C, B–D, and C–D; together with the corresponding snapshots at $t = 60$ ns (on the right side).

consistent with these experimental observations. On the other hand, the interaction of $Gd@C_{82}(OH)_{22}$ with molecular collagen should also participate into a further assembly of collagen fibrils and regulate the morphology and biophysical properties (*e.g.* radius and stiffness) of the collagen fibrous layer. Unlike the effect of enhancing the formation of an oligomer or a microfibril, the "fullerenol-mediated bridge" may disturb the well-ordered interaction among collagen molecules and induce deviation from the native periodic arrangement within collagen fibrils. Such bridging roles may serve as the molecular mechanism for a collagen fibril with an irregular periodic structure and reduced stiffness of the collagen matrix after addition of $Gd@C_{82}(OH)_{22}$.[18] Further research is required to study the impact of fullerene derivatives on the structure and mechanical stability of collagen fibrils.

Conclusion

In this work, we investigated the impact of fullerene derivatives, $Gd@C_{82}(OH)_{22}$ and $C_{60}(OH)_{24}$, on the structure and assembly of the collagen triplex by using all-atom MD simulation. A stable adhesion has been observed for both fullerene derivatives toward the collagen triplex. In systems of a single collagen triplex, the adhesion of both fullerene derivatives can relatively suppress the thermal fluctuation of the protein structure. Similarly, both nanoparticles could facilitate an assembly of collagen triplexes. During the collagen assembly, nanoparticles act as anchors to restrict the relative rotation among the collagen triplexes, favoring parallel configurations and likely stabilizing the collagen complex. $Gd@C_{82}(OH)_{22}$ was shown to further suppress transversal motion among the triplexes. Our results are in good compliance with the experimental results. Although nanoparticles have often been reported to induce a disruptive impact to the native structure and/or assembly of proteins, our simulation reveals that the interaction of a nanoparticle can also enhance the rigidity of the protein structure as well as facilitate the native protein assembly.

The structural impact on the collagen complex is largely attributed to a "fullerenol-mediated bridge" interaction in which the fullerene derivatives mediate multiple collagen peptides, utilizing hydrogen bonds. The "fullerenol-mediated bridge" can enhance the interaction between collagen molecules during the nucleation process and facilitate the formation of oligomers or microfibrils, on the other hand such interference to the interactions among the collagen molecules may disturb the native molecular arrangement within a matured collagen fibril and affect the structure and stiffness of the collagen fibril layer. The mediation of fullerene derivatives largely depends on the surface charge and hydroxyl group distribution on the fullerene derivatives. Compared to neutral $C_{60}(OH)_{24}$, the negatively charged $Gd@C_{82}(OH)_{22}$ are more available to have an electrostatic interaction with collagen. In addition, evenly distributed hydroxyl groups further facilitated effective hydrogen bond formation with collagen as well as intrinsic hydrophobic interactions with non-polar residues, which has a more significant effect on the collagen complex with $Gd@C_{82}(OH)_{22}$. This explains another facet of the molecular mechanisms of how the metallofullerenol $Gd@C_{82}(OH)_{22}$ affects the ECM network rigidity elaborated with binding on MMPs, to inhibit cancer metastasis.

Acknowledgements

References

1 T. Da Ros and M. Prato, *Chem. Commun.*, 1999, 663–669.
2 C. Y. Chen, G. M. Xing, J. X. Wang, Y. L. Zhao, B. Li, J. Tang, G. Jia, T. C. Wang, J. Sun, L. Xing, H. Yuan, Y. X. Gao, H. Meng, Z. Chen, F. Zhao, Z. F. Chai and X. H. Fang, *Nano Lett.*, 2005, **5**, 2050–2057.
3 J. X. Wang, C. Y. Chen, B. Li, H. W. Yu, Y. L. Zhao, J. Sun, Y. F. Li, G. M. Xing, H. Yuan, J. Tang, Z. Chen, H. Meng, Y. X. Gao, C. Ye, Z. F. Chai, C. F. Zhu, B. C. Ma, X. H. Fang and L. J. Wan, *Biochem. Pharmacol.*, 2006, **71**, 872–881.
4 Y. Liu, F. Jiao, Y. Qiu, W. Li, F. Lao, G. Zhou, B. Sun, G. Xing, J. Dong, Y. Zhao, Z. Chai and C. Chen, *Biomaterials*, 2009, **30**, 3934–3945.
5 X.-J. Liang, H. Meng, Y. Wang, H. He, J. Meng, J. Lu, P. C. Wang, Y. Zhao, X. Gao, B. Sun, C. Chen, G. Xing, D. Shen, M. M. Gottesman, Y. Wu, J.-j. Yin and L. Jia, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 7449–7454.
6 H. Meng, G. Xing, B. Sun, F. Zhao, H. Lei, W. Li, Y. Song, Z. Chen, H. Yuan, X. Wang, J. Long, C. Chen, X. Liang, N. Zhang, Z. Chai and Y. Zhao, *ACS Nano*, 2010, **4**, 2773–2783.
7 D. Yang, Y. Zhao, H. Guo, Y. Li, P. Tewary, G. Xing, W. Hou, J. J. Oppenheim and N. Zhang, *ACS Nano*, 2010, **4**, 1178–1186.
8 S. Kang, G. Zhou, P. Yang, Y. Liu, B. Sun, T. Huynh, H. Meng, L. Zhao, G. M. Xing, C. Y. Chen, Y. L. Zhao and R. Zhou, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 15431–15436.
9 H. Meng, G. M. Xing, E. Blanco, Y. Song, L. Zhao, B. Sun, X. Li, P. C. Wang, A. Korotcov, W. Li, X. J. Liang, C. Y. Chen, H. Yuan, F. Zhao, Z. Chen, T. Sun, Z. F. Chai, M. Ferrari and Y. L. Zhao, *Nanomedicine*, 2012, **8**, 136–146.
10 D. Yin, Z. Ge, W. Yang, C. Liu and Y. Yuan, *Cancer Lett.*, 2006, **243**, 71–79.
11 C. Lewis and J. Pollard, *Cancer Res.*, 2006, **66**, 605–612.
12 M. Egeblad and Z. Werb, *Nat. Rev. Cancer*, 2002, **2**, 161–174.
13 J. Myllyharju and K. Kivirikko, *Trends Genet.*, 2004, **20**, 33–43.
14 K. Kadler, D. Holmes, J. Trotter and J. Chapman, *Biochem. J.*, 1996, **316**, 1–11.
15 J. Myllyharju and K. Kivirikko, *Ann. Med.*, 2001, **33**, 7–21.
16 S. Viguet-Carrin, P. Garnero and P. Delmas, *Osteoporosis Int.*, 2006, **17**, 319–336.
17 M. Grabowska, *Nature*, 1959, **183**, 1186–1187.
18 Y. Song, M. Zhang, L. Zhao, X. Yin, J. Zhao, J. Li, R. He, X. Liu, Y. Zhao, J. Li and G. Xing, unpublished data.
19 K. Balamurugan, E. Singam and V. Subramanian, *J. Phys. Chem. C*, 2011, **115**, 8886–8892.
20 K. Balamurugan, R. Gopalakrishnan, S. Raman and V. Subramanian, *J. Phys. Chem. B*, 2010, **114**, 14048–14058.
21 G. Zuo, Q. Huang, G. Wei, R. Zhou and H. Fang, *ACS Nano*, 2010, **4**, 7508–7514.
22 C. Ge, J. Du, L. Zhao, L. Wang, Y. Liu, D. Li, Y. Yang, R. Zhou, Y. Zhao, Z. Chai and C. Chen, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 16968–16973.
23 H. He, L. Zheng, P. Jin and M. Yang, *Comput. Theor. Chem.*, 2011, **974**, 16–20.
24 J. Zhang, F. Li, X. Miao, J. Zhao, L. Jing, G. Yang and X. Jia, *Chem. Phys. Lett.*, 2010, **492**, 68–70.
25 J. Orgel, T. Irving, A. Miller and T. Wess, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 9001–9005.
26 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
27 P. Veld and M. Stevens, *Biophys. J.*, 2008, **95**, 33–39.
28 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
29 N. Dai, X. Wang and F. Etzkorn, *J. Am. Chem. Soc.*, 2008, **130**, 5396.
30 A. Bachmann, T. Kiefhaber, S. Boudko, J. Engel and H. Bachinger, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 13897–13902.
31 J. Engel and H. Bachinger, *Top. Curr. Chem.*, 2005, **247**, 7–33.
32 T. Gurry, P. Nerenberg and C. Stultz, *Biophys. J.*, 2010, **98**, 2634–2643.
33 S. Raman, R. Gopalakrishnan, R. Wade and V. Subramanian, *J. Phys. Chem. B*, 2011, **115**, 2593–2607.
34 U. Freudenberg, S. H. Behrens, P. B. Welzel, M. Müller, M. Grimmer, K. Salchert, T. Taeger, K. Schmidt, W. Pompe and C. Werner, *Biophys. J.*, 2007, **92**, 2108–2119.