ABSTRACT

LIGAND-BASED DRUG DESIGN: I. CONFORMATIONAL STUDIES OF GBR 12909 ANALOGS AS COCAINE ANTAGONISTS; II. 3D-QSAR STUDIES OF SALVINORIN A ANALOGS AS KAPPA OPIOID AGONISTS

by
Deepangi Pandit

Ligand-based drug design (LBDD) techniques are applied when the structure of the receptor is unknown but when a series of compounds or ligands have been identified that show the biological activity of the interest. Generally, availability of a series of compounds with high activity, with no activity, and also with a range of intermediate activities for the desired biological target is required. It is common that structures of membrane-bound proteins (for example, monoamine transporter proteins and opioid receptor proteins) are unknown as these proteins are notoriously difficult to crystallize.

In Part I of this study, analogs of the flexible dopamine reuptake inhibitor, GBR 12909, may have potential usefulness in the treatment of cocaine abuse. As a first step in the 3D-QSAR modeling of the dopamine transporter (DAT)/serotonin transporter (SERT) selectivity of these compounds, conformational analysis of a piperazine and related piperidine analog of GBR12909 is performed. These analogs have eight rotatable bonds and are somewhat easier to deal with computationally than the parent compound. Ensembles of conformers consisting of local minima on the potential energy surface of the molecule were generated in the vacuum phase and implicit solvent (also known as continuum solvent) by random search conformational analysis using the molecular mechanics methods and the Tripos and MMFF94 force fields. These conformer populations were classified by relative energy, molecular shape, and their

behavior in 2D torsional angle space in order to evaluate their sensitivity to the choice of charges and force field. Some differences were noted in the conformer populations due to differences in the treatment of the tertiary amine nitrogen and ether oxygen atom types by the force fields.

In Part II of this study, 3D-QSAR studies of salvinorin A analogs as kappa opioid (κ) receptor agonists were performed. Salvinorin A is a naturally-occurring diterpene from the plant *Salvia divinorum* which activates the kappa opioid receptor (KOR) selectively and potently. It is the only known natural non-nitrogenous agent active at the human KOR. Salvinorin A may represent a novel lead compound with possible potential in the treatment of addiction and pain. The primary aim of the current study was to develop Comparative Molecular Field Analysis (CoMFA) models to clarify the correlation between the molecular features of the 2-position analogs of salvinorin A and their KOR binding affinity. The final, stable CoMFA model has predictivity given by q^2 of 0.62 and fit given by r^2 of 0.86. The steric and electrostatic contributions were 47% and 53%, respectively. The CoMFA contour map indicated that the presence of a negative environment and steric region near the 2-position would lead to improved binding affinity at the KOR. Novel salvinorin A analogs with improved binding affinity were predicted based on the stable and predictive CoMFA model. Novel analogs were synthesized by Dr. Thomas Prisinzano of the University of Iowa and preliminary biological results are available from the Rothman laboratory at the National Institute on Drug Abuse. These novel analogs appear to be KOR selective.

LIGAND-BASED DRUG DESIGN: I. CONFORMATIONAL STUDIES OF GBR 12909 ANALOGS AS COCAINE ANTAGONISTS; II. 3D-QSAR STUDIES OF SALVINORIN A ANALOGS AS KAPPA OPIOID AGONISTS

by
Deepangi Pandit

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Chemistry

Department of Chemistry and Environmental Science

May 2007

APPROVAL PAGE

LIGAND-BASED DRUG DESIGN: I. CONFORMATIONAL STUDIES OF GBR 12909 ANALOGS AS COCAINE ANTAGONISTS; II. 3D-QSAR STUDIES OF SALVINORIN A ANALOGS AS KAPPA OPIOID AGONISTS

Dr. Carol A. Venanzi, Dissertation Advisor — Date
Distinguished Professor of Chemistry and Environmental Science, NJIT

Dr. Joseph Bozzelli, Committee Member — Date
Distinguished Professor of Chemistry and Environmental Science, NJIT

Dr. Tamara Gund, Committee Member — Date
Professor of Chemistry and Environmental Science, NJIT

Dr. Edgardo Farinas, Committee Member — Date
Assistant Professor of Chemistry and Environmental Science, NJIT

Dr. Christopher Van Dyke, Committee Member — Date
Consultant, Molecular Modeling and Computational Chemistry
Former Director, Scientific Training, Tripos, Inc, St. Louis, Missouri

BIOGRAPHICAL SKETCH

Author: Deepangi Pandit

Degree: Doctor of Philosophy

Date: May 2007

Undergraduate and Graduate Education:

- Doctor of Philosophy in Chemistry,
  New Jersey Institute of Technology (NJIT), Newark, NJ, 2007
- Graduate Certificate, Internet Application Development,
  NJIT, Newark, NJ, 2002
- Master of Science in Microbiology,
  Maharaja Sayajirao University, Vadodara, India, 1998
- Bachelor of Science in Biochemistry,
  St. Xavier's College, Ahmadabad, India, 1996

Major: Chemistry

Publications

Pandit, D., So, S.-S., Sun, H. Enhancing specificity and sensitivity of pharmacophore-based virtual screening by incorporating chemical and shape features - A case study of HIV protease inhibitors. Journal of Chemical Information and Modeling, 46, 1236-1244, 2006.

Fiorentino, A., Pandit, D., Gilbert, K. M., Misra, M., Dios, R., Venanzi, C. A. Singular value decomposition of torsional angles of analogs of the dopamine reuptake inhibitor GBR 12909. Journal of Computational Chemistry, 27, 609-620, 2006.

Presentation

Pandit, D., Harding, W., Tidgewell, K., Schmidt, M., Lozama, A., Dersch, C., Skawinski, W., Rothman, R., Prisinzano, T., Venanzi, C. A. 3D-QSAR studies of salvinorin A analogs as kappa opioid receptor agonists. Abstracts of Papers, 233rd ACS National Meeting, Chicago, IL, United States, March 25-29, 2007, MEDI-455.

A Sanskrit shloka from the Bhagavad Gita

कर्मण्येवाधिकारस्ते मा फलेषु कदाचन ।
मा कर्मफलहेतुर्भूः मा ते संगोऽस्त्वकर्मणि ।।

karmany evadhikaras te ma phalesu kadachana
ma karma-phala-hetur bhur ma te sango 'stv akarmani

TRANSLATION

Your right is only to perform your prescribed duty or action, but you are never entitled to its fruits. Do not be motivated by the results of your activities, and do not have any attachment for not doing your duty.

This thesis is dedicated to my husband, Nikhil, for his love, support, encouragement and patience throughout the process of my studies. This thesis would be incomplete without a mention of the emotional support given by my parents, Arun and Divya Bhatt, sister, Hemangi Jani and in-laws, Dr. Deepak and Usha Pandit.

ACKNOWLEDGMENT

In Indian philosophy and tradition, an advisor or teacher is referred to by the word "Guru" which means one who removes the darkness of ignorance. For me my advisor, Professor Carol A. Venanzi, is a complete personification of the word "Guru". I am grateful for her guidance, advice, and support which allowed me to perceive and achieve my research and career goals. She exposed me to the world of professional organizations and meetings which helped me to remain up-to-date with developments in the field of computational chemistry. I am indebted for her candid comments on my English writing which inspired me to improve my writing by taking writing classes at Hunter College and also taking the help of Professor Janet Bodner from the Department of Humanities. I want to convey special thanks for her time and effort.

I am grateful to Dr. Joseph Bozzelli, Dr. Tamara Gund, Dr. Edgardo Farinas and Dr. Christopher Van Dyke for their comments and suggestions as the members of my dissertation committee. I want to thank the Department of Chemistry and Environmental Science and Office of Graduate Studies for providing financial assistance. I am grateful to Ms. Gayle Katz for assisting me with all the administrative procedures.

I want to sincerely thank Professor William J. Skawinski for stimulating discussions regarding my research. I appreciate the help and support of Kathleen Gilbert and Milind Misra. Finally, I want to thank the staff of the Library, University Computing Services, and Office of International Students for their services.

TABLE OF CONTENTS

Part I

CONFORMATIONAL STUDIES OF GRB 12909 ANALOGS AS COCAINE ANTAGONISTS

TABLE OF CONTENTS
(Continued)
Part I

CONFORMATIONAL STUDIES OF GRB 12909 ANALOGS AS COCAINE ANTAGONISTS

Chapter **Page**

TABLE OF CONTENTS
(Continued)
Part II

3D-QSAR STUDIES OF SALVINORIN A ANALOGS AS KAPPA OPIOID AGONISTS

TABLE OF CONTENTS
(Continued)
Part II

3D-QSAR STUDIES OF SALVINORIN A ANALOGS AS KAPPA OPIOID AGONISTS

Chapter **Page**

TABLE OF CONTENTS
(Continued)

Chapter **Page**

LIST OF TABLES

Part I

CONFORMATIONAL STUDIES OF GRB 12909 ANALOGS AS COCAINE ANTAGONISTS

Part II

3D-QSAR STUDIES OF SALVINORIN A ANALOGS AS KAPPA OPIOID AGONISTS

LIST OF TABLES
(Continued)
Part II

3D-QSAR STUDIES OF SALVINORIN A ANALOGS AS KAPPA OPIOID AGONISTS

Table **Page**

LIST OF FIGURES
(Continued)
Part I

CONFORMATIONAL STUDIES OF GRB 12909 ANALOGS AS COCAINE ANTAGONISTS

LIST OF FIGURES
(Continued)

Figure **Page**

Part II

3D-QSAR STUDIES OF SALVINORIN A ANALOGS AS KAPPA OPIOID AGONISTS

Figure **Page**

PART I

CONFORMATIONAL STUDIES OF GBR 12909 ANALOGS AS COCAINE ANATAGONISTS

CHAPTER 1

INTRODUCTION

Currently there are no Food and Drug Administration approved therapeutic agents available for the treatment of cocaine abuse or for the prevention of its relapse. It is accepted that addiction is caused by disorders of the brain, with specific neurobiological, molecular and behavioral characteristics that have environmental, drug-induced and genetic determinants of vulnerability.[1] Cocaine addiction may be compared to a chronic illness like diabetes which may be managed but not always cured.[2] Since the proposal of the "dopamine hypothesis", various compounds have been pursued as possible cocaine abuse therapeutics.[3,4] The hypothesis states that cocaine induces a sense of excitement in the user primarily by blocking the reuptake of the neurotransmitter dopamine (DA) at the dopamine transporter protein (DAT) (Figure 1.1). The dopamine hypothesis is able to explain some aspects of stimulant addiction, but other neurochemical mechanisms appear to be more complicated. Recently the analysis of the (DAT)-, serotonin transporter (SERT)-, and norepinephrine transporter (NET)- knockout (KO) mice has indicated that the SERT and the NET also play a role in the pharmacological effects of cocaine.[5-7] All three monoamines (dopamine, serotonin and norepinephrine) have specific transporter proteins for their transport, but it has been observed that either the DAT or NET can transport both dopamine and norepinephrine. It is now well known that cocaine primarily acts on the dopaminergic system, with some effect on the serotonergic, norepinephrinergic, and opioid systems. Cocaine affects dopamine receptors in an indirect fashion by increasing dopamine levels in the synaptic cleft as the result of

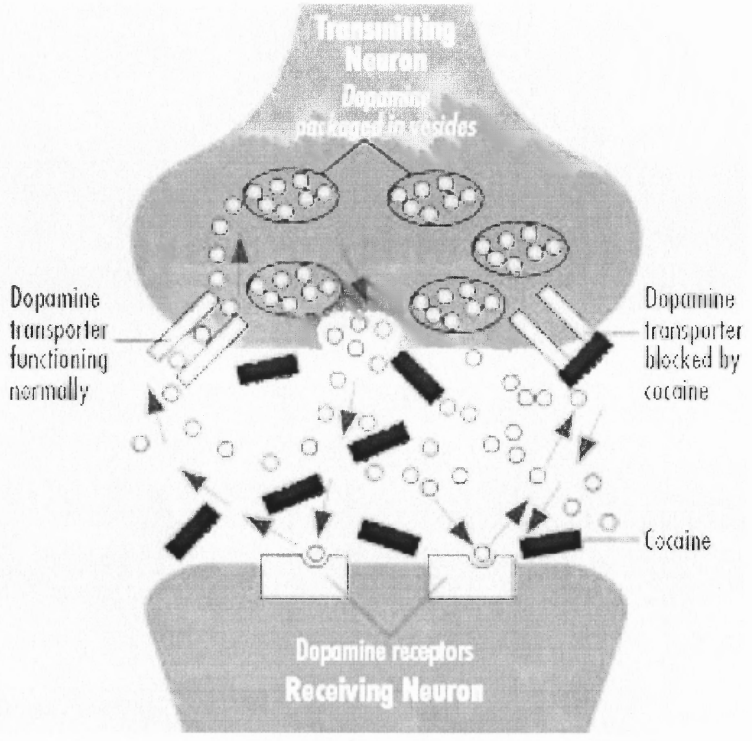


Figure 1.1 Schematic representation of effect of cocaine on DAT.
[Source: National Institutes of Health (NIH) Publication Number 99-4342 Cocaine and Addiction, May 1999]

inhibition of dopamine reuptake through dopamine transporters.[8] The interplay between these systems of monoamine transporter proteins and receptor proteins during cocaine addiction is just beginning to be deciphered. These systems in turn have a profound effect on numerous cellular and molecular targets, including glutamate transmission, GABA transmission and opioid receptors. These factors play an important role in drug addiction which can be divided into three stages: binge/intoxication, withdrawal/negative affect and preoccupation/anticipation.

The primary focus of drug development for the treatment of cocaine addiction has been the design of ligands to prevent the binding of cocaine to the DAT[3] but recently other biological targets also have been explored.[1] DAT ligands are structurally diverse and include tropane, benztropine, 1-[2-[bis(4-fluorophenyl)methoxy]ethyl]-4-(3-phenylpropyl) piperazine (GBR 12909)-analogs, methylphenidate, mazindol and

phencyclidine analogs.[9-13] It is believed that an agent that binds to the DAT with high affinity but slow dissociation rate, preventing cocaine from binding while allowing the reuptake of dopamine, would be able to neutralize the effect of cocaine. GBR 12909 has a unique pharmacological profile which may lend it potential in the treatment of cocaine addiction. Compared to cocaine, it has a higher binding affinity for the DAT and a a slower dissociation rate from the DAT. It has also successfully completed Phase I clinical trials. Our collaborators, the Rice and Rothman groups at the National Institutes of Health, have synthesized and tested the binding affinity of hundreds of GBR 12909 analogs in search of an agent with high selectivity for the DAT, compared to the SERT and NET.[13] Of course, it is clear from the literature that addiction is a complex and complicated phenomenon but GBR 12909 is being used as a tool to test the dopamine hypothesis of stimulant addiction in humans. However, other therapeutic targets are being investigated to determine their role in the treatment and management of stimulant addiction.[13]

1.1 Neurotransmitter Transporter Genes[14]

The DAT, SERT and NET are the members of the sodium- and chloride-dependent neurotransmitter transporter family solute carrier 6 (SLC6). The SLC6 family transports solutes across cell plasma membranes at the same time as it co-transports sodium and chloride down their electrochemical gradients. These transporters have generally 12 transmembrane domains, with cytoplasmic N- and C-terminal tails with glycosylation sites present between transmembranes (TM) 3 and 4. The DAT, SERT and NET have 620, 630 and 619 amino acid residues, respectively. Proteins of this family are regulated

by protein kinases and via protein-protein interactions. Family members also include the transporters for γ-amino butyric acid (GABA), glycine, betaine, taurine, proline and creatine. The transporter proteins of this family have high sequence homology. For example, compared to the DAT, the sequence identity of the other transporters is as follows: NET 67%, serotonin 49%, GABA 45%, glycine 41%, betaine 44%, taurine 42%, proline 44% and creatine 43%. These transporters regulate signaling among neurons in the central and peripheral nervous systems. Below, regulation of the SLC6A3 member (DAT) of the SLC6 transporter proteins expressed in neural cells is briefly discussed.

DAT Gene and Regulation

SLC6A3: DAT The human DAT gene (SLC6A3) is localized[15,16] to chromosome 5p15.3. It spans about 65 kilo base pairs and is divided into 15 exons with no evidence of RNA splice variants. Various approaches have been implemented to map the substrate interaction site of the DAT and its translocation pathway; however a clear picture of the initial docking movement of the substrate across the cell membrane has still not emerged. Although there is no X-ray structure of the DAT, mutational studies have shown that multiple residues interact to form domains for DA and for various inhibitors including cocaine. These residues are separated in the primary structure but may lie near each other in the still unknown tertiary structures.[17] Since dopamine is positively charged at physiological pH, negatively charged aromatic and polar amino acid residues on the DAT are logical sites to be involved in substrate interaction. DA transport across the DAT has three distinct phases: recognition, binding and transport. The DAT can only recognize

and bind cocaine but cannot transport it across the membrane. Conserved residue Asp79 is involved both in DA transport and cocaine binding.[18] Tyr335, which is also a conserved residue among monoamine transporters, appears to be important for DA transport.[18,19] Replacement of Phe105 of TM2 by Ala showed decreased affinity for cocaine, and replacement of Phe155 of TM3 by Ala showed decreased affinity for an analog of cocaine compared to the wild-type transporter. Asp313, Asp436 and Asp376 are also involved in substrate recognition.[20]

Regulation of DAT Protein kinase C (PKC) activation accelerates DAT endocytosis via a clathrin- and dynamin-dependent mechanism.[21] PKC activation also increases phosphorylation of the DAT. The major phosphorylation sites in the DAT are identified as serine residues on the distal cytoplasmic N-terminus. Truncation of the N-terminal of the DAT abolishes PKC-stimulated phosphorylation without impairing transporter internalization.[22] In addition to PKC, other presynaptic proteins also regulate modulation of transporter expression and activity. Endogenous DAT activity is regulated by presynaptic dopamine subtype 2 (D2) receptors through enhanced DAT expression; this modulation is lost in D2-receptor-deficient mice.[23] The oligomerization of the DAT as a dimer due to symmetrical crosslinking between cysteine residues located at the extracellular face of TM6 has been observed. The mutations of TM2 leucine residues eliminate both transporter delivery to the plasma membrane and interaction with wild-type DAT. It has been proposed that TM2 may play an important role in transporter assembly and the oligomerization process, which is essential for the trafficking of the transporter to the cell surface.[24,25]

1.2 GBR 12909 as a Lead Ligand

GBR 12909 (**1,** Figure 1.2) was originally developed and tested in Europe as a potential antidepressant as it was a potent and selective ligand for the DAT.[26] Its ability to produce only a modest increase in extracellular DA and at the same time to attenuate cocaine's ability to increase extracellular DA inspired interest to develop it as a therapeutic agent for stimulant abuse. By blocking cocaine's binding to the DAT but allowing some reuptake of DA, GBR 12909 could act as an antagonist or partial agonist. GBR 12909 differs from cocaine in several important ways: It has higher affinity but a slower dissociation rate from the DAT than cocaine. It has a non-stimulating profile in normal human volunteers as well as the ability to functionally antagonize cocaine.[27-29]

1.2.1 Molecular Modeling of GBR 12909 and its Analogs

Modeling the DAT selectivity of the GBR 12909 analogs is very challenging for three reasons: (1) To date no X-ray crystal structure of the DAT is available, (2) The GBR 12909 analogs are conformationally very flexible, and (3) There are no rigid analogs of GBR 12909. Therefore, a model must be developed using the information about the biological activity of chemically-related structural analogs of GBR 12909. Since it is well-known that a ligand does not necessarily bind to a protein in its global energy minimum (GEM) conformation,[30-34] it becomes imperative that the conformational space of the ligand be thoroughly searched to locate minima that could serve as templates of the ligand in three-dimensional quantitative structure-activity relationships (3D-QSAR) studies. Locating minima becomes especially important in the case of flexible ligands like GBR 12909 and its analogs, as these ligands have a minimum of eight rotatable

bonds. For flexible ligands, the number of minima attained can be very large. Accurate location of minima depends on the algorithm used for searching conformational space as well as on the quality of the molecular mechanics force field and charges used to calculate the energy of the ligand.

The GBR 12909 analogs studied here are the piperazine analog, **2**, and the piperidine analog, **3** (Figure 1.2). Both analogs are somewhat less flexible than **1** and therefore are easier to study computationally. They have fewer rotatable bonds than **1** on the A- (or napthyl) side of the molecule, while the B- (or bisphenyl) side is exactly the same as **1**. The conformer populations of **2** and **3** generated in this work provided the basis for a CoMFA study of the DAT/SERT selectivity of 50 GBR 12909 analogs carried out by the Venanzi group.[35]

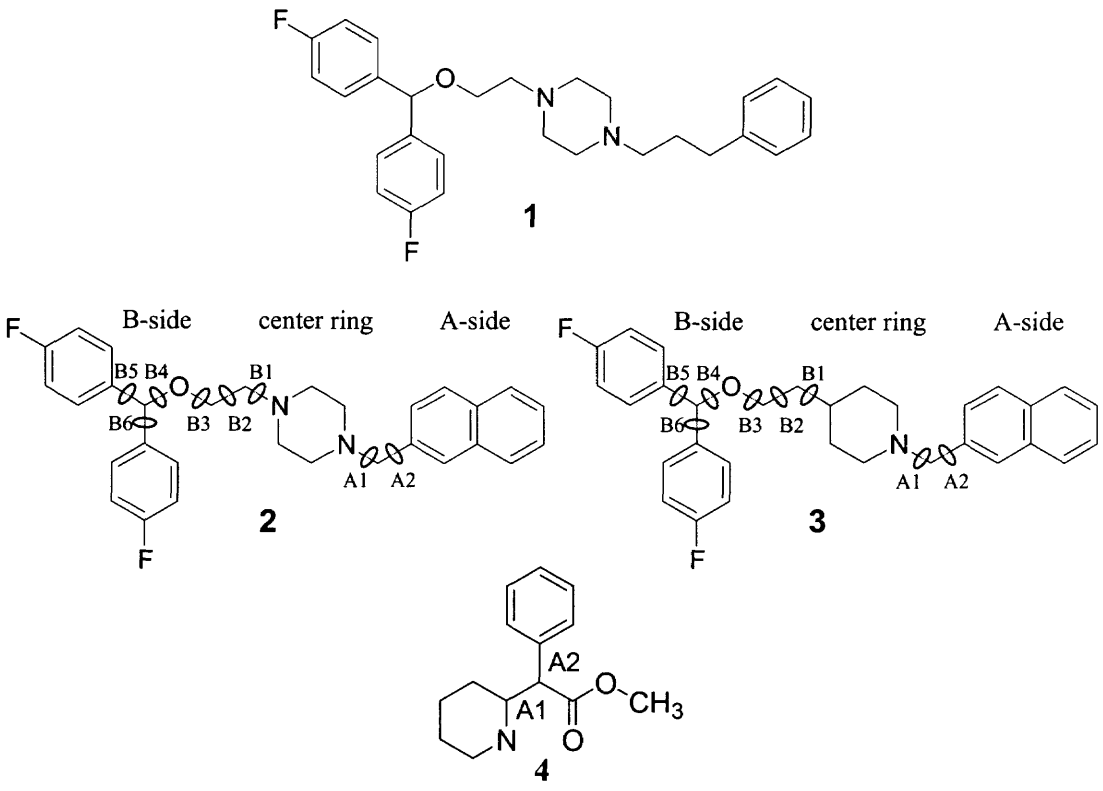


Figure 1.2 Structures of GBR 12909, **1**, and analogs **2** and **3** with marked torsion angles under study and methylphenidate, **4**.

1.2.2 Conformational Analysis

Conformational analysis relates the change in the energy of a molecule to changes in its bond angles, bond lengths, and torsional angles. As the flexibility (defined by the number of rotatable bonds) of a molecule increases, the number of possible conformers with relatively low energy increases dramatically. In fact complete and systematic enumeration of all possible conformers (N) for specified number of rotatable bonds, nbonds, at finite angle increments (δ) can be represented by Equation 1.1.[36]

$$N = \left(\frac{360}{\delta}\right)^{nbonds} \tag{1.1}$$

As the number of rotatable bonds increases, for example as in the present study where the number of rotatable bonds is eight, the problem of investigating all possible conformers in a systematic fashion becomes computationally difficult. An alternative is the use of the random search method proposed by Saunders.[37] In this method, random sampling of the conformational space provides a set of conformations for analysis. Though the sampling is not exhaustive, the completeness of the set of conformations produced can be increased by either increasing the number of iterations in a particular random search or by conducting multiple random searches with a smaller number of iterations and using different starting conformers of the molecule. In this way, the conformational space can be searched to any degree of completeness. The probability of finding all possible conformers during the random search is given by Equation 1.2,[37] where n is the number of times each conformer was found.

$$\text{Probability of finding all conformers} = 1 - (0.5)^n \tag{1.2}$$

1.2.3 Molecular Mechanics Methods

The energy of a conformer obtained with a particular molecular mechanics force field depends on the mathematical formula for the bond stretching, angle bending, torsional, van der Waals and Columbic energy terms, as well as the associated empirical parameters. Force fields are developed to represent structural data for a wide range of compounds, yet they may have certain deficiencies. If one is unaware of these deficiencies, it is possible to draw conclusions about the relationship between molecular structure and biological activity that are based on artifacts in the force field rather than physical reality. The Tripos[38] and MMFF94[39-44] force fields are typically used for ligand modeling and were used in the present study. Since flexible ligands can adopt a wide range of closely-related conformations, it is possible that solvent could affect the relative energy of the local energy minima calculated in the vacuum phase. In order to estimate the solvent effect on ligand conformation and energy, the results of vacuum phase calculations can be compared to results using continuum solvent models such as the ones used in the present study: constant dielectric function and distance dependent dielectric function. Force fields and solvent models are best compared by evaluating their effect on a set of conformers of a ligand rather than by just comparing a single, low-energy conformer. This conformer population approach, first suggested by Boyd and Coner,[45] has the advantage of allowing one to compare a full range of conformers found within a certain energy window rather than just a single structure of each analog, and therefore gives a better picture of the behavior of the different force fields.

1.3 Objectives and Significance

The specific objective of the present work is

> *To study the effect of force fields and solvent models on selected piperazine and piperidine GBR 12909 analogs in order to choose the optimal force field and solvent model for conformational analysis of this class of compounds.*

The significance of this work is that it studies the treatment of the GBR 12909 analogs by two different force fields. The analogs contain piperazine and piperidine moieties, which are common to many pharmaceutically-important compounds. Since these analogs appear to be important lead compounds in the search for a treatment for cocaine abuse, an in-depth examination of the treatment of these molecules by different force fields and solvent models is particularly relevant for accurate drug design. The results of this work provided the basis for the Venanzi group's computational study of the DAT/SERT selectivity of 50 GBR 12909 analogs.[35]

CHAPTER 2

METHODS

2.1 Computational Resources

All calculations were carried out on Silicon Graphics Origin 2000 workstations at New Jersey Institute of Technology using either versions 6.9, 7.1 or 7.2 of the SYBYL molecular modeling program (available from Tripos Inc., St. Louis, MO).

2.2. Analogs 2 and 3

Protonation of Analogs Studies of the pH dependence of dopamine binding to the DAT[46] indicate that dopamine most likely binds in the protonated state. In contrast, similar studies of WIN 35,42847 a cocaine analog, indicate that the protonated and neutral species have similar binding affinity. Our recent studies[48] on the conformational potential energy surface (PES) of the dopamine reuptake inhibitor methylphenidate (**4**, Figure 1.2) indicate that the local minima on the PES of the protonated species are fewer in number but located in the same general region of conformational space as those of the neutral species. It was found that conformers which correspond to the neutral and protonated local energy minima could be grouped roughly into the same conformational families. Therefore, calculations were carried out on the protonated forms of **2** and **3**. Since **2** is a piperazine, it is possible that either nitrogen could be protonated. Molecular orbital theory calculations done by a group member, Dr. William Skawinski, determined that the nitrogen proximal to the naphthalene moiety for **2** and **3** is the preferred site of

the protonation.[49] This is the same nitrogen that Dutta, et al.[50] showed to be required for DAT binding.

2.3 Random Search

In the present study conformational analyses were carried out using the Random Search (RS) option in SYBYL with the force fields, charge sets, and solvent models described below. The RS algorithm is designed to locate the local minima on the conformational PES. The algorithm randomly alters the values of chosen torsional angles and then optimizes the geometry by minimizing the energy of the molecule at each new conformation. The chosen torsional angles (A1, A2, B1-B6) are shown in Figure 1.2 for analogs **2** and **3**. The geometry of the starting conformer for each search was optimized using the Powell[51] minimization method with each respective force field and associated charge set. One thousand search iterations were carried out. At each step in the iteration, the eight torsional angles were randomly altered and the resulting structure was optimized again using the Powell minimization method. The rings of **2** and **3** were held fixed as aggregates with the side chains attached to the central ring in the equatorial position. A convergence threshold of 0.05 and a non-bonded distance cutoff of 8.0 Å were used for each random search. A conformer was accepted into the ensemble of conformers if it met the following energy and root mean square (RMS) criteria: (1) Its RMS distance difference compared to all other conformers was at least 0.20 Å, and (2) Its energy was within 20 kcal/mol of the energy of the conformer identified to have the lowest energy at that particular step in the random search. The random search procedure ended after 1,000 steps. The energy cutoff was purposely set high in order to thoroughly probe the PES of

the molecule. The relative energy of each conformer was calculated by subtracting the absolute energy of the GEM conformer from that of each conformer. The parameters used for the random search are summarized in Table 2.1.

2.4 Force Fields, Charges, Solvent Models

For both **2** and **3**, four different RS runs were made using the following combinations: the Tripos force field and Gasteiger-Hückel atomic charges (in vacuum and solvent) with the default distance-dependent dielectric function, and the MMFF94 force field and associated MMFF94 atomic charges (in vacuum and solvent) with the default constant dielectric function. The dielectric constant value was set equal to one for the vacuum phase and 80 for the solvent calculations. The solvent calculations were performed by William Roosma, supervised by Deepangi Pandit, as part of his master's thesis in Computational Biology.

Table 2.1 Random Search Parameters

Maximum Cycles	1000
Energy Cutoff	20 kcal/mol
Convergence Threshold	0.05
Maximum Iterations	1000
Minimum Hits	6
Check Chirality	Yes
Symmetry	Yes
Force Field	Tripos or MMFF94
Charges	Gasteiger- Hückel or MMFF94
Dielectric Function	Distance Dependent - Tripos Default Constant - MMFF94 Default
Dielectric constant	1.0 (Vacuum) or 80 (Water)
NB Cutoff	8.0 Å
Aggregates	On (Phenyl Rings, Naphthalene Ring, Piperazine/Piperidine Ring)

2.5 Analysis of Conformer Populations

2.5.1 Molecular Shape

The distance of closest approach between the aromatic rings of the A- and B- sides was used to give a gross indication of the extent to which the molecule assumes a "folded" shape. The SYBYL molecular modeling program was used to identify the centroid of each of the phenyl rings and ring 1 of the naphthalene ring (Figure 2.1a). For each conformation the distance between the centroid of each phenyl ring (rings 3 and 4) and that of ring 1 of naphthalene was calculated as D_{13} or D_{14}, respectively. The lesser of these two distances, LD, was defined as the distance of closest approach of the napthalene and bisphenyl moieties. Virtual torsional (VT) angles were defined in order to give an estimate of the degree to which the naphthalene and phenyl rings are offset (or "swung away") from each other. The VT angle was defined by the following four points: (1) the centroid of ring 1 of the naphthalene ring, (2) the nitrogen of the piperazine (or piperidine) ring proximal to the naphthalene ring, (3) the nitrogen (for **2**) or carbon (for **3**) to which the bisphenyl side chain is attached, and (4) the centroid of either phenyl ring. These points are connected by the dotted lines on analog **2** in Figure 2.1b and indicate the virtual torsional angles (VT_{13} and VT_{14}) that define the relative orientation of the napthalene and bisphenyl rings. Conformers were classified into shapes based on the following LD/VT combinations, with LD given in Ångstroms and VT_{13} or VT_{14} given in degrees. Examples of typical shapes are given in Figure 2.2.

C: LD $\leq$ 5, and VT_{13} or $VT_{14} \leq \pm 45$

V: 5 <LD $\leq$ 7 and $\pm 45 < VT_{13}$ or $VT_{14} \leq \pm 75$

U: 7 < LD $\leq$ 12 and $\pm 75 < VT_{13}$ or $VT_{14} \leq \pm 100$

S: 7 < LD $\leq$ 12 and $\pm 100 < VT_{13}$ or $VT_{14} \leq \pm 180$

I: D_{13} or $D_{14} > 12$.

2.5.2 Energy Profiles

The local energy minimum conformations identified by each RS run were assigned to energy "bins" (in units of kcal/mol) on a histogram as follows: 0-4 bin: 0 $\leq$ relative energy < 4, 4-8 bin: 4 $\leq$ relative energy < 8, and so on. Histogram energy profiles were compared for the different RS runs.

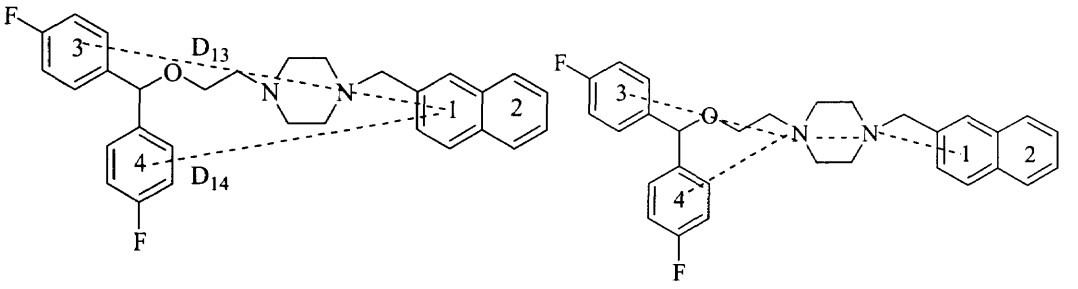


Figure 2.1 Definitions of the closest distance and virtual angles. (a) Distance of closest approach between the A- and B- sides is the lesser of the two distances D_{13} and D_{14}. D_{13} is the distance between the centroids of rings 1 and 3; D_{14} is the distance between the centroids of rings 1 and 4, (b) Points used to define virtual torsion angles: VT_{13} = 1---N---N---ring 3 centroid and VT_{14} = 1---N---N---ring 4 centroid.

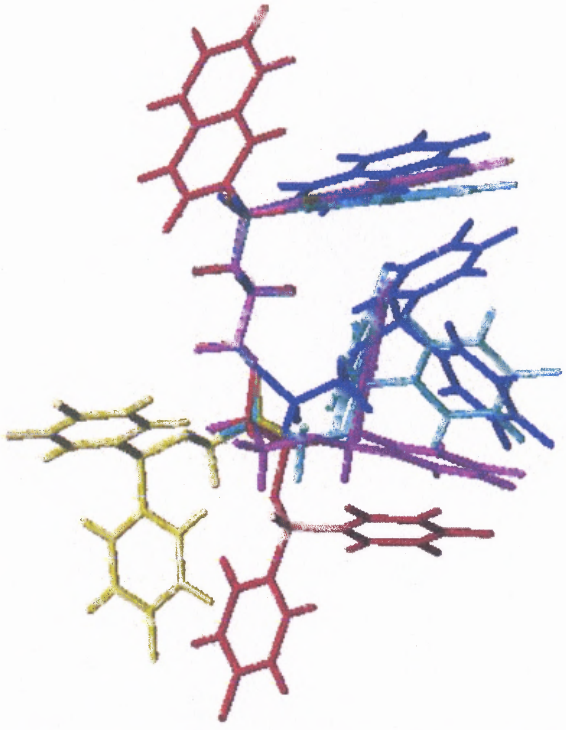


Figure 2.2 Typical conformers representative of each shape: blue - C, cyan - V, magenta - U, yellow - S, Red - I. The yellow S shape overlaps the magenta U shape in the upper right-hand portion of the figure.

2.5.3 Conformer Populations in Torsional Angle Space

Due to the impossibility of viewing the results in eight-dimensional torsional angle space, the local energy minima from each RS run were plotted in two-dimensional torsional angle space for all pairs of consecutive torsional angles: (A1, A2), (B1, B2),...(B4, B5), (B4, B6), and (B5, B6). As will be seen below, the two force fields resulted in a significant difference in the range of B4 values taken on by the conformer populations of **2** and **3**. Note that the torsional angle B4 involves rotation around a $C(sp^3)$-$O(sp^3)$ bond. To examine this difference in detail, models of the B-side of the analogs were constructed and relevant regions of their PES were studied with both molecular mechanics and molecular orbital techniques.

2.6 Model Compounds

Two model compounds (**5** and **6,** Figure 2.3) were constructed to study the influence of the bisphenyl moiety on $C(sp^3)$-$O(sp^3)$ internal rotation. Note that **5** has the $C(sp^3)$-$O(sp^3)$ bond proximal to the bisphenyl group, as in **2** and **3**, so that torsional angle 2 is a model for B4; **6** has one less phenyl ring than **5**.

2.6.1 Generation of Conformer Populations by Random Search Conformational Analysis

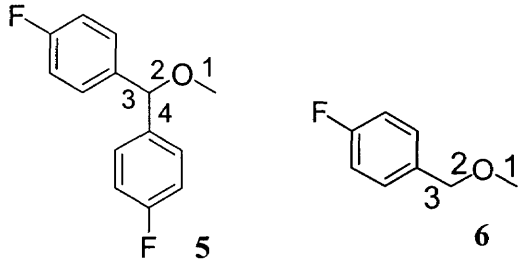


Figure 2.3 Model compounds.

Vacuum phase RS conformational analysis was carried out for **5** and **6** using the protocol described above and the search parameters in Table 2.1. All torsional angles were allowed to vary for **5** and **6**. The conformer populations were plotted in (torsional angle 2, torsional angle 3) space.

2.6.2 Calculation of Potential Energy Surface by Grid Search Conformational Analysis

Since RS only locates the minima on the conformational PES, vacuum phase grid search was also carried out on both **5** and **6** in order to obtain additional details about the PES landscape. For each model compound, torsional angles 2 and 3 were altered in 10° increments forming a grid of torsional angle points. At each grid point, the energy of the molecule was minimized using the chosen force field, holding torsional angles 2 and 3 constant. Each PES grid was plotted using the Origin Pro 7 SR4, Version 7.0552 (B552) package (available from the OriginLab Corporation, Northhampton, MA). The grid searches were conducted by Anuj Kumar, supervised by Deepangi Pandit, as part of his master's project in Computational Biology.

2.6.3 Molecular Orbital Calculations

As will be seen in the Results section, for **5**, the largest difference between the conformational energy calculated by the two force fields was found in the region where torsional angle 2 was between -60° and 60°. Not only were the barriers to internal rotation significantly different, but also the conformational energy minima calculated by the two force fields were found at very different values of torsional angle 2 in this range. The same result was noted in the RS output for **2** and **3** for B4: conformational energy minima calculated by the two force fields were found at very different values of B4 for B4 between -60° and 60°.

To probe these differences, molecular orbital calculations were carried out on **5** using the GAUSSIAN03 program.[52] A "slice" was taken through the potential energy surface of **5** by freezing torsional angle 3 at -30°. Then torsional angle 2 was incremented

in 30° steps and the energy was minimized at each point using the HF/6–31G(d) and B3LYP/6–31G(d) basis sets in vacuum phase. The C-O rotational barrier and the location of the minima with respect to torsional angle 2 in the molecular orbital results were compared to the molecular mechanics grid search results for rotation of torsional angle 2 with torsional angle 3 frozen at -30°. It should be noted that the molecular orbital calculations here are not meant to provide a definitive calculation of C-O internal rotation since MMFF94 was parameterized to a much higher level of theory.[39]

CHAPTER 3

RESULTS

3.1. Analogs 2 and 3

3.1.1 Conformer Populations from Random Search

The RS algorithm identified more than 700 distinct local energy minimum conformations for **2** and **3** with the Tripos force field in both vacuum and solvent (vacuum: 728 for **2** and 718 for **3**, solvent: 735 for **2** and 733 for **3**). With MMFF94, the RS algorithm found more than 600 distinct local energy minimum conformations in vacuum (643 for **2** and 632 for **3**) and more than 750 in solvent (791 for **2** and 780 for **3**).

a. Energy Profiles Figure 3.1 shows the number of conformers in each energy bin for **2** and **3** for the Tripos and MMFF94 random search results. The percent distribution of conformers among energy bins (i.e., the conformer energy profile) is very similar for **2** and **3** in the MMFF94 force field, but somewhat different in the Tripos force field. The MMFF94 vacuum phase conformer profile of **2** is similar to that of **3**, as is the MMFF94 solvent phase profile. However in the Tripos force field, **3** has a significantly larger percentage of conformers in the 0-4 kcal/mol energy bin than **2** in both vacuum and solvent phase. In all cases solvent significantly increases the percentage of conformers in the 0-4 kcal/mol energy bin compared to the vacuum phase results.

b. Molecular Shape Tables 3.1-3.4 give the number of C, V, U, S and I shapes of **2** and **3** in each of the energy bins, as well as the total number of conformers of each shape and the total number in each energy bin. For the Tripos force field in vacuum and solvent (Tables 3.1 and 3.2), **2** has more C shapes than **3** by an order of magnitude. The

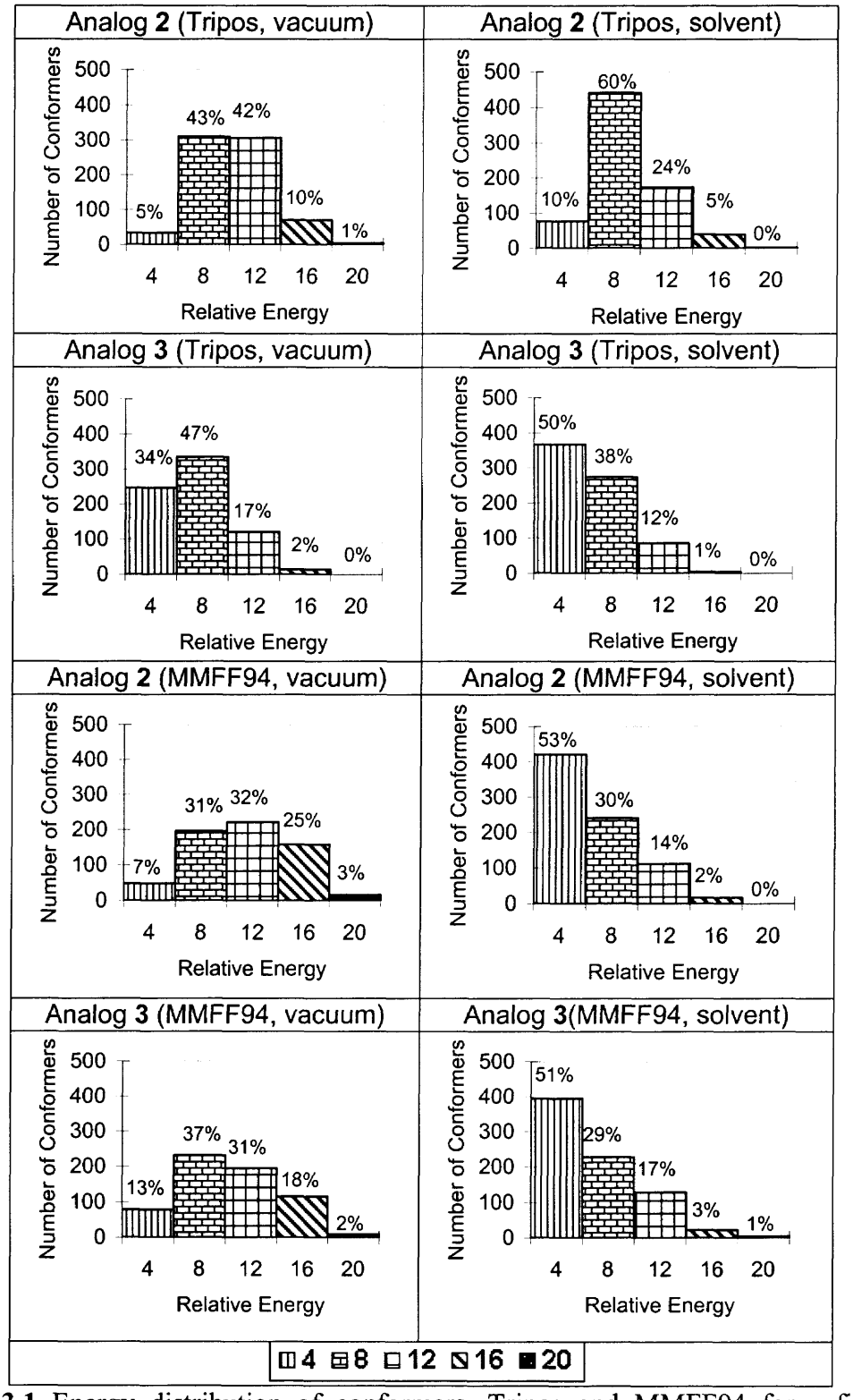


Figure 3.1 Energy distribution of conformers, Tripos and MMFF94 force fields in vacuum and solvent. Relative energy in kcal/mol.

Table 3.1 Energy Distribution of Molecular Shapes, Tripos Force Field, Vacuum Phase

Analog **2**	**0-4**[a]	**4-8**[a]	**8-12**[a]	**12-16**[a]	**16-20**[a]	**Total/shape**[b]
C	17	12	1	0	0	30
V	7	40	9	4	0	60
U	10	111	108	16	3	248
S	0	49	22	3	0	74
I	0	99	167	48	2	316
Total/bin	**34**	**311**	**307**	**71**	**5**	**728**
Analog **3**	**0-4**[a]	**4-8**[a]	**8-12**[a]	**12-16**[a]	**16-20**[a]	**Total/shape**[b]
C	1	0	0	0	0	1
V	27	4	3	1	0	35
U	103	71	20	2	0	196
S	68	53	13	2	0	136
I	48	207	85	10	0	350
Total/bin	**247**	**355**	**121**	**15**	**0**	**718**

[a] kcal/mol. Columns contain number of conformers in each energy bin.
[b] Total number of conformers of each shape. See text for definitions.

Table 3.2 Energy Distribution of Molecular Shapes, Tripos Force Field, Solvent Phase

Analog **2**	**0-4**[a]	**4-8**[a]	**8-12**[a]	**12-16**[a]	**16-20**[a]	**Total/shape**[b]
C	24	6	0	0	0	30
V	24	29	7	1	0	61
U	23	172	68	14	0	277
S	6	52	19	1	2	80
I	0	183	80	24	0	287
Total/bin	**77**	**442**	**174**	**40**	**2**	**735**
Analog **3**	**0-4**[a]	**4-8**[a]	**8-12**[a]	**12-16**[a]	**16-20**[a]	**Total/shape**[b]
C	1	2	0	0	0	3
V	30	9	3	0	0	42
U	135	62	17	4	0	218
S	47	23	2	0	0	72
I	154	179	64	1	0	398
Total/bin	**367**	**275**	**86**	**5**	**0**	**733**

[a] kcal/mol. Columns contain number of conformers in each energy bin.
[b] Total number of conformers of each shape. See text for definitions.

Table 3.3 Energy Distribution of Molecular Shapes, MMFF94 Force Field, Vacuum Phase

Analog 2	0-4[a]	4-8[a]	8-12[a]	12-16[a]	16-20[a]	Total/shape[b]
C	0	3	0	0	0	3
V	17	31	10	8	0	66
U	16	59	38	49	8	170
S	1	24	6	5	0	36
I	14	80	168	96	10	368
Total/bin	**48**	**197**	**222**	**158**	**18**	**643**
Analog 3	**0-4[a]**	**4-8[a]**	**8-12[a]**	**12-16[a]**	**16-20[a]**	**Total/shape[b]**
C	0	0	0	0	0	0
V	18	15	4	2	0	39
U	37	89	42	28	3	199
S	3	16	1	2	0	22
I	21	111	149	84	7	372
Total/bin	**79**	**231**	**196**	**116**	**10**	**632**

[a] kcal/mol. Columns contain number of conformers in each energy bin.
[b] Total number of conformers of each shape. See text for definitions.

Table 3.4 Energy Distribution of Molecular Shapes, MMFF94 Force Field, Solvent Phase

Analog 2	0-4[a]	4-8[a]	8-12[a]	12-16[a]	16-20[a]	Total/shape[b]
C	0	0	0	0	0	0
V	29	14	4	0	0	47
U	96	50	40	6	0	192
S	20	12	8	1	0	41
I	274	165	62	10	0	511
Total/bin	**419**	**241**	**114**	**17**	**0**	**791**
Analog 3	**0-4[a]**	**4-8[a]**	**8-12[a]**	**12-16[a]**	**16-20[a]**	**Total/shape[b]**
C	0	0	0	0	0	0
V	42	7	3	0	0	52
U	111	46	45	8	2	212
S	19	8	5	0	2	34
I	222	168	75	16	1	482
Total/bin	**394**	**229**	**128**	**24**	**5**	**780**

[a] kcal/mol. Columns contain number of conformers in each energy bin.
[b] Total number of conformers of each shape. See text for definitions.

number of C shapes in vacuum and solvent for the MMFF94 force field (Tables 3.3 and 3.4) is negligible for both analogs. The I shape predominates in all cases. Its percentage varies from about 40-50% (Tripos) to 60% (MMFF94) of the total number of conformers for both **2** and **3**. The U shape is the second most favored shape with about 25-35% of the total number of conformers for both force fields for both **2** and **3**. Comparison of Tables 3.1 and 3.2 or Tables 3.3 and 3.4 shows that inclusion of implicit solvent has no effect on the molecular shape profile (i.e. the distribution of conformers among molecular shapes).

c. Conformer Populations in Torsional Angle Space Plots of the conformational energy minima in the torsional angle space of (B2, B3) and (B3, B4) show no significant difference between the force fields and are given in the Appendix. The (A1, A2) plot is discussed below as an example of this case. Since B5 and B6 are correlated due to the constraints of the structure of the bisphenyl group, plots in (B4, B6) and (B5, B6) space give no additional information than plots in (B4, B5) space, so the first two are given in the Appendix.

(1) (A1, A2) Torsional Angle Space: Figure 3.2 plots the conformational energy minima from the random search runs in (A1, A2) torsional angle space. The minima are color-coded by relative energy as described in the figure legend. The figure shows that all the patterns are very similar, indicating that the local minima on the A-side of **2** and **3** are located in very similar regions of (A1, A2) space for the Tripos and MMFF94 results in vacuum and solvent. In all cases the conformational energy minima cluster into groups (i.e. at A1 = $\pm 60°$, $\pm 180°$) that have values of A1 that differ by approximately 120°. From the structure of **2** it can be seen that the three clusters correspond to the staggered conformations which are the conformational energy minima for rotation around the

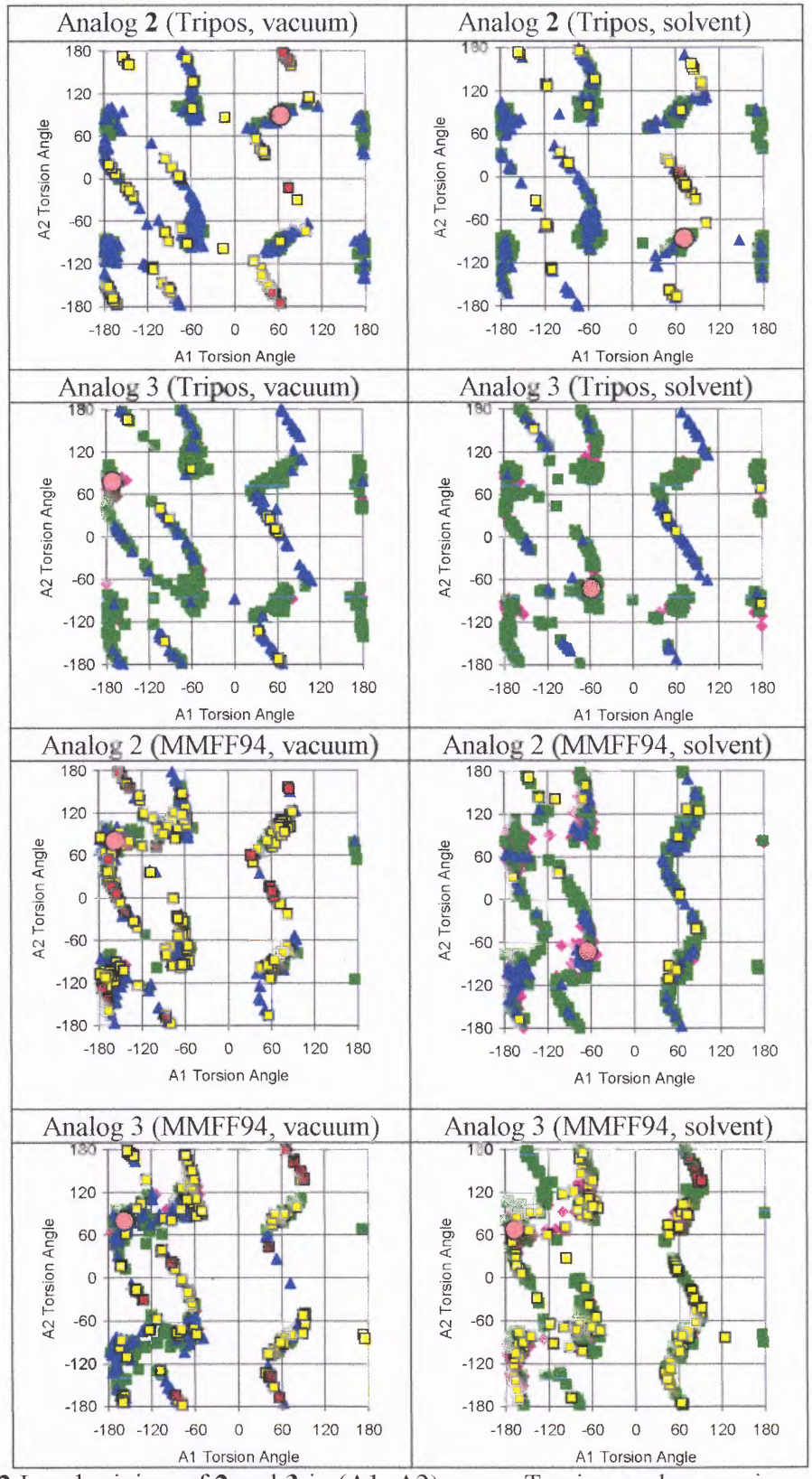


Figure 3.2 Local minima of **2** and **3** in (A1, A2) space. Torsion angles are given in degrees. Minima are color coded by energy in units of kcal/mol. GEM marked by circle. Plot symbols: pink diamond: 0 - 4, green box: 4 - 8, blue triangle: 8 - 12, yellow box: 12 - 16, red box: 16 - 20.

$N(sp^3)$-$C(sp^3)$ bond of the A1 torsional angle. The pattern of minima along the A2 axis is more complex and corresponds to staggered conformations which are the conformational energy minima for rotation around the $C(sp^3)$-$C(sp^2)$ bond of the A2 torsional angle. Similar behavior was seen in the conformational PES of methylphenidate, which also has a piperazine ring separated from an aromatic ring by a methylene group.[48]

Figure 3.2 also shows that the location of the global energy minimum (GEM) conformer (indicated by a large circle) is influenced by both solvent and force field. Except for the MMFF94 results for **3**, for each analog in each force field, the GEM conformer from the solvent study is located in a different region of (A1, A2) space than that from the vacuum phase study. Also, except for the Tripos and MMFF94 vacuum phase results for **3**, the GEM conformer from the Tripos results is located in a different region of torsional angle space than that for the MMFF94 results.

(2) (B1, B2) Torsional Angle Space: Figure 3.3 shows the local minima of **2** and **3** in (B1, B2) torsional angle space. All cases, except the Tripos vacuum and solvent phase results for **2**, show nine well-defined minima which result from combination of the staggered conformations, which are the minima for rotation around the $N(sp^3)$-$C(sp^3)$ bond (in **2**) or the $C(sp^3)$-$C(sp^3)$ bond (in **3**) in the B1 torsional angle (at B1 = $\pm 60°$, $\pm 180°$), with those that are the minima for rotation around the $C(sp^3)$-$C(sp^3)$ bond in B2 (at B2 = $\pm 60°$, $\pm 180°$). In contrast, the Tripos results for **2** give minima that take on a range of B1 values not seen in the Tripos results for **3** or in the MMFF94 results for **2** and **3**. Since **2** and **3** differ only by a tertiary amine nitrogen versus a tetrahedral carbon in the B1 torsional angle, the difference in the conformer populations of **2** and **3** may be due to differences in how the force fields treat a tertiary amine nitrogen atom type and, ultimately, in the effect of that nitrogen on $N(sp^3)$-$C(sp^3)$ internal rotation.

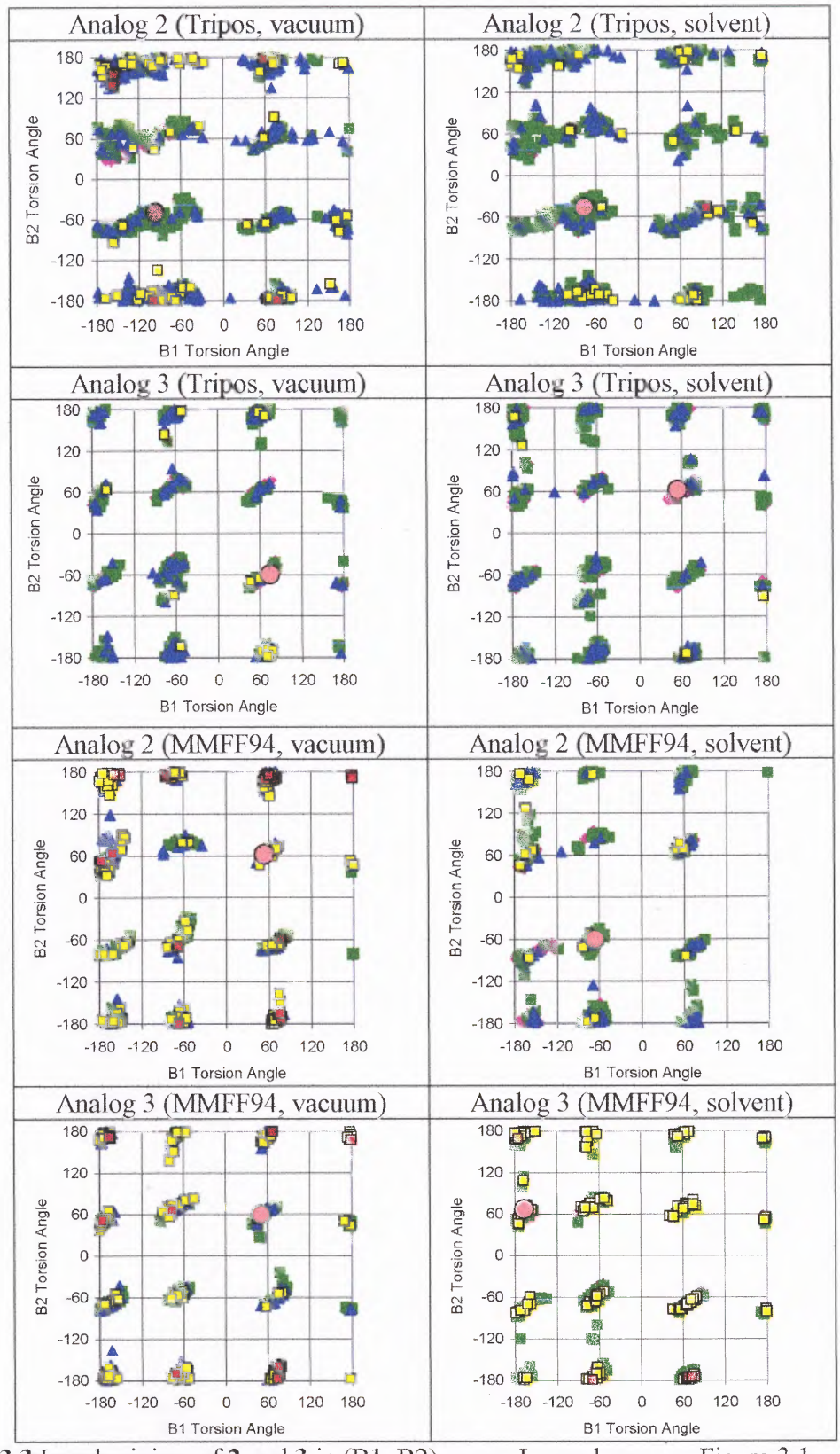


Figure 3.3 Local minima of **2** and **3** in (B1, B2) space. Legend same as Figure 3.1.

This is illustrated in Figure 3.4. Using the conformer populations from the vacuum-phase results for both force fields, for each conformer Figure 3.4 plots the distance of the nitrogen atom in B1 of **2** with respect to a plane formed by three neighboring carbon atoms and compares it to a similar plot for the related carbon atom in **3**. The figure shows a distinctive difference in the pattern for the nitrogen of **2** in the Tripos versus MMFF94 force fields. In the Tripos case (Figure 3.4a), about 15% of the conformers have somewhat a planar nitrogen (arbitrarily defined as a nitrogen with a distance with respect to the plane between -0.30 and 0.30 Å). For the MMFF results for **2** (Figure 3.4c), essentially 100% of the conformers have a tetrahedral nitrogen. This seems to indicate that during the RS procedure, the Tripos force field allows the nitrogen to attempt to "flatten out", even though the B-side side chain was held fixed in the equatorial position. In contrast, the MMFF94 force field keeps the nitrogen tetrahedral. Figures 3.4b and 3.4d show that the related carbon atom in **3** remains tetrahedral during the RS with both force fields. Similar trends are seen in the solvent-phase results (not shown).

(3) (B4, B5) Torsional Angle Space: Figure 3.5 shows the local minima of **2** and **3** in (B4, B5) torsional angle space. There are striking differences between the Tripos and MMFF94 results. In the Tripos case, conformers cluster along the B4 values of -60° and 60°, with few conformers in the B4 range between -60° to 60°. The MMFF94 results present the opposite picture, with most conformers found for values of B4 between -60° and 60°. This difference between the two force fields with respect to B4 is also demonstrated in the plots of the local minima of **2** and **3** in (B3, B4) and (B4, B6) space (see Appendix A).

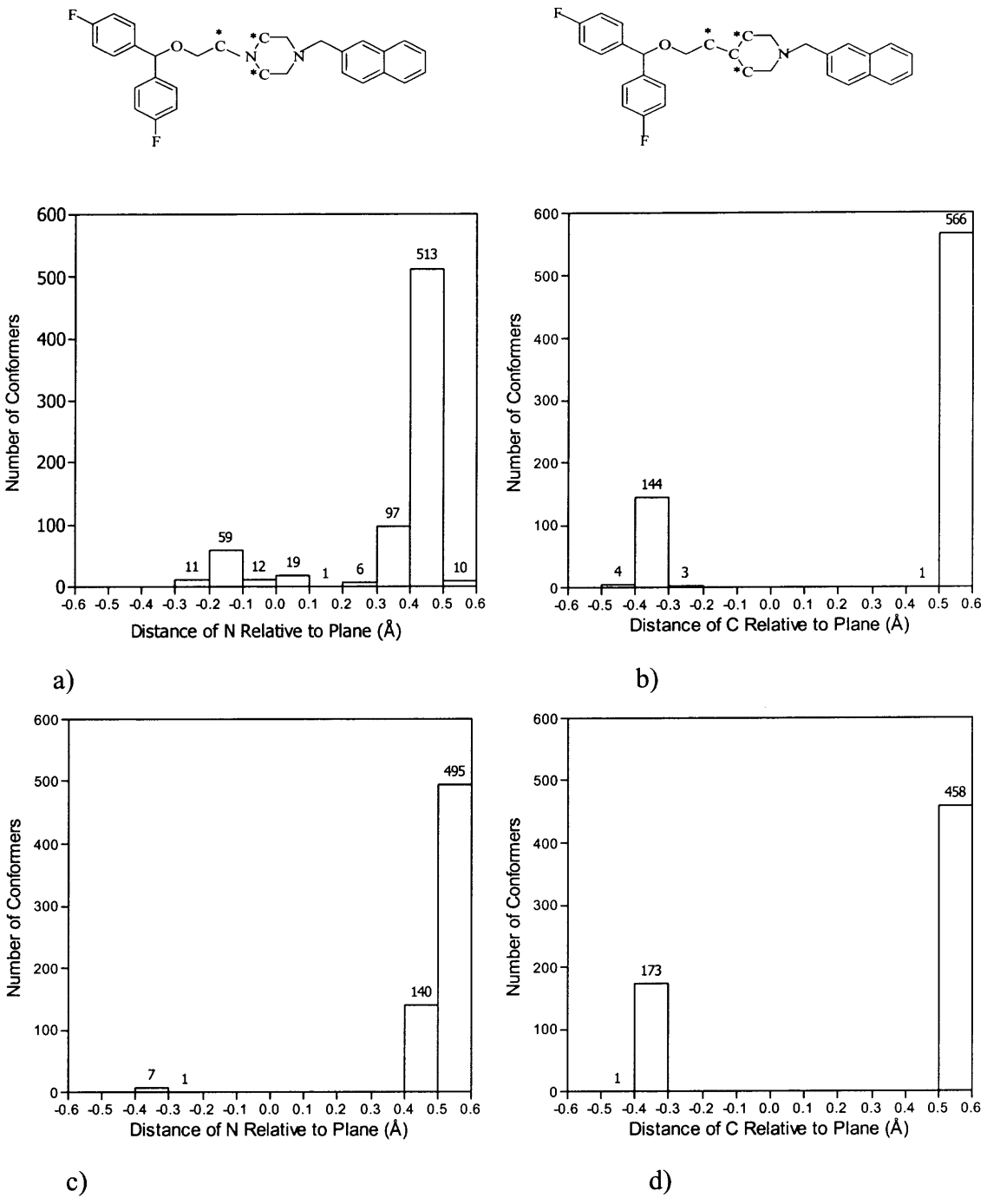


Figure 3.4 Number of conformers with distance of N or C relative to the plane of three carbons identified on molecular structure. (a) **2**, Tripos in vacuum phase, (b) **3**, Tripos in vacuum phase, (c) **2**, MMFF94 in vacuum phase, (d) **3**, MMFF94 in vacuum phase. Conformers having distance greater than 0 and less than or equal to 0.1 are assigned to bin labeled 0.1, and so on.

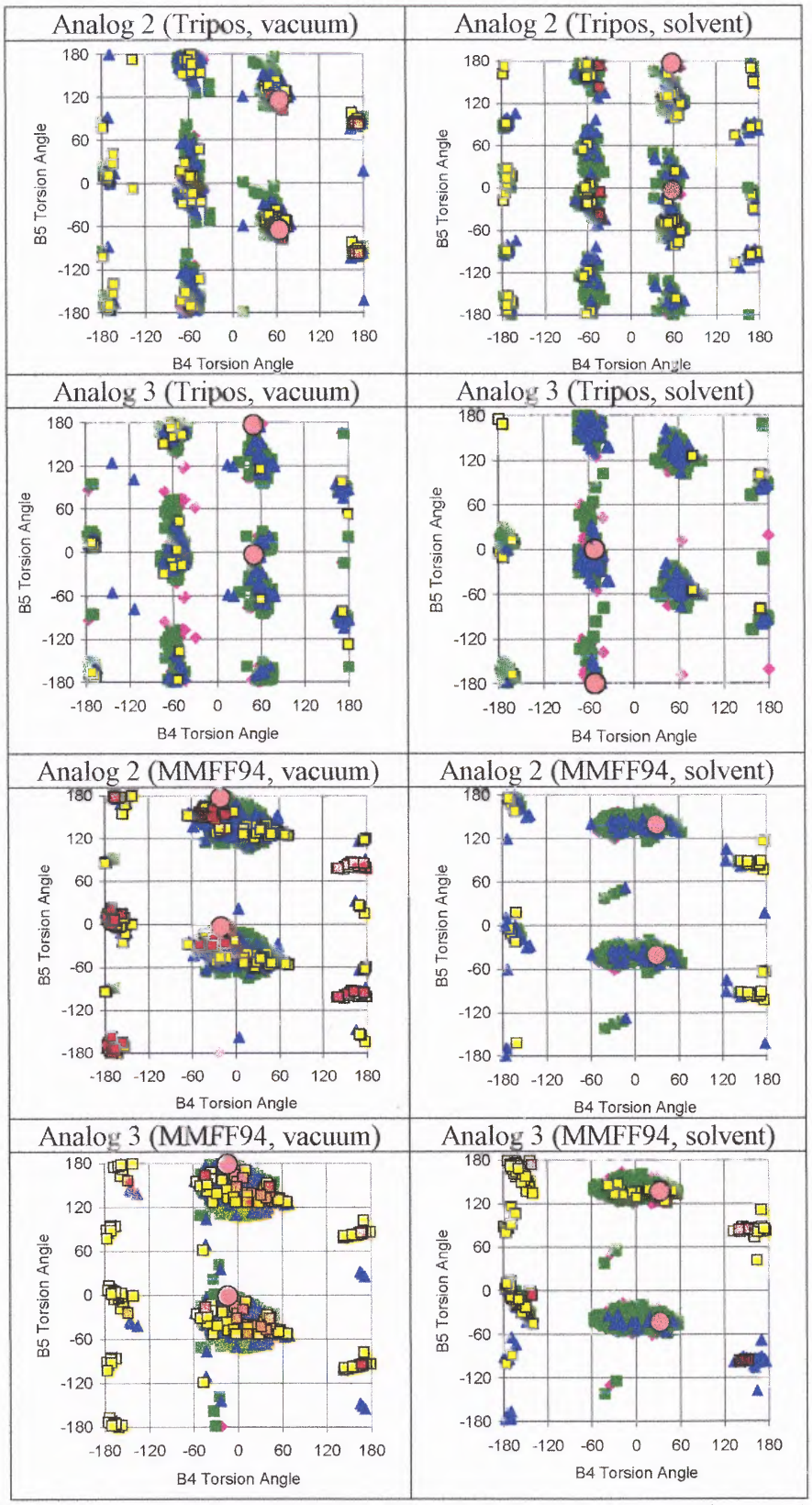


Figure 3.5 Local minima of **2** and **3** in (B4, B5) space. Legend same as Figure 3.1.

Note that the B4 torsional angle involves rotation around the $C(sp^3)$-$O(sp^3)$ bond. Although the B3 torsional angle also involves rotation around the $C(sp^3)$-$O(sp^3)$ bond, in this case the two force fields agree that the minima for **2** and **3** are found for values of B3 close to -60° and 60°, rather than between -60° and 60° (see plots of minima in (B2, B3) and (B3, B4) space in the Appendix). Since B4 is proximal to the bisphenyl moiety whereas B3 is an additional bond length away, it is possible that the presence of the two phenyl rings affects the minima for $C(sp^3)$-$O(sp^3)$ rotation differently in the Tripos and MMFF94 force fields. In order to investigate these issues in more detail, molecular mechanics and molecular orbital calculations were carried out on model compounds. The results are described below.

3.2 Model Compounds

The model compound studies were designed to investigate the effect of the bisphenyl moiety on internal rotation by comparing rotation around the C-O bond in **5** to that in **6**. Torsional angle 2 in **5** is the equivalent of B4 in **2** and **3**. Figure 3.6 displays the results of the molecular mechanics grid search calculation of the PES of **5** and **6**. The low energy conformers identified by the RS calculation are shown as white squares on the grid. Comparison of Figures 3.6a and 3.6b shows a general similarity in the location of regions of high and low energy for **5** in the two force fields. Both force fields find high energy regions for torsional angle 2 between -180° and -120°, as well as between 90° and 180°. Both force fields show regions of low energy for torsional angle 2 equal to $\pm$60°. However, the MMFF94 force field displays a broad region of low energy for **5** for torsional angle 2 between -60° and +60°, while the Tripos force field shows a much higher energy range in this region. Comparison of Figures 3.6a and 3.6c shows that

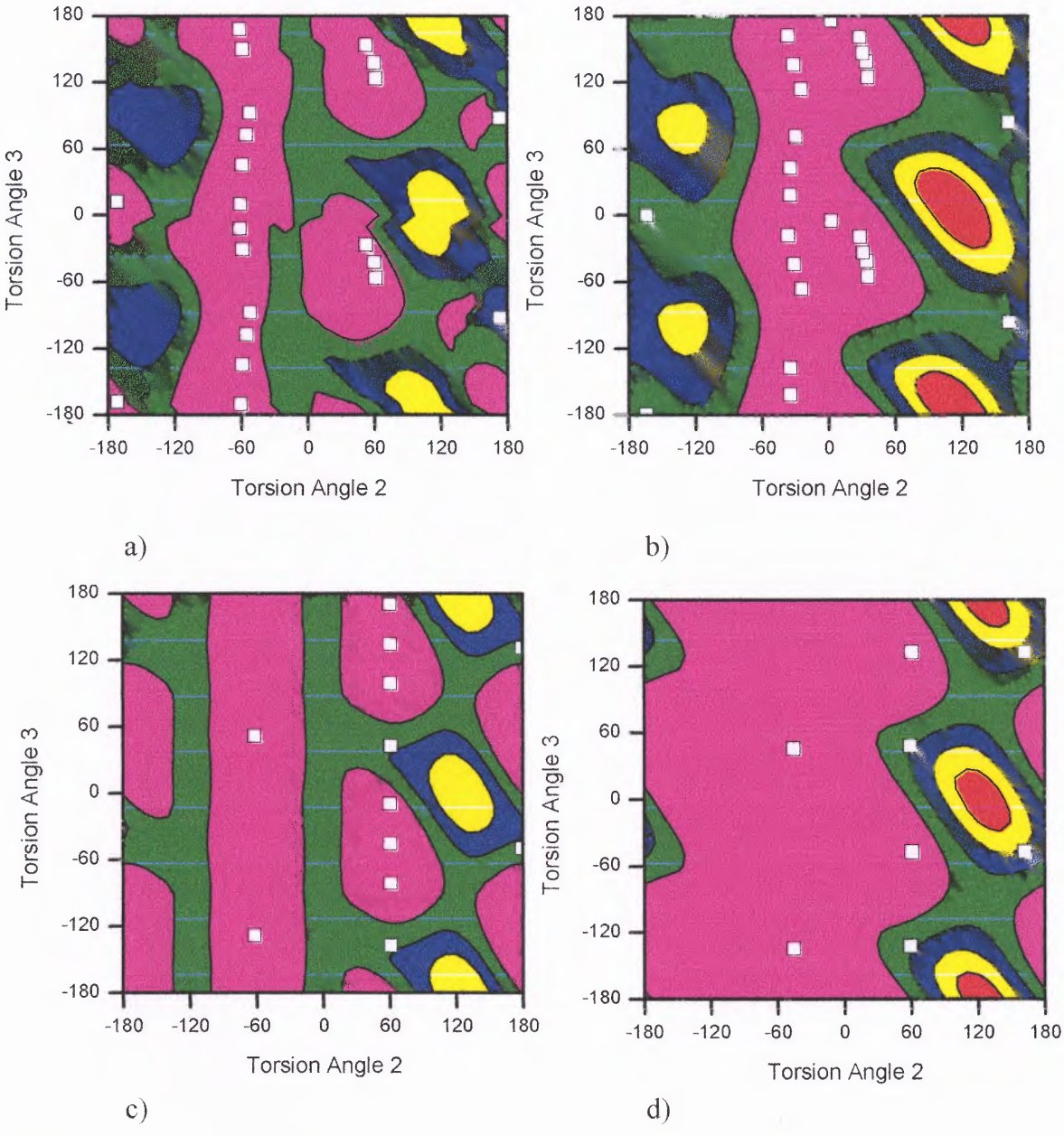


Figure 3.6 Vacuum phase potential energy surfaces of model compounds. Contours are color-coded by relative energy (kcal/mol) as follows: Magenta (0-4); Green (4-8); Blue (8-12); Yellow (12-16); Red (16-20). (a) **5**, Tripos, (b) **5**, MMFF94, (c) **6**, Tripos, (d) **6**, MMFF94. Minima obtained from random search for corresponding model compounds are shown as white squares.

removing one phenyl ring from **5** to give **6** has little effect on the characteristics of the Tripos PES. In contrast, comparison of Figures 3.6b and 3.6d shows that removal of one phenyl ring leads to significant broadening of the low energy region on the MMFF94 PES. Therefore it seems that the bisphenyl moiety has a larger effect on internal rotation of a proximal C-O bond in the MMFF94 force field than in the Tripos force field.

Molecular Orbital Calculations Figure 3.7 plots the results for C-O internal rotation for **5** at the HF/6-31G(d) and B3LYP/6-31G(d) levels with torsional angle 3 set equal to -30°. For comparison purposes, the "slice" through the PES of **5** for torsional angle 3 equal to -30° from Figures 3.6a and 3.6b is also shown on the graph. Figure 3.7 shows that both the Hartree Fock (HF) HF/6-31G(d) and density functional theory (DFT) B3LYP/6-31G(d) techniques give very similar results, with the global energy minimum falling at -60°, local minima at -150° and 30°, a low rotational barrier (1.52 kcal/mol for HF, 1.89 kcal/mol for DFT) at 0°, and a very high barrier at 120°. Overall, the molecular mechanics results are qualitatively similar to the molecular orbital results. However, there are some significant differences in the PES for torsional angle 2 between -60° and 60°. In contrast to the molecular orbital results, the MMFF94 force field locates the minima at -30° and 30°, with a much small barrier (only 0.49 kcal/mol) at 0°. The Tripos force field, however, gives a value at 0° (3.53 kcal/mol) which is significantly higher than the molecular orbital and MMFF94 results, and locates the minima at -90° and 30° with a barrier at -30° of 4.44 kcal/mol. This explains why so many of the MMFF94 RS conformers of **2** and **3** are found in the region with B4 between -60° and 60° in Figure 3.5, whereas the Tripos results tend to cluster around B4 = $\pm$60° with no conformers found at B4 = 0°.

3.5, whereas the Tripos results tend to cluster around B4 = ±60° with no conformers found at B4 = 0°.

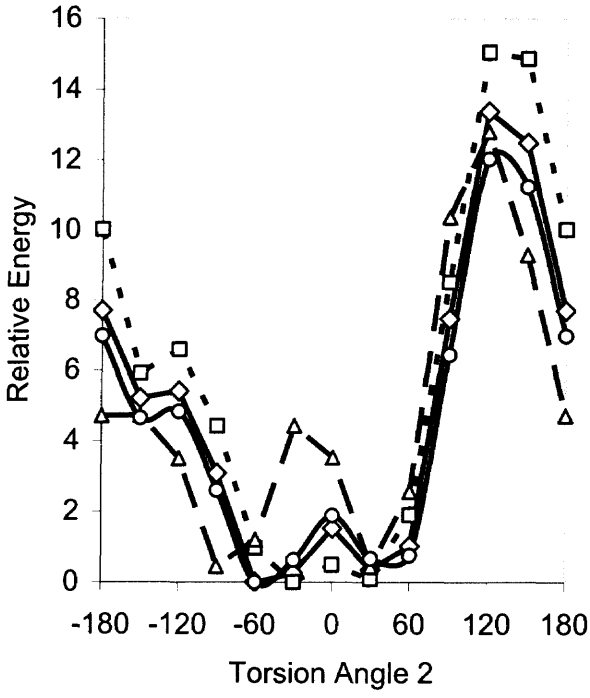


Figure 3.7 Rotational barrier of torsion angle 2 of **5** (vacuum phase) with torsion angle 3 fixed at -30°. Solid line with diamonds: HF/6-31G(d), solid line with circles: B3LYP/6-31G(d), dashed line with squares: MMFF94 force field, broken line with triangles: Tripos force field. Relative energy in kcal/mol.

CHAPTER 4

DISCUSSION

4.1 Effect of Implicit Solvent on Conformer Populations

In all cases, the inclusion of an implicit solvent model significantly increased the percentage of conformers in the 0-4 kcal/mol energy range compared to the vacuum phase results and affected the location of the GEM conformer in torsional angle space, without changing the molecular shape profile of the conformer populations. It should be noted that MMFF94 was parameterized for use in molecular dynamics simulations with discrete water molecules[39] whereas MMFF94s was developed for use in energy minimization studies.[44] The two force fields give identical results for most systems and differ only in their treatment of resonance-delocalized trigonal nitrogen atoms.[44] This atom type is not found in the GBR 12909 analogs, making either force field a suitable choice for the present study. Although the MMFF94 force field was not validated for use with the implicit solvent model employed here, since this options is available in the popular molecular modeling program SYBYL, it is of some interest to note its effect on the conformer populations.

4.2 Effect of Force Field on Conformer Populations

To the best of our knowledge, the results of calculations with the Tripos force field and Gasteiger-Hückel charge set have never before been directly compared to those obtained with the MMFF94 force field for large, flexible molecules containing carbon, hydrogen, nitrogen and oxygen atoms, such as the GBR 12909 analogs considered here. Since the

Tripos force field was constructed and validated without atomic charges,[38] it is useful to compare its behavior with the Gasteiger-Hückel charge set, a frequent choice among SYBYL users, to that of the more recent and more extensively-validated MMFF94 force field.[35, 39-44, 53]

MMFF94 was parameterized for a wide variety of pharmaceutically-relevant chemical systems using both high-level ab initio molecular orbital theory[39-42] and experimental data.[43] The "core" parameterization involved, among other calculations, geometry optimization of 500 molecular structures at the HF/6-31G* level, 475 structures at the MP2/6-31G* level, and 380 structures at a higher level including electron correlation and triple zeta plus polarization basis sets. Conformational energies were calculated in 250 cases at the "MP4SDQ/TZP" level, i.e. triple zeta plus polarization calculations at a defined approximation to the MP4SDQ level of theory. Approximately 1200 torsional profile structures, obtained by rotating a given torsional angle by a specified increment, were calculated at the MP2/TZP level derived from MP2/6-31G*-optimized geometries. The MMFF94 parameters were determined in a "mutually consistent" fashion from all the available data using an iterative procedure in which each type of parameter was optimized while using increasingly well-refined parameters for the other parameter types.[39] This is different than "functional group" approach employed by most force fields, including Tripos, in which certain parameters are fit to a portion of the data, then frozen. MMFF94 also employs a unique functional form for describing the van der Waals interactions.[39]

In contrast, the Tripos force field (without charges) was validated against X-ray structures by minimizing three cyclic hexapeptides, crambin, and 76 small organic molecules. Thermodynamic barriers were calculated for 17 different conformational energies, 12 stereoisomers, and 15 rotational barriers.[38]

4.2.1 Parameterization of Tertiary Amine Nitrogen

Several differences were noticed in the conformer populations generated by the two force fields. For example, MMFF94 gave very similar conformer energy profiles for **2** and **3**, in contrast to the results with the Tripos force field for which **3** had a significantly larger proportion of conformers in the 0-4 kcal/mol energy bin than **2** (Figure 3.1). Although for both force fields most conformers of **2** and **3** were found in the I shape, with the U shape being the second most favored, only the Tripos force field allowed some conformers of **2** to take on the C shape (Tables 3.1-3.4). Comparison of the plots of the conformer populations in torsional angle space shows that both force fields yielded similar plots for **2** and **3** in (A1, A2) space (Figure 3.2), and similar plots for **3** in (B1, B2) space (Figure 3.3). But for **2**, the (B1, B2) plots show that the conformers took on a different range of values for B1 in the Tripos versus the MMFF94 force fields. It should be noted that for **2**, although A1 and B1 both involve internal rotation around an $N(sp^3)$-$C(sp^3)$ bond, the nitrogen in A1 is protonated whereas the nitrogen in B1 is not. For **3**, B1 involves internal rotation around the $C(sp^3)$-$C(sp^3)$ bond.

In previous work,[54] we applied the singular value decomposition technique to all eight torsional angles (A1,..., B6) of the Tripos vacuum phase conformer populations of **2** and **3**. We uncovered differences in how the data separated along certain principal

components and in which torsional angles were the chief contributors to those principal components.[54] This indicated an underlying difference in the conformer populations of **2** and **3** generated by the Tripos force field. Since **2** and **3** only differ by an unprotonated tertiary amine nitrogen versus a tetrahedral carbon at the same location in the B1 torsional angle, the difference in their conformer populations was attributed to the treatment of the tertiary amine nitrogen by the Tripos force field.

In the present work, plots of the distance of the B1 nitrogen of **2** and the corresponding B1 carbon of **3** with respect to the plane of their neighboring carbons (Figure 3.4) show that the Tripos force field allowed a significant percentage (15%) of the conformers of **2** to have tertiary amine nitrogens to be somewhat planar, while both Tripos and MMFF94 kept the B1 carbon tetrahedral. It should be noted that various conformers of 23 different amines were used in the "core" parameterization of MMFF94 and that the root mean square deviation (in degrees) for 96 out of plane angles in the set of saturated amines was only 0.91 for MMFF94 compared to 57.5 for MP2/6-31G* method.[39] Of the 76 small molecules tested with the Tripos force field, about half contained various types of amine nitrogens, but statistics were given only for the rms errors in bond lengths, bond angles, and torsional angles for the dataset as a whole. It should be noted, however, that compared to MMFF94 the Tripos force field overestimates the $N(sp^3)$–$C(sp^3)$ rotational barrier (in kcal/mol): methylamine (1.98-experiment,[55] 2.8-Tripos,[38] 2.36-MMFF94[42]), dimethylamine (3.62-experiment,[55] 4.9-Tripos,[38] 3.52-MMFF94[42]).

The difference in how the force fields treat a tertiary amine nitrogen atom type and, ultimately, the effect of that nitrogen on $N(sp^3)$-$C(sp^3)$ internal rotation appears to be one source of difference in the conformer populations of **2** and **3**.

4.2.2 Proximity of Bisphenyl Moiety to C-O Bond

The effect of the bisphenyl group on $C(sp^3)$-$O(sp^3)$ internal rotation appears to be another source of difference between the force fields. The plot of the conformer populations of **2** and **3** in (B4, B5) space (Figure 3.5) showed striking differences between the Tripos and MMFF94 force fields for behavior with respect to B4, whereas a similar plot in (B2, B3) space (Appendix) showed little difference between the force fields. Torsional angles B3 and B4 both contain a $C(sp^3)$-$O(sp^3)$ bond, but B4 is proximal to the bisphenyl group, whereas B3 is one bond length further away (Figure 1.1). Calculations on model compounds **5** and **6** showed that the MMFF94 force field is sensitive to the effect of the bisphenyl moiety on C-O internal rotation (Figure 3.6). Although 25 types of aromatic and heteroaromatic molecules along with various conformers of 14 types of conjugated systems were used in the core parameterization of MMFF94, none contained the bisphenyl moiety. However, MMFF94 has been shown to reproduce the $C(sp^3)$-$C(sp^2)$ rotational barrier for ethylbenzene (1.16 kcal/mol, experiment; 1.19 kcal/mol, MMFF94; 1.10 kcal/mol, "MP4SDQ/TZP").[42] In the Tripos validation study,[38] no bisphenyl group was contained in the 76 small molecules studied and no $C(sp^3)$-$C(sp^2)$ rotational barrier was contained in the 15 torsional barriers studied. It should be noted that the two force fields use different torsional potential functions (Tripos – one term, MMFF94 – three terms).

4.2.3 C-O Internal Rotation

Calculations of $C(sp^3)$-$O(sp^3)$ internal rotation in the model compound **5** (Figure 3.6) showed that the MMFF94 force field allows a broad region of low energy between -60° and 60°, with a very low barrier (0.49 kcal/mol) at 0°, in contrast to the high barrier (4.44 kcal/mol) at -30° noted with the Tripos force field. The HF/6-31G(d) and B3LYP/6-31G(d) calculations were in qualitative agreement with the MMFF94 results. It has been demonstrated that MMFF94, parameterized to the "MP4SDQ/TZP" level in conjunction with MP2/6-31G*-optimized geometries, reproduces experimental conformational energies more accurately than the HF/6-31G* or MP2/6-31G* methods.[42] Of particular interest is the case of methylethyl ether, where the experimental value (in kcal/mol) for the gauche-anti energy difference is 1.5, the "MP4SDQ/TZP" result is 1.41, and the MMFF94 result is 1.5.[42] MMFF94 also reproduces the experimental barrier for C-O rotation in dimethyl ether, whereas Tripos significantly overestimates it (in kcal/mol): 2.7-experiment,[55] 2.43-MMFF94,[42] and 4.2-Tripos.[38] Various conformations of 14 different ethers were used in the core parameterization of MMFF94.[39]

The difference in how the Tripos and MMFF94 force fields treat $C(sp^3)$-$O(sp^3)$ internal rotation is responsible for the fact that, in the random search calculations, the local minima tend to collect at B4 values equal to $\pm$60° for the Tripos force field, but at B4 values between -60° to 60° for the MMFF94 force field.

4.3 Summary

In summary, from comparison of the conformer populations of **2** and **3** generated by the Tripos and MMFF94 force fields, it seems that there are subtle differences in how a tertiary amine type nitrogen is handled by the force fields. The force fields also differ significantly in their description of C-O internal rotation and the effect of the bisphenyl moiety on C-O internal rotation. These subtle differences affect some characteristics of the conformer populations collected during a random search. Ultimately, these differences could affect the results of a 3D-QSAR analysis which depends on the selection of certain representative conformers from a conformer population to act as templates for the analysis. The present work suggests that it is important to check the applicability of the force field for the types of questions one is attempting to answer by computer modeling. In previous work, we used the Tripos vacuum phase conformer populations of **2** and **3** to select representative conformers[56] as templates for 3D-QSAR analysis of a series of 48 GBR 12909 analogs that differ by changes in the A-side only.[35] This seems an acceptable choice since the differences in the Tripos and MMFF94 force fields were found to occur only for the nitrogen and oxygen on the B-side of the analogs.

4.4 Comparison of Conformer Populations of 2 and 3 to Methylphenidate

Methylphenidate, **4**, and the GBR 12909 analogs **2** and **3** share some pharmacophore features that are typical of most dopamine reuptake inhibitors: an aromatic group in close proximity to a basic nitrogen. Torsional angles A1 and A2 control the relative orientation of these important functional groups in **2**, **3**, and **4**. In order to compare the behavior of the A-sides of **2** and **3** to **4**, the conformer populations of all three (from random search calculations with the vacuum phase Tripos force field, Gasteiger-Hückel charges, and

distance dependent dielectric function with dielectric constant set equal to 1) were plotted in (A1, A2) space. The conformer populations of **4** were taken from our previous work.[48]

Figure 4.1 compares the location of all the local minima of **2** and **3** to those of protonated **4** in (A1, A2) space and shows that methylphenidate is far more conformationally restricted than **2** and **3**. Although **4** has far fewer minima, they are located in roughly the same region of (A1, A2) space as those of **2** and **3**. This suggests that **2**, **3**, and **4** may share a common pharmacophore for DAT binding that involves the aromatic ring and nitrogen. However, the nitrogen in **4** is separated from the aromatic ring by three bonds, while the nitrogen in **2** and **3** is separated from the aromatic ring by only two bonds. For this reason, Figure 4.2 compares the distance of the nitrogen from the centroid of the phenyl ring in **4** to the distance of the nitrogen in **2** and **3** from the centroid of ring 1 of naphthalene. A distance range of 3.75-4.35 Å is common to all the three analogs. For **2** and **3** all the conformers are in this range, while for **4**, 43% of the conformers are in this range. The fact that **2**, **3**, and **4** can orient their common pharmacophore elements in the same way suggests that the A-side of the GBR 12909 analogs may attempt to bind to the DAT in a way similar to that of methylphenidate. However, the long B-side of the GBR 129009 analogs may influence this interaction through a range of additional interactions with the DAT protein. A more definitive understanding of the DAT pharmacophore awaits the study of more rigid dopamine reuptake inhibitors, such as the recent study of rigid analogs of methylphenidate.

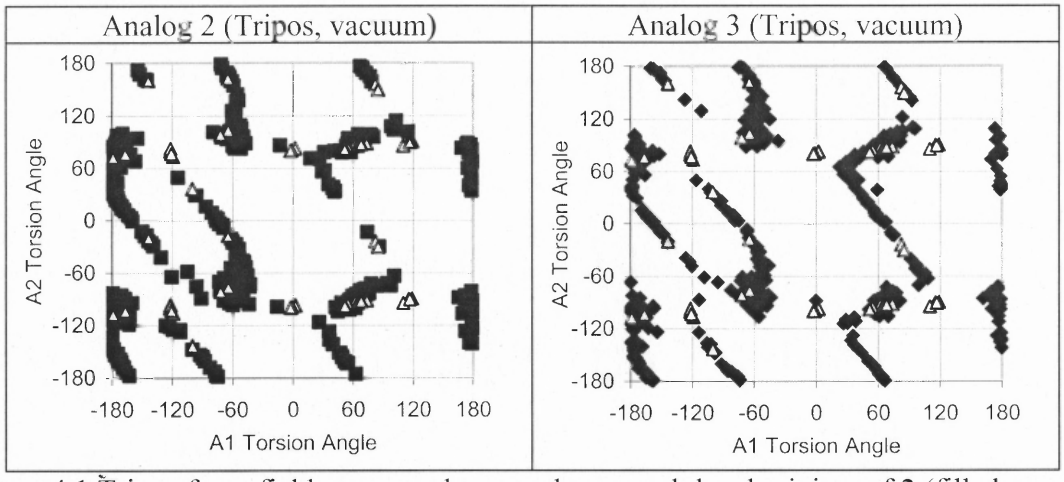


Figure 4.1 Tripos force field vacuum phase random search local minima of **2** (filled squares) and **3** (filled diamonds) compared to local minima of **4** (open triangles) in (A1, A2) space. Torsion angles are given in degrees.

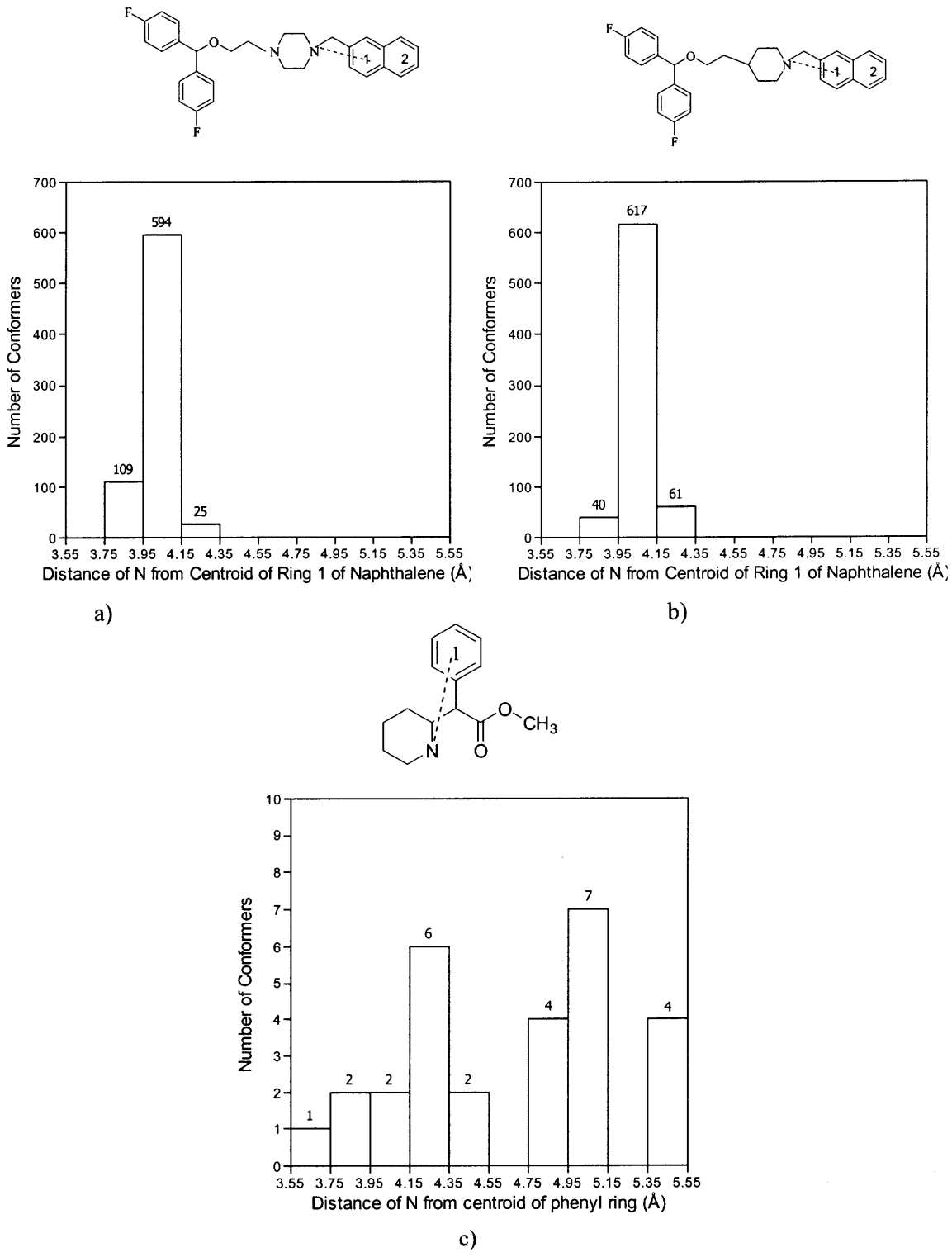


Figure 4.2 Number of conformers with given distance of N from centroid of the aromatic ring. Tripos in vacuum. Conformers having distance greater than 3.55 and less than or equal to 3.60 were assigned to bin labelled as 3.55, and so on. (a) **2**, (b) **3**, (c) **4**.

PART II

3D-QSAR STUDIES OF SALVINORIN A ANALOGS AS KAPPA OPIOID AGONISTS

CHAPTER 1

INTRODUCTION

1.1 Opioid Receptors

Opioid receptors[1] (OR, Figure 1.1) are members of the seven transmembrane-spanning. G-protein coupled receptor (GPCR) superfamily and the subfamily of the rhodopsin receptor. The genes for four opioid receptors (the mu (μ) opioid receptor (MOR), the kappa (κ) opioid receptor (KOR), the delta (δ) opioid receptor (DOR), and the nociceptin/orphanin FQ receptor (NOP-R) also known as ORL-1) have been cloned.[2,3] In general, μ, κ, and δ opioid receptors are about 60% identical to one another in terms of amino acid composition. The transmembrane domains and intracellular loops have the greatest similarity (73–76% and 86–100%, respectively), while the N terminus, extracellular loops and C terminus regions are the most structurally divergent.

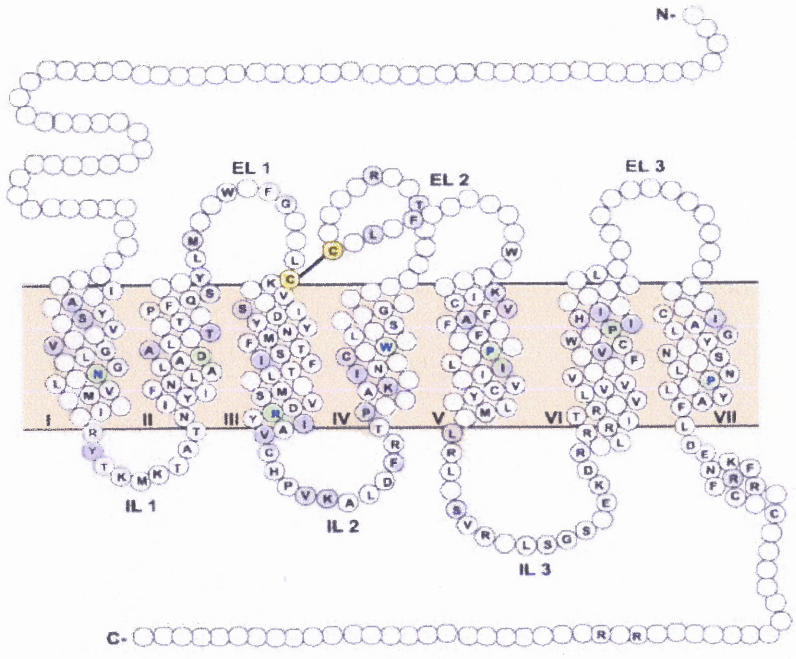


Figure 1.1 Schematic representation of the opioid receptor.[3]

Opioid receptors can regulate the range of effectors, which are small molecules that, when bound to an allosteric site of a macromolecule, cause either a decrease or an increase in the activity of the macromolecule. They can regulate macromolecules such as adenylate cyclase, Ca^{+2} channels, phospholipase C, K^{+2} channels and mitogen-activated protein kinases.[1] By virtue of these effectors, ORs are capable of participating in the process of signal transduction in which a signal from outside the cell (for example, from the binding of OR ligands) can control function within the cell (for example, through an interplay between the OR and effectors). ORs are activated both by endogenous peptides, including dynorphins, enkephalins and endorphins, as well as exogenous opioid drugs such as morphine and heroin.[3] Dynorphins and enkephalins are considered to be of major importance to the central nervous system (CNS) because they are the neurotransmitters involved in pain perception, cognitive function and endocrine functions. Beta (β)-endorphin is not considered to be as important because it is expressed in low levels and is only associated with the neuronal pathway originating from hypothalamic nuclei. Dynorphin A interacts preferentially with the KOR while enkephalin and β-endorphin mainly interact with the MOR and DOR. According to the "address-message"[4,5] concept, the "message" component of the ligand or peptide specifies general OR recognition and the "address" portion confers selectivity by specific recognition at a particular OR. Furthermore, the "message" part is connected to signal transduction, while the "address" part provides additional binding affinity and is not necessary for signal transduction.

Ligands for all three opioid receptors (MOR, DOR and KOR) have some analgesic effect,[6] however all lead to unpleasant side effects such as addiction, sedation, decreased respiratory function, seizure and gastrointestinal complications. It has been

recently noted that KOR agonists of the peripheral nervous system (that is, the nerves and neurons that reside outside the CNS to serve the limbs and organs) would have few side-effects and limited abuse potential.[7] Recently, salvinorin A was identified as a highly-selective (higher binding affinity for the KOR than the MOR or DOR) and efficacious KOR agonist.[8] Salvinorin A can impair perception and coordination but at present limited information is available on such issues and more data is required to understand its effects on humans.[9] Recently, in a comprehensive review,[10] it was noted that KORs appear to be implicated in the modulation of the abuse-related effects of CNS stimulants and may antagonize cocaine's abuse-related effects. For example, KORs appear to modulate DA levels[11] and affect cocaine-induced locomotor activity in rats.[12] KOR agonists may, therefore, present a new direction in the search for a treatment for cocaine abuse. Although these compounds are effective in reducing cocaine self-administration in monkeys, they produce unpleasant side effects.[13] The inclusion of MOR agonist/antagonist activity in the design of novel KOR agonists has been suggested as way to minimize these side effects[14] and possibly lead to new agents for the treatment of cocaine abuse. Thus it is possible that a study of salvinorin A and its analogs may help to illuminate the role of the KOR system in addiction and analgesia in humans.

1.2 Salvinorin A

Salvinorin A has been recently identified as a naturally-occurring hallucinogen, (**1**, Figure 1.2) from the plant *Salvia divinorum* which activates the KOR selectively and potently.[8,15] Salvinorin A is unique as it is the only known natural non-nitrogenous agent that binds to the human KOR.[8] It does not show structural similarity to other known nonpeptidic opioid receptor ligands (such as **2-6**, Figure 1.2) and due to the absence of nitrogen, it also shows qualitatively different binding than other traditional KOR ligands. It does not show any significant activity at the MOR, DOR, ORL-1, or other tested GPCRs, neurotransmitters or ion channels. Due to the role of KORs as a target of psychomimetic agents and because salvinorin A is a selective agent, it may represent a novel lead compound for psychotherapy with possible potential in the treatment of schizophrenia, dementia, and bipolar disease. In the only reported behavioral study of salvinorin A in nonhuman primates,[16] salvinorin A also showed potential to reduce cocaine self-administration. It can also serve as a novel template for the development of non-addictive analgesic opioids.[17,18] Since salvinorin A has only recently been discovered as a KOR agent, limited data on biological effects are available. But initial results seem to indicate that salvinorin A may have potential usefulness in the area of analgesic and addiction research.

1.2.1 Structure-Activity Relationships of Salvinorin A Analogs

The exact nature of the molecular interactions between salvinorin A and the KOR is not known as no crystal structure of the KOR is available. Recent experimental structure-activity relationship (SAR) studies of salvinorin A analogs showed that the methyl ester

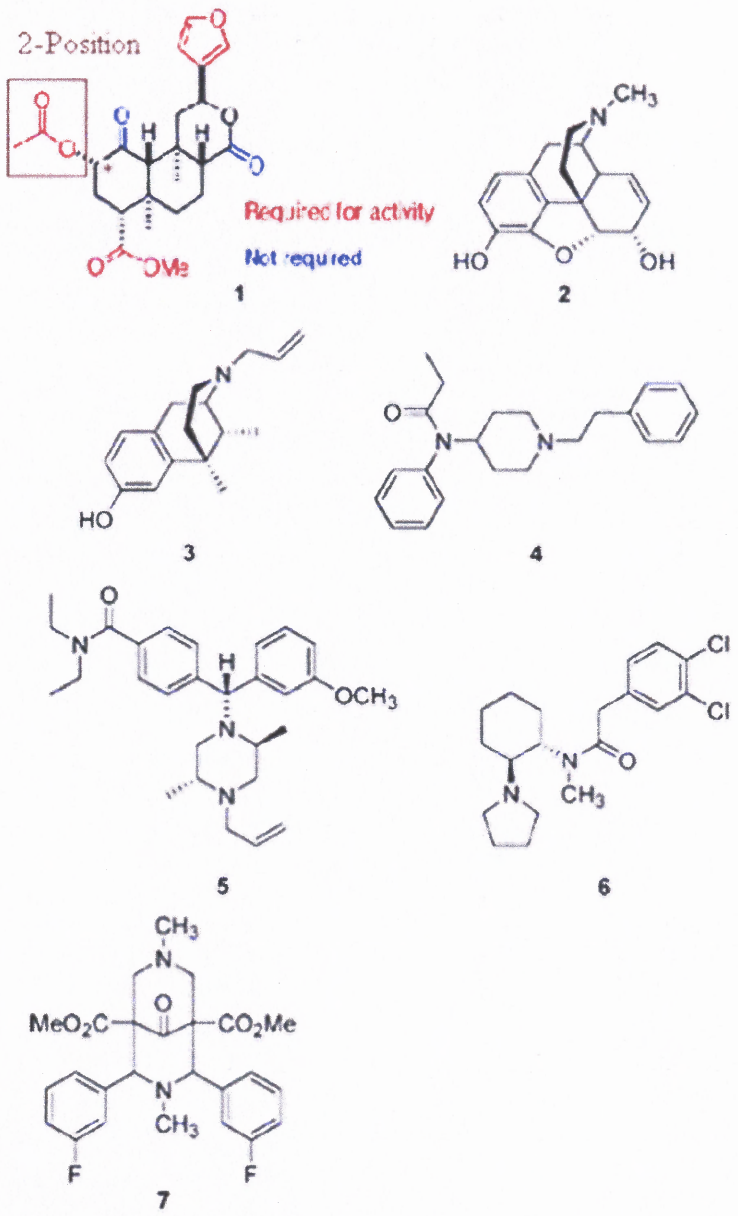


Figure 1.2 Structures of opioid ligands:[25] salvinorin A (**1**) (showing the importance of various moieties to KOR binding affinity), morphine (**2**), cyclazocine (**3**), fentanyl (**4**), SNC 80 (**5**), U50,488H (**6**), and 3FLB (**7**).

and the furan ring are required for activity but that the lactone and ketone functionalities do not play a role in binding (See **1** in Figure 1.2).[19,20] The nature of the 2-position[21] substituents of salvinorin A also plays a critical role in binding affinity (K_i) at the KOR[15,21-24] and MOR.[25,26] However, the primary focus of the present modeling study is to understand the effect of the 2-position substituents of salvinorin A on the KOR binding affinity. Prisinzano, Rothman, and coworkers[20, 23, 24, 26,27] have studied a series of about 55 salvinorin A analogs, of which 33 were used in the present three-dimensional quantitative structure-activity relationships (3D-QSAR) modeling study (Table 1.1). A C-8 hydrogen was found to be favored at the β-position as opposed to the α-position. The major KOR SAR used for this study is summarized here. Comparison of **1-6** in Table 1.1 shows the effect of the 2-position side chain length on the KOR binding affinity. Increasing the length by one or two carbons compared to **1**, as seen in **2** and **4**, had little effect on the KOR affinity, while additional chain length (**5** and **6**) decreased KOR affinity. The effect on KOR affinity was further explored by the addition of an aromatic ring at position 2 as in **8-16** and **21-30**. Increasing the number of methylene spacer units between the carbonyl and phenyl groups (as in **13** and **14**) resulted in decreased affinity at the KOR compared to **8**.

Eight analogs (**9-11** and **21-25**) explored the effect of ring substitution at the 2-position. The effect of halogen substitution on the phenyl ring can be seen by comparing the affinity of **9**, **10**, and **11** (bromine substitution) or **24** and **25** (fluorine substitution) to that of **8**. Substitution of bromine at positions 2 or 3 (**9** and **10**, respectively) had little effect on KOR affinity. In contrast, substitution at the 4-position of the benzene ring (**11**) had the greatest effect in terms of decreasing the KOR affinity. Fluorine substitution at

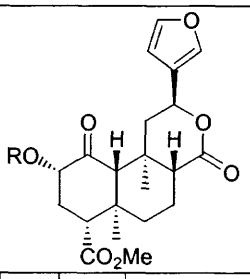
the 2 and 3 positions (**24** and **25**, respectively) paralleled the affinity changes noted with bromine, except that KOR affinity was decreased to a greater extent. The effect of methoxy substitution on the phenyl ring (**21-23**) was to decrease KOR affinity.

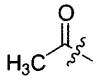
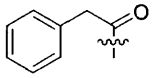
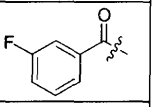
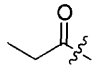
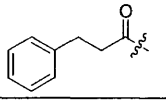
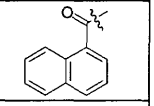
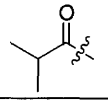
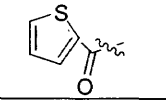
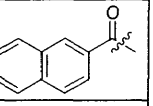
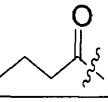
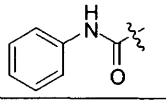
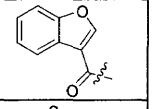
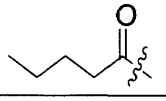
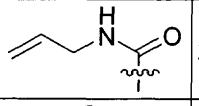
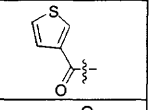
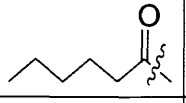
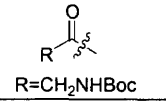
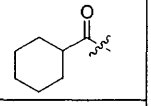
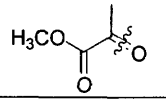
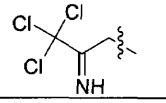
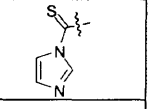
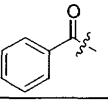
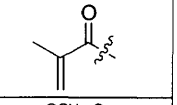
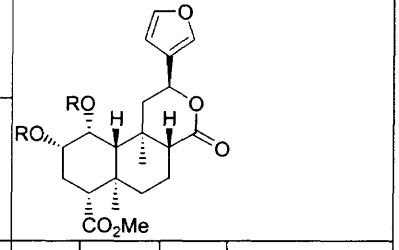
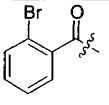
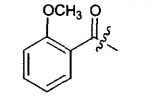
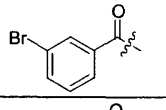
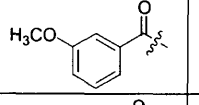
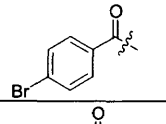
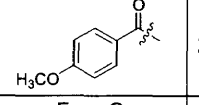
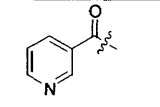
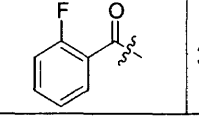
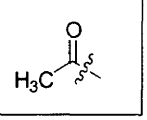
The effect of the size and nature of the aromatic substituent at the 2-position of salvinorin A was explored in analogs **26-30**. The extended aromatic system, 1-naphthalene (**26**) has reasonable binding affinity at the KOR compared to 2-naphthalene (**27**) which has significantly poorer KOR affinity. However, use of a sulfur in the aromatic system, i.e. 2-thiophene and 3-thiophene (**15** and **29** respectively), led to KOR binding affinities on par with **8** for **29** but a slight drop was observed for **15**. Use of an oxygen in the aromatic system (**28**) resulted in almost the same KOR affinity as **8,** but use of a nitrogen in the aromatic system (**12**) significantly decreased KOR affinity. Analogs **32** and **33** show the importance of the methoxy moiety because in its absence, binding affinity drops.

Several analogs studied by the Prisinzano and Rothman groups were not included in the present modeling study and are not shown in Table 1.1. Analogs with 2-position sulfonate substituents were not included because it has been pointed out that they might be binding in a different manner at the KOR.[25] Analogs with changes in the furan ring were excluded because the focus of the present study is on the 2-position substituents. This position has the most SAR data associated with it and the KOR binding affinity appears to be very sensitive to changes at this location. The 2-position trimethyl silicate analog was not included because molecular mechanics parameters for silicon are not available.

In summary, the SAR data in Table 1.1 contain a wealth of information relating KOR affinity to changes in the molecular structure and properties at the 2-position of salvinorin A. Certain substituents provide only a change in the shape (or bulkiness) of the molecule. Others, due to their electron-donating or electron-withdrawing properties, affect the electronic properties of the molecules as whole. Some substituents provide various combinations of steric and electrostatic effects. However, the large amount of information in Table 1.1 is difficult to visualize as a whole. For that reason, use of the 3D-QSAR techniques described in Chapter 2, in combination with further synthetic and pharmacological studies from the Prisinzano and Rothman laboratories, may prove to be useful in the design of novel KOR agents.

Table 1.1 Binding Affinities of Salvinorin A Analogs at the Kappa Opioid Receptor (KOR) Using [^{125} I] IOXY as Radioligand. Reference listed in the superscripts. Test set marked by asterisk.



#	K_i nM	pK_i	R	#	K_i nM	pKi	R	#	K_i nM	pK_i	R
1[25]	1.9	8.72		**13***[24]	290	6.54		**25**[26]	320	6.49	
2[25]	1.8	8.74		**14**[24]	**180**	**6.74**		**26**[26]	410	6.39	
3[25]	19	7.72		**15**[24]	260	6.59		**27**[26]	5490	5.26	
4*[24]	4	8.40		**16**[25]	93	7.03		**28**[26]	70	7.15	
5*[24]	15	7.82		**17**[25]	120	6.92		**29**[26]	80	7.10	
6[24]	70	7.15		**18***[24]	90	7.05	 R=CH_2NHBoc	**30**[27]	2010	5.70	
7[25]	430	6.37		**19***[25]	64	7.19		**31***[26]	2105	5.68	
8[24]	90	7.05		**20**[25]	42	7.38					
9[24]	90	7.05		**21**[26]	230	6.64					
10*[24]	70	7.15		**22**[26]	550	6.26		#	K_i nM	pK_i	R
11[24]	740	6.13		**23**[26]	540	6.27		**32**[25]	3190	5.50	H
12[25]	1930	5.71		**24**[26]	320	6.49		**33**[25]	650	6.19	

1.2.2 Molecular Modeling of the Salvinorin A-KOR Complex

To date, four different models (Figure 1.3) have been proposed for the binding orientation of salvinorin A in the KOR site. The models incorporated, where available, information from mutational and/or chimeric studies of the KOR. Each binding model was based on a different molecular model of the KOR. Each of the KOR models was derived using a different modeling protocol. No X-ray crystal structure of the KOR exists. The opioid receptors are GPCRs and the only X-ray structures of GPCRs available are those of bovine rhodopsin. As will be seen below, subtle differences in the underlying KOR molecular models significantly impacted the postulated binding models of salvinorin A.

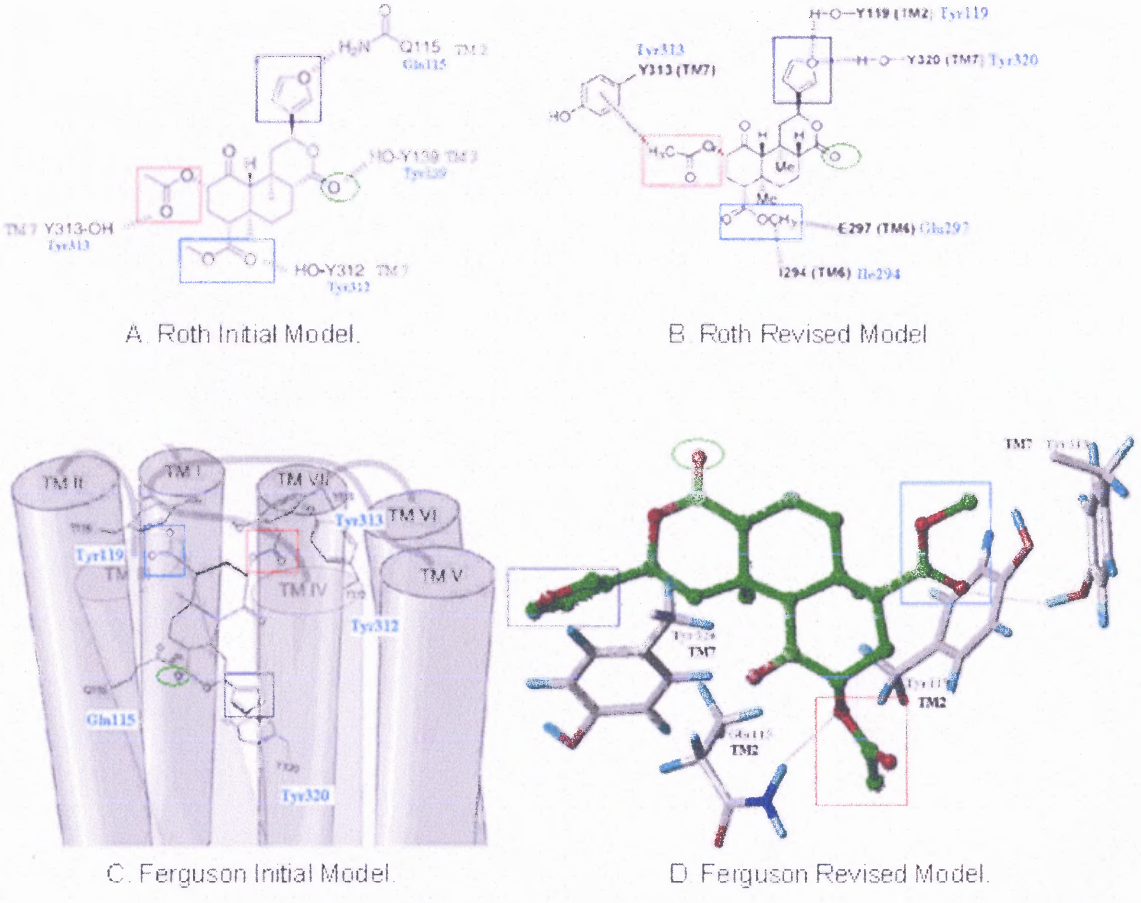


Figure 1.3 Molecular models of the salvinorin A-KOR Binding. Groups postulated to be important for the activity in colored boxes.[8, 28, 29, 31]

Roth and coworkers[8] initially proposed a model (Roth Initial Model, Figure 1.3a) based on a structure of the KOR,[28] derived without any experimental structural data for bovine rhodopsin, by means of a distance geometry algorithm using hydrogen bonding constraints.[28] Subsequently this KOR model was revised[29] (Roth Revised Model, Figure 1.3b) by homology and molecular modeling based on a 2000 X-ray structure of rhodopsin.[30] The initial model of Ferguson and co-workers[31] (Ferguson Initial Model, Figure 1.3c) was derived using their 1996 homology model of the KOR.[32] Their revised model[33] (Ferguson Revised Model, Figure 1.3d) was derived by homology and molecular modeling of the KOR based on a 2004 X-ray structure of rhodopsin with a bound antagonist.[34] Energy minimization and molecular dynamics simulation of salvinorin A in the binding site were carried out to obtain a KOR-agonist bound conformation.[33]

Figure 1.3 shows that the four models differ in several significant aspects. Colored boxes surround those chemical moieties postulated by all four models to be important to salvinorin A-KOR binding: the 2-position substituent (red box), the 4-position substituent (blue box), and the furan ring (black box). The initial Roth model postulated the involvement of the lactone carbonyl group (green circle) in molecular interactions with the KOR (hydrogen bonding to Tyr139), a hydrogen bonding interaction between the 2-position carbonyl oxygen and the hydroxyl hydrogen of Tyr313, a hydrogen bonding interaction between the furan ring and the amino group of Gln115, and a hydrogen bonding interaction between the 4-position carbonyl oxygen and the hydroxyl hydrogen of Tyr312. Since mutation of Tyr139 to Ala had essentially no effect on salvinorin A binding affinity and mutation to Phe had only modest effects,[29] the revised Roth model (Figures 1.3b) no longer postulated a hydrogen-bonding interaction between

the lactone carbonyl and Tyr139 of the KOR. In addition, mutating the Tyr313 residue to Phe resulted in no loss of binding affinity,[29] while mutating it to Ala gave a dramatic decrease in affinity. This indicated that a hydrophobic interaction between the 2-position methyl group and the aromatic ring of Tyr is more likely, as shown in the revised model (Figure 1.3b). Also mutating Try119 or Try320 to Phe or Ala resulted in significantly decreased affinity, and mutation of Tyr312 to Phe or Ala had little affect on binding affinity[29] so the revised Roth model postulated the stabilization of the furan ring by hydrogen bonding interactions with the hydroxyl hydrogens of Tyr119 and Tyr 320 and postulated stabilization of the 4-position substituent by hydrophobic interactions with the lipophilic portions of Glu297 and Ile294.

The initial Ferguson model[31] was developed to explain the results of chimeric opioid receptor and single point mutational studies.[31] As above, mutation of Gln115 to Ala resulted in significant decrease in affinity, but the initial Ferguson model (Figure 1.3c) postulated a hydrogen bonding interaction between the amino group of Gln115 and the lactone carbonyl oxygen of salvinorin A. This is in contrast to the initial Roth model which postulated hydrogen bonding between the furan oxygen and Gln115, as well as the revised Roth model which eliminated the lactone oxygen as a significant participant in the binding interaction. The revised Roth model postulated the stabilization of the furan ring of salvinorin A by both Try119 and Tyr320. However, in the KOR model used for the initial Ferguson model, these residues are 15 Å apart. As a result, the initial Ferguson model postulated a π-stacking interaction between the furan ring of salvinorin A and the aromatic ring of Tyr320 at one end of the binding site, and a hydrophobic interaction between the methyl group of the 4-position of salvinorin A and the aromatic ring of

Tyr119 at the other (see Figure 1.3c). This latter interaction is similar to that postulated to be between the 2-position methyl and Tyr313 in the revised Roth model. In the initial Ferguson model, the 2-position side chain was stabilized by both Tyr312 and Tyr313, whereas in the revised Ferguson model (Figure 1.3d), the ester oxygen of the 2-position side chain was postulated to interact with the amino group of Gln115 and the 4-position side chain was stabilized by interaction with Tyr313. The initial and revised Ferguson models agree in involving the furan ring in a π-stacking interaction with Tyr320, but differ in the other major points of the interaction.

In summary, it is clear that differences in the underlying KOR molecular model can lead to significantly different postulated binding models. The difficulty in constructing an accurate salvinorin A-KOR binding model lies not only in determining an accurate protein structure but also in deciding on the most appropriate orientation of the ligand in the binding site. The modeling of salvinorin A binding to the κOR is not straightforward for two reasons. First, salvinorin A lacks the nitrogen which is used by modelers to anchor the other KOR ligands in the binding site since it is believed that cationic amino charge on the opioid ligand interacts with the side chain carboxyl group of an aspartate residue (Asp138) located in TM III of the opioid receptor.[25] In the absence of a nitrogen, it is not obvious how salvinorin A orients itself in the site. Second, the homology between the κOR and rhodopsin is only in the range of 25-29% which is much less than the 50% homology recommended to obtain a reasonable model using homology modeling techniques.[35] Although the four models were developed in conjunction with the available data from mutational, chimeric, and SAR studies, it appears that there may still be some uncertainty as to the optimal salvinorin A-KOR binding mode. The residues

responsible for binding and the types of interactions involved in the binding differ considerably from one model to another. Clearly, the determination of the precise location of salvinorin A in the KOR binding site is a challenge yet to be met. For this reason, ligand-based modeling techniques, such as the 3D-QSAR methods described below, provide a complementary approach to the protein modeling studies. Since the ligand-based techniques are independent of the protein structure, they do not suffer from the drawbacks associated with protein homology modeling and the uncertainty in the appropriate ligand orientation in the binding site.

1.3 Ligand-Based 3D-QSAR Modeling

There is no X-ray structure of the KOR and homology modeling of the protein would constitute material for a thesis in itself. As discussed above, four conflicting models of salvinorin A-KOR binding interaction based on homology models of the KOR have been proposed. Also, the SAR of the analogs in Table 1.1 does not clearly indicate whether hydrophobic or hydrogen bonding interactions contribute to the formation of the ligand-KOR complex. In this scenario, 3D-QSAR studies may shed some light on the types of interactions with the receptor that are favored at certain locations in the ligand, thereby complementing the techniques used to develop the binding models.

Comparative Molecular Field Analysis (CoMFA)[36,37] is a 3D-QSAR method which relates the shape-dependent steric and electrostatic fields of molecules to their biological activity (here their KOR binding affinity). This technique describes 3D structure-activity relationships in a quantitative manner. CoMFA models in the present work were developed in an attempt to elucidate the molecular properties of the 2-position

salvinorin A analogs responsible for their binding to the KOR. As will be seen in the Results section, CoMFA models identify the location of those steric and electrostatic fields around the analogs that contribute significantly to the binding affinity. This will provide some understanding of the "hot spots" in the molecular substructure which are most involved in the binding and may provide some insight into what chemical changes could lead to optimal interaction between the ligand and the receptor.

CoMFA has been extensively used since its introduction in 1988. It is evident from the fact that the original CoMFA paper[36] has more than a thousand citations, and this does not take into acccount the unpublished CoMFA studies done in drug companies. CoMFA studies have been conducted on ligands for the MOR and DOR as well as other KOR ligands,[38-43] but the present work is the first CoMFA study on salvinorin A analogs. There has been only one ligand-based modeling study of the KOR binding affinity of salvinorin A analogs,[33] but the CoMFA methodology was not used in that study. The weakness of that study lies in the fact that it used SAR data compiled from several different laboratories. This is problematic since the same analogs, when tested in different laboratories, can show a different rank order of binding affinity due to the subtle differences in the pharmacological protocol. The theory and principles behind CoMFA are discussed in the next chapter.

Objective and Significance

The objective of this work is to model the KOR binding affinity of 2-postion salvinorin A analogs in order to propose new analogs with improved binding affinity. The modeling will be carried out using the 3D-QSAR method of CoMFA to identify regions of salvinorin A where changes in steric bulk or electrostatic properties, for example, may to lead to analogs with improved binding affinity for the KOR. Our results will be shared with the Prisinzano and Rothman groups to assist in the identification of novel salvinorin A analogs which will be synthesized and tested for their KOR binding affinity with the ultimate goal of designing drugs with potential usefulness in the area of analgesic and addiction research.

CHAPTER 2

BACKGROUND

2.1 3D-QSAR

The first step in the series of biochemical events which leads to the observed biological activity of a drug is its interaction with the target macromolecule. This interaction is influenced by the degree of steric and electrostatic complementarily between the surface of the drug (or ligand) and that of its protein binding site. The biological activity is some experimentally measurable property of the drug-receptor complex, such as the binding affinity of the drug for the protein. A 3D-QSAR study is the analysis of the quantitative relationship between the biological activity of a set of compounds and their molecular properties (based on steric, electrostatic, hydrophobic and other molecular "fields") calculated in 3D space using statistical methods. In the absence of structural data about the receptor, as in the present study, this type of analysis may help to identify the molecular features of a ligand responsible for its binding affinity and to predict the biological activity of ligands not yet tested The underlying assumptions of 3D-QSAR methods discussed by Oprea and Waller[44] are: 1) The compounds in the modeling study cause the observed biological effect; 2) Only a single conformation of the drug is involved in binding; 3) All modeled compounds bind at the same site; 4) The measured biological activity can be explained mainly as the result of enthalpic processes; 5) Entropic terms are similar for all modeled compounds; 6) The system is considered to be at equilibrium and kinetic parameters are not considered; i.e., the on and off rates are similar for all the compounds; and 7) Solvent effects, diffusion and transport are not

necessary to explain the observed biological activity and so are not included in the modeling calculations.

There is considerable evidence that ligands do not bind to proteins in their lowest energy (or global energy minimum) conformation.[45-49] Flexible ligands may take on a large range of conformations that are very similar in energy. Rigid ligands, in contrast, have only a small number of conformers available at low energy. As a result, rigid compounds give important clues as to the bioactive (or binding) conformation of the drug. If a rigid analog with only one possible low-energy conformation is shown to have high binding affinity for the receptor protein, then it is a reasonable assumption that its conformation is the bioactive conformation. A 3D-QSAR analysis requires that a particular ligand conformation be selected as a template for superposition (or alignment) of the other compounds in the study. The results of the analysis are very sensitive to the conformer chosen as the template. In the optimal situation, the set of compounds under study would contain a one-conformer rigid analog with high binding affinity. This compound would be used as the template, the other molecules in the series would be aligned to this structure, and the 3D-QSAR analysis would be carried out, potentially resulting in a predictive model relating molecular structure and properties to biological activity. However, many series of compounds of interest may not have such an obvious choice for the template, and a more flexible molecule must be chosen. In a study investigating how to best align molecules for 3D-QSAR, Klebe and Abraham[50] pointed out that it is not necessary for structurally similar compounds to superimpose exactly in the same binding site. The alignment used for superposition of the ligands, therefore, should not be assumed to be the conformation in which ligand binds to the

macromolecule. An effective molecular alignment should superimpose molecules to describe the cause of their different biological activities but not the consequence, i.e., it should reflect rather than reproduce reality.[51]

2.2 Comparative Molecular Field Analysis (CoMFA)

CoMFA[52,53] is a 3D-QSAR method which is performed on a series of molecules to compare differences in their molecular potential energy fields (by default steric and electrostatic fields, although other fields can be considered) and to correlate them to changes in biological activity.[44] The Lennard-Jones 6-12 function is used to calculate the steric potential and a simple Coulombic function is used for the electrostatic potential to describe the molecular interaction between each molecule in the series and a probe atom.

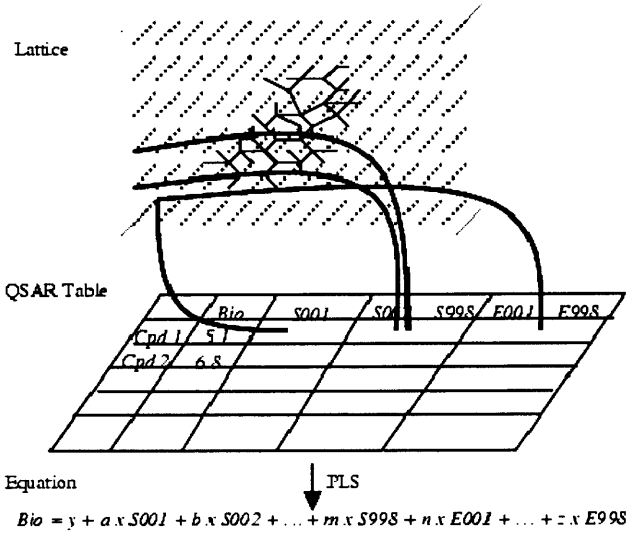


Figure 2.1 The CoMFA Process.
Source:Reprinted from SYBYL QSAR Manual, Tripos, Inc.

The CoMFA process is depicted in Figure 2.1. Each molecule in the series is aligned (or superimposed) according to a designated rule and then enclosed by a 3D lattice of regularly-spaced points. The boundary of the lattice walls extend at least 4 Å beyond the union volume of aligned or superimposed structures. The distance between neighboring grid points is typically 1-2 Å. Each row of the table in Figure 2.1 describes the fields exerted by a particular compound (Cpd1, Cpd2, ...) in the series, in a particular conformation, on any surrounding atoms.[51,53] In other words, each row corresponds to a compound in the study and the collection of columns for each row corresponds to its fields. For a particular ligand in the series, e.g. Cpd1, at the first grid point the total steric (S001) and electrostatic (E001) intermolecular interaction energies of each of the atoms in the ligand with each of the atoms in the probe molecule are calculated. The probe molecule is then moved to the second grid point and the calculations repeated to yield S002 and E002. This process is repeated for all the grid points. The same process is performed for all the other compounds in the series (Cpd2, Cpd3,...). The data table in Figure 2.1 contains the information used as input to the CoMFA calculation. It contains the experimental biological activity data ("Bio", the dependent variable) and the intermolecular interaction energies, or fields ("S001",..."E001",..., the independent variables). The objective of the CoMFA technique, as of any 3D-QSAR method, is to find an equation that relates the biological activity of a series of compounds to their molecular properties, such as steric and electrostatic fields.[53]

2.3 Principal Component Analysis (PCA) and Partial Least-Squares [44,51,54]

Typically, there are relatively few (20-100) molecules in the series, but thousands of molecular property values for each molecule. It is difficult or almost impossible to solve a series of such equations. Multiple linear regression (MLR) does not work as there are more molecular property values than compounds, and the properties are correlated. In this situation, principal component analysis (PCA) is useful. PCA constructs the directions or principal components ordered in terms of the amount of variance in a dataset. For example if the dataset is CoMFA field values (**X** matrix), the principal components capture the variation in the CoMFA fields. These components can be used to explain the biological activity (**Y** matrix). However, there is no guarantee that the principal components derived by PCA using CoMFA fields are relevant for biological activity. Partial least squares (PLS), or projection to latent squares, finds components from **X** that are also relevant for **Y**. This is achieved by searching for a set of components (PLS components or latent variables) which can explain the covariance between **X** and **Y**. In other words, latent variables relevant to the biological activity are constructed by the PLS algorithm. The following constraints are simultaneously satisfied by the PLS algorithm: 1) the latent variables are orthogonal, 2) error is minimized, and 3) the covariance between dependent variable (bioactivity) and the latent variables is maximized.

2.3.1 Cross-Validation

The internal predictive ability of a 3D-QSAR model is assessed by the technique of cross-validation. In this technique one or more compounds are excluded from the input data set and a PLS model is derived from the remaining compounds. Then the activity for the omitted compounds is predicted and the resulting individual squared errors of prediction are accumulated. Sample-distance Partial Least Squares (SAMPLS)[67] vastly speeds up crossvalidation calculations for PLS analyses involving CoMFA fields.

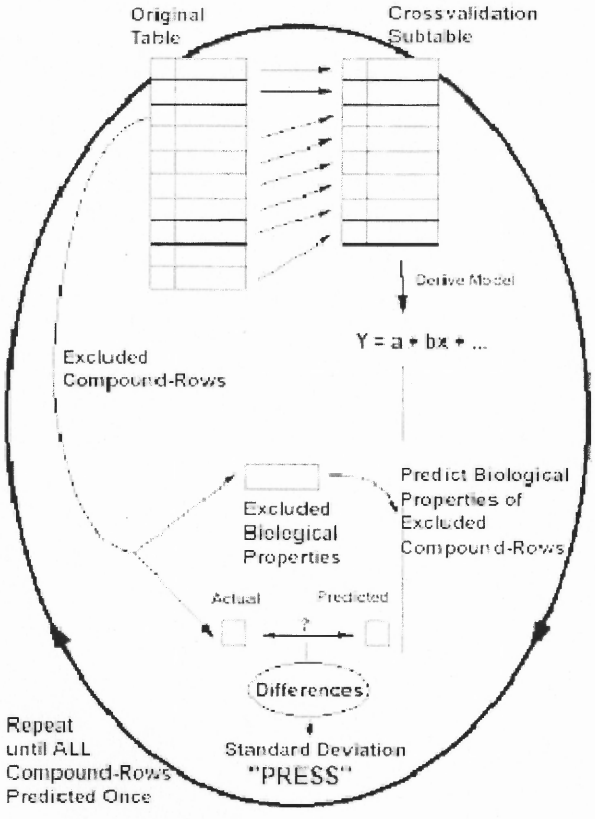


Figure 2.2 The crossvalidation process.
Source: Reprinted from SYBYL QSAR Manual, Tripos, Inc.

The important statistical parameters obtained from the cross-validation are the Predictive Residual Sum of Squares, (PRESS), the "cross-validated r-squared", q^2 or Q^2, and the cross-validated standard error of estimate, $s_{cv}(n)$:

$$\text{PRESS} = \sum_{i=1}^{N} \left(Y_{obs,i} - Y_{pred,i}\right)^2 \tag{2.6}$$

$$q^2 = 1 - \frac{\sum_{Y} \left(Y_{pred} - Y_{obs}\right)^2}{\sum_{Y} \left(Y_{obs} - Y_{mean}\right)^2} \tag{2.7}$$

$$S_{cv(n)} := \sqrt{\frac{PRESS}{n - c - 1}} \tag{2.8}$$

Here, Y_{pred} = predicted value of the biological activity, Y_{obs} = observed value of the biological activity, Y_{mean} = the best estimate of the mean of all values that might be predicted, n = number of rows and c = number of PLS components.

Generally, q^2 increases as the first few PLS components (or latent variables) are added to the model, then reaches a plateau or drops upon addition of more PLS components variables. This means that additional PLS components describe properties that are not related to the biological activity. These irrelevant PLS components should not be included in the model as they can have a detrimental effect on the predictivity of the model. The value of q^2 is generally between 0.0 (corresponding to no model) and 1.0 (corresponding to a perfect prediction). There is a possibility of getting a negative q^2 when one or more unique compounds are present in the dataset and cross-validation relies on the properties of a single compound. In this study the lowest number of latent variables (or components) with a reasonably high q^2 and the lowest standard of error of estimate was selected as the optimal number of PLS components.[55] The optimal number

of components was used to construct a non-cross-validated (NCV) model, or "full PLS" model, in which all the compounds of the dataset are used to develop the model. This full model gives the conventional r^2 (or R^2), calculated as shown in Equation **2.9**, which measures how well a particular model fits the data.

$$r^2 = 1 - \frac{SS(F)}{SS(Y)} \qquad \textbf{(2.9)}$$

Here, SS(F) is the sum of the squares of the y residuals and SS(Y) is sum of the squares of the y variables. The NCV model can be used to further improve q^2 by performing PLS Region Focusing,[56,57] an iterative procedure which refines a model by increasing the weight for those lattice points which are most relevant to the model.

2.3.2 Types of Validation

Though cross-validation is an internal validation method, the very nature of the cross-validation method overestimates the predictive power of the model based on datasets with redundancy in properties and underestimates the predictive power of the model in which each compound is unique in its set of properties.[51,58,59] The drawbacks of the cross-validation technique can be overcome in part by dividing the whole dataset into a "training" set and a "test" set and applying additional techniques to better assess the predictivity of the model and evaluate the stability and robustness of the model. The compounds of the training set are used to construct 3D-QSAR models which are then used to predict the activities of compounds in the test set. In addition, the technique of progressive scrambling (PS),[60] also known as y-scrambling, can be applied to the training set to address redundancy in the dataset, i.e. molecules which have similar CoMFA field

values. Compounds of the test set (compounds not used in developing the model) are used as a tool for external validation of the model.[59,61] The predictivity of the model is tested by seeing how well it predicts the known biological activity of these compounds.

a. Internal Validation: As the first step in PS, rows are sorted based on the biological activity. Then they are partitioned into a number of bins specified by the user. Within each bin, the dependent variables (biological activities) are scrambled or reshuffled a certain number of times specified by the user. Each such scrambling is characterized in terms of the correlation of the scrambled responses with the unperturbed data ($r^2_{yy'}$). SAMPLS[67] is applied to the perturbed data set to obtain the cross-validated correlation coefficient (q^2) and standard error of prediction (SDEP) both as a function of $r^2_{yy'}$. This process is repeated by decreasing the number of bins by 1 for each iteration until a user-defined minimum number of bins or 2 comes first. The value of cSDEP is obtained by calculating SDEP at a user-defined critical point, typically 0.85. The Q^2 (Equation **2.10**) and $\frac{dq^{2'}}{r^2_{yy'}}$ are the important statistics returned. Q^2 is the predictivity of the model after the

$$Q^2 = 1 - (sSDEP)^2 \qquad \textbf{(2.10)}$$

potential effects of redundancy have been removed, sSDEP is the scaled SDEP. $\frac{dq^{2'}}{r^2_{yy'}}$ is the instantaneous slope of the predictivity with respect to the degree of perturbation. This slope is the critical statistic. QSAR models which are unstable, i.e. models which change greatly with small changes in underlying response values, are characterized by slopes greater than 1.20.[60] Stable models, i.e. models which change proportionally with small changes in underlying data, have slopes near unity.

b. External Validation: In external validation, the models developed using the compounds of the training set are used to predict the activity of the compounds in the test set. The following statistical characteristics of the test set can be used to study the stability and robustness of the model:[59, 61]

i) R^2, correlation coefficient between the predicted and observed activities

ii) R_0^2, coefficient of determination between predicted versus observed activities and $R_0'^2$, coefficient of determination between observed versus predicted activities

iii) Slopes k and k' of the regression lines through the origin.

For a QSAR model to be predictive, the conditions listed below have to be satisfied:[59, 61]

$$q^2 > 0.5 \quad \textbf{(2.12)}$$

$$R^2 > 0.6 \quad \textbf{(2.13)}$$

$$\frac{\left(R^2 - R_0^2\right)}{R^2} < 0.1 \quad \text{or} \quad \frac{R^2 - R_0'^2}{R^2} < 0.1 \quad \textbf{(2.14)}$$

$$0.85 \leq k \leq 1.15 \text{ or } 0.85 \leq k' \leq 1.15 \quad \textbf{(2.15)}$$

2.4 Contour Maps

The CoMFA QSAR equation can be depicted as 3D contour maps. Here it is worthwhile to note that CoMFA contour maps do not show the commonality of all molecules but rather map the relationship between structural differences and differences in biological activity.[51] This type of map helps to identify the steric and electrostatic features beneficial to improved biological activity and also shows features which are detrimental to the biological activity. Figure 2.3 shows an example. Regions of favorable steric interactions for biological activity are shown in green, with sterically unfavorable regions in yellow. Red contours show regions where increasing the positive charge at that location on the ligand are correlated with decreased biological activity; blue contours indicate areas where increasing the positive charge on the ligand are correlated with improved biological activity. These maps can be used to suggest new molecules with increased biological activity.

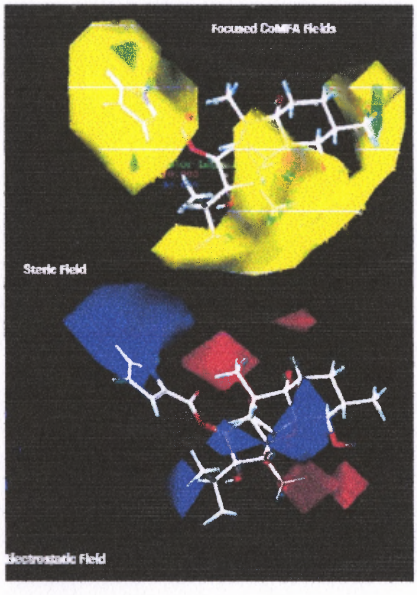


Figure 2.3 CoMFA contour maps.
Source: Reprinted from Advanced CoMFA brochure from Tripos, Inc.

CHAPTER 3

METHODS

3.1 Computational Resources

All calculations were carried out on Silicon Graphics Origin 2000 workstations at New Jersey Institute of Technology using the SYBYL molecular modeling program version 7.2.[62]

3.2 Series of Analogs

A series of salvinorin A analogs with 2-position substituents were selected for the study (Table 1.1). Biological activity was represented as pK_i, where $pK_i = -\log K_i$, and K_i is the binding affinity of the analog at the KOR. The series has a biological activity range spanning 3 log units, an important condition for a reasonable 3D-QSAR analysis. The series was divided into a training set of twenty-six compounds and a test set (marked by asterisks) of seven compounds. It is generally recommended[61] that the test set include no fewer than five compounds. Compounds in the test set were selected to be representative of the range of biological activity of the training set. All the analogs listed in Table 1.1 were created based on the crystal structure of salvinorin A (Cambridge Crystal Database code: DADMOK).[63] The geometry of the crystal structure was optimized using the Powell[64] method of minimization, the Tripos force field[65] and Gasteiger-Hückel charges.[66] Different side chains were added to this template structure at the 2-position and the geometry of the resulting analogs was optimized and stored in a database. The analogs were aligned using the standard database alignment option of SYBYL. The

extended 2-butanoate analog, **4**, was selected as the template for alignment. Before performing alignment, **4** was oriented using the SYBYL command "ORIENT BEST_VIEW". The "heavy" (non-hydrogen) atoms of the rings and oxygen of the 2-position substituents were selected as the common substructure for alignment. As described in Chapter 2, the steric and electrostatic interaction energies of each atom in the analog with a probe atom (sp^3 carbon with charge 1) placed at different points on a grid surrounding the molecule were calculated.

3.3 Preliminary and Focused CoMFA Models

The database of aligned molecules was used to create the molecular spreadsheets (MS). The rows of the spreadsheet contain the analogs while the columns contain the biological activity data (pK_i) and the steric and electrostatic interaction energies. The latter are refered to as CoMFA columns with field class Tripos Standard. Separate CoMFA columns were constructed with interaction energies from nine different steric field/electrostatic field cutoffs (in units of kcal/mol): 60/1, 30/1, 10/1, 60/10, 30/10, 10/10; 60/30, 30/30 (default) and 10/30, with all other CoMFA parameters set to default. Details of CoMFA input parameters are listed in Appendix A. Then the seven compounds of the test set were "hidden" in each MS and the partial least square (PLS) with leave-one-out (LOO)/cross-validation(CV) method using SAMPLS[67] was performed on the twenty-six compounds of the training set. The initial number of components for the LOO/CV method was set to six (default). The cross-validated model provided q^2 (predictivity) and standard error of prediction (SEP) for each component. The steric/electrostatic cutoff of 30/1 kcal/mol gave the best q^2 and the data in the

corresponding MS was used to develop the preliminary model. The optimum number components was selected based on the lowest SEP, which generally corresponded to the highest q^2 or a value greater than 0.5, and was used to create the non-cross-validated (NCV), or preliminary, model. The NCV model gave r^2 (goodness of fit), standard error of estimate (SEE), and the percent contribution of the steric and electrostatic fields.

The preliminary model was refined using Region Focusing,[57] an iterative procedure which increases the weight of those lattice points which are most relevant to the model. This enhances the resolution and predictive power (q^2) of a subsequent PLS analysis. StDev*Coefficient, the product of the variation at each lattice point with the lattice point's regression coefficient from the model, and Discriminant Power, that fraction of the variation in the model's components attributable to each lattice point, were used as weights. Exponential factors of 0.3 (default), 0.5 and 1, recommended by Tripos Bookshelf 7.2,[68] were applied to each weight, and CoMFA calculations were carried out using the 30/1 kcal/mol steric/electrostatic cutoff, giving a total of six focused models. The LOO/CV model with the lowest SEP (and highest q^2) in both the StDev*Coefficient and Discriminant Power cases was used to construct the focused NCV (full) model for each case. Both the focused NCV StDev*Coefficient and Discriminant Power models were subjected to validation testing as described in the next section.

To study hydrogen-bonding behavior, the H-bond field class was used with a probe atom type of H (charge 1) and a probe atom O.3 (i.e. sp^3 oxygen; charge -1). This was done after setting the "TAILOR SET COMFA SWITCH_FCN" setting to "NO". However, it was observed that adding the H-bond CoMFA field class gave models with decreased values of q^2 compared to those obtained with Tripos Standard CoMFA field

class calculations described above. For this reason further calculations with the H-bond field were not pursued.

Internal and External Validation In addition to the internal validation method of CV, progressive scrambling (or y-value scrambling)[60] tests were performed on the focused NCV StDev*Coefficient and Discriminant Power models. This estimates the stability of the model with respect to random noise in the data. The y-value is the biological activity (here, pKi at the KOR). This was done using 100 y-value scramblings with critical point of 0.85, maximum number of bins equal to 8 and minimum number of bins equal to 2. For the sake of reproducibility, the random seed was always chosen as "12345". All other parameters were set to default values. All other details including key statistics are explained in section **2.3.2**. Based on the key statistics, the focused NCV StDev*Coefficient and Discriminant Power models were identified either as stable or unstable. As will be seen in the Results section, both models were found to be stable and were used for external validation by predicting the biological activities of the compounds in the test set using the test set correlation method as explained in the section **2.3.2**. As will be seen in the Results section, only the focused NCV Discriminant Power model (with exponential factor 1) was found to be both stable and predictive. This model was used to predict the biological activity of novel compounds.

3.4 Novel Compounds

A 3D CoMFA contour map was created for the focused NCV Discriminant Power model (with exponential factor 1) and used to identify key locations in 3D space where changing steric and electrostatic characteristics would have an effect on the biological activity. Salvinorin A was used as a template for interpreting the contour maps. The Optimize QSAR procedure of SYBYL was used to replace selected hydrogen atoms with different substituents, their geometry was optimized, and the biological activity of the modified analogs was predicted. Those with the best biological activity (highest pK_i) were selected and recommended to the Prisinzano group for synthesis.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Alignment

The butanoate analog used as a template for the database alignment is shown in Figure 4.1a. The alignment of all the test and training set analogs is shown in Figure 4.1b. Since the structure of the analogs the same except for the 2-position, only that region shows variation.

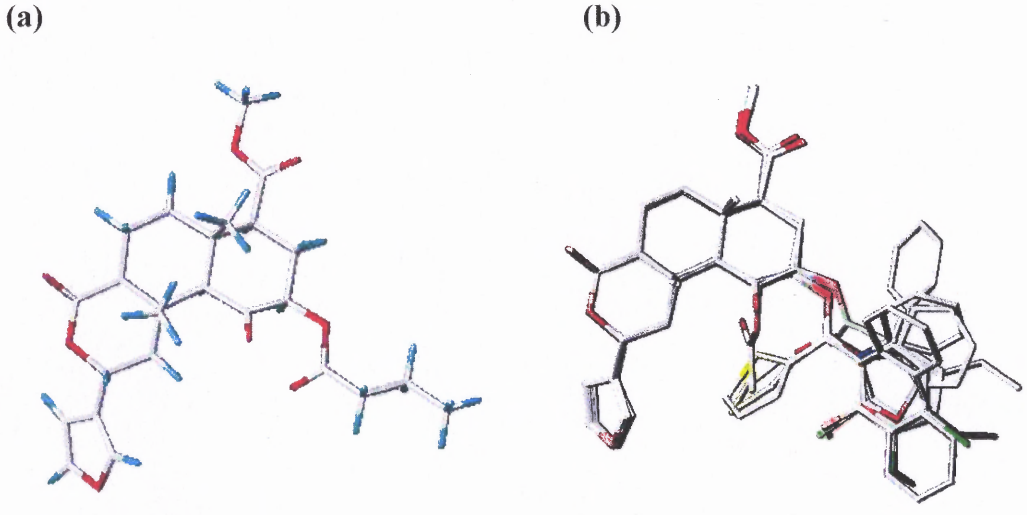


Figure 4.1 Alignment of salvinorin A analogs. a) Butanoate analog used as template shown with hydrogens, b) Aligned salvinorin A analogs shown without hydrogens for clarity

4.2 Preliminary and Focused Models; Internal Validation

The results of the NCV PLS preliminary model and the associated LOO/CV PLS model based on the steric/electrostatic cutoff of 30/1 kcal/mol are summarized in Table 4.1, along with the LOO/CV model for the default steric/electrostatic cutoff (30/30kcal/mol), given for comparison purposes only. Compared to the results obtained with the default cutoff, q^2 increased to 0.408 from 0.255 (15%) and SEP decreased to 0.670 from 0.769.

cutoff, q^2 increased to 0.408 from 0.255 (15%) and SEP decreased to 0.670 from 0.769. Although the q^2 value of 0.408 is below the widely-accepted value of 0.5[61,68,69] for a predictive CoMFA model, it is above the statistical 95% confidence limit of $q^2 = 0.3$.[70] Results of the LOO/CV models with other steric field/electrostatic field cutoffs (in units of kcal/mol: 60/1, 10/1, 60/10, 30/10, 10/10; 60/30, 30/30 and 10/30) are listed in Appendix B.

Table 4.1 Preliminary Results of the QSAR Study on Salvinorin A Analogs

LOO/CV PLS			**LOO/CV PLS**			**NCV PLS(Full Model)[a]**				
S/ES[b] cutoffs: 30/30 kcal/mol			**S/ES[b] cutoffs: 30/1 kcal/mol**			**S/ES[b] cutoffs: 30/1 kcal/mol**				
q^2	C[c]	SEP[d]	q^2	C[c]	SEP[d]	r^2	SEE[e]	F value[f]	% S[g]	% E[h]
0.255	3	0.769	0.408	2	0.670	0.796	0.670	44.806	50%	50%

[a] Non-cross-validated full models developed using the optimal number of components
[b] Steric/electrostatic cutoff parameters
[c] Optimal number of components selected from LOO/CV model
[d] Lowest standard error of prediction obtained from LOO/CV model
[e] Standard error of estimate
[f] Statistical measure of whether the model is significant
[g] Percent steric field contribution to full model
[h] Percent electrostatic field contribution to full model

The LOO/CV focused models developed using different weights and exponential factors are summarized in Table 4.2. The highest q^2 and lowest SEP in the StDev*Coefficient case ($q^2 = 0.529$; SEP = 0.612) occur for exponential factor 0.5; the highest q^2 and lowest SEP in the Discriminant Power case ($q^2 = 0.621$; SEP = 0.549) occur for exponential factor 1. NCV PLS (full) models were only developed for these two cases. As shown in Table 4.3, the r^2 and percent steric and electrostatic contribution for both these focused NCV models is similar. The Progressive Scrambling results are

also summarized in Table 4.3. The stability tests show that both models are robust and stable based on Q^2, cSDEP and $\frac{dq^{2'}}{r^2_{yy'}}$. Q^2 values are expected to be more conservative than LOO/CV q^2 values and even a Q^2 value as low as 0.35 signifies that the original, unperturbed model is robust.[68] Since the focused NCV StDev*Coefficient and Discriminant Power models have Q^2 equal to 0.416 and 0.498, respectively, they are robust. The cross-validated standard deviation of error of prediction (cSDEP) values are similar to SEP values obtained from the original, unperturbed models. The value of $\frac{dq^{2'}}{r^2_{yy'}}$ for both models is below 1.2 and also signifies stability.[60]

Based on these results, both models in Table 4.3 were selected as acceptable models and used to predict the pK_is of the compounds in the training set. Predictions from the models are compared to experimental values in Figure 4.2. The training set analogs are spread out uniformly around the ideal diagonal line for both the models. The training set predictions and residual values (predicted pK_i - experimental pK_i) obtained from both models are listed in Appendix C and are generally small. For the focused NCV StDev*Coefficient and Discriminant Power models, the average and minimum residuals for the training set are 0 and -0.67, respectively; the average absolute value of the errors are 0.26 and 0.25, respectively; and the maximum residuals are 0.64 and 0.55, respectively.

Table 4.2 Focused LOO/CV Models of the QSAR Study on Salvinorin A Analogs

LOO/CV PLS						
Steric/Electrostatic cutoffs: 30/1 kcal/mol						
Weights	**StDev*Coefficient**			**Discriminant Power**		
Weight Factors	**q^2**	**C[a]**	**SEP[b]**	**q^2**	**C[a]**	**SEP[b]**
0.3	0.489	3	0.637	0.515	3	0.620
0.5	0.529	3	0.612	0.588	4	0.585
1	0.524	3	0.614	0.621	3	0.549

[a] Optimal number of components selected from LOO/CV model
[b] Standard error of prediction

Table 4.3 Focused Full Models of the QSAR Study on Salvinorin A Analogs

Weights	**NCV PLS(Full Model)[a]** **Steric/Electostatic cutoffs: 30/1 kcal/mol**						**Progressive Scrambling**		
	C[b]	r^2	SEE[c]	F value[d]	% S[e]	% E[f]	$Q^{2\,g}$	cSDEP[h]	$\frac{dq^{2'i}}{r^2_{yy'}}$
StDev*Coefficient (weight factor: 0.5)	3	0.838	0.358	38.067	45%	55%	0.416	0.698	0.578
Discriminant Power (weight factor: 1)	3	0.857	0.336	44.108	47%	53%	0.498	0.630	0.543

[a] Non-cross-validated full models developed using the optimal number of components selected
[b] Optimal number of components selected from LOO/CV model
[c] Standard error of estimate
[d] Statistical measure of whether the model is significant
[e] Percent steric field contribution to full model
[f] Percent electrostatic field contribution to full model
[g] Equals ((1-(sSDEP)^2). Predictivity of the model using the scaled standard deviation of error of prediction
[h] Cross-validated standard deviation of error of prediction
[i] Slope of q^2

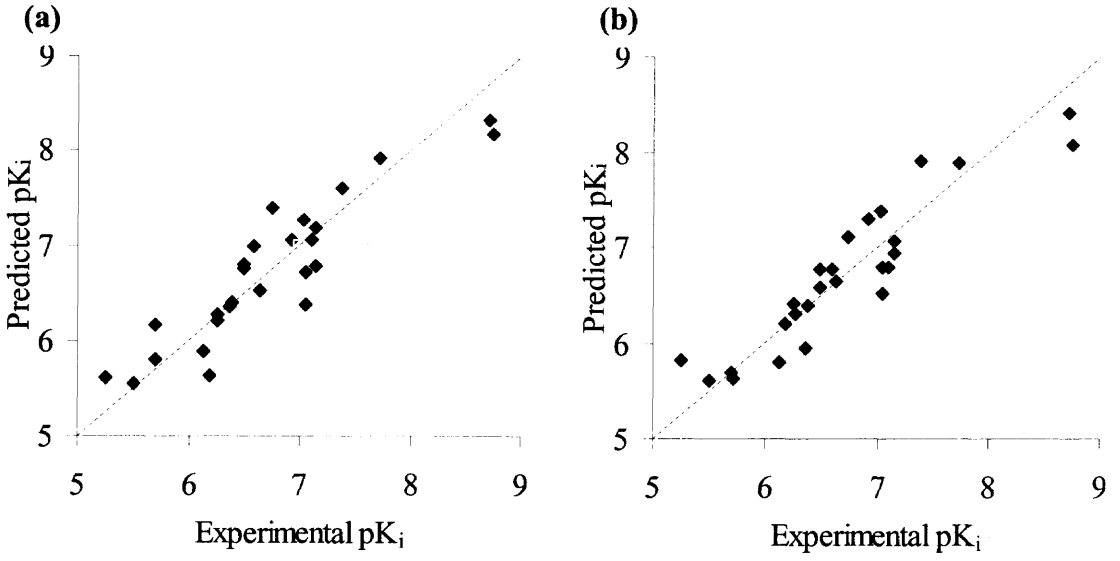


Figure 4.2 Training set predictions of the QSAR study of salvinorin A analogs. (a) Model based on StDev*Coefficient as weights and exponential factor of 0.5, (b) Model based on Discriminant Power as weights and exponential factor of 1.

4.3 External Validation: Prediction of Activity of Test Set Analogs

Since both models are stable, both were used to predict the pK_is of the compounds in the test set. For the focused NCV StDev*Coefficient and Discriminant Power models, the average absolute residuals were 0.53 and 0.54, respectively; the minimum residuals were -0.91 and -1.14, respectively; and the maximum residuals were 0.61 and 0.22, respectively. The test set predictions and residual values (predicted pK_i - experimental pK_i) obtained from both models are listed in Appendix C. The test set predictions were used to validate the models as recommended by Tropsha et al.[61,69] The recommended key characteristics of the test set useful to identify a predictive model are listed in Table 4.4 and plots are shown in Figure 4.3. From the values in the table and plots it is clear that only the focused NCV Discriminant Power model (exponential factor 1) is a stable

and predictive model. It satisfies all the characteristics of a predictive model listed in the CoMFA background section of **2.3.2**. The most notable characteristic is the R^2 value of 0.665, which is greater than 0.6 and is closer to the $R'_0{}^2$ value of 0.650. This value is close to the cross-validated R^2 (q^2) value of 0.621 obtained from the respective LOO/CV model. This model was used to create the CoMFA contour map and to predict the activities of novel compounds as explained in the next section.

Table 4.4 Test Set Correlation Validation

Weights	$R^{2\,a}$	$R_0{}^{2\,b}$	k^{c}	$R'_0{}^{2\,d}$	$k'^{\,e}$	Max of $(R_0{}^2$ or $R'_0{}^2)^{f}$
StDev*Coefficient (weight factor: 0.5)	0.562	-0.109	0.972	0.550	1.029	0.549
Discriminant Power (weight factor: 1)	0.665	0.2673	0.930	0.650	1.071	0.650

[a] Correlation coefficient between the predicted pK_i and experimental pK_i
[c] Coefficient of determination for the zero-intercept line of predicted pK_i vs. experimental pK_i
[c] Slope of the zero-intercept line of predicted vs. experimental pK_i
[d] Coefficient of determination for the zero-intercept line of predicted pK_i vs. experimental pK_i
[e] Coefficient of determination for the zero-intercept line of experimental pK_i vs. predicted pK_i
[f] Maximum of the coefficients of determination for the zero-intercept lines

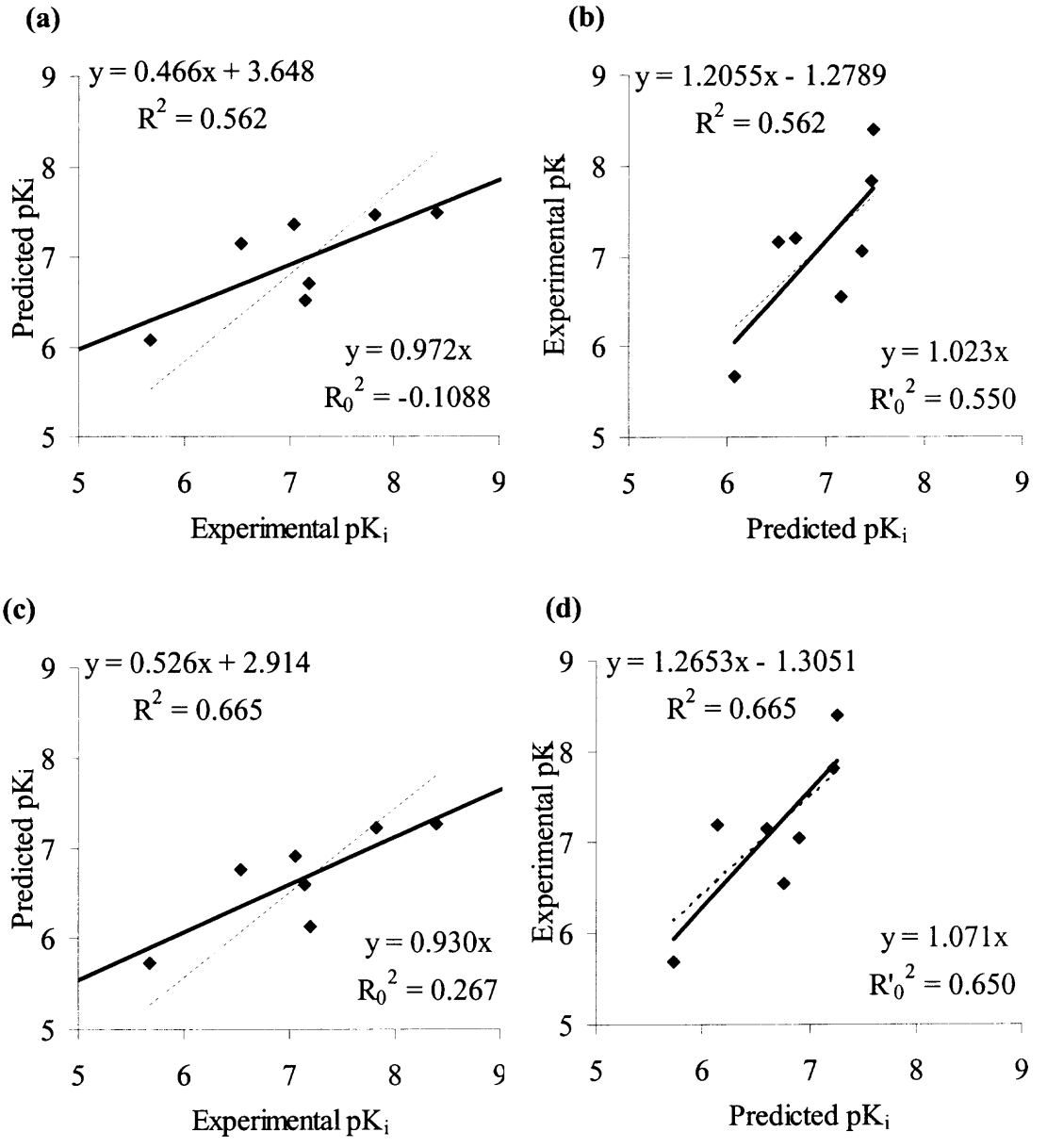


Figure 4.3 Test set predictions of the QSAR Study on salvinorin A analogs. Best-fit line with R^2 and zero-intercept line with R_0^2 for predicted pK_i vs. experimental pK_i and $R'_0{}^2$ for experimental pK_i vs. predicted pK_i . (a) and (b) Model based StDev*Coefficient as weights and exponential factor of 0.5, (c) and (d) Model based on Discriminant Power as weights and exponential factor of 1.

4.4 Novel Compounds

The CoMFA steric/electrostatic contour map for the focused NCV Discriminant Power model is shown in Figure 4.4a with the parent salvinorin A molecule as a reference. In general, the map indicates that increasing the steric bulk near the green areas and decreasing the steric bulk near the yellow areas would lead to a higher pK_i, as would increasing the positive charge near blue areas and increasing the negative charge near the red areas. Specific areas of interest are the green region near the carbonyl carbon, where an increase in bulk would improve pK_i, and the red region near the methyl group where an increase in the negative charge would lead to a higher pK_i, as shown schematically in Figure 4.4b. The other area of interest is the blue region where an increase in positive charge may improve pK_i. The contour map also indicates that steric and electrostatic fields both play an important role in activity.

Based on the above observations, the QSAR Optimization tool of SYBYL was used to find the optimal substitution pattern at X_1, X_2, and X_3 (Figure 4.4b) and to get an idea of the range of activities expected by changing substituents. The complete list of substituents used to make X_1, X_2, and X_3 combinations along with the QSAR Optimization settings is included in Appendix D. Not all combinations gave improved pK_i values. Only the twenty compounds with pK_i greater than 8 are shown in Table 4.5. The complete results are given in Appendix E. These were suggested to Dr. Thomas Prisinzano of the University of Iowa for synthesis. Out of these twenty analogs, preliminary biological results for **2**, **3** and **8** are available from the Rothman laboratory at the National Institute on Drug Abuse. All three analogs appear to be KOR selective. Preliminary biological tests were conducted on human μ, δ and κ opioid receptors

expressed in Chinese Hamster Ovary Cells (CHO cells) using [^{125}I]IOXY as a radioligand.[71] Currently binding studies are underway in the Rothman laboratory and we are awaiting the binding affinity data.

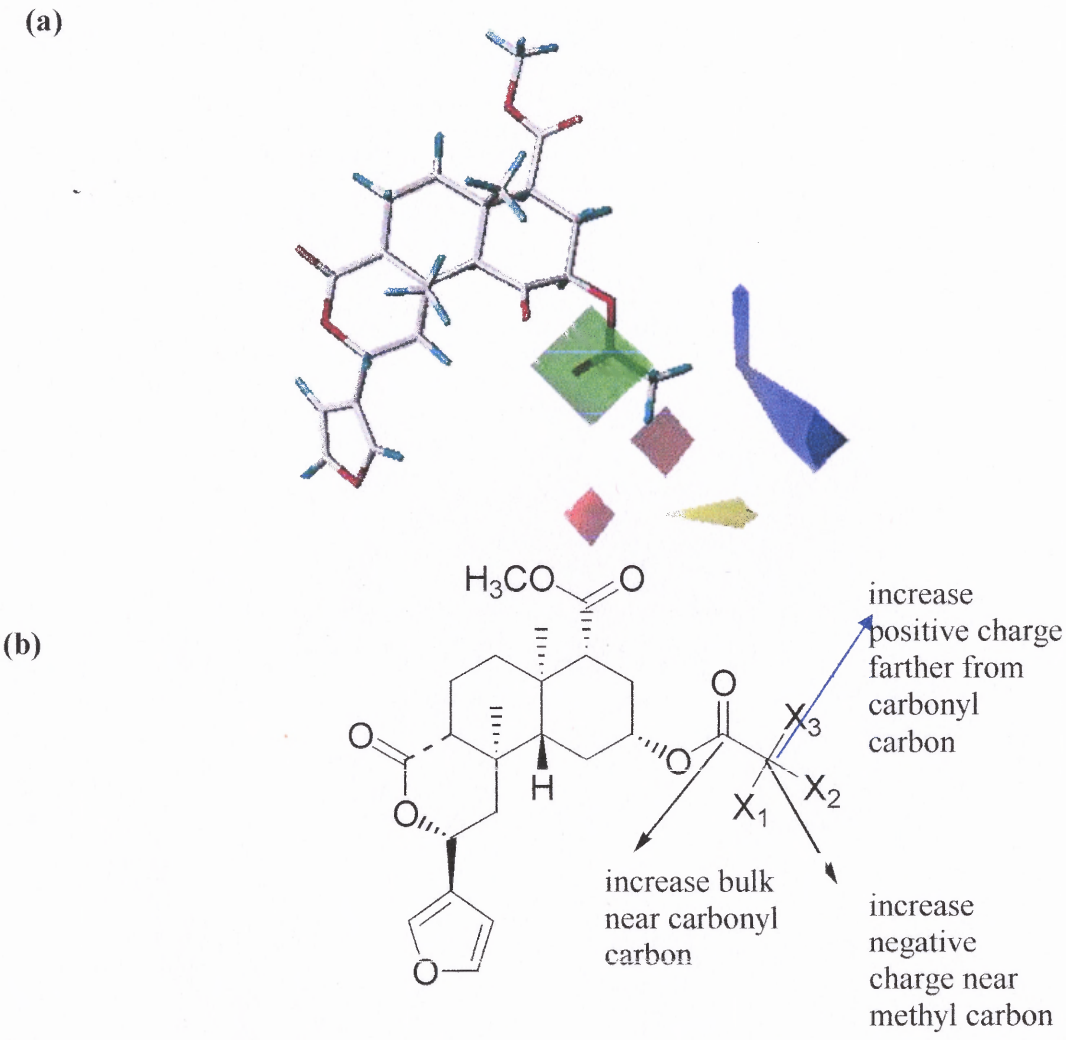
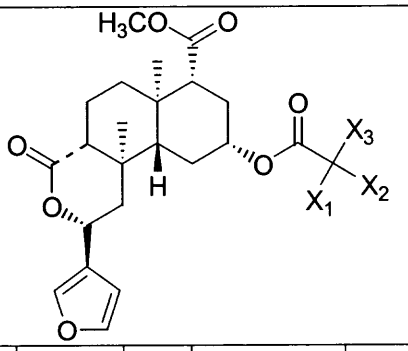


Figure 4.4 (a) Results of CoMFA steric/electrostatic contour maps with salvinorin A as reference, (b) Prescription for changes in regions of interest, shown schematically.

Table 4.5 Novel Compounds Based on Salvinorin A



Analog	X_1	X_2	X_3	Predicted pK_i
1. Salvinorin A	H	H	H	8.42 (Exp = 8.72)
2.	F	H	H	8.49
3.	F	H	F	8.57
4.	F	H	Cl	8.64
5.	F	H	Br	8.56
6.	F	H	CN	8.44
7.	F	H	CH_3	8.19
8.	F	H	CF_3	8.25
9.	F	H	OCH_3	8.27
10.	F	H	NH_2	8.14
11.	F	H	COOH	8.34
12.	F	H	OH	8.28
13.	F	H	CH_2Cl	8.11
14.	F	H	$NHCH_3$	8.19
15.	OH	H	H	8.46
16.	NH_2	H	H	8.45
17.	CF_3	H	H	8.12
18.	Cl	H	H	8.41
19.	Br	H	H	8.37
20.	CN	H	H	8.32

4.5 Discussion

The 3D-QSAR study performed on 2-position salvinorin A analogs highlighted important characteristics required for the better binding affinity (expressed in terms of pK_i) at the KOR. The study showed that both steric (47%) and electrostatic properties (53%) contribute almost equally to the model and hence to the binding affinity. The study showed the importance of trying a range of steric and electrostatic cutoff values in order to optimize the preliminary models.

The 3D-QSAR study performed here further reinforced the fact that both y-value scrambling and test set correlation validation should be applied to evaluate the predictivity of the model. This should be done in addition to the criteria of $q^2 > 0.5$.[61,69,72] The most important part of QSAR model development is the model validation. Performing just leave-one-out (or leave-some-out) cross-validation procedure is not sufficient to evaluate the predictivity of the model. The condition that a test set should have at least five compounds may be problematic for a small set of compounds (for example the present data set with less than 35 compounds), nevertheless an external test set should be used to gauge the predictivity of the model. Using recommended validation measures allowed the development of a stable and predictive model which was used to predict the binding affinity of the novel analogs. Initial biological results show promise but a definitive picture will appear only after the binding affinity data of novel analogs becomes available.

The observation that adding the H-bond CoMFA field class gave models with decreased values of q^2 compared to those obtained with Tripos Standard CoMFA field class calculations is an interesting point to note. It suggests that for the 2-position

salvinorin A analogs, hydrogen bonding may not play a role in explaining their observed biological activity. This possibility was also suggested based on mutagenesis studies of Tyr313 residue of KOR which is proposed to interact with 2-position of salvinorin A. Mutating to Tyr313 to Phe resulted in no loss of binding affinity,[29] while mutating it to Ala gave a dramatic decrease in affinity. However, the final CoMFA model suggests the role of electrostatic properties in addition to steric properties. There is a possibility that electrostatic interactions other than hydrogen bonding might play a role in binding of 2-position of salvinorin A analogs at the KOR.

It is known that salvinorin A is metabolized to an inactive form of salvinorin B which has a hydroxyl group at position 2.[73,74] There is also a difference in the observed in-vivo and in-vitro activity of salvinorin A probably due to the same reason. Because of the presence of the electronegative fluorine (a good leaving group),[75] it remains to be seen whether the fluorinated analogs of Table 4.5 are metabolized at the same rate or more rapidly than the parent salvinorin A. It will be interesting to compare the results of **2** with **9**, **10, 15** and **16** (Table 4.5) since CH_3O, NH_2 and OH are poorer leaving groups than F.

4.6 Conclusion

The 3D-QSAR method of CoMFA, based on ligands (in present study: salvinorin A analogs) can be applied in absence of a reliable structure of the receptor (in present study: κOR) to obtain a predictive and stable model using internal (LOO/CV and y-value scrambling) and external validation (test set correlation validation) tools. The ligand-based approach applied here is specifically valuable because proposed binding models of salvinorin A with the κOR based on homology modeling of proteins with mutagenesis data have not been able to provide a clear picture of the molecular properties playing role in the binding.

Part I - APPENDIX A
Scatter Plots

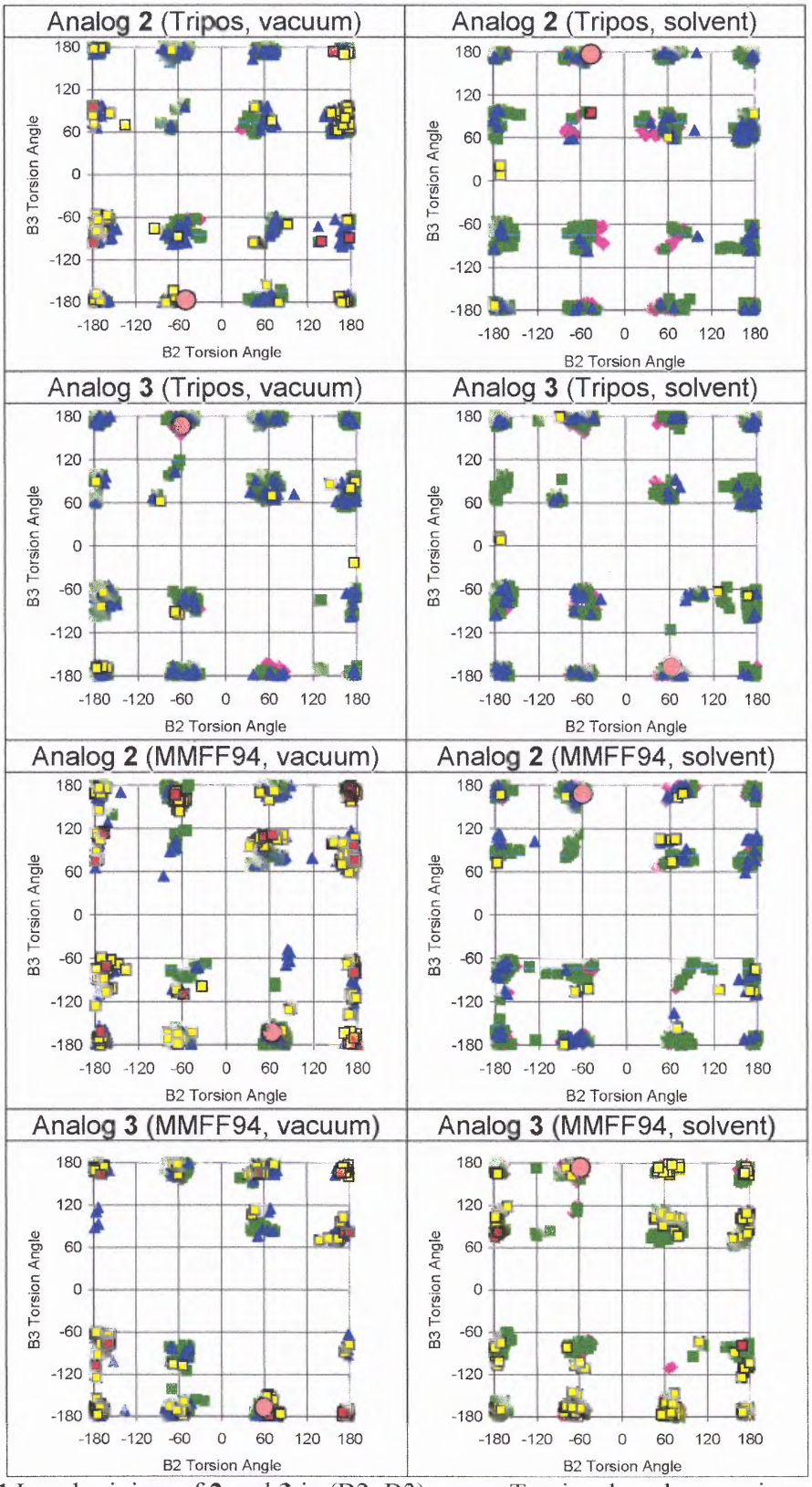


Figure A.1 Local minima of **2** and **3** in (B2, B3) space. Torsional angles are given in degrees. Minima are color coded by energy in units of kcal/mol. GEM marked by circle. Plot symbols: pink diamond: 0 – 4, green box: 4 – 8, blue triangle: 8 – 12, yellow box: 12 – 16, red box: 16 – 20.

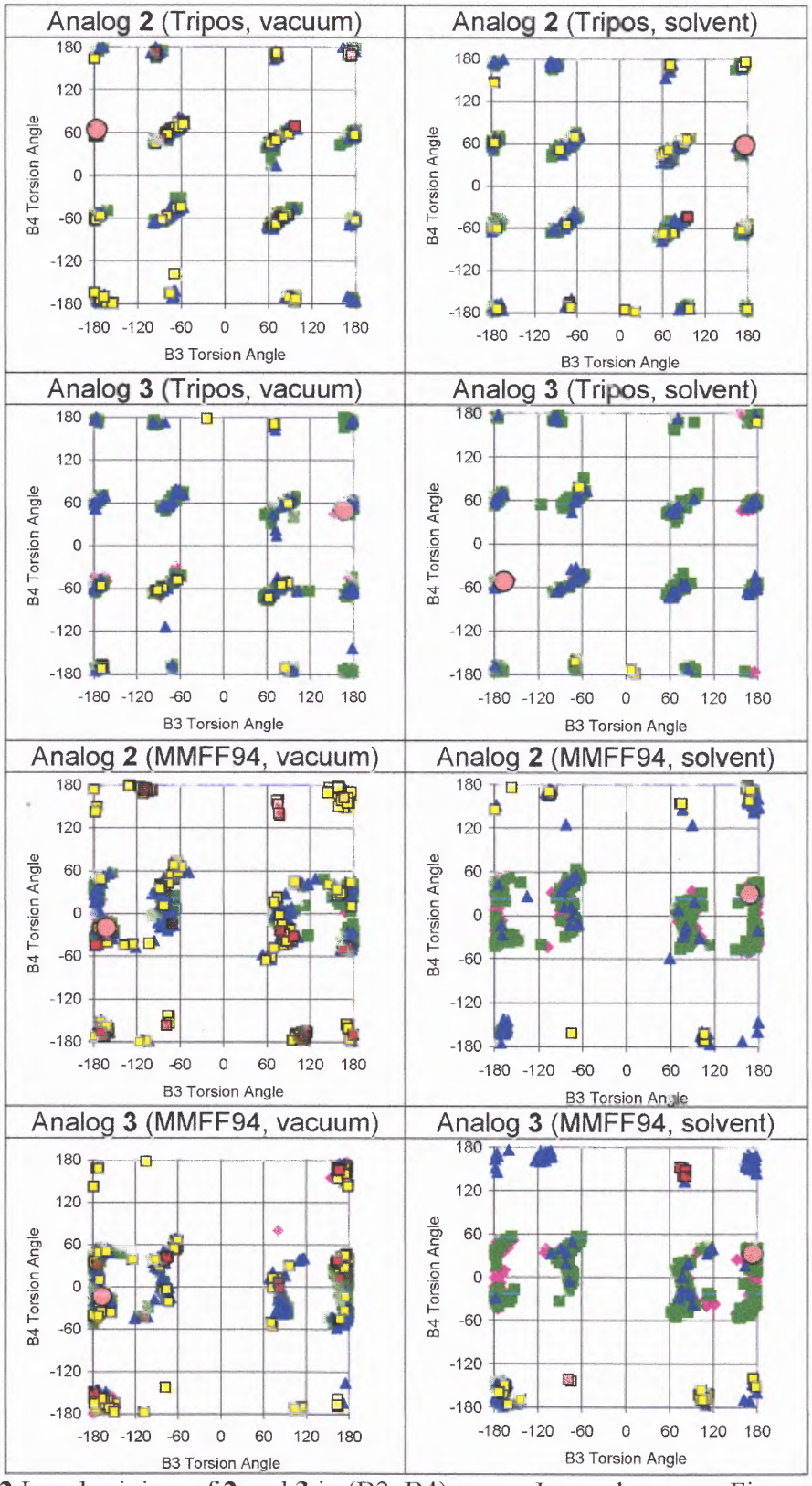


Figure A.2 Local minima of **2** and **3** in (B3, B4) space. Legend same as Figure A1.

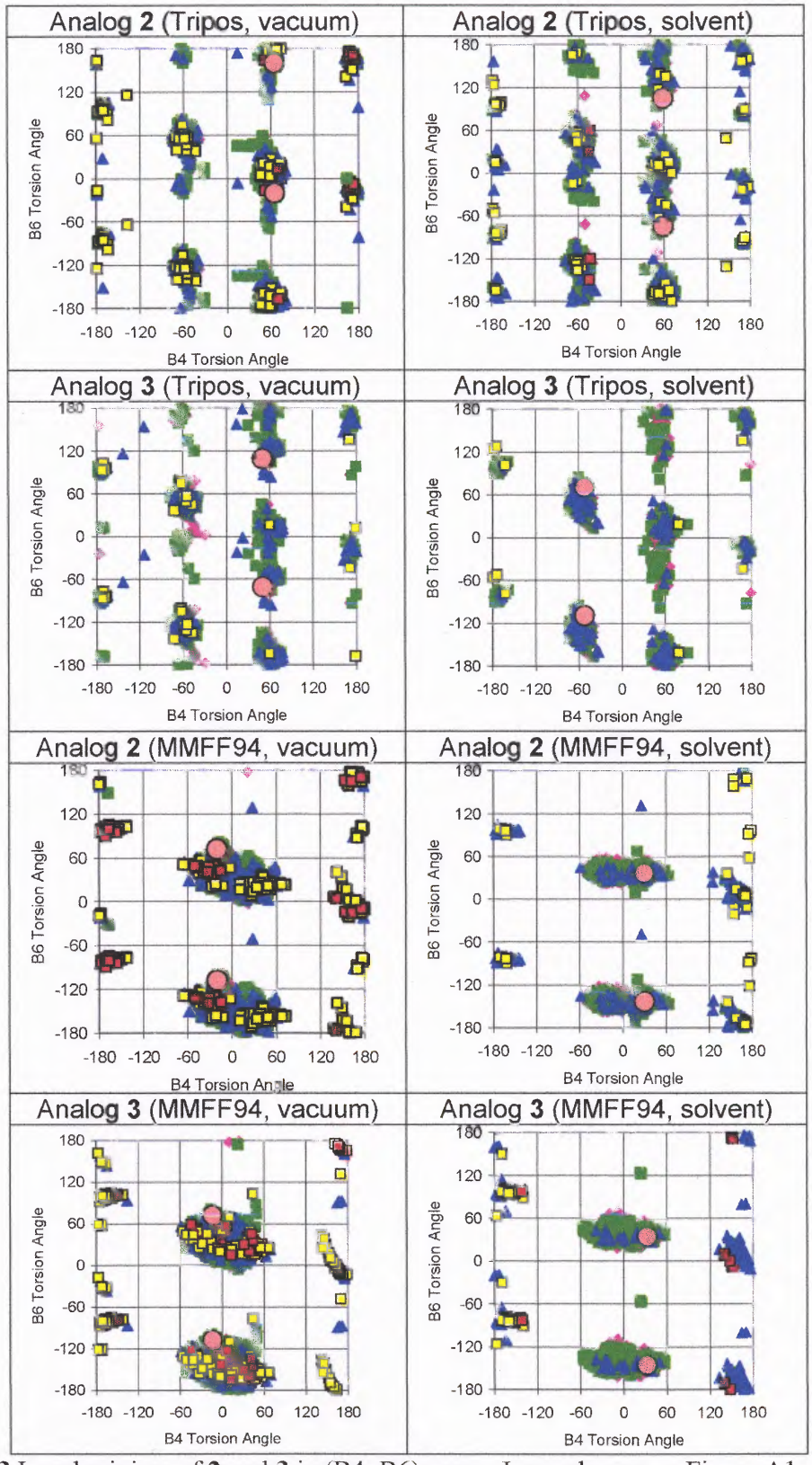


Figure A3 Local minima of **2** and **3** in (B4, B6) space. Legend same as Figure A1. The GEM structure appears twice due to the symmetry of the molecule to rotation of the phenyl ring around the B6 torsional angle.

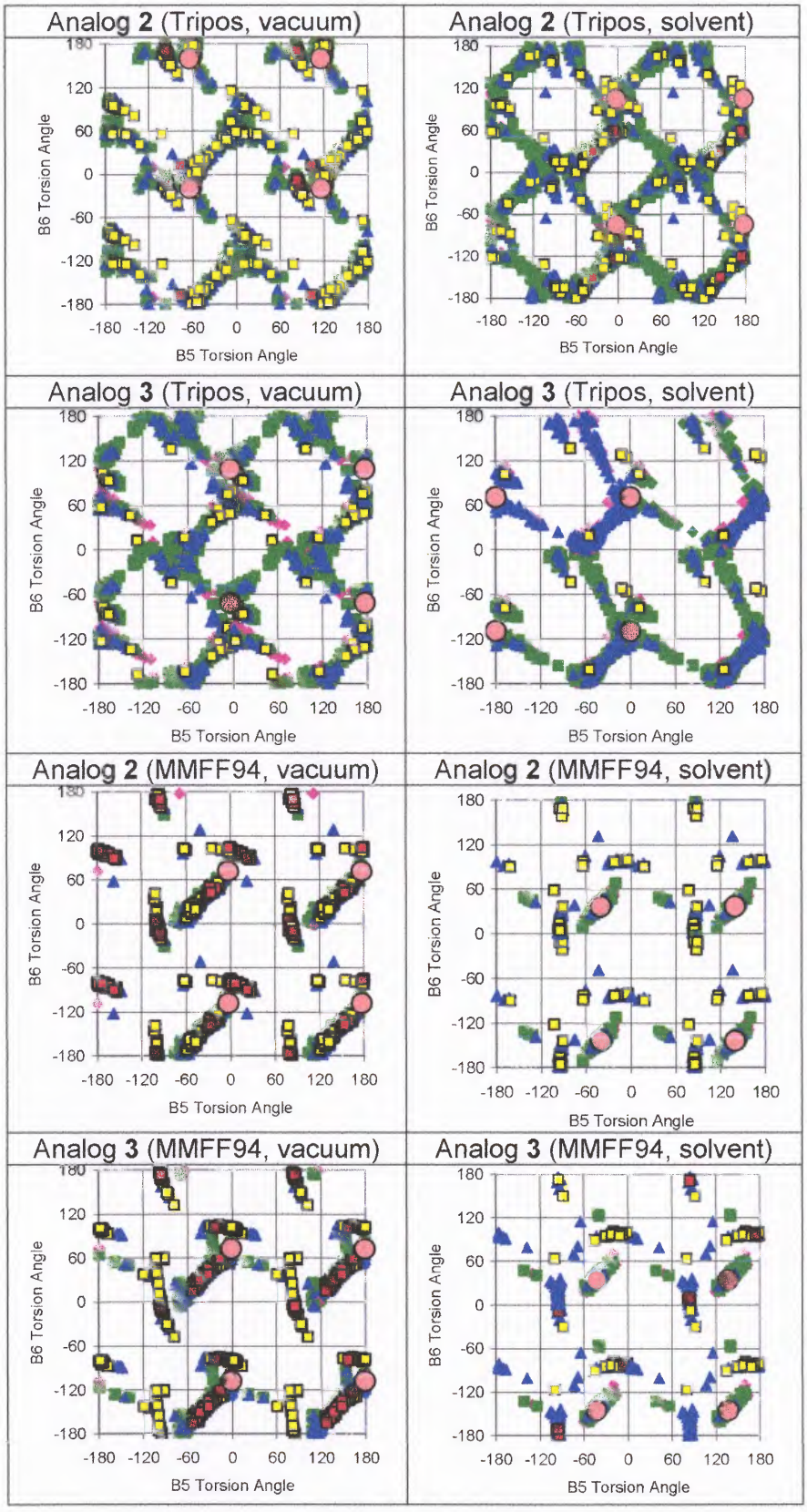


Figure A.4 Local minima of **2** and **3** in (B5, B6) space. Legend same as Figure A1. The GEM structure appears four times due to the symmetry of the molecule to rotation of the phenyl ring around the B5 and B6 torsional angles.

Part II - APPENDIX A

COMFA INPUT PARAMETERS

SYBYL CoMFA Parameters

All settings were kept as default except for changes in steric and electrostatic cutoffs.

CoMFA Field Class: Tripos Standard

Field Values: Type(s): Both

Dielectric: Distance

Smoothing: None

Drop Electrostatics: Within Steric Cutoff for Each Row

Transition: Smooth

Region: Create Automatically

SYBYL PLS Parameters

LOO/CV Model

Column to Use: CoMFA Field, pK_i
Dependent Column: pK_i

Validation: "Leave-One-Out" box checked
"Use SAMPLS" box checked
Components: 6

Scaling: CoMFA Standard

NCV Full Model

Column to Use: CoMFA Field, pK_i
Dependent Column: pK_i

Validation: "No Validation" box checked

Components: Selected based on LOO/CV

Scaling: CoMFA Standard

PART II - APPENDIX B

PRELIMINARY COMFA MODELS - COMBINATIONS OF STERIC AND ELECTROSTATIC CUTOFFS

Table B.1 Preliminary LOO/CV Results of the QSAR Study on Salvinorin A Analogs

LOO/CV PLS			
Field Cutoffs	q^2	C[a]	SEP[b]
Steric: 10 kcal/mol Electrostatic: 1 kcal/mol	0.387	2	0.682
Steric: 10 kcal/mol Electrostatic: 10 kcal/mol	0.225	3	0.784
Steric: 10 kcal/mol Electrostatic: 30 kcal/mol	0.190	3	0.802
Steric: 30 kcal/mol Electrostatic: 1 kcal/mol	0.408	2	0.670
Steric: 30 kcal/mol Electrostatic: 10 kcal/mol	0.270	3	0.761
Steric: 30 kcal/mol Electrostatic: 30 kcal/mol	0.255	3	0.769
Steric: 60 kcal/mol Electrostatic: 1 kcal/mol	0.384	2	0.684
Steric: 60 kcal/mol Electrostatic: 10 kcal/mol	0.271	3	0.761
Steric: 60 kcal/mol Electrostatic: 30 kcal/mol	0.247	3	0.773

[a] Optimal number of components selected from LOO/CV model
[b] Standard error of prediction

PART II - APPENDIX C

RESIDUALS

Table C.1 Training Set and Test Set (shown with asterisks) Predictions

		Model 1 Weights = StDev*Coefficient Exponential factor = 0.5		**Model 2** Weights = Discriminant Power Exponential factor = 1	
Analog	**Experimental Pki**	**Predicted pK_i**	**Residuals**	**Predicted pK_i**	**Residuals**
1	8.72	8.32	-0.40	8.42	-0.30
2	8.74	8.17	-0.57	8.07	-0.67
3	7.72	7.91	0.19	7.88	0.16
4*	8.40	7.49	-0.91	7.26	-1.14
5*	7.82	7.46	-0.37	7.22	-0.60
6	7.15	7.19	0.04	6.94	-0.21
7	6.37	6.35	-0.02	5.95	-0.42
8	7.05	6.38	-0.67	6.52	-0.53
9*	7.05	6.71	-0.34	6.79	-0.26
10	7.15	6.52	-0.63	6.59	-0.56
11	6.13	5.90	-0.23	5.79	-0.34
12	5.71	5.80	0.09	5.64	-0.07
13*	6.54	7.15	0.61	6.76	0.22
14	6.74	7.38	0.64	7.10	0.36
15	6.59	6.99	0.40	6.77	0.18
16	7.03	7.27	0.24	7.38	0.35
17	6.92	7.05	0.13	7.30	0.38
18*	7.05	7.37	0.32	6.90	-0.15
19*	7.19	6.70	-0.49	6.14	-1.05
20	7.38	7.60	0.22	7.90	0.52
21	6.64	6.53	-0.11	6.64	0.00
22	6.26	6.27	0.01	6.41	0.15
23	6.27	6.20	-0.07	6.30	0.03
24	6.49	6.76	0.27	6.77	0.28
25	6.49	6.80	0.31	6.58	0.09
26	6.39	6.39	0.00	6.40	0.01
27	5.26	5.61	0.35	5.81	0.55
28	7.15	6.77	-0.38	7.06	-0.09
29	7.10	7.06	-0.04	6.79	-0.31
30*	5.70	6.16	0.46	5.70	0.00
31	5.68	6.08	0.40	5.73	0.05
32	6.19	5.63	-0.56	6.21	0.02
33	5.50	5.54	0.04	5.62	0.12

PART II - APPENDIX D

QSAR OPTIMIZATION

Table D.1 Substituents Used for Various X_1, X_2, and X_3 Combinations (Figure 4.3b) for QSAR Optimization

-H (Default)	-OH	- CF_3	-NO_2
-F	-OCH_3	-CF_2H	-$CONH_2$
-Cl	-CN	- CH_2F	-$CONHCH_3$
-Br	-NH_2	-CH_2Cl	
-CH_3	-$NHCH_3$	-COOH	

D.2 QSAR Optimization Settings

Source Table <name of the molecular spreadsheet>

Configuration Option

Conformation Refinement: Relax
Computation of Charges: GAST_HUCK

Analysis <name of the PLS analysis> **Optimize**: PK_i

Core Molecule <M1> New Area<M2> Work Area<M3>

Method = Linear
Best = 250

Run Name <name of the run>

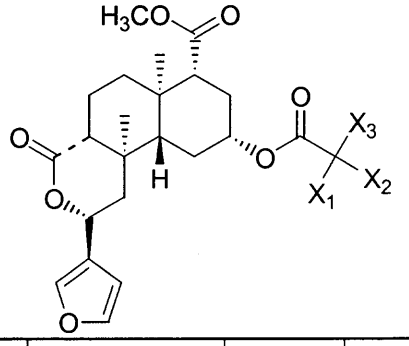
Default Substituent Table <name of the table with substituents>

Assign R Groups <assign appropriate R groups from the substituents table>

PART II - APPENDIX E

PREDICTION OF NOVEL ANALOGS

Table E.1 Prediction of novel analogs using Optimized QSAR based on Focused Model with Weights = Discriminant Power, Exponential factor = 1



Novel Analogs	Predicted pK_i	X_1	X_2	X_3
1	8.64	F	H	Cl
2	8.57	F	H	F
3	8.56	F	H	Br
4	8.5	F	H	H
5	8.5	F	H	H
6	8.49	F	H	H
7	8.46	OH	H	H
8	8.46	OH	H	H
9	8.45	NH_2	H	H
10	8.44	F	H	CN
11	8.44	NH_2	H	H
12	8.42	H	H	H
13	8.42	H	H	H
14	8.42	H	H	H
15	8.42	H	H	H
16	8.42	H	H	H
17	8.41	Cl	H	H
18	8.41	Cl	H	H
19	8.38	H	H	Br
20	8.37	Br	H	H
21	8.37	Br	H	H
22	8.34	F	H	CO_2H
23	8.32	H	H	Cl
24	8.32	CN	H	H
25	8.32	CN	H	H
26	8.28	F	H	OH
27	8.27	F	H	OMe
28	8.25	F	H	CF_3
29	8.25	Methyl	H	H
30	8.24	Methyl	H	H
31	8.2	F	Methyl	H
32	8.19	F	H	Methyl
33	8.19	F	H	NHMe

34	8.16	H	H	Methyl
35	8.15	F	H	NO_2
36	8.14	F	H	NH_2
37	8.12	CF_3	H	H
38	8.12	H	Methyl	H
39	8.12	CF_3	H	H
40	8.11	F	H	CH_2Cl
41	8.08	H	NH_2	H
42	8.06	H	H	F
43	8.02	H	H	CH_2Cl
44	7.98	H	H	OH
45	7.96	H	H	CF3
46	7.94	H	H	OMe
47	7.92	H	H	CH2F
48	7.89	F	Br	H
49	7.88	F	OH	H
50	7.87	H	H	NHMe
51	7.85	F	CO_2H	H
52	7.85	H	H	CN
53	7.84	F	H	CH_2F
54	7.84	H	H	NH_2
55	7.83	CF_2H	H	H
56	7.81	H	H	$CONH_2$
57	7.81	H	H	CONHMe
58	7.79	F	H	CF2H
59	7.79	F	Cl	H
60	7.76	CF_2H	H	H
61	7.63	H	H	NO_2
62	7.62	F	CH_2Cl	H
63	7.6	H	CO2H	H
64	7.59	H	OH	H
65	7.58	CO_2H	H	H
66	7.57	F	NH2	H
67	7.56	CO_2H	H	H
68	7.55	H	Br	H
69	7.48	F	NHMe	H
70	7.47	H	CH2Cl	H
71	7.43	H	Cl	H
72	7.41	F	CF2H	H
73	7.41	$CONH_2$	H	H
74	7.4	H	H	CF_2H
75	7.4	F	CH_2F	H
76	7.35	CH_2F	H	H
77	7.35	F	F	H
78	7.35	CH_2F	H	H
79	7.32	H	CH_2F	H
80	7.29	H	$CONH_2$	H
81	7.29	H	NHMe	H
82	7.28	H	CF2H	H
83	7.24	F	H	CONHMe
84	7.22	H	H	CO_2H

85	7.21	OMe	H	H
86	7.2	F	H	CONH2
87	7.18	F	OMe	H
88	7.11	H	CONHMe	H
89	7.09	F	CF3	H
90	7.04	CONHMe	H	H
91	7.02	NHMe	H	H
92	7.02	OMe	H	H
93	7.02	F	CN	H
94	6.97	H	F	H
95	6.86	NHMe	H	H
96	6.84	H	OMe	H
97	6.66	F	CONHMe	H
98	6.63	H	CN	H
99	6.59	H	CF3	H
100	6.5	CH2Cl	H	H
101	6.48	CH2Cl	H	H
102	6.44	CONH2	H	H
103	6.44	CONHMe	H	H
104	6.33	F	NO2	H
105	6.3	F	CONH2	H
106	6.24	NO2	H	H
107	6.02	H	NO2	H

REFERENCES

Part I

1. Kreek, M. J.; LaForge, K. S.; Butelman, E., Pharmacotherapy of addictions. *Nature Reviews Drug Discovery* 2002, 1, 710-726.

2. Leshner, A. I., Addiction is a brain disease, and it matters. *Science* 1997, 278, 45-47.

3. Dutta, A. K.; Zhang, S.; Kolhatkar, R. B.; Reith, M. E. A., Dopamine transporter as target for drug development of cocaine dependence medication. *European Journal of Pharmacology* 2003, 479, 93-106.

4. Kuhar, M. J.; Ritz, M. C.; Boja, J. W., The dopamine hypothesis of the reinforcing properties of cocaine. *Trends in Neuroscience* 1991, 14, 299-302.

5. Rocha, B. A., Stimulant and reinforcing effects of cocaine in monoamine transporter knockout mice. *European Journal of Pharmacology* 2003, 479, 107-115.

6. Sora, I.; Scott Hall, F.; Andrews, A. M.; Itokawa, M.; Li, X.-F.; Wei, H.-B.; Wichems, C.; Lesch, K.-P.; Murphy, D. L.; Uhl, G. R., Molecular mechanisms of cocaine reward. Combined dopamine and serotonin transporter Knockouts eliminate cocaine place preference. *Proceedings of the National Academy of Sciences of the United States of America* 2001, 98, 5300-5305.

7. Sora, I.; Wichems, C.; Takahashi, N.; Li, X.-F.; Zeng, Z.; Revay, R.; Lesch, K. P.; Murphy, D. L.; Uhl, G. R., Cocaine reward models: Conditioned place preference can be established in dopamine- and in serotonin-transporter knockout mice. *Proceedings of the National Academy of Sciences of the United States of America* 1998, 95, 7699-7704.

8. Nestler, E. J., Molecular basis of long-term plasticity underlying addiction. *Nature Reviews Neuroscience* 2001, 2, 119-128.

9. Chen, N. H.; Reith, M. E. A., Structure-function relationships for biogenic amine neurotransmitter transporters. In *Contemporary neuroscience: Neurotramsmitter transporters: Structure, function, and regulation*, Reith, M. E. A., Ed. Humana Press: Totowa, NJ, 1997; pp. 53-109.

10. Carroll, F. I.; Lewin, A. H.; Mascarella, S. W., Dopamine-Transporter Uptake Blockers. In *Contemporary Neuroscience: Neurotransmitter Transporters: Structure, Function, and Regulation*, Reith, M. E. A., Ed. Humana Press: Totowa, NJ, 1997; pp. 381-432.

11. Singh, S., Chemistry, design, and structure-activity relationship of cocaine antagonists. *Chemical Review* 2000, 100, 925-1024.

12. Carroll, F. I., 2002 Medicinal chemistry division award address: Monoamine transporters and opioid receptors. Targets for addiction therapy. *Journal of Medicinal Chemistry* 2003, 46, 1775-1794.

13. Prisinzano, T.; Rice, K. C.; Baumann, M. H.; Rothman, R. B., Development of neurochemical normalization ("agonist substitution") therapeutics for stimulant Abuse: Focus on the dopamine uptake inhibitor, GBR 12909. *Current Medicinal Chemistry - Central Nervous System Agents* 2004, 4, 47-59.

14. Chen, N.-H.; Reith, M.E.A.; Quick, M.W., Synaptic uptake and beyond: The sodium- and chloride-dependent neurotransmitter transporter family SLC6. *Pflugers Archiv European Journal of Physiology* 2004, 447, 519-531.

15. Giros, B.; El Mestikawy, S.; Godinot, N.; Zheng, K.; Han, H.; Yang-Feng, T.; Caron, M. G., Cloning, pharmacological characterization and chromosome assignment of the human dopamine transporter. *Molecular Pharmacology* 1992, 42, 383-390.

16. Vandenbergh, D. J.; Persico, A. M.; Hawkins, A. L.; Griffin, C. A.; Li, X.; Jabs, E. W.; Uhl, G. R., Human dopamine transporter gene (DAT1) maps to chromosome 5p15.3 and displays a VNTR. *Genomics* 1992, 14, 1104-1106.

17. Chen, N.; Reith, M. E. A., Structure and function of the dopamine transporter. *European Journal of Pharmacology* 2000, 405, 329-339.

18. Kitayama, S.; Shimada, S.; Xu, H.; Markham, L.; Donnovan, D. M.; Uhl, G. R., Dopamine transporter site-directed mutations differentially alter substrate transport and cocaine binding. *Proceedings of the National Academy of Sciences of the United States of America* 1992, 89, 7782-7785.

19. Loland, C. J.; Norregaard, L.; Litman, T.; Gether, U., Generation of an activating Zn^{2+} switch in the dopamine transporter: Mutation of an intracellular tyrosine constitutively alters the conformational equilibrium of the transport cycle. *Proceedings of the National Academy of Sciences of the United States of America* 2002, 99, 1683-1688.

20. Chen, N.; Vaughan, R. A.; Reith, M. E. A., The role of conserved tryptophan and acidic residues in the human dopamine transporter as characterized by site-directed mutagenesis. *Journal of Neurochemistry* 2001, 77, 1116-1127.

21. Sorkina, T.; Hoover, B. R.; Zahniser, N. R.; Sorkin, A., Constitutive and protein kinase C-induced internalization of the dopamine transporter is mediated by a clathrin-dependent mechanism. *Traffic* 2005, 6, 157-170.

22. Foster, J. D.; Pananusorn, B.; Vaughan, R. A., Dopamine transporters are phosphorylated on N-terminal serines in rat striatum. *Journal of Biological Chemistry* 2002, 277, 25178-25186.

23. Zahniser, N. R.; Doolen, S., Chronic and acute regulation of Na^+/Cl^--dependent neurotransmitter transporters: Drugs, substrates, presynaptic receptors, and signaling systems. *Pharmacology and Therapeutics* 2001, 92, 21-55.

24. Hastrup, H. K., A.; Javitch, J. A., Symmetrical dimer of the human dopamine transporter revealed by cross-linking Cys-306 at the extracellular end of the sixth transmembrane segment. *Proceedings of the National Academy of Sciences of the United States of America* 2001, 98, 10055-10060.

25. Torres, G. E.; Carneiro, A.; Seamans, K.; Fiorentini, C.; Sweeney, A.; Yao, W.-D.; Caron, M. G., Oligomerization and trafficking of the human dopamine transporter: Mutational analysis identifies critical domains important for the functional expression of the transporter. *Journal of Biological Chemistry* 2003, 278, 2731-2739.

26. Andersen, P. H., The dopamine uptake inhibitor GBR 12909: selectivity and molecular mechanism of action. *European Journal of Pharmacology* 1989, 166, 493-504.

27. Rothman, R. B.; Mele, A.; Reid, A. A.; Akunne, H.; Greig, N.; Thurkauf, A.; Rice, K. C.; Pert, A., Tight binding dopamine reuptake inhibitors as cocaine antagonists: A strategy for drug development. *FEBS Letters* 1989, 257, 341-344.

28. Rothman, R. B.; Mele, A.; Reid, A. A.; Akunne, H.; Grieg, N.; Thurkauf, A.; de Costa, B. R.; Rice, K. C.; Pert, A., Preliminary evidence that the high affinity dopamine reuptake inhibitor, GBR12909, antagonizes the ability of cocaine to elevate extracellular levels of dopamine. *NIDA Research Monograph* 1991, 105, 359-360.

29. Rothman, R. B.; Mele, A.; Reid, A. A.; Akunne, H. C.; Greig, N.; Thurkauf, A.; De Costa, B. R.; Rice, K. C.; Pert, A., GBR 12909 antagonizes the ability of cocaine to elevate extracellular levels of dopamine. *Pharmacology, Biochemistry, and Behavior* 1991, 40, 387-397.

30. Nicklaus, M. C.; Wang, S.; Driscoll, J.; Milne, G. W. A., Conformational changes of small molecules binding to proteins. *Bioorganic and Medicinal Chemistry* 1995, 3, 411-428.

31. Veith, M.; Hirst, J. D.; Brooks, I., C.L., Do active site conformations of small ligands correspond to low free-energy solution structures? *Journal of Computer-Aided Molecular Design* 1998, 12, 563-572.

32. Boström, J.; Norrby, P.-O.; Liljefors T. Conformational energy penalties of protein-bound ligands. *Journal of Computer-Aided Molecular Design* 1998, 12, 383-396.

33. Debnath, A. K., Three-dimensional quantitative structure-activity relationship study on cyclic urea derivatives as HIV-1 protease inhibitors: Application of Comparative Molecular Field Analysis. *Journal of Medicinal Chemistry* 1999, 42, 249-259.

34. Perola, E.; Charifson, P. S., Conformational analysis of drug-like molecules bound to proteins: An extensive study of ligand reorganization upon binding. *Journal of Medicinal Chemistry* 2004, 47, 2499-2510.

35. Gilbert, K. M.; Boos, T. L.; Dersch, C. M.; Greiner, E.; Jacobson, A. E.; Lewis, D.; Matecka, D.; Prisinzano, T. E.; Zhang, Y.; Rothman, R. B.; Rice, K. C.; Venanzi, C. A., DAT/SERT selectivity of flexible GBR 12909 analogs modeled using 3D-QSAR methods. *Bioorganic and Medicinal Chemistry* 2007, 15, 1146-1159.

36. Beusen, D. D.; Shands, E. F. B.; Karasek, S.F.; Marshall, G. R.; Dammkoehler, R. A., Systematic search in conformational analysis. *Journal of Molecular Structure: THEOCHEM* 1996, 370, 157-171.

37. Saunders, M., Stochastic exploration of molecular mechanics energy surfaces. Hunting for the global minimum. *Journal of the American Chemical Society* 1987, 109, 3150-3152.

38. Clark, M.; Cramer III, R. D.; Van Opdenbosch, N., Validation of the General Purpose TRIPOS 5.2 Force Field. *Journal of Computational Chemistry* 1989, 10, 982-1012.

39. Halgren, T. A., Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry* 1996, 17, 490-519.

40. Halgren, T. A., Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *Journal of Computational Chemistry* 1996, 17, 520-552.

41. Halgren, T. A., Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. *Journal of Computational Chemistry* 1996, 17, 553-586.

42. Halgren, T. A., Merck molecular force field. IV. Conformational energies and geometries for MMFF94. *Journal of Computational Chemistry* 1996, 17, 587-615.

43. Halgren, T. A.; Nachbar, R., Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *Journal of Computational Chemistry* 1996, 17, 616-641.

44. Halgren, T. A., Merck molecular force field VI. MMFF94s option for energy minimization studies. *Journal of Computational Chemistry* 1999, 20, 720-729.

45. Boyd, D. B.; Coner, R. D., Stochastic approach to force field evaluations: Conformational analysis of raloxifene, a potential new therapeutic agent for post-menopausal osteoporosis. *Journal of Molecular Structure: THEOCHEM* 1996, 368, 7-15.

46. Berfield, J. L.; Wang, L. C.; Reith, M. E. A., Which form of dopamine is the substrate for the human dopamine transporter: The cationic or the uncharged species? *Journal of Biological Chemistry* 1999, 274, 4876-4882.

47. Xu, C.; Reith, M. E. A., Modeling the pH dependence of the binding of WIN 35,428 to the dopamine transporter in rat striatal membranes: Is the bioactive form positively charged or neutral? *Journal of Pharmacology and Experimental Therapeutics* 1996, 278, 1340-1348.

48. Gilbert, K. A.; Skawinski, W. J.; Misra, M.; Paris, K.; Naik, N.; Deutsch, H. M.; Venanzi, C. A., Conformational analysis of methylphenidate: Comparison of molecular orbital and molecular mechanics methods. *Journal of Computer-Aided Molecular Design* 2004, 18, 719-738.

49. Pandit, D.; Roosma, W.; Misra, M.; Gilbert, K. M.; Skawinski, W. J.; Matecka, D.; Prisinzano, T.; Rice, K. C.; Venanzi, C. A., Conformational analysis of piperazine and piperadine Analogs of 1-[2-[Bis (4-fluorophenyl) methoxy] ethyl]-4-(3-phenylpropyl) piperazine (GBR12909): Effect of force field and solvent. *Manuscript in Preparation.*

50. Dutta, A. K.; Meltzer, P. C.; Madras, B. K., Positional importance of the nitrogen atom in novel piperidine analogs of GBR 12909: Affinity and selectivity for the dopamine transporter. *Medicinal Chemistry Research* 1993, 3, 209-222.

51. Powell, M. J. D., Restart procedures for the conjugate gradient method. *Mathematical Programming* 1977, 12, 241-254.

52. Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. *Gaussian 03 C.02*, Gaussian 03; Gaussian, Inc., Wallingford CT.: 2004.

53. Halgren, T. A., Merck molecular force field VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. *Journal of Computational Chemistry* 1999, 20, 730-748.

54. Fiorentino, A.; Pandit, D.; Gilbert, K. A.; Misra, M.; Dios, R.; Venanzi, C. A., Singular value decomposition of torsional angles of analogs of the dopamine reuptake inhibitor GBR 12909. *Journal of Computational Chemistry* 2006, 27, 609-620.

55. Lowe, J.P., Barriers to internal rotation about single bonds. *Progress in Physical Organic Chemistry* 1968, 6, 1-80.

56. Gilbert, K. M.; Venanzi, C. A., Hierarchical clustering analysis of flexible GBR 12909 dialkyl piperazine and piperidine analogs. *Journal of Computer-Aided Molecular Design* 2006, 20, 209-225.

Part II

1. Law, P.-Y.; Wong, Y. H; Loh, H.H, Molecular mechanisms and regulation of opioid receptor signaling. Annual Review of Pharmacology and Toxicology 2000, 389 - 430.

2. Dhawan, B. N.; Cesselin, F.; Raghubir, R.; Reisine, T.; Bradley, P.B.; Portoghese, P.S.; Hamon, M., International Union of Pharmacology. XII. Classification of opioid receptors. Pharmacological Reviews 1996, 48, 567-592.

3. Waldhoer, M.; Bartlett, S. E.; Whistler, J. L., Opioid Receptors. Annual Reviews of Biochemistry 2004, 73, 953-90.

4. Chavkin, C.; Goldstein, A., Specific receptor for the opioid peptide dynorphin: Structure-activity relationships. Proceedings of the National Academy of Sciences of the United States of America 1981, 78, 6543-6547.

5. Schwyzer, R., ACTH: a short introductory review. Annals of the New York Academy of Sciences 1977, 297, 3-26.

6. Janecka, A.; Fichna, J.; Janecki, T., Opioid receptors and their ligands. Current Topics in Medicinal Chemistry 2004, 4, 1-17.

7. DeHaven-Hudkins, D. L.; Dolle, R. E., Peripherally restricted opioid agonists as novel analgesic agents. Current Pharmaceutical Design 2004, 10, 743-757.

8. Roth, B. L.; Baner, K.; Westkaemper, R.; Siebert, D.; Rice, K.C.; Steinberg, S.; Ernsberger, P.; Rothman, R.B., Salvinorin A: A potent naturally occurring nonnitrogenous κ opioid selective agonist. Proceedings of the National Academy of Sciences of the United States of America 2002, 99, 11934 - 11939.

9. Yan, F.; Roth, B.L., Salvinorin A: A novel and highly selective kappa-opioid receptor agonist. Life Sciences 2004, 75, 2615-2619.

10. Prisinzano, T. P.; Tidgewell, K.; Harding, W.W., Kappa opioids as potential treatments for stimulant dependence. AAPS Journal 2005, 7, E592-E599.

11. Margolis, E. B.; Hjelmsatd, G.O.; Bonci, A.; Fields, H.L., Kappa opioid agonists directly inhibit midbrain dopaminergic neurons. Journal of Neuroscience 2003, 23, 9981-9986.

12. Zhang, Y.; Butleman, E.R.; Sclussman, S.D.; Ho, A.; Kreek, M. J., Effect of the endogenous κ opioid agonist dynorphin A(1-17) on cocaine-evoked increases in striatal dopamine levels and cocaine-induced place preference in C57BL/6J mice. Psychopharmacology 2004, 172, 422-429.

13. Mello, N. K.; Negus, S.S., Interactions between kappa opioid agonists and cocaine. Preclinical studies. Annals of the New York Academy of Sciences 2000, 909, 104-132.

14. Neumeyer, J. L.; Gu, X.-H.; Van Vliet, L.A.; DeNunzio, N.J.; Rusovici, D.E.; Cohen, D.J.; Negus, S.S.; Mello, N.K.; Bidlack, J.M., Mixed κ agonists and μ agonists/antagonists as potential pharmacotherapeutics for cocaine abuse: Synthesis and opioid receptor binding affinity of N-substituted derivatives of morphinan. Bioorganic and Medicinal Chemistry Letters 2001, 11, 2735-2740.

15. Chavkin, C.; Sud, S.; Jin, W.; Stewart, J.; Zjawiony, J.K.; Siebert, D.J.; Toth, B.A.; Hufeisen, S.J.; Roth, B.L., Salvinorin A, an active component of the hallucinogenic sage Salvia divinorum Is a highly efficacious κ-opioid receptor agonist: structural and functional considerations. Journal of Pharmacology and Experimental Therapeutics 2004, 308, 1197-1203.

16. Butelman, E. R.; Harris, T.J.; Kreek, M.J., The plant-derived hallucinogen, salvinorin A, produces κ-opioid agonist-like discriminative effects in rhesus monkeys. Psychopharmacology 2004, 172, 220-224.

17. McCurdy, C. R.; Sufka, K.J.; Smith, G.H.; Warnick, J.E.; Nieto, M.J., Antinociceptive profile of salvinorin A, a structurally unique kappa opioid receptor agonist. Pharmacology Biochemistry and Behavior 2006, 83, 109-113.

18. John, T. F.; French, L.G.; Erlichman, J.S., The antinociceptive effect of salvinorin A in mice. European Journal of Pharmacology 2006, 545, 129-133.

19. Munro, T. A.; Rizzacasa, M.A.; Roth, B.L.; Toth, B.A.; Yan, F., Studies toward the pharmacophore of salvinorin A, a potent κ opioid receptor agonist. Journal of Medicinal Chemistry 2005, 48, 345-348.

20. Harding, W. W.; Schmidt, M.; Tidgewell, K.; Kannan, P.; Holden, K.G.; Dersch, C.; Rothman, R.R.; Prisinzano, T.P., Synthetic studies of neoclerodane diterpenes from Salvia divinorum: selective modification of the furan ring. Bioorganic and Medicinal Chemistry Letters 2006, 16, 3170-3174.

21. Lee, D. Y. W.; Karnati, V.V.R.; He, M.; Liu-Chen, L.-Y.; Kondaveti, L.; Ma, Z.; Wang, Y.; Chen, Y.; Begiun, C.; Carlezon Jr., W.A.; Cohen, B., Synthesis and in vitro pharmacological studies of new C(2) modified salvinorin A analogues. Bioorganic and Medicinal Chemistry Letters 2005, 15, 3744-3747.

22. Béguin, C.; Richards, M.R.; Li, J.-G.; Wang, Y.; Xu, W.; Liu-Chen, L.-Y.; Carlezon Jr., W.A.; Cohen, B.M., Synthesis and in vitro pharmacological studies of new C(2) modified salvinorin A analogues. Bioorganic and Medicinal Chemistry Letters 2005, 15, 3744-3747.

23. Harding, W. W.; Schmidt, M.; Tidgewell, K.; Kannan, P.; Holden, K.G.; Gilmour, B.; Navarro, H.; Rothman, R.R.; Prisinzano, T.P., Synthetic studies of neoclerdane diterpenes from Salvia divinorum: semisynthesis of salvinicins A and B and other chemical transformations of salvinorin A. Journal of Natural Products 2006, 69, 107-112.

24. Tidgewell, K.; Harding, W.W.; Lozama, A.; Cobb, H.; Shah, K.; Kannan, P.; Dersch, C.M.; Parrish, D.; Deschamps, J.R.; Rothman, R.B.; Prisinzano, T.E.,

Synthesis of salvinorin A analogues as opioid receptor probes. Journal of Natural Products 2006, 69, 914-918.

25. Harding, W. W.; Tidgewell, K.; Byrd, N.; Cobb, H.; Dersch, C.M.; Butelman, E.R.; Rothman, R.B.; Prisinzano, T.E., Neoclerodane diterpenes as a novel scaffold for μ opioid receptor ligands. Journal of Medicinal Chemistry 2005, 48, 4765-4771.

26. Tidgewell, K.; Groer, C.; Harding, W.W.; Lozama, A.; Schmidt, M.; Dersch, C.; Moyer. R.A.; Rothman, R.R.; Bohn, L.M.; Prisinzano, T.P., Mu opioids derived from salvinorin A with altered receptor regulation. Journal of Medicinal Chemistry submitted.

27. Prisinzano, T. P., unpublished data. Personal communication.

28. Pogozheva, I. D.; Lomize, A.L., Mosberg, H.I., Opioid receptor three-dimensional structures from distance geometry calculations with hydrogen bonding constraints. Biophysical Journal 1998, 75, 612-634.

29. Yan, F.; Mosier, P.D.; Westkaemper, R.B.; Stewart, J.; Zjawiony, J.K.; Vortherms, T.A.; Sheffler, D.J.; Roth, B.L., Identification of the molecular mechanisms by which the diterpenoid salvinorin A binds to κ-opioid receptors. Biochemistry 2005, 44, 8643-8651.

30. Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C.A.; Motoshima, H.; Fox, B.A.; Le Trong, I.; Teller, D.C.; Okada, T.; Stenkamp, R.E.; Yamamoto, M.; Miyano, M., Crystal structure of rhodopsin: A G protein-coupled receptor. Science 2000, 289, 739-745.

31. Kane, B. E.; Nieto, M.J.; McCurdy, C.R.; Ferguson, D.M., A unique binding epitope for salvinorin A, a non-nitrogenous kappa opioid receptor agonist. FEBS Journal 2006, 273, 1966-1974.

32. Metzger, T. G.; Paterlini, M.G.; Portoghese, P.S.; Ferguson, D.M., Application of the message-address concept of the docking of naltrexone and selective naltrexone-derived opioid antagonists into opioid receptor models. Neurochemical Research 1996, 21, 1287-1294.

33. Singh, N.; Cheve, G.; Ferguson, D.M.; McCurdy, C.R., A combined ligand-based and target-based drug design approach for G protein-coupled receptors: application to salvinorin A, a selective kappa opioid receptor agonist. Journal of Computer-Aided Molecular Design 2006, 20, 471 - 493.

34. Okada, T.; Sugihara, M.; Bondar, A.-N.; Elstner, M.; Entel, P.; Buss, V., The Retinal conformation and its environment in Rhodopsin in light of a new 2.2 A° crystal structure. Journal of Molecular Biology 2004, 342, 571-583.

35. Pogozheva I. D.; Przydzial, M. J.; Mosberg H. I., Homology modeling of opioid receptor-ligand complexes using experimental constraints. AAPS Journal. 2005, 7, E434-E448.

36. Cramer III, R. D.; Patterson, D.E.; Bunce, J.D., Comparative Molecular Field Analysis (CoMFA) 1. Effect of shape on binding of steroids to carrier proteins. Journal of American Chemical Society 1988, 110, 5959-5967.

37. Cramer III, R. D.; DePriest, S.A.; Patterson, D.E.; Hecht, P, The developing practice of Comparative Molecular Field Analysis. In 3D QSAR in Drug Design: Theory, Methods and Applications, Kubinyi, H., Ed. ESCOM: Leiden, 1993; pp. 443-485.

38. Li, W.; Tang, Y.; Xie, Q.; Sheng, W.; Qiu, Z.-B., 3D-QSAR studies of orvinol analogs as K-opioid agonists. Journal of Molecular Modeling 2006, 12, 877-884.

39. Li, W.; Tang, Y.; Zheng, Y.-L.; Qiu, Z.-B., Molecular modeling and 3D-QSAR studies of indolomorphinan derivatives as kappa opioid antagonists. Bioorganic and Medicinal Chemistry 2006, 14, 601-610.

40. Peng, Y.; Keenan, S.M.; Zhang, Q.; Welsh, W.J., 3D-QSAR comparative molecular field analysis on delta opioid receptor agonist SNC80 and its analogs. Journal of Molecular Graphics and Modelling 2005, 24, 25-33.

41. Peng, Y.; Keenan, S.M.; Zhang, Q.; Kholodovych, V.; Welsh, W.J., 3D-QSAR comparative molecular field analysis on opioid receptor antagonists: Pooling data from different studies. Journal of Medicinal Chemistry 2005, 48, 1620-1629.

42. Podlogar, B. L.; Poda, G.I.; Demeter, D.A.; Zhang, S.-P.; Carson, J.R.; Neilson, L.A.; Reitz, A.B.; Ferguson, D.M., Synthesis and evaluation of 4-(N,N-diarylamino)piperidines with high selectivity to the δ-opioid receptor: A combined 3D-QSAR and ligand docking study. Drug Design and Discovery 2000, 17, 34-50.

43. Rong, S.-B.; Zhu, Y.-C.; Jiang, H.-L.; Wang, Q.-M.; Zhao, S.-R.; Chen, K.-X.; Ji, R.-Y., Interaction models of 3-methylfentanyl derivatives with μ opioid receptors. Acta Pharmacologica Sinica 1997, 18, 128-132.

44. Oprea, T. I.; Waller, C. L., Theoretical and practical aspects of three-dimensional Quantitative structure-activity relationships. In Reviews in Computational Chemistry, Lipkowitz, K. B.; Boyd, D. B., Eds. WILEY-VCH: New York, 1997; Vol. 11, pp. 127-182.

45. Perola, E.; Charifson, P. S., Conformational analysis of drug-like molecules bound to proteins: An extensive study of ligand reorganization upon binding. Journal of Medicinal Chemistry 2004, 47, 2499-2510.

46. Boström, J.; Norrby, P.-O.; Liljefors, T., Conformational energy penalties of protein-bound ligands. Journal of Computer-Aided Molecular Design 1998, 12, 383-396.

47. Debnath, A. K., Three-dimensional quantitative structure-activity relationship study on cyclic urea derivatives as HIV-1 protease inhibitors: Application of Comparative Molecular Field Analysis. Journal of Medicinal Chemistry 1999, 42, 249-259.

48. Nicklaus, M. C.; Wang, S.; Driscoll, J.; Milne, G. W. A., Conformational changes of small molecules binding to proteins. Bioorganic and Medicinal Chemistry 1995, 3, 411-428.

49. Veith, M.; Hirst, J. D.; Brooks, I., C.L., Do active site conformations of small Ligands correspond to low free-energy solution structures? Journal of Computer-Aided Molecular Design 1998, 12, 563-572.

50. Klebe, G.; Abraham, U., On the prediction of binding properties of drug molecules by Comparative Molecular Field Analysis. Journal of Medicinal Chemistry 1993, 36, 70-80.

51. Greco, G.; Novellino, E.; Martin, Y. C., Approaches to three-dimensional quantitative structure-activity relationships. In Reviews in Computational Chemistry, Lipkowitz, K. B.; Boyd, D. B., Eds. WILEY-VCH: New York, 1997; Vol. 11, pp. 183-240.

52. Cramer III, R. D.; Patterson, D. E.; Bunce, J. D., Comparative Molecular Field Analysis (CoMFA).1. Effect of shape on binding of steroids to carrier proteins. Journal of the American Chemical Society 1988, 110, 5959-5967.

53. Cramer III, R. D.; DePriest, S. A.; Patterson, D. E.; Hecht, P., The developing practice of Comparative Molecular Field Analysis. In 3D QSAR in Drug Design: Theory, Methods, and Applications, Kubinyi, H., Ed. ESCOM: Leiden, 1993; Vol. 1, pp. 443-485.

54. Martin, Y. C.; Kim, K-H; Lin, T., Comparative Molecular Field Analysis:CoMFA. In Advances in Quantitative Structure Property Relationships, Charton, M., Ed. JAI Press Inc Elsevier: 1996; Vol. 1, pp. 1-52.

55. Wold, S.; Johansson, E.; Cocchi, M., PLS - Partial least squares projections to latent structures. In 3D QSAR in Drug Design: Theory, Methods, and Applications, Kubinyi, H., Ed. ESCOM: Leiden, 1993; Vol. 1, pp. 523-550.

56. Forina, M.; Casolino, C.; Pizarro Millan, C., Interactive predictor weighting (IPW) PLS: A technique for the elimination of useless predictors in regression problems. 1999, 13, 165-184.

57. Lindgren, F., Geladi, P., Rännar, S., Wold, S., Interactive variable selection (IVS) for PLS. Part 1: Theory and algorithms. Journal of Chemometrics 1994, 8, 349-363.

58. Golbraikh, A.; Tropsha, A., Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. Journal of Computer-Aided Molecular Design 2002, 16, 357-369.

59. Golbraikh, A.; Shen, M.; Xiao, Z.; Xiao, Y.-D.; Lee, K.-H.; Tropsha, A., Rational selection of training and test sets for the development of validated QSAR models. Journal of Computer-Aided Molecular Design 2003, 17, 241-253.

60. Clark, R. D.; Fox, P.C., Statistical variation in progressive scrambling. Journal of Computer-Aided Molecular Design 2004, 18, 563-576.

61. Golbraikh, A.; Tropsha, A., Beware of q^2! Journal of Molecular Graphics and Modeling 2002, 20, 269-276.

62. SYBYL 7.2. Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.

63. Valdes III, L. J.; Butler, W. M.; Hatfield, G.M.; Paul, A. G.; Koreeda, M., Divinorin A, a psychotropic terpenoid, and divinorin B from the hallucinogenic Mexican mint, Salvia divinorum. Journal of Organic Chemistry 1984, 49, 4716-4720.

64. Powell, M. J. D., Restart procedures for the conjugate gradient method. Mathematical Programming 1977, 12, 241-254.

65. Clark, M.; Cramer III, R. D.; Van Opdenbosch, N., Validation of the general purpose Tripos 5.2 force field. Journal of Computational Chemistry 1989, 10, 982-1012.

66. Gasteiger, J.; Marsili, M., Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges. Tetrahedron 1980, 36, 3219-3288.

67. Bush, B. L.; Nachbar, R. B. Jr., Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA. Journal of Computer-Aided Molecular Design 1993, 7, 587-619.

68. SYBYL 7.2 Tripos Bookshelf 7.2. Tripos Inc., 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA.

69. Golbraikh, A.; Tropsha, A., Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. Journal of Computer - Aided Molecular Design 2002, 16, 357-369.

70. Clark, M.; Cramer III, R. D., The probability of chance correlation using partial least squares (PLS). Quantitative Structure-Activity Relationships 1993, 12, 137-145.

71. Ni, Q.; Xu, H.; Partilla, J.S.; De Costa, B.R.; Rice, K.C.; Rothman, R.B., Selective labeling of κ2 opioid receptors in rat brain by [125I]IOXY: Interaction of opioid peptides and other drugs with multiple κ(2a) binding sites. Peptides 1993, 14, 1279-1293.

72. Golbraikh, A.; Shen, M.; Xiao, Z.; Xiao, Y.-D.; Lee, K.-H.; Tropsha, A., Rational selection of training and test sets for the development of validated QSAR models. Journal of Computer - Aided Molecular Design 2003, 17, 241-253.

73. Schmidt, M. D.; Schmidt, M.S.; Butelman, E.R.; Harding, W.W.; Tidgewell, K.; Murry, D.J.; Kreek, M.J.; Prisinzano, T.E., Pharmacokinetics of the plant-derived κ-opioid hallucinogen salvinorin A in nonhuman primates. Synapse 2005, 58, 208-210.

74. Wang, Y.; Tang, K.; Inan, S.; Siebert, D.; Holzgrabe, U.; Lee, D.Y.W.; Huang, P.; Li, J.-G.; Cowan, A.; Liu-Chen, L.-Y., Comparison of pharmacological activities of three distinct κ ligands (Salvinorin A, TRK-820 and 3FLB) on κ opioid receptors in vitro and their antipruritic and antinociceptive activities in vivo. Journal of Pharmacology and Experimental Therapeutics 2005, 312, 220-230.

75. Wade, L. G., Organic Chemistry. 5th ed.; Prentice Hall: 2002.