PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Molecular modeling of hair keratin/peptide complex: Using MM-PBSA calculations to describe experimental binding results

Nuno G. Azoia,[1] Margarida M. Fernandes,[1] Nuno M. Micaêlo,[2] Cláudio M. Soares,[3] and Artur Cavaco-Paulo[1]*

[1] Departamento de Engenharia Têxtil, Universidade do Minho, Campus de Azurém, 4800-058 Guimarães, Portugal
[2] Chemistry Center, Minho University, Campus Gualtar, 4710-057 Braga, Portugal
[3] Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Portugal

ABSTRACT

Molecular dynamics simulations of a keratin/peptide complex have been conducted to predict the binding affinity of four different peptides toward human hair. Free energy calculations on the peptides' interaction with the keratin model demonstrated that electrostatic interactions are believed to be the main driving force stabilizing the complex. The molecular mechanics–Poisson-Boltzmann surface area methodology used for the free energy calculations demonstrated that the dielectric constant in the protein's interior plays a major role in the free energy calculations, and the only way to obtain accordance between the free energy calculations and the experimental binding results was to use the average dielectric constant.

Proteins 2012; 80:1409–1417.

Key words: dielectric constant; free energy; free energy of solvation; keratin dimer; hair fiber.

INTRODUCTION

The interaction of small peptides with protein matrices can be an important factor in the development of new added value products with applications in medicine and cosmetics. The use of peptides in cosmetic formulations to regenerate skin and hair keratin properties is well established. In this work, we report the application of four different peptides, based on the amino-acid sequence of pulmonary-associated surfactant protein D (SP-D) from mammalian lungs, to evaluate their binding affinity toward hair.

SP-D is a collagenous, carbohydrate-binding glycoprotein that is synthesized and secreted into the air space of the lung by alveolar type II cells and also by nonciliated bronchiolar epithelial cells.[1] The α-helical neck region of this protein has been reported to be essential for the binding of the protein carbohydrate recognition domain to lipids[2] and to hair fibers and moreover for recovering its mechanical properties.[3] Electrostatic interactions are believed to play a significant role in hair adsorption mechanisms. However, the presence of a hydrophobic lipid layer hinders the diffusion of compounds into hair.[4] The ability of SP-D to interact with these lipids is thought to assist their penetration onto hair fibers. For that reason, smaller fragments of SP-D protein were applied on normal and bleached hair, because the bleaching process is known to destroy the hair lipid layer. A formulation that joins the ability to penetrate deep in the fiber (small size) with the ability to interact with lipids present in hair was sought.

To obtain information about the binding affinity of these peptides toward hair, molecular dynamics (MD) simulations were conducted. MD simulations on human keratins, both from hair and epithelium cells, have been extensively performed to study the structure of this complex fibrous protein. There is, however, a drawback in the study of keratin fibers by computational methods.[5–9] There are no crystallographic structures of keratins, so it is necessary to build a suitable model to study keratin interaction. There are some different approaches for the construction of these models. Akkermans and Warren[5] did some structural studies using the leucine zipper domain in the GCN4 protein because this zipper domain has a reasonable sequence homology with hair kera-

Grant sponsor: Contract Research Program "Compromisso com a Ciência"; Grant number: C2008-UMINHO-CQ-03; Grant sponsor: FCT "Fundação para a Ciência e Tecnologia" (PhD); Grant number: SFRH/BD/38363/2007
*Correspondence to: Artur Cavaco-Paulo, Universidade do Minho, Departamento de Engenharia Têxtil, Campus de Azurém, 4800 058 Guimarães, Portugal. E-mail: artur@det.uminho.pt.
Received 5 November 2011; Revised 23 December 2011; Accepted 29 December 2011
Published online 18 January 2012 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/prot.24037

tins. This sequence homology was used by others[6–8] to build keratin models by homology alignment with the structure of GCN4 leucine zipper. Buehler and coworkers[10–12] had reported an extensive work on the characterization of vimentin intermediate filaments, and Smith and Parry[9] reported some work using hair keratin structures build by homology alignment with the crystal structures of vimentin.

Herein, we report, for the first time, the MD simulations of the interaction between human hair keratin and small peptides. The keratin model was constructed by homology modeling with the crystal structure of vimentin, and corresponds to the 1A segment of the dimer formed by keratin 32 and keratin 84. Vimentin and keratin belongs to the same family of proteins, the intermediate filaments. It is known that all intermediate filaments have coiled-coil structures, very similar to each other, and they all have similar amino acids sequences. Furthermore, there were recent developments in the X-ray resolution of the structure of vimentin toward its full knowledge,[13] making it a good choice as a model in the homology alignments studies.

MATERIALS AND METHODS

Hair

Natural European blond human hair samples were provided by International Hair Importers & Products (New York). Before use, the hair fibers tresses (0.1 g each) were washed with a commercial shampoo. These hair tresses were used with two different pretreatments. Virgin hair corresponds to hair without any chemical pretreatment. Chemically damaged hair (8×B) corresponds to hair subjected to eight cycles of bleaching (8×B). In each bleaching cycle, the hair tresses was treated with 10% H_2O_2 (v/v) in the presence of 0.1*M* Na_2CO_3/$NaHCO_3$ pH 9.0 buffer at 50°C for 1 h.

Engineered peptides

The peptides used in this study were synthesized by JPT Peptide Technologies GmbH (Berlin, Germany) and were covalently linked by their N-terminal to a fluorescent dye, (5(6)-carboxytetramethyl-rhodamine, succinimidyl ester), that is, 5(6)-TAMRA, with spectral properties of λ_{ex} = 544 nm and λ_{em} = 572 nm, to facilitate the analysis of peptide penetration.

The peptides were dissolved in 0.05*M* Phosphate buffer solution, pH 7.5, containing 10% ethanol, 1.5% propylene glycol, and 0.5% benzyl alcohol, with a final concentration of 0.6 m*M* in terms of peptide content. The peptide characteristics are shown in Figure 1. The sequence of the peptides is described by one letter code which represents one of known amino acid residues.

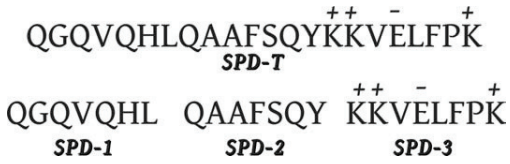

Figure 1

Peptides' nomenclature and amino acid sequence, showing side-chain charges at pH = 7. All peptides are attached to a fluorescent dye TAMRA at N terminus and to NH2 at C terminus.

Hair treatments with peptides

As mentioned above, the peptides were dissolved in an aqueous solution containing a mixture of alcohols. This mixture of solvents was used not only to promote peptide solubility but also because it is known that this solvent is able to ease the interaction of peptides with hair fibers by promoting hair swelling[14] and exposing the hair shaft to the treatment solution. The hair samples were treated with 3.0 mL of 0.05*M* phosphate buffer solution, pH 7.5, containing 600 μL of peptide solution (final concentration 120 μ*M*), for 1 h at 37°C and 100 rpm. Subsequently, all the samples were thoroughly rinsed with tap water and washed with a commercial shampoo.

Color variation

The color variation was measured directly from the hair-peptide solution by monitoring absorbance at 555 nm (corresponding to maximum absorption of the treatment solution) before and after the treatment. As the color of the solution is due solely to the presence of unbound peptide, the quantity of bound peptide that has penetrated the hair is related to the difference in absorbance (Abs). The percentage of color variation was calculated using Eq. (1).

$$\text{Color Variation}(\%) = \frac{\text{Abs}_{\text{initial}} - \text{Abs}_{\text{final}}}{\text{Abs}_{\text{initial}}} \times 100 \qquad (1)$$

Fluorescence microscopy

Transversal cuts of the hair fiber samples were analyzed by fluorescence microscopy. Hair fibers were embedded into an epoxy resin, and 15-μm transversal cuts of the fibers were prepared using a microtome (Microtome Leitz). Fiber cross sections were analyzed using a LEICA DM 5000B fluorescence microscope (Leica), at a magnification of ×40. All fluorescence microscopy images were recorded using identical filter, exposure, brightness, and gain settings. The most representative images were chosen.

Peptide setup

The peptide used was chemically synthesized, and it represents a part of the human SP-D protein neck domain. The lung surfactant proteins are characterized by a trimeric structure, and the neck domain is necessary for the assembly of the trimer. In contrast to the other surfactant proteins that have three heptad repeats in the neck domain, SP-D have four and the extra heptad motif makes this neck domain sufficient for the trimerization.[15,16] The SPD-T peptide used corresponds to only two heptads, which is not sufficient for assure the trimerization of the peptides in solution, making it very unlikely for the peptide to present this type of structure. Even with the assumption that this peptide will not present the coiled-coil structure, its conformation was unknown.

Starting from an X-ray structure of SP-D (PDB code 1pw9[17]), the segments corresponding to SPD-T peptide (residues 215-235) were selected, and one trimeric structure of SPD-T was obtained. Using the same conditions reported below to the simulations of the keratin/SPD complexes, it was observed that this trimeric structure is not stable in water, in pure ethanol, or in the solvent mixture used in the hair treatments. The next step was the evaluation of the structures of SPD-T and SPD 1-3, starting from the corresponding monomeric structures extracted from 1pw9, using the solvent mixture used for the keratin/SPD complexes. This procedure was used not only to evaluate the behavior of the structures over time but also to obtain fully relaxed structures of the four peptides to be used as starting points in the studies of complex formation. As expected, peptides SPD 1-3 do not have any particular structure after 100 ns of simulation time, and the same was observed for SPD-T.

Keratin setup

The keratin fibers are not homogeneous in their composition, so any representation of keratin must be considered as a model. Keratin is known to form α-helices, which aggregate to form filaments. The basic structural unit of these filaments is the keratin heterodimer, consisting in one Type I keratin and the corresponding Type II keratin, forming the α-helical rod domain. Each of these heterodimers are divided into four major segments (1A, 1B, 2A, and 2B).[6]

There are no X-ray structures available for the keratin structure at the atomic level. The only option to simulate keratin is to derive its 3D structures by homology modeling. This is a valid approach because keratin belongs to a very well-known family of proteins, the intermediate filaments, all having a similar structure. The homology modeling studies were made using the SWISS-MODEL tool.[18–20] It was possible to obtain homology alignments for all the sequences from hair keratin, both from Type I and Type II. The structures chosen for this work were the ones having the best alignment against 1A segment, namely, keratin-32 (UniProtKB accession Q14532) and keratin-84 (UniProtKB accession Q9NSB2). The results from the SWISS-MODEL showed that the alignment for both structures was performed with vimentin (PDB code 3g1eA[21]), with 44% of sequence identity for keratin 32 and 54% for keratin-84. This structure (3g1eA) corresponds to chain A of fragment 1A of vimentin.

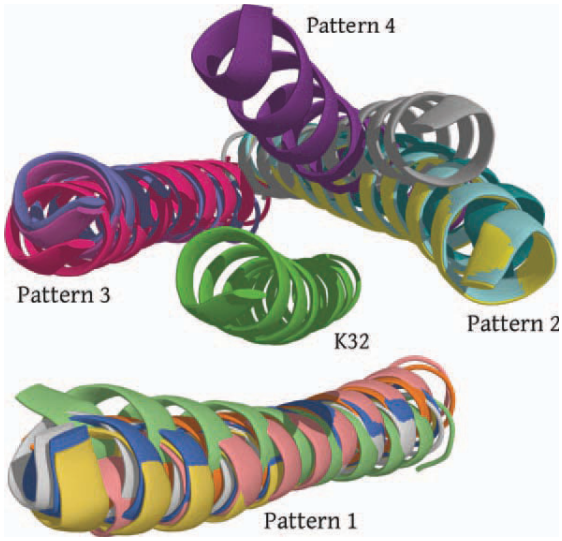

Figure 2

Results from the docking calculations, showing K32 as the central structure and the results from different calculations. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The dimeric structure of keratin was obtained by rigid docking approach using AutoDockVina[22] from the structures for both keratin-32 (K32) and keratin-84 (K84). K32 was used as the receptor, considered as if it is fully rigid. The ligand was K84 with rigid backbone only. Each of these proteins has a volume of roughly 56 × 15 × 15 Å. For the docking calculations, a box of size 60 × 45 × 45 Å was used, centered at the geometric center of K32 structure. In this way, the K84 structure could explore all the positional space surrounding K32. Three independent calculations were performed. Two of them used the standard configurations of AutodockVina and in the other one, the exhaustiveness value was changed to 80 (standard is 8), and the program was allowed to generate 27 binding modes (standard is 9). Keratin dimers present in hair are parallel (N-terminus and C-terminus align with each other), so any docking result showing antiparallel association were discarded, resulting in four possible interaction patterns (see Fig. 2). The orientation that represents the largest number of hits (Pattern 1) was chosen, and from that orientation, the structure that is energetically more favorable.

The choice of the dimer to use as model could be critical, but it was not possible to find any information about the presence of the dimers in human hair. What is known is that all combinations of keratins Type I and II are possible and are present in human keratin. K32 and K84 were chosen because, as mentioned before, these are the monomers that gave the best homology alignment.

Simulations

The simulations were performed with the GROMACS 4.0.7[23] package using the GROMOS 53A6 force field.[24] The keratin dimer obtained from the docking calculations was used without further energy minimization. The peptides were added to the simulation box in a random way. Before solvation benzyl alcohol and propylene glycol molecules were added to the simulation box, in a random way, such that the final concentration of these alcohols match those used in the experimental essays. The system was solvated using a pre-equilibrated box of water containing 10% of ethanol. The system size was chosen according to the minimum image convention taking into account a cutoff of 1.4 nm. The bonds lengths of the proteins were constrained with LINCS[25] and those for water with SETTLE.[26] Nonbonded interactions were calculated using a twin-range method, with short- and long-range cutoffs of 0.8 and 1.4 nm, respectively. Neighbor searching was carried out up to 1.4 nm and updated every five steps. A time step of integration of 2 fs was used. A reaction field correction for the electrostatic interactions was applied using a dielectric constant of 54.[27] The single point charge model[28] was used for water molecules. The initial systems were energy minimized for 2000 steps using the steepest descent method, with all heavy atoms harmonically restrained using a force constant of 10^3 kJ/mol nm^2. The systems were initialized in the canonical ensemble (NVT) for 50 ns, with all heavy atoms harmonically restrained using a force constant of 10^3 kJ/mol nm^2. The simulation was then continued for 50 ns in the isothermal–isobaric ensemble (NPT), with the heavy atoms harmonically restrained with the same force constant. Finally, for allowing the equilibration of the system properties, the simulations were further extended in the NPT ensemble with positional restrains applied to the Cα atoms. Pressure control was implemented using the Berendsen barostat,[29] with a reference pressure of 1 bar, 0.5 ps of relaxation time, and isothermal compressibility of 4.5×10^{-5} bar^{-1}. Temperature control was set using the V-rescale[29,30] thermostat at 300 K. The keratin and the peptide were grouped in the same heat bath and solvent molecules in a separated heat bath, with temperatures coupling constants of 0.025 ps in the first two initialization steps and with 0.1 ps for the rest of the simulations. Five replica simulations of 75 ns in length were carried out using different initial velocities taken from a Maxwell–Boltzman distribution at 300 K, leading to a total simulation time of 375 ns for each system. For the evaluation of the contact surface area and calculations of the entropic contribution for the free energy of binding, one simulation for each system was extended to a total of 100 ns. For each system, only the replica having the highest contact surface area value was chosen.

Free energy calculations

For the free energy calculations, only end-point methods were considered, mainly because the convergence of pathway methods can be difficult. Concerning only end-point approaches, two different methods were considered: the linear interaction energy method and the molecular mechanics–Poisson-Boltzmann surface area (MM-PBSA) method. A full description of these different methods was described in detail elsewhere.[31] MM-PBSA methods uses MD simulations of the free ligand, free protein, and their complex as a basis for calculating the average potential and solvation energies in Eq. (2),

$$G = G_{\mathrm{MM}} + G_{\mathrm{solv}} + G_{\mathrm{np}} - TS \qquad (2)$$

and the free energy on binding is calculated as represented in Eq. (3).

$$\Delta G_{\mathrm{binding}} = G_{\mathrm{complex}} - (G_{\mathrm{protein}} + G_{\mathrm{ligand}}) \qquad (3)$$

Taking into account the previous equations, the free energy of binding was calculated using Eq. (4)

$$\Delta G_{\mathrm{binding}} = \Delta E_{\mathrm{MM}} + \Delta G_{\mathrm{solv}} + \Delta G_{\mathrm{np}} - T\Delta S \qquad (4)$$

in which each term is defined as the difference between the value for the complex and the sum of the values for both protein and ligand.

The "MM" subscript corresponds to the internal energy of the protein or ligand (bond, angle, dihedral, and the electrostatic and van der Waals interactions), the "solv" subscript refers to the free energy of polar solvation, and the "np" subscript the free energy of nonpolar solvation. S is the entropic contribution for the free energy and T is the temperature.

Although early applications of MM-PBSA appeared promising, apparently it is difficult to obtain convergence for the energy averages.[31] Convergence has been accelerated by a single-trajectory approach,[32,33] in which only one MD simulation of the protein-ligand complex is carried out. Conformations of the nominally free ligand are then derived simply by deleting the protein from the resulting snapshots, and likewise for the "free" protein. Lee and Olson[33] had reported a comparative analysis of different methods to calculate protein–ligand binding affinity, and for their particular system, they found that

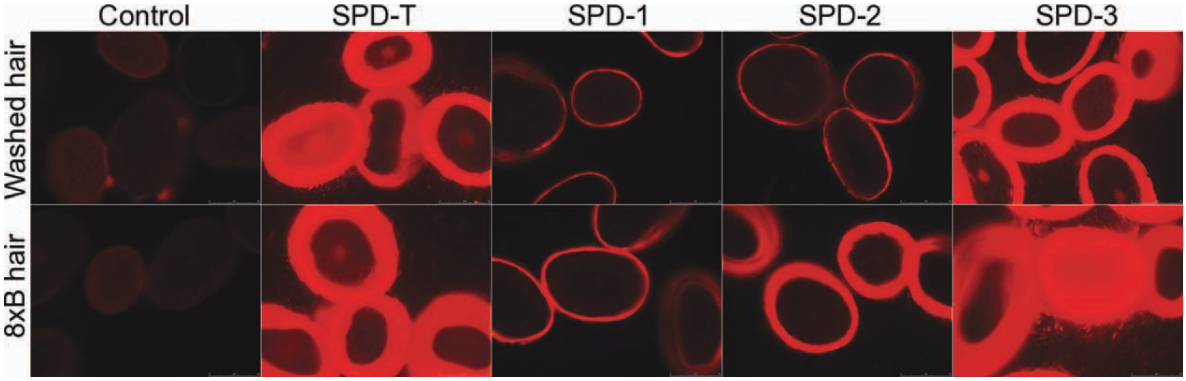

Figure 3

Hair cross sections of fibers without treatment (control) and treated with SPD peptides in washed hair and eight times bleached hair. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the least reliable results were those obtained by methods that simulate both bounded and unbounded states.

Taking all of these into account the free energies of binding was determined with the MM-PBSA approach, using the MD simulation of SPD-Keratin complexes only. Only the last 50 ns of the simulation were used, and 101 frames were extracted from the respective trajectory. ΔE_{MM} and ΔG_{solv} were calculated using these 101 frames.

The electrostatic potential and solvation free energy (ΔG_{solv}) calculations were carried out using the Poisson-Boltzmann method as implemented in the MEAD package version 2.2.8.[34] First, the dielectric constant within the molecules was calculated from the total dipole and its fluctuations, using independent 100-ns MD simulations of each component. This was made in an effort to better relate the calculation of the free energies with the experimental values. The results of the dielectric constant were 2 for the keratin, the complexes and for peptides SPD-1 and SPD-2. For peptides SPD-T and SPD-3, the values of 3 and 7 were obtained, respectively. These calculated values were used for determine the solvation-free energies. The dielectric constant for the solvent was considered as 80.

The nonpolar solvation energy was determined as described elsewhere[35] from the solvent accessible surface area (SASA), calculated with a probe radius of 1.4 Å, and γ constant of 5 (cal/mol)/Å^2 [Eq. (5)].

$$\Delta G_{np} = \gamma \text{ SASA} \quad (5)$$

The entropy was computed based on the quasiharmonic approach, using Schlitter's formula. Schlitter[36] introduced a very elegant formula to calculate absolute entropies from MD trajectories using the covariance matrix of atom-positional fluctuation. Before Schlitter's work, Levy et al.[37] had introduced a method based on the quasiharmonic approximation which also connects the absolute entropy of the atoms to the covariance matrix, using internal coordinates. Latter, Andricioaei and Karplus[38] showed that Levy's method can be extended to estimate the total entropy using the covariance matrix with Cartesian coordinates. Carlsson and Åqvist[39] had proved that Schlitter's formula, and the quasiharmonic analysis do not differ much in the entropy values. In this work, a methodology reported elsewhere[38–40] was used to calculate the entropic contribution, using Schlitter's formula, but assuming that the fluctuations in the motions of the system can be described by a Gaussian probability distribution (quasiharmonic approximation). This contribution was determined using the extended simulations (100 ns), with the entropy calculated for the last 90 ns.

RESULTS

Penetration of compounds into the hair is very important when new approaches for hair care applications are considered. However, the presence of the covalent isopeptide crosslinks as well as the covalently attached lipids (predominantly 18-methyleicosanoic acid) on the hair surface constitute a diffusion barrier to chemicals and other treatment agents, thus impairing penetration.[41,42] Being linked to the proteinaceous epicuticle via thioester linkage,[4] this fatty acid forms a layer that can be removed by alkaline oxidative process, such as those used in the bleaching process for the hair samples. Because of the damage imparted by the bleaching process, the alpha-keratin present in the cortex of hair is thought to be exposed. It is well reported that the bleaching process modifies hair properties along the cuticle.[43–45]

The peptides used in this work were developed based on the pulmonary-associated surfactant protein D (SP-D) from mammals' lungs amino acid residue sequence[46,47] (Fig. 1). A longer peptide (SPD-T) and shorter fragments of it (SPD-1, SPD-2, and SPD-3) were applied on human hair to study their binding affinity and ability to restore

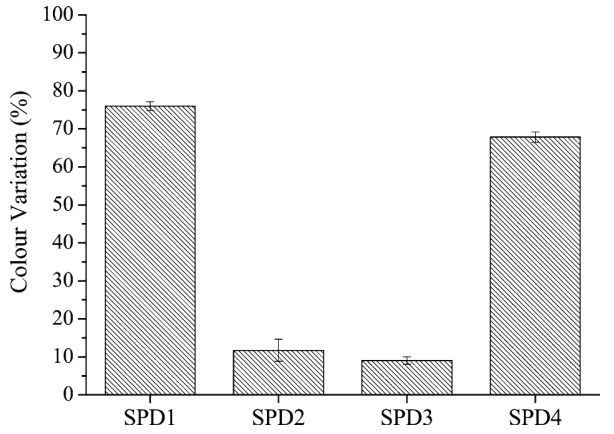

Figure 4

Uptake of SPD peptides on chemically damaged hair, measured by color variation method in bath treatment solution at 555 nm.

damaged parts. This assumption came from the fact that surfactant proteins from mammals' lungs are able to interact with lipids. Fragments or models representing those proteins could, therefore, overcome the lipidic barrier present on hair, increasing the penetration of peptides inside the hair.[3] This fact was indeed proved by the visualization of peptide at cuticle level in the washed hair, which possesses an intact lipid barrier (Fig. 3).

These peptides have in their structure a covalently attached fluorescent dye that facilitates the analysis of its incorporation. Fluorescence micrographs of hair fibers treated with SP-D-based peptide, at the same conditions of brightness, exposure time, gain, and intensity, show the affinity of peptides toward hair (Fig. 3). Chemically damaged hair was shown to possess higher quantity of peptide at its surface. The damage imparted by the chemical bleaching process may constitute a pathway for the penetration of small peptides. In fact, previous studies indicate that the adsorption onto hair is increased both when low molecular weight compounds are applied and when hair is chemically damaged.[48]

SPD-T and SPD-3 attach at higher yield than SPD-1 and SPD-2, in both virgin and damaged hair (Fig. 3). The difference in chemistry should be the reason for this behavior. SPD-3 is a small peptide, which amino acid residue sequence may be responsible for the high affinity of SPD-T. The chemistry of SPD-3 amino acid sequence (Fig. 1) may facilitate the adsorption on human hair, due to the presence of the positively charged lysines, that links to the negatively charged hair (pI = 3.7).[49] These observations were confirmed by the color variation measurements, a method that determines the initial affinity of peptide toward hair while it is being applied. SPD-T and SPD-3 shows high affinity toward hair, while SPD-1 and SPD-2 poorly attach to it (Fig. 4).

Molecular modeling of the keratin-SPD complexes

To understand the molecular mechanism of the interaction between SPD's and the hair fibers, molecular modeling studies were conducted using several methodologies. The keratin structure was modeled using rigid docking methods.[22] The keratin–SPD structural complexes were obtained using molecular dynamics (MD) simulations,[23,24] with the peptide structures placed randomly near the keratin structure, to provide greater physical realism of the interaction between them. Furthermore, the contact surface and binding free energy of each complex was estimated using the MM-PBSA method.

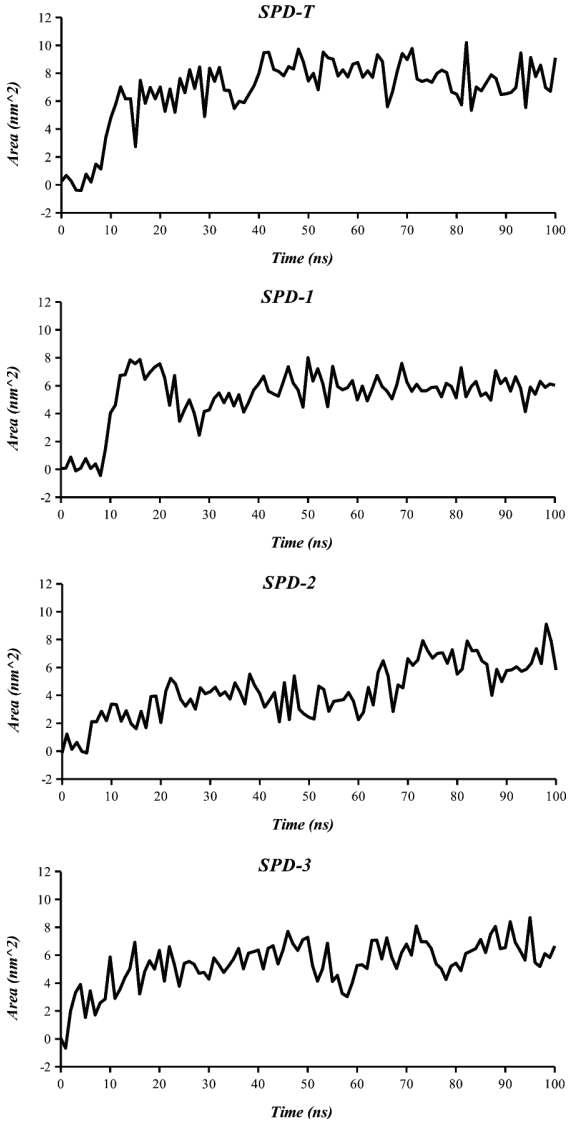

Figure 5

Contact surface area for peptide–keratin complexes. The values are taken from the chosen replica for each peptide.

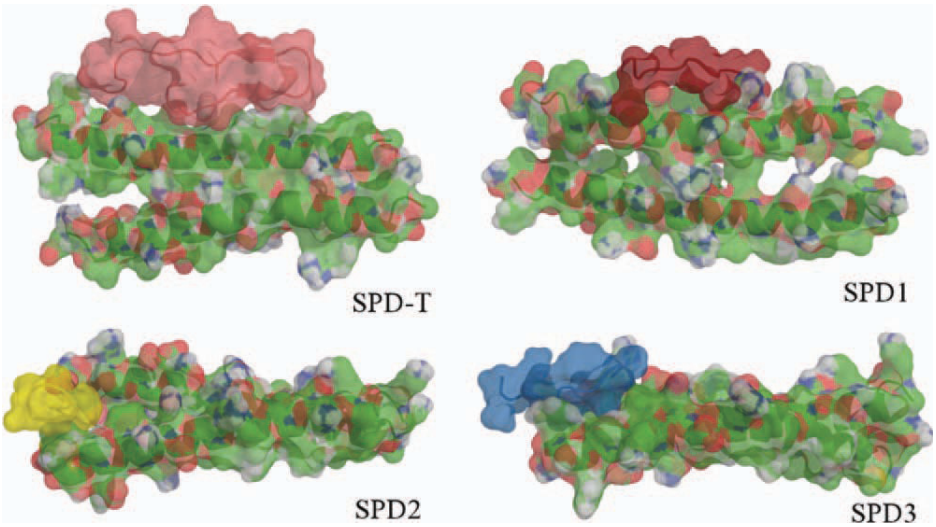

Figure 6
Representation of the peptide–keratin complex after 75 ns.

Analysis of the simulations

Contact surface area

The objective of the simulations was the evaluation of the interaction between the peptides and the keratin structure, so one of the first analysis to be perform was the surface contact area.[50] In this way, it is possible to evaluate, not only the interaction between the molecules but also the stability and convergence of the simulations.

Figure 5 shows the contact surface between keratin dimer (K32 and K84) and the four SPD peptides. For the analysis of the contact surface area, the simulations were extended for an extra 25 ns, to check the stability and convergence of the complex. For each system, it is clear that there are no significant difference between the values at 75 ns and the same values at 100 ns. It is also possible to observe that the behavior of each system is very similar. SPD-T peptide is larger than any other, but the values obtained for the contact surface area are close to the values obtained to the other peptides. The differences in these values can be easily explained looking to the structures representing the configuration of the complexes after 75 ns, shown in Figure 6. In contrast to peptides SDP1-3, SDP-T is not fully extended. The configuration of SPD-T minimizes the surface of the peptide, having the same effect on the surface contact area between SPD-T and the keratin.

Free energy calculations

The results from these calculations are shown in Table I. The dielectric value used for calculating the free energy of polar solvation is a key factor for the accuracy of the MM/PBSA methodology. There are many published works in this field,[51–55] but there is no universal approach about the value to be used. In free energy calculations, the values used for the dielectric constant (ϵ_p) are usually 1, 2, or 4. If the objective is to determine pK_a values for the side-chain residues of proteins, then higher values are used, usually between 10 and 20.[56,57] Considering these approaches, the solvation free energy (ΔG_{solv}) was first determined using ϵ_p equal to 2, 4, and 10. These three values were used to check the differences on the

Table I
Free Energy of Binding

	ΔE_{MM}	ΔG_{solv}	ΔG_{np}	$T\Delta S$	$\Delta G_{binding}$
ϵ_p = 3 for SPD-T, ϵ_p = 7 for SPD-3, and ϵ_p = 2 for all the other species					
SPD-T	−654	236	−17	18	−453
SPD-1	−349	508	−12	31	116
SPD-2	−280	342	−9	46	7
SPD-3	−603	420	−12	36	−231
ϵ_p = 2 for all the species					
SPD-T	−389	444	−17	18	20
SPD-1	−349	508	−12	31	116
SPD-2	−280	342	−9	46	7
SPD-3	−493	798	−12	36	257
ϵ_p = 4 for all the species					
SPD-T	−274	367	−17	18	58
SPD-1	−225	439	−12	31	171
SPD-2	−174	282	−9	46	53
SPD-3	−307	747	−12	36	392
ϵ_p = 10 for all the species					
SPD-T	−204	268	−17	18	29
SPD-1	−151	327	−12	31	133
SPD-2	−110	196	−9	46	31
SPD-3	−195	629	−12	36	386

All values are indicated in kJ/mol.

final results. The results of these calculations are shown in Table I. What became clear is that this approach is not suitable to describe these keratin–peptide systems. Positive values for the binding free energy do not describe the results from MD simulations or the experimental results.

To better describe these systems, the average dielectric constant ($\bar{\epsilon}$) for the molecules, calculated from the total dipole and its fluctuations, was used. In other words, $\bar{\epsilon}$ was used for calculating the polar contribution of the free energy of solvation, instead of ϵ_p. This is not a very common approach. In fact it is normally stated that ϵ_p is completely different from $\bar{\epsilon}$,[58] but what was observed in this work was that $\bar{\epsilon}$ gave a better description of the interaction of keratin with these small peptides, two of them charged. The effect of these charged residues in a small peptide will be much more pronounced than in large globular proteins. Lund[59] has discussed and used a protein model, for the calculation of titration curves, in which the protein is treated as having to dielectrics: one low dielectric core and a high dielectric outer shell where the charged residues are located. In this outer shell, the dielectric was considered to be as high as the solvent dielectric. In this case, it does not seem to be necessary to use such dielectric values, because what is required are free energies and not titration curves, but it still make sense to use different dielectrics to describe a small charged peptide and a bigger protein. The results obtained show that for proteins, more specifically for the keratin dimer and the keratin–peptide complexes, and for small uncharged peptides, $\bar{\epsilon}$ is similar to ϵ_p. For charged peptides, the value is higher, describing the fact that the solvent is able to solvate the charged residues, thus increasing the dielectric value in the protein's interior.

In Table I, it is possible to observe the importance of each component used in the free energy calculations. The interactions come essentially from the mechanical energy component. The polar solvation energy component is always unfavorable, and the nonpolar and entropic contributions are negligible if compared with ΔG_{solv}.

The positive change in the entropy was not fully expected. The convergence of these values was evaluated calculating the entropy using different time intervals, ranging from only the last 10 ns to the full 100-ns trajectory. A good convergence was found.

CONCLUSIONS

SPD-T and SPD-3 have shown more affinity toward hair that SPD-1 and SPD-2. These differences could result from the fact that both SPD-T and SPD-3 possess a positive net charge, while SPD-1 and SPD-2 do not have neither net nor local charges. SPD-1, SPD-2, and SPD-3 represents each one a different segment of SPD-T, and all together represent the totality of SPD-T sequence. It is possible to conclude that the segment responsible for the affinity toward hair keratin, it is the segment represented by SPD-3. This conclusion is not strange, if the charge distribution in each segment is considered.

The charged distribution in these relatively small peptides is also important when considering the dielectric constant in the proteins' interior: the choice of the best value for ϵ_p was a key factor to obtain a good description of our system. It was not possible to get good results while using the standard values for ϵ_p, but the use of $\bar{\epsilon}$ in the calculation of the solvation free energy lead to a good agreement between free energy values and the experimental results.

ACKNOWLEDGMENTS

The authors acknowledge the access to the Minho University GRIUM cluster.

REFERENCES

1. Persson A, Chang D, Rust K, Moxley M, Longmore W, Crouch E. Purification and biochemical characterization of CP4 (SP-D), a collagenous surfactant-associated protein. Biochemistry 1989;28:6361–6367.
2. Kishore U, Wang JY, Hoppe HJ, Reid KB. The alpha-helical neck region of human lung surfactant protein D is essential for the binding of the carbohydrate recognition domains to lipopolysaccharides and phospholipids. Biochem J 1996;318:505–511.
3. Cavaco–Paulo A, Silva CJSM. Formulation containing neck domains and/or carbohydrate recognition domains for cosmetic applications, namely for the treatment of keratin fibres like hair.WO2007136286, 2007.
4. Breakspear S, Smith JR, Luengo G. Effect of the covalently linked fatty acid 18-MEA on the nanotribology of hair's outermost surface. J Struct Biol 2005;149:235–242.
5. Akkermans RLC, Warren PB. Multiscale modelling of human hair. Philos Trans R Soc A 2004;362:1783–1793.
6. Knopp B, Jung B, Wortmann F-J. Modeling of the transition temperature for the helical denaturation of α-keratin intermediate filaments. Macromol Theory Simul 1997;6:1–12.
7. Liovic M, Stojan J, Bowden PE, Gibbs D, Vahlquist A, Lane EB, Komel R. A novel keratin 5 mutation (K5V186L) in a family with EBS-K: a conservative substitution can lead to development of different disease phenotypes. J Invest Dermatol 2001;116:964–969.
8. Danciulescu C, Nick B, Wortmann F-J. Structural stability of wild type and mutated alpha-keratin fragments: molecular dynamics and free energy calculations. Biomacromolecules 2004;5:2165–2175.
9. Smith TA, Parry DAD. Three-dimensional modelling of interchain sequence similarities and differences in the coiled-coil segments of keratin intermediate filament heterodimers highlight features important in assembly. J Struct Biol 2008;162:139–151.
10. Arslan M, Qin Z, Buehler MJ. Coiled-coil intermediate filament stutter instability and molecular unfolding. Comput Methods Biomech Biomed Eng 2011;14:483–489.
11. Qin Z, Buehler MJ. Structure and dynamics of human vimentin intermediate filament dimer and tetramer in explicit and implicit solvent models. J Mol Model 2011;17:37–48.
12. Qin Z, Kreplak L, Buehler MJ. Hierarchical structure controls nanomechanical properties of vimentin intermediate filaments. PloS one 2009;4:e7294.
13. Nicolet S, Herrmann H, Aebi U, Strelkov SV. Atomic structure of vimentin coil 2. J Struct Biol 2010;170:369–376.

14. Breakspear S, Ito T, Koike K, Cavaco-Paulo A. Hair Cosmetic Composition.WO2009041739, 2009.
15. Parry DAD, Fraser RDB, Squire JM. Fifty years of coiled-coils and α-helical bundles: a close relationship between sequence and structure. J Struct Biol 2008;163:258–269.
16. Zhang P, McAlinden A, Li S, Schumacher T, Wang H, Hu S, Sandell L, Crouch E. The amino-terminal heptad repeats of the coiled-coil neck domain of pulmonary surfactant protein d are necessary for the assembly of trimeric subunits and dodecamers. J Biol Chem 2001;276:19862–19870.
17. Shrive AK, Tharia HA, Strong P, Kishore U, Burns I, Rizkallah PJ, Reid KBM, Greenhough TJ. High-resolution structural insights into ligand binding and immune cell recognition by human lung surfactant protein D. J Mol Biol 2003;331:509–523.
18. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 2006;22:195–201.
19. Peitsch MC. Protein modeling by e-mail. Bio/Technol 1995;13:658–660.
20. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T. The SWISS-MODEL repository and associated resources. Nucleic Acids Res 2009;37:D387–D392.
21. Meier M, Padilla GP, Herrmann H, Wedig T, Hergt M, Patel TR, Stetefeld J, Aebi U, Burkhard P. Vimentin coil 1A—a molecular switch involved in the initiation of filament elongation. J Mol Biol 2009;390:245–261.
22. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 2010;31:455–461.
23. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theor Comput 2008;4:435–447.
24. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF. A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. J Comput Chem 2004;25:1656–1676.
25. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997;18:1463–1472.
26. Miyamoto S, Kollman PA. Settle: an analytical version of the SHAKE and RATTLE algorithm for rigid water models. J Comput Chem 1992;13:952–962.
27. Smith PE, van Gunsteren WF. Consistent dielectric properties of the simple point charge and extended simple point charge water models at 277 and 300 K. J Chem Phys 1994;100:3169.
28. Hermans J, Berendsen HJC, Van Gunsteren WF, Postma JPM. A consistent empirical potential for water-protein interactions. Biopolymers 1984;23:1513–1518.
29. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
30. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. J Chem Phys 2007;126:014101.
31. Gilson MK, Zhou H-X. Calculation of protein-ligand binding affinities. Annu Rev Biophys Biomol Struct 2007;36:21–42.
32. Brown SP, Muchmore SW. High-throughput calculation of protein-ligand binding affinities: modification and adaptation of the MM-PBSA protocol to enterprise grid computing. J Chem Inf Model 2006;46:999–1005.
33. Lee MS, Olson MA. Calculation of absolute protein-ligand binding affinity using path and endpoint approaches. Biophys J 2006;90:864–877.
34. Bashford D. An object-oriented programming suite for electrostatic effects in biological molecules. An experience report on the MEAD project. In: Lecture notes in computer science. Ishikawa Y, Oldehoeft R, Reynders J, Tholburn M, Eds., Springer Berlin/Heidelberg; 1997, Vol.1343, pp 233–240.
35. Elcock AH, Sept D, McCammon JA. Computer simulation of protein-protein interactions. J Phys Chem B 2001;105:1504–1518.
36. Schlitter J. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. Chem Phys Lett 1993;215:617–621.
37. Levy RM, Karplus M, Kushick J, Perahia D. Evaluation of the configurational entropy for proteins: application to molecular dynamics simulations of an α-helix. Macromolecules 1984;17:1370–1374.
38. Andricioaei I, Karplus M. On the calculation of entropy from covariance matrices of the atomic fluctuations. J Chem Phys 2001;115:6289–6292.
39. Carlsson J, Åqvist J. Absolute and relative entropies from computer simulation with applications to ligand binding. J Phys Chem B 2005;109:6448–6456.
40. Schäfer H, Mark AE, van Gunsteren WF. Absolute entropies from molecular dynamics simulation trajectories. J Chem Phys 2000;113:7809–7817.
41. Heine E, Höcker H. Enzyme treatments for wool and cotton. Rev Prog Color Relat Top 1995;25:57–70.
42. Nolte H, Bishop DP. Effects of proteolytic and lipolytic enzymes on untreated and shrink-resist-treated wool. J Text Inst 1996;87:212–226.
43. Gray J. Hair care and hair care products. Clin Dermatol 2001;19:227–236.
44. Bolduc C, Shapiro J. Hair care products: waving, straightening, conditioning and coloring. Clin Dermatol 2001;19:431–436.
45. Dawber R. Hair: its structure and response to cosmetic preparations. Clin Dermatol 1996;14:105–112.
46. Moore KJ, D'Amore-Bruno MA, Korfhagen TR, Glasser SW, Whitsett JA, Jenkins NA, Copeland NG. Chromosomal localization of three pulmonary surfactant protein genes in the mouse. Genomics 1992;12:388–393.
47. Pilot-Matias TJ, Kister SE, Fox JL, Kropp K, Glasser SW, Whitsett JA. Structure and organization of the gene encoding human pulmonary surfactant proteolipid SP-B. DNA 1989;8:75–86.
48. Faucher JA, Goddard ED. Sorption of a cationic polymer by stratum corneum. J Soc Cosmet Chem 1976;27:543–553.
49. Regismond STA, Heng Y-M, Goddard ED, Winnik FM. Fluorescence microscopy observation of the adsorption onto hair of a fluorescently labeled cationic cellulose ether. Langmuir 1999;15:3007–3010.
50. Conte LL, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. J Mol Biol 1999;285:2177–2198.
51. Hou T, Wang J, Li Y, Wang W. Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 2011;51:69–82.
52. Simonson T, Archontis G, Karplus M. A Poisson-Boltzmann study of charge insertion in an enzyme active site: the effect of dielectric relaxation. J Phys Chem B 1999;103:6142–6156.
53. Marshall NJ, Payne JW. The importance of electrostatic charge and dielectric constant in conformational analysis of biologically active dipeptides. J Mol Model 2001;7:112–119.
54. Yang T, Wu JC, Yan C, Wang Y, Luo R, Gonzales MB, Dalby KN, Ren P. Virtual screening using molecular simulations. Proteins 2011;79:1940–1951.
55. Warshel A, Papazyan A. Electrostatic effects in macromolecules: fundamental concepts and practical modeling. Curr Opin Struct Biol 1998;8:211–217.
56. Antosiewicz J, McCammon JA, Gilson MK. The determinants of pKas in proteins. Biochemistry 1996;35:7819–7833.
57. Hajjar E, Dejaegere A, Reuter N. Challenges in pKa predictions for proteins: the case of Asp213 in human proteinase 3. J Phys Chem A 2009;113:11783–11792.
58. Schutz CN, Warshel A. What are the dielectric "constants" of proteins and how to validate electrostatic models? Proteins 2001;444:400–417.
59. Lund M, Jönsson B, Woodward CE. Implications of a high dielectric constant in proteins. J Chem Phys 2007;126:225103.