

Review

Not peer-reviewed version

---

# The Study of Molecules and Processes in Solution: An Overview of Questions, Approaches and Applications

---

[Neani Tshilande](#) , [Liliana Mammino](#) <sup>\*</sup> , Mireille Kabuyi Bilonda

Posted Date: 9 January 2024

doi: 10.20944/preprints202401.0704.v1

Keywords: green solvents; implicit solvation models; explicit solvation models; hybrid solvation models; solute-solvent interactions; polarizable continuum model; water molecules in biomolecules' interactions



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

# The Study of Molecules and Processes in Solution: An Overview of Questions, Approaches and Applications

Neani Tshilande <sup>1</sup>, Liliana Mammino <sup>2,\*</sup> and Mireille K.. Bilonda <sup>3</sup>

<sup>1</sup> Department of Chemistry, University of Venda, Thohoyandou, South Africa

<sup>2</sup> Faculty of Science, Engineering and Agriculture, University of Venda, Thohoyandou, South Africa

<sup>3</sup> Chemistry and Industry Department, Faculty of Sciences and Technology, University of Kinshasa, Kinshasa, D.R. Congo

\* Correspondence: sasdestria@yahoo.com

**Abstract:** Many industrial processes, several natural processes involving non-living matter, and all the processes occurring within living organisms take place in solution. This means that the molecules playing active roles in the processes are present within another medium, called solvent. The solute molecules are surrounded by solvent molecules and interact with them. Understanding the nature and strength of these interactions, and the way in which they modify the properties of the solute molecules, is important for a better understanding of the chemical processes occurring in solution, including possible roles of the solvent in those processes. Computational studies can provide a wealth of information on solute-solvent interactions and their effects. Two major models have been developed to this purpose: a model viewing the solvent as a polarizable continuum surrounding the solute molecule, and a model considering a certain number of explicit solvent molecules around a solute molecule. Each of them has its advantages and challenges, and one selects the model that is more suitable for the type of information desired for the specific system under consideration. These studies are important in many areas of chemistry research, from the investigation of the processes occurring within a living organism to drug design and to the design of environmentally benign solvents meant to replace less benign ones in the chemical industry, as envisaged by the green chemistry principles. The paper presents a quick overview of the modelling approaches and an overview of concrete studies, with reference to selected crucial investigation themes.

**Keywords:** green solvents; implicit solvation models; explicit solvation models; hybrid solvation models; solute-solvent interactions; polarizable continuum model; water molecules in biomolecules interactions

---

## 1. Introduction

As W. Ostwald already stated in 1890, "Almost all the chemical processes which occur in nature, whether in animal or vegetable organisms, or in the non-living surface of the earth, and also all the processes which are carried out in the laboratory, take place between substances in solution" [1]. This is true also for most of the production processes in the chemical industry, as being dissolved in a solution provides opportunity for the molecules of different substances to meet and react.

Because of their general importance, the properties of solutions have been objects of intensive studies since the late XIX century, and the interpretation of observations has led to insights about what happens when a substance (solute) dissolves in a given solvent. For instance, the fact that some solutions can conduct electric current led to the inference that the particles dissolved in those solutions are in ionic form, and the fact that the magnitude of the colligative properties of those solutions was greater than would have been expected on the basis of the concentration values led to

the inference that the molecules of those solutes (acids, bases) dissociate into ions (or otherwise give rise to ions) in water solution.

The key features of what happens when a substance dissolves in a solvent are known from basic chemistry: the solvent molecules interact with the solute molecules with which they come into contact, separate them from the solid of which they are initially part, and surround them; the molecules of sufficiently polar solvents make the molecules of polar solutes dissociate into ions and then surround the individual ions; if the solute is an ionic compound and the solvent molecules are polar, they separate the ions from the solid and surround them.

Whether a certain solute dissolves in a certain solvent, and the energetics of the dissolution process, depend on the enthalpy and entropy changes accompanying the process; therefore, the energetics is expressed in terms of free energy of solvation ( $\Delta G_{\text{solv}}$ ). At the molecular level, they depend on the interactions between the solute molecules and the solvent molecules (solute-solvent interactions). The interactions can be of various types, depending on the nature of the solute and the solvent: intermolecular hydrogen bonds, electrostatic interactions, hydrophobic interactions, dispersion interactions, and also repulsion. Hydrogen bonds (H-bonds) are generally the strongest, and can form when both the solute and the solvent molecules contain H-bond donor or acceptor groups. A given solute dissolves in a given solvent if the solute-solvent interactions overcome the interactions among solute molecules in the pure solute.

The solute-solvent interactions determine several changes in the properties of the solute molecules with respect to when the molecule is isolated (gas phase). The changes may concern the geometry parameters (bond lengths, bond angles, torsion angles) of the equilibrium geometry of the molecule's individual conformers, its conformational preferences, its charge distribution, dipole moment, IR vibrational frequencies, ultraviolet/visible signals, electronic transition energies, NMR constants, chemical reactivity, and various others [2–4].

Understanding how a specific solute and a specific solvent interact is fundamental for a better understanding of the processes occurring in living organisms, including the effects of biologically active substances introduced within an organism to obtain desirable results, e.g. for the treatment of diseases. It is also fundamental for the understanding of other processes occurring in solution, including industrial processes. In recent decades, it is also fundamental in the design of more environmentally-benign processes along the patterns envisaged by green chemistry [5–8], as the general reduction of the use of solvents and the selection of more benign ones play crucial roles to increase the sustainability of industrial processes [9,10]. In order to be benign, a solvent should have low toxicity and low volatility, not pose risks such as flammability, be biodegradable in the environment, with non-harmful degradation products, and require low energy costs for its synthesis; in order to be performing, it has to be inert (not reacting with the solutes) and be easily recovered at the end (without contaminating the final product). Reducing the use of harmful solvents entails a variety of approaches, from the design of new, more benign solvents suitable for specific processes [11–15] to the design of processes that can make use of existing benign solvents [16], such as water [17–20] or supercritical  $\text{CO}_2$  [21–25]. Separation processes constitute a major component of many industrial processes and the possibility of carrying them out in green solvents is often focus of specific attention [26–28].

The advances in computational chemistry have enabled the generation of models for the study of solute-solvent interactions as well as continuous enhancements in their appraisal. The next section presents a quick review of the major models for the study of these interactions and their effects. The subsequent sections outline a number of issues for which the study of solute-solvent interactions has proved of interest for research and industry, and provide examples of the information contributed by computational approaches. It is impossible to make a comprehensive review of studies of molecules and processes in solution, because their number is in the range of several thousands. The present work aims at highlighting the variety of research questions and applications that require adequate understanding of what happens in solution for the search of effective answers to be viable, and the role of computational chemistry to facilitate this understanding. Since the main focus concerns the conceptual nature of the questions, relevant works outlining the main characters of this

nature, and of the corresponding search for answers, are included in the review, which thus covers the last two decades (sometimes even earlier) to ensure a comprehensive presentation; basic historical information is also included for the development of the computational models, and when expedient to highlight the research questions more completely.

## 2. Models for the Computational Study of Solvent Effects

Two major approaches have been developed for the study of solute-solvent interactions (including  $\Delta G_{\text{solv}}$ ) and their effects on the properties of the solute molecules, differing by the way in which the solvent is represented: implicit models, where the solvent is represented as a continuum surrounding the solute molecule, and explicit models, where a certain number of solvent molecules are considered individually. Several reviews of the approaches are already available (e.g., [29]); a review giving particular attention to biological systems is included in [4], and a review giving particular attention to green chemistry is included in [30]; therefore, only the main features of the two approaches are recalled here.

### 2.1. Implicit Solvation Models

Implicit solvation models represent the solvent as a continuous polarisable medium characterised by its dielectric constant. The solute molecule is viewed as embedded in a cavity within this continuum and is represented by the charge distribution ( $\rho(r)$ ) on the surface of the cavity. The dissolution process entails the formation of the cavity, with the solute molecule displacing enough solvent molecules to form it [31,32]. The charge distribution of the solute polarises the solvent around the cavity, generating a reaction-field potential in it, which, in turn, polarises the solute charge distribution. Methods utilising this model are therefore often termed *self-consistent reaction field* (SCRF) methods. The solute-solvent interactions are considered as a perturbation with respect to the situation of the solute molecule in the gas phase, and a perturbation term is added to the Hamiltonian operator of the isolated solute molecule, to write the Schrödinger equation for the solute molecule in solution. The simplest way to build the shape of the cavity considers the solute molecule as made of interlocking spheres, each having the van der Waals radius of the corresponding atom, and rolling a sphere with the diameter of a solvent molecule on the surface of the resulting structure, to smooth sharp intersections (solvent accessible surface [33,34]; some illustrations included in [4]).

The standard continuum model is typically represented by the polarizable continuum model (PCM), in which the polarisation of the medium outside the cavity – generated by the charge distribution inside the cavity – is modelled by a system of apparent surface charges (ASC) spread on the surface of the cavity. Through the years after its introduction [35], the model has seen continuous evolution, developing approaches to take into account different solute and solvent characteristics and the resulting effects (e.g., [36–39]), including the possibility of conductor solvents [40]. The integral equation formalism PCM (IEP-PCM, [41–43] enables the study of both isotropic systems (like solutions) and anisotropic systems (like liquid crystals), as well as systems where the liquid contains charged species (like ionic solutions). Subsequent advances are also outlined in review and reflection articles [31,32,44,45].

$\Delta G_{\text{solv}}$  is defined as the change in the free energy of the solute upon going from being isolated (ideal gas phase) to the solution phase. Therefore, it is estimated as the sum of the contributions that arise because of the dissolution process [30,32,46]:

$$\Delta G_{\text{solv}} = G_{\text{el}} + G_{\text{cav}} + G_{\text{dis}} + G_{\text{rep}} \quad (1)$$

where  $G_{\text{el}}$  is the electrostatic contribution and the other terms correspond to non-electrostatic contributions: the free energy of cavitation ( $G_{\text{cav}}$ , which is the reversible work needed to form the cavity where the solute gets embedded); the dispersion contribution ( $G_{\text{dis}}$ ), due to the dispersion interactions between the solute molecule and the solvent; and the repulsion contribution ( $G_{\text{rep}}$ ), which relates to the Pauli repulsion between the solute molecule and the solvent molecules. Some authors (e.g., [29]) add a thermal fluctuation contribution  $G_{\text{tm}}$ .

Refining the evaluation of  $G_{el}$  has been a major focus of attention for the improvement of the evaluation of  $\Delta G_{solv}$  [36]. Since the evaluation of  $G_{el}$  depends on the description of the charge distribution, and the charge distribution is associated with the shape (surface) of the cavity, improving the description of the cavity was a route to improve the evaluation of  $G_{el}$  [34,47–50].

The evaluation of  $G_{el}$  is based on the Poisson equation, which expresses the electrostatic potential ( $\phi$ ) in terms of the dielectric constant ( $\epsilon$ ) and the charge density ( $\rho$ ). In the PCM model, the value of  $\epsilon$  is 1 inside the cavity and takes the value of the specific liquid for the bulk medium representing the solvent. Two options have been developed to solve the equation for continuum solvation models: the Poisson–Boltzmann (PB) model and the Generalized Born (GB) model [29]. The former utilises the Poisson–Boltzmann equation, which expands the Poisson equation to take into account the possible presence of mobile electrolytes in the solution. The GB model utilises an approximation to the Poisson equation which can be solved analytically.

Modified versions of the PCM model have been developed to respond to specific criteria. The conductor-like polarisable continuum model (CPCM, [51]) is considered one of the most successful [29]. The conductor-like screening model (COSMO) and conductor-like screening model for real solvents (COSMO-RS), [52,53] solvation models are respectively variations of Poisson–Boltzmann PCM and CPCM. They combine the ASC formulation for the electrostatic component with a statistical thermodynamic treatment [30]. They consider the solvent as a conductor (thus setting  $\epsilon = \infty$ ) and use a scaling factor to attain a correct description of the considered solvent. COSMO-RS adds a statistical thermodynamic approach to the results of quantum chemical calculations to attain a realistic description of the dissolution mechanism [54]; it can provide very accurate  $\Delta G_{solv}$  estimations its ability to treat mixtures at variable temperatures has made it very popular in chemical engineering and in pharmaceutical chemistry, including for tasks like large-scale solvent screening [30,55].

The solvation model based on electron density (SMD, [56]) separates  $\Delta G_{solv}$  into two main components: the electrostatic contribution, obtained via IEF-PCM, and a cavity-dispersion-solvent-structure term, arising from short-range interactions between the solute and solvent molecules in the first solvation shell [30].

Calculations options based on continuum models are currently present in all the popular computational chemistry softwares; IEP-PCM is the default option for PCM.

## 2.2. *Explicit Solvation Models*

Explicit solvation models consider individual solvent molecules interacting with a solute molecule. Within a purely quantum mechanical (QM) approach, input supermolecular structures (adducts) are built, considering the most favourable arrangements of the solvent molecules around the solute molecule, and optimised with QM procedures. It is expedient to include at least the solvent molecules forming the first solvation layer, intended as those directly ‘attached’ to the solute molecule and those bridging them [4]. The identification of possible arrangements is easier when sufficiently strong interactions between the molecules can be predicted, as is the case of molecules that can form H-bonds (for instance, water molecules surrounding a solute molecule having H-bond donors or acceptors). The number of solvent molecules that can be included is however limited, because of the fast increase in computational costs as their number increases. In addition, it is important to identify a ‘balanced’ (not too high) number of solvent molecules, to prevent their clustering on optimisation if their mutual interactions are strong (as is the case, e.g., of water molecules): if the solvent molecules cluster together during optimisation, they move away from the solute molecule, and the resulting optimised geometry does not provide a realistic description of the first solvation layer and short-range solvation interactions.

The optimised structure does not correspond to a long-life structure in solution, as the solvent molecules attached to a solute molecule interchange fast with molecules from the bulk solvent; it does, however, show probable average arrangements of solvent molecules around the solute molecule. Furthermore, it highlights geometry changes that may occur in the solute molecules because of the presence of the solvent molecules, and it highlights phenomena like the outcome of the competition between intramolecular H-bonds that might be present in the isolated solute

molecule and solute-solvent H-bonds in solution (e.g., [57]). The approach is particularly relevant when directional solute-solvent interactions like H-bonds are possible, because they are not taken into specific account by continuum models. For an adduct containing  $n$  solvent molecules, the energy of the solute-solvent interactions is calculated as:

(energy of solute-solvent interactions) = (energy of the adduct) – (energy of the isolated solute molecule) –  $n$  (energy of an isolated solvent molecule) – (energy of the interactions among solvent molecules)

and corrections [58] for basis set superposition errors are advisable.

If the consideration of a high number of explicit solvent molecules is needed, approaches like Monte Carlo (MC) or Molecular Dynamics (MD, whose theoretical framework was introduced in [59]) are utilised. The solvent molecules are considered in motion. While the free energy contributions stemming from intramolecular components (vibrational and librational motions) can be estimated from the partition function, the contributions from the motion of the molecules through the solution and from the solute-solvent interactions are not so easily identifiable, and additional approaches, such as free energy perturbation methods, become expedient [29].

### 2.3. Quantum Mechanical /Classical or /Continuum Hybrid Approaches

QM methods provide the most accurate descriptions of molecular-level systems. On the other hand, their computational costs increase rapidly as the complexity of the system increases. Hybrid approaches have been developed, for which the portion of the system that is of highest interest is described with best accuracy at the QM level, whereas the rest of the system is described using a less expensive model. In the case of solutions, the portion of highest interest is the solute molecule and the solvent molecules more closely interacting with it (more often, the first solvation layer, as defined in section 2.2); this portion is described with QM approaches, and the rest of the solution with a continuum model; thus, this hybrid method is a QM/continuum method. In the case of the interaction between a biologically active molecule (substrate) and a protein, the portion of highest interest is the part of the protein interacting with the substrate, the substrate molecule itself, and the water molecules which might play a role in the interaction; this portion is described at QM level and the rest of the protein with a classical molecular mechanics (MM) approach; this hybrid method is a QM/MM method [30]; it was first introduced for the study of enzyme-involving reactions [60]. In both QM/continuum and QM/MM hybrid approaches, the part that is outside the QM portion can influence its properties, and this is taken into account through ways of coupling the description of the two portions.

In a QM/MM hybrid method, an effective Hamiltonian operator  $\hat{H}_{\text{eff}}$  is considered, as the sum of the Hamiltonian of the isolated QM subsystem ( $\hat{H}_{\text{QM}}$ ) and a term ( $\hat{H}_{\text{env}}$ ) related to the presence of the classical subsystem [30].

$$\hat{H}_{\text{eff}} = \hat{H}_{\text{QM}} + \hat{H}_{\text{env}}$$

In turn,  $\hat{H}_{\text{env}}$  contains a term ( $\hat{H}_{\text{MM}}$ ) corresponding to the MM force field description of the classical subsystem, and a term ( $\hat{H}_{\text{QM-MM}}$ ) coupling the two descriptions:

$$\hat{H}_{\text{env}} = \hat{H}_{\text{MM}} + \hat{H}_{\text{QM-MM}}$$

Many different approaches for the evaluation of  $\hat{H}_{\text{QM-MM}}$  have been developed during the years. A recent one allows mutual polarization effects between the QM and MM subsystems [30].

It may be interesting to consider the issue of the boundary somewhat in more detail. Molecules are always moving in a liquid, because their intermolecular interactions are not strong enough to freeze their arrangement into a fixed structure. This is true also for solute-solvent interactions, including H-bonds (which are the strongest ones). Therefore, the solvent molecules included in the first solvation layer exchange fast with time: each of them is replaced by another molecule from the bulk solvent, and moves away from the vicinity of the solute to become part of the bulk solvent. This is tantamount to a continuous exchange of solvent molecules between the QM and MM part of the simulation. Shiga and Masia suggested an approach restraining the solvent molecules from departing

from the QM region, while the bounding surface is allowed to fluctuate during the simulation, to enable improvements of the geometrical definition of the QM region [61]. Takahashi and co-workers developed a “boundary constraint with correction” approach to pursue the same objective [62]. Adaptive QM/MM models attempt to introduce greater flexibility for the boundary concept. Zheng and Waller provide a review of adaptive QM/MM approaches, where the ‘adaptive’ term refers to the partition of the system into more than two regions, namely: a QM-core region corresponding to the solute molecule; a QM-adaptive region comprising the solvent molecules that are treated at the QM level of theory dynamically; a transition region comprising the solvent molecules which can be viewed as having partial QM and partial MM character; and an MM region comprising all the solvent molecules beyond the transition region, which are treated with a force field [63]. The review also analyses the main approaches for a better identification of these regions. Duster and co-workers provide a review largely focusing on the merits and pitfalls of the various adaptive treatments, also considering perspectives for future improvements [64].

In a QM/continuum hybrid approach, a supermolecular structure (adduct) with a suitable number of explicit solvent molecules takes the role of solute, the cavity is built around it, and the rest of the solvent is viewed as a continuum with dielectric constant  $\epsilon$ . The problem becomes a typical ASC problem, with the polarisation of the medium outside the cavity modelled by a system of apparent surface charges on the surface of the cavity. The cavity surface and the apparent charges on it realise the coupling between the QM and continuum descriptions (some examples in [4]).

### 3. Applications Relevant to Industry-Related Issues

Considering the properties and effects of solvents is crucial for all the processes that occur in solution. In industrial processes, solvents may have the following major roles: reactants, reaction media, separation agents, and transportation agents for both mass and heat transfer [55]. In living organisms, they provide the medium within which processes occur. In research, they are used for the roles that they play in the processes under investigation. The present section attempts a sufficiently comprehensive overview of applications of computational studies of the properties of solutions for a variety of research and practical questions.

#### 3.1. The Search for Green Solvents

Solvents play a crucial role in organic and inorganic syntheses as well as in extraction processes, being therefore very important for industrial production processes. Careful solvent selection is often essential to reduce process costs and to make a process greener. Volatile organic compounds (VOCs, mostly of petrochemical origin) have been used extensively for decades, as they are comparatively easy to remove from reaction mixtures or extracted materials; for instance, n-hexane has the advantages of low polarity, optimal boiling point, ease of removal, and stability. On the other hand, the nature and high volatility of VOC pose significant risks to human and animal health. Exposure to VOCs, whether through skin contact, inhalation, or contact with mucous membranes, can prompt various health issues, including irritation, nausea, and dizziness; in the long term, it can cause damages to vital organs such as liver, kidneys, and central nervous system [65,66]. This has prompted active search for more benign solvents (green solvents, [16,67–69]), often integrated with the design of more benign processes [11–15].

The traditional trial-and-error approach to solvent selection is highly expensive and time-consuming, also in view of the high number of possible solvents for each task. The continuous improvement of theoretical and modelling methods and the huge growth of computing power enable the integration of computational methods into the selection and design of solvents [55,70].

#### 3.2. Predicting Solubility

The first requirement for a solvent to be suitable is its ability to dissolve the target solute. The solubility thus becomes a key selection criterion. It is usually defined as the maximum concentration

of a specified solute that can be present in a specified solvent at a given temperature. Predicting solubility becomes crucial in the selection and design of suitable solvents.

A solubility parameter ( $\delta$ ) is defined as a property related to the intermolecular interactions of a pure substance; it provides indications of the total cohesive forces holding the molecules together in a given liquid or an amorphous solid [71,72]. It is equal to the square root of its cohesive energy density, CED. CED was initially defined as the ratio of the energy of vaporization,  $\Delta U_{vap}$ , to the molar volume,  $V_m$ ; this definition makes it an empirical quantity, which can be calculated from experimental values; it is suitable for non-polar and non-associating systems [72]. Crowley [73] and later Hansen [71] proposed to split this parameter into three components, related to the types of intermolecular interactions in a liquid:

$$\delta = \sqrt{(\delta_d^2 + \delta_p^2 + \delta_h^2)}$$

where  $\delta_d$  is the dispersive component,  $\delta_p$  is the polar component, and  $\delta_h$  represents the H-bonding component. By reporting each of these components on one of the axes in a 3D diagram, Hansen developed the Hansen Solubility Parameter (HSP) 3D space, in which a material is represented by a point identified by its three components [72]. Two liquids with close HSP parameters are likely miscible. The parameters are now-a-day often obtained from computational simulations [72].

COSMOS-RS is widely used to compare the solubilities of a variety of compounds in different solvents, as well as other solvents' properties [74–78]. Being based on quantum chemistry, it does not require experimental data.

### 3.3. Solvents for Extraction Processes

A variety of compounds are obtained from natural sources (herbs, leaves, barks, roots, fruits, etc) via extraction processes. The compounds to be extracted must have good solubility in the solvent used for the extraction.

Greener solvents, meant to replace those of petrochemical origin like VOC, can be derived from natural sources, including agri-food byproducts (e.g., orange and grape peels, mangosteen pericarps) and their extraction ability can be predicted and subsequently tested [78–81]. The dissolving power and selectivity of these solvents can be predicted using computational methods [82]. Filly and co-workers utilised HSP simulation and experimental studies to evaluate the performance of nine alternative solvents (pinene, MeTHF, ethyl acetate, methyl acetate, ethyl lactate, butanol, isopropanol, ethanol, and CO<sub>2</sub> supercritical fluid) with respect of that of n-hexane (a VOC) for the extraction of aromas from blackcurrant buds (*Ribes nigrum L.*); the results indicated MeTHF as the most promising for n-hexane substitution [83]. They also evaluated the performance of eight solvents that could replace n-hexane for the extraction of food aromas from caraway seeds (*Carum carvi L.*) using both COSMOS-RS studies and experimental solubility profile; the results indicated ethyl acetate and dimethylcarbonate as promising alternative solvents [78].

Extracts from mangosteen (*Garcinia mangostana L.*, a plant common in Eastern and Southern Thailand) exhibit a wide range of pharmacological activities, including antioxidant, anticancer, antimicrobial, anti-inflammatory, and wound-healing properties [84–87], and they have been incorporated into various commercial products, including nutritional supplements, pharmaceuticals, and cosmetics [88]. Traditionally, dichloromethane (a VOC) was the preferred solvent for the extraction of  $\alpha$ -mangostin (the main active compound) from mangosteen pericarps [89]. However, its toxicity recommends its replacement with safer solvents. Bundeeshomchok and co-workers compared the effectiveness of potentially alternative solvents, including d-limonene, DMC, ethanol, ethyl acetate (EtAc), ethyl lactate, and methyltetrahydrofuran (MeTHF), with that of dichloromethane [82], utilising both HSP and COSMO-RS. HSP analysis indicated dichloromethane as the most suitable solvent, whereas the COSMO-RS analysis indicated that  $\alpha$ -mangostin has greater solubility in ethyl lactate, DMC, MeTHF, ethyl acetate, and ethanol. Experimental studies (classical reflux extraction and HPLC analysis) confirmed the COSMO-RS simulation's predictions [82], thus indicating those solvents as promising alternative greener solvents. All this confirms that COSMO-RS is more suitable for screening solvents for the extraction of complex molecules [82].

Rosemary (*Rosmarinus officinalis* L.) contains several bioactive compounds (phenolic compounds, such as carnosol, carnosic acid, and rosmarinic acid, and volatile compounds from essential oil like  $\alpha$ -pinene, camphor, eucalyptol, or 1,8-cineole); it therefore exhibits antioxidant, anticancer, diuretic, antimicrobial, antiproliferative, anti-inflammatory and anti-hyperglycemic properties [90]. Nutrizio and co-workers used both HSP analysis and COSMO-RS simulations to assess the viability of ethyl acetate, methylacetate, ethanol, 1-butanol, isopropanol, methanol, CPME, dimethylcarbonate and MeTHF as possible alternatives to the traditionally used n-hexane [90]. Although the trends from the two sets of results were largely similar, COSMO-RS gave better solubility results than HSPs for extraction with ethanol, and closer-to-experimental results for the solubility of camphor and borneol in ethanol [90].

The determination of lipids is important in food chemistry because of the effects of different types of lipids on human health. Lipids are usually classified into nonpolar lipids {triglycerides (TAGs), diglycerides (DAGs), monoglycerides (MAGs), and sterols} and polar lipids {free fatty acids (FFAs), phospholipids, and sphingolipids} [91]. The long-chain Omega-3 polyunsaturated fatty acids present in salmon fish have received particular attention because they can reduce the risk of cardiovascular events such as sudden cardiac death, coronary heart disease, and congestive heart failure [91]. Cascant and co-workers utilised both HSP analysis and COSMO-RS simulations to assess the solvation properties of different solvents for different lipid classes found in salmon fish oil [91]. Both models indicated that d-limonene and p-cymene had abilities similar to n-hexane for solvating TAGs. HSP predicted that cyclopentyl methyl ether (CPME) could be the most suitable solvent for all the considered compounds. COSMO-RS values indicated that CPME, ethyl acetate (EtAc), and methyltetrahydrofuran (MeTHF) had similar solvation capabilities for TAGs, DAGs, FFAs and ergosterol; however, experimental data showed that TAGs were the predominant lipid class (73-77%), followed by DAGs (15-20%), FFAs (5-6%), and ergosterol (2-3%) in these solvents. On the other hand, p-cymene and limonene extracts contained higher levels of DAGs and lower amounts of TAGs compared to other solvent extracts. The discrepancy could be attributed to lipid degradation caused by the elevated temperatures needed when using p-cymene and limonene – a factor that is not included in the simulation; therefore, p-cymene and limonene should theoretically be the best solvents for dissolving TAGs [91].

In recent times, COSMO-RS is often the only simulation option utilised for solvent screening and assessment. Touaibia and co-workers utilized it to assess the capability of bio-based solvents chloropinane and chloromenthene – which can be obtained from pinene and limonene – to solubilize  $\beta$ -carotenoids, vanillin, and rosmarinic acid [54]. The results indicated that these two solvent have 3.5 to 2 times greater efficiency in solubilizing the target compounds than hexane. Moreover,  $\beta$ -carotene and vanillin exhibited 6 to 20 times greater solubility in chloropinane than in hexane. This highlights potentialities for chloropinane and chloromenthene as green solvents; a final confirmation requires additional investigation of their bioaccumulation potential and their ecotoxicological profile. Yara-Varon and co-workers used HSP and COSMO-RS to evaluate the performance of 2-methyltetrahydrofuran (2-MeTHF), dimethyl carbonate (DMC), cyclopentyl methyl ether (CPME), isopropyl alcohol (IPA) and ethyl acetate, as possible substitutes of n-hexane in the extraction of carotenoids from carrots [92]. HSPs analysis indicated non-polar or slightly polar solvents as the most suitable. COSMO-RS analysis indicated higher probability of solubility of the carotenoids in CPME, 2-MeTHF and ethyl acetate than in n-hexane, and this was confirmed by experimental results.

The wastes from the food industry may contain high valuable compounds. An example is offered by orange peels, which contain fermentable sugars, carbohydrate polymers, flavonoids, polyphenols and essential oils; their extraction would respond to the valorisation of a renewable source of high value-added chemicals (perfectly in line with the green chemistry principles). Ozturk and co-workers used COSMO-RS simulation to conduct preliminary solvent screenings and identified bio-based CPME and 2-MeTHF as promising options, capable of increasing limonene extraction yields from orange peels up to 80% and 40% with respect to hexane [93]. The recyclability of these solvents further increases their potential for the development of sustainable bio-refineries for citrus waste valorisation.

In homogeneous catalysis, the catalyst is dissolved in the reaction mixture, and it needs to be recovered from it at the end of the process. In their search for eco-friendly solvents to replace dimethylformamide (DMF, developmentally toxic) for the catalyst separation, Linke and co-workers [94] utilised COSMO-RS for thermodynamic property prediction; they also performed a screening of the solvents' compatibility with environmental, health, and safety (EHS) criteria using VEGA, a toolbox containing 33 QSAR (quantitative structure activity relationships, [95]) models in association with 15 different EHS properties [96]. Diethyl sulfoxide (DESO) emerged as a highly promising solvent, outperforming DMF, and potentially capable to replace DMF in other applications [94].

### 3.4. Deep Eutectic Solvents

Deep Eutectic Solvents (DES) were introduced by Abbott and co-workers between 2002 and 2003 [97]. They are mixtures of two compounds which, when mixed, form a eutectic system (a system that melts at much lower temperature than each of the starting components). The molecules in a DES form a network held together by intermolecular H-bonds and/or Van der Waals interactions; these interactions force the DES to remain liquid in a wide range of temperatures. Their ability to establish combinations of H-bond donor-acceptor interactions allows the tailoring of the physical and chemical properties and phase behaviours of DES, making them suitable for being employed as versatile solvents [98,99]. When the starting compounds are abundant in natural sources (like sugars, amino acids, organic acids and choline derivatives), their mixtures are termed Natural Deep Eutectic Solvents (NADES, [100]). NADES respond to the green chemistry criterion of favouring materials from natural sources. Their properties (adjustable viscosity, negligible volatility, capacity to dissolve several less polar compounds, ability to remain in the liquid phase at temperatures below 0°C, limited or null toxicity, biocompatibility, and low cost) make them promising green alternatives to environmentally-harmful solvents [99,101,102], and also to ionic liquids, whose toxicity makes their 'greenness' uncertain [103]. Their application has been explored for selective absorption on contaminants such as aromatic compounds, pharmaceuticals and personal care products, heavy metal ions, and other hazardous materials, including gaseous ones [104] and also as solvents for syntheses, biomass processing, electrolytes for energy storage devices, and metal processing [105]. Because of their nature, the interaction in a DES mixture, as well as the interactions with a potential solute, can be investigated with standard QM methods.

El Kantar and co-workers investigated the use of different DES to extract polyphenols (especially naringin) from grapefruit peels [102]. The experimental procedures utilised high voltage electrical discharge and showed that DES, or a mixture of glycerol and water, are promising green solvents to replace VOCs for this extraction. HSP predictions proved consistent with experimental results.

Most DES are mixtures of molecules containing H-bond donors (HBDs) and H-bond acceptors (HBAs). Given the high number of HBAs and HBDs, an experimental trial-and-error approach to identify DESs suitable for a specific purpose would be faced with the consideration of millions of potential combinations [104]. Computational approaches can make the design and selection rational. An emerging strategy utilises Density Functional Theory (DFT) to this purpose. Zhu and co-workers [104] used DFT calculations to prepare a predictive virtual library for a theoretical screening of possible DES to be used for the adsorption of mephedrone (4-methylmethcathinone). On the basis of their binding energies, three types of DES were then selected for synthesis, using lactic acid as HBD and benzyltributylammonium chloride, choline chloride, and tetrabutylammonium chloride in turn as HBAs. These DES were associated with zeolite imidazoline framework-8 (a metal organic framework, MOF) and shrimp shells as a biomass source to obtain a DES-functionalized ZIF-8/biochar with high selectivity for mephedrone. Comprehensive characterizations showed that DES impregnation of the MOF regulated pore space and introduced additional adsorption sites. The DES with benzyltributylammonium chloride showed better performance than the others, and maintained stable adsorption capacity over a wide (5-11) pH range; its greater affinity was attributed to pore-filling, H-bonding, and  $\pi$ - $\pi$  interactions made possible by its aromatic properties. Overall, the study highlights the potential of tailored DES-functionalized materials for selective adsorption properties, and the potential of computational approaches to make the DES design rational.

### 3.5. Ionic Liquids

#### 3.5.1. Nature and Properties of Ionic Liquids

Ionic liquids (ILs) are molten salts formed by an organic cation and an organic or inorganic anion [65]. Due to their characteristics (asymmetrical and often larger cations), the ions do not form crystals as easily as other salts; many of these compounds remain liquid at room temperature, and their melting point are below 100°C. ILs also have negligible vapour pressure for temperatures below 400°C, good ionic conductivity, high thermal and chemical stability, and low flammability. These characteristics make them preferable to organic solvents, and they are currently used in many industrial applications [106]. On the other hand, it is not easy to decide whether they can be considered green solvents, because of some important drawbacks which do not make them environmentally safe. Their synthesis process involves many steps and includes the use of solvents harmful to human health and the environment. The identification of economically viable large-scale recovery options after use is still a challenge. The main currently used ILs are non-biodegradable and toxic in nature, and their toxicities vary considerably for different organisms. All this make ILs not eco-friendly [65].

Trying to obtain ILs with both low toxicity and good biodegradability is a major challenge. Computational modelling, including QSPR (Quantitative Structure-Property Relationship) analysis, can be of great help, making the search for 'greener' ILs rational [65,105,107]. Combinations of several theoretical models and approaches have enabled the prediction of many properties [108]. On the other hand, the nature of ILs poses various challenges to their modelling. In a conventional medium, solvent-solute interactions are generally predominant while, in ILs, interactions within the solvent can become more important. The presence of ion-ion interactions among the ions constituting an IL may generate clusters, making its bulk inhomogeneous; this makes it difficult to transfer continuum-type models to the study of ILs and of molecules dissolved in them [108]. The performance of DFT is not ideal because of the presence of delocalized charges in the constituting molecular ions. Ab initio molecular dynamics (AIMD), which combines electronic structure calculations with conventional MD, appears to be the best approach to predict the intermolecular structure of ILs [108]. The next subsections outline some of the current applications of ILs that have also been objects of computational studies.

#### 3.5.2. Representative Models and Applications of Ionic Liquids

Bruzzone et al. [109] developed QSAR analyses for the prediction of aquatic toxicity (modelled as toxicity to *Vibrio fischeri*) of 33 ILs containing chloride or bromide anions and different cations; this choice was motivated by the observation that ILs with the same cation and different anions do not show any statistical toxicity difference. They determined molecular properties by optimising the cations at the DFT level (including the Natural Bond Orbital method, NBO, [110,111], to obtain partial atomic charges), and performed calculations both in the gas phase and in water solution, using the PCM model for the latter. Then, they used the CODESSA program to calculate the selected molecular descriptors and to derive correlations between the descriptors and the toxicity data. Satisfactory correlations with the same descriptors were found for both phases, but the descriptors calculated in water solution provided better correlation and were inherently more reliable because they referred to the actual medium where the toxicity action occurs. The results are also consistent with the observation that toxicity is related to the cation branching and to the ability of its lipophilic part to intercalate into the cell membrane. The model proved also able to predict the correct order of magnitude of the toxicity.

Catalysis is fundamental in green chemistry, to enable processes to occur under sufficiently benign conditions [112,113]. The 'green' tendency to use natural products or mimic natural processes has prompted the exploration of a variety of biological or biomimetic catalysts. For instance, the enzyme *Candida antarctica* lipase B (CALB) is an efficient biocatalyst for hydrolysis and esterification. Its activity also depends on the solvent used, and the search for benign solvent is active. Several works showed that its activity in suitable ILs is good; for instance, the conversion rate in fatty acid methyl ester synthesis using CALB in [Emim][TfO] IL is greater than in tert-butanol [114]. Subtle

interactions between an enzyme and the IL in which it is dispersed influence the enzyme's conformation, thus determining whether its activity in the given IL increases or decreases with respect to organic solvents [115]. Kim and co-workers investigated the catalytic activity of CALB for the lipase-catalyzed trans-esterification of butyl alcohol with vinyl acetate to produce butyl acetate, in four different solvents – [Bmim][TfO] IL, [Bmim][Cl] IL, tert-butanol, and 0.3 M NaCl solution. The study involved both experimental components and MD simulations [115]. The experimental results demonstrated that the enzymatic activity followed the order: [bmim][TfO] > tert-butanol > [bmim][Cl]. The MD simulations indicated that the structure of the catalytic cavity is solvent-dependent: the cavity can be open or closed in water, is open in [Bmim][TfO] and tert-butanol, and closed in [Bmim][Cl].

Since ILs are meant as solvents, it is important to predict the solubility in them of the compounds of interest. Katritzky and co-workers developed QSPR predictive models for the Ostwald solubility coefficient ( $\log L$ ) and partition coefficients ( $\log P$ ), considering 92 organic solutes and eight different ILs [116]. The QSPR analysis utilised molecular descriptors calculated solely from the molecules' structures – charge-related, geometrical, topological and QM descriptors. The most important descriptors (including H-bond donor/acceptor abilities) in the  $\log L$  models relate to the charge distribution within the molecules, which, in turn, relates to the electrostatic interactions between the solute and the IL. Geometrical descriptors account for the effects of the size and shape of the solute molecule. The QM descriptors reflect the interatomic interactions within molecules. The predictive power of the models proved satisfactory.

### 3.5.3. How Ionic Liquids Dissolve Cellulose

ILs have the ability to dissolve cellulose (the main component of biomass of plant origin), primarily thanks to the formation of H-bonds and van der Waals interactions [117]; this is particularly interesting because cellulose does not dissolve in most other solvents. Several works investigated the possible mechanism of the dissolution process. Initial studies considered the solvation of glucose, the monomer forming the long cellulose chains. Youngs and co-workers used MD simulations to study the solvation of glucose in 1,3-dimethylimidazolium chloride ([dmim][Cl]), and found that the anions interact directly with the OH groups of glucose, while the cations interacted with the anions bonded to glucose; the anions were thus forming the first solvation layer [118,119].

Derecsei and co-workers performed the first MD simulation of the interactions involved, using polysaccharides of different lengths to mimic cellulose and calculating solubility parameters; they found that the solubility parameters for longer oligomers were significantly lower than those of the monomer [120]. Rabideau and co-workers performed an MD simulation focusing on the role of the ions, and specifically of the cations [121]. They considered combinations of chloride, acetate, and dimethylphosphate anions with alkylimidazolium-based cations of increasing tail length, and confirmed that the anions bind to the OH groups of the polysaccharide via H-bonds, while the cations interact through dispersion interactions with the nonpolar regions of cellulose and electrostatically with the anions bound to the OHs. This leads to the formation of networks of alternating cations and anions. The tail length causes only very minor effects on the solvation structure and overall interaction energies, although it modifies the framework of the anions-cellulose H-bonds. The imidazolium ring also allows for the formation of anion-cation chains and networks at the cellulose surface, resulting in strong cellulose-IL interactions without significant disruption of the IL bulk connectivity. Payal and co-workers focused more specifically on the role of the anion, considering cellobiose as a model for cellulose, and utilising ILs containing the [bmim] cation and one of the following anions:  $[OAc]^-$ ,  $[NO_3]^-$ ,  $[Cl]^-$ ,  $[BF_4]^-$ ,  $[PF_6]^-$ ,  $[CF_3SO_3]^-$ , and  $[NTf_2]^-$  [122]. MD simulations confirmed that the dissolution mechanism entails the disruption of the intermolecular and intramolecular cellobiose H-bonding network and that the anions play crucial role by H-bonding to the OHs.

Cao and co-workers [117] modelled the dissolution process utilising cellobiose and 1-butyl-3-methylimidazolium acetate (BmimAc) as model system for the calculations, and compared the results with experimental information. The calculations were performed at the M062X/6-311++G\*\* level, with

the addition of approaches (NBO, AIM [123]) meant to obtain information on charges on the atoms and on non-covalent interactions. The dissolution process entails the removal of intramolecular H-bonds in cellulose and the formation of cellulose-acetate intermolecular H-bonds. The cooperation of H-bonds and van der Waals interactions between cellobiose and the IL determines the IL's dissolution ability.

Li and co-authors [124] provide an extensive review of experimental and computational investigations of the mechanism of the dissolution of cellulose in ILs, also highlighting the challenges. One of the problems is the fact that using glucose, cellobiose or polysaccharides as models for cellulose is not adequately informative because they are too different from cellulose (polysaccharides with DP < 6 can dissolve in water). The stability of natural cellulose in common solvents is due to its crystalline form. Several scientists tried to study larger cellulose microfibrils. The same group used a 500ns MD simulation on a cellulose bunch consisting of 7 glucan chains (DP=8) put in [Emim][OAc] and [Emim][Cl] to investigate the dissolving process [125]. They found that complete dissolution happened in [Emim][OAc], with every single chain separated from each other. The original H-bonding network was destroyed by ILs and replaced by a new anion-cellulose H-bonding network. [OAc]<sup>-</sup> could form three different kinds of H-bonds within cellulose chains, thus being able to separate them, while [Cl]<sup>-</sup> was not effective. The anions initially form H-bonds with the closest hydroxyl groups of cellulose, thus inserting itself into the strand; as the number of anions bound to the cellulose chains increases, the cations start intercalating between the strands, thanks to their electrostatic interactions with the anions and the van der Waals interactions with the cellulose strands; this starts the dissolution process (separation of the strands from each other). In addition, they found that cellulose can only dissolve in ILs containing unsaturated cations [125].

Some works investigated the effect of the addition of a co-solvent on the dissolution of cellulose. Zhao and co-workers used MD simulations and quantum chemistry calculations to study the effects of dimethylsulfoxide (DMSO), DMF, CH<sub>3</sub>OH and water as co-solvents in [Bmim][OAc], [126]. The presence of the co-solvent influences the H-bond interactions between the anions and the OHs of cellulose, thus influencing its solubility. Protic solvents (CH<sub>3</sub>OH and H<sub>2</sub>O) have a strong tendency to solvate [CH<sub>3</sub>COO]<sup>-</sup>: they compete with the cellulose-[CH<sub>3</sub>COO]<sup>-</sup> H-bond interaction, leading to decreased cellulose solubility. Aprotic solvents (DMSO and DMF) can solvate the cation and anion of the IL, weakening the cation-anion interactions and therefore making the anions more available to form H-bonds with the OHs of cellulose. Huo and co-workers obtained similar results through MD simulations [127]; they also added a criterion termed 'pair energy distribution' (PED) to the analysis of H-bond patterns, and evaluated it between cellulose and cations, anions, and co-solvent molecules near the interface. The results showed that the PEDs between anions and cellulose are sensitive to the addition of co-solvents and to the type of anions, and this can be used as an additional indicator of the ability of the 'IL + co-solvent' to dissolve cellulose. Veliogly and co-workers also used MD simulations to study the dissolution of a cellulose I $\beta$  microcrystal in the 1-butyl-3-methylimidazolium acetate IL, and in the same IL with the addition of DMSO as co-solvent [128]. They found that DMSO does not interact strongly with cellulose and does not interfere with the interactions of cellulose with the anions and cations; its presence increases the velocity of mass transport and dissolution because it lowers the viscosity of the medium. Parthasarathi and co-workers used MD simulations to study the dissolution of a cellulose I $\beta$  microcrystal in the 1-ethyl-3-methylimidazolium acetate ([C<sub>2</sub>C<sub>1</sub>Im][OAc]) IL, to which different amounts of water (co-solvent) were added [129]. They found that a comparatively small proportion of water favours the dissolution of cellulose, whereas higher proportions hamper it. Rabideau and Ismail used MD simulations to study the dynamics of the formation of the H-bonds between IL anions and cellulose and the effects of the presence of water in the IL [130]. They examined fifteen different ILs with 1-alkyl-3-methylimidazolium cations ([C<sub>n</sub>mim], n = 1, 2, 3, 4, 5) as cation and chloride, acetate or dimethylphosphate (DMP) as anions, for the study of cellulose-anion H-bonding, and [C<sub>2</sub>mim] in combination with each of the three just-mentioned anions for the presence of water. They found that increased tail length in the cation has only minor effects on the transition of the anions' H-bonding to cellulose into different bonding states but, by tending to slow the transitions' dynamics, it increases the H-bond lifetimes. Each anion can form up

to four H-bonds with cellulose, and the bonding lifetimes of multiply bonded anions are three to four times longer than singly bonded anions. When water molecules are present together with the IL, they surround the H-bond accepting sites of the anions, hampering their interactions with cellulose; if an already formed H-bond between an anion and cellulose breaks, the anion also gets surrounded by water molecules and becomes unavailable for new H-bonds with cellulose. Greater concentration of water corresponds to greater drop in anion-cellulose H-bonding [130].

### 3.6. Nanoparticles in Liquid Media

#### 3.6.1. Nanoparticles and Their Properties

Nanoparticles (NPs) are aggregates of many atoms (metal atoms, metal oxides, or carbon atoms), usually between 1 and 100 nm in size. Their applications (nanomaterials, nanotechnology) are continuously expanding. By 2009, already more than 500 consumer products on the market contained elements of nanoscience and nanotechnology [131]. NPs are characterised by high specific surface area. A comprehensive overview of the different classes of NPs, their physicochemical properties (mechanical, thermal, magnetic, electronic, optical, and catalytic), and their applications in medicine, electronics, water treatment, energy storage, agriculture, and food production is offered in [132].

The increasing utilization of NPs entails increasing presence in the environment, prompting studies of their possible impacts, first of all toxicity. Because of their size, they are not stopped by the protective barriers in the upper airways (such as the cilia in the nose) and can reach the lungs if present in the air; for instance, NPs with an aerodynamic diameter of less than 100 nm are deposited mainly in the alveolar region, while NPs with a smaller diameter can penetrate deeper into the lungs [133]. They can also contaminate media like water and soil. An overview of their toxic effects is included in [134], and an overview of their presence, interactions, transformation processes, possible toxicity forms and the factors affecting the toxicity in aquatic environment is offered in [135]. Computational studies can examine or predict the interactions of NPs with biological systems and with environmental media like water, and are expected to help in the design of safer NPs, as well as appropriate storage systems, to protect nanomaterials from the environment and simultaneously protect the environment (and us) from NPs [136]. Gajewicz and co-workers strongly recommend the use of computational modelling to evaluate the properties of each NP and use the obtained information to conduct risk assessment [137].

Because of their size, NPs are considered more dispersed than dissolved in a liquid medium, and the resulting suspension often has colloidal properties. Biological applications require NPs to be water dispersible and to remain soluble, whereas many catalytic and/or industrial applications require nanoparticles to be dispersed in organic solvents [138]. Understanding the dispersion and stability of NPs in liquid media is often crucial for nanomaterial utilization [139], also because many NPs are synthesised in a liquid and the nature of the used solvent often governs the internal phase structure of the formed NPs [140]. This requires understanding the molecular interactions between suspended NPs and the surrounding solvent [141]. Leekumjor and co-workers used a coarse-grid computational model to investigate the molecular interactions between oleate-capped NPs and various solvents and found that the solvent polarity correlated better than other solvent properties to both simulation and experimental results. Xu and co-workers developed an improved coarse-grained MD simulation for NPs and applied it to the study of gel ink [139]. Electrostatic interactions and London-van der Waals interactions are the most frequently considered interactions for NPs in an electrolytic medium [142]. The presence and nature of a solvent plays a significant role in the interactions between two NPs [143] and several studies investigate these interactions. The development of specific computational models to treat some properties typical of solutions can be important for better utilisations of NPs; for instance, a theoretical or computational model to estimate the osmotic pressure created by the magnetic NPs and the water flux in the forward osmosis process would be important for a more efficient utilisation of NPs in water treatment and desalination [144].

#### 3.6.2. Carbon Nanotubes

Carbon nanotubes (CNTs) consist exclusively of carbon atoms arranged in a series of condensed benzene rings rolled-up into a tubular structure [137]. The toxicity of carbon-based nanomaterials depends on several properties, such as structure, length, surface area, surface charge, aggregation, oxidation, functional groups, manufacturing methods, morphology, concentration, and dosage [131,133,145–149].

CNTs are not soluble in water. On the other hand, they need to be dispersed/dissolved in water for biomedical applications, because water is the main medium in living organisms. The problem may be partly overcome by chemical functionalization of the CNT surface with groups that increase their compatibility with water [150]. Sasidharan and co-workers reported that CNTs functionalised with carboxylic groups showed excellent dispersion in aqueous medium, whereas the pristine (non-functionalised) CNTs do not disperse; this also poses the question of additionally investigating potential toxicity effects of functionalised CNTs [147]. Computational simulations and modelling are highly recommended for objectives like improving the biodistribution, pharmacokinetics, and solubility of CNTs, as well as diminishing their toxicity [150].

CNTs can interact with DNA molecules in the presence of a solvent. Gao and co-workers report the results of MD simulations of the dynamic processes of encapsulating DNA inside a CNT in water environment; the results indicate spontaneous insertion of single-strand DNA oligonucleotide thanks to the combined action of van der Waals and hydrophobic interactions [151].

Single-walled carbon nanotubes (SWCNTs) are largely used in medical research, for efficient drug delivery, and for biosensing methods for disease treatment and health monitoring [152]. Mananghaya and co-workers used DFT calculations to investigate how the SWCNT properties change following covalent functionalization with selected organic acid groups either on the walls or on the tips of SWCNTs. They found that the functionalization is thermodynamically favourable. Calculation of  $\Delta G_{\text{solv}}$  with the PCM model showed that the solubility of the functionalised SWCNT improved. By enhancing SWCNT biocompatibility, suitable functionalization also reduces toxicity and the chance of tissue accumulation [150].

Cisplatin (cis-PtCl<sub>2</sub>(NH<sub>3</sub>)<sub>2</sub>) is an anticancer drug which crosslinks DNA in several different ways, interfering with cell division by mitosis. Monajjemi and Mollaamin used Monte Carlo and DFT calculations, with the Onsager model of solvent (the first SCRF model, initially proposed in [153]) to investigate the thermodynamic properties (Gibbs free energy, enthalpy, energy and entropy) and dielectric effects of a cisplatin-SWCNT combination in different solvents, and found that the combination has more effective action on DNA, thus supporting the suitability of SWCNTs for drug delivery in different media [152].

The use of surfactants is one of the options to favour CNTs' dispersion. Obata and Honda used MD simulations to study the effect of surfactants on CNTs in aqueous environment [154]. They used a biosurfactant (dipalmitoyl phosphatidylcoline, DPPC, a primary component of human lung surfactants) and an artificial surfactant that is often used to disperse CNTs (polysorbate 80, Tween 80). The CNT-DPPC complex was energetically more stable than the CNT-Tween 80 complex. The MD simulations also indicated that the adsorption and desorption of surfactants on the CNT surface occurred on a millisecond timescale in aqueous environment; this suggests that CNT-surfactant complexes can change into more energetically stable complexes within biological systems through a surfactant exchange in which biosurfactants present in greater concentration replace the initial one through adsorption and desorption on the CNT surface. They also suggest that the biological effects of CNTs (including toxicity) are related more to the surface properties of the CNT-surfactants complexes than to the intrinsic surface properties of pristine CNTs [154].

### 3.6.3. Silica Nanoparticles

NPs may also be built from silica (SiO<sub>2</sub>). Carmofur (HCFU), an antineoplastic drug used in cancer treatment, exhibited promising activity against the SARS-CoV-2 (Covid-19) virus. Nivetha and co-workers used theoretical methods to understand the absorption mechanism and interaction between the HCFU molecule and noble metal (Ag/Au/Pt)-loaded silica nanocomposites (HCFU + Ag/Au/Pt-SiO<sub>2</sub>) [155]. They used DFT calculations with the LANL2DZ pseudopotential to obtain

optimized geometries, vibrational frequencies, and interactions of the HCFU molecule and its complexes. They also used theoretical methods such as HOMO-LUMO analysis, Molecular Electrostatic Potential (MEP) mapping, Electron Localized Function (ELF), and Localized Orbital Locator (LOL) to assess chemical reactivity, charge distribution, and other molecular properties. They investigated UV-visible spectra in polar protic and aprotic solvents (ethanol, methanol, water, and DMSO) using Time-Dependent DFT (TD-DFT) and added molecular docking predictions to analyse the interactions between HCFU and its complexes with relevant selected proteins.

### 3.6.4. Multiscale Modelling Options

As already mentioned, many NPs are formed in solution, and the nature of the medium (one solvent, or a solvent with a co-solvent) and the conditions of the process influence the nature, shape and properties of the NPs (e.g., their catalytic function [156]). Modelling all this is a complex task, and may require multiscale options. Multiscale modelling investigates the behaviour of a system, or the properties of a material, at a selected level, using information from a different level, where the 'level' may be referred to computational approaches (e.g., QM and MM, discrete and continuum) or to size-range (e.g., atomic/molecular level, nano-level, macroscopic level) [157]. As already seen for the hybrid levels (section 2.3), the most challenging task is the design of the connection between the two levels concerned. Multiscale modelling may be particularly suitable to relate the properties of the resulting NPs to the features of their preparation. For instance, Lavino and co-workers proposed a multiscale model which accounts for kinetic effects during the NP formation process and can address the relevant length scales (molecular, nano and macro) and used poly-e-caprolactone nanoparticles in acetone–water mixture to validate the model [158].

## 3.7. Representative Examples of Other Types of Investigation Topics

### 3.7.1. Organometallic Catalysis

Organometallic catalysis has been attracting considerable interest because of the broad range of possible applications in academia and the industry [159]. Computational methods can help predict the stability of complexes of organic molecules with transition metals, e.g., by predicting the ligand dissociation reaction enthalpies.

Sperger and co-workers provide an extensive review of the computational studies of Pd-, Ni-, Rh-, and Ir-mediated transformations conducted between 2008 and 2014; they consider different types of organic synthesis reactions, explaining the details of the computational approaches and the comparisons with experimental results [159]. They analysed the performance of DFT with different functionals, without and with the inclusion of Grimme's dispersion correction D3 [160], coupled with CPCM (or, in some cases, with COSMO-RS or SMD) for solvent modelling. DFT-D3 generally performed better than DFT. In their conclusions, they observe that most calculations reviewed in their work were conducted post-experimentally to rationalize the observed chemical reactivities, but that the reviewed results encourage the possibility of using computational studies also for predictive roles.

Jacobsen and Cavallo [161] had used DFT-D3 combined with CPCM to study the ligand dissociation reaction enthalpies of several catalysts containing Fe and Ru, including  $(CO)_3Fe(benzylideneacetone)$ ,  $(Cp)Ru-(Cl)(cyclooctadiene)$ ,  $(CO)_3Fe(PMe_3)$ , and  $(CO)_3Fe(PPh_3)$ . They observed good agreement of DFT-D3 results with experimental results for  $(CO)_3Fe(benzylideneacetone)$  and  $(Cp)Ru-(Cl)(cyclooctadiene)$ . However, for the ligand exchange reactions of phosphine ligands of different sizes (e.g.,  $PM_3$  and  $PPh_3$ ) the agreement was not good, possibly because DFT-D3 may overestimate the binding energy of large ligands and organometallic complexes containing them. Grimme obtained better agreement by using a larger basis set and COSMO-RS for the solvent [60]. This also suggests that some caution may be advisable when interpreting dispersion-corrected dissociation energies.

### 3.7.2. The Medium in Lithium Ion Batteries

Lithium ion batteries are rechargeable batteries with metallic lithium at the anode and  $\text{Li}^+$  ions in the electrolyte. Their performance depends on the efficiency with which the ions move through the material in which the electrolytic solution is dispersed; therefore, the lithium ion solvation and diffusion are the important properties to consider on evaluating the suitability of a solvent. The search for new materials with improved efficiency to solvate and transport  $\text{Li}^+$  is quite active [162]. Organic carbonates are commonly used as solvents in lithium-ion batteries because of their capacity to dissolve lithium salts. Rempe and co-workers studied the solvation and diffusion of  $\text{Li}^+$  in ethylene carbonate (EC) and propylene carbonate (PC), with lithium hexafluorophosphate as lithium salt, using both molecular simulations and experiments [162]. The simulations used standard force field MD (FFMD) to evaluate  $\Delta G_{\text{solv}}$  and the solvation structure of the ion, as well as the solvent properties, and ab initio MD (AIMD, which is more costly) to investigate the solvation structure of a single  $\text{Li}^+$  ion. The results showed that electrostatic forces dominate  $\text{Li}^+$  solvation in carbonate-based solvents and in water. The results for the free energy of the transfer of  $\text{Li}^+$  to the carbonate solvents from water, obtained from simulations using scaled partial charges (by 80% for EC and 90% for PC) for the solvent, were also in agreement with experimental data, suggesting that such scaling may be convenient for analogous systems.

Dimethyl carbonate (DMC) has low volatility and is non-flammable, non-toxic and biodegradable; because of this, it has been used for several chemical reactions as well as for electrochemical and extractive applications [163]. Reddy and Balasubramanian's extensive MD simulations found that most molecule in liquid DMC in ambient conditions are in the cis-cis conformation, and only about 2% are in the cis-trans conformation, with the former having around 1.0 D dipole moment and the latter around 4.5D [163]. DMC is a suitable solvent for lithium salts, but it crystallizes at 2 °C, restraining its use in low-temperature applications. To prevent crystallization, it is often mixed with ethylene carbonate (EC).

A central challenge in the refinement of lithium-ion batteries is to control cathode-induced oxidative decomposition of the solvent. Miller and co-workers investigated the electrochemical stability of EC, DMC, and a 1:1 EC-DMC mixture [164]. They used a wavefunction-in-DFT embedding approach, calculating the vertical ionization energy of individual molecules at the CCSD(T) level of theory (coupled cluster with full treatment singles and doubles and perturbation theory treatment of triples) and explicitly accounting for the solvent using a combination of DFT and MM interactions. They found 1.2 eV for the solvent reorganization energies of both EC and DMC. They verified that simple dielectric continuum models accurately reproduce the reorganization energy of EC, but significantly underestimate that of DMC, because of the important role of quadrupolar interactions in DMC solvation, which are not taken into account by standard dielectric continuum models. The quadrupolar interactions may also favour the coordination of  $\text{Li}^+$  cations by DMC in EC-DMC mixtures [164].

### 3.7.3. The Behaviour of Liquid Mixtures

Understanding and predicting the behaviour of liquid mixtures is important both in research and for practical applications, including process design for the chemical industry [165]. Atilhan and Aparicio used MD simulations of dialkylcarbonates and 1-alkanol mixtures with different proportions and at different temperatures, with particular attention to their behaviour at the liquid/gas or liquid/vacuum interfaces, because of its relevance for industrial operations involving phase equilibria, gas adsorption or absorption, and heat or mass transfer. They found that the situation at the interface is completely different than in the bulk liquid, as the 1-alkanol molecules tend to develop a highly dense layer in the vicinity of the vacuum region, while the dialkylcarbonate molecules remain closer to the bulk liquid [165].

### 3.7.4. Solvents Influencing Chemical Reactions

Solvents influence many aspects of the chemical reactions occurring in them, including thermodynamics, kinetics and product selectivity. The effects have been studied since long (e.g., [166,167]), and substantial information has been accumulated through years. Computational methods

may play key roles to elucidate whether an observed effect is due to the solvent. For instance, Kostal and Jorgensen used QM and statistical mechanics calculations to investigate the experimentally observed rate enhancements for the base-initiated cyclization reactions of 2-chloroethoxide derivatives in aqueous solution when the methylation on C1 is increased [168]. Gas-phase QM calculations at the MP2 and CBS-Q levels indicated little intrinsic reactivity difference. On the other hand, continuum solvation calculations, or MC/FEP calculations in explicit water molecules, showed that the reactivity changes were due to the solvent effect [168].

Varghese and Mushrif analyse the pathways and outcomes of these influences, to suggest routes for optimal solvent selection in relation to the desired outcomes [169]. They consider the main computational techniques for the investigation of solvent effects (implicit models, MD, QM/MM, approaches combining DFT calculations and MD) and provide a review of examples to illustrate the application of these methods, the types of information that they provide, and their relevance to catalytic reactions. They envisage that further development of effective combinations of two or more computational techniques can enhance the investigation of the solid-liquid interface, thus enabling better computational investigation of heterogeneous catalysis [169].

Making production processes greener by recycling anthropogenic CO<sub>2</sub> into fuels and chemicals, or making the production of ammonia and hydrogen more sustainable, still present major selectivity and energy efficiency challenges [170]. Basdogan and co-workers suggest that computational quantum chemistry can help interpret and guide experimental work by providing insights into key aspects of reaction mechanisms. They review applications of implicit, mixed implicit/explicit (QM/continuum), and explicit solvation models for the identification of steps that might make these processes (hydrogenation of CO<sub>2</sub>, oxidation of N<sub>2</sub>, reduction of H<sub>2</sub>, and others) more efficient and sustainable, by identifying solvent usage that can produce the desirable effects [170]. They recommend the integration of solvation energy contributions and other environmental parameters into future high-throughput screening approaches, and consider further improvements of the solvation models as crucial for the design of sustainable processes [170]. Zhu and co-workers investigated the electrocatalytic reduction of CO<sub>2</sub> (one of the promising options to convert excess CO<sub>2</sub> in the atmosphere to industrial feedstocks), aiming at better understanding of how alkali cations increase the selectivity and activity of the process [171]. They found that the CO<sub>2</sub> reduction kinetics is closely correlated with the strength of the Onsager reaction field [153,172], which is induced by polarization of the electrolyte by the solute dipole. They conclude that it is important to develop models capable of handling solvation-mediated Onsager reaction fields and interfacial solvation structure with the same relevance as the electric field produced by the electrochemical double layer [171].

### 3.7.5. Ions in Solution

Understanding the solvation structure of cations in solution is crucial for a variety of purposes. The hydration structure of Na<sup>+</sup> and K<sup>+</sup> in liquid water are assumed to play important roles within living organisms, from selectivity to their motion through their respective channels. Rowley and Roux used a QM/MM simulation, where the ion and a selected number of closer water molecules constituted a dynamical and flexible inner region treated at QM level, while the outer-region water molecule were treated with a polarizable MM force field [173]. The results indicated coordination number ranges of 5.7–5.8 for Na<sup>+</sup> and 6.9–7.0 for K<sup>+</sup> [173].

Similarly, the hydration structure of Ca<sup>2+</sup> and Mg<sup>2+</sup> with the presence of OH<sup>-</sup> ions in water is crucial for understanding their roles in biochemical and chemical processes [174]. Liu and co-workers used AIMD simulations to investigate the solvation structures of the three ions (Ca<sup>2+</sup>, Mg<sup>2+</sup>, and OH<sup>-</sup>) and found that the water molecules in the first solvation shell of Ca<sup>2+</sup> change their preferred orientation faster than those of Mg<sup>2+</sup> and that, if the cation reaches the first solvation layer of OH<sup>-</sup>, the bonds of OH<sup>-</sup> are altered and the water molecules adjacent to OH<sup>-</sup> are squeezed [174].

Taking into account the solvation structure of cations in the solvent of interest is important for the applications in which cations play key roles. Battery technology is one such area, linked to the search for sustainable and affordable energy sources and storage [175,176]. The solvation structures

and networks of the cations (such as  $\text{Li}^+$  or  $\text{Na}^+$ ) in the electrolyte have significant influence on the formation and properties of the solid-electrolyte interphases which, in turn, condition battery performance; understanding them is crucial for the design of more efficient batteries [175]. The concerns about the sustainability, as well as production and extraction costs, of lithium is prompting active search for alternative batteries based on abundant elements, such as calcium and magnesium, and, therefore, also for the study of their solvation structures [176]. A review of the roles of the elucidation of solvation structures for the rational design of electrolytes is offered in [177]. Computational chemistry modelling is expected to play crucial roles in the elucidations.

### 3.7.6. Studies in Solution to Understand the Properties of Natural Materials

Materials of natural origin encounter a vast variety of industrial utilizations. Their interactions with solvents often play crucial roles in the understanding of their properties and the design of applications. For instance, computational modelling has proven a powerful instrument to elucidate the causes of the outstanding mechanical properties of silk (a natural material made exclusively of proteins), to predict the properties of other biomaterials that can be derived from it, and to assist in design of new manufacturing strategies; the modelling included studies in water solution, both with implicit solvation models and with MD simulations [178,179].

## 4. Computational Studies Concerning the Solvent Role in the Interactions and Activities of Biologically Active Molecules and Biomolecules

### 4.1. The complexity of biomolecules and biochemical processes

As mentioned at the beginning, all processes occurring within living organisms occur in a medium; water is the dominant component of living organisms and, therefore, also the main medium within which processes occur. Biologically active molecules are molecules that can cause a response when introduced into a living organism by exerting some action; a typical example is offered by drugs, which are used to treat diseases. Biomolecules are molecules that are part of living organisms and have specific roles in the functioning of the organism; some of them – like proteins or DNA – consist of thousands of atoms.

As Finney recalled already in 1996, "Water has been recognised ... as one of the major structuring factors of biomolecules" and "As the natural solvent of biological macromolecules, water influences many aspects of biological functions" [180]. A number of aspects concerning the behaviour of these molecules in solution can be objects of investigation: how individual molecules interact with the solvent and the effects of these interactions on the molecule and its behaviour; the possible roles of the solvent in the interactions between two biomolecules; and the possible roles of the solvent in the interactions between a biologically active molecule and the biomolecule that is its biological target. Given the size and complexity of proteins and DNA, it becomes important to select the portion of greatest interest for high accuracy computational treatment, while the rest is treated at a less costly level; therefore, hybrid approaches are often the most suitable and affordable options.

### 4.2. Solvents and the Structure of Proteins

#### 4.2.1. Water Molecules in and around Proteins' Structures

The importance of water for proteins' structure and behaviour has been recognised since several decades: "The water-protein interaction has long been recognised as a major determinant of chain folding, conformational stability, internal dynamics, and binding specificity of globular protein" [180]; "In general, a protein molecule is surrounded by layers of solvent which mediate its functional conformation as well as its chemical characteristics" [181]; "Water in close proximity to the protein surface is fundamental to protein folding, stability, recognition and activity" [181]; "Water plays many roles on the surface of proteins, filling in gaps and cavities, fulfilling unsatisfied hydrogen bonds and mediating interactions" [182]; "Most of the decisive molecular events in biology take place at the protein-water interface".

The hydration layer surrounding a protein plays key roles for protein folding and membrane stability [183]. Schoenborn and co-workers provide a review of experimental (spectroscopic) and computational (MD) results available by 1995 on the water molecules layer surrounding a protein [181]. The water molecules bind to the polar or partially charged site of the protein; those that bind to the protein's surface through H-bonds, and those bridging them, may form complex H-bond networks, often further stabilised through cooperativity [184]; those that bind to the protein surface with sufficiently strong H-bonds remain also in the protein's crystal structure [181,184]. Both experimental and computational studies show the presence of these water molecules. For instance, both experimental and computational studies indicate that some nonpolar cavities may be hydrated in the crystal structure of the bovine pancreatic trypsin inhibitor and barnase [185].

The water molecules that remain bonded to a certain part of the protein can be considered spatially fixed 'structural' water molecules constituting an integral part of the protein's structure [183]. They play a crucial role in the stabilization of secondary structure, protein activity, flexibility and ligand binding [186]. Thermal stability, including resistance to higher temperatures (thermophilicity, which might be relevant in industrial processes) depends, directly or indirectly, on the coupling between the protein and surrounding water [187]; suitably modulating the presence of structural water can be relevant to the design of thermophilic enzymes with the desired stability range [188].

Understanding the nature of the hydration layer is crucial to understand the protein's biological functions: the layer must interact sufficiently strongly with the protein to stabilize it, but not so strongly as to block surface sites or inhibit structural change [183]. Ebbinghaus and co-workers investigated the extent of the hydration layer and found that both experiments and MD simulations indicated a long-range dynamical hydration shell [189]. Mattea and co-workers investigated the dynamical properties of proteins' hydration layer [190]. Teixeira investigated the stability of the protein-water H-bonds [191] and Born investigated the solvation dynamics of ubiquitin [192]. Wallnoefer and co-workers used MD to investigate the effect of water molecules on factor Xa – a key enzyme in the blood coagulation process – and highlight the importance of a well-determined set of internal water clusters in the protein's initial structure for the quality of the MD simulation [193]; the study also showed that appropriate internal water clustering is integral to the protein's initial structure required for stable and realistic MD simulations [193].

#### 4.2.2. Water Molecules and Proteins' Folding

Within the cell, proteins are synthesized as linear strings of amino acids, as the ribosome 'reads' the corresponding DNA strand. For most proteins, this linear chain must fold into a unique three-dimensional structure (native structure of the protein) which is responsible for the protein's function. Correct folding is a prerequisite for proper functioning [194]. Incorrectly folded (misfolded) proteins may aggregate into so-called amyloid fibrils, thus becoming responsible for neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, Huntington's disease and others [194–204]. Explicit-solvent all-atom MD simulations of amyloid aggregation have revealed valuable information about this phenomenon; however, some challenges (as analysed in [202]) need further refinement of the approaches – the most impacting being the following: insufficient accuracy of current force fields modelling amyloid aggregation; the fact that protein concentrations in MD simulations are usually orders of magnitude higher than those used *in vitro* or found *in vivo*; and the time-scale limit of MD simulations (mostly in the nanosecond scale, whereas protein folding and unfolding occur in the microsecond to second timescale) [202].

The solvent plays a critical role in protein folding, as most of the free energy for the folding process comes from the maximization of solvent entropy. When two nonpolar side chains residues come together, some of the water molecules are expelled from the hydration layer of that portion of the protein and become part of the bulk solvent; this entails an increase of the net disorder of the water molecules, i.e., an increase in the entropy of the water (which, in turn, makes the total  $\Delta G$  more negative); thus, the hydrophobic effect is the driving factor for protein folding [196,205,206]. Given the importance of the solvent, many studies experimentally investigated what happens in one or

more solvents (e.g., [207–209]). Computational studies started to compare the performance of explicit and implicit solvent models for the analysis of protein folding. Zhou and Berne compared explicit solvent models with the generalized Born continuum model (GB) and found that the free energy landscapes were very different, that the GB lowest free energy state did not correspond to the native state, and that the minimum-energy structure was different in the two cases; on the other hand, GB had mostly provided the native  $\beta$ -hairpin structure [210,211]. In other cases, implicit models had recognised the folded state of peptides as the lowest free energy state [212]. Some works coupled implicit solvent models with MD (e.g., [213]). Although implicit models are less costly, explicit solvent models were viewed as the most suitable [214] and their popularity kept increasing (with MD simulations as frequently preferred option) because they provide a clearer picture of what happens at molecular level, above all for the structural water molecules and the water molecules in the hydration layer [215–217], thus being more informative for the investigation of issues such as the involvement of water molecules in the folding kinetics and mechanism.

A number of works chose to study the 76-amino acid (HP35) compact f-actin binding terminal domain ("headpiece") of the villin protein [218], because its small size and fast folding made it particularly suitable for pilot investigation of the folding mechanism. Their consideration can suitably illustrate the approaches in the first decade of our century. The solvent was treated differently in different works: implicit models [219–221], the GB implicit model combined with MD [213,222,223], explicit models [224–228] and both implicit and explicit models [229]. Results show that simulations with explicit and implicit solvent can produce different ensembles of structures, and even simulations with different implicit solvent models may lead to different ensembles of structures [220]; that the use of implicit models may be related to some incorrect predictions because, when water is implicit, unsatisfied H-bonds cannot be compensated for via H-bonding to water molecules [223], or may lead to incorrect properties for folding intermediates [227]; that the explicit representation of water is ideally suitable to explore the influence of solvent at the molecular level [226]; and that the use of unbiased, explicit solvent atomistic simulations of folding can provide detailed information on the nature of intermediate structures occurring during folding, which might be obscured within non-explicit approximations [227].

The fast increase in computers' power and the continuous refinement of computational models has continued opening new possibilities for the investigation of relevant aspects and questions. Few illustrative examples are recalled here. Oshima and Kinoshita investigated the components contributing to the solvent-entropy change upon protein folding [229]. Wang analysed a variety of physical and geometrical properties of the solvent-excluded surfaces to explore their contributions to protein-solvent interaction [230,231]. Besides the studies in aqueous medium, increasing attention has been given to the effects of non-aqueous solvents, which might be present as co-solvents in the biological medium in the cell. Yu and co-workers investigated the effects of several organic solvents on protein folding [232]. Van der Vegt and Nayar investigated the role of co-solvents known to be present within living cells, where they modulate aqueous solubility, hydrophobic interactions and the stability and function of many proteins; they found that direct interactions of co-solvents with nonpolar solutes can strengthen hydrophobic interactions [233]. Davis and co-workers considered the composition of the cell's internal solution and how it can affect protein folding and binding [234]. Mishra and co-workers investigated the solvent accessibility of aggregation patches and found that it is low for native crystal structures, as protein folding minimizes the solvent accessibility of aggregation-prone residues [235]. Hayashi and co-workers investigated the effects of the specificities of different solvents on the stability of a protein's native structure from the point of view of the free-energy involved and the enthalpy and entropy contributions to it [236]. Co-solvents (including salts, sugars, polyols, amino acids and amines) are known to enhance the folding and stability of proteins and the assembly of macromolecules such as microtubule [237]. Arakawa investigated their roles in protein folding and found that the co-solvents exert considerable effects at high concentrations, indicating that their interactions with proteins are weak; the co-solvents that enhance protein folding and macromolecular interactions get excluded from the protein surface [237]. Bucciarelli and co-workers studied  $\alpha$ -lactalbumin and showed that protein self-assembly pathways are determined by

a subtle balance between H-bonds' formation and hydrophobic interactions. Hydrophobic co-solvents modulate these two factors through a combination of direct solvent-protein and solvent-mediated interactions [238]. Dispersion interactions play an essential role in intraprotein and protein-water interactions. Stöhr and Tkatchenko used an explicit QM approach combining density-functional tight-binding with the many-body dispersion formalism and demonstrated the relevance of many-body van der Waals forces both for protein energetics and for protein-water interactions; they inferred that many-body effects substantially decrease the relative stability of native states in the absence of water whereas, in the presence of water, protein-water dispersion interactions counteract this effect and stabilize native conformations and transition states [239].

#### 4.2.3. Proteins in Non-Aqueous Solvent

The increasing use of enzymes in organic syntheses, as a greener option than traditional approaches, has prompted the need to understand the catalytic activity of enzymes in non-aqueous media. The use of enzymes in non-aqueous media can provide a number of synthetic and processing advantages [240]. Raccatano used MD simulations to investigate the effects of various organic solvents on protein folding and found that the simulations started to clarify several aspects of non-aqueous enzymology related to the peptide and protein dynamics and preferential solvation [241].

MM and MD simulations of proteins in non-aqueous solvents provide insights into what happens in those solvents to the water molecules on the protein surface. In non-aqueous media, the water molecules form clusters preferentially hydrating charged or polar residues; these clusters populate identical enzyme surface regions in different organic solvents; their number and size increase as water is added [242,243]. In non-polar solvents, large water clusters remain tightly bound to the protein surface whereas, in polar solvents, the clusters are smaller and loosely bound [242]. These water molecules are crucial for the efficacy of enzymes' catalytic activities in non-polar solvents. Natural enzymes exhibit very low activities in organic solvents, often 4 or 5 orders of magnitude lower than in aqueous solutions, and this restricts their potential industrial applications. Proteins require the presence of a certain number of water molecule bound to them (essential water) to retain their activity in non-aqueous solvents. It is difficult to elucidate the details of the organic solvent-enzyme-essential water interactions experimentally. MD simulations had been able to provide some insights into the molecular level of these interactions; the QM/MD combination proved considerably more efficient to enable more complete insights [240]. Zhu and co-workers studied  $\gamma$ -chymotrypsin both in acetonitrile with inclusion of 151 crystal water molecules and in water, using MD simulations and QM modelling with the PCM solvent model. The results showed that acetonitrile causes deviations from the native enzyme structure and flexibility loss, and that the structure changes occurring in the active pocket weaken the catalytic H-bond network and increase the proton transfer barriers, leading to a decrease in the enzymatic activity [243]. Meng and co-workers used MD simulations to study trypsin in water, acetonitrile and hexane, and found that it is more compact and less native-like in non-polar hexane than in the other two (polar) solvents [244].

#### 4.3. Solvents and the Structure of DNA

Nuclei acids have been objects of intensive investigation because of their roles in life. The study of their interactions with solvents started quite early, initially only at experimental level and progressively also with computational modelling. Early experimental and spectroscopic studies on DNA dehydration showed that water can be removed from sugars and bases, while the phosphate groups remain hydrated [245–250]. As with proteins, the aqueous solution is critical to the conformation and function of nucleic acids, and dehydration causes a transition from B-DNA to A-DNA [251].

H-bonds pair the bases of the two DNA strands, holding them together. Stacking interactions between consecutive base pairs play important roles for the stabilization of the 3-dimensional structure of DNA and RNA [251,252]. They are sequence dependent, and have been attributed to electrostatic interactions, hydrophobic effects, and dispersion interactions. Norberg and Nilsson investigated base-stacking in aqueous solution and in organic solvents, using nanosecond MD

simulations [252]. The results show that base stacking is mostly favoured in the high dielectric aqueous solution, followed by methanol and dimethyl sulfoxide (with intermediate dielectric constants), and chloroform (with a low dielectric constant) [252]. The DNA interior is mainly hydrophobic; its surface is rich with hydrophilic groups from the phosphates and sugars, promoting a tight hydration shell, whose characteristics depend on the DNA conformation and sequence; the water molecules of this shell shield the electrostatic repulsions between phosphate groups, thus stabilising the double-helix structure [251].

The 'natural' medium for the double-helix structure of DNA is the high-dielectric aqueous medium present in the cells, and DNA has evolved to be stable in this environment. Studies on the stability of DNA in different solvents have rapidly grown with the development of biotechnologies (utilising genetic engineering), because of the industrial significance associated with them; new applications are also envisaged, including a new generation of biocatalysts or chiral scaffolds for metal catalysts [253]. Bonner and Klibanov investigated the structural stability of DNA in non-aqueous solvents and found results recalling those for the proteins' structure and folding, first of all the importance of hydrophobic interactions [254]. Shen and co-workers used MD simulations to assess the effects of different metal ions present in aqueous medium on the conformation of a dodecamer DNA segment, at varying temperatures, and found that the light ions ( $\text{Li}^+$  or  $\text{Na}^+$ ) prefer to interact with the free phosphate oxygen atoms while the heavier ions ( $\text{Rb}^+$  and  $\text{Cs}^+$ ) strongly interact with the base pairs [255]. Arcella and co-workers used an atomistic-molecular simulation to compare the situation of DNA in a highly apolar environment and in aqueous medium, and found that even the neutral form (predicted to be the dominant one in apolar solvents) would surround itself with a small number of highly stable water molecules when moving from water to an apolar environment [256].

Given the size of a DNA molecule, some effects are conveniently investigated on smaller-size models (oligonucleotides). Nakano and Sugimoto present a review of studies of the structural stability and catalytic activity of DNA and RNA oligonucleotides in organic solvents, also in view of practical utilizations [257]. Still keeping practical utilizations in view, Zhao investigated the stability of DNA in ILs and in DES, as possible alternative solvents for DNA preservation and stabilization; however, organic cations may intrude into the DNA minor grooves, and their interaction with the DNA phosphate backbone may become predominant, while anions may form H-bonds with the cytosine, adenine and guanine bases [253]. Nan and co-workers used MD simulations to investigate DNA in ethylene glycol solution and found that its double helix is similar to the structure of DNA in the aqueous solutions but more compact; the similarity may be related to the fact that both solvents can form H-bonds with DNA, and that ethylene glycol molecules have greater capacity than water molecules to H-bond to the partially negative oxygen atoms in the DNA's phosphate groups [258].

#### 4.4. Solvent Roles in Protein-Protein Interactions

Protein-protein interactions (PPIs) are crucial for the proper functioning of the molecular mechanisms underlying cellular life, and are often perturbed in disease states [259,260]. The number of PPIs occurring in a living organism at a certain time is huge. Proteins can interact with each other through hydrophobic interactions, dispersion interactions and salt bridges. An overview of the molecular bases of PPIs is offered in [261] and an overview of the characteristics of PPI interfaces in [262]. Solvent molecules mediate PPIs as well as the interactions between proteins and other biomolecules in living organisms. Co-solvents present in the system may strengthen or weaken PPIs, depending on the nature of the co-solvent and of the proteins; changes in the nature or concentration of co-solvents can induce changes in the PPIs and, consequently, in the processes within a cell [263,264].

Levy and co-workers consider that the binding between two proteins are governed primarily by the protein's native topology and depend on the network of non-covalent residue-residue interactions that can establish between them [265]. They investigated in detail the binding mechanisms of selected protein pairs. They suggested that solvent-mediated H-bond formation and, in general, solvent mediated contacts may facilitate the interaction between two proteins, including

antigen–antibody association; they also recall the relevance of MD simulations, including for scopes such as clarifying the extent and characteristics of the desolvation accompanying the coming together of two solvated proteins [265]. Water molecules are more abundant at the interfaces between proteins and, therefore, they play important role in binding and recognition [251]. In some complexes, the water molecules are only at the interface rim, whereas in others they cover the entire interface area. Levy and Onuchic suggest that simulations of antibody-antigen complex using the topology-based model do not reproduce the transition state energy adequately if water is not included in the model; they infer that water molecules assist the initial stage leading to protein-protein binding and, therefore, water has to be included in computational simulations of the process. In addition, they suggest that water molecules assist two proteins in the identification of the appropriate binding sites among the potential ones, thus contributing to molecular recognition [251]. Ahmad and co-workers used extensive unbiased MD simulations and found that the water molecules in the interfacial gap forms an adhesive hydrogen-bond network between the interfaces, thus stabilizing early intermediates before native contacts are formed [266].

Vagenende and co-workers investigated the molecular origins of the co-solvents' effects on PPIs, using MD simulations to characterize local protein solvation [264]. They found that changes in preferential solvent interactions at the protein-protein interface account for the effect type of the co-solvent; the solvation changes, in turn, depend on the de-hydration extent of the protein-protein contact region, and also on structural changes that alter cooperative solvent-protein interactions at the margins of the protein-protein interface [264].

PPIs may become additional drug-targets for the treatment of diseases. The design of such drugs is challenging, because PPI sites are shallow protein surfaces. The design requires good understanding of the characteristics of the given PPI interface. Ghanakota and co-workers evaluate the ability of mixed-solvent MD (MSMD) simulations to detect spots at PPI interfaces, which could become suitable targets, by considering 21 PPI targets which had already been validated experimentally. They showed that MSMD simulations comprising explicit solvent and full protein flexibility provide more complete information than simulation which do not include these features [267]. Mayol and co-workers also used MSMD simulations to predict protein-drug and PPIs in different solvents [268].

An increasing number of diseases appear to be linked to aggregation of proteins and peptides, including cancer, the previously mentioned neurodegenerative diseases, and amyotrophic lateral sclerosis (ALS). An overview of classical MD studies of both the protein and the solvent involved in protein aggregation and fibril formation, utilising atomistic and coarse-grained models, is offered in [269]. A description of explicit solvent all-atom MD simulations and their applications for the study of the early stages of aggregation processes is offered in [270].

Klimov and co-workers used MD simulations to investigate the stability of oligomers of  $\text{A}\beta_{16-22}$  (KLVFFAE) peptides in aqueous urea solution and found that high urea concentration promotes the formation of  $\beta$ -strand structures, whereas largely compact random coil structures are preferred in water; in other words, urea opposes aggregation [271]. Matubayasi and co-workers investigated the co-solvent effect on peptide aggregation with all-atom MD simulation and free-energy calculation and found that the stability of a flexible solute is modulated by a co-solvent through the solvation free energy; they also found that urea and DMSO inhibit the aggregation because they stabilize the monomer more strongly than the aggregates [272].

A simulation of the association of strongly aggregating proteins (like the amyloid-*b* (*Ab*) peptide) using explicit solvent MD is unaffordable because of the size of the 'proteins + water molecules' system that should be used (such system would contain a huge number of water molecules). Emperador opted for an implicit solvent approach, using a highly detailed coarse-grained representation of the amino acid side chains, while keeping an atomistic representation of the backbone in order to maintain adequate accuracy for the description of secondary structure elements [273]. He coupled this model with the discrete MD (DMD) approach to study both molecular recognition and protein aggregation, and obtained results in complete agreement with the experimental evidence [273]. Stephens and co-workers used both ab initio MD (AIMD) and classical

MD simulations to investigate the differences in the aggregation rate of  $\alpha$ -synuclein ( $\alpha$ Syn, which plays relevant roles in the development of Parkinson's disease and other diseases) in different ionic solutions [274]. They showed that the addition of NaCl (which consists of two small ions with high charge density) reduced the H-bond dynamics of water and increased the aggregation propensity of  $\alpha$ Syn, whereas the addition of CsI (which consists of two large ions with low charge density) increases water mobility, thus contributing to an increase in the protein mobility and, consequently, reducing the proclivity of  $\alpha$ Syn to aggregate [274].

#### 4.5. Solvent Roles in Protein-DNA Interactions

The characteristics of DNA and its association with other molecules in solution have been the object of early exploration of the PCM applicability to the study of biological systems [275,276]. An example is the study of the energetics of the wrapping of DNA around a histone octamer (nucleosome), where the octamer and the portion of DNA molecule wrapping it around take the role of solute and the cavity surface is built around it [276].

The recognition between two biomolecules requires both geometric and chemical complementarity, and leads to the formation of a thermodynamically stable and specific complex [277]. Recognition is crucial for many functions within biological systems, including the binding of enzymes and substrates, the mediation of signal transduction via networks of specific protein pairs, and the regulation of protein expression by nucleic acids [277]. Site-specific associations between DNA and proteins regulate many biological events [277], with key involvement in transcription, replication and recombination. The analysis of the first-obtained protein-DNA crystal structure showed that several contributions lead to formation of the complex, including H-bonds, electrostatic interactions, direct and indirect contacts between amino acids and phosphate, sugars and bases, water-mediated contacts, hydrophobic effects, and others [277].

Understanding the interactions stabilizing biomolecular complexes in aqueous solution, and how small change can influence properties and behaviours, is essential for the design of drugs or other simple molecules that can influence cellular processes in a desired way [278]. DNA-protein interactions are ideal for such studies [278]. The DNA hydration shell is tight and closely associated to the DNA conformation and sequence, and therefore the water molecules of this shell can be viewed as an integral part of DNA [251]. The crystal structure of a complex between the Trp repressor and DNA showed three ordered water molecules at the protein-DNA interface, H-bonding both to the base pairs of DNA and to the protein side-chains; this suggested that water molecules mediate contacts between residues and base pairs that are important for sequence recognition, and that would not interact without the water mediation [279]. MD calculations in a simulated water bath confirmed the presence of water molecules at the protein-DNA interface, in positions consistent with this mediation role [279]. Crystallographic and NMR studies confirmed this role [280]. However, it is not always possible to characterize the localization and dynamics of the water molecules through X-ray or NMR techniques; MD simulations can complement experimental studies and provide insights about structure, dynamics, interactions, and the roles of water molecules [278]. Simulations of various protein-DNA systems showed the presence of both direct protein-DNA H-bonds and water-mediated H-bonds [278,281–286]; this also suggests the importance of water-mediated interactions in the recognition of DNA by many proteins and in the stabilization of the protein-DNA complexes [282,286]. Jayaram and Jain summarise the binding-facilitating roles of water molecules at the protein-DNA interface in the following terms: forming mediated H-bonds or H-bond networks which compensate for the lack of direct H-bonds in certain positions; shielding electrostatic repulsions between electronegative atoms or like charges of the protein and the DNA; and maintaining packing-densities at the interface, by filling spaces that would otherwise not be filled [287].

Single water molecules often constitute determining factors ruling the specificity and selectivity in molecular recognition and enzymatic reactions [288]. Protein-DNA complexes are particularly suitable for illustration purposes: direct interactions between protein side chains and DNA bases often remain scarce; bridging water molecules are present in the protein-DNA interface, and their

occurrence and positions appears to be crucial for specificity and selectivity [288]. Although water-mediated interactions often contribute significantly to the affinity and specificity of molecular interactions, most current protein design programs do not predict the location and contribution of bridging water molecules because of the high computational costs involved [288]. This has prompted the design of software whose algorithm aims at identifying physically ideal positions for explicit water molecules [288].

#### 4.6. Solvent Roles in the Ligand-Protein Interactions

##### 4.6.1. Interactions of Proteins with Small Molecules

Proteins may have active sites that can interact with small molecules (ligands). Particularly important are the cases in which the small molecules can act as drugs, producing effects that lead to diseases' treatment. A protein active site is mostly a 'pocket' into which the drug molecule inserts itself to a greater or lesser extent, according to the 'depth' of the pocket. If the pocket contains hydration water molecules, some or most of these molecules are expelled in the process (they are 'displaceable' [181,289]); simultaneously, the water molecules surrounding the ligand's portion that enters into the pocket are left out (remain in the bulk solvent), and the ligand gets partially desolvated (basic illustration in [290]). Some water molecules may remain in the pocket (are 'conserved') and mediate the ligand-protein interactions (LPIs), thus enhancing specificity and affinity; they may facilitate the formation of H-bond networks, which can be further stabilised through cooperativity effects and can have a role for the biomolecular structure, and for functions such as recognition and specificity [184]. The first realizations of these effects suggested that designing suitable presence of LPIs-mediated water molecules could make drug design more efficient, by enabling more accurate predictions of the ligand-protein binding mode [291].

In order to consider water molecules within drug design, it is necessary to identify those that can effectively mediate LPIs. Two subclasses can be identified among the water molecules that are conserved: those that are not displaced by any of the ligands, and those that are displaced by some ligands [292]. Techniques entailing the explicit inclusion of tightly bound water molecules have been developed for the molecular modelling of hydration effects within computational structure-based drug design [293]; examples have been codes like WaterDock, WaterScore, and PyWATER [184,186,294]. MD simulations – which also consider water molecules explicitly – increasingly became the preferred computational option [293].

A crucial question in the study of proteins and their action is the identification of their active sites. The presence and behaviour of solvent molecules can provide valuable information, and computational approaches (e.g., WaterMap) had been designed to identify active sites on the basis of the differences in thermodynamic and hydration free energies profiles relative to bulk solvent [295]. In many cases, the binding sites are well-defined regions, easily identifiable in the crystal structure as suitable to host substrates, and partially enclosed within the 3D fold of the protein [296]. However, some binding sites (sometimes termed "cryptic sites") can be identified only upon ligand binding (the 'pockets' may be too shallow to be evident if a ligand is not attached to them). These sites may have great biological relevance, e.g., for protein-protein recognition processes. Identifying them and understanding their structure can be interesting for the design of drugs that can target them. The addition of organic co-solvents to water proved expedient to reveal cryptic sites by binding to them (creating greater concentration of solvent molecules near a specific site than in the bulk solvent) [297] or through the conformational changes that they induce on the protein surface [298]. Multiple solvent crystal structures (MSCS) methods have been used to complement computational information for the determination of these binding sites [299]. MD simulations can provide a realistic assessment of the complex kinetics and thermodynamics by integrating the consideration of protein flexibility and of the role of water [296,297,300,301].

The role of water in modulating the stability of drug-receptor complexes has been recognized for various decades. Water-mediated interactions between a ligand and a protein play a key role in biomolecular assembly processes, such as protein-ligand recognition, the binding of the HIV virus or

the dengue virus to human cells, the inhibition of influenza virus infectivity, or the binding of a synthetic drug to a biomolecular guest [302]. Already in 1997, Bohm and Klebe presented a review of works focusing on the physical nature of molecular recognition in protein-ligand complexes and the application of existing computational tools enabling the utilization of available knowledge on LPIs in the design of novel ligands [303]. A number of subsequent works aimed at elucidating the details of the 'binding event' [304]. Setny and co-workers combined explicit-water MD simulations and the variational implicit solvent model (VISM) to investigate a generic pocket-ligand model, and found that the approaching ligand initially stabilizes the wet state in the weakly hydrated pocket, while further approach induces pocket dewetting [302]. Schmidtke and co-workers showed that, when a ligand and a receptor form interactions via H-bonds that are shielded from water by adjacent hydrophobic regions, the resulting complex tends to exhibit higher kinetic stability compared to situations where these H-bonds are less shielded [305]. Dror and co-workers used unbiased MD simulations to study the event, and concluded that several  $\beta$ -blockers or  $\beta$ -agonists initially make contact with a vestibule on each receptor's extracellular surface; this stage often entails the largest energy barrier to binding, because it involves substantial dehydration of the pocket; in the subsequent stage, the ligand enters into the binding pocket by squeezing through a narrow passage [306]. Setny and co-workers used explicit-water MD simulations to investigate the role of water in the mechanisms with which a ligand binds to a prototypical hydrophobic pocket, and confirmed the existence of a dewetting barrier for the first step [307]. Young and co-workers used MD simulations and a solvent analysis technique based on inhomogenous solvation theory to investigate the properties of water molecules that solvate the confined regions of protein active sites, with particular attention molecular recognition patterns in which the displacement of the solvent by the ligand leads to exceptional binding affinities [308]; they concluded that the hydrophobic enclosures aid molecular recognition by perturbing the solvation of the binding cavity, and this leads to a relative stabilization of the bound complex [308]. In general, biomolecular recognition is influenced as much by rearrangements in the water molecules that solvate interacting species as it is by the interactions between those species [309].

Molecular docking is a modelling technique which starts from structural information and searches for the best-energy (minimal energy) way in which two molecules can fit to each other to form a complex. At least one of the molecules is a biomolecule. The interaction may entail DNA-ligand, protein-DNA, protein–protein, ligand-protein, substrate-enzyme, and so on [310]. Given the importance of the roles of water molecules to shape the form of LPIs (or other interactions involving biomolecules), docking or virtual screening softwares increasingly incorporate the possibility of considering explicit water molecules. Hu and co-workers present a review of questions and approaches used in computational drug discovery and drug development, for which the effect of a single water molecule, or of a small network of interacting waters molecules, needs to be considered [311]. They consider issues such as the selection or prediction of hydration sites before starting the simulation, the computational methods that can best predict water positions, the ensuing incorporation of explicit water molecules into the docking procedure, and the methods adopted by various softwares [311]. Several works provide information about the development of specific approaches to the identification of the physically ideal positions for explicit water molecules in a simulation, and their utilization in computational softwares [288,310,312–320].

#### 4.6.2. Enzymes: Proteins That Are Catalysts

Enzymes are proteins that act as catalysts for reactions occurring within living organisms. As catalysts, they bind a specific reactant molecule (substrate) in such a way that the reaction activation energy is lowered, and get released when the transition state gives way to the products. The structural complementarity of the substrate and the enzyme's binding site prompted the widely-used image of key-and-lock. It is possible to design molecules with greater affinity than the 'natural' substrate for a given enzyme, which bind to the enzyme permanently, thus inhibiting further catalytic activity [321]. Such molecules can act as drugs to inhibit vital processes of pathogens.

From a biochemical point of view, the correct structure of an enzyme can be simply defined as the conformation that enables good catalytic activity. Protein hydration is essential for enzyme catalysis to occur; dry enzymes are inactive, and there is a minimum water proportion for the activity to begin [185,322,323]. Verma and Mitchell-Koch provide a review of computational studies (primarily MD simulations) exploring the dynamics and thermodynamics of the participation of the relevant small molecules (solvent, substrate, and cofactor molecules) in enzyme catalytic processes (molecular recognition, substrate binding, catalysis, and product release), as well as the role of protein flexibility, and include information on the development of the relevant theoretical approaches [324].

What happens to the water molecules involved in the catalytic process may be different for different types of enzymes. Glycoside hydrolases conserve internal water molecules; MD simulations for *Thermus thermophilus*  $\beta$ -glycosidase suggested the involvement of two water channels, while another chain of highly conserved water molecules (going from the protein surface to the bottom of the active site cleft) is able to exchange with the bulk at the nanosecond time scale [325]. MD simulations showed that it may also happen that the collective enzyme–substrate–water-coupled motions persist beyond the steady state, indicating that the long-lasting water dynamics contribute to the net enzyme reactivity, impacting substrate binding, positional catalysis, and product release [326].

The human immunodeficiency virus (HIV) attacks the body's immune system and, if not treated, can lead to AIDS (acquired immunodeficiency syndrome), for which the body loses the ability to fight infections and the patient dies from illnesses that would not have been acquired or not been deadly in the absence of AIDS. Around 40 million people have died from AIDS-related illnesses since the start of the epidemic [327]. Three enzymes are fundamental for the life cycle of the virus: protease (PR), reverse transcriptase (RT), and integrase (INT) [328]. The search for molecules that can inhibit these enzymes has been intensive. Rungrotmongkol and co-workers investigated the structure and dynamics of the RT active site by modelling the active conformation of the HIV-1 RT/DNA/deoxythymidine triphosphate (dTTP) ternary complex; they used both MD simulations with the CHARMM27 force field and QM/MM approaches, and examined potentially important H-bonding interactions with amino acids and with water molecules bound to the system [329]. INT is vital for the integration of the viral DNA into the host DNA, and its inhibition had proved capable of slowing down the progression of AIDS [328]. Ribeiro and co-workers built a model of the holo-integrase:DNA complex comprising an entire central core domain, a ssDNA GCAGT substrate, and two magnesium ions; then they used a combination of MD, thermodynamic integration, and high-level QM/MM calculations to investigate the possible pathways for the mechanism of the process catalyzed by INT [328]. They found that the only viable mechanism to hydrolyze the DNA substrate is a nucleophilic attack by an active site water molecule to the phosphorus atom of the scissile phosphoester bond, with the attacking water being simultaneously deprotonated by an  $Mg^{2+}$ -bound hydroxide ion [328].

A broad variety of questions have been investigated in the last decade. Fox and co-workers studied mutants of human carbonic anhydrase (HCAII, a metalloenzyme) to investigate how changes in the organization of the water molecules filling a binding pocket can alter the thermodynamics of the ligand-protein association [309]. They used a combination of calorimetry, crystallography, and computational methods and found that, within the confines of the HCAII binding pocket, binding events associated with enthalpically favourable rearrangements of the water molecules are stronger than those associated with entropically favourable rearrangements of water [309]. Gopal and co-workers used isothermal titration calorimetry (ITC) and MD simulations to investigate the effect of solvent composition on the thermodynamics of protein–ligand binding, selecting the binding of p-aminobenzamidine (PAB) to trypsin in various water/methanol mixtures as a case study [330]. They found that the MD and free energy simulations reproduced the experimental binding free energies and also provided atomic-level insights into the mechanisms underpinning the thermodynamic observations [330]. Full understanding of plant polysaccharide biosynthesis remains arduous because of the challenges in the characterisation of the structure of glycosyltransferase (GT) enzymes. Urbanovicz and co-workers investigated the mechanistic basis for fucosylation in *Arabidopsis*

thaliana, where a glycosyltransferase (fucosyltransferase 1, AtFUT1), catalyzes the regiospecific transfer of terminal 1,2-fucosyl residues to xyloglucan side chains [331]. QM/MM and MD calculations suggest that AtFUT1 may use an atypical water-mediated mechanism with the potential contribution of an H-bonding network for acceptor nucleophile activation [331].

Enzymes are widely used in industrial processes to accelerate chemical reactions and obtain high selectivity and specificity under ambient conditions (ambient conditions being recommended within green chemistry perspectives). However, naturally-occurring enzymes cannot meet the increasing demands of catalysts for green processes' development, as they are not always suitable to act as catalysts for non-natural substrates. To address this problem, native enzymes have been adapted to catalyse non-natural chemical transformations on the basis of information obtained through high-throughput screening or through structure-based computational enzyme design approaches [332]. The latter may entail *de novo* design of new active sites or the redesign of existing active sites. Model accuracy relies on QM/MM and MD simulations to account for the details of LPi, including the role of the solvent. Xue and co-workers quantitatively assessed the computationally designed variants of a *Rhodococcus* sp. cocaine esterase for the hydrolysis of cephadrine, using MM/Poisson–Boltzmann surface area (MM/PBSA) and MM/generalized Born surface area (MM/GBSA) methods; the way in which explicit water molecules around the substrate were considered was based on MD simulations [332].

QM/MM methods are among the most widely used for the computational study of enzymes and their activities. Sousa and co-workers outline relevant applications [333]. Magalhães and co-workers provide a review in which they list the main tasks for which QM/MM is used in the study of enzymes (to validate or disprove different mechanistic hypotheses regarding the catalytic pathway of a specific enzymatic reaction; to obtain an atomic-level analysis of the main interactions formed in the reactants, transition state and products; to identify new scaffolds suitable for drug discovery) as well as the main modelling choices (preparation of the initial structure, choice of the QM/MM boundary, choice of the QM level, choice of the MM level, use of link atoms, inclusion of solvent, use of constraints in the MM region) [334]. Specifically for the solvent, they recommend that the water molecules that are assumed to directly participate in the reaction, or to play an important role in directly stabilizing a specific group or interaction, should be included explicitly in the QM region; the other water molecules can be included in the MM region. The surrounding aqueous environment can be treated with the IEF-PCM or C-PCM models or, alternatively, a 5–10 Å cap of water molecules included in the MM region can be added to the enzyme [334]. Jędrzejewsk and co-workers used a combination of bioinformatics analysis, molecular docking, MD simulations, and QM calculations to investigate the mechanism through which methyltransferase Nep1 catalyses the N1 methylation of pseudouridine during the rRNA processing [335]. The MD simulations identified the active site arrangements with a water molecule bridging the N1 of pseudouridine and the putative aspartate proton acceptor through two consecutive H-bonds. The QM calculations established that the energy barrier for the methylation is lower when water molecules mediate the proton transfer than in pathways in which the OH group of serine/threonine acts as proton–shuttle. [335].

#### 4.7. Solvent Roles in the Interactions between DNA and Other Molecules

Interactions between DNA and proteins are fundamental for many biological processes; the role of solvents has already been considered in section 4.5. The current section considers the interactions between DNA and small molecules, more often having the role of drugs.

DNA was the first target of anticancer drugs, acting on the cancer cells with various mechanisms: antimetabolites, which deplete nucleotides; alkylation agents, which cause direct DNA damage; and intercalators (mostly containing a rigid planar part, such as anthracyclines), which bind to DNA in a way that prevents it from unravelling and, therefore, from duplicating, or damage it in other ways; and other drugs that can damage it [336]. In a similar way, the DNA of pathogens (e.g., *plasmodium falciparum*, causing the most dangerous form of malaria, or DNA-based viruses) is a biological target in the treatment of the corresponding diseases, and RNA becomes a target for RNA-based viruses.

The importance of considering the role of water was acknowledged since early studies of the binding of a drug with DNA (e.g., [337–339]).

Sheng and co-workers list the major components of the interactions between DNA and ligands as hydrophobic packing, direct H-bonding of the ligand to the groove floor before and after the turn, water-mediated H-bonding of the ligand to the groove floor, and conformational pucks that minimize steric interactions of the ligand [340]. They also list the main strategies through which cancer DNA can be targeted by small molecules: Pt-containing compounds forming covalent bonds with DNA (e.g., cisplatin, or some organic compounds); minor or major groove binders; intercalators; multifunctionalised ligands; targeting the DNA quadruplex; targeting unmatched bulge, DNA junctions, and the phosphate backbone. Understanding the details of the ligand-DNA interaction is important for the design of new drugs, or the design of modified drugs with reduced side effects with respect to the current ones [340].

Experimental studies on DNA-ligand interactions in the cellular environment have remained problematic due to the scarcity of suitable biophysical tools [341]. Still in 2009, Ricci and Netz noted the scarcity of ligand-DNA docking studies in comparison with the abundance of ligand-protein docking studies [342]. A major challenge for the design of small molecules that can target DNA stems from the absence, in DNA molecules, of the clearly identifiable binding sites (pockets) present in proteins. The design could focus on molecules that can selectively bind DNA in specific areas and also meet criteria such as water-solubility, cellular and nuclear-uptake, and absence of off-site activities [343]. Carter and co-workers used long timescale MD simulations to analyse the sequence-specific DNA association of a synthetic small molecule; what happened to water (displacement from the DNA minor groove involved in the binding) was a major focus of attention, because of its recognition-related role [343].

Schuurs and co-workers consider an approach in which small molecules target a protein (Human Single-Stranded DNA Binding Protein 1) that binds to DNA and plays an important role in the ability of cancerous cells to survive traditional treatments; they discovered three small molecules that appear to prevent that protein from binding to DNA, by binding to the protein's site that would bind to DNA, and they used MD simulations with co-solvent simulation in the study of the interactions between the small molecules and the protein [344].

The other nucleic acid, RNA, can also become a target for drug action, including against RNA-based viruses. Lang and co-workers compiled a test-set of RNA-ligand complexes to validate the ability of the DOCK suite of programs to recreate experimentally determined binding poses; they noted that the success rate increases to 80% when docking-obtained conformations are rescored with the PBSA and GBSA implicit solvent models in combination with explicit water molecules and sodium counterions [345]. Panai and co-workers designed a computational technique to identify potential binding sites for small-molecule in RNA structural ensembles, and included the solvent consideration in the corresponding software [346].

The extensive industrial uses of DNAs in biotechnologies have made it important to evaluate the interactions with other solvents, besides those present in living organisms. Yusof and co-workers used spectroscopic analysis and MD simulations to investigate the binding characteristics of calf thymus DNA in a tetrabutylammonium bromide based DES [347]. DNA-templated silver clusters with up to about 30 silver atoms have proven to be bright emitters in the visible to near-infrared range, with the phenomenon controlled by the selected DNA sequence [346]. Malola and co-workers studied the DNA-stabilized silver cluster  $\text{Ag}_{16}\text{Cl}_2$  in aqueous solution at the DFT level, using an implicit solvent model, and inferred that the consideration of explicit solvent molecules at the DNA-water interface would be important for the study of transitions in the high-UV region [348].

Knowing how small molecules interact with DNA is important also to prevent damages. Molecules that are identified as being toxic to DNA need to be removed from industrial wastes or any material that reaches the environment or comes in contact with human beings. Li and co-workers investigate the ability of a series of DES to remove traces of substances that are toxic to DNA, and used the information from MD simulations to improve aspects of the practical procedure [349].

## 5. Discussion and Conclusions

The examples considered in the previous sections illustrate the importance of taking into account the properties of solvents, the interactions between solvent and solute and the effects of these interactions on the properties and behaviour of the solute, and – quite often – also the importance of following individual solvent molecules that are particularly relevant for the interactions or for a process. Computational chemistry offers a variety of approaches to investigate these aspects, and can help validate and interpret experimental information, but also predict properties and behaviours, thus guiding further experiments, or find information that is not easily detectable experimentally. The possible applications cover a broad range, from the selection or design of green solvents and production processes that can utilise them efficiently to the design of new materials and to continuous advancements in our understanding of biological processes and of the possibilities of treating diseases. Continuous improvement in the computational modelling-power can be expected ensuing from both the envisaged theoretical refinements on which the models are based and the incessant and fast technological growth of computers' technology. Therefore, the extent and quality of the contributions that computational chemistry can bring to the crucial problems concerning solutions and the phenomena occurring in them can be expected to increase steadily.

**Author Contributions:** design and supervision, L. Mammino; literature search, N. Tshilande, M.K.Bilonda, L. Mammino; writing, N. Tshilande, L. Mammino.

**Funding:** The study for this article did not require any external funding.

**Ethical approval:** N/A. This article did not require any studies with human or animal subjects performed by the authors.

## References

1. From a quote at the end of a Gaussian-16 output.
2. Reichardt, C. Solvents and solvent effects: an introduction. *Org. Process Res. Dev.* 2007, 11(1), 105–113.
3. Reichardt, C.; Welton, T. Solvents and Solvent Effects in Organic Chemistry, 4th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2011.
4. Mammino, L.; Kabanda, M.M. Considering the Medium when Studying Biologically Active Molecules: Motivation, Options and Challenges. In *Frontiers in Computational Chemistry*; Zaheer-ul-Haq, J.D. Madura, Eds.; Bentham Science Publishers: 2014; pp. 197–256.
5. Anastas, P.T.; Williamson, T.C. *Green Chemistry: Designing Chemistry for the Environment*; American Chemical Society: Washington, DC, USA, 1996.
6. Anastas, P. T.; Warner, J. C., *Green Chemistry: Theory and Practice*. Oxford University Press: New York, 1998.
7. Tundo, P.; Anastas, P.T. *Green Chemistry, Challenging Perspectives*; Oxford University Press: Oxford, UK, 2000.
8. Anastas, P.T.; Kirchhoff, M.M. Origins, current status, and future challenges of green chemistry. *Acc. Chem. Res.* 2002, 35(9), 686–694.
9. Welton T. Solvents and sustainable chemistry. *Proc. R. Soc. A* 2015, 471: 20150502. <http://dx.doi.org/10.1098/rspa.2015.0502>
10. Constable, D.J.C.; Jimenez-Gonzalez, C.; Henderson, R.K. Perspective on solvent use in the pharmaceutical industry. *Org. Process Res. Dev.* 2007, 11, 133–137.
11. Papadopoulos, A.I.; Linke, P. Multiobjective molecular design for integrated process-solvent systems synthesis. *AIChE J.* 2006, 52, 1057–1070.
12. Papadopoulos, A.I.; Linke, P. Efficient integration of optimal solvent and process design using molecular clustering. *Chem. Eng. Sci.* 2006, 61, 6316–6336.
13. Bardow, A.; Steur, K.; Gross, J. Continuous-Molecular Targeting for Integrated Solvent and Process Design. *Ind. Eng. Chem. Res.* 2010, 49, 2834–2840.
14. Zhou, T.; Zhou, Y.; Sundmacher, K. A hybrid stochastic-deterministic optimization approach for integrated solvent and process design. *Chem. Eng. Sci.* 2017, 159, 207–216.
15. Keßler, T.; Kunde, C.; Linke, S.; Sundmacher, K.; Kienle, A. Integrated computer-aided molecular and process design: Green solvents for the hydroformylation of long-chain olefines. *Chem. Eng. Sci.* 2022, 249, 117243.
16. Winterton, N. The green solvent: a critical perspective. *Clean Techn Environ Policy* 2021, 23, 2499–2522. <https://doi.org/10.1007/s10098-021-02188-8>

17. Breslow, R. The Principles of and Reasons for Using Water as a Solvent for Green Chemistry, Part 5. *Reactions in Water*, 2010, 1–29. <https://doi.org/10.1002/9783527628698.hgc047>
18. Simon, M.O.; Li, C.J. Green chemistry oriented organic synthesis in water. *Chem. Soc. Rev.* 2012, 41(4), 1415–1427. <https://doi.org/10.1039/C1CS15222J>
19. Majhi, K.C.; Karfa, P.; Kumar, S.; Madhuri, R. Water as the Green Solvent in Organic Synthesis. *Materials Research Foundations* 2019, 54, 182–201. <https://doi.org/10.21741/9781644900314-8>
20. Cao, X.; Zhang, G.; Jiang, L.; Cai, Y.; Gao, Y.; Yang, W.; He, X.; Zeng, Q.; Xing, G.; Jia, Y.; Wei, J. Water, a Green Solvent for Fabrication of High-Quality CsPbBr<sub>3</sub> Films for Efficient Solar Cells. *ACS Appl. Mater. Interfaces* 2020, 12(5), 5925–5931. <https://doi.org/10.1021/acsami.9b20376>
21. Beckman, E.J. Supercritical and near-critical CO<sub>2</sub> in green chemical synthesis and processing. *J. Supercrit. Fluids* 2004, 28, 121–191.
22. Nalawade, S.P.; Picchioni, F.; Janssen, L.P.B.M. Supercritical carbon dioxide as a green solvent for processing polymer melts: Processing aspects and applications. *Prog. Polym. Sci.* 2006, 31(1), 19–43. <https://doi.org/10.1016/j.progpolymsci.2005.08.002>
23. Sánchez-Vicente, Y.; Cabañas, A.; Renuncio, J.A.R.; Pando, C. Supercritical CO<sub>2</sub> as a green solvent for eucalyptus and citrus essential oils processing: role of thermal effects upon mixing. *RSC Adv.* 2013, 3, 6065–6075.
24. Vandeponsele, A.; Draye, M.; Piot, C.; Chatel, G. Subcritical water and supercritical carbon dioxide: efficient and selective eco-compatible solvents for coffee and coffee by-products valorization. *Green Chem.* 2020, 22(24), 8544–8571. DOI: 10.1039/d0gc03146a
25. Soren, S.; Sahoo, T.; Panda, J.; Senapati, D.K.; Sahu, J.R.; Rath, C.K.; Sahu, R. Carbon Dioxide-Based Green Solvents. In *Carbon Dioxide Utilization to Sustainable Energy and Fuels, Advances in Science, Technology & Innovation.*; Inamuddin, Boddula, R. Ahamed, M.I. Khan, A. Eds.; Springer: Cham, 2022; pp. 323–333
26. Petigny, L.; Öznel, M.Z.; Périno, S.; Wajsman, J.; Chemat, F. Water as Green Solvent for Extraction of Natural Products. In *Green Extraction of Natural Products: Theory and Practice*; Chemat, F.; Strube, J. Eds.; Wiley-VCH Verlag GmbH & Co. KGaA, 2015. <https://doi.org/10.1002/9783527676828.ch7>
27. Castro-Puyana, M.; Marina, M.L.; Plaza, M. Water as green extraction solvent: Principles and reasons for its use. *Curr. Opin. Green Sustain. Chem.* 2017, 5, 31–36. <https://doi.org/10.1016/j.cogsc.2017.03.009>
28. Hartonen, K.; Riekola, M.-L. Water as the First Choice Green Solvent. In *The Application of Green Solvents in Separation Processes*; Elsevier 2017; pp. 19–55. <https://doi.org/10.1016/B978-0-12-805297-6.00002-4>
29. Skyner, R.E.; McDonagh, J.L.; Groom, C.R.; van Mourik, T.V.; Mitchell, J.B.O. A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Phys. Chem. Chem. Phys.* 2015, 17(9), 6174–6191.
30. Mennucci, B. Multiscale strategies for describing environment effects: From solvents to biomatrices. In *Green Chemistry and Computational Chemistry – shared lessons in sustainability*, Mammino, L. Ed.; Elsevier: 2021; pp. 263–280.
31. Tomasi, J.; Persico, M. Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent. *Chem. Rev.* 1994, 94, 2027–2094.
32. Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* 2005, 105, 2999–3094.
33. Pascual-Ahuir, J.L.; Silla, E.; Tuñón, I. GEPOL: An improved description of molecular surfaces. III. A new algorithm for the computation of a solvent-excluding surface. *J. Comput. Chem.* 1994, 5, 1127–1138.
34. Pascual-Ahuir, J.L.; Silla, E.; Tomasi, J.; Bonaccorsi, R. Electrostatic interaction of a solute with a continuum. Improved description of the cavity and of the surface cavity bound charge distribution. *J. Comp. Chem.* 1987, 8, 778–787.
35. Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilization of *ab initio* molecular potentials for the prevision of solvent effects. *Chem. Phys.* 1981, 55, 117–129.
36. Miertus, S.; Tomasi, J. Approximate evaluations of the electrostatic free-energy and internal energy changes in solution processes. *Chem. Phys.* 1982, 65, 239–245.
37. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, structures and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* 2003, 24, 669–681.
38. Scalmani, G.; Frisch, M.J. Continuous surface charge polarizable continuum models of solvation. I. General formalism. *J. Chem. Phys.* 2010, 132, 114110.
39. Cramer, C.J.; Truhlar, D.G. Implicit solvation models: equilibria, structure, spectra, and dynamics. *Chem. Rev.* 1999, 99, 2161–2200.
40. Barone, V.; Cossi, M. Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *J. Phys. Chem. A* 1998, 102, 1995–2001.
41. Mennucci, B.; Cancès, E.; Tomasi, J. Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: Theoretical bases, computational implementation, and numerical applications. *J. Phys. Chem. B* 1997, 101, 10506–10517.

42. Cancès, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* 1997, 107, 3032–3041.

43. Soteras, I.; Curutchet, C.; Bidon-Chanal, A.; Orozco, M.; Luque, F.J. Extension of the MST model to the IEF formalism: HF and B3LYP parametrizations. *J. Mol. Struct. (Theochem)* 2005, 727, 29–40.

44. Mennucci, B. Continuum solvation models: What else can we learn from them? *J. Phys. Chem. Lett.* 2010, 1, 1666–1674.

45. Lipparini, F.; Mennucci, B. Perspective: Polarizable continuum models for quantum-mechanical descriptions. *J. Chem. Phys.* 2016, 144, 160901–160909. doi:10.1063/1.4947236.

46. Amovilli, C.; Barone, V.; Cammi, R.; Cancès, E.; Cossi, M.; Mennucci, C.; Pomelli, C.S.; Tomasi, J. Recent advances in the description of solvent effects with the polarisable continuum model. *Adv. Quantum Chem.* 1999, 32, 227–259.

47. Barone, V.; Cossi, M.; Tomasi, J. A new definition of cavities for the computation of solvation free energies by the polarizable continuum model. *J. Chem. Phys.* 1997, 107, 3210–3221.

48. Mennucci, B.; Tomasi, J. Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries. *J. Chem. Phys.* 1997, 106, 5151–5158.

49. Pomelli, C.S.; Tomasi, J.; Cossi, M.; Barone, V. Effective generation of molecular cavities in the polarizable continuum model by the DefPol procedure. *J. Comp. Chem.* 1999, 20, 1693–1701.

50. Tomasi, J. Cavity and reaction field: "robust" concepts. Perspective on "Electric moments of molecules in liquids" - Onsager, L. (1936) *J. Am. Chem. Soc.* 58: 1486. *Theor. Chem. Acc.* 2000, 103(3-4), 196–199.

51. Takano, Y.; Houk, K.N. Benchmarking the Conductor-like Polarizable Continuum Model (CPCM) for Aqueous Solvation Free Energies of Neutral and Ionic Organic Molecules. *J. Chem. Theory Comput.* 2005, 1, 70–77.

52. Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans. 2* 1993, 799–805.

53. Klamt, A. The COSMO and COSMO-RS solvation models. *WIREs Comput Mol Sci.* 2011, 1, 699–709.

54. Touaibia, M.; Fabiano-Tixier, A.-S.; Chemat, F. Chloropinane and Chloromenthene as Novel Solvents for Solubilisation of Natural Substances. *Molbank* 2021, 2021, M1205.

55. Zhou, T.; McBride, K.; Linke, S.; Song, Z.; Sundmacher, K. Computer-Aided Solvent Selection and Design for Efficient Chemical Processes. *Curr. Opin. Chem. Eng.* 2020, 27, 35–44.

56. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B.* 2009, 113, 6378–6396.

57. Mammino, L. Adducts of arzanol with explicit water molecules: an ab initio and DFT study. In *Concepts, Methods and Applications of Quantum Systems in Chemistry and Physics*, vol.31; Wang, Y.A., Thachuk, M., Krems, R., Maruani, J., Eds.; Cham, Springer: 2018; 281–304.

58. Boys, S.F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* 1970, 19(4), 553–566. DOI: 10.1080/00268977000101561

59. Alder, B.J.; Wainwright, T.E. Phase transition for a hard sphere system. *J. Chem. Phys.* 1957, 27, 1208–1209.

60. Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* 1976, 103(2), 227–249.

61. Shiga, M.; Masia, M. Boundary based on exchange symmetry theory for multilevel simulations. I. Basic theory. *J. Chem. Phys.* 2013, 139(4), 044120.

62. Takahashi, H.; Kambe, H.; Morita, A. A simple and effective solution to the constrained QM/MM simulations. *J. Chem. Phys.* 2018, 148(13), 134119.

63. Zheng, M.; Waller, M.P. Adaptive quantum mechanics/molecular mechanics methods. *WIREs: Comput. Mol. Sci.* 2016, 6(4), 369–385.

64. Duster, A.W.; Wang, C.H.; Garza, C.M.; Miller, D.E.; Lin, H. Adaptive quantum/molecular mechanics: What have we learned, where are we, and where do we go from here? *WIREs: Comput. Mol. Sci.* 2017, 7(5), e1310.

65. De Jesus, S.S.; Maciel Filho, R. Are ionic liquids eco-friendly? *Renewable Sustainable Energy Rev.* 2022, 157, 112039.

66. Crinnion, W.J. The CDC fourth national report on human exposure to environmental chemicals: what it tells us about our toxic burden and how it assists environmental medicine physicians. *Altern. Med. Rev.* 2010, 15, 101–109.

67. Quesne, M.G.; Silveri, F.; de Leeuw, N.H.; Catlow, C.R.A. Advances in Sustainable Catalysis: A Computational Perspective. *Front. Chem.* 2019, 7, 14674–23.

68. Bubalo, M.C.; Vidović, S.; Redovniković, I.R.; Jokić, S. Green solvents for green technologies. *J. Chem. Technol. Biotechnol.* 2015, 90, 1631–1639.

69. Gu, Y.; Jérôme, F. Bio-based solvents: An emerging generation of fluids for the design of eco-efficient processes in catalysis and organic chemistry. *Chem. Soc. Rev.* 2013, 42, 9550–9570.

70. Das, S.; Mondal, A.; Balasubramanian, S. Recent advances in modeling green solvents. *Curr. Opin. Green Sustain. Chem.* 2017, 5, 37–43.

71. Hansen, C.M. The universality of the solubility parameter. *Ind. Eng. Chem. Prod. Res. Dev.* 1969, 8(1), 2–11.

72. Faasen, D.P.; Jarray, A.; Zandvliet, H.J.W.; Kooij, E.S.; Kwiecinski, W. Hansen solubility parameters obtained via molecular dynamics simulations as a route to predict siloxane surfactant adsorption. *J. Colloid Interface Sci.* 2020, 575, 326–336. <https://doi.org/10.1016/j.jcis.2020.04.070>

73. Crowley, J.D.; Teague, G.S. Jr.; Lowe, J.W. Jr. A three-dimensional approach to solubility. *J. Paint Technol.* 1966, 38(496), 269–280.

74. Durand, M.; Molinier, V.; Kunz, W.; Aubry, J.M. Classification of organic solvents revisited by using the COSMO-RS approach. *Chem. Eur. J.* 2011, 17, 5155–5164.

75. Garcia-Chavez, L.Y.; Hermans, A.J.; Schuur, B.; de Haan, A.B. COSMO-RS assisted solvent screening for liquid-liquid extraction of mono ethylene glycol from aqueous streams. *Sep. Purif. Technol.* 2012, 97, 2–10.

76. Moity, L.; Durand, M.; Benazzouz, A.; Pierlot, C.; Molinier, V.; Aubry, J.M. Panorama of sustainable solvents using the COSMO-RS approach. *Green Chem.* 2012, 14(4), 1132–1145.

77. Benazzouz, A.; Moity, L.; Pierlot, C.; Molinier, V.; Aubry, J.M. Hansen approach versus COSMO-RS for predicting the solubility of an organic UV filter in cosmetic solvents. *Colloids Surf. A: Physicochem. Eng. Asp.* 2014, 458, 101–109.

78. Filly, A.; Fabiano-Tixier, A.S.; Fernandez, X.; Chemat, F. Alternative Solvents for Extraction of Food Aromas; Experimental and COSMO-RS Study. *LWT Food Sci. Technol.* 2015, 61, 33–40.

79. Del Pilar Sánchez-Camargo, A.; Pleite, N.; Herrero, M.; Cifuentes, A.; Ibáñez, E.; Gilbert-López, B. New Approaches for the Selective Extraction of Bioactive Compounds Employing Biobased Solvents and Pressurized Green Processes. *J. Supercrit. Fluids* 2017, 128, 112–120.

80. Li, Z.; Smith, K.H.; Stevens, G.W. The use of environmentally sustainable bio-derived solvents in solvent extraction applications - A review. *Chin. J. Chem. Eng.* 2016, 24, 215–220.

81. Li, Y.; Fabiano-Tixier, A.S.; Ginies, C.; Chemat, F. Direct green extraction of volatile aroma compounds using vegetable oils as solvents: Theoretical and experimental solubility study. *LWT - Food Sci. Technol.* 2014, 59, 724–73.

82. Bundejomchok, K.; Filly, A.; Rakotomanomana, N.; Panichayupakaranant, P.; Chemat, F. Extraction of  $\alpha$ -Mangostin from *Garcinia mangostana* L. Using Alternative Solvents: Computational Predictive and Experimental Studies. *LWT Food Sci. Technol.* 2016, 65, 297–303.

83. Filly, A.; Fabiano-Tixier, A.S.; Lemasson, Y.; Roy, C.; Fernandez, X.; Chemat, F. Extraction of aroma compounds in blackcurrant buds by alternative solvents: Theoretical and experimental solubility study. *C. R. Chim.* 2014, 17, 1268–1275.

84. Moongkarndi, P.; Kosem, N.; Kaslunka, S.; Luanratana, O.; Pongpan, N.; Neungton, N. Antiproliferation, antioxidation and induction of apoptosis by *Garcinia mangostana* (mangosteen) on SKBR3 human breast cancer cell line. *J. Ethnopharmacol.* 2004, 90, 161–166.

85. Nganlasom, J.; Suttipit, T.; Jirakulsomchok, D.; Puapairoj, A. Effects of *Centella asiatica* Linn. leaves and *Garcinia Mangostana* Linn. hull on the healing of dermal wounds in diabetic rats. *Srinagarind Med. J.* 2008, 23, 402–407.

86. Nualkaew, N.; Natsajee, N.; Morita, H.; Shimokawa, Y.; Kinjo, K.; Kushiro, T.; De-Eknamkul, W.; Ebizuka, Y.; Abe, I. Benzophenone synthase from *Garcinia mangostana* L. pericarps. *Phytochemistry* 2012, 77, 60–69.

87. Sakagami, Y.; Iinuma, M.; Piyasena, K.; Dharmaratne, H. Antibacterial activity of  $\alpha$ -mangostin against vancomycin-resistant enterococci (VRE) and synergism with antibiotics. *Phytomedicine* 2005, 12, 203–208.

88. Pothitirat, W.; Chomnawang, M.T.; Gritsanapan, W. Anti-acne inducing bacteria activity and  $\alpha$ -mangostin content of *Garcinia mangostana* fruit rind extracts from different provenience. *Songklanakarin J. Sci. Technol.* 2009, 31, 41–47.

89. Pothitirat, W.; Chomnawang, M.T.; Gritsanapan, W. Anti-acne-inducing bacterial activity of mangosteen fruit rind extracts. *Med. Princ. Pract.* 2010, 19, 281–286.

90. Nutrizio, M.; Gajdoš Kljusurić, J.; Marijanović, Z.; Dubrović, I.; Viskić, M.; Mikolaj, E.; Režek Jambrak, A. The potential of high voltage discharges for green solvent extraction of bioactive compounds and aromas from rosemary (*Rosmarinus officinalis* L.)—Computational simulation and experimental methods. *Molecules* 2020, 25(16), 711.

91. Cascant, M.M.; Breil, C.; Garrigues, S.; de la Guardia, M.; Fabiano-Tixier, A.S.; Chemat, F. A green analytical chemistry approach for lipid extraction: computation methods in the selection of green solvents as alternative to hexane. *Anal. Bioanal. Chem.* 2017, 409, 3527–3539.

92. Yara-Varon, E.; Fabiano-Tixier, A.S.; Balcells, M.; Canela-Garayoa, R.; Bily, A.; Chemat, F. Is it possible to substitute hexane with green solvents for extraction of carotenoids? A theoretical versus experimental solubility study. *RSC Adv.* 2016, 6, 27750–27759.

93. Ozturk, B.; Winterburn, J.; Gonzalez-Miquel, M. Orange peel waste valorisation through limonene extraction using bio-based solvents. *Biochem. Eng. J.* 2019, 151, 107298.

94. Linke, S.; McBride, K.; Sundmacher, K. Systematic green solvent selection for the hydroformylation of long-chain alkenes. *ACS Sustainable Chem. Eng.* 2020, 8(29), 10795–10811.

95. Neves, B.J.; Braga, R.C.; Melo-Filho, C.C.; Moreira-Filho, J.T.; Muratov, E.N.; Horta Andrade, C. QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. *Front. Pharmacol.* 2018, 9. <https://doi.org/10.3389/fphar.2018.01275>

96. Benfenati, E.; Manganaro, A.; Gini, G. VEGA-QSAR: AI inside a platform for predictive toxicology. *CEUR Workshop Proc.* 2013.

97. Abbott, A.P.; Capper, G.; Davies, D.L.; Rasheed, R.K.; Tambyrajah, V. Novel solvent properties of choline chloride/urea mixtures. *Chem. Comm.* 2003, 70–71.

98. Francisco, M.; van den Bruinhorst, A.; Kroon, M.C. Low-transition-temperature mixtures (LTTMs): A new generation of designer solvents. *Angew. Chem. Int. Ed.* 2013, 52, 3074–3085.

99. Souza, O.A.; Rinaldo, D.; Porto, C.M.; Sambrano, J.R.; Morgan, N.H.; de Souza, A.R. Computer simulation applied to structural analysis and experimental applications of natural deep eutectic solvents. In *Green Chemistry and Computational Chemistry – Shared Lessons in Sustainability*, Mammino, L. Ed. Elsevier: 2021; pp. 281–298.

100. Espino, M.; Fernández, M.A.; Gomez, F.J.V.; Silva, M.F. Natural designer solvents for greening analytical chemistry. *TrAC Trends Anal. Chem.* 2016, 76, 126–136.

101. Dai, Y.; Spronzen, J.V.; Witkamp, G.J.; Verpoorte, R.; Choi, Y.H. Natural deep eutectic solvents as new potential media for green technology. *Anal. Chim. Acta* 2013, 766, 61–68.

102. El Kantar, S.; Rajha, H.N.; Boussetta, N.; Vorobiev, E.; Maroun, R.G.; Louka, N. Green Extraction of Polyphenols from Grapefruit Peels Using High Voltage Electrical Discharges, Deep Eutectic Solvents and Aqueous Glycerol. *Food Chem.* 2019, 295, 165–171.

103. Abbott, A.P.; Boothby, D.; Capper, G.; Davies, D.L.; Rasheed, R.K. Deep Eutectic Solvents formed between choline chloride and carboxylic acids: versatile alternatives to ionic liquids. *J. Am. Chem. Soc.* 2004, 126, 9142–9147.

104. Zhu, R.; Cao, S.; Tian, J.; Luo, M.; You, J.; Chen, Z. Computational design of deep eutectic solvent functionalized ZIF-8/biochar with high selectivity for mephedrone: Experimental validation and microscopic mechanism. *J. Clean. Prod.* 2023, 399, 136687.

105. Shah, F.U.; An, R.; Muhammad, N. Editorial: properties and applications of ionic liquids in energy and environmental science. *Front. Chem.* 2020, 8, 627213. <https://doi.org/10.3389/fchem.2020.627213>.

106. Nasirpour, N.; Mohammadpourfard, M.; Heris, S.Z. Ionic liquids: promising compounds for sustainable chemical processes and applications. *Chem. Eng. Res. Des.* 2020, 160, 264–300. <https://doi.org/10.1016/j.cherd.2020.06.006>.

107. Das, R.N.; Roy, K. Advances in QSPR/QSTR models of ionic liquids for the design of greener solvents of the future. *Mol. Divers.* 2013, 17(1), 151–196. doi:10.1007/s11030-012-9413-y

108. Olivier-Bourbigou, H.; Magna, L.; Morvan, D. Ionic liquids and catalysis: recent progress from knowledge to applications. *Appl. Catal. A* 2010, 373, 1–56. doi.org/10.1016/j.apcata.2009.10.008.

109. Bruzzone, S.; Chiappe, C.; Focardi, S.E.; Pretti, C.; Renzi, M. Theoretical descriptor for the correlation of aquatic toxicity of ionic liquids by quantitative structure-toxicity relationships. *Chem Eng J* 2011, 175, 17–23. doi:10.1016/j.cej.2011.08.073

110. Reed, A.E.; Weinhold, F. Natural Atomic Orbitals and Natural Population Analysis. *J Chem Phys* 1983, 78(6), 4066–4074. <https://doi.org/10.1063/1.445134>

111. Reed, A.E.; Weinstock, R.B.; Weinhold, F. Natural Atomic Orbitals and Natural Population Analysis. *J Chem Phys* 1985, 83(2), 735–747. <https://doi.org/10.1063/1.449486>

112. Anastas, P.T.; Bartlett, L.B.; Kirchhoff, M.M.; Williamson, T.C. The role of catalysis in the design, development, and implementation of green chemistry. *Catalysis Today* 2000, 55(1-2), 11–22.

113. Anastas, P.T.; Kirchhoff, M.M.; Williamson, T.C. Catalysis as a foundational pillar of green chemistry. *Appl Catal Gen* 2001, 221, 3–13.

114. Ha, S.H.; Lan, M.N.; Lee, S.H.; Hwang, S.M.; Koo, Y.M. Enzyme *Microb Technol* 2007, 41, 480–483.

115. Kim, H.S.; Ha, S.H.; Sethaphong, L.; Koo, Y.-M.; Yingling, Y.G. The relationship between enhanced enzyme activity and structural dynamics in ionic liquids: a combined computational and experimental study. *Phys Chem Chem Phys* 2014, 16, 2944–2953.

116. Katritzky, A.R.; Kuanar, M.; Stoyanova-Slavova, I.B.; Slavov, S.H.; Dobchev, D.A.; Karelson, M.; Acree, W.E. Jr. Quantitative structure–property relationship studies on Ostwald solubility and partition coefficients of organic solutes in ionic liquids. *J Chem Eng Data* 2008, 53, 1085–1092. doi:10.1021/je700607b

117. Cao, B.; Dua, J.; Dua, D.; Suna, H.; Zhua, X.; Fu, H. Cellobiose as a model system to reveal cellulose dissolution mechanism in acetate-based ionic liquids: density functional theory study substantiated by NMR spectra. *Carbohydr Polym* 2016, **149**, 348–356.

118. Youngs, T.G.; Holbrey, J.D.; Deetlefs, M.; Nieuwenhuyzen, M.; Costa Gomes, M.F.; Hardacre, C. A molecular dynamics study of glucose solvation in the ionic liquid 1, 3-dimethylimidazolium chloride. *ChemPhysChem* 2006, **7**, 2279–2281.

119. Youngs, T.G.A.; Hardacre, C.; Holbrey, J.D. Glucose solvation by the ionic liquid 1, 3-dimethylimidazolium chloride: A simulation study. *J Phys Chem B* 2007, **111**, 13765–13774.

120. Derecskei, B.; Derecskei-Kovacs, A. Molecular dynamic studies of the compatibility of some cellulose derivatives with selected ionic liquids. *Mol Simulat* 2006, **32**, 109–115.

121. Rabideau, B.D.; Agarwal, A.; Ismail, A.E. The role of the cation in the solvation of cellulose by imidazolium-based ionic liquids. *J Phys Chem B* 2014, **118**, 1621–1629.

122. Payal, R.S.; Bejagam, K.K.; Mondal, A.; Balasubramanian, S. Dissolution of cellulose in room temperature ionic liquids: anion dependence. *J Phys Chem B* 2015, **119**, 1654–1659.

123. Bader, R.F.W. Atoms in Molecules: A Quantum Theory. Clarendon Press: Oxford, 1990.

124. Li, Y.; Wang, J.; Liu, X.; Zhang, S. Towards a molecular understanding of cellulose dissolution in ionic liquids: Anion/cation effect, synergistic mechanism and physicochemical aspects. *Chemical Science* 2018, **9**(17), 4027–4043.

125. Li, Y.; Liu, X.; Zhang, S.; Yao, Y.; Yao, X.; Xu, J.; Lu, X. Dissolving process of a cellulose bunch in ionic liquids: a molecular dynamics study. *Phys Chem Chem Phys* 2015, **17**, 17894–17905.

126. Zhao, Y.; Liu, X.; Wang, J.; Zhang, S. Insight into the cosolvent effect of cellulose dissolution in imidazolium-based ionic liquid systems. *J Phys Chem B* 2013, **117**, 9042–9049.

127. Huo, F.; Liu, Z.; Wang, W. Cosolvent or antisolvent? A molecular view of the interface between ionic liquids and cellulose upon addition of another molecular solvent. *J Phys Chem B* 2013, **117**, 11780–11792.

128. Velioglu, S.; Yao, X.; Devemy, J.; Ahunbay, M.G.; Tantekin-Ersolmaz, S.B.; Dequidt, A.; Costa Gomes, M.F.; Padua, A.A. Solvation of a cellulose microfibril in imidazolium acetate ionic liquids: effect of a cosolvent. *J Phys Chem B* 2014, **18**, 14860–14869.

129. Parthasarathi, R.; Balamurugan, K.; Shi, J.; Subramanian, V.; Simmons, B.A.; Singh, S. Theoretical insights into the role of water in the dissolution of cellulose using IL/water mixed solvent systems. *J Phys Chem B* 2015, **119**, 14339–14349.

130. Rabideau, B.D.; Ismail, A.E. Mechanisms of hydrogen bond formation between ionic liquids and cellulose and the influence of water content. *Phys Chem Chem Phys* 2015, **17**, 5767–5775.

131. Zhao, X.; Liu, R. Recent progress and perspectives on the toxicity of carbon nanotubes at organism, organ, cell, and biomacromolecule levels. *Environ. Int.* 2012, **40**, 244–256.

132. Joudeh, N.; Linke, D. Nanoparticle classification, physicochemical properties, characterization, and applications: a comprehensive review for biologists. *J Nanobiotechnol* 2022, **20**, 262. <https://doi.org/10.1186/s12951-022-01477-8>

133. Hoet, P.H.M.; Brüske-Hohlfeld, I.; Salata, O.V. Nanoparticles known and unknown risks. *J. Nanobiotechnol.* 2004, **2**, 12.

134. Jeevanandam, J.; Barhoum, A.; Chan, Y.S.; Dufresne, A.; Danquah, M.K. Review on nanoparticles and nanostructured materials: history, sources, toxicity and regulations. *Beilstein J. Nanotechnol.* 2018, **9**, 1050–1074. doi:10.3762/bjnano.9.98

135. Turan, N.B.; Erkan, H.S.; Engin, G.O.; Bilgili, M.S. Nanoparticles in the aquatic environment: Usage, properties, transformation and toxicity—A review. *Process Safety and Environmental Protection* 2019, **130**, 238–249.

136. Barnard, A.S. How can ab initio simulations address risks in nanotech? *Nat. Nanotechnol.* 2009, **4**, 332–335.

137. Gajewicz, A.; Rasulev, B.; Dinadayalane, T.C.; Urbaszek, P.; Puzyn, T.; Leszczynska, D.; Leszczynski, J. Advancing risk assessment of engineered nanomaterials: application of computational approaches. *Advanced drug delivery reviews* 2012, **64**, 1663–1693.

138. Cook, E.; Labiento, G.; Chauhan, B.P.S. Fundamental Methods for the Phase Transfer of Nanoparticles. *Molecules* 2021, **26**(20), 6170.

139. Xu, N.; Liu, Z.; Lv, Y.; Liu, S.; Yang, S.; Zhang, W. Improved Coarse-Grained Model for Nanoparticles Based on the Martini Force Field and Its Application in Molecular Dynamics Simulation on Gel Ink. *Langmuir* 2022, **38**, 14172–14184. <https://doi.org/10.1021/acs.langmuir.2c02185>

140. Wagener, P.; Jakobi, J.; Rehbock, C.; Chakravadhanula, V.S.K.; Thede, C.; Wiedwald, U.; Bartsch, M.; Kienle, L.; Barcikowski, S. Solvent-surface interactions control the phase structure in laser-generated iron-gold core-shell nanoparticles. *Sci Rep* 2016, **6**, 23352. <https://doi.org/10.1038/srep23352>

141. Leekumjorn, S.; Gullapalli, S.; Wong, M.S. Understanding the Solvent Polarity Effects on Surfactant-Capped Nanoparticles. *J. Phys. Chem. B* 2012, **116**, 13063–13070. <https://doi.org/10.1021/jp307985c>

142. Filippov, A.V.; Starov, V. Interaction of Nanoparticles in Electrolyte Solutions. *J. Phys. Chem. B* 2023, **127**, 6562–6572. doi.org/10.1021/acs.jpcb.3c01220J

143. Chinthia, D.; Veesam, S. K.; Boattini, E.; Filion, L.; Punnathanam, S. N. Modeling of effective interactions between ligand coated nanoparticles through symmetry functions. *J. Chem. Phys.* 2021, 155(24), 244901. <https://doi.org/10.1063/5.0072272>

144. Hafiz, M.; Hassanein, A.; Talhami, M.; AL-Ejji, M.; Hassan, M.K.; Hawari, A.H. Magnetic nanoparticles draw solution for forward osmosis: Current status and future challenges in wastewater treatment. *J. Environ. Chem. Eng.* 2022, 10(6), 108955. <https://doi.org/10.1016/j.jece.2022.108955>

145. Wörle-Knirsch, J.M.; Krug, H.F. Cause I'm CNT, not dynamite. *Nano Today* 2006, 1(4), 48.

146. Lacerda, L.; Bianco, A.; Prato, M.; Kostarelos, K. Carbon nanotubes as nanomedicines: from toxicology to pharmacology. *Adv. Drug Deliv. Rev.* 2006, 58(14), 1460–1470.

147. Sasidharan, A.; Panchakarla, L. S.; Chandran, P.; Menon, D.; Nair, S.; Rao, C. N. R.; Koyakutty, M. Differential nano-bio interactions and toxicity effects of pristine versus functionalized graphene. *Nanoscale* 2011, 3(6), 2461–2464.

148. Morimoto, Y.; Kobayashi, N.; Shinohara, N.; Myojo, T.; Tanaka, I.; Nakanishi, J. Hazard assessments of manufactured nanomaterials. *J. Occup. Health* 2010, 52, 325–334.

149. Schinwald, A.; Murphy, F.A.; Jones, A.; MacNee, W.; Donaldson, K. Graphene-based nanoplatelets: a new risk to the respiratory system as a consequence of their unusual aerodynamic properties. *ACS Nano* 2012, 6, 736–746.

150. Mananghaya, M.; Rodulfo, E.; Santos, G.N.; Villagracia, A.R. Theoretical investigation on the solubilization in water of functionalized single-wall carbon nanotubes. *J. Nanotechnol.* 2012, 2012, 780815.

151. Gao, H.; Kong, Y.; Cui, D.; Ozkan, C.Z. Spontaneous insertion of DNA oligonucleotides into carbon nanotubes. *Nano Lett.* 2003, 3, 471–473.

152. Monajjemi, M.; Mollaamin, F. Molecular modeling study of drug–DNA combined to single walled carbon nanotube. *J. Clust. Sci.* 2012, 23, 259–272.

153. Onsager, L. Electric Moments of Molecules in Liquids. *J. Am. Chem. Soc.* 1936, 58, 1486–1493. <https://doi.org/10.1021/ja01299a050>

154. Obata, S.; Honda, K. Dynamic behavior of carbon nanotube and bio-/artificial surfactants complexes in an aqueous environment. *J. Phys. Chem. C* 2011, 115, 19659–19667.

155. Nivetha, G.F.; Vettrivelan, V.; Muthu, S.; Prasath, M. Adsorption behavior, different green solvent effect and surface enhanced Raman spectra (SERS) investigation on inhibition of SARS-CoV-2 by antineoplastic drug Carmofur with silver/gold/platinum loaded silica nanocomposites: A combined computational analysis and molecular modelling approach. *Results Chem.* 2023, 6, 101096.

156. Stals, P.J.M.; Cheng, C.-Y.; van Beek, L.; Wauters, A.C.; Palmans, A.R.A.; Han, S.; Meijer, E.W. Surface water retardation around single-chain polymeric nanoparticles: Critical for catalytic function? *Chem. Sci.* 2016, 7, 2011–2015.

157. Odegard, G.M. Computational Multiscale Modeling - Nanoscale to Macroscale. *Comprehensive Composite Materials II*, Elsevier 2018, Volume 6, 28–51.

158. Lavino, A.D.; Di Pasquale, N.; Carbone, P.; Marchisio, D.L. A novel multiscale model for the simulation of polymer flash nano-precipitation. *Chem. Eng. Sci.* 2017, 171, 485–494. <https://doi.org/10.1016/j.ces.2017.04.047>

159. Sperger, T.; Sanhueza, I.A.; Kalvet, I.; Schoenebeck, F. Computational studies of synthetically relevant homogeneous organometallic catalysis involving Ni, Pd, Ir, and Rh: an overview of commonly employed DFT methods and mechanistic insights. *Chem. Rev.* 2015, 115(17), 9532–9586.

160. Risthaus, T.; Grimme, S. Benchmarking of London Dispersion-Accounting Density Functional Theory Methods on Very Large Molecular Complexes. *J. Chem. Theory Comput.* 2013, 9, 1580–1591.

161. Grimme, S. Comment on: “On the Accuracy of DFT Methods in Reproducing Ligand Substitution Energies for Transition Metal Complexes in Solution: The Role of Dispersive Interactions” by H. Jacobsen and L. Cavallo. *ChemPhysChem* 2012, 13, 1407–1409.

162. Chaudhari, M.I.; Nair, J.R.; Pratt, L.R.; Soto, F.A.; Balbuena, P.B.; Rempe, S.B. Scaling atomic partial charges of carbonate solvents for lithium ion solvation and diffusion. *J. Chem. Theory Comput.* 2016, 12, 5709–5718.

163. Reddy, S.K.; Balasubramanian, S. Liquid dimethyl carbonate: a quantum chemical and molecular dynamics study. *J. Phys. Chem. B* 2012, 116, 14892–14902.

164. Barnes, T.A.; Kaminski, J.W.; Borodin, O.; Miller III, T.F. Ab initio characterization of the electrochemical stability and solvation properties of condensed-phase ethylene carbonate and dimethyl carbonate mixtures. *J. Phys. Chem. C* 2015, 119, 3865–3880.

165. Atilhan, M.; Aparicio, S. Properties of dialkylcarbonate + 1-alkanol mixtures at the vacuum interface. *J. Phys. Chem. C* 2016, 120, 29126–29134.

166. Amis, E.S. Solvent Effects on Reaction Rates and Mechanisms; Academic Press: New York, 1966.

167. Amis, E.S.; Hinton, J.F. Solvent influence on rates and mechanisms. In Solvent Effects on Chemical Phenomena; Amis, E.S. Hinton, J.F. Eds.; Academic Press: New York, 1973; Volume 1, pp. 207–449.

168. Kostal, J.; Jorgensen, W.L. Thorpe-ingold acceleration of oxirane formation is mostly a solvent effect. *J. Am. Chem. Soc.* 2010, 132, 8766–8773.

169. Varghese, J.J.; Mushrif, S.H. Origins of Complex Solvent Effects on Chemical Reactivity and Computational Tools to Investigate Them: A Review. *React. Chem. Eng.* 2019, 4(2), 165–206.

170. Basdogan, Y.; Maldonado, A.M.; Keith, J.A. Advances and challenges in modeling solvated reaction mechanisms for renewable fuels and chemicals. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2020, 10(2), e1446.

171. Zhu, Q.S.; Wallentine, S.K.; Deng, G.H.; Rebstock, J.A.; Baker, L.R. The Solvation-Induced Onsager Reaction Field Rather than the Double-Layer Field Controls CO<sub>2</sub> Reduction on Gold. *JACS Au* 2022, 2, 472–482. <https://doi.org/10.1021/jacsau.1c00512>

172. Sorenson, S.A.; Patrow, J.G.; Dawlaty, J.M. Solvation reaction field at the interface measured by vibrational sum frequency generation spectroscopy. *J. Am. Chem. Soc.* 2017, 139, 2369–2378.

173. Rowley, C.N.; Roux, B. The solvation structure of Na<sup>+</sup> and K<sup>+</sup> in liquid water determined from high level ab initio molecular dynamics simulations. *J. Chem. Theory Comput.* 2012, 8(10), 3526–3535.

174. Liu, J.C.; Liu, R.X.; Cao, Y.; Chen, M. Solvation structures of calcium and magnesium ions in water with the presence of hydroxide: a study by deep potential molecular dynamics. *Phys. Chem. Chem. Phys.* 2023, 25, 983–993.

175. Hu, Y.S.; Pan, H. Solvation Structures in Electrolyte and the Interfacial Chemistry for Na-Ion Batteries. *ACS Energy Lett.* 2022, 7(12), 4501–4503.

176. Forero-Saboya, J.D.; Marchante, E.; Araujo, R.B.; Monti, D.; Johansson, P.; Ponrouch, A. Cation Solvation and Physicochemical Properties of Ca Battery Electrolytes. *J. Phys. Chem. C* 2019, 123(49), 29524–29532. DOI: 10.1021/acs.jpcc.9b07308

177. Rajput, N.N.; Seguin, T.J.; Wood, B.M.; Qu, X.; Persson, K.A. Elucidating Solvation Structures for Rational Design of Multivalent Electrolytes—A Review. *Top Curr Chem (Z)* 2018, 376, 79–124. <https://doi.org/10.1007/s41061-018-0195-2>

178. López Barreiro, D.; Yeo, J.; Tarakanova, A.; Martin-Martinez, F.J.; Buehler, M.J. Multiscale Modeling of Silk and Silk-Based Biomaterials—A Review. *Macromol. Biosci.* 2019, 19(3), 1800253.

179. Dinjaski, N.; Ebrahimi, D.; Qin, Z.; Giordano, J.E.M.; Ling, S.; Buehler, M.J.; Kaplan, D.L.J. Predicting rates of in vivo degradation of recombinant spider silk proteins. *Tissue Eng. Regener. Med.* 2018, 12, 97–105.

180. Finney, J.L. Hydration processes in biological and macromolecular systems. *Faraday Discuss.* 1996, 103, 1–18.

181. Schoenborn, B.P.; Garcia, A.; Knott, R. Hydration in protein crystallography. *Prog. Biophys. Mol. Biol.* 1995, 64, 105–119.

182. Raschke, T.M. Water structure and interactions with protein surfaces. *Curr. Opin. Struct. Biol.* 2006, 16, 152–159.

183. Mazur, K.; Heisler, I.A.; Meech, S.R. Ultrafast dynamics and hydrogen-bond structure in aqueous solutions of model peptides. *J. Phys. Chem. B* 2010, 114, 10684–10691.

184. Garcia-Sosa, A.T.; Mancera, R.L.; Dean, P.M. WaterScore: A novel method for distinguishing between bound and displaceable water molecules in the crystal structure of the binding site of protein-ligand complexes. *J. Mol. Model.* 2003, 9, 172–182.

185. Olano, L.R.; Rick, S.W. Hydration free energies and entropies for water in protein interiors. *J. Am. Chem. Soc.* 2004, 126, 7991–8000.

186. Patel, H.; Gruning, B.A.; Gunther, S.; Merfort, I. PyWATER: A PyMOL plug-in to find conserved water molecules in proteins by clustering. *Bioinformatics* 2014, 30, 2978–2980.

187. Rahaman, O.; Kalimeri, M.; Melchionna, S.; Hénin, J.; Sterpone, F. On the Role of Internal Water on Protein, Thermal Stability: The Case of Homologous G-domains. *J. Phys. Chem. B* 2015, 119, 8939–8949.

188. Chakraborty, D.; Taly, A.; Sterpone, F. Stay Wet, Stay Stable? How Internal Water Helps Stability of Thermophilic Proteins. *J. Phys. Chem. B* 2015, 119, 12760–12770.

189. Ebbinghaus, S.; Kim, S.J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. An extended dynamical hydration shell around proteins. *Proc. Natl. Acad. Sci. USA* 2007, 104, 20749–20752.

190. Mattea, C.; Qvist, J.; Halle, B. Dynamics at the Protein-Water Interface from (17)O Spin Relaxation in Deeply Supercooled Solutions. *Biophys. J.* 2008, 95, 2951–2963.

191. Teixeira, J. Dynamics of hydration water in proteins. *Gen. Physiol. Biophys.* 2009, 28, 168–173.

192. Born, B.; Kim, S.J.; Ebbinghaus, S.; Gruebele, M.; Havenith, M. The terahertz dance of water with the proteins: The effect of protein flexibility on the dynamical hydration shell of ubiquitin. *Faraday Discuss.* 2009, 141, 161–173.

193. Wallnoefer, H.G.; Handschuh, S.; Liedl, K.R.; Fox, T. Stabilizing of a Globular Protein by a Highly Complex Water Network: A Molecular Dynamics Simulation Study on Factor Xa. *J. Phys. Chem. B* 2010, 114, 7405–7412.

194. Clark, P.L.; Ugrinov, K.G. Measuring Cotranslational Folding of Nascent Polypeptide Chains on Ribosomes. In *Methods in Enzymology*; Elsevier: 2009; Volume 466, pp. 567–590. DOI: 10.1016/S0076-6879(09)66024-9

195. Fersht, A. R. Structure and mechanism in protein science, 3rd ed.; W.H. Freeman and Company: New York, USA, 1999.

196. Jonák, J. Protein folding and misfolding, diseases associated with protein misfolding & aggregation. *Curr. Opin. Struct. Biol.* 2004, 14, 616–621.

197. Reynaud, E. Protein Misfolding and Degenerative Diseases. *Nature Education* 2010, 3(9), 28.

198. Matus, S.; Glimcher, L.H.; Hetz, C. Protein folding stress in neurodegenerative diseases: a glimpse into the ER. *Curr. Opin. Cell Biol.* 2011, 23, 239–252.

199. Sweeney, P.; Park, H.; Baumann, M.; Dunlop, J.; Frydman, J.; Kopito, R.; McCampbell, A.; Leblanc, G.; Venkateswaran, A.; Nurmi, A.; Hodgson, R. Protein misfolding in neurodegenerative diseases: implications and strategies. *Transl. Neurodegener.* 2017, 6, 1–13. doi: 10.1186/s40035-017-0077-5.

200. Fare, C.M.; Shorter, J. (Dis)Solving the problem of aberrant protein states. *Disease Models & Mechanisms* 2021, 14(5), dmm048983. doi:10.1242/dmm.048983

201. Khanam, H.; Ali, A.; Asif, M. Neurodegenerative diseases linked to misfolded proteins and their therapeutic approaches: A review. *European Journal of Medicinal Chemistry* 2016, 124, 1121–1141. <https://doi.org/10.1016/j.ejmech.2016.08.006>

202. Strodel, B. Amyloid aggregation simulations: challenges, advances and perspectives. *Current Opinion in Structural Biology* 2021, 67, 145–152. <https://doi.org/10.1016/j.sbi.2020.10.019>

203. Ochneva, A.; Zorkina, Y.; Abramova, O.; Pavlova, O.; Ushakova, V.; Morozova, A.; Zubkov, E.; Pavlov, K.; Gurina, O.; Chekhonin, V. Protein Misfolding and Aggregation in the Brain: Common Pathogenetic Pathways in Neurodegenerative and Mental Disorders. *Int. J. Mol. Sci.* 2022, 23(22), 14498. <https://doi.org/10.3390/ijms232214498>

204. Baker, J.D.; Webster, J.M.; Shelton, L.B.; Koren III, J.; Uversky, V.N.; Blair, L.J.; Dickey, C.A. Neurodegenerative Diseases as Protein Folding Disorders. In *The Molecular and Cellular Basis of Neurodegenerative Diseases, Underlying Mechanisms*; Elsevier, 2018; pp. 243–267. <https://doi.org/10.1016/B978-0-12-811304-2.00009-2>

205. Del Lucent, V.; Vishal, V.; Pande, V.S. Protein folding under confinement: A role for solvent. *Biophysics and Computational Biology* 2007, 104, 10430–10434. <https://doi.org/10.1073/pnas.0608256104>

206. Onuchic, J.N.; Luthey-Schulten, Z.; Wolynes, P.G. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* 1997, 48, 545–600.

207. Blanco, F.J.; Rivas, G.; Serrano, L. A short linear peptide that folds in a native stable  $\beta$ -hairpin in aqueous solution. *Nat Struct Biol* 1994, 1(9), 584–590.

208. Blanco, F.J.; Serrano, L. Folding of protein g b1 domain studied by the conformational characterization of fragments comprising its secondary structure elements. *Eur J Biochem.* 1995, 230(2), 634–649.

209. Munoz, V.; Henry, E.R.; Hofrichter, J.; Eaton, W.A. A statistical mechanical model for  $\beta$ -hairpin kinetics. *Proc Natl Acad Sci USA* 1998, 95(11), 5872–5879.

210. Zhou, R.; Berne, B.J. Can a continuum solvent model reproduce the free energy landscape of a  $\beta$ -hairpin folding in water? *Proc. Natl. Acad. Sci. U.S.A.* 2002, 99, 12777–12782.

211. Zhou, R. Free energy landscape of protein folding in water: explicit vs. implicit solvent. *Proteins: Structure, Function, and Bioinformatics* 2003, 53(2), 148–161.

212. Freddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K. Tenmicrosecond MD simulation of a fast-folding WW domain. *Biophys. J.* 2008, 94, L75–L77.

213. Jang, S.; Kim, E.; Shin, S.; Pak, Y. Ab initio folding of helix bundle proteins using molecular dynamics simulations. *J. Am. Chem. Soc.* 2003, 125, 14841–14846.

214. Vaiana, S.M.; Manno, M.; Emanuele, A.; Palma-Vittorelli, M.B.; Palma, M.U. The Role of Solvent in Protein Folding and in Aggregation. *J. Biol. Phys.* 2001, 27, 133–145.

215. Juraszek, J.; Bolhuis, P.G. Sampling the multiple folding mechanisms of Trp-cage in explicit solvent. *Proc. Natl. Acad. Sci. U.S.A.* 2006, 103, 15859–15864.

216. Paschek, D.; Hempel, S.; Garcia, A.E. Computing the stability diagram of the Trp-cage miniprotein. *Proc. Natl. Acad. Sci. U.S.A.* 2008, 105, 17754–17759.

217. Sanbonmatsu, K.Y.; Garcia, A.E. Structure of Met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics. *Proteins* 2002, 46, 225–234.

218. McKnight, C.J.; Doering, D.S.; Matsudaira, P.T.; Kim, P.S. A thermostable 35-residue subdomain within villin headpiece. *J. Mol. Biol.* 1996, 260, 126–134.

219. Fernandez, A.; Shen, M.Y.; Colubri, A.; Sosnick, T.R.; Berry, R.S.; Freed, K.F. Large-scale context in protein folding: villin headpiece. *Biochemistry* 2003, 42, 664–671.

220. Lei, H.; Duan, Y. Two-stage folding of HP-35 from ab initio simulations. *J. Mol. Biol.* 2007, 370, 196–206.

221. Yang, J.S.; Wallin, S.; Shakhnovich, E.I. Universality and diversity of folding mechanics for three-helix bundle proteins. *Proc. Natl. Acad. Sci. U.S.A.* 2008, 105, 895–900.

222. Zagrovic, B.; Snow, C.D.; Shirts, M.R.; Pande, V.S. Simulation of folding of a small  $\alpha$ -helical protein in atomistic detail using worldwide-distributed computing. *J. Mol. Biol.* 2002, 323, 927–937.

223. Lei, H.; Wu, C.; Liu, H.; Duan, Y. Folding free-energy landscape of villin headpiece subdomain from molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* 2007, 104, 4925–4930.

224. Duan, Y.; Kollman, P.A. Pathways to a protein folding intermediate observed in a 1-ms simulation in aqueous solution. *Science* 1998, 282, 740–744.

225. Kubelka, J.; Eaton, W.A.; Hofrichter, J. Experimental tests of villin subdomain folding simulations. *J. Mol. Biol.* 2003, 329, 625–630.

226. Lei, H.; Deng, X.; Wang, Z.; Duan, Y. The fast-folding HP35 double mutant has a substantially reduced primary folding free energy barrier. *J. Chem. Phys.* 2008, 129, 155104–155107.

227. Freddolino, P.L.; Schulten, K. Common structural transitions in explicit-solvent simulations of villin headpiece folding. *Biophys. J.* 2009, 97(8), 2338–2347.

228. Jayachandran, G.; Vishal, V.; Pande, V.S. Using massively parallel simulation and Markovian models to study protein folding: examining the dynamics of the villin headpiece. *J. Chem. Phys.* 2006, 124(16), 164902.

229. Oshima, H.; Kinoshita, M. Essential roles of protein-solvent many-body correlation in solvent-entropy effect on protein folding and denaturation: Comparison between hard-sphere solvent and water. *J. Chem. Phys.* 2015, 142(14), 145103. <https://doi.org/10.1063/1.4917075>

230. Wang, L. The contributions of surface charge and geometry to protein-solvent interaction. *arXiv:1605.01155v1* [q-bio.BM], 2016.

231. Wang, L. The solvent-excluded surfaces of water-soluble proteins. *bioRxiv* 2018, 294082. <https://doi.org/10.1101/294082>

232. Yu, Y.; Wang, J.; Shao, Q.; Shi, J.; Zhu, W. The effects of organic solvents on the folding pathway and associated thermodynamics of proteins: a microscopic view. *Sci. Rep.* 2016, 6, 19500. DOI: 10.1038/srep19500

233. van der Vegt, N.F.A.; Nayar, D. The Hydrophobic Effect and the Role of Cosolvents. *J. Phys. Chem. B.* 2017, 121, 9986–9998.

234. Davis, C.M.; Gruebele, M.; Sukenik, S. How does solvation in the cell affect protein folding and binding? *Curr. Opin. Struct. Biol.* 2018, 48, 23–29.

235. Mishra, A.; Ranganathan, S.; Jayaram, B.; Sattar, A. Role of solvent accessibility for aggregation-prone patches in protein folding. *Sci. Rep.* 2018, 8, 12896. <https://doi.org/10.1038/s41598-018-31289-6>

236. Hayashi, T.; Inoue, M.; Yasuda, S.; Petretto, E.; Škrbić, T.; Giacometti, A.; Kinoshita, M. Universal effects of solvent species on the stabilized structure of a protein. *J. Chem. Phys.* 2018, 149(4), 1–23.

237. Arakawa, T. Protein-solvent interaction. *Biophys. Rev.* 2018, 10(2), 203–208. doi: 10.1007/s12551-017-0339-6

238. Bucciarelli, S.; Sayedi, E.S.; Osella, S.; Trzaskowski, B.; Vissing, K.J.; Vestergaard, B.; Foderà, V. Disentangling the role of solvent polarity and protein solvation in folding and self-assembly of  $\alpha$ -lactalbumin. *J. Colloid Interface Sci.* 2020, 561, 749–761

239. Stöhr, M.; Tkatchenko, A. Quantum mechanics of proteins in explicit water: The role of plasmon-like solute-solvent interactions. *Sci. Adv.* 2019, 5. <https://www.science.org>

240. Zhu, L.; Yang, W.; Meng, Y.Y.; Xiao, X.; Guo, Y.; Pu, X.; Li, M. Effects of organic solvent and crystal water on gamma-chymotrypsin in acetonitrile media: Observations from molecular dynamics simulation and DFT calculation. *J. Phys. Chem. B* 2012, 116, 3292–3304.

241. Roccatano, D. Computer simulations study of biomolecules in non-aqueous or cosolvent/water mixture solutions. *Curr. Protein Pept. Sci.* 2008, 9, 407–426.

242. Micaelo, N.M.; Soares, C.M. Modeling hydration mechanisms of enzymes in nonpolar and polar organic solvents. *FEBS J.* 2007, 274, 2424–2436.

243. Yang, L.; Dordick, J.S.; Garde, S. Hydration of enzyme in nonaqueous media is consistent with solvent dependence of its activity. *Biophys. J.* 2004, 87, 812–821.

244. Meng, Y.; Yuan, Y.; Zhu, Y.; Guo, Y.; Li, M.; Wang, Z.; Pu, X.; Jiang, L. Effects of organic solvents and substrate binding on trypsin in acetonitrile and hexane media. *J. Mol. Model.* 2013, 19, 3749–3766.

245. Falk, M.; Hartman, K.A.; Lord, R.C. Hydration of deoxyribonucleic acid. I. A gravimetric study. *J. Am. Chem. Soc.* 1962, 84, 3843–3846.

246. Falk, M.; Hartman, K.A.; Lord, R.C. Hydration of deoxyribonucleic acid. II. An infrared study. *J. Am. Chem. Soc.* 1963, 85, 387–391.

247. Falk, M.; Hartman, K.A.; Lord, R.C. Hydration of deoxyribonucleic acid. III. A spectroscopic study of the effect of hydration on the structure of deoxyribonucleic acid. *J. Am. Chem. Soc.* 1963, 85, 391–394.

248. Tunis, M-J.B.; Hearst, J.E. On the hydration of DNA. I. Preferential hydration and stability of DNA in concentrated trifluoracetate solution. *Biopolymers* 1968, 6, 1325–1344.

249. Tunis, M-J.B.; Hearst, J.E. On the hydration of DNA. II. Base composition dependence of the net hydration of DNA. *Biopolymers* 1968, 6, 1345–1353.

250. Falk, M.; Poole, A.G.; Goymom, C.G. Infrared study of the state of water in the hydration shell of DNA. *Can. J. Chem.* 1970, 48, 1536–1542

251. Levy, Y.; Onuchic, J.N. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct.* 2006, 35, 389–415.

252. Norberg, J.; Nilsson, L. Solvent Influence on Base Stacking. *Biophys. J.* 1998, 74, 394–402. [https://doi.org/10.1016/S0006-3495\(98\)77796-3](https://doi.org/10.1016/S0006-3495(98)77796-3)

253. Zhao, H. DNA Stability in Ionic Liquids and Deep Eutectic Solvents. *J. Chem. Technol. Biotechnol.* 2015, 90, 19–25. doi: 10.1002/jctb.4511.

254. Bonner, G.; Klibanov, A.M. Structural stability of DNA in nonaqueous solvents. *Biotechnology and Bioengineering*, 2000, 86, 339–344.

255. Shen, X.; Gu, B.; Che, S.A.; Zhang, F.S. Solvent effects on the conformation of DNA dodecamer segment: A simulation study. *J. Chem. Phys.* 2011, 135(3), 034509. <https://doi.org/10.1063/1.3610549>

256. Arcella, A.; Portella, G.; Colleparo-Guevara, R.; Chakraborty, D.; Wales, D.J.; Orozco, M. Structure and Properties of DNA in Apolar Solvents. *J. Phys. Chem. B* 2014, 118, 8540–8548. <https://doi.org/10.1021/jp503816r>

257. Nakano, S.I.; Sugimoto, N. The structural stability and catalytic activity of DNA and RNA oligonucleotides in the presence of organic solvents. *Biophys. Rev.* 2016, 8, 11–23. doi: 10.1007/s12551-015-0188-0.

258. Nan, Z.; Ming-Ru, L.; Feng-Shou, Z. Ethylene glycol solution-induced DNA conformational transitions. *Chinese Physics B* 2018, 27(11), 113102.

259. Kuzmanov, U.; Emili, A. Protein-protein interaction networks: probing disease mechanisms using model systems. *Genome Med.* 2013, 5, 5(4), 1–12 <https://doi.org/10.1186/gm441>

260. Garland, W.; Benezra, R.; Chaudhary, J. Chapter Fifteen - Targeting Protein–Protein Interactions to Treat Cancer—Recent Progress and Future Directions. *Annu. Rep. Med. Chem.* 2013, 48, 227–245. <https://doi.org/10.1016/B978-0-12-417150-3.00015-6>

261. Seychell, B.C.; Beck, T. Molecular basis for protein–protein interactions. *Beilstein J. Org. Chem.* 2021, 17, 1–10. <https://doi.org/10.3762/bjoc.17.1>

262. Schreiber, G. Protein–Protein Interaction Interfaces and their Functional Implications. In *Protein – Protein Interaction Regulators*, Roy, S. Fu, H. Eds.; The Royal Society of Chemistry, 2020, pp. 1–24.

263. Ebel, C. Solvent Mediated Protein-Protein Interactions. In *Protein Interactions: Biophysical Approaches for the Study of Complex Reversible Systems*; Springer US: Boston, MA, USA, 2007; Volume 5, pp. 255–287. [https://doi.org/10.1007/978-0-387-35966-3\\_9](https://doi.org/10.1007/978-0-387-35966-3_9)

264. Vagenende, V.; Han, A.X.; Pek, H.B.; Loo, B.L.W. Quantifying the Molecular Origins of Opposite Solvent Effects on Protein-Protein Interactions. *PLOS Comput. Biol.* 2013. <https://doi.org/10.1371/journal.pcbi.1003072>

265. Levy, Y.; Cho, S.S.; Onuchic, J.N.; Wolynes, P.G. A survey of flexible protein binding mechanisms and their transition states using native topology based energy landscapes. *J. Mol. Biol.* 2005, 346, 1121–1145.

266. Ahmad, M.; Gu, W.; Geyer, T.; Helms, V. Adhesive water networks facilitate binding of protein interfaces. *Nat. Commun.* 2011, 2, 261.

267. Ghanakota, P.; van Vlijmen, H.; Sherman, W.; Beuming, T. Large-Scale Validation of Mixed-Solvent Simulations to Assess Hotspots at Protein–Protein Interaction Interfaces. *J. Chem. Inf. Model.* 2018, 58, 4, 784–793.

268. Mayol, G.F.; Defelipe, L.A.; Arcon, J.P.; Turjanski, A.; Martí, M.A. Solvent Sites Improve Docking Performance of Protein–Protein Complexes and Protein–Protein Interface-Targeted Drugs. *J. Chem. Inf. Model.* 2022, 62, 15, 3577–3588.

269. Morrissey-Andrews, A.; Shea, J.-E. Computational Studies of Protein Aggregation: Methods and Applications. *Annu. Rev. Phys. Chem.* 2015, 66, 643–666. doi: 10.1146/annurev-physchem-040513-103738

270. Kouza, M.; Kolinski, A.; Buhimschi, I.A.; Kloczkowski, A. Explicit-Solvent All-Atom Molecular Dynamics of Peptide Aggregation. In *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes*; Liwo, A. Ed.; Springer Series on Bio- and Neurosystems: Cham, 2019; Volume 8, pp. 541–558. [https://doi.org/10.1007/978-3-319-95843-9\\_16](https://doi.org/10.1007/978-3-319-95843-9_16)

271. Klimov, D.K.; Straub, J.E.; Thirumalai, D. Aqueous urea solution destabilizes A-beta(16–22) oligomers. *Proc. Natl. Acad. Sci. USA* 2004, 101, 14760–14765.

272. Matubayasi, N.; Masutani, K. Energetics of co-solvent effect on peptide aggregation. *Biophys. Physicobiol.* 2019, 16, 185–195. doi: 10.2142/biophysico.16.0\_185

273. Emperador, A. Accurate Description of Protein–Protein Recognition and Protein Aggregation with the Implicit-Solvent-Based PACSAB Protein Model. *Polymers* 2021, 13(23), 4172, DOI:10.3390/polym13234172

274. Stephens, A.D.; Kölbel, J.; Moons, R.; Chung, C.W.; Ruggiero, M.T.; Mahmoudi, N.; Shmool, T.A.; McCoy, T.M.; Nietlispach, D.; Routh, A.F.; Sobott, F.; Zeitler, J.A.; Kaminski Schierle, G.S. Decreased water mobility contributes to increased  $\alpha$ -Synuclein aggregation. *AngewandteChemie* 2023, 62(7), e202212063 <https://doi.org/10.1002/anie.202212063>

275. Bonaccorsi, R.; Scrocco, E.; Tomasi, J. Simple theoretical models for biochemical systems, with applications to DNA. *Proc. Int. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci.* 1985, 8, 627–634.

276. Tomasi, J. Effective and practical ways of introducing the effect of the solvent in the theoretical evaluation of conformational properties of biomolecules. In QSAR in drug design and toxicology, Proceedings of the Sixth European Symposium on Quantitative Structure-Activity Relationships, Portorož - Portorose, Yugoslavia, September 22-26, 1986; Hadži, D., Jerman-Blažič, B. Eds.; Elsevier: 1987.

277. Spyros, F.; Cozzini, P.; Bertoli, C.; Marabotti, A.; Kellogg, G.E.; Mozzarelli, A. Energetics of the protein-DNA-water interaction. *BMC Struct Biol* 2007, 7(1), 1–18. <https://doi.org/10.1186/1472-6807-7-4>

278. Serf, S.; Nilsson, I. Structure, interaction, dynamics and solvent effects on the DNA-EcoRI complex in aqueous solution from molecular dynamics simulation. *Biophys. J.* 1999, 77, 1782–800.

279. Schwabe, J.W. The role of water in protein-DNA interactions. *Curr. Opin. Struct. Biol.* 1997, 7, 126–34.

280. Woda, J.; Schneider, B.; Patel, K.; Mistry, K.; Berman, H.M. An analysis of the relationship between hydration and protein-DNA interactions. *Biophys. J.* 1998, 75, 2170–77.

281. Harris, L.F.; Sullivan, M.R.; Popken-Harris, P.D. Molecular dynamics simulations in solvent of the bacteriophage 434 cl repressor protein DNA binding domain amino acids (r1-69) in complex with its cognate operator (OR1) DNA sequence. *J. Biomol. Struct. Dyn.* 1999, 19, 1–17.

282. Suenaga, A.; Yatsu, C.; Komeiji, Y.; Uebeyasi, M.; Meguro, T.; Yamato, I. Molecular dynamics simulation of trp repressor-operator complex. Analysis of hydrogen bond patterns of protein-DNA interactions. *J. Mol. Struct.* 2000, 526, 209–18.

283. Tsui, V.; Radhakrishnan, I.; Wright, P.E.; Case, D.A. NMR and molecular dynamics studies of the hydration of a zinc finger DNA complex. *J. Mol. Biol.* 2000, 302, 1101–17.

284. Giudice, E.; Lavery, R. Simulations of nucleic acids and their complexes. *Acc. Chem. Res.* 2002, 35, 350–57.

285. Yang, L.; Beard, W.A.; Wilson, S.H.; Roux, B.; Broyde, S.; Schlick, T. Local deformations revealed by dynamics simulations of DNA polymerase  $\beta$  with DNA mismatches at the primer terminus. *J. Mol. Biol.* 2002, 321, 459–78.

286. Marco, E.; Garcia-Nieto, R.; Gago, F. Assessment of molecular dynamics simulations of the structural determinants of DNA-binding specificity for transcription factor Sp1. *J. Mol. Biol.* 2003, 328, 9–32.

287. Jayaram, B.; Jain, T. The role of water in protein-DNA recognition. *Annu. Rev. Biophys. Biomol. Struct.* 2004, 33, 343–361.

288. Kriegel, M.; Muller, Y.A. De novo prediction of explicit water molecule positions by a novel algorithm within the protein design software MUMBO. *Sci Rep* 2023, 13, 16680.

289. Mobley, D.L.; Dill, K.A. Binding of Small-Molecule Ligands to Proteins: “What You See” Is Not Always “What You Get”. *Structure* 2009, 17, 489–498.

290. Mammino, L. Computational chemistry: Studying the properties and behaviours of molecules. In Green Chemistry and Computational Chemistry – shared lessons in sustainability; Mammino, L. Ed.; Elsevier: 2021; pp. 1–40.

291. Ladbury, J.E. Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chem. Biol.* 1996, 3(12), 973–980.

292. Barillari, C.; Taylor, J.; Viner, R.; Essex, J.W. Classification of water molecules in protein binding sites. *J. Am. Chem. Soc.* 2007, 129, 2577–2587.

293. Mancera, R.L. Molecular modeling of hydration in drug design. *Curr. Opin. Drug Discov. Dev.* 2007, 10, 275–280.

294. Ross, G.A.; Morris, G.M.; Biggin, P.C. Rapid and accurate prediction and scoring of water molecules in protein binding sites. *PLoS ONE* 2012, 7, e32036.

295. Han, S.; Zaniewski, R.P.; Marr, E.S.; Lacey, B.M.; Tomaras, A.P.; Evdokimov, A.; Miller, J.R.; Shanmugasundaram, V. Structural Basis for Effectiveness of Siderophore-Conjugated Monocarbams against Clinically Relevant Strains of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A.* 2010, 107(51), 22002–22007.

296. Kimura, S.R.; Hu, H.P.; Ruvinsky, A.M.; Sherman, W.; Favia, A.D. Deciphering cryptic binding sites on proteins by mixed-solvent molecular dynamics. *J. Chem. Inf. Model.* 2017, 57(6), 1388–1401.

297. Arcon, J.P.; Defelipe, L.A.; Modenutti, C.P.; López, E.D.; Alvarez-Garcia, D.; Barril, X.; Turjanski, A.G.; Martí, M.A. Molecular dynamics in mixed solvents reveals protein-ligand interactions, improves docking, and allows accurate binding free energy predictions. *J. Chem. Inf. Model.* 2017, 57(4), 846–863.

298. Mattos, C.; Ringe, D. Locating and characterizing binding sites on proteins. *Nat. Biotechnol.* 1996, 14, 595–599.

299. Mattos, C.; Bellamacina, C.R.; Peisach, E.; Pereira, A.; Vitkup, D.; Petsko, G.A.; Ringe, D. Multiple solvent crystal structures: Probing binding sites, plasticity and hydration. *J. Mol. Biol.* 2006, 357, 1471–1482.

300. Ghanakota, P.; Carlson, H.A. Driving Structure-Based Drug Discovery through Cosolvent Molecular Dynamics. *J. Med. Chem.* 2016, 59, 10383–10399.

301. Ghanakota, P.; Carlson, H.A. Moving Beyond Active-Site Detection: MixMD Applied to Allosteric Systems. *J. Phys. Chem. B* 2016, 120, 8685–8695.

302. Setny, P.; Wang, Z.; Cheng, L.T.; Li, B.; McCammon, J.A.; Dzubiella, J. Dewetting-controlled binding of ligands to hydrophobic pockets. *Phys. Rev. Lett.* 2009, 103, 187801.

303. Bo'hm, H.-J.; Klebe, G. What can we learn from molecular recognition in protein–ligand complexes for the design of new drugs? *Angew. Chem. Int. Ed. Engl.* 1996, 35, 2588–2614.

304. Ahmad, M.; Gu, W.; Helms, V. Mechanism of fast peptide recognition by SH3 domains. *Angew. Chem. Int. Ed. Engl.* 2008, 47, 7626–7630.

305. Schmidtke, P.; Luque, F.J.; Murray, J.B.; Barril, X. Shielded hydrogen bonds as structural determinants of binding kinetics: application in drug design. *J. Am. Chem. Soc.* 2011, 133, 18903–18910.

306. Dror, R.O.; Pan, A.C.; Arlow, D.H.; Borhani, D.W.; Maragakis, P.; Shan, Y.; Xu, H.; Shaw, D.E. Pathway and mechanism of drug binding to G-protein-coupled receptors. In *Proceedings of the National Academy of Sciences*; 2011; 108(32), 13118–13123.

307. Setny, P.; Baron, R.; Michael Kekenes-Huskey, P.; McCammon, J.A.; Dzubiella, J. Solvent fluctuations in hydrophobic cavity–ligand binding kinetics. In *Proceedings of the National Academy of Sciences*; 2013; 110(4), 1197–1202.

308. Young, T.; Abel, R.; Kim, B.; Berne, B.J.; Friesner, R.A. Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *Proc. Natl. Acad. Sci.* 2007, 104(3), 808–813.

309. Fox, J.M.; Kang, K.; Sastry, M.; Sherman, W.; Sankaran, B.; Zwart, P.H.; Whitesides, G.M. Water-Restructuring Mutations Can Reverse the Thermodynamic Signature of Ligand Binding to Human Carbonic Anhydrase. *Angew. Chem. Int. Ed.* 2017, 56, 3833–3837.

310. Mohanty, M.; Mohanty, P.S. Molecular docking in organic, inorganic, and hybrid systems: a tutorial review. *Monatsh Chem* 2023, 154, 683–707. <https://doi.org/10.1007/s00706-023-03076-1>

311. Hu, X.; Maffucci, I.; Contini, A. Advances in the treatment of explicit water molecules in docking and binding free energy calculations. *Curr. Med. Chem.* 2018, 26(42), 7598–7622. doi: 10.2174/0929867325666180514110824

312. Pavlovic, R.E.; Park, H.; DiMaio, F. Efficient consideration of coordinated water molecules improves computational protein–protein and protein–ligand docking discrimination. *PLoS Comput Biol* 2020, 16, e1008103. <https://doi.org/10.1371/journal.pcbi.1008103> (2020).

313. Xiao, W.; Wang, D.; Shen, Z.; Li, S.; Li, H. Multi-Body Interactions in Molecular Docking Program Devised with Key Water Molecules in Protein Binding Sites. *Molecules* 2018, 23(9), 2321. <https://doi.org/10.3390/molecules23092321>

314. Cuzzolin, A.; Deganutti, G.; Salmaso, V.; Sturlese, M.; Moro, S. AquaMMapS: An Alternative Tool to Monitor the Role of Water Molecules During Protein–Ligand Association. *ChemMedChem* 2018, 13, 522–531.

315. Mitusińska, K.; Raczyńska, A.; Bzówka, M.; Bagrowska, W.; Góra, A. Applications of water molecules for analysis of macromolecule properties. *Comput. Struct. Biotechnol. J.* 2020, 18, 355–365. <https://doi.org/10.1016/j.csbj.2020.02.001> 2001-0370

316. Tuccinardi, T. What is the current value of MM/PBSA and MM/GBSA methods in drug discovery?. *Expert Opin. Drug Discov.* 2021, 16(11), 1233–1237.

317. Stanzione, F.; Giangreco, I.; Cole, J.C. Use of molecular docking computational tools in drug discovery. *Prog. Med. Chem.* 2021, 60, 273–343. <https://doi.org/10.1016/bs.pmch.2021.01.004>

318. Smith, S.T.; Shub, L.; Meiler, J. PlaceWaters: Real-time, explicit interface water sampling during Rosetta ligand docking. *PLoS ONE* 2022, 17(5), e0269072. <https://doi.org/10.1371/journal.pone.0269072>

319. Bálint, M.; Zsidó, B.Z.; van der Spoel, D.; Hetényi, C. Binding Networks Identify Targetable Protein Pockets for Mechanism-Based Drug Design. *Int. J. Mol. Sci.* 2022, 23, 7313. <https://doi.org/10.3390/ijms23137313>

320. Zsidó, B.Z.; Bayarsaikhan, B.; Börzsei, R.; Szél, V.; Mohos, V.; Hetényi, C. The Advances and Limitations of the Determination and Applications of Water Structure in Molecular Engineering. *Int. J. Mol. Sci.* 2023, 24, 11784. <https://doi.org/10.3390/ijms241411784>

321. Wolfenden, R. Conformational aspects of inhibitor design: enzyme-substrate interactions in the transition state. *Bioorg. Med. Chem.* 1999, 7(5), 647–652. [https://doi.org/10.1016/S0968-0896\(98\)00247-8](https://doi.org/10.1016/S0968-0896(98)00247-8)

322. Dunn, R.V.; Daniel, R.M. The use of gas-phase substrates to study enzyme catalysis at low hydration. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 2004, 359, 1309–1320.

323. Oleinikova, A.; Smolin, N.; Brovchenko, I.; Geiger, A.; Winter, R. Formation of spanning water networks on protein surfaces via 2D percolation transition. *J. Phys. Chem. B* 2005, 109, 1988–1998.

324. Chianella, I.; Karim, K.; Piletska, E.V.; Preston, C.; Piletsky, S.A. Computational design and synthesis of molecularly imprinted polymers with high binding capacity for pharmaceutical applications-model case: Adsorbent for abacavir. *Anal. Chim. Acta* 2006, 559(1), 73–78.

325. Teze, D.; Hendrickx, J.; Dion, M.; Tellier, C.; Woods, V.L.; Tran, V.; Sanejouand, Y.-H. Conserved water molecules in family 1 glycosidases: A DXMS and molecular dynamics study. *Biochemistry* 2013, 52, 5900–5910.

326. Dielmann-Gessner, J.; Grossman, M.; Conti Nibali, V.; Born, B.; Solomonov, I.; Fields, G.B.; Havenith, M.; Sagi, I. Enzymatic turnover of macromolecules generates long-lasting protein–water-coupled motions beyond reaction steady state. *Proc. Natl. Acad. Sci. USA* 2014, 111 (50), 17857–17862.

327. UNAIDS. Fact Sheet. Available online: <https://www.unaids.org/en/resources/fact-sheet> (accessed on 20 December 2023).

328. Ribeiro, A.J.M.; Ramos, M.J.; Fernandes, P.A. The catalytic mechanism of HIV-1 integrase for DNA 3'-end processing established by QM/MM calculations. *J. Am. Chem. Soc.* 2012, 134, 13436–13447.

329. Rungrotmongkol, T.; Mulholland, A.J.; Hannongbua, S. Active site dynamics and combined quantum mechanics/molecular mechanics (QM/MM) modelling of a HIV-1 reverse transcriptase/DNA/dTTP complex. *J. Mol. Graph. Model.* 2007, 26(1), 1–13.

330. Gopal, S. M.; Klumpers, F.; Herrmann, C.; Schäfer, L.V. Solvent effects on ligand binding to a serine protease. *Phys. Chem. Chem. Phys.* 2017, 19, 10753–10766.

331. Urbanowicz, B.R.; Bharadwaj, V.S.; Alahuhta, M.; Peña, M.J.; Lunin, V.V.; Bomble, Y.J.; Wang, S.; Yang, J.Y.; Tuomivaara, S.T.; Himmel, M.E.; Moremen, K.W. Structural, mutagenic and in silico studies of xyloglucan fucosylation in *Arabidopsis thaliana* suggest a water-mediated mechanism. *The Plant J.* 2017, 91, 931–949.

332. Xue, J.; Huang, X.; Zhu, Y. Using molecular dynamics simulations to evaluate active designs of cephradine hydrolase by molecular mechanics/Poisson–Boltzmann surface area and molecular mechanics/generalized Born surface area methods. *RSC Adv.* 2019, 9, 13868–13877.

333. van der Kamp, M.W.; Mulholland, A.J. Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology. *Biochemistry* 2013, 52, 2708–2728.

334. Magalhães, R.P.; Fernandes, H.S.; Sousa, S.F. Modelling enzymatic mechanisms with QM/MM approaches: current status and future challenges. *Isr. J. Chem.* 2020, 60, 655–666.

335. Jedrzejewski, M.; Belza, B.; Lewandowska, I.; Sadlej, M.; Perlinska, A.P.; Augustyniak, R.; Christian, T.; Hou, Y.M.; Kalek, M.; Sulkowska, J.I. Nucleolar Essential Protein 1 (Nep1): Elucidation of enzymatic catalysis mechanism by molecular dynamics simulation and quantum mechanics study. *Comp. Struct. Biotechnol. J.* 2023, 21, 3999–4008.

336. Gurova, K. New hopes from old drugs: revisiting DNA-binding small molecules as anticancer agents. *Future Oncol.* 2009, 5, 1685–1704. doi: 10.2217/fon.09.127.

337. Misra, V.K.; Honig, B. On the magnitude of the electrostatic contribution to ligand-DNA interactions. In *Proceedings of the National Academy of Sciences*; 1995; 92(10), 4691–4695.

338. Chaires, J.B. Energetics of drug-DNA interactions. *Biopolymers* 1997, 44, 201–215.

339. Harris, S.A.; Gavathiotis, E.; Searle, M.S.; Orozén, M.; Laughton, C.A. Cooperativity in drug-DNA recognition: molecular dynamics study. *J. Am. Chem. Soc.* 2001, 123, 12658–12663.

340. Sheng, J.; Gan, J.; Huang, Z. Structure-based DNA-targeting strategies with small molecule ligands for drug discovery. *Med Res Rev.* 2013, 33(5), 1119–73. doi: 10.1002/med.21278.

341. Krafcikova, M.; Dzatko, S.; Caron, C.; Granzhan, A.; Fiala, R.; Loja, T.; Teulade-Fichou, M.-P.; Fessl, T.; Hanse-Hertsch, R.; Mergny, J.-L.; Foldynova-Trantirkova, S.; Trantirek, L. Monitoring DNA–Ligand Interactions in Living Human Cells Using NMR Spectroscopy. *J. Am. Chem. Soc.* 2019, 141, 13281–13285.

342. Ricci, C. G.; Netz, P. A. Docking Studies on DNA-Ligand Interactions: Building and Application of a Protocol To Identify the Binding Mode. *J. Chem. Inf. Model.* 2009, 49(8), 1925–1935. doi: 10.1021/ci9001537.

343. Carter, E. K.; Laughlin-Toth, S.; Dodd, T.; Wilson, W. D.; Ivanov, I. Small molecule binders recognize DNA microstructural variations via an induced fit mechanism. *Phys Chem Chem Phys.* 2019, 21(4), 1841–1851. doi: 10.1039/c8cp05537h.

344. Schuurs, Z. P.; Martyn, A. P.; Soltau, C. P.; Beard, S.; Shah, E. T.; Adams, M. N.; Croft, L. V.; O’Byrne, K. J.; Richard, D. J.; Gandhi, N. S. An Exploration of Small Molecules That Bind Human Single-Stranded DNA Binding Protein 1. *Biology* 2023, 12(11), 1405. doi: 10.3390/biology12111405

345. Lang, P. T.; Brozell, S. R.; Mukherjee, S.; Pettersen, E. F.; Meng, E. C.; Thomas, V.; Rizzo, R. C.; Case, D. A.; James, T. L.; Kuntz, I. D. DOCK 6: combining techniques to model RNA-small molecule complexes. *RNA* 2009, 15(6), 1219–30. doi: 10.1261/rna.1563609.

346. Panei, F. P.; Gkekka, P.; Bonomi, M. Identifying small-molecules binding sites in RNA conformational ensembles with SHAMAN. *bioRxiv* preprint. 2023. doi: <https://doi.org/10.1101/2023.08.08.552403>.

347. Yusof, R.; Jumbri, K.; Ahmad, H.; Abdulmalek, E.; Abdul Rahman, M. B. Binding of tetrabutylammonium bromide based deep eutectic solvent to DNA by spectroscopic analysis. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 2021, 253, 119543. doi: 10.1016/j.saa.2021.119543.

348. Malola, S.; Matus, M. F.; Häkkinen, H. Theoretical Analysis of the Electronic Structure and Optical Properties of DNA-Stabilized Silver Cluster Ag<sub>16</sub>Cl<sub>2</sub> in Aqueous Solvent. *J. Phys. Chem. C* 2023, 127, 16553–16559.

349. Li, X.; Zhang, S.; Wu, Y.; Jiang, L.; Zhang, W.; Qiao, X.; Yan, H.; Zhou, H.; Tang, B. Removal of trace DNA toxic compounds using a Poly(deep eutectic solvent) @Biomass based on multi-physical interactions. *J. Hazard. Mater.* 2021, 418, 126369. doi: 10.1016/j.jhazmat.2021.126369.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.