J Comput Aided Mol Des (2012) 26:9–10
DOI 10.1007/s10822-011-9516-z

PERSPECTIVE

Computational chemistry in 25 years

Ruben Abagyan

Received: 18 November 2011 / Accepted: 5 December 2011 / Published online: 13 December 2011

Abstract Here we are making some predictions based on three methods: a straightforward extrapolations of the existing trends; a self-fulfilling prophecy; and picking some current grievances and predicting that they will be addressed or solved. We predict the growth of multicore computing and dramatic growth of data, as well as the improvements in force fields and sampling methods. We also predict that effects of therapeutic and environmental molecules on human body, as well as complex natural chemical signalling will be understood in terms of three dimensional models of their binding to specific pockets.

Keywords Computational chemistry · Structure prediction · Force field · Drug discovery · Ligand docking · Green chemistry

Predicting future is a lot of fun. No responsibility and a chance to inspire, scare, or misguide people (provided someone really cares). However the Editors deprived us from the Nostradamus/Steven Hawking time scale and apocalyptic intensity by insisting on a relatively short twenty-five year projection. This time span is comfortably close to being somewhat predictable by the following three methods:

- a straightforward extrapolations of the existing trends;
- a self-fulfilling prophecy: meaning that the writer is planning to do exactly that during the next 25 years;
- picking some current grievances and predicting that they will be addressed or solved.

The first two items are relatively easy, while the last one has a certain inevitable randomness about it. We should also be conscious about factors external to the field. Will the funding, educational, science-geopolitical environment of our research be favorable of hostile? Our ability to reach our goals clearly depends on it. If most of the brains will continue to be siphoned into Wall Street offices and the funding dries up or moves to China the field will have to hibernate or learn another language. Our still energetic 39 years old community (arbitrarily counting from the birth of PDB circa 1971) should ask the Public "Will you still need me will you still feed me when I am 64?"

Assuming that we will still be fed, powered and networked, and any of the assumptions may break if an unexpected limit is reached, let us embark on a thought experiment.

Billion transistors per microprocessor. We all know Moore's law. So, if in the last 25 years the transistor count grew from 10^5th per microprocessor to 10^9th, so in 2036 it will grow to about 10^{12}–10^{13}.

Thousands of cores for processor. We know that the number of cores in processors went up from one to now fifty cores in Intel Knights Ferry, hence in 2036 the chemistry software applications will have to be re-engineered for processors with around a thousand cores. However, I am not so sure about GPU computing since it is not really designed or software-populated for our kind of computing.

Lo-o-o-ng dynamic trajectories for larger systems. As a simple derivative of the previous two rules we can predict that molecular dynamics runs will be routinely generating millisecond trajectories for medium size proteins or seconds for smaller systems on mainstream high-performance hardware.

R. Abagyan (✉)
Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, CA, USA
e-mail: rabagyan@ucsd.edu

Mobile device interface. More software will be developed for iPad-like mobile devices that interact with a powerful compute server and distributed data sources.

Dramatic growth of data. Explosive growth of experimental data about molecules, biomacromolecules and their assemblies and methods/models based on these data.

The fast growth of useful experimental data on chemical matter, biological and environmental will lead to a different type of computing, where the data and the models get mature in a few powerful servers with distributed file systems. The future models will in part be derived from and heavily depend upon these data.

The structural databases. The Protein Data bank will contain several million entries and with tens of thousands of unique proteins. It finally will be served as an animated interactive entity in which the electron density in critical parts, dynamics, ambiguities and multiple solutions will be presented and the functions described as a story, …—on second thought,—no,—the PDB will still be served the same limited way.

The funding agencies will wake up and ban the deposition of small molecule crystallographic information into an unavailable private database,—the small molecule structures will finally become as public as the PDB.

Realistic atomic models. Fully protonated atomic models of proteins, macromolecules and assemblies with dynamic protonation and tautomerization states and with explicit representation of conformational diversity and induced fit will become important starting models for drug discovery and structure-based medicinal chemistry. The treatment of various transitions in small molecules and biomacromolecules including switches between stereo isomers, tautomers, cis–trans states, (de)protonation, ring puckering, will become routine and much more realistic.

Larger role of 3D models in medicinal chemistry. Computational chemistry will play a larger role in small molecule drug discovery and medicinal chemistry, which in turn, will become predominantly structure-based.

QM and statistical mechanics will be sufficient as the theoretical basis. The quantum mechanics and its improved approximations and implementations, will provide a sufficient theoretical and software basis for understanding molecular structure, interactions, transformations and reactions. The mechanical models and force fields for molecular mechanics will be re-parameterized using the high level QM calculations and will become much more realistic in describing the relative interaction energies and statistical properties of conformational ensembles.

Some self-fulfilling prophecies

The following is partially contingent upon our mini-community being able to stay healthy and funded. Twenty years ago we developed the internal coordinate mechanics, ICM, approach and an efficient global stochastic conformation sampler (including collective-variable/fragment moves and the weighted square root sampling) to be able to predict how molecules fold or dock and indeed over this period ICM and ICM-like methods fold peptides and small proteins, dock compounds and peptides, and predict how proteins associate. Most of these goals were achieved. In the next 25 years we will keep improving the force field and develop our internal coordinate sampling methods to find the conformational distributions and their relative energies for the critical events in the lives of our molecular machines.

Activity predicting Pocketome. During the next 25 years we will develop and refine a comprehensive encyclopedia of the binding pockets on various macromolecules, and interfaces together with their conformations, tautomeric, protonation and functional states, as well as cellular concentrations and localizations. These flexible pocket models will be able to predict with high accuracy if a new chemical matter can bind to this pocket and with what affinity. Combined with the experimental data it will provide a powerful framework for understanding biology.

Understanding metabolites and what they do in our body. We will be able to understand and predict what approximately a hundred thousand man-made or nature-made chemicals in our environment and some in our body do to us and to other living creatures on the planet. We will be able to predict how our body generates new metabolites from food and molecules in the environment. We will be able to predict their adverse effects and be more efficient in preventing or at least warning politicians, who will be even more corrupt in 25 years, about the existential dangers of some of the new chemicals.

In conclusion, here is an equally valid perspective from my friend to whom I mentioned about writing this commentary. His opinion was : "… nothing will happen in 25 years, I have been in this field since 1994 and there has been no major change".

Acknowledgments The author had a pleasure to work and talk about some of these matters with several inspired individuals including but not limited to Max Totrov, Eugene Raush, Irina Kufareva, Andrew Bordner, Brian Marsden, Pete Myers and Bruce Blumberg.