Coarse-Grained Monte Carlo Simulations of Mucus: Structure, Dynamics, and Thermodynamics

Pawel Gniewek and Andrzej Kolinski*
Laboratory of Theory of Biopolymers, Faculty of Chemistry, University of Warsaw, Warsaw, Poland

ABSTRACT A simple coarse-grained model of mucus structure and dynamics is proposed and evaluated. The model is based on simple cubic, face-centered lattice representation. Mucins are simulated as lattice chains in which each bead of the model chains represents a mucin domain, equivalent to its Kuhn segment. The remaining lattice sites are considered to be occupied by the solvent. Model mucins consist of three types of domains: polar (glycosylated central segments), hydrophobic, and cysteine-rich, located at the terminal part of the mucin chains. The sequence of these domains mimics the sequence of real mucins. Static and dynamic properties of the system were studied by means of Monte Carlo dynamics. It was shown that the model system undergoes sol-gel transition and that the interactions between hydrophobic domains are responsible for the transition and characteristic properties of the dynamic network in the gel phase. Cysteine-rich domains are essential for frictional properties of the system. Structural and dynamic properties of the model mucus observed in simulations are in qualitative agreement with known experimental facts and provide mechanistic explanation of complex properties of real mucus.

INTRODUCTION

Mucus, which covers many organs and tissues in the human body, is a major component of an innate defense system against environmental elements (1). The main role of mucus is to immobilize external bodies (e.g., viruses) or nanoparticles (e.g., dust) and then remove them from the organs or body. It is recognized that the structure and composition of mucus is similar among many organs and tissues. It is also commonly recognized that this system consists of two distinct parts: the luminal mucus layer (LML) and the adhesive mucus layer (AML) (1–4). The function and composition of these two layers are qualitatively different. Each layer has a different composition of water, ions, nucleic acids, proteins, lipids, and glycoproteins, the most important structural components of mucus. As a result, each layer has different physicochemical properties. The layers do not penetrate one another, but rather interact through a narrow interface (2,3).

Because of their different properties, the two layers of mucus play different roles in the defense system (1). The AML, which comprises very long and charged glycoproteins, is tightly packed and adsorbed at the cell surface. As a result, the mobility of AML molecules is restricted. Thus, when an external body comes in contact with this layer, it is retained there for a long time. The cleaning time for this layer is long (2–4).

In contrast to the AML the LML has a gel-like structure. The LML is more mobile than the AML, and consequently more rapidly cleared. Additionally, the LML contains some immersed species of antibodies and therefore constitutes the main part of the mucosal defense system. Despite this important role in the defense system, the mucus surface causes problems when medical treatment is needed. A number of medicines are ineffective because they cannot penetrate the LML in time short enough before being removed from the luminal layer, or even if they can, there is a need for high medicine concentration, so treatment costs rapidly increase. These issues of drug delivery make experimental and theoretical studies of the LML very important and interesting (1).

The LML is composed of ~90–95% water and 5% glycoproteins; the remainder is made up of ions, lipids, smaller proteins, and nucleic acids. Despite its large water content, viscous and elastic properties of the LML are controlled by large molecules called glycoproteins (mucins) and interpenetrating solvent. The protein composition of mucus slightly depends on organs and location in the body, although existing data show that MUC5AC and MUC5B are the most abundant mucins, whereas others play a less important or marginal role in mucus. MUC5B and MUC5AC are very large glycosylated proteins. The length of these molecules is variable, but usually it is ~5000 amino acids. MUC5B and MUC5AC sequences are highly modular and consist of many repeating domains. For secreted mucins, the central part is composed mostly of serine, threonine, and proline, which are densely covered by sugars (up to 20 sugar units/amino acid). This glycosylation makes the central part of mucins very rigid and highly charged. Polar domains are flanked by hydrophobic domains that are naked with respect to glycosylation. These hydrophobic domains are separated by cysteine-rich domains (1–3).

Due to the system complexity, experimental studies of mucus are difficult (5,6). Most investigations have been

Submitted July 29, 2010, and accepted for publication September 16, 2010.
*Correspondence: kolinski@chem.uw.edu.pl
Editor: Ruth Nussinov.

0006-3495/10/12/3507/10 $2.00
doi: 10.1016/j.bpj.2010.09.047

performed in vitro, with only a small amount of data related to in vivo studies. In vitro studies can provide some picture of mucus, but even so, the relation between in vitro and in vivo findings for mucus is not clear, since the environment (pH, ion concentration, etc.) can significantly change upon extraction from a living system. Since the AML is directly connected to the surface of a living cell, it is more difficult to study than the LML.

Despite much experimental work on gel-like mucus, the role of hydrophobic and cysteine-rich domains in gel formation, and whether their presence is essential for the sol-gel transition, remains unclear. Another yet unanswered question is what the role of internal cysteine-rich domains is in the formation of cross-links between mucin chains. From a medical point of view, one of the most important questions is the role of Ca^{2+} in mucus secretion and mucosal diseases and causes and mechanisms of these diseases (7,8). In this work, we provide a model-based description of thermodynamics, dynamics, and the molecular mechanism of mucus gel formation.

Coarse-grained mucus model

Soluble mucus (LML) consists mainly of water (~95%), although large glycoproteins are responsible for its viscous and elastic properties. There are many kinds of glycoproteins inside the mucus phase, but the most abundant are MUC5B and MUC5AC. In many different tissues, the length and domain composition of these mucins may vary, but the physicochemical behavior is very similar in all cases (1). Therefore, we assume in this study that all modeled mucins have the same length and the same domain structure/composition.

It is assumed that mucin structure can be represented by three types of distinct domains: cysteine-rich, hydrophobic, and polar. In the case of the cysteine-rich domains, this assumption is well justified because of strong sequence conservation in these domains. For the hydrophobic domains, the interactions are rather nonspecific, making them quite robust. Thus, they also could be treated as a single type. The sequences of the polar mucin domains, despite being commonly rich in serine, threonine, and proline, are not conserved among various species, which supports the idea that essential functionalities are governed by carbohydrates. Due to the apparent lack of well-defined patterns of amino acid glycosylation, it seems to be reasonable to assume that gel formation is a generic feature of mucins, regardless of glycosylation details (1). Accordingly, we assume the following generic primary structure of mucins: SHSHSPPPPPPPPPPPPHHS (System I), where S denotes a cysteine-rich domain (referred to henceforth as Cys), H is a hydrophobic domain, and P is a polar domain. This assumed generic sequence of model mucins is justified by inspection of the sequence of the chromosome 11p15.5, where MUC5AC and MUC5B are localized (9,10). It can be determined that polar segments comprise ~60% of the sequence, and the rest is made up of hydrophobic and cysteine-rich domains. Hydrophobic domains are larger then cysteine-rich domains, but because hydrophobic domains have denser packing than cysteine-rich domains, a larger number of amino acids in hydrophobic domains correspond to one Kuhn segment. Because of the nature of evolution of many mucins, which is governed by duplication of sequence fragments, polar parts of mucins are flanked by SH repetitions.

Biophysical studies of mucins have quite a long history (5,6). It is well established that the length of mucins in the extract is ~450–500nm (11,12). It has been found that a Kuhn's length of mucin chains varies from 25 to 30 nm (6,13), leading to the assumption that mucin chains can be represented by a random chain composed of 20 freely connected segments. Even with such simplification, simulation studies of the long-time behavior of many particles could be too expensive. To facilitate large-scale simulation, we confined the mucus system in a simple cubic face-centered (FCC) lattice. Such simplification has been extensively employed in studies of thermodynamics and dynamics of polymeric solutions and melts (14,15). The FCC lattice has a high coordination number, $z = 12$ (corresponding to close packing of spherical objects (16,17)), and lattice chain flexibility enables efficient simulation of their stochastic dynamics. The lattice representation allows a very simple form of interaction potential between chain segments. The excluded volume of mucin chains is modeled by exclusion of double occupancy of lattice sites. The neighboring sites interact via simple contact-type potentials, and interaction is considered only between segments that are not consecutive in the mucin sequence. Similar highly idealized models of proteins are very popular in studies of small proteinlike systems (18–20). The robustness of these simplified models is demonstrated in terms of a specific choice of parameters; changes of parameters in a reasonable range do not affect the basic physics of such systems (20). Longer-distance (electrostatic) interactions are neglected, which is justified by the large size of the equivalent segments of the model chains.

The following magnitudes of contact potentials between P-, H-, and S-type segments are assumed:

$E_{PP} = 0.0$ (neutral)
$E_{H/CysP} = 0.5$ (repulsive)
$E_{HH} = -0.5$ (attractive)
$E_{HCys} = -0.1$
$E_{CysCys} = -2.0$ (to mimic the covalent aspect, only one Cys-Cys interaction per segment is counted).

The total configurational energy of the model system is the sum of all contact interactions. The values of contact potentials are expressed in conventional dimensionless units of k_BT. Thus, the absolute values of the interaction parameters are not important. What is relevant is the relations

between particular parameters. Interactions between polar domains are set to 0, as for an athermal solution. The reason for the last assumption is that the carbohydrate side chains are hydrated and the charges at the glycosylated surface are neutralized by ions (e.g., Na^+, K^+). The strength of interactions between cysteine-rich domains has been set only four times higher than the strength of interactions between hydrophobic domains. The reason for this is that the specific level of the oxidation-reduction potential within the mucus environment enables reversible Cys-Cys cross-link formation. Reversibility of cross-linking, in which a fraction of Cys residues are reduced, is essential for mucus function, which requires a nonzero diffusion coefficient of mucins.

The dynamics of model chains is simulated via a dynamic Monte Carlo scheme, controlled by the Metropolis acceptance criterion. Fig. 1 illustrates conformational updates implemented in the Monte Carlo algorithm. These micro-modifications preserve chain connectivity and prevent chain-crossing events (21–23). The Monte Carlo process provides a solution to the master equation of motion and thereby delivers a picture of long-time dynamics of the model system.

To address the roles of hydrophobic and cysteine-rich domains in mucus gel formation, we introduced and simulated two hypothetical systems. In the first system (System II), we replaced all cysteine-rich domains from System I with hydrophobic domains, and in the second (System III), we replaced hydrophobic domains with polar domains to find out whether the cysteine-rich domains are sufficient for mucus gel formation. To answer the question of the possible Ca^{2+} effect on mucus structure and dynamics, we kept the sequence of System I, modified by a possibility of attractive interactions between serine/threonine-rich domains via Ca^{2+} salt bridges between charged hydrocarbons (System IV). In this system, $E_{PP} = -0.2$. The nature of these weak attractive interactions comes from the observed mechanism of mucus goblet secretion, where densely packed mucus goblets, with their water-deprived interior, are enriched in Ca^{2+}. This behavior is counterintuitive, because polar segments are collapsed in the presence of water. It is assumed that positive polyions shield the negative charges of polar segments of mucins. This effect is not observed for single positive ions like K^+,Na^+, Thus, the double-positive charges, and the resulting salt bridges, lead to net attractions of the polar segments, although the exact nature of this shielding is still not known (7,24,25). These interactions are relatively weak, and after mucus goblet secretion and Ca^{2+} ion removal, the polar segments rapidly swell, despite the fact that cysteine bonds and hydrophobic interactions remain intact (7).

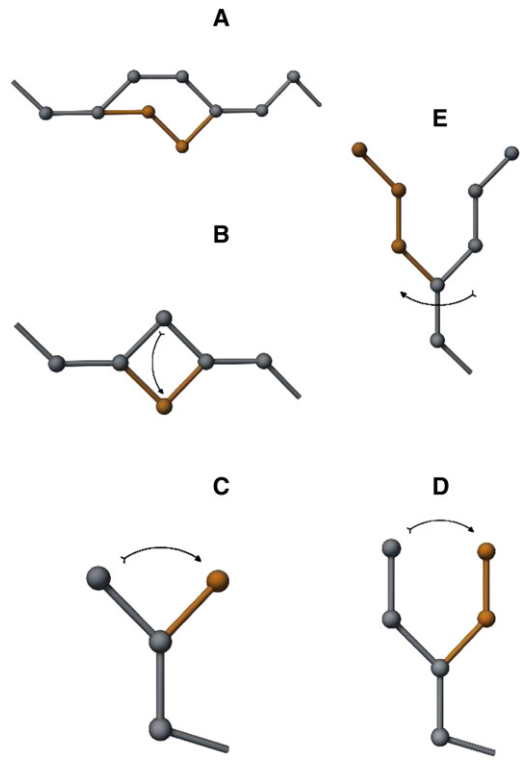

FIGURE 1 Local chain modifications in the Monte Carlo sampling scheme. (*A*) Three-bond permutation. (*B*) One-bead kink. (*C*) Pivot for a single terminal segment. (*D*) Pivot for two terminal segments. (*E*) Pivot for three terminal segments.

METHODS

Lattice model

The model systems were restricted to the FCC lattice with superimposed periodic boundary conditions. In most simulations, the cubic periodic box had an edge of $L = 39$ (Fig. 2). Different concentrations were obtained by changing the number of model chains, assuming $N = 77$, 154, and 230 for $C = 5$, 10, and 15% volume fractions, respectively. The exception was the case where $C = 2\%$ with box size $L = 45$ and number of chains $N = 47$. Unoccupied lattice sites represent solvent. The length of a single polymer segment in the FCC in the lattice units is equal to $\sqrt{2}$ for Cartesian coordinates, which corresponds to 25 nm according to the Kuhn-length-based scaling described above. Stochastic dynamics in the NVT ensemble of model chains was modeled by long random sequences of local conformational transitions. The time unit of the model dynamics corresponds to the time needed for one attempt at conformational transition per chain segment on average. The set of local moves consists of one bead kink, pivots for one, two, and three terminal segments, and three-bond permutations (Fig. 1). The random sequence of such sets of moves assures ergodicity of the sampling scheme, at least in a practical sense (some nontypical, very dense conformations could be not accessible). The design of local moves prevents chain-crossing (or phantomization).

Annealing

Simulated annealing was performed starting from temperature $T_{start} = 1.5$, and ending at $T_{end} = 0.01$. The system was periodically cooled down every 50,000 time units. The first 10,000 time units after each cooling event were disregarded in the calculation of observables and treated as an equilibration cycle system. The cooling schedule was designed as follows:

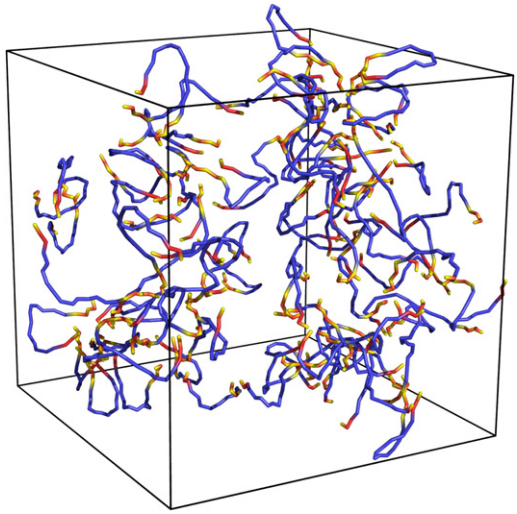

FIGURE 2 A snapshot of 77 mucin chains in a periodic Monte Carlo box. For clarity, the chains are smoothened by replacing coordinates of lattice polymer units with average coordinates of three subsequent segments. Blue segments correspond to polar domains, red segments to hydrophobic domains, and yellow segments to cysteine-rich domains.

$T_i = \gamma T_{i-1}$ when $T_{i-1} > T_{lin}$, and $T_i = T_{i-1} - \Delta T$ otherwise. The settings used were $T_{lin} = 1.0$, $\gamma = 0.96$, and $\Delta T = 0.015$ (26).

To calculate observables we collected data in intervals of 500 time units.

Heat capacity

Heat capacity as a function of temperature could be calculated using the formula

$$C_v(T_i) = \frac{\left\langle \left(E^{(i)} - \langle E^{(i)} \rangle \right)^2 \right\rangle}{T_i^2},$$

where $E^{(i)}$ means energy in the ith annealing step. Due to the data scattering, the heat capacity readings were averaged over five neighboring temperature readings.

Contacts

Patterns of intra- and interchain contacts were stored every 500 time units. Only nonbonded contacts were included. The coordination number of the FCC lattice is $z = 12$. Thus, the maximum number of nonbonded contacts (except for the chain ends) is equal to 10. Due to the specific design of the interaction scheme for cysteine-rich domains, only one Cys-Cys contact was counted for each Cys segment. Contacts between cysteine-rich domains and other segments were counted as for other types of contacts. For each temperature, the average ratio of observed contacts to the maximum number of contacts was calculated and stored for further analysis.

Topology

Similar to the other statistics, the network topology was analyzed every 500 time units. Distinct topologies were considered as schematically depicted in Fig. 3. Free-chain topology includes chains that have only intrachain contacts (Fig. 3 *A*). Loop topology was assigned to chains having intrachain contacts between the N- and C-terminal segments and an additional hydrophobic contact with another chain (Fig. 3 *D*). For chains with hydrophobic contacts at both termini with different termini of other chains, and without loop-type intrachain contacts, the topology is called a bridge (Fig. 3 *B*). Dangling topology (Fig. 3 *C*) was assigned to chains interacting with other chains but without loop or bridge topology.

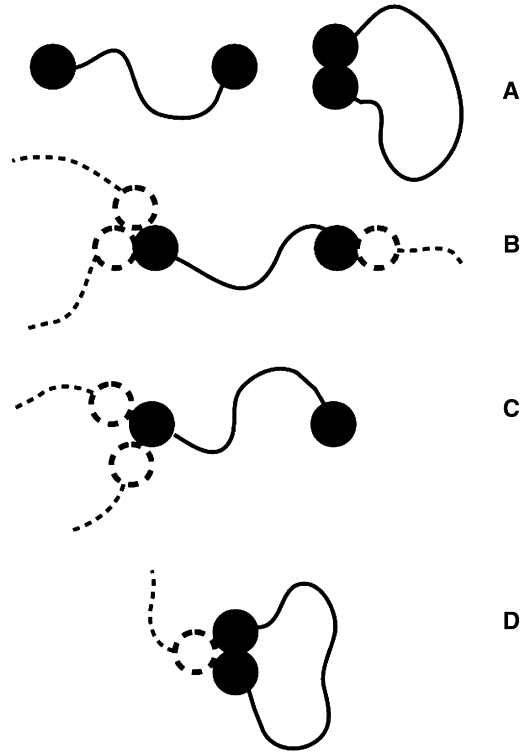

FIGURE 3 Possible mucin chain topologies in a mucus system. (*A*) Free chain. (*B*) Bridge. (*C*) Dangling chain. (*D*) Loop.

Gyration tensor, end-to-end distance, and radius of gyration

Radius of gyration and end-to-end distance were calculated for each chain and its hydrophobic and polar parts at 500-time-unit intervals. Averaged values were then calculated at various temperatures of the annealing protocol. In addition, for some systems, asphericity characteristics of entire chains were monitored (27).

The end-to-end distance is defined in a conventional way:

$$\mathbf{R}_e = |\mathbf{R}_N - \mathbf{R}_1|,$$

where $\mathbf{R}_N$ and $\mathbf{R}_1$ are Cartesian coordinates of the last and the first segments, respectively. In a similar way, the end-to-end distances for subchains (hydrophobic and polar) were monitored. Asphericity and radius of gyration were calculated from the gyration tensor, defined as

$$S_{mn} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{r}_m^{(i)} \mathbf{r}_n^{(i)},$$

where $r^{(i)}{}_m$ ($r^{(i)}{}_n$) is the mth (nth) coordinate of the position of the ith segment of the polymer chain. The origin of the coordinate system for $\mathbf{r}^{(i)}$ vectors is placed at the center of gravity of the molecule. To calculate the center of gravity for each mucin, the mass of segments was assumed to be 1.0. For

such established centers of gravity, the m, n index for $\mathbf{r}^{(i)}$ varies from 1 to 3 (which corresponds to x, y, and z components of the $\mathbf{r}^{(i)}$ vector).

The tensor of gyration could be diagonalized to extract its principal components, $\lambda_x^2 \geq \lambda_y^2 \geq \lambda_z^2$:

$$\mathbf{S} = \begin{pmatrix} \lambda_x^2 & 0 & 0 \\ 0 & \lambda_y^2 & 0 \\ 0 & 0 & \lambda_z^2 \end{pmatrix}$$

Radius of gyration and asphericity could then be calculated as

$$\mathbf{R}_G^2 = \lambda_x^2 + \lambda_y^2 + \lambda_z^2$$

and

$$b = \lambda_x^2 - \frac{1}{2}\left(\lambda_y^2 + \lambda_z^2\right).$$

Radial distribution functions and time autocorrelation functions

Radial distribution was calculated for pairs of hydrophobic and cysteine-rich domains, pairs of polar domains, and all types of domains together:

$$\mathrm{G}(\mathbf{r}) = \frac{\mathrm{V}}{\mathrm{N}^2}\sum_{\mathrm{i}}\sum_{\mathrm{j}\neq\mathrm{i}}\left\langle\delta\left(\mathbf{r}-\mathbf{r}_{\mathrm{ij}}\right)\right\rangle,$$

where V is the volume of the system, N is the number of polymer segments in the system, δ function is an operator which counts histogram bins for the occurrence of distances between ith and jth segments. The radial distribution functions are normalized so that uniform distribution corresponds to 1.0. For selected temperatures, radial distribution functions were calculated from long (10^6 time units) isothermal simulations at 10,000-time-unit intervals and then averaged over all snapshots.

To monitor chain autorelaxation, the formula used was

$$\mathrm{c_{AA}(t)} = \frac{\langle\mathbf{R}_\mathrm{i}(t)\mathbf{R}_\mathrm{i}(0)\rangle}{\langle\mathbf{R}(0)\mathbf{R}(0)\rangle},$$

where t is time, $\mathbf{R}$(t) is the vector connecting the first and last segments of a chain at time t. The averages were computed for all chains and sliding windows of simulation time at various time intervals. Due to statistical uncertainty, for very long times the correlation functions were analyzed only for $\mathrm{t} < 1/2\mathrm{t_{max}}$, where $\mathrm{t_{max}}$ is the simulation length.

For the transport autocorrelation function, the formula used was

$$\mathrm{g(t)} = \left\langle\left\langle\left|\mathbf{r}(\mathrm{t})-\mathbf{r}(0)\right|^2\right\rangle\right\rangle,$$

where $\mathbf{r}$(t) is the position of a segment at time t. The averaging was done in the same fashion as for the end-to-end autocorrelation function. Diffusion coefficients were estimated from log-log plots of autocorrelation functions. At a short timescale, mucins follow a subdiffusion behavior, typical of polymers. A straight line was fitted to the long-timescale regime where displacements are larger than an average radius of gyration and extrapolated to time t = 0 to obtain a log(D) value, where D is the diffusion coefficient (28).

RESULTS AND DISCUSSION

Thermodynamics

Simulated annealing computational experiments were performed for the model systems. To check possible concentration effects for System I, four mucin concentrations, C = 2, 5, 10, and 15%, were simulated. These simulations showed that the transition temperature ($\mathrm{T_f}$) of System I (see Fig. 4) does not change significantly with concentration. Moreover, the average energies per chain at different concentrations do not differ by much (see Fig. 5). The $\mathrm{T_f}$ values obtained for System II were much higher than those for System I at

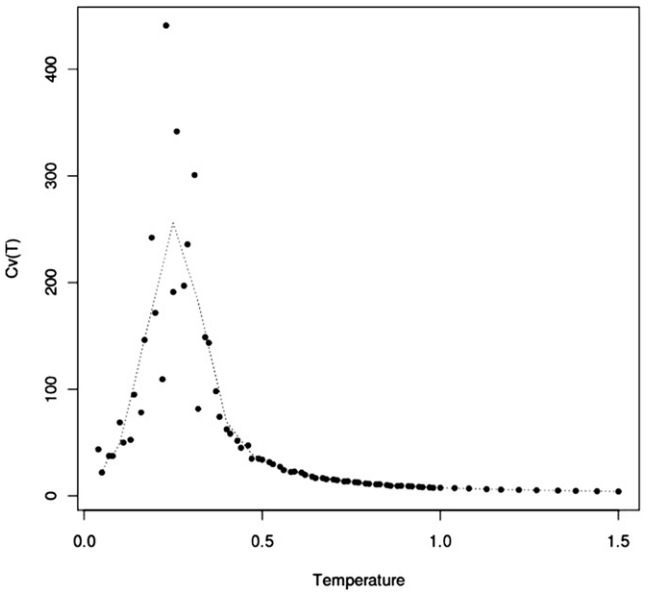

FIGURE 4 Heat capacity as a function of temperature for C = 5% concentration of mucins and System I. The dotted line represents smoothened data obtained by averaging of five consecutive solid circles. The transition temperature from the picture is ~0.25. For Systems II and IV, the transition temperatures from analogous pictures are 0.77 and 0.55, respectively (data not shown).

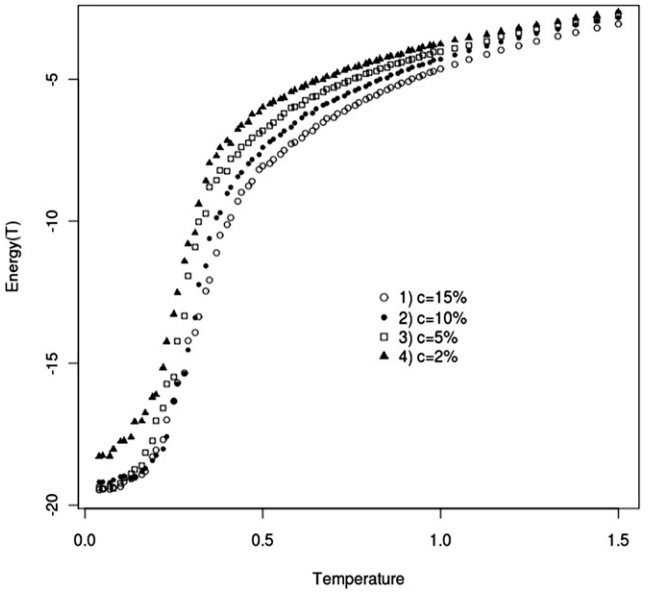

FIGURE 5 Average energy per mucin chain as a function of temperature for the four different concentrations in System I.

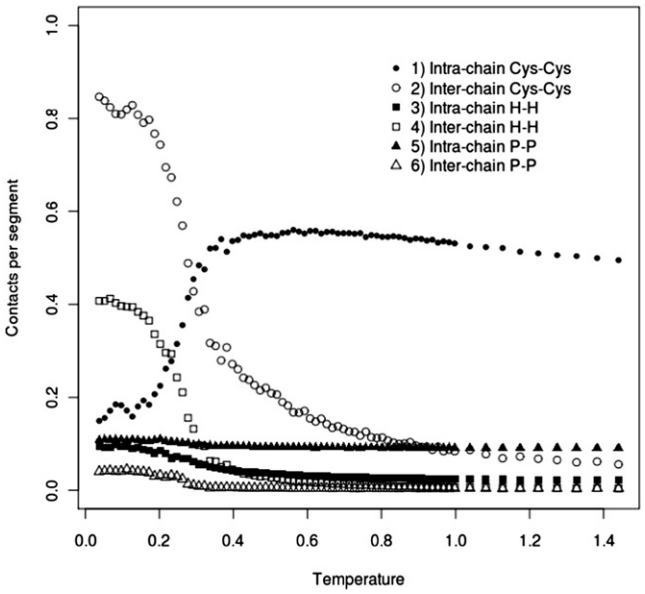

FIGURE 6 Average number of intrachain (*solid symbols*) and interchain (*open symbols*) interactions per chain segment (see Methods) as a function of temperature in System I at C=5% concentration of mucins.

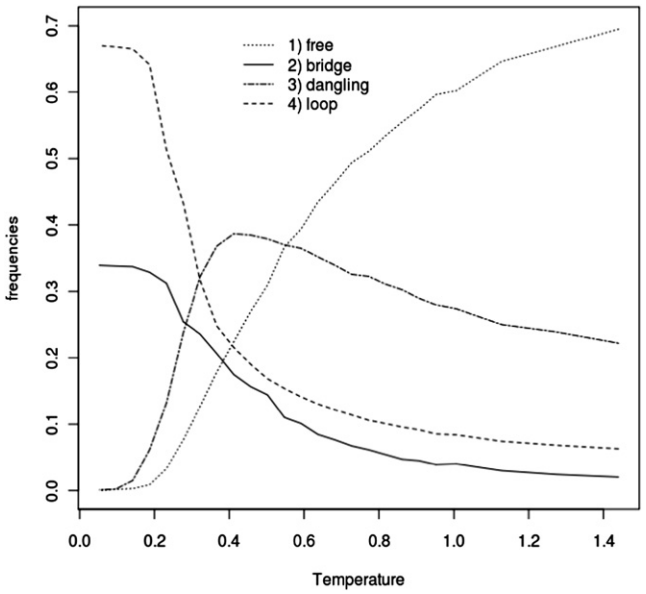

FIGURE 7 Frequency of various types of mucin network topologies (see Methods) as a function of temperature for System I at C = 5% concentration of mucins.

any concentration, and for System III, no clear phase transition was observed. For System IV, the T_f values fell between those observed for Systems I and II. Analysis of the number of contacts per chain as a function of temperature showed that in a sol phase there are almost no intra- and interchain interactions other than intrachain Cys-Cys interactions and a small number of intrachain polar-polar interactions (Fig. 6). Polar-polar interactions probably originate from the random-coil-contact contribution, and the number of these interactions does not change noticeably with temperature. At temperatures above the T_f, the contribution from intrachain Cys-Cys interactions is large, whereas the number of hydrophobic interactions remains low. This situation changes dramatically upon cooling of the system, when the Cys-Cys intrachain interactions are converted into interchain interactions. The interchain hydrophobic-hydrophobic contribution also increases during the cooling process (Fig. 6). For System II, we observed a very similar situation, although without the Cys-Cys contribution. In System IV, we observed almost the same picture as for System I, with some additional contribution of polar-polar interchain interactions accompanied by phase transition (for details regarding Systems II–IV, see the Supporting Material).

Structure

The sol-gel transition is accompanied by significant changes in topology of the cross-linked network of mucins. In the case of System III (for which we did not obtain sol-gel transition), at low temperatures, system topology is mostly limited to free chains and dangling chains (probably two head-to-tail connected chains) with a small number of bridges and loops. A qualitatively different network topology was observed for Systems I and II, although between these two systems it was very similar (Figs. 7 and 8), which leads to the conclusion that hydrophobic domains play a crucial role in gel formation. In these cases, we observe a main contribution from loops and a smaller one from bridges. Dangling topology and free chains are

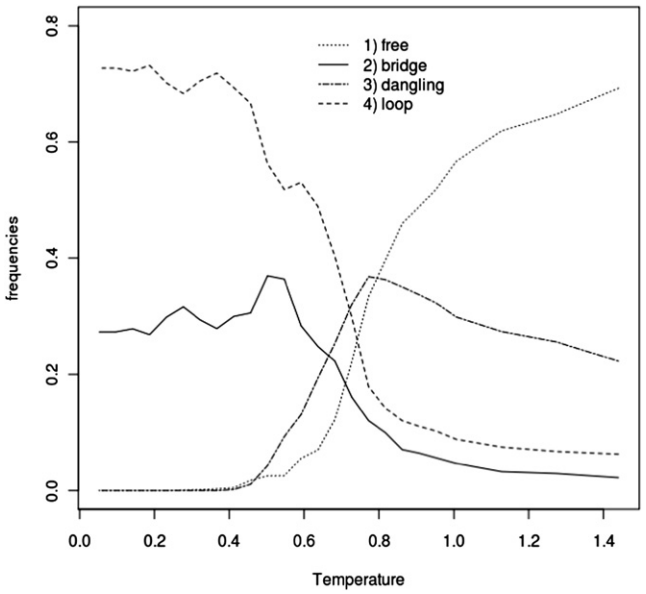

FIGURE 8 Frequency of various types of mucin network topologies (see Methods) as a function of temperature for System II at C = 5% concentration of mucins.

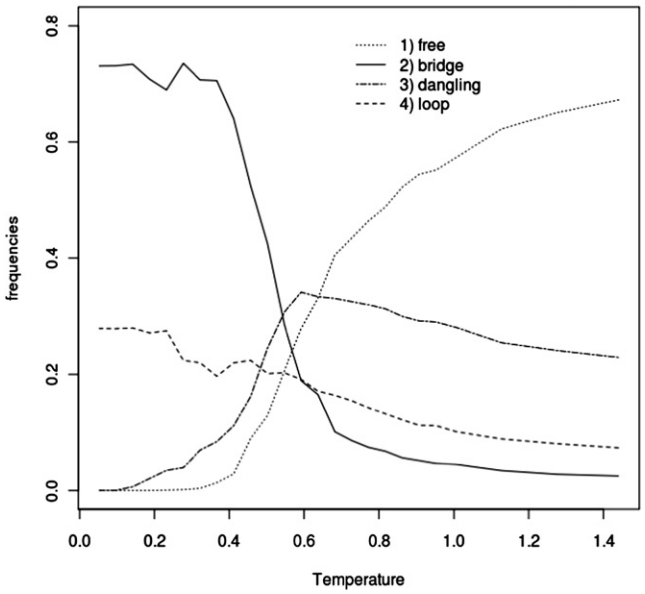

FIGURE 9 Frequency of various types of mucin network topologies (see Methods) as a function of temperature for System IV at C = 5% concentration of mucins.

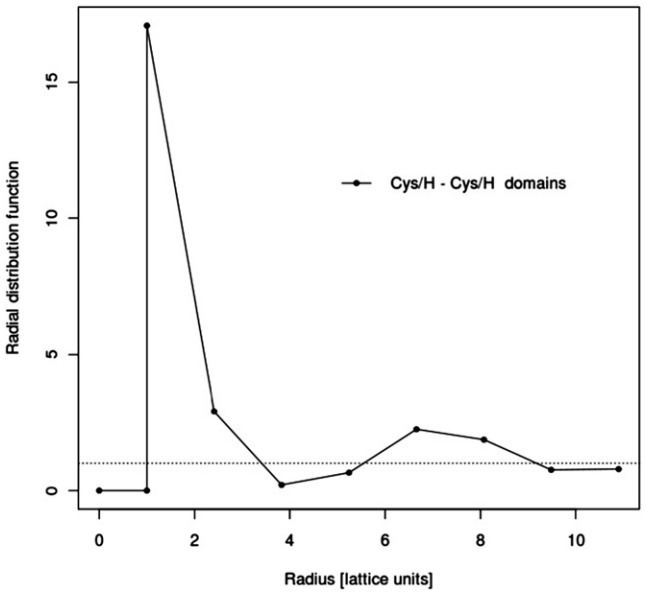

FIGURE 10 Radial distribution function for cysteine-rich and hydrophobic domains at temperature T = 0.2 for System I at C = 5% concentration of mucins. A radial distribution function of 1.0 indicates uniform, random distribution. One lattice unit corresponds to ~25 nm in a real system.

absent in the gel phase. The emerging overall picture is that hydrophobic bundles (or blobs) are connected by bridges. In the case of System IV, the picture of topology is quite different. The role of loops and bridges with respect to Systems I and II is inverted. Bridges play a major role in the gel network structure (see Fig. 9).

To compare the shape of mucin molecules forming the gel in System I and IV, tensors of gyration were analyzed as a function of temperature. In the case of System I, a disklike picture shape of mucins is observed, whereas for System IV an asymmetric oval-like shape is evident.

Other characteristics of mucin molecules, such as the average radius of gyration and end-to-end distances, also change significantly upon cooling. For System I, both measures indicate highly extended conformations of entire chains, and especially of their polar parts, which determine the overall shape of molecules. For hydrophobic domains, a systematic decrease of end-to-end distances and radii of gyration is observed during cooling until the T_f, after which chain dimensions increase. This reentrant increase in the dimensions of hydrophobic sections is caused by the stretching effect of the network.

The analysis of radial distribution function for hydrophobic and cysteine-rich domains at low temperature, T = 0.2, is quite insightful (Fig. 10). There is a well-defined signal from the network pattern with distances between peaks of distribution equivalent to approximately six chain units, which corresponds to 180 nm in real systems. Conversely, the radial distribution function for polar segments does not exhibit any specific patterns: the glycosylated segments are distributed uniformly in space, forming connections between the hydrophobic blobs. In the case of radial distribution function for all monomers, we also observe two peaks separated by approximately six lattice units, but the network pattern is less clear because of the shadowing effect of polar segments.

Dynamics

Dynamic properties of the model system were studied in long isothermal Monte Carlo runs. The simulations were performed for System I at C = 5%, which is a typical concentration for mucins in mucus. Short time (and small distance) mobility of various parts of chains is similar, with somewhat faster local relaxation of polar sections. On a longer timescale, the picture is qualitatively different. In general, hydrophobic sections of model mucins relax their conformations much faster, at least at temperatures around the transition temperature (or just below). The decay of end-to-end vectors has a single exponential character, which is typical for flexible polymers in solutions. During a single simulation run (10^6 time units), the hydrophobic chains change their global conformations many times. A similar picture emerges from relaxation times for polar parts of mucins and for a full chain: they also relax many times during the simulation. If we compare the relaxation time of the full chain at different temperatures below and above the T_f, we observe sharp changes in relaxation time at temperatures just above the T_f, which reflect the loss of network connectivity after heating to temperatures above the T_f (see Fig. 11).

We also monitored characteristic relaxation times for Systems I, II, and IV at corresponding T_f values. Analysis

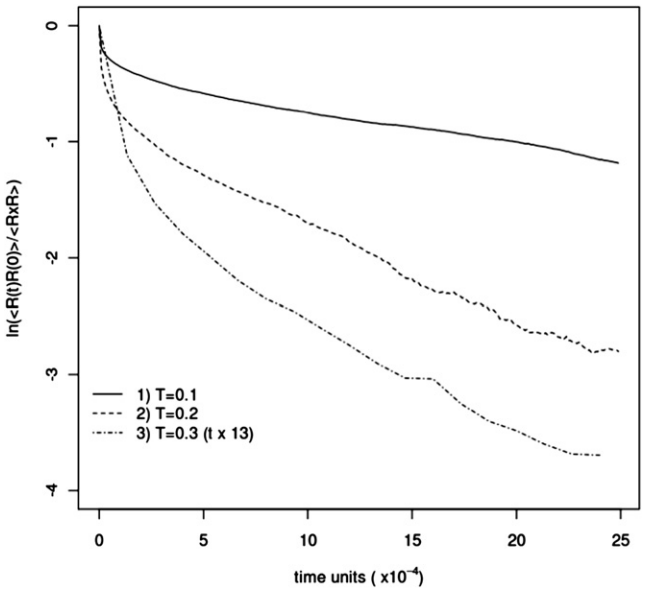

FIGURE 11 Relaxation of the end-to-end vectors of the full mucin chains for three different temperatures: T = 0.1 (*solid line*), T = 0.2 (*dashed line*), and T = 0.3 (*dash-dotted line*) in System I. For temperature T = 0.3, time is multiplied by 13 because of very fast relaxation. A Monte Carlo scheme time unit of the simulation process corresponds to the time required for, on average, one attempt at conformational transition per polymer segment. Log at the vertical axis means natural logarithm.

of the simulation results shows that among the three cases, System II relaxes the fastest. Due to high local density and a large number of interchain polar interactions, System IV is the slowest-relaxing system.

To compare the transport properties of these systems, we calculated diffusion transport coefficients from autocorrelation displacement functions. Differences in diffusion coefficients (see Table 1) follow the different local mobility of chains monitored via relaxation curves.

Role of cysteine-rich and hydrophobic domains

To answer the question about the role of hydrophobic and cysteine-rich domains in gel formation, we performed simulations for Systems I–III. Analysis of the $C_v(T)$ curve for the case of System II shows that the presence of hydrophobic domains is sufficient and necessary for gel formation. Moreover, the topology of the dynamic network formed by mucins so modified is similar to that observed in System I (compare Fig. 7 with Fig. 8). Such a similarity shows clearly

TABLE 1 Ratios of diffusion coefficients for three different systems at their transition temperatures

D_H/D_0	D_0/D_0	D_0/D_A
3.1	1	2.9

D_0, D_H, and D_A denote diffusion coefficients for systems I, II, and IV, respectively.

that the structure of the gel is controlled by generic hydrophobic interactions. For System III, the sol-gel transition is not observed, probably due to the effect of repulsive interactions between polar and cysteine-rich domains. It seems reasonable to speculate that cysteine-rich domains located between the hydrophobic domains make the mucin chains more flexible. In literature, another hypothesis exists on the role of cysteine-rich domains in the mucin network. Specifically, it is postulated that only terminal cysteine-rich domains create interchain bonds, whereas the remaining cysteine-rich domains scattered between naked hydrophobic sections form intrachain links that maintain suitable shape and positions of hydrophobic domains within the interaction network between the chains (5). Our simulations clearly show that this is not the case. Analysis of contributions from different types of cross-links at various reduced temperatures shows that intrachain cross-links occur only in the sol phase, where these bonds determine the overall shape of naked, hydrophobic parts due to a very small number of hydrophobic interactions. Upon sol-gel transition, the fraction of these intrachain contacts quickly decays (Fig. 6). To further resolve the role of cysteine-rich domains, we compared transport coefficients (Table 1) for Systems I and II at their transition temperatures. From these data, we can see that mucins, which have cysteine-rich domains, have lower transport coefficients, probably caused by interchain disulfide bonds. Thus, it could be concluded that despite the crucial role of hydrophobic domains in gel formation, viscosity and elastic properties of mucus are controlled by cysteine-rich domains.

Network structure in the gel phase

An interesting insight into the structure of mucus is gained by analysis of radial distribution functions for domains of various types. For hydrophobic and cysteine-rich domains, one can observe an internal structure, i.e., two peaks in the pairwise radial distribution functions, which indicates partial long-range spatial ordering of these domains (Fig. 10). The polar domains do not exhibit such ordering: in these domains, there is only one peak, resulting mostly from neighbors along the chain (the height of this peak is about half that of the peak for hydrophobic domains). Assuming that the peaks of the radial distribution function for hydrophobic and cysteine-rich domains correspond to nodes of a fluctuating network of mucins and the intervening regions of lower density correspond to pores in that network, the size of such a pore is ~180 nm. This estimation is in good qualitative agreement with experimental data, showing a pore size of 100–200 nm (2,3,29,30). It is important to note that this size pore is strongly dependent on mucin concentration, ionic strength, etc. (6).

It is known that Ca^{2+}-ion concentration has a strong influence on mucus structure and its frictional properties. This

issue has important physiological connotations, since elevated Ca^{2+} concentration is typical for patients with respiratory diseases, e.g., cystic fibrosis (8). To model the effect of Ca^{2+} ions on mucus properties, we modified System I by changing the nature of the repulsive interactions between the polar domains to weakly attractive. The attractive interactions between the polar segments mimic the effect of salt bridges between the OH groups of glycoproteins. Simulated annealing of such a modified model of mucins leads to qualitative changes in gel topology, resulting from an additional contribution from interchain polar interactions. The observed shapes of mucin molecules differ significantly in Systems I and IV. Mucus structure in the presence of Ca^{2+} ions is highly packed and has granular structure. A comparison of the transport properties of Systems I and IV also shows strong effects of ionic strength. The diffusion coefficient calculated for System IV is around three times lower than the coefficient for System I, which is in good agreement with experimental evidence showing an approximately two- to fourfold decrease in the presence of Ca^{2+} ions (5,8).

CONCLUSIONS

The proposed coarse-grained model of mucus structure and dynamics explains many properties of real mucus. The system undergoes sol-gel transition. In the gel phase, mucin chains form a dynamic network with complex topology of chain interactions. The observed size of pores in the network is in qualitative agreement with experimental estimations. Computational study of various hypothetical systems allowed us to elucidate the role of different domains in mucus properties. In particular, simulations of mucin chains lacking the cysteine-rich domains show that hydrophobic interactions are mainly responsible for the sol-gel transition, whereas cysteine-rich domains control the frictional properties of mucus. Modeling a system in which polar domains attract one another (as is possible at high concentrations of Ca^{2+} ions, where salt bridges are formed between OH groups of glycosylated segments) shows that at such conditions, mucus loses its gel-like properties.

In future research, the model proposed here will be applied in studies of diffusion of various nanoparticles throughout mucus layers, which may provide some new insights into the biological function of mucus. In particular, such studies could be helpful in designing new means of drug delivery across the mucus.

SUPPORTING MATERIAL

Fourteen figures are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(10)01197-5.

The computational part of this work was carried out using the computer cluster at the Computing Centre of the Faculty of Chemistry, University of Warsaw, Warsaw, Poland.

This work was partially supported by the Polish Ministry of Science and Higher Education (grant N209150036).

REFERENCES

1. Thornton, D. J., K. Rousseau, and M. A. McGuckin. 2008. Structure and function of the polymeric mucins in airways mucus. *Annu. Rev. Physiol.* 70:459–486.
2. Lai, S. K., Y. Y. Wang, …, J. Hanes. 2009. Micro- and macrorheology of mucus. *Adv. Drug Deliv. Rev.* 61:86–100.
3. Lai, S. K., Y. Y. Wang, and J. Hanes. 2009. Mucus-penetrating nanoparticles for drug and gene delivery to mucosal tissues. *Adv. Drug Deliv. Rev.* 61:158–171.
4. Hattrup, C. L., and S. J. Gendler. 2008. Structure and function of the cell surface (tethered) mucins. *Annu. Rev. Physiol.* 70:431–457.
5. Verdugo, P. 1990. Goblet cells secretion and mucogenesis. *Annu. Rev. Physiol.* 52:157–176.
6. Bansil, R., E. Stanley, and J. T. LaMont. 1995. Mucin biophysics. *Annu. Rev. Physiol.* 57:635–657.
7. Kesimar, M., A. M. Makhov, …, J. K. Sheehan. 2010. Unpacking a gel-forming mucin: a view of MUC5B organization after granular release. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 298:L16–L22.
8. Raynal, B. D. E., T. E. Hardingham, …, D. J. Thornton. 2003. Calcium-dependent protein interactions in MUC5B provide reversible cross-links in salivary mucus. *J. Biol. Chem.* 278:28703–28710.
9. Desseyn, J. L., V. Guyonnet-Duperat, …, A. Laine. 1997. Human mucin gene MUC5B, the 10.7-kb large central exon encodes various alternate subdomains resulting in a super-repeat. Structural evidence for 11p15.5 gene family. *J. Biol. Chem.* 272:3168–3178.
10. Escande, F., J. P. Aubert, …, M. P. Buisine. 2001. Human mucin gene MUC5AC: organization of its 5′-region and central repetitive region. *Biochem. J.* 358:763–772.
11. Sheehan, J. K., K. Oates, and I. Carlstedt. 1986. Electron microscopy of cervical, gastric and bronchial mucus glycoproteins. *Biochem. J.* 239: 147–153.
12. Sheehan, J. K., and I. Carlstedt. 1990. Electron microscopy of cervical-mucus glycoproteins and fragments therefrom. The use of colloidal gold to make visible 'naked' protein regions. *Biochem. J.* 265: 169–177.
13. Perez-Vilar, J. 2007. Mucin granule intraluminal organization. *Am. J. Respir. Cell Mol. Biol.* 36:183–190.
14. Kolinski, A., J. Skolnick, and R. Yaris. 1987. Does reptation describe the dynamics of entangled polymer systems? A model simulation. *J. Chem. Phys.* 86:1567–1585.
15. Binder, K., M. Müller, and J. Baschnagel. 2004. Polymers models on the lattice. *In* Simulation Methods for Polymers. M. J. Kotelyanskii, and D. N. Theodorou, editors. Marcel Dekker, New York. 125–146.
16. Cipra, B. 1998. Mathematics: packing challenge mastered at last. *Science.* 281:1267.
17. Sloane, N. J. A. 1998. Kepler's conjecture confirmed. *Nature.* 395: 435–436.
18. Dill, K. A., S. Bromberg, …, H. S. Chan. 1995. Principles of protein folding—a perspective from simple exact models. *Protein Sci.* 4:561–602.
19. Pokarowski, P., K. Droste, and A. Kolinski. 2005. A minimal protein-like lattice model: an α-helix motif. *J. Chem. Phys.* 122:214915–214916.
20. Pokarowski, P., A. Kolinski, and J. Skolnick. 2003. A minimal physically realistic protein-like lattice model: designing an energy landscape that ensures all-or-none folding to a unique native state. *Biophys. J.* 84:1518–1526.
21. Kolinski, A., J. Skolnick, and R. Yaris. 1985. On the short time dynamics of dense polymeric system and the origin of the glass transition: a model system. *J. Chem. Phys.* 84:1922–1931.

22. Kolinski, A., J. Skolnick, and R. Yaris. 1987. Monte Carlo studies on the long time dynamic properties of dense cubic lattice multichain systems. II. Probe polymer in a matrix of different degrees of polymerization. *J. Chem. Phys.* 86:7174–7180.

23. Kolinski, A., and J. Skolnick. 1996. Lattice Models of Protein Folding. Dynamics and Thermodynamics. Chapman & Hall, Boca Raton, FL.

24. Verdugo, P., I. Deyrup-Olsen, …, D. Johnson. 1987. Molecular mechanism of mucin secretion: I. The role of intragranular charge shielding. *J. Dent. Res.* 66:506–508.

25. Paz, H. B., A. S. Tisdale, …, I. K. Gipson. 2003. The role of calcium in mucin packaging within goblet cells. *Exp. Eye Res.* 77:69–75.

26. Ferguson, D. M., J. I. Siepmann, and D. G. Truhlar. 1999. Monte Carlo Methods in Chemical Physics. John Wiley & Sons, New York.

27. Theodorou, D. N., and U. W. Suter. 1985. Shape of unperturbed linear polymers: polypropylene. *Macromolecules.* 18:1206–1214.

28. Allen, M. P., and D. J. Tildesley. 1987. Computer Simulations of Liquids. Oxford University Press, New York.

29. Lai, S. K., D. E. O'Hanlon, …, J. Hanes. 2007. Rapid transport of large polymeric nanoparticles in fresh undiluted human mucus. *Proc. Natl. Acad. Sci. USA.* 104:1482–1487.

30. Saltzman, W. M., M. L. Radomsky, …, R. A. Cone. 1994. Antibody diffusion in human cervical mucus. *Biophys. J.* 66:508–515.