International Journal of Biological Macromolecules 317 (2025) 144813

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

AI, docking, and molecular dynamics to track the binding of structural peptides to different keratin models

Tiago Ferreira [a,c], Francisco Pereira [a], Filipa Gonçalves [c], André da Costa [c], Artur Cavaco-Paulo [a,b,c,*], Tarsila G. Castro [c,*]

[a] CEB – Center of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
[b] LABBELS – Associate Laboratory, Braga, Guimarães, Portugal
[c] SOLFARCOS – Soluções Farmacêuticas e Cosméticas, Lda., 4700-034 Braga, Portugal

ARTICLE INFO

Keywords:
Artificial intelligence
Hair-binding peptides
Keratin models
Molecular dynamics simulations
Molecular docking
Peptide-keratin interactions

ABSTRACT

The present work shows a computational approach to assess the interactions of different nature-inspired peptides with hair keratin models. An updated keratin model was validated, and comparisons with previous models were traced, thereby highlighting the time-evolution of both the model and the in silico methods. Importantly, the computational methodology presented here allows for the study of peptide-protein interactions between unknown or unsolved structures. For that, artificial intelligence, molecular docking, and molecular dynamics simulations were used to predict the peptides' structural conformation, stability, binding, and interaction patterns with keratin. The keratin models were designed to include a conserved form and two hyperoxidized forms, where a few thiols were replaced by sulphonic acid in cysteine residues. Most peptides displayed strong affinity and stability over time, established by distinct interactions with the three keratin models, but peptides P34 (resilin-like) and P36 (abductin-like) stood out due to their amino acid variety, the different types of interactions they performed and the sustained binding proximity within the keratin helices. AlphaFold also demonstrated high reliability and accuracy when constructing the 'virgin' keratin model and the peptide structures. This study furthers many others on peptide-protein interactions whilst paving another setting stone in the use of computational methods on previously under-explored biotechnological fields such as cosmetics.

1. Introduction

In recent years, the cosmetic industry has witnessed the employment of small peptides as key ingredients in various formulations, particularly in hair care products. These molecules have garnered attention for their ability to address specific hair health and appearance concerns [1–3]. Amongst their numerous applications, their interaction with keratin (KRT), the primary structural protein of hair fibers, is particularly remarkable [4–6]. Due to their unique biochemical properties, small peptides can penetrate the hair shaft, and interact with KRT fibrils on a structural level, and by targeting specific regions, they can improve the overall health and appearance of the hair [4,7–9].

Human hair fibers are about 50–100 μm in diameter and are composed of three distinct structures: cuticle, cortex, and medulla. The cuticle is the outermost layer that protects the inner structures from external agents. Approximately 65–95 % of the cuticle is amorphous KRT, with the remaining composition being mostly lipids (1–9 %) and pigments (0.3–0.9 %). This structure is also rich in cystines, which increase the toughness of the hair fiber, and is one of the most attractive targets for new actives [10,11]. The cortex contains the intermediate filaments, which comprise several keratin tetramers or protofilaments, which in turn are composed of α-helical keratin proteins [12]. These keratin filaments are exposed when the hair cuticle is damaged. Consequently, its intra- and intermolecular interactions, as well as the disulfide bonds between chains, are equally sensitive to chemical and thermal manipulation.

Hair bleaching, dyeing, and perming (curling) are techniques widely used in beauty salons to create various hairstyles. However, these treatments involve redox steps that alter the structure of the hair fiber, causing significant damage [10,13]. Oxidizing agents like hydrogen peroxide can act directly on hair cystines, where strong and successive oxidation reactions lead to their unwanted conversion to cysteic acid (CYA), a major hallmark of hair damage [6]. Overall, these treatments can cause permanent hair damage by making the neutral and

* Corresponding authors.
E-mail addresses: artur@deb.uminho.pt (A. Cavaco-Paulo), tgc@solfarcos.com (T.G. Castro).

https://doi.org/10.1016/j.ijbiomac.2025.144813
Received 29 January 2025; Received in revised form 28 May 2025; Accepted 29 May 2025
Available online 30 May 2025
0141-8130/© 2025 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

hydrophobic hair surface become acidic and hydrophilic [14–16].

In the past, Cavaco-Paulo's group developed a successful commercial peptide, based on *Homo sapiens sapiens* (human) hair KRT type II cuticular Hb5 sequence (Uniprot ID P78386, KRT85) and named it as *KP* peptide. This molecule is currently used in cosmetics to improve hair's mechanical properties [7,17,18]. The results obtained from the KP peptide study ignited the search for novel peptides with physicochemical or organoleptic properties targeting hair health and appearance.

Recently, we gathered efforts to design and developed a new series of peptides that are now patented [WO/2024/166050– Fusion Proteins/Peptides, Methods and uses Thereof] [19]. In silico methods have been employed to shed light on the molecular features and interactions between some of these peptides and hair-KRT models, both damaged and undamaged. The aim was to identify key interactions and binding affinities, to enable their further design and optimization as ingredients in hair cosmetics.

The present work focuses on the study of peptide sequences derived from important structural proteins with remarkable mechanical properties, that don't have cysteine in their sequence, such as collagen, elastin, resilin, abductin, and silk-fibroin. The main aim is to evaluate the binding established by other types of interaction besides the typical attraction and reaction between Cys on the peptide and Cys of keratin. This collection of proteins was recently reviewed by Miserez and colleagues [20], where it was determined that this set of proteins could be repurposed in the design of new materials and technology due to their elastomeric profile.

A previous KRT model [9] was updated using the generative AI technique implemented by AlphaFold (AF) 2.0 [21], which predicted the structure of a KRT tetramer from the KRT34-KRT86 heterodimer, and evolved the computational protocol in contrast to previously employed techniques. This tetramer represented our 'virgin' KRT fiber. The oxidized forms were designed from this structure by replacing 3 to 4 random cysteine residues for cysteic acid. Similarly, AI was used to predict the secondary structure of the peptides under study, providing our initial structures to dock against KRT. Our findings indicate the preferable binding regions and interactions, and molecular dynamics (MD) simulations estimate how stable these complexes are over time, considering both the maintenance of docked interactions and the establishment of new ones.

2. Materials and methods

2.1. Hair-keratin tetramer: Description, modifications, and AI

In the past, Cavaco-Paulo's group generated 2 KRT models by using homology modelling and rigid docking [8,9,22]. To date and to our knowledge, there are no X-ray structures for human hair KRT at the atomic level, thus, we used AI technology, via Alphabet's AF 2.0 [23,24], to generate an updated model based on KRT34-KRT86 heterodimers (UniProt IDs O76011 and O43790, respectively). For this, the coil 2 sequence-segment of type I and type II KRTs was selected as an input, which is a well-defined α-helical portion, able to represent the α-helical rod domain (Fig. 1). The damaged KRT forms were designed by randomly replacing 3 or 4 out of the 12 cysteines by cysteic acids (thiol replaced by sulfonic acid). These two forms, named KRT_{oxi3} and KRT_{oxi4} (Fig. 1), represent the average percentage of oxidation found in chemically damaged hair [25]. These conversions alter the electrostatic profile of KRT by adding negative charges to the surface.

A thorough analysis of the KRT model was accomplished by determining its electrostatic potential surface via the Adaptive Poisson–Boltzmann Solver (APBS) software [26]. In addition, EMBOSS's sequence statistics dispatcher was used to describe the amino acid properties of KRT, such as charged residues [27].

2.2. AI-based structure prediction of peptide sequences

Our selection consists of 7 peptides (Table 1) with different sizes, polarities, and secondary structures. These peptides represent our new

Table 1
Peptide classes and sequences, where peptides are numbered according to the patent WO/2024/166050. Classes represent their respective domains that are comprised of elastin-like peptides (ELP), silk-like peptides (SLP), collagen-like peptides (CLP), resilin-like peptides (RLP), abductin-like peptides (ALP) and Linker/Spacer-like peptides.

Peptide Number	Class	Sequence
22	ELP	GVGVPGVGVPGLGVPGVGVPGVG
32	SLP	GAGAGSGAGAGSGAGAGSGAGAGSGAGAGS
33	CLP	GPTGPTGPAGPRGLQGLQGLQGERGEQGPT
34	RLP	GGRPSDSYGAPGGGN
36	ALP	GGFGGMGGGS
98	Charged linker	AKAKAKAKAKAKAKAKAKAKAKAKAKAK
99	Charged linker	LKLKLKLKLKLKLKLKLKLKLKLKLKLK

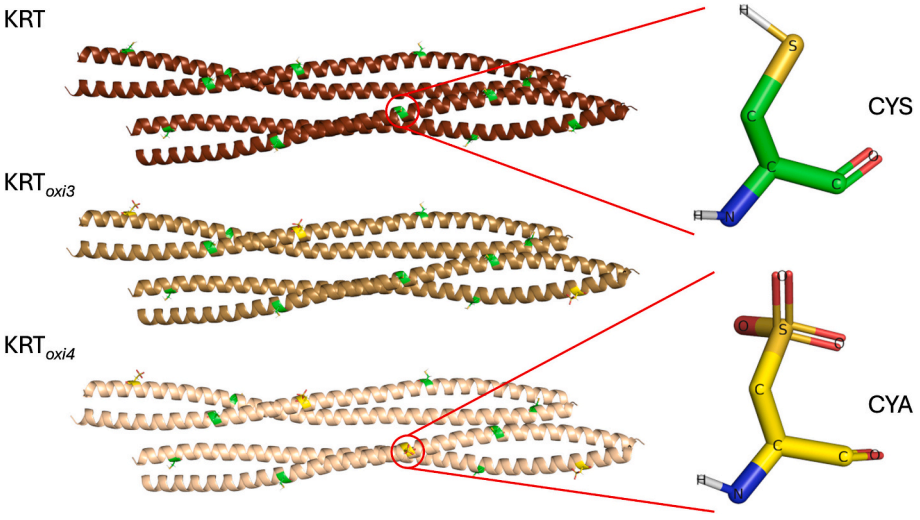


Fig. 1. Designed and predicted KRT models in different states of damage. KRT represents a virgin keratin fiber, KRT_{oxi3} refers to a keratin with 3 oxidized cysteines, and KRT_{oxi4} represents 4 oxidized cysteines. All cysteine residues (CYS) are highlighted in green, whilst modified ones (CYA) are in yellow sticks.

nature-inspired designs, where repeating motifs of elastin (ELP), collagen (CLP), resilin (RLP), silk-fibroin (SLP), abductin (ALP), or linker/spacer-like sequences were considered. These peptide classes refer to structural/mechanical proteins, which may confer many properties for cosmetic applications in haircare, such as toughness, adhesion, elasticity, strength, and elastomeric properties [20].

The 7 peptides also come from a list of over 500 patented peptide sequences and combinations of sequences, advanced in experimental testing amongst our research team for their incorporation in hair-care formulations. The numbering of peptides follows the numbering attribution on the patent (WO/2024/166050).

The different sequences were used as input into AF 2.0, which employs deep learning techniques to model the physical interactions between amino acids in a sequence, allowing it to predict the three-dimensional conformation of a protein with outstanding accuracy. It leverages vast amounts of genomic data and evolutionary information, using neural networks trained on known protein structures to infer the spatial arrangement of amino acids in an unknown sequence, thus providing insights into the protein's functional mechanisms and potential interactions [21,28].

The resulting structures were collected and used in this work.

2.3. MD simulations on keratin model and peptides

The generated KRT model was submitted to MD simulations for 50 ns to assess its stability in solution, using the GROMACS 2021 series and the GROMOS54a7 force field [29,30]. Similarly, all peptide structures were evaluated in solution for 50 ns to follow the maintenance of the predicted secondary structure.

All systems were energy minimized using the steepest descent algorithm for 50,000 steps. Subsequently, a double initialization procedure took place, first equilibrating temperature in a canonical ensemble (NVT; constant number of particles (N), volume (V), and temperature (T)), then equilibrating density by using an isothermal-isobaric ensemble (NPT; constant number of particles (N), pressure (P), and temperature (T)), with each step running for 100 ps to relax hydrogen atoms and water molecules. The V-rescale thermostat algorithm [31,32] was employed to set the systems to room temperature (298 K, approximately 25 °C), and pressure was controlled at 1 atm using the Parrinello-Rahman barostat [32,33]. The coupling constants used were $\tau_T = 0.10$ ps for temperature and $\tau_P = 2.0$ ps for pressure. Once the systems were adequately prepared, production runs began without position restraints using the NPT ensemble. The particle mesh Ewald (PME) method [34] was used for the treatment of electrostatic interactions in a single cut-off scheme of 1.4 nm. The van der Waals interactions were simply truncated from 1.4 nm.

Following the same conditions, the damaged KRT forms underwent MD simulations for 50 ns, but at 310 K (37 °C). These simulations had a higher temperature to fully assess the stability under chemical modifications.

2.4. Molecular docking of peptide to keratin tetramer

The peptide-protein docking was employed via HDOCK server by the Huang Laboratory [35], which performs a rigid-body docking approach for protein-protein interactions. Although flexibility is a very important factor in docking simulations and for peptide-protein systems, HDOCK has been shown to be more effective than other docking tools such as HADDOCK, PyDock, and even Rosetta [36]. The Citation Report from Clarivate's Web of Science shows that HDOCK has been widely cited and used over the past five years. HDOCK is a versatile, open-access program that is both fast and practical. The user only has to input two PDB structures and indicate the receptor and the ligand. Importantly, unlike PyDOCK, it can deal with non-canonical amino acids such as cysteic acid (oxidized cysteine). Here, the structures were considered in their entirety, and the options were kept standard, as there were no specific docking site residues, and no symmetric multimer docking was employed. LigPlot+ was used to disclose the interaction pattern between peptide-protein pair [37].

2.5. Molecular dynamics simulations and analysis

To evaluate the stability of the complexes generated in the previous step, MD simulations were performed for 30 ns, using the GROMACS 2021 series and the GROMOS54a7 force field [29,38]. The parameters described in section 2.3 were also used here.

The systems were all assembled into a simulation box of 2400 nm^3 ($10 \times 20 \times 12$ nm) that was filled with water molecules using the simple point charge (SPC) water model. Ions were then added to balance the overall systems' charges, thus emulating their natural environment. The extremities of the KRT tetramer were placed near the box's limits to account for fiber behavior across periodic boundary conditions (PBC) copies.

Once the systems were adequately prepared, production runs began, and short simulations, with no position restraints, were conducted to follow peptide-keratin dynamics. Due to the size of the systems and considering the previous stabilization of both the keratin and the peptides individually, in solution, the time frame of 30 ns was determined. Several analyses were conducted to evaluate the peptide-KRT interaction, such as RMSD (Root Mean Square Displacement), RDF (Radial Distribution Function), clustering, and H-BOND (Hydrogen Bond analysis). Clustering analysis was performed by using the single-linkage method and following the RMSD variance amongst sampled peptide structures over time. Similar peptide conformations are grouped up to a 0.15 nm cut-off.

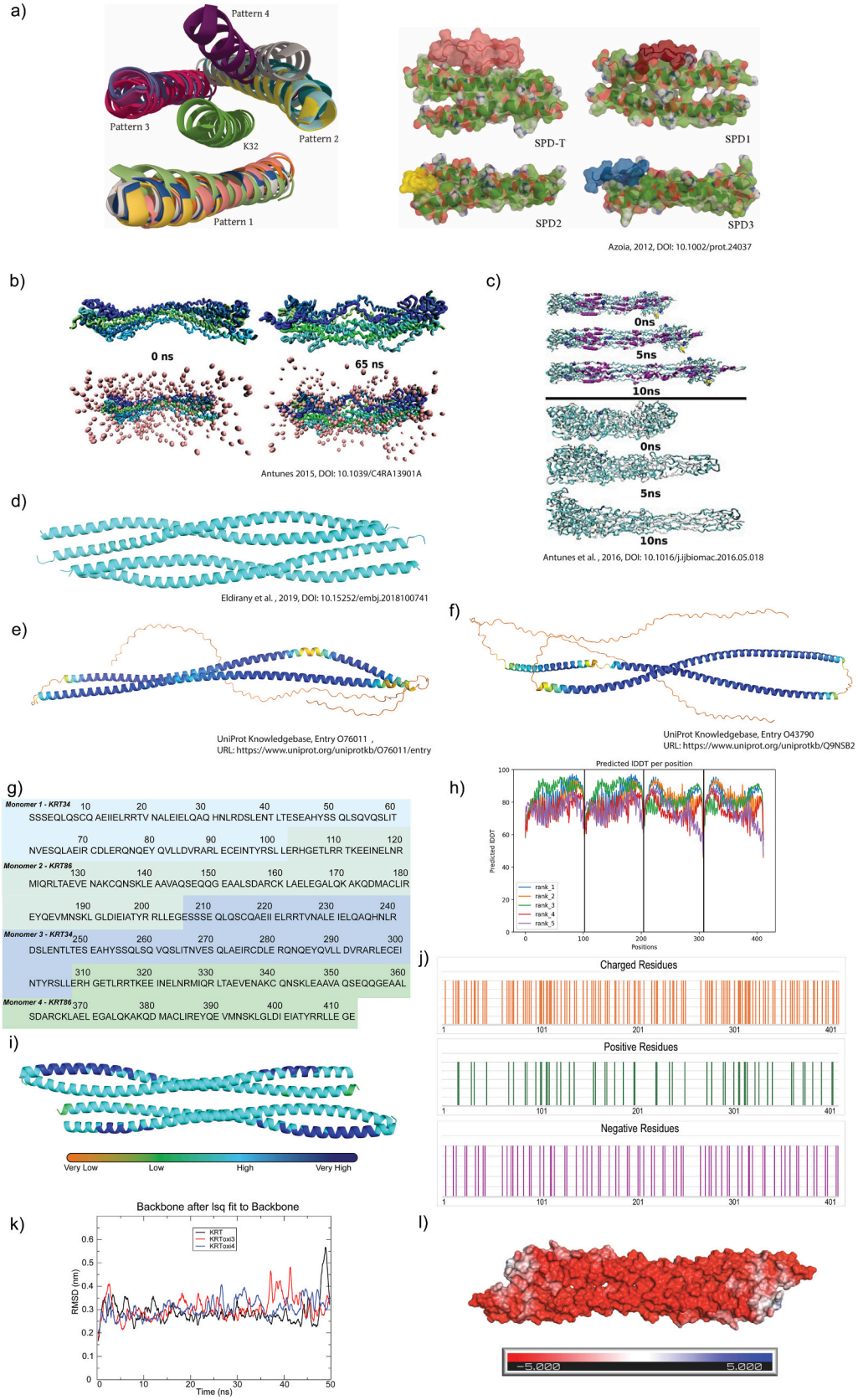
2.6. Principal component analysis

Principal component analysis (PCA) was conducted using the SPSS software and the dimension reduction function with varimax rotation [39]. This analysis was used to understand the role of several intrinsic and calculated properties of peptides in their binding towards the keratin types. Thus, the data collected includes the peptides' binding scores on each of the three targets, KRT, KRT_{oxi3}, and KRT_{oxi4}, and intrinsic and physicochemical properties of the peptides, namely: the number of acidic amino acids (aAA), number of basic amino acids (bAA), number of aromatic amino acids (aroAA), peptide's net charge (charge), number of hydrophobic amino acids (hAA), number of residues having alcohol groups (OH), number of Gly, isoelectric point (pI), grand average of hydropathicity index (GRAVY) and molecular weight (MW), all calculated using the ProtParam tool from ExPasy web server [40].

3. Results and discussion

3.1. AI structure prediction

KRT fibers are formed by assembling multiple KRT heterodimers, which in turn combine cuticular type I and type II KRT. The most abundant KRT heterodimers in human hair shafts are KRT31-KRT85, KRT34-KRT86, and KRT32-KRT84, which are very similar in sequence and structure [41]. The construction of the present keratin model evolved from previous models, starting from the work of Azoia et al. (2012) [9], where a short keratin heterodimer was built using monomers generated by homology modelling based on the vimentin structure. Keratin 32 (KRT32) and keratin 84 (KRT84) were selected due to their high sequence identity with the template, and docking was performed using KRT32 as the receptor and KRT84 as the ligand (Fig. 2, a). This work was later extended by assembling a keratin protofibril composed of four KRT34/KRT86 heterodimers (an octamer), modelled using coarse-grained MARTINI simulations of the 2B domain in an antiparallel configuration. These models were further refined through 500 ns of MD simulations (Fig. 2, b) [8].



(caption on next page)

Fig. 2. Keratin model generation process from the a) KRT32-KRT84 assembly pattern and subsequent docking of a peptide to the short heterodimer model b) coarse-grained KRT34-KRT86 octamer simulated by Antunes et al. (2014) c) United-atom steered KRT34-KRT86 octamer also predicted by Antunes et al. (2015) d) KRT1-KRT10 tetramer experimentally solved by X-ray diffraction, e) KRT34 AlphaFold full structure with UniProt entry O76011, f) KRT86 AlphaFold full structure with UniProt entry O43790, g) input sequence composed by the Coil 2 segment of KRT34 and KRT86 monomers; h) the plDDT scores per position of each predicted structures or ranks; i) AlphaFold 2.0 predicted structure with plDDT confidence coloring; j) amino acids' properties throughout the 412 total residues, k) RMSD of each keratin model over 50 ns of MD simulation and l) the electrostatics potential surface determined by APBS.

The search for an accurate keratin protofibril representation continued in a subsequent study, where the coarse-grained KRT34/KRT86 model was converted into a united-atom representation and subjected to steered molecular dynamics simulations. This remains the largest non-coarse-grained model of a keratin complex available, but the secondary and tertiary structures are not fully helical or well-defined (Fig. 2, c) [22]. Nonetheless, the high-resolution crystal structure of the KRT1/KRT10 tetramer by Eldirany et al. (2019) [42] provided essential spatial constraints and interaction motifs applicable to type I–type II keratin pairings (Fig. 2, d). Lastly, on the UniProt entries for human KRT34 (O76011) [43] and KRT86 (O43790) [44], the full structure derived from AlphaFold are presented with high confidence on the helical portions, (Fig. 2, e, f), which, together with Eldinary's model, enables the prediction of the novel keratin model presented here.

Our updated KRT model, constructed with machine learning, was based on KRT34-KRT86 heterodimer coil 2 sequence-segment (Fig. 2, i). The resemblance to Eldinary's experimental structure enhances our confidence in the present AlphaFold model, rather than following the previous methodologies used by our group in the past. Our confidence in the AI-generated model also stems from the current validation of the AlphaFold itself, which has been transversal to various types of proteins and folding, and is intensively applied in research today.

To sum up, the role of AF in the prediction of such complex macromolecules was very important in the present work, both for reliability and simplicity, as in the past, previous models were constructed by using a combination of homology modelling, docking, and molecular dynamics simulations [9]. Here, AF estimates the tetrameric structure directly, replacing the necessity of traditional homology and docking steps to assemble the tetramer. MD simulations was only used to relax the obtained structure. The predicted KRT structure coloring points towards a high to very high confidence, as the AF color pattern suggests (Fig. 2, i). In addition, the predicted local-distance difference test (plDDT) in Fig. 2(h) indicates the per-residue accuracy of the structure, showing that the first rank is predicted with a high degree of certainty, along with some uncertainty for the termini.

The KRT34 portion, type I and acidic, has 102 amino acids, the most common being leucine, glutamate, serine, and glutamine, with a charge of −10 (monomers 1 and 3 of Fig. 2 (g)), whereas the portion of KRT86, type II, has 104 amino acids (monomers 2 and 4 of Fig. 2 (g)), with the most common residues being glutamate, alanine, and leucine, with a − 3 charge. The electrostatic potential surface reflects this overall negative charge (Fig. 2, l) for these coil-2 segments of heterodimers KRT34-KRT86, confirmed by the statistics on the amino acids' profile of charged residues (Fig. 2, j). This surface is also very similar to the reported for KRT1-KRT10 heterodimer [42]. The final net charge of each monomer is balanced by the great amount of hydrophobic amino acids present in KRT.

The RMSD analysis of the native and oxidized KRTs showed very stable keratin models, with low conformational displacements of around 0.3 nm, indicative of minimal structural deviations from the predicted conformations (Fig. 2, h).

Regarding peptide classes, P22, elastin-like, was thought to carry the same elastin properties: remarkable resilience and structural stability [45–48], and some hydrophobicity due to the GVGVP repeat. The P32, silk-like, through the GAGAGS [49] block, aims to bring resistance and elasticity, and assure some hydrogen bonds from Ser residues [45,46]. Silk fibroin is known to exhibit mechanical properties superior to synthetic polymers [47]. Collagen-inspired oligopeptides are already in use in cosmetics. Here, the sequence GPTGPTGPAGPRGLQGLQGLQGERGEQGPT can act as a shield against daily stressors on hair. Resilin is a highly deformable elastomeric protein with high resistance to stretching and the ability to restore its original form [50–52]. The motif GGRPSDSYGAPGGGN was considered and represents the most versatile sequence in terms of amino acid types. Abductin is an elastomeric protein like elastin and resilin [53] with rubber-like properties, being more elastic than elastin but less than resilin [51,52]. The amino acid sequence of our abductin-like peptide is rich in glycine and carries one aromatic residue [55], GGFGGMGGGS [56]. Lastly, the linker/spacer sequences, P98 and P99, were thought to study the positive charge role in peptide-keratin interaction and their availability to be coupled to other sequences in the future.

Observing the peptides' structure prediction, the peptide sequences have diverse folding patterns (Fig. 3), with 4 out of 7 having α-helical portions (P33, P34, P98, and P99), 2 presenting a β-sheet tendency (P22 and P32), and 1 being highly unstructured (P36). All the structure predictions were considered to have a high degree of fidelity, proven by their plDDT color grading, which only points to less certainty for the helical structures P33, P98 and P99.

Information on amino acid types and physicochemical properties was also collected for the set of peptides in order to relate their profile to keratin binding. For that, we compute the number of acidic amino acids (aAA), number of basic amino acids (bAA), number of aromatic amino acids (aroAA), peptide's net charge (charge), number of hydrophobic amino acids (hAA), number of residues having alcohol groups (OH), number of Gly, isoelectric point (pI), Grand average of hydropathicity index (GRAVY) and molecular weight (MW). Data was collected on the Expasy ProtParam web tool [40].

From the physicochemical properties collected, GRAVY values indicate that P33, P34, P98, and P99 are considered hydrophilic, yet the high content of charged amino acids on P99 is balanced by the apolar and bulky Leu residues in comparison with Ala of P98. Importantly, P34 is the most hydrophilic from the series and populates all but one of the properties listed, being the most versatile peptide. On the other hand, the positive GRAVY values of P22, P32, and P36 point to hydrophobicity. In fact, P22 and P32 exhibit a considerable number of hydrophobic residues. Regarding pI, most peptides need a slightly acidic pH to assume the neutral form in solution, exceptions are the highly positively charged P98 and P99. Again, all peptides except P98 and P99, present a medium to high content of Gly residues, assuring conformational flexibility. The peptides also vary in MW, P98 and P99 being the ones with higher molecular weight and a greater number of residues. P22, P32, and P33 are of intermediate size/weight, while P34 and P36 are the smallest peptides with low MW.

MD simulations were then applied to infer the stability of the predicted structures along simulation time. The RMSD of the peptides pointed to a high variation, ranging from 2.8 to 3.8 nm (Table S1). Free in solution, the peptide's secondary structure is expected to be different than when near a macromolecule because the non-bonded contributions involved in its neighborhood are massively different. Nevertheless, P33, P98, and P98, although predicted with less confidence, are stabilized in solution, keeping a great portion of the predicted α-helical secondary structure (Fig. S1). Peptides P22, P32, P34, and P36 are highly mobile in solution, exploring diverse conformational states. On the other hand, the KRT tetramer and its oxidized forms are expected to introduce a strong negative charge into their environment, which will greatly affect the behavior and structure of the peptides, due to changes in polarity and

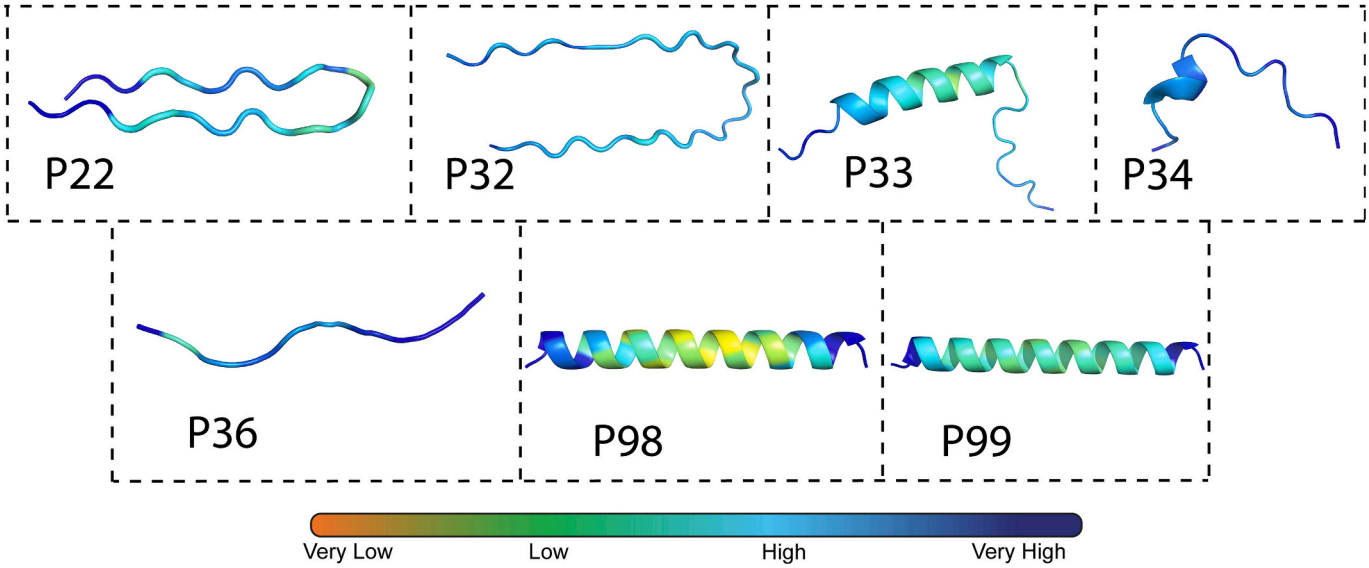


Fig. 3. Structural prediction of peptide sequences according to AF 2.0. Peptides are numbered accordingly to the patent WO/2024/166050. P22 peptide sequence is based on tropoelastin consensus sequence, P32 is based on silk fibroin monomer, P33 is based on typical collagen repeating motif, P34 is based on resilin consensus sequence, P36 is based on abductin consensus sequence, and P98 and P99 are designed as linker/spacer-like peptides. Coloring schemes follow AF's plDDT color pattern describing the confidence of AF's prediction according to each residue.

electrostatics.

3.2. Molecular docking and MD simulation of keratin:Peptide complexes

Table 3 shows the dimensionless docking scores obtained by the HDOCK server, with the most negative values representing the best binding predictions. The entirety of the *KRT* structures was considered in a blind docking procedure. HDOCK employs a hierarchical docking approach involving global rigid-body docking followed by semi-flexible refinement. Its scoring function incorporates shape complementarity and uses a knowledge-based scoring function, improving prediction reliability, to evaluate protein-protein interactions [35]. Additionally, Table 3 shows the number of H-bonds occurring between the peptide and tetramer at the predicted docking site. Here, the ligand efficiency (LE) corresponds to the ratio between the binding score and the number of peptide residues (typically, is the binding energy/number of heavy atoms).

A first look at the scores indicates that, in general, a given peptide binds with similar affinity to all KRT models. P34 stands out as the one with the most similar scores despite the target (lowest Std Dev). Along with P36, this peptide also had great ligand efficiency (score/n° of residues) due to their small size and amino acid types. Additionally, P36 has 7 glycines and only 3 residues with “alkyl” side chains, which gives it great flexibility that augments the ligand efficiency. P34 has 6 glycines, but 9 “alkyl” side chains, which should impact the interactions established, conformation, and binding site.

Having many Gly evenly spaced doesn't seem to be advantageous in longer peptides like P32. Despite the high number of residues capable of interacting with KRT and the good number of H-bonds with KRT_{oxi3} and KRT_{oxi4}, this peptide has a low ligand efficiency. In this case, the OH groups from Ser amino acids must contribute to the H-bonds. P22 also has a balanced Gly distribution, but its small size appears to give the peptide a better score. P22 and P33, which have intermediate molecular weights, had the starkest docking scores towards the three KRT tetramers, but the Std Dev amongst KRT models was more accentuated in these cases.

Looking globally at the H-bonds for the first docking poses, one can conclude that although important to the binding, they are not always the leading interaction. P22 and P33 have very negative docking scores, but the number of H-bond interactions they establish with KRT is divergent. Again, a combination of structural flexibility and different types of interactions must influence the final binding score. One would expect higher binding energy for the highly positive peptides P98 and P99 with the oxidized KRT forms (which are negatively charged), but due to steric hindrance, repulsion near positively charged amino acids, and other factors, some penalty must occur. Furthermore, these two peptides, although longer than the remaining ones, presented a low number of H-bonds in proportion to their size and considering the number of hydrogen bond donors from the 15 Lys side chains on each peptide. The Lys possibly prioritized ionic interactions towards the negative/oxidized KRTs forms rather than H-bonds (distance and angle cut-off not satisfied for hydrogen bonds).

Fig. 4 helps to visualize the structural versatility of P34. This peptide acquires different binding locations consonant with the KRT form but settles in a region where the keratin amino acids are mostly polar and/or charged, allowing for strong hydrogen bonds and salt bridges.

Analyzing now the peptide-KRT complexes after MD simulations, the RMSD values for all *KRT* models were very similar, stable, and low (Table 4; Fig. S2). They were mainly around 0.28 and 1.00 nm, indicating that both the oxidation and the presence of peptides do not impair the KRT structure. The exception was P98, which caused a significant deviation in the oxidized forms, but generally, the systems are well-equilibrated for the time window the simulations were conducted.

Remarkably, the peptides' RMSD complexed with KRT was much less pronounced than their RMSD in solution (Table S1). Although there are still some stark differences, the values ranged mostly from 0.2 to 1.2 nm. P98 and P99 consistently had the lowest RMSD values, oscillating between 0.14 and 0.28 nm, probably due to their well-defined α-helix structure (Fig. 2).

P32, a glycine-rich sequence, presented the highest deviations along the simulations as expected due to its highly flexible structure. P33 also showed relevant RMSD, but in this case, the reason is less clear, as this peptide had a remarkable score in all three KRTs and a reasonable number of H-bonds. Despite carrying two Glu residues, repulsion is probably not the reason because it didn't affect the previous quantities. This peptide remains well-bonded to the KRTs but with great plasticity on its conformation.

P34 and P36 are quite mobile due to their small size, causing moderate RMSD variations in all KRT models.

Overall, the RMSD pattern (Fig. S2) indicates that the 30 ns time-

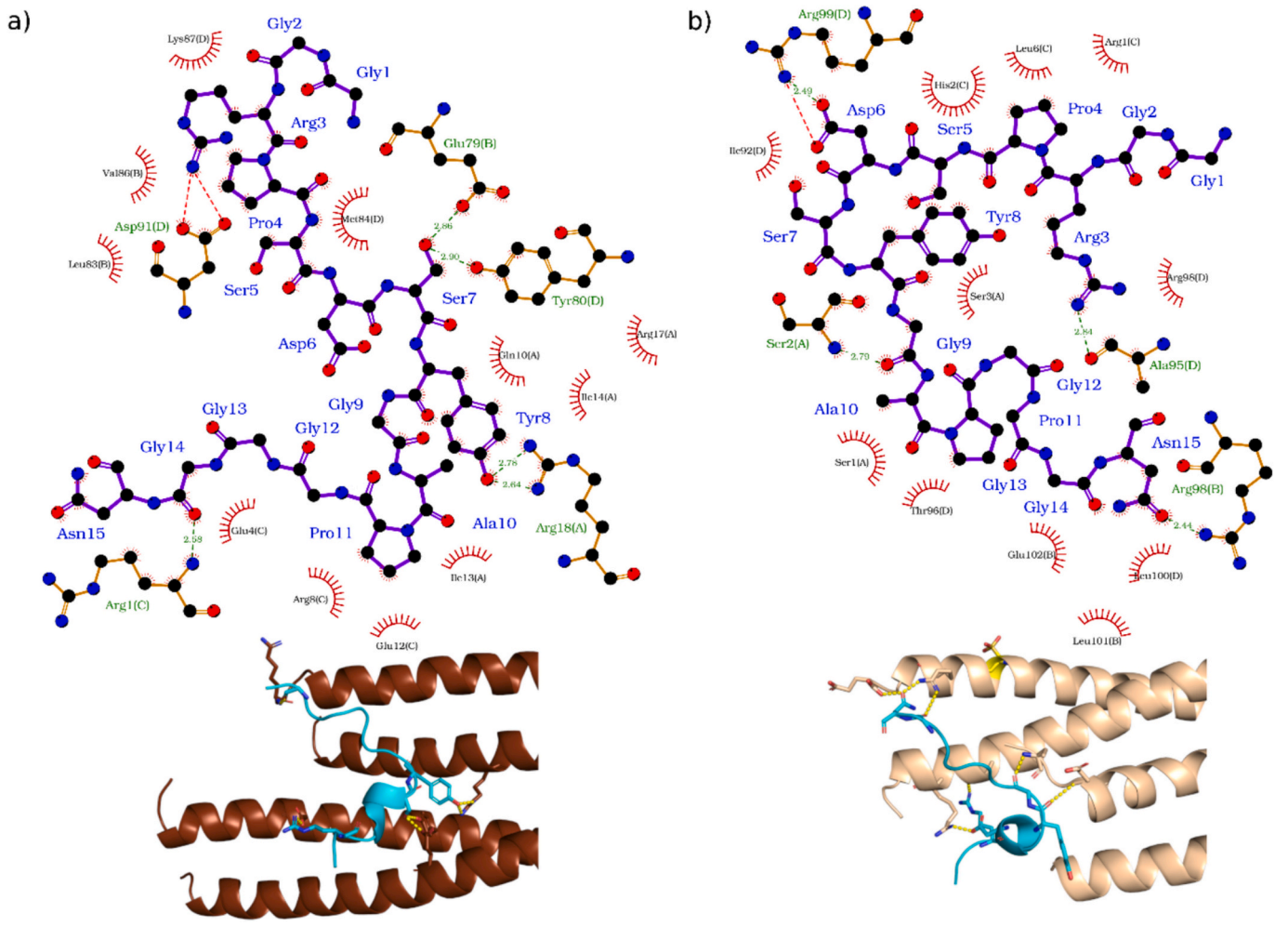


Fig. 4. 2D interaction diagrams for peptide P34 interacting with KRT (a) and KRT$_{oxi4}$ (b) models, where hydrogen bonds are represented in dashed green traces and salt bridges in red dashed traces; blue atoms are Nitrogen, red are Oxygen and black are Carbon. Below, is the respective docking configuration of the first pose of P34 (cyan) on both targets, with hydrogen bonds represented in yellow dots. The keratin colors are related to the oxidation level; virgin hair is dark brown, and the oxidated KRT$_{oxi4}$ is light brown, with yellow sticks highlighting CYA residues.

window was sufficient to stabilize the systems and to capture valuable insights into the peptide-protein dynamics. Although some peptides could benefit from a longer simulation time, such as P32, it is observed that in the last 10 ns of simulation, the RMSD of most peptides varies in a range of up to 0.1 nm, indicating stabilization in the keratin vicinity.

Afterwards, the average number of H-bond interactions established between the peptides and KRT throughout the post-docking MD Simulations was computed (Table 5). The average number of H-bonds during simulation was lower for some peptides compared to the docking pose, such as for P98 and P99, but consolidated with additional interactions for others, such as for P32. A similar number of interactions was observed for peptides P34 and P36, both in docking and during simulation. This diverse behavior proves that the peptides interact dynamically with KRT. They are not tied to the docking position but undergo a constant adsorption-detachment process, revealing a transient nature of the interaction towards the keratin surface during the simulation, which is somehow expected for MD simulations. Nonetheless, most peptides keep a close distance to the fiber despite the interaction type preference.

Particularly, P32 acquired between 16 to 19H-bonds when interacting with the oxidized KRT forms, which is a good indicative for the binding of this longer peptide in simulation conditions.

For the remaining peptides, a slight decrease of the average number of H-bonds is observed towards the oxidized forms, from which one can infer that another type of interaction, such as ionic, takes place. Also, the peptides' proximity to the KRT may reinforce weak interactions, like van der Waals and other hydrophobic forces.

Hydrogen bonds existence maps were also calculated using the GROMACS' *hbonds* module. This result offers a detailed, comprehensive visualization of hydrogen bonding dynamics, indicating all the possible H-bond pairs throughout the frames but also revealing the persistence of each pair over time. All H-bond maps are presented as Supplementary Data (Figs. S3-S4), from which one can calculate the global number of H-bond pairs, identify the interacting pairs, and look at the occupancy (%) of how often a particular H-bond occurs.

For P34, 87H-bond pairs are observed towards KRT and 175 pairs to KRT$_{oxi4}$. P36 presents 77 different H-bonds pairs on KRT and 120 on KRT$_{oxi4}$. Interestingly, P32 has the opposite tendency, having 180H-bond pairs towards KRT and 150 with the oxidized form (Table S2).

In addition, the maintenance of the identified pairs was estimated over time. Peptides P32, P34 and P36 are highlighted in Table 6 due to the remarkable number of hydrogen bond pairs between these peptides and the keratins mentioned above. Thus, the five preferable interactions/pairs, i.e. the ones which persist throughout a higher number of frames, were tracked. One can conclude that peptides P32, P34, and P36 interact mainly with polar and charged amino acids, and although keratin is richer in negatively charged amino acids, peptide interactions with the side chain of positive amino acids are favored and more stable over time. Remarkably, these interactions are mediated mostly through

Ser and Gly amino acids present on peptides. In P36, a noteworthy tendency is for Ser10 to interact in all KRT forms, with Arg in the case of KRT and KRT$_{oxi4}$ and a Lys on KRT$_{oxi3}$. The remaining peptides also tend to interact mostly with Arg, nonetheless, hydrogen bonds are also established with Lys, Glu, Ser, Thr, Asp, and Gln (Table S3).

Regarding the occupancy, although the number of H-bond pairs increased for P34 and P36 with the oxidized forms, the stability of these interactions is almost always inferior compared to the ones established with KRT (the exception is P36 in KRT$_{oxi3}$). This, again, indicates a very dynamic interaction between parts, with H-bonds continuously forming and breaking apart. Importantly, amongst the identified pairs, one can track the participation of all keratin chains (A to D), which indicates great internalization of the peptides.

The RDF analysis of the peptides near the KRT surface was conducted (Fig. 5). These results are shown at close distances, when the non-bonding interactions, such as H-bonds, ionic, and Van der Waals, are maximized. Thus, we present the maximum RDF up to 0.5 nm from the KRT surface. The results showed that P32, P34, and P36 had a higher probability of being placed at close distances from all KRT models, followed by P22 and P33. Of note, P36 and P34 have higher ligand efficiencies against all KRT tetramers, while P32, although with a good score, has a low ligand efficiency in proportion to its size. These two peptides, P34 and P36, have the lowest molecular weight, a balanced number of Gly, and aromatic amino acids (Table 2). Therefore, one can

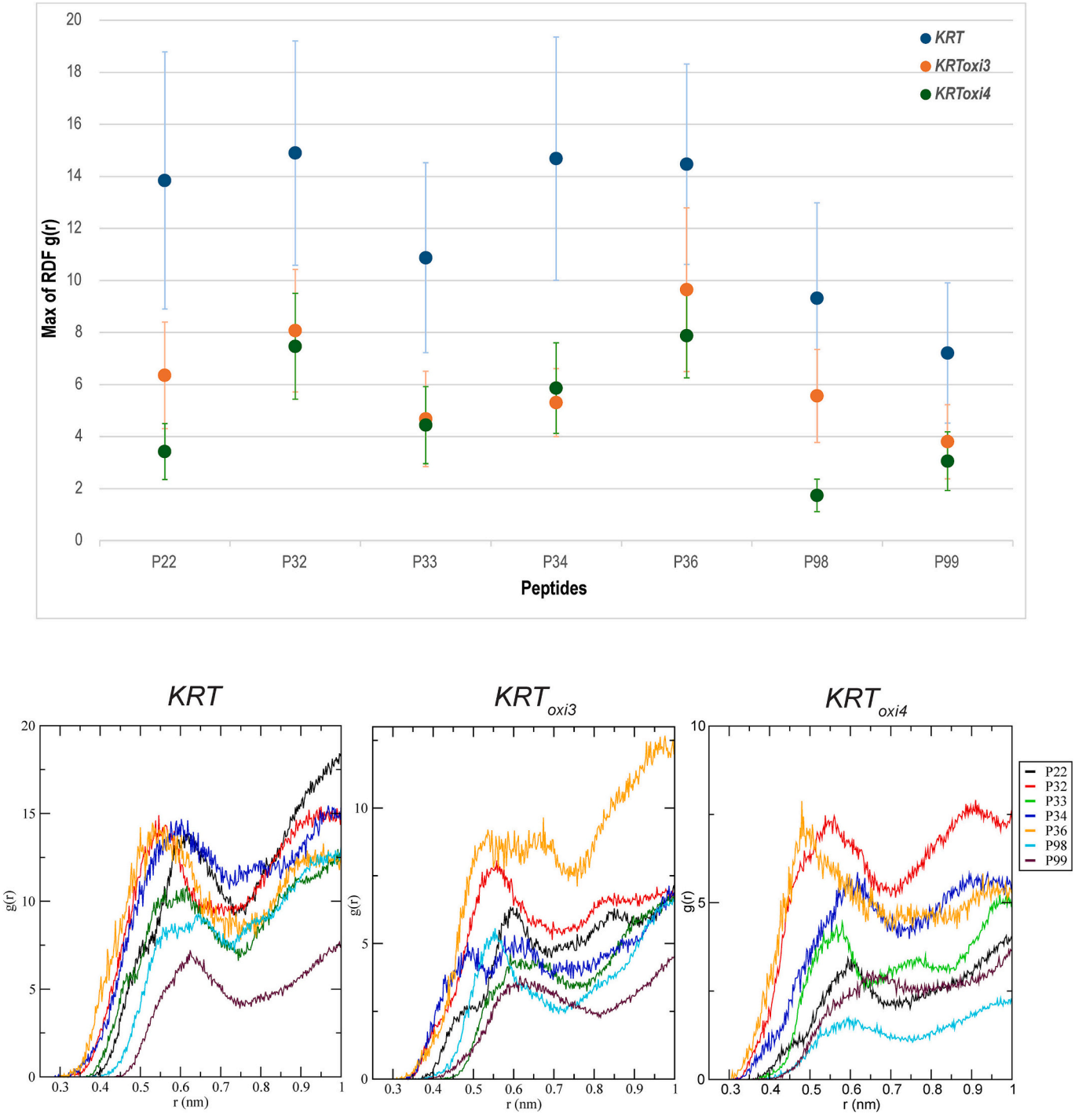


Fig. 5. Maximum and Radial Distribution Function (RDF) of peptides on the vicinity of KRT, KRT$_{oxi3}$ and KRT$_{oxi4}$ surfaces.

Table 2
Peptides' profile described by their physicochemical properties and amino acid types.

Peptide	aAA	bAA	charge	aroAA	hAA	OH	Gly	pI	GRAVY	MW (g/mol)
22	0	0	0	0	13	0	10	5.52	1.174	1883.22
32	0	0	0	0	10	5	15	5.52	0.267	2019.97
33	2	2	0	0	7	3	10	6.14	−1.03	2871.12
34	1	1	0	1	3	2	6	5.84	−1.213	1348.35
36	0	0	0	1	2	1	7	5.52	0.11	782.83
98	0	15	15	0	15	0	0	11.15	−1.05	3006.81
99	0	15	15	0	15	0	0	11.15	−0.05	3638.02

aAA = acidic amino acids; bAA = basic amino acids; charge = peptide's net charge; aroAA = aromatic amino acids; hAA = hydrophobic amino acids; OH = number of OH groups; Gly = number of Gly residues; pI = isoelectric point; GRAVY = Grand average of hydropathicity index; MW = Molecular Weight.

deduce that these factors might be the prime characteristics that lead to a better interaction with the native and oxidized keratins. To support this statement, it is necessary to look again at the types of amino acids and the physicochemical properties of each peptide. In this sense, Table 2 is now presented in the form of a histogram combined with a Principal Components Analysis (PCA) which relates these characteristics to the binding score on each keratin target (Fig. 6).

As previously mentioned, a thorough survey of several physicochemical properties of the tested peptides was conducted via ProtParam from Expasy [40] (Table 2). These variables were input into a dataset and normalized through the scaling of each property's value into a uniform range (from zero to 1), to ensure comparability across the entire range of variables (Table S4). Following this rationale, Fig. 6 gives a more summarized representation of the intrinsic properties of all sequences. The chart allows a visual representation of these, and together with PCA, infers the most promising peptides (P32, P33, P34 and P36) regarding the previous affinity data (Score, H-bonds and RDF). For example, the ones more prone to binding and maintaining interactions and stability at the KRT surface are the ones with more alcohol moieties, similar molecular weights, and similarly-sized sequences. P33 and P34 are the only peptides carrying acidic amino acids, which might be an important profile at low pH, as hair acidification is sometimes desirable to seal hair cuticles and lock in nutrients [57].

The overall picture points towards a diversity of characteristics influencing the binding. P33 is the most acidic, but at simulation pH (7.4), the acidic amino acids are in their negative form, which would impair a tightening binding to the oxidized KRTs. This characteristic is nonetheless balanced by its great variety of residues, relatively low hydrophobicity, and aromatic amino acids, assuring a very good score. P98 and P99 are markedly different from the rest, having high molecular weight, positively charged amino acids, and no Gly residues. They were, however, less prone to bind to KRT than P33 and P34 because their size didn't allow a good fitting. The pI is a variable that indicates that all peptides contribute different electrostatics towards their environment and thus interact differently. But then again, a marked difference is found on P98 and P99, which in turn return higher values.

Complementary, PCA allows the reduction of the dimensionality of the above data while increasing interpretability. This is achieved through the creation of new uncorrelated variables that successively maximize variance. Essentially, the analysis finds new variables, principal components, which are linear functions of those in the original dataset, that successively maximize variance whilst being uncorrelated with each other [58]. A PCA was generated to understand the relationship between the peptides' physicochemical properties and the binding scores obtained towards each keratin model.

PCA biplot is the final confirmation that P32, P34, and P36 are the ones with characteristics that result in a good binding on the keratin models, especially on the oxidized forms in the case of P34 and P36. On PCA, correlated variables present a small angle between vectors or an angle approaching 180°. For the mentioned peptides, the hydrophobic-hydrophilic balance (GRAVY) is important for the binding, as well as the presence of at least one aromatic residue for P34 and P36. On the other hand, the molecular weight, the number of basic amino acids, and the number of hydrophobic residues show to have less impact on the obtained scores (angles near 90°), which place the profiles of peptides P98 and P99 less proper for keratin binding. Of note, a low molecular weight will be preferable for the binding.

A visual representation of peptides' internalization or placement near keratin helices comes from clustering analysis, where the last 10 ns of simulation time was considered, to determine a middle structure amongst conformationally sampled peptides. In this sense, Fig. 7 shows the most representative structure of peptides P34 and P99 in interaction with oxidized keratin, for comparison of their placement and average distance to the keratin surface. While P34 is closely adsorbed into the KRT helices, P99 is placed at 14.8 Å (~1.49 nm) from the KRT surface, possibly indicating a repulsive mechanism involving Lys residues. Moreover, this explains the low RDF probability of P99 as well as the decreasing number of H-bonds during simulation.

The interaction and biocompatibility of some of our peptide classes and keratin have been studied experimentally, but here, the interaction patterns are disclosed in molecular detail. Silk-fibroin, for instance, was used in a combination blend with human hair keratin to generate wound-healing substrates [59]. Importantly, regarding silk peptide application in hair care, testing revealed effects on thickness, moisture, and absorbance rates. Also, good penetration into damaged hair was observed, followed by the recovery of hair cuticle [60,61]. The use of natural collagen was also investigated in hair styling and proved to form a protective film that closes the exposed cuticles and the cortex of the hair [62]. For elastin, high molecular weight peptides were tested and only demonstrated to have a secondary effect on hair against aggressive substances [63]. Importantly, resilin and abductin have been less explored for their application to hair fibers. These proteins or derived peptides are less commonly studied in the context of hair treatments compared to collagen or silk fibroin. Still, the present work sheds light on the role and potential of the P34 and P36 peptides (resilin-like and abductin-like) on hair treatment, paving the way for the exploitation of these two structural proteins in cosmetics.

4. Conclusions

The incorporation of functional peptides in cosmetics, particularly for hair care, has become increasingly popular due to their proven effectiveness. This study aimed to build on existing knowledge by exploring specific peptide sequences for their potential as KRT-binding motifs meant to stabilize and protect virgin, damaged, or oxidized hair.

Our findings demonstrate the promising role of artificial intelligence (AI) in the computational study of amino acid chain structures. Notably, AF 2.0 proved to be an invaluable tool, providing efficient and accurate predictions of protein and peptide conformations and making the study of unknown structures possible. However, AF 2.0 failed to simultaneously predict the keratin multimer structure and its interaction with a given peptide, i.e., it was unable to predict the peptide-keratin complex

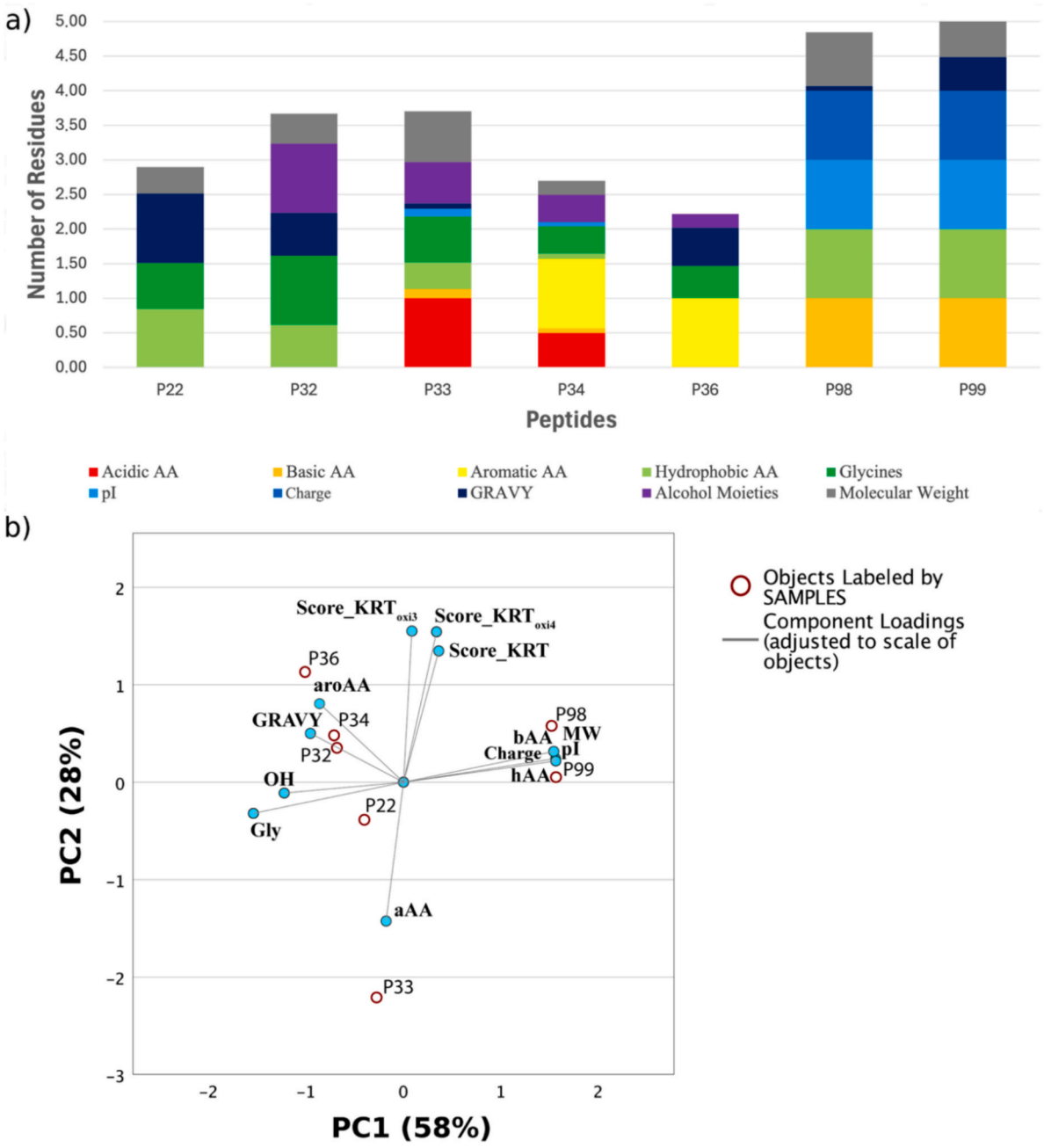


Fig. 6. a) Normalization of the physicochemical parameters from all seven studied peptides including acidic amino acids (aAA) basic amino acids (bAA), net charge (charge), aromatic amino acids (aroAA), hydrophobic amino acids (hAA), number of OH groups (OH), number of glycine residues (Gly), isoelectric point (pI), grand average of hydropathy index (GRAVY), and molecular weight (MW). b) Principal Component Analysis (PCA) biplot depicting the relationship between peptides' properties and binding score.

Table 3
Docking scores and Ligand Efficiency (LE) results for all seven studied peptides to *KRT* models obtained with HDOCK, followed by the number of H-bond interactions between the peptide-KRT pairs for the docking first pose, tracked with PyMOL.

		KRT			*KRT_{oxi3}*			*KRT_{oxi4}*		*Std Dev*
Peptides	*Score*	*LE*	*H-bond*	*Score*	*LE*	*H-bond*	*Score*	*LE*	*H-bond*	*Score*
22	−198.0	−8.6	5.0	−174.1	−7.6	6.0	−179.8	−7.8	5.0	12.5
32	−168.1	−5.6	9.0	−161.5	−5.4	8.0	−174.1	−5.8	10.0	6.3
33	−197.7	−6.6	11.0	−223.1	−7.4	9.0	−198.3	−6.6	9.0	14.5
34	−178.0	−11.9	6.0	−172.7	−11.5	9.0	−178.5	−11.9	7.0	3.2
36	−137.8	−13.8	5.0	−156.6	−15.7	3.0	−141.7	−14.2	4.0	9.9
98	−161.4	−5.4	7.0	−146.8	−4.9	3.0	−149.2	−5.0	7.0	7.8
99	−179.5	−6.0	5.0	−178.3	−5.9	7.0	−167.3	−5.6	5.0	6.7

LE = *score/number of peptide's residues.*

Table 4
Average results from RMSD analysis for each peptide-KRT complex, from MD simulations using HDOCK prediction.

Peptides	Average RMSD (nm)		Average RMSD (nm)		Average RMSD (nm)	
	KRT	*peptide*	*KRT_{oxi3}*	*peptide*	*KRT_{oxi4}*	*peptide*
22	0.283 ± 0.027	0.300 ± 0.037	0.342 ± 0.326	0.437 ± 0.071	0.301 ± 0.037	0.506 ± 0.094
32	0.310 ± 0.049	0.767 ± 0.141	2.232 ± 2.113	1.177 ± 0.123	0.377 ± 0.454	0.940 ± 0.125
33	0.920 ± 1.630	0.520 ± 0.046	1.304 ± 1.761	0.846 ± 0.106	0.706 ± 0.863	0.700 ± 0.174
34	0.341 ± 0.069	0.536 ± 0.056	0.888 ± 1.212	0.625 ± 0.081	0.873 ± 1.040	0.658 ± 0.094
36	0.281 ± 0.035	0.461 ± 0.088	0.757 ± 1.020	0.545 ± 0.127	1.953 ± 3.042	0.562 ± 0.135
98	0.717 ± 1.188	0.275 ± 0.031	4.435 ± 2.365	0.326 ± 0.052	5.106 ± 4.129	0.311 ± 0.079
99	0.287 ± 0.032	0.241 ± 0.049	1.261 ± 1.409	0.476 ± 0.100	3.011 ± 2.790	0.322 ± 0.045

Table 5
Average number of hydrogen bonds (H-BOND) observed throughout simulation time for the peptide-KRT complexes.

Peptides	Average no. of H-bonds		
	KRT	*KRT_{oxi3}*	*KRT_{oxi4}*
22	5.16 ± 1.23	4.05 ± 1.66	1.57 ± 1.02
32	14.32 ± 3.69	15.89 ± 3.50	19.45 ± 3.64
33	8.55 ± 1.75	6.30 ± 1.97	12.65 ± 4.46
34	9.01 ± 2.60	7.72 ± 2.56	6.62 ± 2.35
36	5.22 ± 1.77	5.92 ± 1.97	4.31 ± 1.68
98	7.14 ± 2.30	4.81 ± 2.19	1.61 ± 1.30
99	2.84 ± 1.56	4.19 ± 1.95	2.65 ± 1.49

so that the tetramer remained intact. Thus, the computational protocol employed included protein-protein interaction analysis through the use of a web-based docking tool, which generated reliable results despite known limitations. HDOCK presents limited flexibility handling, being less effective for highly flexible peptides or large conformational changes. However, this server is a fast and automated tool, featuring a knowledge-based scoring function, which improves prediction reliability. In addition, simulation analysis supported many of the predicted peptide conformations and docking sites, validating the computational approach.

Amongst the peptide sequences studied, P34 and P36 stand out as the most promising candidates for KRT binding. Led by their size and amino acid types, these peptides were the ones that internalized best between the keratin helices, with good binding efficiency and many hydrogen bonds and polar interactions established in dynamic conditions. These peptides contain motifs related to resilin and abductin proteins, respectively, and demonstrate potential for stabilizing KRT and improving hair health. The resilin motif is rich in proline and glycine residues, which have been associated with enhanced elasticity, flexibility and resilience [64]. Moreover, P36, being rich in glycines, even without the presence of prolines, with such a small size and good internalization, is also a promising presence within the vicinity of KRT as an effective quencher of potential damage.

Amongst the larger peptides, P32, a silk-fibroin inspired peptide, proved to consolidate its interactions during simulation with a remarkable RDF and number of hydrogen bonds. Still, P22, and P33 had a balanced combination of binding score, RDF, and H-bonds values.

Furthermore, Solfarcos' intellectual property surrounding the existent KP::KRT binding offers an excellent platform for further optimization. By incorporating elements from the identified peptides into fusion sequences, it may be possible to enhance the performance of this already effective peptide and develop even more potent solutions for hair care and repair.

Molecular modelling techniques for cosmetic applications are less commonly used or publicly reported, and here, we have combined the innovation of AI with docking and molecular dynamics simulations to streamline the research process for new ingredients. Noteworthy, this is the first in silico study to include and compare all of these structural peptide classes at the same time - abductin, collagen, elastin, resilin and silk fibroin - on the interaction with hair keratin.

CRediT authorship contribution statement

Tiago Ferreira: Writing – original draft, Methodology, Formal analysis. **Francisco Pereira:** Writing – original draft, Formal analysis. **Filipa Gonçalves:** Writing – review & editing, Validation. **André da Costa:** Writing – review & editing, Validation. **Artur Cavaco-Paulo:** Supervision, Funding acquisition, Conceptualization. **Tarsila G. Castro:** Writing – review & editing, Supervision, Conceptualization.

Funding

Cavaco-Paulo, Ferreira, and Pereira thank FCT under the scope of the strategic funding of UIDB/04469/2020 unit, and by LABBELS –

Table 6
H-bonds Occupancy (%) of the most persistent hydrogen bonds throughout 3000 frames. Letters A-D correspond to residues on the keratin monomers, while E indicates amino acids on the peptides.

Peptide	H-bond pairs	Occupancy (%)	H-bond pairs	Occupancy (%)	H-bond pairs	Occupancy (%)
	KRT		KRT_{oxi3}		KRT_{oxi4}	
32	Arg98(C) - Ser6(E)	69.41	Gln78(B) - Ala20(E)	64.85	Arg17(B) - Ala8(E)	67.54
	Arg98(C) - Gly23(E)	58.6	Glu24(A) - Ser6(E)	62.65	Glu11(D) - Ser30(E)	57.78
	Arg98(C) - Gly7(E)	55.95	Arg70(B) - Ser24(E)	58.15	Arg75(A) - Gly11(E)	57.31
	Glu93(A) - Ser24(E)	50.92	Ser64(B) - Ala4(E)	56.43	Arg75(A) - Ser12(E)	56.18
	Arg98(C) - Ser30(E)	49.22	Arg17(A) - Ala10(E)	49.11	Asp91(C) - Ser18(E)	55.55
34	Gln78(B) - Gly2(E)	78.24	Arg17(B) - Gly1(E)	46.72	Thr5(C) - Ser7(E)	22.63
	Tyr97(B) - Asn15(E)	61.41	Arg17(B) - Gly13(E)	45.12	Arg1(C) - Ser5(E)	21.35
	Gln10(A) - Asp6(E)	52.52	Arg22(D) - Ser5(E)	43.75	Arg98(B) - Asn15(E)	19.49
	Arg1(C) - Gly13(E)	49.62	Arg34(B) - Tyr8(E)	42.15	Leu101(B) - Gly14(E)	18.49
	Asp91(D) - Arg3(E)	39.05	Arg17(B) - Gly2(E)	42.02	His2(C) - Asn15(E)	17.43
36	Arg75(A) - Ser10(E)	29.29	Gln71(D) - Ser10(E)	44.69	Arg98(C) - Gly5(E)	26.36
	Arg70(A) - Gly2(E)	28.56	Lys68(D) - Ser10(E)	37.84	Arg98(C) - Met6(E)	24.59
	Arg75(A) - Phe3(E)	26.36	Arg56(D) - Gly5(E)	33.92	Arg98(C) - Ser10(E)	16.96
	Arg70(A) - Phe3(E)	26.06	Arg56(D) - Gly4(E)	30.96	Arg98(C) - Gly7(E)	16.63
	Arg75(A) - Gly2(E)	25.95	Tyr48(A) - Gly1(E)	30.42	Arg99(C) - Gly1(E)	15.13

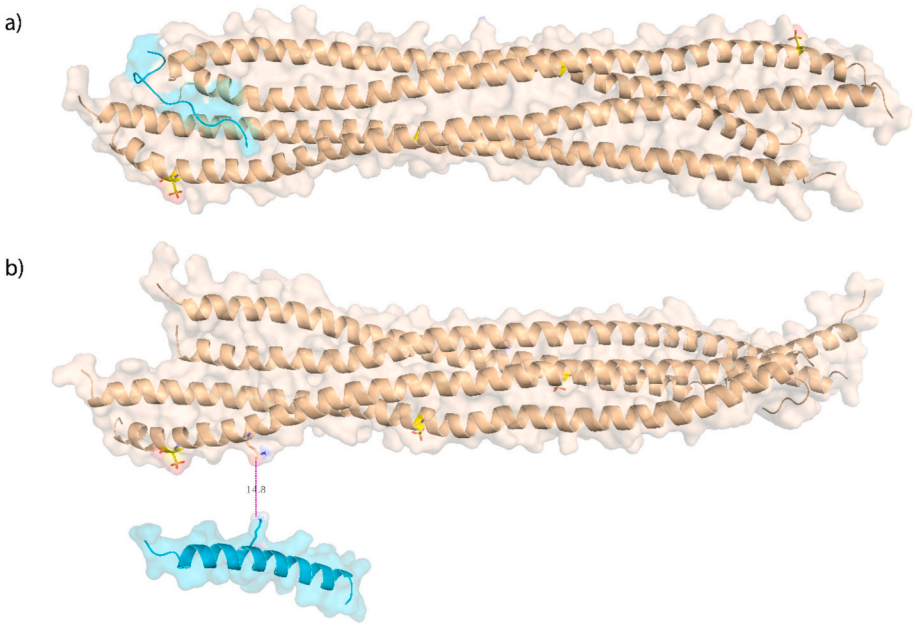


Fig. 7. a) Clustering middle structure of P34 well-embedded into keratin helices, and **(b)** the middle structure of peptide P99, where detachment occurs due to lysine repulsion. These representative frames were captured from peptides' interactions towards oxidized KRT_{oxi4}, where CYA are highlighted in yellow.

Associate Laboratory in Biotechnology, Bioengineering and Microelectromechanical Systems, LA/P/0029/2020. Ferreira thanks the SOLFARCOS' PhD grant which supports him as PhD student of the Doctoral Program in Chemical and Biological Engineering at the Centre of Biological Engineering of the University of Minho. Pereira thanks the grant CEB-BI-06-2024/UMINHO/BIM-CNCG/2024/72.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Artur Cavaco-Paulo has patent Fusion Proteins/Peptides, Methods and uses Thereof issued to WO/2024/166050. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Ferreira, Pereira and Castro greatly thank the computational infrastructure provided by FCT/INCD CIRRUS HPC through the project 2023.10823.CPCA.A2, which granted additional computational infrastructure for this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2025.144813.

Data availability

Data will be made available on request.

References

[1] C. Fernandes, B. Medronho, L. Alves, M.G. Rasteiro, On hair care Physicochemistry: from structure and degradation to novel biobased conditioning agents, Polymers 15 (2023) 608, https://doi.org/10.3390/polym15030608.
[2] N. Sadgrove, S. Batra, D. Barreto, J. Rapaport, An Updated Etiology of Hair Loss and the New Cosmeceutical Paradigm in Therapy: Clearing 'the Big Eight Strikes', Cosmetics 10 (2023) (2023) 106 https://doi.org/10.3390/cosmetics10040106.
[3] M. Ash, M. Zibitt, O. Shauly, A. Menon, A. Losken, D. Gould, The innovative and evolving landscape of topical exosome and peptide therapies: a systematic review of the available literature, Aesthet. Surg. J. Open Forum 6 (2024) ojae017, https://doi.org/10.1093/asjof/ojae017.
[4] C.F. Cruz, N.G. Azoia, T. Matamá, A. Cavaco-Paulo, Peptide-protein interactions within human hair keratins, Int. J. Biol. Macromol. 101 (2017) 805–814, https://doi.org/10.1016/j.ijbiomac.2017.03.052.
[5] N. Reihani, H. Ghourchian, Synergy of keratin peptides and natural oils embedded in nanostructured lipid carriers to protect and restore human hair structure, J. Mol. Liq. (2025) 126957, https://doi.org/10.1016/j.molliq.2025.126957.
[6] A. Tinoco, M. Martins, A. Cavaco-Paulo, A. Ribeiro, Biotechnology of functional proteins and peptides for hair cosmetic formulations, Trends Biotechnol. 40 (2022) 591–605, https://doi.org/10.1016/j.tibtech.2021.09.010.
[7] M.M. Fernandes, C.F. Lima, A. Loureiro, A.C. Gomes, A. Cavaco-Paulo, Keratin-based peptide: biological evaluation and strengthening properties on relaxed hair: hair strengthening with keratin peptide, Int. J. Cosmet. Sci. 34 (2012) 338–346, https://doi.org/10.1111/j.1468-2494.2012.00727.x.
[8] E. Antunes, C.F. Cruz, N.G. Azoia, A. Cavaco-Paulo, The effects of solvent composition on the affinity of a peptide towards hair keratin: experimental and molecular dynamics data, RSC Adv. 5 (2015) 12365–12371, https://doi.org/10.1039/C4RA13901A.
[9] N.G. Azoia, M.M. Fernandes, N.M. Micaêlo, C.M. Soares, A. Cavaco-Paulo, Molecular modeling of hair keratin/peptide complex: using MM-PBSA calculations to describe experimental binding results, Proteins Struct. Funct. Bioinforma. 80 (2012) 1409–1417, https://doi.org/10.1002/prot.24037.
[10] C. Robbins, Chemical and physical behavior of human hair, 2012, https://doi.org/10.1007/978-3-642-25611-0_2.
[11] T. Hirai, M. Ikeda-Imafuku, N. Tasaka, V.T.G. Chuang, M. Xian, T. Ishida, T. Akaike, Y. Ishima, Human hair keratin responds to oxidative stress via reactive sulfur and supersulfides, Adv. Redox Res. 10 (2024) 100091, https://doi.org/10.1016/j.arres.2023.100091.
[12] F.-C. Yang, Y. Zhang, M.C. Rheinstädter, The structure of people's hair, PeerJ 2 (2014) e619, https://doi.org/10.7717/peerj.619.
[13] K. Watanabe, K. Nagami, K. Suzuta, T. Maeda, L. Ito, Cysteic acid formation behaviors in bleached hair of southeast Asian characterized by infrared spectroscopy, Adv. Life Sci. 5 (4) (2015) 85–89, https://doi.org/10.5923/j.als.20150504.02.
[14] M.F. Gavazzoni Dias, Hair cosmetics: an overview, Int. J. Trichology 7 (2015) 2, https://doi.org/10.4103/0974-7753.153450.
[15] T.C. de Sá Dias, A.R. Baby, T.M. Kaneko, M.V. Robles Velasco, Relaxing/straightening of afro-ethnic hair: historical overview, J. Cosmet. Dermatol. 6 (2007) 2–5, https://doi.org/10.1111/j.1473-2165.2007.00294.x.
[16] M. Di Foggia, P. Taddei, C. Boga, B. Nocentini, G. Micheletti, Interactions between damaged hair keratin and Juglone as a possible restoring agent: a vibrational and scanning Electron microscopy study, Molecules 29 (2024) 320, https://doi.org/10.3390/molecules29020320.
[17] M.M. Fernandes, A.C. Gomes, A. Vasconcelos, F.-D. Munteanu, T. Tzanov, M.S. T. Gonçalves, N. End, K.-U. Schoening, G.M. Guebitz, A. Cavaco-Paulo, Protein disulphide isomerase-assisted functionalization of keratin-based matrices, Appl. Microbiol. Biotechnol. 90 (2011) 1311–1321, https://doi.org/10.1007/s00253-011-3194-6.

[18] C.F. Cruz, M. Martins, J. Egipto, H. Osório, A. Ribeiro, A. Cavaco-Paulo, Changing the shape of hair with keratin peptides, RSC Adv. 7 (2017) 51581–51592, https://doi.org/10.1039/C7RA10461H.
[19] Solfarcos, FUSION PROTEINS/PEPTIDES, METHODS AND USES THEREOF, WO/2024/166050, n.d.
[20] A. Miserez, J. Yu, P. Mohammadi, Protein-based biological materials: molecular design and artificial production, Chem. Rev. 123 (2023) 2049–2111, https://doi.org/10.1021/acs.chemrev.2c00621.
[21] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A. A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589, https://doi.org/10.1038/s41586-021-03819-2.
[22] E. Antunes, C.F. Cruz, N.G. Azoia, A. Cavaco-Paulo, Insights on the mechanical behavior of keratin fibrils, Int. J. Biol. Macromol. 89 (2016) 477–483, https://doi.org/10.1016/j.ijbiomac.2016.05.018.
[23] J. Skolnick, M. Gao, H. Zhou, S. Singh, AlphaFold 2: why it works and its implications for understanding the relationships of protein sequence, structure, and function, J. Chem. Inf. Model. 61 (2021) 4827–4831, https://doi.org/10.1021/acs.jcim.1c01114.
[24] L.M.F. Bertoline, A.N. Lima, J.E. Krieger, S.K. Teixeira, Before and after AlphaFold2: an overview of protein structure prediction, Front. Bioinforma. 3 (2023) 1120370, https://doi.org/10.3389/fbinf.2023.1120370.
[25] A.J. Grosvenor, S. Deb-Choudhury, P.G. Middlewood, A. Thomas, E. Lee, J. A. Vernon, J.L. Woods, C. Taylor, F.I. Bell, S. Clerens, The physical and chemical disruption of human hair after bleaching – studies by transmission electron microscopy and redox proteomics, Int. J. Cosmet. Sci. 40 (2018) 536–548, https://doi.org/10.1111/ics.12495.
[26] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L.E. Felberg, D.H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D.W. Gohara, T. Dolinsky, R. Konecny, D.R. Koes, J.E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G.-W. Wei, M. J. Holst, J.A. McCammon, N.A. Baker, Improvements to the APBS biomolecular solvation software suite, Protein Sci. 27 (2018) 112–128, https://doi.org/10.1002/pro.3280.
[27] F. Madeira, N. Madhusoodanan, J. Lee, A. Eusebi, A. Niewielska, A.R.N. Tivey, R. Lopez, S. Butcher, The EMBL-EBI job dispatcher sequence analysis tools framework in 2024, Nucleic Acids Res. 52 (2024) W521–W525, https://doi.org/10.1093/nar/gkae241.
[28] M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G.R. Lee, J. Wang, Q. Cong, L.N. Kinch, R.D. Schaeffer, C. Millán, H. Park, C. Adams, C. R. Glassman, A. DeGiovanni, J.H. Pereira, A.V. Rodrigues, A.A. van Dijk, A. C. Ebrecht, D.J. Opperman, T. Sagmeister, C. Buhlheller, T. Pavkov-Keller, M. K. Rathinaswamy, U. Dalwadi, C.K. Yip, J.E. Burke, K.C. Garcia, N.V. Grishin, P. D. Adams, R.J. Read, D. Baker, Accurate prediction of protein structures and interactions using a three-track neural network, Science 373 (2021) 871–876, https://doi.org/10.1126/science.abj8754.
[29] H. Bekker, H. Berendsen, E. Dijkstra, S. Achterop, R. Vondrumen, D. Vanderspoel, A. Sijbers, H. Keegstra, B. Reitsma, M. Renardus, Gromacs - a parallel computer for molecular-dynamics simulations, in: R.A. DeGroot, J. Nadrchal (Eds.), Phys. Comput. 92, World Scientific Publ Co Pte Ltd, Singapore, 1993, pp. 252–256. https://www.webofscience.com/wos/woscc/full-record/WOS:A1993BA17B00037. (Accessed 17 April 2024).
[30] GROMACS 2021.2 Release Notes — GROMACS 2021.2 Documentation, (n.d.). https://manual.gromacs.org/2021.2/release-notes/2021/2021.2.html (accessed April 17, 2024).
[31] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007) 014101, https://doi.org/10.1063/1.2408420.
[32] G. Bussi, T. Zykova-Timan, M. Parrinello, Isothermal-isobaric molecular dynamics using stochastic velocity rescaling, J. Chem. Phys. 130 (2009) 074101, https://doi.org/10.1063/1.3073889.
[33] R. Martoňák, A. Laio, M. Parrinello, Predicting crystal structures: the Parrinello-Rahman method revisited, Phys. Rev. Lett. 90 (2003) 075503, https://doi.org/10.1103/PhysRevLett.90.075503.
[34] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092, https://doi.org/10.1063/1.464397.
[35] Y. Yan, D. Zhang, P. Zhou, B. Li, S.-Y. Huang, HDOCK: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy, Nucleic Acids Res. 45 (2017) W365–W373, https://doi.org/10.1093/nar/gkx407.
[36] Y. Yan, H. Tao, J. He, S.-Y. Huang, The HDOCK server for integrated protein–protein docking, Nat. Protoc. 15 (2020) 1829–1852, https://doi.org/10.1038/s41596-020-0312-x.
[37] R.A. Laskowski, M.B. Swindells, Ligplot+: multiple ligand–protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786, https://doi.org/10.1021/ci200227u.
[38] A.T.N. Vo, M.A. Murphy, T.W. Stone, P.K. Phan, M.I. Baskes, R.K. Prabhu, Molecular dynamics simulations of phospholipid bilayer mechanoporation under different strain states—a comparison between GROMACS and LAMMPS, Model. Simul. Mater. Sci. Eng. 29 (2021) 055015, https://doi.org/10.1088/1361-651X/abfeaf.
[39] IBM SPSS Statistics. 2024. https://www.ibm.com/products/spss-statistics (accessed November 6, 2024).
[40] Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A. Protein identification and analysis tools on the ExPASy server, in: Walker, J.M. (Ed.), Proteomics Protoc. Handb., Humana Press, Totowa, NJ, 2005: pp. 571–607. https://doi.org/10.1385/1-59259-890-0:571.
[41] N.R. Barthélemy, A. Bednarczyk, C. Schaeffer-Reiss, D. Jullien, A. Van Dorsselaer, N. Cavusoglu, Proteomic tools for the investigation of human hair structural proteins and evidence of weakness sites on hair keratin coil segments, Anal. Biochem. 421 (2012) 43–55, https://doi.org/10.1016/j.ab.2011.10.011.
[42] S.A. Eldirany, M. Ho, A.J. Hinbest, I.B. Lomakin, C.G. Bunick, Human keratin 1/10-1B tetramer structures reveal a knob-pocket mechanism in intermediate filament assembly, EMBO J. 38 (2019) e100741, https://doi.org/10.15252/embj.2018100741.
[43] UniProt O76011, UniProt · KRT34HUMAN (n.d.). https://www.uniprot.org/uniprotkb/O76011/entry (accessed May 8, 2025).
[44] UniProt O43790, UniProt (n.d.). https://www.uniprot.org/uniprotkb/O43790/entry (accessed May 7, 2025).
[45] H.-J. Jin, D.L. Kaplan, Mechanism of silk processing in insects and spiders, Nature 424 (2003) 1057–1061, https://doi.org/10.1038/nature01809.
[46] A. Motta, L. Fambri, C. Migliaresi, Regenerated silk fibroin films: thermal and dynamic mechanical analysis, Macromol. Chem. Phys. 203 (2002) 1658–1665, https://doi.org/10.1002/1521-3935(200207)203:10/11<1658::AID-MACP1658>3.0.CO;2-3.
[47] G.H. Altman, F. Diaz, C. Jakuba, T. Calabro, R.L. Horan, J. Chen, H. Lu, J. Richmond, D.L. Kaplan, Silk-based biomaterials, Biomaterials 24 (2003) 401–416, https://doi.org/10.1016/S0142-9612(02)00353-8.
[48] W. DAAMEN, J. VEERKAMP, J. VANHEST, T. VANKUPPEVELT, Elastin as a biomaterial for tissue engineering, Biomaterials 28 (2007) 4378–4398, https://doi.org/10.1016/j.biomaterials.2007.06.025.
[49] C. Zhou, F. Confalonieri, M. Jacquet, R. Perasso, Z. Li, J. Janin, Silk fibroin: structural implications of a remarkable amino acid sequence, Proteins Struct. Funct. Bioinforma. 44 (2001) 119–122, https://doi.org/10.1002/prot.1078.
[50] K. Bailey, T. Weis-Fogh, Amino acid composition of a new rubber-like protein, resilin, Biochim. Biophys. Acta 48 (1961) 452–459, https://doi.org/10.1016/0006-3002(61)90043-9.
[51] T. Weis-Fogh, A rubber-like protein in insect cuticle, J. Exp. Biol. 37 (1960) 889–907, https://doi.org/10.1242/jeb.37.4.889.
[52] M. Burrows, S.R. Shaw, G.P. Sutton, Resilin and chitinous cuticle form a composite structure for energy storage in jumping by froghopper insects, BMC Biol. 6 (2008) 41, https://doi.org/10.1186/1741-7007-6-41.
[53] R.E. Kelly, R.V. Rice, Abductin: a rubber-like protein from the internal triangular hinge ligament of Pecten, Science 155 (1967) 208–210, https://doi.org/10.1126/science.155.3759.208.
[55] Q. Cao, Y. Wang, H. Bayley, Sequence of abductin, the molluscan 'rubber' protein, Curr. Biol. 7 (1997) R677–R678, https://doi.org/10.1016/S0960-9822(06)00353-8.
[56] B. Bochicchio, A. Pepe, A.M. Tamburro, Circular dichroism studies on repeating polypeptide sequences of abductin, Chirality 17 (2005) 364–372, https://doi.org/10.1002/chir.20172.
[57] M.F.R. Gavazzoni Dias, A.M. de Almeida, P.M.R. Cecato, A.R. Adriano, J. Pichler, The shampoo pH can affect the hair: myth or reality? Int. J. Trichology 6 (2014) 95–99, https://doi.org/10.4103/0974-7753.139078.
[58] I.T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 374 (2016) 20150202, https://doi.org/10.1098/rsta.2015.0202.
[59] N. Bhardwaj, W.T. Sow, D. Devi, K.W. Ng, B.B. Mandal, N.-J. Cho, Silk fibroin–keratin based 3D scaffolds as a dermal substitute for skin tissue engineering, Integr. Biol. 7 (2015) 53–63, https://doi.org/10.1039/c4ib00208c.
[60] P.J. Babu, L. Suamte, Applications of silk-based biomaterials in biomedicine and biotechnology, Eng. Regen. 5 (2024) 56–69, https://doi.org/10.1016/j.engreg.2023.11.002.
[61] Hyun, J.-W., Lee, K., Yeo, J., Choe, T. Hair care effects of hair cosmetics including low molecular weight silk peptide component and micro structure analysis, KSBB J. (2008). https://www.semanticscholar.org/paper/Hair-Care-effects-of-hair-cosmetics-including-Low-Hyun-Lee/e99b633c62b7701b0a6cf2e36296ff2ba552a537 (accessed March 25, 2025).
[62] J. Igielska-Kalwat, E. Kilian-Pięta, S. Połoczańska-Godek, The use of natural collagen obtained from fish waste in hair styling and care, Polymers 14 (2022) 749, https://doi.org/10.3390/polym14040749.
[63] F. Langmaier, M. Mládek, K. Kolomazník, S. Sukop, Isolation of elastin and collagen polypeptides from long cattle tendons as raw material for the cosmetic industry, Int. J. Cosmet. Sci. 24 (2002) 273–279, https://doi.org/10.1046/j.1467-2494.2002.00151.x.
[64] N.C. Abascal, L. Regan, The past, present and future of protein-based materials, Open Biol. 8 (2018) 180113, https://doi.org/10.1098/rsob.180113.