



## The physics-AI dialogue in drug design†

Cite this: *RSC Med. Chem.*, 2025, **16**, 1499

Pablo Andrés Vargas-Rosales and Amedeo Caflisch \*

Received 7th November 2024,  
Accepted 16th January 2025

DOI: 10.1039/d4md00869c

[rsc.li/medchem](http://rsc.li/medchem)

A long path has led from the determination of the first protein structure in 1960 to the recent breakthroughs in protein science. Protein structure prediction and design methodologies based on machine learning (ML) have been recognized with the 2024 Nobel prize in Chemistry, but they would not have been possible without previous work and the input of many domain scientists. Challenges remain in the application of ML tools for the prediction of structural ensembles and their usage within the software pipelines for structure determination by crystallography or cryogenic electron microscopy. In the drug discovery workflow, ML techniques are being used in diverse areas such as scoring of docked poses, or the generation of molecular descriptors. As the ML techniques become more widespread, novel applications emerge which can profit from the large amounts of data available. Nevertheless, it is essential to balance the potential advantages against the environmental costs of ML deployment to decide if and when it is best to apply it. For hit to lead optimization ML tools can efficiently interpolate between compounds in large chemical series but free energy calculations by molecular dynamics simulations seem to be superior for designing novel derivatives. Importantly, the potential complementarity and/or synergism of physics-based methods (e.g., force field-based simulation models) and data-hungry ML techniques is growing strongly. Current ML methods have evolved from decades of research. It is now necessary for biologists, physicists, and computer scientists to fully understand advantages and limitations of ML techniques to ensure that the complementarity of physics-based methods and ML tools can be fully exploited for drug design.

## 1. Introduction

### 1.1 The path to protein structure prediction

More than 60 years ago, the first protein structures were determined experimentally. The three-dimensional conformations of myoglobin and hemoglobin were described by scientists at Cambridge and appeared published in 1960.<sup>1,2</sup> First hand accounts of this momentous event help us understand the difficulty and work that went into these discoveries which today are routine work.<sup>3–5</sup> This set the course for the beginning of the structural biology revolution and protein structure-based drug design (Fig. 1). Nevertheless, most protein structures remained unknown and biochemical analyses were the main method to obtain information about protein function and behavior. In 1961, Anfinsen *et al.*, showed that a ribonuclease could be reversibly denatured, and regain function after renaturing.<sup>6</sup> Levinthal *et al.*, proceeded in similar way using alkaline phosphatase, from *Escherichia coli* and *Serratia marescens*, and not only found they could obtain active enzymes

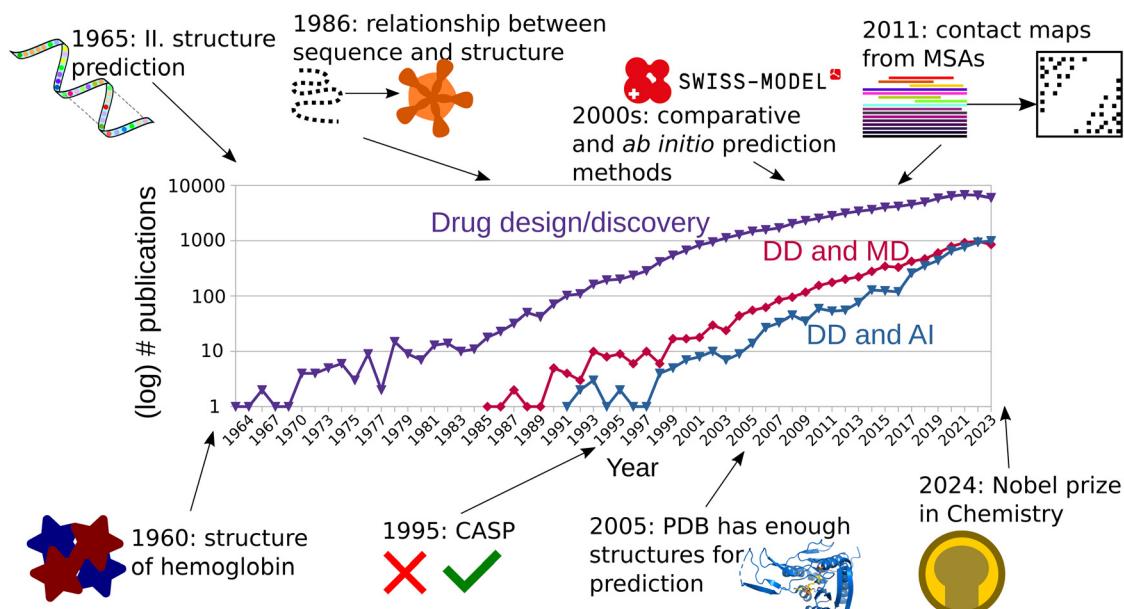
after renaturing them, but also that the interspecific dimer of the two was active as well. They therefore theorized that both must share a conserved active site, and a configuration which allows for active heterodimers.<sup>7</sup> Thanks to the advances in the availability of protein sequence information, Perutz *et al.* proposed in 1965 that despite poor sequence conservation, the structure of globins was similar across all vertebrates.<sup>8</sup> The foundational advances of the first half of the 1960 decade enabled Guzzo to postulate in 1965 that there was enough evidence that in proteins, “sequence implies structure”, and a thermodynamically most stable form must be the native and active one.<sup>9</sup> These were the first building blocks for the successful protein structure prediction methods of today.

Already in the 1960s scientists tried to predict the structure of proteins from their sequence. In his paper, Guzzo did not only task himself with the understanding of protein function and its relationship to structure, but mostly with the prediction of protein structure.<sup>9</sup> Due to the complexity of the task, he focused instead on predicting the secondary structure of proteins, with the hope that solving a smaller part of the problem might prove easier than predicting the whole tertiary structure, while still giving valuable insight into the fold of the protein. He analyzed the sequences and structures of myoglobin and the  $\alpha$ - and  $\beta$ -hemoglobin to predict that  $\alpha$  helices are disrupted by “critical residues”: P, D, E. and H. He later applied this

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057, Zürich, Switzerland. E-mail: [caflisch@bioc.uzh.ch](mailto:caflisch@bioc.uzh.ch)

† This review is dedicated to the memory of Martin Karplus who has been a pioneer of physics-based methods for simulating the dynamics of biological macromolecules.





**Fig. 1** The number of publications mentioning drug design or drug discovery (DD) has continually increased since the 1960s (violet curve). The discovery of the structure of hemoglobin in 1960 opened the door to structure-based drug design. In the 1990s, the number of DD publications that mentioned MD simulations (red curve) or AI tools (blue curve) started to grow steadily. Since the late 2010s, the gap of DD publications based on AI methods versus MD simulations has narrowed.

prediction to TMV capsid and lysozyme.<sup>9</sup> Such prediction was later expanded by Prothero, who gave more complex and complete rules on the influence of residues to secondary structure.<sup>10</sup> The turn of the decade brought further research, incorporating more data, and more advanced analyses. Pain and Robson proposed a new approach in 1970, where pairs of residues were screened to understand the “helix-forming power” of each residue. They had the advantage of having more structures available.<sup>11</sup> It was not until 1973 that Nagano proposed a statistical analysis based on 95 available proteins to predict helices, but also loops and  $\beta$ -strands. They did not only focus on pairs of adjacent residues, but also recognized the long-range influence of other residues away from the position considered.<sup>12</sup> These years saw many more attempts at the prediction of protein secondary structure, with varying levels of success.<sup>13,14</sup> Ten more years would still pass, until enough structures were available to generate a detailed and unambiguous definition of secondary structure.<sup>15</sup>

Since protein sequence encodes structure,<sup>9,16</sup> it was theorized that the sequences of an entire protein family may also contain information about its tertiary structure.<sup>17</sup> Pazos *et al.*<sup>18</sup> showed in 1997 that the evolution of a protein must be, in some way, constrained by the sequence of that of its interaction partners, and these correlated mutations could be discovered in multiple sequence alignments (MSA). In the end they used this information to predict the interfaces between interacting proteins.<sup>18</sup> Afterwards, Fariselli *et al.* used neural networks to predict the contact maps from a database of 173 proteins, with at least 15 sequences in each MSA. They obtained a relatively low accuracy, albeit the best for its time.<sup>19</sup> Hopf *et al.* continued this line of work, using

the higher number of sequences available 20 years later.<sup>20,21</sup> In 2011, Morcos *et al.* used direct-coupling analysis to predict contact maps from MSAs.<sup>22</sup> Coevolution methods continued to ripen in the decade of 2010 with new methodologies and proposed applications at the level of structures and interfaces, but also others such as binding site prediction.<sup>23,24</sup>

## 1.2 Comparative and *ab initio* modelling

A huge step in the protein prediction history can be traced to the introduction of the critical assessment of methods of protein structure prediction (CASP)<sup>25</sup> in 1995. The standardized experiments, occurring every two years, started yielding information about the main bottlenecks in protein prediction.<sup>26</sup> One category of such prediction methods was the *ab initio* method, where protein sequence is the main input for the prediction. *Ab initio* modelling usually follows physical principles, using techniques such as Monte Carlo sampling,<sup>27,28</sup> threading (fold recognition),<sup>29</sup> fragment based prediction,<sup>30</sup> or stepwise secondary structure and then fold prediction.<sup>31,32</sup> ROSETTA started as an *ab initio* prediction software based on the assembly of small fragments. It managed root mean square deviations of around 6 to 4 Å with respect to the native structure in CASP3.<sup>33</sup> Distance based prediction methods also showed promise, when scientists realized that a subset of native inter-C $\alpha$  distances could be used as additional restraints to generate native-like conformations using *ab initio* methods.<sup>34,35</sup>

Furthermore, the genomic explosion of the early 2000s generated a huge gap between the number of sequenced



genes and their solved structures. For all the proteins for which their genetic sequences were known in 2004, only around 1% had their structure experimentally determined.<sup>36</sup> Nevertheless, already since the 1980s, and thanks to previous observations such as those by Levinthal *et al.*,<sup>7</sup> it had been proposed that there exists a strong relationship between sequence identity and fold conservation.<sup>37</sup> This is why homology (or comparative) modelling also surged as an important category of protein structure prediction for proteins with known sequence but yet undetermined structure. Homology modelling usually consists of the following steps: identification of a template based on evolutionary closeness, alignment of target and template, modelling of conserved regions, modelling of divergent regions, assignment of sidechain rotamers, and refinement.<sup>38</sup> For targets with strong evolutionary relationship to the templates, model building was usually simple except for nonconserved regions and loops, while more distant evolutionary relationships required more advanced alignment methods for finding templates.<sup>39</sup> Early homology methods were able to correctly align models to templates with sequence similarities above 50%, while quality deteriorated for sequence identities lower than 35%.<sup>40</sup> Despite their poor performance for distantly related proteins, comparative methods were pioneering in the use of evolutionary data at large scale. These homology modelling programs incorporated a combination of MSAs and structural methods to improve the alignment methods.<sup>41</sup> One of the pioneering automated servers for homology modelling is SWISS-MODEL,<sup>42</sup> which is still in use today.<sup>43</sup> The current engine of SWISS-MODEL, ProMod3, implements homology modelling, as well as the loop and sidechain reconstruction, and inclusion of ligands in the binding pocket.<sup>44</sup> A stand-alone program for comparative modelling is MODELLER, which implements all the steps of homology modelling described above. Unlike SWISS-MODEL, MODELLER is based on a probabilistic description of spatial restraints which guide the structural prediction.<sup>45</sup> The alignment of templates is used to generate a probabilistic density function on which the template is aligned.<sup>46</sup> The modelling of loops is achieved through a combination of structural restraints, the probabilistic information from the alignment, and force field information.<sup>47</sup> This serves as an example that, as the years passed, the division between *ab initio* and comparative (data-driven) methods started to dilute.<sup>48</sup> Another example of the fusion of different modelling methods comes from TASSER, where threading was used for template identification and followed by refinement.<sup>49</sup> The increase in protein structures deposited in the PDB led Zhang and Skolnick to declare in 2005 that the protein folding problem could be solved based on the entries available at the time, given efficient fold recognition algorithms that could be used to assign templates to the sequence being predicted.<sup>50</sup> The following years continued with incremental advances being reported for *ab initio* and comparative methods at the successive CASPs, albeit at a more modest pace than before,<sup>51–53</sup> until the introduction of the first AlphaFold model at CASP13.<sup>54,55</sup>

Today, artificial intelligence (AI) and deep learning (DL) based methods for protein structure prediction and design are widely regarded as a revolution in life sciences, to the point that we talk of structural biology in terms of the times “before and after AlphaFold2”.<sup>56</sup> The main developers of AlphaFold, John Jumper and Demis Hassabis, have shared the 2024 Nobel award in Chemistry with David Baker who has pioneered computational protein design (Fig. 1).

## 2. AI applications in structure prediction

### 2.1 Deep learning-based prediction

Today, after the explosion of machine learning (ML) in the 2010s, computer and biological scientists alike have worked to transfer the advances of AI, creating data-driven methods for protein structure prediction with a high level of success.<sup>57</sup> The dream of scientists from the early 2000s of achieving proteome-scale prediction, uncovering new protein folds, and helping predict new functions,<sup>58,59</sup> was finally achieved in 2023 (ref. 60) thanks to the establishment of the AlphaFold Protein Structure Database.<sup>61</sup> It has been claimed that neural networks seem to have “largely solved” the protein folding problem at the domain level,<sup>62</sup> although as Bowman correctly points, fields are advanced, not solved.<sup>63</sup> Such claims can derive from a lack of understanding of the background of AlphaFold, and the work behind it which explains how it achieved such remarkable accuracy at the prediction of tertiary structure. Recapitulating some points discussed above, the biochemical studies of 1960s elucidated the structural similarity of proteins related by evolution. This was exploited by scientists who tried to predict contacts based on evolutionary information, the ones who tried to model novel structures based on already available structures of homologous proteins, or the ones who tried to recognize folds from primary sequence. In parallel, decades of experimental structure determination yielded a database of globular proteins which is complete enough for an advanced data mining strategy to take advantage of it. The success of AlphaFold2 can then be understood as the result of an excellent fold recognition procedure, which exploits the completeness of the library of single-domain proteins in the PDB.<sup>64</sup> Therefore this revolution did not occur in isolation, half a century of research paved the way for the AI-based methods of protein structure prediction and design.

A common feature of all the AlphaFold models,<sup>55,65,66</sup> and indeed of other related methods such as RosettaFold,<sup>67,68</sup> is the use of MSAs to find evolutionary relationships that can be used to predict inter-residue contacts.<sup>22</sup> The first AlphaFold model used the MSAs to bias a statistical potential of inter-residue interactions<sup>69</sup> that better satisfies these contacts.<sup>55</sup> AlphaFold2 uses the transformer architecture<sup>70</sup> to integrate the information from the MSA and structural templates together.<sup>65</sup> The MSA is so important to AlphaFold2, that it has been pointed out that AlphaFold2 has learnt a MSA-structure relationship, not a sequence-structure



relationship as claimed.<sup>71</sup> AlphaFold3 still uses the MSA to find information on close-by residue pairs, but the MSA is not used as input to the network directly.<sup>66</sup> The MSA is completely ignored in ESMFold, a language model which captures evolutionary relatedness by learning dependencies between aminoacids at the sequence level, and then using this information to predict the contact map.<sup>72</sup> The architecture of the ML models also reflects the directions of research in the deep learning community in general. While the transformer is present in the AlphaFold2 and ESMFold structure modules,<sup>73</sup> AlphaFold3 and RosettFoldAllAtom use diffusion models to generate the final structures.<sup>66,68</sup> Diffusion models are increasingly being used in structural biology and drug discovery,<sup>74</sup> finding applications in protein design,<sup>75,76</sup> conformer generation,<sup>77</sup> and small molecule binder design.<sup>78</sup>

## 2.2 The challenges of DL-based methods

Despite the claims of protein structure prediction being “solved”,<sup>63</sup> several challenges remain. Even some globular proteins can be predicted incorrectly,<sup>79</sup> and extreme care must be taken when using AI structure prediction tools for disordered proteins.<sup>80</sup> For example, AlphaFold2 overestimates the confidence of structure predictions in thousands of intrinsically disordered regions that fold upon binding or modification.<sup>81</sup> Several strategies have been implemented to make “safer” predictions of these disordered regions. Bret *et al.* used fragments of disordered regions and different MSA schemes to predict interfaces of interacting disordered proteins.<sup>82</sup> In recent work, we combined AlphaFold predictions of amyloid  $\beta$  dimers with molecular dynamics (MD) simulations to validate the predicted structures.<sup>83</sup>

Another important challenge in protein structure prediction is the generation of conformational ensembles. The prediction of single structures has been pointed as the current main limitation of these models.<sup>84</sup> The generation of multiple structural models has been achieved in AlphaFold2 for example by activating dropout layers during prediction,<sup>85</sup> subsampling the MSAs,<sup>86,87</sup> MSA subsampling in combination with enhanced-sampling MD,<sup>88</sup> *in silico* mutagenesis of the MSA,<sup>89</sup> flow matching,<sup>90</sup> and others.<sup>91</sup> Novel diffusion models are also emerging which are able to generate conformational ensembles, even of novel proteins.<sup>92</sup> Finally, Cfold is an implementation of an architecture similar to AlphaFold, which was specifically trained with different conformations of the same sequence.<sup>93</sup>

A deeper understanding of the physics behind these models is also crucial to make the most of AI-powered protein structure prediction. Outeiral *et al.* found that the AI-based models are not appropriate tools to investigate folding, as the folding pathways they produce are inconsistent with experimental data.<sup>94</sup> For fold-switching proteins it was found that AlphaFold2 assumes a “most-probable” fold while missing the other. Additionally, the chosen fold is predicted with an overestimated confidence due to the high conservation of these proteins.<sup>95</sup> Later research theorized that prevalence of a single

conformation is due to a memorization of the structures in its training set and not due to learning of a biophysical energy function. This renders the models unable to predict alternate conformations even in the presence of their binding partners.<sup>96</sup> Indeed, it was shown that the performance of AlphaFold suffers with proteins that adopt diverse conformations.<sup>97</sup> Such evidence is in clear contrast with previous claims that AlphaFold has learned an approximate biophysical energy function.<sup>98,99</sup> It has been proposed that while AlphaFold and related methods learn the contacts between residues at the minimum of the free energy funnel characteristic of globular proteins, the shallow or multi-funnelled landscapes of disordered and fold switching proteins counter this principle.<sup>100</sup> A recent study in which perturbations were introduced to binding sites showed that AlphaFold3 does not predict binding based on molecular interactions, but based on general protein patterns. Thus non-physical predictions are possible because of overfitting to specific subsets of structural data.<sup>101</sup> It is important to point out that the study was limited by the fact that AlphaFold3 was only available as a web server with limited capability for small molecule prediction. The fully open access to these deep learning models is not only essential to use them efficiently, but also to find new ways to improve them.

## 3. AI and the physics-based methods

### 3.1 Small molecule docking

Two physics-based methods essential to protein structure-based ligand design are MD simulations,<sup>102–104</sup> and small molecule docking.<sup>105–110</sup> Docking relies on scoring functions to describe protein–ligand interactions. The scoring functions have been classified into three main categories: force field-based, empirical, and knowledge-based.<sup>111</sup> Force fields are analytical functions that make use of classical physics approximations of the potential energy of (macro)molecules which is calculated by the sum of bonding and non-bonding (van der Waals and Coulomb energy) contributions.<sup>112</sup> The bonding interactions are calculated for pairs of atoms separated by one, two, or three covalent bonds, *e.g.*, Hooke’s law is employed for the covalent bonds which does not allow the rupture of bonds or formation of new ones. The parameters of the force field are derived either from quantum mechanical calculations (*e.g.*, the partial charges for the Coulomb term) or by fitting to experimental data.<sup>113,114</sup> The force fields might also include some desolvation terms, usually based on an implicit representation of solvent effects.<sup>115–117</sup> For docking large libraries of compounds, the binding free energy is usually approximated by the difference between the energy of the protein/ligand complex and the energy of the unbound protein and ligand. Most frequently, the flexibility of the protein is ignored and entropic effects are neglected or approximated coarsely. Force field-based energy functions include those available in the docking programs SEED<sup>118</sup> and AutoDock Vina.<sup>119</sup> An additional sub-category, which is related to the force field-family, is the use of quantum mechanical descriptors for scoring. One example is the use of quantum mechanical “probes” which approximate a subset of



the polar groups in the binding pocket of the target protein.<sup>120</sup> Zhou and one of us screened a large library of compounds by the interaction energy with the probes calculated at a semi-empirical level of theory. In this way a novel and selective low micromolar inhibitor of the EphB4 tyrosine kinase was identified from a large library of compounds.<sup>120</sup> Quantum mechanics-based scoring methods are less approximated than classical force fields, but are computationally more expensive.<sup>121-123</sup>

Unlike force field-based scoring functions, empirical scoring functions approximate the binding affinity directly. Analogous to force fields, they contain individual interaction descriptors of binding, trained using a regression model to fit the descriptors to the experimental binding affinity. Such descriptors can include intermolecular interactions like van der Waals and Coulomb terms, electrostatic desolvation penalty, ligand entropy and torsion, *etc.*<sup>124</sup> Glide<sup>125</sup> and ChemScore<sup>126</sup> are examples of empirical scoring functions. Empirical scoring functions can also be employed for positioning small molecules in electron density maps determined by cryogenic electron microscopy (CryoEM).<sup>127</sup>

Knowledge-based scoring functions calculate the frequency of occurrence of the diverse atom pairs in a database from which, using the inverse Boltzmann relation, they obtain an approximation of the potential of mean force.<sup>69,116</sup> An example of this type of functions is DrugScore.<sup>128</sup> Interestingly, the concept of predicting and minimizing the potential of pairwise interactions was used as the basis of the first AlphaFold model,<sup>55</sup> while the idea itself was already published in 1990.<sup>129</sup> Due to the rapid changes and the different hybrid forms, Liu and Wang proposed in 2015 a new classification of scoring functions: physics-based methods (force field and quantum mechanics), empirical scoring functions, knowledge-based potentials, and descriptor-based scoring functions (such as those derived from ML).<sup>130</sup>

ML methods are becoming frequent in diverse aspects of scoring functions, and have performed well even when using simple methods such as a random forest.<sup>131</sup> Two examples of ML-based scoring functions are PointVS<sup>132</sup> and GNina.<sup>133</sup> Guedes *et al.* have used linear and nonlinear ML methods to fit the coefficients of the physics-based terms of DockTScore, an empirical scoring function.<sup>134</sup> Fujimoto *et al.* used molecular fingerprints of the protein-ligand interactions to build a regression model to approximate the potential mean force.<sup>135</sup> As mentioned above, AlphaFold also incorporates concepts from knowledge-based scoring to predict protein structure.<sup>55</sup> While not a knowledge-based model *per se*, Isert *et al.* used deep learning and quantum mechanics hand in hand for predicting protein-ligand binding affinity from CryoEM maps, giving strong emphasis to the study of interatomic interactions.<sup>136</sup> Indeed with the deep learning explosion, came new and more data-hungry methods which unfortunately do not necessarily perform better than simpler “traditional” ML methods.<sup>137,138</sup> Some of these deep learning methods were actually found to be even worse at generalizing than traditional docking methods.<sup>139</sup> An important factor to

keep in mind is that many ML scoring functions are applied at a postprocessing stage, with only select ones (such as GNina) being integrated into docking workflows.<sup>140</sup> The use of different paradigms for sampling poses of the ligand and scoring them is not optimal. As an example, a force field-based sampling engine might not reach protein/ligand structures close to poses with optimal ML-based scores. The co-folding of proteins with their binding partners, for example as proposed by AlphaFold3, uses ML for both posing and scoring but can be affected by overfitting.<sup>101</sup>

### 3.2 Simulating the motion of atoms

Simulations are another important technique in drug design. They yield insights on the time-resolved behavior of biomolecules on an atomic scale. There are several types of atomistic simulations, such as Monte Carlo,<sup>141</sup> MD,<sup>142</sup> and quantum mechanics calculations.<sup>143</sup> Monte Carlo simulations make use of random perturbations for iteratively evolving a molecular system. They can sample a thermodynamic ensemble but usually do not preserve the kinetic properties. In contrast, MD simulations are based on the classical Newtonian equations of motion (solved numerically) and thus not only reproduce a thermodynamic ensemble but also correctly reproduce the kinetics.<sup>142</sup> Quantum mechanical simulations solve a system's electronic structure. This means they are very accurate and can describe processes such as chemical reactions, but are too slow to be applied to whole systems.<sup>144</sup> Therefore, they are usually employed in combination with MD as multiscale simulations.<sup>145</sup> MD has been used since long to obtain thermodynamics and kinetics of small molecule binding to proteins, validate predicted binding modes, to generate conformations for docking, identification of cryptic pockets, or (relative) binding free energy calculations.<sup>146,147</sup> Simulations have been used to study the folding pathway of the cellular prion protein, from which druggable pockets were identified and targeted using small molecules to arrest folding.<sup>148</sup> In another translational study, umbrella sampling MD simulations were successfully employed to predict the relative binding free energy of a series of anti-prion compounds which were then validated *in vivo*.<sup>149</sup> Many simulation studies have been launched by different groups to analyze the self-assembly process of amyloid (poly) peptides.<sup>150</sup> MD has also been used to try to open new avenues of treatment for amyloid diseases, by subjecting either small amyloidogenic fragments,<sup>151</sup> or dimers of  $\text{A}\beta42$  (ref. 83) to external electric fields. In our group, we have used MD and quantum mechanics (semi-empirical level) simulations to propose a catalytic mechanism for the human methyltransferase METTL3.<sup>152</sup> We have also used MD to find structural information about binders for which no bound structure could be determined experimentally.<sup>153</sup>

MD simulations offer a means to evaluate the interaction free energy between a small-molecule ligand and its protein target, and rank ligands by relative affinity.<sup>154</sup> In contrast to docking, MD simulations can take into account the full



flexibility of the protein target, ligand, and surrounding solvent. They make use of the full force field, *i.e.*, including the bonded terms which is essential for reproducing the strain in the ligand upon binding. Free energy methods usually rely on a thermodynamic cycle to calculate the free energy differences between the states of interest. Since the direct calculation of the transformation of interest is usually difficult to obtain, a series of transformations is constructed that yields the same energy difference through simpler calculations.<sup>155</sup> Two alchemical transformation protocols which can be used for free energy calculations are thermodynamic integration and free energy perturbation.<sup>156–158</sup> Thermodynamic integration calculates the free energy difference between two states by numerically integrating the thermodynamic path between them. The path corresponds to an interpolation between the two end states' Hamiltonians, and is controlled by a coupling parameter.<sup>159</sup> Free energy perturbation<sup>160</sup> is based on the conversion of a molecule to another passing through unphysical intermediates of the two molecules.<sup>161</sup> Instead of direct integration, the differences between small steps is used.<sup>155</sup> If the systems are carefully prepared, which is time consuming and requires an in-depth knowledge of simulation protocols, free energy perturbation calculations yield an accuracy almost comparable to experimental measurement errors for relative binding free energy determination.<sup>162,163</sup> Some challenges faced by these methods are the accuracy of the force fields used (which results in a systematic error) and the convergence of sampling (statistical error).<sup>154,155</sup> Thermodynamic integration and free energy perturbation can be used for calculating both relative and absolute binding free energies.<sup>164</sup> They have diverse applications in drug design, such as derivatization of ligands, scaffold-hopping, and binding pose validation.<sup>165</sup> It is also possible to perform binding free energy calculations by Monte Carlo simulations in implicit solvent, using a thermodynamic cycle between the complex and the free protein and ligand. In a recent study, Monte Carlo sampling in implicit solvent with explicit ions as competitors, and the integration over multiple protonation states of protein and ligand, were assessed as a tool for virtual screening, and for the ranking of derivatives of hits obtained by docking.<sup>166</sup>

Several challenges remain in biomolecular modelling.<sup>167</sup> One challenge is improving the physical models behind the simulations such as adjusting force fields to better represent disordered proteins<sup>114</sup> or nucleic acids,<sup>168</sup> in particular RNA.<sup>169,170</sup> A second challenge of simulations is the timescales that can be reached. Even with recent advances in computing hardware, all-atom simulations remain prohibitive beyond the microsecond timescale. One possible solution is to leverage multiscale simulations to explore larger conformational spaces.<sup>145,171,172</sup> Another interesting simulation protocol is enhanced sampling,<sup>173,174</sup> for example by using swarms of trajectories to rebuild a reaction coordinate,<sup>113,175,176</sup> or by using a(n approximate) reaction coordinate to reseed trajectories in a diverse manner.<sup>177</sup> A good source of enhanced sampling protocols is PLUMED. The PLUMED library is a modular, open-source initiative which provides algorithms for enhanced

sampling MD, free energy methods, and analysis tools. Finally, steps are also being taken to optimize the different simulation packages to take advantage of current hardware architectures such as GPUs.<sup>178,179</sup> A third challenge of biomolecular simulation is the integration of experimental data into the simulations.<sup>167</sup> These integrative approaches incorporate data from different sources to understand biomolecules.<sup>180</sup> Experimental data such as NMR or CryoEM has been used in conjunction with MD to understand RNA conformational diversity and dynamics.<sup>181</sup> In the case of enhanced sampling protein simulations, it is very important to validate the obtained data against experimental data, due to the bias introduced.<sup>182</sup> An example of an integrative modelling approach is metainference,<sup>183</sup> which allows the construction of an ensemble of models consistent with experimental data by introducing the measurements as part of the energy function of the system.<sup>184</sup>

ML is also entering the world of biomolecular simulations. In the area of enhanced sampling, ML techniques have been applied to calculate reaction coordinates or collective variables for biased sampling.<sup>185</sup> Designing or learning these reaction coordinates is difficult, and the simulation must be biased to sample the Boltzmann distribution appropriately. A more efficient approach would be to sample directly from the Boltzmann distribution to obtain the different conformations of the system. This is the idea behind Boltzmann generators. Boltzmann generators use neural networks to learn a transformation from a normal to a Boltzmann distribution, such that sampling from the normal can be used to generate many independent Boltzmann-distributed samples. Unlike enhanced sampling, they are not dependent on trajectory-based methods such as long simulations to obtain the samples.<sup>186</sup> Another way to convert between distributions is flow matching, which has been used together with AlphaFold and ESMFold to predict protein ensembles. Similar to Boltzmann generators, flow matching uses a generative neural network to approximately transform a prior distribution to a Boltzmann one. Jing *et al.* proposes then to change AlphaFold from a regression model into a generative one, by feeding it with the noisy conformation generated by sampling from the prior and converting it to an approximately Boltzmann-distributed conformation. AlphaFold then “denoises” these generated coordinates and produces a high quality model based on this sample.<sup>90</sup> In the force field development area ML applications have been extensively reviewed by Unke *et al.* and Chen *et al.*<sup>187,188</sup> An example of a deep learning force field is shown by Majewski *et al.* They used MD data to construct an ML coarse-grained force field to recreate protein dynamics.<sup>189</sup> Also, tools such as TorchMD enable researchers to run simulations using both classical force fields and ML potentials.<sup>190</sup> A different application of deep learning are convolutional neural networks for reintroducing atomic detail into coarse grained models.<sup>191</sup> Another approach employed a generative adversarial network to solve the backmapping problem, using as an analogy the image-to-image problem of going from a low to a high resolution image.<sup>192</sup> Flow matching has also been used to describe coarse-grained force fields which match all-atom



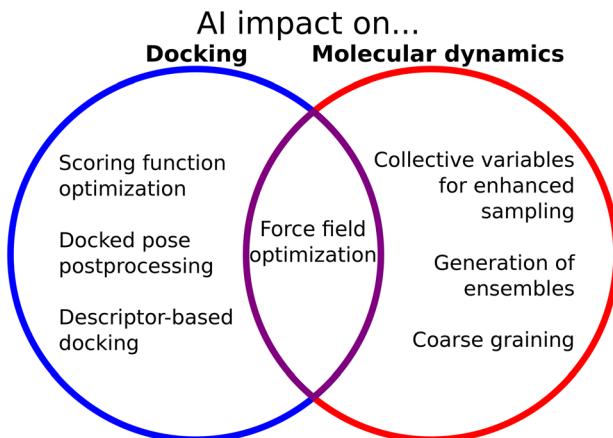


Fig. 2 AI influence on docking and MD simulations models and methods.

ones.<sup>193</sup> Although progress has been made, these ML force fields are not yet considered mature enough to be used in production simulations and are mostly applied to small molecules or single elements.<sup>194</sup> In the future, optimization of these force fields is likely to better approximate the interactions between (candidate) drugs and their targets resulting in improved accuracy in virtual screening and MD simulations, but further research is needed.<sup>188</sup>

The interface between biophysical knowledge and ML methods is of utmost importance to expand the capabilities of these models and understand their limitations (Fig. 2). Domain scientists have used and expanded the ML methods, to a great extent. In contrast to CASP14, where AlphaFold2 had a clear dominance, during CASP15 many groups incorporated the ideas from AlphaFold into their pipelines, and the difference in performance was less pronounced.<sup>195</sup> Structural biologists have also improved ML prediction of proteins by associating different depths of the MSAs with different folds, for example for fold switchers,<sup>96</sup> disordered proteins,<sup>80,100</sup> or proteins with different conformations.<sup>97</sup> Also, knowledge of the biological and biophysical behavior of macromolecules has led scientists to propose ways to incorporate ML into physical methods such as docking or simulations. Janel and Bajorath in particular call for an integration of computational studies into well-planned experimental evaluations to assess the predictive capacity of the different ML methods which are being increasingly proposed.<sup>138</sup> This is evidence that the interplay between the “hard” ML computer science and the domain application is necessary to find good applications and solve the shortcomings of the original models. It also highlights the importance of the open-source code, which allows scientists to build upon previous work to improve it or find new instances to use it.

## 4. AI applications in drug discovery

An important question is if the recent computational advances can help in hit discovery and/or lead optimization. This question has been asked 20 years ago by Hillisch *et al.*

while assessing possible applications of homology modelling in the drug discovery process. Some applications proposed back then included the prediction of binding pockets in homology models of clinically-relevant target proteins, site-directed mutagenesis to (de)sensitize a target to a compound, design of ligands based on the homologous modelled structures, prediction of drug metabolism and toxicity, *etc.*<sup>36</sup> Modern drug discovery is profiting from AI in several ways, such as studies of structure–activity relationship and data integration.<sup>196</sup> In contrast, generative drug design does not seem (yet) to be of genuine utility for designing novel molecules in medicinal chemistry campaigns. Thus, for discriminating between hype and real utility it is essential to follow the guidelines formulated by Walters and Murcko for publications of results of generative modeling.<sup>197</sup> Despite the lack of novel molecules, we hope to see new applications of generative AI in medicinal chemistry. An example is its recent application to scaffold hopping.<sup>198,199</sup>

### 4.1 Describing and quantifying molecules and their interactions

Quantitative structure–activity relationship (QSAR) studies are a natural subject for deep learning integration. Traditionally, linear equations were used to correlate the functional groups and compound properties with the activity observed, but deep learning has also been increasingly used to find relationships between properties and activity.<sup>200</sup> Unlike “traditional” QSAR, deep QSAR learns the embeddings of the molecules directly, and can also be pretrained with large unlabelled datasets.<sup>201</sup> An example is provided by Li and Fourches, who trained a general domain model using the ChEMBL database, which then is fine-tuned with target-specific experimental data, to finalize by transferring these pre-trained model weights to a final QSAR network which is used for the final predictions.<sup>202</sup> Also related to QSAR is the generation of novel molecules using ML. Initially, *de novo* generation was done by fragment based approaches or evolutionary algorithms.<sup>203,204</sup> Today, fragment based methods are still in use, having the advantage that the synthetic accessibility can be easily predicted when using rule-based fragment joining.<sup>205,206</sup> Other generative AI methods include variational autoencoders, generative adversarial networks, flow-based methods, transformer models, diffusion-based, and others, working on different molecular representations such as SMILES or graphs.<sup>207</sup> An interesting example is provided by Munson *et al.*, who use a variational autoencoder to target two proteins at the same time. They present the problem in an analogous way as networks trained to generate images along different variables, such as age or mood. This allowed them to target pairs of proteins which are together relevant to disease.<sup>208</sup> Schneider and Clark described several compounds which have been designed *de novo*, albeit sometimes as part of a longer design process usually involving some level of expert input.<sup>209</sup> To ensure the novelty and to be able to better understand the design process of the AI-generated compounds, Walters and Murcko have called for transparent reporting of the datasets



used for training, showing the most similar molecule in the training set, and evaluating the molecules with the same criteria as those generated by medicinal chemists.<sup>197</sup> Nonetheless, new AI-generated compounds have already entered clinical trials as treatments against diverse illnesses including atopic dermatitis, neurofibromatosis type 2, and others.<sup>210</sup>

The way in which small molecules are digitally represented is essential for QSAR studies, and also cheminformatics. Initially, “traditional”, or “bespoke” descriptors, as called by McGibbon *et al.*, were the main types of molecular representations.<sup>211</sup> The advent of ML in cheminformatics means that new representations can now be “learned” from data. The input representation and the type of ML method used for embedding the represented molecule determine the type of encoding needed.<sup>212</sup> McGibbon *et al.* describe three main types of learned representations. First, convolutional encodings have a high tolerance for many different inputs, but their main limitation is the lack of rotation-invariance. Second, graph encodings represent molecules and their features as a graph, and can be used by a variety of neural network architectures. Finally, the string encodings are traditionally used with transformer architectures.<sup>211</sup> An example of a string-based representation designed specifically for ML-based methods is SELFIES.<sup>213</sup>

Molecular descriptors are representations which encode the physicochemical information of the molecule. They can be derived from experimental data, such as the solubility or the octanol/water partition coefficient, or be theoretically defined. Theoretical descriptors can also vary in the level of abstraction, ranging from adimensional descriptors such as molecular weight and heavy atom number, up to four-dimensional descriptors encoding the interactions with binding partners.<sup>214</sup> Fingerprints are a type of representation based on encoding descriptors into a vector.<sup>215</sup> The importance of accurate descriptors was highlighted by van Tilborg *et al.*, who showed that SAR predictors based on molecular descriptors outperformed deep learning models based on SMILES or graphs.<sup>216</sup> Therefore, the use of learned descriptors should be carefully considered, for example Capecchi *et al.* describe a molecular fingerprint based on substructures which performs well, without the need of a ML-based encoding.<sup>217</sup> It is important to note that in general the prediction of bio-activity data (e.g., binding potency for the target) is a more challenging task than learning physicochemical properties (e.g., aqueous solubility) or ADME (absorption, distribution, metabolism, and excretion) properties. Furthermore, bio-activity data is usually sparser which is a strong limitation for ML methods.

#### 4.2 AI beyond descriptors

AI also has a role in helping medicinal chemists plan their synthetic activities. Traditional retrosynthesis prediction relies heavily in chemical knowledge to set the rules of reactions. Language models can exploit the analogy between language and organic chemistry<sup>218</sup> to predict synthetic

precursors.<sup>219</sup> Apart from predicting the reactions themselves, it would be valuable to predict their yield. To this effect, Schwaller *et al.* built a transformer model to predict reaction yields based on SMILES representations. They have achieved this by combining a reaction SMILES encoder with a reaction regression to predict the yield, and speculate this could be applied to other regression tasks such as activation energies.<sup>220</sup> In other study, Schwaller *et al.* tackled the problem of reaction classification, also using a transformer.<sup>221</sup> Nevertheless, it is important to note that reaction fingerprinting based on *k*-nearest neighbors can achieve comparable accuracy in reaction classification and yield prediction with much less complexity.<sup>222</sup>

Other applications of deep learning in drug discovery include for example drug repurposing. An example comes from Zhang *et al.*, who used a transformer based on SMILES-protein sequence pairs to predict commercially available antiviral drugs which could be used against SARS-CoV-2.<sup>223</sup> A similar study was performed by Beck *et al.*<sup>224</sup> A completely different approach was taken by Yan *et al.*, who prompted ChatGPT with the task of proposing approved drugs which could be useful against Alzheimer's disease. They theorized that the model's ability to efficiently parse literature could be a reason for the plausibility of the suggestions it generated.<sup>225</sup> These are examples of what Vincent *et al.* describe as one category of ML studies in the area of phenotypic drug discovery. In this case, pharmacology data from other studies is transferred and used to predict new scaffolds for a given disease. The other category of studies is those that use phenotypical data, for example training them directly on data on cellular perturbations or gene expression changes.<sup>226</sup> Returning to target-based drug discovery, ML can be a useful help in finding new targets. ML models can be a useful tool for drug target prediction, meaning instead of finding a good binder for a specific target, existing molecules can be screened and their protein target predicted.<sup>227</sup> An example of drug-target interaction prediction is MolTrans, which based on a transformer classifier, predicts whether a drug-protein pair will interact.<sup>228</sup> A coupling of these techniques could be applied also in basic research. For example, given phenotypical data, drug-target interaction prediction could be used to find the mode of action of a drug.

Another active line of research concerns foundation models, which are usually large language models which are pretrained on large amounts of data to be later finetuned for a specific task.<sup>229</sup> These are similar to transfer learning, which is generally applied in drug discovery.<sup>230</sup> An example of transfer learning was presented by Tysinger *et al.*, where the authors pretrain a transformer model based on pairs of bioactive molecules from ChEMBL, and then use it to predict new molecules using known hits as input.<sup>231</sup> Unfortunately, this paper does not include any perspective, *i.e.*, experimental validation of the transformer model. A foundation model was used by Chenthamarakshan *et al.* to predict new binders using pretrained molecular and protein representations to classify molecules as binders or not, while also taking into account off target effects and synthesizability.<sup>232</sup> Chang and



Ye presented a bidirectional model linking SMILES and property prediction.<sup>233</sup> Finally, a concept that is recently expanding is that of digital twins, where systems of various complexities, cells for example, are represented virtually, enabling *in silico* experiments to be performed on them. A dialogue is then set between the digital twin and the experimental data, such that the twin can be fine tuned with the experimental data, and the predictions used to inform further experiments.<sup>234</sup>

Deep learning and related methods can have a strong impact in drug discovery, for example in aiding the sampling of novel chemical space and the virtual screening of these myriad new compounds, for generating quantum mechanics-level descriptors of molecular interactions, and to accelerate virtual screening.<sup>201</sup> In a time when target-based drug discovery has been described as inefficient, ML techniques could open the door to integrative data modelling which can yield not only the binding affinity to a single target, but also a prediction of the phenotypical effect of the screened molecules.<sup>235,236</sup> Challenges remain such as data curation and availability, or the increasing complexity in the types of data available. Additionally, molecular generation models need to be validated to ensure their output is sensible.<sup>237</sup> Still, ML methods could accelerate the drug discovery process not only by finding molecules that bind a target with high affinity, but also yield the desired effect based on omics data, phenotypical observation, selectivity prediction, or pharmacokinetics. Furthermore, test cycles could be reduced using predicted data, or the synthesis of new compounds made easier with retrosynthesis prediction tools.<sup>238</sup> In the end, the most important aspect will be to find a balance between the outputs of the ML models, and the human creativity of the medicinal chemists and structural bioinformaticians applying them.<sup>239</sup> AI applications in biology will continue to expand these coming years, in drug discovery and other areas,<sup>240</sup> but although several steps of the process have profited from data-driven augmentation, for now, human intervention remains essential.<sup>241</sup>

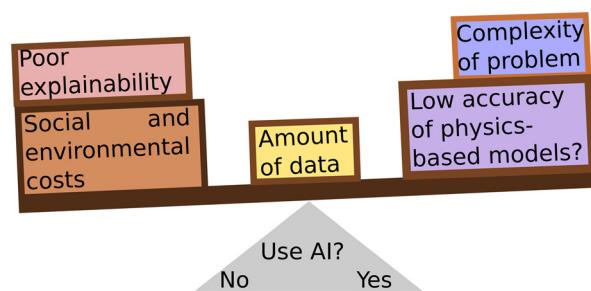
## 5. Is AI really what we need?

### 5.1 AI imposes a burden on resources

The ML explosion of the past few years has provided potential improvements in different stages of the drug discovery process, but it is also accompanied by some concerning trends following the wide application of AI. An usually forgotten factor when using ML models, and indeed also when running simulations,<sup>242,243</sup> is the energy consumption involved in the training and use of the models. Recent reports show that the “AI boom” of the past years is already threatening the climate goals of tech companies.<sup>244,245</sup> The International Energy Agency estimates that data centers, cryptocurrencies and AI have consumed a 2% of the total energy used in 2022. The projected growth of these sectors until 2026 means their electricity needs will be equivalent to the energy consumption of Germany.<sup>246</sup> The different sources of electricity used in the places where AI

models are trained, the cost (economic and environmental) of manufacturing the needed devices, and the diverse infrastructures where they are housed add increased complexity to the calculation of the environmental impact of AI models. The electricity and carbon cost of training one iteration of a model, *e.g.*, GPT-3 (around 1.2 GWh and 588 tons of CO<sub>2</sub> equivalents) seems negligible on its own,<sup>247</sup> but adds up quickly when considering the number of training runs needed to obtain a final version and the current explosion in the number of AI methods. Thus frameworks for quantifying and regulating the emissions due to AI (training and inference) are urgently needed.<sup>248</sup> Water is another equally important natural resource which is put under pressure for training and deploying AI systems. Water is used in the cooling towers of power plants generating electricity for AI, but also cooling systems of data centers themselves, and also during the manufacturing of computing infrastructure (chips). A recent estimate suggests that the training of the GPT-3 model requires on the order of 700 000 liters of water.<sup>249</sup> Obtaining accurate AI models while reducing the carbon output can be achieved through careful selection of the model used and has already been done by scientists.<sup>250</sup> Although individual action on its own is not enough to stop the detrimental advance of climate change,<sup>251</sup> we should as scientists be conscious of the impacts of our development and use of AI, and push for new paradigms in the deployment of such systems.<sup>252,253</sup> Using a combination of metrics such as efficiency and interpretability, in addition to accuracy, could be an answer to have better-designed and -implemented deep learning models, depending on fewer parameters, and therefore also reducing their economic and environmental cost.<sup>254</sup>

Careful consideration must be taken when choosing to use an AI solution in drug discovery, and indeed in any case. The goal of any scientific application should be to solve a problem in the most efficient way, not necessarily using the most advanced model (Fig. 3). Simpler models can be deployed for appropriate tasks, such as fingerprinting and regression for chemical properties,<sup>217</sup> conformer generation,<sup>255,256</sup> docking,<sup>257,258</sup> molecular descriptors,<sup>259</sup> reaction classification and yield prediction,<sup>222</sup> or others. Meanwhile, AI can be



**Fig. 3** The decision whether to deploy an AI model must be a careful consideration of several factors: can the problem be solved with simpler, *e.g.*, less data-hungry, methods? Is there enough data to train an AI model? Is explainability necessary for understanding the problem? Does the problem justify the costs of training and deploying an AI model?



reserved to those tasks that merit the cost of deploying it. As scientists, we should take care to consider the environmental, social, and ethical costs of AI systems while making this choice.

## 5.2 The knowledge of AI is becoming concentrated

Another concerning tendency that is manifesting itself, not only in biology, is the privatization of science. Deep learning systems, like AlphaFold,<sup>65</sup> make use of public datasets such as the Protein Data Bank for training.<sup>260</sup> This data has been collected over many decades and financed mostly through public money. Indeed the replacement cost of the PDB has been estimated at around USD 20 billion.<sup>261</sup> Nevertheless, AlphaFold3,<sup>66</sup> the latest iteration of the greatly successful deep learning protein structure prediction model, has been published without making its code freely available, angering scientists.<sup>262</sup> This decision even went against the editorial policy of its journal, which justified itself by claiming AlphaFold3 was privately funded and the service is still open and publicly available.<sup>263</sup> Only during revision of this manuscript, *i.e.*, about six months after its initial release, the code of AlphaFold3 was made public, with strong licensing restrictions. AlphaFold3 was developed by Alphabet's subsidiaries DeepMind and Isomorphic Labs, the latter of which will use it for its rational drug design campaigns.<sup>66,264</sup> Indeed, Fernández Pinto argues, this is a common characteristic of Open Science, where data generation is financed by the public and shared openly, but the private actors are not compelled to maintain the same standards and end profiting from the system.<sup>265</sup> Rikap shows how the assetization of knowledge and data to generate intellectual monopolies is a common strategy of tech giants.<sup>266</sup> Preventing intellectual monopolies could allow diverse smaller companies access and adopt the AI drug discovery workflows at a more competitive rate, instead of concentrating the technology on the tech giants with the money and IP to deploy them. A more critical standpoint might even argue that it is valid to oppose to profits being made from publicly sourced science. From the scientific point of view, the monopolization and platformization of AI research makes it more difficult to benchmark and improve these methods. Therefore, safeguards and regulations must be implemented to ensure that knowledge generated by AI is kept open and public, and to prevent that public infrastructures end being abused and profited from by private companies.<sup>267</sup> A framework moving towards this goal is needed and work on it is underway, to ensure not only public and open AI, but also that the society that contributed to the models can benefit from it.<sup>268,269</sup>

## 6. The future of ML in biology

What are the outstanding issues and next research directions in protein structure prediction and design? Current ML methods for protein structure prediction are trained on secondary data, *i.e.*, data derived usually by fitting from the raw (primary) data. Networks like AlphaFold are trained on

the protein structures available on the PDB, to obtain a probabilistic model of protein structure, and generate (a set of) single structures as output. The prediction of a single structure has been proposed as the current main limitation of these methods.<sup>84</sup> Nevertheless, the structures employed for training have also been fitted onto an electronic density (from X-ray crystallography or CryoEM), which is also modelled from the collected data. So far, researchers have used deep learning methods to reconstruct the protein structure into CryoEM maps,<sup>270–272</sup> predict flexibility,<sup>273</sup> or to solve the phase problem in X-ray crystallography for short peptide sequences.<sup>274</sup> Zhong *et al.* take an interesting direction, using variational autoencoders that encode CryoEM images and decode density maps, taking heterogeneity into account.<sup>275</sup> A similar method has been proposed by Rosenbaum *et al.*<sup>276</sup> The use of ML methods in CryoEM has expanded a lot and will probably continue to expand itself in the coming years.

To date, no method for protein structure prediction seems to be based on anything but static structures. Therefore, a possible paradigm shift could be to stop training the networks on the atomic coordinates of a single structure, and instead try to assign a probability distribution to their electronic densities, on which the deep neural networks can then be trained on. A possible difficulty arising from this is the (nearly impossible) assignment of continuous densities to the electronic cloud of single atoms, but this could be modelled based on the (constructed) structures which have been deposited. Nevertheless, and continuing with the translation from computational linguistics to biology, this could be seen as an analogous problem to the recognition of handwritten cursive text in computer vision.<sup>277–279</sup> Indeed, recent work has been done on generating the “MNIST of aminoacids” dataset, which could be an initial step towards assigning the continuous electron density to discrete residues.<sup>280</sup> Using this type of data could be a new way to tackle the lack of conformational ensembles from AlphaFold and related models.<sup>63</sup> Another possibility is to use raw data from NMR to augment ML networks. As discussed earlier, coevolutionary restraints from the MSAs applied in DL-based structure prediction work analogously to NMR structure determination using distance restraints. NMR data has been used to enhance the predictions of AlphaFold and obtain structures that comply with the experimental restraints,<sup>281</sup> or to assess the AlphaFold predictions.<sup>79,282</sup> However, training a network directly on NMR data has only been attempted for small molecules.<sup>283</sup> Of course, a problem which could arise comes from the fact that NMR experiments usually yield very few restraints, which are sometimes redundant. Still, an interesting avenue of research would be to incorporate different experimental data into the training of the deep learning model such that it can use the different structures, CryoEM electron densities, and NMR restraints to predict ensembles of structures.

Finally, something we consider as extremely interesting would be to try to abstract the physical rules of protein folding,



and indeed biomolecular interaction, from deep learning models for structure prediction. As discussed above, decades of research were invested into *ab initio* protein structure prediction with moderate success. It has been widely established that folding follows physical “rules”. At folding conditions, the native conformation is the thermodynamically most stable state, located at the bottom of the free energy funnel.<sup>284</sup> Is it possible that the physical models used for protein structure prediction until now were too simple to capture all the interactions needed to determine the three-dimensional conformation? An advantage of AI-based methods is their ability to learn representations of, or embed, complex data. If AlphaFold has learnt an energy function for protein structure, as proposed by some authors,<sup>98,99</sup> there should be a way to extract this information from the model to try to understand the physical interactions which govern protein structure and folding. While this claim has been disputed,<sup>94,96</sup> it is undeniable that the model predicts globular proteins with high accuracy. One of the newer trends in AI research is explainable AI (xAI).<sup>285</sup> There is no exact definition of xAI, but it can be understood as the set of methodologies which help users understand and believe the models and predictions from AI.<sup>286</sup> It would be interesting to apply the xAI framework, for example by examining the attention mechanism in AlphaFold, to try to understand how it predicts inter-residue interactions. Such an approach would offer a bridge between the “black box” approach of deep learning, and physics-based methods. The high-dimensional abstraction of the AI model could be used to better capture the interactions that govern protein folding and thus improve *ab initio* methods. A pitfall for this would be if AlphaFold has not learned the energy function of protein structure but rather it bases its predictions in memorization,<sup>96</sup> if the prediction is actually relating MSAs to structure instead of the sequence,<sup>71</sup> and indeed if the MSAs and the coevolutionary information they encode are not enough to predict different conformations. More research is needed to understand if AlphaFold and related methods can actually be interpreted and used to understand the physics behind protein folding.

## 7. Conclusions

We have reviewed the recent progresses of ML tools and their influence on protein structure prediction and computer-aided drug design. A long journey from the determination of the first crystal structure of a protein (myoglobin, more than 60 years ago) has culminated into half of the 2024 Nobel prize in chemistry which was awarded to the main developers of the deep learning programs for protein structure prediction. It is evident that these ML tools have predictive ability mainly because they are trained on the more than 200 000 experimentally determined structures of proteins. In turn, such a rich data set of protein structures exists because of impressive progress in the (mainly bacterial) production of pure proteins and the hard work of many research groups most of which are affiliated with not-for-profit institutions. The interplay and synergism between physics and AI are

expected to grow in the near future. The data-driven methods for modeling protein structures based on sequence homology have evolved into very powerful deep learning platforms. At the same time, the knowledge-based docking functions are being enhanced by using ML methods. Even force fields are being improved thanks to the large availability of data which can be exploited by ML models. The rich intersection between scientific disciplines such as the integration of language models into biology, or the physical inspiration of AI,<sup>287</sup> are opening new ways to tackle the shortcomings of current ML models. This exchange needs to be promoted to obtain better systems for protein structure prediction, docking of small molecules, or generation of ensembles. These advances did not happen from one day to the other, and it is important to recognize all the scientific progresses which enabled these technologies to be implemented and which are often forgotten.

Another important factor to consider is the cost of ML from the point of view of the natural resources needed to train and deploy it. This consideration should be central to decide whether it is justifiable to use an ML system in research, or if there are simpler (physics-based) solutions that can yield comparable and more explainable results with less energy consumption.

We should also be aware of the tendency for ML models to become services which scientists can use but have no full control over. This is exemplified by the most recent AlphaFold3. The emergence of “Science as a Service” already is hindering our understanding of the working of AlphaFold3 and other deep learning models. Therefore we should fight for control and access to these ML models which would not exist without the decades of work of generations of scientists supported mainly by public funding. As an example, during the past decade the research group of the senior author of this review has released in the PDB database more than 300 high-resolution crystal structures of proteins of pharmacological interest in complex with small molecule-ligands. Interestingly, most of these 300+ ligands were identified by high-throughput docking and force field-based binding energy evaluation.<sup>110,288–290</sup> Hence, we can state that physics-based docking, at least in the Caflisch group, and more generally protein X-ray crystallography have contributed substantially to the data used for training AlphaFold3. Thus, we should be prepared to discuss and set the rules of the game for open access to ML tools trained on open access data.<sup>291</sup>

Several challenges lie ahead on the road. Successful prediction of single protein structures is possible efficiently and routinely, but the generation of conformational ensembles requires additional effort. The most-widely used methods for protein structure prediction, and AI models in general, often lack explainability, making it difficult to understand their workings. For hit identification, physics based methods have been used in (high-throughput) docking programs for long, but their scoring functions can be further improved for reducing the number of false positives and thus increase the hit rates. Concerning hit to lead optimization, it



has been noted that ML tools, *e.g.*, generative modeling, are useful for interpolating within a known chemical series but are not able to extrapolate to new chemical matter.<sup>197,292,293</sup> The latter seems possible (at least for some chemical series and protein targets) by MD free energy-calculations which have improved substantially in accuracy thanks to continuous optimization of force fields (during the past decade particularly for small molecules) and faster hardware. Physics-based models use analytical energy functions (force fields) which achieve high accuracy, speed, and extensive coverage of chemical space by employing a fraction of the parameters used by ML models.<sup>294</sup> In conclusion, ML methods have evolved by taking advantage of decades of previous research in computational and structural biology. Now it is up to us to promote the physics-AI dialogue to successfully combine ML tools and physics-based methods for designing new drugs.

## Data availability

No primary research results, software or code have been included and no new data were generated or analyzed as part of this review.

## Author contributions

PV and AC wrote this review.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We thank Dr. J. R. Marchand for interesting discussions. Financial support was provided by the Swiss National Science Foundation (grant number 310030-212195). We thank the University Library Zurich for making this article Open Access.

## References

- 1 M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will and A. C. North, *Nature*, 1960, **185**, 416–422.
- 2 J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips and V. C. Shore, *Nature*, 1960, **185**, 422–427.
- 3 B. Strandberg, *J. Mol. Biol.*, 2009, **392**, 2–10.
- 4 R. E. Dickerson, *J. Mol. Biol.*, 2009, **392**, 10–23.
- 5 M. G. Rossmann, *J. Mol. Biol.*, 2009, **392**, 23–32.
- 6 C. B. Anfinsen, E. Haber, M. Sela and F. H. White, *Proc. Natl. Acad. Sci. U. S. A.*, 1961, **47**, 1309–1314.
- 7 C. Levinthal, E. R. Signer and K. Fetherolf, *Proc. Natl. Acad. Sci. U. S. A.*, 1962, **48**, 1230–1237.
- 8 M. F. Perutz, J. C. Kendrew and H. C. Watson, *J. Mol. Biol.*, 1965, **13**, 669–678.
- 9 A. V. Guzzo, *Biophys. J.*, 1965, **5**, 809–822.
- 10 J. W. Prothero, *Biophys. J.*, 1966, **6**, 367–370.
- 11 R. H. Pain and B. Robson, *Nature*, 1970, **227**, 62–63.
- 12 K. Nagano, *J. Mol. Biol.*, 1973, **75**, 401–420.
- 13 E. A. Kabat and T. Wu, *Biopolymers*, 1973, **12**, 751–774.
- 14 M. Froimowitz and G. D. Fasman, *Macromolecules*, 1974, **7**, 583–589.
- 15 W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577–2637.
- 16 C. J. Epstein, *Nature*, 1964, **203**, 1350–1352.
- 17 I. Shindyalov, N. Kolchanov and C. Sander, *Protein Eng., Des. Sel.*, 1994, **7**, 349–358.
- 18 F. Pazos, M. Helmer-Citterich, G. Ausiello and A. Valencia, *J. Mol. Biol.*, 1997, **271**, 511–523.
- 19 P. Fariselli, O. Olmea, A. Valencia and R. Casadio, *Protein Eng., Des. Sel.*, 2001, **14**, 835–843.
- 20 T. A. Hopf, C. P. Schärfe, J. P. Rodrigues, A. G. Green, O. Kohlbacher, C. Sander, A. M. Bonvin and D. S. Marks, *eLife*, 2014, **3**, e03430.
- 21 D. S. Marks, T. A. Hopf and C. Sander, *Nat. Biotechnol.*, 2012, **30**, 1072–1080.
- 22 F. Morcos, A. Pagnani, B. Lunt, A. Bertolino, D. S. Marks, C. Sander, R. Zecchina, J. N. Onuchic, T. Hwa and M. Weigt, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, E1293–E1301.
- 23 D. De Juan, F. Pazos and A. Valencia, *Nat. Rev. Genet.*, 2013, **14**, 249–261.
- 24 J. Andreani and R. Guerois, *Arch. Biochem. Biophys.*, 2014, **554**, 65–75.
- 25 J. Moult, J. T. Pedersen, R. Judson and K. Fidelis, *Proteins: Struct., Funct., Bioinf.*, 1995, **23**, ii–iv.
- 26 J. Moult, K. Fidelis, A. Zemla and T. Hubbard, *Proteins: Struct., Funct., Bioinf.*, 2001, **45**, 2–7.
- 27 R. Srinivasan and G. D. Rose, *Proteins: Struct., Funct., Bioinf.*, 2002, **47**, 489–495.
- 28 C. Thachuk, A. Shmygelska and H. H. Hoos, *BMC Bioinf.*, 2007, **8**, 1–20.
- 29 J. Skolnick and D. Kihara, *Proteins: Struct., Funct., Bioinf.*, 2001, **42**, 319–331.
- 30 K. E. Han, C. Bystroff and D. Baker, *Protein Sci.*, 1997, **6**, 1587–1590.
- 31 J. L. Klepeis and C. A. Floudas, *Biophys. J.*, 2003, **85**, 2119–2146.
- 32 J. L. Klepeis, Y. Wei, M. H. Hecht and C. A. Floudas, *Proteins: Struct., Funct., Bioinf.*, 2005, **58**, 560–570.
- 33 K. T. Simons, R. Bonneau, I. Ruczinski and D. Baker, *Proteins: Struct., Funct., Bioinf.*, 1999, **37**, 171–176.
- 34 E. S. Huang, J. W. Ponder and R. Samudrala, *Protein Sci.*, 1998, **7**, 1998–2003.
- 35 J. Skolnick, A. Kolinski and A. R. Ortiz, *J. Mol. Biol.*, 1997, **265**, 217–241.
- 36 A. Hillisch, L. F. Pineda and R. Hilgenfeld, *Drug Discovery Today*, 2004, **9**, 659–669.
- 37 C. Chothia and A. M. Lesk, *EMBO J.*, 1986, **5**, 823–826.
- 38 A. Tramontano and V. Morea, *Proteins: Struct., Funct., Bioinf.*, 2003, **53**, 352–368.
- 39 J. Moult, *Curr. Opin. Struct. Biol.*, 2005, **15**, 285–289.
- 40 J. Kopp and T. Schwede, *Pharmacogenomics*, 2004, **5**, 405–416.
- 41 J. A. Dalton and R. M. Jackson, *Bioinformatics*, 2007, **23**, 1901–1908.



42 T. Schwede, J. Kopp, N. Guex and M. C. Peitsch, *Nucleic Acids Res.*, 2003, **31**, 3381–3385.

43 A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore and T. Schwede, *Nucleic Acids Res.*, 2018, **46**, W296–W303.

44 G. Studer, G. Tauriello, S. Bienert, M. Biasini, N. Johner and T. Schwede, *PLoS Comput. Biol.*, 2021, **17**, e1008667.

45 B. Webb and A. Sali, *Curr. Protoc. Bioinf.*, 2016, **54**, 5.6.1–5.6.37.

46 A. Šali and T. L. Blundell, *J. Mol. Biol.*, 1993, **234**, 779–815.

47 A. Fiser, R. K. G. Do and A. Šali, *Protein Sci.*, 2000, **9**, 1753–1773.

48 J. Skolnick, A. Kolinski, D. Kihara, M. Betancourt, P. Rotkiewicz and M. Boniecki, *Proteins: Struct., Funct., Bioinf.*, 2001, **45**, 149–156.

49 Y. Zhang and J. Skolnick, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 7594–7599.

50 Y. Zhang and J. Skolnick, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 1029–1034.

51 A. Kryshtafovych, K. Fidelis and A. Tramontano, *Proteins: Struct., Funct., Bioinf.*, 2011, **79**, 91–106.

52 Y. J. Huang, B. Mao, J. M. Aramini and G. T. Montelione, *Proteins: Struct., Funct., Bioinf.*, 2014, **82**, 43–56.

53 J. Moult, K. Fidelis, A. Kryshtafovych, T. Schwede and A. Tramontano, *Proteins: Struct., Funct., Bioinf.*, 2018, **86**, 7–15.

54 A. Kryshtafovych, T. Schwede, M. Topf, K. Fidelis and J. Moult, *Proteins: Struct., Funct., Bioinf.*, 2019, **87**, 1011–1020.

55 A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Židek, A. W. Nelson and A. Bridgland, *et al.*, *Nature*, 2020, **577**, 706–710.

56 L. M. Bertoline, A. N. Lima, J. E. Krieger and S. K. Teixeira, *Front. Bioinform.*, 2023, **3**, 1120370.

57 C.-X. Peng, F. Liang, Y.-H. Xia, K.-L. Zhao, M.-H. Hou and G.-J. Zhang, *J. Chem. Inf. Model.*, 2023, **64**, 76–95.

58 D. Baker and A. Sali, *Science*, 2001, **294**, 93–96.

59 J. Skolnick, J. S. Fetrow and A. Kolinski, *Nat. Biotechnol.*, 2000, **18**, 283–287.

60 J. Durairaj, A. M. Waterhouse, T. Mets, T. Brodiazhenko, M. Abdullah, G. Studer, G. Tauriello, M. Akdel, A. Andreeva and A. Bateman, *et al.*, *Nature*, 2023, **622**, 646–653.

61 K. Tunyasuvunakool, J. Adler, Z. Wu, T. Green, M. Zielinski, A. Židek, A. Bridgland, A. Cowie, C. Meyer and A. Laydon, *et al.*, *Nature*, 2021, **596**, 590–596.

62 R. Pearce and Y. Zhang, *Curr. Opin. Struct. Biol.*, 2021, **68**, 194–207.

63 G. R. Bowman, *Annu. Rev. Biomed. Data Sci.*, 2024, **7**, 51–57.

64 J. Skolnick, M. Gao, H. Zhou and S. Singh, *J. Chem. Inf. Model.*, 2021, **61**, 4827–4831.

65 J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Židek and A. Potapenko, *et al.*, *Nature*, 2021, **596**, 583–589.

66 J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard and J. Bambrick, *et al.*, *Nature*, 2024, **630**, 1–3.

67 M. Baek, F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G. R. Lee, J. Wang, Q. Cong, L. N. Kinch and R. D. Schaeffer, *et al.*, *Science*, 2021, **373**, 871–876.

68 R. Krishna, J. Wang, W. Ahern, P. Sturmels, P. Venkatesh, I. Kalvet, G. R. Lee, F. S. Morey-Burrows, I. Anishchenko and I. R. Humphreys, *et al.*, *Science*, 2024, **384**, eadl2528.

69 A. Godzik, *Structure*, 1996, **4**, 363–366.

70 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. U. Kaiser and I. Polosukhin, *Adv. Neural Inf. Process Syst.*, 2017, **30**, 5998–6008.

71 G. Ahdritz, N. Bouatta, C. Floristean, S. Kadyan, Q. Xia, W. Gerecke, T. J. O'Donnell, D. Berenberg, I. Fisk and N. Zanichelli, *et al.*, *Nat. Methods*, 2024, **21**, 1–11.

72 Z. Lin, H. Akin, R. Rao, B. Hie, Z. Zhu, W. Lu, N. Smetanin, R. Verkuil, O. Kabeli and Y. Shmueli, *et al.*, *Science*, 2023, **379**, 1123–1130.

73 B. Moussad, R. Roche and D. Bhattacharya, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**, e2303499120.

74 J. Yim, H. Stark, G. Corso, B. Jing, R. Barzilay and T. S. Jaakkola, *WIREs Comput. Mol. Sci.*, 2024, **14**, e1711.

75 J. L. Watson, D. Juergens, N. R. Bennett, B. L. Trippe, J. Yim, H. E. Eisenach, W. Ahern, A. J. Borst, R. J. Ragotte and L. F. Milles, *et al.*, *Nature*, 2023, **620**, 1089–1100.

76 K. E. Wu, K. K. Yang, R. van den Berg, S. Alamdari, J. Y. Zou, A. X. Lu and A. P. Amini, *Nat. Commun.*, 2024, **15**, 1059.

77 B. Jing, G. Corso, J. Chang, R. Barzilay and T. Jaakkola, *Adv. Neural Inf. Process Syst.*, 2022, 24240–24253.

78 A. Schneuing, Y. Du, C. Harris, A. Jamasb, I. Igashov, W. Du, T. Blundell, P. Lio, C. Gomes and M. Welling, *et al.*, *arXiv*, 2022, preprint, arXiv:2210.13695, DOI: [10.48550/arXiv.2210.13695](https://doi.org/10.48550/arXiv.2210.13695).

79 J. P. Bonin, J. M. Aramini, Y. Dong, H. Wu and L. E. Kay, *J. Magn. Reson.*, 2024, **364**, 107725.

80 K. M. Ruff and R. V. Pappu, *J. Mol. Biol.*, 2021, **433**, 167208.

81 T. R. Alderson, I. Pritišanac, D. Kolarić, A. M. Moses and J. D. Forman-Kay, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**, e2304302120.

82 H. Bret, J. Gao, D. J. Zea, J. Andreani and R. Guerois, *Nat. Commun.*, 2024, **15**, 597.

83 P. A. Vargas-Rosales, A. D'Addio, Y. Zhang and A. Caflisch, *ACS Phys. Chem. Au*, 2023, **3**, 456–466.

84 T. J. Lane, *Nat. Methods*, 2023, **20**, 170–173.

85 B. Wallner, *Bioinformatics*, 2023, **39**, btad573.

86 G. Monteiro da Silva, J. Y. Cui, D. C. Dalgarno, G. P. Lisi and B. M. Rubenstein, *Nat. Commun.*, 2024, **15**, 2464.

87 D. del Alamo, D. Sala, H. S. Mchaourab and J. Meiler, *eLife*, 2022, **11**, e75751.

88 B. P. Vani, A. Aranganathan, D. Wang and P. Tiwary, *J. Chem. Theory Comput.*, 2023, **19**, 4351–4354.

89 R. A. Stein and H. S. Mchaourab, *PLoS Comput. Biol.*, 2022, **18**, 1–16.

90 B. Jing, B. Berger and T. Jaakkola, *arXiv*, 2024, preprint, arXiv:2402.04845, DOI: [10.48550/arXiv.2402.04845](https://doi.org/10.48550/arXiv.2402.04845).

91 D. Sala, F. Engelberger, H. Mchaourab and J. Meiler, *Curr. Opin. Struct. Biol.*, 2023, **81**, 102645.



92 N. Anand and T. Achim, *arXiv*, 2022, preprint, arXiv:2205.15019, DOI: [10.48550/arXiv.2205.15019](https://doi.org/10.48550/arXiv.2205.15019).

93 P. Bryant and F. Noé, *Nat. Commun.*, 2024, **15**, 7328.

94 C. Outeiral, D. A. Nissley and C. M. Deane, *Bioinformatics*, 2022, **38**, 1881–1887.

95 D. Chakravarty and L. L. Porter, *Protein Sci.*, 2022, **31**, e4353.

96 D. Chakravarty, J. W. Schafer, E. A. Chen, J. F. Thole, L. A. Ronish, M. Lee and L. L. Porter, *Nat. Commun.*, 2024, **15**, 7296.

97 T. Saldaño, N. Escobedo, J. Marchetti, D. J. Zea, J. Mac Donagh, A. J. Velez Rueda, E. Gonik, A. García Melani, J. Novomisky Nechcoff, M. N. Salas, T. Peters, N. Demitroff, S. Fernandez Alberti, N. Palopoli, M. S. Fornasari and G. Parisi, *Bioinformatics*, 2022, **38**, 2742–2748.

98 J. P. Roney and S. Ovchinnikov, *Phys. Rev. Lett.*, 2022, **129**, 238101.

99 J. A. Gut and T. Lemmin, *bioRxiv*, 2024, preprint, DOI: [10.1101/2024.03.14.585076v2](https://doi.org/10.1101/2024.03.14.585076v2).

100 B. Strodel, *J. Mol. Biol.*, 2021, **433**, 167182.

101 M. R. Masters, A. H. Mahmoud and M. A. Lill, *bioRxiv*, 2024, preprint, DOI: [10.1101/2024.06.03.597219v1](https://doi.org/10.1101/2024.06.03.597219v1).

102 M. De Vivo, M. Masetti, G. Bottegoni and A. Cavalli, *J. Med. Chem.*, 2016, **59**, 4035–4061.

103 J. D. Durrant and J. A. McCammon, *BMC Biol.*, 2011, **9**, 1–9.

104 P. Śledź and A. Caflisch, *Curr. Opin. Struct. Biol.*, 2018, **48**, 93–102.

105 I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge and T. E. Ferrin, *J. Mol. Biol.*, 1982, **161**, 269–288.

106 D. S. Goodsell and A. J. Olson, *Proteins: Struct., Funct., Bioinf.*, 1990, **8**, 195–202.

107 A. Caflisch, S. Fischer and M. Karplus, *J. Comput. Chem.*, 1997, **18**, 723–743.

108 M. A. Phillips, M. A. Stewart, D. L. Woodling and Z.-R. Xie, *Molecular Docking*, IntechOpen, Rijeka, 2018, ch. 8.

109 B. Shaker, S. Ahmad, J. Lee, C. Jung and D. Na, *Comput. Biol. Med.*, 2021, **137**, 104851.

110 J.-R. Marchand and A. Caflisch, *Eur. J. Med. Chem.*, 2018, **156**, 907–917.

111 R. Wang, Y. Lu and S. Wang, *J. Med. Chem.*, 2003, **46**, 2287–2303.

112 K. Vanommeslaeghe and A. MacKerell, *Biochim. Biophys. Acta, Gen. Subj.*, 2015, **1850**, 861–871.

113 W. Hwang, S. L. Austin, A. Blondel, E. D. Boittier, S. Boresch, M. Buck, J. Buckner, A. Caflisch, H.-T. Chang and X. Cheng, *et al.*, *J. Phys. Chem. B*, 2024, **128**, 9976–10042.

114 J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. De Groot, H. Grubmüller and A. D. MacKerell Jr, *Nat. Methods*, 2017, **14**, 71–73.

115 U. Haberthür and A. Caflisch, *J. Comput. Chem.*, 2008, **29**, 701–715.

116 S.-Y. Huang, S. Z. Grinter and X. Zou, *Phys. Chem. Chem. Phys.*, 2010, **12**, 12899–12908.

117 F. Ahmed and C. L. Brooks III, *J. Chem. Inf. Model.*, 2023, **63**, 7219–7227.

118 N. Majeux, M. Scarsi, J. Apostolakis, C. Ehrhardt and A. Caflisch, *Proteins: Struct., Funct., Bioinf.*, 1999, **37**, 88–105.

119 J. Eberhardt, D. Santos-Martins, A. F. Tillack and S. Forli, *J. Chem. Inf. Model.*, 2021, **61**, 3891–3898.

120 T. Zhou and A. Caflisch, *ChemMedChem*, 2010, **5**, 1007–1014.

121 T. Zhou, D. Huang and A. Caflisch, *Curr. Top. Med. Chem.*, 2010, **10**, 33–45.

122 C. N. Cavasotto and M. G. Aucar, *Front. Chem.*, 2020, **8**, 246.

123 A. A. Adeniyi and M. E. Soliman, *Drug Discovery Today*, 2017, **22**, 1216–1223.

124 I. A. Guedes, F. S. Pereira and L. E. Dardenne, *Front. Pharmacol.*, 2018, **9**, 1089.

125 R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley and J. K. Perry, *et al.*, *J. Med. Chem.*, 2004, **47**, 1739–1749.

126 M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini and R. P. Mee, *J. Comput.-Aided Mol. Des.*, 1997, **11**, 425–445.

127 A. Sweeney, T. Mulvaney, M. Maiorca and M. Topf, *J. Med. Chem.*, 2023, **67**, 199–212.

128 H. Gohlke, M. Hendlich and G. Klebe, *J. Mol. Biol.*, 2000, **295**, 337–356.

129 M. J. Sippl, *J. Mol. Biol.*, 1990, **213**, 859–883.

130 J. Liu and R. Wang, *J. Chem. Inf. Model.*, 2015, **55**, 475–482.

131 H. M. Ashtawy and N. R. Mahapatra, *IEEE/ACM Trans. Comput. Biol. Bioinf.*, 2014, **12**, 335–347.

132 J. Scantlebury, L. Vost, A. Carbery, T. E. Hadfield, O. M. Turnbull, N. Brown, V. Chenthamarakshan, P. Das, H. Grosjean and F. Von Delft, *et al.*, *J. Chem. Inf. Model.*, 2023, **63**, 2960–2974.

133 A. T. McNutt, P. Francoeur, R. Aggarwal, T. Masuda, R. Meli, M. Ragoza, J. Sunseri and D. R. Koes, *J. Cheminf.*, 2021, **13**, 43.

134 I. A. Guedes, A. M. Barreto, D. Marinho, E. Krempser, M. A. Kuenemann, O. Sperandio, L. E. Dardenne and M. A. Miteva, *Sci. Rep.*, 2021, **11**, 3198.

135 K. J. Fujimoto, S. Minami and T. Yanai, *ACS Omega*, 2022, **7**, 19030–19039.

136 C. Isert, K. Atz, S. Riniker and G. Schneider, *RSC Adv.*, 2024, **14**, 4492–4502.

137 C. Shen, J. Ding, Z. Wang, D. Cao, X. Ding and T. Hou, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2020, **10**, e1429.

138 T. Janelia and J. Bajorath, *Cell Rep. Phys. Sci.*, 2024, **5**, 101988.

139 M. Buttenschoen, G. M. Morris and C. M. Deane, *Chem. Sci.*, 2024, **15**, 3130–3139.

140 R. Meli, G. M. Morris and P. C. Biggin, *Front. bioinform.*, 2022, **2**, 639349.

141 A. Vitalis and R. V. Pappu, *Annu. Rep. Comput. Chem.*, 2009, **5**, 49–76.

142 M. Karplus and J. A. McCammon, *Nat. Struct. Biol.*, 2002, **9**, 646–652.

143 T. Ginex, J. Vázquez, C. Estarellas and F. Luque, *Curr. Opin. Struct. Biol.*, 2024, **87**, 102870.

144 C. M. Clemente, L. Capece and M. A. Martí, *J. Chem. Inf. Model.*, 2023, **63**, 2609–2627.

145 D. R. Salahub, *Phys. Chem. Chem. Phys.*, 2022, **24**, 9051–9081.



146 H. Zhao and A. Caflisch, *Eur. J. Med. Chem.*, 2015, **91**, 4–14.

147 X. Liu, D. Shi, S. Zhou, H. Liu, H. Liu and X. Yao, *Expert Opin. Drug Discovery*, 2018, **13**, 23–37.

148 G. Spagnolli, T. Massignan, A. Astolfi, S. Biggi, M. Rigoli, P. Brunelli, M. Libergoli, A. Ianeselli, S. Orioli and A. Boldrini, *et al.*, *Commun. Biol.*, 2021, **4**, 62.

149 U. S. Herrmann, A. K. Schütz, H. Shirani, D. Huang, D. Saban, M. Nuvolone, B. Li, B. Ballmer, A. K. O. Åslund, J. J. Mason, E. Rushing, H. Budka, S. Nyström, P. Hammarström, A. Böckmann, A. Caflisch, B. H. Meier, K. P. R. Nilsson, S. Hornemann and A. Aguzzi, *Sci. Transl. Med.*, 2015, **7**, 299ra123.

150 I. M. Ilie and A. Caflisch, *Chem. Rev.*, 2019, **119**, 6956–6993.

151 S. Kalita, H. Bergman, K. D. Dubey and S. Shaik, *J. Am. Chem. Soc.*, 2023, **145**, 3543–3553.

152 I. Corbeski, P. A. Vargas-Rosales, R. K. Bedi, J. Deng, D. Coelho, E. Braud, L. Iannazzo, Y. Li, D. Huang, M. Ethève-Quelquejeu, Q. Cui and A. Caflisch, *eLife*, 2024, **12**, RP92537.

153 F. Nai, M. P. Flores Espinoza, A. Invernizzi, P. A. Vargas-Rosales, O. Bobileva, M. Herok and A. Caflisch, *ACS BioMed Chem Au*, 2024, **4**, 100–110.

154 V. A. Adediwura, K. Koirala, H. N. Do, J. Wang and Y. Miao, *Expert Opin. Drug Discovery*, 2024, **19**, 671–682.

155 A. Pohorille, C. Jarzynski and C. Chipot, *J. Phys. Chem. B*, 2010, **114**, 10235–10253.

156 J. Gao, K. Kuczera, B. Tidor and M. Karplus, *Science*, 1989, **244**, 1069–1072.

157 T. Simonson, G. Archontis and M. Karplus, *Acc. Chem. Res.*, 2002, **35**, 430–437.

158 R. Qian, J. Xue, Y. Xu and J. Huang, *J. Chem. Inf. Model.*, 2024, **64**, 7214–7237.

159 T. J. Giese and D. M. York, *J. Chem. Theory Comput.*, 2018, **14**, 1564–1582.

160 R. W. Zwanzig, *J. Chem. Phys.*, 1954, **22**, 1420–1426.

161 W. Jespers, J. Åqvist and H. Gutiérrez-de Terán, in *Free Energy Calculations for Protein–Ligand Binding Prediction*, ed. F. Ballante, Springer US, New York, NY, 2021, pp. 203–226.

162 G. A. Ross, C. Lu, G. Scarabelli, S. K. Albanese, E. Houang, R. Abel, E. D. Harder and L. Wang, *Commun. Chem.*, 2023, **6**, 222.

163 C. E. Schindler, D. Kuhn and I. V. Hartung, *Nat. Rev. Chem.*, 2023, **7**, 752–753.

164 H.-J. Woo and B. Roux, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6825–6830.

165 I. Muegge and Y. Hu, *ACS Med. Chem. Lett.*, 2023, **14**, 244–250.

166 J.-R. Marchand, T. Knehans, A. Caflisch and A. Vitalis, *J. Chem. Inf. Model.*, 2020, **60**, 5188–5202.

167 S. Bottaro and K. Lindorff-Larsen, *Science*, 2018, **361**, 355–360.

168 J. Šponer, G. Bussi, M. Krepl, P. Banáš, S. Bottaro, R. A. Cunha, A. Gil-Ley, G. Pinamonti, S. Poblete and P. Jurečka, *et al.*, *Chem. Rev.*, 2018, **118**, 4177–4338.

169 J. Widmer, A. Vitalis and A. Caflisch, *J. Biol. Chem.*, 2024, **300**, 107998.

170 F. Panai, P. Gkeka and M. Bonomi, *Nat. Commun.*, 2024, **15**, 5725.

171 A. Bartocci, A. Grazzi, N. Awad, P.-J. Corringer, P. C. Souza and M. Cecchini, *bioRxiv*, 2024, preprint, DOI: [10.1101/2023.04.19.537578v2](https://doi.org/10.1101/2023.04.19.537578v2).

172 I. M. Ilie, W. K. den Otter and W. J. Briels, *J. Chem. Phys.*, 2017, **146**, 115102.

173 J. Hénin, T. Lelièvre, M. R. Shirts, O. Valsson and L. Delemonette, *Living J. Comput. Mol. Sci.*, 2022, **4**, 1583.

174 R. C. Bernardi, M. C. Melo and K. Schulten, *Biochim. Biophys. Acta, Gen. Subj.*, 2015, **1850**, 872–877.

175 S. V. Krivov, *J. Chem. Theory Comput.*, 2021, **17**, 5466–5481.

176 B. Roux, *J. Phys. Chem. A*, 2021, **125**, 7558–7571.

177 M. Bacci, A. Vitalis and A. Caflisch, *Biochim. Biophys. Acta, Gen. Subj.*, 2015, **1850**, 889–902.

178 C. Kutzner, S. Pál, M. Fechner, A. Esztermann, B. L. de Groot and H. Grubmüller, *J. Comput. Chem.*, 2015, **36**, 1990–2008.

179 S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374–380.

180 S. Hoff, M. Zinke, N. Izadi-Pruneyre and M. Bonomi, *Curr. Opin. Struct. Biol.*, 2024, **84**, 102746.

181 M. Bernetti and G. Bussi, *Curr. Opin. Struct. Biol.*, 2023, **78**, 102503.

182 J. R. Allison, *Curr. Opin. Struct. Biol.*, 2017, **43**, 79–87.

183 M. Bonomi, C. Camilloni, A. Cavalli and M. Vendruscolo, *Sci. Adv.*, 2016, **2**, e1501177.

184 M. Bonomi and C. Camilloni, *Bioinformatics*, 2017, **33**, 3999–4000.

185 S. Mehdi, Z. Smith, L. Herron, Z. Zou and P. Tiwary, *Annu. Rev. Phys. Chem.*, 2024, **75**, 347–370.

186 F. Noé, S. Olsson, J. Köhler and H. Wu, *Science*, 2019, **365**, eaaw1147.

187 O. T. Unke, S. Chmiela, H. E. Sauceda, M. Gastegger, I. Poltavsky, K. T. Schütt, A. Tkatchenko and K.-R. Müller, *Chem. Rev.*, 2021, **121**, 10142–10186.

188 M. Chen, X. Jiang, L. Zhang, X. Chen, Y. Wen, Z. Gu, X. Li and M. Zheng, *Med. Res. Rev.*, 2024, **44**, 1147–1182.

189 M. Majewski, A. Pérez, P. Thölk, S. Doerr, N. E. Charron, T. Giorgino, B. E. Husic, C. Clementi, F. Noé and G. De Fabritiis, *Nat. Commun.*, 2023, **14**, 5739.

190 S. Doerr, M. Majewski, A. Pérez, A. Kramer, C. Clementi, F. Noe, T. Giorgino and G. De Fabritiis, *J. Chem. Theory Comput.*, 2021, **17**, 2355–2363.

191 E. Christofi, P. Bacova and V. A. Harmandaris, *J. Chem. Inf. Model.*, 2024, **64**, 1853–1867.

192 W. Li, C. Burkhardt, P. Polińska, V. Harmandaris and M. Doxastakis, *J. Chem. Phys.*, 2020, **153**, 041101.

193 J. Kohler, Y. Chen, A. Kramer, C. Clementi and F. Noé, *J. Chem. Theory Comput.*, 2023, **19**, 942–952.

194 S. Röcken and J. Zavadlav, *npj Comput. Mater.*, 2024, **10**, 69.

195 A. J. Simpkin, S. Mesdaghi, F. Sánchez Rodríguez, L. Elliott, D. L. Murphy, A. Kryshtafovych, R. M. Keegan and D. J. Rigden, *Proteins: Struct., Funct., Bioinf.*, 2023, **91**, 1616–1635.

196 M. Pandey, M. Fernandez, F. Gentile, O. Isayev, A. Tropsha, A. C. Stern and A. Cherkasov, *Nat. Mach. Intell.*, 2022, **4**, 211–221.

197 W. P. Walters and M. Murcko, *Nat. Biotechnol.*, 2020, **38**, 143–145.

198 L. Rossen, F. Sirockin, N. Schneider and F. Grisoni, *ChemRxiv*, 2024, preprint, DOI: [10.26434/chemrxiv-2024-gd3j4](https://doi.org/10.26434/chemrxiv-2024-gd3j4).

199 K. Yoo, O. Oertell, J. Lee, S. Lee and J. Kang, *arXiv*, 2024, preprint, arXiv:2410.20660, DOI: [10.48550/arXiv.2410.20660](https://doi.org/10.48550/arXiv.2410.20660).

200 L. K. Tsou, S.-H. Yeh, S.-H. Ueng, C.-P. Chang, J.-S. Song, M.-H. Wu, H.-F. Chang, S.-R. Chen, C. Shih and C.-T. Chen, *et al.*, *Sci. Rep.*, 2020, **10**, 16771.

201 A. Tropsha, O. Isayev, A. Varnek, G. Schneider and A. Cherkasov, *Nat. Rev. Drug Discovery*, 2024, **23**, 141–155.

202 X. Li and D. Fourches, *J. Cheminf.*, 2020, **12**, 1–15.

203 G. Schneider and U. Fechner, *Nat. Rev. Drug Discovery*, 2005, **4**, 649–663.

204 N. Budin, N. Majeux, C. Tenette-Souaille and A. Caflisch, *J. Comput. Chem.*, 2001, **22**, 1956–1970.

205 L. Batiste, A. Unzue, A. Dolbois, F. Hassler, X. Wang, N. Deerain, J. Zhu, D. Spiliopoulos, C. Nevado and A. Caflisch, *ACS Cent. Sci.*, 2018, **4**, 180–188.

206 Y. Tang, R. Moretti and J. Meiler, *J. Chem. Inf. Model.*, 2024, **64**, 1794–1805.

207 C. Pang, J. Qiao, X. Zeng, Q. Zou and L. Wei, *J. Chem. Inf. Model.*, 2023, **64**, 2174–2194.

208 B. P. Munson, M. Chen, A. Bogosian, J. F. Kreisberg, K. Licon, R. Abagyan, B. M. Kuenzi and T. Ideker, *Nat. Commun.*, 2024, **15**, 3636.

209 G. Schneider and D. E. Clark, *Angew. Chem., Int. Ed.*, 2019, **58**, 10792–10803.

210 C. Arnold, *Nat. Med.*, 2023, **29**, 1292–1295.

211 M. McGibbon, S. Shave, J. Dong, Y. Gao, D. R. Houston, J. Xie, Y. Yang, P. Schwaller and V. Blay, *Briefings Bioinf.*, 2023, **25**, bbad422.

212 D. Baptista, J. Correia, B. Pereira and M. Rocha, *J. Integr. Bioinform.*, 2022, **19**, 20220006.

213 M. Krenn, F. Hase, A. Nigam, P. Friederich and A. Aspuru-Guzik, *Mach. Learn.: Sci. Technol.*, 2020, **1**, 045024.

214 P. Carracedo-Reboredo, J. Liñares-Blanco, N. Rodríguez-Fernández, F. Cedrón, F. J. Novoa, A. Carballal, V. Maojo, A. Pazos and C. Fernandez-Lozano, *Comput. Struct. Biotechnol. J.*, 2021, **19**, 4538–4558.

215 K. Gao, D. D. Nguyen, V. Sresht, A. M. Mathiowetz, M. Tu and G.-W. Wei, *Phys. Chem. Chem. Phys.*, 2020, **22**, 8373–8390.

216 D. van Tilborg, A. Alenicheva and F. Grisoni, *J. Chem. Inf. Model.*, 2022, **62**, 5938–5951.

217 A. Capecchi, D. Probst and J.-L. Reymond, *J. Cheminf.*, 2020, **12**, 1–15.

218 A. Cadeddu, E. K. Wylie, J. Jurczak, M. Wampler-Doty and B. A. Grzybowski, *Angew. Chem., Int. Ed.*, 2014, **53**, 8108–8112.

219 J. Jiménez-Luna, N. Grisoni, F. Weskamp and G. Schneider, *Expert Opin. Drug Discovery*, 2021, **16**, 949–959.

220 P. Schwaller, A. C. Vaucher, T. Laino and J.-L. Reymond, *Mach. Learn.: Sci. Technol.*, 2021, **2**, 015016.

221 P. Schwaller, D. Probst, A. C. Vaucher, V. H. Nair, D. Kreutter, T. Laino and J.-L. Reymond, *Nat. Mach. Intell.*, 2021, **3**, 144–152.

222 D. Probst, P. Schwaller and J.-L. Reymond, *Digital Discovery*, 2022, **1**, 91–97.

223 Y. Zhang, T. Ye, H. Xi, M. Juhas and J. Li, *Front. Microbiol.*, 2021, **12**, 739684.

224 B. R. Beck, B. Shin, Y. Choi, S. Park and K. Kang, *Comput. Struct. Biotechnol. J.*, 2020, **18**, 784–790.

225 C. Yan, M. E. Grabowska, A. L. Dickson, B. Li, Z. Wen, D. M. Roden, C. Michael Stein, P. J. Embí, J. F. Peterson and Q. Feng, *et al.*, *NPJ Digit. Med.*, 2024, **7**, 46.

226 F. Vincent, A. Nueda, J. Lee, M. Schenone, M. Prunotto and M. Mercola, *Nat. Rev. Drug Discovery*, 2022, **21**, 899–914.

227 A. Mayr, G. Klambauer, T. Unterthiner, M. Steijaert, J. K. Wegner, H. Ceulemans, D.-A. Clevert and S. Hochreiter, *Chem. Sci.*, 2018, **9**, 5441–5451.

228 K. Huang, C. Xiao, L. M. Glass and J. Sun, *Bioinformatics*, 2020, **37**, 830–836.

229 M. Eisenstein, *Nat. Biotechnol.*, 2024, **42**, 1323–1325.

230 C. Cai, S. Wang, Y. Xu, W. Zhang, K. Tang, Q. Ouyang, L. Lai and J. Pei, *J. Med. Chem.*, 2020, **63**, 8683–8694.

231 E. P. Tysinger, B. K. Rai and A. V. Sinit斯基, *J. Chem. Inf. Model.*, 2023, **63**, 1734–1744.

232 V. Chenthamarakshan, S. C. Hoffman, C. D. Owen, P. Lukacik, C. Strain-Damerell, D. Fearon, T. R. Malla, A. Tumber, C. J. Schofield, H. M. Duyvesteyn, W. Dejnirattisai, L. Carrique, T. S. Walter, G. R. Sreaton, T. Matviuk, A. Mojsilovic, J. Crain, M. A. Walsh, D. I. Stuart and P. Das, *Sci. Adv.*, 2023, **9**, eadg7865.

233 J. Chang and J. C. Ye, *Nat. Commun.*, 2024, **15**, 2323.

234 M. Bordukova, N. Makarov, R. Rodriguez-Esteban, F. Schmid and M. P. Menden, *Expert Opin. Drug Discovery*, 2024, **19**, 33–42.

235 A. Sadri, *J. Med. Chem.*, 2023, **66**, 12651–12677.

236 Y.-C. Lo, S. E. Rensi, W. Torng and R. B. Altman, *Drug Discovery Today*, 2018, **23**, 1538–1546.

237 J. T. Bush, P. Pogany, S. D. Pickett, M. Barker, A. Baxter, S. Campos, A. W. Cooper, D. Hirst, G. Inglis and A. Nadin, *et al.*, *J. Med. Chem.*, 2020, **63**, 11964–11971.

238 J. Arús-Pous, D. Probst and J.-L. Reymond, *Chimia*, 2018, **72**, 70–70.

239 P. Schneider, W. P. Walters, A. T. Plowright, N. Sieroka, J. Listgarten, R. A. Goodnow Jr, J. Fisher, J. M. Jansen, J. S. Duca and T. S. Rush, *et al.*, *Nat. Rev. Drug Discovery*, 2020, **19**, 353–364.

240 Embedding AI in biology, *Nat. Methods*, 2024, **21**, 1365–1366, DOI: [10.1038/s41592-024-02391-7](https://doi.org/10.1038/s41592-024-02391-7).

241 C. Hasselgren and T. I. Oprea, *Annu. Rev. Pharmacol. Toxicol.*, 2024, **64**, 527–550.

242 A. Poghosyan, H. Astsatryan, W. Narsesian and Y. Mamasakhlisov, *Cybern. Inf. Technol.*, 2017, **17**, 68–80.

243 E. Dlinnova, S. Biryukov and V. Stegailov, *Parallel Computing: Technology Trends*, IOS Press, 2020, pp. 574–582.

244 D. Milmo, A. Hern and J. Ambrose, *Can the climate survive the insatiable energy demands of the AI arms race?*, 2024, <https://www.theguardian.com/business/article/2024/jul/04/can-the-climate-survive-the-insatiable-energy-demands-of-the-ai-arms-race>, (accessed January 2025).

245 D. Milmo, *Google's emissions climb nearly 50% in five years due to AI energy demand*, 2024, <https://www.theguardian.com/technology/article/2024/jul/02/google-ai-emissions>, (accessed January 2025).



246 E. Çam, Z. Hungerford, N. Schoch, F. Pinto Miranda and C. D. Yáñez de León, *Electricity 2024 - Analysis and forecast to 2026, International Energy Agency technical report*, 2024.

247 D. Patel, D. Nishball and J. E. Ontiveros, *AI Datacenter Energy Dilemma - Race for AI Datacenter Space*, 2024, <https://www.semanalysis.com/p/ai-datacenter-energy-dilemma-race>, (accessed January 2025).

248 L. H. Kaack, P. L. Donti, E. Strubell, G. Kamiya, F. Creutzig and D. Rolnick, *Nat. Clim. Change*, 2022, **12**, 518–527.

249 P. Li, J. Yang, M. A. Islam and S. Ren, *arXiv*, 2023, preprint, arXiv:2304.03271, DOI: [10.48550/arXiv.2304.03271](https://doi.org/10.48550/arXiv.2304.03271).

250 S. Ali, E. T. Fapi, B. Jaumard and A. Planche, *Intelligent Methods, Systems, and Applications (IMSA)*, 2024, pp. 524–529.

251 E. Flanagan and D. Raphael, *Human Geography*, 2023, **16**, 244–259.

252 A. Van Wynsberghe, *AI Ethics*, 2021, **1**, 213–218.

253 S. Robbins and A. van Wynsberghe, *Sustainability*, 2022, **14**, 4829.

254 D. Probst, *Nat. Rev. Chem.*, 2023, **7**, 227–228.

255 A. T. McNutt, F. Bisiriyu, S. Song, A. Vyas, G. R. Hutchison and D. R. Koes, *J. Chem. Inf. Model.*, 2023, **63**, 6598–6607.

256 T. Seidel, C. Permann, O. Wieder, S. M. Kohlbacher and T. Langer, *J. Chem. Inf. Model.*, 2023, **63**, 5549–5570.

257 G. Durant, F. Boyles, K. Birchall, B. Marsden and C. Deane, *bioRxiv*, 2023, preprint, DOI: [10.1101/2023.10.30.564251v1](https://doi.org/10.1101/2023.10.30.564251v1).

258 Y. Yu, S. Lu, Z. Gao, H. Zheng and G. Ke, *arXiv*, 2023, preprint, arXiv:2302.07134, DOI: [10.48550/arXiv.2302.07134](https://doi.org/10.48550/arXiv.2302.07134).

259 D. Jiang, Z. Wu, C.-Y. Hsieh, G. Chen, B. Liao, Z. Wang, C. Shen, D. Cao, J. Wu and T. Hou, *J. Cheminf.*, 2021, **13**, 1–23.

260 wwPDB consortium, *Nucleic Acids Res.*, 2018, **47**, D520–D528.

261 S. K. Burley, C. Bhikadiya, C. Bi, S. Bittrich, H. Chao, L. Chen, P. A. Craig, G. V. Crichlow, K. Dalenberg and J. M. Duarte, *et al.*, *Nucleic Acids Res.*, 2023, **51**, D488–D508.

262 S. Wankowicz, P. Beltrao, B. Cravatt, R. Dunbrack, A. Gitter, K. Lindorff-Larsen, S. Ovchinnikov, N. Polizzi, B. Shoichet and J. Fraser, *AlphaFold3 Transparency and Reproducibility*, 2024, DOI: [10.5281/zenodo.11391920](https://doi.org/10.5281/zenodo.11391920), (accessed January 2025).

263 AlphaFold3 – why did Nature publish it without its code?, *Nature*, 2024, **629**, 728, DOI: [10.1038/d41586-024-01463-0](https://doi.org/10.1038/d41586-024-01463-0).

264 M. Jaderberg, A. Stecula and P. Savy, *Rational drug design with AlphaFold 3*, 2024, <https://www.isomorphiclabs.com/articles/rational-drug-design-with-alphaFold-3>, (accessed January 2025).

265 M. Fernández Pinto, *Front. Res. Metr. Anal.*, 2020, **5**, 588331.

266 C. Rikap, *Econ. Soc.*, 2023, **52**, 110–136.

267 F. Ferrari, J. V. Dijck and A. V. D. Bosch, *Nat. Mach. Intell.*, 2023, **5**, 818–820.

268 G. S. Saidakhrarovich, S. Gulyamov, I. Rustambekov, S. Zolea, E. Juchniewicz, P. Pokhariyal and A. Rodionov, *SSRN*, 2024, preprint, DOI: [10.2139/ssrn.4826900](https://doi.org/10.2139/ssrn.4826900).

269 W. Blau, V. G. Cerf, J. Enriquez, J. S. Francisco, U. Gasser, M. L. Gray, M. Greaves, B. J. Grosz, K. H. Jamieson, G. H. Haug, J. L. Hennessy, E. Horvitz, D. I. Kaiser, A. J. London, R. Lovell-Badge, M. K. McNutt, M. Minow, T. M. Mitchell, S. Ness, S. Parthasarathy, S. Perlmutter, W. H. Press, J. M. Wing and M. Witherell, *Proc. Natl. Acad. Sci. U. S. A.*, 2024, **121**, e2407886121.

270 N. Giri, R. S. Roy and J. Cheng, *Curr. Opin. Struct. Biol.*, 2023, **79**, 102536.

271 D. Si, S. A. Moritz, J. Pfab, J. Hou, R. Cao, L. Wang, T. Wu and J. Cheng, *Sci. Rep.*, 2020, **10**, 4282.

272 X. Zhang, B. Zhang, P. L. Freddolino and Y. Zhang, *Nat. Methods*, 2022, **19**, 195–204.

273 X. Song, L. Bao, C. Feng, Q. Huang, F. Zhang, X. Gao and R. Han, *Nat. Commun.*, 2024, **15**, 5538.

274 T. Pan, C. Dun, S. Jin, M. D. Miller, A. Kyriolidis, J. Phillips and N. George, *Struct. Dyn.*, 2024, **11**, 044701.

275 E. D. Zhong, T. Bepler, B. Berger and J. H. Davis, *Nat. Methods*, 2021, **18**, 176–185.

276 D. Rosenbaum, M. Garnelo, M. Zielinski, C. Beattie, E. Clancy, A. Huber, P. Kohli, A. W. Senior, J. Jumper and C. Doersch, *et al.*, *arXiv*, 2021, preprint, arXiv:2106.14108v1, DOI: [10.48550/arXiv.2106.14108](https://doi.org/10.48550/arXiv.2106.14108).

277 A. A. Rangari, S. Das and D. Rajeswari, *2023 International Conference on Artificial Intelligence and Knowledge Discovery in Concurrent Engineering (ICECONF)*, 2023, pp. 1–6.

278 G. Muehlberger, L. Seaward, M. Terras, S. A. Oliveira, V. Bosch, M. Bryan, S. Colutto, H. Dejean, M. Diem and S. Fiel, *et al.*, *J. Doc.*, 2019, **75**, 954–976.

279 A. Choudhary, R. Rishi and S. Ahlawat, *AASRI Procedia*, 2013, **4**, 306–312.

280 Y. Zhang and A. Vitalis, *Patterns*, 2025, **6**, 101147.

281 G. T. Montelione and Y. J. Huang, *bioRxiv*, 2024, preprint, DOI: [10.1101/2024.06.26.600902v1](https://doi.org/10.1101/2024.06.26.600902v1).

282 E. H. Li, L. E. Spaman, R. Tejero, Y. J. Huang, T. A. Ramelot, K. J. Fraga, J. H. Prestegard, M. A. Kennedy and G. T. Montelione, *J. Magn. Reson.*, 2023, **352**, 107481.

283 H. W. Kim, C. Zhang, R. Reher, M. Wang, K. L. Alexander, L.-F. Nothias, Y. K. Han, H. Shin, K. Y. Lee and K. H. Lee, *et al.*, *J. Cheminf.*, 2023, **15**, 71.

284 K. A. Dill and J. L. MacCallum, *Science*, 2012, **338**, 1042–1046.

285 J. Jiménez-Luna, F. Grisoni and G. Schneider, *Nat. Mach. Intell.*, 2020, **2**, 573–584.

286 I. Ponzoni, J. A. Páez Prosper and N. E. Campillo, *WIREs Comput. Mol. Sci.*, 2023, **13**, e1681.

287 H. Levine and Y. Tu, *Proc. Natl. Acad. Sci. U. S. A.*, 2024, **121**, e2403580121.

288 F. Nai, R. Nachawati, F. Zálešák, X. Wang, Y. Li and A. Caflisch, *ACS Med. Chem. Lett.*, 2022, **13**, 1500–1509.

289 R. K. Bedi, D. Huang, Y. Li and A. Caflisch, *ACS Bio Med Chem Au*, 2023, **3**, 359–370.

290 F. Zálešák, F. Nai, M. Herok, E. Bochenkova, R. K. Bedi, Y. Li, F. Errani and A. Caflisch, *J. Med. Chem.*, 2024, **67**, 9516–9535.

291 C. Prunkl, *Nat. Methods*, 2024, **21**, 1407–1408.

292 A. Gangwal, A. Ansari, I. Ahmad, A. K. Azad and W. M. A. Wan Sulaiman, *Comput. Biol. Med.*, 2024, **179**, 108734.

293 F. Kretschmer, J. Seipp, M. Ludwig, G. W. Klau and S. Böcker, *bioRxiv*, 2024, preprint, DOI: [10.101/2023.03.27.534311v2](https://doi.org/10.101/2023.03.27.534311v2).

294 A. Croitoru, A. Kumar, J. Lambry, J. Lee, S. Sharif, W. Yu, A. D. MacKerell and A. Aleksandrov, *ChemRxiv*, 2025, preprint, DOI: [10.26434/chemrxiv-2025-zkg2q](https://doi.org/10.26434/chemrxiv-2025-zkg2q).

