

Diss. ETH Nr. 12603

**Molecular Dynamics Simulation of
Biomolecules: Software Engineering,
Parallelisation and Methodology**

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH

for the degree of
Doctor of Natural Sciences

presented by
Walter Robert Peter Scott
Dipl. Informatik Ing. ETH

born 12th October 1964
British nationality

accepted by recommendation of
Prof. Dr. W.F. van Gunsteren, examiner
Prof. Dr. U.W. Suter, co-examiner



CatE

Kurzfassung

Die Simulation von molekularen Systemen ist ein Gebiet von wachsender Bedeutung. Trotz grosser Erfolge in der Modellierung von Systemen praktischer Bedeutung ist die Anwendung von Simulationstechniken jedoch in vielen Fällen immer noch nicht möglich. In dieser Arbeit werden Aspekte der Softwareentwicklung, der Parallelisierung und der simulationstechnischer Methodenentwicklung für Molekulardynamik von Biomolekülen beleuchtet.

Kapitel 2 stellt die Entwicklung eines biomolekularen Simulationspaketes, GROMOS96, dar. Einzelheiten betreffend die Simulationsalgorithmen, das Kraftfeld, die Zuverlässigkeit und Effizienz werden beschrieben, unter besonderer Berücksichtigung der gegenüber dem Vorgängerpaket, GROMOS87, neuimplementierten Funktionalität. Dazu dienen zwei Anwendungsbeispiele: Local Elevation Suche einer Schlaufenregion von Ribonuklease A und eine Konformationssuche in vier Raumdimensionen von Cyklosporin A.

Die Aufteilung eines mit hohem Berechnungsaufwand verbundenen Problemes auf mehrere Prozessoren ist ein Ansatz es zu lösen. Auf dem durch die Flieskompleistung beschränkten Gebiet der Molekulardynamik erlaubt die gleichzeitige Benutzung von mehreren Prozessoren die Grenze der möglichen Systemgrösse und/oder Simulationsdauer nach oben zu verschieben. Die effiziente Ausnutzung von Parallelarchitekturen ist jedoch mit Problemen behaftet. In Kapitel 3 wird der an der ETH Zürich entwickelte Parallelrechner MUSIC und die mit diesem erzielte Leistung eines Molekulardynamik Simulationsprogrammes beschrieben.

Kapitel 4 befasst sich mit der Berechnung der Differenz der Freien Energie zwischen Konformationen molekularer Systeme. Die Zuverlässigkeit von Berechnungen dieser Art leidet häufig unter ungenügendem Sampling des Konformationsraumes wegen der kleinen Wahrscheinlichkeit, Energiebarrieren in der Energiefläche zu überwinden. Sampling-Verbesserungsmethoden werden oft angewandt, um diesem Problem zu begegnen. Eine solche kürzlich entwickelte Methode ist die Local Elevation Methode, die aber als Nachteil für die Berechnung freier Energiedifferenzen hat, dass das System sich nicht im Gleichgewicht befindet. Ein allgemeines Konzept, mit dessen Hilfe adaptive Nicht-Gleichgewichts Sampling-Verbesserungsmethoden auf die Berechnung freier Energiedifferenzen angewandt werden können, wird vorgeschlagen und anhand einer Local Elevation Studie von Butan aufgezeigt.

In Kapitel 5 wird das Problem der Bestimmung der Molekularstruktur aus experimentell bestimmten Grössen angegangen. Um die Übereinstimmung zwischen Computersimulationen auf der einen Seite und experimentellen Bestimmungsgrössen (z.B. Overhauser Intensitäten oder Abstände, 3J Kopplungskonstanten, chemische Verschiebungen, kristallographische B-Faktoren) auf der anderen Seite zu verbessern, werden häufig Kopplungsbedingungen in Simulationen eingebracht. Die erzwungenen Kopplungsbedingungen sollten der Art und Wei-

se, mit der die experimentellen Grösse bestimmt worden sind, Rechnung tragen. Insbesondere muss dem Umstand besondere Beachtung geschenkt werden, dass experimentelle Werte Zeit- und Raummittelwerte darstellen. Früher vorgestellte zeitmittelnde Strukturbestimmungsmethoden mit 3J Kopplungskonstanten weisen künstlich induzierte Fluktuationen der Struktur als Nachteil auf. Die in dieser Arbeit entwickelten Methoden begegnen diesem Problem, was anhand eines zyklischen Dekapeptids, Antamanid, aufgezeigt wird.

Die Arbeit schliesst mit einem kurzen Blick auf die Zukunftsperspektiven der Molekulardynamik.

Summary

Molecular simulation has become a field of increasing importance over the last decades. However, despite the increasing number of successful simulation studies of practical relevance, the application of molecular simulation is in many cases still impractical. In this work, aspects of software engineering, parallelisation and methodology in molecular dynamics simulations of biomolecules are presented.

In chapter 2, the development of the GRONingen MOlecular Simulation program package, GROMOS96, is outlined. Details pertaining to the simulation algorithm, the force field, the reliability and efficiency of the program are described. The new functionality with respect to its predecessor, GROMOS87, is demonstrated on two examples: local elevation search in a loop region of ribonuclease A and conformational search in four spatial dimensions of cyclosporin A.

Bringing the power of multiple processors to bear on a computationally difficult problem is one approach to solving an otherwise intractable task. In the field of molecular dynamics, where applications are limited by the floating point performance of computers, the computing power of multiple processors can be harnessed to increase the size of molecular system and/or time scale amenable to simulation. However, the efficient use of multiple processors poses a difficult problem. In chapter 3, the parallel computer architecture MUSIC developed at the ETH Zürich is described. The implementation of a molecular dynamics simulation program and the achieved performance is discussed.

In chapter 4, the calculation of free energy differences between conformations of molecular systems is addressed. These calculations often suffer from poor sampling of the relevant configurational space due to the system's reluctance to overcome barriers in the energy surface. For this reason, sampling enhancement methods are often applied. One such recently developed method, local elevation, has the disadvantage that it results in non-equilibrium dynamics. A scheme by which non-equilibrium adaptive sampling enhancement methods can be used for free energy calculations is detailed and demonstrated using local elevation on butane as a test system.

Chapter 5 is concerned with the refinement of molecular structure. In order to improve agreement between simulations and experimentally determined parameters (e.g. nuclear Overhauser intensities or distances, 3J coupling constants, chemical shifts, crystallographic structure factors), restraints are often introduced into molecular simulations. The restraints should be introduced in a way that is consistent with the way in which the experimental values were determined. In particular, due regard must be given to the fact that experimentally determined values are averages over time and space. Previous time-averaging refinement methods proposed for structure refinement using experimental 3J coupling constants can suffer from large artificially induced structural fluctuations. The newly developed methods, outlined in Chapter 5, alleviate this problem. Examples are

given on the cyclic decapeptide antamanide.

In conclusion, prospects for future developments in the field of molecular dynamics simulations are briefly touched upon.