Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

ABSTRACT

COMPARATIVE MOLECULAR FIELD ANALYSIS (CoMFA) OF PHENYL RING SUBSTITUTED METHYLPHENIDATES

Comparative Molecular Field Analysis (CoMFA) was performed on 30 methylphenidate analogues having phenyl ring substituents at the 2-, 3-, and 4-positions in an attempt to explain the structure-activity relationships of the analogues. Several CoMFA studies were carried out using different conformations of methylphenidate as the template structures for the corresponding alignment rules. The identification of the bioactive conformation was variously based on the interfeature distances in WIN 35,428, on cluster analysis, and on a comparison of active methylphenidate analogues. The sensitivity of the models to various CoMFA parameters was compared. These parameters included altering the size and spacing of the CoMFA grid, and changing different run-time CoMFA standards such as the steric and electrostatic energy cutoffs, and the column filtering value. Predictions of "novel", yet-to-be-synthesized methylphenidate analogues and of test sets of analogues were carried out using different models. The results showed that the q^2 of a model is affected most by the template structure, followed (in order of decreasing effect on q^2) by the energy cutoffs, the column filtering value, and grid spacing and region box parameters. It was found that the 2-position analogues have poor IC_{50} values because they protrude into a sterically unfavorable region. Based on the CoMFA studies performed, a novel compound, 3,4-Br methylphenidate, has been suggested for synthesis and testing.

COMPARATIVE MOLECULAR FIELD ANALYSIS (CoMFA) OF PHENYL RING SUBSTITUTED METHYLPHENIDATES

by
Milind Misra

A Master's Thesis
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Applied Chemistry

Department of Chemical Engineering, Chemistry and Environmental Science

August 1999

APPROVAL PAGE

COMPARATIVE MOLECULAR FIELD ANALYSIS (CoMFA) OF PHENYL RING SUBSTITUTED METHYLPHENIDATES

Milind Misra

Dr. Carol A. Venanzi, Thesis Advisor — Date
Distinguished Professor of Chemistry, NJIT

Dr. David S. Kristol, Committee Member — Date
Professor of Chemistry, NJIT

Dr. Tamara M. Gund, Committee Member — Date
Professor of Chemistry, NJIT

BIOGRAPHICAL SKETCH

Author: Milind Misra

Degree: Master of Science in Applied Chemistry

Date: August 1999

Undergraduate and Graduate Education:

- Master of Science in Applied Chemistry,
  New Jersey Institute of Technology, Newark, NJ, 1999

- Bachelor of Science in Engineering Science,
  New Jersey Institute of Technology, Newark, NJ, 1998

Presentations and Publications:

Joseph S. Hinksmon, Milind Misra, Ronald A. Buono, Carol A. Venanzi, Howard M. Deutsch, Qing Shi, Margaret M. Schweri, "Computer-Aided Molecular Design of Molecules Potentially Useful in the Treatment of Cocaine Abuse," AIChE Undergraduate Research Symposium, Columbia University, April 24, 1999.

Milind Misra, Ronald A. Buono, Carol A. Venanzi, Howard M. Deutsch, Qing Shi, Margaret M. Schweri, "Comparative Molecular Field Analysis of Methylphenidate Analogs with Phenyl Ring Substituents," Quantitative Structure-Activity Relationship (QSAR) Gordon Conference, July 1999.

To my teachers

ACKNOWLEDGMENT

I would like to thank my advisor, Carol Venanzi, for her constant and genuine support during the course of my master's degree. If not for Dr. Venanzi, I would have been unable to initiate, persevere through, and complete my degree. Indeed, she probably spent as much time editing my thesis as I took to write it.

I am grateful to the Department of Chemical Engineering, Chemistry and Environmental Science and the Office of Graduate Studies for their comprehensive financial assistance that sustained me for over a year.

Special thanks to Dr. David Kristol and Dr. Tamara Gund for agreeing to be my thesis committee members and providing suggestions for improvement.

This research was supported by National Institute of Health grant DA11541. Data presented in this work was provided by Dr. Howard Deutsch (Georgia Institute of Technology) and Dr. Margaret Schweri (Mercer University School of Medicine).

I also wish to acknowledge the guidance of Ron Buono (who, incidentally, was the first person to suggest pursuing a master's), Bill Skawinski, and Dmitry Bondarev who provided valuable input whenever there was a need, and together contributed to an atmosphere conducive for research. Joseph Hinksmon helped create some graphs and tables presented in this thesis.

Finally, I would like to thank David Perel and Ram Reddy of Engineering Computing for their prompt and usually effective hardware and network support.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF TABLES
(Continued)

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Perspective

Cocaine abuse continues to be a major health problem in the United States and has, over the years, also become a serious threat to a stable society internationally. This study forms part of the ongoing attempt to develop a pharmacotherapeutic agent for the treatment of cocaine dependence. Such a compound would be a selective dopamine reuptake inhibitor in that it would exhibit a high binding affinity for the dopamine transporter and simultaneously permit some magnitude of dopamine reuptake. One class of compounds which exhibits these characteristics is the methylphenidate analogues. The mechanism of action of methylphenidate (Ritalin®) is similar to that of cocaine.[1-4] Studies of methylphenidate analogues have shown that some *threo*-methylphenidate derivatives with halogen substituents on the aromatic ring have increased binding affinity.[1,5-10] Furthermore, for most aromatic ring substituted compounds there was a clear correlation between their binding affinity and their ability to substitute for cocaine in drug discrimination studies. However, as the three-dimensional structure of the cocaine receptor site is unknown, the design of new ligands must be based on structure-activity relationships derived from a set of known ligands. Comparative Molecular Field Analysis (CoMFA) is at present the most widely used approach for three-dimensional quantitative structure-activity relationship (3D-QSAR) studies.

The objective of this research is twofold. First, to use CoMFA to elucidate structure-activity relationships of methylphenidate analogues with phenyl ring

substituents. Second, to use CoMFA to predict the activity of novel methylphenidate analogues. This research is one component of the iterative cycle of concerted drug development using molecular modeling, synthesis, binding and behavioral studies. Thus, molecular modelers use available experimental data to propose a novel compound that is synthesized and then tested for biological activity. The feedback obtained from testing is used to fine-tune the predictive model and the cycle is repeated until a satisfactory compound is realized. The methylphenidate analogues examined in this study were synthesized in the laboratory of Dr. Howard Deutsch at Georgia Institute of Technology, while testing was carried out in the laboratory of Dr. Margaret Schweri at Mercer University School of Medicine. This work was supported in part by NIH grant DA11541. By determining exploratory CoMFA models using phenyl ring substituted methylphenidate analogues, this research marks a beginning in molecular modeling studies on methylphenidate analogues in particular and on dopamine reuptake inhibitors in general.

1.2 Outline

The remainder of Chapter 1 presents the rationale of this research including a discussion of the hypothesized mechanism of cocaine action, the importance of the dopamine transporter and the "Dopamine Hypothesis", and the significance of methylphenidate analogues. Chapter 2 introduces the development of CoMFA as a powerful predictive tool, and describes its important components such as the statistical Partial Least Squares technique and the Leave-One-Out cross-validation. Conformational analysis, the selection of a template structure (for alignment) that could reflect the bioactivity of the

methylphenidate analogues, and the generation of a suitable molecular alignment of the 30 analogues considered in this study form important steps in the preparation for actual CoMFA experiments and are considered in Chapter 3. Several CoMFA experiments were carried out, using different alignments, to compare the sensitivity of the results to variations in the CoMFA parameters used. Some of these studies included altering the size and spacing of the CoMFA grid, changing different run-time CoMFA parameters, and prediction of "novel" compounds and of "test set" compounds. Chapter 4 illustrates results obtained from these experiments. Finally, Chapter 5 discusses the conclusions.

1.3 Background

1.3.1 About Cocaine

An alkaloid extracted from the leaves of the coca plant, cocaine is a powerful local anesthetic with vasoconstrictor properties. It is also used and abused as a mental stimulant that is capable of generating a series of epinephrine-like reactions due to the stimulation of the sympathetic nervous system and release of epinephrine. The desired effects of cocaine are feelings of euphoria, alertness, excitement, and rapid flow of thought but these could be expressed as hyperactivity, increased confidence, talkativeness, and mood-shifts. As a stimulant, it helps to combat fatigue and instill feelings of increased capacity to do work, great physical strength, and mental superiority.

In the brain, cocaine inhibits reuptake of dopamine (DA), norepinephrine, and serotonin in central and peripheral nerve synapses[11-15] and thus effectively prolongs and augments their effects. The inhibition of dopamine reuptake is thought to cause euphoria and is the main reason why users continue to take the drug. Mice that have been

genetically altered to lack a central dopamine reuptake mechanism at synapses do not respond to cocaine at all.[16]

1.3.2 The Dopamine Transporter

The dopamine transporter (DAT) is thought to contain a specific binding site for cocaine.[17] Dopamine nerve endings have unique, high-affinity dopamine uptake sites that play a key role in terminating transmitter action and in maintaining transmitter homeostasis. A membrane carrier that is capable of bi-directional transmission of dopamine is responsible for uptake.[18] The dopamine transporter recycles dopamine in the synaptic cleft by recycling extracellular dopamine (ECDA) back into the nerve terminal. The structural and pharmacological features of the dopamine transporter have been reviewed,[19] and rat,[20-22] bovine,[23] and human[24] dopamine transporters have been cloned and characterized. The human DAT is 92% identical to the rat DAT and 84% identical to the bovine DAT. The human DAT is a protein made of 620 amino acids and has three glycosylation sites. It is hypothesized to contain 12 hydrophobic putative transmembranal domains and is a member of the family of Na^+/Cl^- -dependent plasma membrane transporters. Chimeric dopamine-norepinephrine transporters have been employed to describe distinct domains for ionic dependence and uptake mechanisms (domains 1-5), cocaine and tricyclic antidepressant interactions (domains 6-8), and substrate affinity and stereoselectivity (domains 9-12) on the DAT.[25] In the seventh transmembrane segment, there are several closely spaced serine residues that could act to bind the hydroxyl groups of DA. Kitayama, et al.[26] have shown that replacement of some

of these serine residues by glycine or alanine in the rat DAT completely impairs DA uptake ability.

1.3.3 The "Dopamine Hypothesis"

According to the "Dopamine Hypothesis",[27] the reinforcing properties of cocaine arise when cocaine binds to the dopamine site on the DAT inhibiting the reuptake of DA by the DAT. This raises the ECDA levels and eventually leads to euphoric and addictive effects.[28] The finding that binding of cocaine in some of the chimeric dopamine-norepinephrine transporters can be virtually eliminated without affecting the uptake properties suggests that a specific determinant for cocaine binding must exist independently of the binding of substrate. This implies that an antagonist of cocaine action without uptake blockade activity could be developed for clinical use for controlling drug dependence. Such a compound would bind to some part of the cocaine site on the DAT but would not inhibit DA reuptake. The drawback of this approach is that the patient could annul the effect of treatment simply by self-administering more cocaine, leading perhaps to dangerous side effects such as cardiac arrhythmias. Rothman, et al.[29,30] suggest a different strategy by seeking to develop a DAT-selective drug that binds with high affinity to, but dissociates slowly from, the DAT, such that it behaves like a noncompetitive inhibitor to cocaine. By partially inhibiting DA reuptake, the ideal agent would provide sufficient ECDA to minimize cocaine craving, yet insufficient to produce euphoria. The ideal agent would also be insoluble in aqueous solution to preempt patients from covertly self-administering it intravenously.

Pharmacologically, a number of drugs have been used to reduce cocaine use, decrease craving, and alleviate dependence or symptoms of withdrawal. Research has concentrated mainly on two approaches. In the first, antidepressants[31,32] have been used because cocaine abstinence produces symptoms similar to depression but these invariably produce troublesome side effects.[33,34] The second popular approach has been to employ drugs that interact with DA because of this neurotransmitter's important role in the positive reinforcement produced by cocaine. Drug development targeting the DAT has resulted in the generation of some very potent and selective molecules of diverse structural backgrounds.[35-37] The identification of radioligands, such as nomifensine,[38] methylphenidate,[39] 1-[2-(diphenylmethoxy)-ethyl]-4-(3-phenylpropyl)-piperazine (GBR 12935),[40,41] mazindol,[12] *N*-[1-(2-benzo(*b*)thiophenyl)cyclohexyl] piperidine (BTCP),[42] benztropine[11,37,43] and its analogues,[37,44-50] and cocaine and its analogues[17,38-42,51-58] which bind to the DAT and influence the reuptake of DA has led to a wealth of structure-activity data, much of which has been summarized recently.[35,59-63] To date, a selective dopamine uptake inhibitor with the characteristics of the ideal agent outlined above has not been identified.

1.3.4 The Pharmacophore Model

According to Kier[64] who popularized the term in the late 1960s, a pharmacophore consists of those structural features that impart to a molecule a particular pharmacological activity. A pharmacophore model for the ideal agent defines the three-dimensional geometric arrangement between chemical moieties (such as a carbonyl group or aromatic ring) or chemical features (such as a hydrogen bond donor or acceptor) that have been

identified from the experimental structure-activity data as being required for biological activity. As additional structure-activity data becomes available, the pharmacophore model for a particular set of molecules can be refined further. The available structure-activity data for methylphenidate analogues are discussed below.

1.3.5 Structure-Activity Relationships (SAR) of Methylphenidate Analogues

Methylphenidate (MP) is a central nervous system stimulant that is the drug of choice for the treatment of Attention Deficit Hyperactivity Disorder in children. It was first synthesized over 50 years ago[65] and was recognized as a stimulant ten years later.[66] Recently, because of the similarity of its mechanism of action to that of cocaine,[1-4] there have been several studies of MP analogues.[1,5-10] All of the structure-activity studies to date originate from the laboratories of Drs. Howard Deutsch and Margaret Schweri[67-72] and have been recently reviewed.[61] The data presented here come from these studies (both published and unpublished) and reflect the analogues' ability to inhibit the binding of [^{3}H]WIN 35,428 to rat striatal tissue membrane preparations, as well as the uptake of [^{3}H]dopamine into rat striatal synaptasomes.[67] The studies indicate that some *threo*-MP derivatives with halogen substituents on the aromatic ring have increased binding affinity compared to the unsubstituted *threo*-MP. Electron-donating substituents caused little change or a small loss of affinity. *N*-substituents generally caused a marked loss of affinity, except for the benzyl moiety that has slightly increased affinity. For the unsubstituted amine, reduction of the methyl ester to a hydroxyl sharply reduces affinity, whereas conversion of the hydroxyl to a methyl ether restores affinity. For *N*-benzyl (tertiary amine) derivatives, the same changes increase affinity. Although the activity of

the MP derivatives was highly correlated in the binding and uptake assays,[67] a number of compounds were 4- to 8-fold more effective at inhibiting [^{3}H]WIN 35,428 binding than [^{3}H]dopamine uptake. For most aromatic ring substituted compounds, there was a clear correlation between their binding affinity and their ability to substitute for cocaine in drug discrimination studies. This correlation did not hold for *N*-substituted analogues which were generally much less potent in the drug discrimination test than predicted by their binding affinity. A pharmacophore model has been proposed to explain the similar pharmacological properties of cocaine, MP, and the WIN analogue CFT.[73] This model identifies the protonated amine, the ester group, and the aromatic ring as important structural elements. Superposition of the sequence of atoms from the amine group through the ester group of the (+)-(R, R)-*threo*-MP (calculated from the X-ray structure), CFT, and cocaine showed the similarity of these positions in space, with the phenyl rings of the three compounds oriented in very different regions of space. Froimowitz and coworkers[73] concluded that the DAT can accommodate a wide variety of positions of the phenyl ring. However, it is possible that the DAT site, which accommodates the C-2 position substituent, is a large lipophilic pocket, or that the DAT site may accommodate all three phenyl rings in the same orientation, thereby forcing the 2-position substituents into different parts of the lipophilic pocket. The authors did not propose a detailed pharmacophore model for MP nor did they investigate superposition of other low energy conformers of the ligands or alternative superpositions, or attempt to explain the structure-activity relationships of the MP analogues. It is here that molecular modeling studies incorporating CoMFA of the MP analogues may be useful. The research

presented here is an attempt to generate CoMFA models based on an analysis of the structure-activity relationships of MP analogues with phenyl ring substituents.

The published[61,67] and unpublished binding and uptake data for the 30 MP analogues indicate that the more active compounds, defined as those with the higher affinity for the binding site (lower IC_{50}), are 3- or 4- substituted, such as for example, 3-Br (binding IC_{50} = 4.2 nM), 3-Cl (5.1 nM), 3,4-Cl (5.3 nM), and 4-Br (6.9 nM). The unsubstituted MP analogue has a binding IC_{50} value of 83.0 nM and an uptake IC_{50} value of 224.0 nM. In general, 3- and 4- substituted compounds have comparable IC_{50} values while the 2-substituted compounds are relatively much less active, having much higher IC_{50} values. This is probably because of increased steric effects introduced in the compound by the 2-position substituent. The use of CoMFA may help establish such a correlation between differences in bioactivity and the SAR data for the MP analogues.

CHAPTER 2

CoMFA THEORY

2.1. CoMFA 3D-QSAR

Comparative Molecular Field Analysis (CoMFA) is one of the most powerful tools in the area of three-dimensional quantitative structure activity relationships (3D-QSAR) studies. It was developed by Cramer, et al.[74] for predicting the biological activity of a drug and is based on the assumption that the interactions between the ligand and its receptor site are primarily non-covalent in nature. Since most drugs exhibit biological activity first by binding to a target receptor molecule, CoMFA assumes that changes in the biological activities of drug molecules correspond to changes in their steric and electrostatic energies, which, in turn, are considered to be the chief components of drug-receptor interactions. The method itself contains six fundamental steps: (1) selection of a bioactive conformation of each compound under study; (2) superposition of selected structures in an alignment that best approximates the putative alignment of the active drugs at the receptor site; (3) generation of grid points with regular spacing around the molecules under investigation; (4) sampling of the steric and electrostatic interaction between the investigated molecule and the probe atoms at each grid point; (5) construction of the model equation between the interaction energies and biological activities of the molecules in the series using Partial Least Squares (PLS) analysis; and (6) display of the coefficient contour maps of the PLS model equation by computer graphics. Of these six steps, the first two (i.e. the selection of a suitable template structure and the subsequent alignment rule) are described for MP analogues in Chapter

3. The remaining four steps are incorporated into the CoMFA Studies described in Chapter 4.

The CoMFA technique has been successfully used for the prediction of the biological activity of many different types of drugs, such as cocaine analogues,[75-77] benzodiezapine inverse agonists,[78] drugs that act at the cannabinoid receptor,[79] anticoccidial triazines,[80] drugs that bind to the human rhinovirus, thermolysin, and renin,[81] 5-HT_{5A} receptor agents,[82] imidazoline receptor ligands,[83] N-methyl-D-protease inhibitors,[84] phenothiazines and other multidrug resistance modifiers,[85] dopamine receptor agonists,[86] and HIV-1 protease inhibitors.[87] About 400 CoMFA papers were published during 1993-1997 alone.[88] Recent reviews of the methodology have pointed out its scope and limitations.[89-92]

Since the molecular structure and properties of the cocaine receptor site are unknown, CoMFA is performed by calculating and comparing the molecular steric (Lennard Jones) and electrostatic (Coulombic) "fields" of both active and inactive analogues. These fields are calculated at each lattice point around each analogue using a probe atom, such as a sp^3 carbon atom with +1 charge, at regularly-spaced points on the three-dimensional CoMFA grid. The energy values thus calculated are entered into columns in a CoMFA QSAR table. Application of special multivariate statistical analyses routines, such as Partial Least Squares analysis, cross-validation, and bootstrapping, produces a statistically significant final CoMFA equation. The result of this analysis is a "cross-validated r^2" (called q^2), which is a quantitative indication of the quality of the final CoMFA model. A q^2 value close to 1.0 suggests that the model equation is most predictive, and when it is small or negative, the model equation is not

good enough for prediction. A q^2 value of greater than 0.5 is considered significant. The optimum number of components generated by the cross-validated analysis (see Sections 2.2 and 2.3) is used to perform a non-validated analysis. The results of this non-validated run are displayed as three-dimensional color-coded contour plots that show differences in the steric and electrostatic fields and, therefore, illustrate the differences in binding affinity due to particular substituents. These maps help the user to easily identify the properties in particular regions of space that correspond to increased biological activity and vice-versa. It should be kept in mind, however, that the accuracy of the pharmacophore model in predicting activity of new compounds depends on the quality and availability of experimental structure-activity data.

The two biggest hurdles in generating a CoMFA model with a diverse set of molecules are the selection of the bioactive conformation and the need for a suitable superposition rule. Often these two procedures are subjective and time-consuming, and could compound the problem by being interdependent. For most studies, superposition so that the most important pharmacophore features (such as hydrogen bond donors and acceptors, aromatic rings, and hydrophobic regions) are aligned is sufficient for good results.[90] This is a straightforward procedure in the case of structurally-related analogues, such as the methylphenidate analogues considered here. It may be more difficult to define the important pharmacophore features (and thereby the superposition rule) for comparison of molecules from several different structural classes. Most studies seem to choose the global energy minimum gas phase conformation as the bioactive conformation[76,77,93], even though there is evidence from ligand-protein binding studies

that the bound conformation of the substrate is different (sometimes dramatically so) from the gas phase global energy minimum conformation.[94,95]

2.2 Partial Least Squares Technique

Partial Least Squares (PLS) is a popular regression method developed by H. Wold and S. Wold and is used to solve the multivariate structure-activity relationships in QSAR studies.[96] The three-dimensional information that is obtained from calculating the steric and electrostatic field energy values (i.e. descriptors) at each point on the CoMFA grid is placed in a two-dimensional CoMFA QSAR table. This table contains hundreds or thousands of such descriptors for each molecule. PLS assumes that these descriptors capture the dominant effects due to changes in molecular substructure. Depending on the probe used in the molecular calculations, these descriptors may be electrostatic and steric energy values (as mentioned above) or lipophilic or other contributions to the ligand-receptor interaction. PLS also assumes the presence of a small number of "intrinsic" variables (called latent variables) that contain all the information relevant to the biological activity. The result of a PLS analysis is an equation that describes or predicts the differences in the values of one or more table columns (the dependent variables, or target properties, such as the biological activity given by the IC_{50} value) from differences in the values in other table columns (the descriptors).

In PLS, the relationship between the parameters of the given samples (**X**) and the corresponding responses (**y**) is described using a linear model equation with latent variable **t**. The latent variable **t** is expressed by the linear combination of parameters (equation 1).

$$\mathbf{t} = \mathbf{Xw} \qquad (1)$$

where **w** is the weight factor for calculation of the latent variable. Bilinear model equations are shown in equations 2 and 3.

$$\mathbf{X} = \sum_{i=1}^{A} \mathbf{t}_i\mathbf{p}_i + \mathbf{E} \qquad (2)$$

$$\mathbf{Y} = \sum_{i=1}^{A} \mathbf{u}_i\mathbf{q}_i + \mathbf{F} \qquad (3)$$

where: **u** = latent variable for the response variable

p = loadings corresponding to **t**

q = loadings corresponding to **u**

E = model residuals for **X**

F = model residuals for **Y**

A = number of components in the PLS model equation

2.3 Leave-One-Out Cross-Validation

The number of PLS components (*A*) is usually determined by means of the Leave-One-Out cross-validation method. In this procedure, the data for one of the ligands is deleted and a CoMFA model is constructed and used to predict the biological activity of the deleted ligand. The predicted error sum of squares (PRESS), the cross-validated correlation coefficient (q^2), and the cross-validated standard error of estimate (s_{cv}) are computed as shown in equations 4-7. The data for that ligand is then restored to the set and the data for the next ligand is deleted. The CoMFA model is derived again and used to predict the activity of the second ligand and so on for each ligand in the set.

$$PRESS = \sum_{Y} (Y_{pred} - Y_{actual})^2 \qquad (4)$$

$$SS = \sum_{Y} (Y_{actual} - Y_{mean})^2 \qquad (5)$$

$$q^2 = 1 - [\, PRESS / SS \,] \qquad (6)$$

$$s_{cv} = \sqrt{\frac{PRESS}{n - A - 1}} \qquad (7)$$

where: Y = set of all samples

Y_{pred} = a predicted value

Y_{actual} = an actual or experimental value

Y_{mean} = the best estimate of the mean of all values that might be predicted

n = number of rows

A = number of components

SS = sum of squared residuals (actual)

s_{cv} = cross-validated standard error of estimate

The number of components to be used in the non-validated run (i.e. the optimum number of components) is selected as the number of components that significantly reduces s_{cv} (or increases q^2), or that corresponds to the minimum s_{cv} (or to the maximum q^2). After the non-validated run, the resulting CoMFA model is then used to predict the activity of a new molecule. It should be noted, however, that the new molecule cannot differ too much from those used to derive the CoMFA equations; usually, the difference is restricted to a small region of the molecule such as a substituent. This new molecule is added to the set, superimposed on the template, and its steric and electrostatic fields are calculated. The results are displayed as three-dimensional contour plots that correlate

steric and electrostatic features in particular regions of space with increased or decreased bioactivity. Finally, the predicted r^2 and s_{cv} values are calculated to get a quantitative idea of the bioactivity of the new molecule.

CHAPTER 3

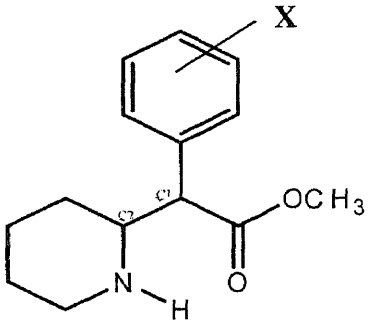
CoMFA METHODS

3.1 Software and Hardware

The conformational analysis and the CoMFA studies described here were carried out using SYBYL (version 6.5), a pharmacophore modeling package available from Tripos, Inc. Calculations were carried out on a Silicon Graphics O2 workstation (4.2 GB hard drive, 128 MB RAM) at New Jersey Institute of Technology.

3.2 Conformational Analysis

This study investigated 30 neutral MP analogues with aromatic ring substitution at the 2-, 3-, or 4-position as shown in Figure 3.1. Methylphenidate has two asymmetric carbon atoms (C1 and C2 in Figure 3.1) and can therefore exist as two pairs of diastereomeric enantiomers. Most pharmacological activities of MP seem to be associated with the *threo* isomer whose (+)-enantiomer is believed to be responsible for the stimulant properties.[6,68,97-102] The absolute configuration of this active enantiomer has been determined to be (R,R).[103] It was thus decided to use for this study the *threo* analogues having the (R,R) absolute configuration.



X:

None				
2-Br	3-Br	4-Br		
2-Cl	3-Cl	4-Cl	3,4-Cl	3,5-Cl
2-F	3-F	4-F		
2-OCH_3	3-OCH_3	4-OCH_3	3,4-OCH_3	
2-OH	3-OH	4-OH		
2-naphthyl				
	3-CH_3	4-CH_3		3,5-CH_3
	3-NH_2	4-NH_2		
		4-I		
		4-CF_3		
		4-C_2H_5		
		4-NO_2		
		4-*t*-butyl		

Figure 3.1 The Methylphenidate Analogues Investigated in This Study

Investigation of the conformational space of a molecule is inherently a difficult process[104-106] but is necessary for many computer-aided analyses such as CoMFA. Methylphenidate is a flexible molecule as it contains readily rotatable single bonds, and is structurally somewhat similar to cocaine, which is shown in Figure 3.2. Conformational analysis to search the conformational potential energy surface of the MP analogues and to generate a set of low-energy conformers for use in the pharmacophore model was performed using the RANDOMSEARCH utility and the SYBYL Maximin2 force field. A cutoff value of 5 kcal/mol above the global energy minimum (GEM) was used to choose the conformers to be considered for analysis. Four torsional angles affecting the relative orientation of putative pharmacophoric elements (i.e. the nitrogen, the ester group, and the phenyl ring) were selected to be randomly altered. The four bonds involved were the three between the central carbon atom and the phenyl ring, the piperidine ring, and the carbomethoxy side chain, and that between the carbonyl carbon and methoxy oxygen. These are indicated for unsubstituted MP in Figure 3.3.

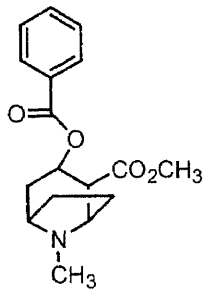


Figure 3.2 Cocaine

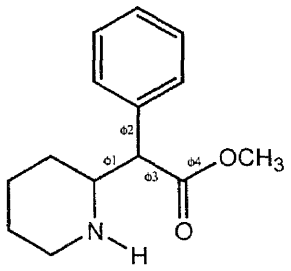


Figure 3.3 Torsional Angles in MP for Conformation Searching

Appendix A provides the SYBYL parameters used for the RANDOMSEARCH procedure. RANDOMSEARCH randomly alters the selected torsional angles to find the

various energy minima available to a given molecule and minimizes the energy of the resulting geometry. After minimization, the conformation is compared with those already located and is saved if it is unique. The RMS superposition threshold value (0.2 Å here) defined during RANDOMSEARCH setup distinguishes between two minima: a RMS value greater than the threshold represents two different conformations. Minimization was carried out using the Powell method with a maximum number of 1,000 iterations allowed. The Gasteiger-Hückel method of charge computation was used during minimization. This method combines the Gasteiger-Marsili method (to calculate the σ component of the atomic charge) and the Hückel method (to calculate the π component of the atomic charge).[107-111] The total charge is the sum of the charges computed by the two methods. A distance-dependent dielectric function was used with dielectric constant of 1.0. A cutoff of 8 Å was set on the non-bonded interactions. To ensure a comprehensive search of 3D space, the RANDOMSEARCH routine visited 10,000 conformations of each analogue and took 2-3 days per analogue to run on the workstation. During random searching, the (R,R) configuration of the *threo* analogues was preserved (see Appendix B) as was the low-energy chair conformation of the piperidine ring. In trial RANDOMSEARCH runs, it was found that the piperidine ring tended to undergo "ring inversion", changing the bulky substituent (i.e. the carbon atom attached to the phenyl ring and to the ester group) from the equatorial position to axial. Since the equatorial position is known to be more stable (assuming the absence of other interactions such as hydrogen bonding that could stabilize an axial conformation), the piperidine ring was held rigid in the chair conformation during the random search. For the same reason, the

hydrogen atom on the nitrogen was also retained in its equatorial position during random searching.

Table 3.1 Conformational Analysis of Methylphenidate Analogues

Substituent	Number of Unique Conformers [a]	Conformer Selected for Superposition [b]	Energy of Selected Conformer (kcal/mol)	State of Selected Conformer [c]	Relative Energy [d] (kcal/mol)
MP	20	6	0.53	GEM	0
2-Br	20	6	-0.06	GEM	0
2-Cl	22	9	-0.36	GEM	0
2-F	32	10	-1.54	GEM	0
2-naphthyl	47	9	1.45	GEM	0
2-OCH_3	76	23	0.12	GEM	0
2-OH	45	2	-0.75	GEM + 1	0.27
3-Br	34	15	-0.30	GEM + 1	0.04
3-CH_3	36	6	0.21	GEM + 1	0.06
3-Cl	36	3	-0.23	GEM + 1	0.05
3-F	32	20	-0.31	GEM + 1	0.06
3-NH_2	39	6	9.12	GEM + 7	1.08
3-OCH_3	113	38	1.90	GEM + 3	0.15
3-OH	35	7	0.31	GEM + 1	0.03
4-Br	19	3	-0.18	GEM	0
4-C_2H_5	59	13	0.31	GEM	0
4-CF_3	79	12	0.94	GEM	0
4-CH_3	22	5	0.36	GEM	0
4-Cl	21	1	-0.10	GEM	0
4-F	22	11	-0.08	GEM	0
4-I	22	8	-0.31	GEM	0
4-NH_2	29	2	8.53	GEM + 1	0.37
4-NO_2	23	7	3.53	GEM	0
4-OCH_3	67	22	2.02	GEM	0
4-OH	22	14	0.53	GEM	0
4-*t*-butyl	67	4	1.28	GEM	0
3,4-Cl	38	20	-0.23	GEM + 1	0.04
3,4-OCH_3	43	6	6.98	GEM + 1	0.09
3,5-CH_3	22	8	-0.18	GEM	0
3,5-Cl	26	4	-1.45	GEM	0

a) Within 5 kcal above GEM. b) For CoMFA study I. c) "State" refers to the relationship of the conformer selected for superposition to the GEM of analogue. d) Energy of selected conformer relative to GEM of analogue.

Table 3.1 lists the number of unique conformers (as defined from the RMS superposition threshold of 0.2 Å) found by RANDOMSEARCH for each analogue. The table shows that random search of the MP analogues produced a low of 19 conformations (for the 4-Br analogue) and a high of 113 conformations (for the 3-OCH_3 analogue) within just 5 kcal/mol of their respective global energy minima. The large number of conformations found for the methoxy-substituted analogues are due to rotation of the methyl group. To get a good predictive CoMFA model, it is necessary to choose for alignment conformers that are representative of the bioactivity of the MP analogues, as discussed in the next section.

3.3 Selection of Bioactive Conformer

Most compounds of biological importance are usually large enough to possess several rotatable bonds that allow the compound to exist in several different low-energy conformations. For example, the most active conformer of a certain analogue may not be the gas-phase global energy minimum, which could be an intuitive first choice and was used by several researchers.[76,77,93] Others have used a higher-energy conformation, up to 12 kcal/mol above the GEM conformation, as the bioactive conformation.[112] However, it should be noted that there may be several reasons for selecting conformers not based on Boltzmann distribution and conformational energies. In this context, according to Yliniemela et al.,[113] first, the molecular mechanical or semi-empirical conformational energies may not be very accurate. Second, the selected low-energy conformer may not properly reflect solvent and physiological effects. Third, if its energy is not too high above the GEM, even a non-optimal conformer may have a significant presence. Recent

CoMFA investigators, such as Debnath,[87] have used their knowledge of molecular structure and properties of receptor sites in identifying suitable conformers of their ligand analogues. Using this knowledge, they were able to restrict their choice of conformers to only those that could possibly bind with the receptor. However, in the absence of knowledge of the receptor site structure, as is the case here, the problem is more difficult.

3.3.1 Identification of Bioactive Conformer Common Only to Active Analogues

In general for the MP analogues, when all the conformers generated for a particular analogue were superimposed using the piperidine ring and the central chiral carbon atom, roughly three different "classes" of phenyl ring orientation in space were identified. This is shown for MP in Figures 3.4 and 3.5.

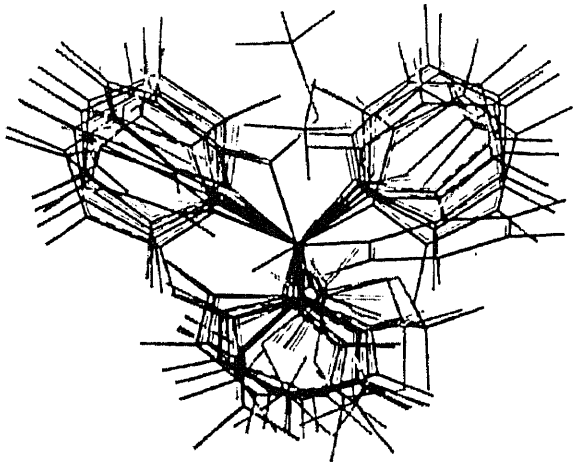


Figure 3.4 Superposition of RANDOMSEARCH Conformers of Methylphenidate

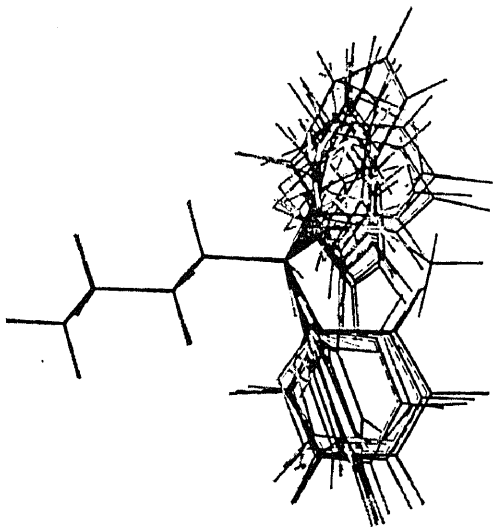


Figure 3.5 The Same Superposition Rotated to Show the Near-Perfect Alignment of Piperidine Rings of the Conformers

Theoretically, if the bioactive conformer for MP were within 5 kcal/mol of MP's GEM conformation, it would fall into one of the three classes shown in Figure 3.4. If this conformer could be identified, it could be used as a template for the alignment necessary for CoMFA of the 30 analogues. The assumption was made that the 2-position analogues are relatively inactive (with high IC_{50} values) because they *cannot* easily attain the bioactive conformation, whereas the 3- and 4-position analogues are active because they can easily attain the bioactive conformer (i.e. that the bioactive conformer would be found among the local minima located by RANDOMSEARCH). Therefore, an analysis of the four randomly-altered torsional angles was carried out for all conformations of MP and the 2-, 3-, and 4-substituted Br, Cl, F, OCH_3, and OH analogues in Table 3.1. For each analogue, scatter graphs (see Appendix C) of $\phi 1$ vs. $\phi 2$, $\phi 1$ vs. $\phi 3$, and $\phi 2$ vs. $\phi 3$ were created (see Figure 3.3 for torsion angle definitions). The aim was to establish whether there was a region or point in torsion space (i.e. within 5 kcal/mol of each

analogue's GEM conformation) that was accessible to those analogues with low IC_{50} values (i.e. MP and the 3- and 4-analogues) but not to those analogues with high IC_{50} values (i.e. the 2-analogues). The search for the presence in the scatter graphs of MP and the 3- and 4-analogues and absence in those of the 2-analogues of a conformer/point was initially based on the reasoning that the 2-substituent may have poorer steric interactions with the rest of the molecule costing it more than 5 kcal/mol to attain the bioactive conformation. As such, it was useful to look first at the $\phi 2$ vs. $\phi 3$ scatter plots for the analogues since this would pit the 2-substituent against the carbonyl oxygen (for greater steric hindrance). As expected from the size and location of the substituent, the 2-analogues appear to have a much more conformationally restricted torsional angle space. From all three sets of graphs, however, no conformation was identifiable that was common to MP and all 3- and 4-substituents but not to the 2-analogues.

Since this result could stem from the possibility that the bioactive conformer might not fall within the 5 kcal/mol energy cutoff, the 30 analogues were subjected to RANDOMSEARCH again using a higher 20 kcal/mol energy cutoff value. This time the new sets of conformations for MP and all the 2-position analogues (Br, Cl, F, OCH_3, and OH) were plotted in 3D with $\phi 1$, $\phi 2$, and $\phi 3$ as the three axes. First, all the local minima for MP within 20 kcal/mol of its GEM conformation were plotted in green. Then all the local minima for *all* the 2-position analogues (within 20 kcal/mol of their respective GEM conformations) were plotted in red. The two plots were superimposed to produce the figure shown in Plate 1 in Appendix D. The actual value of the energy of each conformer is not important here; rather just the fact that each point identifies a local minimum with the 20 kcal/mol cutoff. An attempt was made to locate a green point that

had no red neighbors in 3D torsional angle space. Such a point could signify a bioactive conformation that is allowed for MP but not for the 2-analogues and could be used as a template for alignment of the 30 analogues for CoMFA. Seven such points were identified: MP conformation numbers 27 (energy = 12.72 kcal/mol), 39 (18.85 kcal/mol), 40 (19.82 kcal/mol), 42 (12.19 kcal/mol), 45 (19.95 kcal/mol), 59 (7.5 kcal/mol), and 65 (12.67 kcal/mol). If the 3- and 4-position analogues are active, one would expect them to have the same bioactive conformation as MP. The scatter plots for selected 3- and 4-analogues were investigated to see whether a conformation could be found that was common to MP and 3- and 4-position analogues but not found for the 2-position analogues. The results, as shown in Table 3.2, convey that no consistent pattern emerged.

Table 3.2 Trial Bioactive Conformations of MP Analogues

Analogue	IC_{50}	MP conformation number						
		27	39	40	42	45	59	65
3-Br	4.2	N	Y	Y	Y	Y	Y	Y
3-F	40	N	Y	Y	Y	Y	N	Y
3-OCH_3	287	Y	N	N	N	N	N	N
4-Br	6.9	Y	N	N	N	N	N	N
4-Cl	20	Y	N	N	N	N	N	N

Y = yes, analogue has MP conformation; N = no

3.3.2 Identification of Bioactive Conformer Based on Interfeature Distances in WIN 35,428

As a next step, it was felt that the bioactive conformer of the MP analogues might be defined from that of the highly active cocaine analogue WIN 35,428 used by Schweri and coworkers to radio-label the cocaine binding site. Subsequently, RANDOMSEARCH was performed on WIN (Figure 3.6) and conformers within 20 kcal/mol of the GEM of WIN

were investigated. The distances between the nitrogen and the aromatic centroid (represented by the 2- or 5-position hydrogen, H1 and H2 respectively), the nitrogen and the two oxygens (O1: carbonyl oxygen and O2: methoxy oxygen), and the oxygens and the centroid were calculated for each of the WIN conformers. These distances along with the energy values of the 16 RANDOMSEARCH WIN conformers are given in Table 3.3, with the WIN GEM conformation (number 4) listed in bold. The distance calculation was repeated for MP and some of the 3- and 4-analogues and the distances and energies found for the 67 RANDOMSEARCH MP conformers are given in Table 3.4, with the MP GEM conformation (number 14) values listed in bold. The idea was to see if any conformer of MP, 3-, and 4-analogues had the same distances between relevant features as one or more conformers of WIN since this conformer might be bioactive (and therefore suitable to use for CoMFA alignment). This method of identifying the bioactive conformer for MP analogues was also not successful, primarily because of the different size of the MP analogues and WIN, with the MP analogues being much "shorter" than WIN in all the inter-feature distances.

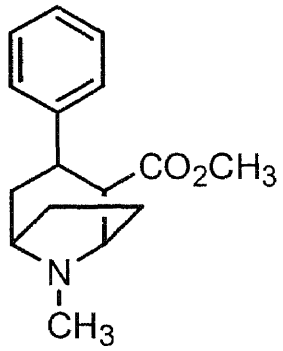


Figure 3.6 WIN

Table 3.3 Energy and Interfeature Distances for WIN Conformers

Conformer	Energy	N_H1	N_H2	O1_H1	O1_H2	O2_H1	O2_H2	N_O1	N_O2
1	12.65	4.66	5.30	3.04	4.15	5.14	4.45	3.33	3.55
2	12.57	4.20	5.59	3.11	5.58	2.83	4.19	2.78	4.29
3	12.93	4.60	5.34	4.20	5.01	3.54	3.42	2.76	4.23
4	**11.69**	**5.56**	**4.25**	**4.63**	**3.46**	**5.30**	**2.55**	**4.09**	**2.68**
5	20.38	4.58	5.34	2.84	4.47	4.72	4.47	3.07	3.89
6	19.60	4.16	5.62	2.76	5.61	2.86	4.77	2.64	4.29
7	26.54	5.18	4.80	2.80	5.56	3.90	4.66	4.36	3.62
8	12.19	4.67	5.27	4.73	3.61	3.49	4.54	4.07	2.80
9	14.01	5.13	4.94	4.43	2.91	6.14	3.18	3.25	3.54
10	22.00	5.17	4.93	4.59	3.02	6.14	2.94	3.03	3.80
11	24.83	5.63	4.12	4.18	3.38	5.44	2.66	4.52	3.55
12	26.11	5.30	4.65	2.92	4.53	4.55	4.10	4.51	3.55
13	21.11	4.36	5.48	2.57	4.99	3.86	4.66	2.77	4.10
14	15.04	5.32	4.70	5.98	2.66	5.40	2.69	3.91	2.68
15	14.01	4.32	5.52	2.52	5.10	3.83	4.62	2.84	4.09
16	14.41	4.65	5.28	3.50	4.58	4.64	3.58	2.82	4.17

Table 3.4 Energy and Interfeature Distances for MP Conformers

Conformer	Energy	N_H1	N_H2	O1_H1	O1_H2	O2_H1	O2_H2	N_O1	N_O2
1	7.00	4.57	2.85	3.04	5.12	3.05	4.63	4.22	5.00
2	8.66	4.32	3.28	5.10	2.89	4.68	3.05	2.74	4.35
3	5.33	4.40	4.66	2.98	5.07	3.39	4.35	2.82	4.21
4	11.22	4.22	5.05	5.02	3.28	4.02	3.77	3.21	3.74
5	16.90	5.31	2.95	3.30	4.80	2.64	5.20	4.17	4.60
6	12.32	2.44	4.78	5.10	3.47	4.36	3.47	4.46	4.76
7	11.18	4.52	3.20	3.16	5.03	2.62	5.09	4.47	3.60
8	15.14	4.95	4.41	3.41	4.45	2.60	5.22	3.70	3.58
9	1.76	2.87	4.54	5.13	3.22	4.39	2.77	4.26	4.90
10	2.86	2.97	4.40	4.48	2.53	5.11	3.67	4.87	4.22
11	3.19	4.74	2.77	3.34	4.60	2.60	5.10	4.72	4.35
12	2.84	4.21	3.45	4.70	3.39	5.05	2.50	4.22	2.68
13	1.40	4.79	4.35	4.05	4.04	4.88	2.54	3.75	3.40
14	**0.54**	**4.70**	**4.40**	**5.06**	**3.13**	**4.03**	**3.22**	**2.83**	**4.25**
15	3.45	4.34	3.31	5.12	2.96	4.45	2.94	2.81	4.31
16	4.57	2.35	4.98	4.40	2.70	5.20	3.55	4.63	4.37
17	1.31	4.37	4.73	2.70	4.24	3.94	4.90	4.21	2.69
18	4.25	2.35	5.02	5.21	3.19	4.34	2.83	4.50	4.54
19	11.62	5.10	2.35	2.61	4.53	3.36	5.37	4.54	4.58
20	5.43	4.87	2.95	2.96	4.23	3.30	5.21	3.00	3.76
21	4.23	4.85	2.98	2.99	5.17	3.16	4.16	3.86	2.74
22	4.28	4.61	3.15	2.95	5.20	2.85	4.53	4.03	2.76
23	5.70	3.25	5.23	4.22	4.03	2.53	4.72	3.66	2.67
24	12.02	4.33	3.08	2.32	4.81	4.24	4.89	4.62	4.62
25	6.61	3.37	5.20	2.65	4.32	4.21	4.66	3.47	2.89
26	7.09	3.35	4.98	2.56	5.01	3.73	4.22	2.53	4.07
27	12.72	2.67	4.32	2.89	4.04	4.54	4.34	4.27	4.74
28	11.69	4.91	3.62	2.84	5.10	3.22	4.51	3.95	4.72
29	12.80	2.98	4.80	4.24	3.08	5.33	3.02	2.74	4.02
30	14.07	5.26	3.22	4.46	2.53	4.33	4.39	2.91	3.53
31	16.13	2.65	4.42	2.41	4.65	4.27	4.47	3.67	4.73
32	14.43	5.13	2.95	3.39	5.13	3.46	4.39	3.79	3.37
33	6.48	4.68	4.05	4.69	3.50	5.04	2.42	3.73	2.55
34	1.62	4.72	4.38	4.09	3.12	5.11	3.25	4.20	2.72
35	4.51	4.32	2.71	3.93	4.50	4.30	2.80	4.72	4.18
36	6.16	4.25	2.79	4.07	2.97	4.35	4.48	4.40	4.54
37	7.86	2.67	4.32	2.52	4.43	4.28	4.49	4.02	4.53
38	4.30	2.35	5.14	5.20	3.17	4.64	2.60	4.47	4.53
39	18.85	2.50	4.94	4.89	3.30	3.76	4.21	4.71	4.12
40	19.82	2.65	4.96	4.72	3.35	3.66	4.55	4.71	3.91
41	12.97	3.21	4.55	4.59	2.82	5.34	2.96	2.75	4.19
42	12.19	5.15	2.44	2.43	4.86	3.40	5.21	4.44	4.64
43	6.34	3.66	4.90	5.11	2.93	4.24	3.12	3.95	4.64
44	6.32	5.24	2.97	3.75	3.34	4.20	4.65	3.27	3.10
45	19.95	4.26	5.17	5.05	2.97	5.14	2.86	3.39	3.69
46	6.93	3.28	5.21	2.55	4.68	4.30	4.12	2.76	3.64
47	6.41	3.62	4.95	4.49	3.58	5.06	2.46	4.30	4.15
48	15.59	4.24	4.83	4.69	2.55	5.32	3.22	2.87	3.42
49	13.01	3.61	3.56	2.41	4.57	4.28	4.76	4.46	4.77
50	5.62	2.98	5.25	4.56	3.74	2.90	4.30	3.60	2.68
51	5.51	3.21	4.63	2.35	4.78	4.24	4.77	3.35	3.67
52	7.69	2.86	5.14	4.44	4.28	4.59	2.50	2.62	4.14
53	12.94	3.48	4.91	2.79	5.08	3.37	4.42	2.46	4.19
54	18.55	3.76	5.29	4.66	3.51	5.18	2.47	3.76	4.27
55	7.45	2.63	4.39	4.24	4.17	2.46	4.64	4.68	3.79
56	5.04	2.37	5.17	4.81	2.47	5.16	3.41	4.54	4.41
57	7.26	3.64	4.83	4.53	2.46	4.94	4.01	4.73	3.62
58	9.12	4.89	4.28	4.54	2.52	4.32	4.37	3.65	3.50
59	7.51	2.89	5.13	4.65	4.17	4.20	2.69	2.82	3.87
60	16.15	3.15	3.86	2.78	5.04	3.53	4.24	3.96	4.98
61	14.07	2.94	5.24	3.02	4.14	4.73	3.78	2.69	3.71
62	8.75	4.23	4.90	4.74	2.91	5.16	2.77	2.68	3.67
63	7.59	4.20	4.98	5.18	2.75	4.59	3.01	3.52	2.78
64	13.17	3.38	3.78	4.46	2.41	4.89	4.18	4.62	4.74
65	12.67	2.45	4.61	3.41	3.61	4.92	3.81	4.34	4.77
66	11.78	2.35	4.95	4.21	2.90	5.32	3.33	4.50	4.68
67	5.81	3.23	5.03	4.05	4.33	2.43	4.92	3.79	2.60

3.3.3 Identification of Bioactive Conformer by Cluster Analysis

An alternative definition of the different classes of the MP conformers was attempted by using hierarchical cluster analysis on the MP conformers obtained by RANDOMSEARCH. Hierarchical cluster analysis is a means to locate various groupings that might be present in a given data set. It can be used to see which structures are most similar in terms of the selected distance metric, as well as to choose a compound or a subset of compounds that could be considered representatives of the larger data set. The cluster analysis results are displayed as an inverted tree, or a dendrogram. Based on the selected criterion for clustering (such as torsional angles, distances between atoms or features, etc.), the analysis moves from the bottom of the dendrogram to the top, progressively merging pairs of clusters. Each node at the bottom represents a row (or conformer), while the top-most cluster represents the entire table (such as the molecular spreadsheet containing the different RANDOMSEARCH conformers of an analogue). The lengths of the vertical lines in the dendrogram provide qualitative information about the separation between various clusters with a shorter branch signifying greater similarity between the clusters. Clusters represented by long unbranched strands are strongly separated from other clusters. A dendrogram that has well-formed clusters has most of the original nodes merging at its bottom.

The cluster analysis was done on the 67 MP conformers generated by using the energy cutoff value of 20 kcal/mol during RANDOMSEARCH. The four torsional angles $\phi 1$, $\phi 2$, $\phi 3$, and $\phi 4$ shown in Figure 3.3 were chosen as the criterion for clustering. The resulting dendrogram is illustrated in Figure 3.7 and the torsional angles in Table 3.5. One way to select subsets is to select all points at a particular vertical level and then

choose one member from each resulting group. Accordingly, five points were chosen near the top of the dendrogram in order to divide the 67 MP conformations into five families, as shown in Figure 3.7. The lowest energy structure from each of the five families was then selected as a potential bioactive conformer with which five different CoMFA experiments could be carried out. The conformer chosen from the first family was conformer number 1 (with energy = 7.00 kcal/mol), from the second was conformer number 3 (5.33 kcal/mol), from the third was conformer number 14 (0.54 kcal/mol), from the fourth was conformer number 17 (1.31 kcal/mol), and from the fifth was conformer number 20 (5.43 kcal/mol). Of these, conformer number 14 was the GEM structure for the 67 MP conformations. This structure (energy = 0.54 kcal/mol, $\phi 1$ = 175.6, $\phi 2$ = 74.1, $\phi 3$ = 3.6, $\phi 4$ = 179.3) was very similar to the GEM structure (energy = 0.53 kcal/mol, $\phi 1$ = 177.1, $\phi 2$ = 74.6, $\phi 3$ = 3.2, $\phi 4$ = 178.9) of the 20 MP conformers obtained earlier using the 5 kcal/mol RANDOMSEARCH energy cutoff value. By the time cluster analysis was performed on the 20 kcal/mol RANDOMSEARCH MP conformers, CoMFA Study I had already been carried out using the 5 kcal/mol MP GEM structure as the template for alignment (which is described in the next section). Given the similarity between the two MP GEM structures, it was decided to leave out conformer number 14 from the CoMFA studies. In all, five CoMFA studies (using the 5 kcal/mol cutoff GEM conformer, and conformer numbers 1, 3, 17, and 20 from the 20 kcal/mol cutoff data set) were done and are described in the next chapter.

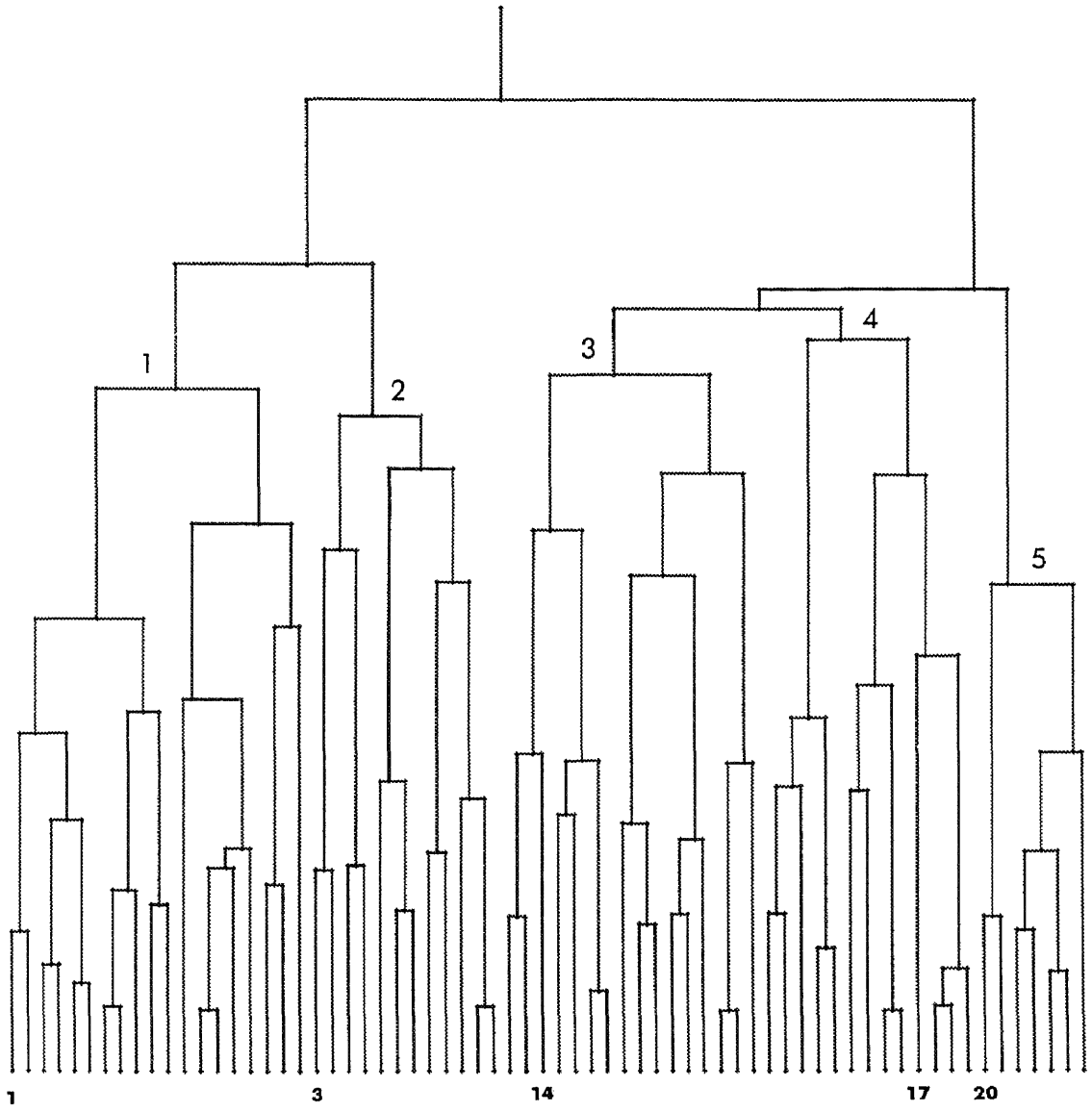


Figure 3.7 Dendrogram for MP Conformers (20 kcal/mol Cutoff)

Table 3.5 Torsional Angles for MP Conformers

Conformer	Torsion 1	Torsion 2	Torsion 3	Torsion 4
1	67.40	-85.00	-6.00	-2.50
2	-69.00	98.40	-13.00	-1.70
3	176.00	-103.10	-5.10	2.90
4	179.10	-157.60	98.00	-10.10
5	99.70	-52.30	112.70	7.30
6	40.20	-165.40	81.50	-12.40
7	-50.50	-39.10	108.50	11.80
8	173.40	-74.60	127.10	11.20
9	63.90	94.70	5.80	179.00
10	58.50	90.70	-151.50	-179.50
11	72.40	-79.30	142.20	-179.20
12	-66.80	103.50	134.40	179.40
13	172.30	63.30	133.80	-179.80
14	**175.60**	**74.10**	**3.60**	**179.30**
15	-70.20	98.90	-9.40	178.60
16	54.90	172.10	-105.40	-179.00
17	171.40	-111.90	-130.20	-180.00
18	58.90	167.00	92.00	-179.90
19	65.90	-24.00	-89.30	7.40
20	-72.70	-5.70	-89.40	-179.10
21	-70.40	-9.10	106.20	179.50
22	-60.60	-27.90	111.10	179.10
23	-119.10	-121.70	122.20	179.70
24	58.30	-93.80	-92.80	8.20
25	-124.10	-115.90	-110.80	-178.50
26	-109.00	-99.20	-35.50	179.30
27	17.50	-132.40	-97.70	4.50
28	116.40	-89.50	-15.60	-0.80
29	-68.00	167.70	-71.00	5.20
30	-118.90	54.80	-73.70	0.50
31	-22.70	-112.50	-74.70	4.50
32	-84.80	13.20	85.10	-19.00
33	-123.40	102.80	129.30	178.80
34	176.00	73.40	-170.90	-179.70
35	19.60	47.10	101.10	179.60
36	18.20	51.10	-106.50	-178.90
37	-14.90	-113.50	-98.30	-178.40
38	69.30	149.50	99.40	-179.50
39	-26.30	-136.50	101.70	-12.30
40	-42.80	-125.10	113.00	-16.90
41	-59.40	148.70	-63.60	2.40
42	76.50	-53.50	-68.30	-0.40
43	116.70	89.30	-10.30	178.90
44	-102.10	37.00	-109.10	-178.20
45	157.40	135.90	113.60	-29.00
46	-120.40	-119.40	-62.80	-178.80
47	116.20	91.10	134.30	-179.20
48	-131.40	141.70	-71.10	-1.50
49	29.40	-105.60	-97.80	6.90
50	-103.40	-140.20	111.20	179.10
51	-81.00	-90.30	-104.10	-178.90
52	-85.80	-161.60	3.90	179.90
53	-113.20	-94.10	-15.10	1.20
54	133.40	114.80	110.80	18.90
55	-13.80	-116.70	107.00	-179.90
56	72.00	139.50	-80.80	-179.70
57	110.90	85.10	-131.10	179.90
58	174.30	55.60	-70.30	2.10
59	-89.60	-160.30	32.50	-179.80
60	-7.70	-98.50	-18.60	2.90
61	-100.70	-144.80	-68.90	3.40
62	-136.30	146.50	-51.90	-179.40
63	-139.50	152.70	123.00	-179.30
64	33.20	79.30	-113.00	-8.70
65	28.30	-152.70	-93.30	4.70
66	53.50	174.30	-90.50	8.30
67	-103.10	-104.70	120.60	179.40

3.4 Alignment

Molecular alignment plays an essential role in areas such as 3D-QSAR, pharmacophore identification, molecular diversity, and structural database searching.[114,115] The results of CoMFA also are strongly dependent on the alignment of the analogues.[74] The structures for the first CoMFA study were aligned by first choosing the GEM structure of unsubstituted MP (from the 5 kcal/mol data set) as the template and then superimposing the GEM structure of each analogue on the template by superimposing the six aromatic carbons and the piperidine nitrogen. The alignment was carried out in SYBYL using its "Align Database" option, which superimposes the structures by minimizing the RMS distance between each aligned analogue and the template structure (see Appendix B). As shown in Figure 3.8, this superposition produced an alignment in which the piperidine ring, the ester side chain, and the phenyl ring overlap almost exactly for all analogues except 2-OH, 3-NH_2, and 4-NH_2. In addition, in this alignment, the 2-position substituents on the phenyl ring lie on the opposite side of the 3-position substituents, essentially turning the 3-position substituents into 5-position substituents relative to the 2-position substituents. For a coherent analysis, it was decided to have, as a requirement for the alignment for CoMFA, all analogues with substituents on the same side as the 2-position analogues. Table 3.1 lists for each analogue the conformer number selected for the new alignment and also its relationship to the GEM structure of the analogue. For example, for the unsubstituted MP analogue, 20 conformers were found that were within 5 kcal/mol of the GEM of this analogue. The GEM conformer of this analogue turned out to be the sixth conformer located by RANDOMSEARCH and its energy was 0.53 kcal/mol. The "state" of a conformer selected for superposition alludes to its relationship

to the GEM of the particular analogue, in terms of increasing relative energy: GEM < GEM+1 < GEM+2, and so on. This alignment is displayed in Figure 3.9 and was used for all subsequent experiments in the first CoMFA study.

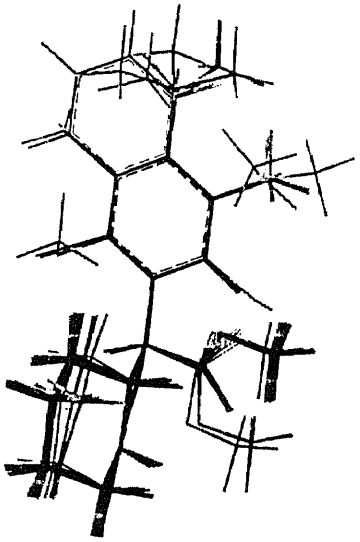


Figure 3.8 Alignment of All GEM Structures

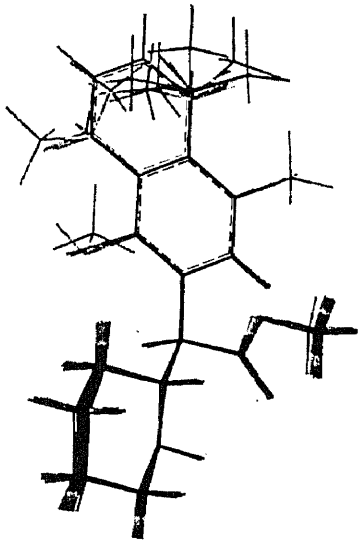


Figure 3.9 Alignment #1

For each of the other four CoMFA studies, the template MP conformer was used to construct (using SYBYL's Build/Edit option) all other analogues by adding substituents at the right position on the phenyl ring. Since the template structure was kept unchanged during construction of the other analogues, and since no energy minimizations were performed on the constructed structures (see notes in Appendix B), the process produced automatic alignments for each of these four CoMFA studies. These alignments (Alignments #2, #3, #4, and #5 for the corresponding CoMFA Studies II – V) are shown

in Figures 3.10 – 3.13. After these preparatory steps, the CoMFA procedure described in Chapter 2 was carried out on the 30 MP analogues.

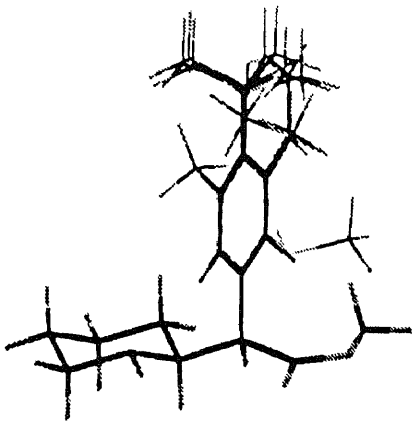


Figure 3.10 Alignment #2

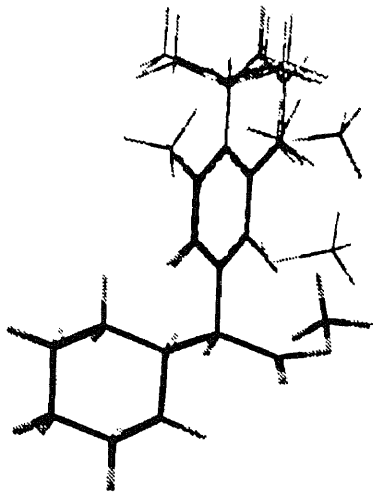


Figure 3.11 Alignment #3

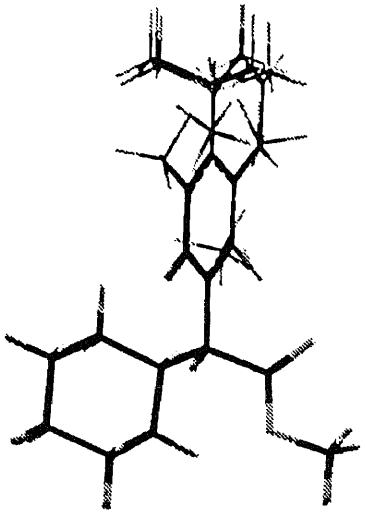


Figure 3.12 Alignment #4

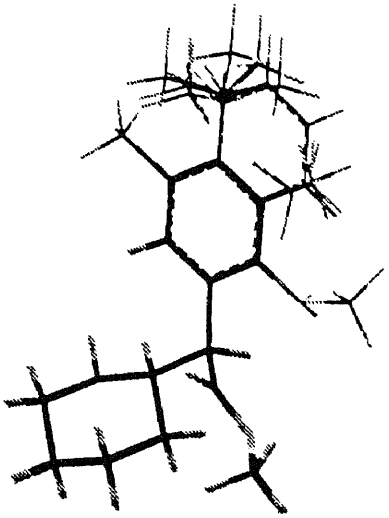


Figure 3.13 Alignment #5

The results of a CoMFA run include residual values, defined as the (Actual – Predicted) value of (-log IC_{50}), for each analogue. As discussed in Chapter 4, if warranted by the quality of the experimental data, the predictability of the CoMFA model can be improved (i.e. the q^2 can be raised) by deleting from the analysis compounds with large residuals. Chapter 4 also describes the sensitivity of the CoMFA models to various run-time parameters such as the region box, grid spacing, steric and electrostatic energy cutoff values, and the column filtering (σ) value. The CoMFA region box is the 3D space surrounding the alignment of molecules and contains lattice points at which the steric and electrostatic fields are calculated using a probe atom, such as a sp^3 carbon atom with +1 charge. Grid spacing is the distance between two lattice points and is 2.00 Å by default. During the calculation of field energies at the lattice points, some values may exceed the selected cutoffs for the steric and electrostatic energy. These are replaced by the chosen energy cutoff values, thus making a plateau of the steric fields close to the center of any atom. The column filtering value, σ, is the minimum level of energy variation allowed for a lattice point to be included in the PLS analysis; it is 2.0 kcal/mol by default.

CHAPTER 4

RESULTS AND DISCUSSION

The CoMFA experiments in the studies presented here use the (published[61,67] and unpublished) nanomolar binding and uptake IC_{50} values for MP analogues of Schweri and coworkers. The PLS analyses were performed using the [-log (molar IC_{50})] values. Five CoMFA studies (I-V) were carried out. CoMFA Study I used the 5 kcal/mol cutoff GEM conformer as the template for the alignment rule. CoMFA Studies II-V used conformer numbers 1, 3, 17, and 20, obtained after cluster analysis of the 20 kcal/mol cutoff data set, as the templates for the respective alignment rules.

4.1 CoMFA Study I

Three sets of experiments were carried out: a) separate CoMFA runs using binding and uptake data with default grid size and spacing; b) changing the grid spacing and CoMFA region box size; and c) changing the steric and electrostatic (ES) cutoffs and the column filtering (σ) parameter.

4.1.1 CoMFA Using Default Grid Size and Spacing

Table 4.1 shows the automatically generated default specifications for the CoMFA region box for the particular alignment used in this study. The x, y, and z coordinates are given in Angstroms. Automatically-generated region boxes extend by an extra 2 Å in all directions around the aligned structures, thus effectively covering the entire alignment with sufficient spatial margin. The steric and ES field points were calculated at each of

the grid points of each of the analogues using the default 2 Å grid spacing. The Leave-One-Out cross-validation method was used to calculate q^2, advancing through all 30 analogues and leaving one analogue at a time out of the calculation to predict its (-log IC_{50}) value. The optimum number of components was then used to carry out a non-validated CoMFA run. The results are summarized in Table 4.2 for both cross-validated and non-validated runs. In the table, "both" refers to inclusion of both steric and ES fields in the calculations. These CoMFA runs produced barely significant q^2 values of about 0.5 using binding or uptake data for all 30 analogues (Runs 1 and 3 in Table 4.2), indicating an acceptable model. The optimum number of components for both runs was four and the non-validated r^2 was about 0.8-0.9. A comparison of the steric and electrostatic contributions (about 80% and 20% respectively) for the two runs points to a preponderance of steric influences. The similarity of the results obtained for binding and uptake models is probably a consequence of the fact that the binding and uptake data are highly correlated.[61]

Table 4.1 Default Region Box

# of lattice points:	800	# of Boxes:	1
		Box Center (x, y, z):	(0.8, 1.5, -0.3)
	X	Y	Z
Lower Corner (Å):	-8.79	-7.61	-8.27
High Corner (Å):	10.33	10.52	7.68
Step size (Å):	2.0	2.0	2.0
# of steps:	10	10	8
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

Table 4.2 Selected CoMFA Results for Binding and Uptake Data

Run				Cross Validation		No Validation			
#	Data	Type of Field(s)	# of Analogues	Optimum # of Components	q^2	Standard Error of Estimate	% Steric	% ES	r^2
1	Binding	Both	30	4	0.477	0.419	78.9	21.1	0.885
2	Binding	Both	28	5	0.590	0.299	74.8	25.2	0.934
3	Uptake	Both	30	4	0.494	0.388	79.9	20.1	0.865
4	Uptake	Both	29	6	0.444	0.267	75.6	24.4	0.942
5	Binding	Steric	30	4	0.470				
6	Binding	ES	30	2	0.281				
7	Uptake	Steric	30	3	0.444				
8	Uptake	ES	30	2	0.301				

Runs 2 and 4 were made with only 28 and 29 analogues, respectively. In Run 2, 4-CF_3 and 2-OCH_3 were dropped; in Run 4, 4-CF_3 was dropped due to high predictive errors (see discussion below and Table 4.3). Run 2 gave the best q^2 (0.590) and associated r^2 (0.934). Runs 5-8 test the use of steric fields only in modeling the binding and uptake data (Runs 5 and 7) versus only ES fields (Runs 6 and 8) for all 30 analogues. Since Runs 1 and 3 showed that steric effects contribute about 80% to the binding and uptake data, it is not surprising that Runs 6 and 8, based on ES fields only, show very poor values of q^2 (around 0.3). Similarly, Runs 5 and 7, based on steric fields alone, give q^2 values close to Runs 1 and 3, which are based on both steric and ES fields. Non-validated runs were not performed for Runs 5-8 due to low value of q^2.

Table 4.3 Actual Binding and Uptake Data and CoMFA Residuals

		Actual Data *				Binding Residuals		Uptake Residuals		Binding Residuals		Uptake Residuals	
		Binding		Uptake		*With all 30 analogues*				*Without 4-CF_3, 2-OCH_3*		*Without 4-CF_3*	
#	Substituent	IC_{50}	- log IC_{50}	IC_{50}	- log IC_{50}	Cross Validated	Non Validated	Cross Validated	Non Validated	Cross Validated	Non Validated	Cross Validated	Non Validated
1	MP	83.0	7.081	224.0	6.650	-0.103	0.045	0.099	0.087	-0.035	0.105	0.174	0.199
2	2-Br	1865.0	5.729	3410.0	5.467	0.860	0.436	0.597	0.339	0.638	0.490	0.609	0.461
3	2-Cl	1946.7	5.711	2660.0	5.575	0.519	0.221	0.709	0.323	0.380	0.167	0.703	0.261
4	2-F	1415.0	5.849	2900.0	5.538	-0.457	-0.399	-0.624	-0.379	-0.283	-0.097	-0.408	-0.113
5	2-naphthyl	11.0	7.959	53.0	7.276	0.578	-0.034	-0.562	-0.087	0.383	0.254	-0.107	0.161
6	2-OCH3	100666.7	3.997	81000.0	4.092	-1.789	0.208	-1.609	0.204			-1.738	-0.157
7	2-OH	23050.0	4.637	35750.0	4.447	-1.065	-0.779	-1.151	-0.757	-1.094	-0.420	-1.239	-0.420
8	3-Br	4.2	8.377	12.8	7.893	0.466	0.455	0.622	0.489	0.532	0.431	0.516	0.411
9	3-CH3	21.4	7.670	100.0	7.000	0.014	0.062	-0.888	-0.006	-0.580	-0.024	-0.675	-0.063
10	3-Cl	5.1	8.292	23.0	7.638	0.837	0.489	0.655	0.322	0.516	0.318	0.258	0.151
11	3-F	40.5	7.392	160.0	6.796	-0.177	0.110	-0.397	0.041	-0.384	-0.173	-0.376	-0.187
12	3-NH2	265.0	6.577	577.5	6.238	-0.805	-0.177	-0.089	-0.142	0.320	0.026	0.918	0.026
13	3-OCH3	287.5	6.541	635.0	6.197	-0.378	-0.244	0.764	-0.124	0.764	0.172	0.678	0.228
14	3-OH	321.0	6.493	790.0	6.102	-1.548	-0.979	-1.632	-0.904	-1.374	-0.805	-1.296	-0.653
15	3,4-Cl	5.3	8.276	7.0	8.155	-0.136	-0.219	0.498	0.105	0.062	-0.148	0.414	0.052
16	3,4-OCH3	810.0	6.092	1760.0	5.754	-1.082	-0.265	0.080	-0.221	-1.115	0.110	-0.553	-0.027
17	3,5-CH3	4685	5.329	3041.5	5.517	-1.051	0.005	-0.589	0.208	-1.297	-0.074	-0.918	0.001
18	3,5-Cl	65.6	7.183	113.0	6.947	0.378	0.163	0.021	0.108	0.288	0.158	0.149	0.045
19	4-Br	6.9	8.161	26.3	7.580	0.314	0.324	0.090	0.301	0.355	0.118	0.341	0.075
20	4-C2H5	736.7	6.133	1580.0	5.801	-0.098	0.073	-0.527	0.044	-0.224	0.017	-0.210	-0.027
21	4-CF3	615.0	6.211	2365.0	5.626	-2.053	-0.973	-2.201	-0.974				
22	4-CH3	33.0	7.481	125.5	6.901	0.950	0.481	0.698	0.395	1.024	0.240	0.987	0.179
23	4-Cl	20.6	7.686	73.8	7.132	-0.334	-0.150	-0.652	-0.187	-0.174	-0.271	-0.418	-0.313
24	4-F	35.0	7.456	142.0	6.848	0.514	0.391	0.509	0.277	0.138	0.279	0.092	0.209
25	4-I	14.0	7.854	64.5	7.190	0.428	0.240	0.378	0.164	-0.500	-0.244	-0.457	-0.268
26	4-NH2	34.5	7.462	114.5	6.941	-0.121	0.290	-1.744	0.266	-1.708	-0.136	-1.661	-0.168
27	4-NO2	493.8	6.306	1610.5	5.793	-0.012	-0.115	1.127	-0.076	0.352	-0.122	0.454	0.054
28	4-OCH3	83.0	7.081	292.5	6.534	0.959	0.225	0.930	0.113	0.503	-0.180	0.365	-0.085
29	4-OH	98.0	7.009	340.0	6.468	-0.261	-0.001	-0.227	-0.074	0.003	-0.015	0.026	0.018
30	4-*t*-butyl	13450.0	4.871	9350.0	5.029	-0.906	0.118	0.137	0.145	-1.386	-0.174	-1.045	-0.048

* The IC_{50} values are shown in nM units and the -log IC_{50} values are based on molar IC_{50} values. [61,67]

Table 4.3 provides the observed binding and uptake data[61,67] for the 30 MP analogues, along with the values of (-log IC_{50}) that were used as the biological activity measurement in the CoMFA analyses. It also lists the residuals, calculated as (Actual – Predicted) value of (-log IC_{50}), generated in these runs. For the binding run (Run 1 in Table 4.2), the analogues with the largest (cross-validated) residuals are: 4-CF_3 (-2.053), 2-OCH_3 (-1.789), and 3-OH (-1.548). For the uptake run (Run 2 in Table 4.2), the analogues with the largest (cross-validated) residuals are 4-CF_3 (-2.201), 2-OCH_3 (-1.609), 3-OH (-1.632), and 4-NH_2 (-1.744). Since the (cross-validated) q^2 is a better indicator of the usefulness of the model than the (non-validated) r^2, one way to improve q^2 is to delete those analogues that are poorly predicted. For this reason, a binding CoMFA run without 4-CF_3 and 2-OCH_3, and an uptake CoMFA run without 4-CF_3 were performed. These results also are shown in Tables 4.2 (Runs 2 and 4) and 4.3 with the blank spaces in Table 4.3 indicating the analogue that was left out. The q^2 value for the binding run (Run 2) increased somewhat to about 0.6 without 4-CF_3 and 2-OCH_3, but that for the uptake run (Run 4) decreased slightly to about 0.4 without 4-CF_3. However, the standard error of estimate decreased for these two runs, while the non-validated r^2 improved to over 0.9. The steric contribution was again favored over the electrostatic contribution. However, some large residuals still remain in the 28- and 29-analogue runs. For the binding run, Run 2, deleting 4-CF_3 and 2-OCH_3 diminished the (cross-validated) residual for 3-OH (from -1.548 to -1.374), but actually increased it for 4-NH_2 (from -0.121 to -1.744), 3,5-CH_3 (from -1.051 to -1.297), and 4-*t*-butyl (from -0.906 to -1.386). Similarly, deleting 4-CF_3 from the uptake run, Run 4, decreased the (cross-validated) residuals for 3-OH (from -1.632 to -1.296) and 4-NH_2 (from -1.744 to -1.661), but

increased those for 2-OCH_3 (from -1.609 to -1.738) and 2-OH (from -1.151 to -1.239). Since deleting analogues with large residuals from subsequent CoMFA runs did not necessarily improve the predictive power of those runs, and since there is no reason that the experimental data for those analogues with large residuals would be in error, this procedure was not continued in CoMFA Studies II – V.

The CoMFA steric and electrostatic coefficient contour maps based on the above PLS analyses are provided as 3D plots in Plate 2 in Appendix D. The contours were generated with the following settings: green 80, yellow 35, blue 95, and red 20. For clarity, only one structure, that of the 2-OCH_3 analogue, is depicted in the plots. These coefficient contour plots may help to identify important regions where any change in the steric and/or electrostatic fields may affect the bioactivity. It can be seen that there is no substantive difference between the plots based on binding versus uptake data, nor between the plots for studies based on different number of analogues.

The plots are interpreted as follows. The steric contours are indicated by the green and yellow polyhedra while the electrostatic contours are represented by the blue and red polyhedra. The green steric contours indicate that increasing the bulk in those areas may help to increase the binding or uptake activity of the MP analogues. The yellow regions correspond to steric contours which indicate that increasing bulk in those areas may decrease bioactivity. As shown in the plots, there is a large yellow region just off the 2-position of the aromatic ring indicating that increasing bulk in this region may lead to lowering of activity. This is most clearly shown for the displayed 2-OCH_3 analogue which penetrates directly into this disallowed yellow region and has the highest IC_{50} value (about 10^5 nM) among all 30 analogues studied. Other 2-position analogues

also have very high IC_{50} values: 2-Br (1,865 nM), 2-Cl (1,947 nM), 2-F (1,415 nM), and 2-OH (23,050 nM). For the electrostatic coefficient contours, more negative charge in the red regions may increase bioactivity, while more negative charge in the blue regions may decrease it. In this way, the color-coded plots are a useful tool for relating the steric and electrostatic features of the analogues to their biological activity.

4.1.1.1 Predicting Bioactivity of "Novel" Compounds: The whole exercise of CoMFA is undertaken to generate a useful model that can accurately predict the bioactivity of a yet-to-be-synthesized compound. For this purpose, the non-validated CoMFA run with 30 analogues using binding data, both steric and ES fields, default region box, and the GEM superposition (Run 1 in Table 4.2), was tested by predicting the bioactivity of some "novel" compounds that were generated by editing a suitable analogue already aligned on the template. For example, the 2-I structure was made by replacing the chlorine atom of the 2-Cl analogue in the alignment with an iodine atom on screen. The predictions were made using the "Predict Property" option in SYBYL. Table 4.4 gives the predicted (-log IC_{50}) values and the corresponding nanomolar IC_{50} values for these compounds.

Table 4.4 Predicted Bioactivity of "Novel" Compounds

Compound	Predicted -log IC_{50}	IC_{50} (nM)
2-I	4.503	31390
3-I	7.596	25.6
3,4-Br	8.513	3.1
3,5-Br	6.729	187
3,4,5-Br	7.560	27.8
3,4,5-Cl	7.813	15.4
4,5-Cl	7.147	71.3
3-NO_2	7.638	23.0

To test the robustness of the model the predicted IC_{50} values of these "novel" compounds can be evaluated based on the known activity trends of the 30 analogues from Table 4.3. For example, the relatively inactive 2-position analogues have high IC_{50} values and, as noted above, stick out into the sterically unfavorable yellow region in the contour map. The chimeric 2-I analogue is the bulkiest of all the halogen substituted analogues and it is reasonable that it would have the highest IC_{50} value amongst them. The model predicts an IC_{50} of 31,390 nM for 2-I, which is much higher than the IC_{50} values of all 2-position halogens and comparable to that of the 2-OH analogue (23,050 nM). Similarly, the chimeric 3-I analogue (25.4 nM) is predicted to have an IC_{50} value comparable to 3-CH_3 (21.4 nM) and higher than both 3-Br (4.2 nM) and 3-Cl (5.1 nM). The brominated MP analogues consistently have a better activity than do the chlorinated ones: 2-Br (1,865 nM) vs. 2-Cl (1,947 nM), 3-Br (4.2 nM) vs. 3-Cl (5.1 nM), and 4-Br (6.9 nM) vs. 4-Cl (20.6 nM). In keeping with this trend, the 3,4-Br compound (3.1 nM) is predicted to have a lower IC_{50} value than the 3,4-Cl analogue (5.3 nM). Indeed, the 3,4-Br compound may be a good choice to synthesize and provide experimental verification of the model. It has been suggested to Howard Deutsch, Georgia Institute of Technology, as a potential synthetic target. The preceding line of reasoning, however, fails when comparing 3,5-Br (predicted 187 nM) with 3,5-Cl (actual 65.6 nM). No clear correlations can be made for the remaining "novel" compounds except that the order of activity for 3,4,5-Br, 3,4,5-Cl, and 4,5-Cl appears to be correct.

4.1.1.2 Predicting Bioactivity of "Test Set": A standard procedure to test the robustness of the model is to divide the analogues into a training set and a test set, derive

a model based on the training set, and use it to predict the biological activity of the test set. The test set usually contains about 10% of the compounds and is chosen to reflect a wide range of bioactivity. An initial test set was created that consisted of six analogues (20 percent of the total number of analogues studied): 2-Br (1,865 nM), 2-OCH_3 (10^5 nM), 3-CH_3 (21.4 nM), 3,5-Cl (65.6 nM), 4-Br (6.9 nM), and 4-NO_2 (493.8 nM). This CoMFA run was carried out using the binding data, both steric and ES field calculations, the Leave-One-Out cross-validation method, and the default region box (Table 4.1). The cross-validated and non-validated CoMFA results are summarized in Table 4.5 as Test 1A. The actual and predicted (-log IC_{50}) values for the compounds of the test set, along with the corresponding residuals, are provided in Table 4.6 under the column heading "Test Set 1". A second CoMFA run was performed as above but using a user-defined CoMFA region box (see Table 4.16 below). The results of this run are listed in Table 4.5 as Test 1B and in Table 4.7 under the column heading "Test Set 1". Since the q^2 values of both these runs were much less than 0.5 (0.289 for the default region box and 0.293 for the user-defined region box) and therefore not significant, a second test set was created consisting of three analogues (10 percent of the total number of analogues studied). Two more CoMFA runs predicting the (-log IC_{50}) values of the compounds in the second test set were carried out using the default or user-defined region box. These runs are listed in Table 4.5 as Tests 2A and 2B and in Tables 4.6 and 4.7 under the column heading "Test Set 2". Though Test Set 2 produced improved q^2 values over those produced by Test Set 1, they were still below the 0.5 significance level (0.365 using the default region box and 0.408 using the user-defined region box). These results indicate that the present CoMFA model based on both steric and ES fields, default region box, and Alignment #1 (Figure

3.9) may not be the best predictive model. Therefore alterations in the box region and the grid spacing were made, as described below.

Table 4.5 CoMFA Using Two Training Sets with Two Different Region Boxes

Test				Cross Validation		No Validation			
#	Size (%)	# of Training Set Analogues	Region Box	Optimum # of Components	q^2	Standard Error of Estimate	% Steric	% ES	r^2
1A	20	24	Default [a]	4	0.289	0.440	74.8	25.2	0.859
1B	20	24	User Defined [b]	6	0.293	0.343	73.5	26.5	0.924
2A	10	27	Default [a]	4	0.365	0.420	77.2	22.8	0.860
2B	10	27	User Defined [b]	4	0.408	0.448	77.4	22.6	0.841

[a] Table 4.1; [b] Table 4.16

Table 4.6 Bioactivity Predictions of Test Sets Using Default CoMFA Region Box

Region Box: Table 4.1		Test Set 1		Test Set 2	
Analogue	**Actual -logIC_{50}**	**Predicted -logIC_{50}**	**Residual**	**Predicted -logIC_{50}**	**Residual**
2-Br	5.73	5.40	0.33		
2-OCH_3	4.00	5.84	-1.84	5.84	-1.84
3-CH_3	7.67	7.64	0.03	7.55	0.12
3,5-Cl	7.18	6.84	0.34		
4-Br	8.16	7.66	0.50	7.67	0.49
4-NO_2	6.31	6.61	-0.30		

Table 4.7 Bioactivity Predictions of Test Sets Using User-Defined CoMFA Region Box

Region Box: Table 4.16		Test Set 1		Test Set 2	
Analogue	**Actual -logIC_{50}**	**Predicted -logIC_{50}**	**Residual**	**Predicted -logIC_{50}**	**Residual**
2-Br	5.73	4.96	0.77		
2-OCH_3	4.00	4.92	-0.92	5.39	-1.39
3-CH_3	7.67	6.87	0.80	6.85	0.82
3,5-Cl	7.18	7.12	0.06		
4-Br	8.16	7.57	0.59	7.63	0.53
4-NO_2	6.31	6.97	-0.66		

4.1.2 Resizing CoMFA Region Box and Grid Spacing

Cross-validated CoMFA runs were carried out after altering the size of the CoMFA region box surrounding the 30 aligned structures and changing the distance between lattice points in the x, y, and z directions. The binding data and both steric and electrostatic field calculations were used for these analyses. Three sets of three (for a total of nine) CoMFA runs were performed as shown in Table 4.8. Three grid spacings were used at 1.00 Å, 1.50 Å, and 2.00 Å. For each grid spacing, three sizes of the CoMFA region box were used: "base", "base + 2 Å", and "base – 2 Å". The "base" size signifies the minimum whole number end-coordinates beyond the coordinates of the automatically generated default CoMFA region box (Table 4.1) that would ensure an exact "fit" for the number of steps with a particular grid spacing. For example, in Table 4.10, the selection of the low and high x-coordinates as -9.0 and 11.0 respectively ensures that the 1.00 Å grid spacing will cover the x-axis in exactly 21 steps. Addition or subtraction of 2 Å to/from the low and high coordinates of the CoMFA region box yields the "base + 2 Å" or the "base – 2 Å" region boxes. The results are summarized in Table 4.8 while the residuals are listed in Table 4.9. The nine region boxes generated are shown in Tables 4.10 through 4.18.

Resizing of the region box for a given set of CoMFA runs (i.e., at a given grid spacing) produced virtually no change in the cross-validated q^2 of each run of the set. For example, for the 1.00 Å grid spacing, the "base", "base + 2 Å", and "base – 2 Å" region boxes, all gave a q^2 value of 0.57. Upon varying the grid spacing, however, a minor change in q^2 was effected, with the 1.00 Å grid spacing producing the highest q^2. Nevertheless, the overall change for the 1.00 Å, 1.50 Å, and 2.00 Å grid spacings is barely

significant. Consequently, it can be surmised that changing the grid spacing for this particular alignment had negligible effect on the predictability of the models.

It can be seen from Table 4.8 that the q^2 values for the 1.50 Å grid spacing were the lowest at 0.43, while those for the 1.00 Å and 2.00 Å were similar at 0.57 and 0.55 respectively. Perhaps this is because the center of the CoMFA region box for the 1.50 Å grid spacing (at 1.5, 2.5, 0.0, see Table 4.13) is different from that for the 1.00 Å and 2.00 Å grid spacings (at 1.0, 2.0, 0.0, see Table 4.10). Not enough data is available to corroborate this finding.

Since changing the box region and grid spacing had essentially no effect on the predictive quality of the CoMFA model, the default values for these were used in the remaining CoMFA studies in this section, as well as for CoMFA studies II – V.

Table 4.8 Cross-Validated CoMFA Results of Region Box and Grid Spacing Resizing

Test				Cross Validation	
#	Grid Spacing	Region Box Name	Region Box Table #	Optimum # of Components	q^2
1a	1.00	Base	4.9	6	0.572
1b	1.00	Base + 2 Å	4.10	6	0.572
1c	1.00	Base – 2 Å	4.11	6	0.571
2a	1.50	Base	4.12	3	0.433
2b	1.50	Base + 1.5 Å	4.13	3	0.433
2c	1.50	Base – 1.5 Å	4.14	5	0.433
3a	2.00	Base	4.15	5	0.545
3b	2.00	Base + 2 Å	4.16	5	0.545
3c	2.00	Base – 2 Å	4.17	5	0.547

Table 4.9 Binding Residuals upon Changing Grid Spacing and Region-Box Size

		Grid Spacing = 1.00 A			Grid Spacing = 1.50 A			Grid Spacing = 2.00 A		
#	Substituent	Base	Base + 2 A	Base – 2 A	Base	Base + 1.5 A	Base – 1.5 A	Base	Base + 2 A	Base – 2 A
1	MP	0.299	0.299	0.302	0.527	0.527	0.524	0.479	0.479	0.478
2	2-Br	1.212	1.212	1.209	1.190	1.190	1.194	1.562	1.562	1.545
3	2-Cl	-0.290	-0.290	-0.285	-0.735	-0.736	-0.718	-1.085	-1.085	-1.091
4	2-F	-0.703	-0.704	-0.698	-0.997	-0.998	-0.994	0.928	0.928	0.903
5	2-naphthyl	1.164	1.165	1.162	0.762	0.762	0.754	2.146	2.146	1.976
6	2-OCH3	-2.680	-2.679	-2.679	-2.846	-2.847	-2.838	-1.374	-1.374	-1.372
7	2-OH	-1.425	-1.424	-1.435	-1.662	-1.662	-1.668	-1.647	-1.647	-1.663
8	3-Br	0.260	0.259	0.267	0.435	0.435	0.436	0.286	0.286	0.284
9	3-CH3	0.293	0.292	0.301	-0.230	-0.230	-0.233	1.355	1.355	1.302
10	3-Cl	0.314	0.313	0.325	0.626	0.626	0.629	0.185	0.185	0.185
11	3-F	0.024	0.023	0.030	0.669	0.670	0.668	0.201	0.201	0.221
12	3-NH2	0.790	0.786	0.830	0.485	0.484	0.523	3.585	3.585	3.683
13	3-OCH3	-0.856	-0.856	-0.855	-1.094	-1.094	-1.097	-1.767	-1.767	-1.755
14	3-OH	-1.218	-1.218	-1.212	-1.256	-1.256	-1.256	-1.166	-1.166	-1.156
15	3,4-Cl	-0.426	-0.426	-0.429	-0.410	-0.410	-0.410	-0.991	-0.991	-0.998
16	3,4-OCH3	-0.604	-0.601	-0.622	0.868	0.868	0.889	-1.977	-1.977	-1.975
17	3,5-CH3	-0.914	-0.917	-0.882	-1.085	-1.085	-1.087	-1.450	-1.450	-1.446
18	3,5-Cl	0.476	0.475	0.481	0.834	0.834	0.837	0.966	0.966	0.975
19	4-Br	0.446	0.447	0.445	0.452	0.452	0.450	0.736	0.736	0.726
20	4-C2H5	0.050	0.050	0.047	-0.110	-0.110	-0.110	0.638	0.638	0.631
21	4-CF3	-1.250	-1.249	-1.251	-1.766	-1.766	-1.771	-1.188	-1.188	-1.181
22	4-CH3	0.824	0.824	0.823	1.257	1.257	1.255	0.607	0.607	0.589
23	4-Cl	0.016	0.016	0.015	-0.273	-0.273	-0.275	0.373	0.373	0.362
24	4-F	0.314	0.314	0.314	0.228	0.229	0.224	0.646	0.646	0.649
25	4-I	-0.146	-0.146	-0.144	0.558	0.558	0.552	-0.705	-0.705	-0.713
26	4-NH2	0.349	0.326	0.338	-0.598	-0.598	-0.612	1.088	1.088	1.082
27	4-NO2	-0.602	-0.602	-0.602	-1.200	-1.200	-1.206	-0.553	-0.553	-0.557
28	4-OCH3	-0.003	-0.003	-0.002	0.713	0.713	0.713	0.512	0.512	0.528
29	4-OH	-0.518	-0.518	-0.522	-0.374	-0.374	-0.379	-1.095	-1.095	-1.077
30	4-*t*-butyl	-0.511	-0.511	-0.515	-1.115	-1.115	-1.115	-1.012	-1.012	-1.006

A. CoMFA region boxes for 1.00 Å grid spacing

Table 4.10 Base for 1.00 Å Grid Spacing

# of lattice points:	8379	# of Boxes:	1
		Box Center (x, y, z):	(1.0, 2.0, 0.0)
	X	Y	Z
Lower Corner (Å):	-9.0	-8.0	-9.0
High Corner (Å):	11.0	12.0	9.0
Step size (Å):	1.0	1.0	1.0
# of steps:	21	21	19
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

Table 4.11 Base + 2 Å for 1.00 Å Grid Spacing

# of lattice points:	14375	# of Boxes:	1
		Box Center (x, y, z):	(1.0, 2.0, 0.0)
	X	Y	Z
Lower Corner (Å):	-11.0	-10.0	-11.0
High Corner (Å):	13.0	14.0	11.0
Step size (Å):	1.0	1.0	1.0
# of steps:	25	25	23
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

Table 4.12 Base – 2 Å for 1.00 Å Grid Spacing

# of lattice points:	1210	# of Boxes:	1
		Box Center (x, y, z):	(1.0, 2.0, 0.0)
	X	Y	Z
Lower Corner (Å):	-7.0	-6.0	-7.0
High Corner (Å):	9.0	10.0	7.0
Step size (Å):	1.0	1.0	1.0
# of steps:	17	17	15
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

B. CoMFA region boxes for 1.50 Å grid spacing

Table 4.13 Base for 1.50 Å Grid Spacing

# of lattice points:	2925	# of Boxes:	1
		Box Center (x, y, z):	(1.5, 2.5, 0.0)
	X	Y	Z
Lower Corner (Å):	-9.0	-8.0	-9.0
High Corner (Å):	12.0	13.0	9.0
Step size (Å):	1.5	1.5	1.5
# of steps:	15	15	13
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

Table 4.14 Base + 1.5 Å for 1.50 Å Grid Spacing

# of lattice points:	4335	# of Boxes:	1
		Box Center (x, y, z):	(1.5, 2.5, 0.0)
	X	Y	Z
Lower Corner (Å):	-10.5	-9.5	-10.5
High Corner (Å):	13.5	14.5	10.5
Step size (Å):	1.5	1.5	1.5
# of steps:	17	17	15
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

Table 4.15 Base – 1.5 Å for 1.50 Å Grid Spacing

# of lattice points:	1859	# of Boxes:	1
		Box Center (x, y, z):	(1.5, 2.5, 0.0)
	X	Y	Z
Lower Corner (Å):	-7.5	-6.5	-7.5
High Corner (Å):	10.5	11.5	7.5
Step size (Å):	1.5	1.5	1.5
# of steps:	13	13	11
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

C. CoMFA region boxes for 2.00 Å grid spacing

Table 4.16 Base for 2.00 Å Grid Spacing

# of lattice points:	1210	# of Boxes:	1
		Box Center (x, y, z):	(1.0, 2.0, 0.0)
	X	Y	Z
Lower Corner (Å):	-9.0	-8.0	-9.0
High Corner (Å):	11.0	12.0	9.0
Step size (Å):	2.0	2.0	2.0
# of steps:	11	11	10
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

Table 4.17 Base + 2 Å for 2.00 Å Grid Spacing

# of lattice points:	2028	# of Boxes:	1
		Box Center (x, y, z):	(1.0, 2.0, 0.0)
	X	Y	Z
Lower Corner (Å):	-11.0	-10.0	-11.0
High Corner (Å):	13.0	14.0	11.0
Step size (Å):	2.0	2.0	2.0
# of steps:	13	13	12
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

Table 4.18 Base – 2 Å for 2.00 Å Grid Spacing

# of lattice points:	648	# of Boxes:	1
		Box Center (x, y, z):	(1.0, 2.0, 0.0)
	X	Y	Z
Lower Corner (Å):	-7.0	-6.0	-7.0
High Corner (Å):	9.0	10.0	7.0
Step size (Å):	2.0	2.0	2.0
# of steps:	9	9	8
Probe Atom:	sp^3 Carbon		
Charge:	+ 1.0		

4.1.3 Changing Steric and ES Field Cutoffs and σ

All previous CoMFA runs were made with both the steric and electrostatic cutoffs set to 30 kcal/mol and a column filtering value of 2. The sensitivity of q^2 to variation of these parameters was tested for the MP analogues, producing the results shown in Table 4.19. All CoMFA runs in the table were carried out using all 30 analogues, binding data only, both steric and ES fields, the Leave-One-Out cross-validation method, and the default CoMFA region box (Table 4.1) and grid spacing (2.0 Å). The combinations which result in the best q^2 are highlighted in bold.

Table 4.19 CoMFA Parameter Optimization to Maximize q^2 for CoMFA Study I

Steric [a]	ES [a]	σ [a]	# of Components	q^2
30 [b]	**30** [b]	**2** [b]	**4**	**0.477**
Change Steric Cutoff				
100	30	2	5	0.517
115	30	2	5	0.546
125	**30**	**2**	**5**	**0.555**
135	30	2	5	0.546
150	30	2	5	0.542
175	30	2	5	0.526
200	30	2	5	0.478
Change ES Cutoff				
125	50	2	5	0.557
125	100	2	5	0.563
125	150	2	5	0.565
125	**200**	**2**	**5**	**0.573**
125	225	2	5	0.562
125	250	2	5	0.562
Change σ				
125	200	1	5	0.555
125	200	3	5	0.568
125	200	4	5	0.566
125	200	5	5	0.567
125	200	6	5	0.571
125	**200**	**7**	**5**	**0.584**
125	200	8	6	0.583
125	200	9	6	0.580

[a] In kcal/mol [b] Defaults; Run 1 in Table 4.2

Changing the steric energy cutoff while keeping the ES energy cutoff at its default value of 30 kcal/mol increased q^2 from 0.477 (for 30/30 Steric/ES energy cutoffs) to 0.555 (for 125/30). Maintaining the steric energy cutoff at 125 kcal/mol while changing the ES cutoff, further increased the q^2 to 0.573 (for 125/200). Altering the σ value next maximized the q^2 to 0.584 (for 125/200 and $\sigma = 7$). A non-validated CoMFA run was done on this model producing $r^2 = 0.930$, standard error of estimate = 0.341, 68.4% steric contribution, and 31.6% ES contribution. This indicates an increase in r^2 from that (0.885) produced by using the default energy cutoff and σ values. The ES contribution was significantly increased (from 21.1%) in agreement with the increased ES energy cutoff value (from 30 kcal/mol to 200 kcal/mol). The CoMFA steric and electrostatic coefficient contour map for this model (not shown) was generated with the following settings: green 80, yellow 35, blue 95, and red 35, and was similar to that generated from the run using default energy cutoff and σ values (Run 1, Table 4.2).

In order to attempt to improve the value of q^2 even further, alternative superposition rules were attempted as described below, in CoMFA studies II – V.

4.2 CoMFA Study II

In this experiment, the MP conformer number 1 from the 20 kcal/mol cutoff data set of 67 random search MP conformers was selected as the template for the alignment for CoMFA (Alignment #2, see Figure 3.10). CoMFA was done on all 30 analogues using the binding data only, both steric and ES fields, the Leave-One-Out cross-validation method, and the default CoMFA region box (Table 4.1) and grid spacing (2.0 Å). The results are given in Table 4.20. CoMFA of all 30 analogues produced a q^2 of 0.466. An

examination of the residuals for this run indicated that the following analogues had high (greater than 1.5) values: 2-Naphthyl (7.03), 2-OH (-1.839), 4-OCH_3 (1.782), and 4-C_2H_5 (1.586). It should be noted here that the 2-Naphthyl analogue does not actually have a substituent at the 2-position. Rather the second ring of the naphthyl group protrudes into the region between the 3- and 4- positions. The large residual for 2-Naphthyl seems to indicate that the data do not predict substituents in this region well. Since the residual value for 2-Naphthyl was so high, another CoMFA run was done on 29 analogues, with the 2-Naphthyl omitted. This run raised the q^2 value to 0.546. Coupled with the fact that both runs produced an inadequate r^2 value of less than 0.8, it would seem, therefore, that these models are unsuitable for predictions of bioactivity and further experiments based on altering the CoMFA parameters were not pursued.

Table 4.20 Selected Results for CoMFA Study II

Run	Cross Validation		No Validation			
	Optimum # of Components	q^2	Standard Error of Estimate	% Steric	% ES	r^2
With all 30 analogues	3	0.466	0.565	97.4	2.6	0.783
Without 2-Naphthyl	3	0.546	0.544	99.2	0.8	0.798

Of note, however, is the near-total domination of the steric contributions over the ES contributions in both the 30-analogue and the 29-analogue CoMFA runs. This represents a significant increase from the results obtained in CoMFA Study I where the steric contributions amounted to about 80% and the ES to about 20%. Yet, with a q^2 of about 0.5, the models in this study as well as those in CoMFA Study I are about the same in terms of predictability (based on q^2).

4.3 CoMFA Study III

In this study, conformer number 3 from the 67 MP conformers of the 20 kcal/mol data set was chosen as the template for alignment (Alignment #3, see Figure 3.11). CoMFA was performed on all 30 analogues using the same parameters as those used in CoMFA Study II. This run produced a model with a q^2 of 0.459 with optimum number of components at three. The greatest residual value was -2.099, that of the 4-*t*-butyl analogue. Other high residuals were those of 4-CF_3 (-1.405), 3,4-OCH_3 (1.409), 4-NO_2 (-1.379), 4-NH_2 (1.362), and 2-Naphthyl (1.342). A non-validated run was carried out which gave standard error of estimate = 0.529, r^2 = 0.809, and steric contribution of 99.3 %. The results are thus very similar to those obtained in CoMFA Study II and further parameter variation was not performed.

4.4 CoMFA Study IV

This experiment utilized MP conformer number 17 from the 20 kcal/mol data set for the template for alignment (Alignment #4, see Figure 3.12). CoMFA was again performed on all 30 analogues using the parameters as before. A q^2 of 0.531 with optimum number of components four were obtained. Some analogues with relatively high residuals in this run were 2-OCH_3 (1.968), 3,5-CH_3 (-1.833), 2-Naphthyl (1.618), 4-C_2H_5 (-1.575), 3-NH_2 (-1.574), and 3,5-Cl (1.519). The non-validated run produced standard error of estimate of 0.530, r^2 of 0.816, and steric contribution of 98.1%. Thus, a relatively high (cross-validated) q^2 value did not generate a high r^2 value. Consequently, this model can also be deemed inadequate for prediction purposes.

Since the q^2 value in this study was relatively high (0.531), several different combinations of the steric and ES energy cutoffs were tried with the 30 analogues to optimize the q^2. The column filtering value, σ, was kept at the default 2.0 kcal/mol. The results are shown in Table 4.21. Upon changing the steric energy cutoff from 30 to 60 kcal/mol, the q^2 dropped sharply to 0.360 and further to 0.310 for cutoff values 100 kcal/mol and above. This meant that the steric cutoff of 30 kcal/mol was already at an optimal level for this model. Raising the ES cutoff values, with the steric cutoff at 30 kcal/mol, did not affect the original q^2 of 0.531. This result is in accord with the non-validated CoMFA run indication of only 1.9% ES contribution for this model.

Table 4.21 Parameter Optimization for CoMFA Study IV

Steric [a]	ES [a]	# of Components	q^2
30 [b]	**30** [b]	**4**	**0.531**
Change Steric Cutoff			
60	30	4	0.360
100	30	5	0.310
150	30	5	0.310
200	30	5	0.310
Change ES Cutoff			
30	50	4	0.531
30	75	4	0.531
30	150	4	0.531

[a] In kcal/mol [b] Defaults

4.5 CoMFA Study V

The final CoMFA experiment had MP conformation number 20 from the 20 kcal/mol RANDOMSEARCH data set as the template for alignment (Alignment #5, see Figure 3.13). CoMFA was carried out on all 30 analogues using the binding data only, both steric and ES fields, the Leave-One-Out cross-validation method, and other defaults: default

CoMFA region box (Table 4.1), grid spacing (2.0 Å), steric and ES energy cutoffs (both 30 kcal/mol), and σ (2 kcal/mol). This run produced the best (cross-validated) q^2 (0.662) obtained so far using the initial parameters listed above. The optimum number of components for this run was four. Only three analogues had residuals larger than 1.5: 2-Naphthyl (-1.587), 4-*t*-butyl (1.897), and 2-OCH_3 (-2.569). The non-validated CoMFA run gave $r^2 = 0.901$, standard error of estimate = 0.388, and steric contributions = 98.2%. A CoMFA coefficient contour map (not shown) was generated for this model (with green 80, yellow 35, blue 90, and red 80). As before, a sterically unfavorable yellow region was present off the 2-position (which correlates with the inactivity of the 2-position analogues) and between the 4- and 5-positions (which correlates with the inactivity of the 4-*t*-butyl analogue).

4.5.1 Changing Steric and ES Field Cutoffs and σ

Since a good initial q^2 was obtained above, and since optimization of some CoMFA parameters in CoMFA Study I did lead to an increase in q^2, several runs were conducted using various steric and electrostatic energy cutoffs and σ values. The results are given in Table 4.22. Increasing the steric energy cutoff to 40 kcal/mol (keeping the ES energy cutoff at the default 30 kcal/mol) decreased the q^2 to 0.632. Further increase in the steric cutoff resulted in progressive decrease in q^2. Increasing the ES cutoff (keeping the steric cutoff at 30 kcal/mol) decreased the q^2 to 0.557, which increased marginally upon further increasing ES cutoff but which remained well below the original 0.662 value.

As Table 4.22 shows, altering the σ value (at the default steric and ES energy cutoffs since these gave the highest q^2) mostly led to small differences in q^2, as in

CoMFA study I. However, for $\sigma = 9$, the q^2 shot up to 0.765, clearly giving the best q^2 value obtained in any of the studies described here. The optimum number of components for this model, however, was 11, indicating a very complex model. A non-validated CoMFA run was carried out on this model giving a high r^2 of 0.972, a relatively low standard error of estimate of 0.242, and a steric contribution of 96.3%. The CoMFA coefficients contour map generated for this model (Plate 3, Appendix D) was essentially the same as the one produced by the 0.662 q^2 model (not shown) indicating that alteration of σ does not affect the map for this model. In addition, the qualitative information obtained from this map is the same as that obtained from the maps in CoMFA Study I; both indicate sterically unfavorable yellow regions corresponding to the inactive 2-OCH_3 and 4-*t*-butyl analogues.

Table 4.22 Parameter Optimization for CoMFA Study V

Steric [a]	ES [a]	σ [a]	# of Components	q^2
30 [b]	**30** [b]	**2** [b]	**4**	**0.662**
Change Steric Cutoff				
40	30	2	3	0.632
60	30	2	5	0.600
100	30	2	5	0.541
Change ES Cutoff				
30	50	2	5	0.557
30	100	2	5	0.563
30	150	2	5	0.565
Change σ				
30	30	1	4	0.663
30	30	3	4	0.666
30	30	4	4	0.657
30	30	5	4	0.664
30	30	6	4	0.647
30	30	7	4	0.636
30	30	8	14	0.673
30	**30**	**9**	**11**	**0.765**
30	30	10	4	0.353

[a] In kcal/mol [b] Defaults

4.5.2 Predicting Bioactivity of "Test Set"

The predictability of a CoMFA model based on Alignment #5 was compared to that of the CoMFA model based on Alignment #1 of CoMFA Study I using the same test and training sets as in Study I (see Section 4.1.1.2). The steric and ES energy cutoffs and the σ value used were those that maximized q^2 to 0.765, i.e. 30/30 kcal/mol and 9 respectively. CoMFA was carried out using the binding only data, both steric and ES fields, the Leave-One-Out cross-validation method, and the default region box (Table 4.1). The cross-validated and non-validated CoMFA results are given as Test 1 (for the 20% Test Set 1) and as Test 2 (for the 10% Test Set 2) in Table 4.23. Table 4.24 lists the residuals for the compounds in Test Set 1 under the column heading "Test Set 1" and those for the compounds in Test Set 2 under the column heading "Test Set 2".

Table 4.23 CoMFA Using Two Training Sets for CoMFA Study V

Test				Cross Validation		No Validation			
#	Size (%)	# of Training Set Analogues	Region Box	Optimum # of Components	q^2	Standard Error of Estimate	% Steric	% ES	r^2
1	20	24	Default [a]	4	0.279	0.328	98.1	1.9	0.922
2	10	27	Default [a]	4	0.746	0.287	96.9	3.1	0.947

[a] Table 4.1

Table 4.24 Bioactivity Predictions of Test Sets for CoMFA Study V

Region Box: Table 4.1		Test Set 1		Test Set 2	
Analogue	**Actual -logIC_{50}**	**Predicted -logIC_{50}**	**Residual**	**Predicted -logIC_{50}**	**Residual**
2-Br	5.73	5.10	0.63		
2-OCH_3	4.00	5.48	-1.48	6.58	-2.58
3-CH_3	7.67	7.66	0.01	7.72	-0.05
3,5-Cl	7.18	7.20	-0.02		
4-Br	8.16	7.55	0.61	7.58	0.58
4-NO_2	6.31	6.83	-0.52		

Test 1 produced a low q^2 of 0.279 with optimum number of components being four. The non-validated CoMFA run for this model, however, gave a high r^2 value of 0.922. This may explain the marginally better predictability of this model over its counterpart in CoMFA Study I (i.e. Test 1A in Table 4.5) which had a slightly higher q^2 of 0.289 but a lower r^2 of 0.859. In Table 4.24, two compounds in Test Set 1 (3-CH_3 and 3,5-Cl) came within a hundredth of their actual (-log IC_{50}) values as against only one (3-CH_3) in Test Set 1A (Table 4.5). Test 2 had a high q^2 value of 0.746 as well as a high r^2 value of 0.947 (both higher than their counterparts for Test Set 2A in CoMFA Study I, Table 4.5 with q^2 of 0.365 and r^2 of 0.860). CoMFA Study I gave about the same results whether the test set consisted of 10 or 20% of the analogues. CoMFA Study V gave much better results for the smaller training set and seems to be a reasonably predictive model.

4.5.3 Predicting Bioactivity of "Novel" Compounds

The CoMFA model with $q^2 = 0.765$ was used to predict the activity of the same "novel" compounds as CoMFA Study I. The predicted (-log IC_{50}) values and the corresponding nanomolar IC_{50} values for these compounds are listed in Table 4.25. Most of the compounds were predicted to have a lower IC_{50} value than those predicted in CoMFA Study I. By and large, the pattern remained the same with 2-I being predicted to have a high IC_{50} of 2630 nM and 3,4-Br being predicted to have an even lower (than in CoMFA Study I where it was 3.1 nM) IC_{50} of 0.9 nM. The ultimate test of the predictability of the models will be the synthesis and testing of some of these analogues by the Deutsch and Schweri groups.

Table 4.25 Predicted Bioactivity of "Novel" Compounds for CoMFA Study V

Compound	Predicted -logIC_{50}	IC_{50} (nM)
2-I	5.58	2630
3-I	8.03	9.3
3,4-Br	9.05	0.9
3,5-Br	6.88	132
3,4,5-Br	7.62	24.0
3,4,5-Cl	7.80	15.8
4,5-Cl	6.88	132
3-NO_2	6.55	282

CHAPTER 5

CONCLUSION

Several conclusions can be drawn from the CoMFA studies carried out on the MP analogues in this research. CoMFA results that were based on the binding data were similar to those based on the uptake data reflecting the highly correlated nature of the two sets of data. Changing the grid spacing and region box parameters did not significantly affect q^2. Changing the steric and ES energy cutoff and column filtering (σ) values had a much larger effect on q^2, though varying the σ value alone usually produced a smaller change in q^2. Changing the template structure (i.e. the hypothesized bioactive conformation) for superposition produced the biggest effect in q^2, underlining the importance of using a thoroughly investigated alignment rule for useful CoMFA results. Thus, in order of decreasing effect on q^2 are: the template structure > energy cutoff value > σ value > grid spacing and region box parameters. In addition, the steric and ES energy cutoffs can affect the steric and ES contributions in the non-validated runs. The non-validated runs indicated the dominance of steric effects in all the analyses.

The best (cross-validated) q^2 for the GEM conformation of MP (conformer number 6 from the 5 kcal/mol data set or conformer number 14 from the 20 kcal/mol data set) was 0.584 and was obtained using the default region box, steric/ES cutoffs of 125/200 kcal/mol, and a σ value of 7. The (non-validated) r^2 for this model was 0.930, while the steric and ES contributions were 68.4% and 31.6% respectively.

The best overall (cross-validated) q^2 was 0.765 and was achieved using the MP conformer number 20 from the 20 kcal/mol data set, the default region box, and the

default steric and ES cutoffs (30/30 kcal/mol), but with a σ value of 9. The (non-validated) r^2 for this model was 0.972, while the steric and ES contributions were 96.3% and 3.7% respectively. Thus, by q^2 and r^2, MP conformer number 20 produced a better model than did the MP GEM conformer.

By predictability of test sets, CoMFA based on MP conformer number 20 was slightly better than that based on the MP GEM conformer for the 20 percent test set but about the same for the 10 percent test set.

The CoMFA coefficients contour maps generated for the two models, i.e., those based either on the MP conformer 20 or on the MP GEM conformer, produced qualitatively similar results even though the q^2 were different (0.765 vs. 0.584). In both maps, the sterically unfavorable yellow regions around the 2- and the 4-position on the phenyl ring correlated with the poor IC_{50} values of the 2-OCH_3 (10^5 nM) and the 4-*t*-butyl analogues (13,450 nM). In addition, the 2-position analogues have poor IC_{50} values because they protrude into a sterically unfavorable region.

The model using MP conformer number 20 as the template predicted the "novel" compounds to be more active than the model that used the MP GEM conformer as the template. However, the ultimate test of these models will be the synthesis and testing of these compounds by Drs. Howard Deutsch and Margaret Schweri. A novel compound for synthesis and testing is 3,4-Br MP.

APPENDIX A

SYBYL PARAMETERS USED FOR RANDOMSEARCH

Minimize Details:

Method: Powell	Termination: Gradient
Max. Iterations: 1000	Non-bonded reset: 10
Max. Displacement: 0.01	RMS Displacement: 0.001
Min. Energy change: 0.05	Gradient: 0.05
Simplex threshold: 1000	Simplex Iterations: 20
LS accuracy: 0.001	Derivative reset: 100
LS step size: 0.001	Color Option: Potential
Status update: 1	Checkpoint Interval: 0
Graphics update: 1	List terms threshold: No

Energy Setup:

Force Field: Tripos

Charges: Gasteiger-Hückel

Dielectric function: Distance

Dielectric constant: 1.0

NB Cutoff: 8.0 Å

Aggregates: On (*Eliminated piperidine ring inversion*)

Random Search Details:

Maximum cycles: 10,000 (*Ensured comprehensive search of 3D space*)

Energy cutoff: 5.0 kcal/mol

RMS threshold: 0.2 Å

Convergence threshold: 0.05

Minimum hits: 500 (*Ensured completion of 10,000 cycles*)

Check chirality: Yes (*Ensured consistency of (R, R) configuration*)

Symmetry: No

APPENDIX B

NOTES ON METHODOLOGY

A. The ninety-eight 2D neutral methylphenidates (of which 30 were studied in this research) of Howard Deutsch were provided by him as a ".sdf" file and were optimized using SYBYL to give rise to their 3D structures.

B. This collection of compounds consisted of mostly *threo* (R,R or S,S configuration of the two chiral carbons in the analogs) and some *erythro* (R,S configuration) structures. Since SYBYL assigned random configurations during minimization of the 2D structures, all the *threo* structures were modified, wherever necessary, to have the (R,R) configuration.

C. The ninety-eight analogues were divided into seven general "classes" and the DISCO features assigned to each *threo* analogue were also identified and recorded. Conformational analysis was performed using RANDOMSEARCH on the 30 *threo* structures belonging to class "A." This class was selected because it had the largest number of analogues to work with.

D. Each analogue was submitted for random search with its piperidine ring as an "aggregate".

E. During the course of this study, it was noted that the analogues used in the early, trial computations had been assigned by SYBYL the wrong carbon atom-type in the aromatic ring. This was due to a probable bug in the software that didn't recognize the correct carbon atom-type while reading the molecules from the original ".sdf" file

F. that contained them. Subsequently, the analogues were modified to have the correct carbon atom-type and then subjected to RANDOMSEARCH again.

G. The superpositions of the MP analogues that produced the five alignments considered in this study were carried out in SYBYL using the "Align Database" option. This option aligns all molecules present in a database and that contain the indicated substructure, with the selected template molecule by minimizing the RMS distance to the substructure atoms in that template molecule.

H. For CoMFA Studies II – V, systematic search was performed on the analogues containing substituents with one or more rotatable bonds and on the 2-Naphthyl analogue. This was done to account for the rotations around such bonds in these analogues and the least energetic conformer was selected for the superposition.

Details

A. Conversion of the ninety-eight 2D neutral methylphenidates to their 3D structures: The "sdf" file sent by Howard Deutsch, hdtest1.sdf, contained 2D structures of ninety-eight methylphenidate analogues. This file was opened in SYBYL as a molecular spreadsheet (MSS) and each row was treated as follows:

   a) Highlight the row.

   b) MSS >> File >> Put rows into molecular areas.

   c) Compute >> Minimize. The defaults were used during minimization.

   d) Build/Edit >> Name molecule.

   e) Save as...[molecule name]

The molecules were saved in a database (compounds.mdb) and this was used as the data source for a spreadsheet (compounds.tbl). The IC_{50} values for WIN binding and uptake and other values included by Howard Deutsch were entered in this spreadsheet.

B. Conversion of all *threo* structures to their (R, R) configuration:

a) Find chirality:
   1. Build/Edit >> Other tools >> Find Chirality
   2. Choose RS
   3. Choose All (The chirality for all chiral centers will be listed in the console window)

b) Change chirality:
   1. Build/Edit >> Other tools >> Invert
   2. Click on the desired individual atom(s)
   3. Compute >> Minimize

The (R, R) *threo* compounds were saved in a new database (threos.mdb) and this was used as the data source for a new spreadsheet (threos.tbl).

D. Consistency was maintained in the following conditions:

a) All structures submitted for random search had the (R, R) configuration.

b) The piperidine ring was in the chair form.

c) The bulky group attached to the piperidine ring and also the hydrogen on the nitrogen in the piperidine ring were both in the equatorial position. (The Build/Edit >> Sketch molecule >> Move atom command was used, where necessary, to get the equatorial position.)

d) To maintain the above consistency, the piperidine ring of each analogue was submitted as an active aggregate (see below for procedure). This was done so that

during random search this ring may remain rigid to avoid ring inversion. This would enable RANDOMSEARCH to search through 10,000 cycles using only the conformations that have the bulky group in the equatorial position and not in the more energetic axial position.

To define an aggregate:

1. Read the molecule on screen
2. Build/Edit >> Aggregates
3. Choose New
4. Pick desired atoms and press OK
5. Enter a name for the aggregate and an optional comment string

E. Modification of aromatic carbon atom type:

a) Read analogue on screen

b) Build/Edit >> Modify >> Atom >> Type

c) Pick atom(s) to modify

d) Choose the correct atom type

e) Compute >> Minimize

F. Using the "Align Database" option in SYBYL: Once the database containing the 30 analogues in the hypothesized conformation was created, it was subjected to database alignment as a final preparatory step before CoMFA. The following steps describe the SYBYL commands used for this purpose:

a) File >> Align Database...

b) Select the database that contains the structures to be aligned

c) Select the template molecule structure from the structures in the database

d) Select the core substructure of the template that will be used for alignment

e) Select "Inertial" orientation (inertially aligns the template molecule to get a more robust CoMFA with a good q^2)

f) Select "Existing" database (for storage of aligned molecules)

g) Select "Align all molecules"

G. Systematic search: Systematic search was performed as follows on 2-OCH_3, 2-OH, 3-CH_3, 3-OCH_3, 3-OH, 3,4-OCH_3, 3,5-CH_3, 4-C_2H_5, 4-CF_3, 4-CH_3, 4-OCH_3, 4-OH, 4-*t*-butyl, and 2-Naphthyl.

a) Read a molecule on screen

b) Compute >> Search >> Systematic Search

c) Select bond(s)

d) Select increment size (10° if one bond was selected and 30° if more than one bonds were selected)

e) Set all van der Waals radius scale factors to zero

f) Select "Compute energy" = On

g) Select "Use electrostatics" = On

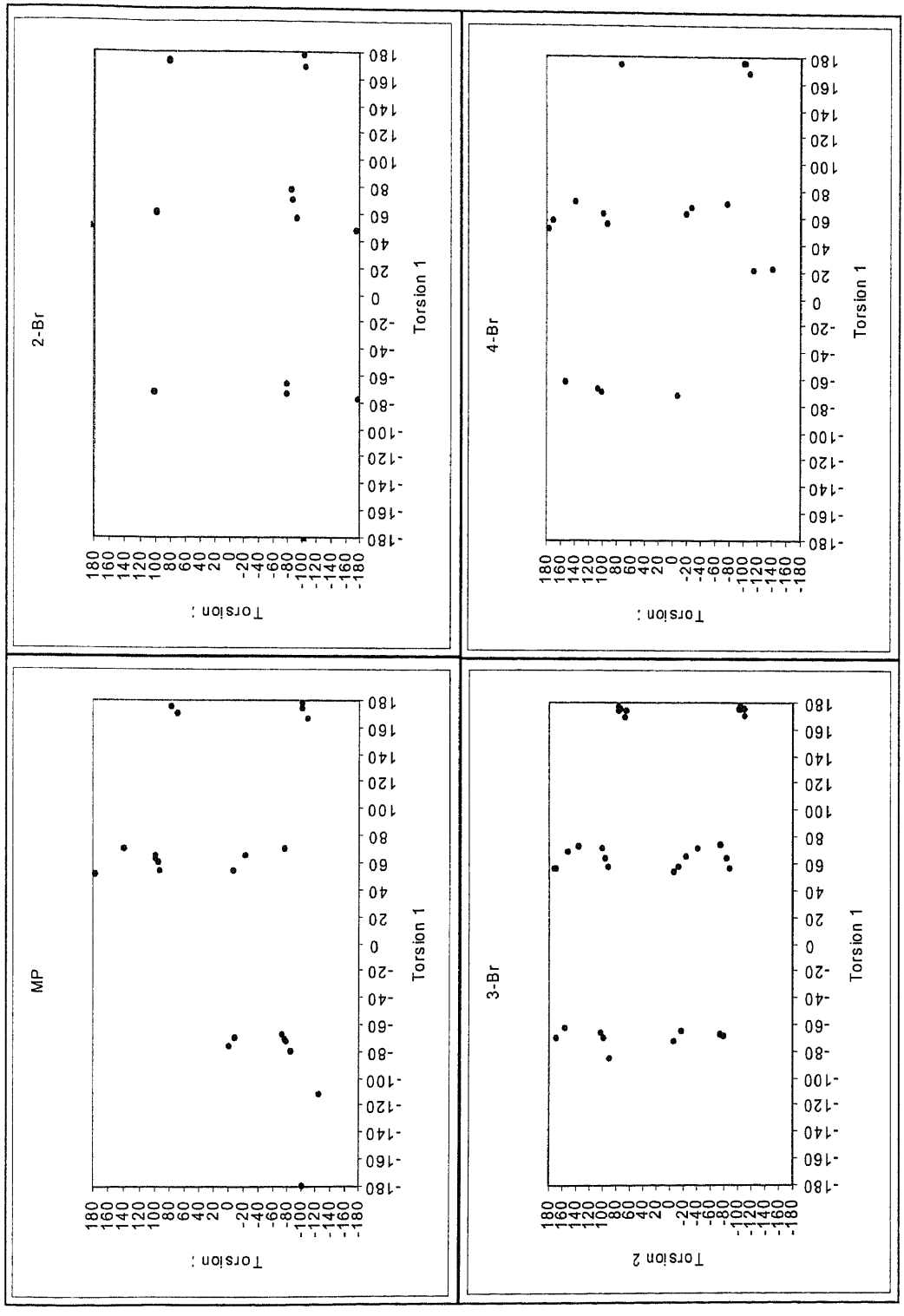
h) Select "Write angle file" = On

Upon completion of the search, a database containing the new conformers is created. Then use Analyze >> Search >> Systematic Search to select the appropriate conformer.

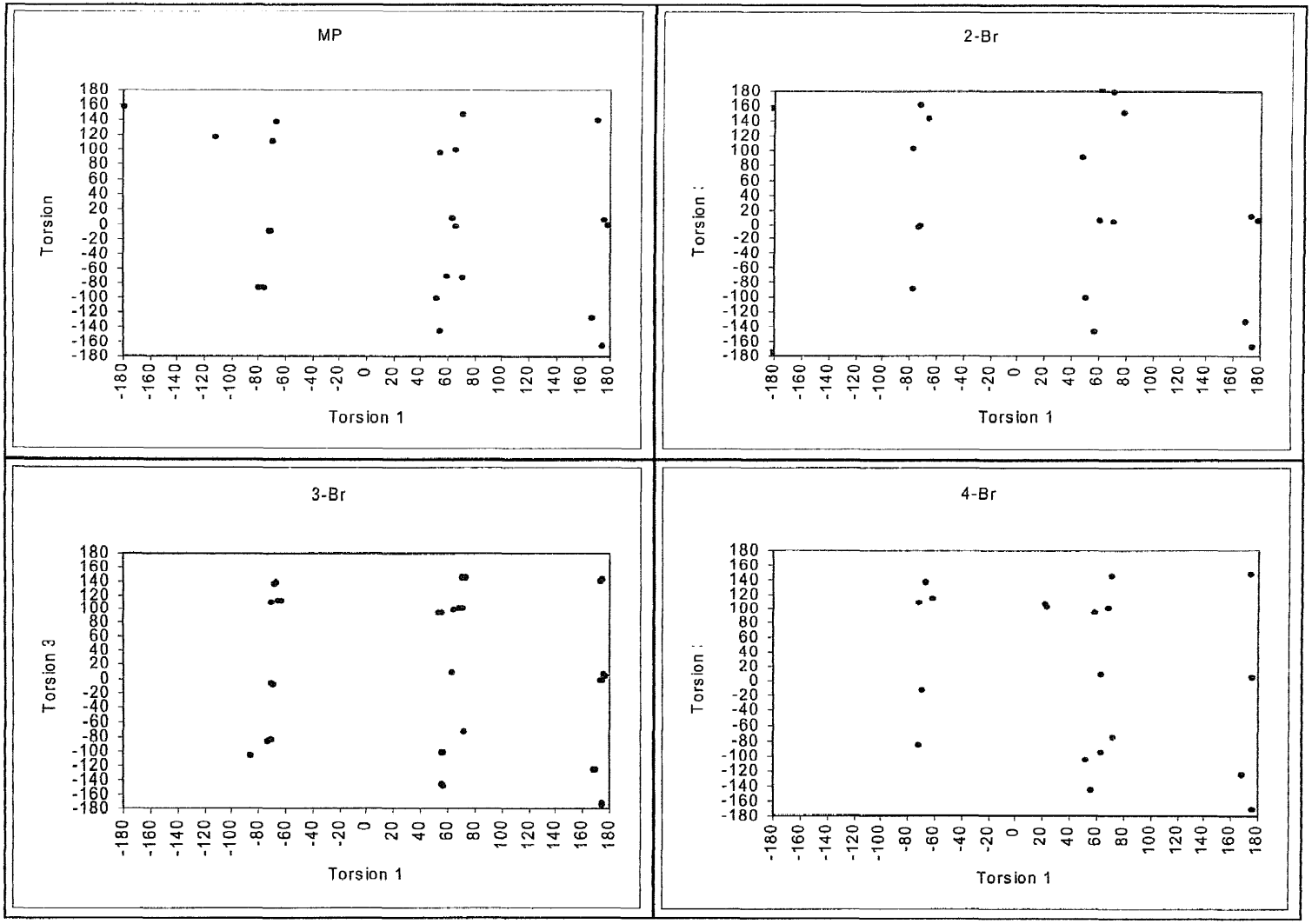
APPENDIX C

SCATTER GRAPHS FOR TORSIONAL ANGLE ANALYSIS

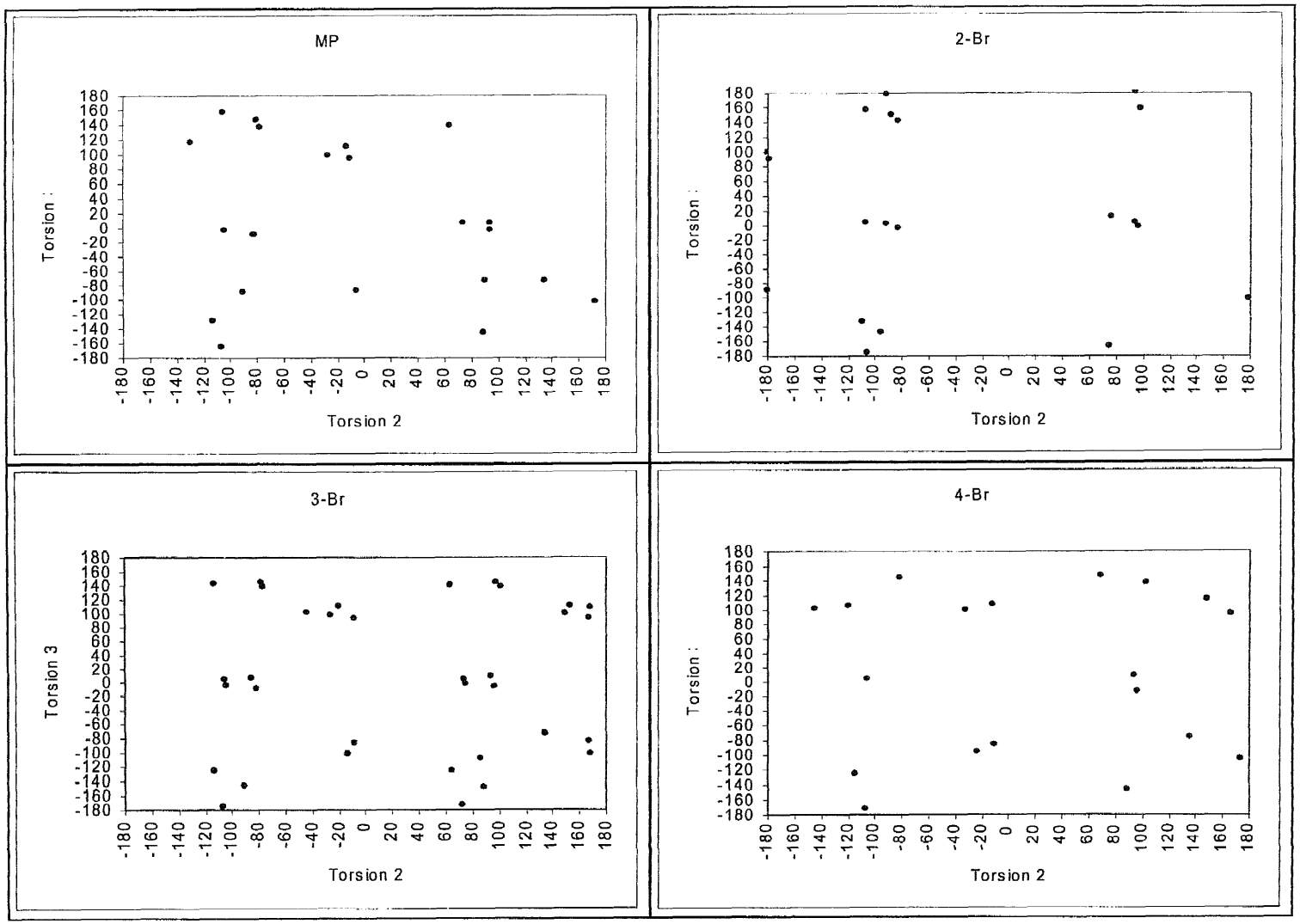
Sets of scatter plots for conformers of MP and of the 2-, 3, and 4-position Br, Cl, F, OCH_3, and OH analogues are provided (courtesy: Joseph Hinksmon) in this appendix. Each set shows, for a particular pair of torsional angles ($\phi 1$ vs. $\phi 2$, $\phi 1$ vs. $\phi 3$, or $\phi 2$ vs.$\phi 3$), four plots: that of MP, and those of the 2-, 3-, and 4-substitutions of an analogue. The points indicate the local minima corresponding to the conformers.



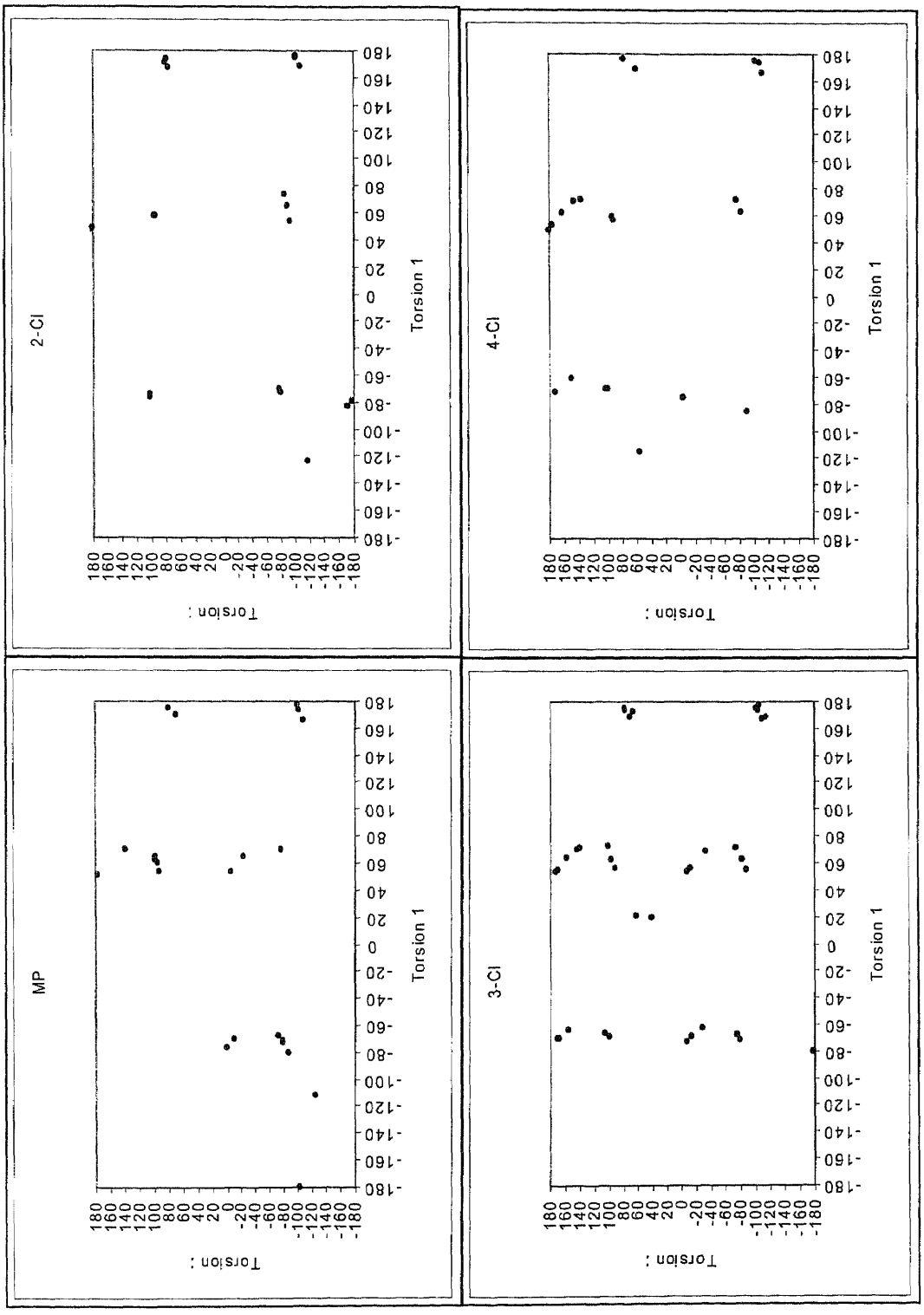
Scatter Plot Set 1a. Torsion 1 vs. Torsion 2 for MP and 2-, 3-, and 4-Br



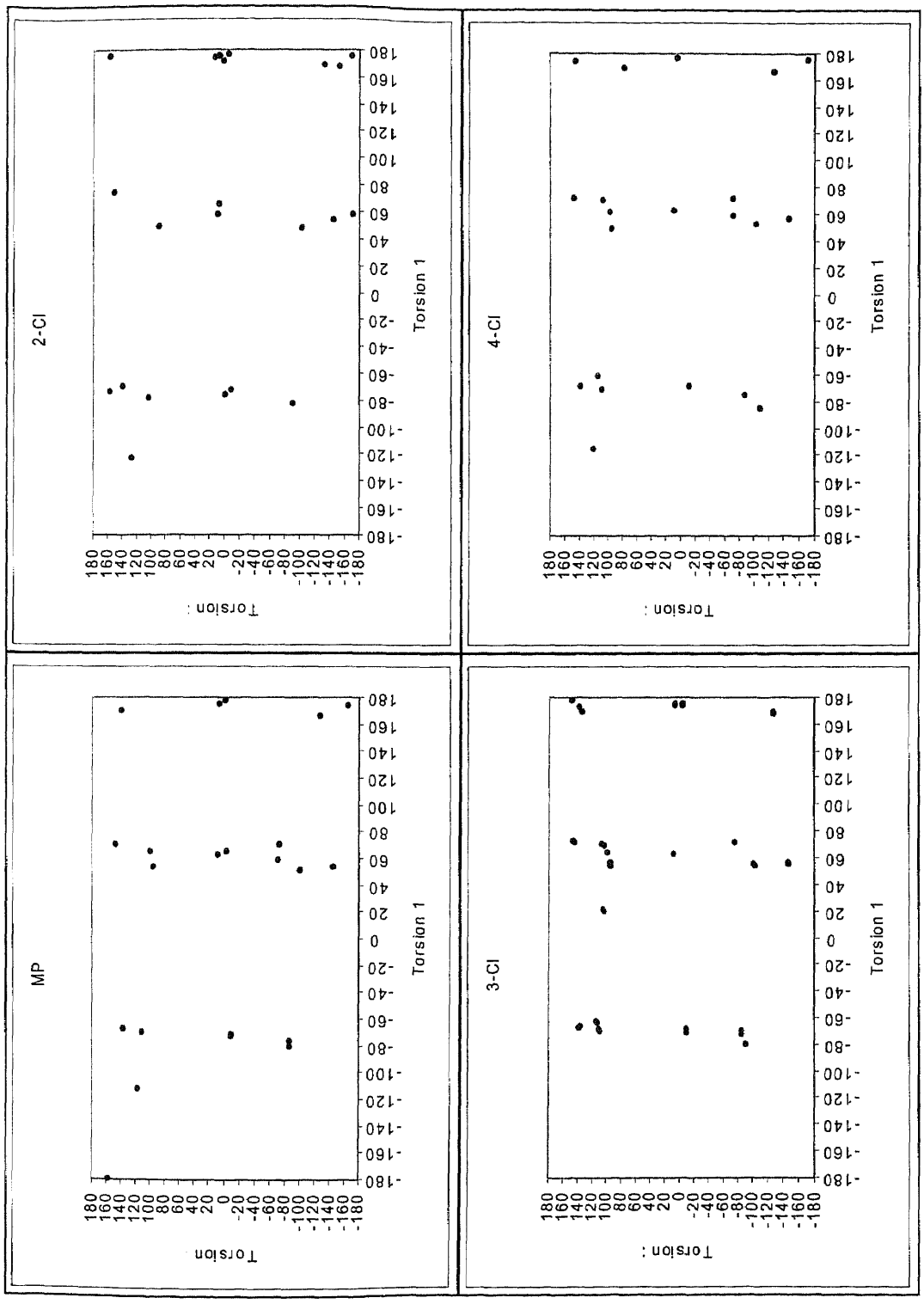
Scatter Plot Set 1b. Torsion 1 vs. Torsion 3 for MP and 2-, 3-, and 4-Br



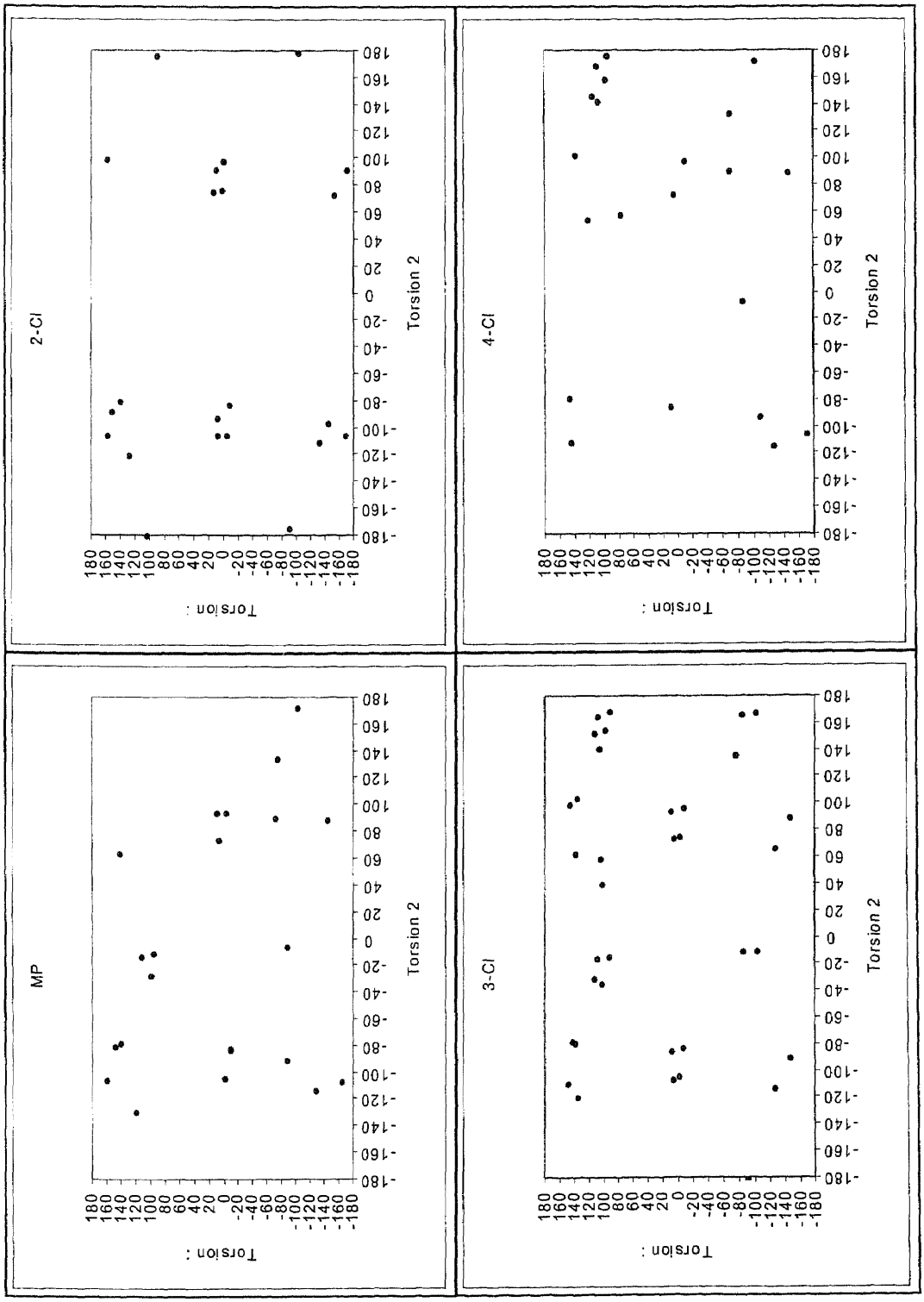
Scatter Plot Set 1c. Torsion 2 vs. Torsion 3 for MP and 2-, 3-, and 4-Br



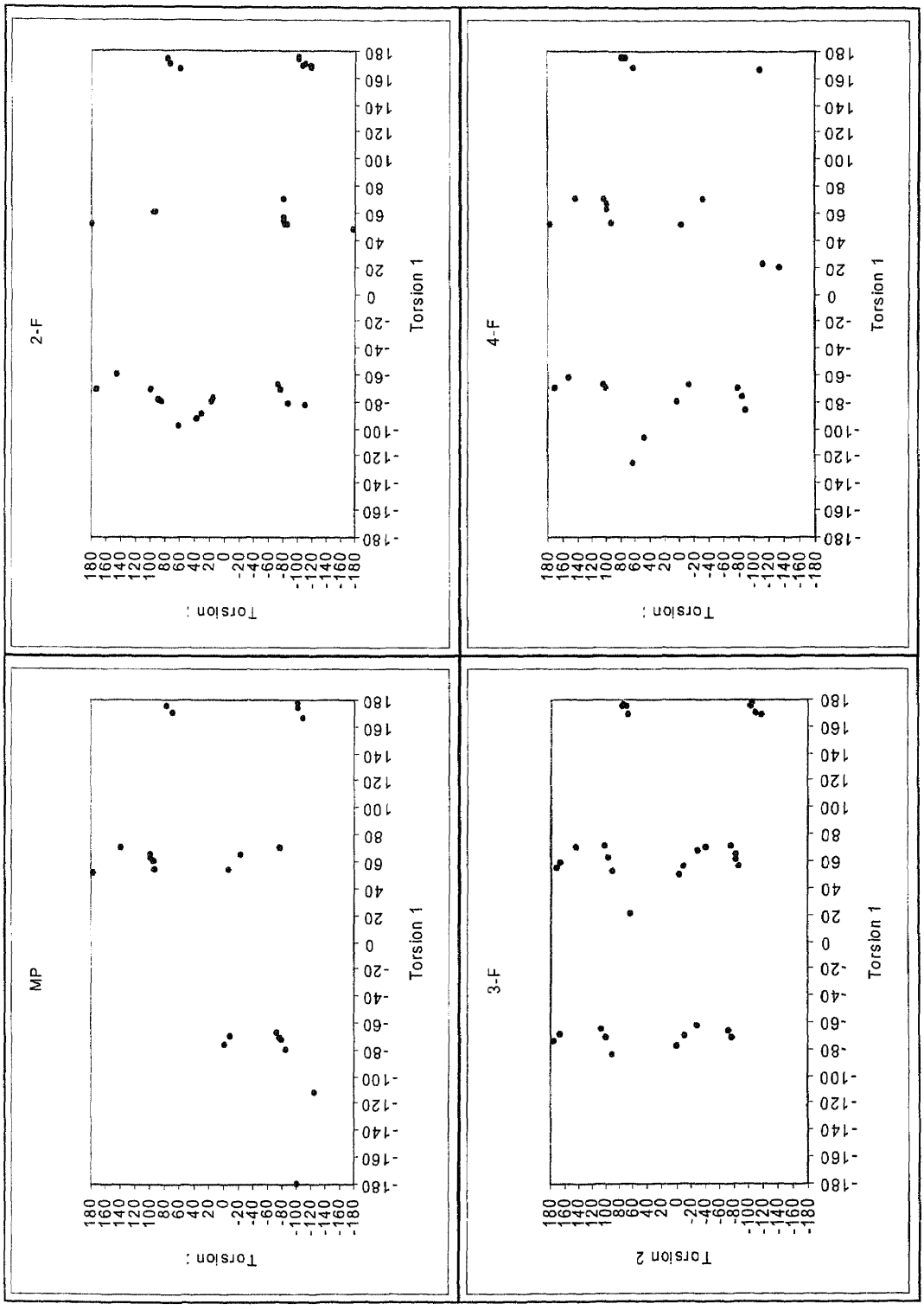
Scatter Plot Set 2a. Torsion 1 vs. Torsion 2 for MP and 2-, 3-, and 4-Cl



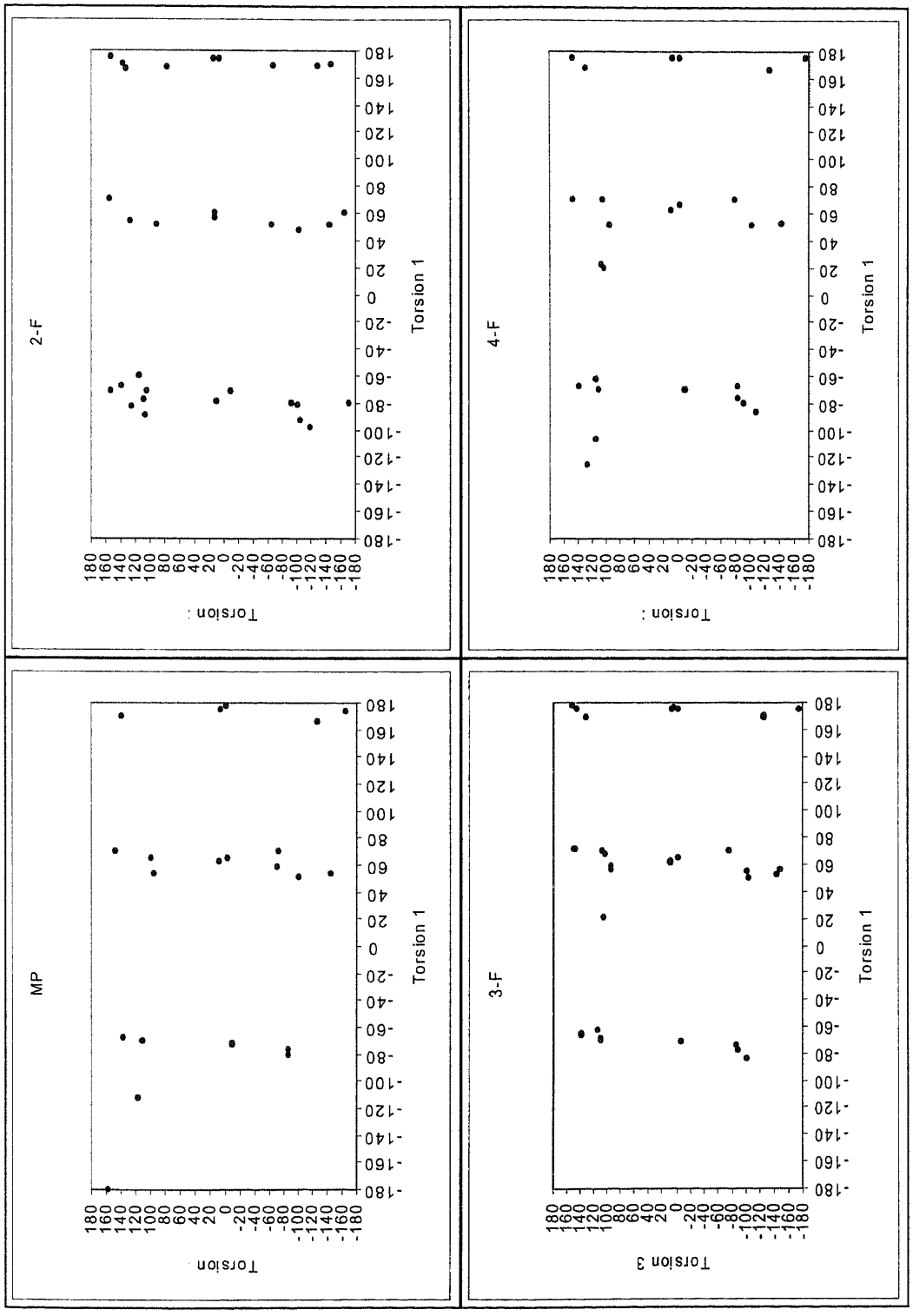
Scatter Plot Set 2b. Torsion 1 vs. Torsion 3 for MP and 2-, 3-, and 4-Cl



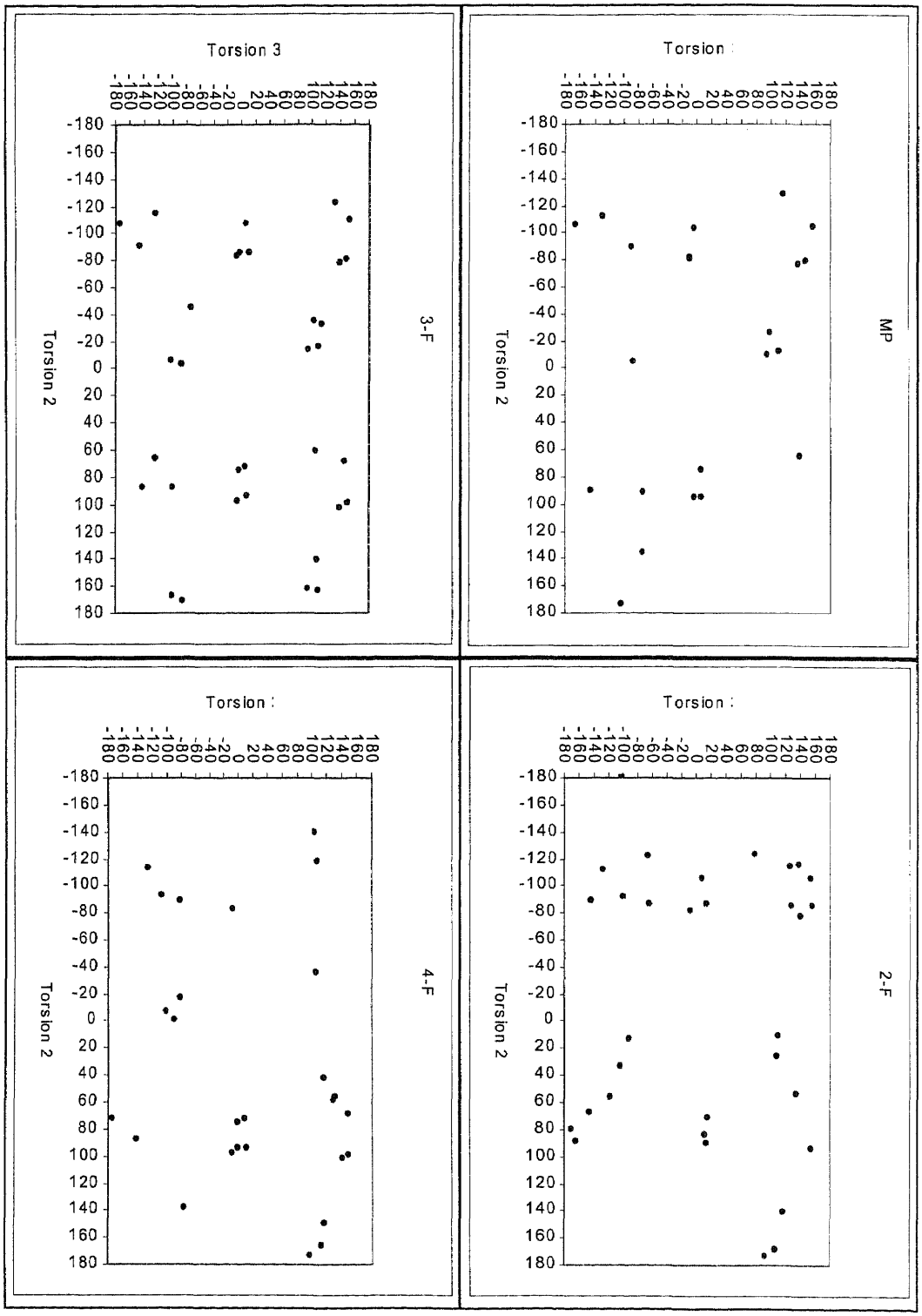
Scatter Plot Set 2c. Torsion 2 vs. Torsion 3 for MP and 2-, 3-, and 4-Cl



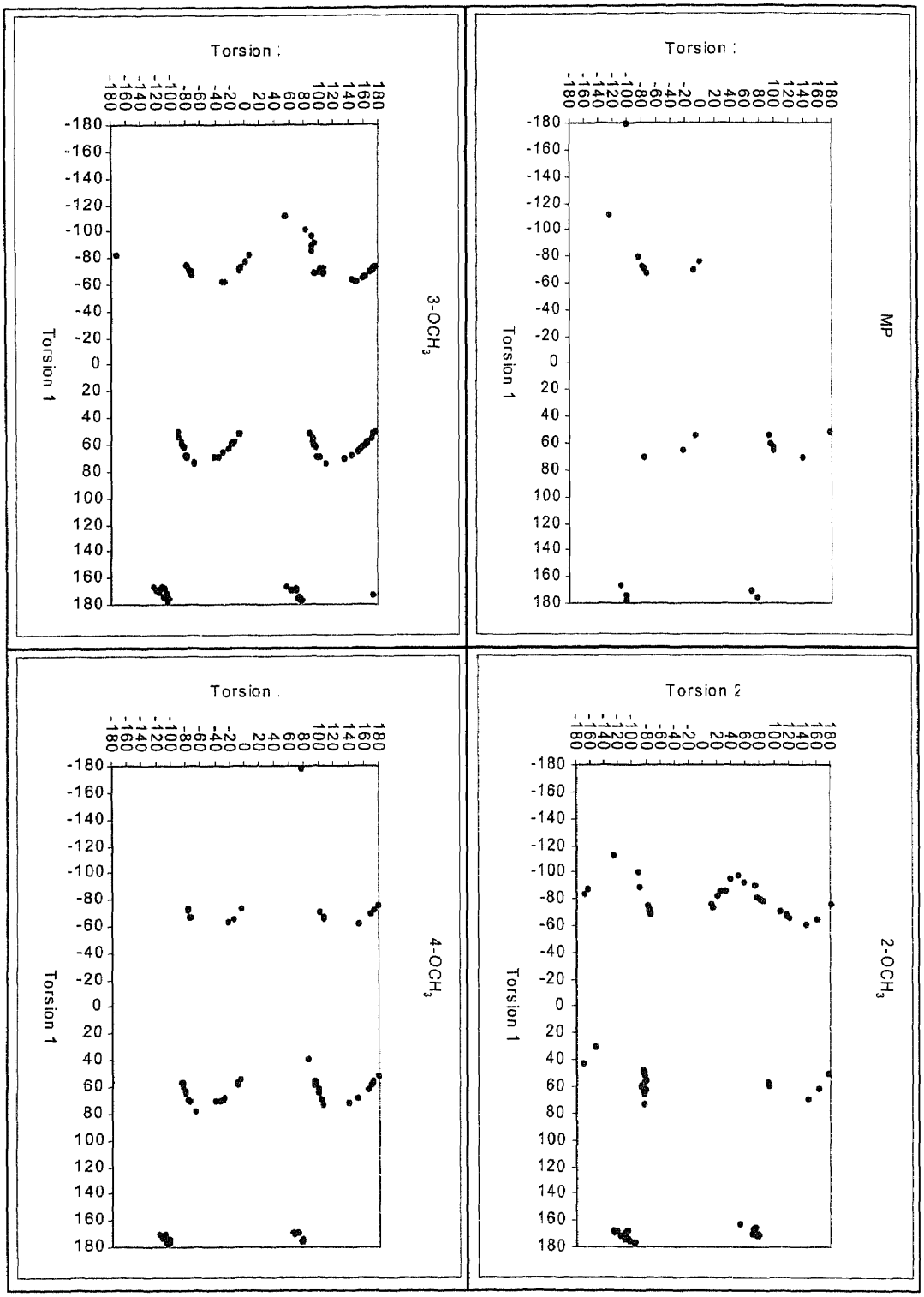
Scatter Plot Set 3a. Torsion 1 vs. Torsion 2 for MP and 2-, 3-, and 4-F



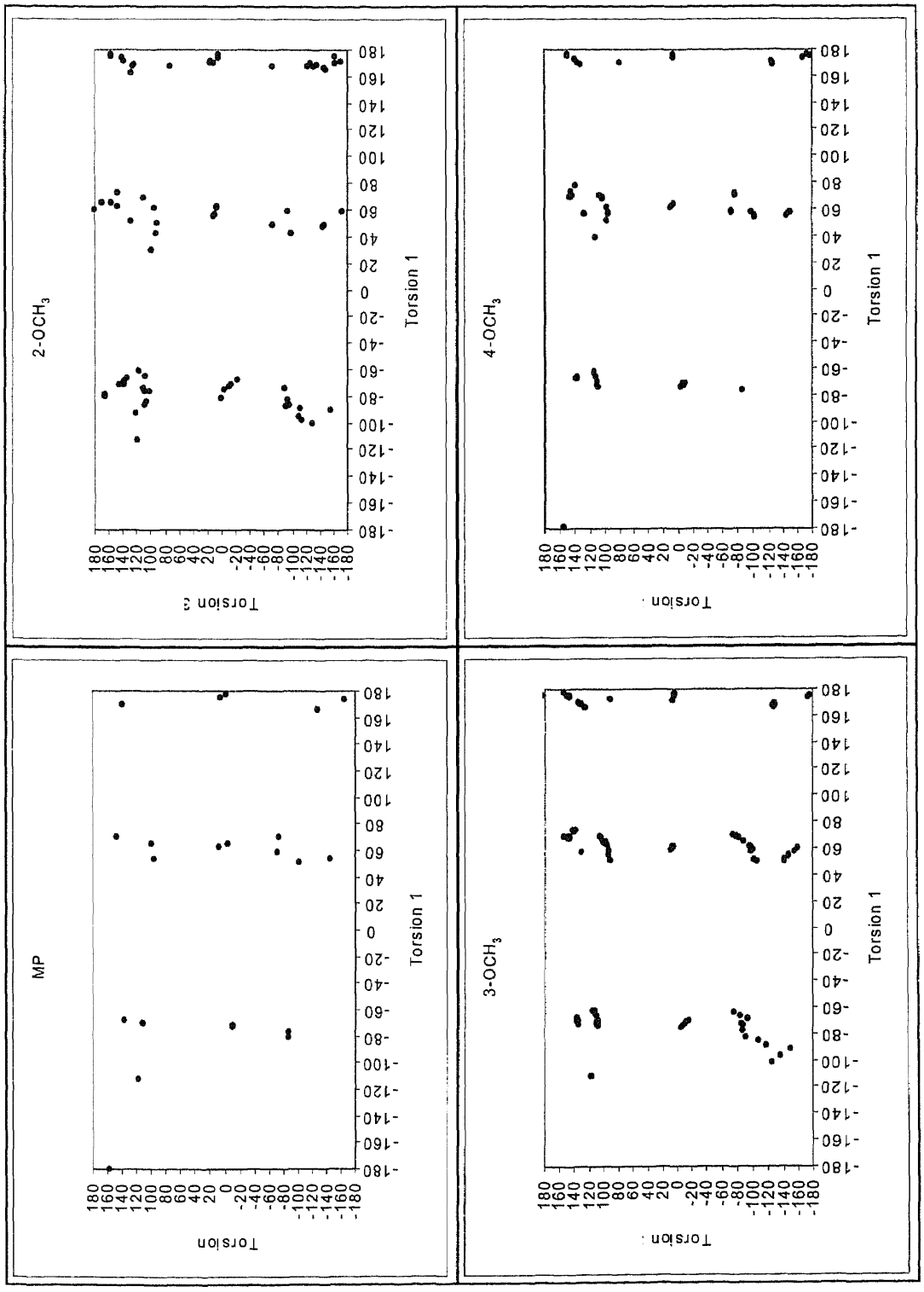
Scatter Plot Set 3b. Torsion 1 vs. Torsion 3 for MP and 2-, 3-, and 4-F



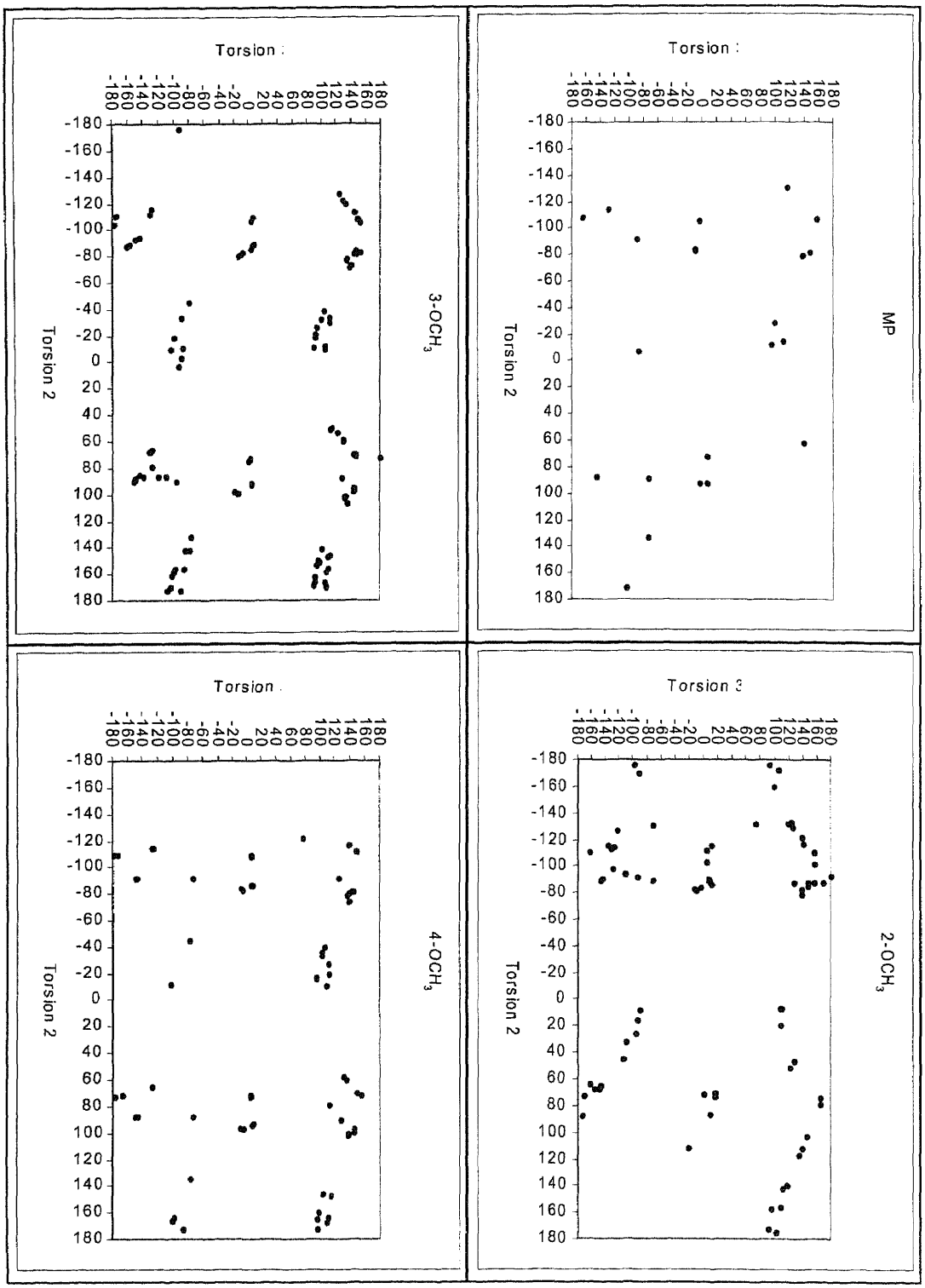
Scatter Plot Set 3c. Torsion 2 vs. Torsion 3 for MP and 2-, 3-, and 4-F



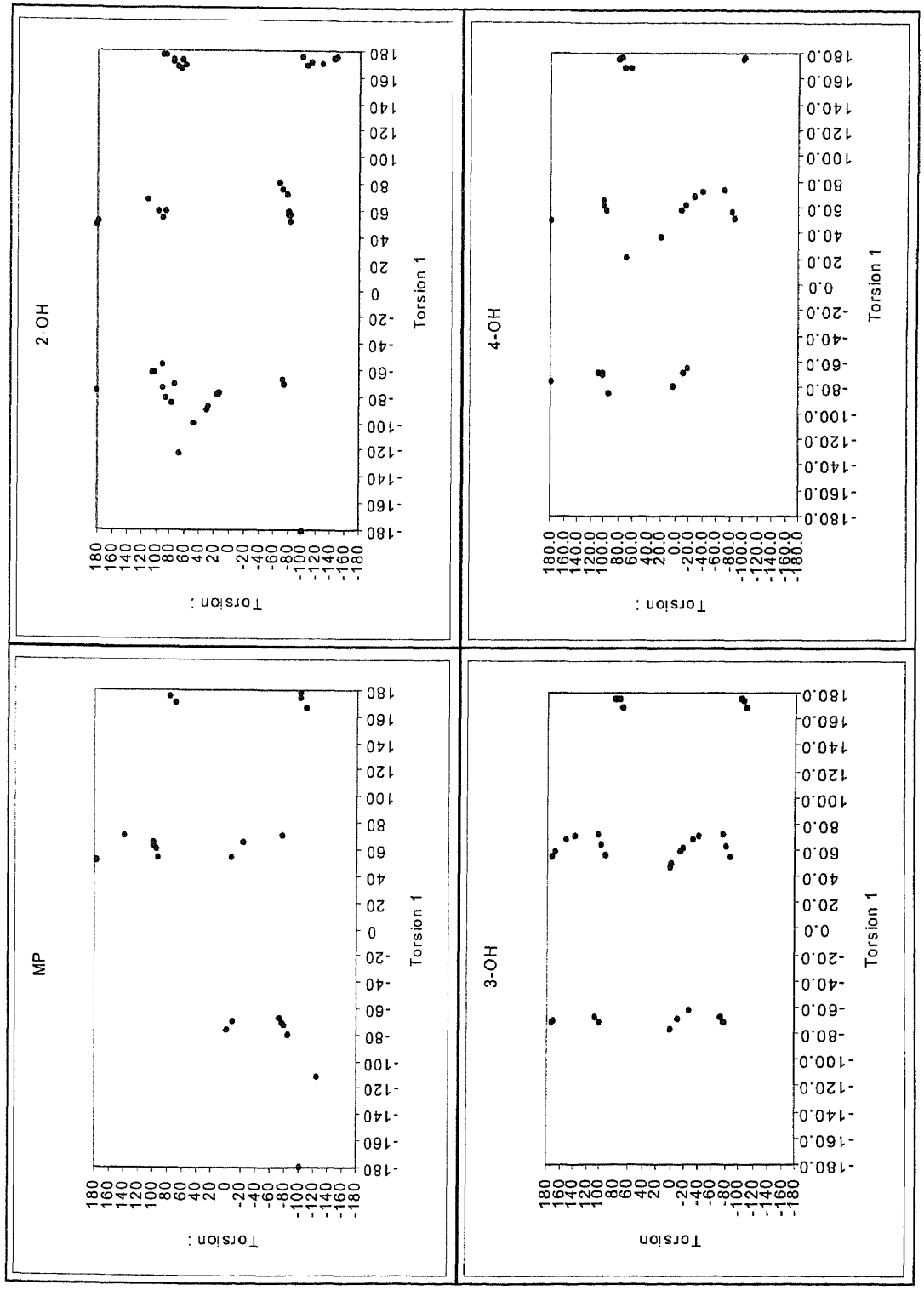
Scatter Plot Set 4a. Torsion 1 vs. Torsion 2 for MP and 2-, 3-, and 4-OCH_3



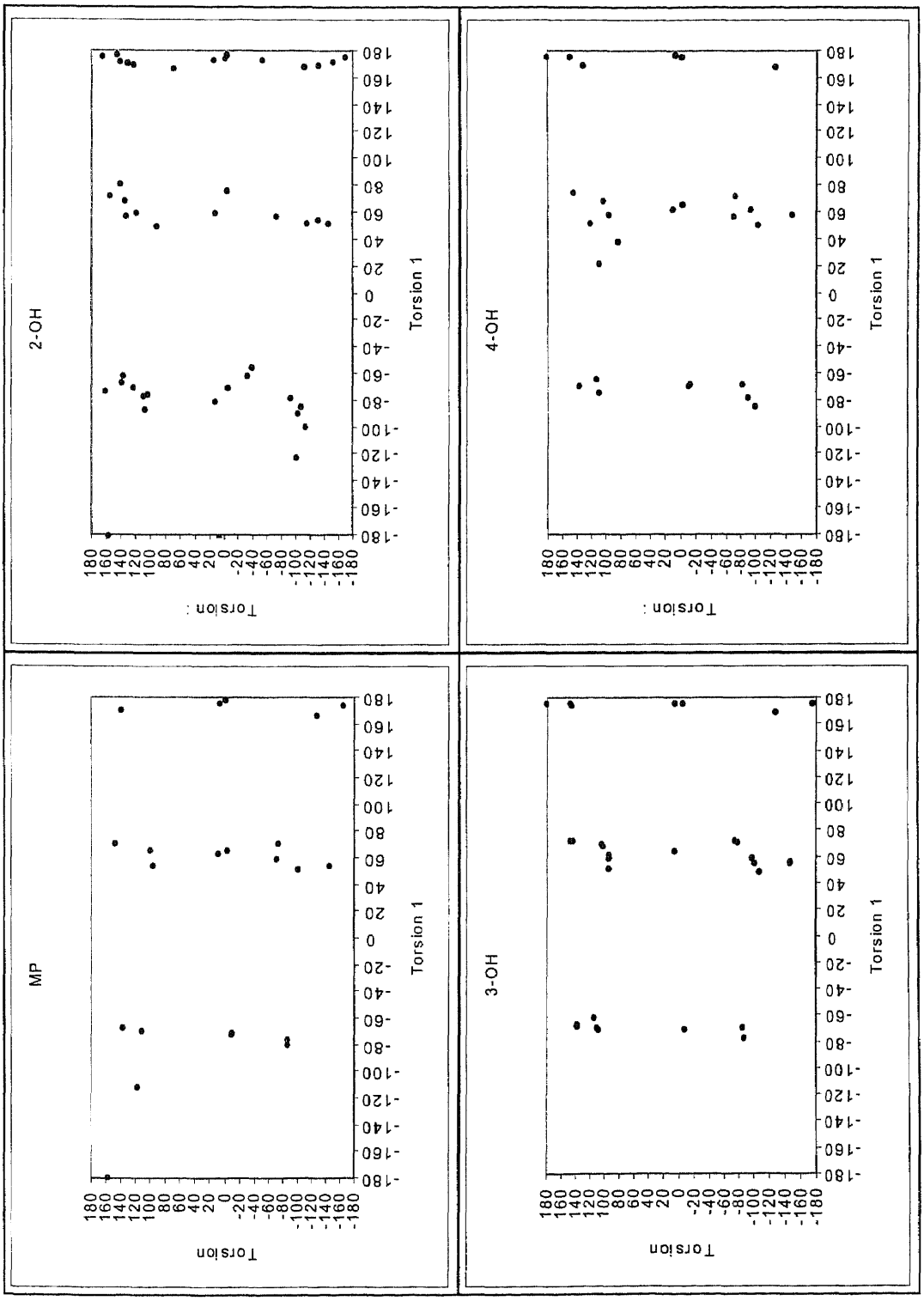
Scatter Plot Set 4b. Torsion 1 vs. Torsion 3 for MP and 2-, 3-, and 4-OCH_3



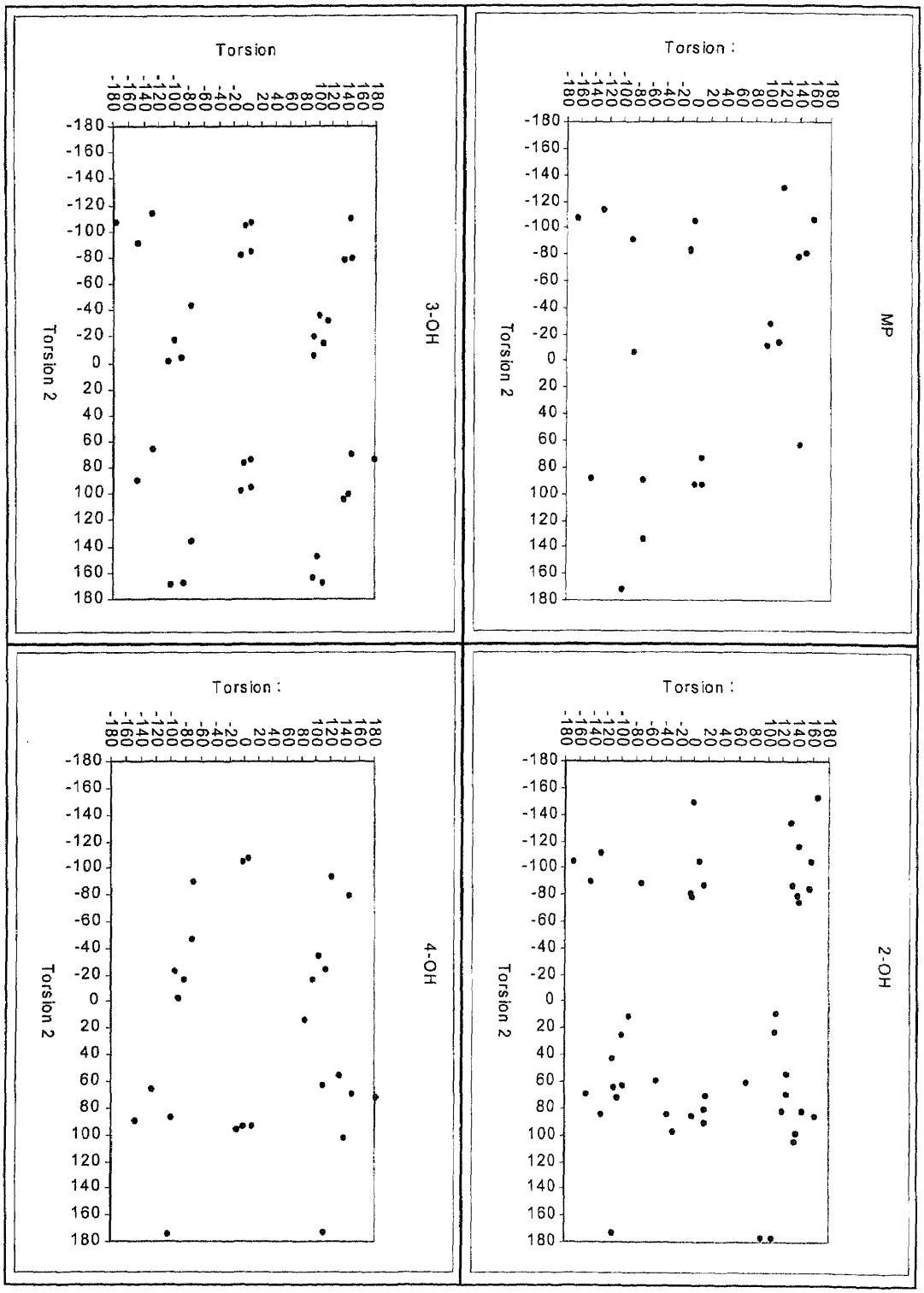
Scatter Plot Set 4c. Torsion 2 vs. Torsion 3 for MP and 2-, 3-, and 4-OCH_3



Scatter Plot Set 5a. Torsion 1 vs. Torsion 2 for MP and 2-, 3-, and 4-OH



Scatter Plot Set 5b. Torsion 1 vs. Torsion 3 for MP and 2-, 3-, and 4-OH



Scatter Plot Set 5c. Torsion 2 vs. Torsion 3 for MP and 2-, 3-, and 4-OH

APPENDIX D

COLOR FIGURES

This appendix contains three color figures for which the legend is given below.

Legend:

Plate 1 Local Energy Minima (2-Position Analogues Vs. MP). The green points signify the local energy minima of the MP conformers located at the torsional angle values indicated for the plot. The red points are the local energy minima for all conformers of all 2-position analogues.

Plate 2 CoMFA Coefficient Contour Plots (Study I, using MP conformer 14 as template). 1A: Map for CoMFA of all 30 analogues using binding data ($q^2 = 0.477$, $r^2 = 0.885$; see Table 4.2, Run 1). 1B: Map with 2-OCH_3 and 4-CF_3 removed from analysis; also using binding data ($q^2 = 0.590$, $r^2 = 0.934$; see Table 4.2, Run 2). 2A: Map for CoMFA of all 30 analogues using uptake data ($q^2 = 0.494$, $r^2 = 0.865$; see Table 4.2, Run 3). 2B: Map with 4-CF_3 removed; using uptake data ($q^2 = 0.444$, $r^2 = 0.942$; see Table 4.2, Run 4). The contours were generated using green (80), yellow (35), blue (95), and red (20). The yellow and green regions indicate areas where more steric bulk is unfavorable and favored respectively, while the blue and red regions indicate areas where more negative charge may worsen or improve bioactivity respectively.

Plate 3 CoMFA Coefficient Contour Plot (Study V, using MP conformer 20 as template). Map for CoMFA of all 30 analogues using binding data, steric/ES cutoffs of 30/30 kcal/mol, and $\sigma = 9$ ($q^2 = 0.765$, $r^2 = 0.972$; see Table 4.22). The contours were generated using green (80), yellow (35), blue (90), and red (80).

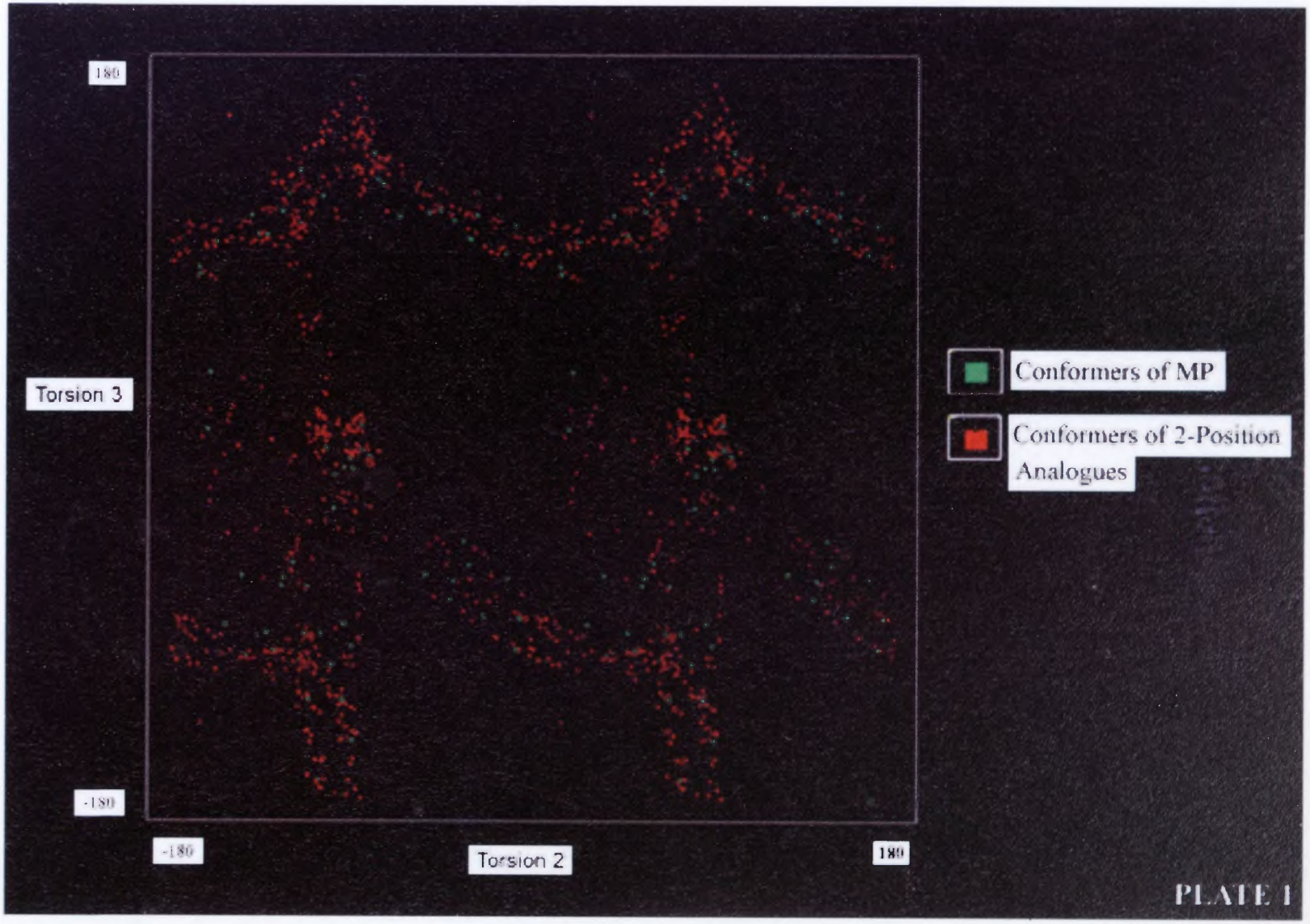

180
Torsion 3
-180
-180
Torsion 2
180
Conformers of MP
Conformers of 2-Position Analogues

PLATE 1

PLATE 2

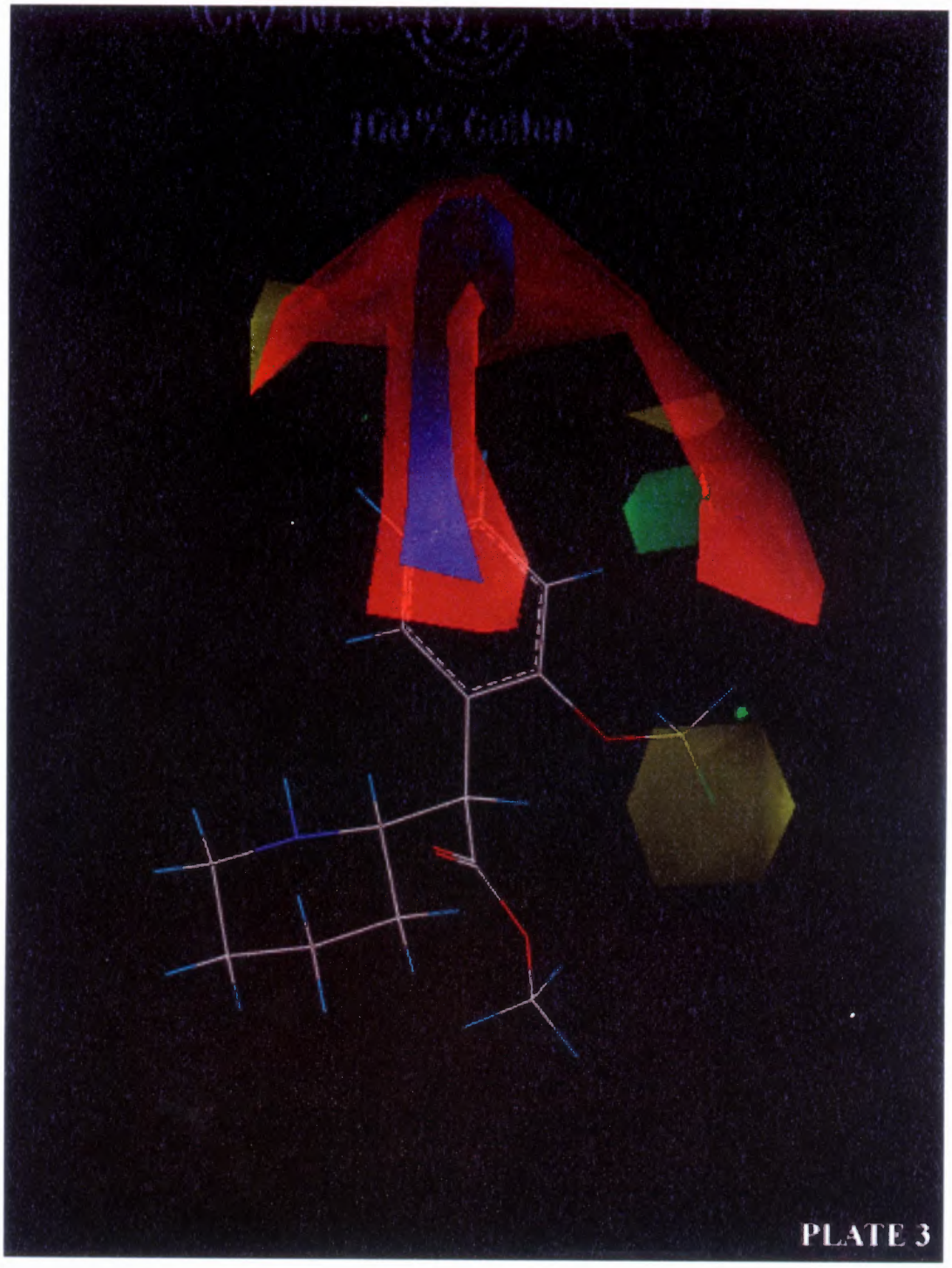


PLATE 3

REFERENCES

1. Volkow, N.D., Ding, Y.-S., Fowler, J.S., Wang, G.-J., Logan, J., Gatley, S.J., Dewey, S., Ashby, C., Liebermann, J., Hitzemann, R., and Wolf, A.P. (1995) Is Methylphenidate Like Cocaine? Studies on Their Pharmacokinetics and Distribution in the Human Brain, *Arch. Gen. Psych.* **52**, 456-463.

2. Volkow, N.D., Wang, G.-J., Gatley, S.J., Fowler, J.S., Ding, Y.-S., Logan, J., Hitzemann, R., Angrist, B., and Liebermann, J. (1996) Temporal Relationships Between the Pharmacokinetics of Methylphenidate in the Human Brain and its Behavioral and Cardiovascular Effects, *Psychopharm.* **123**, 26-33.

3. Volkow, N.D., Wang, G.-J., Fowler, J.S., Gatley, S.J., Ding, Y.-S., Logan, J., Dewey, S.L., Hitzemann, R., and Liebermann, J. (1996) Relationship Between Psychostimulant-Induced "High" and Dopamine Transporter Occupancy., *Proc. Natl. Acad. Sci. USA* **93**, 10388-10392.

4. Volkow, N.D., Wang, G.-J., Fowler, J.S., Logan, J., Angrist, B., Hitzemann, R., Liebermann, J., and Pappas, N. (1997) Effects of Methylphenidate on Regional Brain Glucose Metabolism in Humans: Relationship to Dopamine D2 Receptors, *Am. J. Psychiatry* **154**, 50-55.

5. Volkow, N.D., Ding, Y.-S., Fowler, J.S., Wang, G.-J., Logan, J., Gatley, S.J., Schlyer, D.J., and Pappas, N. (1995) A New PET Ligand for the Dopamine Transporter: Studies in the Human Brain, *J. Nucl. Med.* **36**, 2162-2168.

6. Patrick, K.S., Caldwell, R.W., Ferris, R.M., and Breese, G.R. (1987) Pharmacology of the Enantiomers of *threo*-Methylphenidate, *J. Pharm. Expt. Ther.* **241**, 152-158.

7. Ding, Y.-S., Fowler, J.S., Volkow, N.D., Logan, J., Gatley, S.J., and Sugano, Y. (1995) Carbon-11-*d-threo*-Methylphenidate Binding to Dopamine Transporter in Baboon Brain, *J. Nucl. Med.* **36**, 2298-2305.

8. Gatley, S.J., Ding, Y.-S., Volkow, N.D., Chen, R., Sugano, Y., and Fowler, J.S. (1995) Binding of *d-threo*-[^{11}C]methylphenidate to the Dopamine Transporter In Vivo: Insensitivity to Synaptic Transmission, *Eur. J. Pharmacol.* **281**, 141-149.

9. Thai, D.L., Sapko, M.T., Reiter, C.T., Bierer, D.E., and Perel, J.M. (1998) Asymmetric Synthesis and Pharmacology of Methylphenidate and Its Para-Substituted Derivatives, *J. Med. Chem.* **41**, 591-601.

10. Volkow, N.D., Wang, G.-J., Fowler, J.S., Gatley, S.J., Logan, J., Ding, Y.-S., Dewey, S.L., Hitzemann, R., Gifford, A.N., and Pappas, N.R. (1999) Blockade of Striatal Dopamine Transporters by Intravenous Methylphenidate is Not Sufficient to Induce Self-Reports of "High", *J. Pharm. Expt. Ther.* **288**, 14-20.

11. Coyle, J.Y. and Snyder, S.H. (1969) Antiparkinsonian Drugs: Inhibition of Dopamine Uptake in the Corpus Striatum as a Possible Mechanism of Action, *Science* **166**, 899-901.

12. Javitch, J.A., Blaustein, R.O., and Snyder, S.H. (1984) [^{3}H]Mazindol Binding Associated With Neuronal Dopamine and Norepinephrine Uptake Sites, *Molecular Pharmacol.* **26**, 33-44.

13. Ross, S.B. and Renyi, A.L. (1969) Inhibition of the Uptake of Tritiated 5-Hydroxy-Tryptamine in Brain Tissues, *Eur. J. Pharmacol.* **7**, 270-277.

14. Blackburn, K.J., French, P.C., and Merrils, R.J. (1967) 5-Hydroxytryptamine Uptake by Rat Brain In Vitro, *Life Sci.* **6**, 1653.

15. Horn, A.S., Cuello, C., and Miller, R.J. (1974) Dopamine in the Mesolimbic System of the Rat Brain: Endogenous Levels and the Effects of Drugs on the Uptake Mechanism and Stimulation of Adenylate Cyclase Activity, *J. Neurochem.* **22**, 265-270.

16. Leshner, A.I. (1996) Molecular Mechanisms of Cocaine Addiction, *New Engl. J. Med.* **335**, 128-129.

17. Madras, K., Fahey, M.A., Bergman, J., Canfield, D.R., and Spealman, R.D. (1989) Effects of Cocaine and Related Drugs in Nonhuman Primates. I. [^{3}H]Cocaine Binding Sites in Caudate Putamen, *J. Pharm. Expt. Ther.* **251**, 131-141.

18. Cooper, J.R., Bloom, F.E., and Roth, R.H. (1996) *The Biochemical Basis of Neuropharmacology*, 7th Ed. ed. Oxford University Press, New York.

19. Giros, B. and Caron, M.G. (1993) Molecular Characterization of the Dopamine Transporter, *TIPS* **14**, 43-49.

20. Giros, B., El Mestikawy, S., Bertrand, L., and Caron, M.G. (1991) Cloning and Functional Characterization of a Cocaine-Sensitive Dopamine Transporter, *FEBS Letters* **295**, 149-154.

21. Kilty, J.E., Lorang, D., and Amara, S.G. (1991) Cloning and Expression of a Cocaine-Sensitive Rat Dopamine Transporter, *Science* **254**, 578-579.

22. Shimada, S., Kitayama, S., Lin, C.-L., Patel, A., Nanthakumar, E., Gregor, P., Kuhar, M., and Uhl, G. (1991) Cloning and Expression of a Cocaine-Sensitive Dopamine Transporter Complementary DNA, *Science* **254**, 576-578.

23. Usdin, T.B., Mezey, E., Chen, C., Brownstein, M.J., and Hoffman, B.J. (1991) Cloning of the Cocaine-Sensitive Bovine Dopamine Transporter, *Proc. Natl. Acad. Sci. USA* **88**, 11168-11171.

24. Giros, B., El Mestikawy, S., Godinot, N., Zheng, K., Han, H., Yang-Feng, T., and Caron, M.G. (1992) Cloning, Pharmacological Characterization and Chromosome Assignment of the Human Dopamine Transporter, *Molecular Pharmacol.* **42**, 383-390.

25. Giros, B., Wang, Y.-M., Suter, S., McLeskey, S.B., Pifl, C., and Caron, M.C. (1994) Delineation of Discrete Domains for Substrate, Cocaine, and Tricyclic Antidepressant Interactions Using Chimeric Dopamine-Norepinephrine Transporters, *J. Biol. Chem.* **269**, 15985-15988.

26. Kitayama, S., Shimada, S., Xu, H., Markham, L., Donnovan, D.M., and Uhl, G.R. (1992) Dopamine Transporter Site-Directed Mutations Differentially Alter Substrate Transport and Cocaine Binding, *Proc. Natl. Acad. Sci. USA* **89**, 7782-7785.

27. Kuhar, M.J., Ritz, M.C., and Boja, J.W. (1991) The Dopamine Hypothesis of the Reinforcing Properties of Cocaine, *TINS* **14**, 299-302.

28. Ritz, M.C., Lamb, R.J., Goldberg, S.R., and Kuhar, M.J. (1987) Cocaine Receptors on Dopamine Transporters Are Related to Self-Administration of Cocaine, *Science* **237**, 1219-1223.

29. Rothman, R.B. (1990) High Affinity Dopamine Reuptake Inhibitors as Potential Cocaine Antagonists: A Strategy for Drug Development, *Life Sci.* **46**, PL17-PL21.

30. Rothman, R.B., Mele, A., Reid, A.A., Akunne, H.C., Greig, N., Thurkauf, A., De Costa, B.R., Rice, K.C., and Pert, A. (1991) GBR 12909 Antagonizes the Ability of Cocaine to Elevate Extracellular Levels of Dopamine, *Pharmacol. Biochem. Behav.* **40**, 387-397.

31. Gawin, F.H., Kleber, H.D., Byck, R., and al., e. (1989) Desipramine Facilitation of Initial Cocaine Abstinence, *Arch. Gen. Psychiatr.* **46**, 117-121.

32. Schottenfeld, R., Caroll, K., and Rounsaville, B. (1993) Comorbid Psychiatric Disorders and Cocaine Abuse, *NIDA Res. Monogr.* **135**, 31-47.

33. Mendelson, J.H. and Mello, N.K. (1996) Management of Cocaine Abuse and Dependence, *New Engl. J. Med.* **334**, 965-972.

34. Kleber, H.D. (1995) Pharmacotherapy, Current and Potential, for the Treatment of Cocaine Dependence, *Clinical Neuropharmacology* **18, Suppl. 1**, S96-S109.

35. Carroll, F.I., Lewin, A.H., Boja, J.W., and Kuhar, M.J. (1992) Cocaine Receptor: Biochemical Characterization and Structure-Activity Relationships of Cocaine Analogs at the Dopamine Transporter, *J. Med. Chem.* **35**, 969-981.

36. Chaudieu, I., Vignon, J., Chicheportiche, M., Kamenka, J.M., Trouiller, G., and Chicheportiche, R. (1989) Role of the Aromatic Group in the Inhibition of Phencyclidine Binding and Dopamine Uptake by PCP Analogues., *Pharmacol. Biochem. Behav.* **32**, 699-705.

37. Van der Zee, P., Koger, H.S., Gootjes, J., and Hespe, W. (1980) Aryl 1,4-dialk(en)ylpiperizines as Selective and Very Potent Inhibitors of Dopamine Uptake, *Eur. J. Med. Chem.* **15**, 363-370.

38. Dubocovich, M.L. and Zahniser, N.R. (1985) Binding Characteristics of the Dopamine Uptake Inhibitor [^{3}H]Nomifensine to Striatal Membranes, *Biochem. Pharmacol.* **34**, 1137-1144.

39. Janowsky, A., Schweri, M.M., Berger, P., Long, R., Skolnick, P., and Paul, S.M. (1985) The Effects of Surgical and Chemical Lesions on Striatal [^{3}H]*threo*-(±)-Methylphenidate Binding: Correlation with [^{3}H]Dopamine Uptake, *Eur. J. Pharmacol.* **108**, 187-191.

40. Janowsky, A., Berger, P., Vocci, F., Labarca, R., Skolnick, P., and Paul, S.M. (1986) Characterization of Sodium-Dependent [^{3}H]GBR-12935 Binding in Brain: a Radioligand for Selective Labelling of the Dopamine Transport System, *J. Neurochem.* **46**, 1272-1276.

41. Andersen, P.H. (1987) Biochemical and Pharmacological Characterization of [^{3}H]GBR 12935 Binding In Vitro to Rat Striatal Membranes: Labeling of the Dopamine Uptake Complex, *J. Neurochem.* **48**, 1887-1896.

42. Vignon, J., Pinet, V., Cerruti, C., Kamenka, J.-M., and Chicheportiche, R. (1988) [^{3}H]*N*-[1-(2-Benzo(*b*)thiophenyl)cyclohexyl] Piperidine ([^{3}H]BTCP: a New Phencyclidine Analog Selective for the Dopamine Uptake Complex, *Eur. J. Pharmacol.* **148**, 427-436.

43. Van der Zee, P. and Hespe, W.A. (1978) Comparison of the Inhibitory Effects of Aromatic Substituted Benzhydryl Ethers on the Uptake of Catecholamines and Serotonin into Synaptosomal Preparations in the Rat Brain, *Neuropharmacol.* **17**, 483-490.

44. Meltzer, P.C., Liang, A.Y., and Madras, B.K. (1996) 2-Carbomethoxy-3-(diarylmethoxy)-1αH, 5αH-tropane Analogs: Synthesis and Inhibition of Binding at the Dopamine Transporter and Comparison with Piperazines of the GBR Series, *J. Med. Chem.* **39**, 371-379.

45. Meltzer, P.C., Liang, A.Y., and Madras, B.K. (1994) The Discovery of an Unusually Selective and Novel Cocaine Analog: Difluoropine. Synthesis and Inhibition of Binding at Cocaine Recognition Sites, *J. Med. Chem.* **37**, 2001-2010.

46. Newman, A.H., Allen, A.C., Izenwasser, S., and Katz, J.L. (1994) Novel 3α-(Diphenylmethoxy)tropane Analogs: Potent Dopamine Uptake Inhibitors Without Cocaine-Like Behavioral Profiles, *J. Med. Chem.* **37**, 2258-2261.

47. Newman, A.H., Kline, R.H., Allen, A.C., Izenwasser, S., George, C., and Katz, J.L. (1995) Novel 4'-Substituted and 4',4"-Disubstituted 3α-(Diphenylmethoxy)tropane Analogs as Potent and Selective Dopamine Uptake Inhibitors, *J. Med. Chem.* **38**, 3933-3940.

48. Katz, J.L., Newman, A.H., and Izenwasser, S. (1997) Relations Between Heterogeneity of Dopamine Transporter Binding and Function and the Behavioral Pharmacology of Cocaine, *Pharmacol. Biochem. Behav.* **57**, 505-512.

49. Kline, R.H., Izenwasser, S., Katz, J.L., Joseph, D.B., Bowen, W.D., and Newman, A.H. (1997) 3'-Chloro-3α-(diphenylmethoxy)tropane but Not 4'-Chloro-3α–(diphenylmethoxy)tropane Produces a Cocaine-Like Behavioral Profile, *J. Med. Chem.* **40**, 851-857.

50. Agoston, G.E., Wu, J.H., Izenwasser, S., George, C., Katz, J., Kline, R.H., and Newman, A.H. (1997) Novel N-Substituted 3α–[Bis(4'-fluorophenyl)methoxyltropane Analogues: Selective Ligands for the Dopamine Transporter, *J. Med. Chem.* **40**, 4329-4339.

51. Reith, M.E.A., Sershen, H., and Lajtha, A. (1980) [^{3}H]Cocaine Binding in the Central Nervous System of Mouse, *Life Sci.* **27**, 1055-1062.

52. Kennedy, L.T. and Hanbauer, I. (1983) Sodium-Sensitive Cocaine Binding to Rat Striatal Membrane: Possible Relationship to Dopamine Uptake Sites, *J. Neurochem.* **41**, 172-178.

53. Schoemaker, H., Pimoule, C., Arbilla, S., Scatton, B., Javoy-Agid, F., and Langer, S.Z. (1985) Sodium Dependent [^{3}H]Cocaine Binding Associated with Dopamine Uptake Sites in the Rat Striatum and Human Putamen Decrease after Dopaminergic Denervation and in Parkinson's Disease, *Naunyn-Schmiedeberg's Ach. Pharmacol.* **329**, 227-235.

54. Calligaro, D.O. and Eldefrawi, M.E. (1987) Central and Peripheral Cocaine Receptors, *J. Pharm. Expt. Ther.* **243**, 61-67.

55. Calligaro, D.O. and Eldefrawi, M.E. (1988) High Affinity Stereospecific Binding of [^{3}H]Cocaine in Striatum and Its Relationship to the Dopamine Transporter, *Memb. Biochem.* **7**, 87-106.

56. Madras, B.K., Spealman, R.D., Fahey, M.A., Neumeyer, J.L., Saha, J.K., and Milius, R.A. (1989) Cocaine Receptors Labeled by [^{3}H]2β–Carbomethoxy-3β–(4–fluorophenyl)tropane, *Molecular Pharmacol.* **36**, 518-524.

57. Ritz, M.C., Boja, J.W., Grigoriadis, D.E., Zacek, R., Carroll, F.I., Lewin, A.H., and Kuhar, M.J. (1990) [^{3}H]WIN 35,065-2: A Ligand for Cocaine Receptors in Striatum, *J. Neurochem.* **55**, 1556-1562.

58. Boja, J.W., Patel, A., Carroll, F.I., Rahman, M.A., Phillip, A., Lewin, A.H., Kopajtic, T.A., and Kuhar, M.J. (1991) [^{125}I]RTI-55: A Potent Ligand for Dopamine Transporters, *Eur. J. Pharmacol.* **194**, 133-134.

59. Vaughan, R.A., Brown, V.L., McCoy, M.T., and Kuhar, M.J. (1996) Species- and Brain Region-Specific Dopamine Transporters: Immunological and Glycosylation Characteristics, *J. Neurochem.* **66**, 2146-2152.

60. Carroll, F.I., Lewin, A.H., and Kuhar, M.J. (1998) 3β-Phenyl-2β-Substituted Tropanes - An SAR Analysis, *Med. Chem. Res.* **8**, 59-65.

61. Deutsch, H.M. (1998) Structure-Activity Relationships for Methylphenidate Analogs and Comparisons to Cocaine and Tropanes, *Med. Chem. Res.* **8**, 91-99.

62. Meltzer, P.C., Blundell, P., and Madras, B.K. (1998) Structure Activity Relationships of Inhibition of the Dopamine Transporter by 3-Arylbicyclo[3.2.1]octanes, *Med. Chem. Res.* **8**, 12-34.

63. Newman, A.H. and Agoston, G.E. (1998) Novel Benztropine [3α-(diphenylmethoxy)tropane] Analogs as Probes for the Dopamine Transporter, *Curr. Med. Chem.* **5**, 305-319.

64. Kier, L.B. (1971) *Molecular Orbital Theory in Drug Research*, Academic, New York.

65. Pannizzon, L. (1944) *Hel. Chim. Acta.* **27**, 1748.

66. Meier, R., Gross, F., and Tripod, J. (1954) Ritalin, A New Synthetic Compound with Specific Central Nervous System Activity, *Klin. Wochenschr.* **32**, 445-450.

67. Deutsch, H.M., Shi, Q., Gruszecka-Kowalik, E., and Schweri, M. (1996) Synthesis and Pharmacology of Potential Cocaine Antagonists. 2. Structure-Activity Relationship Studies of Aromatic Ring-Substituted Methylphenidate Analogs, *J. Med. Chem.* **39**, 1201-1209.

68. Schweri, M.M., Skolnick, P., Rafferty, M.F., Rice, K.C., Janowsky, A.J., and Paul, S.M. (1985) [^{3}H]*Threo*-(±)-Methylphenidate Binding to 3,4-Dihydroxyphenylethylamine Uptake Sites in Corpus Striatum: Correlation with the Stimulant Properties of Ritalinic Acid Esters, *J. Neurochem.* **45**, 1062-1070.

69. Schweri, M.M. (1990) *N*-ethylmaleimide Irreversibly Inhibits the Binding of [^{3}H]*threo*-(±)-Methylphenidate to the Stimulant Recognition Site, *Neuropharmacol.* **29**, 901-908.

70. Schweri, M.M. (1994) Mercuric Chloride and *p*-Chloromercuriphenylsulfonate Exert a Biphasic Effect on the Binding of the Stimulant [^{3}H]Methylphenidate to the Dopamine Transporter, *Synapse* **16**, 188-194.

71. Froimowitz, M., Deutsch, H.M., Shi, Q., Wu, K.-M., Glaser, R., Adin, I., George, C., and Schweri, M.M. (1997) Further Evidence for a Dopamine Reuptake Pharmacophore. The Effect of N-methylation on *Threo*-Methylphenidate and Its Analogs, *Bioorganic & Medicinal Chemistry Letters* **7**, 1213-1218.

72. Glaser, R., Adin, I., Shiftan, D., Shi, Q., Deutsch, H.M., George, G., Wu, K.-M., and Froimowitz, M. (1998) Solution and Solid-State Conformational and Structural Analysis of the *N*-Methyl Derivatives of (±)-*Threo*-Methylphenidate, (±)-*Erythro*-Methylphenidate and (±)-*Threo-p*-methyl-methylphenidate HCl Salts, *J. Org. Chem.* **63**, 1785-1794.

73. Froimowitz, M., Patrick, K.S., and Cody, V. (1995) Conformational Analysis of Methylphenidate and Its Structural Relationship to Other Dopamine Blockers Such as CFT, *Pharm. Res.* **12**, 1430-1434.

74. Cramer, I., R.D., Patterson, D.E., and Bunce, J.D. (1988) Comparative Molecular Field Analysis (CoMFA).1. Effect of Shape on Binding of Steroids to Carrier Proteins, *J. Am. Chem. Soc.* **110**, 5959-5967.

75. Carroll, F.I., Gao, Y., Rahman, M.A., Abrams, P., Parham, K., Lewin, A.H., Boja, J.W., and Kuhar, M.J. (1991) Synthesis, Ligand Binding, QSAR, and CoMFA Study of 3β-(*p*-Substituted phenyl)tropane-2β–carboxylic Acid Methyl Esters, *J. Med. Chem.* **34**, 2719-2725.

76. Lieske, S.F., Yang, B., Eldefrawi, M.E., MacKerell, J., A.D., and Wright, J. (1998) (-)-3ß-Substituted Ecgonine Methyl Esters as Inhibitors for Cocaine Binding and Dopamine Uptake, *J. Med. Chem.* **41**, 864-876.

77. Carroll, F.I., Mascarella, S.W., Kuzemko, M.A., Gao, Y., Abraham, P., Lewin, A.H., Boja, J.W., and Kuhar, M.J. (1994) Synthesis, Ligand Binding, and QSAR (CoMFA and Classical) Study of 3β-(3'-Substituted phenyl)-, 3β-(4'-Substituted phenyl)-, and 3β-(3',4'-Disubstituted phenyl)tropane-2β–carboxylic Acid Methyl Esters, *J. Med. Chem.* **37**, 2865-2873.

78. Allen, M.S., Tan, Y.C., Trudell, M.L., Narayanan, K., Schindler, L.R., Martin, M.J., Shultz, C., Hagen, T.J., Koehler, K.F., Codding, P.W., Skolnick, P., and Cook, J.M. (1990) Synthetic and Computer-Assisted Analyses of the Pharmacophore for the Benzodiazepine Receptor Inverse Agonist Site, *J. Med. Chem.* **33**, 343-2357.

79. Thomas, B.F., Compton, D.R., Martin, B.R., and Semus, S.F. (1991) Modeling the Cannabinoid Receptor - A 3-dimensional Quantitative Structure-Activity Analysis, *Molecular Pharmacol.* **40**, 656-665.

80. McFarland, J.W. (1992) Comparative Molecular Field Analysis of Anticoccidial Triazines, *J. Med. Chem.* **35**, 2543-2550.

81. Klebe, G. and Abraham, U. (1993) On the Prediction of Binding Properties of Drug Molecules by Comparative Molecular Field Analysis, *J. Med. Chem.* **36**, 70-80.

82. Teitler, M., Scheick, C., Howard, P., Sullivan III, J.E., Iwamura, T., and Glennon, R.A. (1997) 5-HT_{5A} Serotonin Receptor Binding: A Preliminary Structure-Affinity Investigation, *Med. Chem. Res.* **7**, 207-218.

83. Carrieri, A., Brasili, L., Leonetti, F., Pigini, M., Gianella, M., Bousquet, P., and Carotti, A. (1997) 2-D and 3-D Modeling of Imidazoline Receptor Ligands: Insights Into Pharmacophore, *Bioorg. Med. Chem.* **5**, 843-856.

84. Kroemer, R.T., Koutsilieri, E., Hecht, P., Leidel, K.R., Riederer, P., and Kornhuber, J. (1998) Quantitative Analysis of the Structural Requirements for Blockade of the *N*-Methyl-D-aspartate Receptor at the Phencyclidine Binding Site, *J. Med. Chem.* **41**, 393-499.

85. Pajeva, I. and Weise, M. (1998) Molecular Modeling of Phenothiazines and Related Drugs as Multidrug Resistance Modifiers: A Comparative Molecular Field Analysis Study, *J. Med. Chem.* **41**, 1815-1826.

86. Wilcox, R.E., Tseng, T., Brusniak, M.-Y.K., Ginsburg, B., Pearlman, R.S., Teeter, M., DuRand, C., Starr, S., and Neve, K.A. (1998) CoMFA-Based Prediction of Agonist Affinities at Recombinant D1 vs D2 Dopamine Receptors, *J. Med. Chem.* **41**, 4385-4399.

87. Debnath, A.K. (1999) Comparative Molecular Field Analysis (CoMFA) of a Series of Symmetrical Bis-Benzamide Cyclic Urea Derivatives as HIV-1 Protease Inhibitors, *J. Chem. Inf. Comp. Sci.* **38**, 761-767.

88. Kim, K.H. (1998) List of CoMFA References, 1993-1997, in Kubinyi, H., Folkers, G., and Martin, Y.C., (eds.), *3D QSAR in Drug Design: Recent Advances,* Kluwer Academic, Dordrecht, pp. 317-333.

89. Kim, K.H. (1995) Comparative Molecular Field Analysis (CoMFA), in Dean, P.M., (eds.), *Molecular Similarity in Drug Design,* Blackie, London, pp. 291-331.

90. Martin, Y.C., Kim, K.H., and Lin, C.T. (1996) Comparative Molecular Field Analysis: CoMFA, in Charton, M., (eds.), *Advances in Quantitative Structure-Property Relationships,* JAI Press, Greenwich, CT, pp. 1-52.

91. Kim, K.H., Greco, G., and Novellino, E. (1998) A Critical Review of Recent CoMFA Applications, in Kubinyi, H., Folkers, G., and Martin, Y.C., (eds.), *3D QSAR in Drug Design: Recent Advances,* Kluwer Academic, Dordrecht, pp. 257-315.

92. Martin, Y.C. (1998) 3D QSAR: Current State, Scope, and Limitations, in Kubinyi, H., Folkers, G., and Martin, Y.C., (eds.), *3D QSAR in Drug Design: Recent Advances,* Kluwer Academic, Dordrecht, pp. 3-23.

93. Corelli, F., Manetti, F., Tafi, A., Campiani, G., Nacci, V., and Botta, M. (1997) Diltiazem-Like Calcium Entry Blockers: A Hypothesis of the Receptor-Binding Site Based on a Comparative Molecular Field Analysis, *J. Med. Chem.* **40**, 125-131.

94. Debnath, A.K. (1999) Three-Dimensional Quantitative Structure-Activity Relationship Study on Cyclic Urea Derivatives as HIV-1 Protease Inhibitors: Application of Comparative Molecular Field Analysis, *J. Med. Chem.* **42**, 249-259.

95. Vieth, M., Hirst, J.D., and Brooks, I., C.L. (1998) Do Active Site Conformations of Small Ligands Correspond to Low Free-Energy Solution Structures?, *J. Comp.-Aided Mol. Des.* **12**, 563-572.

96. Miyashita, Y., Li, Z., and Sasaki, S. (1993) Chemical Pattern Recognition and Multivariate Analysis for QSAR Studies, *Trends Anal. Chem.* **12**, 50-60.

97. Szporny, L. and Görõg, P. (1961) Investigations into the Correlations between Monoamine Oxidase Inhibition and Other Effects due to Methylphenidate and its Stereoisomers., *Biochem. Pharmacol.* **8**, 253-268.

98. Rometsch, R. (1958)*Process for the Conversion of Stereoisomers*. U.S.

99. Buckner, C.K., Patil, P.N., Tye, A., and Malspeis, L. (1969) Steric Aspects of Adrenergic Drugs. XII. Some Peripheral Effects of (+/-)-erythro- and (+/-)-threo-Methylphenidate, *J. Pharm. Exp. Ther.* **166**, 308-319.

100. Maxwell, R.A., Chaplin, E., Eckhardt, S.B., Soares, J.R., and Hite, G. (1970) Conformational Similarities Between Molecular Models of Phenethylamine and of Potent Inhibitors of the Uptake of Tritiated Norepinephrine by Adrenergic Nerves in Rabbit Aorta, *J. Pharm. Expt. Ther.* **173**, 158-165.

101. Patrick, K.S., Kilts, C.D., and Breese, G.R. (1981) Synthesis and Pharmacology of Hydroxylated Metabolites of Methylphenidate, *J. Med. Chem.* **29**, 1237-1240.

102. Eckerman, D.A., Moy, S.S., Perkins, A.N., Patrick, K.S., and Breese, G.R. (1991) Enantioselective Behavioral Effects of Methylphenidate in Rats, *Parmacol. Biochem. Behav.* **40**, 875-880.

103. Shafi'ee, A. and Hite, G. (1969) The Absolute Configuration of the Pheniramines, Methylphenidates, and Pipradrols, *J. Med. Chem.* **12**, 266-270.

104. Saunders, M., Houk, K.N., Wu, Y.D., Still, W.C., Lipton, M., Chang, G., and Guida, W.C. (1990) Conformations of Cycloheptadecane - A Comparison of Methods for Conformational Searching, *J. Am. Chem. Soc.* **112**, 1419-1427.

105. Ngo, J.T. and Karplus, M. (1997) Pseudosystematic Conformational Search. Application to Cycloheptadecane, *J. Am. Chem. Soc.* **119**, 5657-5667.

106. Goodman, J.M. and Still, W.C. (1991) An Unbounded Systematic Search of Conformational Space, *J. Comput. Chem.* **12**, 1110-1117.

107. Gasteiger, J. and Marsili, M. (1980) *Tetrahedron* **36**, 3219-3228.

108. Gasteiger, J. and Marsili, M. (1981) *Organ. Magn. Reson.* **15**, 353-360.

109. Marsili, M. and Gasteiger, J. (1980) *Croat. Chem. Acta.* **53**, 601-614.

110. Streitwieser, A. (1961) *Molecular Orbital Theory for Organic Chemists*, Wiley, NY.

111. Purcel, W.P. and Singer, J.A. (1967) *J. Chem. Eng. Data* **12**, 235-246.

112. Myers, A.M., Charifson, P.S., Owens, C.E., Kula, N.S., McPhail, A.T., Baldessarini, R.J., Booth, R.G., and Wyrick, S.D. (1994) Conformational Analysis Pharmacophore Identification and Comparative Molecular Field Analysis of Ligands for the Neuromodulatory Sigma3 Receptor, *J. Med. Chem.* **37**, 4109-4117.

113. Yliniemela, A., Konschin, H., Neagu, C., Pajunn, A., Hase, T., Brunow, G., and Teleman, O. (1995) Design and Sythesis of a Transition State Analog for the Ene Reaction Between Maleimide and 1-Alkenes, *J. Am. Chem. Soc.* **117**, 5120-5126.

114. Klebe, G. (1993) Structural Alignment of Molecules, in Kubinyi, H., (eds.), *3D-QSAR in Drug Design. Theory, Methods and Applications,* ESCOM Science Publishers, Leiden, The Netherlands, pp. 173-199.

115. Bures, M.G. (1997) Recent Techniques and Applications in Pharmacophore Mapping, in Chariforn, P.S., (eds.), *Practical Applications of Computer Aided Drug Design,* Marcel Dekker, New York, pp. 39-72.