A PRACTICAL INTRODUCTION TO THE SIMULATION OF MOLECULAR SYSTEMS Second Edition

Molecular simulation is a powerful tool in materials science, physics, biophysics, chemistry, drug design and many other areas of research. This updated edition provides a pragmatic introduction to a wide range of techniques for the simulation of molecular systems at the atomic level. The first part of the book concentrates on methods for calculating the potential energy of a molecular system, with new chapters on quantum chemical, molecular mechanical and hybrid potential techniques. The second part covers ways of investigating the conformational, dynamical and thermodynamical properties of systems, discussing such techniques as geometry-optimization, normal-mode analysis, and molecular dynamics and Monte Carlo simulation.

Now employing Python, the second edition includes a wealth of examples and program modules for each simulation technique. This allows readers to carry out the calculations and to appreciate the inherent difficulties involved in each. This is a valuable resource for researchers and graduate students wanting to know how to perform atomic-scale molecular simulations.

Additional resources for this title, including the program library, technical information and instructor-solutions, are available online at www.cambridge.org/ 9780521852524.

MARTIN J. FIELD is Group Leader of the Laboratoire de Dynamique Moléculaire at the Institut de Biologie Structurale – Jean-Pierre Ebel, Grenoble. He was awarded his Ph.D. in Quantum Chemistry from the University of Manchester, UK, in 1985. His areas of research include using molecular modeling and simulation techniques to study biological problems. More specifically, his current interests are in the development and application of hybrid potential techniques to study enzymatic reaction mechanisms and other condensed phase processes.

Reviews of the first edition:

"... a valuable teaching aid for those presenting this topic. It should be of interest not only to the physical chemist, but also to those involved in computational biophysics, biochemistry or molecular physics." *Scientific Computing World*

'... this book is a valuable addition to my shelf and one that I must make sure doesn't disappear because my research group has taken off with it!' *Nell L. Allan. Chemistry and Industry*

A PRACTICAL INTRODUCTION TO THE SIMULATION OF MOLECULAR SYSTEMS

Second Edition

MARTIN J. FIELD

Institut de Biologie Structurale – Jean-Pierre Ebel, Grenoble, France



> CAMBRIDGE UNIVERSITY PRESS Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

> > Cambridge University Press The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org Information on this title: www.cambridge.org/9780521852524

© Martin J. Field 2007

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2007

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-85252-4 hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

Preface to the first edition		<i>page</i> ix	
Preface to the second edition			xi
	_		
1	Pre	liminaries	1
	1.1	Introduction	1
	1.2	Python	2
	1.3	Object-oriented programming	5
	1.4	The pDynamo library	8
	1.5	Notation and units	9
2	Che	emical models and representations	14
	2.1	Introduction	14
	2.2	The System class	14
	2.3	Example 1	17
	2.4	Common molecular representations	18
	2.5	Example 2	27
3	Coo	ordinates and coordinate manipulations	31
	3.1	Introduction	31
	3.2	Connectivity	31
	3.3	Internal coordinates	35
	3.4	Example 3	38
	3.5	Miscellaneous transformations	41
	3.6	Superimposing structures	45
	3.7	Example 4	47
4	Qua	antum chemical models	51
	4.1	Introduction	51
	4.2	The Born–Oppenheimer approximation	51
	4.3	Strategies for obtaining energies on a potential energy surface	53

vi	Contents	
	 4.4 Molecular orbital methods 4.5 The Hartree–Fock approximation 4.6 Analysis of the charge density 4.7 Example 5 4.8 Derivatives of the potential energy 4.9 Example 6 	54 56 67 70 74 78
5	Molecular mechanics5.1Introduction5.2Typical empirical energy functions5.3Calculating a molecular mechanics energy5.4Example 75.5Parametrizing potential energy functions5.6Soft constraints	81 81 93 101 103 105
6	Hybrid potentials6.1Introduction6.2Combining QC and MM potentials6.3Example 86.4Covalent bonds between QC and MM atoms6.5Example 9	110 110 110 114 116 120
7	 Finding stationary points and reaction paths on potential energy surfaces 7.1 Introduction 7.2 Exploring potential energy surfaces 7.3 Locating minima 7.4 Example 10 7.5 Locating saddle points 7.6 Example 11 7.7 Following reaction paths 7.8 Example 12 7.9 Determining complete reaction paths 7.10 Example 13 	122 122 126 129 130 134 136 139 140 144
8	 Normal mode analysis 8.1 Introduction 8.2 Calculation of the normal modes 8.3 Rotational and translational modes 8.4 Generating normal mode trajectories 8.5 Example 14 8.6 Calculation of thermodynamic quantities 8.7 Example 15 	148 148 148 153 156 158 161 165

		Contents	vii
9	Mole	ecular dynamics simulations I	170
	9.1	Introduction	170
	9.2	Molecular dynamics	170
	9.3	Example 16	178
	9.4	Trajectory analysis	182
	9.5	Example 17	184
	9.6	Simulated annealing	186
	9.7	Example 18	189
10	More	e on non-bonding interactions	195
	10.1	Introduction	195
	10.2	Cutoff methods for the calculation of non-bonding interactions	195
	10.3	Example 19	205
	10.4	Including an environment	209
	10.5	Periodic boundary conditions	212
	10.6	Example 20	215
	10.7	Ewald summation techniques	217
	10.8	Fast methods for the evaluation of non-bonding interactions	223
11	Mole	ecular dynamics simulations II	225
	11.1	Introduction	225
	11.2	Analysis of molecular dynamics trajectories	225
	11.3	Example 21	233
	11.4	Temperature and pressure control in molecular dynamics	
		simulations	235
	11.5	Example 22	244
	11.6	Calculating free energies: umbrella sampling	246
	11.7	Examples 23 and 24	252
	11.8	Speeding up simulations	258
12	Mon	te Carlo simulations	262
	12.1	Introduction	262
	12.2	The Metropolis Monte Carlo method	262
	12.3	Monte Carlo simulations of molecules	266
	12.4	Example 25	277
	12.5	Calculating free energies: statistical perturbation theory	280
	12.6	Example 26	286
Aj	ppend	ix 1 The pDynamo library	294
Aj	ppend	ix 2 Mathematical appendix	298
	A2.1	The eigenvalues and eigenvectors of a matrix	298
	A2.2	The method of Lagrange multipliers	300

viii Contents	
Appendix 3 Solvent boxes and solvated molecules	302
A3.1 Example 27	302
A3.2 Example 28	305
Bibliography	307
Author index	326
Subject index	330

Preface to the first edition

The reason that I have written this book is simple. It is the book that I would have liked to have had when I was learning how to carry out simulations of complex molecular systems. There was certainly no lack of information about the theory behind the simulations but this was widely dispersed in the literature and I often discovered it only long after I needed it. Equally frustrating, the programs to which I had access were often poorly documented, sometimes not at all, and so they were difficult to use unless the people who had written them were available and preferably in the office next door! The situation has improved somewhat since then (the 1980s) with the publication of some excellent monographs but these are primarily directed at simple systems, such as liquids or Lennard-Jones fluids, and do not address many of the problems that are specific to larger molecules.

My goal has been to provide a practical introduction to the simulation of molecules using molecular mechanical potentials. After reading the book, readers should have a reasonably complete understanding of how such simulations are performed, how the programs that perform them work and, most importantly, how the example programs presented in the text can be tailored to perform other types of calculation. The book is an *introduction* aimed at advanced undergraduates, graduate students and confirmed researchers who are newcomers to the field. It does not purport to cover comprehensively the entire range of molecular simulation techniques, a task that would be difficult in 300 or so pages. Instead, I have tried to highlight some of the basic tasks that can be done with molecular simulations and to indicate some of the many exciting developments which are occurring in this rapidly evolving field. I have chosen the references which I have put in carefully as I did not want to burden the text with too much information. Inevitably such a choice is subjective and I apologise in advance to those workers whose work or part of whose work I did not explicitly acknowledge.

There are many people who directly or indirectly have helped to make this book possible and whom I would like to thank. They are: my early teachers in the

х

Preface to the first edition

field of computational chemistry, Nicholas Handy at Cambridge and Ian Hillier at Manchester; Martin Karplus and all the members of his group at Harvard (too numerous to mention!) during the period 1985–9 who introduced me to molecular dynamics simulations and molecular mechanics calculations; Bernie Brooks and Rich Pastor, at the NIH and FDA, respectively, whose lively discussion and help greatly improved my understanding of the simulations I was doing; and all the members of my laboratory at the IBS, past and present, Patricia Amara, Dominique Bicout, Celine Bret, Laurent David, Lars Hemmingsen, Konrad Hinsen, David Jourand, Flavien Proust, Olivier Roche and Aline Thomas. Finally, special thanks go to Patricia Amara and to Dick Wade at the IBS for comments on the manuscript, to Simon Capelin and the staff of Cambridge University Press for their guidance with the production of the book, to the Commissariat à l'Energie Atomique and the Centre National de la Recherche Scientifique for financial support and to my wife, Laurence, and to my sons, Mathieu and Jeremy, for their patience.

> Martin J. Field Grenoble, 1998

Preface to the second edition

This edition of *A Practical Introduction* has two major differences from the previous one. The first is a discussion of quantum chemical and hybrid potential methods for calculating the potential energies of molecular systems. Quantum chemical approaches are more costly than molecular mechanical techniques but are, in principle, more 'exact' and greatly extend the types of phenomena that can be studied with the other algorithms described in the book. The second difference is the replacement of FORTRAN 90 by Python as the language in which the DYNAMO module library and the book's computer programs are written. This change was aimed to make the library more accessible and easier to use. As well as these major changes, there have been many minor modifications, some of which I wanted to make myself but many that were inspired by the suggestions of readers of the first edition.

Once again, I would like to acknowledge my collaborators at the Institut de Biologie Structurale in Grenoble and elsewhere for their comments and feedback. Special thanks go to all members, past and present, of the Laboratoire de Dynamique Moléculaire at the IBS, to Konrad Hinsen at the Centre de Biophysique Moléculaire in Orléans and to Troy Wymore at the Pittsburgh Supercomputing Center. I would also like to thank Michelle Carey, Anna Littlewood and the staff of Cambridge University Press for their help during the preparation of this edition, Johny Sebastian from TechBooks for answers to my many LATEX questions, and, of course, my family for their support.

Martin J. Field Grenoble, 2006

xi