

# Combining Experimental and Computational Methods

Subjects: Biochemistry & Molecular Biology

Contributor: Carlos Amaro

The integration of experimental and computational methods can assist and enrich the interpretation of the experimental results, providing new detailed molecular understanding of the systems.

Keywords: Integrative structural biology ; Experimental techniques ; Computational methods.

---

## 1. Introduction

One of the main aims in molecular biochemistry is to obtain mechanistic insights into the function of biomolecules. To accomplish this, researchers must design experiments that provide new information about the molecule in question by using a variety of biochemical and biophysical techniques. Subsequently, the experimental data have to be correlated with the specific characteristics of the molecule under study. This process is sometimes a straightforward interpretation, but in many others cases, it is difficult to decipher the molecular meaning of the data. Consequently, one of the main roles of an experimentalist is to interpret the data to obtain new information on a specific molecular mechanism based on the results.

With the advent of new computational methods, one of the experimentalist desires is to be able to incorporate the experimental data into a detailed representation of the different mechanisms using in-silico modeling to assist and enrich the interpretation. This conjunction could provide researchers with a new detailed molecular understanding and allow for the proposal of more complete mechanisms.

## 2. Basic Strategies to Integrate Experiments and Computational Methods

Although the combination of computational methods and experiments has been used for a long time, for instance, computational approaches have relied on experimental data to calibrate the force fields, while experiments have used computing power to process and analyze the data, this review focused on the use of computational methods to assist in the interpretation of experimental results.

This combination of methods can use four major different strategies:

### 2.1. Independent approach.

Experimental and computational protocols are performed independently, and then the results of both methods are compared. The first step in a molecular simulation would consist of sampling different conformations, which can be performed using a detailed atomic or coarse-grained representation (less detailed). The sampling protocols can be molecular dynamics (MD), Monte Carlo simulation (MC), or any other sampling technique<sup>[2]</sup>. In the best case scenario, computational models and experimental data correlate and complement each other. However, on some occasions, the biomolecular process under investigation is a "rare" event, and therefore, successfully sampling this event using a simulation technique requires a global search of the entire conformational space, which could be challenging. To solve this problem, several variations to enhance the sampling of conformations such as replica exchange molecular dynamics, metadynamics, and accelerated MD have been developed. However, even with these advanced techniques, the sampling and accuracy of the generated structures are still bound by the limits of the force field and the theoretical model used, and sometimes the experimental data and the simulation do not correlate.

This independent approach has been by far the most explored method, and although extremely powerful, we are more interested in a more integrated approach.

## 2.2. Guided simulation (restrained) approach.

In a guided simulation, data obtained through experiments are used to effectively guide the three-dimensional conformation sampling in the computational method. This is usually done by the addition of external energy terms related to the experimental data into the computational protocol (restraints). Each restraint has its target value (experimental distribution), against which the back-calculated values would be compared in each simulation step. Since the guided methods involve evaluating the models during the simulation, they need to be implemented directly in the software. This type of guided simulation has been used in several programs<sup>[3]</sup>.

## 2.3. Search and select (reweighting) approach.

In a conceptually different strategy, the computation method is performed first to generate a large number of different conformation molecules (large ensemble), and then the experimental data are used to filter (search and select) the results. Only conformations that correlate with the experimental data are selected. The generation of the initial pool of conformations can be performed by any of the simulation sampling techniques already mentioned. Sometimes, even less computational demanding protocols are used, such as generating a large pool of random conformations or simulated annealing. Then, different protocols based on maximum entropy or maximum parsimony are used to select conformations that fit the data<sup>[4]</sup>.

## 2.4. Guided docking.

A different category of computational method would be molecular docking, which refers to methodologies that predict the final structure of a complex, starting with the structure of the two free molecules. Docking protocols are composed of two basic steps, a sampling algorithm to generate different binding conformations (poses) and a scoring process that assesses the quality of each pose<sup>[5]</sup>. In guided docking, the experimental data are used to help define the binding sites. In principle, the experimental data can be used either in the sampling or the scoring process<sup>[6]</sup>.

# 3. Integrative modelling

It should be noted that many of the experimental biophysical techniques report average values over many molecules and long periods of time. Consequently, a better correlation has often been observed with back-calculated data from an ensemble of conformations than with data from just a single conformer. All of the strategies listed above can be used to obtain an ensemble of conformations that are compatible with the set of experimental average values. Hence, a large number of programs have been created to select ensembles that fit the experimental data, differing in the way in which the initial ensemble is generated as well as in the algorithm used to search and select the final ensemble<sup>[7]</sup>.

The use of one strategy over the others would depend on the specific characteristics of each study. However, we can list some of the advantages and disadvantages that would make it more likely to choose one approach over the other. The computational sampling in the independent approach is not restricted to sample a specific region of the conformational space and therefore can provide information on “unexpected” conformations. Additionally, if one is interested in the specific sequential pathways of a process, unbiased sampling can provide a plausible pathway based on the physical model in which the computational method is based. On the other hand, one of the main advantages of the guided simulation approach is that the restraints considerably limit the conformational space and, in principle, the “observed experimental” conformations are sampled more efficiently. The main disadvantage of this approach is that the experimental data have to be implemented as a restraint during the sampling, and this could be a difficult task and in most cases would require certain computational knowledge.

In the search and select approach, the sampling process is uncoupled and is performed independently of the experimental data, and consequently the integration of different methodologies and more than one experimental restraint is simpler. Furthermore, it is possible to incorporate new experimental data without the need to generate a new conformational ensemble. One of the drawbacks would be that the initial pool must contain the “correct” conformations, and therefore it also requires a large sampling of the conformational space, however, several programs that easily generate a large pool of structure have been developed. Finally, if what one wants is to understand the formation of a complex, the best approach would probably be the use of guided docking.

In order to be able to integrate the experimental results into these approaches, it is necessary to compare the experimental data with a back-calculated value from the computational method.

## 4. Conclusion

Since the conformation of biomolecules undergoes variations with time and functional state, providing a detailed molecular description that incorporates these changes based solely on experimental results is a difficult task.

The integration of experimental data with computational techniques allows us to obtain a detailed interpretation of the results that would not be achievable using only experimental methods.

We are certain that the integration and applicability of some experimental techniques with computational methods are going to continue, and we anticipate new developments and integration with other experimental techniques.

---

## References

1. Reinier Cárdenas; Javier Martínez-Seoane; Carlos Amero; Combining Experimental Data and Computational Methods for the Non-Computer Specialist. *Molecules* **2020**, 25, 4783, [10.3390/molecules25204783](https://doi.org/10.3390/molecules25204783).
2. Tatiana Maximova; Ryan Moffatt; Buyong Ma; Ruth Nussinov; Amarda Shehu; Principles and Overview of Sampling Methods for Modeling Macromolecular Structure and Dynamics. *PLOS Computational Biology* **2016**, 12, e1004619-e1004619, [10.1371/journal.pcbi.1004619](https://doi.org/10.1371/journal.pcbi.1004619).
3. Stewart A. Adcock; J. Andrew McCammon; Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chemical Reviews* **2006**, 106, 1589-1615, [10.1021/cr040426m](https://doi.org/10.1021/cr040426m).
4. Simone Orioli; Andreas Haahr Larsen; Sandro Bottaro; Kresten Lindorff-Larsen; How to learn from inconsistencies: Integrating molecular simulations with experimental data. *Progress in Molecular Biology and Translational Science* **2020**, 170, 123-176, [10.1016/bs.pmbts.2019.12.006](https://doi.org/10.1016/bs.pmbts.2019.12.006).
5. Veronica Salmaso; Stefano Moro; Bridging Molecular Docking to Molecular Dynamics in Exploring Ligand-Protein Recognition Process: An Overview. *Frontiers in Pharmacology* **2018**, 9, 923, [10.3389/fphar.2018.00923](https://doi.org/10.3389/fphar.2018.00923).
6. P.I. Koukos; Alexandre M. J. J. Bonvin; Integrative Modelling of Biomolecular Complexes. *Journal of Molecular Biology* **2020**, 432, 2861-2881, [10.1016/j.jmb.2019.11.009](https://doi.org/10.1016/j.jmb.2019.11.009).
7. Massimiliano Bonomi; Gabriella T. Heller; Carlo Camilloni; Michele Vendruscolo; Principles of protein structural ensemble determination. *Current Opinion in Structural Biology* **2017**, 42, 106-116, [10.1016/j.sbi.2016.12.004](https://doi.org/10.1016/j.sbi.2016.12.004).

---

Retrieved from <https://encyclopedia.pub/entry/history/show/15175>