# Hydrophilic Linkers and Polar Contacts Affect Aggregation of FG Repeat Peptides

Nicole Dölker,* Ulrich Zachariae, and Helmut Grubmüller
Max-Planck-Institute for Biophysical Chemistry, Department of Theoretical and Computational Biophysics, Göttingen, Germany

ABSTRACT Transport of large proteins into the nucleus involves two events, binding of the cargo protein to a transport receptor in the cytoplasm and passage of the cargo-transporter complex through the selective permeability barrier of the nuclear pore complex. The permeability barrier is formed by largely disordered polypeptides, each containing a number of conserved hydrophobic phenylalanine-glycine (FG) sequence motifs, connected by hydrophilic linkers of varying sequence (FG nups). How the motifs interact to form the permeability barrier, however, is not yet known. We have, therefore, carried out molecular dynamics simulations on various model FG repeat peptides to study the aggregation propensity of FG nups and the specific roles of the hydrophobic FG motifs and the hydrophilic linkers. Our simulations show spontaneous aggregation of the model nups into hydrated aggregates, which exhibit structural features assumed to be part of the permeability barrier. Our simulations suggest that short $\beta$-sheets are an important structural feature of the aggregates and that Phe residues are sufficiently exposed to allow rapid binding of transport receptors. A surprisingly large influence of the amino acid composition of the hydrophilic linkers on aggregation is seen, as well as a major contribution of hydrogen-bonding patterns.

## INTRODUCTION

In eukaryotic cells, all traffic into and out of the nucleus proceeds through nuclear pore complexes (NPCs). NPCs are large protein assemblies consisting of multiple copies of ~30 different nucleoporins (nups) (1,2). Although the exact architecture of the NPC is still unknown, its three-dimensional geometry has been revealed by electron tomography (3,4), and attempts have been made to reconstruct its structure by merging numerous strands of experimental data (5,6).

Whereas small molecules and proteins can pass through NPCs by free diffusion, larger macromolecules are strictly excluded from crossing the nuclear envelope unless they are bound to nuclear transport receptors (NTRs). Transport receptor/cargo complexes, however, are transferred rapidly through the NPC, defying concentration gradients (for recent reviews, see (7,8)).

The mechanism of selective admission to the nucleus, taking place within the NPC, remains one of the major and most controversially debated questions around nuclear transport. It has been found that a specific class of nucleoporins plays a leading role in establishing the selectivity of NPCs. These nups have large intrinsically disordered domains and are located both in the central channel of NPCs and at their periphery (9–13). A unique richness in phenylalanine-glycine (FG) repeat motifs is the most conspicuous feature within the sequence of these nups. These FG motifs are highly conserved, which has led to the generally accepted view that the properties of nuclear transport within the NPC are largely determined by interactions between transport receptors and the FG units (14–22).

Conversely, the blockage of proteins from passage through the NPC is still a subject of dispute. Proposed models for the selectivity barrier range from virtual gating or polymer-brush models to selective-phase models. The former are based on freely moving, noninteracting FG repeats, which impede the passage of inert macromolecules by entropic exclusion and steric hindrance (23–28). According to the latter, FG repeats interact with each other, thereby forming a hydrogel or a meshwork (27,29,30). Recently, it has also been proposed that the permeability barrier could consist of a combination of both an entropic barrier and a meshwork, and that, according to their specific amino acid composition, FG repeat nups adopt diverse structures and fulfill distinct roles in different zones of the nuclear pore (10).

In this context, Frey et al. (30) performed experiments demonstrating that the FG repeat domain from the yeast nucleoporin Nsp1p forms a hydrogel in vitro. Moreover, this hydrogel showed the same selective transport properties as NPCs (31). For example, the FG-hydrogel reproduced import signal-dependent and importin-mediated cargo influx, allowing importin-$\beta$ to accelerate the diffusion into the gel of a large cognate cargo >20,000-fold. Passage times of NTRs and NTR-cargo complexes through the nuclear pore between 5.8 ms and 10 ms (32,33) were also reproduced in the hydrogel. Mutating phenylalanine to serine in the FG motifs prevented gel formation (30).

Although the FG motifs have been the subject of numerous studies, and a coarse-grained molecular dynamics (MD) study has been carried out on the global aggregation properties of

*Submitted November 15, 2009, and accepted for publication February 12, 2010.*

*Correspondence: ndoelke@gwdg.de



Editor: Gregory A. Voth.



0006-3495/10/06/2653/9

doi: 10.1016/j.bpj.2010.02.049

nups (34), the abundant hydrophilic spacer sequences located between the FG motifs are viewed as largely inactive linker sections, and, therefore, have not yet been examined in atomic detail, and have often been viewed as largely inactive linker sections. A recent study, however, pointed toward a more active role in the formation of a possible selective phase by observing that the cohesiveness of FG nups varies widely and depends not only on the composition of the FG motifs, but also on the number of charged residues in the linkers (10).

In this work, we study the aggregation properties of FG repeat motifs by atomistic simulations, with an emphasis on the structures formed, the interactions stabilizing these structures, the interactions with transport receptors, and especially the role of the linker regions between FG repeats. To this end, we performed MD simulations of the spontaneous aggregation of model peptides inspired by the sequence of the FG repeats from the C-terminal part of yeast nucleoporin Nsp1p, and corresponding Phe to Ser mutants. Simulations were carried out in aqueous solution, as well as in presence of the nuclear import receptor importin $\beta$.

## METHODS

FG nucleoporins are natively disordered, and crystal structures are, thus, only available for fragments bound to transport receptors. To enhance conformational sampling and avoid biasing by the starting structure, all models were simulated starting from a completely extended structure, which was constructed with PyMOL (35) by imposing an all-*trans* geometry on the backbone dihedrals. All peptides were capped with acetyl groups at the N-terminus and N-methylamide groups at the C-terminus.

All MD simulations were carried out with the GROMACS program package (36) (CVS version from 2007-07-20), using the OPLS-AA force field (37,38) and the SPC water model (39). All simulations were performed in the NPT ensemble. The temperature was kept constant using Berendsen coupling at $T = 300$ K with a coupling time of $\tau_T = 0.1$ ps (40). The pressure was coupled to a Berendsen barostat with $\tau_p = 1.0$ ps and an isotropic compressibility of $4.5 \times 10^{-5}$ bar$^{-1}$ in the $x$, $y$, and $z$ directions (40). All bonds were constrained by using the LINCS algorithm (36). An integration time step of 2 fs was used. Lennard-Jones interactions were calculated with a cutoff of 10 Å. Electrostatic interactions were calculated explicitly at a distance smaller than 10 Å; long-range electrostatic interactions were calculated by particle-mesh Ewald summation with a grid spacing of 0.12 nm and fourth-order B-spline interpolation. Structures were written out every 1 ps for subsequent analysis.

The extended starting structures were subjected to 8 ns of stochastic-dynamics simulations in the gas phase at 500 K, and snapshots from the trajectory were clustered using the program NMRCLUST (41). Subsequently the three most representative structures were equilibrated and simulated for 3 ns at 300 K in water, with an $Na^+/CL^-$ concentration of 150 mmol/L.

The most representative structures from each production run were used for the subsequent aggregation simulations. In the case of the four amino acid models, 10 monomers were placed randomly in a periodic cubic box with box vectors of 65 Å length; for the study of the 19-amino acid models, four monomers and a cubic $80 \times 80 \times 80$ Å$^3$ box were used. In all cases, the NaCl concentration was 150 mmol/L.

Solvent-accessible surface (SAS) areas were calculated with the g_sas program, using the double cubic lattice method (42) with a 0.14-nm probe radius. In case of the 19-amino acid peptides, an exponential function was fitted to the data by the least-square method,

$$f(x) = (S_0 - \alpha) \times \exp(-t/\tau) + \alpha,$$

$$S_0 = \mathrm{SAS(monomers)/SAS(extended)}.$$

SAS (extended) is the SAS of four monomers of a specific peptide in completely extended conformation, as built in PyMOL. SAS (monomers) is the average SAS of the representative structures from MD simulations of a monomer in water. The values $\tau$ and $\alpha$ are fitting parameters.

## RESULTS

### Aggregation propensity of FSFG motifs

To study the aggregation propensities of FG motifs and hydrophilic stretches of FG nups separately, we carried out simulations of small model peptides in solution. As a minimal FG-repeat system, 10 FSFG fragments were simulated in a water-filled cell, at a concentration of 60.5 mmol/L. Three 100-ns simulations were carried out, in all of which fast peptide aggregation was observed. Fig. 1 *A* shows their cohesive behavior in terms of a drastic drop in the SAS of the peptides (*blue curves*). The SAS decreased by ~45% already within the first 20 ns of the simulation. All trajectories ultimately led to the formation of a single peptide aggregate comprising all 10 monomers (Fig. 1 *C*). In two simulations, the aggregation process was completed after 10 and 25 ns, respectively. In the third simulation, two separate aggregates formed initially, which subsequently merged into one. In all cases, the aggregates remained stable throughout the rest of the simulation.

Next, we simulated the FSFG to SSSG mutation studied in the experiments by Frey et al. (30), which prevented hydrogel formation. As for the wild-type, we studied the aggregation of 10 individual tetrapeptides. Quite unexpectedly, although the nature of these fragments is very hydrophilic, partial spontaneous aggregation was seen here as well (Fig. 1 *A*, *red curves*). However, the total SAS decreased by only ~30%, and the SSSG peptides formed smaller aggregates comprising only up to four monomers. Larger associates proved to be unstable. To test whether aggregation is merely slowed down or, alternatively, SSSG peptides are not capable of forming larger aggregates, a cluster of FSFG peptides was taken out of a simulation snapshot after ~70 ns and mutated into SSSG. Fig. 1 *B* shows that the initial aggregate started to dissociate almost immediately and break up into smaller aggregates in the subsequent simulations. During the course of the 100-ns simulation, larger aggregates were formed only transiently. This result supports the view that the markedly reduced aggregation tendency of SSSG indeed reflects a decrease in intrinsic aggregation propensity and is not only a kinetic effect. The observed differences between FSFG and SSSG point to aromatic or hydrophobic interactions providing the main driving force for aggregation. Thus, one would expect a hydrophobic core, mainly composed of phenylalanine side chains, at the center of the aggregates.

### Structural properties of FSFG aggregates

Quite in contrast, however, closer structural analysis of the aggregates revealed that most of the phenyl rings are solvent-accessible. The aggregates in fact display an organization

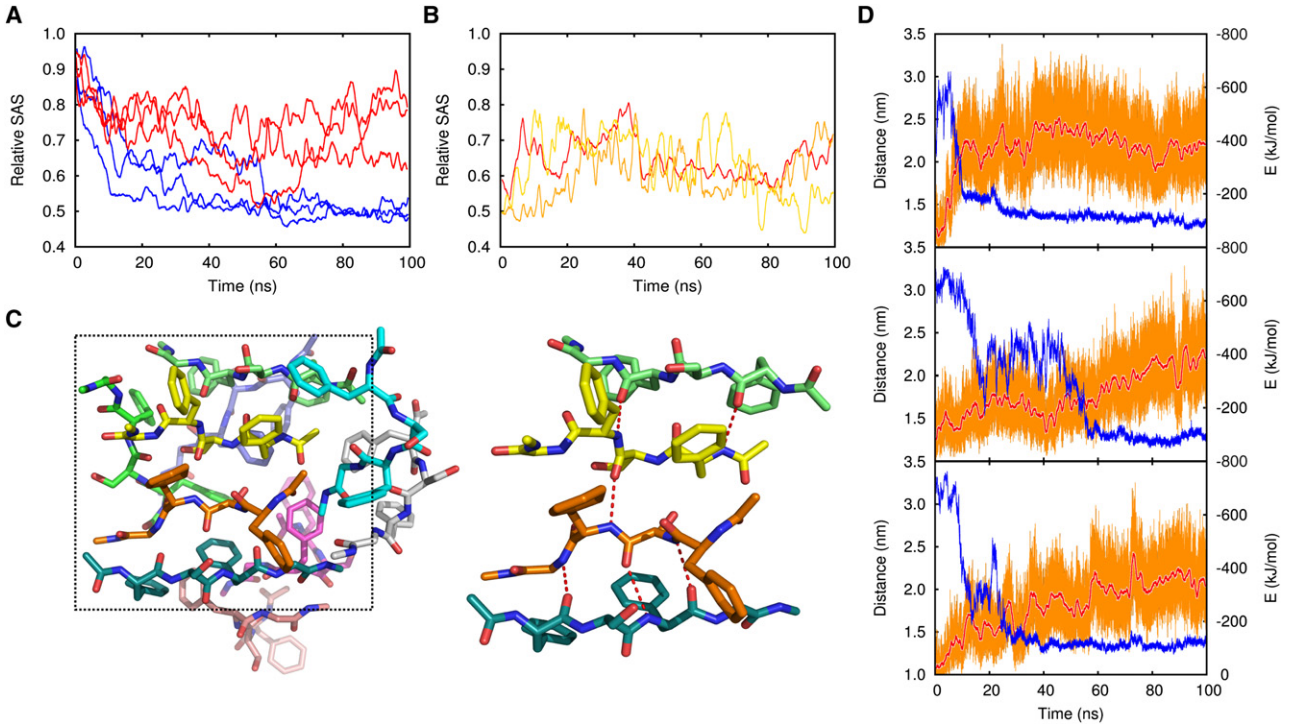


FIGURE 1 Aggregation of FSFG motifs. (*A*) Running averages of the solvent-accessible surface (SAS) during aggregation of 10 four-amino-acid peptides, normalized with the SAS of 10 monomers in fully extended conformation. Results of three independent simulations of FSFG (*blue*) and SSSG (*red*) are show. (*B*) Running averages of SAS after Phe to Ser mutation in an aggregate. (*C*) Snapshot after 100 ns of the aggregate formed by FSFG with a detailed view of a small $\beta$-sheet comprising four monomers. (*D*) Evolution of the average intermolecular Phe-Phe distance (*blue*) and the total backbone-backbone hydrogen bond energy (*orange*) during aggregation (running averages in *red*).

into parallel or antiparallel substructures, comprising between two and four monomers each, which are connected by $\beta$-bridges. Our simulations thus suggest that FG repeat aggregation is not governed by hydrophobic Phe-Phe interactions alone, but rather that the formation of secondary structural elements via backbone-backbone hydrogen bonding also plays an important role in the aggregation process. The evolution of the average intermolecular distance between Phe rings and the total energy of all backbone hydrogen bonds, calculated according to Espinosa et al. (43), is shown in Fig. 1 *D*. As expected, the average Phe-Phe distance decreased rapidly as aggregation occurred. Simultaneously, however, interpeptide hydrogen bonds started to form in all three simulations.

Most strikingly, a structural reorganization is seen after the initial phase of cluster formation, during which the mean Phe-Phe distance is not reduced further, but the backbone hydrogen-bond energy continues to increase. The picture emerging from this observation is that the hydrophobic Phe side chains indeed drive initial aggregation, but that the final structure of FG repeat aggregates is largely determined by interrepeat backbone hydrogen bonds that bring the cluster into a proto-$\beta$ strand state.

This structure, leaving the Phe side chains largely exposed to solvent, could play a significant role in nuclear transport, as it has been shown, crystallographically as well as by simulations, that the Phe side chains constitute the principal binding sites for nuclear transport receptors (8,20). To ensure rapid translocation through the NPC, it appears favorable that these Phe groups are readily accessible for interactions with receptor-cargo complexes.

To test whether the observed structure of the clusters is indeed able to support and/or facilitate binding, we performed a simulation in which an aggregate of FSFG repeats, taken from a previous simulation, was placed at a distance of 20 Å from a receptor-cargo complex consisting of importin $\beta$ bound to the IBB domain of importin $\alpha$. We observed that the FG cluster approached and bound to the convex surface of importin $\beta$ within a timescale of only 10 ns through insertion of the freely accessible Phe side chains into hydrophobic clefts on the outer surface of importin $\beta$ (Fig. 2). These clefts were previously suggested to be FG binding sites (8,20). Thus, this result shows that the specific structure we found for FG aggregates is not only consistent with the role of FG motifs in nuclear transport, but, by leaving the Phenylalanine side chains exposed, it also facilitates and probably accelerates binding to transport receptors.

## Do hydrophilic linker sections influence aggregation?

Despite the low sequence variability in the FG motifs between different FG nucleoporins, the observed aggregation

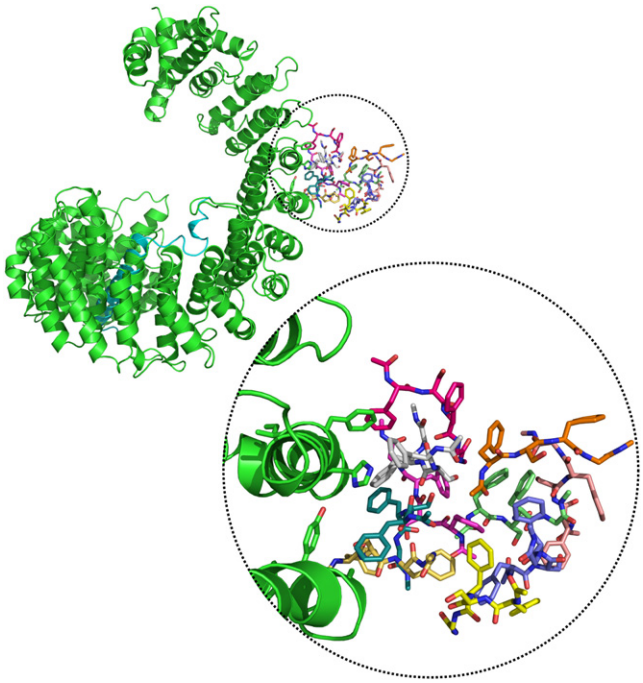

FIGURE 2 Interaction of an aggregate formed by 10 FSFG peptides with importin $\beta$.

tendencies vary widely. Although certain FG nucleoporins bind readily to each other, others are generally noncohesive (10). Interactions between FG domains are therefore unlikely to be controlled by hydrophobic FG motifs alone, but should also depend on the amino acid composition of the hydrophilic linkers connecting them.

To study the influence of the hydrophilic linker sequences on aggregation, five different linkers were constructed, based on the consensus sequence of the C-terminal end of the FG repeat domain of Nsp1p (Table 1). These sequences display a large sequence homology, and all of them possess a zero net charge. They differ, however, in their total number of charged residues. This set of linkers is therefore well suited to analyze the impact on aggregation of specific charged residues. For each peptide, three independent simulations of four monomers each were carried out.

Fig. 3 *A* shows the SAS of these peptides relative to four monomers in completely extended conformation during the simulations. As can be seen, all five model peptides aggregated quickly and formed stable tetrameric clusters. However, marked differences are seen for the differently charged

**TABLE 1 Sequences of FSFG peptides with different hydrophilic linkers**

| Linker | Sequence | Charged residues |
|---|---|---|
| 1 | NGTSTPAFSFGASNNNSTN | +0/−0 |
| 2 | DGTSKPAFSFGASNNNSTN | +1/−1 |
| 3 | NGTSTPAFSFGAKNNESTN | +1/−1 |
| 4 | DGTSKPAFSFGAKNNESTN | +2/−2 |
| 5 | DGTSKPAFSFGAKNNEDKN | +3/−3 |

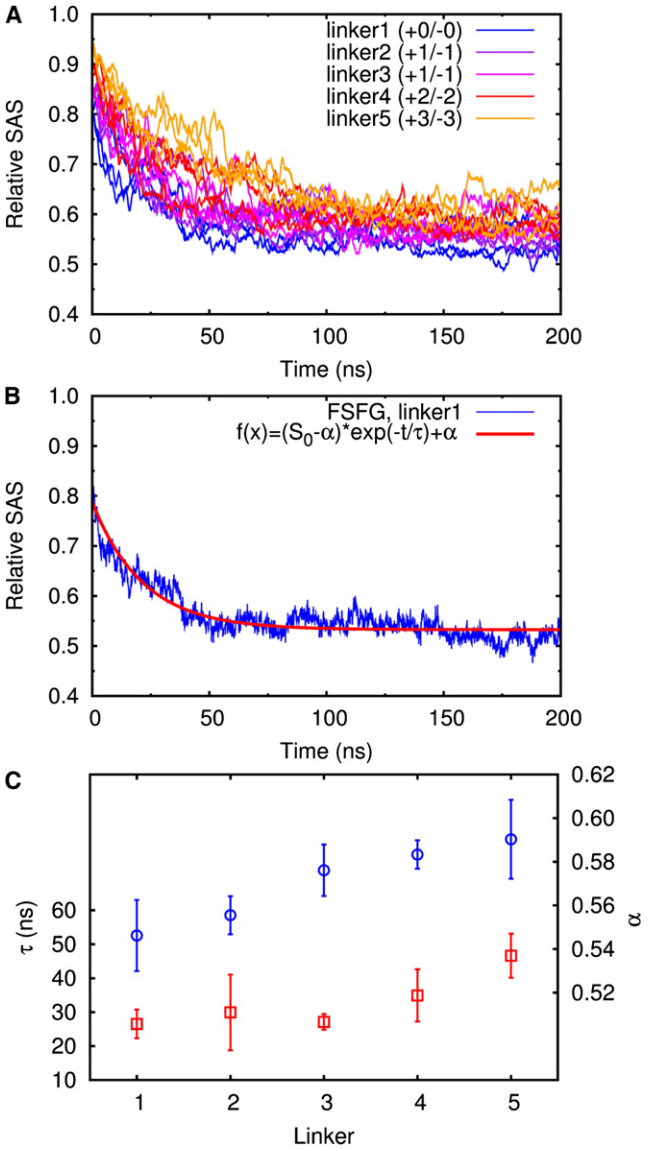


FIGURE 3 Dependence of FG repeat aggregation on the linker sequence. (*A*) Running averages of the solvent-accessible surface (SAS) of four copies of 19-amino acid FSFG peptides. Peptides with five different linkers are shown (*blue*, linker 1, *purple*, linker 2, *magenta*, linker 3, *red*, linker 4, and *orange*, linker 5). The SAS was normalized with the value for four monomers in extended conformation. For each model, three independent simulations are shown. (*B*) Example for an exponential fit to the normalized SAS. (*C* and *D*) Average values for fitting parameters $\tau$ (*red squares*) and $\alpha$ (*blue circles*).

peptides, with peptides possessing no or few charged residues aggregating considerably faster. The aggregation of the different peptides was quantified by fitting an exponential function to the SAS curves (Fig. 3 *B*)

$$f(x) = (S_0 - \alpha) \times \exp(-t/\tau) + \alpha,$$

where $S_0$ is a parameter that is specific for each model peptide. The parameters $\tau$ and $\alpha$ were optimized during the fitting process (see Methods). Table 2 shows the mean values of $\tau$ and $\alpha$ for the differently charged peptides.

**TABLE 2 Solvent-accessible surface of four free monomers, relative to the extended conformation ($S_0$), and mean values for time constant $\tau$ (ns) and aggregate density $\alpha$ for the different peptides ($\tau$, $\alpha$ with standard error of the mean)**

| | FSFG | | | SSSG | | |
|---|---|---|---|---|---|---|
| Linker (charge) | $S_0$ | $\tau$ | $\alpha$ | $S_0$ | $\tau$ | $\alpha$ |
| 1 (+0/−0) | 0.79 | 27 (4) | 0.55 (0.02) | 0.84 | 26 (8) | 0.56 (0.01) |
| 2 (+1/−1) | 0.88 | 30 (11) | 0.56 (0.01) | | | |
| 3 (+1/−1) | 0.87 | 28 (2) | 0.58 (0.01) | | | |
| 4 (+2/−2) | 0.88 | 35 (8) | 0.58 (0.01) | | | |
| 5 (+3/−3) | 0.93 | 47 (7) | 0.59 (0.02) | 0.87 | 23 (5) | 0.57 (0.02) |

The time constant $\tau$ gives a measure for the speed of aggregation of a specific peptide (Fig. 3 *C*, *red squares*). Peptide 1, which does not contain any charged amino acids, formed a single aggregate in ~50 ns in two out of three simulations. Subsequently, the change in SAS was observed to be very small. For peptide 5, displaying three positive and three negative residues, the reduction in SAS was found to take >100 ns. This is reflected by the different values of $\tau$. While for peptides containing no or one positively and negatively charged residues, $\tau$ lies between 27 ns and 30 ns, it increases to 47 ns for peptides with +3/−3 charge.

The parameter $\alpha$ describes the compactness of the tetramers. In addition, here, slight differences, following the same tendency of the time constant $\tau$, were seen (Fig. 3 *C, blue circles*). The relative SAS of the aggregates after 200 ns of elapsed simulation time was significantly smaller for peptides without charged residues ($\alpha = 0.55$) than for peptides with +3/−3 charge ($\alpha = 0.59$). These structural differences were also observed directly from snapshots of the aggregates formed by peptides 1 and 5 (Fig. 4, *A* and *B*), which show that the tetramer of uncharged peptide 1 is more tightly packed than that of peptide 5.

In line with what was found for the FSFG tetrapeptides, in all aggregates short $\beta$-sheets or single $\beta$-bridges were observed. However, contrary to the small system, where the monomers formed $\beta$-sheets through intermolecular hydrogen bridges, in the larger systems backbone-backbone hydrogen bonds are predominantly intramolecular, resulting in $\beta$-hairpins, which are connected to each other mostly by interactions involving polar side chains.

To obtain further insight into the interplay between the hydrophobic FxFG motifs and the hydrophilic linker sections, we analyzed the dependence of intermolecular Phe-Phe distances in the aggregate on the nature of the linker sequence. Fig. 5, *A* and *B*, show the results for peptides 1 and 5, respectively, as histograms with a normal distribution fit. The reduced cohesive tendency of peptides containing charged residues shows up here as larger Phe-Phe distances. While the mean Phe-Phe distance in the aggregate of peptide 5 is 2.20–2.53 nm, it is reduced to 1.84–1.99 nm for peptide 1. The number of short Phe-Phe contacts (with a distance of <0.7 nm between phenyl rings) is 4.7 times larger for peptide 1 than for peptide 5. In addition, the distance distribution is generally narrower in simulations of peptide 1, with a standard deviation of 0.62–0.68, compared to 0.67–0.89 for peptide 5. Both distributions are nearly symmetric with an average skew of −0.06 and −0.04 for peptides 1 and 5, respectively. While the distance distribution for peptide 5 is mesokurtic, the distribution for peptide 1 is platykurtic (with an average excess kurtosis of −0.49). Our results suggest, thus, that the amino acid composition of the hydrophilic linkers determines the compactness of the FG repeat aggregates and controls the spacing of the FG motifs and, thereby, also the distribution of binding sites for transport receptors.

Our simulations also showed that the number of charged residues in the linker sequence affects the conformation of the free monomers. The normalized SAS values at $t = 0$, when the four monomers are not in contact, show that

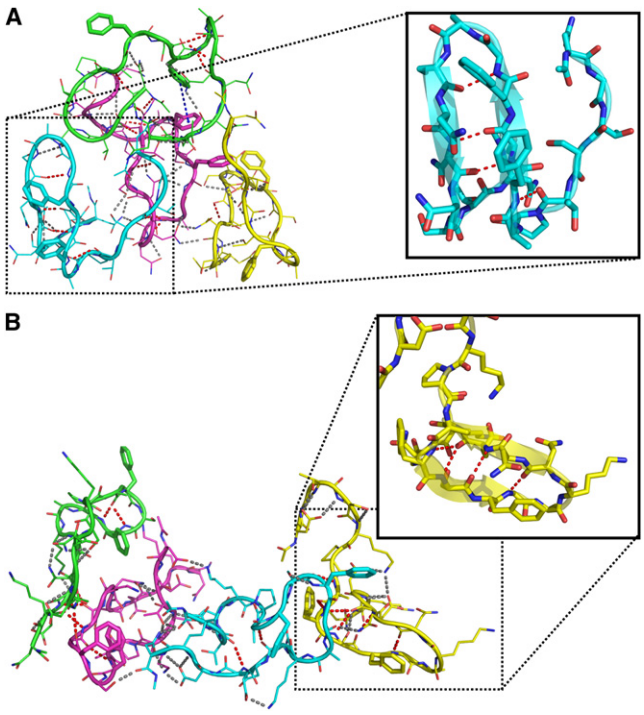


FIGURE 4 Snapshots of aggregates with detailed view of an intramolecular $\beta$-hairpin. Intermolecular Phe-Phe interactions are shown in blue, backbone-backbone H-bonds in red, and polar contacts involving side chains in gray. Structures of the uncharged peptide 1 (*A*) and peptide 5 with +3/−3 charge (*B*) are shown.

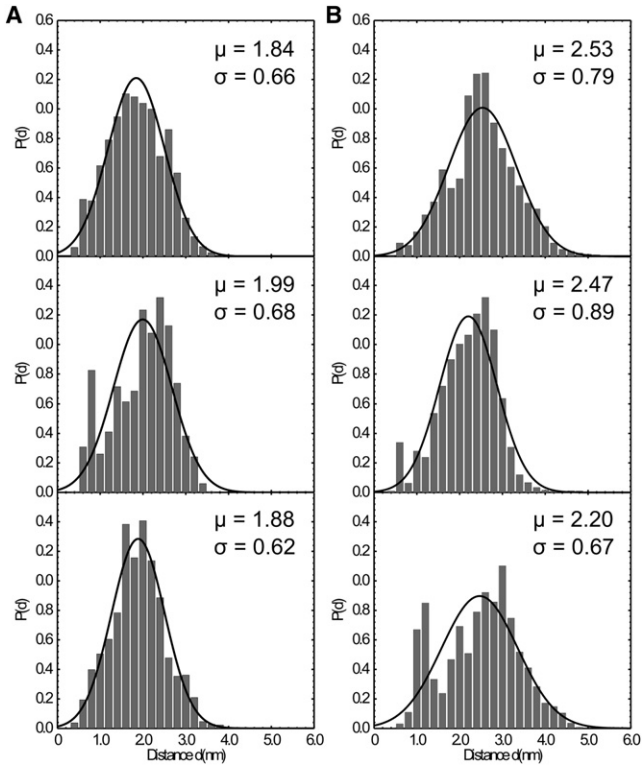


FIGURE 5 Distribution of intermolecular Phe-Phe distances in the tetrameric aggregates. Distances were calculated between centers of mass of phenyl rings. Results from the second half of the simulations are shown with fitted normal distribution. Comparison between peptide 1 (*A*), with uncharged linkers and peptide 5 with +3/−3 charges (*B*). Results are shown for three independent simulations each.

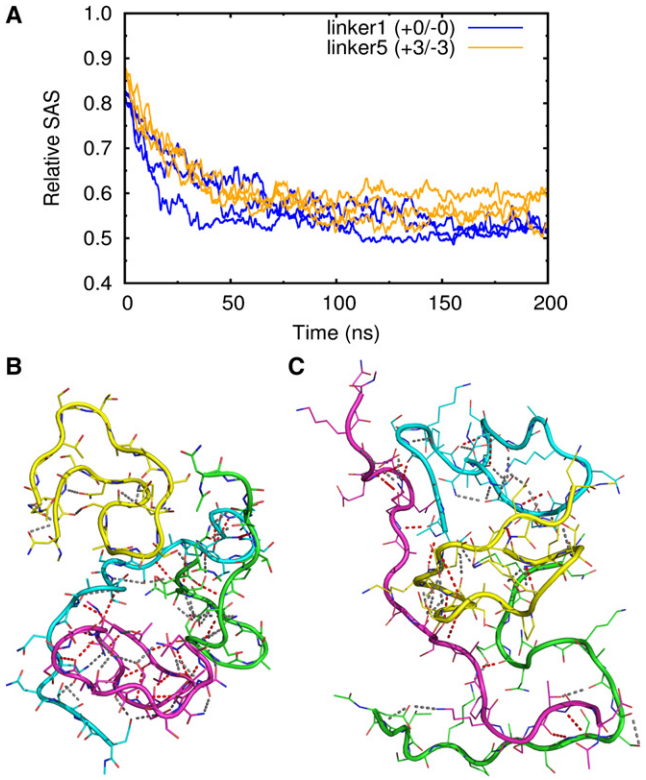


FIGURE 6 Aggregation of Phe to Ser mutants. (*A*) Solvent-accessible surface (SAS) of four 19-amino acid SSSG mutant peptides with hydrophilic linkers, containing no (*blue lines*) and +3/−3 (*orange lines*) charge. The SAS was normalized with the value for four monomers in extended conformation. Three independent simulations are shown for each model. (*B* and *C*) Snapshots of aggregates. Backbone-backbone H-bonds are shown as red dashed lines and polar contacts involving side chains in gray. Structures of the mutant of peptide 1 with uncharged linkers (*B*) and of the mutant of peptide 5 with +3/−3 charge (*C*) are shown.

peptides consisting only of neutral amino acids adopt a more compact conformation. For peptide 1, the total SAS of the four monomers reaches only 80% of the value that would be expected for fully extended peptides. In contrast, peptide 5, which contains six charged residues, reaches 93% of this value.

Overall, the hydrophilic linkers contribute substantially to the aggregation of FG repeat peptides. The aggregation process and the structure of the aggregates are modulated by the amino acid composition of the linkers. The number of charged residues, especially, determines the spacing between Phe residues and the flexibility of the aggregates.

## Phe to Ser mutation in FG repeats

Similar to the simulations of the wild-type (wt) peptides, Phe to Ser mutants of peptides 1 (without charged residues) and 5 (with +3/−3 charge) were subjected to 200 ns of free MD simulation. Fig. 6 *A* depicts the total SAS of the mutant peptides, which shows that both systems underwent rapid aggregation. Contrary to what was observed for the wild-type, the mutants exhibited virtually no differences in aggregation speed. The time constant of the solvent-accessible surface $\tau$ is 26 ± 8 nm and 23 ± 5 nm for SSSG peptides with linker 1 and linker 5, respectively (see Table 2). The compactness of the aggregates is nearly identical, as shown by $\alpha$-values of 0.56 and 0.57 for peptides with linker 1 and 5, respectively. Furthermore, comparison with the wt peptides shows that aggregation of peptide 1, which contains no charged amino acids, was not altered by the Phe to Ser mutation. Surprisingly, for peptide 5, with six charged residues, the aggregation of the mutant was found to occur even faster than that of the wt ($\tau = 23 \pm 5$ nm for the mutant, compared to 47 ± 7 nm for the wt).

The structures of the wt and mutant peptides also differed, both in simulations of the monomers and in the tetramers. For peptide 1, the free monomers of the wt and the mutant adopted compact geometries with a relative SAS of only 79% and 84%, respectively. For peptide 5, the wt exhibited a more extended conformation, and the SAS reached 93% of the value of the fully elongated state. Here, Phe to Ser mutation slightly increased the compactness, and reduced the relative SAS to 87%. The effect of Phe to Ser mutation on the conformation of the monomers is also observed in terms of changes of the radius of gyration ($R_g$) of individual

monomers within the aggregate. The mutants exhibit a slightly larger average $R_g$. For peptides with uncharged linker 1, the $R_g$ of the mutants is 8.4 ± 0.2 Å, compared to 8.1 ± 0.1 Å for the wild-type. For peptides with linker 3, it is 8.7 ± 0.1 Å and 8.5 ± 0.1 Å for the mutant and the wild-type, respectively.

In summary, the aggregation behavior of 19-amino acid peptides differs from that of the short FSFG motifs alone, and the influence of the aromatic phenylalanine residues is less dominant. The dependency of aggregate structure and composition on the hydrophilic part of the peptides shows that both the properties of monomers and aggregates are indeed affected by the linker sequence. These results underline our finding that the presence of charged residues influences the aggregate structure, as they modulate not only the structure of the monomers, but also reduce the effect of Phe to Ser mutation on aggregation.

## DISCUSSION

The ongoing debate about the structure and function of the permeability barrier in nuclear pore complexes has recently largely focused on the question how and to what extent FG-rich nucleoporins, lining the inner pore, interact, which structures they form, and which role the conserved FG motifs play in creating these structures.

If the permeability barrier is formed by a meshwork of FG repeats, its structure will be controlled, by specific interactions between two or more FG repeats and by the density of such interaction sites and their spatial distribution. While the time span covered by all-atom MD is still too short to reproduce the formation of a large meshwork, these simulations can yield insight into the nature of specific interactions and the influence of certain amino acid motifs. It is, therefore, not possible to draw direct correlation between the aggregation propensity of model peptides and the macroscopic structure of FG nucleoporins. However, the study of short FG repeat sequences can give valuable information about the types of interactions responsible for the formation of the permeability barrier.

The results presented here show that the FG motifs alone, as well as linked to different hydrophilic spacers, exhibit a strong intrinsic disposition to aggregate. This cohesive tendency suggests that inter- and intramolecular interactions of FG-rich nucleoporins determine the structure and dynamic behavior of the permeability barrier. Thus, our simulation results disfavor models of a purely noninteracting anatomy of the permeability barrier.

Frey and Görlich (31) observed that the FG repeat domain of yeast nucleoporin Nsp1p can form a hydrogel that reproduces the permeability and selectivity properties of the nuclear pore complex. The ability to form such a gel has been found to be markedly reduced when all Phe residues are mutated to Ser (30). Aggregation of the FG repeat domain of Nsp1p therefore depends upon the presence of hydrophobic FG motifs. Hydrophobic interactions or $\pi$-$\pi$-stacking of the aromatic phenylalanine side chains had initially been proposed to be the dominant structural motif of the hydrogel and, by extension, also of the permeability barrier itself. Our MD simulations indeed show that, although FSFG tetrapeptides readily aggregate, the corresponding, strongly hydrophilic SSSG mutant does not form comparably stable clusters. However, whereas the presence of hydrophobic Phe residues is obviously paramount for the cohesive behavior of the tetrapeptides, the predominant structural feature in the resulting cluster is that of backbone-backbone hydrogen bonding. This interaction even leads to the formation of short $\beta$-sheets that are comprised of up to four monomers. Such small $\beta$-aggregates have been widely described as key structural motifs of various types of peptide aggregates, such as spider silk (44), elastomeric proteins (45), and early precursors of amyloid fibers (46). In these structures, the transition between rigid fibrils and highly elastic structures is smooth and is determined by small changes in the amino acid composition (45,47). In some cases, amyloid fibril formation and protein aggregation are mediated by glutamine/asparagine-rich domains. Such domains have been identified in a number of proteins with diverse biological functions, including nucleoporins Nup116, Nup110, Nup57, and Nup49 (48). It is, therefore, conceivable that the permeability barrier of the nuclear pore contains small $\beta$-structures that determine the three-dimensional structure of the peptide network, surrounded by disordered domains that provide the necessary flexibility.

Although the FG sequence motifs found in nucleoporins show only little variation, and are most often represented by the sequence motifs FxFG and GLFG, the interstitial hydrophilic linkers, comprising usually 5–20 amino acids, show large variability and lack prominent motifs. It has therefore been assumed that their role is confined to that of passive spacers, only required to keep the FG motifs apart at a certain distance. By contrast, our simulations suggest that the inclusion of these sequences in the simulation altered the aggregation process markedly.

The structure of the aggregates formed by the larger model peptides is much less ordered than that of the tetrapeptides. After 200 ns of simulation, only little $\beta$-sheet content was found. Hydrogen-bonding was still found to be the most important structural determinant in the aggregate, but interactions between polar side chains dominated the intermolecular energetics. These results indicate that interactions between FG repeats are determined by hydrogen bonding and electrostatic interactions rather than $\pi$-$\pi$ bonding, and that the hydrophilic linkers play a much larger role in the formation of the permeability barrier than previously thought. However, it has to be taken into account that, due to the slow timescale of the formation of a hydrogel, in our simulations we are not viewing a converged structural ensemble. It can, therefore, not be definitively excluded that the lack of Phe-Phe contacts observed in the aggregates

results from a sampling problem of the simulations. Nevertheless, the fact that the aggregation rate is sensitive to the amino acid composition of the hydrophilic parts of the model peptides underlines the importance of the hydrophilic linkers. Our findings agree with a recent experimental study that suggests an influence of hydrophilic sections on nup cohesion (10).

Although our results do not support models based solely on noninteracting FG repeats, it is conceivable that differences in the linker sequences between nups located in different zones of the pore result in areas of more or less interacting FG nups. The permeability barrier can then be thought of as comprising areas of denser meshworks as well as regions with more freely moving FG repeats, depending on the function of the specific nups in the transport process. In this way, the amino acid composition of the linkers, and specifically the amount and distribution of charged residues, could regulate the inner structure of the barrier.

We found that, for the larger model peptides, Phe to Ser mutation has a smaller impact on aggregation. Loss of the hydrophobic Phe residues, however, leads to an increase in conformational flexibility, which is consistent with a recent study on the FG repeat domain of yeast nucleoporin nup116, which concluded that the FG motifs impart order on the domain and determine its compact structure (49). In a macroscopic ensemble of FG nucleoporins, the increased flexibility of the Phe to Ser mutants could then lead to disruption of the meshwork.

In the context of the permeability barrier, composed of intermingled FG nups, each containing a large number of FG motifs connected by hydrophilic linkers, interactions between the hydrophilic linkers would ensure a certain distance between FG motifs. During nucleocytoplasmic transport, NTRs are supposed to bind simultaneously to several FG repeats. The spacing between proposed adjacent FG binding spots on the NTRs importin $\beta$ and Cse1p is ~14 $\pm$ 3 Å and ~16 $\pm$ 4 Å, respectively (22). The distance of ~18–25 Å between FG motifs, found in our simulations, is, thus, close to the optimal distance for NTR binding. Binding of NTRs to multiple FG motifs could also stabilize the network formed by the FG nups.

The picture emerging from our study is that the hydrophilic linkers appear to fine-tune the density of the peptide aggregates as well as the average distance between FG motifs, while the FG motifs drive aggregation and provide a stabilization of the permeability barrier of the nuclear pore.

A recent computational study using coarse-grained simulation methods to study the nature of FG nucleoporin interactions concluded that a less cooperative entropic brush structure is more likely formed than a selective or gel phase (34). In these simulations, coagulation was seen for individual FG repeat domains, resulting in cluster formation. When multiple nup domains were fixed with their termini on one side (i.e., brush), partial coagulation was found, reflected in a reduction of the observed brush height by approximately fourfold. The authors suggested that a brush-like structure would be biologically preferred, as entropic brushes exhibit surface-exposed FG-repeats; allow rapid entry of proteins; and because the biological density of nups in the NPC was reproduced (34).

Our results show that a meshwork-like structure with strongly interacting FG repeats does allow for easily accessible binding sites for transport receptors. The FSFG clusters seen in our all-atom MD simulations provide an even larger percentage of solvent-exposed Phe side chains than that seen in the coarse-grained simulations, while density and structure of the aggregates formed varied with the sequence of the linker sections. This variation might explain functional differences between different FG nups in cargo translocation. It should be noted, however, that the more accurate, but more computationally demanding all-atom simulations we used here preclude the simulation of large polypeptides on microsecond timescales. Despite this limitation, however, the good agreement with experimental data we found here and the apparent functionality of the observed aggregates in transport factor sequestration suggest that aggregation of unstructured FG-rich nups into a meshwork is a relevant structural motif for nucleocytoplasmic transport.

We thank Steffen Frey for designing the peptide sequences used in this study, and Steffen Frey and Dirk Görlich for helpful discussions and carefully reading the manuscript.

## REFERENCES

1. Rout, M. P., J. D. Aitchison, …, B. T. Chait. 2000. The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J. Cell Biol.* 148:635–651.
2. Tran, E. J., and S. R. Wente. 2006. Dynamic nuclear pore complexes: life on the edge. *Cell.* 125:1041–1053.
3. Beck, M., F. Förster, …, O. Medalia. 2004. Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. *Science.* 306:1387–1390.
4. Beck, M., V. Lucić, …, O. Medalia. 2007. Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. *Nature.* 449:611–615.
5. Alber, F., S. Dokudovskaya, …, M. P. Rout. 2007. The molecular architecture of the nuclear pore complex. *Nature.* 450:695–701.
6. Brohawn, S. G., J. R. Partridge, …, T. U. Schwartz. 2009. The nuclear pore complex has entered the atomic age. *Structure.* 17: 1156–1168.
7. Lim, R. Y. H., K. S. Ullman, and B. Fahrenkrog. 2008. Biology and biophysics of the nuclear pore complex and its components. *Int. Rev. Cell Mol. Biol.* 267:299–342.
8. Stewart, M. 2007. Molecular mechanism of the nuclear protein import cycle. *Nat. Rev. Mol. Cell Biol.* 8:195–208.
9. Allen, N. P. C., S. S. Patel, …, M. Rexach. 2002. Deciphering networks of protein interactions at the nuclear pore complex. *Mol. Cell. Proteomics.* 1:930–946.
10. Patel, S. S., B. J. Belmont, …, M. F. Rexach. 2007. Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell.* 129:83–96.
11. Denning, D. P., S. S. Patel, …, M. Rexach. 2003. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proc. Natl. Acad. Sci. USA.* 100:2450–2455.

12. Bayliss, R., A. H. Corbett, and M. Stewart. 2000. The molecular mechanism of transport of macromolecules through nuclear pore complexes. *Traffic*. 1:448–456.

13. Strawn, L. A., T. Shen, …, S. R. Wente. 2004. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nat. Cell Biol.* 6:197–206.

14. Bayliss, R., K. Ribbeck, …, M. Stewart. 1999. Interaction between NTF2 and xFxFG-containing nucleoporins is required to mediate nuclear import of RanGDP. *J. Mol. Biol.* 293:579–593.

15. Bayliss, R., T. Littlewood, and M. Stewart. 2000. Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. *Cell*. 102:99–108.

16. Bayliss, R., S. W. Leung, …, M. Stewart. 2002. Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. *EMBO J.* 21:2843–2853.

17. Bayliss, R., T. Littlewood, …, M. Stewart. 2002. GLFG and FxFG nucleoporins bind to overlapping sites on importin-$\beta$. *J. Biol. Chem.* 277:50597–50606.

18. Liu, S. M., and M. Stewart. 2005. Structural basis for the high-affinity binding of nucleoporin Nup1p to the *Saccharomyces cerevisiae* importin-$\beta$ homologue, Kap95p. *J. Mol. Biol.* 349:515–525.

19. Patel, S. S., and M. F. Rexach. 2008. Discovering novel interactions at the nuclear pore complex using bead halo: a rapid method for detecting molecular interactions of high and low affinity at equilibrium. *Mol. Cell. Proteomics*. 7:121–131.

20. Isgro, T. A., and K. Schulten. 2005. Binding dynamics of isolated nucleoporin repeat regions to importin-$\beta$. *Structure*. 13:1869–1879.

21. Isgro, T. A., and K. Schulten. 2007. Association of nuclear pore FG-repeat domains to NTF2 import and export complexes. *J. Mol. Biol.* 366:330–345.

22. Isgro, T. A., and K. Schulten. 2007. Cse1p-binding dynamics reveal a binding pattern for FG-repeat nucleoporins on transport receptors. *Structure*. 15:977–991.

23. Rout, M. P., J. D. Aitchison, …, B. T. Chait. 2003. Virtual gating and nuclear transport: the hole picture. *Trends Cell Biol.* 13:622–628.

24. Pyhtila, B., and M. Rexach. 2003. A gradient of affinity for the karyopherin Kap95p along the yeast nuclear pore complex. *J. Biol. Chem.* 278:42699–42709.

25. Macara, I. G. 2001. Transport into and out of the nucleus. *Microbiol. Mol. Biol. Rev.* 65:570–594.

26. Lim, R. Y. H., N.-P. Huang, …, U. Aebi. 2006. Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport. *Proc. Natl. Acad. Sci. USA*. 103:9512–9517.

27. Ribbeck, K., and D. Görlich. 2002. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J.* 21:2664–2671.

28. Kramer, A., I. Liashkovich, …, V. Shahin. 2007. Atomic force microscopy visualizes a hydrophobic meshwork in the central channel of the nuclear pore. *Pflugers Arch.* 456:155–162.

29. Ribbeck, K., and D. Görlich. 2001. Kinetic analysis of translocation through nuclear pore complexes. *EMBO J.* 20:1320–1330.

30. Frey, S., R. P. Richter, and D. Görlich. 2006. FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science*. 314:815–817.

31. Frey, S., and D. Görlich. 2007. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell*. 130:512–523.

32. Kubitscheck, U., D. Grünwald, …, R. Peters. 2005. Nuclear transport of single molecules: dwell times at the nuclear pore complex. *J. Cell Biol.* 168:233–243.

33. Yang, W., J. Gelles, and S. M. Musser. 2004. Imaging of single-molecule translocation through nuclear pore complexes. *Proc. Natl. Acad. Sci. USA*. 101:12887–12892.

34. Miao, L., and K. Schulten. 2009. Transport-related structures and processes of the nuclear pore complex studied through molecular dynamics. *Structure*. 17:449–459.

35. DeLano, W. L. 2008. The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA.

36. Van Der Spoel, D., E. Lindahl, …, H. J. Berendsen. 2005. GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26:1701–1718.

37. Jorgensen, W. L., D. S. Maxwell, and J. Tirado-Rives. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118:11225–11236.

38. Kaminski, G. A., R. A. Friesner, …, W. L. Jorgensen. 2001. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B*. 105:6474–6487.

39. Berendsen, H. J. C., J. P. M. Postma, …, J. Hermans. 1981. Interaction models for water in relation to protein hydration. *In* Intermolecular Forces. B. Pullman, editor. Reidel, Dordrecht, The Netherlands.

40. Berendsen, H. J. C., J. P. M. Postma, …, J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81: 3684–3690.

41. Kelley, L. A., S. P. Gardner, and M. J. Sutcliffe. 1996. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. *Protein Eng.* 9: 1063–1065.

42. Eisenhaber, F., P. Lijnzaad, …, M. Scharf. 1995. The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *J. Comput. Chem.* 16:273–284.

43. Espinosa, E., E. Molins, and C. Lecomte. 1998. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chem. Phys. Lett.* 285:170–173.

44. Gatesy, J., C. Hayashi, …, R. Lewis. 2001. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science*. 291:2603–2605.

45. Rauscher, S., S. Baud, …, R. Pomès. 2006. Proline and glycine control protein self-organization into elastomeric or amyloid fibrils. *Structure*. 14:1667–1676.

46. Chiti, F., and C. M. Dobson. 2006. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75:333–366.

47. Rousseau, F., J. Schymkowitz, and L. Serrano. 2006. Protein aggregation and amyloidosis: confusion of the kinds? *Curr. Opin. Struct. Biol.* 16:118–126.

48. Michelitsch, M. D., and J. S. Weissman. 2000. A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. *Proc. Natl. Acad. Sci. USA*. 97: 11910–11915.

49. Krishnan, V. V., E. Y. Lau, …, M. F. Rexach. 2008. Intramolecular cohesion of coils mediated by phenylalanine-glycine motifs in the natively unfolded domain of a nucleoporin. *PLOS Comput. Biol.* 4:e1000145.