Current Research in Structural Biology 3 (2021) 187–191

Contents lists available at ScienceDirect

Current Research in Structural Biology

journal homepage: www.journals.elsevier.com/current-research-in-structural-biology

Current approaches to flexible loop modeling

Amélie Barozet[a], Pablo Chacón[b], Juan Cortés[a,*]

[a] LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France
[b] Department of Biological Physical Chemistry, Rocasolano Physical Chemistry Institute C.S.I.C., Madrid, Spain

ARTICLE INFO

Keywords:
Loop modeling
Protein flexibility
Conformational sampling
Structure prediction
Energy landscapes

ABSTRACT

Loops are key components of protein structures, involved in many biological functions. Due to their conformational variability, the structural investigation of loops is a difficult topic, requiring a combination of experimental and computational methods. This paper provides a brief overview of current computational approaches to flexible loop modeling, and presents the main ingredients of the most standard protocols. Despite great progress in recent years, accurately modeling the conformational variability of long flexible loops remains a challenging problem. Future advances in this field will likely come from a tight coupling of experimental and computational techniques, which would enable a better understanding of the relationships between loop sequence, structural flexibility, and functional roles. *In fine*, accurate loop modeling will open the road to loop design problems of interest for applications in biomedicine and biotechnology.

Loops are protein fragments with an irregular structure connecting two secondary structure elements (typically α-helices and β-strands) and are main actors in many biological functions. They are often located at the surface of globular proteins, and their conformational versatility plays key roles in processes such as molecular recognition, allosteric regulation or signaling. On average, loops correspond to about one third of a protein sequence, and up to one half of the residues of enzyme active sites are located in loops (Bartlett et al., 2002).

Protein loops can have very different dynamic behaviors. In some cases, loops adopt a relatively stable conformation in the folded state of the protein, being part of the overall protein scaffold. Based on this static perspective, some efforts have been made for the classification of loop structures (Oliva et al., 1997). However, conformational changes in loop regions are frequently observed to drive enzyme function or regulation (Malabanan et al., 2010). For example, flexible loops (P-loop and T-loop) modulate protein kinases activity. Loops are also essential in many protein-ligand/cofactor interactions. For instance, extracellular loops of G-Protein-Coupled Receptors (GPCRs) are crucial for ligand recognition and binding (Hilger et al., 2018). Note that GPCRs and protein kinases are by far the major drug targets for pharmaceutical industry. Together with their variable motion amplitude, loop conformational transitions can take place in a broad range of timescales. A comprehensive list of cases demonstrating the functional importance of protein loops can be found in a recent review (Papaleo et al., 2016). Fig. 1 is an illustrative summary of loop functions and of the interplay between their structural properties and functional roles.

To better understand the dynamic nature of loops, a suitable representation of their conformational space is based on the concept of energy landscape (Wales, 2003). Although one loop can exist in multiple conformations, they are not equally likely to be adopted. More precisely, the probability of each possible state depends on its associated free energy. The energy landscape provides an overall view by displaying the energy at each point of the conformational space. The topology of this landscape gives insight into the main conformational and dynamic properties of the loop: the basins constitute meta-stable conformations, while saddle regions are associated with transition states. A visual representation can then be obtained by projecting this landscape along two meaningful coordinates, but other representations based on stationary points of the landscape can also be utilized. This is illustrated in Fig. 2.

The structural investigation and characterization of flexible loops is challenging. X-ray crystallography, which is the most widely used experimental method to determine high-resolution protein structures, only provides static snapshots in particular experimental conditions (e.g. presence/absence of a ligand). Furthermore, flexible regions are often missing in X-ray structures, or even in electron microscopy maps, because of lacking electron density. According to Djinovic-Carugo and Carugo (2015), 69% of the structures deposited in the Protein Data Bank (PDB) have missing fragments, and this percentage increases up to 80% for structures solved with a resolution above 2.0 Å. In more than 90% of the cases, these missing fragments are located in loops or unstructured

[*] Corresponding author.
E-mail address: juan.cortes@laas.fr (J. Cortés).

https://doi.org/10.1016/j.crstbi.2021.07.002
Received 30 September 2020; Received in revised form 30 June 2021; Accepted 25 July 2021

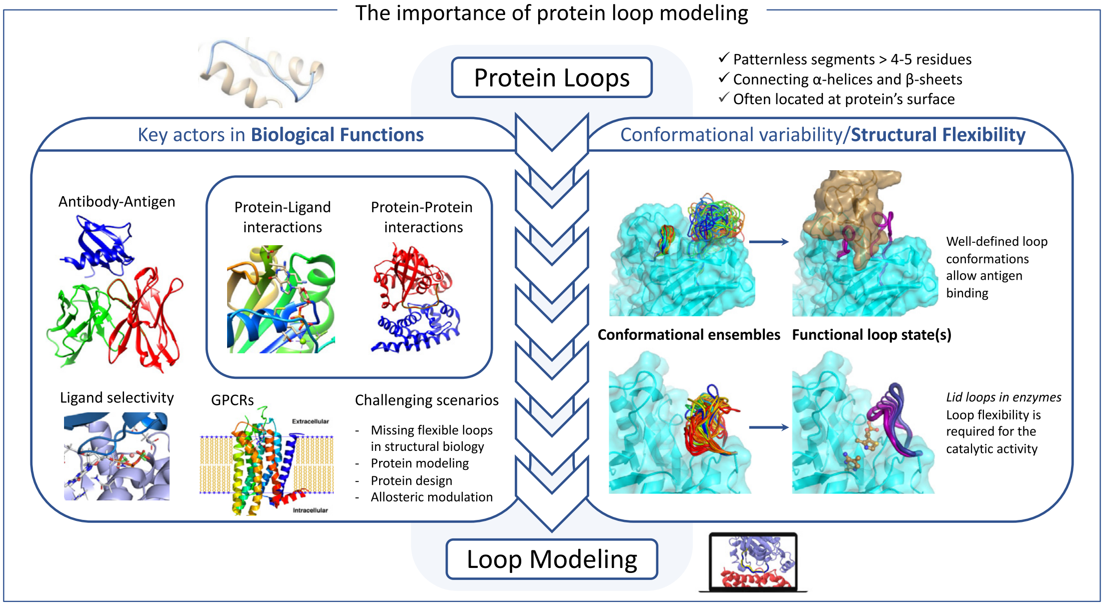

The importance of protein loop modeling
Protein Loops
✓ Patternless segments > 4-5 residues
✓ Connecting α-helices and β-sheets
✓ Often located at protein's surface
Key actors in Biological Functions
Antibody-Antigen
Protein-Ligand interactions
Protein-Protein interactions
Ligand selectivity
GPCRs
Extracellular
Intracellular
Challenging scenarios
- Missing flexible loops in structural biology
- Protein modeling
- Protein design
- Allosteric modulation
Conformational variability/Structural Flexibility
Well-defined loop conformations allow antigen binding
Conformational ensembles
Functional loop state(s)
Lid loops in enzymes
Loop flexibility is required for the catalytic activity
Loop Modeling

Fig. 1. Illustration of important functional roles of protein loops. Their location at the protein surface together with their conformational versatility make them key actors in intermolecular interactions. Their flexibility constitutes a challenge for their structural investigation, which requires a combination of experimental and computational methods.

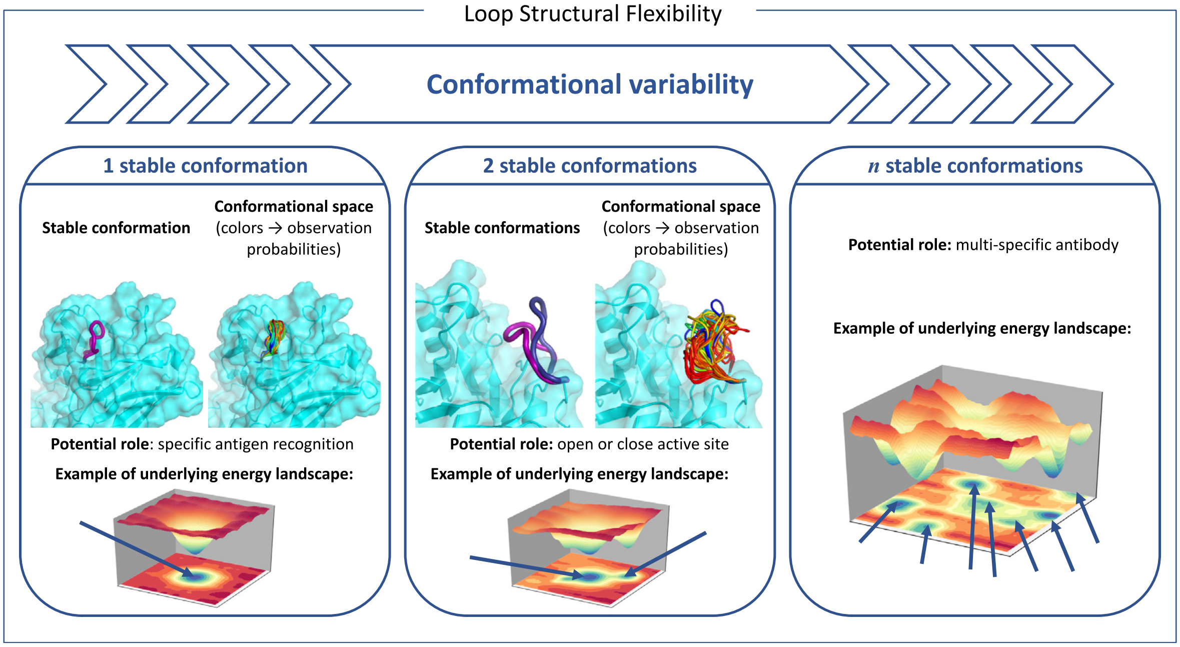

Loop Structural Flexibility
Conformational variability
1 stable conformation
Stable conformation
Conformational space (colors → observation probabilities)
Potential role: specific antigen recognition
Example of underlying energy landscape:
2 stable conformations
Stable conformations
Conformational space (colors → observation probabilities)
Potential role: open or close active site
Example of underlying energy landscape:
n stable conformations
Potential role: multi-specific antibody
Example of underlying energy landscape:

Fig. 2. Loops exhibit different levels of conformational variability. In some cases, they fluctuate around a main conformational state. For example, some CDR loops in antibodies adopt a single conformation suited to the recognition of a specific antigen. In other cases, a loop can adopt two main conformational states. For instance, this may correspond to a loop opening and closing the access to the active site of an enzyme. In general, the energy landscape of a loop is more complex, involving several basins corresponding to different meta-stable states. These multiple states can also be functionally important, for example in the case of multi-specific antibodies.

terminal regions. Other experimental techniques, such as X-ray solution scattering (Petoukhov et al., 2002) or nuclear magnetic resonance (NMR) (Boehr et al., 2006), can provide very relevant but limited structural and kinetic information regarding these flexible regions. Therefore, since an accurate atomistic representation of the diverse conformations adopted by the loop from experimental measurements is extremely difficult, computational methods are an essential complement to study flexible protein loops.

From a practical point of view, loop modeling techniques usually implement a multi-stage protocol, which varies depending on the underlying modeling approach. Fig. 3 aims to provide a general scheme in which most of the existing loop modeling techniques can fit. In brief, modeling can be divided into three major stages: (1) conformational sampling or search, (2) scoring and clustering, and (3) a final post-processing or refinement stage.

The core of loop modeling is the exploration of the vast

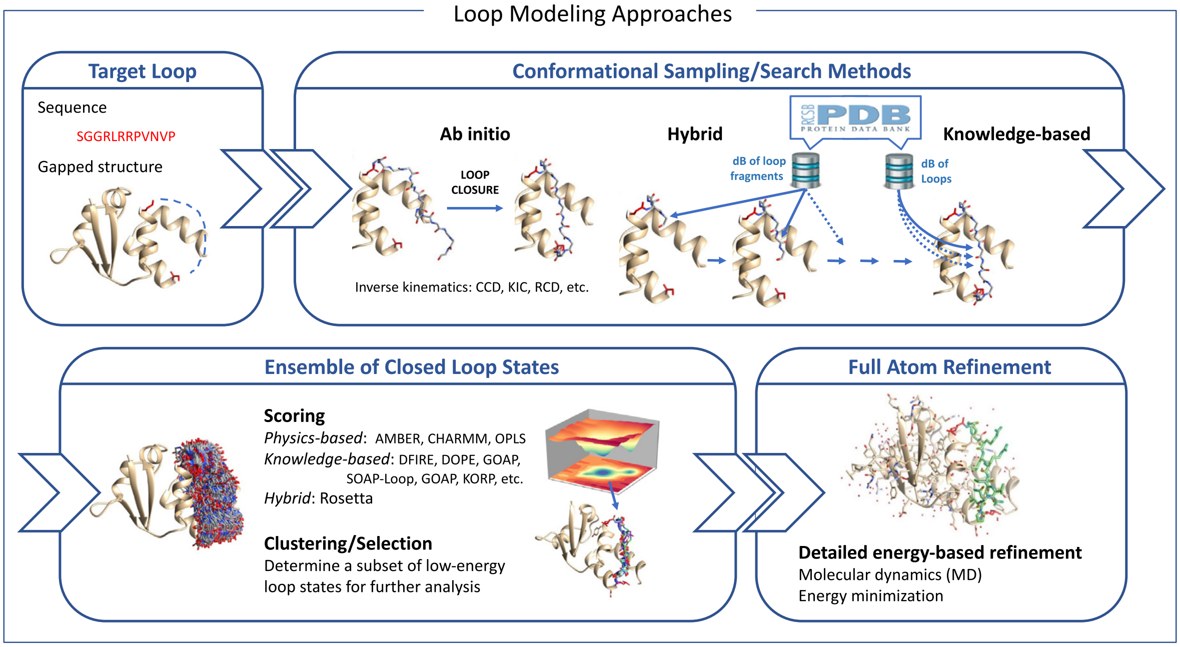

Loop Modeling Approaches
Target Loop
Sequence
SGGRLRRPVNVP
Gapped structure
Conformational Sampling/Search Methods
Ab initio
LOOP CLOSURE
Inverse kinematics: CCD, KIC, RCD, etc.
Hybrid
RCSB PDB PROTEIN DATA BANK
dB of loop fragments
dB of Loops
Knowledge-based
Ensemble of Closed Loop States
Scoring
Physics-based: AMBER, CHARMM, OPLS
Knowledge-based: DFIRE, DOPE, GOAP, SOAP-Loop, GOAP, KORP, etc.
Hybrid: Rosetta
Clustering/Selection
Determine a subset of low-energy loop states for further analysis
Full Atom Refinement
Detailed energy-based refinement
Molecular dynamics (MD)
Energy minimization

Fig. 3. Schematic workflow of a prototype loop modeling protocol. It involves several stages. Given a three-dimensional model of a protein and the loop sequence (required if the loop is missing in the input structure), the first stage generates possible states of the loop. This can be based on *ab initio* sampling, searches in structural repositories, or a combination of both approaches. When the number of sampled states is large, scoring and clustering methods are useful to select a reduced number of the most likely loop models. These models can be subsequently refined at the all-atom level, considering the flexibility of the surrounding parts of the protein, and possibly including a model of the solvent.

conformational space to be performed at the sampling/search step. Note that the volume of the conformational space grows exponentially with loop length. Very different approaches to protein loop modeling have been proposed over the years depending on the sampling/search protocol (Shehu and Kavraki, 2012; Papaleo et al., 2016; Kundert and Kortemme, 2019). Overall, they can be divided into three main classes: knowledge-based, *ab initio* and hybrid approaches. Knowledge-based approaches, also called template-based or homology-based, exploit structural repositories to retrieve observed loop conformations for a given set of sequence and geometric descriptors about the anchoring points (Choi and Deane, 2010; Messih et al., 2015; Shirvanizadeh et al., 2018; Karami et al., 2018). These methods are in general computationally fast since they do not rely on expensive simulations, but they are limited by the availability of suitable loop conformations from known protein structures, and only provide a small set of solutions. Especially for long loops, there is not enough structural data currently to cover the whole conformational space.

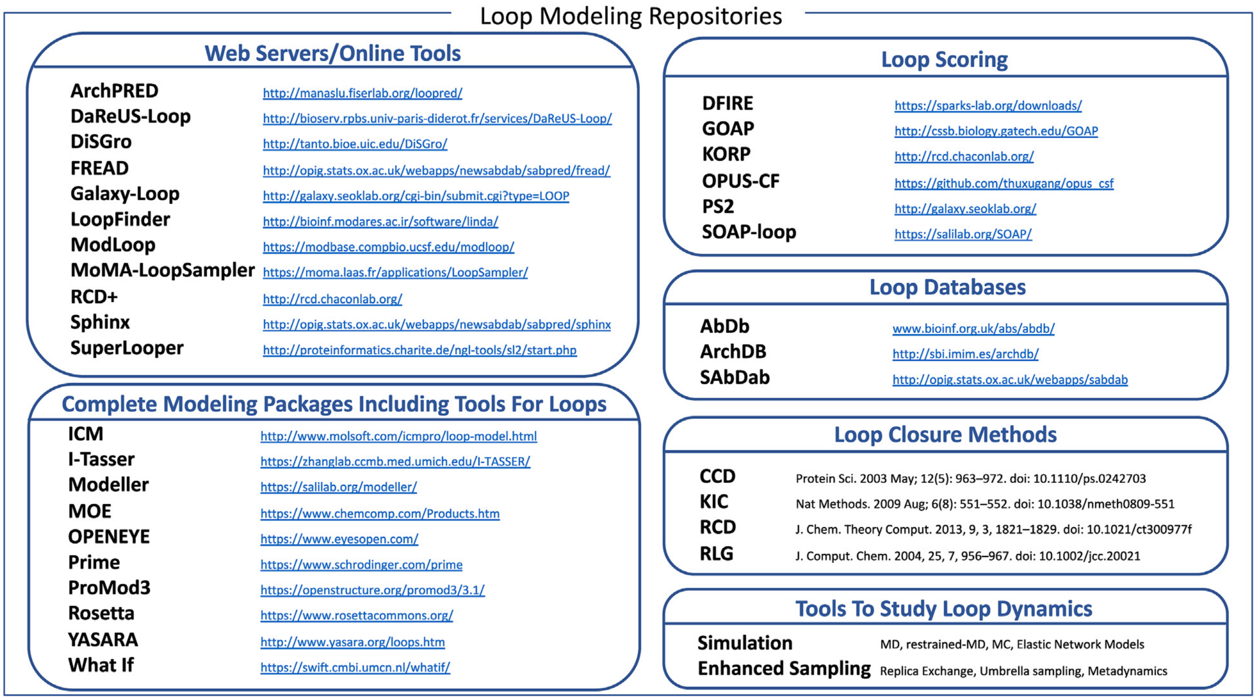
On the other end of the spectrum, *ab initio* approaches can perform a wider sampling of the conformational space, for example by exhaustively varying the torsional angles of the loop (Jacobson et al., 2004; López-Blanco et al., 2016). Their computational cost is highly variable, mainly depending on the underlying sampling method, but overall they are more demanding than knowledge-based approaches. To a lesser extent, they are also limited when dealing with long loops, as sampling coverage decreases with loop length.

Numerous loop modeling methods combine ideas from knowledge-based and *ab initio* approaches. Hybridization is performed at different levels: some methods apply a consensus between results obtained from different types of approaches (Deane and Blundell, 2001), while other methods apply a more intricate combination of strategies to explore the conformational space. In particular, many hybrid approaches use small fragments from structural databases within an *ab initio* sampling technique (Stein and Kortemme, 2013; Tang et al., 2014; Park et al., 2014; Marks et al., 2017; Barozet et al., 2019). The interest of one type of approach over the others may depend on the specific application and on the nature of the loop. For instance, knowledge-based approaches can be a suitable choice to predict the most probable conformations of more structurally conserved L1, L2, L3, H1, and H2 antibody CDR loops, whereas the more exhaustive sampling performed by *ab initio* and hybrid approaches will be more adequate to the highly variable H3 loop.

One of the main ingredients of *ab initio* and hybrid sampling approaches are loop-closure methods. They are required to ensure that N-terminal and C-terminal ends of the loop keep bonded to the anchoring points in the protein, while correct bond geometry is preserved. They are usually based on the concept of *inverse kinematics*, very common in robotics. Two classes of loop-closure methods have been proposed: those based on numerical optimization techniques (Canutescu and Dunbrack, 2003; Chys and Chacón, 2013), and those based on analytical solutions (Coutsias et al., 2004; Cortés et al., 2004). Numerical methods are conceptually simpler and directly applicable to long loops, but analytical methods are successfully used within iterative algorithms for the same purpose.

Scoring methods are another key ingredient, required to evaluate the quality of the sampled loop states. Scoring methods can also be classified into knowledge-based (López-Blanco and Chacón, 2019; Dong et al., 2013), physics-based (generally based on molecular mechanics approaches, not quantum physics models) and hybrid methods (Park et al., 2014; Alford et al., 2017). As before, knowledge-based scoring functions, typically linked to a coarse-grain representation of the loop, are the most computationally efficient alternative. However, their implicit simplifications prevent a faithful representation of the rugged energy surface of the all-atom conformational space. Conversely, the computational cost of detailed physics-based energy functions prevents a thorough exploration of the space. Hybrid approaches try to balance the best of these two worlds. A common practice is to use less accurate scoring functions in an initial search, and an atomistic detailed potential to refine only the high-scoring loop states. A suitable combination of sampling protocols and scoring functions is the major burden for loop modeling, and the key for future progress in the field.

In addition to the aforementioned loop modeling methods, Monte Carlo (MC) and molecular dynamics (MD) simulations can be used to investigate thermodynamic and kinetic properties of loops. Note



Loop Modeling Repositories

Web Servers/Online Tools	
ArchPRED	http://manaslu.fiserlab.org/loopred/
DaReUS-Loop	http://bioserv.rpbs.univ-paris-diderot.fr/services/DaReUS-Loop/
DiSGro	http://tanto.bioe.uic.edu/DiSGro/
FREAD	http://opig.stats.ox.ac.uk/webapps/newsabdab/sabpred/fread/
Galaxy-Loop	http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=LOOP
LoopFinder	http://bioinf.modares.ac.ir/software/linda/
ModLoop	https://modbase.compbio.ucsf.edu/modloop/
MoMA-LoopSampler	https://moma.laas.fr/applications/LoopSampler/
RCD+	http://rcd.chaconlab.org/
Sphinx	http://opig.stats.ox.ac.uk/webapps/newsabdab/sabpred/sphinx
SuperLooper	http://proteinformatics.charite.de/ngl-tools/sl2/start.php

Complete Modeling Packages Including Tools For Loops	
ICM	http://www.molsoft.com/icmpro/loop-model.html
I-Tasser	https://zhanglab.ccmb.med.umich.edu/I-TASSER/
Modeller	https://salilab.org/modeller/
MOE	https://www.chemcomp.com/Products.htm
OPENEYE	https://www.eyesopen.com/
Prime	https://www.schrodinger.com/prime
ProMod3	https://openstructure.org/promod3/3.1/
Rosetta	https://www.rosettacommons.org/
YASARA	http://www.yasara.org/loops.htm
What If	https://swift.cmbi.umcn.nl/whatif/

Loop Scoring	
DFIRE	https://sparks-lab.org/downloads/
GOAP	http://cssb.biology.gatech.edu/GOAP
KORP	http://rcd.chaconlab.org/
OPUS-CF	https://github.com/thuxugang/opus_csf
PS2	http://galaxy.seoklab.org/
SOAP-loop	https://salilab.org/SOAP/

Loop Databases	
AbDb	www.bioinf.org.uk/abs/abdb/
ArchDB	http://sbi.imim.es/archdb/
SAbDab	http://opig.stats.ox.ac.uk/webapps/sabdab

Loop Closure Methods	
CCD	Protein Sci. 2003 May; 12(5): 963–972. doi: 10.1110/ps.0242703
KIC	Nat Methods. 2009 Aug; 6(8): 551–552. doi: 10.1038/nmeth0809-551
RCD	J. Chem. Theory Comput. 2013, 9, 3, 1821–1829. doi: 10.1021/ct300977f
RLG	J. Comput. Chem. 2004, 25, 7, 956–967. doi: 10.1002/jcc.20021

Tools To Study Loop Dynamics	
Simulation	MD, restrained-MD, MC, Elastic Network Models
Enhanced Sampling	Replica Exchange, Umbrella sampling, Metadynamics

Fig. 4. Non-exhaustive list of available tools for protein loop modeling, classified in different categories.

however that in this context, MC-derived approaches are mainly employed as a sampling engine within more sophisticated loop prediction techniques, e.g. (Stein and Kortemme, 2013; Tang et al., 2014). Enhanced MD methods (Bernardi et al., 2015) are suitable tools to study loop dynamics (Papaleo et al., 2016). For example, MD simulations coupled with a higher-level representation such as Markov state models are a relevant approach to investigate the role of loop conformational changes in enzyme catalysis (Liao et al., 2018). Nevertheless, due to their very high computational cost, MD simulations are often only applied in a post-processing stage to refine solutions provided by much faster loop sampling methods (Lee et al., 2016).

Fig. 4 provides a (non-exhaustive) list of loop modeling protocols, related techniques for sampling and scoring, and structural repositories. Note also that other methods (in addition to MD) have been proposed to model loop conformational transitions rather that loop conformations, but they are not covered by this review.

Despite huge progress over the last two decades, flexible loop modeling remains an open problem (Marks et al., 2018; Barozet et al., 2021). State-of-the-art methods are able to predict stable conformations of relatively short loops (up to 12 residues). However, accurately sampling, scoring and representing the great diversity of loop conformations and transitions between them is still a computational challenge, in particular for long loops. Nowadays, Machine/Deep-Learning (ML/DL) approaches are showing huge potential in many areas, including structural bioinformatics (Gao et al., 2020; Gkeka et al., 2020; Pakhrin et al., 2021). In the context of loop modeling, ML/DL-based approaches have already provided good results as scoring methods (Ruffolo et al., 2020). To the best of our knowledge, the ability of ML/DL methods to generate meaningful conformational ensemble models of flexible loops remains to be demonstrated, but it is undoubtedly a promising research direction.

The main limitation for the development of more accurate and general loop modeling methods is the lack of experimental data. As mentioned before, loop flexibility is a challenge for biophysical methods: X-ray crystallography provides only static and possibly biased snapshots, NMR methods have difficulties dealing with large proteins, other methods (such as small-angle X-ray scattering or Förster resonance energy transfer) only provide coarse-grained constraints to build models. Therefore, new integrated approaches, tightly coupling several experimental and computational methods, are necessary for advances in this field.

Future improvements in loop modeling will help us to better understand the relationships between loop sequence, structural flexibility and biological function. Meanwhile, current approaches, with their limitations, are ready to use in challenging problems such as loop design in enzymes and antibodies (Kundert and Kortemme, 2019). Encouraging results in these fields bring us closer to the ultimate goal of designing loops able to recognize partner molecules with high specificity and affinity, or to enhance the catalytic activity of enzymes thanks to optimized dynamic properties.

Authorship contribution statement

All the authors contributed equally to this work.

CRediT authorship contribution statement

Amélie Barozet: Writing – original draft, Writing – review & editing, Visualization. **Pablo Chacón:** Writing – original draft, Writing – review & editing, Visualization. **Juan Cortés:** Writing – original draft, Writing – review & editing, Visualization.

Declaration of competing interest

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 paper.

Acknowledgments

This work has been supported by the French National Research Agency (ANR) through grant ANR-19-PI3A-0004 and by the Spanish grants BFU2016-76220-P and PID2019-109041 GB-C21 (AEI,10.13039/501100011033).

References

Alford, R., Leaver-Fay, A., Jeliazkov, J.R., O'Meara, M.J., DiMaio, F.P., Park, H., Shapovalov, M.V., Renfrew, P.D., Mulligan, V.K., Kappel, K., Labonte, J.W., Pacella, M.S., Bonneau, R., Bradley, P., Dunbrack, R., Das, R., Baker, D., Kuhlman, B., Kortemme, T., Gray, J.J., 2017. The Rosetta all-atom energy function for macromolecular modeling and design. J. Chem. Theor. Comput. 13, 3031–3048.

Barozet, A., Bianciotto, M., Vaisset, M., Siméon, T., Minoux, H., Cortés, J., 2021. Protein loops with multiple meta-stable conformations: a challenge for sampling and scoring methods. Proteins 89 (2), 218–231. https://doi.org/10.1002/prot.26008.

Barozet, A., Molloy, K., Vaisset, M., Siméon, T., Cortés, J., 2019. A reinforcement-learning-based approach to enhance exhaustive protein loop sampling. Bioinformatics 36, 1099–1106.

Bartlett, G.J., Porter, C.T., Borkakoti, N., Thornton, J.M., 2002. Analysis of catalytic residues in enzyme active sites. J. Mol. Biol. 324, 105–121.

Bernardi, R.C., Melo, M.C., Schulten, K., 2015. Enhanced sampling techniques in molecular dynamics simulations of biological systems. Biochim. Biophys. Acta Gen. Subj. 1850, 872–877.

Boehr, D.D., McElheny, D., Dyson, H.J., Wright, P.E., 2006. The dynamic energy landscape of dihydrofolate reductase catalysis. Science 313, 1638–1642.

Canutescu, A.A., Dunbrack, R.L., 2003. Cyclic coordinate descent: a robotics algorithm for protein loop closure. Protein Sci. 12, 963–972.

Choi, Y., Deane, C.M., 2010. FREAD revisited: accurate loop structure prediction using a database search algorithm. Proteins 78, 1431–1440.

Chys, P., Chacón, P., 2013. Random coordinate descent with spinor-matrices and geometric filters for efficient loop closure. J. Chem. Theor. Comput. 9, 1821–1829.

Cortés, J., Siméon, T., Remaud-Siméon, M., Tran, V., 2004. Geometric algorithms for the conformational analysis of long protein loops. J. Comput. Chem. 25, 956–967.

Coutsias, E., Seok, C., Jacobson, M., Dill, K., 2004. A kinematic view of loop closure. J. Comput. Chem. 25, 510–528.

Deane, C.M., Blundell, T.L., 2001. CODA: a combined algorithm for predicting the structurally variable regions of protein models. Protein Sci. 10, 599–612.

Djinovic-Carugo, K., Carugo, O., 2015. Missing strings of residues in protein crystal structures. Intrinsically Disord. Proteins 3, e1095697.

Dong, G.Q., Fan, H., Schneidman-Duhovny, D., Webb, B., Sali, A., 2013. Optimized atomic statistical potentials: assessment of protein interfaces and loops. Bioinformatics 29, 3158–3166.

Gao, W., Mahajan, S.P., Sulam, J., Gray, J.J., 2020. Deep learning in protein structural modeling and design. Patterns 1, 100142.

Gkeka, P., Stoltz, G., Barati Farimani, A., Belkacemi, Z., Ceriotti, M., Chodera, J.D., Dinner, A.R., Ferguson, A.L., Maillet, J.B., Minoux, H., Peter, C., Pietrucci, F., Silveira, A., Tkatchenko, A., Trstanova, Z., Wiewiora, R., Lelièvre, T., 2020. Machine learning force fields and coarse-grained variables in molecular dynamics: application to materials and biological systems. J. Chem. Theor. Comput. 16, 4757–4775.

Hilger, D., Masureel, M., Kobilka, B.K., 2018. Structure and dynamics of GPCR signaling complexes. Nat. Struct. Mol. Biol. 25, 4–12.

Jacobson, M.P., Pincus, D.L., Rapp, C.S., Day, T.J.F., Honig, B., Shaw, D.E., Friesner, R.A., 2004. A hierarchical approach to all-atom protein loop prediction. Proteins 55, 351–367.

Karami, Y., Guyon, F., De Vries, S., Tufféry, P., 2018. DaReUS-Loop: accurate loop modeling using fragments from remote or unrelated proteins. Sci. Rep. 8, 13673.

Kundert, K., Kortemme, T., 2019. Computational design of structured loops for new protein functions. Biol. Chem. 400, 275–288.

Lee, G.R., Heo, L., Seok, C., 2016. Effective protein model structure refinement by loop modeling and overall relaxation. Proteins 84, 293–301.

Liao, Q., Kulkarni, Y., Sengupta, U., Petrović, D., Mulholland, A.J., van der Kamp, M.W., Strodel, B., Kamerlin, S.C.L., 2018. Loop motion in triosephosphate isomerase is not a simple open and shut case. J. Am. Chem. Soc. 140, 15889–15903.

López-Blanco, J.R., Canosa-Valls, A.J., Li, Y., Chacón, P., 2016. RCD+: fast loop modeling server. Nucleic Acids Res. 44, W395–W400.

López-Blanco, J.R., Chacón, P., 2019. KORP: knowledge-based 6D potential for fast protein and loop modeling. Bioinformatics 35 (17), 3013–3019. https://doi.org/10.1093/bioinformatics/btz026.

Malabanan, M.M., Amyes, T.L., Richard, J.P., 2010. A role for flexible loops in enzyme catalysis. Curr. Opin. Struct. Biol. 20, 702–710.

Marks, C., Nowak, J., Klostermann, S., Georges, G., Dunbar, J., Shi, J., Kelm, S., Deane, C.M., 2017. Sphinx: merging knowledge-based and ab initio approaches to improve protein loop prediction. Bioinformatics 33, 1346–1353.

Marks, C., Shi, J., Deane, C.M., 2018. Predicting loop conformational ensembles. Bioinformatics 34, 949–956.

Messih, M.A., Lepore, R., Tramontano, A., 2015. LoopIng: a template-based tool for predicting the structure of protein loops. Bioinformatics 31, 3767–3772.

Oliva, B., Bates, P.A., Querol, E., Avilés, F.X., Sternberg, M.J., 1997. An automated classification of the structure of protein loops. J. Mol. Biol. 266, 814–830.

Pakhrin, S.C., Shrestha, B., Adhikari, B., Kc, D.B., 2021. Deep learning-based advances in protein structure prediction. Int. J. Mol. Sci. 22.

Papaleo, E., Saladino, G., Lambrughi, M., Lindorff-Larsen, K., Gervasio, F.L., Nussinov, R., 2016. The role of protein loops and linkers in conformational dynamics and allostery. Chem. Rev. 116, 6391–6423.

Park, H., Lee, G.R., Heo, L., Seok, C., 2014. Protein loop modeling using a new hybrid energy function and its application to modeling in inaccurate structural environments. PloS One 9, 1–18.

Petoukhov, M.V., Eady, N.A., Brown, K.A., Svergun, D.I., 2002. Addition of missing loops and domains to protein models by x-ray solution scattering. Biophys. J. 83, 3113–3125.

Ruffolo, J.A., Guerra, C., Mahajan, S.P., Sulam, J., Gray, J.J., 2020. Geometric potentials from deep learning improve prediction of CDR H3 loop structures. Bioinformatics 36, i268–i275.

Shehu, A., Kavraki, L.E., 2012. Modeling structures and motions of loops in protein molecules. Entropy 14, 252–290.

Shirvanizadeh, N., Vriend, G., Arab, S.S., 2018. Loop modelling 1.0. J. Mol. Graph. Model. 84, 64–68.

Stein, A., Kortemme, T., 2013. Improvements to robotics-inspired conformational sampling in Rosetta. PloS One 8, e63090.

Tang, K., Zhang, J., Liang, J., 2014. Fast protein loop sampling and structure prediction using distance-guided sequential chain-growth Monte Carlo method. PLoS Comput. Biol. 10, e1003539.

Wales, D., 2003. Energy Landscapes: Applications to Clusters, Biomolecules and Glasses. Cambridge University Press.