Molecular Orbital Calculations for Amino Acids and Peptides

Anne-Marie Sapse

Molecular Orbital Calculations for Amino Acids and Peptides

With 32 Figures

Springer Science+Business Media, LLC

Anne-Marie Sapse
John Jay College and Graduate School
City University of New York
New York, NY 10019
and
Rockefeller University
New York, NY 10021
USA

Library of Congress Cataloging-in-Publication Data
Sapse, Anne-Marie.
Molecular orbital calculations for amino acids and peptides / Anne-Marie Sapse.
p. cm.
Includes bibliographical references and index.
ISBN 978-1-4612-7109-3 ISBN 978-1-4612-1354-3 (eBook)
DOI 10.1007/978-1-4612-1354-3
1. Amino acids. 2. Peptides. 3. Molecular orbitals. I. Title.
QD431.S257 1999
547'.750448—dc21 99-26375
CIP

Printed on acid-free paper.

Originally published by Birkhäuser Boston in 2000
Softcover reprint of the hardcover 1st edition 2000

ISBN 978-1-4612-7109-3
SPIN 19901572

Typeset by Best-set Typesetter Ltd., Hong Kong.

9 8 7 6 5 4 3 2 1

To my husband, Marcel Sapse, and to my daughter, Danielle Sapse, without whose support I could not have written this book.

Contents

Preface

This book is intended mainly for biochemists who would like to augment experimental research in the domain of amino acids and small peptides with theoretical calculations at the ab initio level.

The book does not require a profound knowledge of mathematics and quantum chemistry. It teaches one rather how to use computer software such as the Gaussian programs and gives examples of problems treated in this manner.

Chapter 1 describes the calculations and one of the programs used for ab initio work.

Chapter 2 describes calculations on small amino acids, such as glycine and alanine.

Chapter 3 discusses the biochemical properties of GABA (gamma amino butyric acid), which is one of the most important amino acids of the nervous system. Ab initio calculations performed in order to study the structure of GABA are presented.

Chapter 4 discusses an amino acid related to GABA, namely DABA (diaminobutyric acid), presenting information about its structure and transport properties.

A number of amino acids, essentials in the biochemistry of organisms, are discussed in Chapter 5. These acids have been subjected to ab initio investigation. Proline, a special amino acid as far as structure is concerned, is discussed in Chapter 6.

Chapter 7 discusses two sulfur-containing amino acids, taurine and hypotaurine, presenting some experimental studies on their mode of action and an ab initio study of their structure.

Starting with Chapter 8, small peptides of great importance are discussed. Glucagon, a small peptide that plays a role in diabetes, is the subject of Chapter 8.

Chapter 9 discusses the pheromone alpha factor, from an experimental and theoretical point of view.

Chapter 10 presents calculations on tight turns in proteins.

Chapter 11 discusses some small peptides that have been studied with ab initio methods.

Oligopeptides that feature anticancer activity, such as lexitropsins, are discussed in Chapter 12.

The book is addressed to graduate and postgraduate students as well as other researchers in the amino acid and peptide area.

New York, NY Anne-Marie Sapse

Introduction

Knowledge about the origin of life requires the recapitulation of the steps of archaic molecular evolution. According to the protenoid model, proteinoids (copolyaminoacids) arose on earth from mixtures of self-sequencing amino acids. The structure of amino acids, of the peptides formed by their polymerization via the formation of peptidic bonds, as well as the structure of the proteins that are polypeptide chains in various numbers and conformations, have formed the subject of an enormous number of experimental and theoretical studies.

At present, both theoretical and experimental methods are taken seriously as useful sources of information. They compare results and confirm or dispute structural findings. While experimental results are usually not doubted, and computational results depend on such parameters as the quality of the basis sets used, there have been instances in which computational results have contradicted experimental ones regarding structural determination. However, in most instances the two types of methods complement each other. For instance, a laboratory search for intermediates in certain reactions can be avoided once large basis-set calculations show the intermediates not to be a stationary state, more exactly, a minimum on the energy hypersurface.

The application of computational methods to biological systems dates from the 1950s, when the pioneering work of Bernard and Alberte Pullman was first published. The biological systems studied with the quantum-chemical methods available at that time had to be small, and not all the conclusions derived were correct. However, this work opened the door to a whole new area of research.

The basic problem in the determination of the structure of biological systems is their size. In order to be able to handle such molecules as the nucleic acids or the proteins, new theoretical methods had to be developed, and the quantum-chemical methods, ab initio and semi-empirical, were augmented by the molecular mechanics method, which uses experimental parameters in order to determine the force fields of the systems.

Huge strides have been made in the development of computer programs that handle larger systems. Researchers are striving to find the optimum combination of accuracy and expediency, with the ultimate goal being the reduction of computational effort with no loss of accuracy.

All three of these types of theoretical methods are used in the description of amino acids and peptides. The size of proteins precludes the use of ab initio or semiempirical methods, so they are mainly described with computer modeling, with programs such as Sybil, Quanta, and Insight, augmented by energy calculations with the Charmm program and other molecular-mechanic calculations.

The primary structure of proteins, characterized by the amino acid composition and sequence, is determined experimentally by degradation via hydrolysis of the peptidic bonds. The classic method of determining the sequence involves Edman degradation, which is an end-labeling procedure. Physical methods used include mass spectrometry and nuclear magnetic resonance (NMR). Since the 1980s, sequencing of proteins has been performed by sequencing its mRNA or gene.

The three-dimensional structures of about 800 proteins have been determined by Max Perutz and John Kendrew using X-ray crystallography. Recently, NMR methods have also been used. The secondary structure of proteins, with 60% alpha helices or beta sheets and the rest random coils and turns, is determined by the propensity of the amino acids constituting the given protein to form either alpha helices or beta sheets. It is recognized now that the sequence of a protein determines its three-dimensional structure.

Given the size of proteins, quantum-chemical conformational and energy calculations are at present impossible. Some calculations on proteins are being performed at present in Dr. Lothar Schafer's laboratory. Undoubtedly, the increase in computer capacity and progress in computer algorithms will make it possible to perform many such calculations in the not too distant future. The theoretical methods used so far for proteins include molecular-mechanics methods that neglect electrons and describe the motion of nuclei under the influence of an empirical or quantum-mechanically calculated potential energy function, methods that do not use energy functions except in terms of stereochemical principles, computer graphics methods, and molecular-dynamic methods.

Smaller peptides have also been described by the above-mentioned methods, especially the empirical conformational energy program for peptides (ECEPP), written by Sheraga and his group, which has been applied to a large number of small peptides.

In recent years it has become possible to treat amino acids and small peptides with quantum-chemical calculations, as will be described in the next chapters.