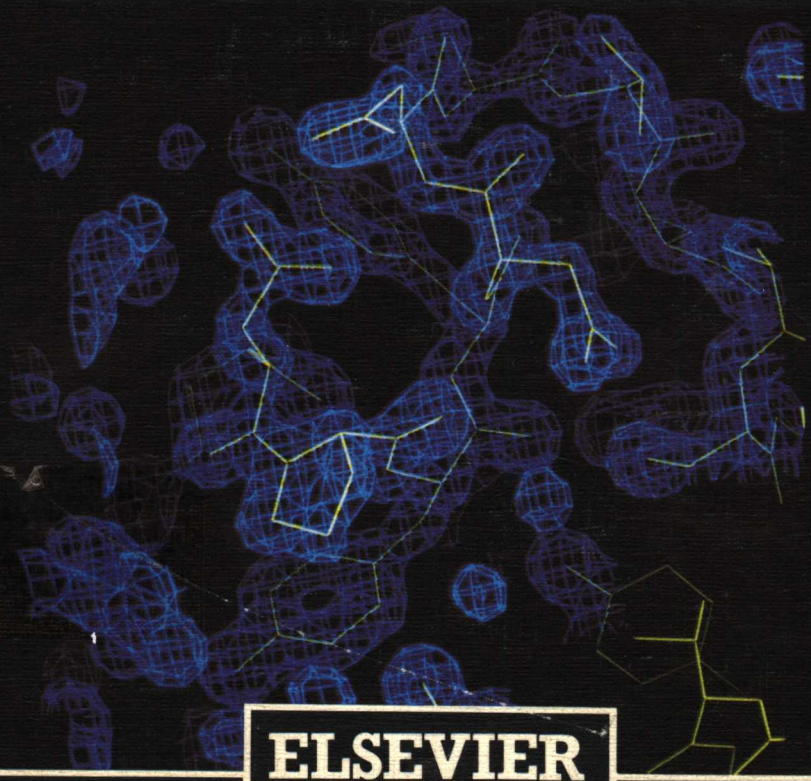
SITE-DIRECTED MUTAGENESIS AND PROTEIN ENGINEERING

EDITOR

M.R. EL-GEWELY



ELSEVIER

SITE-DIRECTED MUTAGENESIS AND PROTEIN ENGINEERING

Proceedings of the International Symposium
on Site-Directed Mutagenesis and Protein Engineering,
Tromsø, 27–30 August 1990

Editors:

M. Rafaat El-Gewely
Department of Biotechnology
Institute of Medical Biology
University of Tromsø
Tromsø, Norway

1991
ELSEVIER SCIENCE PUBLISHERS
AMSTERDAM – NEW YORK – OXFORD

ISBN 0 444 81431 0

This book is printed on acid-free paper.

Published by:
Elsevier Science Publishers B.V.
(Biomedical Division)
P.O. Box 211
1000 AE Amsterdam
The Netherlands

Sole distributors for the USA and Canada:
Elsevier Science Publishing Company Inc.
655 Avenue of the Americas
New York, NY 10010
USA

Library of Congress Cataloging in Publication Data:

International Symposium on Site-Directed Mutagenesis and Protein Engineering (1990 : Tromsø, Norway)
Site-directed mutagenesis and protein engineering : proceedings of the International Symposium on Site-Directed Mutagensis and Protein Engineering, Tromsø, 20-30 August 1990 / editor, M. Rafaat El-Gewely.
p. cm.
Includes indexes.
ISBN 0-444-81431-0 (alk. paper)
1. Protein engineering--Congresses. 2. Site-specific mutagenesis--Congresses. I. El-Gewely, M. Rafaat. II. Title.
[DNLM: 1. Mutagenesis, Site-Directed--congresses. 2. Protein Engineering--congresses. 3. Proteins--genetics--congresses. QU 55 I672s 1990]
TP248.65.P76I58 1990
660'.65--dc20
DNLM/DLC
for Library of Congress
91-9561
CIP

Printed in The Netherlands

INTRODUCTION

One of the biggest surprises in molecular biology is that the primary structure of proteins, as dictated by the genetic code, does not necessarily dictate a unique tertiary structure. Several proteins whose genes have been cloned and expressed by heterologous gene expression did not necessarily have the proper folding (tertiary structure). Therefore these proteins were not active. The question what are the factors controlling the formation of the active form of proteins is really the subject of our symposium.

There are no simple answers. There is no simple or unique approach to find the answer. The promise is, if and when we know the rules of protein folding, the field will be open not only to produce active proteins heterologously, but also to design new protein molecules that never existed before.

The field of protein engineering is multi-disciplinary and it is emerging as a result of interaction between fields of research such as: X-ray Crystallography, Biochemistry, Molecular Biology, and Genetics.

My biasness to Genetics for its involvement in Protein engineering is based not only on the introduction of modern techniques of molecular genetics and recombinant DNA as tools for investigation, but it is also based on the similar basic approach and philosophical goals between the two areas. The science of Genetics was established on the basis of studies dealing with genes through the study of their variation, an important aspect of protein engineering, namely structure-function relationship, is based on the study of a given protein through the study of its variation. 'Variations' in genetics research or in protein engineering could be naturally occurring mutants affecting the phenotype/protein function, or they could be introduced. If the science of Genetics evolved rapidly allowing us to be able to do various genetic manipulations and control, the hope is that protein engineers of the future will be able to design proteins of any needed function.

M. Raafat El-Gewely
Tromsø, Norway

ACKNOWLEDGEMENTS

This Symposium would not have been possible without the generous support of *PHARMACIA BIOSYSTEMS A/S*, and *The Erna and Olav Aakre Foundation for Cancer Research.*

We would also like to acknowledge the support of the following organizations:

The University of Tromsø
The Norwegian Cancer Society
International Union of Biochemistry (IUB)
The Norwegian Society for Science and Humanities (NAVF)
Perkin-Elmer Cetus Analytical Instruments, Germany
The City of Tromsø

A great number of people helped in organizing this symposium. We would like to extend our thanks to Tove Eriksen Heen, Nils Peder Willassen, Norunn Storbakk and to Inge Nilsen for this remarkable help in all the relevant details.

CONTENTS

Molecular structure and dynamics

Protein engineering

MOLECULAR STRUCTURE AND DYNAMICS

© 1991 Elsevier Science Publishers B.V. (Biomedical Division)
Site-Directed Mutagenesis and Protein Engineering
M.R. El-Gewely, editor.

MOLECULAR STRUCTURE AND DYNAMICS
INTRODUCTION

ASBJØRN HORDVIK
Protein Crystallography Group, IMR, University of Tromsø, N-9000 Tromsø, (Norway)

Why are we trying to determine the detailed three-dimensional structure of important biological macromolecules, - like enzymes for example? - Simply because we are convinced that in order to really understand their function and behaviour it is imperative to be able to know them and study them at molecular level.[1,2]

Once the 3D structure of a macromolecule is known in sufficient detail, one may be able to change it a little bit, chemically or by site-directed mutagenesis, to taylor its function for specific purposes.

Such work may, when combined with Molecular Dynamics[3] simultations, reveal valuable information about the interaction between enzyme and substrate for example, and by the latter method one can also seek possible solutions to drug binding problems.

Although one realizes that it is possible to determine the 3D structure of proteins by NMR methods, it should be underlined that the method which so far is most widely used, and far superior with respect to accuracy, is X-ray crystallography[4,5].

Reports from structure work on biological macromolecules will now be presented, and it is encouraging to see how structure chemistry and biology have met on the molecular level in order to disclose the secrets of biological mechanisms.

REFERENCES

1. Rees, A. R., Sternberg, M. J. E. (eds) From Cells to Atoms. Blackwell Scientific Publications, Oxford, (1987).
2. The Molecules of Life. Readings from Scientific American, W. H. Freeman and Company, New York, (1986).
3. Karplus, M., Petsko, G. A. (1990) Nature 347:631-639.
4. Stryer, L. (ed) Biochemistry. W. H. Freeman and Company, New York, (1988), pp 59-62.
5. Richardson, J. S. (1981) In: Anfinsen, C. B., Edsall, J. T., Richards, F. M. (eds), Advances in Protein Chemistry. Academic Press, New York, pp 167-339.

Site-Directed Mutagenesis and Protein Engineering
M.R. El-Gewely, editor.

MOLECULAR STRUCTURE OF NEUROTRANSMITTER RECEPTORS AND LIGANDS

SVEIN G. DAHL
Department of Pharmacology, Institute of Medical Biology, University of Tromsø, N-9001 Tromsø, Norway

MOLECULAR MODELLING

Computer graphics and molecular modelling techniques have undergone a rapid development over the last 5 years, and many laboratories involved in molecular biology research or drug design have recently acquired computer graphics equipment with software for molecular modelling. This is reflected in the increasing number of reports on a wide range of applications of molecular modelling in medicinal chemistry and biology, which has lead to development of specific guidelines for such publications (1). It has been anticipated that worldwide expenditure in computational chemistry will grow from a 1987 level of $125 million to $770 million in 1992, a large proportion of which will be funded by the drug industry (2).

Applications of molecular modelling techniques have become feasible by the recent development in computer technology, which has provided high-performance workstations with increasing power at steadily lower prices. Modern workstations offer high quality raster or vector colour graphics, perspective, depth cueing, three dimensional clipping and real-time translation and rotation. Combined with computational chemistry methods, these techniques have a large potential for providing insight into the spatial arrangements of atoms in molecules, the charge distribution over molecules, and the dynamics of molecular interactions. Molecular modelling methods may thus provide important information about the molecular mechanisms of action of drugs and other biologically active compounds.

Molecular Mechanics and Molecular Dynamics Calculations

X-ray diffraction techniques have been and still are the most widely used experimental methods for determination of three dimensional molecular structures. Many computational methods used in molecular modelling are based on atomic coordinates from crystal structures, as indicated in Fig. 1.

In our recent studies on the molecular conformations and dynamics of psychotropic drug and neurotransmitter molecules, we

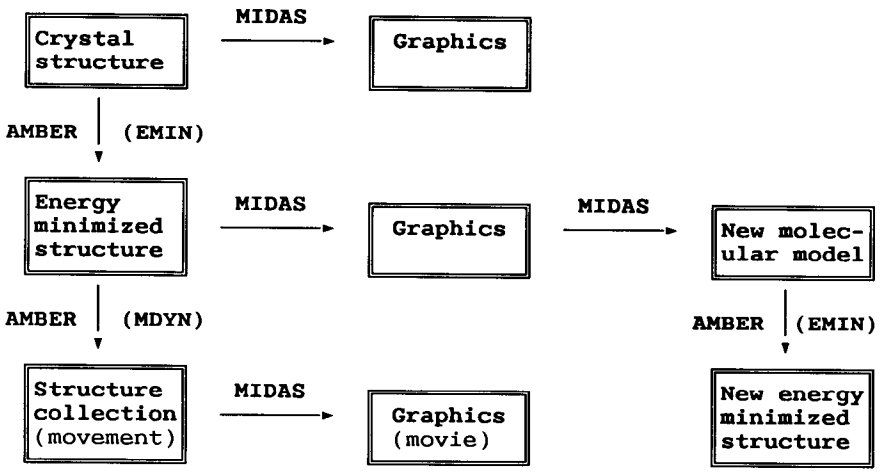
have used the Molecular Interactive Display And Simulation (MIDAS) programs (3) for molecular graphics, with an Evans and Sutherland PS390 workstation and a DEC Microvax II/Ultrix system as the host machine. Water-accessible molecular surfaces and electrostatic potentials 1.4 Å outside the surfaces were calculated with the MIDAS programs, and the potentials were illustrated by colour coding of the surfaces as shown in Fig. 2.

Molecular mechanical geometry optimization and molecular dynamics simulations were performed with the AMBER 3.0 programs (4,5), using the all atom force field. Molecular mechanical calculations provide energy minimized atomic coordinates from a starting structure. Several different force fields are available, and the resulting structure may somewhat depend on the force field and the parameters used for bond lengths, angles, torsional barriers and non-bonded interactions between atoms.

The internal movements of biologically active molecules in solution occur on a picosecond time scale or even faster. Molecular dynamics simulations, which are used to investigate the trajectories of molecular movements across conformational barriers, have provided new insight into the molecular movements and functioning of neurotransmitter receptor ligands (6-12). Combined with modern computer graphics techniques, such simulations provide a new view of the mechanisms of molecular interactions, which may be more valid than static geometric concepts derived from crystal structures and potential molecular energy calculations alone.

As indicated in Fig. 1, crystal structures of drugs and neurotransmitter molecules were initially refined by molecular mechanical energy minimization, and the refined structures were used as starting coordinates for molecular dynamics simulations. For some compounds where no crystal structure was available, an initial model was constructed from the structure of a similar molecule, using the computer graphics system and the MIDAS programs, as indicated in the right part of Fig. 1. The new model was then refined by energy minimization. This strategy was used to construct models of *trans*(E)-chlorprothixene (9), amitriptyline (11) and nortriptyline (11), for which no crystal structures have been reported.

Fig. 1. Molecular modelling procedure. Crystal structures were initially refined by molecular mechanical energy minimization (EMIN), based on the AMBER force field (4,5). Energy minimized structures were used as starting points for molecular dynamics simulations (MDYN), based on the AMBER force field. The MIDAS programs (3) were used for molecular graphics.



MOLECULAR STRUCTURE AND DYNAMICS OF RECEPTOR LIGANDS

The molecular modelling techniques described above were used to study the three dimensional structures, low-energy conformations and molecular dynamics of the neurotransmitters acetylcholine (6), serotonin (7) and dopamine (8), and a series of tricyclic drugs including neuroleptics of the *cis*(Z)-thioxanthene type and their inactive *trans*(E)-isomers (9,10), various tricyclic antidepressants (11), and the four isomers of a metabolite of one of these, 10-hydroxy nortriptyline (12).

As shown by an example in Fig. 2, computer graphics techniques were used to illustrate three dimensional molecular structures and the distribution of positive and negative molecular electrostatic potentials over molecular surfaces. The electrostatic potentials surrounding the dopamine molecule in the energy minimized *anti* conformation, were strongly positive around the terminal part of the side chain containing a protonated dimethylamino group (Fig. 2). Neutral and slightly positive electrostatic potentials surrounded most of the aromatic ring,

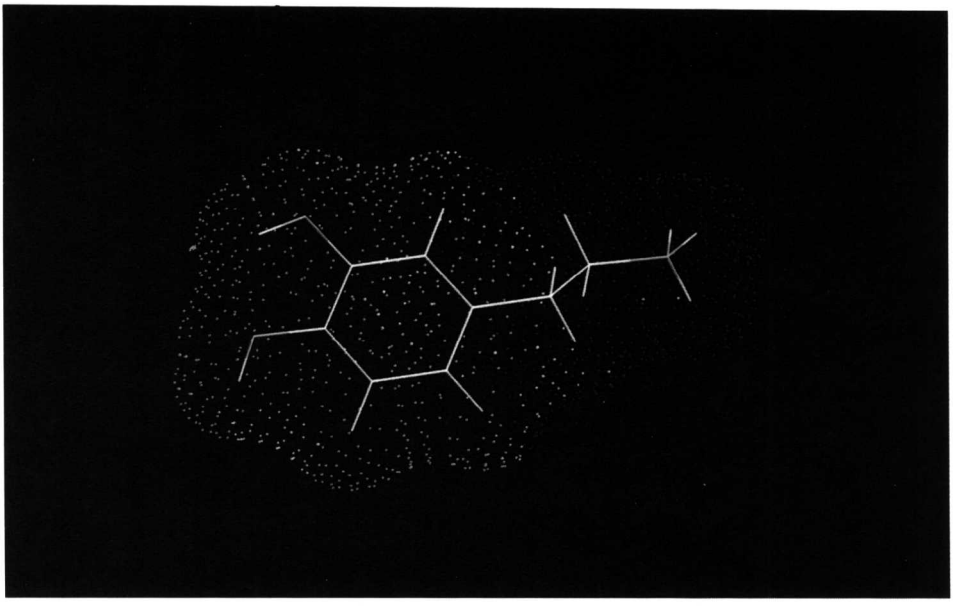


Fig. 2. Energy minimized *anti* conformation of dopamine with water accessible molecular surface, coloured according to molecular electrostatic potentials (e, kcal/mol) 1.4 Å outside the surface: Red e > 15, white 7 $\leq$ e $\leq$ 15, blue 0 $\leq$ e < 7, yellow e < 0. Colour coding of atoms: nitrogen - green, oxygen - blue, hydrogen and carbon - white.

and a small area near the oxygen atom of the m-hydroxyl group had negative electrostatic potentials as low as -5 kcal/mol (8).

Specific binding of ligands to macromolecules usually involves enclosure of most of the ligand molecule by a binding pocket. Electrostatic interactions may contribute to the stabilization of such complexes, and molecular electrostatic potentials therefore provide useful information about the mechanisms of ligand-receptor interactions.

Various side chain conformations of the drug and neurotransmitter molecules were observed during the molecular dynamics simulations, which were carried out *in vacuo* and in aqueous solution. The motions between different side chain conformations followed quite unexpected trajectories, with twisting of the whole molecule such that most of its mass was kept in place, rather than rotating around single bonds. The tricyclic ring systems showed unexpectedly high flexibility, with variations between 100° and 180° in the angle between the phenyl rings.

TABLE I. Neurotransmitter receptors with primary structure known from cloning experiments (19,20)

Superfamily	
Ion channel gating	G protein coupled
$GABA_A$	Dopamine D_1, D_2, D_3
Serotonin $5\text{-}HT_3$	Serotonin $5\text{-}HT_{1A}$, $5\text{-}HT_{1C}$, $5\text{-}HT_{1D}$, $5\text{-}HT_2$
Nicotinic acetylcholine	M_1, M_2, M_3, M_4, M_5 Muscarinic acetylchol.
	α_{1A}-, α_{1B}-, α_{2A}-, α_{2B}- Adrenergic
	β_1-, β_2-, β_3- Adrenergic

These studies have provided new insight into the molecular structure and dynamics of compounds acting as agonists or antagonists on different neurotransmitter receptors.

STRUCTURE OF NEUROTRANSMITTER RECEPTORS

A series of neurotransmitter receptors have been cloned and their amino acid sequence deduced from the DNA sequence. This way modern molecular biology has further clarified and confirmed the subtypes within various families of neuroreceptors, which previously were classified based on the potencies and affinities of various agonists and antagonists. The cloning and sequencing of neurotransmitter receptors has also demonstrated that they may be divided into different superfamilies, as indicated in Table I. Furthermore, that several segments of their peptide chain traverse the cell membrane, connected by intra- and extracellular loops of different lengths.

However, so far there are no available experimental data on the detailed three dimensional structure of any neurotransmitter receptor molecule. The only membrane protein for which a detailed three dimensional crystal structure has been reported, is the photosynthetic reaction center of *Rhodopseudomonas viridis*, which has three sub-units containing a total of 11 membrane spanning alpha helices (13).

The receptors of the superfamily which transfer signals into cells via guanine nucleotide binding regulatory proteins (G proteins), have several common structural features, including seven membrane spanning helices. There is relatively high

sequence homology in the membrane spanning domains between the different families of G protein coupled receptors. Site-directed mutagenesis experiments have demonstrated that two aspartate residues in transmembrane segments 2 and 3 represent binding sites for agonists and antagonists, respectively, in the β_2-adrenergic (14) and M_1 muscarinic acetylcholine (15) receptors. These residues are conserved in all the G protein coupled receptors listed in Table I, including the dopamine D_1, D_2 and D_3 receptors.

The identification of ligand binding sites near the middle of the second and third membrane spanning domains, has lead to the hypotheses that the β_2-adrenergic (16) and M_1 muscarinic acetylcholine receptors (15) form a hydrophilic core by an circular arrangement of the seven transmembrane helices, similar to that observed in the purple membrane of bacteriorhodopsin (17).

We have postulated a similar arrangement of the seven transmembrane helices in the dopamine D_2 receptor, which is assumed to be the primary target of action of antipsychotic drugs. A receptor model was constructed from the amino acid sequence of the rat dopamine D_2 receptor (18), and refined by molecular mechanical energy minimization and molecular dynamics simulations. Its electrostatic potentials suggest that the primary interaction of cationic agonist and antagonist ligands with the receptor is electrostatic, with negatively charged aspartate residues at the synaptic side and in helix 2 and 3.

The virtually inactive *trans*(E)-thioxanthenes have strong negative electrostatic potentials around the 2-substituent on the ring system, contrary to the pharmacologically active *cis*(Z)-isomers (9,10). It seems likely, therefore, that the negative molecular electrostatic potentials around a part of the *trans*(E)-thioxanthene molecules may weaken their initial electrostatic interactions with the dopamine D_2 receptor. This provides a possible explanation of why *trans*(E)-isomers of thioxanthenes are so much less active than *cis*(Z)-isomers in dopamine receptor binding and related tests.

As shown by this example, molecular modelling techniques may extend our knowledge on the molecular function of receptor proteins, even in the absence of detailed x-ray crystallographic or other experimental data.