Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Quantum mechanical-based strategies in drug discovery: Finding the pace to new challenges in drug design

Tiziana Ginex[1], Javier Vázquez[1,2,3], Carolina Estarellas[2,4] and F. Javier Luque[2,3,4]

Abstract

The expansion of the chemical space to tangible libraries containing billions of synthesizable molecules opens exciting opportunities for drug discovery, but also challenges the power of computer-aided drug design to prioritize the best candidates. This directly hits quantum mechanics (QM) methods, which provide chemically accurate properties, but subject to small-sized systems. Preserving accuracy while optimizing the computational cost is at the heart of many efforts to develop high-quality, efficient QM-based strategies, reflected in refined algorithms and computational approaches. The design of QM-tailored physics-based force fields and the coupling of QM with machine learning, in conjunction with the computing performance of supercomputing resources, will enhance the ability to use these methods in drug discovery. The challenge is formidable, but we will undoubtedly see impressive advances that will define a new era.

Addresses

[1] Pharmacelera, Parc Científic de Barcelona (PCB), Baldiri Reixac 4-8, 08028 Barcelona, Spain
[2] Departament de Nutrició, Ciències de l'Alimentació i Gastronomia, Universitat de Barcelona, Institut de Biomedicina (IBUB), 08921 Santa Coloma de Gramenet, Spain
[3] Institut de Biomedicina (IBUB), 08921 Santa Coloma de Gramenet, Spain
[4] Institut de Química Teòrica i Computacional (IQTCUB), 08921 Santa Coloma de Gramenet, Spain

Corresponding author: Luque, F. Javier (fjluque@ub.edu)

Current Opinion in Structural Biology 2024, **87**:102870

This review comes from a themed issue on **New Concepts in Drug Discovery (2024)**

Edited by **Andrea Cavalli** and **Alessio Ciulli**

For complete overview of the section, please refer the article collection - New Concepts in Drug Discovery (2024)

Available online 23 June 2024

https://doi.org/10.1016/j.sbi.2024.102870

Introduction

Electronic structure methods provide a robust, versatile framework to explore chemical reactivity in biomolecular systems. The quantum mechanical (QM) cluster and the multiscale quantum mechanics/molecular mechanics (QM/MM) approaches have disclosed the factors that underlie the catalytic efficiency of enzymes [1–4]. Beyond the reaction mechanisms, applications involve a variety of problems, such as the design of covalent inhibitors, the enzymatic and light-induced activation of prodrugs, and the engineering of enzymes [5,6].

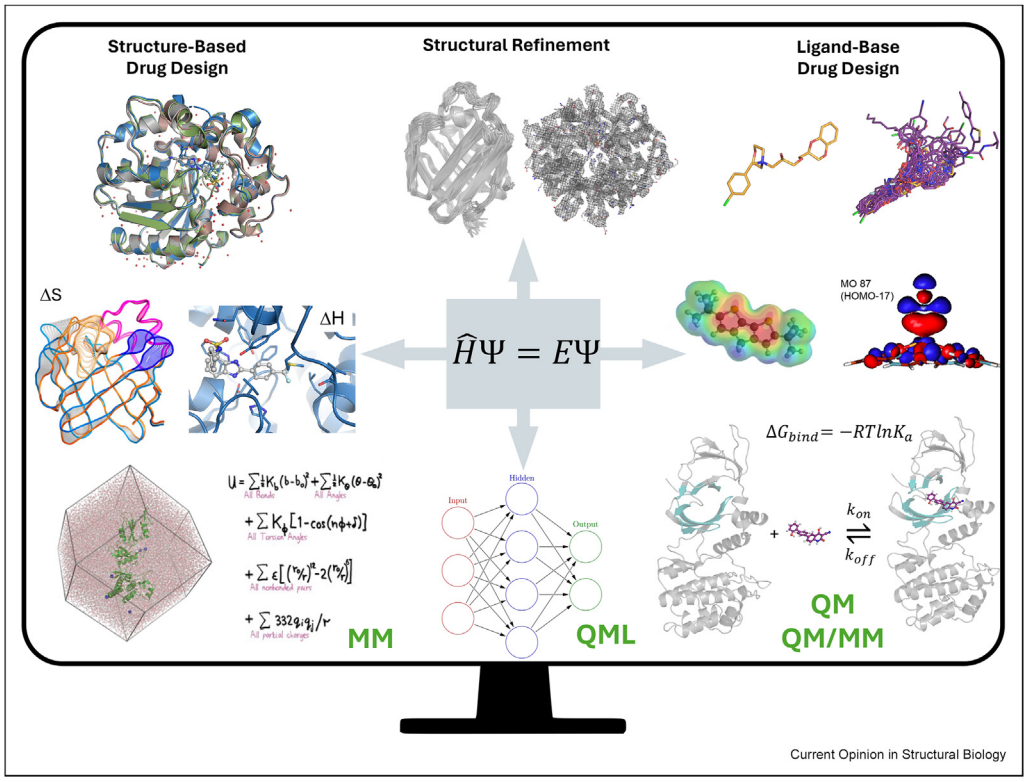
QM methods have also permeated drug discovery over the past decades. While the computational cost limited the early application to small molecules, the development of novel algorithms, including the combination with machine learning (ML), and the advances in computer performance enabled the blooming of QM-based strategies in structure-based drug design. The reader is addressed to the review by Manathunga et al. [7] for a description of methodological refinements in optimized QM/MM methods and the emergence of quantum machine learning (QML). Similarly, a critical analysis of the merits of QM/MM techniques relative to traditional computer-aided drug design methods has been recently reported [8].

Our aim here is to summarize relevant advances in the past couple of years in the application of QM-based methods in areas related to the characterization of the bioactive species, the structure-guided *hit*-to-*lead* optimization, and the identification of molecular features of bioactivity (Figure 1).

Conformational sampling of ligands

Assessment of torsional preferences of small molecules and identification of low-energy rotamers are valuable to characterize the physicochemical properties of drug-like compounds and rationalize their activities [9]. Likewise, generation of accessible conformer ensembles permits to identify the bioactive species, to define pharmacophore models and assist the choice of templates in 3D ligand-based screenings [10,11]. Furthermore, the analysis of conformer populations in distinct solvent environments is relevant to predict the solubility and permeability of 'chameleonic' compounds categorized beyond the rule-of-5 [12].

Figure 1



Schematic representation of the permeation of QM-based strategies in computer-aided drug design promoted by the development of QM semiempirical and multiscale QM/MM methods, QM-tailored physics-based force fields, and the integration between QM and ML techniques.

Despite the efforts in force field refinements, the accuracy of MM methods is still challenged by the functional form of the torsional energy and incomplete parameterization [13]. This limits the suitability of MM methods to explore the conformational space of compounds, especially with novel chemical motifs [14,15]. Nevertheless, the development of high-quality force fields with a broad coverage of chemical motifs is an active area of research, as noted in the development of the XFF force field, which was trained with a large set of functional fragments and QM energies [16]. In this context, QM methods emerge as a suitable strategy to predict the conformational preferences of (bio)organic molecules. This is illustrated by the systematic conformational analysis performed by Rulisek et al. for di- and tripeptides by combining DFT and implicit solvation calculations to discern the propensities to adopt secondary structural elements in proteins [17,18]. Furthermore, an additional example is the characterization by means of QM/MM calculations of the atropisomerism of PI3Kδ inhibitors [19].

A key achievement has been the development of dedicated software with greater computational efficiency to explore the conformational sampling of ligands. This is exemplified with the open-source conformer-rotamer ensemble sampling (CREST) program, which enables the automated conformer exploration of drug-like molecules at the extended tight-binding (xTB) level, including calculation of configurational entropy and solvation effects [20,21]. Boz and Stein used this approach to examine the impact of conformational flexibility in the binding of ligands to macrocyclic synthetic receptors [22]. The binding-free energy calculated using the automated QM sampling with CREST led to a mean error of ∼3 kcal/mol relative to the experimental data. The efficacy is supported by the comparison with the results obtained from meta-GGA PW6B95 calculations with the COSMO-RS solvation model at TPSS-D3(COSMO) optimized geometries, which showed a deviation of ∼2 kcal/mol.

Endeavors have also been conducted to implement ML tools trained with accurate QM data. In this regard, Liu et al. [23] developed the Python-based Auto3D package for generating the stereoisomeric conformational space of small compounds. Auto3D is based on the ANI-2xt neural network model, which was trained to mimic the energy and force output from DFT calculations using a tautomer-rich dataset. Likewise, TorsionNet illustrates the efforts to develop a fast and accurate ML-based assessment of small-molecule dihedral energetics [24].

It is a deep neural network model designed to predict torsion energy profiles with QM-level accuracy. To this end, a dataset of 1.2 million DFT energies were compiled for torsional scans of chemical fragments. They were used in the training of the model, leading to a root-mean square deviation of 1.3 kcal/mol relative to the DFT energies. TorsionNet is expected to capture efficiently the high-dimensional potential energy surface of flexible molecules, although the performance may degrade for molecules containing novel chemistries that are not adequately represented in the training set. On the other hand, a distinct strategy to estimate the internal energy of molecules is the QDπ model [25]. This method combines a QM description that exploits a fast third-order self-consistent density-functional tight-binding (DFTB3/3OB) with a deep-learning potential trained against reference data computed at the ωB97X/6-31G* level. Besides providing a reasonable description of the conformational potential energy landscape, this method model is also suitable for describing changes in the charge, protonation, and spin states in a size-consistent manner.

Finally, Das et al. [26] have reported an *ab initio* workflow for the prediction of collisional cross section (CCS) data obtained from ion mobility-mass spectrometry studies of small molecules. The workflow incrementally refines generated structures by combining MM, deep learning potentials, conformational clustering and QM calculations. Recently, Das et al. [27] have compared the conformational space generated by Auto3D, CREST and force field-based approaches (Ballon and ETKDG). While Auto3D was found to exhibit a better performance in global minima identification, CREST showed higher accuracy in predicting the CCS data.

Refinement of the experimental binding pose

The definition of the ligand pose in the binding pocket is crucial for the success of medicinal chemistry studies. In this regard, QM and QM/MM are well suited to assist the ligand's geometry refinement under the electrostatic field created from the protein environment, improving the description of internal parameters and intermolecular distances. For the sake of brevity, we limit ourselves to cite as illustrative examples the identification of the productive orientation of the *O*-arylcarbamate URB524 inhibitor of fatty acid amide hydrolase, as QM/MM computations enabled the discrimination between two putative binding poses [28], and the elucidation of the mechanism leading to the formation of the covalent adduct of dipeptidyl nitriles with cruzain [29].

The usage of QM and QM/MM methods as a powerful strategy to assist the structure refinement through the comparison of experimental and calculated properties, thus enabling the distinction between native poses from decoys, has been recently exploited in distinct studies.

Conventional crystallographic refinement of protein–ligand complexes uses stereochemistry restraints coupled with energy functional to ensure the correct geometry of the ligand within the experimental X-ray density. Borbulevych et al. [30] adopted a QM/MM scheme to assist the refinement protocol of the curated set of structures included in the community structure activity resource, paying attention to the correct assignment of tautomer/protomer states. More recently, the Quantum Mechanical Restraints (QMR) procedure has been proposed to optimize the ligand geometry *in situ* during the macromolecular crystallographic structure refinement [31]. Accordingly, the ligand geometry is determined in the binding pocket considering the effect of neighboring protein residues and waters on the local energy minima of the ligand. The refinement of more than 1700 protein–ligand models showed that the QMR-restrained parameters lead to geometries that fit better to the protein environment compared to the use of conventionally generated restraints, providing accurate torsion restraints for flexible ligands.

The automated fragmentation QM/MM (AF-QM/MM) method has been shown to be an efficient platform for protein–ligand binding structure prediction based on NMR data [32]. In this approach the ligand and/or the protein binding pocket are automatically divided into capped fragments for DFT calculations of NMR chemical shifts, including the electrostatic polarization effect from the solvent. The perturbation of the ^{1}H chemical shifts triggered upon ligand binding is then used to complement the docking score. The AF-QM/MM has also been used to develop the empirical HECSP method, which provides a fast, accurate tool to predict the chemical shift perturbation of protein protons induced by ligand binding [33]. Finally, the extension to ligands bound to membrane proteins has recently been studied by Zhang et al. [34], considering the effect of the local environment (cross-protomer interactions, phospholipid composition, and solvent) on the ^{15}N and ^{13}C chemical shifts of retinal bound to the light-driven sodium pump rhodopsin 2.

On the other hand, comparison of the experimental chemical shifts with those computed from the X-ray structure using QM/MM has been valuable to highlight disagreements for parts of the ligand between the two experimental techniques. The discrepancies observed for protons involved in hydrogen bonding interactions support the inadequacy of considering a single structure to represent the solution state of the protein-ligand complex. In this regard, Platzer et al. have shown that QM/MM molecular dynamics ensembles can be used to refine the initial X-ray co-crystal structure, resulting in a

better agreement with experimental ligand chemical shift values [35].

Prediction of binding affinity

One of the largest challenges in drug discovery is the accurate calculation of the binding free energy between ligand and target. Since the binding affinity is a cornerstone in structure-guided drug design, there is continued interest in developing robust, accurate physics-based models for application in drug discovery. Different QM methods have been employed, ranging from semiempirical QM calculations, *via* DFT, to strict coupled-cluster calculations [36].

Cavasotto and Aucar [37] showed that the choice of a QM-based scoring function that combines PM7 calculations for ligand-protein interactions with changes in the solvation and entropy components led to much higher enrichment than a traditional docking method in the high-throughput docking against 10 distinct targets. Recently, Pecina et al. have reported SQM2.20 [38], which is a semiempirical QM-based scoring function that extents the previous SQM/COSMO tool [39]. The SQM2.20 score combines the gas-phase interaction energy determined at the PM6-D3H4X level, the change of solvation free energy upon complex formation, the conformational contribution of the ligand to the binding, the free energy of proton transfer between ligand and buffer, and the loss of ligand conformational entropy upon binding. The score is evaluated on a model of $\sim$2000 atoms, comprising residues within 10 Å around all the overlaid ligands in each target protein, with an average time of $\sim$20 min per protein−ligand complex.

Several alternative strategies have been proposed within the framework of the linear response approximation. By exploiting the '"divide-and-conquer' fragment molecular orbital (FMO) method, SophosQM [40] computes the interaction between ligand and target from the addition of the pairwise interaction energy between ligand and fragments, whereas nonenthalpic components are captured through the octanol/water partition coefficient (logP). Thus, the binding affinity is expressed as a multilinear regression based on the interaction energy and logP, where the coefficients are derived from the fitting to known binding affinities of ligands. The method assumes the preservation of the binding mode for ligands interacting at the same target, excluding the occurrence of large conformational rearrangements in the pocket. The linear fitting of the FMO-determined interaction energy, deformation energy of the ligand, and solvation free energy also underlies the development of FMOScore, which was applied to the lead optimization against Src homology-2-containing protein tyrosine phosphatase 2 [41]. Finally, the QMH-L approach exploits a single target−ligand complex structure pre-optimized at the MM level, taking into account the interfacial water molecules between ligand and target [42]. Then, the binding free energy is estimated combining i) the QM interaction energy between the fragments of the target, ligand and interfacial waters, and ii) a descriptor of the normalized molecular size of the ligand, which was introduced as a measure of the entropic term. The results determined for protein complexes bound to small molecules and peptides support the balance between accuracy, applicability, and computational cost.

QM/MM calculations can also be used to correct the binding free energy determined from classical simulations using the 'so-called' "indirect" approach, in which a QM/MM "correction" is introduced at the classical endpoints of interest. This strategy reduces the computational expensiveness that would require a complete QM/MM free energy simulation and corrects deficiencies of the classical description, provided that there is sufficient configurational overlap between the MM and QM/MM levels of theory. In this regard, a recent study has examined the suitability of force-matching procedures to predict the binding free energy of several narcotics within the SAMPL blind challenge [43].

The expensiveness of QM and QM/MM calculations has stimulated the search of alternative strategies that exploit the development of i) tailored force fields and ii) ML-based methods for predicting the binding affinity of ligands.

The former category is illustrated by the ARROW force field [44], which was conceived to mitigate the deficiencies of standard force fields to provide an accurate description of electrostatic and exchange interactions, non-additivity effects of atomic interactions, and low transferability of parameters. ARROW is an advanced physics-based model that includes multipolar electrostatics and anisotropic polarization, with parameters fitted exclusively to high-level QM data. Indeed, MP2/aug-cc-pVQZ calculations were performed for monomeric model compounds, and MP2/complete basis set were used for dimers with amino acid fragments and water. Furthermore, nonbonded interactions include electrostatic and exchange-repulsion, which are treated via multipolar expansion, dispersion terms, which are described via spherical (C6 and C8) expressions, and many-body effects described with anisotropic-induced dipoles. A refined version of ARROW includes a short-range neural network correction to the intermolecular interaction energy has been recently reported [45].

A distinct strategy to improve the quality of the force field is represented by the work of González et al. [46] and Macaya et al. [47] regarding the implementation of the Minimal Basis Iterative Stockholder method. In this approach the aim is to derive atomic charges and Lennard-Jones parameters from the polarized electron

density of the ligand in conjunction using various configurations of the bound and unbound state.

Rizzi et al. [48] have reported the implementation of the targeted free energy perturbation to calculate the binding free energy at a target high level (QM) of theory from a cheap reference potential (MM). To this end, a neural network is used to train a mapping function that enhances the overlap between target and reference distributions in conjunction with a limited number of QM calculations of energy and gradients. The application to binding free energy of 22 ligands led to close agreement with the benchmark results.

Another example of the QM-guided efforts to develop ML-based methods is the work by Isert et al. [49], where the critical points (CPs) of the electron density are used as a fundamental representation of the intermolecular interactions formed between ligand and protein. Thus, the ML model was trained using the CPs determined at the GFN2-xTB level in conjunction with a linearized Poisson–Boltzmann water model. The results showed a modest predictive performance relative to benchmark methods, which may reflect limitations due to the QM method, the CP-centered representations, and the inability to account for other contributions, such as ligand strain and entropy.

Beyond the binding affinity, it is also worth emphasizing the efforts in predicting the dissociation rate constant (k_{off}) of the ligand, since the residence time of the ligand-target complex is a key parameter to analyze the drug activity [50]. A review of recent progresses in molecular simulation methodologies for the prediction of k_{off} has been reported by Ahmad et al. [51], paying attention to the use of massively parallel DFT-QM/MM calculations, often complemented by ML techniques, to account for electronic polarizability and charge transfer effects in the unbinding kinetics constant. For our purposes here, we limit ourselves to the work by Ojha et al. [52], who implemented QMrebind in the framework of the Simulation Enabled Estimation of Kinetic Rates v.2 (SEEKR2) method, which exploits MD simulations to predict the unbinding kinetics. QMrebind incorporates QM methods into generating system-specific force fields through re-parameterizing ligand partial charges in the bound state. The results obtained for host–guest complexes and Hsp90-inhibitor complexes highlight the potential of the improved force field parameters to enhance the simulation accuracy.

Molecular determinants of bioactivity

Understanding the structural and chemical features that determine the bioactivity profile of a ligand, including both pharmacodynamics and pharmacokinetics, is fundamental to assist the rational design of drug-like candidates.

QM permits to examine the contribution of specific interactions that cannot be properly assessed by MM force fields, such as halogen-bonds [53,54], polarized CH/n, CH/π, and XH/π interactions [55], and peptide amide-π or alkene-π stacking contacts [56,57]. The limitations of MM methods are also reflected in the binding to metal centers [58]. Similarly, QM-derived properties, such as reactivity indexes of covalent inhibitors, have been used in structure-activity relationships studies [59].

Efforts have also been devoted to address the inaccuracies of molecular descriptors, thus complementing classical approaches in similarity-guided, pharmacophore-based and docking screenings [60]. In this context, using similarity-based approaches, COSMO-RS σ-profiles derived from the QM(B3LYP/DNP)/MM-optimized conformation of remdesivir were used in the screening of analogues targeting the SARS-CoV2 RNA-dependent RNA polymerase [61]. Recently, a virtual screening campaign targeting the soluble epoxide hydrolase enzyme showed how 3D-hydrophobic atomic descriptors derived from QM continuum solvation calculations played a major role in the prioritization of the top-ranked compounds in docking screenings [62,63]. On the other hand, QM FMO calculations have been used to identify hotspots at the interface between two interacting proteins, thus disclosing clues for the design of protein-protein modulators [64].

QM cluster GFN2-xTB calculations have also been used to examine the suitability of bioisosteric replacements, which can be used to tune properties related to activity, bioavailability, and metabolism of ligands [65]. Finally, QM-based descriptors derived from COSMO-RS have been used to explore bioisosteric replacements on the passive permeability for a series of protein arginine methyl transferase 5 inhibitors [66]. Interestingly, the results highlighted the hydrogen-bond acceptor propensity as a key parameter to optimize the transport properties of these compounds.

Although ML may be instrumental in many areas of drug design, the current challenge lies in the absence of curated comprehensive collections of data. To alleviate this limitation, the open-access -mechanical properties of drug-like molecules (Q-Mugs) dataset [67] has been developed to include QM (GFN2-xTB) properties of more than 665,000 molecules extracted from the ChEMBL database. In parallel, Merck & Co. proposed QM9-extended [68], a dataset comprising 153,716 compounds commonly found in drug discovery focusing on QM (B3LYP/6-31G(2df,p)) properties for ADME predictions. Similarly, Isert et al. [69] established an *in-house* training set to develop a ML QM-based protocol for predicting the octanol/water partition coefficient. Among the algorithms tested, message passing neural

network model Chemprop [70] emerged as the best model with mean absolute errors of 0.34 log units.

Conclusions

High-throughput screening is the dominant strategy to identify novel small molecules that modulate the biological activity of selected targets [71]. The explosive growth of the available chemical space, which reaches billions of molecules, opens opportunities for finding hits with novel chemical scaffolds, less biased toward *bio-like* molecules, and better fitting to the target binding site [72]. However, this also challenges the capability of computational approaches for prioritizing the best hits, avoiding the occurrence of biasing artifacts in the screening of chemical libraries [73].

In this scenario, the success of QM-based strategies in drug design will be determined by the development in both algorithms and hardware. Current efforts reflect three major tendencies, which involve (i) the development of accurate, but computationally effective semi-empirical methods as well as the calibration of multiscale QM/MM methods, (ii) the redefinition of QM-tailored physics-based force-fields suitably refined to provide an accurate description of the complex network of intermolecular interactions, and (iii) the generation of QM-assisted ML models. Regarding hardware, the advent of novel supercomputing architectures [74,75] and especially quantum computing [76,77], although requiring the implementation of current codes, will pave the way for a significant leap in quantum chemistry simulations. Overall, the flourishing of algorithmic progresses in the past years in conjunction with the enhanced resources provided by novel hardware platforms suggests that we will witness an exciting evolution of quantum chemistry in drug discovery.

Author contributions

All authors contributed to conceptualization, writing - original draft, and writing - review and editing of the final manuscript.

Funding

This work was supported by the Spanish Ministerio de Ciencia e Innovación [grants PID2020-117646RB-I00 and PID2022-142623OA-I00, MICIN/AEI/10.13039/501100011033; Maria de Maetzu (grant CEX2021-001202-M], and Agència de Gestió d'Ajuts Universitaris i de Recerca [grant 2021SGR00671].

Declaration of competing interest

TG and JV are employees of Pharmacelera S.L. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Data availability

No data was used for the research described in the article.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Salahub DR: **Multiscale molecular modelling: from electronic structure to dynamics of nanosystems and beyond**. *Phys Chem Chem Phys* 2022, **24**:9051–9081, https://doi.org/10.1039/d1cp05928a.

2. Sheng X, Himo F: **The quantum chemical cluster approach in biocatalysis**. *Acc Chem Res* 2023, **56**:938–947, https://doi.org/10.1021/acs.accounts.2c00795.

3. Clemente CM, Capece L, Martí MA: **Best Practices on QM/MM simulations of biological systems**. *J Chem Inf Model* 2023, **63**:2609–2627, https://doi.org/10.1021/acs.jcim.2c01522.

4. Ferreira P, Fernandes PA, Ramos MJ: **Modern computational methods for rational enzyme engineering**. *Chem Catal* 2022, **2**:2481–2498.

5. Ray D, Das S, Raucci U: **Kinetic view of enzyme catalysis from enhanced sampling QM/MM simulations**. *J Chem Inf Model* 2024, **64**:3953–3958, https://doi.org/10.1021/acs.jcim.4c00475.

6. Amaro R, Mulholland AJ: **Multiscale methods in drug design bridge chemical and biological complexity in the search for cures**. *Nat Rev Chem* 2018, **2**:148, https://doi.org/10.1038/s41570-018-0148.

7. * Manathunga M, Götz AW, Merz Jr KM: **Computer-aided drug design, quantum-mechanical methods for biological problem**. *Curr Opin Struct Biol* 2022, **75**, 102417, https://doi.org/10.1016/j.sbi.2022.102417.
This paper reports a detailed review of the advances in the development of accurate, efficient QM methods for computer-aided drug design, emphasizing the methodological progresses in multiscale QM/MM methods and the potential impact of combining QM and ML models.

8. Kar RK: **Benefits of hybrid QM/MM over traditional classical mechanics in pharmaceutical systems**. *Drug Discov Today* 2023, **28**:1–9, https://doi.org/10.1016/j.drudis.2022.103374.

9. Hawkins PCD: **Conformation generation: the state of the art**. *J Chem Inf Model* 2017, **57**:1747–1756, https://doi.org/10.1021/acs.jcim.7b00221.

10. McNutt AT, Bisiriyu F, Song S, Vyas A, Hutchinson GR, Koes DR: **Conformer generation for structure-based drug design: how many and how good?** *J Chem Inf Model* 2023, **63**:6598–6607, https://doi.org/10.1021/acs.jcim.3c01245.

11. Vázquez J, García R, Linares P, Luque FJ, Herrero E: **On the relevance of query definition in the performance of 3D ligand-based virtual screening**. *J Comput Aided Mol Des* 2024, **38**:18, https://doi.org/10.1007/s10822-024-00561-5.

12. Poongavanam V, Wleske LHE, Pelntner S, Erdélyl M, Kihlberg J: **Molecular chameleons in drug discovery**. *Nat Rev Chem* 2024, **8**:45–60, https://doi.org/10.1038/s41570-023-00563-1.

13. Raniolo S, Limongelli V: **Improving small-molecule force field parameters in ligand binding studies**. *Front Mol Biosci* 2021, **8**, 760283, https://doi.org/10.3389/fmolb.2021.760283.

14. Gürsoy O, Smiesko M: **Searching for bioactive conformations of drug-like ligands with current force fields: how good are we?** *J Cheminf* 2017, **9**:29, https://doi.org/10.1186/s13321-017-0216-0.

15. Stylianakis I, Zervos N, Lii J-H, Pantazis DA, Kolocouris A: **Conformational energies of reference organic molecules: benchmarking of common efficient computational methods against couple cluster theory**. *J Comput Aided Mol Des* 2023, **37**:607–656, https://doi.org/10.1007/s10822-023-00513-5.

16. Xue B, Yang Q, Zhang Q, Wan X, Fang D, Lin X, Sun G, Gobbo G, Cao F, Mathiowetz AM, Burke BJ, Kumpf RA, Rai BK, Wood GPF, Pickard FCIV, Wang J, Zhang P, Ma J, Jiang YA, Wen S, Hou X, Zou J, Yang M: **Development and comprehensive benchmark of a high-quality AMBER-consistent small molecule force field with broad chemical space coverage for molecular modeling and free energy calculation**. *J Chem Theor Comput* 2024, **20**:799–818, https://doi.org/10.1021/acs.jctc.3c00920.

17. Kaldova T, Culka M, Rulisek L, Andris E: **Exhaustive mapping of the conformational space of natural dipeptides by the DFT-D3//COSMO-RS method**. *J Phys Chem B* 2022, **126**:5949–5958, https://doi.org/10.1021/acs.jpcb.2c02861.

18. Osifová Z, Kalvoda T, Galgonek J, Culka M, Vondrášek J, Bouř P, Bednárová L, Andrushchenko V, Dračínský M, Rulíšek L: **What are the minimal folding seeds in proteins? Experimental and theoretical assessment of secondary structure propensities of small peptide fragments**. *Chem Sci* 2023, **15**:594–608, https://doi.org/10.1039/d3sc04960d.

19. Lodola A, Bertolini S, Biagetti M, Capacchi S, Facchinetti F, Gallo PM, Pappani A, Mor M, Pala D, Rivara S, Visentini F, Corsi M, Capelli AM: **Atropisomerism and conformational equilibria: impact on PI3Kδ inhibition of 2-((6-amino-9*H*-purin-9-yl)methyl)-5-methyl- 3-(*o*-tolyl)quinazolin-4(3*H*)-one (IC87114) and its conformationally restricted analogs**. *J Med Chem* 2017, **60**:4304–4315, https://doi.org/10.1021/acs.jmedchem.7b00247.

20. Pracht P, Grimme S, Bannwarth C, Bohle F, Ehlert S, Feldmann G, Gorges J, Müller M, Neudecker T, Plett C, Spicher S, Steinbach P, Wesolowski PA, Zeller F: **Crest - a program for the exploration of low-energy molecular chemical space**. *J Chem Phys* 2024, **160**, 114110, https://doi.org/10.1063/5.0197592.

21. Pracht P, Bohle F, Grimme S: **Automated exploration of the low-energy chemical space with fast quantum chemical methods**. *Phys Chem Chem Phys* 2020, **22**:7169–7192, https://doi.org/10.1039/C9CP06869D.

22. Boz E, Stein M: **Accurate receptor-ligand binding free energies from fast QM conformational chemical space sampling**. *Int J Mol Sci* 2021, **22**:3078, https://doi.org/10.3390/ijms22063078.

23. ** Liu Z, Zubatiuk T, Roitberg A, Isayev O: **Auto3D: automatic generation of the low-energy 3D structures with ANI neural network potentials**. *J Chem Inf Model* 2022, **62**:5373–5382, https://doi.org/10.1021/acs.jcim.2c00817.
Conceived as an extension of the ANI model, ANI-2xt defines the algorithm encoded in Auto3D, which was developed as an efficient tool to explore the conformational space of flexible compounds and to identify the stereoconfiguration that yields the lowest-energy conformer.

24. ** Rai BK, Sresht V, Yang Q, Unwalla R, Tu M, Mathiowetz AM, Bakken GA: **TorsionNet: a deep neural network to rapidly predict small-molecule torsional energy profiles with the accuracy of quantum mechanics**. *J Chem Inf Model* 2022, **62**:785–800, https://doi.org/10.1021/acs.jcim.1c01346.
This manuscript reports the development of a neural network model to explore the conformational space of small molecules, where the training procedure was focused on the prediction of the torsional energy rather than learning a general potential energy function.

25. Zeng J, Tao Y, Giese TJ, York DM: **QDπ: a quantum deep potential interaction model for drug discovery**. *J Chem Theor Comput* 2023, **19**:1261–1275, https://doi.org/10.1021/acs.jctc.2c01172.

26. Das S, Dinpazhoh L, Tanemura KA, Merz Jr KM: **Rapid and automated ab initio metabolite collisional cross section prediction from SMILES input**. *J Chem Inf Model* 2023, **63**:4995–5000, https://doi.org/10.1021/acs.jcim.3c00890.

27. Das S, Merz Jr KM: **Molecular gas-phase conformational ensembles**. *J Chem Inf Model* 2024, **64**:749–760, https://doi.org/10.1021/acs.jcim.3c01309.

28. Lodola A, Mor M, Rivara S, Christov C, Tarzia G, Piomelli D, Mulholland AJ: **Identification of productive inhibitor binding orientation in fatty acid amide hydrolase (FAAH) by QM/MM mechanistic modelling**. *Chem Comm* 2008:214–216, https://doi.org/10.1039/b714136j.

29. Dos Santos AM, Cianni L, De Vita D, Rosini F, Leitao A, Laughton CA, Lameira J, Montanari CA: **Experimental study and computational modelling of cruzain cysteina protease inhibition by dipeptidyl nitriles**. *Phys Chem Chem Phys* 2018, **20**:24317–24328, https://doi.org/10.1039/c8cp0332j.

30. Borbulevych OY, Martin RI, Westerhoff LM: **The critical role of QM/MM X-ray refinement and accurate tautomer/protomer determination in structure-based drug design**. *J Comput Aided Mol Des* 2021, **35**:433–451, https://doi.org/10.1007/s10822-020-00354-6.

31. Liebschner D, Moriarty NW, Poon BK, Adams PD: **In situ ligand restraints from quantum-mechanical methods**. *Acta Crystallogr* 2023, **D79**:100–110, https://doi.org/10.1107/S2059798323000025.

32. Xinsheng J, Zhu T, Zhang JZH, He X: **Automated fragmentation QM/MM calculation of NMR chemical shifts for protein-ligand complexes**. *Front Chem* 2018, **6**:150, https://doi.org/10.3389/fchem.2018.00150.

33. Yu Z, Li P, Merz Jr KM: **Using ligand-induced protein chemical shift perturbations to determine protein-ligand structures**. *Biochemistry* 2017, **56**:2349–2362, https://doi.org/10.1021/acs.biochem.7b00170.

34. ** Zhang J, Kriebel CN, Wan Z, Shi M, Glaubitz C, He X: **Automated fragmentation quantum mechanical calculation of ^{15}N and ^{13}C chemical shifts in a membrane protein**. *J Chem Inf Model* 2023, **19**:7405–7422, https://doi.org/10.1021/acs.jctc.3c00621.
This manuscript reports the adaptation of the AF-QM/MM method to calculate the chemical shifts of membrane-embedded proteins accounting for the complexity of the environment, thus assisting the structure validation and refinement of large biomolecules.

35. Platzer G, Ptaszek AL, Böttcher J, Fuchs JE, Geist L, Braun D, McConnell DB, Konrat R, Sánchez-Murcia PA, Mayer M: **Ligand ^{1}H NMR chemical shifts as accurate reporters for protein-ligand binding interfaces in solution**. *ChemPhysChem* 2024, **25**, e202300636, https://doi.org/10.1002/cphc.202300636.

36. Ryde U, Söderhjelm P: **Ligand-binding affinity estimates supported by quantum-mechanical methods**. *Chem Rev* 2016, **116**:5520–5566, https://doi.org/10.1021/acs.chemrev.5b00630.

37. Cavasotto CN, Aucar MG: **High-throughput docking using quantum mechanical scoring**. *Front Chem* 2020, **8**:246, https://doi.org/10.3389/fchem.2020.00246.

38. ** Pecina A, Fanfrlík J, Lepšík M, Řezáč J: **SQM2.20: semi-empirical quantum-mechanical scoring function yields DFT-quality protein–ligand binding affinity predictions in minutes**. *Nat Commun* 2024, **15**:117, https://doi.org/10.1038/s41467-024-45431-8.
This is a contemporary study that reports a semiempirical QM computational scheme SQM2.20. This method outperforms other scoring methods. Reaching a level of accuracy similar to much more expensive DFT calculations, the consistent performance for a variety of macromolecular targets supports the application to lead optimization studies.

39. Pecina A, Eyrilmez SM, Köprülüoğlu C, Miriyala VM, Lepšík M, Fanfrlík J, Řezáč J, Hobza P: **SQM/COSMO Scoring Function: reliable quantum-mechanical tool for sampling and ranking in structure-based drug design**. *Chempluschem* 2020, **85**:2362–2371, https://doi.org/10.1002/cplu.202000120.

40. Guareschi R, Lukac I, Hilbert IH, Zucotto F: **SophosQM: accurate binding affinity prediction in compound optimization**. *ACS Omega* 2023, **8**:15083–15098, https://doi.org/10.1021/acsomega.2c08132.

41. Yuan Z, Chen X, Fan S, Chang L, Chu L, Zhang Y, Wang J, Li S, Xie J, Hu J, Miao R, Zhu L, Zhao Z, Li H, Li S: **Binding free energy calculation based on the fragment molecular orbital method and its application in designing novel SHP-2 allosteric inhibitors**. *Int J Mol Sci* 2024, **25**:671, https://doi.org/10.3390/ijms25010671.

42. Szél V, Zsidó BZ, Jeszenői N, Hetényi C: **Target-ligand binding affinity from single point enthalpy calculation and elemental composition**. *Phys Chem Chem Phys* 2023, **25**:31714–31725, https://doi.org/10.1039/D3CP04483A.

43. Hudson PS, Aviat F, Meana-Pañeda R, Warrensford L, Pollard BC, Prasad S, Jones MR, Woodcock HL, Brooks BR:

Obtaining QM/MM free energies in the SAMPL8 drugs of abuse challenge: indirect approaches. *J Comput Aided Mol Des* 2022, **36**:263–277, https://doi.org/10.1007/s10822-022-00443-8.

44. ** Nawrocki G, Leontyev I, Sakipov S, Darkhovskiy M, Kurnikov I, Pereyaslavets L, Kamath G, Voronina E, Butin O, Illarianov A, Olevanov M, Kostikov A, Ivahnenko I, Patel DS, Sankaranarayanan KRS, Kurnikova MG, Lock C, Crooks GE, Levitt M, Korngerg RD, Fain B: **Protein-ligand binding free-energy calculations with ARROW - a purely first-principles parametrized polarizable force field**. *J Chem Theor Comput* 2022, **18**:7751–7763, https://doi.org/10.1021/acs.jctc.2c00930.
This manuscript reports the definition of a QM-fitted, physics-based polarizable force field aimed to provide an accurate description of intermolecular interactions, achieving almost chemical accuracy in simulations of solvation free energies and ligand-target binding free energies.

45. Illarianov A, Sakipov S, Pereyaslavets L, Kurnikov IV, Kamath G, Butin O, Voronina E, Ivahnenko I, Leontyev I, Nawrocki G, Darkhovskiy M, Olevanov M, Cherniavskyi YK, Lock C, Greenslade S, Sankaranarayanan KRS, Kurnikova MG, Potoff J, Kornberg RD, Fain B: **Combining force fields and neural networks for an accurate representation of chemically diverse molecular interactions**. *J Am Chem Soc* 2023, **145**: 23620–23629, https://doi.org/10.1021/jacs.3c07628.

46. González D, Macaya L, Castillo-Orellana C, Verstraelen T, Vogt-Geisse S, Vöhringer-Martínez E: **Nonbonded force field parameters from minimal basis iterative stockholder partitioning of the molecular electron density improve CB7 host–guest affinity predictions**. *J Chem Inf Model* 2022, **62**: 4162–4174, https://doi.org/10.1021/acs.jcim.2c00316.

47. Macaya L, González D, Vöhringer-Martínez E: **Nonbonded force field parameters from MBIS partitioning of the molecular electron density improve binding affinity predictions of the T4-lysozyme double mutant**. *J Chem Inf Model* 2024, **64**: 3269–3277, https://doi.org/10.1021/acs.jcim.3c01912.

48. Rizzi A, Carloni P, Parrinello M: **Free energies at QM accuracy from force fields via multimap targeted estimation**. *Proc Natl Acad Sci USA* 2023, **120**, e2304308120, https://doi.org/10.1073/pnas.2304308120.

49. Isert C, Atz K, Riniker S, Schneider G: **Exploring protein-ligand binding affinity prediction with electron density-based geometric deep-learning**. *RSC Adv* 2024, **14**:4492, https://doi.org/10.1039/d3ra08650j.

50. Copeland RA: **Evolution of the drug–target residence time model**. *Expet Opin Drug Discov* 2021, **16**:1441–1451, https://doi.org/10.1080/17460441.2021.1948997.

51. * Ahmad K, Rizzi A, Capelli R, Mandelli D, Lyu W, Carloni P: **Enhanced-sampling simulations for the estimation of ligand binding kinetics: current status and perspective**. *Front Mol Biosci* 2022, **9**, 899895, https://doi.org/10.3389/fmolb.2022.899805.
An interesting manuscript that reports the advances made in molecular simulation methodologies for the prediction of (un)binding kinetics. It includes a discussion about the limitations of force fields, and the potential impact of QM/MM and QML methods in the upcoming era of exascale computing.

52. Ojha AA, Votapka LW, Amaro RE: **QMrebind: incorporating quantum mechanical force field reparameterization at the ligand binding site for improved drug-target kinetics through milestoning simulations**. *Chem Sci* 2023, **14**:13159–13175, https://doi.org/10.1039/D3SC04195F.

53. Vaas S, Zimmermann MO, Schollmeyer D, Stahlecker J, Engelhardt MU, Rheinganz J, Droftleff B, Olfert M, Lämmerhofer M, Kramer M, Stehle T, Boeckler FM: **Principles and applications of CF_2X moieties as unconventional halogen bond donors in medicinal chemistry, chemical biology and drug discovery**. *J Med Chem* 2023, **66**: 10202–10225, https://doi.org/10.1021/acs.jmedchem.3c00634.

54. Li J, Zhou L, Han Z, Wu L, Zhang J, Zhu W, Xu Z: **Impact of halogen bonds on protein–peptide binding and protein structural stability revealed by computational approaches**. *J Med Chem* 2024, **67**:4782, https://doi.org/10.1021/acs.jmedchem.3c02359.

55. Xiao Y, Woods RJ: **Protein–ligand CH–π interactions: structural informatics, energy function development, and docking implementation**. *J Chem Theor Comput* 2023, **19**:5503–5515, https://doi.org/10.1021/acs.jctc.3c00300.

56. Krone MW, Travis CR, Lee GY, Eckvahl HJ, Houk KN, Waters ML: **More than π–π–π stacking: contribution of amide–π and CH–π interactions to crotonyllysine binding by the AF9 YEATS domain**. *J Am Chem Soc* 2020, **142**: 17048–17056, https://doi.org/10.1021/jacs.0c06568.

57. Togo T, Tram L, Denton LG, ElHilali-Pollard X, Gu J, Jiang J, Liu C, Zhao Y, Zhao Y, Zheng Y, Zheng Y, Yang J, Fan P, Arkin MR, Härmä H, Sun D, Canan SS, Wheeler SE, Renslo AR: **Systematic study of heteroarene stacking using a congeneric set of molecular glues for procaspase-6**. *J Med Chem* 2023, **66**:9784–9796, https://doi.org/10.1021/acs.jmedchem.3c00590.

58. Gervasoni S, Spencer J, Hinchliffe P, Pedretti A, Vairoletti F, Mahler G, Mulholland AJ: **A multiscale approach to predict the binding mode of metallo beta-lactamase inhibitors**. *Proteins* 2022, **90**:372–384, https://doi.org/10.1002/prot.26227.

59. Bonatto V, Batista PHJ, Cianni L, De Vita D, Silva DG, Cedron R, Tezuka DY, De Albuquerque S, Moraes CB, Franco CH, Lameira J, Leitao A, Montanari CA: **On the intrinsic reactivity of highly potent trypanocidal cruzain inhibitors**. *RSC Med Chem* 2020, **11**:1275–1284, https://doi.org/10.1039/d0md00097c.

60. Gorgulla C, Jayaraj A, Fackeldey K, Arthanari H: **Emerging frontiers in virtual drug discovery: from quantum mechanical methods to deep learning approaches**. *Curr Opin Chem Biol* 2022, **69**, 102156, https://doi.org/10.1016/j.cbpa.2022.102156.

61. Gusarov S, Stoyanov SR: **Cosmo-RS-based descriptors for the machine learning-enabled screening of nucleotide analogue drugs against SARS-CoV-2**. *J Phys Chem Lett* 2020, **11**: 9408–9414, https://doi.org/10.1021/acs.jpclett.0c02836.

62. Vázquez J, Deplano A, Herrero A, Gibert E, Herrero E, Luque FJ: **Assessing the performance of mixed strategies to combine lipophilic molecular similarity and docking in virtual screening**. *J Chem Inf Model* 2020, **60**:4231–4245, https://doi.org/10.1021/acs.jcim.9b01191.

63. Vázquez J, Ginex T, Herrero A, Morisseau C, Hammock BD, Luque FJ: **Screening and biological evaluation of soluble Epoxide Hydrolase inhibitors: assessing the role of hydrophobicity in the pharmacophore-guided search of novel hits**. *J Chem Inf Model* 2023, **63**:3209–3225, https://doi.org/10.1021/acs.jcim.3c00301.

64. Monteleone S, Fedorov DG, Townsend-Nicholson A, Southey M, Bodkin M, Heifetz A: **Hotspot identification and drug design of protein-protein interaction modulators using the fragment molecular orbital method**. *J Chem Inf Model* 2022, **62**: 3784–3799, https://doi.org/10.1021/acs.jcim.2c00457.

65. Losev TV, Gerasimov IS, Panova MV, Lisov AA, Abdyusheva YR, Rusina PV, Zaletskaya E, Stroganov OV, Medvedev MG, Novikov FN: **Quantum mechanical-cluster approach to solve the bioisosteric replacement problem in drug design**. *J Chem Inf Model* 2023, **63**:1239–1248, https://doi.org/10.1021/acs.jcim.2c01212.

66. Argikar U, Blatter M, Bednarczyk D, Chen Z, Cho YS, Doré M, Dumouchel JL, Ho S, Hoegenauer K, Kawanami T, Mathieu S, Meredith E, Möbitz H, Murphy SK, Parthasarathy S, Soldermann CP, Santos J, Silver S, Skolnik S, Stojanovic A: **Paradoxical increase of permeability and lipophilicity with the increasing topological polar surface area within a series of PRMT5 inhibitors**. *J Med Chem* 2022, **65**:12386–12402, https://doi.org/10.1021/acs.jmedchem.2c01068.

67. * Isert C, Atz K, Jiménez-Luna J, Schneider G: **QMugs, quantum mechanical properties of drug-like molecules**. *Sci Data* 2022, **9**:273, https://doi.org/10.1038/s41597-022-01390-7.
By compiling QM properties of more than 665,000 molecules, this manuscriptaims to facilitate the development of ML models to gain insight into the relationships between molecular structure and biological activity.

68. Lim MA, Yang S, Mai H, Cheng AC: **Exploring deep learning of quantum chemical properties for absorption, distribution, metabolism, and excretion predictions**. *J Chem Inf Model* 2022, **62**:6336–6341, https://doi.org/10.1021/acs.jcim.2c00245.

69. Isert C, Kromann JC, Stiefl N, Schneider G, Lewis RA: **Machine learning for fast, quantum mechanics-based approximation of drug lipophilicity**. *ACS Omega* 2023, **8**:2046–2056, https://doi.org/10.1021/acsomega.2c05607.

70. Heid E, Greenman KP, Chung Y, Li SC, Graff DE, Vermeire FH, Wu H, Green WH, McGill CJ: **Chemprop: a machine learning package for chemical property prediction**. *J Chem Inf Model* 2024, **64**:9–17, https://doi.org/10.1021/acs.jcim.3c01250.

71. Zhao H: **The science and art of structure-based virtual screening**. *ACS Med Chem Lett* 2024, **15**:436–440, https://doi.org/10.1021/acsmedchemlett.4c00093.

72. Lyu J, Irwin JJ, Schoichet BK: **Modeling the expansion of virtual screening libraries**. *Nat Chem Biol* 2023, **19**:712–718, https://doi.org/10.1038/s41589-022-01234-w.

73. Kuan J, Radaeva M, Avenido A, Cherkasov A, Gentile F: **Keeping pace with the explosive growth of chemical libraries with structure-based virtual screening**. *WIREs Comput Mol Sci* 2023, **13**, e1678, https://doi.org/10.1002/wcms.1678.

74. Acharya A, Agarwal R, Baker MB, Baudry J, Bhowmik D, Boehm S, Byler KG, Chen SY, Coates L, Cooper CJ, Demerdash O, Daidone I, Eblen JD, Ellingson S, Forli S, Glaser J, Gumbart JC, Gunnels J, Hernandez O, Irle S, Kneller DW, Kovalevsky A, Larkin J, Lawrence TJ, LeGrand S, Liu S-H, Mitchell JC, Park G, Parks JM, Pavlova A, Petridis L, Poole D, Pouchard L, Ramanathan A, Rogers DM, Santos-Martins D, Scheinberg A, Sedova A, Shen Y, SmitH JC, Smith MD, Soto C, Tsaris A, Thavappiragasam M, Tillack AF, Vermaas JV, Vuong VQ, Yin J, Yoo S, Zahran M, Zanetti-Polzi L: **Supercomputer-based ensemble docking drug discovery pipeline with application to Covid-19**. *J Chem Inf Model* 2020, **60**:5832–5852, https://doi.org/10.1021/acs.jcim.0c01010.

75. Raghavan B, Paulikat M, Ahmad K, Callea L, Rizzi A, Ippoliti E, Mandelli D, Bonati L, De Vivo M, Carloni P: **Drug design in the exascale era: a perspective from massively parallel QM/MM simulations**. *J Chem Inf Model* 2023, **63**:3647–3658, https://doi.org/10.1021/acs.jcim.3c00557.

76. Santagati R, Aspuru-Guzik A, Babbush R, Degroote M, Gonzalez L, Kyoseva E, Moll N, Oppel M, Parrish RM, Rubin NC, Streif M, Tautermann CS, Weiss H, Wiebe N, Utschig-Utschig C: **Drug design on quantum computers**. *Nat Phys* 2024, **20**:549–557, https://doi.org/10.1038/s41567-024-02411-5.

77. Pyrkov A, Aliper A, Bezrukov D, Lin Y-C, Polykovskiy D, Kamya P, Ren F, Zhavoronkov A: **Quantum computing for near-term applications in generative chemistry and drug discovery**. *Drug Discov Today* 2023, **28**:1–12, https://doi.org/10.1016/j.drudis.2023.103675.