Hydrophobic mismatch-induced clustering as a primer for protein sorting in the secretory pathway

Ulrich Schmidt, Matthias Weiss *

Cellular Biophysics Group (BIOMS), German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

A B S T R A C T

Sorting of transmembrane proteins is a central task of the secretory pathway. Due to the lack of an organizing mastermind, the decision whether a protein participates in anterograde/retrograde transport or rather stays in its compartment has to be made by a self-organization process on the molecular scale. Minimizing the hydrophobic mismatch between a protein and the surrounding lipid bilayer has been shown to be an important determinant of protein localization. It has remained elusive, however, how mislocalized proteins sense that remote organelles may provide a better lipid environment, i.e. how proteins differentially control their journey along the secretory pathway. Here we show by coarse-grained membrane simulations that proteins partition into the lipid phase with the smallest hydrophobic mismatch on heterogeneous membranes while they cluster and even segregate as homo-oligomers according to their hydrophobic mismatch on a homogeneous bilayer. We propose that protein sorting is facilitated by stabilizing coat proteins at clusters of mislocalized proteins that experience a hydrophobic mismatch.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Protein sorting is a central task of the secretory pathway. While anterograde cargo molecules are exported from the endoplasmic reticulum (ER) at specialized domains ('ER exit sites') via COPII vesicles [1], unfolded and ER-resident proteins are retained [2]. At the level of the Golgi apparatus, cargo proteins and Golgi-resident enzymes are sorted differentially into the distinct cisternae by means of the COPI machinery [3]. As a consequence, Golgi enzymes have been observed to show gradient-like distributions across the Golgi stack in steady state [4].

The observation of distinct steady-state localizations of a multitude of transmembrane proteins along the secretory pathway (from the ER, via the Golgi apparatus towards the plasma membrane) has led to the postulation of two (not mutually exclusive) hypotheses how transmembrane proteins could determine their localization. The *kin recognition* hypothesis [5] invokes the (transient) formation of larger protein hetero-oligomers due to specific motifs in the stalk and transmembrane region of the proteins, i.e. the emerging oligomers are simply too large to enter COPI vesicles for further transport and hence stay in 'their' compartment. In contrast, the *membrane thickness* hypothesis [6] predicts that the length of a protein's transmembrane

domain (TMD) determines the position by sensing the varying lipid thickness along the secretory pathway [7], i.e. a protein will stop its journey when the TMD matches best the local lipid environment. Indeed, both hypotheses have been supported experimentally [8,9].

The key aspect underlying the membrane thickness model has been highlighted in the seminal work by Bloom and Mouritsen [10]: Retaining a TMD that is too long/short to be accommodated by the surrounding lipid bilayer, i.e. if there is a *hydrophobic mismatch*, is energetically unfavourable and consequently the protein partitions (via diffusion) into a membrane domain that yields a lower mismatch. Indeed, very recent reports have further supported the central role of the hydrophobic mismatch as a key element in protein sorting [11–13]. In these studies the partitioning of proteins according to their TMD was either directly observed or indirectly derived from the transport behaviour of the investigated constructs.

While the partitioning of a protein towards a more adequate phase on the same membrane via diffusion is easy to imagine, it has remained elusive how proteins may sense that another remote organelle may offer an environment with a smaller hydrophobic mismatch. To put it differently: How should proteins facilitate their transport towards the organelle with the best-matching lipid bilayer?

Here, we have used coarse-grained membrane simulations to address this point. We show that transmembrane proteins not only partition into a membrane environment with the best hydrophobic mismatch on heterogeneous membranes, but they also segregate and cluster according to their type of hydrophobic mismatch on membranes with a homogeneous thickness. Hydrophobic mismatch-induced clustering,

* Corresponding author. Cellular Biophysics Group (BIOMS), B085, German Cancer Research Center, c/o BIOQUANT (BQ0019), Im Neuenheimer Feld 267, D-69120 Heidelberg, Germany. Tel.: +49 6221 5451304; fax: +49 6221 5451484.
E-mail address: m.weiss@dkfz.de (M. Weiss).

however, can stabilize coat proteins (e.g. COPI/II) on the respective membrane patches and therefore facilitate transport of the protein cluster towards a membrane environment with a lower hydrophobic mismatch by supporting the formation of a transport carrier.

2. Results and discussion

2.1. Proteins partition into lipid phases that match best their hydrophobic length

We first tested whether our simulations would reproduce the commonly anticipated protein partitioning to that part of a lipid bilayer that yields the lowest hydrophobic mismatch. To this end we chose two lipid species of different lengths that were expected to induce a spontaneous segregation (cf. Materials and methods). The formation of coexisting lipid phases with different thicknesses has been reported repetitively for a variety of ternary lipid mixtures [14] and it is assumed to also occur on cellular membranes. To mimic the potentially different composition of the leaflets of typical biomembranes, we chose a configuration in which the lower leaflet was only made of short lipids, while the upper leaflet consisted to 50% of long lipids. As can be seen from Fig. 1A, the chosen setup indeed showed a nice segregation of lipids with the emerging domains having a slight difference in their hydrophobic thickness.
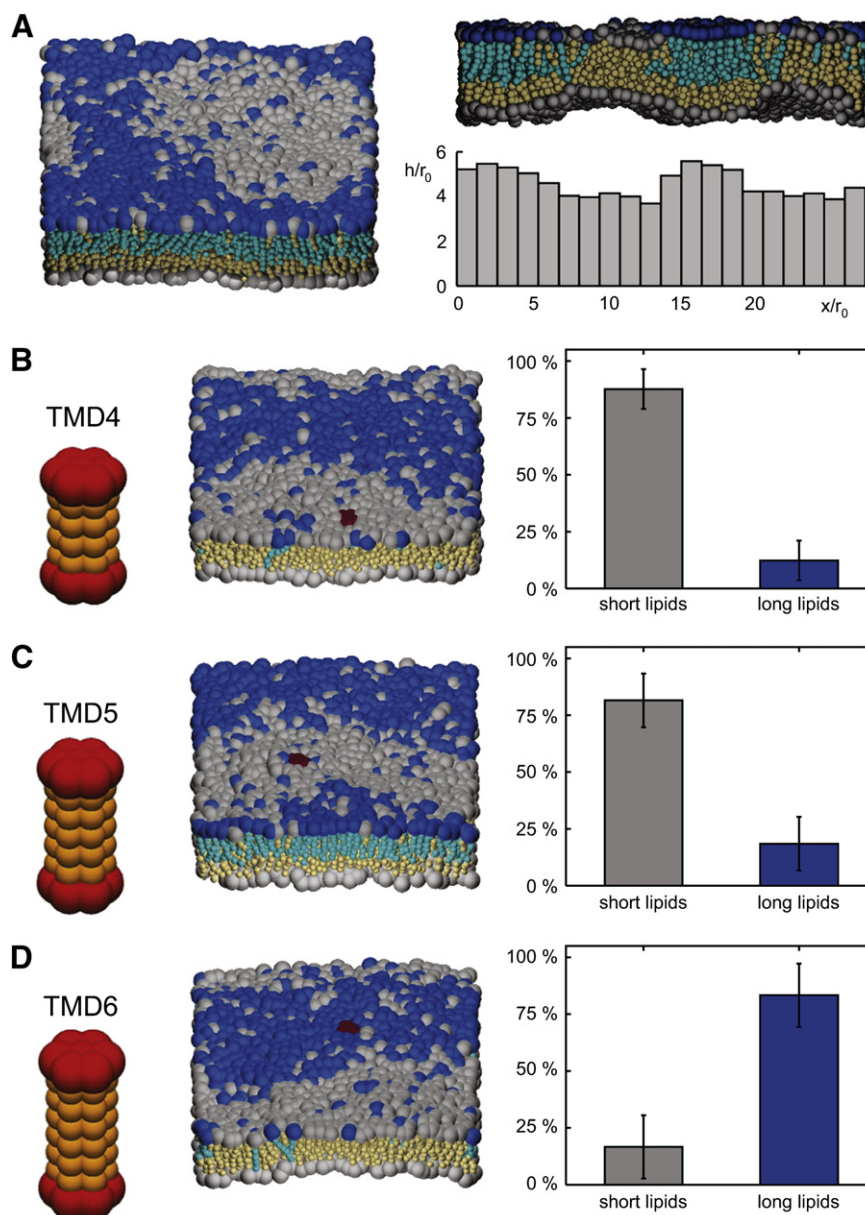
Fig. 1. Transmembrane proteins partition into the lipid phase that yields the smallest hydrophobic mismatch. (A) Snapshot (top and side view; taken at $t \approx 10\,\mu s$) of the steady-state configuration of a phase-separating bilayer with $L = 30 r_0$. Long (blue) and short (grey) lipids have unmixed to form fairly homogeneous phases of different thickness. Indeed, the thickness h of the two lipid phases is slightly different; the change is about 20% of the maximum bilayer thickness. The length of the cross-section and the bilayer thickness are given in simulation units ($r_0 \approx 1$ nm; cf. Materials and methods). (B–D) Images of the proteins TMD4, TMD5, and TMD6 (red hydrophilic groups are shown slightly larger for better visibility) and representative snapshots of their respective steady-state configuration in a phase-separated membrane (blue: long lipids, grey: short lipids). TMD4 and TMD5 partition with high a preference to the thin domain of the bilayer while TMD6 prefers the complementary domain. This is also reflected by the steady-state fraction of short (grey bar) and long (blue bar) lipids in the vicinity (radius of $4 r_0$) of the embedded proteins, respectively. Determining the number of short and long lipids involved only lipids in the heterogeneous upper leaflet. Time series of both lipid species confirmed that TMD4 and TMD5 reside inside the domain of short lipids whereas TMD6 prefers the thicker membrane region (cf. Supplement). They also demonstrate that the averaging was performed over sufficient time.

We next embedded model proteins of increasing TMD length (named TMDn with n being the number of hydrophobic units in the TMD) into the membrane and monitored the amount of long/short lipids in their vicinity as a function of time (cf. Supplement). By construction, TMD4 should favour short lipids in its vicinity due to its considerable hydrophobic mismatch with the long lipids. The strength of the mismatch may be expressed via the difference of the bilayer thickness (on the phase-separated membrane) far away (h_0) and right at the embedded protein (h_p), $\Delta = h_p - h_0$. For TMD4 in the domain of long and short lipids we found $\Delta_l = -1.90 r_0$ and $\Delta_s = -0.55 r_0$, respectively, i.e. TMD4 had a negative mismatch in both domains. Here $r_0 \approx 1$ nm is the intrinsic length of the simulations (cf. Materials and methods). As expected from the values of the mismatches, TMD4 associated predominantly with short lipids (Fig. 1B), i.e. the protein partitioned via diffusion into the phase of shorter lipids where the hydrophobic mismatch was nearest to zero. For TMD5 the environment with short lipids was most favourite (Fig. 1C), since the protein here experienced an almost vanishing mismatch ($\Delta_s = -0.10 r_0$) while in the domain of long lipids it had a strong negative mismatch ($\Delta_l = -1.36 r_0$). In contrast, TMD6 preferred the phase of long lipids (Fig. 1D) albeit the corresponding mismatches nearly coincided ($\Delta_s = 0.65 r_0$ and $\Delta_l = -0.69 r_0$). This, at first sight unexpected, result can be rationalized by comparing the local perturbations associated with the mismatches: The positive mismatch $\Delta_s = 0.65 r_0$ is too large to be compensated by short lipids via a local widening of the membrane. The protein therefore has to tilt with respect to the bilayer normal [15] in this domain. This perturbation is larger than a local compression of (long) lipids and consequently TMD6 partitions into the thicker domain. In any case, these results demonstrate that proteins diffusively partition into the lipid phase with the smallest hydrophobic mismatch and the smallest perturbation of the lipid bilayer.
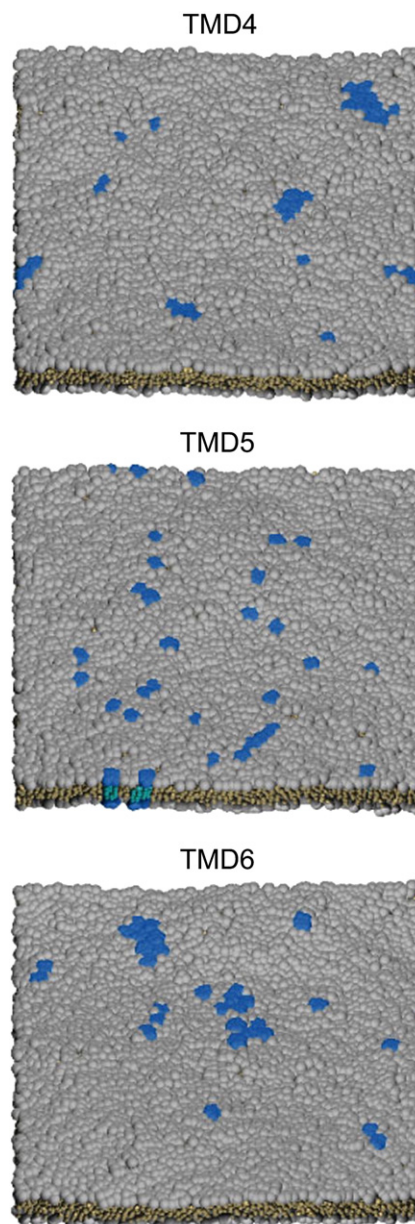
2.2. Transmembrane proteins with hydrophobic mismatch show clustering

To gain insights into the behaviour of proteins having different degrees of hydrophobic mismatching with the surrounding homogeneous membrane environment (representing, for example, various types of putative cargo in the ER), we restricted the membrane composition to the short lipid species. We thus had no pre-structuring of the membrane. We then introduced several proteins with the same TMD length. While TMD3 and TMD4 had negative mismatches, TMD5 was almost neutral; TMD6 and TMD7 showed positive mismatches in this bilayer.

In agreement with previous field-theoretical predictions [16,17] and simulations [15,18], we observed a clustering of proteins according to the strength of their hydrophobic mismatch (Fig. 2). While TMD5 had an almost vanishing mismatch, hence showing only transient dimer formation, TMD4 and TMD6 showed fairly stable oligomers. The same was observed for TMD3 and TMD7 (data not shown). Indeed, it has been shown earlier that such clusters can be stable over extended periods of time, depending on the effective pair potential that is induced by the hydrophobic mismatch [15]. We would like to emphasize at this point that the oligomerization is a result of entropic rather than enthalpic forces, similar to the formation of oil droplets in water: Minimizing the contact area between lipids and proteins by protein clustering reduces the number of lipids with constrained configurations (e.g. altered tilting angles [15]). As a consequence, the configuration space of lipids is larger and hence the entropy of the entire system is increased.

2.3. Proteins show a hydrophobic mismatch-dependent segregation

Given the observation that hydrophobic mismatching induces clustering, the obvious question arises whether this effect also allows


Fig. 2. Transmembrane proteins cluster on homogeneous membranes according to their hydrophobic mismatch. Representative snapshots of the steady-state configuration of TMD4, TMD5, and TMD6 in a homogeneous bilayer made from short lipids (from top). Proteins are colored in blue while lipids are shown in grey. While TMD4 and TMD6 show clustering due to their hydrophobic mismatch, TMD5 is mostly monomeric and only shows few dimers.

for a differential sorting. Indeed, embedding TMD4 together with TMD6 in a homogeneous membrane, we observed a segregation of the proteins into separate homo-oligomers (Fig. 3A). We quantified the degree of unmixing by the pair cross-correlation function (PCCF) which basically measures the amount of particles of species B (here: TMD4) in a distance r from proteins of species A (here: TMD6). As expected from the simulation snapshots, the PCCF vanished for small distances and became constant for larger distances. Hence, the probability of finding TMD4 next to TMD6 is negligible and a strong unmixing is observed.

The separation of TMD5 and TMD3,7 was similarly good (Fig. 3A) as TMD5 only has a weak tendency to form clusters at all (cf. Fig. 2). Indeed, since only TMD3 and TMD7 formed clusters, while TMD5 was mostly monomeric, the PCCF vanished for small distances, thus
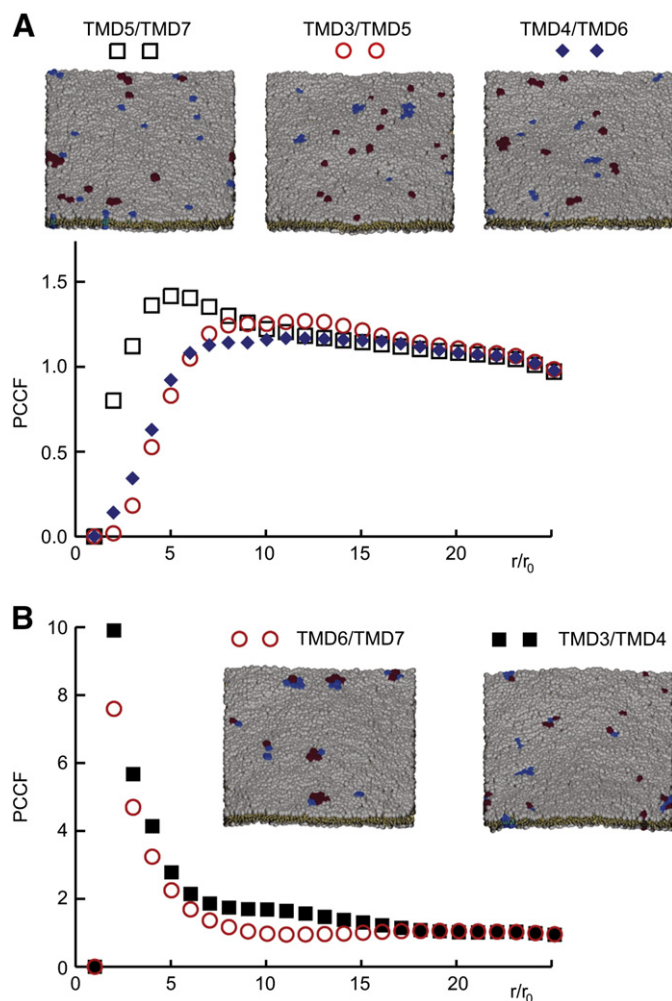
Fig. 3. Clusters of transmembrane proteins with different hydrophobic mismatches can segregate. (A) Snapshots of TMD3–TMD5, TMD5–TMD7, and TMD4–TMD6 mixtures. Shorter proteins are colored in blue, longer proteins in red, lipids are shown in grey. The apparent unmixing of the different proteins seen in the snapshots is confirmed by the pair cross-correlation function (PCCF), that shows a dip for small distances r. The distance is given in simulation units ($r_0 \approx 1$ nm; cf. Materials and methods). (B) In contrast, TMD3–TMD4, TMD6–TMD7 mixtures do not show any segregation but rather colocalization is observed. The formation of hetero-oligomers is confirmed by a strong peak in the PCCF for small distances.

indicating a good unmixing. In contrast, proteins with the same kind of hydrophobic mismatch, i.e. TMD3–TMD4 or TMD6–TMD7 mixtures, did not show an unmixing (Fig. 3B). Unlike the previous cases, the PCCF showed here a strong peak for small distances, reflecting the formation of hetero-oligomers that are also observed in the simulation snapshots. Thus, only if the hydrophobic mismatch is sufficiently different, an unmixing into distinct populations of homo-oligomers is observed which may be used for large-scale protein sorting. We also have tested whether the phenomenon of unmixing according to hydrophobic mismatch is observed on heterogeneous membranes. Using the same mixtures of proteins (TMD4–TMD5 and TMD4–TMD6) on a phase-separated bilayer (cf. Fig. 1A), we did not find any evidence that hetero-oligomers will form at the boundary of the different lipid phases (data not shown). Although TMD6 proteins approached the phase boundary, their shorter counterparts remained surrounded by short lipids, preventing the formation of hetero-oligomers in the boundary region. This result is anticipated as the proteins induce local perturbations in the bilayer (compression vs. widening) that repel each other on short length scales [19].

2.4. Model for hydrophobic mismatch-driven protein sorting between compartments

Having observed the clustering and potential segregation of proteins due to their hydrophobic mismatch, we discuss now how cells may use this phenomenon to promote protein sorting and trafficking along the secretory pathway (cf. sketch in Fig. 4). Given that organellar membranes along the secretory pathway show slightly different membrane thicknesses due to varying contents of cholesterol and sphingolipids, it is conceivable that a given transmembrane protein, e.g. a Golgi-resident glycosylation enzyme, may experience a negative mismatch in the ER after having passed the quality control machinery. It will hence form clusters that can modulate the local turnover rate of COPII proteins [20] to facilitate the formation of a COPII vesicle. Having pinched off, the vesicle eventually arrives at the Golgi apparatus and fuses with the cis-most cisterna. Here, the lipid environment may match well the desire of the transported protein. Upon cisternal maturation, the lipid thickness may increase, resulting in a positive hydrophobic mismatch. Again the protein will cluster and now modulate the COPI machinery [21,22] to stimulate the budding process that enables retrograde transport of the protein to another cisterna with a more appropriate lipid composition. Coat specificity, e.g. for COPI or COPII, in this model may be mediated by specific cytoplasmic residues [23].

It is worthwhile noting that cisternal maturation is not a necessary ingredient for this scenario. Using the vesicular transport model instead, some proteins may accidentally escape their preferred compartment (e.g. the cis-most cisterna). COPI-derived vesicles may then (using the above outlined process) help to retrieve the mislocalized proteins. Given the universality of the membrane-mediated interaction, it is also tempting to speculate that the oligomer formation by kin recognition may be traced back to a specific amino acid sequence in the TMD that enforces a hydrophobic mismatch that is specific for this kin. The previously proposed models would thus be traced back to the same physical origin.

The outlined scenario for facilitating the transport of mislocalized proteins by hydrophobic mismatching is consistent with very recent experimental observations on the transport behaviour of model proteins with variable transmembrane domains [13] and mutant forms of the lipoprotein receptor-related proteins LRP6 [11]. Testing the oligomeric status of such proteins in vivo, e.g. via fluorescence
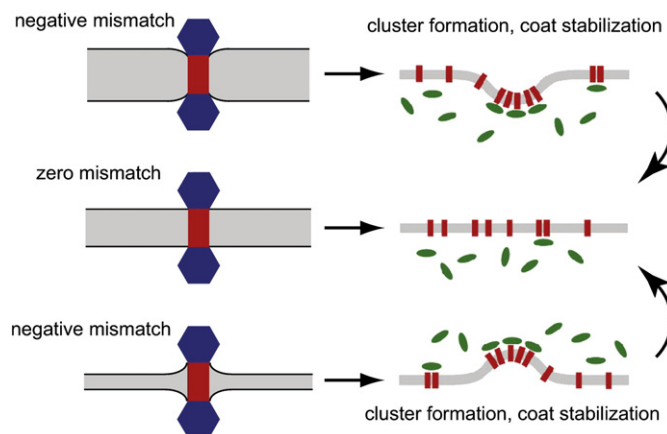
Fig. 4. Model for hydrophobic mismatching facilitated transport along the secretory pathway. Model of how a hydrophobic mismatch may facilitate transport of transmembrane proteins. If a mismatch (top: negative; bottom: positive) exists, clustering of transmembrane proteins is observed. These clusters can modulate the turnover rate of coat components (green ellipses) that shape an emerging vesicle at the locus of the protein cluster. This vesicle will be transported to a nearby compartment with a different membrane thickness, where the protein may not experience a hydrophobic mismatch any more (middle).

fluctuations spectroscopy techniques, will provide a direct test of the phenomena described here.

3. Materials and methods

3.1. Simulations

For our simulations, we have used coarse-grained membrane simulations, so-called dissipative particle dynamics, that are frequently used to perform simulations beyond the limitations of molecular dynamics [24]. In contrast to molecular dynamics, dissipative particle dynamics neglects molecular details on the sub-nanometer level but allows one to simulate large, fully hydrated membranes over periods of 10–100 µs. Thus, this technique is several orders of magnitude faster than molecular dynamics while still having sufficient details to be on the molecular level and to maintain all necessary physical phenomena, e.g. hydrodynamics. For technical details of the simulation, i.e. concerning the equations of motion and their integration as well as the definition of parameters and potentials, we refer the reader to Ref. [15]. In the following paragraph we restrict ourselves to a brief description of the simulations and the used parameters (in dimensionless simulation units, conversion to SI units yields roughly $r_0 \leftrightarrow 1$ nm and $\Delta t \leftrightarrow 80$ ps; cf. [15].

Lipids were taken as linear polymers HT_m with a single hydrophilic head and m hydrophobic tail beads (soft-core radius $r_0 = 1$), all connected by Hookean springs (spring constant $k = 100$, relaxation distance $l_0 = 0.45$). Hence, short lipids were HT_3 while long lipids were HT_6 constructs. In addition, lipids were given a bending rigidity (bending constant $\kappa = 20$). Proteins were modelled as hexagons of HT_nH chains (one central chain surrounded by six outer chains) using the same Hookean springs and bending stiffness. In addition to next-neighbour connections within each chain, each bead of the central chain was connected to its six counterparts of the outer chains. Moreover, beads of neighbouring outer chains were connected. The protein's cross-section hence consisted of 12 bonds, namely a hexagon in which the center and all vertices are fully connected.

The hydrophobic mismatch was varied by changing the number of beads in the transmembrane domain (n). Repulsion parameters between the individual beads (by which one tunes the degree of hydrophobicity) were set to $a_{WW} = a_{WH} = a_{HH} = a_{TT} = 25$, $a_{WT} = a_{HT} = 200$. Here the indices denote water (W), lipid head group (H), and lipid tail group (T). For heterogeneous membranes half of the short lipids in the upper leaflet was replaced by long lipids, and the interaction between heads of the short lipid and heads of the long lipids was set to 30 to allow for a more efficient phase separation. The time increment was $\Delta t = 0.01$, dissipation and noise of the thermostat were set to 3 and 4.5, respectively. In each run, we equilibrated the system for 50,000 time steps using a barostat, then we fixed the equilibrated system size and monitored the system's behaviour for another 500,000 time steps. The size of the membrane patches with area $L \times L$ was $L = 30r_0$ for the partitioning of the heterogeneous lipid system (cf. Fig. 1), and $L = 50r_0$ for observing the clustering and unmixing of proteins (15 proteins per species; cf. Figs. 2 and 3). A single simulation run with $L = 50r_0$ took about 9 weeks on a single CPU (AMD Opteron, 2.4 GHz).

3.2. Calculation of the pair cross-correlation function

To quantify the degree of co-clustering, we followed the approach outlined in [25] that basically employs a variant of the radial distribution function for two species. For any given radius r around a protein of species A we determined the density $\rho_{AB}(r)$ of proteins of species B within a circle of radius r. This was done for all proteins of species A for 500 equidistant snapshots of the time series. The ratio of $\rho_{AB}(r)$ and the total concentration of proteins of species B ($\rho_B = 15/L^2$) yields the pair cross-correlation function, $PCCF = \rho_{AB}(r)/\rho_B$.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bpc.2010.04.009.

Reference

[1] C. Barlowe, COPII-dependent transport from the endoplasmic reticulum, Current Opinion in Cell Biology 14 (2002) 417–422.
[2] L. Ellgaard, A. Helenius, Quality control in the endoplasmic reticulum, Nature Reviews 4 (2003) 181–191.
[3] M. Elsner, H. Hashimoto, T. Nilsson, Cisternal maturation and vesicle transport: join the band wagon! (Review), Molecular Membrane Biology 20 (2003) 221–229.
[4] C. Rabouille, N. Hui, F. Hunte, R. Kieckbusch, E.G. Berger, G. Warren, T. Nilsson, Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides, Journal of Cell Science 108 (Pt 4) (1995) 1617–1627.
[5] T. Nilsson, P. Slusarewicz, M.H. Hoe, G. Warren, Kin recognition. A model for the retention of Golgi enzymes, FEBS Letters 330 (1993) 1–4.
[6] M.S. Bretscher, S. Munro, Cholesterol and the Golgi apparatus, Science 261 (1993) 1280–1281 (New York, N.Y.).
[7] K. Mitra, I. Ubarretxena-Belandia, T. Taguchi, G. Warren, D.M. Engelman, Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol, Proceedings of the National Academy of Sciences of the United States of America 101 (2004) 4083–4088.
[8] S. Munro, A comparison of the transmembrane domains of Golgi and plasma membrane proteins, Biochemical Society Transactions 23 (1995) 527–530.
[9] T. Nilsson, M.H. Hoe, P. Slusarewicz, C. Rabouille, R. Watson, F. Hunte, G. Watzele, E.G. Berger, G. Warren, Kin recognition between medial Golgi enzymes in HeLa cells, The EMBO Journal 13 (1994) 562–574.
[10] O.G. Mouritsen, M. Bloom, Mattress model of lipid–protein interactions in membranes, Biophysical Journal 46 (1984) 141–153.
[11] L. Abrami, B. Kunz, I. Iacovache, F.G. van der Goot, Palmitoylation and ubiquitination regulate exit of the Wnt signaling protein LRP6 from the endoplasmic reticulum, Proceedings of the National Academy of Sciences of the United States of America 105 (2008) 5384–5389.
[12] G.H. Patterson, K. Hirschberg, R.S. Polishchuk, D. Gerlich, R.D. Phair, J. Lippincott-Schwartz, Transport through the Golgi apparatus by rapid partitioning within a two-phase membrane system, Cell 133 (2008) 1055–1067.
[13] P. Ronchi, S. Colombo, M. Francolini, N. Borgese, Transmembrane domain-dependent partitioning of membrane proteins within the endoplasmic reticulum, The Journal of Cell Biology 181 (2008) 105–118.
[14] S.L. Veatch, S.L. Keller, Seeing spots: complex phase behavior in simple membranes, Biochimica et Biophysica Acta 1746 (2005) 172–185.
[15] U. Schmidt, G. Guigas, M. Weiss, Cluster formation of transmembrane proteins due to hydrophobic mismatching, Physical Review Letters 101 (2008) 128104.
[16] N. Dan, P. Pincus, S.A. Safran, Membrane-induced interactions between inclusions, Langmuir 9 (1993) 2768–2771.
[17] N. Dan, A. Berman, P. Pincus, S.A. Safran, Membrane-induced interactions between inclusions, J Phys Ii 4 (1994) 1713–1725.
[18] F.J. de Meyer, M. Venturoli, B. Smit, Molecular simulations of lipid-mediated protein–protein interactions, Biophysical Journal 95 (2008) 1851–1865.
[19] J.N. Israelachvili, Refinement of the fluid-mosaic model of membrane structure, Biochimica et Biophysica Acta 469 (1977) 221–225.
[20] R. Forster, M. Weiss, T. Zimmermann, E.G. Reynaud, F. Verissimo, D.J. Stephens, R. Pepperkok, Secretory cargo regulates the turnover of COPII subunits at single ER exit sites, Current Biology 16 (2006) 173–179.
[21] J. Lanoix, J. Ouwendijk, A. Stark, E. Szafer, D. Cassel, M. Weiss, T. Nilsson, Sorting of Golgi resident proteins into different subpopulations of COPI vesicles: a role for ArfGAP1, The Journal of Cell Biology 155 (2001) 1199–1212.
[22] M. Weiss, T. Nilsson, A kinetic proof-reading mechanism for protein sorting, Traffic 4 (2003) 65–73 (Copenhagen, Denmark).
[23] T. Nilsson, M. Jackson, P.A. Peterson, Short cytoplasmic sequences serve as retention signals for transmembrane proteins in the endoplasmic reticulum, Cell 58 (1989) 707–718.
[24] J.C. Shillcock, Insight or illusion? Seeing inside the cell with mesoscopic simulations, HFSP Journal 2 (2008) 1–6.
[25] A.A. Philimonenko, J. Janacek, P. Hozak, Statistical evaluation of colocalization patterns in immunogold labeling experiments, Journal of Structural Biology 132 (2000) 201–210.