J Comput Aided Mol Des (2012) 26:11–12
DOI 10.1007/s10822-011-9488-z

PERSPECTIVE

Computational chemistry in pharmaceutical research: at the crossroads

Jürgen Bajorath

Received: 20 October 2011 / Accepted: 8 November 2011 / Published online: 15 November 2011

Abstract Computational approaches are an integral part of pharmaceutical research. However, there are many of unsolved key questions that limit the scientific progress in the still evolving computational field and its impact on drug discovery. Importantly, a number of these questions are not new but date back many years. Hence, it might be difficult to conclusively answer them in the foreseeable future. Moreover, the computational field as a whole is characterized by a high degree of heterogeneity and so is, unfortunately, the quality of its scientific output. In light of this situation, it is proposed that changes in scientific standards and culture should be seriously considered now in order to lay a foundation for future progress in computational research.

Keywords Computational chemistry · Molecular design · Key questions · Scientific heterogeneity · Scientific quality standards

Contemplating on the future of computational chemistry, broadly defined, is a formidable task. Perhaps it is already difficult enough to understand where we are right now. For example, let's consider the heterogeneity of this still evolving research field: theoretical chemistry, molecular modeling, computer-aided drug design, data mining, chem(o)informatics, and so on—where are the boundaries between these (sub)disciplines? How do we define one or the other? What are their distinguishing features? It would be difficult to clearly answer these questions.

For the sake of simplicity, my comments are limited to computational approaches in pharmaceutical research, without making further distinctions.

If one wants to look ahead, one should probably also look back and, in addition, reflect on the status quo. It is evident that some of the major unsolved computational problems in pharmaceutical research that we face today date back to the 1980s and 1990s. For example, it is still not possible to accurately model, from first principles, three-dimensional structures of protein–ligand complexes. Neither is it possible to accurately and consistently compute free energies of binding; nor to accurately predict bioactive conformations of ligands. One could go on. These are not exotic issues, but fundamental problems in drug design. They have been on the agenda already 20 years ago, are still essentially unsolved, and will likely be there for years to come. Solving "only" these exemplary problems would revolutionize the drug design field—and its impact on drug development.

Should we be hopeful? One might want to consider that we have not witnessed many true breakthroughs in the computational arena in recent years. Progress has essentially been incremental in most areas. However, it is rather unclear at present whether substantial methodological advances will still be possible, if the "right" ideas and concepts will be developed and implemented, or whether we might already have approached the limits of how accurately we can model the physical and chemical processes underlying, for example, protein–ligand interactions. Such questions might be difficult to answer.

How might one then be able to catalyze further scientific progress in this field? In any scientific area, progress will, first and foremost, depend on a rigorous scientific

J. Bajorath (✉)
Department of Life Science Informatics, B-IT, LIMES Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universität, Dahlmannstr. 2, 53113 Bonn, Germany
e-mail: bajorath@bit.uni-bonn.de

approach. This will in no way guarantee that breakthroughs will be made and key problems ultimately be solved. However, if scientific rigor is lacking, nothing substantial should be expected.

From this point of view, one should also reflect on the current status of our computational science. Unfortunately, the heterogeneity referred to above is not only a characteristic feature of this field as a whole; rather, it is also a characteristic of the quality of computational research. The absence of generally applicable community standards for the design and evaluation of computational studies has frequently been pointed out as well as the need for proper statistical validation of methodological benchmark investigations. Moreover, those of us who frequently review or edit computational papers can attest to the fact that well-meant computational studies often approach the realm of science fiction: unrealistic system setups provide perfect answers, calculation details are over-interpreted, results of hypothetical computations are treated like facts, highly complex computational recipes yield trivial findings; to name a few observations that can often be made. For example, let's consider the use of "double-hypothetical" models, e.g., ligands are docked into putative binding sites of receptor models and the resulting hypothetical complexes are subjected to energy calculations. What does it mean scientifically if free energies are accurately produced under such conditions? An exercise, we frequently witness.

Maybe computational chemistry and modeling have become too easy at the technical level; maybe it has become readily possible to carry out advanced computations without the need to critically assess their scientific foundations and limitations. Whatever the reasons might be for the often questionable scientific quality of computational studies and the apparent absence of rigor, these issues negatively affect the field as a whole and most definitely work against its further progress.

Taking these problems into account, I would propose that short- to mid-term progress in further increasing quality standards for computational studies will be of critical importance to lay a foundation for significant future advances, despite all inherent uncertainties. And the news is not all bad. Several community initiatives have evolved over the years to assess computational predictions on a larger scale, define uniform evaluation criteria, and formulate goals for incremental improvements of computational methods. Moreover, journals and their editors are called upon to further evolve scientific publication standards, educational institutions to refine teaching programs, and -last but not least- individual investigators to strive for quality improvements, at all levels, including doctoral and postdoctoral education.

Certainly, increasingly competitive funding situations and grant pressure do not necessarily help to ensure high scientific quality of computational research. These days it is much easier to carry out -and report- calculations applying pre-formulated computational recipes than to be innovative and develop novel concepts and methods. Nevertheless, raising the bar will be essential for the further development of computational chemistry and molecular design. From this point of view, the field is at the crossroads.

As long as there is evidence that further increasing scientific rigor in the development and application of computational methods might provide a basis for the field to reach new levels, there should be hope for true breakthroughs to occur. Again, many open scientific questions that hamper computational analysis and design and their impact on discovery are not new and one should hence probably be careful not to expect too much too soon. However, if one would indeed be able to, for example, accurately predict ligand binding conformations or complex structures, regardless of their specific features, and accurately calculate binding energies in, let's say, 10 or 15 years from now, we would very likely experience a revolution in our field, with unprecedented impact on the life sciences. To get there, the computational chemistry community will probably need to follow an evolutionary path characterized by incremental advances and steadily increasing scientific rigor and quality. Innovation will of course play a key role during this process, provided it will be based on solid scientific foundations. From my point of view, it would be truly worth doing some homework now and trying to push the limits as we go along. Let's hope for good things to come.