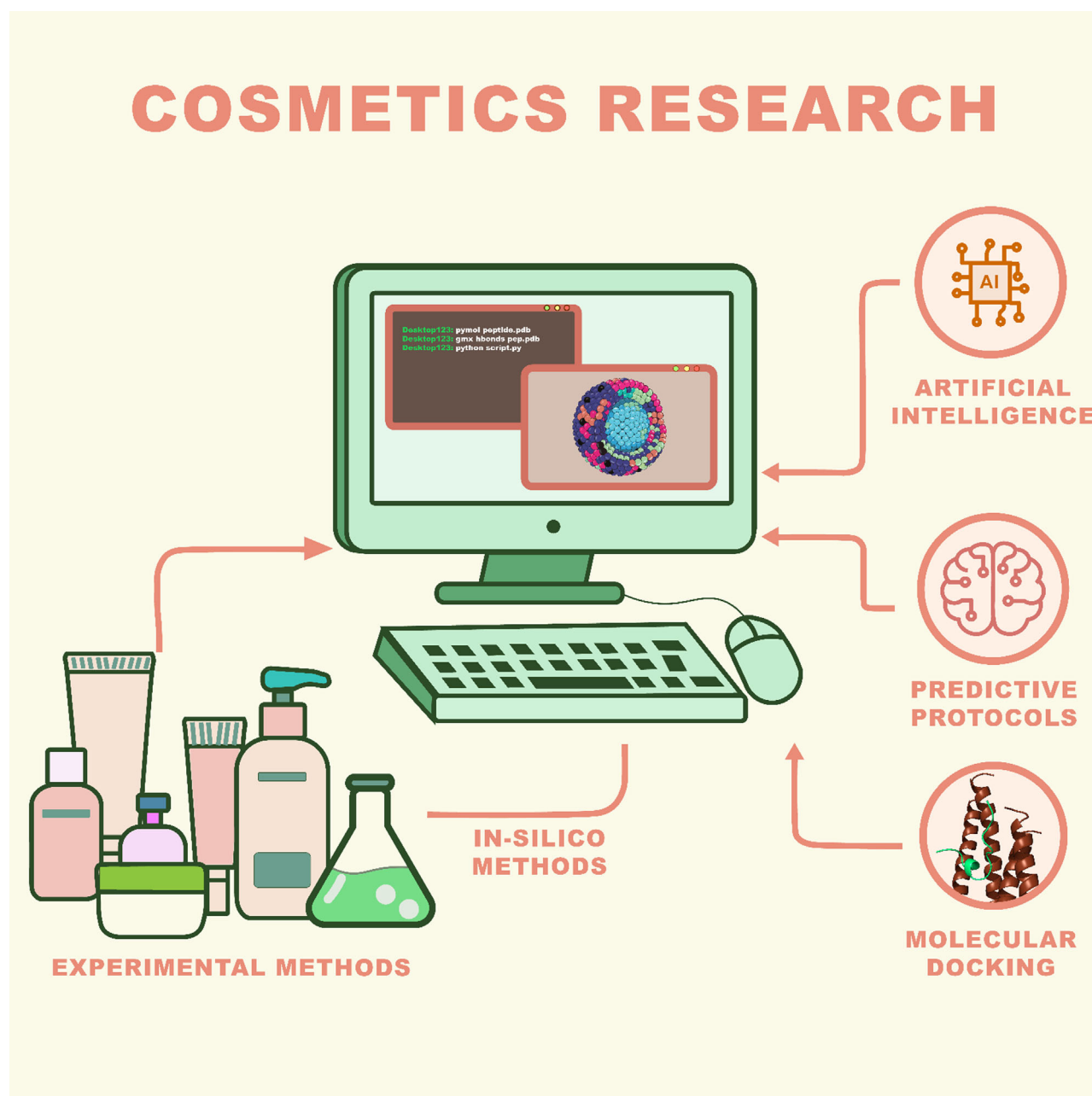


Artificial Intelligence, Molecular Dynamics, and Beyond: Computational Insights In Cosmetics Research and Formulation Design

Tiago Ferreira, Artur Cavaco-Paulo,* and Tarsila G. Castro*



The vast field of cosmetics is mainly explored through experimental methods, while computational tools find broader application in structural biology. The world of formulations remains relatively untouched or nondisclosed due to commercial interests. This work aims to better understand the applications of *in silico* methods in cosmetics research, at the molecular level, and how the recent work is advancing the state of the art. The integration

of AI, molecular docking, and molecular dynamics simulations in cosmetics research enables the efficient discovery and optimization of ingredients, leading to innovative and effective cosmetic products. Importantly, the establishment of predictive computational protocols suitable for different cosmetic applications optimizes experimental processes and simplify testing, being economically important for the industry.

1. Introduction

Research in the field of cosmetic formulations is largely dominated by traditional experimental methods used to understand properties such as ingredient interactions and organoleptic characteristics, as well as to optimize product stability and ensure product efficacy. Techniques such as rheological assessments, electron microscopy, and spectroscopy methods have been extensively employed to study the physical characteristics of emulsions, surfactants, and bioactive ingredients.^[1–5] While these techniques provide valuable insights into formulation behavior, their ability to predict interactions on a molecular scale remains limited.

The cosmetics industry is a rapidly growing sector that continuously innovates in response to consumer demand for products that are effective, safe, and environmentally friendly. This requires R&D processes to include more efficient and ecofriendly methods to also provide a less wasteful environment within scientific research.

Advancements in the field of computational chemistry have significantly expanded the scope of research into various types of liquid formulations and product design, which in turn opens the pathway to studying cosmetic formulations.^[6–8] Molecular dynamics (MD) simulations, molecular docking, and artificial intelligence (AI) models are becoming increasingly vital in the study of ingredient behavior and formulation design, as demonstrated by several recent studies.^[9,10] When combined with experimental approaches, these methods provide detailed insights into molecular interactions within complex formulations, facilitating more efficient development cycles and enhancing product performance. Nonetheless, optimization strategies, mainly mathematically model-based, but moving toward broader and more

integrative approaches, also have a place and a say in product formulation.^[11,12]

In the past years, we have developed a computational protocol to study formulations, both simple and complex (multiingredient types), through MD simulations. More recently, we used AI and Deep Learning (DL), via AlphaFold, to predict the structure of human hair keratin, followed by molecular docking experiments to study the interaction of bioinspired peptides with hair. Thus, our main goal is to pave the way for the generation of important molecular details using atomistic methods, although other methods and strategies are also briefly considered here.

According to the Web of Science (**Figure 1**), studies contemplating the integration of cosmetics and *in silico*/modeling techniques have been relatively limited over the past 20 years. However, it is an emerging field, with MD simulations, docking, and AI gaining increasing relevance in the past five years. Among these methods, molecular docking is the most reported and cited. Still, the search also reveals that the keyword “cosmetics” is often associated with fields such as dermatology and plastics. This suggests that the direct application of computational methods to purely cosmetic targets remains limited, especially for AI implementation.

In the following sections, we will review several studies and highlight their importance in the field of cosmetics formulations. Additionally, we will present a conceptual protocol to follow while using key computational tools to achieve highly accurate results. Importantly, significant contributions recently reported in this field will be reviewed and compared to our efforts.

2. Molecular Dynamics Simulations

By modeling the interactions between atoms and molecules over time, MD simulations can reveal the dynamic behavior of emulsifiers/surfactants, oils, emollients, stabilizers, and active ingredients in solutions.^[13] This is especially useful for studying emulsification or formulation stability. On the other hand, MD can also track the interaction of active ingredients with the skin's lipid barrier and/or hair keratin fibers, and the behavior of hydrophilic and lipophilic phases in cosmetic products.^[14,15] MD simulations enable the examination of molecular assemblies with varying complexity and provide insights into thermodynamic properties, phase transitions, and the structural stability of formulations. These capabilities help reduce the process, time and cost

T. Ferreira, A. Cavaco-Paulo
CEB—Center of Biological Engineering
University of Minho
Campus de Gualtar, 4710–057 Braga, Portugal
E-mail: artur@deb.uminho.pt

T. Ferreira, A. Cavaco-Paulo, T. G. Castro
Modelling Department
SOLFARCOS—Soluções Farmacêuticas e Cosméticas Lda
4700–034 Braga, Portugal
E-mail: tgc@solfarcos.com

A. Cavaco-Paulo
LABBELS—Associate Laboratory
University of Minho
4700–034 Braga/Guimarães, Portugal

of trial-and-error in formulation development by guiding experimental design. Next, a few case studies will highlight the employment of MD simulations in cosmetics research and discuss the advantages and limitations when combined or not with other in silico methods.

2.1. Replacing Silicone in Formulations

Silicones, particularly dimethicone, cyclopentasiloxane, and amodimethicone, are lauded for enhancing the sensory feel of formulations, reducing tackiness, and improving emulsion stability. Their nongreasy texture, water resistance, and ability to form smooth films make them essential in personal care products designed for long-lasting effects and high performance.^[16] However, their environmental persistence and increasing consumer demand for natural and sustainable products have led to growing scrutiny over their continued use. This has triggered research into silicone alternatives, especially as pressure mounts from regulatory bodies and ecoconscious consumers.

Replacing silicones in cosmetic formulations presents significant technical challenges. Silicones play critical roles, especially in surfactant systems, where they stabilize formulations and influence the foam, spreadability, and texture of emulsions.^[17,18] Surfactants, which are pivotal in both cleansing and conditioning products, work synergistically with silicones to form stable emulsions, especially in water-in-oil systems.^[19] Without either of these compounds, maintaining these desirable properties becomes challenging, as alternative ingredients often fail to replicate the same sensory and performance qualities of silicones or stabilize emulsions like surfactants.

Several silicone alternatives, such as plant-based oils, esters, and biodegradable polymers, are emerging as potential substitutes.^[20–22] However, these ingredients often fall short in

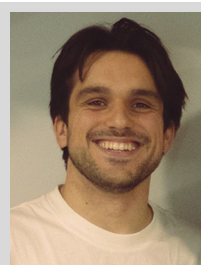
mimicking the high performance of silicones, particularly in terms of water resistance, sensory feel, and the creation of lightweight yet occlusive films on the skin or hair.

Furthermore, silicone alternatives can introduce formulation difficulties, such as increased greasiness, reduced spreadability, and poor stability, making it challenging to fully replicate the attributes that make silicones so effective.^[23] Nevertheless, the potential of a number of esters was demonstrated in a few works from our team, which evaluated them both experimentally and using modeling methods. Importantly, in some cases, the esters were produced by biosynthesis lipase-catalysis.^[24,25] Their chemical structures are presented in **Figure 2**, which compares them with common silicones found in cosmetics and commercial alternatives. All structures are highly flexible, with the ester having a hydrophobic chain of varied size.

In a first computational approach,^[26] Ferreira and collaborators proposed a protocol designed to enable the characterization and evaluation of silicone substitutes, specifically by elucidating their structural and interactive performance within formulations.

Central to this approach is the development of a novel equilibration protocol, which tests the alternatives under temperature stimuli, leveraging MD simulations to assess their distribution and densities, thereby providing insights into their behavior relative to traditional silicones (**Figure 3**).

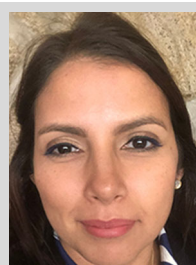
This methodology for tracking distribution and density behavior has already been applied in the study of other esters with experimentally verified emulsifying properties.^[24] The study provided valuable insights for optimizing formulations across various industries, including cosmetics, pharmaceuticals, and food science, through predictive modeling and tailored molecular design. This work emerges as a hybrid paper combining both MD techniques and wet lab studies.



Tiago Ferreira obtained his M.S. in bioinformatics in 2020 at the School of Engineering at the University of Minho and his bachelor's degree in applied biology at the same institution. The very same year, he enrolled in the Ph.D. in chemical and biological engineering, focusing on molecular dynamics and computational chemistry and its applications in cosmetics research and development.



Artur Cavaco-Paulo has an academic experience of over 35 years in chemistry and enzymology applied to textiles, with the last 20 years focused on biotechnology for cosmetics and pharmaceuticals. As a professor at the School of Engineering from the University of Minho, he has mentored more than 35 PostDoc researchers and over 50 Ph.D. students. His expertise in enzymatic processes and nanotechnology led to the founding of Solfarcos, a spin-off of University of Minho. Now a well-established company, Solfarcos specializes in developing biotechnology-driven solutions for cosmetic and pharmaceutical formulation.



Tarsila G. Castro has a Ph.D. in science since 2016, specializing in chemistry and molecular modeling. Her experience in computational chemistry began during her undergraduate studies in industrial chemistry in Brazil and was consolidated during her Master's studies in medicinal chemistry at the University of Minho in Portugal. After her Ph.D., she joined the Aurel Vlaicu University of Arad, Romania, as a postdoctoral researcher in molecular modeling. She is currently a researcher at Solfarcos LDA, as coordinator of the modeling team, and working on the modeling of bio ingredients for cosmetic formulations, and functional peptides/proteins.

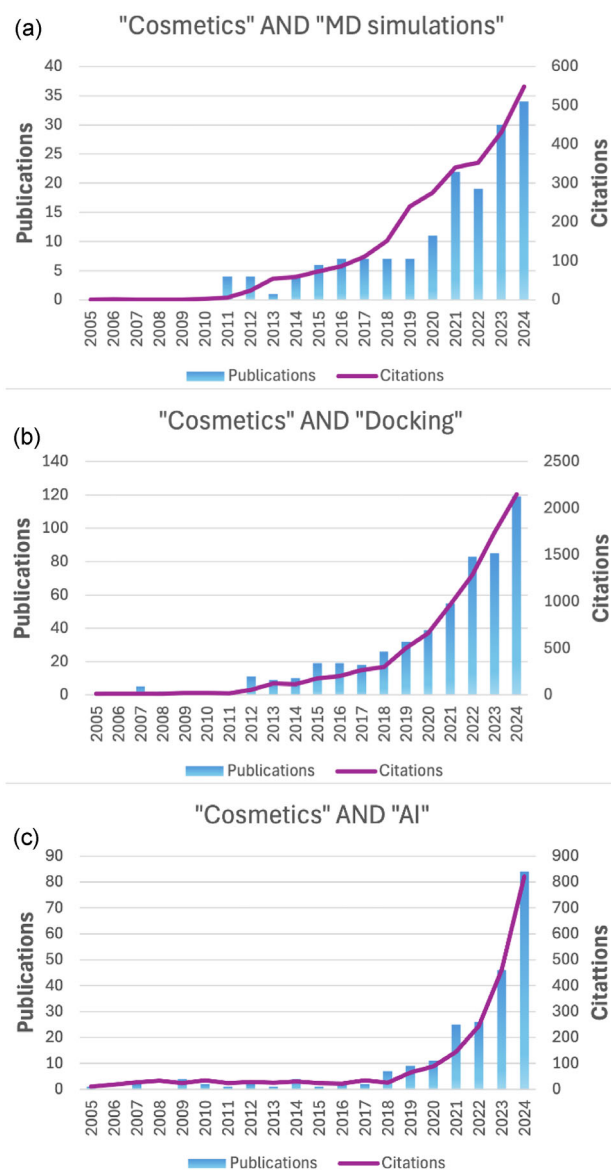


Figure 1. Web of Science number of publications and citations per year presenting the keywords a) "Cosmetics" and "Molecular Dynamics Simulations", b) "Cosmetics" and "Docking", and c) "Cosmetics" and "AI". Data included herein are derived from Clarivate (Web of Science). Copyright Clarivate 2025. All rights reserved.

It introduces five new alkyl esters (Figure 2c) that, besides the emulsifier character, are suitable for replacing synthetic silicones in resemblance to other esters. Through a similar computational protocol, it was observed that they were able to form micelles, and their overall structural stability through density distribution assessment was evaluated. Such esters can function both as emollients and as nonionic emulsifiers.^[24] Of note, this fact is a great advantage in product formulation: to design esters that, while serving as biocompatible emollients, can also act as an emulsifier or be able to aid emulsification, possibly replacing the need for silicones and surfactants simultaneously.^[27] Nonetheless, the obtention of esters that fully mimic the sensory features of silicones is still challenging, but alkyl esters are the most promising candidates.

Further computational protocol and analysis were applied to a more complex formulation.^[14] Here, novel silicone-like ester alternatives were synthesized and tested using a number of analytical methods to assess physicochemical properties (pH, density, conductivity, and refractive index), texture analysis (spreadability, back extrusion), stability under sunlight, stability under temperature stimuli, and rheological properties (viscosity).

The modeling protocol examined the dynamic behavior of emollient molecules, providing insights into the mechanisms of arrangement and interaction with the formulation's ingredients. Thus, the simulations assessed patterns of aggregation and the interaction of the system's components with the solvent. The degree of complexity of the formulation is also increased in comparison with the previous study, but simulations still prove to yield results comparable to the experimental data.

The MD simulations revealed that the amphiphilic esters spontaneously self-assembled into stable micelle-like structures (Figure 4), where the ester is placed at the water interface of such structures, along with octyldodecanol and Xanthan gum, whereas hydrophobic molecules such as triolein and ethylexylglycerin were found more internalized. Of note, Xanthan gum is a recognized and widely used additive for binding and stabilizing emulsions,^[28] thus, the external placement in interaction with the esters observed throughout simulations is very accurate. Moreover, this assembling pattern is a further indication that esters can act both as emollients and aid emulsification. Experimental studies indicate the obtention of dispersed phase droplets forming stable emulsions in all cases, but more stable ones for the tested esters. These findings underscore the potential for sustainable ingredients to deliver comparable, if not superior performance in cosmetic formulations.

Importantly, the previous works indicate that the choice between preorganized assemblies and the random distribution of ingredients in the simulation box, as structural inputs, significantly impacts molecular simulations, with force field selection playing a key role in modulating their behavior.

As discussed in our study, GROMOS MOlecular Simulation (GROMOS) and Optimised Potentials for Liquid Simulations (OPLS) differ in their parameterization strategies, particularly in handling nonbonded interactions and solvation effects, which influence self-assembly pathways. Historically, GROMOS was widely used in implicit solvent simulations with a tendency to favor folded or preorganized states. However, recent advancements over the past 20 years have demonstrated its robust performance in explicit solvent environments, where it has been increasingly used alongside refined water models to achieve more accurate and balanced conformational sampling.^[29]

Other force fields, however, have been developed to better sample and represent molecules with improved and more refined resolution. OPLS, as one of those force fields, with its explicit solvent description and finely tuned torsional parameters, allows for a more dynamic exploration of self-assembly.^[30] Simulation time and temperature further regulate these processes—extended timescales enable the system to overcome kinetic barriers to assembly, while temperature variations affect the entropic contributions to self-organization. As demonstrated in our work,

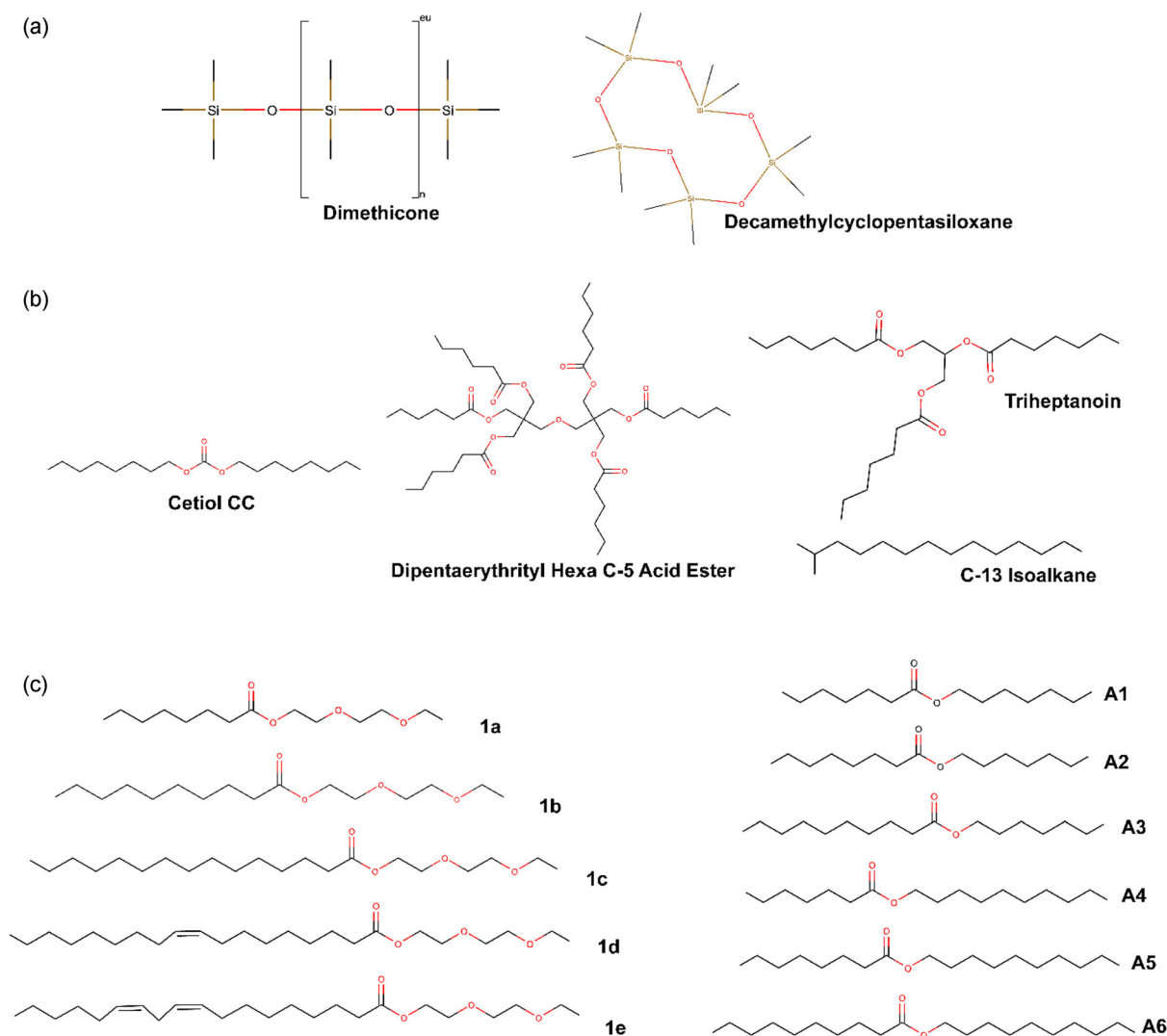


Figure 2. Structural depiction of a) common silicones used in cosmetics, b) commercially available alternatives, and c) molecules synthesized by our team to replace silicones.

systematic force field comparisons are essential for understanding the balance between initial conformational bias and induced self-assembly in complex formulations.

GROMACS, although a widely used and tested tool for MD simulations and trajectories analysis, providing excellent density estimations and reliable results regarding structural deviations over time and conformational changes/behavior, also has its downsides, namely in its capability to evaluate and correctly assess the rheology of a liquid solution. This is clearly explained in the work of Berk Hess, where a comparison is made between two equilibrium methods for shear viscosity estimation, and another nonequilibrium method.^[31]

Equilibrium methods, based on pressure or momentum fluctuations, often suffer from high statistical noise and require long simulations for convergence. Nonequilibrium methods, such as periodic shear flow, offer improved precision but introduce artificial constraints that may not fully capture realistic liquid behavior. Despite advances in computational techniques, no single

method provides a universally reliable solution, and viscosity estimation continues to be an active area of research, requiring careful methodological choices depending on the system studied. This postulate is coherent with our findings, as we were unable to fully assess a viscosity value that was consistent with any of the results, or even within our pool of data derived from the same base formulation.

LAMMPS, another classical MD simulation code focused on materials modeling, appears to perform shear viscosity calculations with high accuracy in both equilibrium and nonequilibrium MD simulations.^[32,33] From our research, however, there appears to be no relevant literature describing the successful use of this tool on an emulsion such as the one we study in our current work. This, however, is a noteworthy tool to investigate in the future to better assess these formulations and address a significant limitation in this field.

A few similar efforts to disclose silicone properties *in silico* have been made by Goussard and coworkers, using techniques

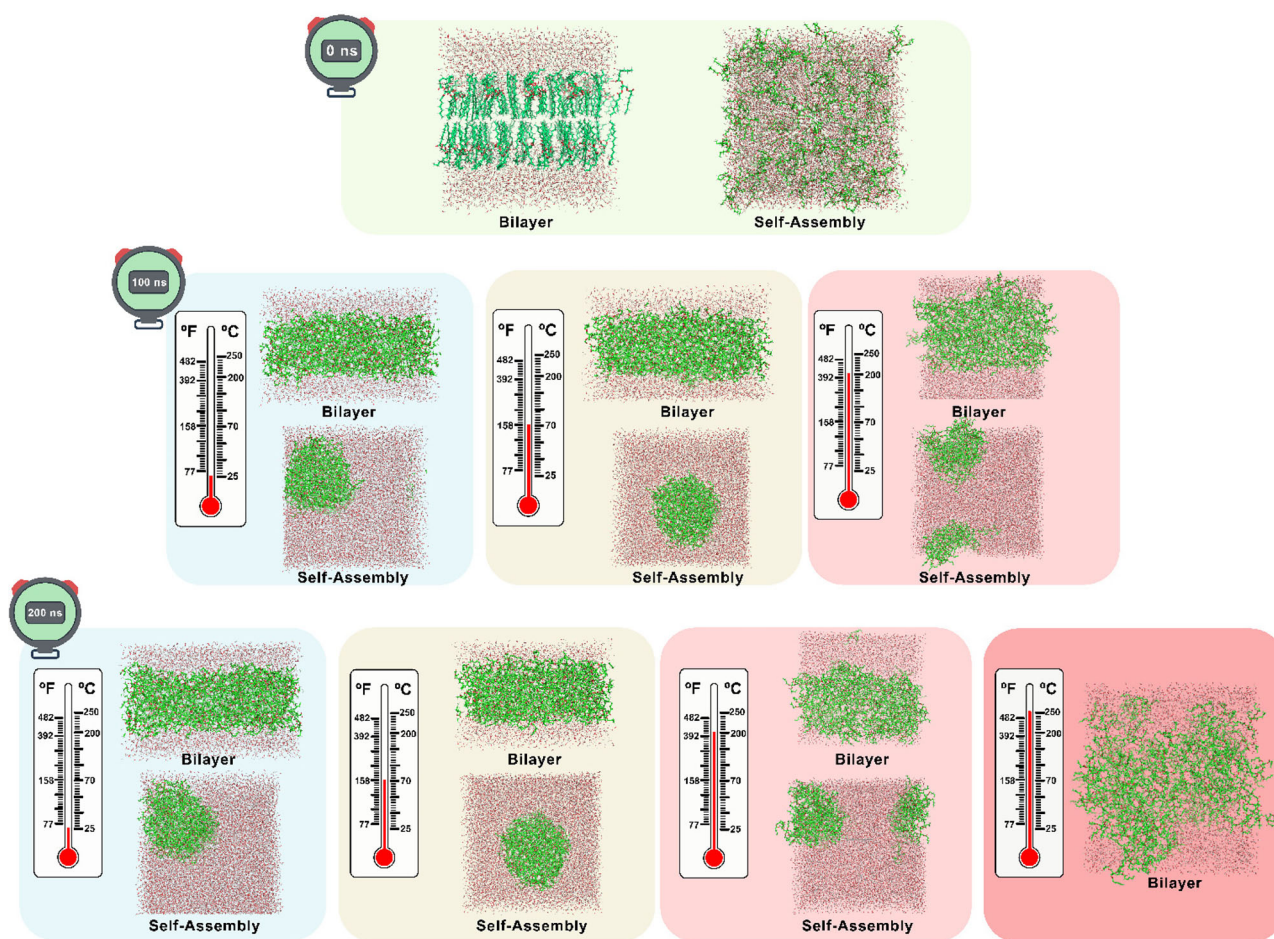


Figure 3. Visual representation of bilayer-like and self-assembly systems through time and at different temperatures. Green molecules are the ester-based systems, whereas the reddish spots are water molecules.

like COSMO-RS (conductor-like screening model for real solvents) and QSPR (quantitative structure-property relationship).^[20] Although different, both methods rely on mathematical derivations to assess the physicochemical properties of studied molecules and, through a computational approach, predict the physicochemical profiles of specific silicones/emollients. More specifically, they have studied linear dimethicones with low molecular weight. QSPR methods use machine learning (ML) approaches where a “learning” stage takes place from a given dataset, enabling the model to be trained and subsequently draw conclusions regarding essential properties, such as viscosity. A major difference, stemming from its algorithmic approach and ultimately beyond the capability scope of COSMO-RS, is that the graph machine uses the molecular structure as input, making it a highly useful tool for researchers studying these types of molecules.

3. AI and Docking in Formulation Design

AI and ML are powerful tools for advancing research in cosmetic formulation. Here, we focus on the AI capability to generate

molecular structures with a higher degree of confidence. The advent of AlphaFold revolutionized the field of Protein Structure Prediction (PSP) by enabling the generation of highly accurate structural models of proteins/peptides from an amino acid sequence alone, thereby accelerating the development of novel formulations. This breakthrough addresses a longstanding challenge in biology, facilitating a deeper understanding of protein function and accelerating advancements in drug discovery and disease research.^[34] Its transformative impact on structural biology and chemical fields ultimately led to the awarding of the 2024 Nobel Prize in Chemistry to Demis Hassabis and John Jumper for the development of AlphaFold, highlighting its significant contributions to biochemistry and molecular biology.^[35]

Molecular docking, traditionally applied to drug discovery, has found utility in cosmetic science, particularly in studying the interactions between bioactive compounds and skin and hair proteins.^[36–39] This technique helps predict how cosmetic ingredients, such as nutraceuticals, peptides or antioxidants, may bind to specific molecular targets, offering insights into their affinity, efficacy and molecular interaction patterns. This, in turn, enables the identification of potential active compounds and facilitates the design of more effective cosmetic products.

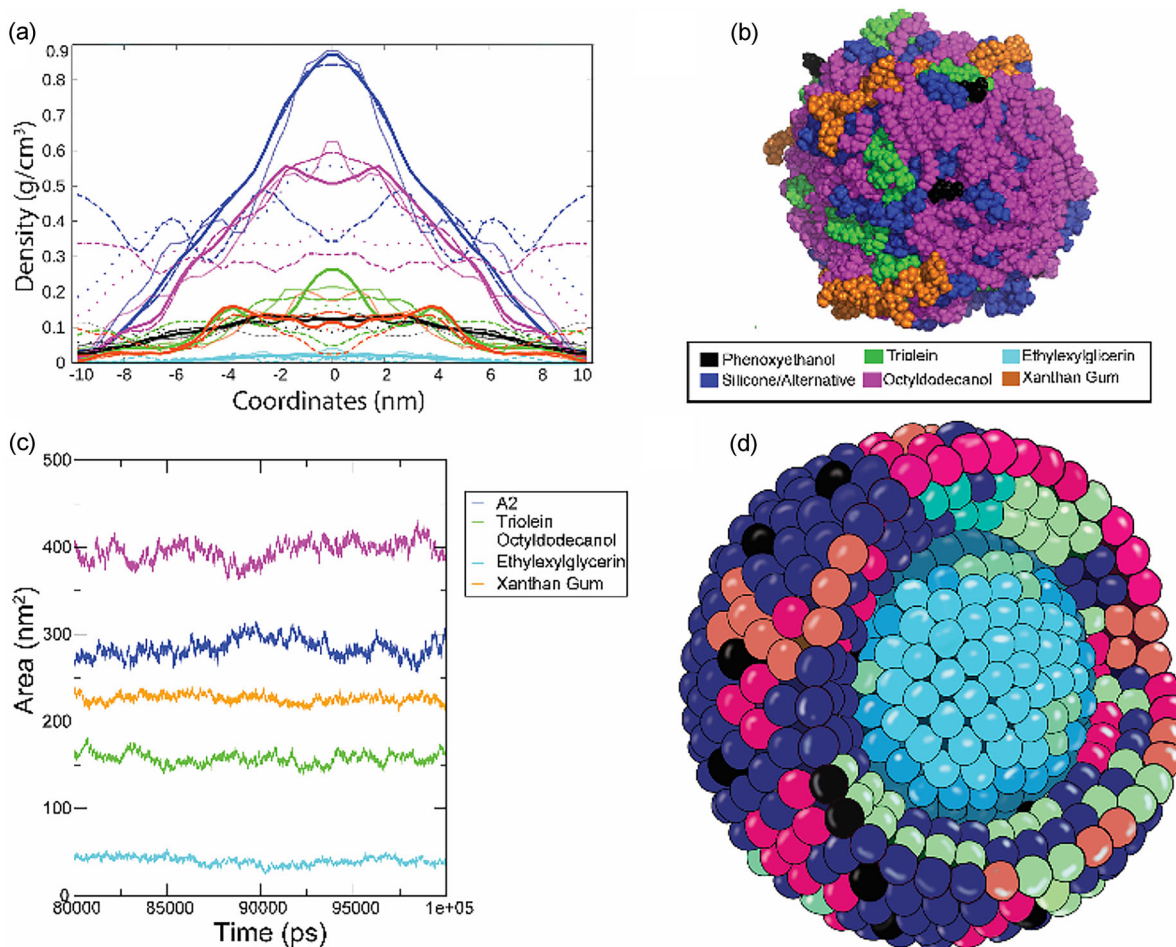


Figure 4. Results from the aforementioned work. In a) is represented the density graph, in b) the predicted micelles, without water and the solvents around, for clarity purposes, in c) a graph of the solvent accessible surface of the components within a tested micelle, and in d) a cross-sectional view of said micelle, in order to better understand the organizational pattern of affinity/hydrophobicity of the components. Ingredients here are phenoxyethanol, in black; the silicone or alternative, in dark blue, triolein, in green, octyldodecanol, in magenta, ethylexyglycerin, in cyan, and xanthan gum, represented in orange.

Additionally, companies like In Silico Discovery utilize advanced computational methods based on AI, a proprietary MD tool from a partner company (Nostrum Biodiscovery) and docking from pyDock^[40] to design active molecules for the cosmetic and biotechnological industries. These methods streamline the discovery process and reduce reliance on experimental approaches. Docking studies can, for example, help assess how natural or synthetic molecules interact with collagen, elastin, or enzymes responsible for skin aging, providing a rational basis for ingredient selection and formulation design.^[41]

Very recently, we explored the interaction of bioinspired peptides with different keratin models (KRT).^[42] This study presents a wide range of peptides' sequences derived from organisms present in nature, namely, abductin-based, collagen-based, resilin-based, elastin-based, and silk-based, in line with the trend of using natural-based products. Structures for each peptide were generated using Alphabet's AlphaFold 2.0. The aim of constructing these sequences was to evaluate the binding of elastomeric-derived motifs to hair fibers. To achieve this, an updated keratin

tetramer model was generated by AlphaFold, surpassing traditional steps of homology modeling, PSP, and docking, to obtain the final 3D structure (Figure 5).

Docking was performed using HDOCK, a semi-rigid docking web-server suited toward proteinprotein interactions. This tool not only provides an excellent docking solution but, according to the authors, is a better solution for the specific testing conditions, as more commonly used docking platforms, such as AutoDock Vina, would introduce excessive error and bias for such a large system. Further testing was conducted using the GROMACS' toolbox, where data from the MD simulations of these peptide:KRT systems were collected and analyzed for root mean square deviation (RMSD), radial distribution function (RDF), (Figure 5c) and number and frequency of hydrogen bonds. Additionally, online bioinformatics tools such as adaptive Poisson-Boltzmann solver (APBS) (Figure 5d) and ProtParam^[43] from ExPASy were employed, with IBM SPSS Statistics also used to generate a PCA biplot of parameters gathered from these resources.^[42]

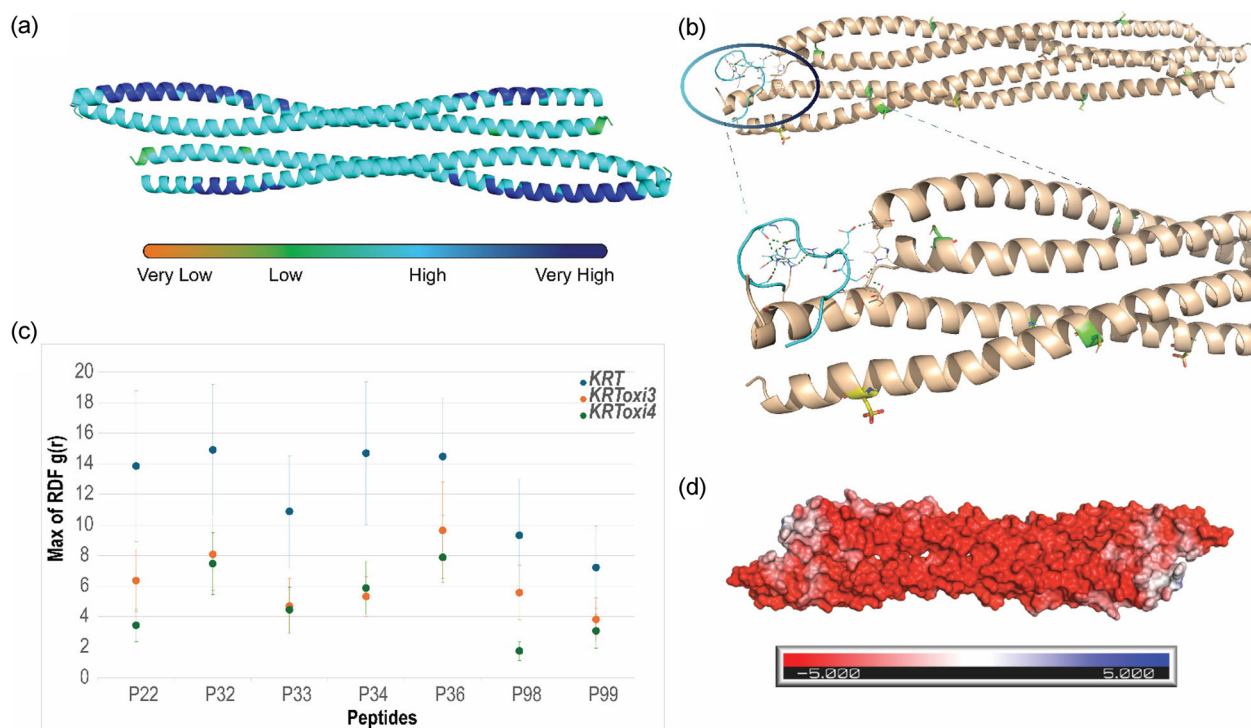


Figure 5. Results from Ferreira et al. showing a) AlphaFold prediction of a heterodimer of combined KRT34-KRT86 sequences with pLDDT coloring, b) a successful docking of an unorganized peptide onto the predicted tetramer, c) the max of RDF of peptide::KRT in a 0.3–1.0 nm distance interval, and d) the electrostatics potential surface determined by APBS.

The data gathered were conclusive regarding the intrinsic properties of these peptides in detecting the presence of KRT in their respective vicinities. All tested peptides showed moderate to very high affinity toward all tested KRT variants, with different degrees of damage/oxidation. Special emphasis was placed on three peptides, P32, P34, and P36, which correspond to silk, resilin, and abductin building blocks. From an interaction standpoint, these results show a clear interface between the peptides and the KRT, where hydrogen bond, van der Waals, ionic interactions, or even salt bridge formations may occur, even in damaged hair. This stable interaction is crucial for developing a product aimed at repairing the fiber. All these points strongly suggest a significant interaction between resilin, abductin, and silk domains with damaged human hair, which could contribute to the development of intellectual property for new cosmetic products. This work represents another important contribution of MD simulations to the computational chemistry of materials, a field that remains underexplored, especially in the context of personal care products, particularly hair care.

4. Optimization Methods in Formulation Design

The use of optimization methods in product design, particularly in cosmetic formulations, has been increasing over the past decade. Formulated cosmetic products are created by mixing selected ingredients into a formula, which includes the key

ingredients (active ones that perform the function) and supporting ingredients (solvents). When combined, these ingredients are subject to physicochemical phenomena, which can lead to synergistic or antagonistic effects.^[44] Ingredients and their related internal product structure (microstructure) are the two main factors influencing product performance.^[12] To speed up the formulation design, model-based optimization strategies have emerged and given rise to computer-aided molecular design (CAMD). This approach relies on diverse mathematical models and database-driven information to predict successful blends.^[45] More recently, optimization methods have evolved from standard model-based CAMD by integrating additional methods and a multiscale approach.^[46,47] This strategy involves coupling macroscopic properties such as rheological and textural characteristics with microscopical ones, such as droplet size distribution, often through models such as the Population Balance Model. A Python-based optimization modeling language, is also discussed and employed to implement these complex multiscale optimization problems.^[47]

The methods previously discussed in this concept, such as molecular docking, MD simulations, and ML, can be contemplated in such integrative and multiscale optimization strategies.^[48] Most of these studies focus on product formulation regarding oil/water emulsions or solvent choice.^[8,46,47,49]

The main take-home message from these works is that cosmetics formulation should move beyond trial-and-error experimentation by embracing multiscale and model-based design. Such methods can even integrate consumer insights and

economic evaluation, which enables faster, more sustainable, and market-relevant product development. Nevertheless, success in modern cosmetic formulation design depends on multiscale modeling—from molecular to macroscopic—with integration to structure-property relationships, sustainability, process simulation, AI-driven methods, consumer satisfaction, and profitability to rapidly design innovative, high-performance, and ecoconscious products.

5. Other Machine Learning, AI Initiatives, and Bioinformatics Tools

The current field has advanced to the point where we can analyze large datasets from experimental studies, consumer preferences, and ingredient libraries. AI models can predict product performance, stability, dynamics, and ingredient properties. ML algorithms, for instance, can predict the outcomes of mixing specific ingredients based on known molecular properties, aiding the design of new formulations that meet the manufacturer's desired criteria. These methods also enable the development of "virtual formulation libraries" where various ingredient combinations can be tested and optimized computationally before being physically produced.^[50]

In silico models built on ML and AI have been developed to predict the sensitizing potential (the ability for a product to induce an allergic reaction) of cosmetic ingredients based on their physicochemical characteristics, achieving accuracy comparable to traditional animal and in vitro tests.^[51]

A particular approach has been adopted for the in silico predictive study of skin sensitization. One such approach, by Urbisch and his colleagues, involved testing a wide range of nonanimal-based methods to assess the potential hazard of skin sensitization to human skin. They employed in chemico, in vitro, and in silico strategies, comparing them to a standard testing using murine animal models.^[52] The interesting limitation of this study is that, although the alternatives to animal models performed well, the computational approach still relied on data extracted from animal tests. This, in turn, led to discrepancies when compared to human data, indicating that there are limitations to in silico methods, largely stemming from the way data is fed into the tool. Another common occurrence in AI models/algorithms is that these tend to incur a phenomenon commonly referred to as hallucination, where the model tries to feed the gap between certain unknown aspects and just infers on things with no actual scientific basis.^[53] However, the authors conclude that, with complementary data from human assays and validation strategies using classical in vitro and/or in chemico methods, these approaches represent valid and promising alternatives for studying and ultimately replacing animal tests for skin sensitization hazards.

Furthermore, the work of Johnson and colleagues^[54] introduces an in silico skin sensitization protocol approach, taking into account the previously mentioned limitations, and presenting an important risk assessment framework aimed at mitigating flaws arising from poor input data or lack of information present

on some models. They also rely on bioinformatics tools like Quantitative Structure-Activity Relationship (QSAR) as a basis for predicting how a given molecule will interact with human skin. These QSAR methods use mathematical and algorithmic frameworks to replace traditional experimental approaches, such as animal testing, with virtual experimentation. These models analyze molecular structures holistically or focus on specific substructures to predict activity or toxicity.^[55]

Likewise, efforts have been made to develop important tools regarding pharmacokinetics and toxicology. Emerging from a combination of ML and AI, ADMETLab,^[56] mainly its new iteration ADMETLab2.0, is a tool that has come to address, as its acronym indicates: adsorption, distribution, metabolism, excretion, and toxicology properties of several drug candidates and other small molecules (up to 128 atoms). Based on models trained from a carefully curated yet vast database of molecules mapped to experimentally determined ADMET properties, the tools allow researchers to input novel molecular structures and receive rapid predictions on a wide array of crucial parameters, including oral bioavailability and blood-brain barrier penetration to cytochrome P450 inhibition and potential toxicities like hERG blockage or mutagenicity. The relevance of such a tool is imperative in modern R&D processes, presenting a cost and time-saving alternative for obtaining intermediate results across broader research pipelines. This enables quicker and more effective drug development and, when applied to cosmetics, facilitates the creation of better and more carefully designed products. It allows a wider range of companies and researchers to introduce new products and technologies, reducing the need to spend valuable resources on measurements by simply having a web-based tool at hand. However, such tools come with their limitations. One must always contextualize their outputs as predictive estimates rather than definitive experimental results. The reliability of these predictions is inherently linked to the quality, diversity, and relevance of the training data. Furthermore, the underlying models are simplifications of complex biological systems, requiring careful interpretation and subsequent experimental validation for critical decision-making in drug or, in this case, ingredient development.

Through the use of novel DL approaches, cosmetic research at the molecular level can be greatly advanced. In a recent work, DeepMol,^[57] an open-source automated ML framework, was explored. This tool simplifies the research process for computational chemists who may not be well-versed in ML tools and pipelines, automating feature engineering, model selection, and hyperparameter tuning. It streamlines complex tasks like toxicity assessment, stability prediction, and ingredient efficacy, which are significant challenges in cosmetics research. Its compatibility with cheminformatics tools (e.g., RDKit, DeepChem) allows seamless processing of molecular structures, while interpretability features (e.g., SHAP values) validate safety and performance for regulatory compliance. This makes it a transformative tool for advancing sustainable and efficient cosmetics innovation. Building on deeper learning strategies, the work of Sim and coworkers^[58] also introduces a novel system that combines DL techniques to analyze cosmetic ingredients and assess

individual skin conditions for personalized skincare product recommendations. In this study, researchers developed a deep neural network model to evaluate the efficacy of cosmetic products based on a given ingredients list. The model treats these lists as sequential data to predict potential skin benefits. Additionally, they implemented AI-driven skin analysis models to assess users' skin conditions from facial images, focusing on concerns such as pores, redness, acne, wrinkles, pigmentation, and dark circles. By integrating the results from both the ingredient analysis and skin assessments, the system provides tailored product suggestions that align with the user's specific skin needs. The effectiveness of this approach was demonstrated through reliable evaluation metrics and practical examples, indicating its potential to enhance personalized skincare routines.

The findings reported herein, alongside ongoing advancements in modern research and development, strongly indicate that AI-driven methodologies represent a significant evolution for both *in silico* and experimental research. This paradigm shift offers substantial advantages over previous, more classical approaches. AI algorithms, often highly optimized, enhance consistency and efficiency by automating complex intermediate tasks, thereby reducing reliance on manual control and mitigating specific forms of human error. In contrast to some earlier classical computational methods that could exhibit slower processing, less robust performance, and a propensity to introduce artifacts during extended runtimes due to hardware limitations or algorithmic sensitivities,^[59,60] AI leverages modern high-performance computing infrastructure for rapid and reliable data processing. Crucially, when employed symbiotically with experimental validation, or classical modeling methods, these AI techniques significantly streamline research workflows.^[61] They facilitate more efficient hypothesis generation, reduce dependence on extensive, time-consuming trial-and-error experimentation, and can lead to substantial savings in reagent consumption and intermediate experimental steps, thus accelerating discovery.

6. Summary and Outlook

While molecular computational methods and product optimization strategies offer significant advantages in predicting formulation behavior and important molecular details, their accuracy and applicability should also be compared/validated against experimental results whenever possible. This integration of computational and experimental approaches is not an afterthought but an essential feedback loop, enabling a more robust and efficient formulation process. For example, MD simulations of emulsion systems can be compared with experimental rheology measurements or microscopy results to ensure that the simulated interactions correspond to real-world macroscopic phenomena. In other words, MD simulations can provide insights regarding material strength, thermal properties, diffusion rates, phase transition, conductivity, surface properties, among other measures, that correlate with micro and macroscopic observations.^[62]

Similarly, molecular docking predictions on ingredients' affinity to a skin or hair target can be comparable to *in vitro* or *in vivo* studies, to assess K_a and binding Free Energy.^[63] One good example of the translation of atomistic data to the macroscopic world is the examination of hydrogen bond patterns, lifetimes and strengths between molecules (e.g., glycerol, hyaluronic acid) and water, given information regarding the moisturizer effect that a formulation can present.^[64,65]

The integration of experimental and computational approaches allows for a more robust and efficient formulation process, where simulations guide experimentation, and experimental data refines computational models. In this symbiotic relationship, simulations guide targeted experimentation, while empirical data, in turn, refines and improves the predictive power of computational models, ultimately leading to validated and highly performant product designs.

Few studies proposing hair and skin structures as scaffold for modeling studies can be found in the literature. Moreover, these structures must be richer in molecular detail (atomistic) to ensure the proper simulation of interactions and the environment. Only from these structures is it possible to optimize computational protocols for studying cosmetic actives/ingredients.

Nowadays, computational power has advanced to the point where MD simulations of larger and more complex systems can be a hindrance, as the potential for even greater computational power comes with a significant cost. Our own outputs, along with many others cited herein, consistently highlight the importance of these approaches in providing cost-saving solutions for modern research. More importantly, in our view, they also address the environmental impact of large computational facilities. The financial cost of CPUs, the infrastructure required for cooling facilities and the well-documented environmental harm caused by AI infrastructure present significant challenges in modern society.^[66,67] Additionally, despite the massive discharge of contaminants into natural environments driven by the ever-growing consumer demand for cosmetics, along with the impact of residue creation and costly assays, spending many reagents, *in silico* approaches still offer a net positive environmental benefit by replacing more traditional approaches. However, these challenges must be acknowledged and not ignored, and we must continue to improve as both a scientific community and a responsible society.

Acknowledgements

T.F. and A.C.-P thank FCT under the scope of the strategic funding of UIDB/04469/2020 unit, and by LABBELS – Associate Laboratory in Biotechnology, Bioengineering, and Microelectromechanical Systems, LA/P/0029/2020. T.F. 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. T.F. also thanks Dr. Diana Guimarães and Dr. Ana Tinoco for their valuable review of the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: bioinspired molecules · computational chemistry · hair-binding · molecular docking · molecular modelling · peptides · silicone alternatives

- [1] A. Franzol, T. M. Banin, T. R. Brazil, M. C. Rezende, *J. Sol-Gel Sci. Technol.* **2021**, 99, 469.
- [2] V. Klang, N. B. Matsko, C. Valenta, F. Hofer, *Micron* **2012**, 43, 85.
- [3] C. M. Ribeiro, M. Souza, B. L. Pelegrini, R. S. Palacios, S. M. Lima, F. Sato, A. C. Bento, M. L. Baesso, M. M. S. Lima, *J. Photochem. Photobiol. B* **2022**, 236, 112587.
- [4] M. A. Vratsanos, N. C. Gianneschi, *ACS Nano* **2022**, 16, 7783.
- [5] A. S. Barna, C. Maxim, A. Trifan, A. C. Blaga, R. Cimpoesu, D. Turcov, D. Suteu, *Int. J. Mol. Sci.* **2023**, 24, 7004.
- [6] H. Xin, A. S. Virk, S. S. Virk, F. Akin-Ige, S. Amin, *Curr. Opin. Colloid Interface Sci.* **2024**, 73, 101847.
- [7] E. Conte, R. Gani, Y. S. Cheng, K. M. Ng, *Ch E J.* **2012**, 58, 173.
- [8] F. Calvo, J. M. Gómez, L. Ricardez-Sandoval, O. Alvarez, *Chem. Eng. Res. Des.* **2020**, 161, 279.
- [9] L. E. Bouamri, M. Bouachrine, S. Chtita, *Curr. Top. Med. Chem.* n.d., 24, 1.
- [10] Y. Xu, H. Ni, F. Zhao, Q. Gao, Z. Zhao, C.-H. Chang, Y. Huo, S. Hu, Y. Zhang, R. Grovu, H. He, J. Z. H. Zhang, Y. Wang, *AIP Adv.* **2025**, 15, 010601.
- [11] E. Türedi, N. Acarali, *Ind. Crops Prod.* **2022**, 187, 115303.
- [12] F. Mushtaq, X. Zhang, K. Y. Fung, K. M. Ng, *Front. Chem. Sci. Eng.* **2021**, 15, 1033.
- [13] H. A. L. Filipe, L. M. S. Loura, *Molecules* **2022**, 27, 2105.
- [14] T. Ferreira, D. Rocha, D. Freitas, J. Noro, M. de Castro, C. Roque, D. Guimarães, A. Loureiro, C. Silva, A. Cavaco-Paulo, T. G. Castro, *Ind. Eng. Chem. Res.* **2024**, 63, 9715.
- [15] K. M. Gupta, S. Das, P. S. Chow, *Nanoscale* **2021**, 13, 12916.
- [16] L. T. N. Ngoc, J.-Y. Moon, Y.-C. Lee, *Cosmetics* **2023**, 10, 111.
- [17] E. V. Ivanova, E. O. Minyaylo, M. N. Temnikov, L. G. Mukhtorov, Y. M. Atroshchenko, *Polym. Sci. Ser. B* **2023**, 65, 578.
- [18] E. V. Ivanova, E. O. Minyaylo, M. N. Temnikov, L. G. Mukhtorov, Y. M. Atroshchenko, *Polym. Sci. Ser. B* **2023**, 65, 578.
- [19] A. J. O'Lenick, *Silicones for Personal Care*, Allured Publishing Corporation, USA **2008**.
- [20] V. Goussard, J.-M. Aubry, V. Nardello-Rataj, *Adv. Colloid Interface Sci.* **2022**, 304, 102679.
- [21] C. A. R. Picken, O. Buensoz, P. D. Price, C. Fidge, L. Points, M. P. Shaver, *Chem. Sci.* **2023**, 14, 12926.
- [22] M. Giustra, G. Sinesi, F. Spena, B. De Santes, L. Morelli, L. Barbieri, S. Garbujo, P. Galli, D. Prosperì, M. Colombo, *ChemSusChem* **2024**, 17, e202401065.
- [23] A. Samir, F. H. Ashour, A. A. A. Hakim, M. Bassyouni, *Npj Mater. Degrad.* **2022**, 6, 68.
- [24] C. S. Roque, M. de Castro, T. G. Castro, A. Ribeiro, C. Silva, A. Cavaco-Paulo, J. Noro, *J. Mol. Liq.* **2025**, 420, 126818.
- [25] C. S. Roque, M. de Castro, T. G. Castro, C. Silva, A. Cavaco-Paulo, J. Noro, *ChemBioChem* **2024**, 25, e202300843.
- [26] T. Ferreira, A. Loureiro, J. Noro, A. Cavaco-Paulo, T. G. Castro, *Polymers* **2023**, 15, 796.
- [27] M. B. Ansorge-Schumacher, O. Thum, *Chem. Soc. Rev.* **2013**, 42, 6475.
- [28] V. Krstonošić, L. Dokić, I. Nikolić, M. Milanović, *Food Hydrocoll.* **2015**, 45, 9.
- [29] J. Zhang, H. Zhang, T. Wu, Q. Wang, D. van der Spoel, *J. Chem. Theory Comput.* **2017**, 13, 1034.
- [30] A. M. Fluit, J. J. de Pablo, *Biophys. J.* **2015**, 109, 1009.
- [31] B. Hess, *J. Chem. Phys.* **2002**, 116, 209.
- [32] Y. Zhang, A. Otani, E. J. Maginn, *J. Chem. Theory Comput.* **2015**, 11, 3537.
- [33] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolinteanu, W. M. Brown, P. S. Crozier, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, S. J. Plimpton, *Comput. Phys. Commun.* **2022**, 271, 108171.
- [34] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek, 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, *Nature* **2021**, 596, 583.
- [35] H. Devlin, *Google DeepMind Scientists and Biochemist Win Nobel Chemistry Prize* **2024**, <https://www.theguardian.com/science/2024/oct/09/google-deepmind-scientists-win-nobel-chemistry-prize>
- [36] R. Hafner, N. Wolgramm, P. Klein, H. M. Urbassek, *J. Phys. Chem. B* **2024**, 128, 45.
- [37] P. C. Agu, C. A. Afiukwa, O. U. Orji, E. M. Ezech, I. H. Ofoke, C. O. Ogbu, E. I. Ugwuja, P. M. Aja, *Sci. Rep.* **2023**, 13, 13398.
- [38] H. Zhang, W. Li, Q. Zhang, R. Zhong, C. Li, Y. Chen, T. Xia, M. Peng, Z. Ren, H. Zhao, Y. Wang, Z. Shu, *Biomed. Pharmacother.* **2022**, 150, 112990.
- [39] S. Zheng, R. Deng, G. Huang, Z. Ou, Z. Shen, *Biogerontology* **2024**, 25, 1115.
- [40] B. Jiménez-García, C. Pons, J. Fernández-Recio, *Bioinformatics* **2013**, 29, 1698.
- [41] X. Yu, J. Li, M. Yang, C. Chen, S. Munir, J. You, T. Yin, R. Liu, S. Xiong, Y. Hu, *Food Chem.* **2021**, 360, 130068.
- [42] T. Ferreira, F. Pereira, F. Gonçalves, A. da Costa, A. Cavaco-Paulo, T. G. Castro, *Int. J. Biol. Macromol.* **2025**, 317, 144813.
- [43] E. Gasteiger, C. Hoogland, A. Gattiker, Séverine Duvaud, M. R. Wilkins, R. D. Appel, A. Bairoch, in *Proteomics Protocols Handbook*, Humana Press, Totowa, NJ **2005**, p. 571.
- [44] X. Zhang, T. Zhou, K. M. Ng, *AIChE J.* **2021**, 67, e17064.
- [45] K. M. Ng, R. Gani, *Comput. Chem. Eng.* **2019**, 126, 147.
- [46] J. J. Torres, C. D. Tinjaca, O. A. Alvarez, J. M. Gómez, *Chem. Eng. Sci.* **2020**, 212, 115326.
- [47] F. Calvo, O. Álvarez, J. M. Gómez, L. Ricardez-Sandoval, *Chem. Eng. Process.-Process Intensif.* **2024**, 197, 109711.
- [48] L. Zhang, H. Mao, Q. Liu, R. Gani, *Curr. Opin. Chem. Eng.* **2020**, 27, 22.
- [49] J. W. Chong, L. Y. Ng, O. A. Aboagwa, S. Thangalazhy-Gopakumar, K. Muthoosamy, N. G. Chemmangattuvalappil, *Processes* **2021**, 9, 2159.
- [50] H. Narayanan, F. Dingfelder, I. C. Morales, B. Patel, K. E. Heding, J. R. Bjelke, T. Egebjerg, A. Butté, M. Sokolov, N. Lorenzen, P. Arosio, *Mol. Pharm.* **2021**, 18, 3843.
- [51] J. Kalicińska, B. Wiśniowska, S. Polak, R. Spiewak, *Int. J. Mol. Sci.* **2023**, 24, 6801.
- [52] D. Urbisch, A. Mehling, K. Guth, T. Ramirez, N. Honarvar, S. Kolle, R. Landsiedel, J. Jaworska, P. S. Kern, F. Gerberick, A. Natsch, R. Emter, T. Ashikaga, M. Miyazawa, H. Sakaguchi, *Regul. Toxicol. Pharmacol.* **2015**, 71, 337.
- [53] O. Kovalevskiy, J. Mateos-Garcia, K. Tunyasuvunakool, *Proc. Natl. Acad. Sci.* **2024**, 121, e2315002121.
- [54] C. Johnson, E. Ahlberg, L. T. Anger, L. Beilke, R. Benigni, J. Bercu, S. Bobst, D. Bower, A. Brigo, S. Campbell, M. T. D. Cronin, I. Crooks, K. P. Cross, T. Doktorova, T. Exner, D. Faulkner, I. M. Fearon, M. Fehr, S. C. Gad, V. Gervais, A. Giddings, S. Glowienke, B. Hardy, C. Hasselgren, J. Hillegass, R. Jolly, E. Krupp, L. Lomnitski, J. Magby, J. Mestres, et al., *Regul. Toxicol. Pharmacol.* **2020**, 116, 104688.
- [55] G. Selvestrel, F. Robino, M. Z. Russo, *Methods Mol. Biol. Clifton NJ.* **2022**, 2425, 291.
- [56] G. Xiong, Z. Wu, J. Yi, L. Fu, Z. Yang, C. Hsieh, M. Yin, X. Zeng, C. Wu, A. Lu, X. Chen, T. Hou, D. Cao, *Nucleic Acids Res.* **2021**, 49, W5.
- [57] J. Correia, J. Capela, M. Rocha, *J. Cheminformatics* **2024**, 16, 136.
- [58] J. H. Sim, J. Yoo, M. L. Lee, S. H. Han, S. K. Han, J. Y. Lee, S. W. Yi, J. Nam, D. S. Kim, Y. S. Yang, *ACS Appl. Mater. Interfaces* **2024**, 16, 25825.
- [59] J. D. Durrant, J. A. McCammon, *BMC Biol.* **2011**, 9, 71.
- [60] D. Perez, A. Thompson, S. Moore, T. Oppelstrup, I. Sharapov, K. Santos, A. Sharifian, D. Z. Kalchev, R. Schreiber, S. Pakin, E. A. Leon, L. J. H. Ill, M. James, S. Rajamanickam, *J. Chem. Phys.* **2025**, 162, 072501.
- [61] L. Nolte, S. Tomforde, *Appl. Sci.* **2025**, 15, 5208.
- [62] G. Ciccotti, S. Decherchi, S. Meloni, *Riv. Nuovo Cimento* **2025**, 48, 1.

- [63] S. T. Ngo, N. M. Tam, M. Q. Pham, T. H. Nguyen, *J. Chem. Inf. Model.* **2021**, *61*, 2302.
- [64] F. O. Akinkunmi, D. A. Jahn, N. Giovambattista, *J. Phys. Chem. B.* **2015**, *119*, 6250.
- [65] M. Morita, F. Matsumura, T. Shikata, Y. Ogawa, N. Kondo, K. Shiraga, *J. Phys. Chem. B.* **2022**, *126*, 9871.
- [66] U. Gupta, Y. G. Kim, S. Lee, J. Tse, H.-H. S. Lee, G.-Y. Wei, D. Brooks, C.-J. Wu, **2020**, arXiv, Cornell University, <https://doi.org/10.48550/arXiv.2011.02839>.
- [67] D. Patterson, J. Gonzalez, Q. Le, C. Liang, L.-M. Munguia, D. Rothchild, D. So, M. Texier, J. Dean, **2021**, arXiv, Cornell University, <https://doi.org/10.48550/arXiv.2104.10350>.

Manuscript received: May 20, 2025
Revised manuscript received: July 4, 2025
Version of record online: