



ISSN : 2347 - 2243

Indo - American Journal of Life Sciences and Biotechnology



<https://iajlb.org>

Email : editor@iajlb.com or iajlb.editor@gmail.com



Homology Modelling in Structure-Based Drug Design: Tools, Validation and Pitfalls

Beatriz Teixeira¹●, Kiran Bose²●, Tyler Dunleavy³●

¹Scientist, Department of Bioinformatics, Minho Valley Institute of Health Research, Viana do Castelo, Portugal.

²Associate Professor, Department of Computational Biology, Nallamala Institute of Life Sciences, Kurnool, India.

³Senior Scientist, Department of Bioinformatics, Cumberland Plateau Institute of Life Sciences, Cookeville, USA.

Email: ¹b.teixeira@mvihr.pt, ²k.bose@nils.ac.in, ³t.dunleavy@cpils.edu

ORCID: ¹0000-0001-2900-1910, ²0000-0001-3973-3605, ³0000-0001-2468-7781

Received: June 27, 2024 Accepted: October 4, 2024 Published: February 19, 2025

DOI: [10.62644/iajlb.2025.v22.i1.pp11-18](https://doi.org/10.62644/iajlb.2025.v22.i1.pp11-18)

ABSTRACT

Homology modelling builds a three-dimensional model of a protein from its sequence and the experimentally solved structure of a related protein, and it is often the only route to a structure for targets that crystallography or cryo-electron microscopy has not yet reached. This narrative mini review asks how far such models can be trusted when they are used for structure-based drug design, and what happens to that trust now that deep-learning predictors deliver models for almost any sequence. It draws on published work on template detection, alignment, model building, validation, docking and the treatment of flexibility, and it reports no new data. The evidence points to a conditional answer. Models built on close templates can support the analysis of a binding site and, in well-known receptor examples, the discovery of new ligands, but accuracy at the level of individual side chains and loops, which is the level that matters for docking, depends on the alignment, on the state of the template and on how the pocket moves. Global validation scores describe the whole fold and say little about the pocket. Predicted confidence from deep-learning methods is useful but does not remove the need for checks tied to the question being asked. The review ends with a short list of practices that make model-based design more reliable and more honestly reported.

Keywords: Homology modelling; Structure-based drug design; Model validation; Molecular docking; Protein structure prediction; AlphaFold; Narrative review

Citation: Teixeira, Bose and Dunleavy [2025]. Homology Modelling in Structure-Based Drug Design: Tools, Validation and Pitfalls. *Indo-Am. J. Life Sci. Biotechnol.* 22(1):11-18. DOI: [10.62644/iajlb.2025.v22.i1.pp11-18](https://doi.org/10.62644/iajlb.2025.v22.i1.pp11-18)

Article Class: Mini Review

1. Introduction

Structure-based drug design starts from a picture of the target: where the pocket is, which residues line it, and how a small molecule might sit there (1, 2). Experimental structures remain the best

picture available, and the Protein Data Bank has grown into a vast archive of them (3). Even so, a great many targets of pharmaceutical interest, among them membrane receptors and enzymes from pathogens, have no experimental structure



at the point when a project begins. Teams in that position build a model from a related protein and design against it (4, 5). Homology modelling rests on a simple observation: protein structure changes more slowly than protein sequence, so related sequences fold alike (6). The practical recipe is to find a solved relative, align the target to it, copy and adjust the backbone, rebuild what differs, and judge the result (7, 8). Since 2021 deep-learning predictors have changed the landscape, delivering models of striking accuracy for many sequences (9, 10). They have not retired the underlying question. Whether a model is good enough for the purpose at hand is still a judgement that has to be made, and made again for each target.

This mini review sets out what the literature says about that judgement. It follows a model from template choice to docking, noting where error enters and which checks can detect it. No local or primary data are reported, and the article says plainly where the evidence is thin. The aim is a fair account of what has been shown, as against what has merely been repeated.

2. Literature Review

2.1 From Template to Model

The first decision is the template. Below a certain level of sequence similarity, the so-called twilight zone, a match can be real and still be hard to tell from chance (11). BLAST-type searches find close relatives readily (12), while profile methods reach further (13). Figure 1 shows the stages that follow. A multiple alignment is prepared, a model is built under spatial restraints or by assembling fragments (8, 14), loops that differ from the template are rebuilt (15), and the result is checked and refined.

Of these stages the alignment is the one that most often decides the outcome. An alignment that shifts a strand by one

residue yields a model with every side chain in the wrong place, however elegant the later steps (7). Alignment, in other words, is where most of the judgement in the whole exercise is spent.

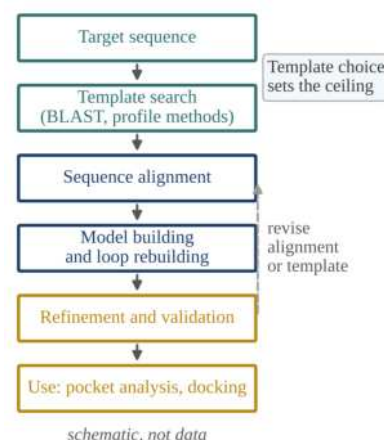


Figure 1. Typical stages of a homology-modelling workflow (schematic, not data).

2.2 The Toolbox

The available tools differ in philosophy more than in outward behaviour. MODELLER satisfies restraints derived from the alignment (8, 16). SWISS-MODEL packages template search, building and quality estimation into an automated pipeline (14, 17). Servers built around threading and fragment assembly, such as Phyre2 and I-TASSER, aim at targets where no close template exists (18, 19). Table 1 lists the main options and their usual caveats.

2.3 Deep-Learning Predictors

The community-wide CASP experiments have tracked progress in structure prediction for decades (22), and they recorded a sharp step forward when deep-learning methods entered the field. AlphaFold 2 and RoseTTAFold produced models that, for many targets, approached experimental accuracy in the fold (9, 10). Figure 2 places these developments on a timeline. The AlphaFold database made predicted models of whole proteomes freely available (23), and ColabFold



lowered the cost of running the method (20).

Table 1. Representative modelling approaches and their usual caveats.

Approach	Basis	Usual caveat
MODELLER	Spatial restraints from an alignment (8, 16)	Quality follows the alignment
SWISS-MODEL	Automated template-based server (14, 17)	Few user decisions; template-bound
Phyre2, I-TASSER	Remote templates, fragment assembly (18, 19)	Weaker where no template fits
RoseTTAFold, AlphaFold 2	Deep learning on sequences (9, 10)	Apo-like pocket; ligand state absent
ColabFold	Fast alignments feeding AlphaFold (20)	Inherits the same limits
AlphaFold 3	Includes ligands and ions (21)	Still a prediction

Note: Qualitative summary of the cited tool papers; no performance figures are given because they depend on the target and the benchmark.

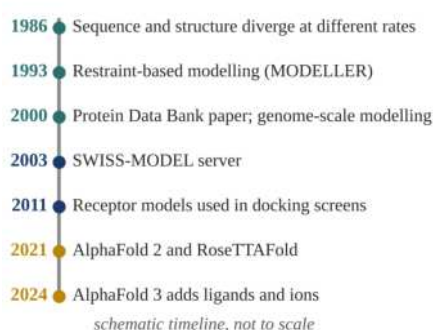


Figure 2. Milestones in template-based and deep-learning structure prediction (schematic timeline, not to scale).

What these methods deliver is a prediction of a single, most probable conformation, with a per-residue confidence estimate. That suits the fold less well when the pocket depends on a bound ligand, cofactor or metal ion. Predictions are valuable hypotheses that accelerate, but do

not replace, experimental determination (24). AlphaFold 3 extends the approach to ligands and ions (21), an advance that will need the same independent testing.

2.4 Models in Design Campaigns

The case for models in drug design rests on a handful of well-studied receptors. Docking against the crystal structure of the beta2-adrenergic receptor found new ligands (25), and a dopamine D3 receptor study screened a homology model and the crystal structure solved afterwards, finding ligands in both (26). The GPCR Dock assessment showed both promise and spread in how well receptor models and ligand poses could be predicted (27), and work on comparative models found that enrichment of actives falls off as the model departs from the true structure (28). Docking programs rank poses with approximate scoring functions (29, 30).

2.5 Flexibility and the State of the Template

Binding sites are not rigid. Proteins adjust to ligands, an idea that goes back to the induced-fit proposal (31) and that matters greatly for design (32). A model inherits the conformation of its template, so a template solved with a different ligand, or none, can give a pocket that is too closed, too open or shaped for another molecule. Models from different templates may therefore disagree about the pocket without either being wrong in the fold. Screens against AlphaFold models have been reported to do less well on average than screens against experimental structures (33). That is why the same model can look convincing to one group and unconvincing to another.

3. Materials and Methods

3.1 Scope and Question

The review asks one question in three parts. How reliable is a homology model,



or a predicted model, as the basis for structure-based design? Which steps contribute most to its errors? And which checks, applied before docking, give the best warning of a problem? The scope is limited to template-based and deep-learning models of single protein targets, used in docking and pocket analysis.

3.2 Literature Assembly

The review draws on peer-reviewed methods papers, benchmark assessments and a few case studies in drug discovery, chosen for relevance, standing in the field and clarity of method. Work postdating the issue year was not considered. No formal systematic search was run, and none is claimed. The article is a narrative synthesis whose selection reflects the authors' judgement, so it guides the structure of the problem and not the size of any reported effect, and it quotes no benchmark figures.

3.3 Appraisal Framework

To compare reports, we asked four questions of each. Was the template, and its state, described? Was the alignment reported and examined (7)? Were validation checks applied to the binding site and not only to the whole fold (34)? And was the docking evaluated against known actives and inactives, as opposed to being shown as a single pose (35)? It is an old lesson, and it keeps needing to be relearned each time a new method arrives.

4. Results

4.1 What Model Quality Depends On

Model quality tracks template closeness, but it tracks alignment accuracy even more closely, and it varies from region to region (7, 15). Figure 3 expresses the common reading in qualitative terms. With a close template, backbone-level questions and pocket location are usually safe, and ligand docking becomes defensible. With a

distant template, the fold may be right and the pocket still too uncertain for anything beyond a first look (4, 5). Low-confidence regions of predicted models deserve the same caution as loops in any other model (9, 23).

4.2 What Validation Can and Cannot Show

A large toolbox exists for checking models (Table 2). Backbone torsions are compared with the Ramachandran map (36, 37), clashes and rotamers are listed (38), residue environments are scored (39), and statistical potentials and composite scores give a global view (40, 41, 42, 43).

	Fold and pocket site	Pocket shape	Ligand docking
Close template	generally reasonable	generally reasonable	use with caution
Moderate template	generally reasonable	use with caution	poorly supported
Distant template	use with caution	poorly supported	poorly supported

schematic reading of the literature, not data

Figure 3. Qualitative reading of how far a model can be pushed by template closeness (schematic, not data).

Table 2. Validation measures and what each can and cannot show.

Measure	Examines	Blind spot
Ramachandran check	Backbone torsion angles (36, 37)	Says little about the pocket
MolProbity	Clashes, rotamers, geometry (38)	Can pass a wrong fold
Verify3D	Residue environment fit (39)	Coarse; fold-level
ProSA, DOPE	Statistical potentials (40, 41)	Global; pocket averaged out
QMEAN, QMEANDisCo	Composite quality estimate (42, 43)	Estimate, not a measurement

Note: The last column gives the authors' reading of the cited methods papers.



These checks answer a general question and are poor at the specific one. A model can have excellent geometry and still carry the wrong fold or a wrong pocket, because stereochemical checks measure whether the model looks like a protein, not whether it is the right shape. Work on identifying correct models showed that telling good from bad candidates is hard and that no single score does it reliably (34). Confidence in a fold, in short, is not confidence in a pocket.

4.3 Evidence from Docking and Screening

The successes in the literature share features: receptors with several experimental relatives, close templates and laboratory testing of the predicted ligands (25, 26). In the D3 comparison both the model and the later structure led to ligands, so models can be useful for finding molecules even when they are imperfect for describing how those molecules bind. Performance on predicted structures is variable and often below that for experimental ones (33, 44), and programs and scoring functions add their own spread (35). A single clean pose from one model is a hypothesis, not a measurement.

4.4 Recurring Pitfalls

Several pitfalls recur (Table 3). The alignment may be wrong in the pocket, the template may be in a different state from the one the project cares about, and loops lining the site may be placed by a method least reliable exactly where it is needed (7, 15). A good global score may be offered as reassurance when the pocket has not been examined (34). A single model may also be treated as the structure, not as one member of a plausible ensemble (32). A model that looks tidy in a viewer has not thereby earned the right to be docked against.

Table 3. Recurring pitfalls in model-based design and ways to reduce them.

Pitfall	Consequence	Mitigation
Alignment error	Shifted residues in the pocket (7, 34)	Inspect and compare alignments
Wrong template state	Pocket closed or open (31, 32)	Use several templates
Poorly modelled loops	Misplaced lining residues (15)	Treat loop atoms with caution
Trust in global scores	A good fold, a poor site (34)	Add pocket-level checks
Docking to one model	Over-confident ranking (28, 35)	Dock to an ensemble

Note: Pitfalls are those discussed in the cited literature; they are not attributed to any single study.

These are the ordinary risks of a method that infers a structure instead of measuring it, and the remedies are equally ordinary: compare several templates and alignments, examine the pocket residue by residue, state the template and its ligand, and dock into more than one conformation where the evidence suggests that the pocket moves (28, 32). None of this makes the method unreliable, only dependent on choices that a reader should be able to see.

5. Discussion

5.1 Interpretation

The main finding of the review is conditional. Models can support design, and in celebrated cases they have done so (25, 26), provided that the template is close and well chosen or the prediction confident, the alignment has been examined, the binding site has been checked in its own right, and docking is evaluated on known actives and inactives (34, 35). Deep-learning predictors lower the barrier to starting a project (23) but do not supply the conformation of a pocket bound to a ligand (24, 44), and newer methods that model ligands directly (21) still need independent testing. Seen this



way, a model is less an answer than a well-informed starting point. Nothing here suggests the modelling route is a poor one, only that its limits are easy to forget.

5.2 Limitations

This is a narrative review and shares the limits of that form. Selection of sources reflects the authors' judgement, no formal search or quality scoring was done, and the article does not quote benchmark numbers from individual studies. The evidence also leans toward a few well-studied receptor families and may not transfer to other target classes. The standard is modest, and a good deal of the published work does not yet meet it.

5.3 Future Work

Three needs stand out. The first is more prospective tests in which models are built and used for design before the experimental structure is known, then compared with it afterwards, as in the dopamine D3 case (26). The second is validation measures aimed at the pocket and not at the fold. The third is routine reporting of the template, its ligand state and the docking controls. Progress will probably come from patient comparison of models with new experimental structures, not from larger prediction sets alone.

6. Conclusion

Homology modelling remains a useful route to a structure for targets that experiments have not reached, and it has contributed to real discovery work. Its reliability rests on a short chain of decisions about template, alignment, pocket conformation and docking, and global validation scores do not certify the weakest link. Deep-learning predictors improve the fold and leave the central difficulty, a pocket in the right shape for the right ligand, largely where it was. Reviews of this kind are most useful when they say where the evidence stops.

For the practitioner, the message is practical. Build more than one model, look at the pocket directly, state what the template was bound to, and test docking against known ligands before trusting any new prediction.

7. Acknowledgements

The authors thank colleagues who commented on earlier drafts of this review.

8. Funding

This work received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

9. Conflict of Interest

The authors declare no conflict of interest.

10. Author Contributions

B.T. conceived the review and wrote the sections on template selection and model building. K.B. wrote the sections on validation and structure prediction. T.D. wrote the sections on docking and drug design. All authors revised the manuscript and approved the final version.

11. Ethics Statement

This is a review of published literature. No human participants or animals were involved and no ethical approval was required.

12. Data Availability

No new data were generated. All sources are cited in the reference list.

13. References

1. Anderson, A. C. (2003). The process of structure-based drug design. *Chemistry & Biology*, 10(9), 787–797. DOI: [10.1016/j.chembiol.2003.09.002](https://doi.org/10.1016/j.chembiol.2003.09.002)
2. Batool, M., Ahmad, B., & Choi, S. (2019). A structure-based drug discovery paradigm. *International Journal of Molecular Sciences*, 20(11), 2783. DOI: [10.3390/ijms20112783](https://doi.org/10.3390/ijms20112783)
3. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., &



- Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, 28(1), 235–242. DOI: [10.1093/nar/28.1.235](https://doi.org/10.1093/nar/28.1.235)
4. Cavasotto, C. N., & Phatak, S. S. (2009). Homology modeling in drug discovery: Current trends and applications. *Drug Discovery Today*, 14(13–14), 676–683. DOI: [10.1016/j.drudis.2009.04.006](https://doi.org/10.1016/j.drudis.2009.04.006)
5. Hillisch, A., Pineda, L. F., & Hilgenfeld, R. (2004). Utility of homology models in the drug discovery process. *Drug Discovery Today*, 9(15), 659–669. DOI: [10.1016/S1359-6446\(04\)03196-4](https://doi.org/10.1016/S1359-6446(04)03196-4)
6. Chothia, C., & Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. *The EMBO Journal*, 5(4), 823–826. DOI: [10.1002/j.1460-2075.1986.tb04288.x](https://doi.org/10.1002/j.1460-2075.1986.tb04288.x)
7. Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F., & Sali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*, 29, 291–325. DOI: [10.1146/annurev.biophys.29.1.291](https://doi.org/10.1146/annurev.biophys.29.1.291)
8. Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3), 779–815. DOI: [10.1006/jmbi.1993.1626](https://doi.org/10.1006/jmbi.1993.1626)
9. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. DOI: [10.1038/s41586-021-03819-2](https://doi.org/10.1038/s41586-021-03819-2)
10. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., ... Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557), 871–876. DOI: [10.1126/science.abj8754](https://doi.org/10.1126/science.abj8754)
11. Rost, B. (1999). Twilight zone of protein sequence alignments. *Protein Engineering*, 12(2), 85–94. DOI: [10.1093/protein/12.2.85](https://doi.org/10.1093/protein/12.2.85)
12. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, 25(17), 3389–3402. DOI: [10.1093/nar/25.17.3389](https://doi.org/10.1093/nar/25.17.3389)
13. Söding, J. (2005). Protein homology detection by HMM–HMM comparison. *Bioinformatics*, 21(7), 951–960. DOI: [10.1093/bioinformatics/bti125](https://doi.org/10.1093/bioinformatics/bti125)
14. Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Research*, 31(13), 3381–3385. DOI: [10.1093/nar/gkg520](https://doi.org/10.1093/nar/gkg520)
15. Fiser, A., Do, R. K. G., & Sali, A. (2000). Modeling of loops in protein structures. *Protein Science*, 9(9), 1753–1773. DOI: [10.1110/ps.9.9.1753](https://doi.org/10.1110/ps.9.9.1753)
16. Webb, B., & Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Current Protocols in Bioinformatics*, 54, 5.6.1–5.6.37. DOI: [10.1002/cpbi.3](https://doi.org/10.1002/cpbi.3)
17. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., ... Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, 46(W1), W296–W303. DOI: [10.1093/nar/gky427](https://doi.org/10.1093/nar/gky427)
18. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), 845–858. DOI: [10.1038/nprot.2015.053](https://doi.org/10.1038/nprot.2015.053)
19. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2015). The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12(1), 7–8. DOI: [10.1038/nmeth.3213](https://doi.org/10.1038/nmeth.3213)
20. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: Making protein folding accessible to all. *Nature Methods*, 19(6), 679–682. DOI: [10.1038/s41592-022-01488-1](https://doi.org/10.1038/s41592-022-01488-1)
21. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., ... Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500. DOI: [10.1038/s41586-024-07487-w](https://doi.org/10.1038/s41586-024-07487-w)
22. Kryshchak, A., Schwede, T., Topf, M., Fidelis, K., & Mout, J. (2021). Critical assessment of methods of protein structure prediction (CASP): Round XIV. *Proteins*, 89(12), 1607–1617. DOI: [10.1002/prot.26237](https://doi.org/10.1002/prot.26237)
23. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., ... Velankar, S. (2022). AlphaFold Protein Structure Database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, 50(D1), D439–D444. DOI: [10.1093/nar/gkab1061](https://doi.org/10.1093/nar/gkab1061)
24. Terwilliger, T. C., Liebschner, D., Croll, T. I., Williams, C. J., McCoy, A. J., Poon, B. K., ... Adams, P. D. (2024). AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination. *Nature Methods*, 21(1), 110–116. DOI: [10.1038/s41592-023-02087-4](https://doi.org/10.1038/s41592-023-02087-4)
25. Kolb, P., Rosenbaum, D. M., Irwin, J. J., Fung, J. J., Kobilka, B. K., & Shoichet, B. K. (2009). Structure-based discovery of beta2-adrenergic receptor ligands. *Proceedings of the National Academy of Sciences*, 106(16), 6843–6848. DOI: [10.1073/pnas.0812657106](https://doi.org/10.1073/pnas.0812657106)
26. Carlsson, J., Coleman, R. G., Setola, V., Irwin, J. J., Fan, H., Schlessinger, A., ... Shoichet, B. K. (2011). Ligand discovery from a dopamine D3



- receptor homology model and crystal structure. *Nature Chemical Biology*, 7(11), 769–778. DOI: [10.1038/nchembio.662](https://doi.org/10.1038/nchembio.662)
27. Kufareva, I., Rueda, M., Katritch, V., Stevens, R. C., & Abagyan, R. (2011). Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure*, 19(8), 1108–1126. DOI: [10.1016/j.str.2011.05.012](https://doi.org/10.1016/j.str.2011.05.012)
 28. Fan, H., Irwin, J. J., Webb, B. M., Klebe, G., Shoichet, B. K., & Sali, A. (2009). Molecular docking screens using comparative models of proteins. *Journal of Chemical Information and Modeling*, 49(11), 2512–2527. DOI: [10.1021/ci9003706](https://doi.org/10.1021/ci9003706)
 29. Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., ... Shenkin, P. S. (2004). Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, 47(7), 1739–1749. DOI: [10.1021/jm0306430](https://doi.org/10.1021/jm0306430)
 30. Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), 455–461. DOI: [10.1002/jcc.21334](https://doi.org/10.1002/jcc.21334)
 31. Koshland, D. E. (1958). Application of a theory of enzyme specificity to protein synthesis. *Proceedings of the National Academy of Sciences*, 44(2), 98–104. DOI: [10.1073/pnas.44.2.98](https://doi.org/10.1073/pnas.44.2.98)
 32. Teague, S. J. (2003). Implications of protein flexibility for drug discovery. *Nature Reviews Drug Discovery*, 2(7), 527–541. DOI: [10.1038/nrd1129](https://doi.org/10.1038/nrd1129)
 33. Scardino, V., Di Filippo, J. I., & Cavasotto, C. N. (2023). How good are AlphaFold models for docking-based virtual screening? *iScience*, 26(1), 105920. DOI: [10.1016/j.isci.2022.105920](https://doi.org/10.1016/j.isci.2022.105920)
 34. Wallner, B., & Elofsson, A. (2003). Can correct protein models be identified? *Protein Science*, 12(5), 1073–1086. DOI: [10.1110/ps.0236803](https://doi.org/10.1110/ps.0236803)
 35. Warren, G. L., Andrews, C. W., Capelli, A.-M., Clarke, B., LaLonde, J., Lambert, M. H., ... Head, M. S. (2006). A critical assessment of docking programs and scoring functions. *Journal of Medicinal Chemistry*, 49(20), 5912–5931. DOI: [10.1021/jm050362n](https://doi.org/10.1021/jm050362n)
 36. Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, 7(1), 95–99. DOI: [10.1016/S0022-2836\(63\)80023-6](https://doi.org/10.1016/S0022-2836(63)80023-6)
 37. Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26(2), 283–291. DOI: [10.1107/S0021889892009944](https://doi.org/10.1107/S0021889892009944)
 38. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., ... Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D*, 66(1), 12–21. DOI: [10.1107/S0907444909042073](https://doi.org/10.1107/S0907444909042073)
 39. Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, 356(6364), 83–85. DOI: [10.1038/356083a0](https://doi.org/10.1038/356083a0)
 40. Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 35(W1), W407–W410. DOI: [10.1093/nar/gkm290](https://doi.org/10.1093/nar/gkm290)
 41. Shen, M., & Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Science*, 15(11), 2507–2524. DOI: [10.1110/ps.062416606](https://doi.org/10.1110/ps.062416606)
 42. Benkert, P., Biasini, M., & Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, 27(3), 343–350. DOI: [10.1093/bioinformatics/btq662](https://doi.org/10.1093/bioinformatics/btq662)
 43. Studer, G., Rempfer, C., Waterhouse, A. M., Gumienny, R., Haas, J., & Schwede, T. (2020). QMEANDisCo: Distance constraints applied on model quality estimation. *Bioinformatics*, 36(6), 1765–1771. DOI: [10.1093/bioinformatics/btz828](https://doi.org/10.1093/bioinformatics/btz828)
 44. Karelina, M., Noh, J. J., & Dror, R. O. (2023). How accurately can one predict drug binding modes using AlphaFold models? *eLife*, 12, RP89386. DOI: [10.7554/eLife.89386](https://doi.org/10.7554/eLife.89386)