

SpringerBriefs in Applied Sciences and Technology

Silvano Coletti · Gabriella Bernardi *Editors*

Exscalate4CoV

High-Performance
Computing for COVID Drug
Discovery

SpringerBriefs in Applied Sciences and Technology

SpringerBriefs present concise summaries of cutting-edge research and practical applications across a wide spectrum of fields. Featuring compact volumes of 50 to 125 pages, the series covers a range of content from professional to academic.

Typical publications can be:

- A timely report of state-of-the art methods
- An introduction to or a manual for the application of mathematical or computer techniques
- A bridge between new research results, as published in journal articles
- A snapshot of a hot or emerging topic
- An in-depth case study
- A presentation of core concepts that students must understand in order to make independent contributions

SpringerBriefs are characterized by fast, global electronic dissemination, standard publishing contracts, standardized manuscript preparation and formatting guidelines, and expedited production schedules.

On the one hand, **SpringerBriefs in Applied Sciences and Technology** are devoted to the publication of fundamentals and applications within the different classical engineering disciplines as well as in interdisciplinary fields that recently emerged between these areas. On the other hand, as the boundary separating fundamental research and applied technology is more and more dissolving, this series is particularly open to trans-disciplinary topics between fundamental science and engineering.

Indexed by EI-Compendex, SCOPUS and Springerlink.

Silvano Coletti · Gabriella Bernardi
Editors

Exscalate4CoV

High-Performance Computing for COVID
Drug Discovery

Editors

Silvano Coletti
Chelonia SA
Lugano, Ticino, Switzerland

Gabriella Bernardi
Lugano, Ticino, Switzerland

ISSN 2191-530X

ISSN 2191-5318 (electronic)

SpringerBriefs in Applied Sciences and Technology

ISBN 978-3-031-30690-7

ISBN 978-3-031-30691-4 (eBook)

<https://doi.org/10.1007/978-3-031-30691-4>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

*To Thales the primary question was not
‘What do we know?’ but ‘How do we
know it?’*

—Aristotle

Foreword

Computers as Primary Tools in Drug Discovery

The ability to correlate the structure of protein–ligand complexes with their binding affinity has been a long-standing dream and challenge of computational chemists. This challenge reflects the need for effective tools in the rational design of drugs that would block the activity of proteins involved in devastating diseases. Obviously, reaching accurate predictions in calculations of binding energies is one of the most important aims of computational biology. However, progressing in this direction has been far from simple. The starting point for addressing this challenge has been the progress in structural evaluations of protein–ligand complexes. The enormous advances on this front call for corresponding progress in the use of computer modeling to estimate the binding affinity of protein–ligand complexes.

The corresponding progress in estimating binding energies has been very slow. Qualitative quantitative structure–activity relationship (QSAR) approaches were reported quite early, followed by very primitive attempts to compute the energetics of ligand–protein complexes. Some of the attempts involved problematic focus on studies of the ligand without the protein. Others included the protein but did not consider the solvent around the protein and the energetics of the ligand in water. The use of thermodynamics in cycles in computations of protein substrate systems emerged in 1981 and began to be applied with simplified solvent models. More rigorous free energy perturbation (FEP) studies of protein–ligand interactions emerged in 1984 but progressed very slowly, except in the calculation of enzymatic reactions. The problem has been that such calculations face enormous convergence problems when applied to absolute binding free energies. The situation is more reasonable when one focuses on evaluating substitutions of small parts of the ligand.

As far as evaluations of absolute binding energies are concerned, it appeared that, in many cases, macroscopic and semi-macroscopic models gave more stable results than the corresponding microscopic models did. These models included approaches that are based on the linear response approximation, such as the protein–dipole Langevin–dipole model with the linear response approximation treatment with a

scaled non-electrostatic term (PDL/D/S-LRA/ β) and the significantly less consistent molecular mechanics generalized Born surface area (MM/GBSA) approach.

Another major problem is the evaluation of the effect of water penetration. Not only is it difficult in such cases to identify the position of the water molecules, but their positions also change drastically during the binding process. In principle, one can try to insert water molecules in a grand canonical Monte Carlo treatment, but the chance of water molecules being accepted is extremely small because, in most insertion attempts, we will not have a large enough cavity in the protein site. Our solution to this problem has been the “water flooding” approach that effectively reproduces grand canonical results. Other important directions are the use of computations to determine protein–protein interactions, whereas at present, the best option is to use coarse-grained (CG) models. Another promising new direction is the use of machine learning. This direction is “orthogonal” to the use of physical-based modeling because the physics is hidden in the data, but the potential of exploiting big data is very promising.

Despite the current problems, there is no doubt that computer-aided calculations will eventually become the key quantitative tool in rational drug design. Once a more quantitative level for calculating protein–ligand binding is achieved, we will see several promising directions. For example, it will be possible to augment automated screening procedures and greatly accelerate the refinement of the final drug candidates.

Major advances in computer-aided drug design will clearly emerge in the future. The only question is how long it will take to make such approaches quantitative. We believe that strategies that go below the 1 kcal/mol error limit would be sufficiently quantitative and advance the field toward a new era.

The enormous increase in computer power in the past decade is leading to a paradigm shift in computer-aided drug design. We are probably approaching a stage in which the computing power will be translated to reliable performance, and drug screening by computers will be more effective than experimental screening. Here, the Exscalate4COV project and other projects described in this book mark the emergence of such directions.

Open-access collaboration is another promising path wherein the effort of a large scientific community can greatly help push the field forward. Educating large parts of the community regarding the power of computers in drug design also holds great value. Educating future medical professionals and future decision-makers will substantially contribute to realizing the aim of using computers as the key tool in developing new medications.

Arieh Warshel
2013 Nobel Prize winner in Chemistry
University of Southern California
Los Angeles, CA, USA

Reference

1. A. Warshel, Theoretical studies of drug receptor interaction. Trends in Biochem. Sci. **1**, N105–N106 (1976)