ABSTRACT

ANALYSIS OF MOLECULAR CONFORMATIONS USING RELATIVE PLANES

by

Deepa S Pai

Ring substructures of a drug usually participate actively in binding to the receptor. It is necessary to study the spatial relationship of these molecular recognition features in order to determine the pharmacophore of the drug. This is a particularly difficult problem when the drug is a flexible molecule with many energetically accessible conformations.

In this research an innovative approach to calculate the relative displacement and orientation of every possible pair of rings in a given molecule was designed, tested, and implemented in the "Planes" program. Planes were defined from each of the ring substructures and the displacement and rotation of one ring with respect to the other was calculated. This approach was derived from the guidelines of 3DNA, a robust program that calculates base-pair and base-step parameters of nucleic acids. The Planes program was subsequently utilized to analyze the conformations of DM324, a flexible analog of GBR 12909, a drug potentially useful in the treatment of cocaine abuse.

The present work suggests that the Planes program could be potentially useful in the description of the relative orientation of pharmacophore features if the molecular conformations could first be classified by their orientation relative to the central piperazine ring.

ANALYSIS OF MOLECULAR CONFORMATIONS USING RELATIVE PLANES

by
Deepa S Pai

A 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 Computational Biology

Department of Computer Science

May 2004

APPROVAL PAGE

ANALYSIS OF MOLECULAR CONFORMATIONS USING RELATIVE PLANES

Deepa S Pai

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

Dr. Barry Cohen, Committee Member — Date
Assistant Professor,
Department of Computer Science, NJIT

Dr. Qun Ma, Committee Member — Date
Assistant Professor,
Department of Computer Science, NJIT

BIOGRAPHICAL SKETCH

Author:	Deepa S Pai
Degree:	Master of Science
Date:	May 2004

Undergraduate and Graduate Education:

- Master of Science in Computational Biology,
  New Jersey Institute of Technology, Newark, NJ, 2004

- Bachelor of Engineering in Computers
  Manipal Institute of Technology, Manipal, India, 1998

Major: Computational Biology

To my Amma and Aanu

ACKNOWLEDGMENT

I am extremely grateful to my thesis advisor Dr. Carol Venanzi for her valuable support, guidance and assistance throughout my research project. I would especially like to thank her for her improvements and suggestions that have enriched this document. I am also thankful to Dr. Barry Cohen and Dr. Qun Ma for being on my thesis committee. Immense thanks and appreciation to Dr. Wilma Olson of Department of Chemistry at Rutgers, The State University of New Jersey, and her group: Dr. A. R. Srinivasan, Dr. A. Matsumoto, Dr. Xiang-Jun Lu for providing us with the codes for the 3DNA program, for their precious time, thoughtful insight and suggestions that helped towards the completion of the project.

I am thankful to Rohan Woodley for modifying the 3DNA program. I would like to express my deepest appreciation to Milind Misra and Deepangi Pandit for all their help. I also wish to thank Dr. William Skawinski for his assistance.

Thanks to the faculty and staff of the Computational Biology program and Office of Graduate Studies for their assistance. My sincere gratitude to my family: my husband and our parents for their constant encouragement, reassurance and patience.

TABLE OF CONTENTS

TABLE OF CONTENTS
(continued)

LIST OF FIGURES

Chapter **Page**

LIST OF FIGURES
(continued)

LIST OF FIGURES
(continued)

Chapter Page

LIST OF FIGURES
(continued)

LIST OF TABLES

LIST OF TABLES
(continued)

Chapter **Page**

CHAPTER 1

INTRODUCTION

1.1 Motivation

Identifying the conformation of a drug that binds optimally to the receptor (i.e. the bioactive conformation) plays a vital role in drug design. Many conformational searching methods generate a large number of conformations for a flexible, large molecule. Usually one uses techniques to find a particular quality of interest such as the lowest energy conformation or the free energy of the system. It has been shown by the Venanzi group (unpublished data) and others (Bernard et al., 2003) that conformations other than the lowest energy conformation have to be considered to determine a pharmacophore model of the molecule. Analyzing this vast set of conformations turns out to be very challenging. It becomes important to understand how these conformations are related to each other in order to derive a set of putative bioactive conformations. Multivariate statistical methods such as principal component analysis and data reduction techniques such as clustering are often used to identify families of conformations. A representative member from each family is then chosen as a possible candidate for the bioactive conformation. These representative conformations are used as input to pharmacophore modeling procedures such as Comparative Molecular Field Analysis (CoMFA) (Cramer et al., 1988) that calculate the relationship between molecular properties and biological activity. Since the results of CoMFA analysis can vary widely depending on the conformation used for the calculation of molecular properties, the conformation that gives the best fit to the biological data is assumed to be the bioactive (or binding) conformation. This information can then be used to design more rigid analogs that are "frozen" into the optimal binding conformation.

Drugs usually contain one or more ring structures, such as aromatic or heterocyclic rings. These ring structures often contain important molecular recognition features (or pharmacophore features), such as a hydrophobic group or a hydrogen bond donor, that actively participate in binding by controlling the alignment of complementary steric and electrostatic features of the drug and the receptor (see the Pharmacophore Modeling section below). Understanding the spatial relationship of these molecular recognition features in the bioactive conformation can aid in the design of optimal binding agents.

The objective of this research is to implement a novel technique for partitioning large number of conformations into families based on the relative orientation of the rings. The significance of this work is that it describes and classifies conformations of drug molecules by the most important chemical features for binding: the relative orientation of the pharmacophore features.

The procedure is broken down as follows:

(1) Adapt the 3DNA technique, which analyzes the base pair geometry in nucleic acid structures, to evaluate the relative geometry of ring structures in *any* molecule. This allows one to analyze molecular conformations based on the relative displacement and orientation of the pharmacophore features contained in the rings. The generalized version of the 3DNA(Lu et al., 2003) program is

dubbed the Planes program because it calculates the distances and angles between molecular planes that contain the rings.

(2) Use the Planes program on conformations of DM324, an analog of GBR 12909, a dopamine reuptake inhibitor currently in Phase II clinical trials as a possible treatment for cocaine dependence.

(3) Analyze the results to see if these values help in identifying families of conformations. Representative conformers from these families will be used by others in the Venanzi group, Department of Chemistry and Environmental Science, NJIT for pharmacophore modeling studies.

This thesis is divided into four chapters and four appendices

Chapter 1 gives a brief overview of molecular modeling to introduce a novice reader to the main concepts that are applied in this thesis. This is followed by background information on the dopamine reuptake inhibitors, cocaine and GBR 12909.

Chapter 2 describes the 3DNA program and the modifications made to generalize the algorithm in order to calculate the planes parameters for any molecule.

Chapter 3 presents the implementation and verification of the Planes program followed by its application to analyze DM324 conformations.

Chapter 4 presents the results, analysis, conclusions, and suggestions for further study.

1.2 Molecular Modeling

Molecular modeling techniques are widely utilized in drug discovery. Development of a new drug is both extensive and expensive. A new compound must produce the desired effects with minimal side effects and must be reasonably better than any existing treatment. A lead compound is first chosen that is active in an assay. It is then modified to enhance its potency and selectivity, to ensure that it is non-toxic and posses required transport characteristics that enable it to pass through the cell membrane to reach its target and bind to it. (Leach, 1996).

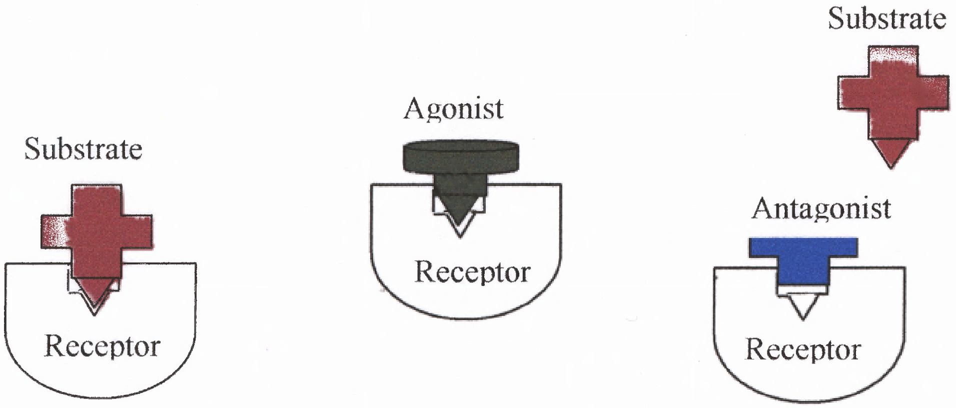


Figure 1.1 Agonist Vs Antagonist

The lock and key concept is used to give an analogy that a drug fits into a protein receptor cavity to produce an ideal biological response. Drugs that interact with receptor proteins are termed agonists or antagonists. Agonists bind to the receptor and produce the same biological activity as a natural substrate, say a hormone or a neurotransmitter. Antagonists also bind to the receptor but do not produce any response. By binding to the receptor, an antagonist prevents the natural substrate from binding and hence inhibits the effect of a natural substrate. Figure 1.1 illustrates the concept schematically.

The most popular approaches used in drug research are pharmacophore modeling and structure-based drug design.

1.2.1 Structure-Based Drug Design

Structure-based design is performed if the three-dimensional structure of the target macromolecule is available – for example through X-ray crystallography or Nuclear Magnetic Resonance (NMR) or homology modeling. The structural information is used for creating new compounds (de novo design) or to search molecular databases for molecules that fit the protein binding site. Potential drug molecules are evaluated in terms of their steric and electrostatic complementarity for the binding site using molecular docking.

1.2.2 Pharmacophore Modeling

The "pharmacophore" refers to the orientation in space of a set of chemical features that is common to a series of active compounds in a study. A three-dimensional pharmacophore gives the spatial relationship between these important groups in terms of distances and angles in the bioactive conformation. Since pharmacophore modeling focuses on a series of drug analogs, it can be used even when the structure of the receptor site is not known. Figure 1.2 shows a schematic of a pharmacophore. The figure shows the distances between chemical features that, in this case, are assumed to be important for drug-receptor binding. Pharmacophore models are created based on compounds of known biological activity. The features are chosen from experimental structure-activity data, which illustrate how the biological activity (i.e. receptor binding affinity) changes as chemical groups are added or subtracted from the drug. The models are further improved as new structure-activity data is obtained. It is relatively easy to derive a pharmacophore for a rigid molecule. For flexible molecules, it becomes challenging as they can have more than one conformation. The present work uses the Planes program to calculate not only the distance between but also the relative orientation of important pharmacophore features of DM324, a very flexible analog of GBR 12909.

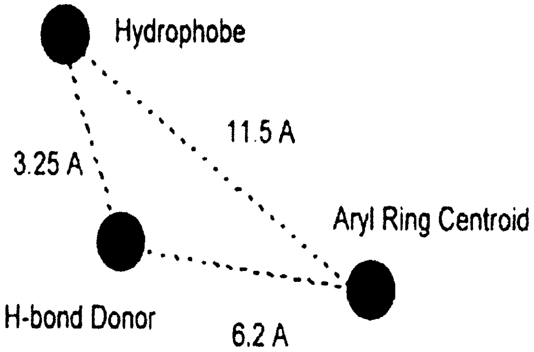


Figure 1.2 Schematic of a pharmacophore

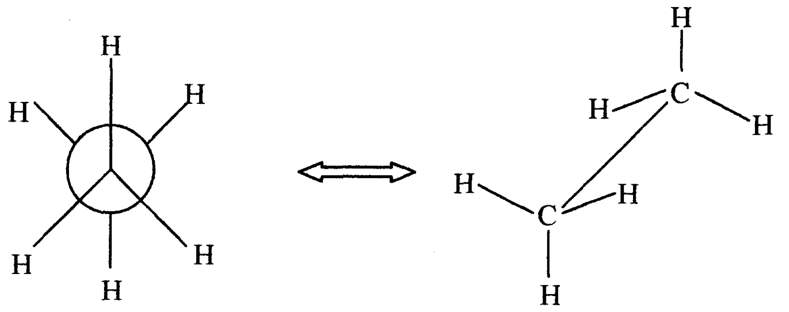
1.3 Conformational Analysis

Each molecular conformation is a particular arrangement of the atoms in space. Rotation around a bond alters that arrangement and leads to a different conformation. The terms conformation and conformer are used interchangeably in this document. In the most stable conformation, the potential energy of the molecule is at a minimum.

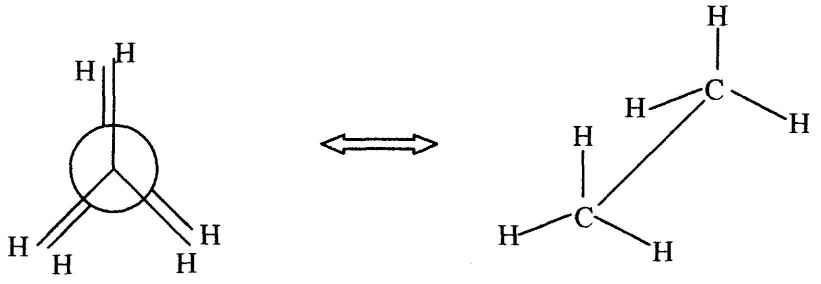
For example in ethane, $H_3C\text{-}CH_3$, the potential energy of the molecule changes as the hydrogen atoms (H) are rotated around the carbon-carbon bond. The energy required to rotate ethane about the carbon-carbon bond is called the torsional energy and the angle of rotation, defined by H-C-C-H, is the torsional angle. Figure 1.3 demonstrates the energy as a function of torsional angle for ethane. The figure shows the potential energy surface (PES) of the molecule. The minimum energy occurs at the staggered conformation (H-C-C-H = 60°, 180°, and 300°) where the H's are furthest apart; the maximum energy occurs in the eclipsed conformation (H-C-C-H = 0°, 120°, and 240°) where the H's are closest together and "lined up".

1.3.1 Conformational Search Procedures

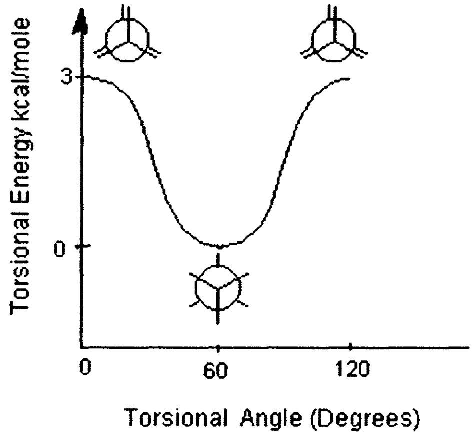
Conformational search procedures are used to generate conformations of a given molecule by exploring the potential energy surface of the molecule to locate minima. This is performed by alteration of the torsional angles of the molecule followed by energy minimization.



A. Lowest energy staggered conformation of Ethane



B. Higher energy eclipsed conformation of Ethane



C. Torsional energy as a function of torsional angle for Ethane

Figure 1.3 Conformational energy profile for ethane

Some popular methods are systematic search, distance geometry, random search, stochastic search and Monte-Carlo search (G.Chang et al., 1989). Of the many conformations generated, it is assumed that only one (the bioactive conformation) can actually bind to the receptor. But the bioactive conformation is not necessarily the one with the lowest energy (Bernard et al., 2003). Therefore a selection of conformers spanning a chosen energy range is chosen as input to separate pharmacophore modeling (CoMFA) studies. Hence data analysis techniques are employed to identify patterns in subgroups of conformations that span the selected energy range. Therefore the technique proposed here, a novel approach to calculating the geometrical parameters that give the relation between ring structures of the drug molecule, may be useful in providing some clues to identify the bioactive conformation for the flexible GBR 12909 analog, DM324, and other flexible molecules.

1.3.2 Additional Concepts

Superposition or alignment of molecules is performed to understand the similarities and differences among a group of compounds being studied as a potential drug. If there is a rigid biologically active compound then this serves as a template for superposition. Apart from defining the common binding groups manually, a clique detection method based on interatomic distances is employed to match 3D substructures (Brint et al., 1987). A more automated method of fitting involves using the program DISCO (Distance Comparisons)(Y.C.Martin et al., 1993) that has the ability to propose superposition rules for structurally-diverse compounds.

Knowledge of the binding cavity is important to study the receptor-ligand interactions. When the structure of the receptor is not available, a receptor-map can be made using the receptor-ligand complementarities. Here possible bioactive conformations are obtained for active and inactive analogs and each set is superimposed. The active analog volume is subtracted from the inactive volume and the result is a receptor map with a cavity where the active analog fits.

1.4 Cocaine and GBR 12909

Cocaine addiction and abuse continues to affect people's health in United States and around the world. Cocaine use is associated with life-threatening cardiovascular problems, spreading the Human Immunodeficiency Virus (HIV) due to the sex-for-drug barters and needle-sharing that is common among drug users, along with potential mental and health disorders in babies born to women who abuse cocaine during pregnancy.

In its pure form, cocaine is a white crystalline powder extracted from the leaves of the South American coca plant, *Erythroxylon coca*. Cocaine users most often inhale the powder sharply through the nose, where it is quickly absorbed into the bloodstream. Users usually feel an initial "rush" or sense of well-being, having more energy, and being more alert. This effect quickly wears off, leading to a craving for more cocaine and hence the addiction. The long-term effects of cocaine usage include irritability, mood shifts, restlessness, paranoia, stroke and even death.

Cocaine is a strong central nervous system stimulant. It inhibits the reuptake of dopamine, a chemical messenger associated with providing pleasure and movement. In the brain, dopamine is released by a pre-synaptic neuron into the synapse (the small gap between two neurons), where it binds to proteins called dopamine receptors on the post-synaptic neuron, thereby sending a signal to that neuron. Normally excess dopamine in the synaptic cleft is then recycled back into the pre-synaptic neuron by the dopamine transporter. However, cocaine binds to the dopamine transporter and blocks the normal reuptake of dopamine, resulting in a build-up of dopamine in the synapse subsequently giving a prolonged feeling of high. Figure 1.4 demonstrates the process.

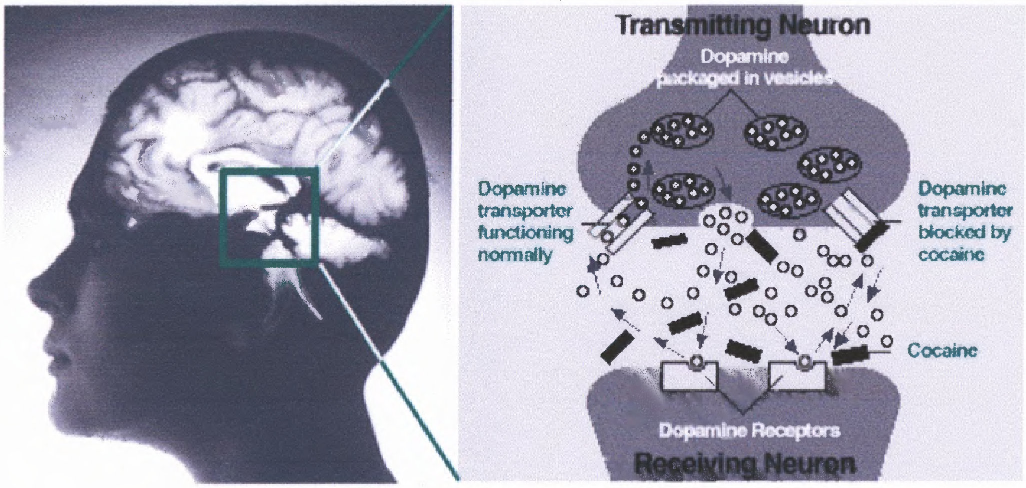


Figure 1.4 How cocaine works
http://www.drugabuse.gov/ResearchReports/Cocaine/Cocaine.html
(As of April 8, 2004)

As research evidence implicates the dopamine transporter in cocaine abuse and addiction, this transporter is a target of drug development against cocaine dependence (Singh, 2000; Smith et al., 1999). The structure-based design approach cannot be applied because the three-dimensional structure of the dopamine transporter is not known. Hence pharmacophore modeling techniques are applied to a series of dopamine reuptake inhibitors. One of the strategies of drug design is the development of a cocaine agonist that can be used as a non-addictive substitution treatment agent. Substitution therapies with full and partial agonists are being successfully used in the treatment of opioid, heroine and nicotine addiction. A potential drug for cocaine substitution therapy should possess these following qualities (Gorelick, 1998):

1. It should enter the brain slowly
2. The drug should produce minimal euphoria thus reducing the cravings for cocaine.
3. It should exhibit its effects for a suitably longer period of time.
4. The drug should have minimum or no side effects.
5. The drug should have reduced or no abuse liability as compared to cocaine.

A drug useful in the treatment of cocaine abuse would be a dopamine reuptake inhibitor with high binding affinity for the dopamine transporter but would simultaneously allow some amount of dopamine reuptake. Different classes of dopamine reuptake inhibitors that are currently under research are tropane, 1-[2-[bis(4-flourophenyl)methoxy]ethyl]-4-(3-phenylprophyl)piperazine

(GBR12909) analogs, methylphenidate, benzotropine, mazindol, and phencyclidine analogs (Dutta et al., 2003). Figure 1.5 shows the representative structures of various classes of dopamine reuptake inhibitors. Although these molecules have very different molecular structures they have common pharmacophore elements: N (nitrogen) and aromatic rings.

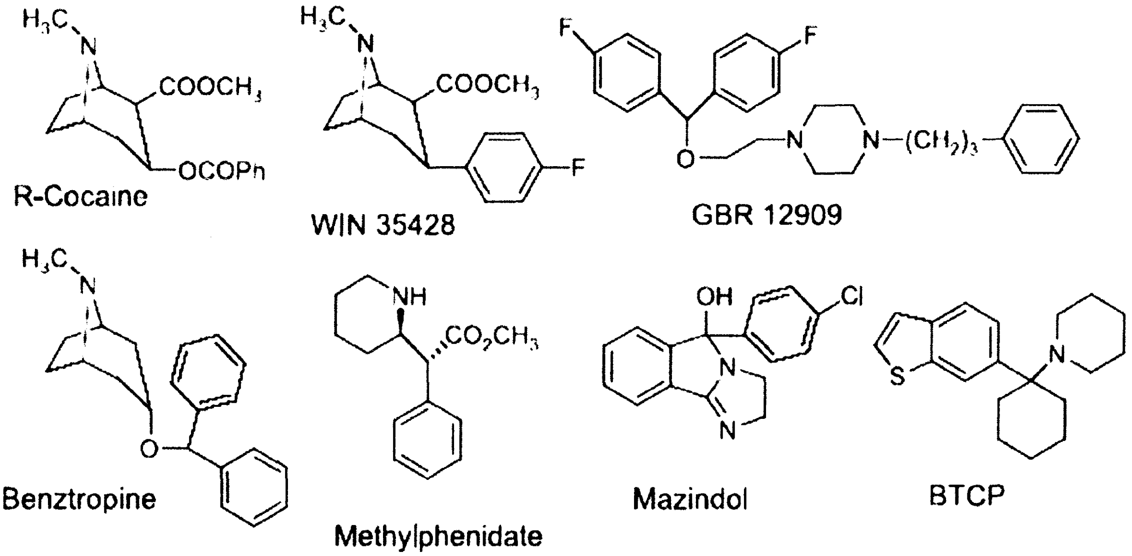


Figure 1.5 Representative structures of different classes of dopamine reuptake inhibitors

This thesis forms a part of ongoing research in the Venanzi group, Department of Chemistry and Environmental Science, NJIT to determine the pharmacophore model for the binding of analogs of GBR 12909 to the dopamine transporter. Much of the experimental structure-activity and behavioral studies of GBR 12909 analogs has been recently reviewed (Prisinzano et al., 2004). A short summary follows: GBR 12909 has much higher binding affinity and slower disassociation rate from dopamine transporter than cocaine. Behavioral studies in humans have shown that GBR 12909 and many of its analogs exhibit decreased cocaine intake without affecting food intake. They also appear to have reduced abuse potential. Initial clinical studies in humans have shown that GBR12909, when orally administered, is not harmful, is well tolerated and achieves moderate binding to the dopamine transporter. GBR12909 is currently in Phase II clinical trails.

This thesis presents an analysis of conformations of DM324, a very flexible analog of GBR 12909. The Planes program was applied to conformations of DM324 to understand the relative spatial orientation of the benzene, naphthalene and piperazine ring, i.e. the rings which contain the pharmacophore features hypothesized to be important for the binding of GBR analogs to the dopamine transporter. The structure of DM324 is shown in figure 1.6

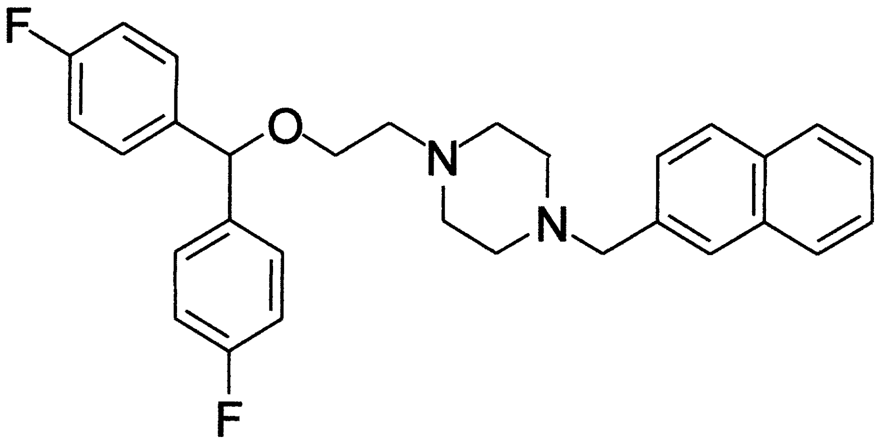
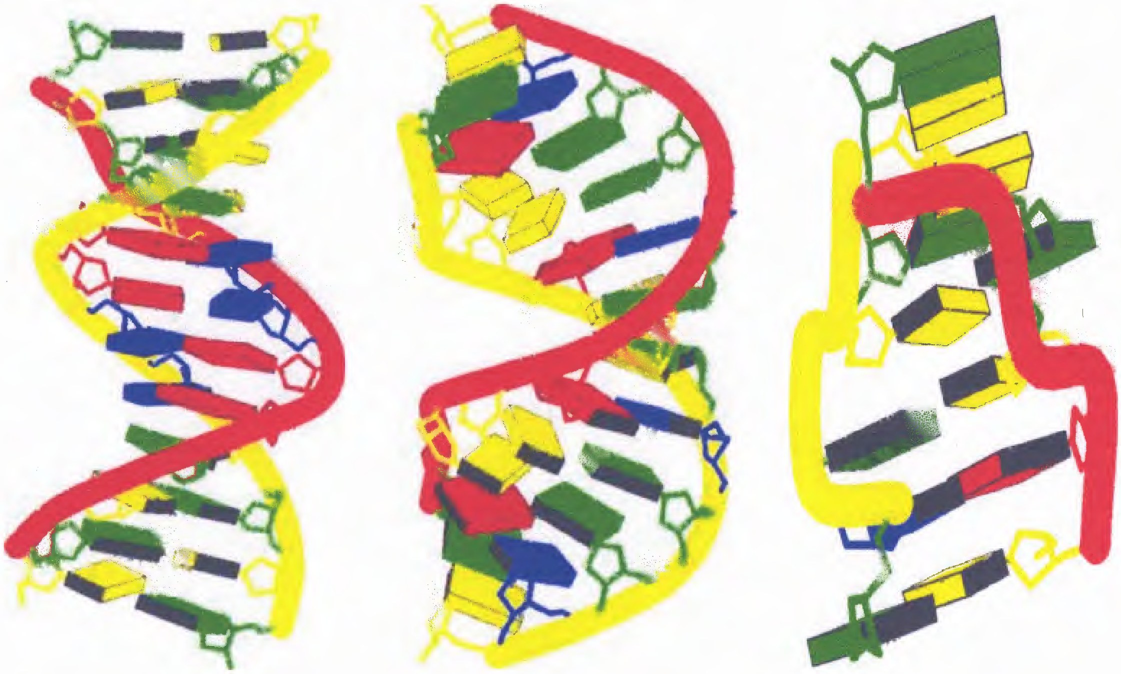


Figure 1.6 DM324, an analog of GBR

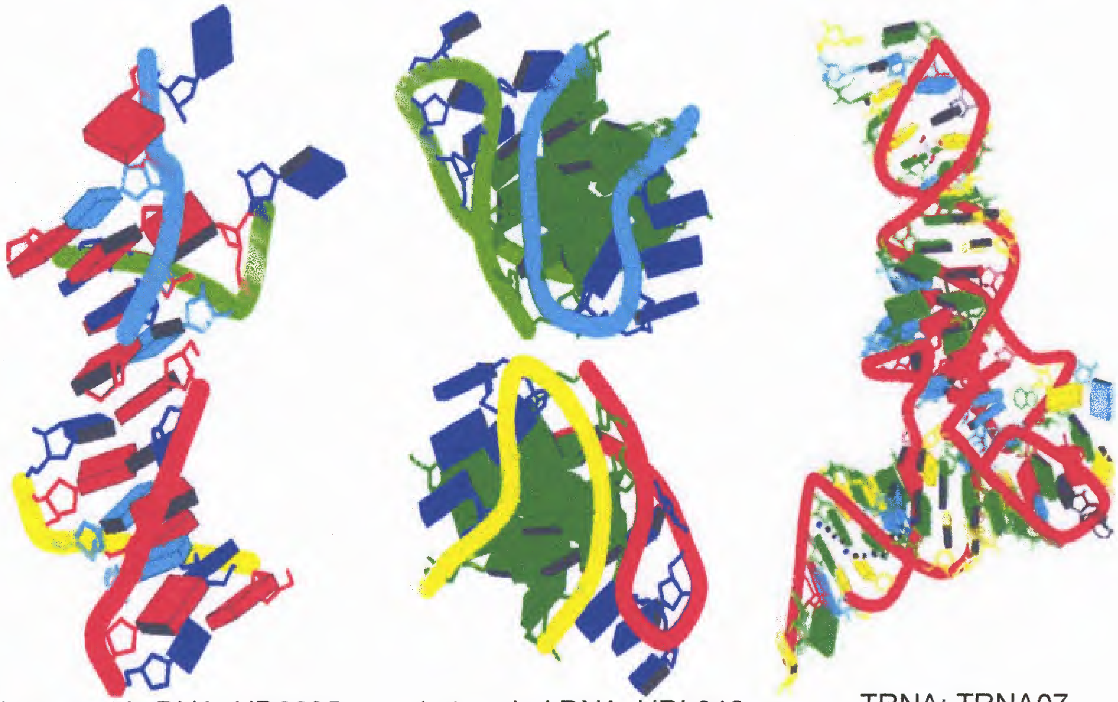
CHAPTER 2

THEORY: THE 3DNA APPROACH

The DNA molecule does not always exist as a perfect helical structure. It is quite flexible on its own. It can be bent, kinked, knotted and unknotted, unwound and rewound by the proteins that interact with it (Saenger, 1984). The results of fiber and crystallographic studies have shown that DNA can have several conformations. The most common form, B-DNA is a right-handed double helix with a wide (major) and a narrow (minor) groove. The bases are perpendicular to the helix axis. Another form of DNA, known as A-DNA, has a very deep major groove and a shallow minor groove. A very unusual form of DNA is the left-handed Z-DNA. There also exist a few asymmetric forms of DNA that are very unusual in that the end pairs are flipped out or there are bulges. Figure 2.1 shows some examples of nucleic acid structures taken from the Rutgers Nucleic Acid Database (NDB) (http://ndbserver.rutgers.edu/) along with their NDB identification codes. In the figure, the molecular structure of the purine and pyrimidine bases is modeled by a block to indicate their planar structures. In these nucleotide block models, adenine is red, thymine is blue, cytosine is yellow, guanine is green, and uracil is cyan. Understanding how the structure of the nucleic acids affects its function is becoming important as more genomic data is being discovered. Understanding the structures is also important to interpreting and predicting drug-DNA and protein-DNA interactions.



BDNA: BDL001 ADNA: AD0005 ZDNA: ZDF059



Asymmetric DNA: UD0035 4 stranded DNA: UDL018 TRNA: TRNA07

Figure 2.1 Nucleic acid structures

Several popular approaches have been employed to analyze the nucleic acid structures. Comparative studies on some of these programs: CEHS (Hassan et al., 1995; Lu et al., 1997), CompDNA (Gorin et al., 1995; Kosikov et al., 1999) Curves (Lavery et al., 1989; Lavery et al., 1988), FREEHELIX (Dickerson, 1998), NGEOM (Soumpasis et al., 1988; Tung et al., 1994), NUPARM (Bansal et al., 1995), RNA (Babcock et al., 1994a; Babcock et al., 1994b; Pednault et al., 1993) have shown that the choice of reference frame rather than the mathematical calculation has led to the discrepancies in the parameters evaluated using different programs (Lu et al., 1998; Lu et al., 1999). A common point of reference is recommended to describe the three dimensional arrangements of bases and base-pairs in nucleic acid structures (Olson et al., 2001).

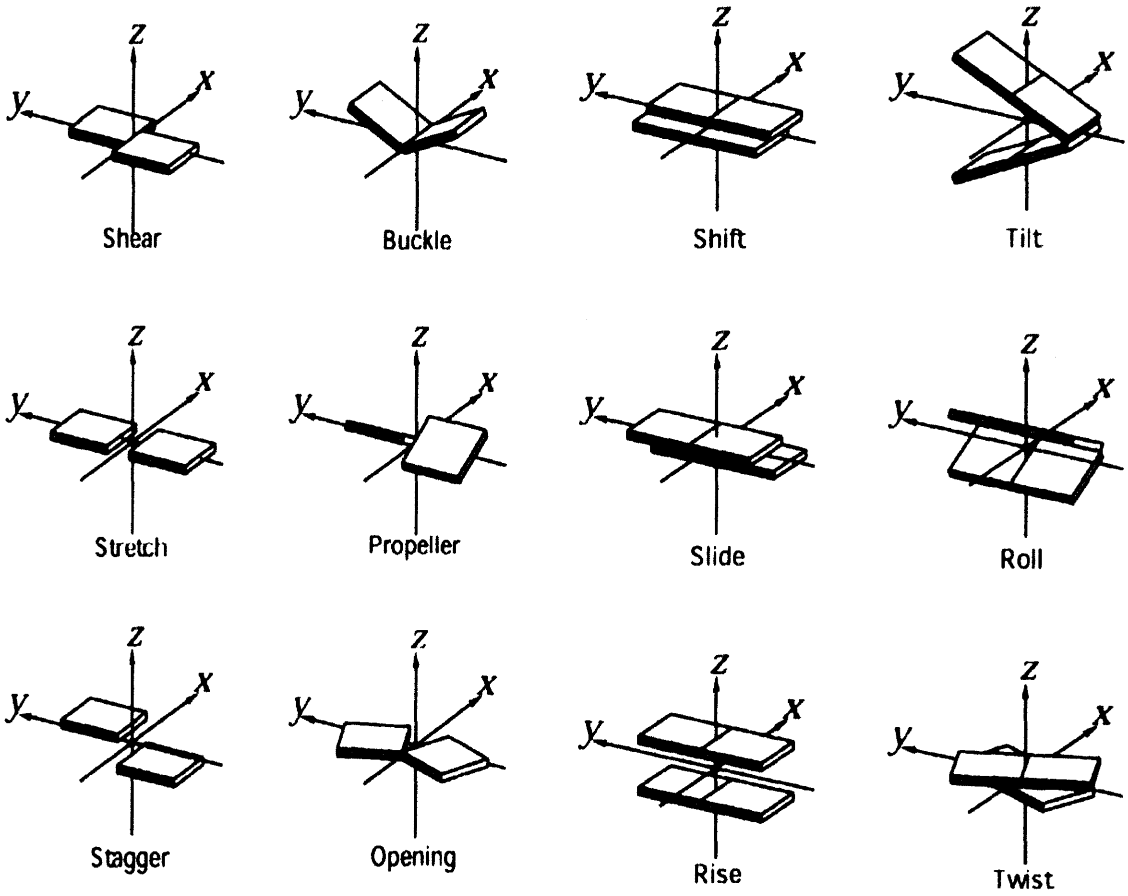
2.1 OVERVIEW OF 3DNA PROGRAM

3DNA (Lu et al., 2003) is a versatile software package for the analysis, reconstruction and visualization of the three-dimensional nucleic acid structures. The program can be used on parallel and anti-parallel double helices, single-stranded nucleic acid structures, multi-stranded helices and complex tertiary folding substructures found in both DNA and RNA.

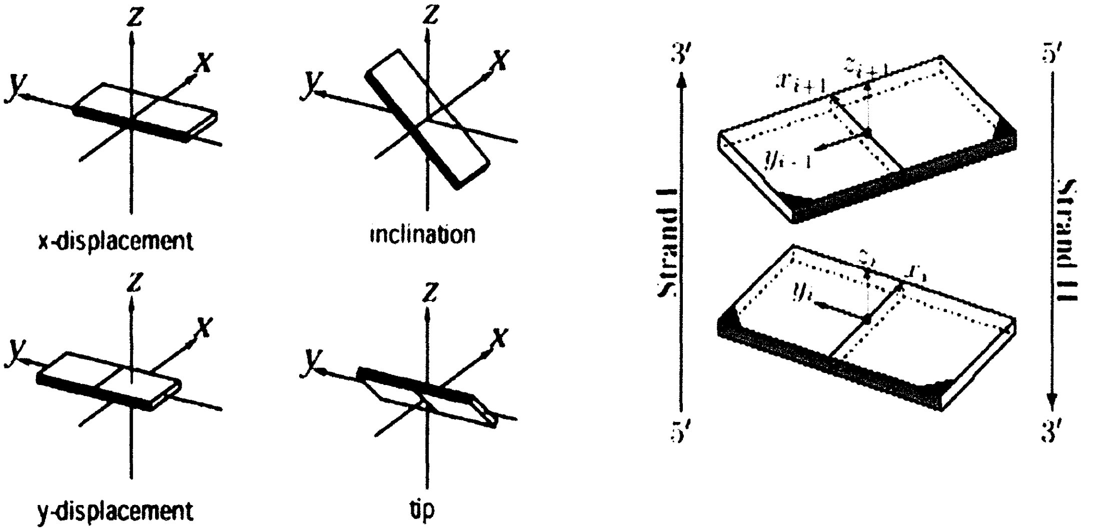
The 3DNA program uses the coordinate reference frame for the description of the nucleic acid base pair geometry and a rigorous matrix-based algorithm to evaluate the local conformational parameters. Calculations of the parameters in 3DNA follow the Cambridge University Engineering Department Helix Computation Scheme (CEHS) definitions (Hassan et al., 1995) as implemented in the Structure and Conformation of Helical Nucleic Acids Analysis Program (SCHNAaP) (Lu et al., 1997). Babcock et al (Babcock et al., 1994b) explain the basic concepts, theorems and proofs of the mathematics behind the analysis of the three dimensional nucleic acid structures.

For DNA, a base along with its complimentary base (A with T, C with G) form a base pair. One base pair along with its adjacent base pair (either above it or below it along the helical axis) form a base step. A coordinate reference frame defined by the planar nucleotides adenine, guanine, cytosine, thymine and uracil is utilized to calculate the base-pair parameters and step parameters. Base pair parameters describe the relative orientation and position of one base with respect to its complimentary base in a base pair. Step parameters describe the relative position and orientation of consecutive bases along a single strand of DNA. These rotational and translational parameters are rotations and displacements about the x-, y- and z-axis of the reference frame. Figure 2.2 gives a pictorial definition of the parameters calculated. A significant and novel aspect of the 3DNA calculation method is the use of the concept of mid-step triad to calculate a set of "absolute" helical parameters.

In reality not all bases are exactly planar. To take this into account in the calculation of the base pair and base step parameters, an ideal structure is defined for each of the purine and pyrimidine rings. This standard structure is created by placing the atoms in a desired orientation with respect to the origin at (0,0,0).



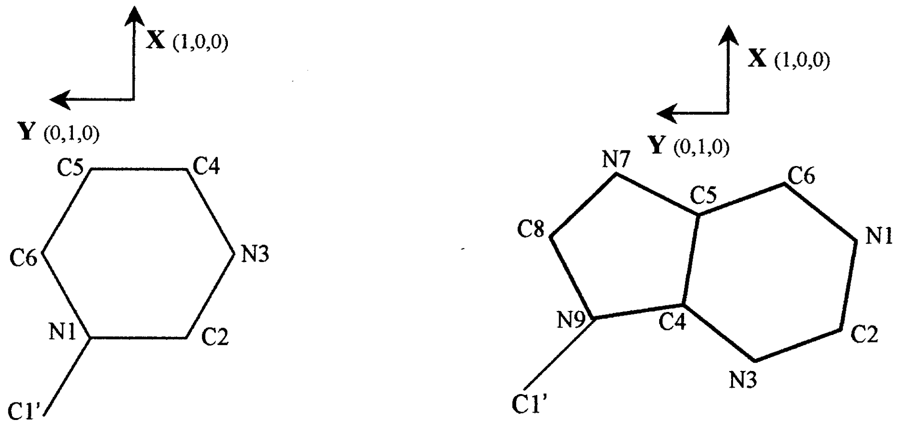
Base pair Parameters **Step Parameters**



Helical Parameters **Reference frame**

Figure 2.2 Pictorial definitions of parameters calculated by 3DNA program

Figure 2.3 illustrates a standard purine and pyrimidine. The x-axis is defined by the vector (1,0,0), the y-axis by the vector (0,1,0) and the z-axis by the vector (0,0,1). Each standard is defined by its carbon and nitrogen framework.



A. Schematic of a standard pyrimidine

B. Schematic of a standard purine

Figure 2.3 Standard bases expressed in a standard reference frame

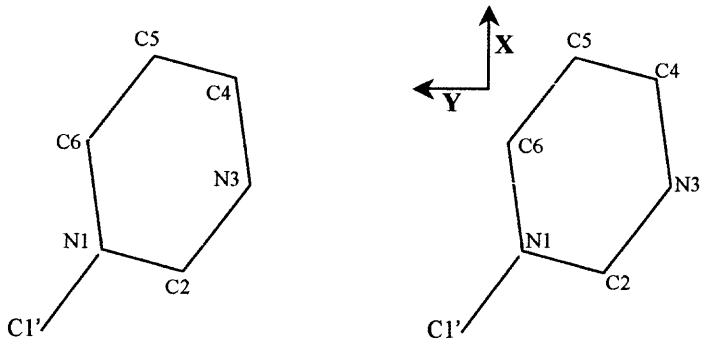


Figure 2.4 Least Square fitting of standard base onto an experimental base

The first step in the analysis is the superimposing of the coordinate reference frame of the ideal base onto the "experimental" (or real) base. Here "experimental" refers to the DNA structure input data set which may come from

X-ray crystallography or from computations such as molecular dynamics simulations. The superposition is performed by least-square fitting (see Figure 2.4) of the correct standard base onto the experimental base, i.e. by minimizing the sum of the squares of the distances of each atom in the experimental base to the corresponding atom on the standard. Least-squares fitting in 3DNA makes use of only the ring atoms: i.e. nine ring atoms for purine that are planar and six ring atoms for pyrimidines. A closed-form solution developed by Horn(Horn, 1987) is utilized in the program for least-squares fitting. This fitting is a crucial step and maintains consistency among the structural parameters.

Once the ideal coordinate frame is superimposed on the experimental, only the unit vectors X, Y and Z that define the coordinate frame are needed for the calculations. Atomic coordinates of both the standard and experimental base are not required. This process is repeated until all experimental bases have been imposed with a coordinate reference frame using its corresponding standard structure. Once the coordinate frame is determined for each base, only these vectors are used for the calculation of parameters.

For step parameters, only Twist is treated as primary and the other two angular rotations Roll and Tilt, are considered secondary. This is because Tilt is almost always negligible and magnitude of Roll is small, typically <20°. A "RollTilt" (Γ) about a Roll-Tilt axis, in the x-y plane of the mid-step triad is defined. The Roll-Tilt axis, also termed the hinge axis is defined such that it is inclined at an angle (Φ) to the y-axis and lies on the x-y plane of the middle reference frame.

Since Φ is usually small, Roll (ρ) and Tilt (τ) are defined from Roll-Tilt (Γ) and Φ approximately as

$$\left.\begin{aligned} \rho &= \Gamma \cos(\Phi) \\ \tau &= \Gamma \sin(\Phi) \end{aligned}\right\} \quad (1)$$

This simply means Roll-Tilt(Γ) is the vector sum of Roll(ρ) and Tilt (τ).

The procedure for the calculation of the step parameters and base-pair parameters is exactly the same. Figure 2.5 demonstrates the overall method of calculation for a schematic base pair and base step in a double helical structure. The direction of the positive z-axis(not shown) is determined by the cross product of the x- and y-axes. Consider Base-1A and Base-2A to represent the two adjacent bases on one strand and Base-1A, Base-1B and Base-2A, Base-2B to represent the pairs of complimentary bases. For double helical DNA, the program first calculates the base pair parameters between, say, Base-1A and Base-1B. 1M is the middle reference frame for the calculation of base pair parameters between Base-1A and Base-1B. Similarly 2M is the middle reference frame used for the calculation of base-pair parameters between Base-2A and Base-2B. These mid-frames 1M and 2M then serve as the base-pair reference frames to calculate the step parameters. M is the middle reference frame of M_1 and M_2 and is used for the calculation of step parameters. Calculation of the step parameters follows the procedure described below.

From the position and orientation of base-pair reference frames M_i and base-pair M_{i+1},

1. Take the cross product of the z-axes of the two base pairs M_i and M_{i+1}. This gives the direction of the hinge, or Roll-Tilt axis, in the middle reference frame M. In other words the hinge axis is the intersection line between the two planes.
2. Rotate the two base-pairs about the hinge with angles of equal magnitude but opposite signs until their x-y planes are parallel to each other. This gives the x-y plane of the middle frame M.
3. The z-axis now obtained is the z-axis of the middle reference frame M.
4. The Roll-Tilt (Γ) angle is calculated as the amount of the relative rotation of the two base-pairs.
5. The hinge axis does not bisect the angle between two y-axes. The hinge's offset from the bisector of the two Y-axes is the angle Φ. Angular rotations about the y- and z-axes, i.e. Tilt (τ) and Roll (ρ) respectively, are calculated using equation 1. A view along the z-axis of the middle reference frame of the current y- (or x-) axis gives the angle Twist (ω).
6. Translational parameters Shift (D_x), Slide (D_y) and Rise (D_z) are the projections of the vector joining the origins of the two base-pairs in step-1 to the x-, y- and z-axes of the middle frame.

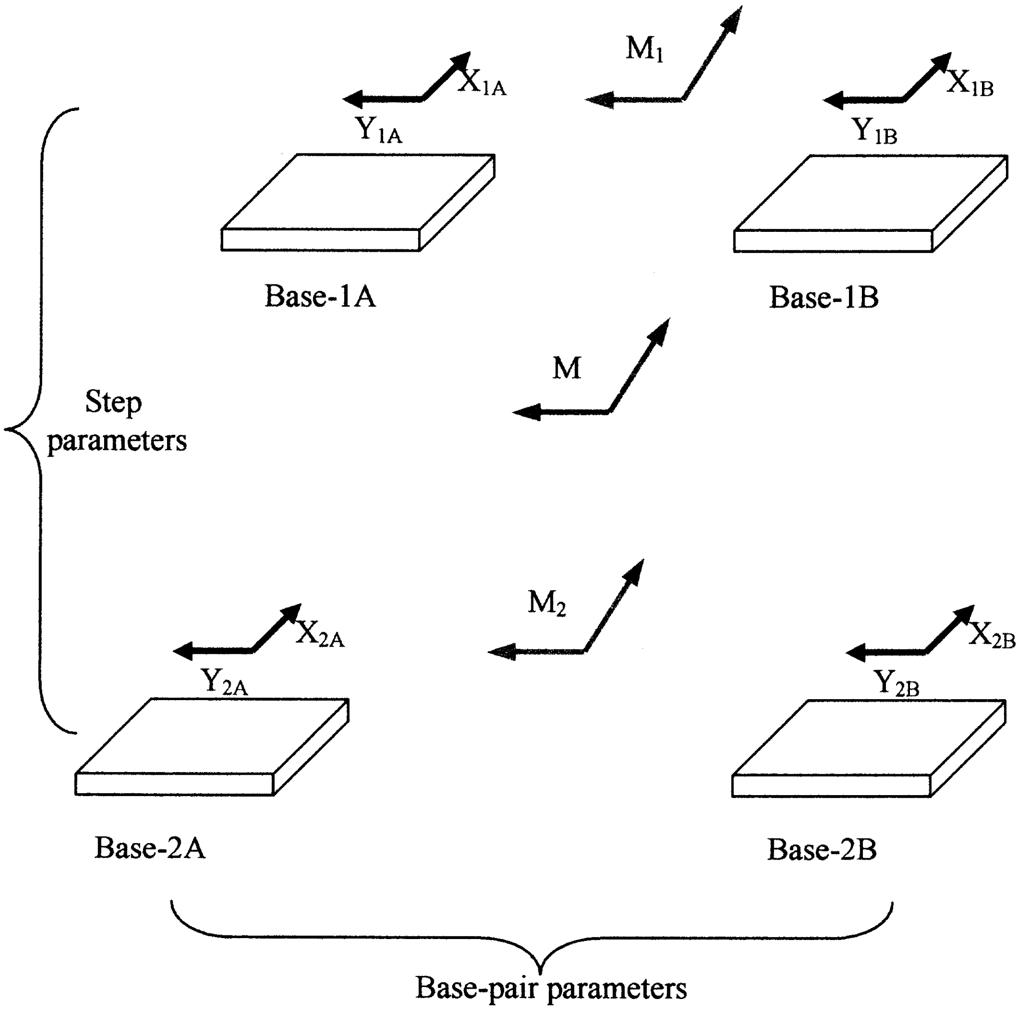


Figure 2.5 Calculation of local base-pair and step parameters in double helical DNA

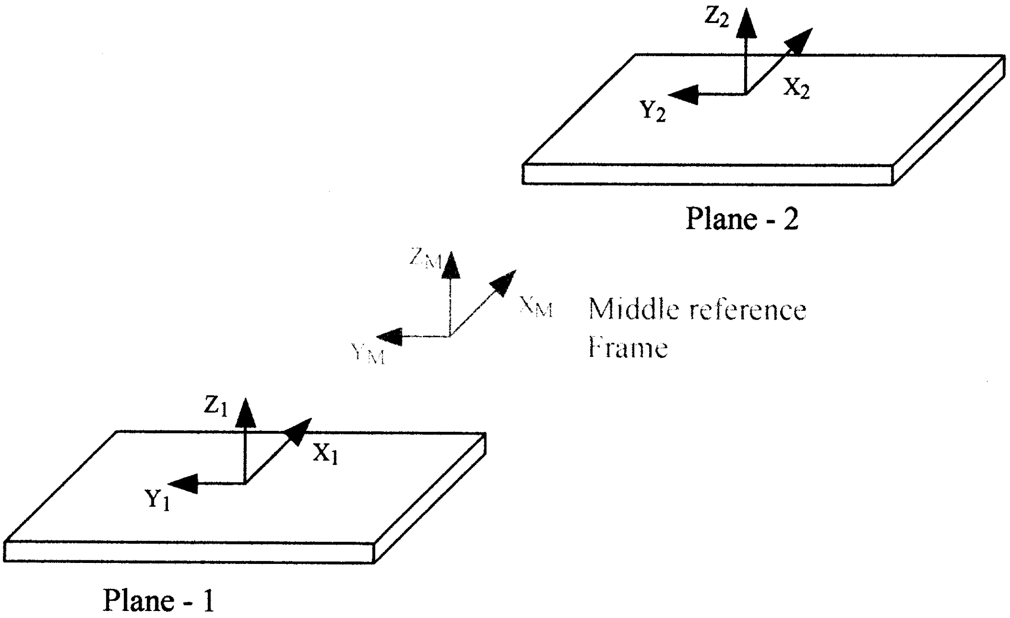
2.2 The Planes Program

The main component of this research has been the design of the Planes program and its application to the novel classification of conformers of the flexible GBR 12909 analog, DM324. Although many molecular modeling programs such as SYBYL are capable of calculating certain angles and distances between molecular fragments, none are able to provide the full range of detailed information on the relative orientation of molecular features available in the Planes program. This innovative approach is derived from the concepts of the 3DNA program.

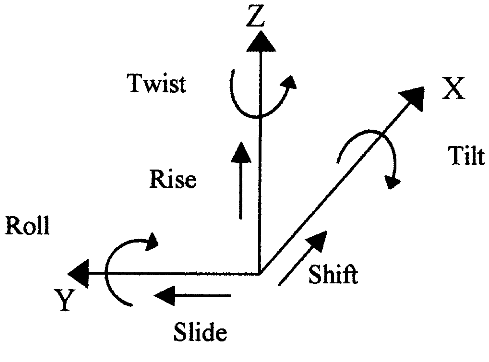
In an arbitrary molecule, viewed as a single strand of DNA, the position and orientation of one fragment with respect to the other can be completely defined using parameters defined from six degrees of freedom: three angles and three distances. The Planes program uses the terminology of the base step (Shift, Slide, Rise, Tilt, Roll, and Twist: see Figure 2.6) rather than the base pair to characterize these parameters. This is an arbitrary choice since, in the 3DNA program, the step parameters and base-pair parameters are calculated exactly the same way.

But, unlike DNA, which contains planes defined by either a purine or a pyrimidine ring, the planes in an arbitrary molecule can be defined from any ring fragment. Due to the helical structure of DNA, the relative position and orientation of either a base pair or base step is somewhat restricted. The same cannot be assumed for an arbitrary molecule. For a flexible molecule, the relative displacement and orientation of the ring fragments can be either small or large, limited only by the molecule's chemical and structural characteristics. The Planes program calculates the parameters of one ring relative to every other ring in the molecule unlike in 3DNA, where step parameters are calculated for only the adjacent base pairs.

The 3DNA program uses the Brookhaven Protein Data Bank (PDB) (Bernstein et al., 1977) file format for input nucleic acid structure. This is a standard format used by molecular modeling programs for description of macromolecular structures. PDB files are plain text ASCII files that store residue names, atom names, and x, y, z coordinates in a Cartesian coordinate system of every atom, along with other detailed structural information of a protein or nucleic acid. Similarly the mol2 file (.mol2) format is commonly used by molecular modeling programs for the description of small molecule structures. A mol2 file is a complete, portable description of a molecule. Like a PDB file, the mol2 file stores the structure and atomic coordinate information of the molecule. This format has the advantage of storing all the necessary information such as atom features, positions, and connectivity. It can also be used to describe the atoms that constitute the planes of interest in the molecule. In other words a PDB or mol2 file can be viewed using a text editor. This text file gives the atomic coordinate information, bond information and other definitions describing the molecule (See Appendix B for sample mol2 files). These same files can be viewed as a structural molecule using either molecular display programs such as Rasmol or molecular modeling software such as SYBYL. The Planes program takes a mol2 file format as input and follows the 3DNA procedure for calculating the "base step" parameters for an arbitrary molecule.



A. Schematic of a plane defined for the ring fragments with reference frames and calculation of parameters using middle frame



B. Positive sense of the parameters calculated going from Plane-1 to Plane-2.

Figure 2.6 Planes Program Parameters

Dr. Xiang-Jun Lu and Dr. Wilma K Olson of the Chemistry Department, Rutgers University, Piscataway, NJ released the MATLAB source code of the 3DNA program to the Venanzi group, Department of Chemistry and Environmental Science, NJIT solely for this research. The initial version of the Planes program was prepared by Rohan Woodley as part of his research project in the Computational Biology M.S. program at NJIT.

CHAPTER 3

METHODS

3.1 Overview

Any generalized program has its own limitations. The burden of providing the input in a specific format is upon the user of the program. Molecule input to the Planes program is in mol2 format and is of two types: input for the molecule under study (here, DM324) and input for the molecular ring fragment "standards" (here, benzene and piperazine rings). The next section of this chapter provides the explanation of the input molecule, the ring and plane definitions, followed by the specifications of the standard for the corresponding ring fragments.

The Planes program reads a mol2 file, mines for information on the plane definitions and atoms that lie on the plane, extracts the corresponding atomic coordinates and calculates the translational and rotational parameters for every pair of planes. A configuration file that specifies the standard, input (.mol2) and output filenames, along with some necessary definitions to evaluate the parameters, serves as an input file to the Planes program.

The Planes program works as follows

1. The Planes program reads the input mol2 file that specifies the molecule of interest.
    a. First it identifies a ring fragment.
    b. It then mines the mol2 file for the plane defined for this ring and extracts the atoms that lie on the plane.

   c. The program then identifies the standard corresponding to the ring fragment.
   d. The program superimposes the standard reference coordinate frame onto the ring by performing a least-squares fit of the atoms that lie on the plane. This gives the origin and the reference frame for the plane.
   e. Steps (a) to (d) are performed for all ring fragments of the molecule.
2. The program then calculates the translational parameters Shift (D_x), Slide (D_y) and Rise(D_z), and rotational parameters Tilt (τ), Roll (ρ) and Twist (ω) for every possible combination of pairs of planes.
3. Steps 1 and 2 are repeated for all the input conformations.

3.2 Calculation of Planes Parameters for DM324

3.2.1 Input Conformations of DM324

The Planes program was used to calculate the planes parameters (i.e. the "base step"-type parameters for a non-DNA molecule) for conformations of DM324, a GBR 12909 analog (see Figure 3.1.)

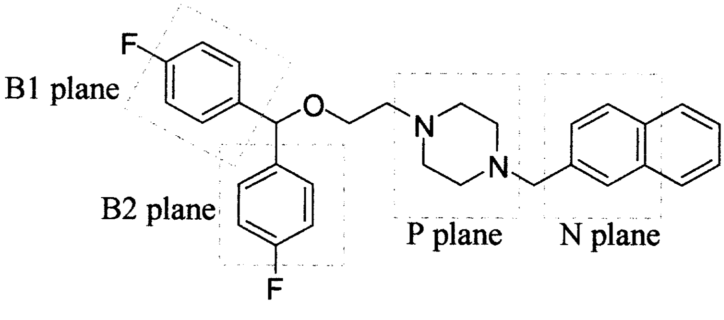


Figure 3.1 DM324 Conformation with planes considered for calculation of parameters.

Each DM324 conformation was stored in mol2 file format (see Appendix C for a sample input conformation file for the rs2_00001 conformer of DM324). The mol2 file was prepared from the SYBYL database containing the Random Search output by using a SYBYL macro (Appendix B) to define the planes of interest in the molecule. Each mol2 file had four molecular planes defined, one for each ring fragment: two benzene planes (shown as B1 and B2 in Figure 3.1) defined from the six aromatic carbons of the ring, one napthalene plane defined from the 10 aromatic carbons of the napthalene ring, and one piperazine plane (P plane in Figure 3.1) defined from two carbons and two nitrogens of the piperazine ring. More details are given in the sections below. There are over 350 analogues of GBR 12909. Many of them have a benzene-like substituent in place of the napthalene ring. In order to make the application of the program as general as possible and to compare the plane parameters with conformations of other analogs of GBR 12909, only the "lower", benzene-like portion of the napthalene was used in the calculation of the planes parameters. This means that it was not necessary to create a napthalene standard. The benzene standard was used instead. Since napthalene is perfectly planar, relative angles calculated using a 10-atom plane or a 6-atom plane would be the same.

The input conformations of DM324 were generated by Milind Misra, using the Random Search conformational analysis function of SYBYL. A total of 728 unique conformations of DM324 with energies within 20kcal/mol of the conformation of lowest energy were identified. During the calculation the benzene-1, benzene-2, piperazine, and napthalene rings were defined as aggregates (i.e. held fixed) and only the eight torsional angles of DM324 were allowed to rotate. During the search, the piperazine ring was held fixed in a chair conformation with the substituents on each nitrogen fixed in the equatorial position.

In order to compare the structures of the 728 conformations, they were oriented relative to the central piperazine ring. The piperazine nitrogen to which the bisphenyl side chain is attached was placed at the origin (0,0,0) of the Cartesian coordinate frame. All 728 conformations were superimposed by fitting to the four piperazine ring atoms (see Section 3.2.2) used to define the piperazine ring plane. The result shows the range of orientations available to the side chains (see Figure 3.2).

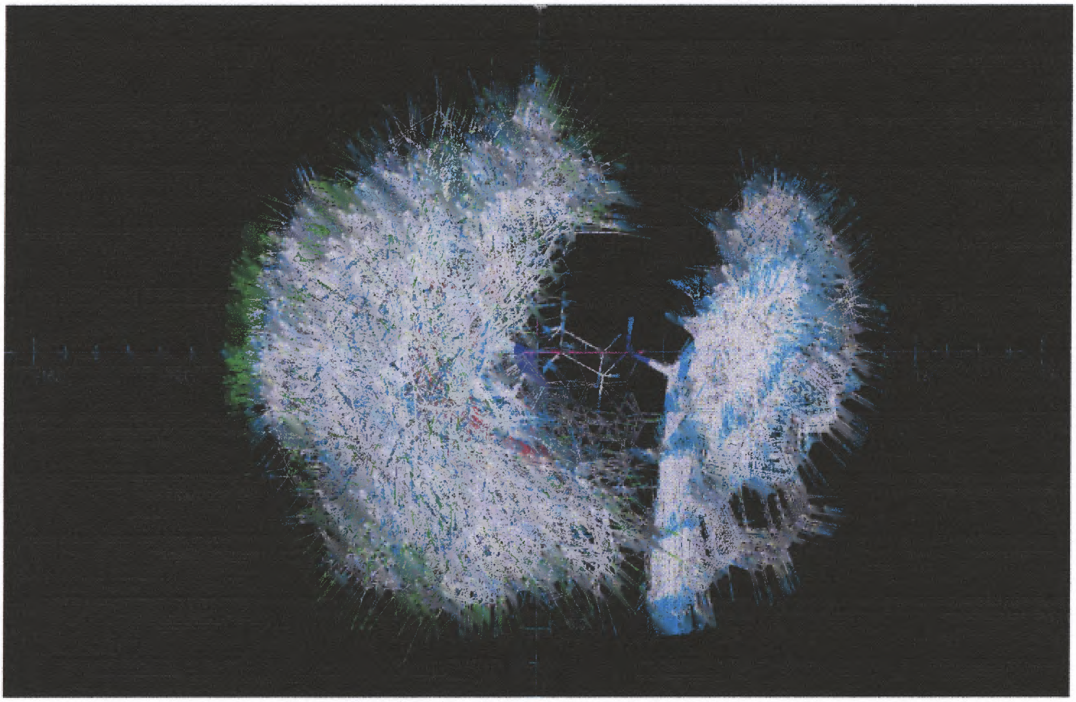
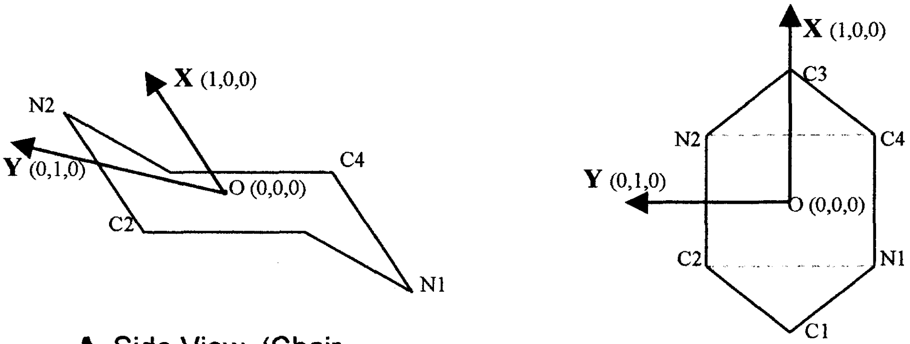


Figure 3.2 provided by Milind Misra. Side view of 728 DM324 conformations aligned on the central piperazine ring

3.2.2 Definition of Standards

3.2.2.1 Piperazine : A piperazine molecule from the SYBYL fragment library was used to define the standard (or "ideal") piperazine ring. Since the piperazine ring itself is nonplanar, a plane defined by four atoms of the ring had to be selected to represent the ring. There are three possible planes containing four atoms each in piperazine: one plane formed with four carbon atoms, and two planes created with the two nitrogens and two carbon atoms on one side or the other of the nitrogen. For this research a plane formed by the two nitrogens and two carbon atoms: N2, C2, N1, C4 was chosen (see Figure 3.3). The symbol "O" identifies the centroid that is calculated as the average of the Cartesian coordinates of these four atoms and lies on the same plane. The right-handed coordinate reference frame attached to an ideal piperazine is shown in Figure 3.3. The x-axis runs parallel to the N2-C2 bond and passes through the centroid O. The y-axis is perpendicular to the x-axis and lies on the plane. The z-axis is defined by the right hand rule i.e. $z = x \times y$. This standard molecule is positioned with its centroid O at the origin (0,0,0) of its reference coordinate frame. The x-axis is defined by the vector (1,0,0), the y-axis by vector (0,1,0) and the z-axis by vector (0,0,1) as shown. See Appendix C for the standard piperazine input file.

3.2.2.2 Benzene and Napthalene : The benzene molecule from the SYBYL fragment library was taken as a standard for both benzene and napthalene for the reasons described above. The right-handed coordinate reference frame for standard benzene is shown in Figure 3.4. Benzene, a rigid aromatic ring, has six carbon atoms on a single plane. O represents the centroid of the molecule, calculated from the Cartesian coordinates of the six carbons. The x-axis is chosen such that atoms C3 and C6 lie along the x-axis which passes through the centroid O. The y-axis is perpendicular to the x-axis and lies on the plane. The z-axis is defined by the right hand rule, i.e. $z = x \times y$. This molecule is positioned with centroid O at the origin (0,0,0) of its reference coordinate frame. The x-, y-, and z-axes are defined by the vectors (1,0,0), (0,1,0), and (0,0,1), respectively, as shown in Figure 3.4. See Appendix C for the standard benzene input file.



A. Side View (Chair conformation of piperazine)

B. Top View

Figure 3.3 Assignment of a coordinate reference frame for a non-planar piperazine standard (---- represents the plane)

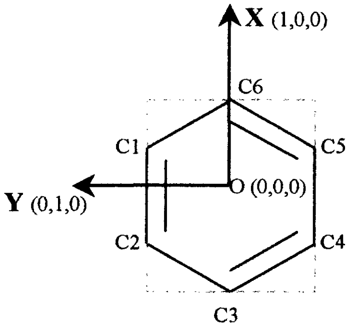


Figure 3.4 Assignment of a coordinate reference frame for planar standard benzene (---- represents the Plane)

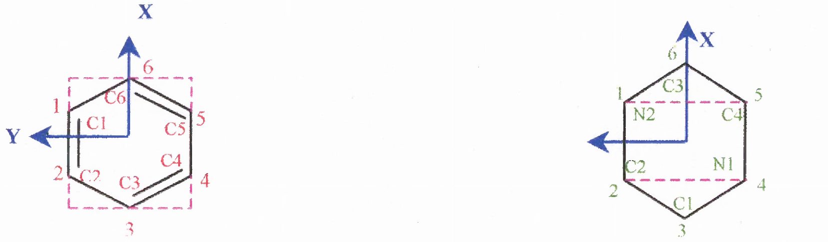
3.2.3 Configuration File

The configuration file for DM324 was created and given as input to the Planes program. See Appendix C for the sample configuration file. This file contains

1. A name with the path (location) of the two standard files in mol2 format: one for benzene and one for piperazine.
2. A name with the path of the 728 input files in mol2 format, one for each DM324 conformation.
3. A declaration statement (FIT) to specify that a benzene standard needs to be used for a naphthalene plane.
4. A declaration statement (OVERRIDE) to override the previous definition of napthalene plane with 10 atoms to a new definition of napthalene plane with six atoms.
5. A name with the location of the output file where the output of the program needs to be printed.

3.2.4 Mapping of Standard Reference Frames to DM324

One of the most important steps in the Planes program is the correct assignment of the standard reference frames to the corresponding ring planes of the molecule. Figure 3.5 illustrates the mapping of the standards to the DM324 planes. The dotted lines show the atoms used to define the planes of the standards and the planes to which they were fit in DM324. The atoms of DM324 are numbered in black, the atoms of the standards in red or green.



Benzene standard

Piperazine standard

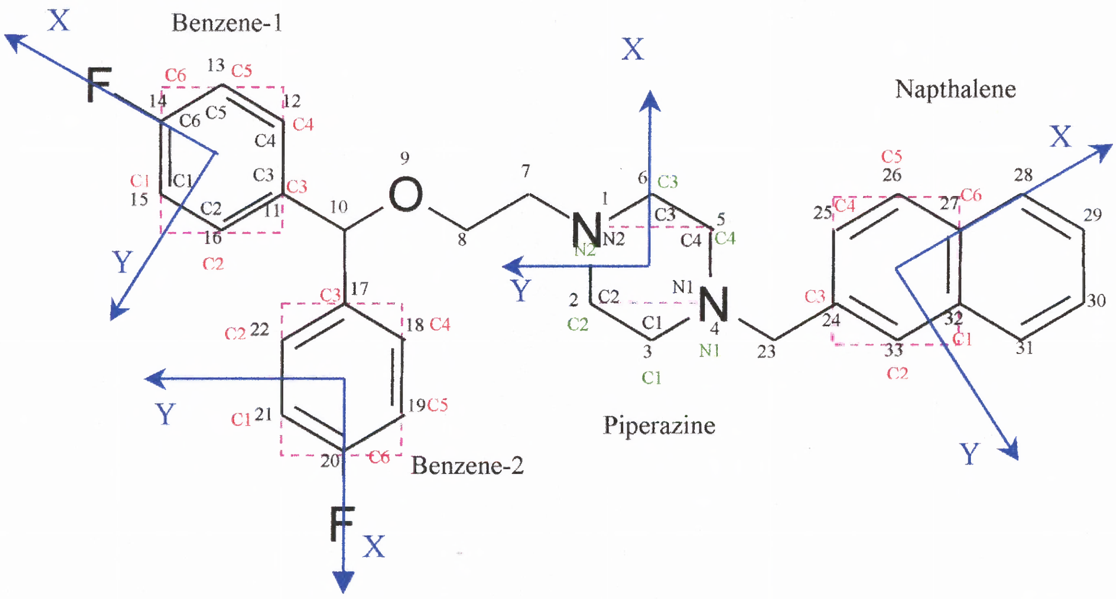


Figure 3.5 Mapping of the reference coordinate frame onto planar rings of DM324.
C1,C2,C3,C4,C5,C6 represents atom names of benzene standard
N1,C1,C2,N2,C3,C4 represents atom names of piperazine standard
------- represents the plane defined

3.2.4.1 Definition of Planes for DM324 : As stated above the benzene-1, benzene-2, piperazine, and napthalene planes were defined for DM324 by running a SYBYL macro on the SYBYL database containing the Random Search output. Appendix B gives the macro that defines the planes and calculates the centroid of each plane. As can be seen from Appendix B, the piperazine plane

(C-plane in the Appendix C conformation file) was defined from the DM324 atoms 1, 2, 4, and 5; the benzene-1 plane (P1_plane in the Appendix C, Conformation file) from atoms 11, 12, 13, 14, 15, and 16; the benzene-2 plane (P2_plane) from atoms 17, 18, 19, 20, 21, and 22; and the napthalene plane (N_plane) from atoms 24, 25, 26, 27, 28, 29, 30, 31, 32, 33. For this calculation only six atoms of naphthalene were used to define a naphthalene plane. This is specified in the OVERRIDE definition of configuration file (See Appendix C, Configuration file) by the statement "OVERRIDE PLANE N_PLANE 6 24 25 26 27 32 33". This statement tells the program to discard the previous definition (with 10 atoms) and gives a new definiton for N_PLANE as the six atoms with atom ID's 24, 25, 26, 27, 32, 33. These planes are shown on the DM324 structure in Figure 3.5.

Appendix C gives the sample input conformation file produced by the macro for one of the DM324 conformations, rs2_00001. Each line after the "@<TRIPOS>ATOM" line contains the following information for an atom in the molecule: atom ID, atom name, x-, y-, and z-coordinates, atom type, fragment number, fragment name, and atomic point charge. The file shows that DM324 contains 66 atoms, with the first 33 being "heavy" (non-hydrogen) atoms. The first atom (atom ID equal to 1), has atom name N2, the second (atom ID 2) has atom name C2, and so on.

Note that there are four fragment names (1 piperazine, 2 napthalene, 3 benzene, and 4 benzene) and that each atom in the molecule is assigned to a fragment. The fragments contain the rings plus hydrogens attached to the carbons on the rings (not shown in Figure 3.5) plus non-ring atoms that are located next to a ring. For example "1 piperazine" contains not only the six atoms of the piperazine ring (atoms 1-6) but also the neighboring atoms 7 and 23. The "4 benzene" fragment contains not only the six atoms used to define the plane (atoms 11-16) but also the neighboring atoms 8, 9 and 10. It should be emphasized that the fragment names are simply a means of grouping the atoms of the molecule and are not used to define rings or planes. As was seen above, subsets of the fragments are used to define the planes. The section of the input conformation file after "@<TRIPOS>SET" lists the atoms that define the planes (see C_PLANE, P1_PLANE, P2_PLANE, and N-PLANE). Note that in Figures 3.1 and 3.5, respectively, the P1_PLANE is referred to as B1 or benzene-1 and the P2_PLANE is referred to as B2 or benzene-2.

3.2.4.2 Location of Coordinate Frame Axes for Planes in DM324 : Before running the Planes program, the user must choose how to orient the molecule and how to define the axes so that the planes parameters are (hopefully) easy to interpret. For DM324 it was decided that the plane of the piperazine (defined by atoms 1, 2, 4 and 5 of DM324) would be located in the x-y plane at z equal to zero and that N2 would be located at the origin of the coordinate system. This choice can be seen from the sample input conformation file given for the rs2_00001 conformation of DM324 in Appendix C. The average z-value of the N2, C2, N1, and C4 atoms that define the ring plane is zero and the N2 atom is at (0, 0, 0). All the 728 conformers were superimposed in this orientation.

It was decided that the x-axes for the benzene fragments in DM324 should run along the C-F bond, passing through carbons 14 and 11 in benzene-1 and carbons 20 and 17 of benzene-2 to intersect at carbon 10. The y-axes were assigned in the following manner. Figure 3.5 shows that the atoms of benzene-1 in DM324 are numbered with atom ID's ranging from 11 to 16. The sample input conformation file in Appendix C shows that the atom names corresponding to these atom ID's are (in parentheses): 11 (C3), 12 (C4), 13 (C5), 14 (C6), 15 (C1) and16 (C2). These atom names are shown in black on the inside of the benzene-1 ring in Figure 3.5. If the x-axis for benzene-1 is chosen to run through atoms 14 (C6) and 11 (C3) and the x-axis of the benzene standard also runs through atoms C6 and C3, then to be consistent with the benzene standard, the y-axis of benzene-1 should be defined in a similar way, i.e. to be a perpendicular bisector of the C1-C2 bond as shown in Figure 3.5. The x- and y-axes for benzene-2 were defined in a similar fashion. The x-axis was set along the C-F bond. The y-axis was defined using the convention of benzene-1. In benzene-1 the atom ID's run consecutively from 11 to 16, with C11 being attached to C10. In benzene-2 the atom ID's run consecutively from 17 to 22, with C17 being attached to C10. Since in benzene-1 the y-axis is the perpendicular bisector of the side of the ring with the higher atom ID's (i.e. 15 and 16), the same convention was chosen for benzene-2: the y-axis was set as the perpendicular bisector of the side of the ring with the higher atom ID's (i.e. 21 and 22). The same convention was used for locating the y-axis in napthalene after the x-axis was set along the atom ID’s 24 and 27.

In DM324 the nonplanar piperazine ring is defined by atom ID's 1 to 6. The sample input conformation file in Appendix C shows that the atom names corresponding to these atom ID's are (in parentheses): 1 (N2), 2 (C2), 3 (C1), 4 (N1), 5 (C4), and 6 (C3). These atom names are shown in black on the inside of the piperazine ring of DM324 in Figure 3.5. It was decided that the x-axis of the piperazine ring would be in the plane defined by atoms 1, 2, 4 and 5, and would pass through the projections of carbons with atom ID's 6 and 3 onto that plane. Since this is the orientation of the x-axis in the piperazine standard, the y-axis was defined as in the piperazine standard to be the perpendicular bisector of the N2-C2 bond.

In all cases the z-axis is defined by the right-hand rule.

3.2.4.3 Mapping : Once the coordinate frames are assigned to the planes of the ring fragments in DM324, the mapping is straight-forward. Figure 3.5 shows how the benzene standard (in red) is mapped onto the benzene-1, benzene-2 rings and napthalene rings of DM324 (in black) and the piperazine standard (in green) is mapped onto the DM324 piperazine (in black). Note that the mapping of the benzene standard follows a particular convention. C3 of the standard is mapped to the atom of the DM324 ring that is the point of attachment to the rest of the molecule: 11 in benzene-1, 17 in benzene-2, and 24 in napthalene. Then the atoms of the standard are mapped in the following order: C3, C4, C5, C6, C1, C2 to the atoms of the DM324 ring in order of their increasing atom ID's as shown in Table 3.1. The least-squares fitting procedure (Horn, 1987) is then performed to obtain the origin and unit vectors X, Y, Z, representative of the reference coordinate frame of the DM324 ring plane. Table 3.2 shows the alignment of the piperazine atom ID's. Only the four atoms used to define the piperazine plane are used for the alignment and least-squares fitting procedure.

STANDARD BENZENE		BENZENE-1	BENZENE-2	NAPTHALENE
Atom ID	**Atom name**	**Atom ID**	**Atom ID**	**Atom ID**
3	C3	11	17	24
4	C4	12	18	25
5	C5	13	19	26
6	C6	14	20	27
1	C1	15	21	32
2	C2	16	22	33

Table 3.1 Mapping of benzene standard to DM324 ring planes

STANDARD PIPERAZINE		PIPERAZINE
ATOM ID	**ATOM NAME**	
1	N2	1
2	C2	2
4	N1	4
5	C4	5

Table 3.2 Mapping of piperazine standard to DM324 ring plane

3.2.5 Output

The four planes of DM324 are represented as N, B1, B2, and P for napthalene benzene-1, benzene-2 and piperazine respectively (Figure 3.1). For each conformation there are 6 pairs of planes: {(N X B1), (N X B2), (N X P), (B1 X B2), (B1 X P), (B2 X P)}. Displacement parameters Shift (D_x), Slide (D_y) and Rise (D_z), and rotational parameters Tilt (τ), Roll (ρ) and Twist (ω) were calculated for each pair of planes. The output parameters are calculated such that one plane is given a priority over the other. The convention followed here is that the first plane is primary and the second is secondary. For example, the three rotational and translational parameters for N X B1 represent the relative rotation and displacement of plane B1 with respect to N. In other words the B1 X N parameters have the same magnitude but opposite sign as the N X B1 parameters. The Planes program was run both ways to test the validity of this statement.

Overall a total of 728 x 6 x 6 parameters were calculated. The program output can be in two formats based on user requirements. One type is the descriptive format. The other type is the spreadsheet style where output is printed out as a space-delimited text and can be directly imported using Excel, Matlab, or any other programs for analysis. See Appendix C for a part of both formats of the program output. For such a huge dataset as that of DM324, the spreadsheet style output is preferable.

3.3 Verification of the Planes Algorithm

Geometrically the distance between the centroids of the pair of planes is the net change in displacement parameters, $D = (D_x^2 + D_y^2 + D_z^2)^{1/2}$. The angle between the planes, $A = (\tau^2 + \rho^2)^{1/2}$, is the net change in out-of-plane bending parameters. These quantities can be calculated using the SYBYL molecular modeling program. The conformations of DM324 were stored in SYBYL's Molecular Spreadsheet™. SYBYL's "INSERT COLUMN" function was used to populate the Spreadsheet™ with the distances between the centroids and the angles between each pair of planes. The distances and angles from the Spreadsheet™ were compared to the values obtained from the Planes program as follows.

Step 1: Consider a pair of planes from one conformation, e.g. the naphthalene plane **N** and piperazine plane P.

Step 2: Evaluate $D = (D_x^2 + D_y^2 + D_z^2)^{1/2}$ and $A = (\tau^2 + \rho^2)^{1/2}$, where D_x, D_y, D_z, τ and ρ are obtained from the Planes program. Identify these quantities as PD and PA to indicate that they are calculated with the Planes program.

Step 3: Use SYBYL to measure D, the distance between the centroids of the planes and A, the angle between the planes. Identify these quantities as SD and SA to indicate that they are calculated with SYBYL.

Step 4: Compare SD with PD and SA with PA. Since the angles are complementary, PA=SA or PA=180°-SA.

CHAPTER 4

RESULTS

For each conformation of DM324, the local coordinate reference frame assigned to each plane is shown in Figure 4.1. As mentioned above the z-axis (not shown) is perpendicular to the x-y plane. The direction of the positive z-axis is determined by the cross product of the x- and y-axes. The figure shows only the non-hydrogen atoms (nitrogen, oxygen, fluorine, and carbon). The carbon atom locations are not labeled with a "C", but they are found at the intersection of bond lines (i.e. at positions 2, 3, 5, 6, 7, 8, 10, 11-16, 17-22, and 23-33). A total of 728 X 6 X 6 parameters were evaluated by the Planes program: 728 conformations, six pairs of planes for each conformation and six parameters (three translation and three rotation) for every pair of planes. These parameters give the relative distance and orientation of the planes containing important pharmacophore features for each minimum energy conformation of DM324.

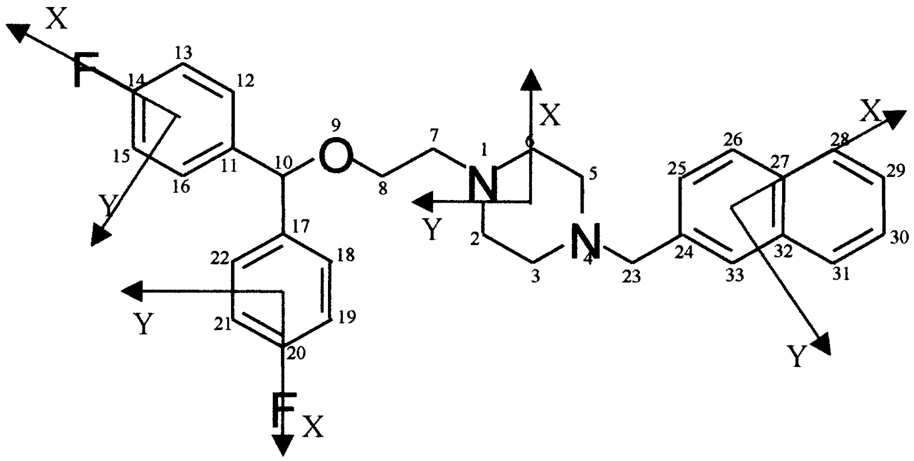


Figure 4.1 DM324 with local coordinate frames assigned by Planes program

Parameters evaluated using the Planes program are verified as shown in Table 4.1. The table shows verification samples for 3 conformations rs2_00001, rs2_00002 and rs2_00003. Similar verification was performed over all 728 conformations. The Planes program evaluates 3 displacement parameters Shift (D_x), Slide (D_y), Rise (D_z) in Angstrom units and 3 rotational parameters Tilt (τ), Roll (ρ) and Twist (ω) in degrees. For NXB1 in rs2_00001, PD represents the distance between the centroids of N (naphthalene plane) and B1 (benzene-1 plane) calculated using the displacement parameters as $PD = (D_x^2 + D_y^2 + D_z^2)^{½}$. SD represents the distance between the two centroids computed using SYBYL. For NXB1 in rs2_00001, $PD=(0.250^2 + 14.158^2 + 0.747^2)^{1/2}=14.179$ A^{o} and SD= 14.180 A^{o}. The values are almost identical thus verifying the correctness of the displacement parameters evaluated by the Planes program. Similarly PA represents the angle between planes calculated using Planes program parameters by applying the formula $PA= (\tau^2 + \rho^2)^{1/2}$. SA gives the angle between the planes calculated using SYBYL. For NXB2 in rs2_00001, $PA=(-2.320^2 + 11.269^2)^{1/2}=11.505$ and is almost equal to SA=11.500 showing that the rotational parameters evaluated by the Planes program are correct. It should be noted that either PA=SA or PA=180-SA can be used to verify the rotational parameters. For NXB2 in rs2_00001 where PA=87.869^0, SA=92.130 and hence 180-SA=87.870^0 which is almost equal to PA. For easy comparison, values to be compared are represented using same color: PD, SD values in blue and PA, SA or 180-SA values in red.

CONFORMATION	PARAMETERS CALCULATED USING PLANES PROGRAM									CALCULATED USING SYBYL		
rs2_00001		Shift (D_x) [a]	Slide (D_y) [a]	Rise (D_z) [a]	Tilt (τ) [b]	Roll(ρ) [b]	Twist(ω) [b]	PD[a]	PA[b]	SD[a]	SA[a]	(180-SA)[b]
	N X B1	0.250	14.158	0.747	-2.320	11.269	130.923	14.179	11.505	14.180	11.500	
	N X B2	-5.830	-1.202	10.525	-4.850	-87.735	-40.111	12.090	87.869	12.092	92.130	87.870
	N X P	-2.410	4.715	-0.408	-16.890	64.334	82.960	5.310	66.515	5.312	66.530	
	B1 X B2	0.710	4.633	-0.016	76.260	39.458	178.241	4.690	85.862	4.686	94.140	85.860
	B1 X P	-6.340	-6.248	0.026	45.420	32.499	-45.418	8.900	55.848	8.900	55.860	
	B2 X P	-5.390	-2.112	-4.117	-23.750	-95.939	-166.738	7.100	98.836	7.102	81.190	98.810
rs2_00002												
	N X B1	1.230	13.476	2.756	-40.420	-135.340	120.898	13.810	141.246	13.810	141.250	
	N X B2	-5.720	-2.901	-9.607	31.660	71.572	-63.315	11.550	78.261	11.550	78.260	
	N X P	-2.450	4.589	-0.790	-26.810	63.172	71.432	5.260	68.627	5.264	68.650	
	B1 X B2	0.030	2.900	3.707	-2.380	-93.363	57.932	4.710	93.393	4.707	93.390	
	B1 X P	-6.330	-4.429	-3.917	-112.460	-96.130	-108.809	8.660	147.947	8.663	147.910	
	B2 X P	-2.070	-2.554	6.089	-107.450	-1.465	68.235	6.920	107.462	6.919	107.500	
rs2_00003												
	N X B1	1.850	13.592	-0.590	-51.120	-85.406	149.871	13.730	99.538	13.730	99.540	
	N X B2	-5.440	-6.415	-7.836	-5.100	45.536	-105.861	11.500	45.822	11.497	45.820	
	N X P	-2.500	4.459	-1.059	-38.270	61.301	64.013	5.220	72.263	5.219	72.290	
	B1 X B2	-0.030	2.598	3.912	3.840	-96.152	51.404	4.700	96.229	4.696	96.230	
	B1 X P	-5.020	-3.932	-5.976	130.730	90.462	28.798	8.740	158.974	8.741	158.970	
	B2 X P	2.180	-1.145	6.159	-86.480	-30.059	133.967	6.630	91.551	6.633	91.590	

[a] Angstroms
[b] Degrees

$PD = (D_x^2 + D_y^2 + D_z^2)^{1/2}$
$PA = (\rho^2 + \tau^2)^{1/2}$

N: Napthalene P: Piperazine B1: Benzene-1 B2: Benzene-2

Table 3.3 Verification of the Planes Program

4.1 Scatter Plot Analysis for Overview of Molecular Conformations

In order to obtain an overview of the range of the Planes parameters, three-dimensional scatter plots (Figures 4.2 to 4.7) were constructed for the translational and rotational parameters for every pair of planes for all 728 conformations. The following shorthand notation is used to describe the planes: N (napthalene, plane defined by atom ID's 24, 25, 26, 27, 32,33), P (piperazine, plane defined by atom ID's 1, 2, 4, and 5), B1 (benzene 1, plane defined by atom ID's 11-16) and B2 (benzene 2, planes defined by atom ID's 17-22). For the displacement parameters, Shift is plotted along the x-axis, Slide along the y-axis and Rise along the z-axis. For the rotational parameters, Tilt is plotted along the x-axis, Roll along the y-axis and Twist along the z-axis. A MATLAB program was written to import the spreadsheet-style output of the Planes program and create the three-dimensional plots.

Scatter plots of the Planes parameters for N X B1 (Figure 4.2) and N X B2 (Figure 4.3) show a great deal of similarity, as expected. Figure 4.1 shows that the two benzene rings are located close to each other in space, separated only by the carbon at position 10. The sp^3 hybridization of C10 forces each of the functionalities to which it is bonded (i.e. B1, B2, H, and O9) to the corners of a tetrahedron, resulting in a relatively restricted spatial relationship of B1 with respect to B2. Table 4.1 shows the minimum, maximum, mean and the standard deviation of the parameters for all six pairs of planes over 728 conformations. The distance between the centroids of the two benzene rings, $D = (D_x^2 + D_y^2 + D_z^2)^{1/2}$, is from 4.6 to 5.1 A° with the average value of 4.8 A° as shown in Table 4.2 It is therefore reasonable that B1 and B2 should display the same general range of distances and orientations compared to N. For the same reason, the B1 X P (Figure 4.4) and B2 X P (Figure 4.5) Planes parameter scatter plots are also similar.

The N X P scatter plots (Figure 4.6) demonstrate the restricted range of locations of the naphthalene plane with respect to the piperazine plane in the minimum energy conformations of the molecule. This is due to the three-fold rotational symmetry of the conformational energy minima around the N4-C23 bond (i.e. see the example of staggered conformations in Figure 1.3) combined with the two-fold rotational symmetry of minima around the C23-C24 bond. Table 4.2 shows that the distance between the centroids of the napthalene and piperazine rings, D, is 4.4 to 5.3 A^o with the average distance of 4.9 A^o.

Table 4.2 Range of the Planes parameter values over 728 conformations

N: Napthalene B1: Benzene-2 B2: Benzene-2 P: Piperazine
N X B1 Relative displacement and rotation of B1 with respect to N
N X B2 Relative displacement and rotation of B2 with respect to N
N X P Relative displacement and rotation of P with respect to N
B1 X B2 Relative displacement and rotation of B2 with respect to B1
B1 X P Relative displacement and rotation of P with respect to B1
B2 X P Relative displacement and rotation of P with respect to B2
Shift(D_x) : Displacement along x-axis
Slide(D_y): Displacement along y-axis
Rise(D_z): Displacement along z-axis
Tilt (τ): Rotation about x-axis
Roll(ρ): Rotation about y-axis
Twist(ω): Rotation about z-axis
Cent-Dist: $(D_x^2 + D_y^2 + D_z^2)^{½}$ (Distance between the centroids of the planes)
Plane-Ang: $(\tau^2 + \rho^2)^{½}$ (Angle between the planes)

Plane Pair Parameter	Minimum	Maximum	Mean	Standard Deviation
NxB1 Shift	-10.31	9.86	-0.46	4.03
NxB1 Slide	-13.57	14.16	-0.76	7.58
NxB1 Rise	-11.63	13.38	1.77	6.03
NxB1 Tilt	-172.79	170.04	-1.06	68.31
NxB1 Roll	-154.63	156.16	-9.49	69.94
NxB1 Twist	-176.42	179.42	-6.31	106.75

Plane Pair Parameter	Minimum	Maximum	Mean	Standard Deviation
NxB1 Cent-Dist	3.99	14.18	10.45	2.16
NxB1 Plane-Ang	1.74	176.22	90.02	39.18
NxB2 Shift	-10.89	9.87	-0.61	3.69
NxB2 Slide	-13.16	13.67	0.21	7.77
NxB2 Rise	-12.54	12.81	0.59	6.17
NxB2 Tilt	-163.94	169.82	5.95	69.52
NxB2 Roll	-156.48	165.53	3.74	71.11
NxB2 Twist	-179.72	179.71	-1.64	103.55
NxB2 Cent-Dist	3.99	13.97	10.41	2.10
NxB2 Plane-Ang	7.81	175.91	90.27	42.17
NxP Shift	-3.80	2.75	-1.79	1.38
NxP Slide	-4.18	5.18	2.74	2.43
NxP Rise	-4.96	4.60	0.95	2.19
NxP Tilt	-111.15	97.64	6.94	55.36
NxP Roll	-125.20	143.97	-2.08	68.09
NxP Twist	-179.98	179.84	58.97	100.08
NxP Cent-Dist	4.41	5.31	4.91	0.25
NxP Plane-Ang	33.55	146.54	84.92	23.06
B1xB2 Shift	-1.48	1.48	-0.05	0.68
B1xB2 Slide	-4.86	4.93	0.53	4.03
B1xB2 Rise	-4.67	4.70	0.33	2.40
B1xB2 Tilt	-79.63	79.05	-7.13	52.57
B1xB2 Roll	-108.31	107.43	-5.50	75.21
B1xB2 Twist	-179.82	179.94	20.17	131.03
B1xB2 Cent-Dist	4.56	5.08	4.78	0.12
B1xB2 Plane-Ang	53.72	126.70	90.52	17.23
B1xP Shift	-8.53	5.48	-3.99	3.20
B1xP Slide	-6.81	4.89	-2.30	2.12
B1xP Rise	-8.61	7.83	-0.30	3.66
B1xP Tilt	-157.92	156.48	1.29	74.02
B1xP Roll	-143.58	169.43	0.12	58.58
B1xP Twist	-179.83	179.28	-12.74	100.03
B1xP Cent-Dist	3.81	8.90	6.93	1.17
B1xP Plane-Ang	1.88	170.93	85.96	38.89
B2xP Shift	-8.43	4.80	-4.09	2.94
B2xP Slide	-7.05	4.73	-2.57	2.09
B2xP Rise	-8.11	8.26	0.24	3.58
B2xP Tilt	-159.36	156.68	-3.75	70.28
B2xP Roll	-174.87	173.92	5.67	67.72
B2xP Twist	-179.48	179.33	-15.82	97.97
B2xP Cent-Dist	3.79	8.89	6.93	1.08
B2xP Plane-Ang	1.30	176.17	90.20	37.75

Table 4.2 (continued) Range of the Planes parameter values over 728 conformations

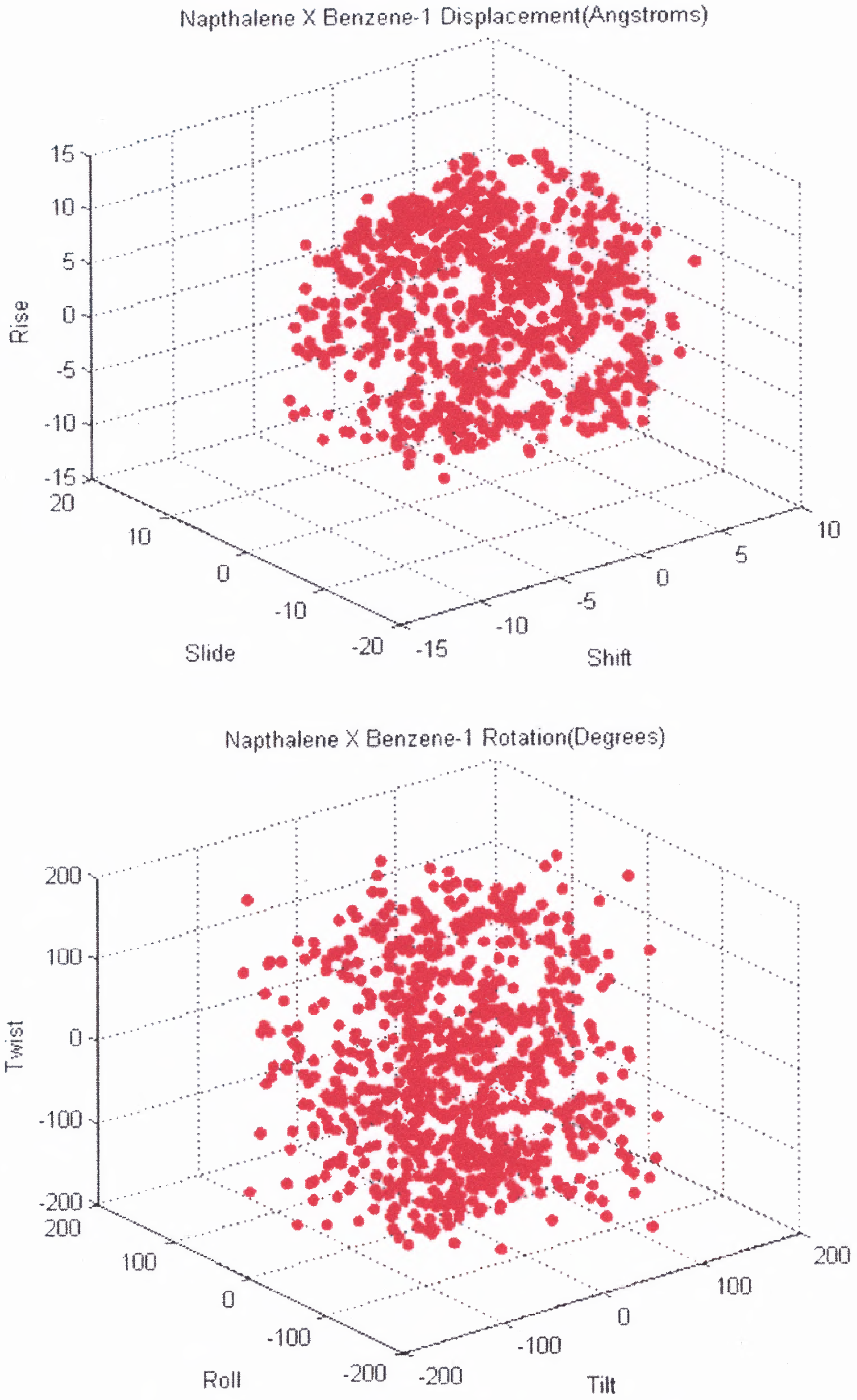


Figure 4.2 3D scatter plots Napthalene X Benzene-1

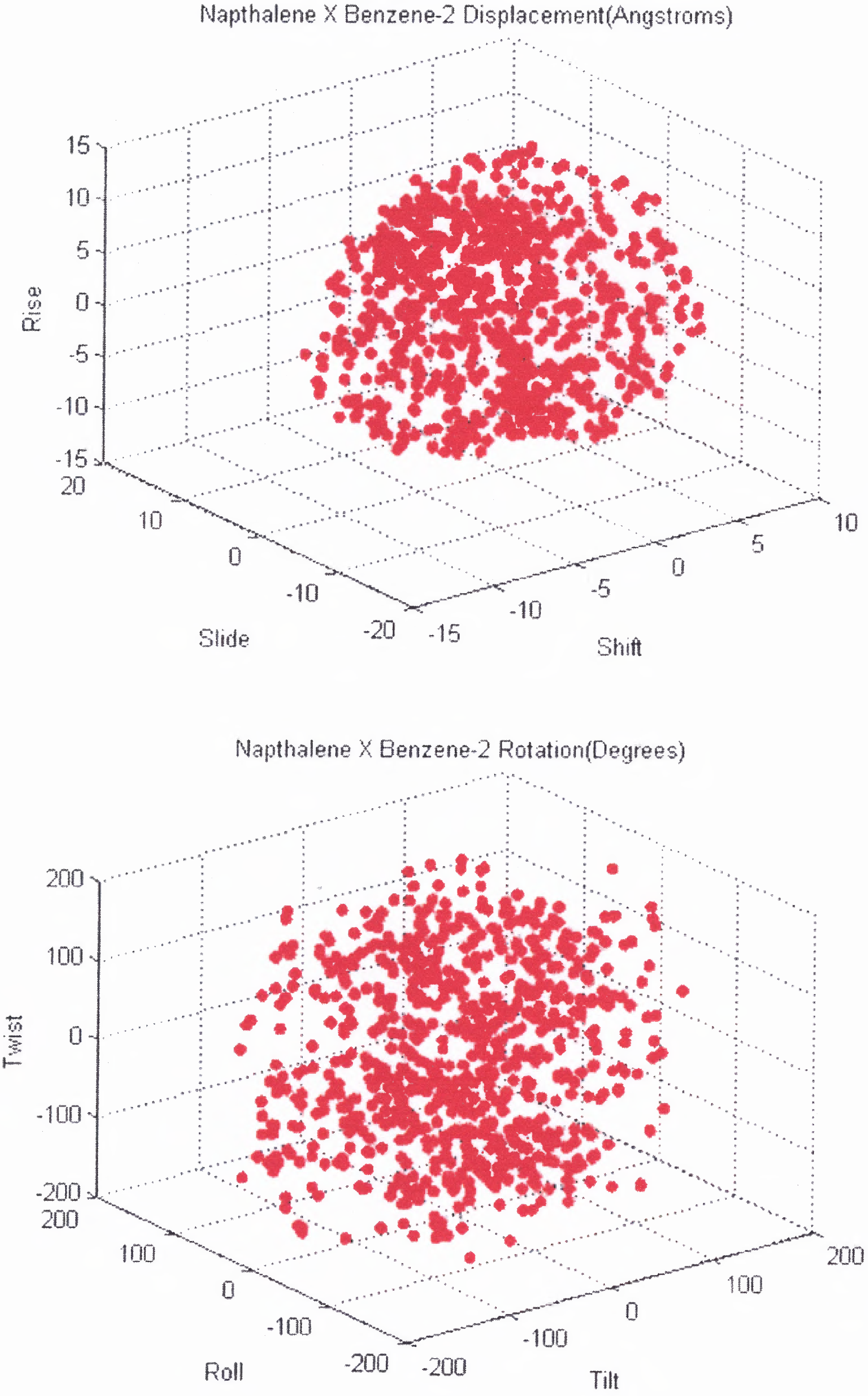


Figure 4.3 3D scatter plots Napthalene X Benzene-2

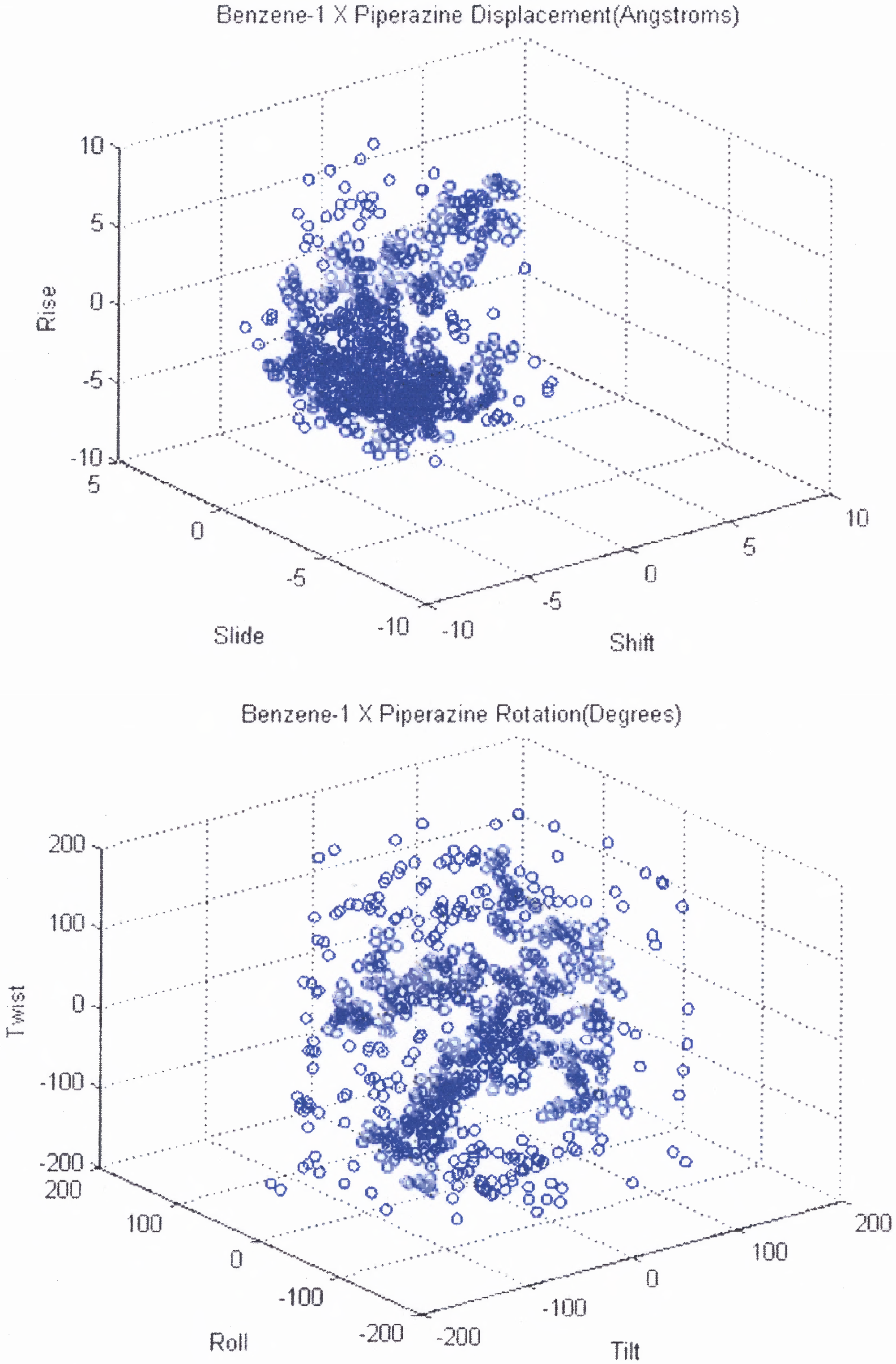
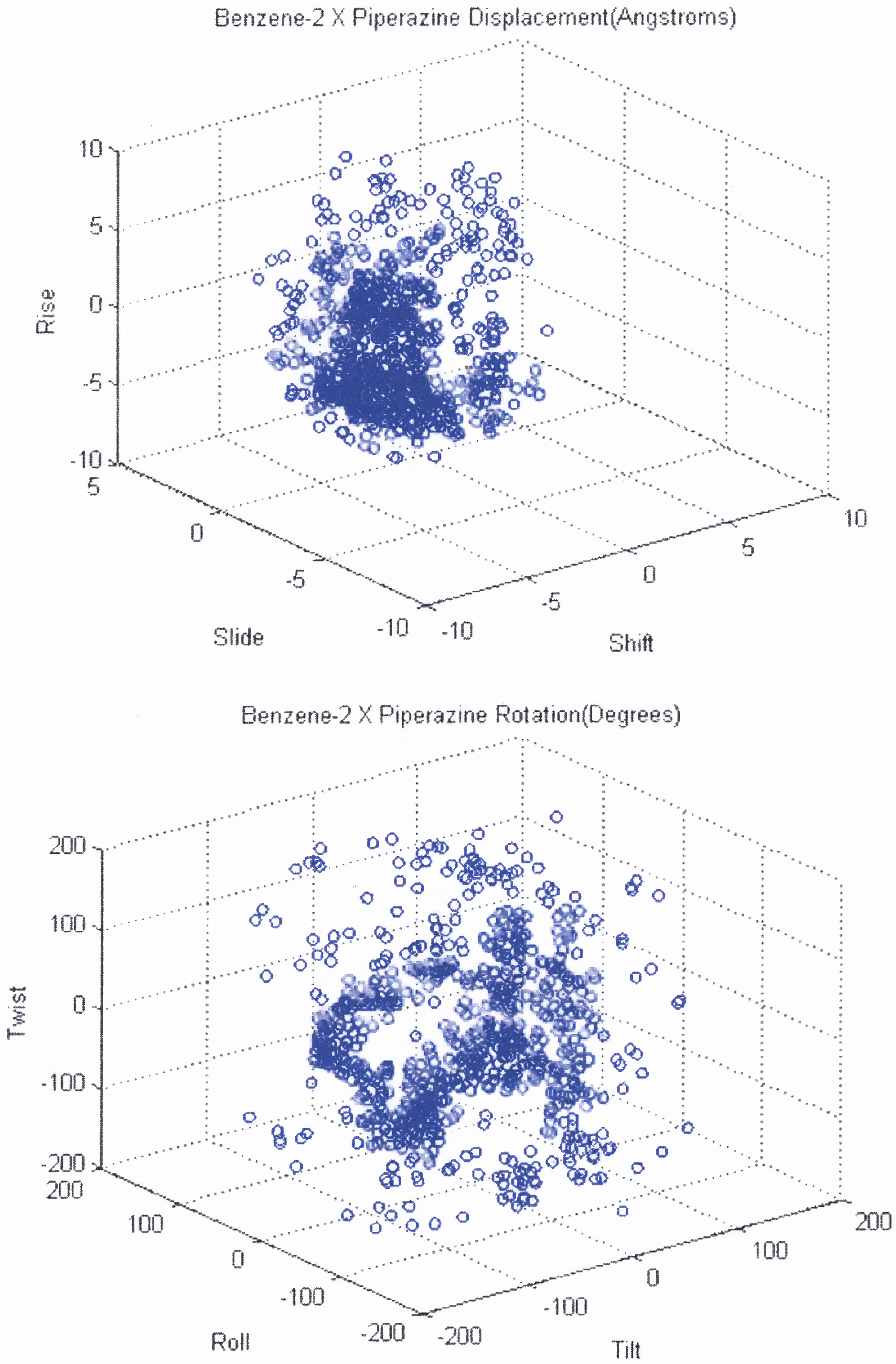


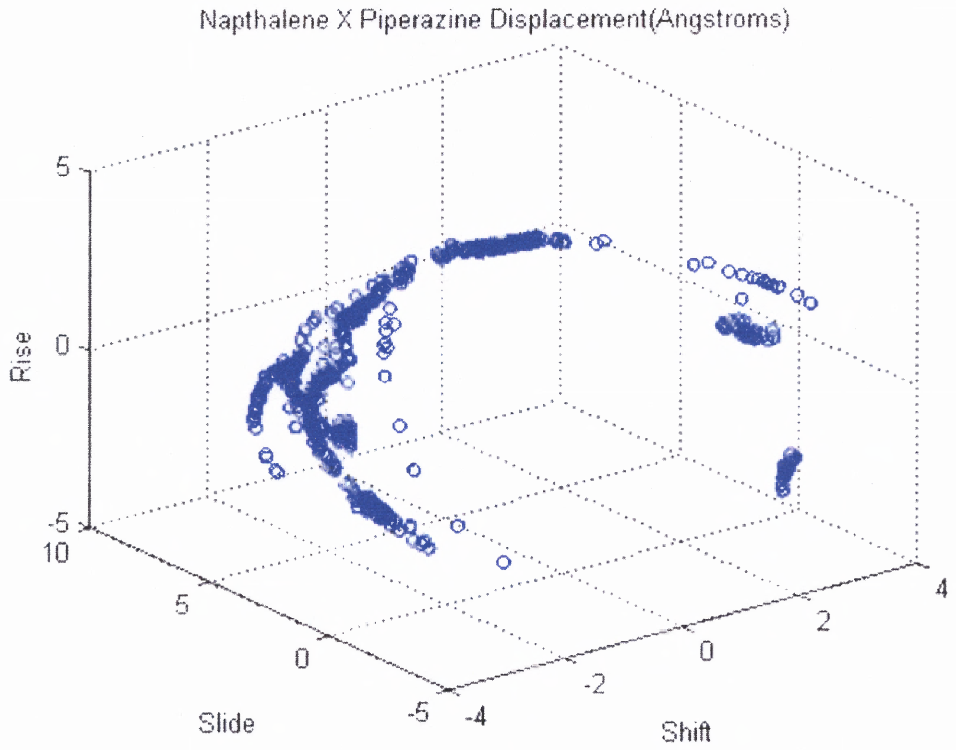
Figure 4.4 3D scatter plots Benzene-1 X Piperazine



Benzene-2 X Piperazine Displacement(Angstroms)

Benzene-2 X Piperazine Rotation(Degrees)

Figure 4.5 3D scatter plots Benzene-2 X Piperazine



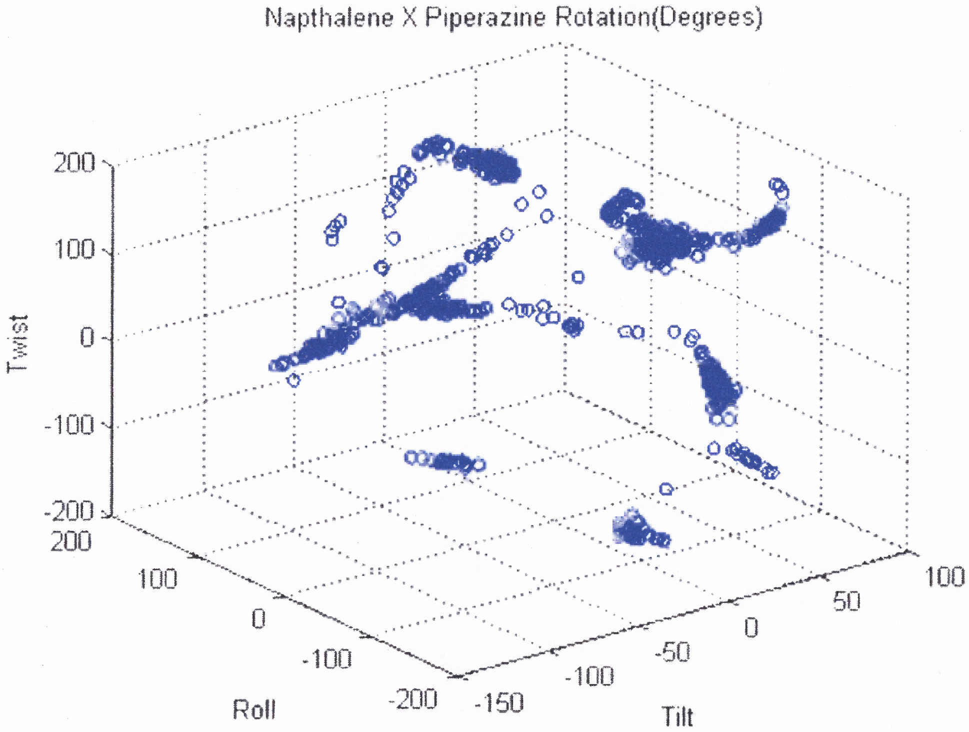


Figure 4.6 3D scatter plots Napthalene X Piperazine

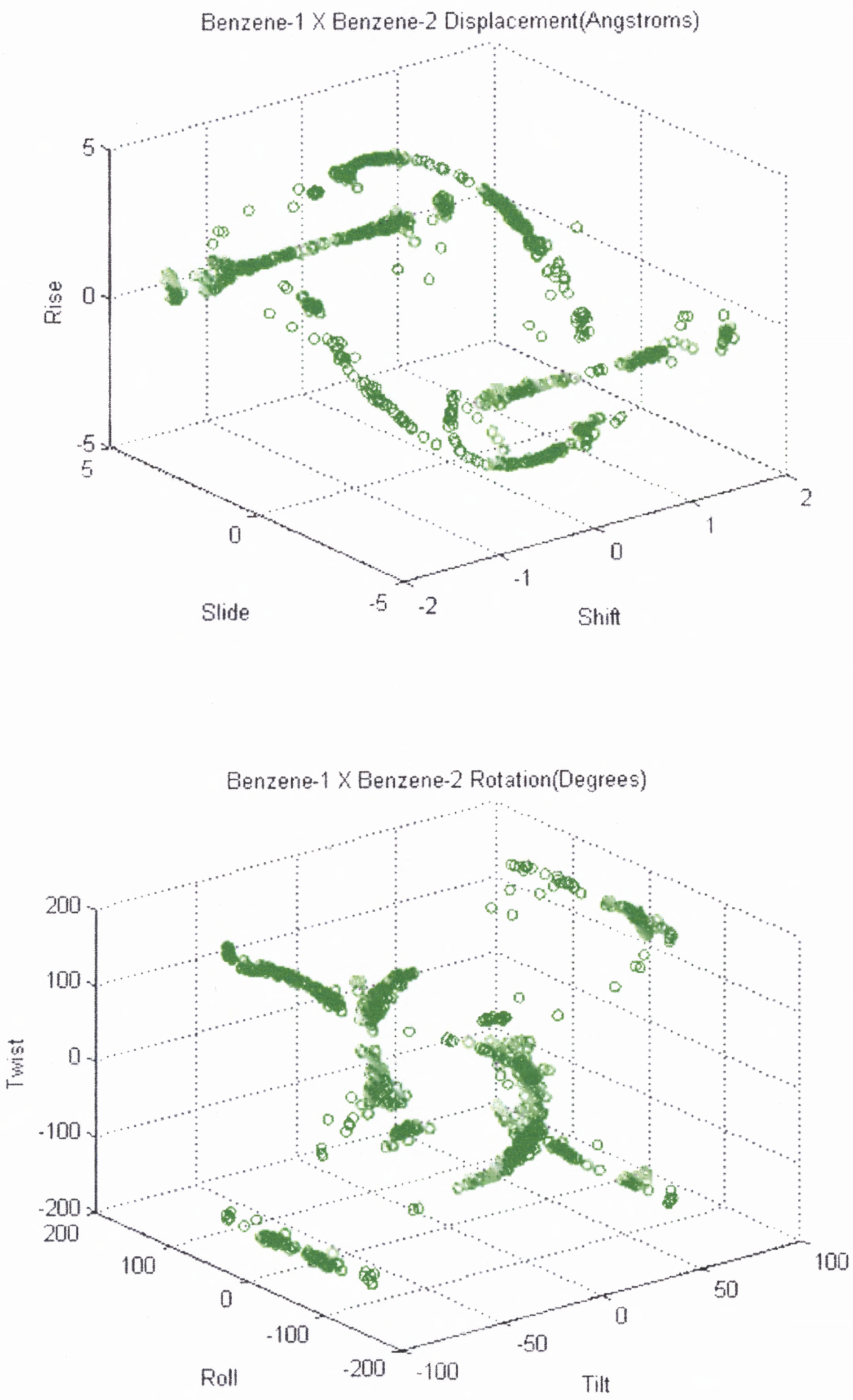


Figure 4.7 3D scatter plots Benzene-1 X Benzene-2

An additional level of data analysis was provided by color-coding the conformational energy minima data points by the relative energy of the conformations. As described in the Methods section, the Random Search procedure generated a set of 728 conformations of DM324 with energies within 20 kcal/mol of the conformation found by the search to have the global energy minimum (GEM) value. The Relative Energy (RE) for each of the conformations, i.e. relative to the GEM conformer, was calculated by subtracting the GEM value from the energy of each conformation. In other words the energy of each conformer was expressed relative to that of the GEM conformer taken as 0 kcal/mol. Conformations were then assigned to energy bins as follows: 0-4 bin: 0 kcal/mol $\leq$ RE < 4 kcal/mol, 4-8 bin: 4 kcal/mol $\leq$ RE< 8 kcal/mol, and so on. The data in the scatter plots of Figures 4.2 to 4.7 were color-coded by energy bin. The results are shown in Appendix D. The 3D data plots did not reveal any specific patterns for the parameter values when color coded by energy bins.

4.2 Analysis by Molecular Shapes

A method for defining the shapes of the DM324 conformations was suggested by Deepangi Pandit of the Venanzi group. The distance of closest approach between the centroids of each of the phenyl rings and each of the rings in the naphthalene moiety was used to define the shapes as follows. The shortest distance was calculated between the following points: the centroid of one of the napthalene rings (NL - the "lower" part of the naphthalene moiety consisting of atoms 24, 25, 26, 27, 32, and 33) or NU (the "upper" part of the naphthalene moiety consisting of atoms 27, 28, 29, 30, 31, and 32) and the centroid of one of the benzene rings. There are actually four combinations: NL with B1, NL with B2, NU with B1 and NU with B2. The lowest of these four values was selected and used to classify conformations into shapes. Conformations having the closest distance of approach were classified into the C (cup), I (intermediate between C and V), V (V-shaped or open cup) and E (extended) shapes as defined in Table 4.2. The numbers in each cell of the table are the number of conformers in each shape and energy range.

DM324	0-4[a]	4-8[a]	8-12[a]	12-16[a]	16-20[a]	Total/shape[b]
C (0<D[c]<4.5)	15	6	0	0	0	21
I (4.5<D<5.0)	5	10	1	0	0	16
V (5.0<D<7.0)	4	38	9	4	0	55
E (7.0<D)	10	257	297	67	5	636
Total/bin[d]	34	311	307	71	5	728 total

[a] kcal/mol

[b] Total number of conformers of each shape

[c] Distance of closest approach (in Angstroms) (see text)

[d] Total number of conformers in each energy bin

Table 4.3 DM324 conformations binned by Energy and Shape

The conformers were divided into groups according to the shape descriptions above. Tables in Appendix E show the Planes parameters for each pair of planes binned by energy and by angle and distance range. This data is shown graphically in the histogram plots in Figures 4.8 to 4.13 and in Appendix F. Some of the more obvious trends in the data are discussed below.

The trends in the B1 x B2 histograms (as shown in Figure 4.8 to 4.11) are independent of shape. The spread of planes parameter values (Shift: - 3 Å to 3 Å, Slide: -6 Å to 6 Å, Rise: -6 Å to 6 Å, Tilt: -90° to 90°, Roll: -135° to 135°, and Twist: -180° to 180°) is the same whether the molecule is in the C, I, V, or E shape. This is reasonable because the shapes are defined in terms of the minimum distance between N and B1 or B2, not the distance between B1 and B2. The B1xB2 histograms show, as do the scatter plots in Figure 4.7, the constrained nature of the distance and orientation of B1 with respect to B2 in the bisphenyl unit. This is an important relationship because the bisphenyl unit is found in many drugs that are active in the central nervous system.

The N x P histograms (Figure 4.12 and 4.13) show a somewhat larger dependence on shape which is consistent with the fact that the molecule is "opening up" as it goes from C to E shape. For example the range of Shift values is from -3 Å to 0 Å for the C shape, but -6 Å to 3 Å for the E shape. Similar trends are seen for Slide (C: 0 Å to 6 Å, E: -6 Å to 6 Å) and Rise (0 Å to 6 Å, -6 Å to 6 Å). Similarly the Tilt angle range is -135° to 45° and 45° to 90° in the C shape, but -135° to 135° in the E shape. Roll goes from -90° to 90° in the C shape to -135° to 180° in the E shape. Twist ranges from -90° to 0° and from 90° to 135° in the C, but from -180° to 180° in the E. In general the NxP histograms show that the two rings are constrained in their relative orientation, as can also be seen in their scatter plots (Figure 4.6). This is significant because it is hypothesized that this side of the molecule (rather than the bisphenyl side) is important for the binding of GBR 12909 analogues to the cocaine binding site on the dopamine transporter. In their studies of methylphenidate, the Venanzi group found the piperadine and phenyl (benzene) rings of many of the phenyl-substituted analogues to be similarly constrained.

The B1 x P and B2 x P histograms show similar trends. Inspection of the translational parameters (Shift, Slide, and Rise) shows that most conformations fall into somewhat restricted regions of space for the C shape and take on a wider range of values for as they "open up" to the E shape. This is particularly obvious in the range of Rise values for the C versus E shapes. There doesn't seem to be any particular trend for the angles values of Tilt, Roll, and Twist. The N x B1 and N x B2 histograms show behavior similar to B1 x P and B2 x P with a slightly larger distance range for all shapes.

All the histograms show that the relative number of conformations that fall in each range as well as the energy of those conformations differs for each of the shapes.

4.3 Discussion

The objective of this work was to implement a novel technique for partitioning large numbers of conformations of a molecule into families based on the relative orientation of their pharmacophore features. A program (Planes) based on the approach used to analyze local helical DNA structure was implemented by generalizing the 3DNA program to treat an arbitrary molecule. The Planes program was applied to the analysis of 728 conformations of the flexible GBR 12909 analog, DM324. Planes parameters, defined in the same way as the DNA base-step parameters Shift, Slide, Rise, Tilt, Roll, and Twist, were calculated and plotted for each pair of planes for each of the minimum energy conformers of the molecule. The results show that the data plots were able to reproduce physical trends in the case of ring pairs which, based on the molecular structure of DM324, are known to have a restricted orientation in space: benzene/benzene or napthalene/piperazine. This supports the correctness of the approach.

However, the planes parameters appear to be less useful in classifying the conformations based on orientation of the planes of the rings (napthalene/benzene) which are related to the overall shape. The histograms for N X B1 and N X B2 (Figures 4.12 & 4.13) show that the translational and rotational parameters take on a wide range of values, especially for the E (extended) conformers. One of the difficulties in interpreting the data is that the conformations of DM324 take on a much larger range than those of DNA, which is restricted by the A, B, Z form of the double helix. Analysis of DNA structure is simplified by the fact that the double helix can be considered to be oriented along the z-axis of the molecule and that the planes containing the base pairs and base steps deviate only slightly from being parallel to the molecule's x-y plane. It is easy to visualize the meaning of the translational and rotational parameters by reference to the well-defined orientation of the base pairs and base steps along the DNA double helix. (However, it should be noted that the base-pair and base step parameters are calculated using the x-, y- and z-axes local to the base-pair and base step under consideration and *not* with respect to the molecule's global axis. In other words the z-axis of the base step calculation is not the same as the helix axis of DNA.)

In contrast, in DM324 the napthalene/benzene planes vary widely in their relative spatial orientation as can be seen from Figure 3.2 which shows all 728 conformers superimposed by the atoms of the central piperazine ring. As Figure

3.2 illustrates, it is very difficult to visualize what the Shift, Slide, Rise, Roll, Tilt, and Twist parameters mean when the napthalene and benzene planes from which they are calculated take on such a variety of orientations. The present work suggests that, in order to more easily interpret the Planes parameters, the conformers should first be classified in a way that allows the user to more easily visualize their differences: for example, by their relationship to the central piperazine ring plane. This plane was situated in the x-y plane at z = 0, with N1 at the origin of the coordinate system (see Chapter 3). A suggestion for future work is to classify the conformations by whether the centroids of their napthalene and benzene rings fall "above" (i.e. have positive z-coordinate values) or "below" (negative z-coordinate values) the piperazine ring plane. For example a conformation with positive z coordinates for the napthalene and two benzene ring centroids might be classified as U-shaped. A conformation with a napthalene centroid with a positive z-coordinate and the two benzene centroids with negative z-coordinates (or vice-versa) might be considered to be S-shaped. Some thought would need to be given to the conformers that fall in "grey" areas such as those with one positive z centroid and one negative z centroid for the benzene rings.

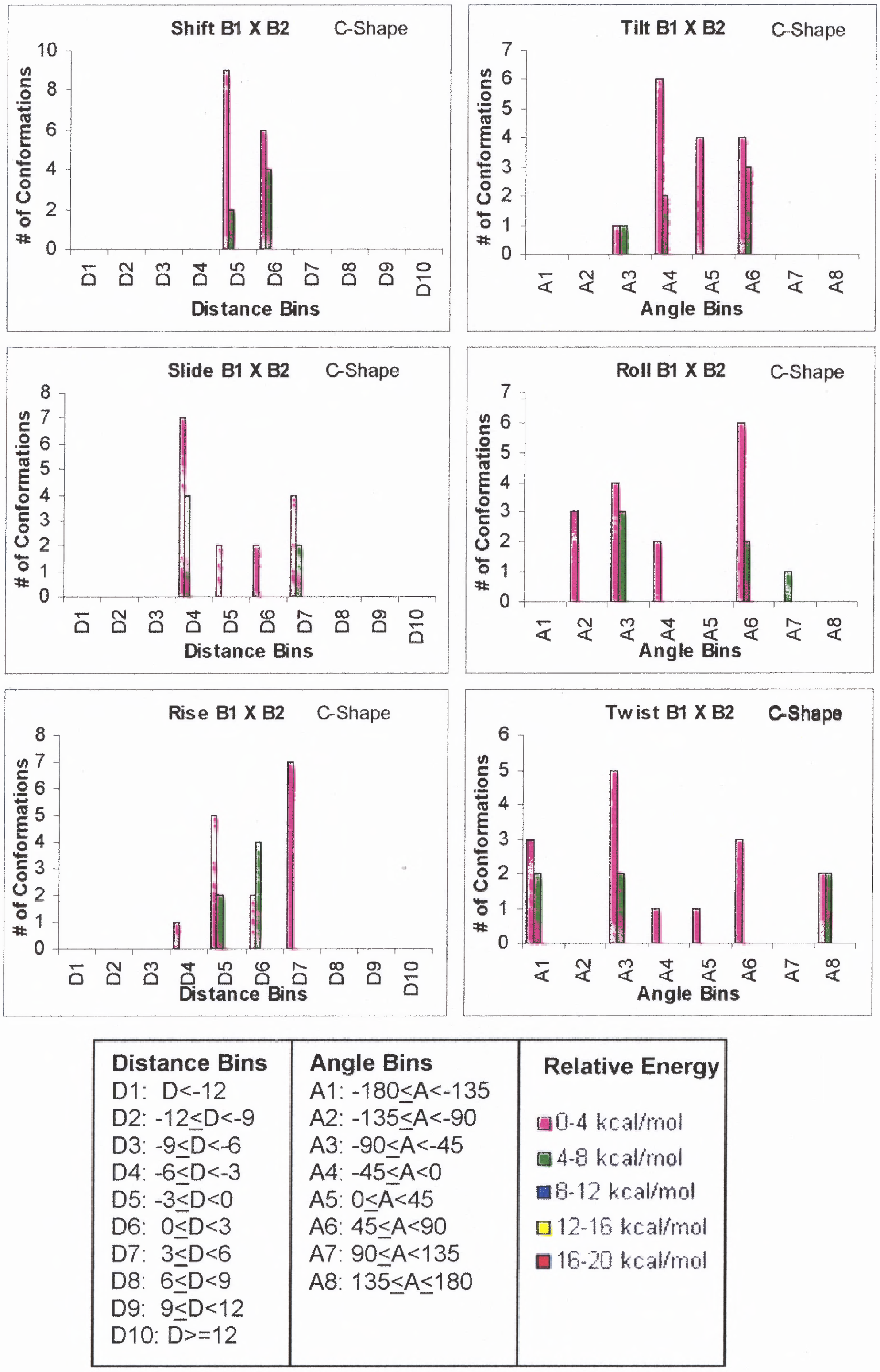
A second suggestion for future work is to redefine the x-axis of the standard piperazine ring. Figure 3.3 shows that in the present work the plane is defined by N2, C2, N1, and C4 (using the labeling scheme of the piperazine ring fragment) with the x-axis along C1...C3 and the y-axis perpendicular to the x-axis. Figure 4.1 shows that this corresponds to the x-axis along atom ID's 3 and 6 (using the numbering of the DM324 molecule). If the x-axis were redefined to contain the two nitrogens (and the y-axis defined perpendicular to that), then the conformers could be classified not only by the position of their napthalene and benzene rings as being above or below the piperazine ring, but by their position to the "right" (negative y-coordinate values) or to the "left" (positive y-coordinate values). This would allow the conformers of DM324 to be viewed in a more standard orientation (similar to the orientation of DNA bases) and would potentially make the interpretation of the Planes parameters easier and more meaningful.

The most difficult aspect of trying to interpret the Planes parameters is the fact that, because of the flexibility of the molecule, the z-axes of the four ring planes are not necessarily pointing in the same direction. In fact the relative orientation of the z-axes varies widely among the conformers. In contrast, for DNA the z-axes of the bases in a Watson-Crick base pair are pointing in opposite directions because they are on two anti-parallel strands. For single-stranded DNA and Hoogsteen base pairs, the bases are oriented in the same directions (with z-axes pointing in same directions). This makes the interpretation of the DNA parameter values somewhat easier. Although the Planes program aligns the coordinate frames of the individual ring planes in the process of calculating the parameters, it is difficult for the user to easily visualize the Shift, Slide, Rise, Tilt, Roll, and Twist parameters and this makes their use in *a priori* classification of molecular conformations somewhat questionable.

However, the Planes parameters may be particularly useful in comparing drugs from different structural classes. For example, methylphenidate (Figure 1.5) is also a dopamine reuptake inhibitor which has some pharmacophore features in common with DM324: a nitrogen in a piperadine ring and a benzene (phenyl) ring. Studies by the Venanzi group in collaboration with Howard Deutsch, Georgia Institute of Technology, and Margaret Schweri, Mercer University Medical School, have lead to a pharmacophore model for methylphenidate binding to the dopamine transporter. The pharmacophore model consists of distances and orientations between the phenyl and piperadine rings. It has been hypothesized that DM324 and methylphenidate bind to the same site on the dopamine transporter and hence have the same pharmacophore for binding. As noted above in the Results section, the scatter plots and histograms show a constrained spatial relationship between the naphthalene (which contains the benzene pharmacophore feature) and piperazine rings in DM324. The Venanzi group found a similar relationship in methylphenidate. Another suggestion for further work is to more completely analyze this similarity by using the Planes program to calculate the parameters from the proposed pharmacophore of methylphenidate and then search the 728 DM324 conformations to find those which have relatively the same pharmacophore as methylphenidate. This would be an excellent way to select DM324 conformers for additional pharmacophore modeling using the CoMFA program. The selected conformation which gave the best CoMFA model (i.e. best described the experimental binding data) would be assumed to contain the optimal orientation of the pharmacophore features for the binding of DM324 to the dopamine transporter. Other more rigid analogs could be constructed to "freeze out" this orientation and improve binding even further.

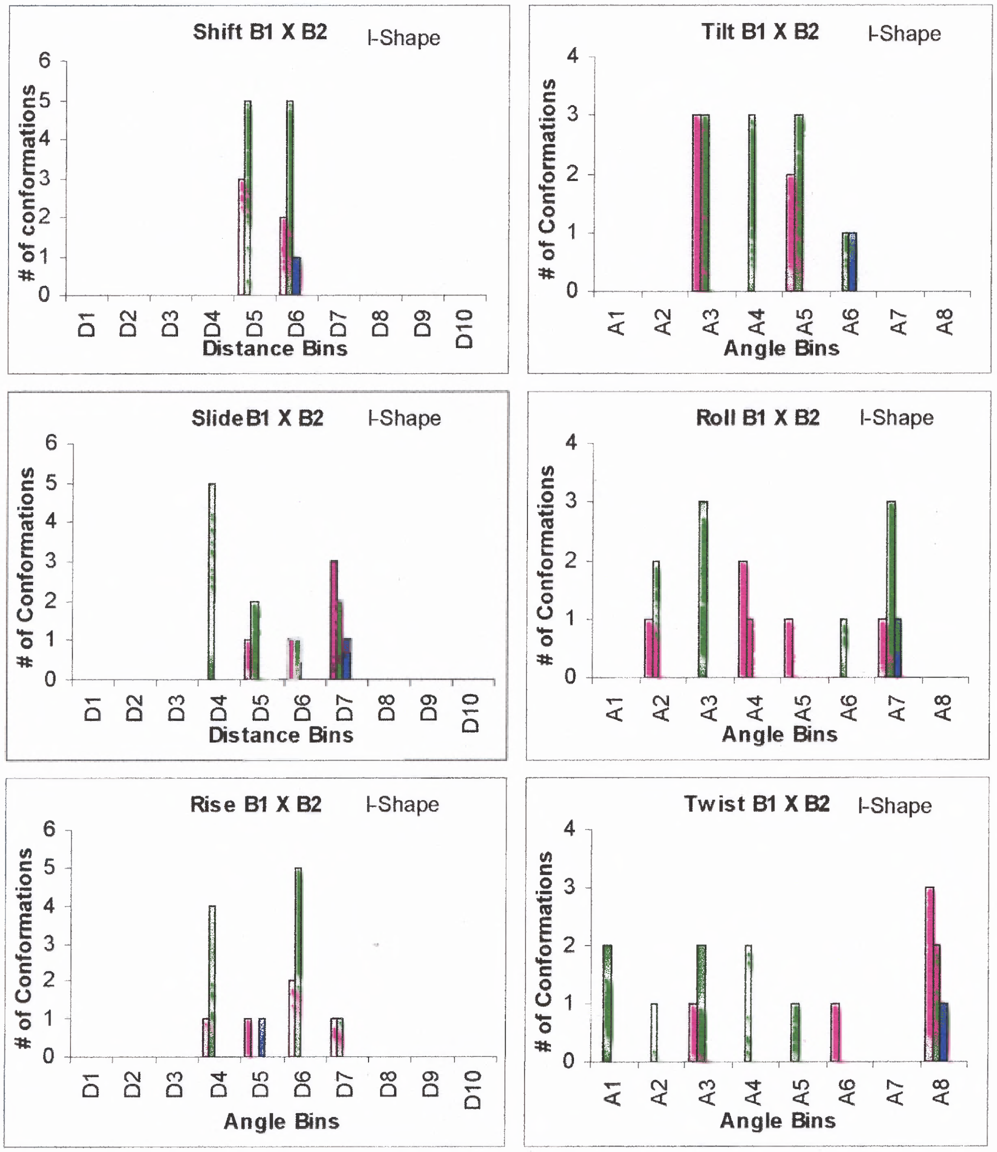
In addition, as noted above, the planes parameters were able to quantify the constrained spatial relationship of the benzene units in the bisphenyl group. Since this group is typical of many drugs that work on the central nervous system, this information may be useful in future work on other drugs that involve the bisphenyl pharmacophore unit.

In summary, the present work suggests that the Planes program could be potentially useful in the description of the relative orientation of pharmacophore features if the molecular conformations could first be classified by their orientation relative to the central piperazine ring. The Planes program may be shown to be particularly useful in comparing the pharmacophore features of different classes of dopamine reuptake inhibitors.



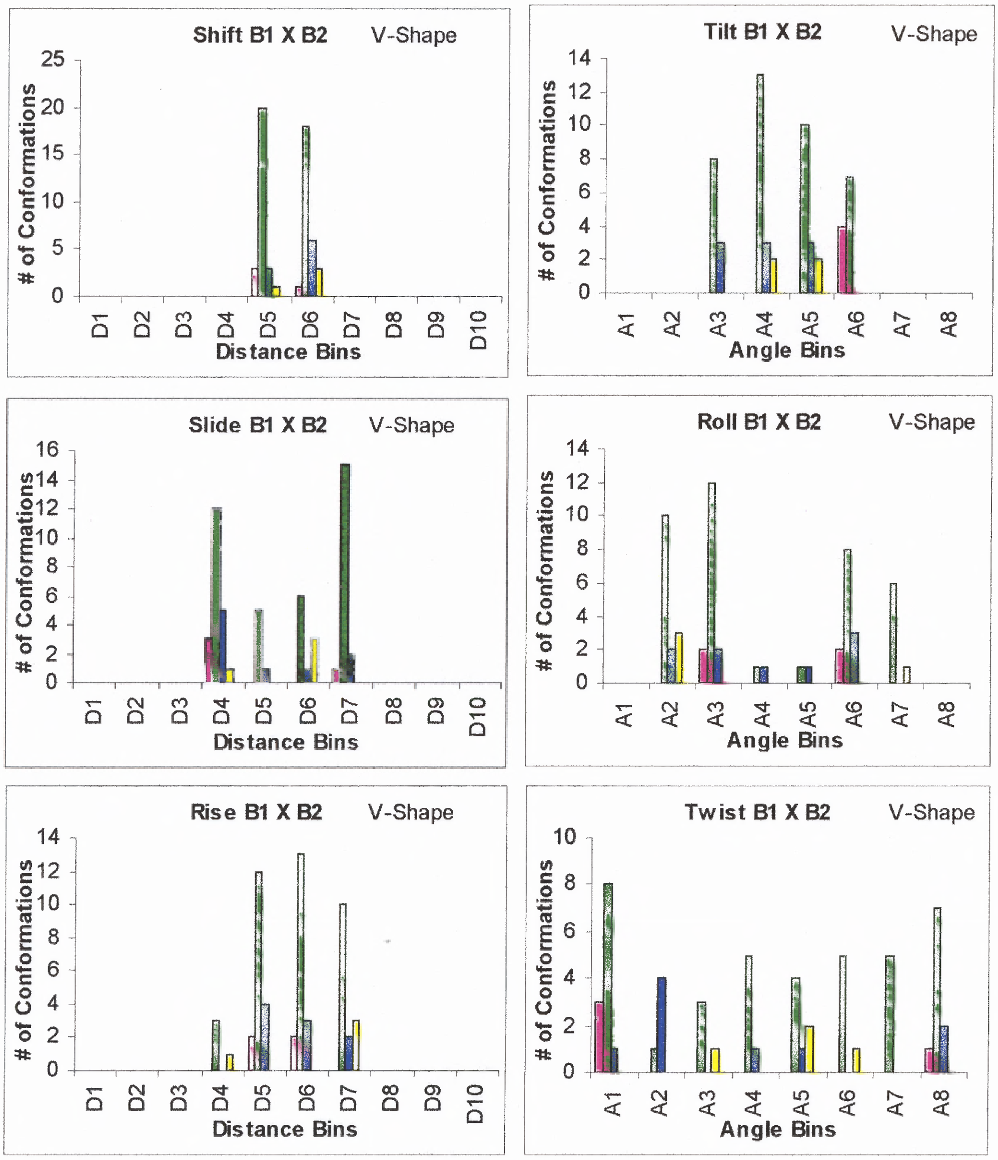
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	🟪 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟩 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟦 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟨 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	🟥 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure 4.8 Histograms of displacement and rotational parameters for B1 X B2 in C-Shaped conformations



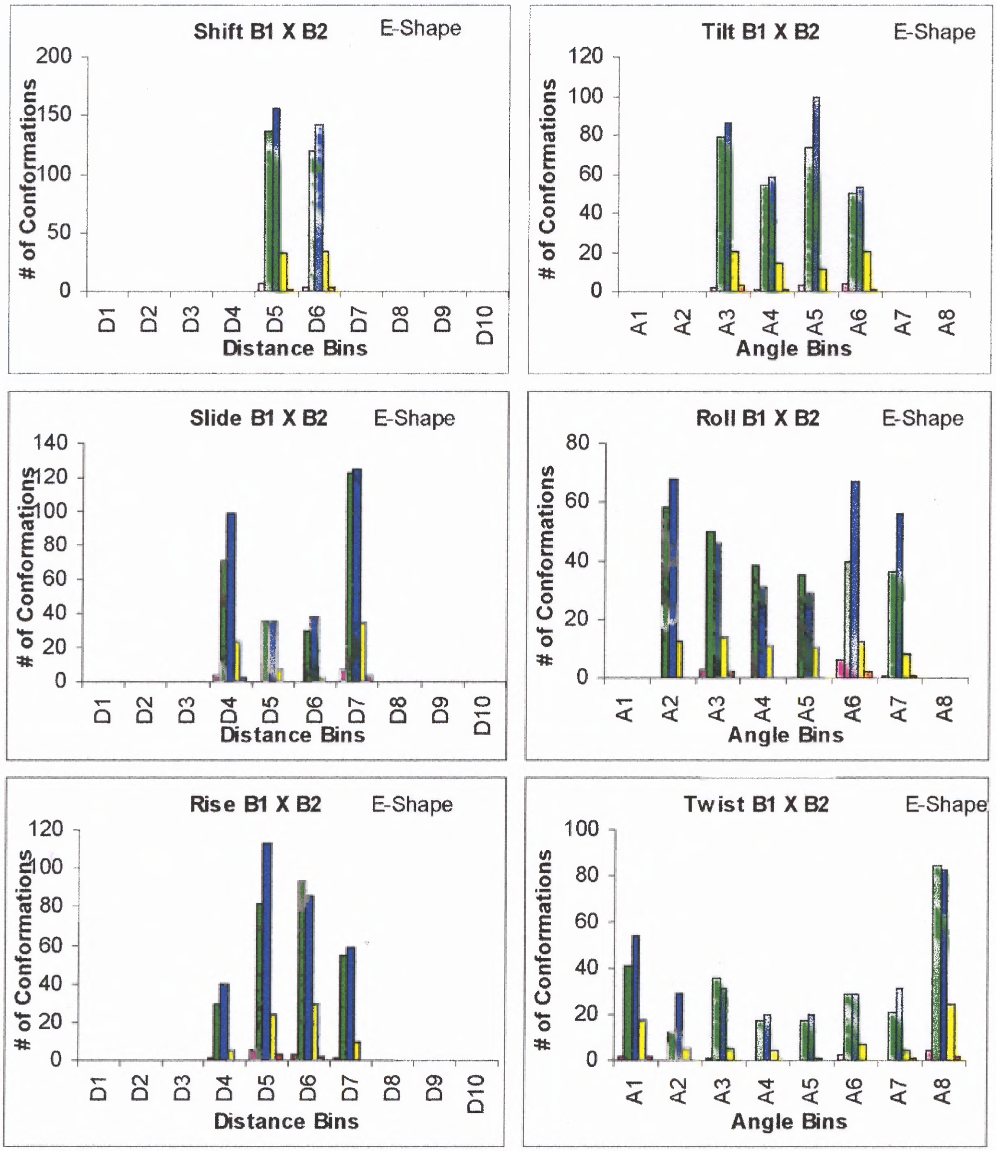
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	🟥 0-4 kcal/mol
D2: -12≤D<-9	A2: -135≤A<-90	🟩 4-8 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟦 8-12 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟨 12-16 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	▨ 16-20 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure 4.9 Histograms of displacement and rotational parameters for B1 X B2 in I-Shaped conformations



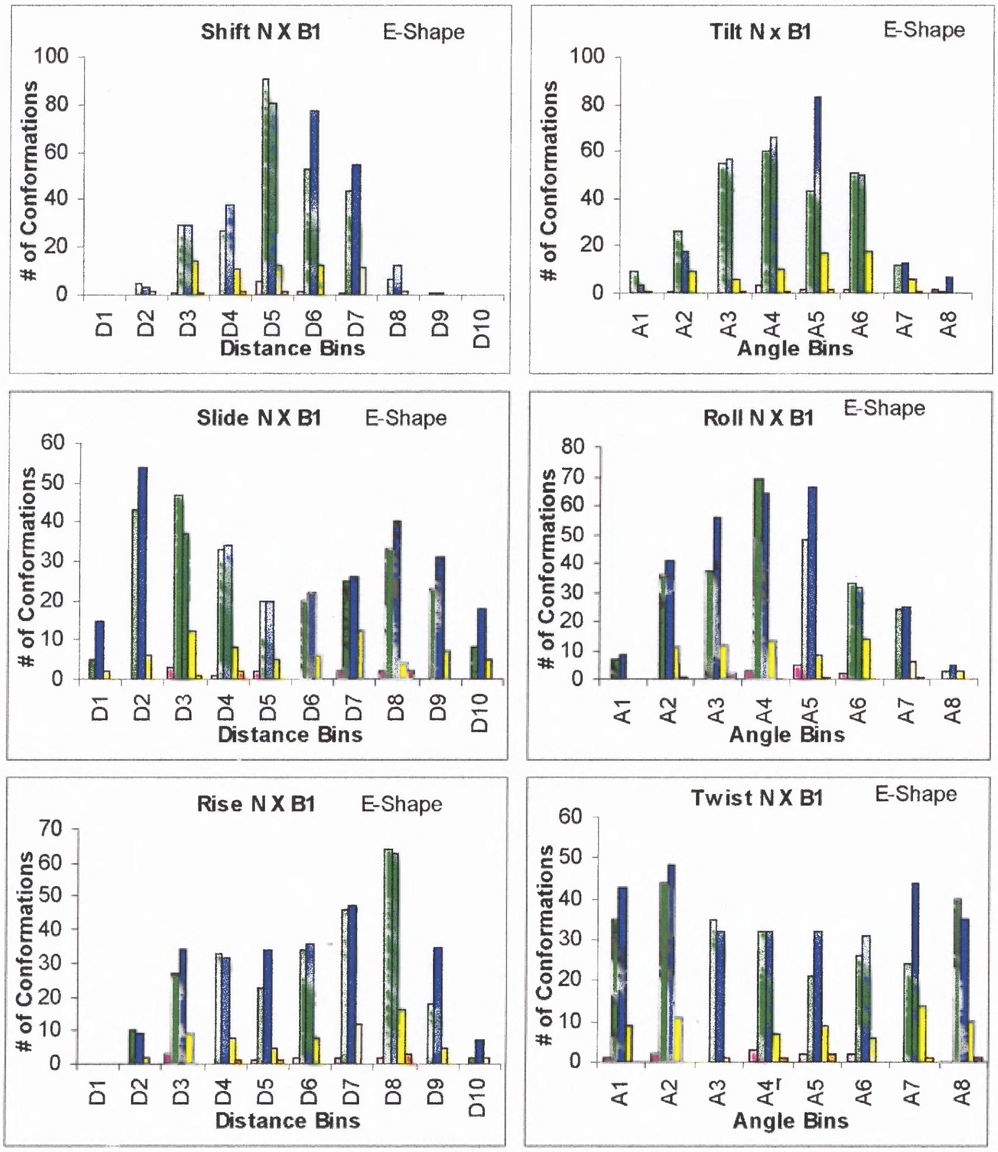
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	▨ 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	▨ 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	▨ 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	▨ 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	▨ 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure 4.10 Histograms of displacement and rotational parameters for B1 X B2 in V-Shaped conformations



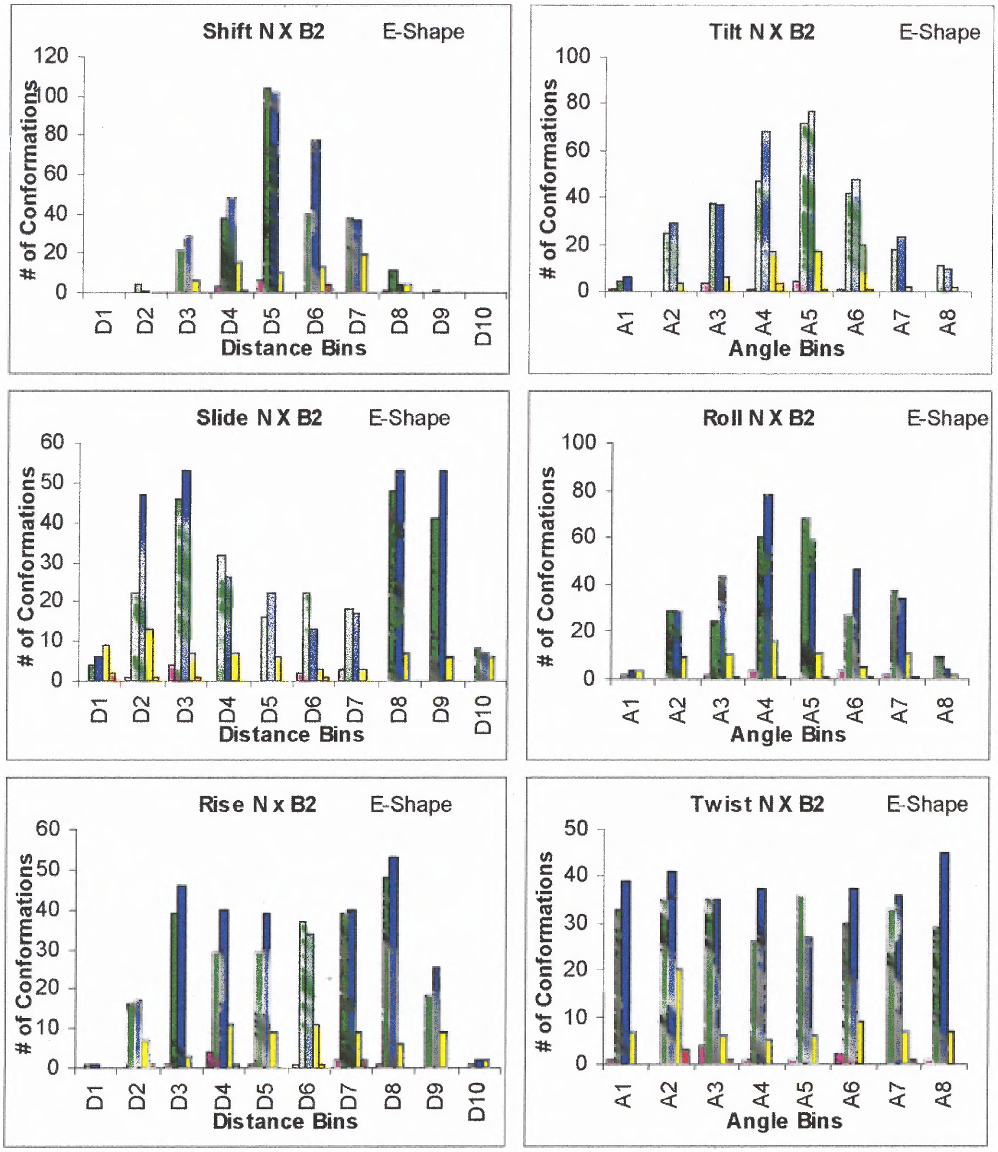
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	▨ 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	▨ 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	■ 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	□ 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	▨ 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure 4.11 Histograms of displacement and rotational parameters for B1 X B2 in E-Shaped conformations



Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	🟪 0-4 kcal/mol
D2: -12≤D<-9	A2: -135≤A<-90	🟩 4-8 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟦 8-12 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟨 12-16 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟥 16-20 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure 4.12 Histograms of displacement and rotational parameters for N X B1 in E-Shaped conformations



Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	▨ 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	▨ 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	■ 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	☐ 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	■ 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure 4.13 Histograms of displacement and rotational parameters for N X B2 in E-Shaped conformations

APPENDIX A

SOFTWARE AND SOURCE DATA INFORMATION

Software used for Planes program: MATLAB 6.0

All work performed in this research is located in Dr. Carol Venanzi's research area on the AFS system. Paths to research directory and other data files are specified below.

Research directory: /afs/cad/research/chem/venanzi/9
Planes Program: /afs/cad/research/chem/venanzi/9/planes
Conformation mol2 files: /afs/cad/research/chem/venanzi/9/planes/rs2files
Standards: /afs/cad/research/chem/venanzi/9/planes/standards
Output file of the Planes Program:
/afs/cad/research/chem/venanzi/9/planes/SprdShtNapLowerData.OUT
3D Scatter Plots: /afs/cad/research/chem/venanzi/9/scatters/3dplots
Colorcoded 3D scatter plots:
/afs/cad/research/chem/venanzi/9/scatters/colorcoded
Histograms: /afs/cad/research/chem/venanzi/9/Histograms

APPENDIX B

SYBYL MACRO TO DEFINE PLANES AND CENTROID

```
###################################################################
# Macro Name: definePlanes.spl                                    #
# Author: Milind Misra (New Jersey Institute of Technology)       #
# Date Created: Sep 12, 2003                                      #
# Modified by: Deepa Pai, Feb 10, 2004                            #
# Functions:                                                      #
#    1.  Center the molecule on N1 atom                           #
#    2.  Define planes in all dm324 molecules in a database       #
#    3.  Define static sets for planes defined in 2               #
#    4.  Define normal and align central plane in xy-orientation  #
# Remark: Align database first (on atoms 1, 2, 4, and 5)          #
###################################################################

# UIMS DEFINE MACRO modifyType SYBYLBASIC YES

default m1 >$NULLDEV
setvar dbname %promptif("$1" DATABASE_FILE " " "Database of molecules"
"Name of database")
if $STATUS
   return
endif

database open $dbname update

if %eq(%count(%database(*)) 0)
   %dialog_message(ERROR "No molecules in database" No molecules")
>$NULLDEV
   return
endif

for i in %range(1 %count(%database(*)))
   zap m1
   database get "%arg($i %database(*))" m1
   CENTER M1(1)
   DEFINE STATIC_SET ATOM M1((((1)+2)+4)+5) C_SET """"
   DEFINE STATIC_SET ATOM M1((((((11)+12)+13)+14)+15)+16) P1_SET """"
   DEFINE STATIC_SET ATOM M1((((((17)+18)+19)+20)+21)+22) P2_SET """"
   DEFINE STATIC_SET ATOM
M1((((((((((24)+25)+26)+27)+28)+29)+30)+31)+32)+33) N_SET """"
   DEFINE PLANE M1({C_SET}) C_PLANE """"
   DEFINE PLANE M1({P1_SET}) P1_PLANE """"
   DEFINE PLANE M1({P2_SET}) P2_PLANE """"
   DEFINE PLANE M1({N_SET}) N_PLANE """"
   DEFINE NORMAL M1(1) C_PLANE C_NORMAL """"
   ORIENT USER_VIEW M1(1) M1(4) M1(84)
   EVALUATE PLANE M1 C_PLANE
   EVALUATE PLANE M1 P1_PLANE
   EVALUATE PLANE M1 P2_PLANE
   EVALUATE PLANE M1 N_PLANE
   EVALUATE NORMAL M1 C_NORMAL
   DEFINE CENTROID M1({C_SET}) C_CENT """"
```

SYBYL MACRO TO DEFINE PLANES AND CENTROID

(continued)

```
   DEFINE CENTROID M1({P1_SET}) P1_CENT """"
   DEFINE CENTROID M1({P2_SET}) P2_CENT """"
   DEFINE CENTROID M1(((((24)+25)+26)+27)+32)+33) N_CENT """"
   DATABASE ADD "M1" REPLACE
endfor

zap m1
database close
# Done!
```

APPENDIX C

STANDARD BENZENE

```
#     File: /afs/…/venanzi/9/planes/standards/benzene_std.MOL
#     Creating user name:     dp36 (Deepa Pai)
#     Creation time:          Tue Feb 17 16:13:24 2004

@<TRIPOS>MOLECULE
benzene
   12    12     1     1     0
SMALL
NO_CHARGES

@<TRIPOS>ATOM
      1 C1     0.6991    1.2108    0.0000 C.ar       1 BENZENE     0.0000
      2 C2    -0.6991    1.2108    0.0000 C.ar       1 BENZENE     0.0000
      3 C3    -1.3981    0.0000    0.0000 C.ar       1 BENZENE     0.0000
      4 C4    -0.6991   -1.2108    0.0000 C.ar       1 BENZENE     0.0000
      5 C5     0.6991   -1.2108    0.0000 C.ar       1 BENZENE     0.0000
      6 C6     1.3981    0.0000    0.0000 C.ar       1 BENZENE     0.0000
      7 H1     2.4838    0.0000    0.0000 H          1 BENZENE     0.0000
      8 H2     1.2419    2.1510    0.0000 H          1 BENZENE     0.0000
      9 H3    -1.2419    2.1510    0.0000 H          1 BENZENE     0.0000
     10 H4    -2.4838    0.0000    0.0000 H          1 BENZENE     0.0000
     11 H5    -1.2419   -2.1510    0.0000 H          1 BENZENE     0.0000
     12 H6     1.2419   -2.1510    0.0000 H          1 BENZENE     0.0000
@<TRIPOS>BOND
     1    1    2 ar
     2    1    6 ar
     3    2    3 ar
     4    3    4 ar
     5    4    5 ar
     6    5    6 ar
     7    6    7 1
     8    1    8 1
     9    2    9 1
    10    3   10 1
    11    4   11 1
    12    5   12 1
@<TRIPOS>SUBSTRUCTURE
     1 BENZENE     1 PERM              0 ****  ****    0 ROOT
@<TRIPOS>NORMAL
@<TRIPOS>FF_PBC
FORCE_FIELD_SETUP_FEATURE Force Field Setup information
v1.0  0 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 NONE 2 0
0 0 0 0 0 0 0 0 0 0 0
```

STANDARD PIPERAZINE

```
#     File: /afs/…/venanzi/9/planes/standards/piperazine_std.MOL
#     Creating user name:     dp36 (Deepa Pai)
#     Creation time:          Wed Feb 18 11:38:54 2004

@<TRIPOS>MOLECULE
piperazine
   21    21     2     3     3
SMALL
GAST_HUCK

@<TRIPOS>ATOM
    1 N2          0.7159    1.2360   -0.0012 N.3 1 PIPERAZINE   -0.3136
    2 C2         -0.7635    1.2363    0.0004 C.3 1 PIPERAZINE    0.0085
    3 C1         -1.2809   -0.0524   -0.6841 C.3 1 PIPERAZINE    0.0085
    4 N1         -0.7159   -1.2361    0.0004 N.3 1 PIPERAZINE   -0.3136
    5 C4          0.7635   -1.2361    0.0004 C.3 1 PIPERAZINE    0.0085
    6 C3          1.2804    0.0530    0.6846 C.3 1 PIPERAZINE    0.0085
    7 H1          1.0974    2.1503    0.4404 H   1 PIPERAZINE    0.1220
    8 H2         -1.1467    1.2782    1.0340 H   1 PIPERAZINE    0.0436
    9 H3         -1.1367    2.1185   -0.5462 H   1 PIPERAZINE    0.0436
   10 H4         -2.3820   -0.0809   -0.6299 H   1 PIPERAZINE    0.0436
   11 H5         -0.9836   -0.0445   -1.7465 H   1 PIPERAZINE    0.0436
   12 H6         -1.0966   -2.1501   -0.4431 H   1 PIPERAZINE    0.1220
   13 H7          1.1361   -2.1179    0.5482 H   1 PIPERAZINE    0.0436
   14 H8          1.1480   -1.2786   -1.0328 H   1 PIPERAZINE    0.0436
   15 H9          0.9820    0.0458    1.7467 H   1 PIPERAZINE    0.0436
   16 H10         2.3816    0.0815    0.6317 H   1 PIPERAZINE    0.0436
   17 C_CENT      0.0000    0.0000    0.0000 Du  1 PIPERAZINE    0.0000
   18 C_PLANE1    0.7566   -1.6793    0.0002 Du     2 <2>       0.0000
   19 C_PLANE2    0.7997    1.2106   -0.0008 Du     2 <2>       0.0000
   20 C_PLANE3   -0.7571    1.6789   -0.0002 Du     2 <2>       0.0000
   21 C_PLANE4   -0.8002   -1.2109    0.0008 Du     2 <2>       0.0000
@<TRIPOS>BOND
     1    4    3 1
     2    3    2 1
     3    2    1 1
     4    1    6 1
     5    6    5 1
     6    4    5 1
     7    1    7 1
     8    2    8 1
     9    2    9 1
    10    3   10 1
    11    3   11 1
    12    4   12 1
    13    5   13 1
    14    5   14 1
    15    6   15 1
    16    6   16 1
    17    1   17 du
    18   18   19 du
    19   19   20 du
    20   20   21 du
    21   21   18 du
```

STANDARD PIPERAZINE

(continued)

```
@<TRIPOS>SUBSTRUCTURE
     1 PIPERAZINE     4 PERM              0 ****  ****    0 ROOT
     2 ****        18 TEMP             0 ****  ****    0 ROOT
@<TRIPOS>SET
C_SET           STATIC     ATOMS    <user>   **** ""
4 1 2 4 5
cent$C_CENT     STATIC     ATOMS    CENTROID SYSTEM|DELETE_EMPTY
4 1 2 4 5
plane$C_PLANE   STATIC     ATOMS    LSPLANE  SYSTEM|DELETE_EMPTY
5 1 2 4 5 17
@<TRIPOS>LSPLANE
C_PLANE        ""
18 19 20 21 3 7.451940e-01 -5.376306e-01 -3.945114e-01 -1.558428e-17
@<TRIPOS>NORMAL
@<TRIPOS>CENTROID
C_CENT         ""
17 2
@<TRIPOS>FF_PBC
FORCE_FIELD_SETUP_FEATURE Force Field Setup information
v1.0  0 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 NONE 0 0
0 0 1 0 0 0 0 0 0 0 0
```

SAMPLE INPUT CONFORMATION FILE

```
#  File: /afs/cad/research/chem./venanzi/9/planes/rs2_files/rs2_00001
#  Creating user name:  mxm0528
#  Creation time:       Tue Oct 14 16:16:18 2003
#  Modifying user name: mxm0528
#  Modification time:   Tue Oct 14 16:16:18 2003

@<TRIPOS>MOLECULE
rs2_00001
   84    88     8    17    13
SMALL
GAST_HUCK
****
23.275
@<TRIPOS>ATOM
  1 N2  0.000000   0.000000    0.000000 N.3    1 PIPERAZINE   -0.290200
  2 C2  0.714843  -0.003850    1.298071 C.3    1 PIPERAZINE    0.039400
  3 C1  2.086784  -0.684186    1.082004 C.3    1 PIPERAZINE   -0.017300
  4 N1  2.854432   0.000000    0.000000 N.4    1 PIPERAZINE    0.249100
  5 C4  2.089522  -0.003849   -1.281048 C.3    1 PIPERAZINE   -0.017300
  6 C3  0.714951   0.686437   -1.104307 C.3    1 PIPERAZINE    0.039400
  7 C6 -1.444721   0.361772    0.075908 C.3    1 PIPERAZINE    0.023200
  8 C1 -2.287160  -0.617303    0.935425 C.3    4 BENZENE     0.065700
  9 O2 -3.690205  -0.357469    0.716778 O.3    4 BENZENE    -0.352000
 10 C3 -4.565473  -1.240505    1.454884 C.3    4 BENZENE     0.122300
 11 C3 -6.047054  -0.932022    1.197453 C.ar   4 BENZENE    -0.007700
 12 C4 -6.437930   0.174170    0.436833 C.ar   4 BENZENE    -0.057600
```

INPUT CONFORMATION FILE
(continued)

```
13 C5 -7.795883   0.432920    0.225268 C.ar   4 BENZENE    -0.034500
14 C6 -8.763111  -0.414613    0.774459 C.ar   4 BENZENE     0.125100
15 C1 -8.372307  -1.520926    1.535063 C.ar   4 BENZENE    -0.034500
16 C2 -7.014261  -1.779645    1.746605 C.ar   4 BENZENE    -0.057600
17 C2 -4.343911   -1.082957   2.966132 C.ar   3 BENZENE    -0.007700
18 C1 -4.034328   -2.182893   3.772265 C.ar   3 BENZENE    -0.057600
19 C6 -3.829147   -2.008717   5.144327 C.ar   3 BENZENE    -0.034500
20 C5 -3.933448   -0.734269   5.710348 C.ar   3 BENZENE     0.125100
21 C4 -4.242880    0.365584   4.904236 C.ar   3 BENZENE    -0.034500
22 C3 -4.448177    0.191210   3.532258 C.ar   3 BENZENE    -0.057600
23 C1  4.207467   -0.619517  -0.165497 C.3    1 PIPERAZINE    0.002300
24 C3  5.394442    0.244891  -0.647223 C.ar   2 NAPHTHALENE    0.002200
25 C2  6.660047   -0.321364  -0.476348 C.ar   2 NAPHTHALENE   -0.051600
26 C1  7.800090    0.392758  -0.854724 C.ar   2 NAPHTHALENE   -0.056500
27 C9  7.681286    1.674766  -1.405179 C.ar   2 NAPHTHALENE   -0.022200
28 C8  8.821192    2.393728  -1.785695 C.ar   2 NAPHTHALENE   -0.056800
29 C7  8.698227    3.673027  -2.334803 C.ar   2 NAPHTHALENE   -0.063400
30 C6  7.432734    4.239238  -2.505750 C.ar   2 NAPHTHALENE   -0.063700
31 C5  6.292517    3.525075  -2.127206 C.ar   2 NAPHTHALENE   -0.057800
32 C10 6.411322    2.243066  -1.576750 C.ar   2 NAPHTHALENE   -0.025700
33 C4  5.271510    1.524139  -1.196251 C.ar   2 NAPHTHALENE   -0.049900
34 H4  0.830365    1.028341   1.668945 H      1 PIPERAZINE    0.047900
35 H5  0.193253   -0.588948   2.069772 H      1 PIPERAZINE    0.047900
36 H2  2.615602   -0.664087   2.052734 H      1 PIPERAZINE    0.082700
37 H3  1.914196   -1.740919   0.812739 H      1 PIPERAZINE    0.082700
38 H11 2.996329    1.017744   0.352094 H      1 PIPERAZINE    0.204900
39 H9  2.639576    0.487880  -2.099905 H      1 PIPERAZINE    0.082700
40 H10 1.888247   -1.041096  -1.600442 H      1 PIPERAZINE    0.082700
41 H7  0.840446    1.756114  -0.867585 H      1 PIPERAZINE    0.047900
42 H8  0.160855    0.608434  -2.055471 H      1 PIPERAZINE    0.047900
43 H15   -1.569041 1.388184   0.458112 H      1 PIPERAZINE    0.044900
44 H14   -1.890713 0.320636  -0.931300 H      1 PIPERAZINE    0.044900
45 H3    -2.059588 -0.471726    2.001227 H    4 BENZENE     0.057700
46 H2   -2.030587  -1.648593    0.642807 H    4 BENZENE     0.057700
47 H5   -4.392574  -2.276529    1.120808 H    4 BENZENE     0.090600
48 H4   -5.697904   0.840573    0.006504 H    4 BENZENE     0.055200
49 H5   -8.105980   1.289637   -0.363857 H    4 BENZENE     0.058700
50 F6   -10.055224 -0.168613    0.573975 F    4 BENZENE    -0.189100
51 H1   -9.127423   -2.174689    1.959172 H   4 BENZENE     0.058700
52 H2   -6.717707   -2.639495    2.339372 H   4 BENZENE     0.055200
53 H1   -3.951247   -3.176190    3.343041 H   3 BENZENE     0.055200
54 H6   -3.588631   -2.857798    5.775769 H   3 BENZENE     0.058700
55 F5   -3.735818   -0.568015    7.015544 F   3 BENZENE    -0.189100
56 H4   -4.322221    1.351927    5.349267 H   3 BENZENE     0.058700
57 H3   -4.687502    1.046182    2.907450 H   3 BENZENE     0.055200
58 H13   4.552605   -0.993588    0.813530 H   1 PIPERAZINE    0.098200
59 H12   4.158552   -1.497938   -0.830383 H   1 PIPERAZINE    0.098200
60 H2    6.779461   -1.316157   -0.056562 H   2 NAPHTHALENE    0.058800
61 H1    8.778703   -0.057767   -0.717284 H   2 NAPHTHALENE    0.062000
62 H8    9.810971    1.965828   -1.658506 H   2 NAPHTHALENE    0.061900
63 H7    9.585064    4.226733   -2.628441 H   2 NAPHTHALENE    0.059800
64 H6    7.336158    5.233210   -2.932392 H   2 NAPHTHALENE    0.061100
65 H5    5.316614    3.979864   -2.267055 H   2 NAPHTHALENE    0.060300
```

INPUT CONFORMATION FILE

(continued)

```
66 H4   4.300792    1.980346   -1.336747 H    2 NAPHTHALENE    0.053300
67 C_PLANE1    -0.500000    -0.001962    -1.781048 Du    5 <5>      0.000000
68 C_PLANE2    -0.500000    -0.001962     1.798071 Du    5 <5>      0.000000
69 C_PLANE3     3.354432    -0.001887     1.798071 Du    5 <5>      0.000000
70 C_PLANE4     3.354432    -0.001887    -1.781048 Du    5 <5>      0.000000
71 P1_PLANE1    -9.263111    -2.279645     2.026002 Du   6 <6>      0.000000
72 P1_PLANE2    -9.263111     0.932920    -0.152757 Du   6 <6>      0.000000
73 P1_PLANE3    -5.547054     0.932920    -0.054095 Du   6 <6>      0.000000
74 P1_PLANE4    -5.547054    -2.279645     2.124665 Du   6 <6>      0.000000
75 P2_PLANE1    -4.177248    -2.682893     2.466132 Du   7 <7>     0.000000
76 P2_PLANE2    -3.542439    -2.682893     6.210348 Du   7 <7>     0.000000
77 P2_PLANE3    -4.100062     0.865584     6.210348 Du   7 <7>     0.000000
78 P2_PLANE4    -4.734871     0.865584     2.466132 Du   7 <7>     0.000000
79 N_PLANE1     4.771510    -0.821364    -0.146434 Du    8 <8>     0.000000
80 N_PLANE2     4.771510     4.739238    -2.564741 Du    8 <8>     0.000000
81 N_PLANE3     9.321192     4.739238    -2.835557 Du    8 <8>     0.000000
82 N_PLANE4     9.321192    -0.821364    -0.417251 Du    8 <8>     0.000000
83 C_NORMAL1  0.000019   -1.000000  0.000000 Du 1 PIPERAZINE  0.000000
84 C_NORMAL2 -0.000019    1.000000  0.000000 Du 1 PIPERAZINE  0.000000
@<TRIPOS>BOND
     1    4    3 1
     2    3    2 1
     3    2    1 1
     4    1    6 1
     5    6    5 1
     6    4    5 1
     7    4   23 1
     8    3   36 1
     9    3   37 1
    10    2   34 1
    11    2   35 1
    12    1    7 1
    13    6   41 1
    14    6   42 1
    15    5   39 1
    16    5   40 1
    17    4   38 1
    18   31   30 ar
    19   31   32 ar
    20   31   65 1
    21   32   33 ar
    22   32   27 ar
    23   27   26 ar
    24   27   28 ar
    25   28   62 1
    26   28   29 ar
    27   29   63 1
    28   29   30 ar
    29   30   64 1
    30   26   61 1
    31   26   25 ar
    32   25   60 1
    33   25   24 ar
```

INPUT CONFORMATION FILE

(continued)

```
34   24   33 ar
35   33   66 1
36   23   24 1
37   23   59 1
38   23   58 1
39   18   17 ar
40   18   19 ar
41   18   53 1
42   19   54 1
43   19   20 ar
44   20   55 1
45   20   21 ar
46   21   56 1
47   21   22 ar
48   22   57 1
49   22   17 ar
50   15   51 1
51   15   14 ar
52   15   16 ar
53   16   52 1
54   16   11 ar
55   11   10 1
56   11   12 ar
57   12   48 1
58   12   13 ar
59   13   49 1
60   13   14 ar
61   14   50 1
62   10   47 1
63   10    9 1
64    9    8 1
65    8   45 1
66    8   46 1
67   10   17 1
68    8    7 1
69    7   44 1
70    7   43 1
71   67   68 du
72   68   69 du
73   69   70 du
74   70   67 du
75   71   72 du
76   72   73 du
77   73   74 du
78   74   71 du
79   75   76 du
80   76   77 du
81   77   78 du
82   78   75 du
83   79   80 du
84   80   81 du
85   81   82 du
86   82   79 du
```

INPUT CONFORMATION FILE

(continued)

```
    87    1   83 du
    88    1   84 du
@<TRIPOS>SUBSTRUCTURE
     1 PIPERAZINE     4 PERM              0 ****  ****    2
     2 NAPHTHALENE    31 PERM              0 ****  ****    1
     3 BENZENE    18 PERM             0 ****  ****    1
     4 BENZENE    15 PERM             0 ****  ****    2 ROOT
     5 ****       67 TEMP             0 ****  ****    0 ROOT
     6 ****       71 TEMP             0 ****  ****    0 ROOT
     7 ****       75 TEMP             0 ****  ****    0 ROOT
     8 ****       79 TEMP             0 ****  ****    0 ROOT
@<TRIPOS>SET
NAPHTHALENE_AGGREGATE STATIC     ATOMS    AGGREGATE SYSTEM|ACTIVE_AGG
""
10 26 25 24 33 31 30 29 28 27 32
BENZENE_1_AGGREGATE STATIC     ATOMS    AGGREGATE SYSTEM|ACTIVE_AGG ""
6 15 16 11 12 13 14
BENZENE_2_AGGREGATE STATIC     ATOMS    AGGREGATE SYSTEM|ACTIVE_AGG ""
6 18 17 22 21 20 19
PIPERAZINE_RING_AGGREGATE STATIC     ATOMS    AGGREGATE
SYSTEM|ACTIVE_AGG ""
6 1 2 3 4 5 6
SEARCH_BONDS    DYNAMIC    BONDS    <user>   **** ""
4=23,1=7,23=24,11=10,10=9,9=8,10=17,8=7
C_SET           STATIC     ATOMS    <user>   **** ""
4 1 2 4 5
P1_SET          STATIC     ATOMS    <user>   **** ""
6 11 12 13 14 15 16
P2_SET          STATIC     ATOMS    <user>   **** ""
6 17 18 19 20 21 22
N_SET           STATIC     ATOMS    <user>   **** ""
10 24 25 26 27 28 29 30 31 32 33
PLANE$C_PLANE   STATIC     ATOMS    LSPLANE  SYSTEM|DELETE_EMPTY
4 1 2 4 5
PLANE$P1_PLANE  STATIC     ATOMS    LSPLANE  SYSTEM|DELETE_EMPTY
6 11 12 13 14 15 16
PLANE$P2_PLANE  STATIC     ATOMS    LSPLANE  SYSTEM|DELETE_EMPTY
6 17 18 19 20 21 22
PLANE$N_PLANE   STATIC     ATOMS    LSPLANE  SYSTEM|DELETE_EMPTY
10 24 25 26 27 28 29 30 31 32 33
@<TRIPOS>ROTATABLE_BOND
    7     6    36  1 1     0  1  30    0 359
   12     4    68  2 1     0  1  30    0 359
   67    63    49  3 1     0  1  30    0 359
@<TRIPOS>ANCHOR_ATOM
4
@<TRIPOS>LSPLANE
C_PLANE         ""
67 68 69 70 10 1.936551e-05 -1.000000e+00 8.399540e-08 1.952147e-03
P1_PLANE        ""
71 72 73 74 11 2.196824e-02 -5.611551e-01 -8.274191e-01 -6.006128e-01
P2_PLANE        ""
75 76 77 78 12 -9.743056e-01 -1.531065e-01 1.651876e-01 4.888059e+00
```

INPUT CONFORMATION FILE

(continued)

```
N_PLANE         ""
79 80 81 82 13 -5.450453e-02 -3.982239e-01 -9.156675e-01 2.011027e-01
@<TRIPOS>NORMAL
C_NORMAL        C_PLANE         ""
83 84 1 5
@<TRIPOS>FFCON_TORSION
1 2 3 4 1.500000e+01 5.600200e+01
2 3 4 5 1.500000e+01 -5.812200e+01
3 4 5 6 1.500000e+01 5.802300e+01
4 5 6 1 1.500000e+01 -5.579300e+01
2 1 6 5 1.500000e+01 5.794500e+01
6 1 2 3 1.500000e+01 -5.805400e+01
@<TRIPOS>FF_PBC
FORCE_FIELD_SETUP_FEATURE Force Field Setup information
v1.0  0 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 NONE 0 0
0 0 1 0 0 0 0 0 0 0 0
@<TRIPOS>SEARCH_OPTS
1 0 1 0 9.999000e+03 0 2.500000e-01 2.500000e-01 2.500000e-01 0 0
```

INPUT CONFIGURATION FILE

```
% FILE: /afs/cad/research/chem/venanzi/9/planes/conf/plane.par

% STANDARD specifies the name of the standard molecule
% Format is: STANDARD name fileName
%           where 'name' is the name of the molecule
%                 'fileName' the file that contains the definition of
%                  the Standard.
% Multiple STANDRARDs can be defined on separate lines
% BENZENE and PIPERAZINE standards are specified.
% BENZENE is the name of the molecule in the file
STANDARD    BENZENE
/afs/cad/research/chem/venanzi/9/planes/standards/benzene_std.MOL
STANDARD    PIPERAZINE
/afs/cad/research/chem/venanzi/9/planes/standards/piperazine_std.MOL

% FIT defines the rings that are to be fitted. You can use this to
% specify which standard molecules are to be fitted to an appropriate
% ring
% Format is: FIT std_molecule ring {atom ids}
%               where 'std_molecule' is the STANDARD molecule
%                     'ring' is the name of the ring system
%                     '{atom_ids} is a list of the atoms (in order) to
%                     correspond to the std_molecule atoms
% Multiple FITs can be defined.
% Conformations of DM324 have plane defined on all 10 ring atoms but
% only lower 6 atoms that form a phenyl are considered as a plane here.
% The Following statement says use BENZENE standard for NAPTHALENE
%  on atom numbers 24 25 26 27 32 33
FIT BENZENE  TO NAPHTHALENE 24 25 26 27 32 33
```

CONFIGURATION FILE
(continued)

```
% OVERRIDE specifies the plane in the MOL2 file to be overridden.
% You can specify multiple planes to override
% Format is: OVERRIDE PLANE name N atom{1} atom{2} .... atom{N}
%                     where 'name' is the name of the atom
%                           'N' is the number of atoms
%                           'atoms{n}' list of atoms which should be
%                                   equal to N
% Conformations of DM324 have plane defined on all 10 ring atoms but
% only lower 6 atoms that form a phenyl are considered as a plane here.
% The following statement specifies to override a N_PLANE(Napthalene)
% defined with 10 atoms on conformations with a new definition of only
% 6 atoms.
OVERRIDE PLANE N_PLANE 6 24 25 26 27 32 33

% Optional output file to store the results.
% If not specified, program stores output in file output.txt in current
% directory
OUTPUT
/afs/cad/research/chem/venanzi/9/planes/standards/NapLowerPlaneParamSpr
dsht.OUT

% Input files.
% Following are DM324 conformations.
FILE /afs/cad/research/chem/venanzi/9/planes/rs2_files/rs2_00001.mol2
FILE /afs/cad/research/chem/venanzi/9/planes/rs2_files/rs2_00002.mol2
FILE /afs/cad/research/chem/venanzi/9/planes/rs2_files/rs2_00003.mol2
…
…
FILE /afs/cad/research/chem/venanzi/9/planes/rs2_files/rs2_00728.mol2
```

APPENDIX D

3D SCATTER PLOTS COLOR CODED BY ENERGY

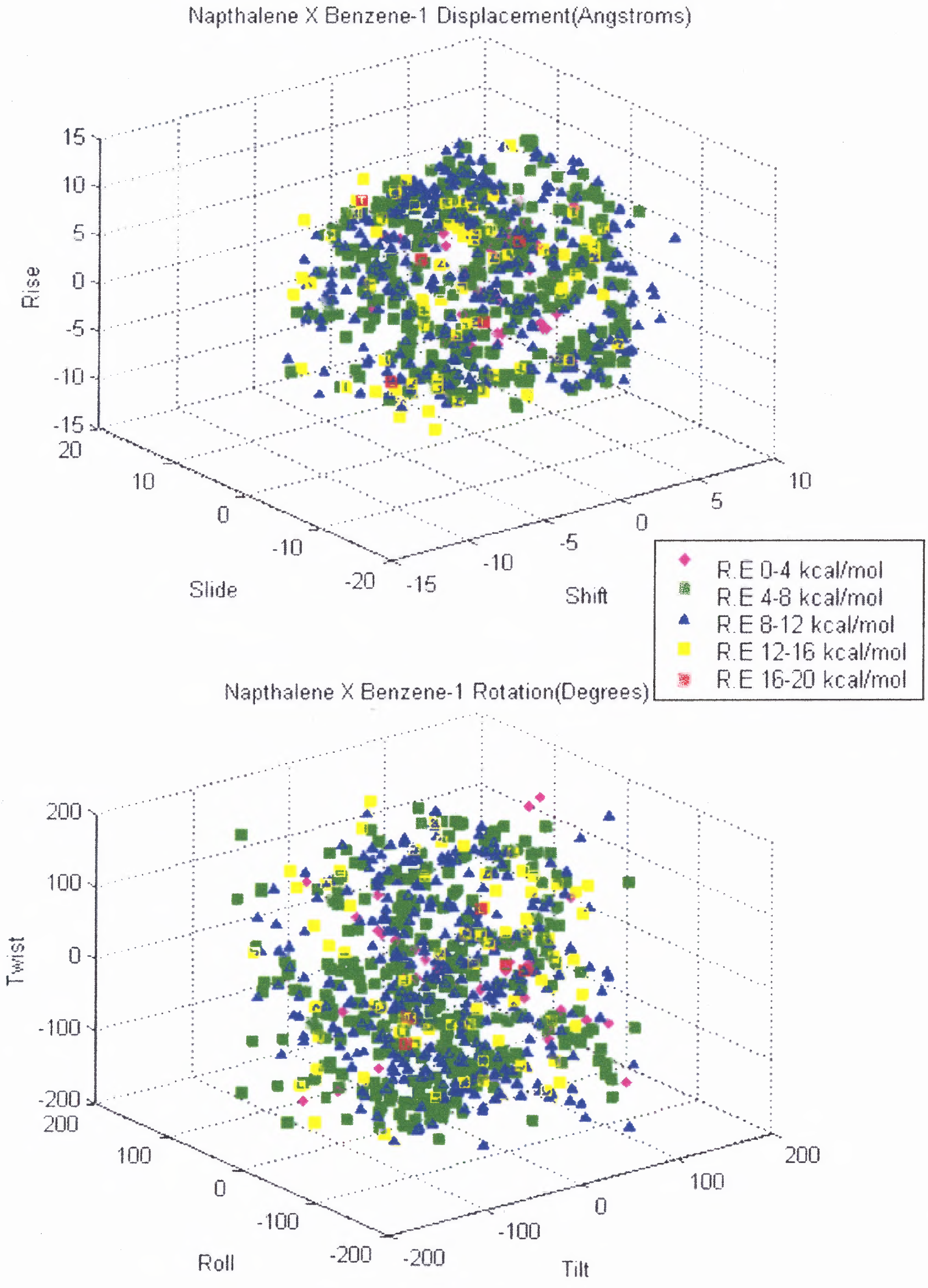


Figure D.1 3D scatter plots Color-coded by energy. Napthalene X Benzene-1.

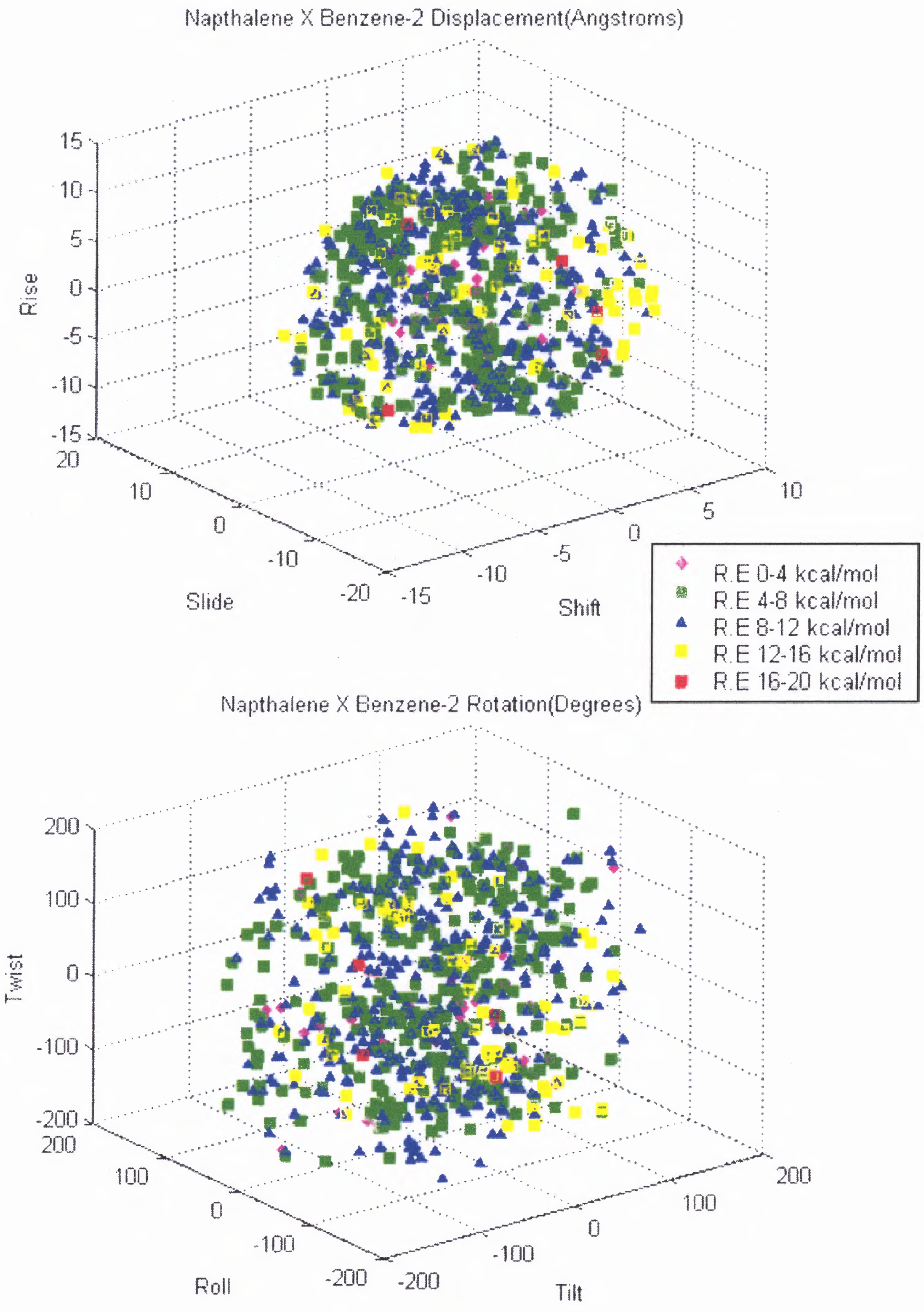


Figure D.2 3D scatter plots Color-coded by energy. Napthalene X Benzene-2

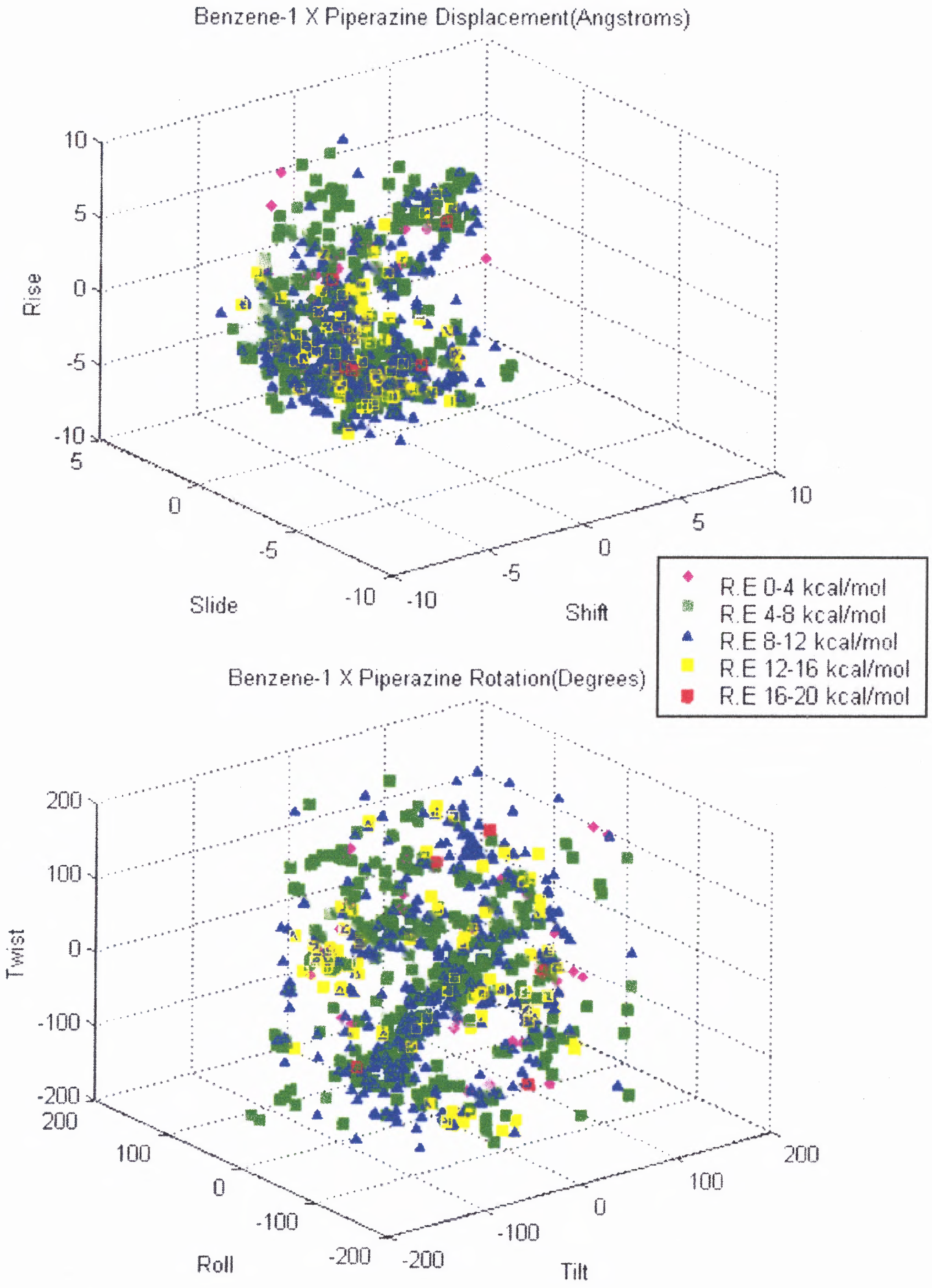


Figure D.3 3D scatter plots Color-coded by energy. Benzene-1 X Piperazine

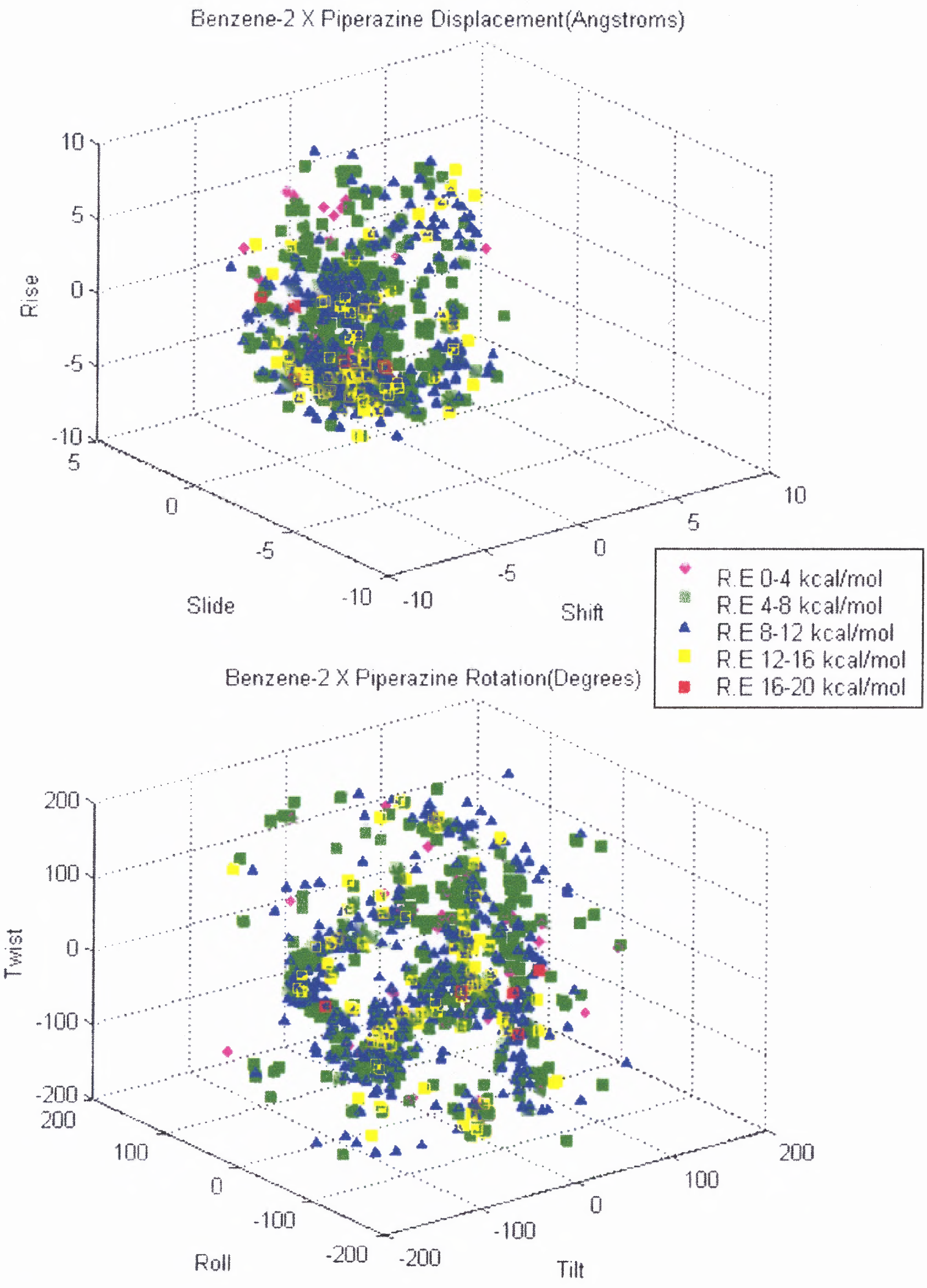


Figure D.4 3D scatter plots Color-coded by energy. Benzene-2 X Piperazine

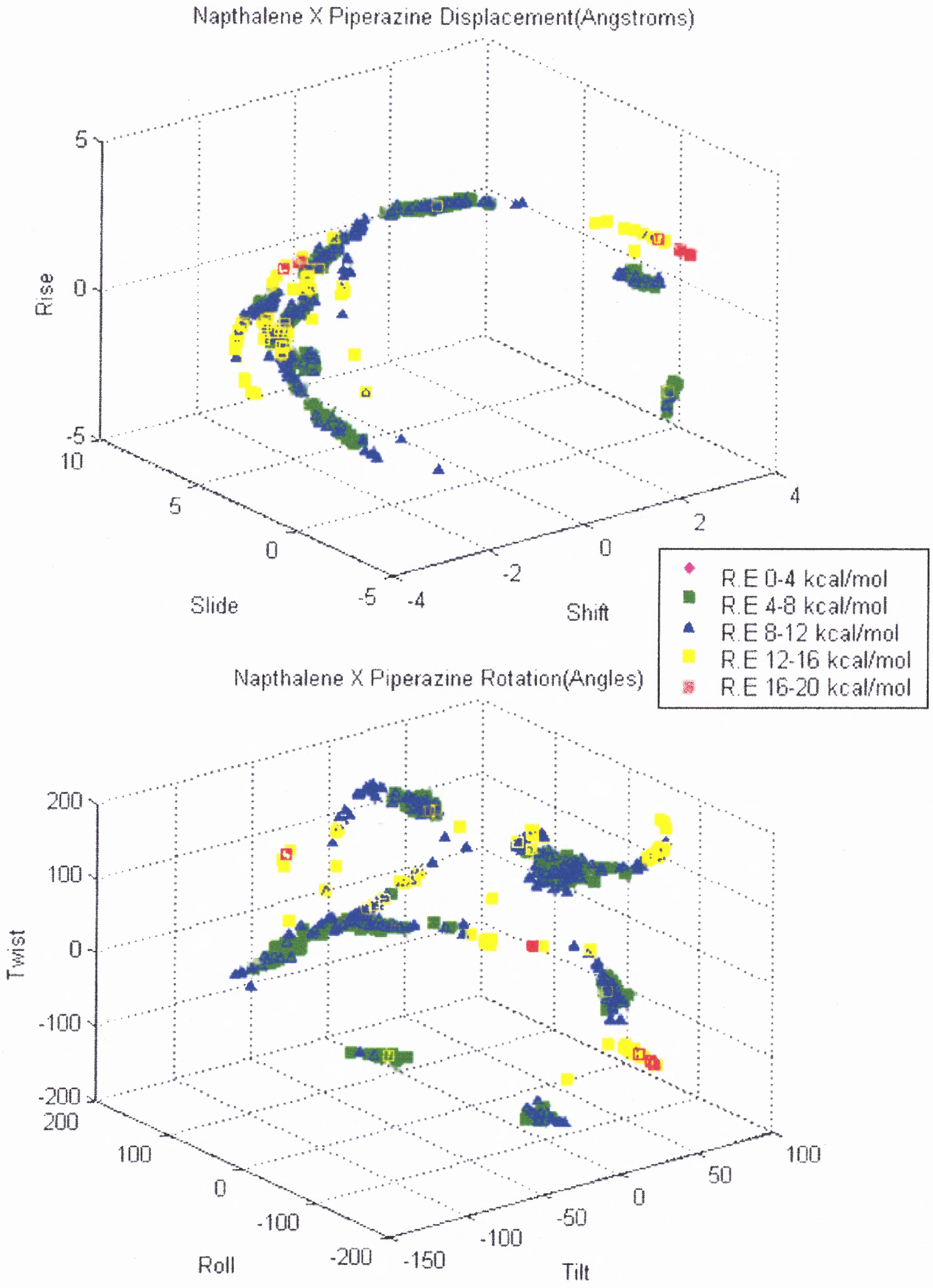


Figure D.5 3D scatter plots Color-coded by energy. Napthalene X Piperazine

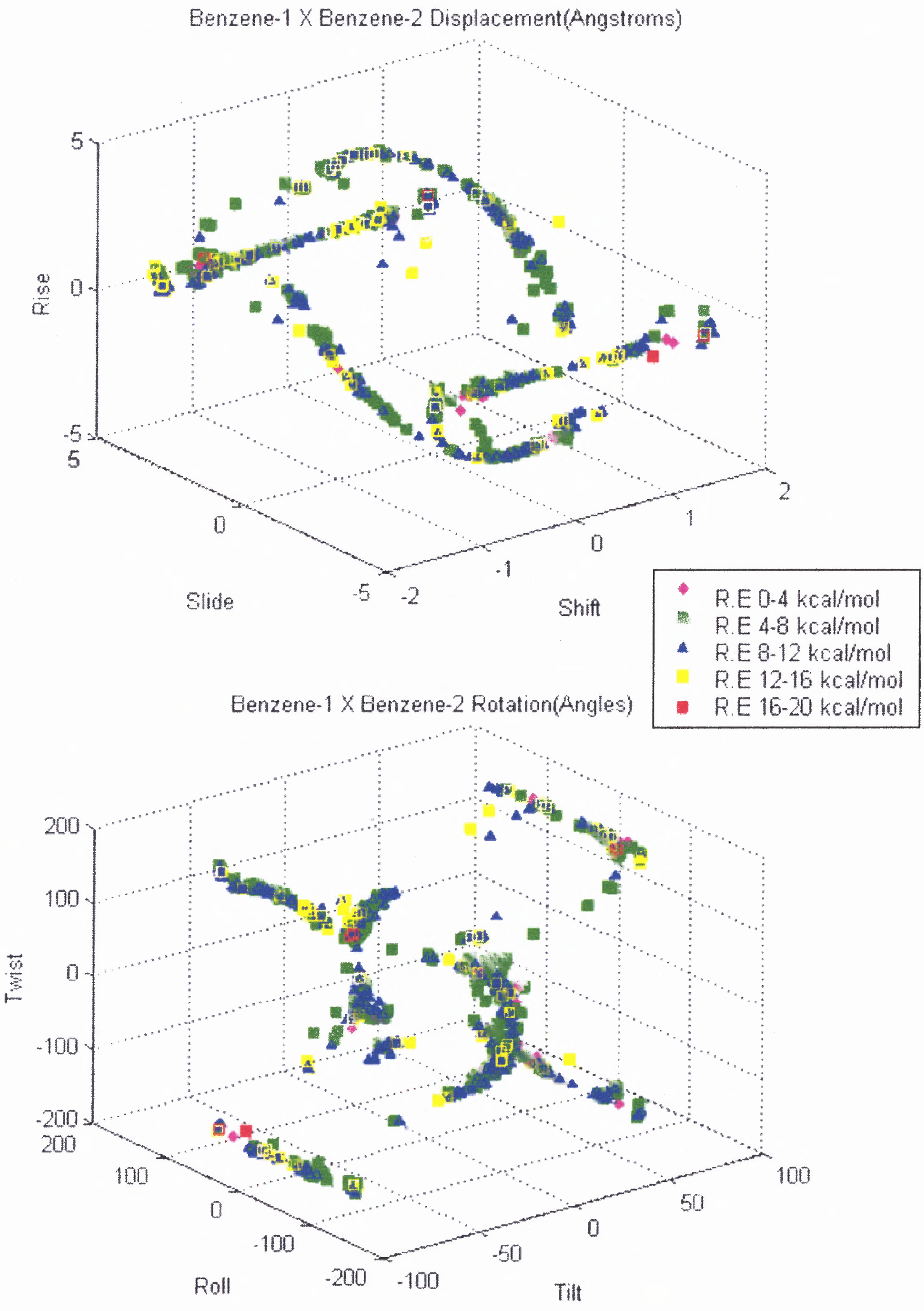


Figure D.6 3D scatter plots Color-coded by energy. Benzene-1 X Benzene-2

APPENDIX E
DISTRIBUTION TABLES FOR MOLECULAR SHAPES

Shift NXB1	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	2	7	6	0	0	0	15
	4-8	0	0	0	0	3	2	1	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	0	5	9	7	0	0	0	21
Slide NXB1												
	0-4	0	0	0	4	4	3	4	0	0	0	15
	4-8	0	0	2	1	1	0	2	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	2	5	5	3	6	0	0	0	21
Rise NXB1												
	0-4	0	0	2	4	2	2	5	0	0	0	15
	4-8	0	0	1	1	3	0	1	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	3	5	5	2	6	0	0	0	21

Tilt N X B1	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	1	2	0	3	3	3	0	3	15
	4-8	0	2	0	0	0	0	3	1	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	1	4	0	3	3	3	3	4	21
Roll N X B1										
	0-4	0	0	5	2	4	4	0	0	15
	4-8	0	0	2	1	1	2	0	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	0	7	3	5	6	0	0	21
Twist N X B1										
	0-4	1	3	2	2	4	0	0	3	15
	4-8	4	1	0	0	0	0	1	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	5	4	2	2	4	0	1	3	21

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.1 Distribution of displacement and rotational parameters for N X B1 in C-Shaped conformations

Shift NXB1	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	1	3	1	0	0	0	5
	4-8	0	0	0	0	4	3	3	0	0	0	10
	8-12	0	0	0	0	0	0	1	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	0	5	6	5	0	0	0	16
Slide NXB1												
	0-4	0	0	1	1	1	1	1	0	0	0	5
	4-8	0	0	1	1	2	2	3	1	0	0	10
	8-12	0	0	0	1	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	2	3	3	3	4	1	0	0	16
Rise NXB1												
	0-4	0	0	0	0	1	1	1	2	0	0	5
	4-8	0	0	2	0	1	3	3	1	0	0	10
	8-12	0	0	1	0	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	3	0	2	4	4	3	0	0	16

Tilt N X B1	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	1	0	2	0	0	1	1	5
	4-8	0	2	0	0	2	0	6	0	10
	8-12	0	0	0	0	0	1	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	3	0	2	2	1	7	1	16
Roll N X B1										
	0-4	0	0	1	3	0	1	0	0	5
	4-8	0	0	4	3	3	0	0	0	10
	8-12	0	0	0	0	1	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	0	5	6	4	1	0	0	16
Twist N X B1										
	0-4	0	1	0	1	2	0	1	0	5
	4-8	1	4	1	2	0	1	0	1	10
	8-12	0	1	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	1	6	1	3	2	1	1	1	16

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.2 Distribution of displacement and rotational parameters for N X B1 in I-Shaped conformations

Shift NXB1	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	4	0	0	0	0	0	4
	4-8	0	0	0	4	14	10	7	3	0	0	38
	8-12	0	0	0	2	2	0	4	1	0	0	9
	12-16	0	0	0	1	1	1	1	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	7	21	11	12	4	0	0	55
Slide NXB1												
	0-4	0	0	0	0	2	1	1	0	0	0	4
	4-8	0	0	6	10	5	5	7	4	1	0	38
	8-12	0	0	0	2	2	0	3	2	0	0	9
	12-16	0	0	0	4	0	0	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	6	16	9	6	11	6	1	0	55
Rise NXB1												
	0-4	0	0	0	2	0	1	1	0	0	0	4
	4-8	0	0	3	10	7	4	11	3	0	0	38
	8-12	0	0	1	1	1	1	2	3	0	0	9
	12-16	0	1	0	0	0	1	2	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	1	4	13	8	7	16	6	0	0	55

Tilt N X B1	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	0	0	0	3	1	0	4
	4-8	2	3	9	6	5	8	5	0	38
	8-12	0	0	0	2	1	6	0	0	9
	12-16	0	0	0	1	2	1	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	2	3	9	9	8	18	6	0	55
Roll N X B1										
	0-4	0	0	1	1	2	0	0	0	4
	4-8	0	2	8	11	9	6	1	1	38
	8-12	0	0	0	4	3	1	1	0	9
	12-16	0	0	0	1	1	2	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	2	9	17	15	9	2	1	55
Twist N X B1										
	0-4	0	0	1	0	3	0	0	0	4
	4-8	4	6	8	3	11	3	3	0	38
	8-12	0	0	1	3	2	0	1	2	9
	12-16	0	0	3	0	1	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	4	6	13	6	17	3	4	2	55

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.3 Distribution of displacement and rotational parameters for N X B1 in V-Shaped conformations

Shift NXB1	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	1	0	6	2	1	0	0	0	10
	4-8	0	5	29	27	91	53	44	7	1	0	257
	8-12	0	3	29	38	81	77	55	13	1	0	297
	12-16	0	2	14	11	13	13	12	2	0	0	67
	16-20	0	0	1	2	2	0	0	0	0	0	5
	Sum[e]	0	10	74	78	193	145	112	22	2	0	636
Slide NXB1												
	0-4	0	0	3	1	2	0	2	2	0	0	10
	4-8	5	43	47	33	20	20	25	33	23	8	257
	8-12	15	54	37	34	20	22	26	40	31	18	297
	12-16	2	6	12	8	5	6	12	4	7	5	67
	16-20	0	0	1	2	0	0	0	2	0	0	5
	Sum	22	103	100	78	47	48	65	81	61	31	636
Rise NXB1												
	0-4	0	0	3	0	1	2	2	2	0	0	10
	4-8	0	10	27	33	23	34	46	64	18	2	257
	8-12	0	9	34	32	34	36	47	63	35	7	297
	12-16	0	2	9	8	5	8	12	16	5	2	67
	16-20	0	0	0	1	1	0	0	3	0	0	5
	Sum	0	21	73	74	64	80	107	148	58	11	636

Tilt N X B1	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	1	0	3	2	2	0	2	10
	4-8	9	26	55	60	43	51	12	1	257
	8-12	3	18	57	66	83	50	13	7	297
	12-16	1	9	6	10	17	18	6	0	67
	16-20	0	0	1	1	2	0	1	0	5
	Sum[f]	13	54	119	140	147	121	32	10	636
RollN X B1										
	0-4	0	0	0	3	5	2	0	0	10
	4-8	7	36	37	69	48	33	24	3	257
	8-12	8	41	56	64	66	32	25	5	297
	12-16	0	11	12	13	8	14	6	3	67
	16-20	0	1	2	0	1	0	1	0	5
	Sum	15	89	107	149	128	81	56	11	636
Twist N X B1										
	0-4	1	2	0	3	2	2	0	0	10
	4-8	35	44	35	32	21	26	24	40	257
	8-12	43	48	32	32	32	31	44	35	297
	12-16	9	11	1	7	9	6	14	10	67
	16-20	0	0	0	1	2	0	1	1	5
	Sum	88	105	68	75	66	65	83	86	636

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.4 Distribution of displacement and rotational parameters for N X B1 in E-Shaped conformations

Shift NXB2	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	8	6	1	0	0	0	15
	4-8	0	0	0	0	0	5	1	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	0	8	11	2	0	0	0	21
Slide NXB2												
	0-4	0	0	1	3	3	2	5	1	0	0	15
	4-8	0	0	0	0	1	2	2	1	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	1	3	4	4	7	2	0	0	21
Rise NXB2												
	0-4	0	0	0	7	2	3	2	1	0	0	15
	4-8	0	0	0	2	0	0	4	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	9	2	3	6	1	0	0	21

Tilt N X B2	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	1	2	1	2	2	0	6	1	15
	4-8	0	0	0	4	0	0	2	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	1	2	1	6	2	0	8	1	21
Roll N X B2										
	0-4	0	0	4	4	3	3	1	0	15
	4-8	0	0	1	2	2	1	0	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	0	5	6	5	4	1	0	21
Twist N X B2										
	0-4	1	5	3	2	1	0	1	2	15
	4-8	0	0	0	3	0	1	2	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	1	5	3	5	1	1	3	2	21

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.5 Distribution of displacement and rotational parameters for N X B2 in C-Shaped conformations

Shift NXB2	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	0	4	1	0	0	0	5
	4-8	0	0	0	0	1	8	1	0	0	0	10
	8-12	0	0	0	0	1	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	0	2	12	2	0	0	0	16
Slide NXB2												
	0-4	0	0	0	1	2	0	2	0	0	0	5
	4-8	0	0	0	1	2	3	3	1	0	0	10
	8-12	0	0	0	1	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	3	4	3	5	1	0	0	16
Rise NXB2												
	0-4	0	0	0	0	1	2	1	1	0	0	5
	4-8	0	0	0	2	1	4	3	0	0	0	10
	8-12	0	0	0	0	0	1	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	2	2	7	4	1	0	0	16

Tilt N X B2	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	0	0	1	1	3	0	5
	4-8	0	0	1	1	3	0	4	1	10
	8-12	0	0	0	0	0	0	1	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	0	1	1	4	1	8	1	16
Roll N X B2										
	0-4	0	0	0	3	2	0	0	0	5
	4-8	0	0	0	1	6	2	1	0	10
	8-12	0	0	0	1	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	5	8	2	1	0	16
Twist N X B2										
	0-4	1	1	1	1	0	0	0	1	5
	4-8	1	0	2	0	0	6	0	1	10
	8-12	0	0	0	0	0	0	0	1	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	2	1	3	1	0	6	0	3	16

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.6 Distribution of displacement and rotational parameters for N X B2 in I-Shaped conformations

Shift NXB2	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	1	2	0	1	0	0	0	4
	4-8	0	0	0	6	8	13	10	1	0	0	38
	8-12	0	0	0	1	2	3	3	0	0	0	9
	12-16	0	0	0	1	1	1	1	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	9	13	17	15	1	0	0	55
Slide NXB2												
	0-4	0	0	1	0	0	0	1	2	0	0	4
	4-8	0	1	1	6	10	6	6	6	2	0	38
	8-12	0	2	1	0	0	4	1	1	0	0	9
	12-16	0	0	1	1	2	0	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	3	4	7	12	10	8	9	2	0	55
Rise NXB2												
	0-4	0	0	0	1	0	2	0	1	0	0	4
	4-8	0	0	5	6	4	9	9	4	1	0	38
	8-12	0	0	1	1	1	2	2	2	0	0	9
	12-16	0	0	0	1	0	0	3	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	6	9	5	13	14	7	1	0	55

Tilt N X B2	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	2	1	1	0	0	0	0	4
	4-8	0	6	10	5	9	3	5	0	38
	8-12	0	1	1	4	0	2	1	0	9
	12-16	0	0	0	2	1	1	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	9	12	12	10	6	6	0	55
Roll N X B2										
	0-4	0	2	0	1	0	0	1	0	4
	4-8	1	4	5	7	9	7	3	2	38
	8-12	0	0	1	3	4	0	1	0	9
	12-16	0	0	1	3	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	1	6	7	14	13	7	5	2	
Twist N X B2										
	0-4	0	0	1	0	0	2	1	0	4
	4-8	2	3	4	6	6	10	5	2	38
	8-12	1	2	0	3	0	2	1	0	9
	12-16	0	0	2	1	0	1	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	3	5	7	10	6	15	7	2	55

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.7 Distribution of displacement and rotational parameters for N X B2 in V-Shaped conformations

Shift NXB2	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	3	6	0	0	1	0	0	10
	4-8	0	4	21	38	104	40	38	11	1	0	257
	8-12	0	1	28	48	102	77	37	4	0	0	297
	12-16	0	0	6	15	10	13	19	4	0	0	67
	16-20	0	0	0	1	0	4	0	0	0	0	5
	Sum[e]	0	5	55	105	222	134	94	20	1	0	636
Slide NXB2												
	0-4	0	1	4	0	0	2	3	0	0	0	10
	4-8	4	22	46	32	16	22	18	48	41	8	257
	8-12	6	47	53	26	22	13	17	53	53	7	297
	12-16	9	13	7	7	6	3	3	7	6	6	67
	16-20	2	1	1	0	0	1	0	0	0	0	5
	Sum	21	84	111	65	44	41	41	108	100	21	636
Rise NXB2												
	0-4	0	0	1	4	1	1	2	1	0	0	10
	4-8	1	16	39	29	29	37	39	48	18	1	257
	8-12	1	17	46	40	39	34	40	53	25	2	297
	12-16	0	7	3	11	9	11	9	6	9	2	67
	16-20	0	1	0	1	0	1	2	0	0	0	5
	Sum	2	41	89	85	78	84	92	108	52	5	636

Tilt N X B2	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	1	0	3	1	4	1	0	0	10
	4-8	4	25	38	47	72	42	18	11	257
	8-12	6	29	37	68	77	48	23	9	297
	12-16	0	3	6	17	17	20	2	2	67
	16-20	0	0	0	3	1	1	0	0	5
	Sum[f]	11	57	84	136	171	112	43	22	636
Roll N X B2										
	0-4	0	0	2	3	0	3	2	0	10
	4-8	2	29	25	60	68	27	37	9	257
	8-12	3	29	43	78	59	47	34	4	297
	12-16	3	9	10	16	11	5	11	2	67
	16-20	0	0	1	1	1	1	1	0	5
	Sum	8	67	81	158	139	83	85	15	636
Twist N X B2										
	0-4	1	0	4	1	1	2	0	1	10
	4-8	33	35	35	26	36	30	33	29	257
	8-12	39	41	35	37	27	37	36	45	297
	12-16	7	20	6	5	6	9	7	7	67
	16-20	0	3	1	0	0	0	1	0	5
	Sum	80	99	81	69	70	78	77	82	636

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.8 Distribution of displacement and rotational parameters for N X B2 in E-Shaped conformations

Shift NXP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	15	0	0	0	0	0	15
	4-8	0	0	0	0	6	0	0	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	0	21	0	0	0	0	0	21
Slide NXP												
	0-4	0	0	0	0	0	6	9	0	0	0	15
	4-8	0	0	0	0	0	4	2	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	0	0	10	11	0	0	0	21
Rise NXP												
	0-4	0	0	0	0	0	9	6	0	0	0	15
	4-8	0	0	0	0	0	2	4	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	0	0	11	10	0	0	0	21

Tilt N X P	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	1	8	0	0	6	0	0	15
	4-8	0	0	2	0	0	4	0	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	1	10	0	0	10	0	0	21
Roll N X P										
	0-4	0	0	6	9	0	0	0	0	15
	4-8	0	0	5	1	0	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	0	11	10	0	0	0	0	21
Twist N X P										
	0-4	0	0	3	3	0	0	9	0	15
	4-8	0	0	1	3	0	0	2	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	0	4	6	0	0	11	0	21

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.9 Distribution of displacement and rotational parameters for N X P in C-Shaped conformations

Shift NXP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	5	0	0	0	0	0	5
	4-8	0	0	0	1	9	0	0	0	0	0	10
	8-12	0	0	0	0	1	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	1	15	0	0	0	0	0	16
Slide NXP												
	0-4	0	0	0	0	0	4	1	0	0	0	5
	4-8	0	0	0	0	0	7	3	0	0	0	10
	8-12	0	0	0	0	0	0	1	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	0	0	11	5	0	0	0	16
Rise NXP												
	0-4	0	0	0	0	0	1	4	0	0	0	5
	4-8	0	0	0	0	1	2	7	0	0	0	10
	8-12	0	0	0	0	0	1	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	0	1	4	11	0	0	0	16

Tilt N X P	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	1	0	0	4	0	0	5
	4-8	0	1	2	0	0	7	0	0	10
	8-12	0	0	1	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	1	4	0	0	11	0	0	16
Roll N X P										
	0-4	0	0	4	1	0	0	0	0	5
	4-8	0	0	7	2	1	0	0	0	10
	8-12	0	0	0	1	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	0	11	4	1	0	0	0	16
Twist N X P										
	0-4	0	0	4	0	0	0	1	0	5
	4-8	0	0	2	5	0	0	3	0	10
	8-12	0	0	0	0	0	0	1	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	0	6	5	0	0	5	0	16

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.10 Distribution of displacement and rotational parameters for N X P in I-Shaped conformations

Shift NXP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	4	0	0	0	0	0	4
	4-8	0	0	0	4	26	8	0	0	0	0	38
	8-12	0	0	0	2	6	1	0	0	0	0	9
	12-16	0	0	0	0	3	1	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	6	39	10	0	0	0	0	55
Slide NXP												
	0-4	0	0	0	0	0	1	3	0	0	0	4
	4-8	0	0	0	8	0	13	17	0	0	0	38
	8-12	0	0	0	1	0	2	6	0	0	0	9
	12-16	0	0	0	1	1	0	2	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	10	1	16	28	0	0	0	55
Rise NXP												
	0-4	0	0	0	0	0	3	1	0	0	0	4
	4-8	0	0	0	2	5	18	13	0	0	0	38
	8-12	0	0	0	0	3	4	2	0	0	0	9
	12-16	0	0	0	0	2	0	2	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	2	10	25	18	0	0	0	55

Tilt N X P	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	2	0	0	2	0	0	4
	4-8	0	2	8	4	6	17	1	0	38
	8-12	0	1	2	0	3	3	0	0	9
	12-16	0	0	0	1	1	2	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	3	12	5	10	24	1	0	55
Roll N X P										
	0-4	0	0	1	3	0	0	0	0	4
	4-8	0	2	20	10	0	3	3	0	38
	8-12	0	0	4	2	1	1	1	0	9
	12-16	0	0	2	1	0	1	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	2	27	16	1	5	4	0	55
Twist N X P										
	0-4	0	0	0	1	0	1	2	0	4
	4-8	8	0	3	10	2	4	8	3	38
	8-12	1	0	1	1	0	3	1	2	9
	12-16	1	0	1	0	0	1	0	1	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	10	0	5	12	2	9	11	6	55

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.11 Distribution of displacement and rotational parameters for N X P in V-Shaped conformations

Shift NXP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	2	8	0	0	0	0	0	10
	4-8	0	0	0	38	198	21	0	0	0	0	257
	8-12	0	0	0	41	239	17	0	0	0	0	297
	12-16	0	0	0	16	38	13	0	0	0	0	67
	16-20	0	0	0	1	1	3	0	0	0	0	5
	Sum[e]	0	0	0	98	484	54	0	0	0	0	636
Slide NXP												
	0-4	0	0	0	0	0	0	10	0	0	0	10
	4-8	0	0	0	20	5	45	187	0	0	0	257
	8-12	0	0	0	16	9	46	226	0	0	0	297
	12-16	0	0	0	10	3	2	52	0	0	0	67
	16-20	0	0	0	3	0	0	2	0	0	0	5
	Sum	0	0	0	49	17	93	477	0	0	0	636
Rise NXP												
	0-4	0	0	0	3	1	6	0	0	0	0	10
	4-8	0	0	0	24	41	150	42	0	0	0	257
	8-12	0	0	0	18	79	156	44	0	0	0	297
	12-16	0	0	0	1	23	31	12	0	0	0	67
	16-20	0	0	0	0	0	4	1	0	0	0	5
	Sum	0	0	0	46	144	347	99	0	0	0	636

Tilt N X P	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	3	1	2	4	0	0	10
	4-8	0	7	50	48	49	101	2	0	257
	8-12	0	15	40	68	69	92	13	0	297
	12-16	0	4	1	18	15	20	9	0	67
	16-20	0	1	0	0	0	2	2	0	5
	Sum[f]	0	27	94	135	135	219	26	0	636
Roll N X P										
	0-4	0	1	0	4	0	5	0	0	10
	4-8	0	14	54	79	3	85	22	0	257
	8-12	0	16	78	73	10	86	26	8	297
	12-16	0	1	36	8	6	11	1	4	67
	16-20	0	0	3	1	1	0	0	0	5
	Sum	0	32	171	165	20	187	49	12	636
Twist N X P										
	0-4	0	0	0	0	3	4	0	3	10
	4-8	20	0	17	25	30	60	33	72	257
	8-12	16	1	21	26	20	70	61	82	297
	12-16	10	2	1	1	3	16	13	21	67
	16-20	3	0	0	0	0	1	0	1	5
	Sum	49	3	39	52	56	151	107	179	636

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.12 Distribution of displacement and rotational parameters for N X P in E-Shaped conformations

Shift B1XB2	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	9	6	0	0	0	0	15
	4-8	0	0	0	0	2	4	0	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	0	11	10	0	0	0	0	21
Slide B1XB2												
	0-4	0	0	0	7	2	2	4	0	0	0	15
	4-8	0	0	0	4	0	0	2	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	11	2	2	6	0	0	0	21
Rise B1XB2												
	0-4	0	0	0	1	5	2	7	0	0	0	15
	4-8	0	0	0	0	2	4	0	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	1	7	6	7	0	0	0	21

Tilt B1XB2	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	1	6	4	4	0	0	15
	4-8	0	0	1	2	0	3	0	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	0	2	8	4	7	0	0	21
Roll B1XB2										
	0-4	0	3	4	2	0	6	0	0	15
	4-8	0	0	3	0	0	2	1	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	3	7	2	0	8	1	0	21
Twist B1XB2										
	0-4	3	0	5	1	1	3	0	2	15
	4-8	2	0	2	0	0	0	0	2	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	5	0	7	1	1	3	0	4	21

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.13 Distribution of displacement and rotational parameters for B1X B2 in C-Shaped conformations

Shift B1XB2	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	3	2	0	0	0	0	5
	4-8	0	0	0	0	5	5	0	0	0	0	10
	8-12	0	0	0	0	0	1	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	0	8	8	0	0	0	0	16
Slide B1XB2												
	0-4	0	0	0	0	1	1	3	0	0	0	5
	4-8	0	0	0	5	2	1	2	0	0	0	10
	8-12	0	0	0	0	0	0	1	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	5	3	2	6	0	0	0	16
Rise B1XB2												
	0-4	0	0	0	1	1	2	1	0	0	0	5
	4-8	0	0	0	4	0	5	1	0	0	0	10
	8-12	0	0	0	0	1	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	5	2	7	2	0	0	0	16

Tilt B1XB2	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	3	0	2	0	0	0	5
	4-8	0	0	3	3	3	1	0	0	10
	8-12	0	0	0	0	0	1	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	0	6	3	5	2	0	0	16
Roll B1XB2										
	0-4	0	1	0	2	1	0	1	0	5
	4-8	0	2	3	1	0	1	3	0	10
	8-12	0	0	0	0	0	0	1	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	3	3	3	1	1	5	0	16
Twist B1XB2										
	0-4	0	0	1	0	0	1	0	3	5
	4-8	2	1	2	2	1	0	0	2	10
	8-12	0	0	0	0	0	0	0	1	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	2	1	3	2	1	1	0	6	16

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.14 Distribution of displacement and rotational parameters for B1X B2 in I-Shaped conformations

Shift B1XB2	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	3	1	0	0	0	0	4
	4-8	0	0	0	0	20	18	0	0	0	0	38
	8-12	0	0	0	0	3	6	0	0	0	0	9
	12-16	0	0	0	0	1	3	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	0	27	28	0	0	0	0	55
Slide B1XB2												
	0-4	0	0	0	3	0	0	1	0	0	0	4
	4-8	0	0	0	12	5	6	15	0	0	0	38
	8-12	0	0	0	5	1	1	2	0	0	0	9
	12-16	0	0	0	1	0	3	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	21	6	10	18	0	0	0	55
Rise B1XB2												
	0-4	0	0	0	0	2	2	0	0	0	0	4
	4-8	0	0	0	3	12	13	10	0	0	0	38
	8-12	0	0	0	0	4	3	2	0	0	0	9
	12-16	0	0	0	1	0	0	3	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	4	18	18	15	0	0	0	55

Tilt B1XB2	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	0	0	0	4	0	0	4
	4-8	0	0	8	13	10	7	0	0	38
	8-12	0	0	3	3	3	0	0	0	9
	12-16	0	0	0	2	2	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	0	11	18	15	11	0	0	55
Roll B1XB2										
	0-4	0	0	2	0	0	2	0	0	4
	4-8	0	10	12	1	1	8	6	0	38
	8-12	0	2	2	1	1	3	0	0	9
	12-16	0	3	0	0	0	0	1	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	15	16	2	2	13	7	0	55
Twist B1XB2										
	0-4	3	0	0	0	0	0	0	1	4
	4-8	8	1	3	5	4	5	5	7	38
	8-12	1	4	0	1	1	0	0	2	9
	12-16	0	0	1	0	2	1	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	12	5	4	6	7	6	5	10	55

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.15 Distribution of displacement and rotational parameters for B1X B2 in V-Shaped conformations

Shift B1XB2	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	6	4	0	0	0	0	10
	4-8	0	0	0	0	137	120	0	0	0	0	255
	8-12	0	0	0	0	155	142	0	0	0	0	294
	12-16	0	0	0	0	32	35	0	0	0	0	67
	16-20	0	0	0	0	2	3	0	0	0	0	5
	Sum[e]	0	0	0	0	332	304	0	0	0	0	636
Slide B1XB2												
	0-4	0	0	0	3	0	0	7	0	0	0	10
	4-8	0	0	0	71	35	29	122	0	0	0	257
	8-12	0	0	0	99	35	38	125	0	0	0	297
	12-16	0	0	0	24	7	2	34	0	0	0	67
	16-20	0	0	0	2	0	0	3	0	0	0	5
	Sum	0	0	0	199	77	69	291	0	0	0	636
Rise B1XB2												
	0-4	0	0	0	1	5	3	1	0	0	0	10
	4-8	0	0	0	29	81	93	54	0	0	0	257
	8-12	0	0	0	40	113	86	58	0	0	0	297
	12-16	0	0	0	5	24	29	9	0	0	0	67
	16-20	0	0	0	0	3	2	0	0	0	0	5
	Sum	0	0	0	75	226	213	122	0	0	0	636

Tilt B1XB2	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	2	1	3	4	0	0	10
	4-8	0	0	79	54	74	50	0	0	257
	8-12	0	0	86	58	100	53	0	0	297
	12-16	0	0	21	14	11	21	0	0	67
	16-20	0	0	3	1	0	1	0	0	5
	Sum[f]	0	0	191	128	188	129	0	0	636
Roll B1XB2										
	0-4	0	0	3	0	0	6	1	0	10
	4-8	0	58	50	38	35	40	36	0	257
	8-12	0	68	46	31	29	67	56	0	297
	12-16	0	12	14	11	10	12	8	0	67
	16-20	0	0	2	0	0	2	1	0	5
	Sum	0	138	115	80	74	127	102	0	636
Twist B1XB2										
	0-4	2	0	1	0	0	3	0	4	10
	4-8	41	12	36	17	17	29	21	84	257
	8-12	54	29	31	20	20	29	31	83	297
	12-16	17	5	5	4	1	7	4	24	67
	16-20	2	0	0	0	0	0	1	2	5
	Sum	116	46	73	41	38	68	57	197	636

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.16 Distribution of displacement and rotational parameters for B1X B2 in E-Shaped conformations

Shift B1XP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	1	8	6	0	0	0	0	0	15
	4-8	0	0	2	1	2	1	0	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	3	9	8	1	0	0	0	0	21
Slide B1XP												
	0-4	0	0	0	1	10	2	2	0	0	0	15
	4-8	0	0	0	0	2	3	1	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	1	12	5	3	0	0	0	21
Rise B1XP												
	0-4	0	0	0	0	0	7	8	0	0	0	15
	4-8	0	0	0	0	0	0	6	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	0	0	7	14	0	0	0	21

Tilt B1XP	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	2	1	1	3	2	5	1	0	15
	4-8	0	1	3	0	2	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	2	2	4	3	4	5	1	0	21
Roll B1XP										
	0-4	0	3	5	4	2	1	0	0	15
	4-8	0	0	1	0	4	1	0	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	3	6	4	6	2	0	0	21
Twist B1XP										
	0-4	3	1	1	3	5	1	0	1	15
	4-8	0	0	1	1	0	3	0	1	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	3	1	2	4	5	4	0	2	21

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.17 Distribution of displacement and rotational parameters for B1X P in C-Shaped conformations

Shift B1XP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	3	1	0	1	0	0	0	0	5
	4-8	0	0	5	2	3	0	0	0	0	0	10
	8-12	0	0	0	1	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	8	4	3	1	0	0	0	0	16
Slide B1XP												
	0-4	0	0	0	2	3	0	0	0	0	0	5
	4-8	0	0	0	0	7	3	0	0	0	0	10
	8-12	0	0	0	1	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	3	10	3	0	0	0	0	16
Rise B1XP												
	0-4	0	0	0	0	0	2	3	0	0	0	5
	4-8	0	0	0	0	0	4	5	1	0	0	10
	8-12	0	0	0	0	0	0	1	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	0	0	6	9	1	0	0	16

Tilt B1XP	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	0	2	3	0	0	0	5
	4-8	1	1	2	2	3	1	0	0	10
	8-12	0	0	0	0	1	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	1	1	2	4	7	1	0	0	16
Roll B1XP										
	0-4	0	0	3	2	0	0	0	0	5
	4-8	0	0	3	2	4	1	0	0	10
	8-12	0	0	0	0	0	1	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	0	6	4	4	2	0	0	16
Twist B1XP										
	0-4	0	0	3	0	0	0	1	1	5
	4-8	1	0	1	1	1	5	1	0	10
	8-12	0	1	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	1	1	4	1	1	5	2	1	16

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.18 Distribution of displacement and rotational parameters for B1XP in I-Shaped conformations

Shift B1XP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	0	2	2	0	0	0	0	4
	4-8	0	0	4	16	10	7	1	0	0	0	38
	8-12	0	0	3	1	3	1	1	0	0	0	9
	12-16	0	0	1	1	1	1	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	8	18	16	11	2	0	0	0	55
Slide B1XP												
	0-4	0	0	0	0	4	0	0	0	0	0	4
	4-8	0	0	0	9	20	9	0	0	0	0	38
	8-12	0	0	0	2	4	3	0	0	0	0	9
	12-16	0	0	0	0	3	1	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	11	31	13	0	0	0	0	55
Rise B1XP												
	0-4	0	0	0	1	0	0	3	0	0	0	4
	4-8	0	0	0	2	2	10	23	1	0	0	38
	8-12	0	0	1	1	3	2	2	0	0	0	9
	12-16	0	0	0	1	1	0	2	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	1	5	6	12	30	1	0	0	55

Tilt B1XP	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	2	2	0	0	0	0	4
	4-8	0	4	3	13	8	3	5	2	38
	8-12	0	2	1	2	1	0	3	0	9
	12-16	0	1	0	1	1	0	1	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	7	6	18	10	3	9	2	55
Roll B1XP										
	0-4	0	0	0	1	3	0	0	0	4
	4-8	1	1	11	7	14	2	1	1	38
	8-12	0	1	3	0	3	1	0	1	9
	12-16	0	0	1	2	1	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	1	2	15	10	21	3	1	2	55
Twist B1XP										
	0-4	0	0	0	0	0	2	0	2	4
	4-8	6	4	3	2	7	6	2	8	38
	8-12	1	3	0	0	1	1	1	2	9
	12-16	0	0	0	1	1	1	0	1	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	7	7	3	3	9	10	3	13	55

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.19 Distribution of displacement and rotational parameters for B1X P in V-Shaped conformations

Shift B1XP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	4	4	0	2	0	0	0	10
	4-8	0	0	82	111	25	26	13	0	0	0	257
	8-12	0	0	111	114	26	29	17	0	0	0	297
	12-16	0	0	23	27	8	8	1	0	0	0	67
	16-20	0	0	0	2	0	3	0	0	0	0	5
	Sum[e]	0	0	216	258	63	66	33	0	0	0	636
Slide B1XP												
	0-4	0	0	0	4	2	2	2	0	0	0	10
	4-8	0	0	9	107	106	32	3	0	0	0	257
	8-12	0	0	6	120	137	31	3	0	0	0	297
	12-16	0	0	2	27	34	4	0	0	0	0	67
	16-20	0	0	0	1	3	1	0	0	0	0	5
	Sum	0	0	17	259	282	70	8	0	0	0	636
Rise B1XP												
	0-4	0	0	2	2	6	0	0	0	0	0	10
	4-8	0	0	2	67	82	48	47	11	0	0	256
	8-12	0	0	14	79	88	72	33	11	0	0	297
	12-16	0	0	4	20	12	22	7	2	0	0	67
	16-20	0	0	1	0	2	0	2	0	0	0	5
	Sum	0	0	23	168	190	142	89	24	0	0	636

Tilt B1XP	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	4	2	0	1	3	0	10
	4-8	7	32	49	44	44	54	21	6	257
	8-12	4	40	41	55	57	58	38	4	297
	12-16	0	4	15	10	14	17	7	0	67
	16-20	0	0	1	1	1	2	0	0	5
	Sum[f]	11	76	110	112	116	132	69	10	636
Roll B1XP										
	0-4	0	0	2	1	3	4	0	0	10
	4-8	0	9	55	63	49	70	9	2	257
	8-12	1	11	63	90	53	60	16	3	297
	12-16	0	6	10	13	16	17	5	0	67
	16-20	0	1	2	1	1	0	0	0	5
	Sum	1	27	132	168	122	151	30	5	636
Twist B1XP										
	0-4	0	0	0	1	3	4	0	2	10
	4-8	36	24	74	31	23	22	21	26	257
	8-12	37	29	78	39	31	28	26	29	297
	12-16	9	2	15	14	6	11	6	4	67
	16-20	1	0	0	0	1	1	0	2	5
	Sum	83	55	167	85	64	66	53	63	636

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.20 Distribution of displacement and rotational parameters for B1X P in E-Shaped conformations

Shift B2XP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	8	7	0	0	0	0	0	15
	4-8	0	0	0	6	0	0	0	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	0	14	7	0	0	0	0	0	21
Slide B2XP												
	0-4	0	0	0	0	7	6	2	0	0	0	15
	4-8	0	0	0	1	3	2	0	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	1	10	8	2	0	0	0	21
Rise B2XP												
	0-4	0	0	0	0	1	5	9	0	0	0	15
	4-8	0	0	0	0	0	1	5	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	0	1	6	14	0	0	0	21

Tilt B2XP	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	1	1	2	5	3	3	0	0	15
	4-8	0	0	2	2	1	1	0	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	1	1	4	7	4	4	0	0	21
Roll B2XP										
	0-4	0	1	7	6	0	1	0	0	15
	4-8	0	3	2	1	0	0	0	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	0	4	9	7	0	1	0	0	21
Twist B2XP										
	0-4	0	2	4	0	2	1	4	2	15
	4-8	1	0	1	0	0	1	3	0	6
	8-12	0	0	0	0	0	0	0	0	0
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	1	2	5	0	2	2	7	2	21

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.21 Distribution of displacement and rotational parameters for B2X P in C-Shaped conformations

Shift B2XP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	1	2	2	0	0	0	0	0	5
	4-8	0	0	0	10	0	0	0	0	0	0	10
	8-12	0	0	0	1	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	1	13	2	0	0	0	0	0	16
Slide B2XP												
	0-4	0	0	0	0	5	0	0	0	0	0	5
	4-8	0	0	0	1	8	1	0	0	0	0	10
	8-12	0	0	0	0	1	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	1	14	1	0	0	0	0	16
Rise B2XP												
	0-4	0	0	0	0	1	2	2	0	0	0	5
	4-8	0	0	0	1	3	1	5	0	0	0	10
	8-12	0	0	0	0	0	1	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	1	4	4	7	0	0	0	16

Tilt B2XP	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	0	2	1	1	1	0	5
	4-8	0	1	3	2	2	1	1	0	10
	8-12	0	0	0	0	0	1	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	0	1	3	4	3	3	2	0	16
Roll B2XP										
	0-4	0	1	1	1	2	0	0	0	5
	4-8	1	4	3	1	1	0	0	0	10
	8-12	0	1	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	1	6	4	2	3	0	0	0	16
Twist B2XP										
	0-4	0	0	0	1	1	2	1	0	5
	4-8	0	2	1	0	1	3	3	0	10
	8-12	1	0	0	0	0	0	0	0	1
	12-16	0	0	0	0	0	0	0	0	0
	16-20	0	0	0	0	0	0	0	0	0
	Sum	1	2	1	1	2	5	4	0	16

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.22 Distribution of displacement and rotational parameters for B2X P in I-Shaped conformations

Shift B2XP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	2	2	0	0	0	0	0	0	4
	4-8	0	0	3	24	6	2	3	0	0	0	38
	8-12	0	0	1	4	3	0	1	0	0	0	9
	12-16	0	0	0	2	2	0	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum[e]	0	0	6	32	11	2	4	0	0	0	55
Slide B2XP												
	0-4	0	0	0	2	2	0	0	0	0	0	4
	4-8	0	0	2	9	18	7	2	0	0	0	38
	8-12	0	0	0	3	3	2	1	0	0	0	9
	12-16	0	0	0	2	2	0	0	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	2	16	25	9	3	0	0	0	55
Rise B2XP												
	0-4	0	0	0	0	4	0	0	0	0	0	4
	4-8	0	0	0	0	6	8	24	0	0	0	38
	8-12	0	0	0	1	2	3	1	2	0	0	9
	12-16	0	0	0	1	0	2	1	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0	0	0
	Sum	0	0	0	2	12	13	26	2	0	0	55

Tilt B2XP	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	0	2	0	2	0	0	4
	4-8	1	1	6	9	8	8	5	0	38
	8-12	0	1	1	2	3	1	1	0	9
	12-16	0	1	1	0	1	1	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum[f]	1	3	8	13	12	12	6	0	55
Roll B2XP										
	0-4	0	0	1	2	0	0	1	0	4
	4-8	4	2	6	8	9	6	3	0	38
	8-12	1	2	1	3	2	0	0	0	9
	12-16	1	0	1	1	1	0	0	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	6	4	9	14	12	6	4	0	55
Twist B2XP										
	0-4	2	0	0	0	0	2	0	0	4
	4-8	4	7	4	3	6	6	3	5	38
	8-12	2	1	2	0	0	1	1	2	9
	12-16	0	3	0	0	0	0	1	0	4
	16-20	0	0	0	0	0	0	0	0	0
	Sum	8	11	6	3	6	9	5	7	55

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

Table E.23 Distribution of displacement and rotational parameters for B2X P in V-Shaped conformations

Shift B2XP	R.E[a]	<-12	-12≤D[b] <-9	-9≤D <-6	-6≤D <-3	-3≤D <0	0≤D <3	3≤D <6	6≤D <9	9≤D <12	≥12	Total[d]
	0-4	0	0	0	6	2	1	1	0	0	0	10
	4-8	0	0	69	125	37	20	6	0	0	0	257
	8-12	0	0	92	128	35	33	9	0	0	0	297
	12-16	0	0	29	23	3	6	6	0	0	0	67
	16-20	0	0	3	2	0	0	0	0	0	0	5
	Sum[e]	0	0	193	284	77	60	22	0	0	0	636
Slide B2XP												
	0-4	0	0	0	3	4	2	1	0	0	0	10
	4-8	0	0	19	94	107	31	6	0	0	0	257
	8-12	0	0	10	125	150	11	1	0	0	0	297
	12-16	0	0	4	33	24	5	1	0	0	0	67
	16-20	0	0	0	2	3	0	0	0	0	0	5
	Sum	0	0	33	257	288	49	9	0	0	0	636
Rise B2XP												
	0-4	0	0	0	5	4	1	0	0	0	0	10
	4-8	0	0	12	52	68	65	49	11	0	0	257
	8-12	0	0	9	53	79	76	65	15	0	0	297
	12-16	0	0	3	16	22	15	8	3	0	0	67
	16-20	0	0	0	1	2	2	0	0	0	0	5
	Sum	0	0	24	127	175	159	122	29	0	0	636

Tilt B2XP	R. E[a]	-180≤A[c] <-135	-135≤A <-90	-90≤A <-45	-45≤A <0	0≤A <45	45≤A <90	90≤A <135	135≤A ≤180	Total[d]
	0-4	0	0	0	7	3	0	0	0	10
	4-8	5	34	37	47	71	42	18	3	257
	8-12	7	39	48	60	52	65	23	3	297
	12-16	1	11	10	14	10	12	8	1	67
	16-20	0	1	1	0	0	2	0	1	5
	Sum[f]	13	85	96	128	136	121	49	8	636
Roll B2XP										
	0-4	1	0	0	1	2	3	1	2	10
	4-8	2	13	43	57	51	64	17	10	257
	8-12	1	17	56	63	54	75	29	2	297
	12-16	0	3	18	9	14	17	4	2	67
	16-20	0	0	1	1	2	1	0	0	5
	Sum	4	33	118	131	123	160	51	16	636
Twist B2XP										
	0-4	5	0	0	0	2	0	1	2	10
	4-8	38	21	52	29	51	30	11	25	257
	8-12	42	27	70	47	29	26	27	29	297
	12-16	5	4	27	10	4	3	5	9	67
	16-20	1	1	1	2	0	0	0	0	5
	Sum	91	53	150	88	86	59	44	65	636

[a] Relative Energy in kcal/mol

[b] Displacement range in Angstroms

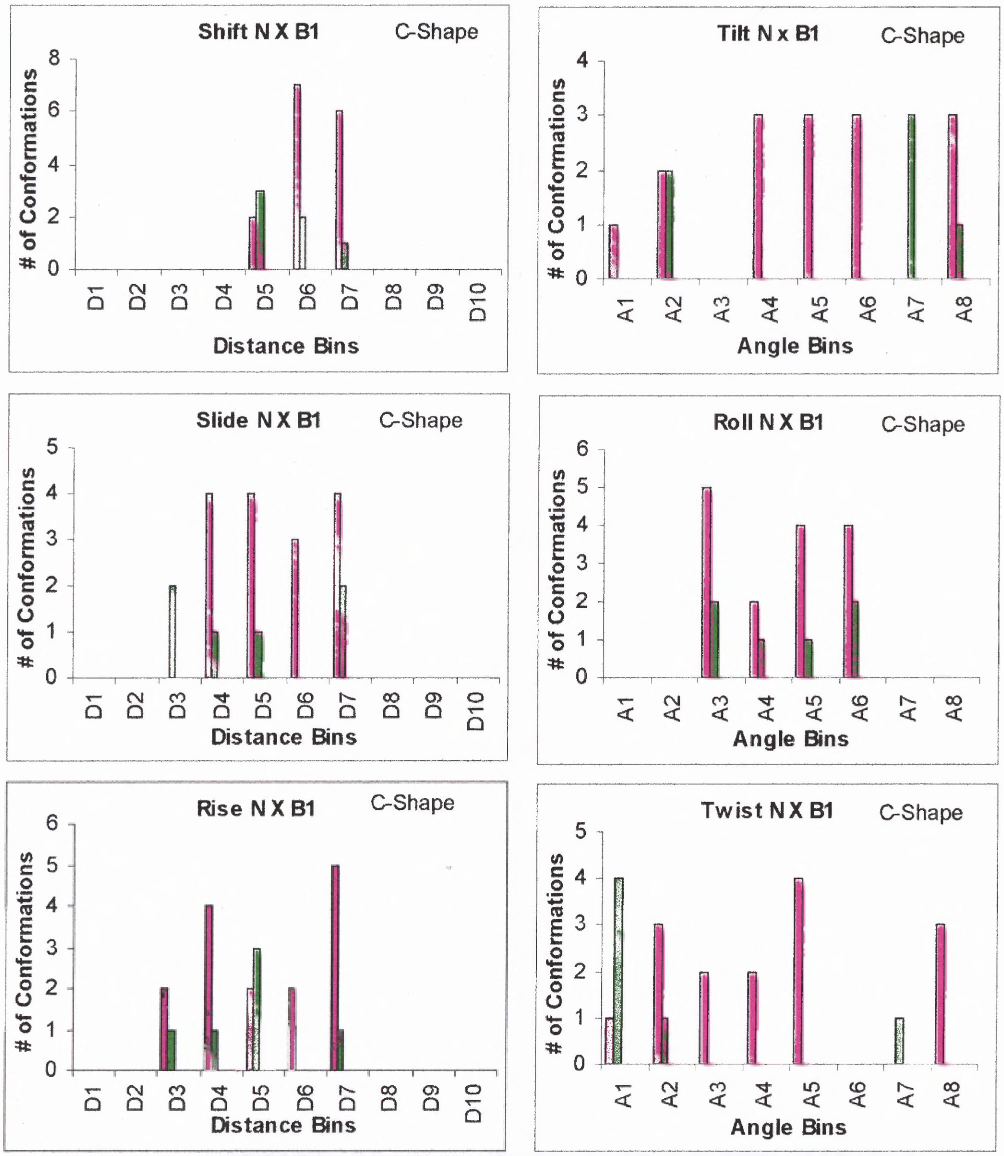
[c] Angle range in degrees

[d] Total number of conformations in each energy bin

[e, f] Total number of conformations in each displacement range and angle bin respectively

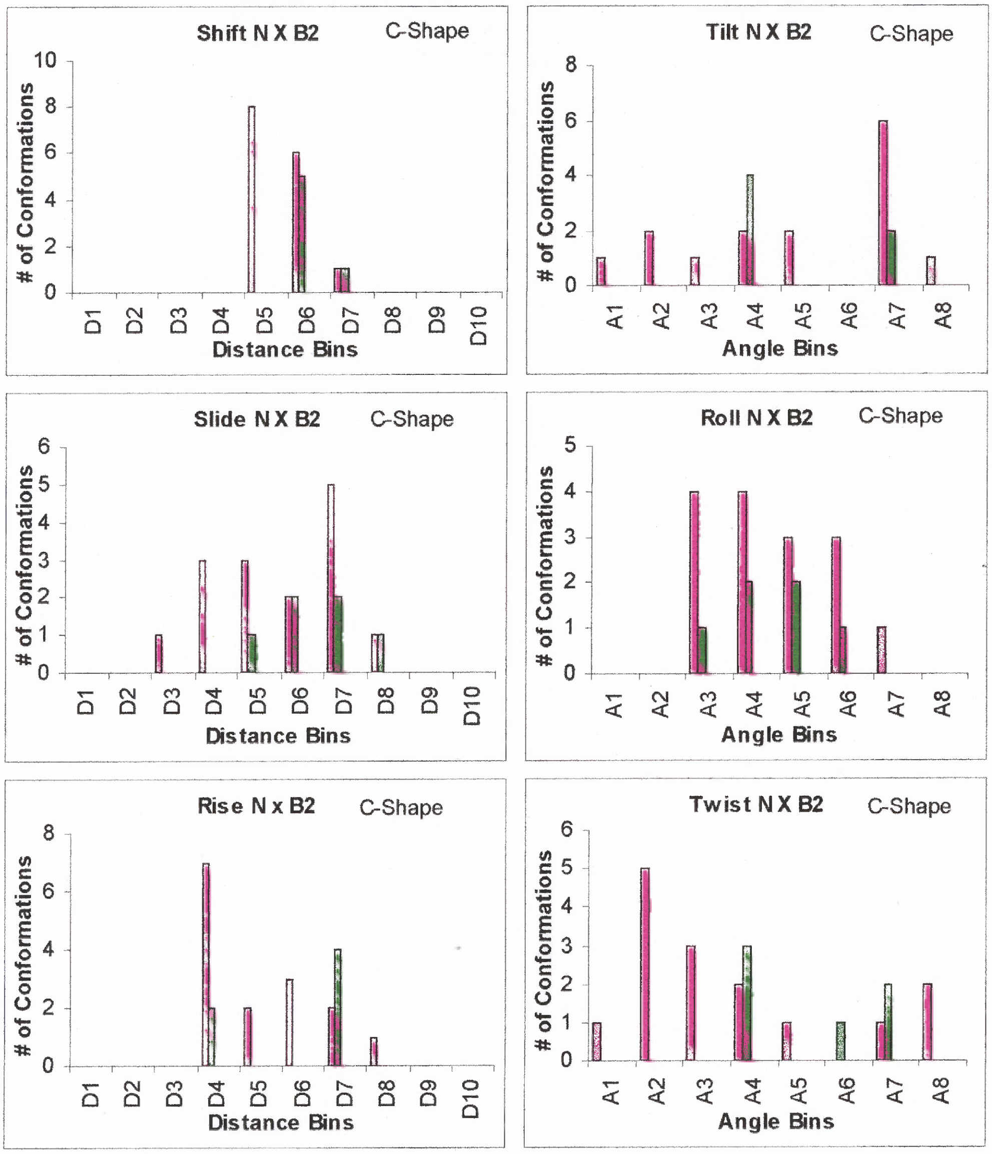
Table E.24 Distribution of displacement and rotational parameters for B2X P in E-Shaped conformations

APPENDIX F
HISTOGRAMS FOR ANALYSIS OF CONFORMATIONS BY MOLECULAR SHAPES



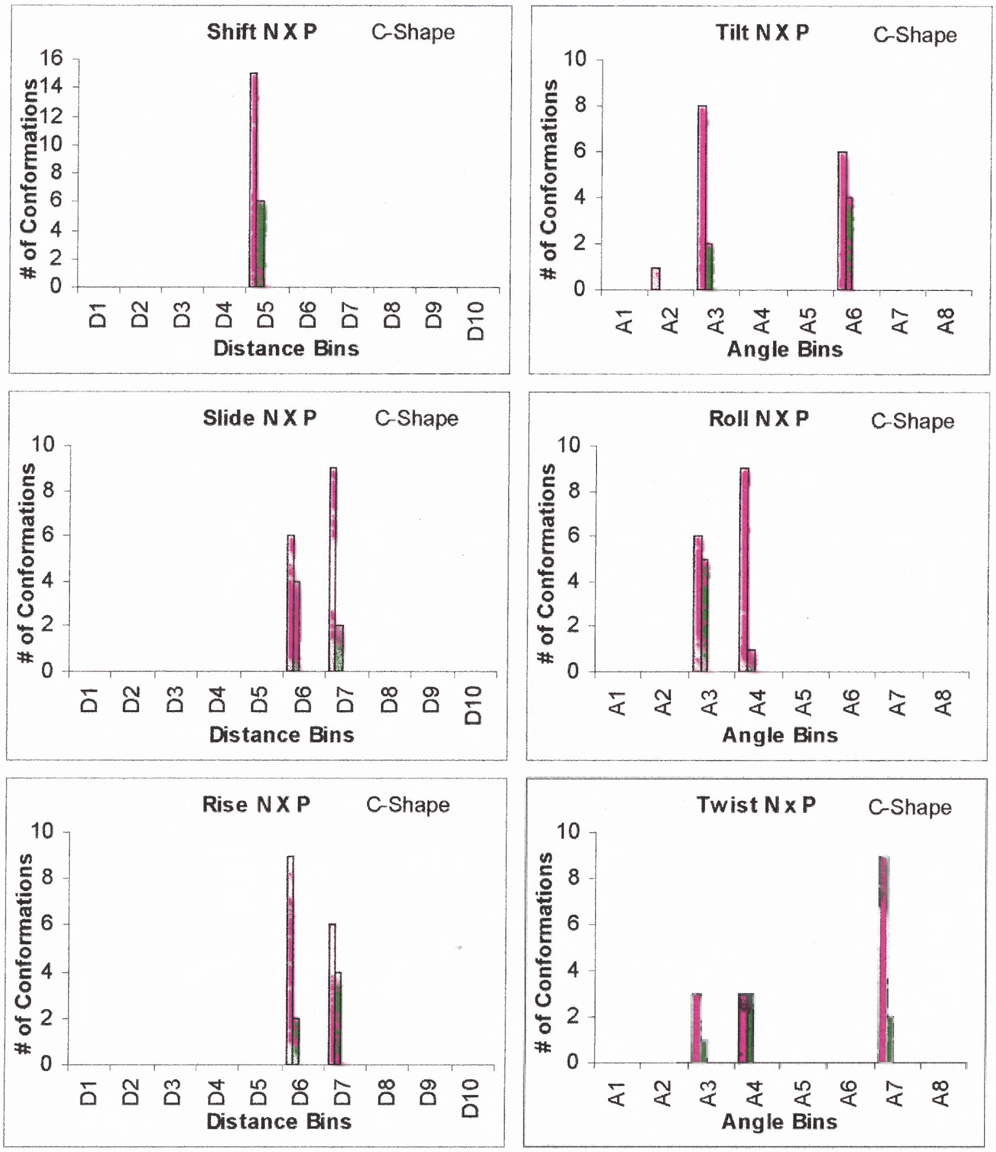
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	🟪 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟩 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟦 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟨 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	🟥 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.1 Histograms of displacement and rotational parameters for N X B1 in C-Shaped conformations



Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	🟥 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟩 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟦 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟨 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	▨ 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.2 Histograms of displacement and rotational parameters for N X B2 in C-Shaped conformations

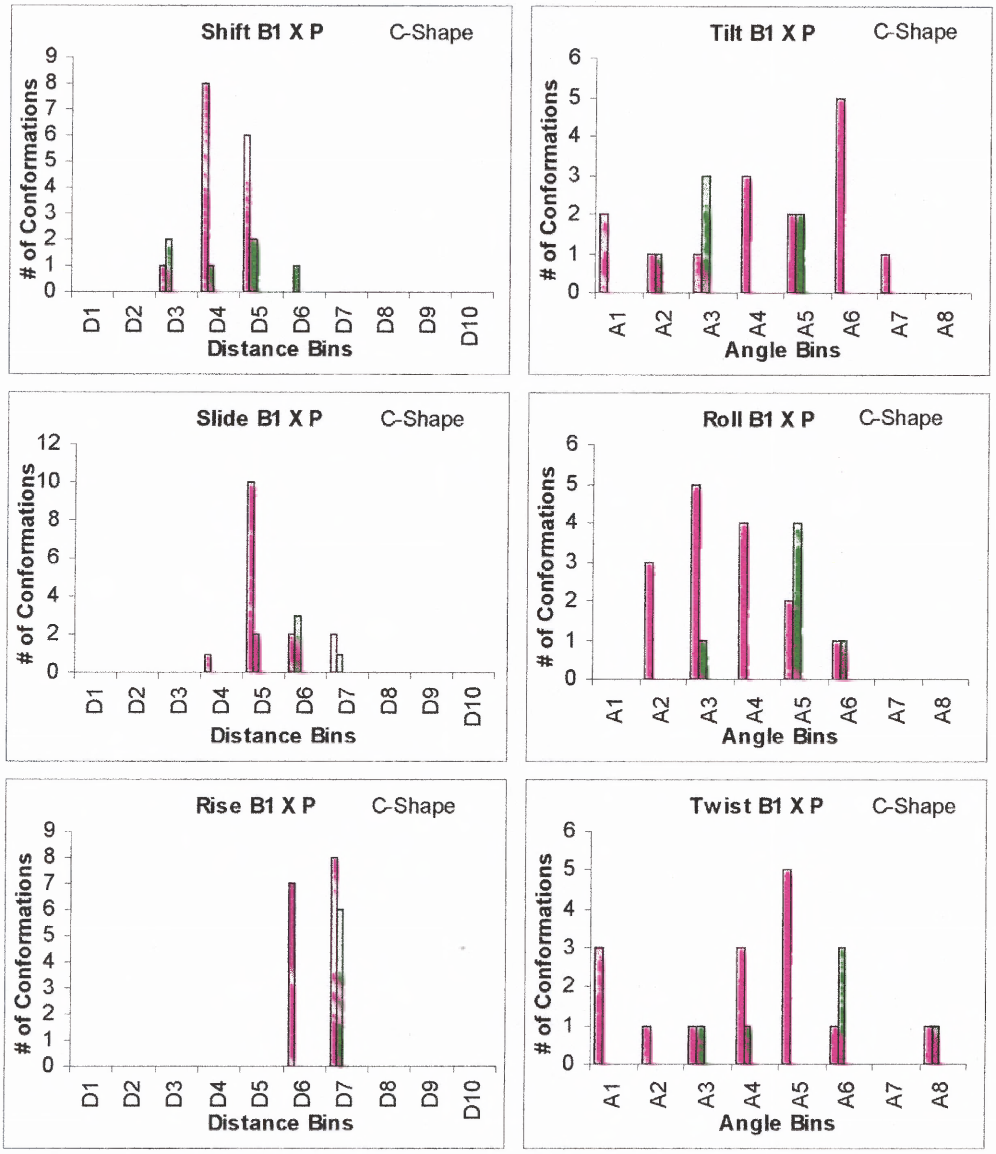


Distance Bins
D1: D<-12
D2: -12≤D<-9
D3: -9≤D<-6
D4: -6≤D<-3
D5: -3≤D<0
D6: 0≤D<3
D7: 3≤D<6
D8: 6≤D<9
D9: 9≤D<12
D10: D>=12

Angle Bins
A1: -180≤A<-135
A2: -135≤A<-90
A3: -90≤A<-45
A4: -45≤A<0
A5: 0≤A<45
A6: 45≤A<90
A7: 90≤A<135
A8: 135≤A≤180

Relative Energy
🟪 0-4 kcal/mol
🟩 4-8 kcal/mol
🟦 8-12 kcal/mol
🟨 12-16 kcal/mol
🟥 16-20 kcal/mol

Figure F.3 Histograms of displacement and rotational parameters for N X P in C-Shaped conformations

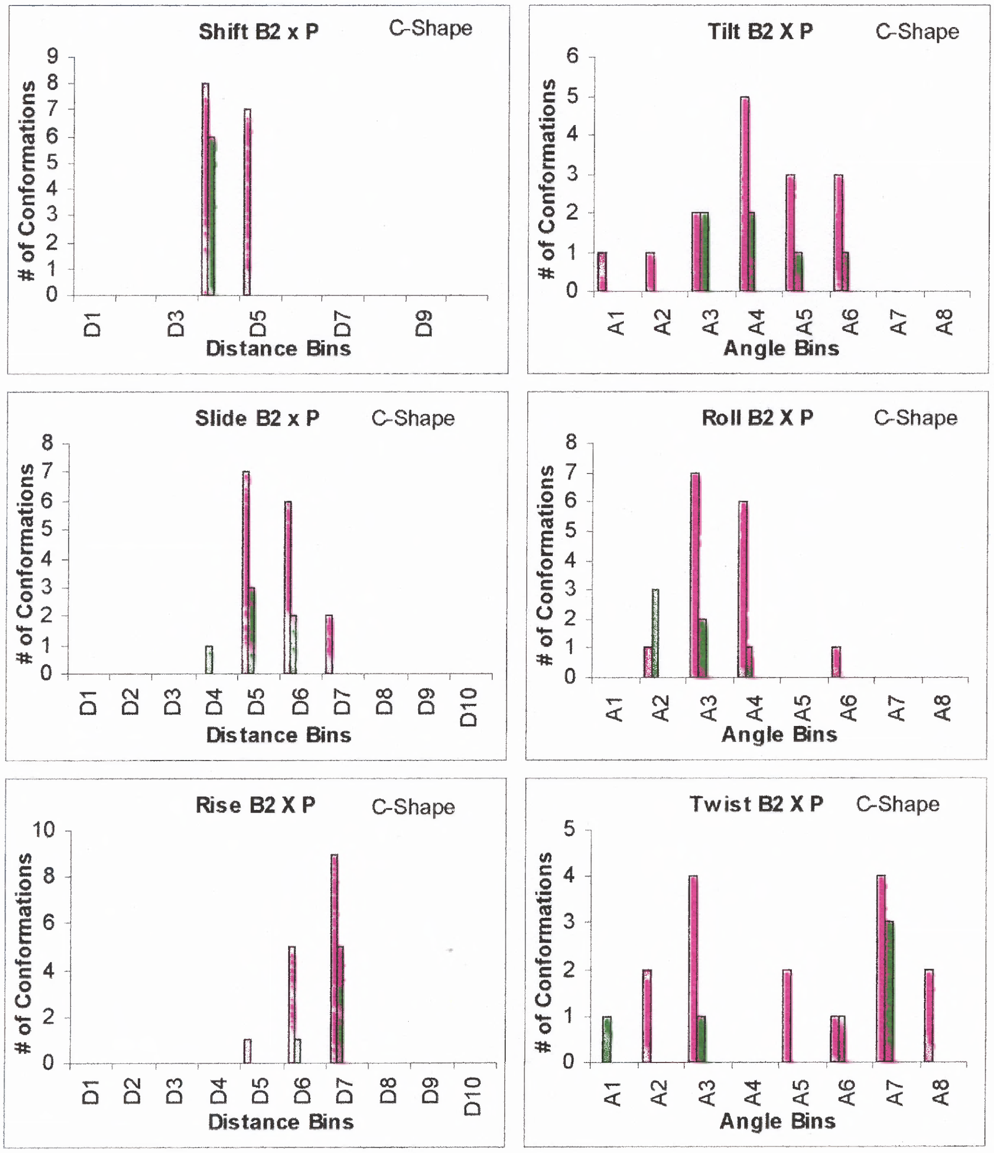


Distance Bins
D1: D<-12
D2: -12≤D<-9
D3: -9≤D<-6
D4: -6≤D<-3
D5: -3≤D<0
D6: 0≤D<3
D7: 3≤D<6
D8: 6≤D<9
D9: 9≤D<12
D10: D>=12

Angle Bins
A1: -180≤A<-135
A2: -135≤A<-90
A3: -90≤A<-45
A4: -45≤A<0
A5: 0≤A<45
A6: 45≤A<90
A7: 90≤A<135
A8: 135≤A≤180

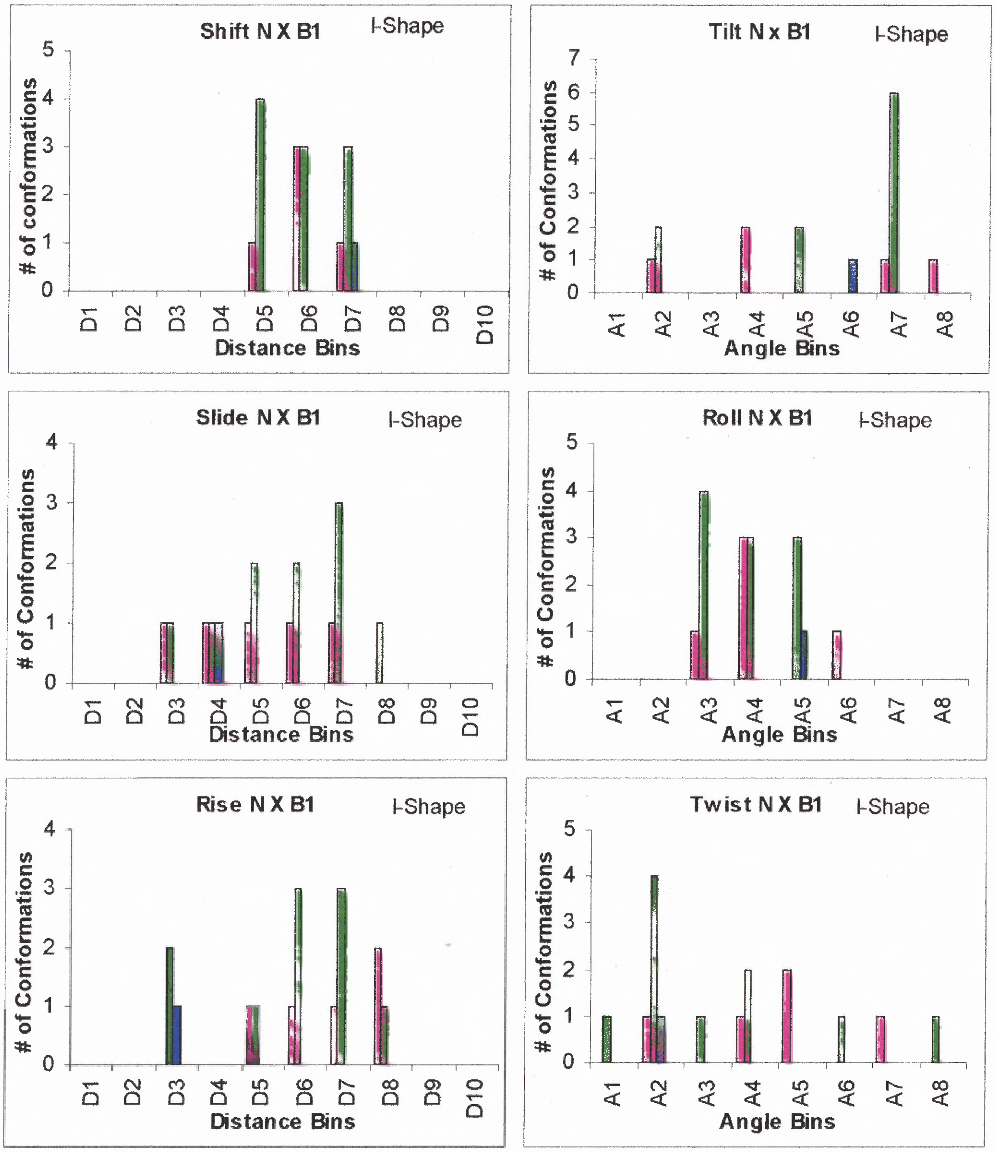
Relative Energy
🟪 0-4 kcal/mol
⬜ 4-8 kcal/mol
🟦 8-12 kcal/mol
🟨 12-16 kcal/mol
🟥 16-20 kcal/mol

Figure F.4 Histograms of displacement and rotational parameters for B1 X P in C-Shaped conformations



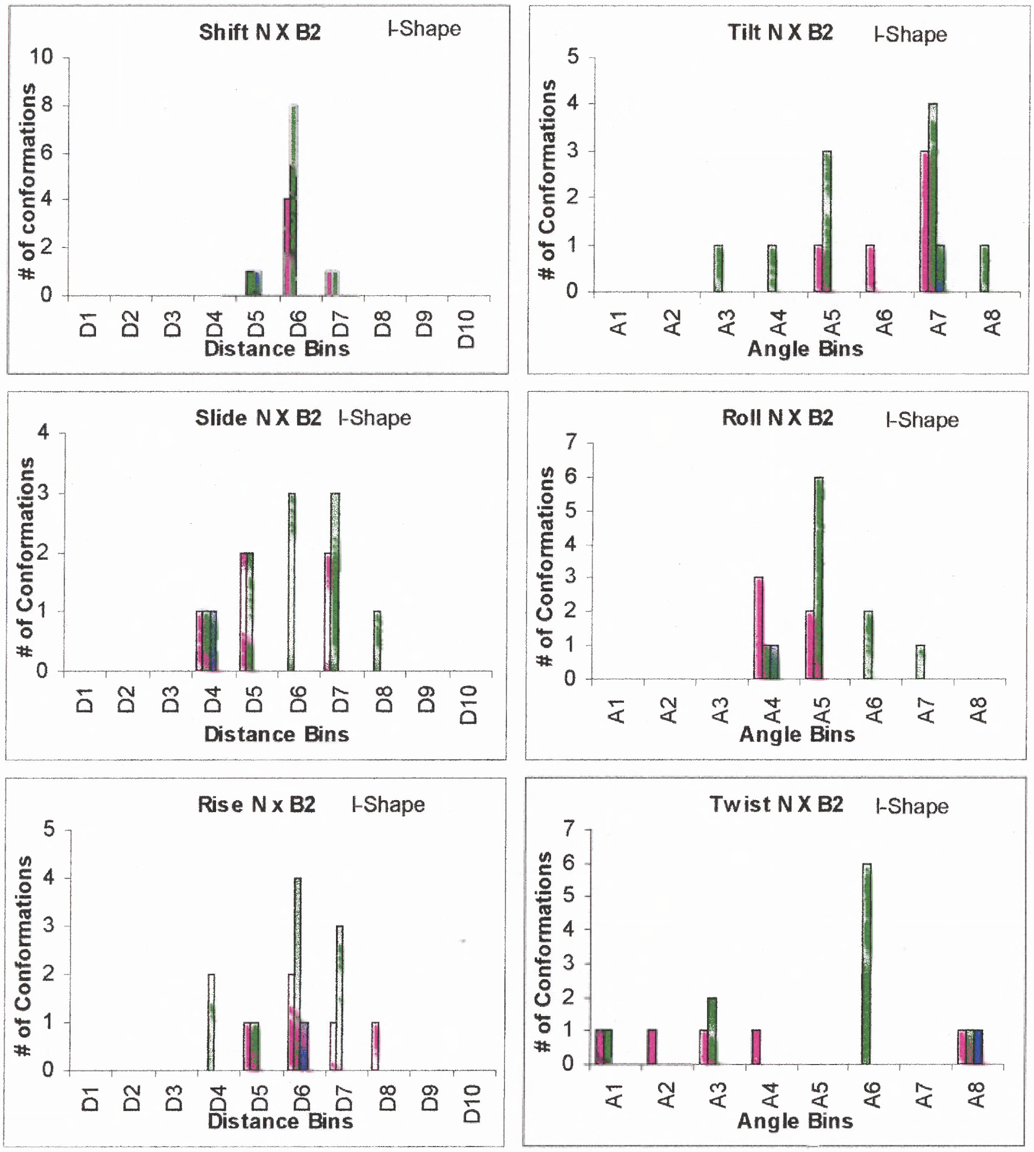
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	🟥 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	▨ 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟦 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟨 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	▨ 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.5 Histograms of displacement and rotational parameters for B2 X P in C-Shaped conformations



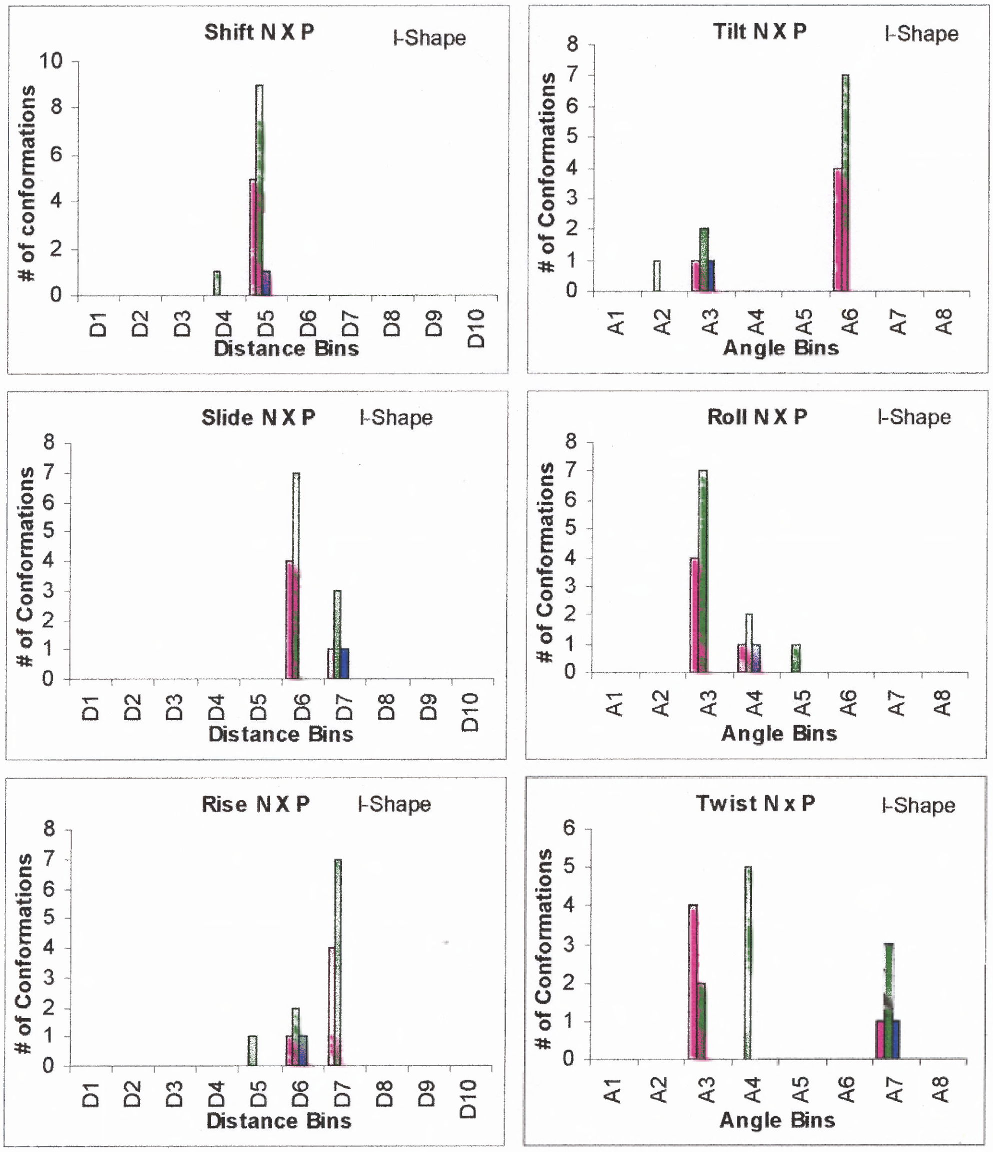
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	🟪 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟩 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟦 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟨 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	🟥 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.6 Histograms of displacement and rotational parameters for N X B1 in I-Shaped conformations



Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	🟪 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟩 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟦 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟨 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	🟥 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.7 Histograms of displacement and rotational parameters for N X B2 in I-Shaped conformations

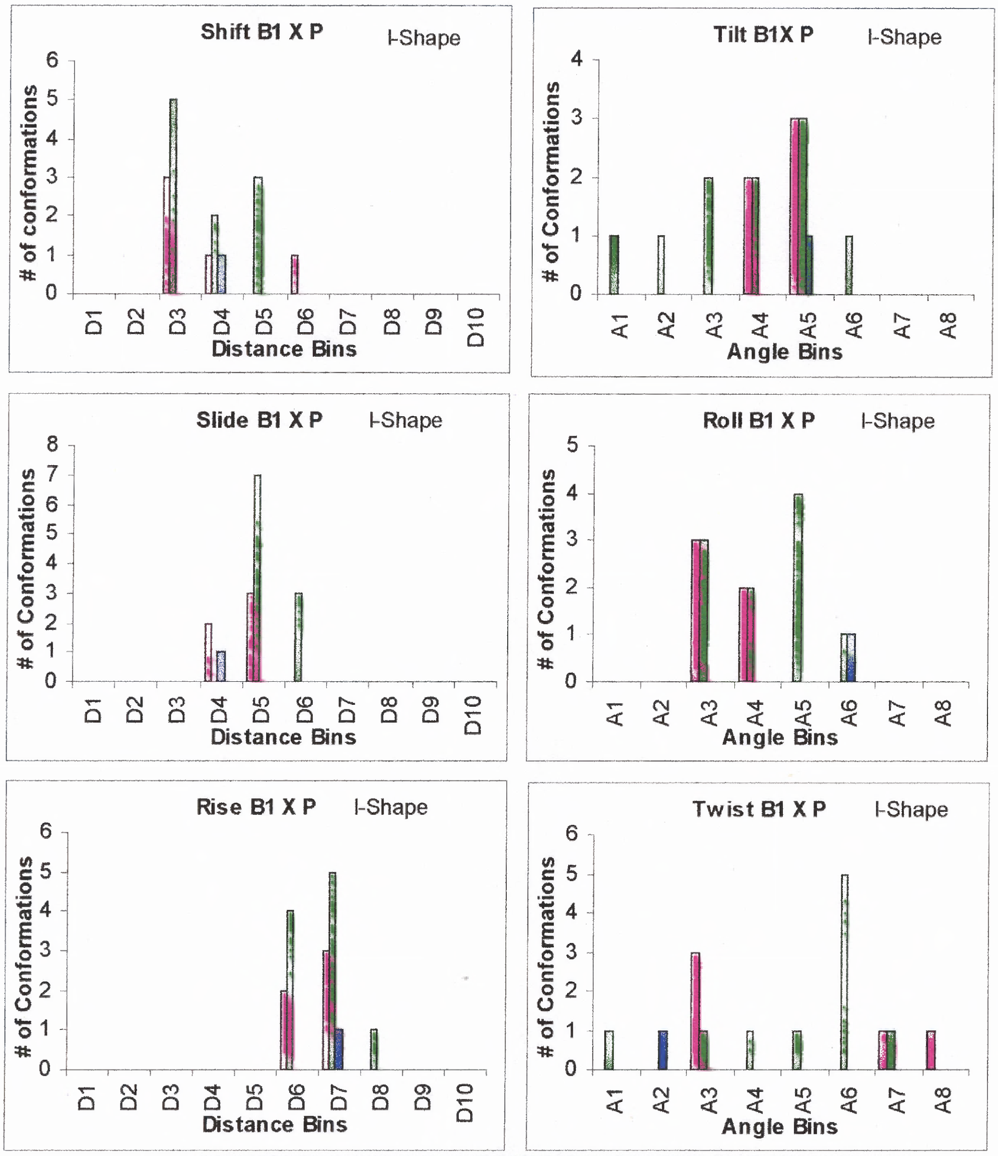


Distance Bins
D1: D<-12
D2: -12≤D<-9
D3: -9≤D<-6
D4: -6≤D<-3
D5: -3≤D<0
D6: 0≤D<3
D7: 3≤D<6
D8: 6≤D<9
D9: 9≤D<12
D10: D>=12

Angle Bins
A1: -180≤A<-135
A2: -135≤A<-90
A3: -90≤A<-45
A4: -45≤A<0
A5: 0≤A<45
A6: 45≤A<90
A7: 90≤A<135
A8: 135≤A≤180

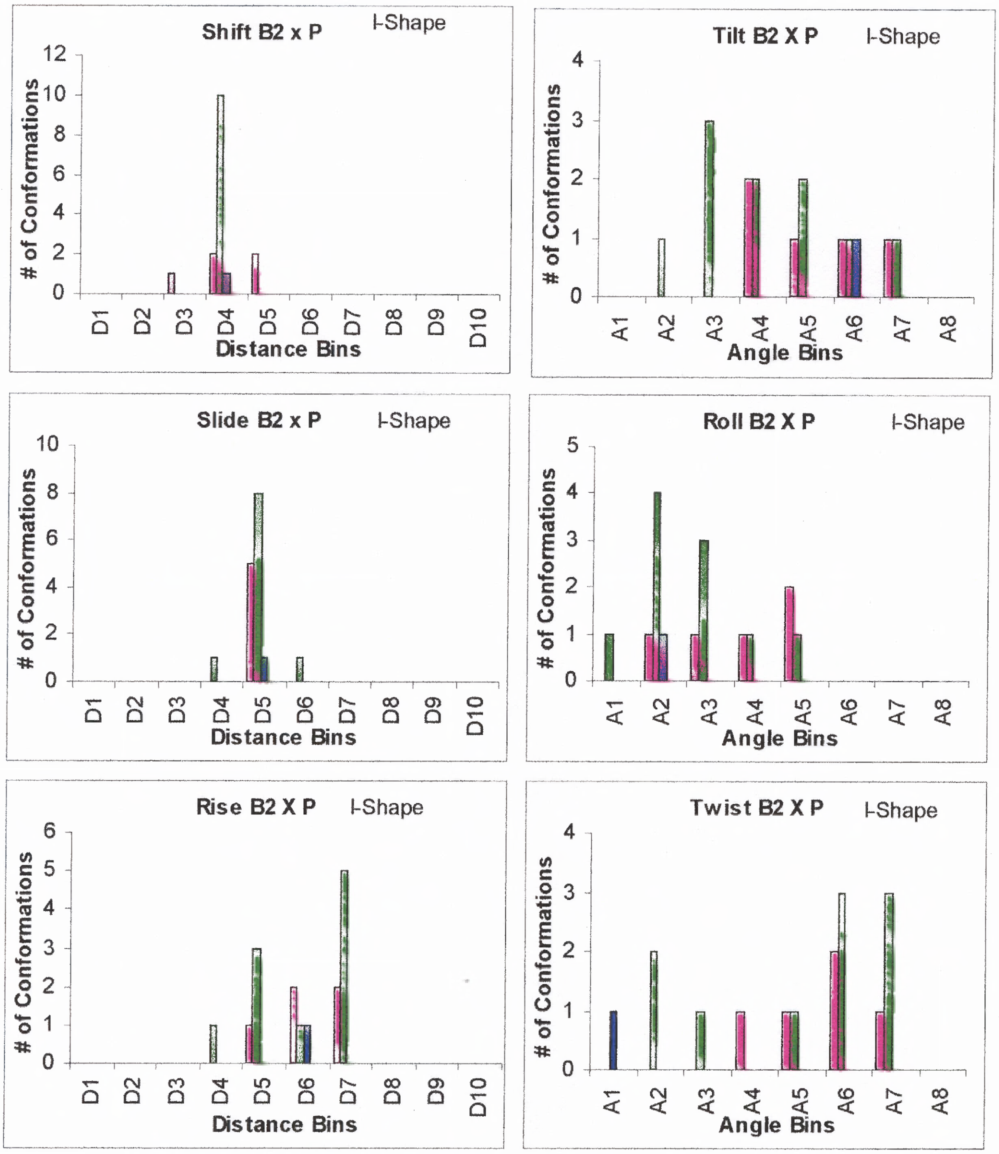
Relative Energy
🟥 0-4 kcal/mol
🟩 4-8 kcal/mol
🟦 8-12 kcal/mol
🟨 12-16 kcal/mol
🟥 16-20 kcal/mol

Figure F.8 Histograms of displacement and rotational parameters for N X P in I-Shaped conformations



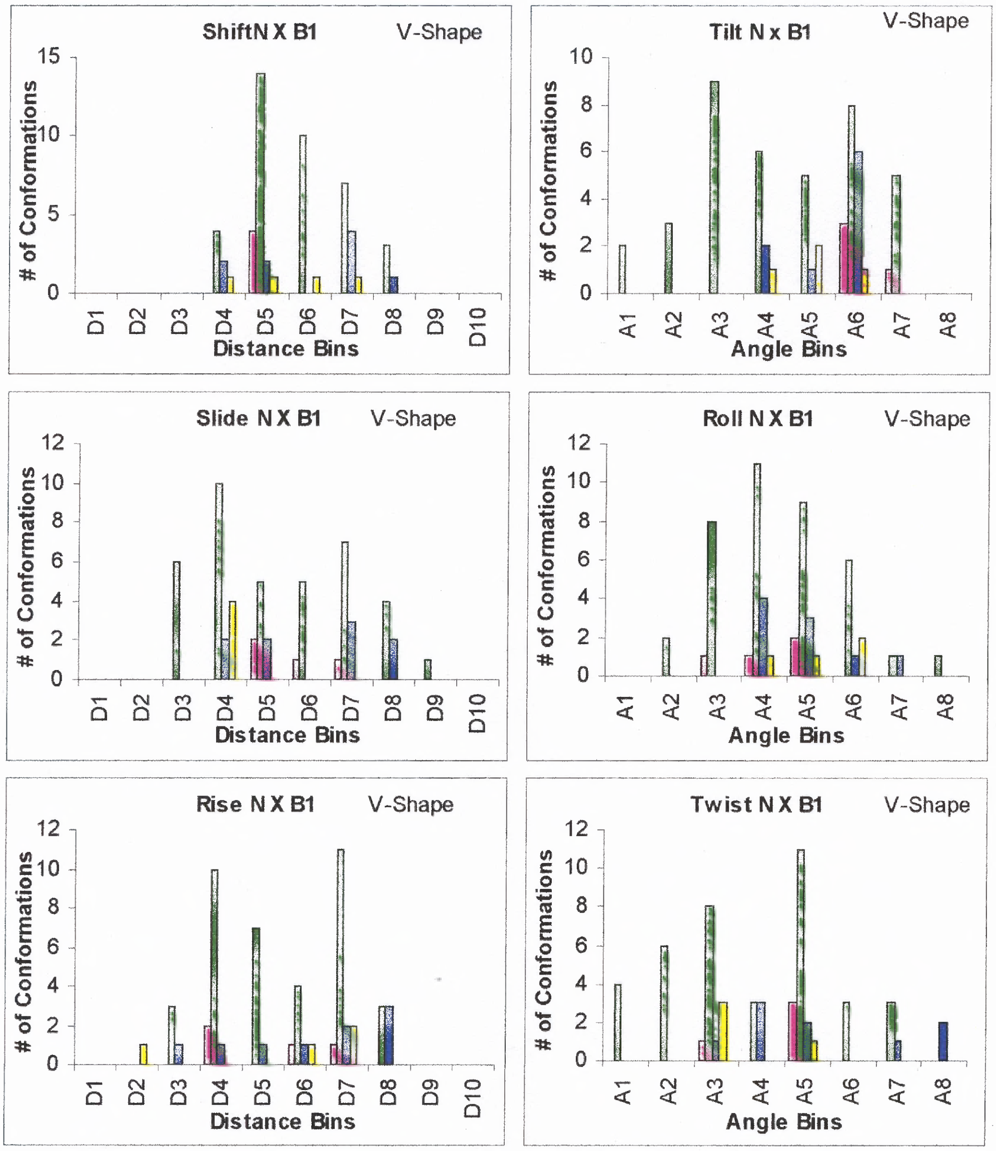
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	🟪 0-4 kcal/mol
D2: -12≤D<-9	A2: -135≤A<-90	🟩 4-8 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟦 8-12 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟨 12-16 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟥 16-20 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.9 Histograms of displacement and rotational parameters for B1 X P in I-Shaped conformations



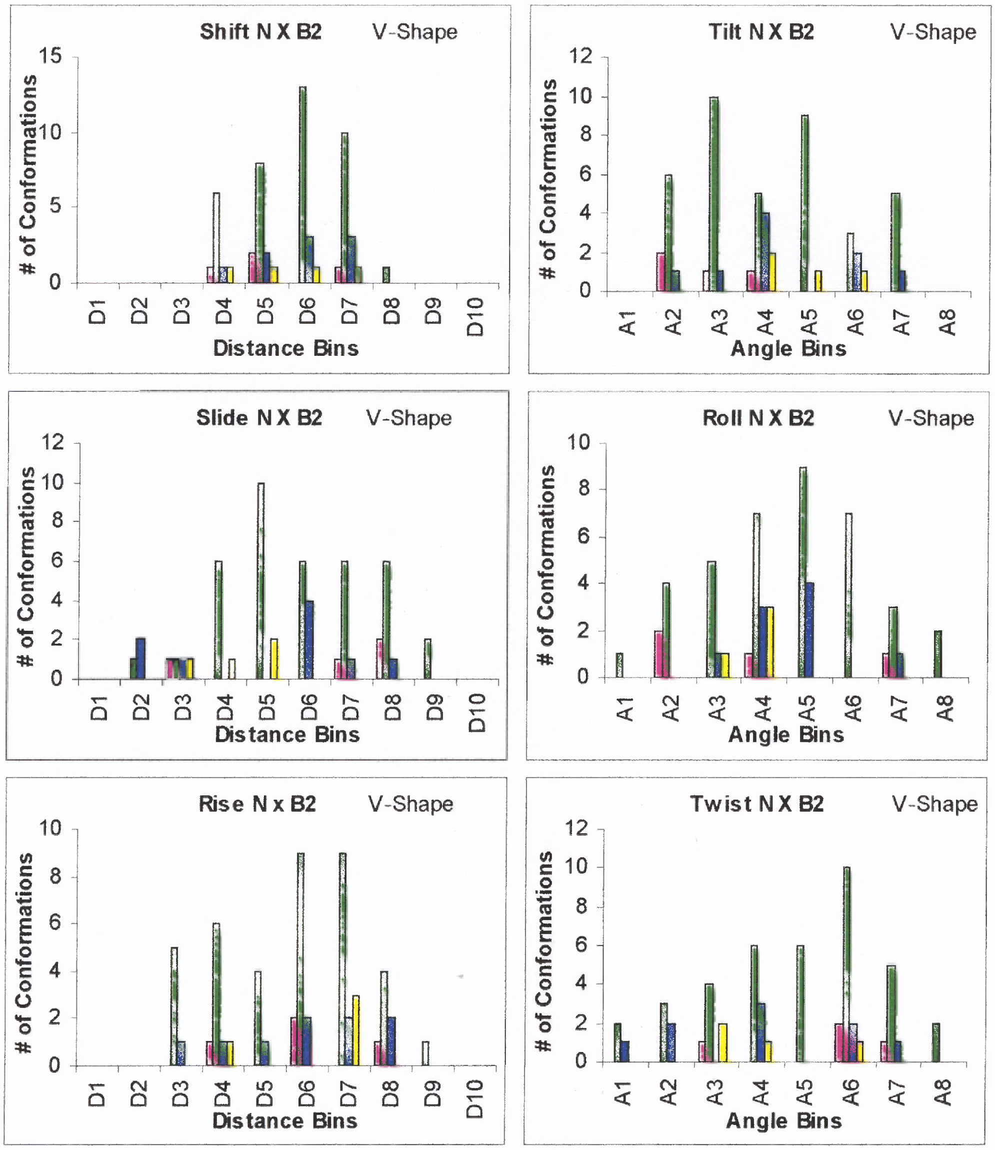
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	🟪 0-4 kcal/mol
D2: -12≤D<-9	A2: -135≤A<-90	🟩 4-8 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟦 8-12 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟨 12-16 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟥 16-20 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.10 Histograms of displacement and rotational parameters for B2 X P in I-Shaped conformations



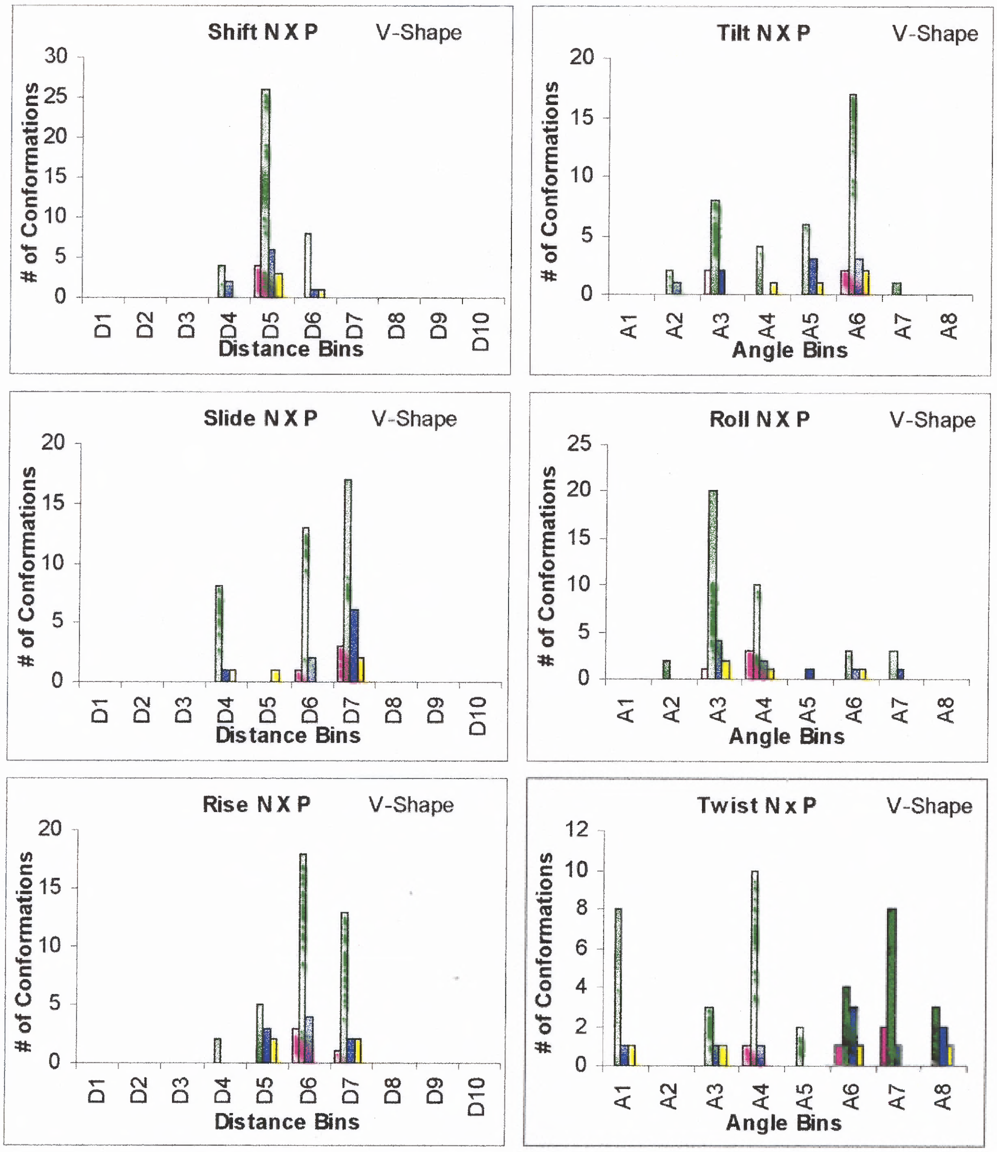
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	🟪 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟩 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟦 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟨 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	🟥 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.11 Histograms of displacement and rotational parameters for N X B1 in V-Shaped conformations



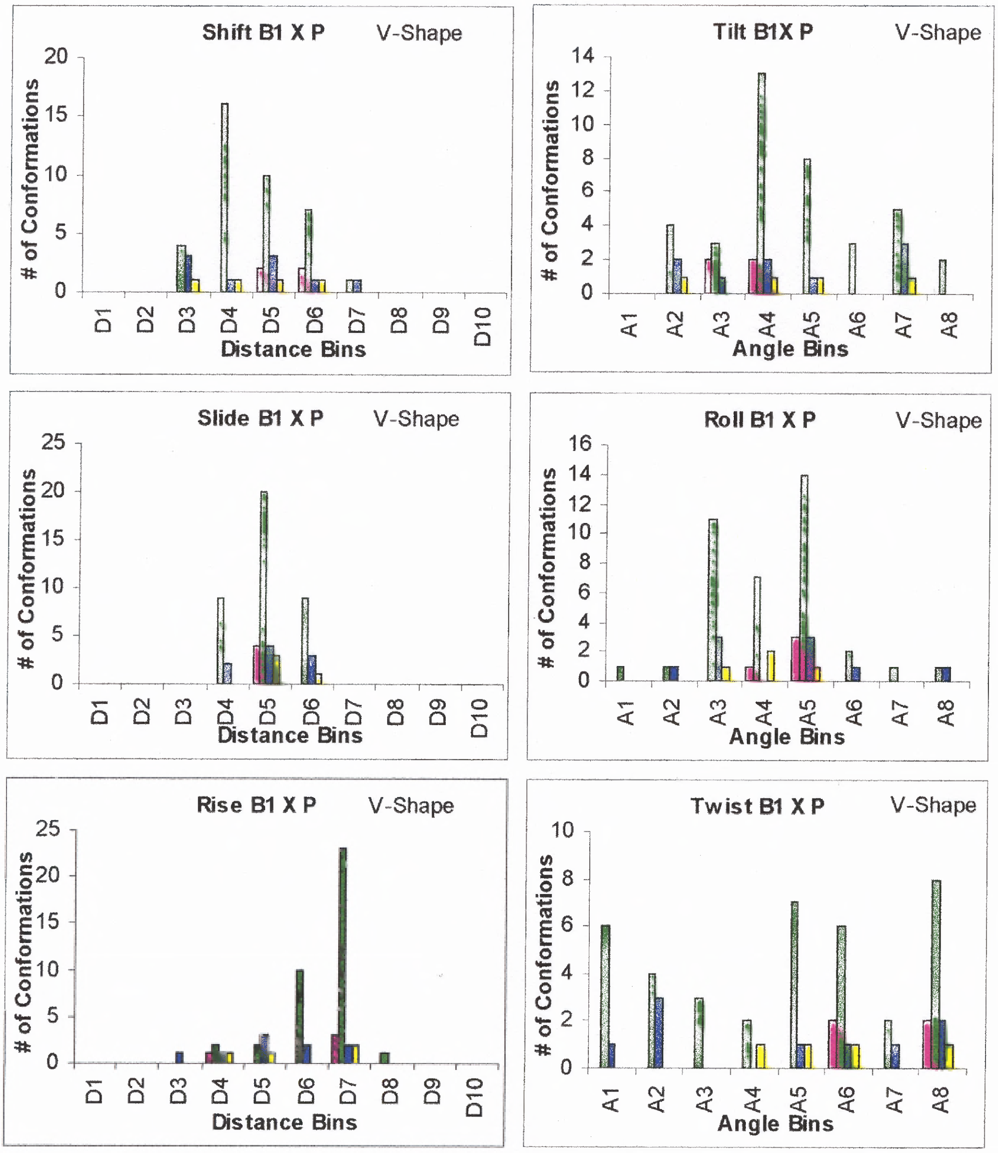
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	🟪 0-4 kcal/mol
D2: -12≤D<-9	A2: -135≤A<-90	▨ 4-8 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟦 8-12 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟨 12-16 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟥 16-20 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.12 Histograms of displacement and rotational parameters for N X B2 in V-Shaped conformations



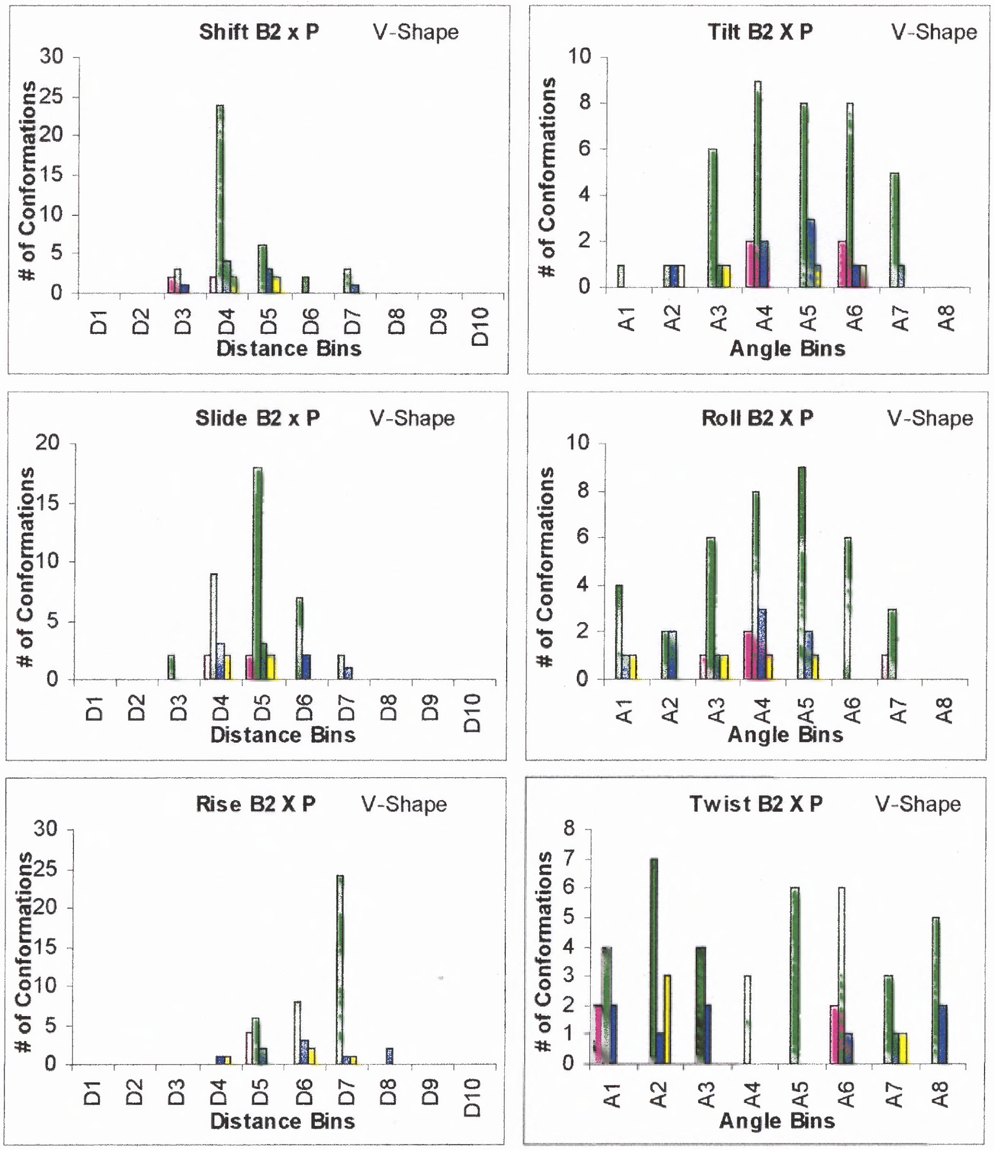
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	🟥 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟩 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟦 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟨 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	▨ 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.13 Histograms of displacement and rotational parameters for N X P in V-Shaped conformations



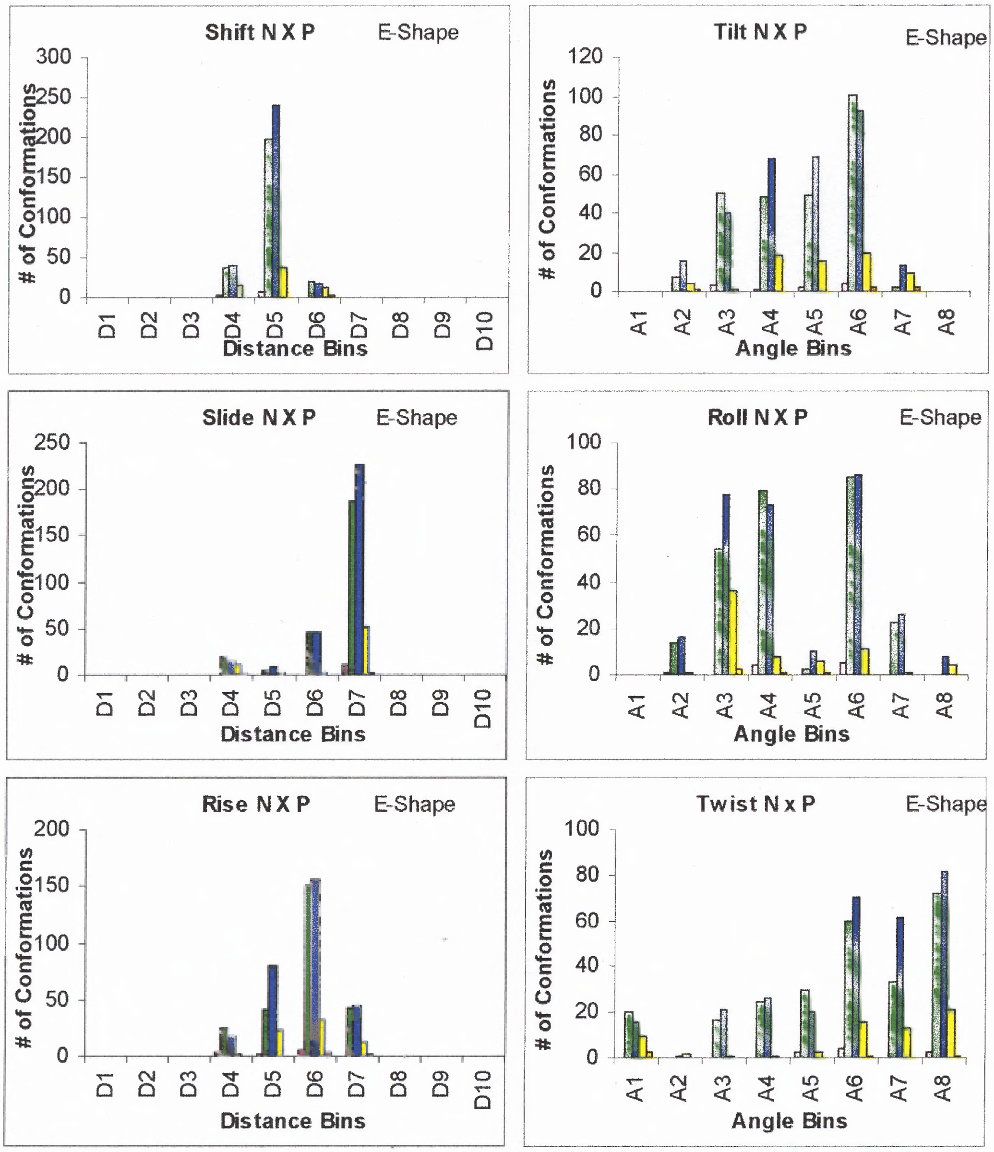
Distance Bins	Angle Bins	Relative Energy
D1: D<-12 D2: -12≤D<-9 D3: -9≤D<-6 D4: -6≤D<-3 D5: -3≤D<0 D6: 0≤D<3 D7: 3≤D<6 D8: 6≤D<9 D9: 9≤D<12 D10: D>=12	A1: -180≤A<-135 A2: -135≤A<-90 A3: -90≤A<-45 A4: -45≤A<0 A5: 0≤A<45 A6: 45≤A<90 A7: 90≤A<135 A8: 135≤A≤180	🟪 0-4 kcal/mol 🟩 4-8 kcal/mol 🟦 8-12 kcal/mol 🟨 12-16 kcal/mol 🟥 16-20 kcal/mol

Figure F.14 Histograms of displacement and rotational parameters for B1 X P in V-Shaped conformations



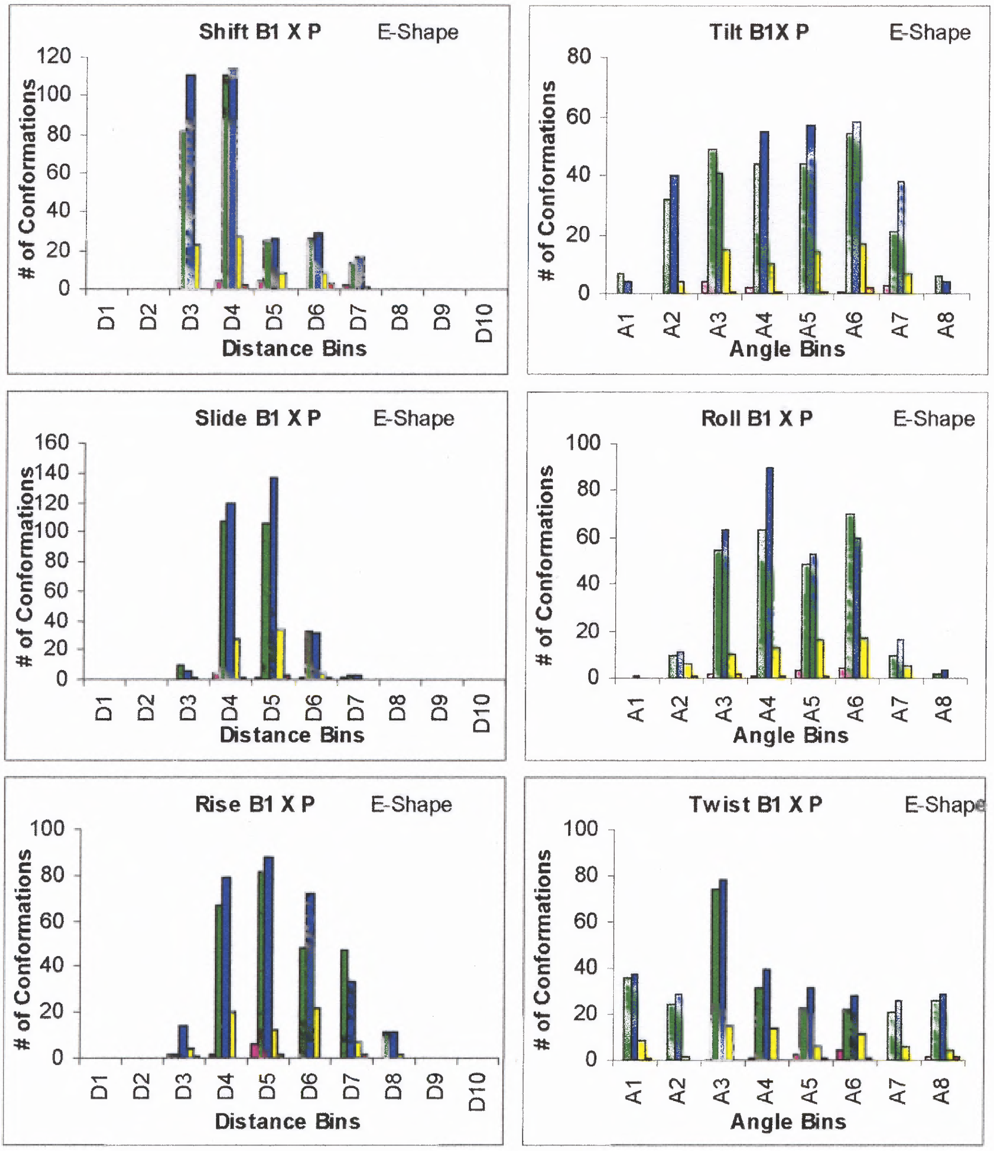
Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	
D2: -12≤D<-9	A2: -135≤A<-90	🟪 0-4 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	🟩 4-8 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	🟦 8-12 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	🟨 12-16 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	🟥 16-20 kcal/mol
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.15 Histograms of displacement and rotational parameters for B2 X P in V-Shaped conformations



Distance Bins	Angle Bins	Relative Energy
D1: D<-12 D2: -12≤D<-9 D3: -9≤D<-6 D4: -6≤D<-3 D5: -3≤D<0 D6: 0≤D<3 D7: 3≤D<6 D8: 6≤D<9 D9: 9≤D<12 D10: D>=12	A1: -180≤A<-135 A2: -135≤A<-90 A3: -90≤A<-45 A4: -45≤A<0 A5: 0≤A<45 A6: 45≤A<90 A7: 90≤A<135 A8: 135≤A≤180	🟪 0-4 kcal/mol 🟩 4-8 kcal/mol 🟦 8-12 kcal/mol 🟨 12-16 kcal/mol 🟥 16-20 kcal/mol

Figure F.16 Histograms of displacement and rotational parameters for N X P in E-Shaped conformations

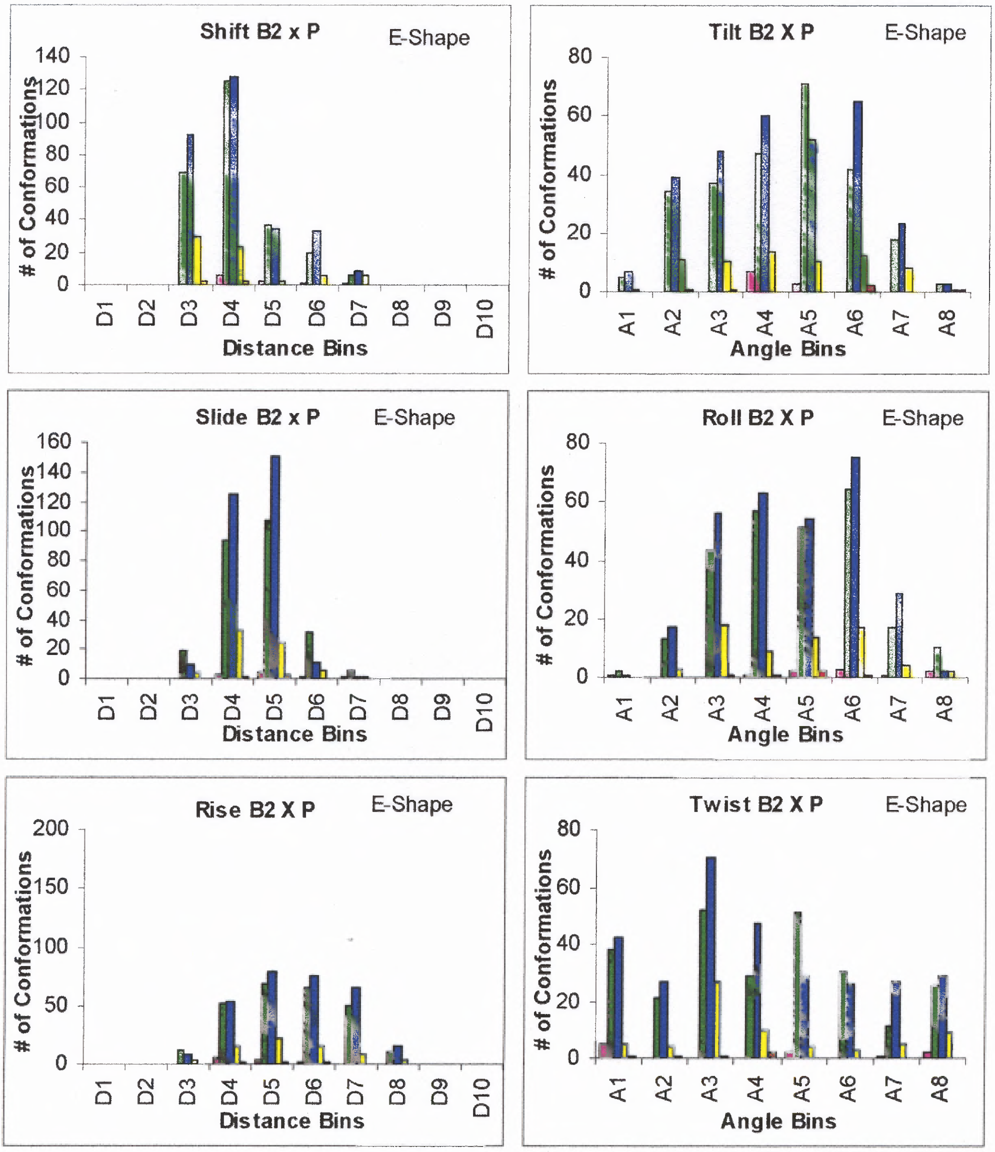


Distance Bins
D1: D<-12
D2: -12≤D<-9
D3: -9≤D<-6
D4: -6≤D<-3
D5: -3≤D<0
D6: 0≤D<3
D7: 3≤D<6
D8: 6≤D<9
D9: 9≤D<12
D10: D>=12

Angle Bins
A1: -180≤A<-135
A2: -135≤A<-90
A3: -90≤A<-45
A4: -45≤A<0
A5: 0≤A<45
A6: 45≤A<90
A7: 90≤A<135
A8: 135≤A≤180

Relative Energy
🟪 0-4 kcal/mol
🟩 4-8 kcal/mol
🟦 8-12 kcal/mol
🟨 12-16 kcal/mol
🟥 16-20 kcal/mol

Figure F.17 Histograms of displacement and rotational parameters for B1 X P in E-Shaped conformations



Distance Bins	Angle Bins	Relative Energy
D1: D<-12	A1: -180≤A<-135	■ 0-4 kcal/mol
D2: -12≤D<-9	A2: -135≤A<-90	■ 4-8 kcal/mol
D3: -9≤D<-6	A3: -90≤A<-45	■ 8-12 kcal/mol
D4: -6≤D<-3	A4: -45≤A<0	■ 12-16 kcal/mol
D5: -3≤D<0	A5: 0≤A<45	■ 16-20 kcal/mol
D6: 0≤D<3	A6: 45≤A<90	
D7: 3≤D<6	A7: 90≤A<135	
D8: 6≤D<9	A8: 135≤A≤180	
D9: 9≤D<12		
D10: D>=12		

Figure F.18 Histograms of displacement and rotational parameters for B2 X P in E-Shaped conformations

REFERENCES

Babcock, M. S., and Olson, W. K. (1994a). "The Effect of Mathematics and Coordinate System on Comparability and "Dependencies" of Nucleic Acid Structure Parameters." *J. Mol. Biol.*, **237**, 98-124.

Babcock, M. S., Pednault, E. P. D., and Olson, W. K. (1994b). "Nucleic Acid Structure Analysis: Mathematics for Local Cartesian and Helical Structure Parameters That Are Truly Comparable Between Structures." *J. Mol. Biol.*, **237**, 125-156.

Bansal, M., Bhattacharyya, D., and Ravi, B. (1995). "NUPARM and NUCGEN: software for analysis and generation of sequence dependent nucleic acid structures." *Compu. Appl. Biosci*, **11**, 281-287.

Bernard, D., Coop, A., and MacKerell, J. (2003). "2D Conformationally Sampled Pharmacophore: A Ligand-Based Pharmacophore to Differentiate δ Opiod Agonists from Antagonists." *J. Am. Chem. Soc.*, **125**, 3101-3107.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Jr, E. F. M., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977). "The Protein Data Bank: a computer-based archival file for macromolecular structures." *J. Mol. Biol.*, **112**, 535-542.

Blow, D. M. (1960). "To Fit a Plane to a Set of Points by Least Squares." *Acta. Cryst.*, **13**, 168-169.

Brint, A., and Willet, P. (1987). "Algorithms for the Identification of three-dimensional maximal common substructures." *J.Chem. Inf. Comput. Sci.*, **27**, 152-158.

Cramer, I., Patterson, D. E., J.D.Bunce, and R.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.

Dickerson, R. E. (1998). "DNA bending: the prevalence of kinkiness and the virtues of normality." *Nucleic Acids Res.*, **26**, 1906-1926.

Dutta, A. K., Zhang, S., Kolhatkar, R., and Reith, M. E. A. (2003). "Dopamine transporter as target for drug development of cocaine dependence medications." *Eur. J. of Phar.*, **479**, 93-106.

G.Chang, Guida, W. C., and W.C.Still. (1989). "An internal coordinate Monte Carlo method for searching conformational space." *J. Am. Chem. Soc.*, **111**, 4379-4386.

Gorelick, D. A. (1998). "The rate hypothesis and agonist substitution approaches to cocaine abuse treatment." *Adv. Pharmacol (NY)*, **42**, 995-997.

Gorin, A. A., Zhurkin, V. B., and Olson, W. K. (1995). "B-DNA Twisting Correlates with Base-Pair Morphology." *J. Mol. Biol.*, **247**, 34-48.

Hassan, M. A. E., and Calladine, C. R. (1995). "The Assessment of the Geometry of Dinucleotide Steps in Double-Helical DNA: A New Local Calculation Scheme." *J. Mol. Biol.*, **251**, 648-664.

Horn, B. K. P. (1987). "Closed-form Solution of Absolute Orientation Using Unit Quaternions." *J. Opt. Soc. A.*, **4**, 629-642.

Kosikov, K. M., Gorin, A. A., Zhurkin, V. B., and Olson, W. K. (1999). "DNA Stretching and Compression: Large scale Simulations of Double Helical Structures." *J. Mol. Biol.*, **289**,1301-1326.

Lavery, R., and Skelner, H. (1989). "Defining the structure of irregular nucleic acids: conventions and principles." *J. Biomol. Struct. Dynam.*, **6**, 655-667.

Lavery, R., and Sklenar, H. (1988). "The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids." *J. Biomol. Struct. Dynam*, **6**, 63-91.

Leach, A. R. (1996). *Molecular modelling: principles and applications*, Harlow, England: Longman.

Lu, X. J., Babcock, M. S., and Olson, W. K. (1998). "Mathematical Overview of Nucleic Acid Analysis Programs." *J. Biomol. Struct. Dynam*, **16**, 833-843.

Lu, X. J., Hassan, M. A. E., and Hunter, C. A. (1997). "Structure and Conformation of Helical Nucleic Acids: Analysis Program (SCHNAaP)." *J. Mol. Biol.*, **273**, 668-680.

Lu, X. J., and Olson, W. K. (1999). "Resolving the Discrepancies Among Nucleic Acid Conformational Analyses." *J. Mol. Biol.*, **285**, 1563-1575.

Lu, X. J., and Olson, W. K. (2003). "3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures." *Nucleic Acids Res.*, **31**, 5108-5121.

Olson, W. K., Bansal, M., Burley, S. K., Dickerson, R. E., Gerstein, M., Harvey, S. C., Heinemann, U., Lu, X.-J., Neidle, S., Shakked Z. ee al. (2001). "A standard reference frame for the description of nucleic acid base-pair geometry." *J. Mol. Biol.*, **313**, 229-237.

Pednault, E. P. D., Babcock, M. S., and Olson, W. K. (1993). "Nucleic acids structure analysis: a users guide to a collection of new analysis programs." *J. Biomol. Struct. Dynam.*, **11**, 597-628.

Prisinzano, T., Rice, K. C., Baumann, M. H., and Rothman, R. B. (2004). "Development of Neurochemical Normalization ("Agonist Substitution") Therapeutics for Stimulant Abuse: Focus on the Dopamine Uptake Inhibitor, GBR12909." *Curr. Med.Chem-Central Nrevous System Agents.*, **4**, 47-59.

Saenger, W. (1984). *Principles of Nucleic acid structure*, Springer-Verlag, New York.

Schomaker, V., Waser, J., and Bergman, R. E. M. G. (1959). "To Fit a Plane or a Line to a Set of Points by Least Squares." *Acta. Cryst.*, **12**, 600-604.

Singh, S. (2000). "Chemistry, Design and Structure-Activity Relationship of Cocaine Antagonists." *Chem. Rev.*, **100**, 925-1024.

Smith, M. P., Hoepping, A., Johnson, K. M., Trzcinska, M., and Kozikowski, A. P. (1999). "Dopaminergic agents for the treatment of cocaine abuse." *Drug Discovery Today.*, **4**, 322-332.

Soumpasis, D. M., and Tung, C. S. (1988). "A rigorous base-pair oriented description of DNA structures." *J. Biomol. Struct. Dynam*, **6**, 397-420.

Tung, C. S., Soumpasis, D. M., and Hummer, G. (1994). "An extension of the rigorous base-unit oriented description of nucleic acid structures." *J. Biomol. Struct. Dynam.*, **11**, 1327-1344.