Report

SVETLANA KIRILLOVA

postdoctoral fellow
project AmoRo

Combined Molecular Mechanics and Motion Planning approach: modeling of protein flexibility and calculations of atom-atom interactions

Laboratory for Analysis and Architecture of System CNRS
7, avenue du Colonel Roche
31077 Toulouse
2005
1 Introduction
   1.1 Motion-planning and Molecular Mechanics approaches for molecular modeling
   1.2 Structure of report

2 Protein flexibility in the docking simulations
   2.1 Experimental determination of protein structures and dynamics
   2.2 Classification of protein motions
   2.3 Computational methods for presentation protein flexibility
   2.4 Molecular Dynamics
   2.5 Normal mode analysis
      2.5.1 Comparison of molecular dynamic and normal mode analysis methods
      2.5.2 Representations of protein motions by normal modes calculated with all-atom model
      2.5.3 Simplified force field
      2.5.4 Representations of protein motions by normal modes calculated with simplified potentials
   2.6 Conclusions

3 Normal modes as collective degrees of freedom for Motion Planning
   3.1 Principle
   3.1.2 Rotation Translation Block (RTB) method
   3.2 Robot motion planning
      3.2.1 Sampling-Based Motion Planning Algorithms
      3.2.2 Applications for Computational Biology
      3.2.3 Geometric View Of Molecular Constraints
   3.3 Combined Motion Planning and Normal Mode Analysis approach
      3.3.1 Main idea of approach
      3.3.2 Collective degrees of freedom in Internal coordinates corresponding to torsion angles
      3.3.3 Representation of protein motions using one collective degree of freedom
      3.3.4 Exploration of multidimensional space of several collective degrees of freedom
      3.3.5 Definition of amplitudes of normal vibrations
      3.3.6 Different ways to represent backbone and sidechains motions of proteins
      3.3.7 Nonlinearity of the protein conformational changes by an iterative usage of normal mode calculations
   3.4 Conclusions

4 Molecular Mechanics block of BioMove3D package
   4.1 Structure of ElNemo package
      4.1.1 Calculations of potential energy of protein using Elastic Network Model
      4.1.2 Transition from Cartesian coordinates to Rotation-Translation coordinates and diagonalization of Hessian matrix
      4.1.3 Additional programs of ElNemo package
   4.2 Interface of Move3D to AMBER
      4.2.1 Input files
      4.2.2 Preparatory part for calculations of atom-atom interactions
4.2.3 Specification of force field
4.2.4 Running of energy calculations, minimization of energy and molecular dynamics calculations
4.2.5 Parameters of energy minimization
4.2.6 Parameters of the molecular potential energy function
4.3 The main capacities of Molecular Mechanics block of BioMove3

References
Appendix A Studies of the dynamical features of macromolecules by normal mode analysis
Appendix B Example: Calculations of energy of atom-atom interactions and energy minimization for molecular system composed of two proteins (~10 000 atoms) using BioMove3d and AMBER
Introduction

1.1 Motion-planning and Molecular Mechanics approaches for molecular modelling

At present the motion-planning approach is successfully used to describe the binding process between protein and ligand [1-3]. This modeling is not computationally expensive and makes possible in reasonable computing time to represent large-amplitude conformational transitions of protein and to calculate ligand trajectories. The molecular construction made of spheres linked by joints is the first stage to create a molecular model in the BioMove3d package. Van der Waals radii of atoms determine the geometrical constraints to avoid the atom clashes. The docking process is treated as motions of two geometrical forms composed by spheres. However, this level of approximation is not sufficient to investigate the changes of molecular structures during the docking process because the atomic motions are determined by the complex system of inter- and intra-molecular interactions. More full representations of a molecule are based on the Molecular Mechanics method. Molecular mechanics techniques take a classical approach to calculate the energy of a structure. The molecule is treated essentially as a set of charged point masses (atoms), which are coupled together with springs. The total energy of a structure is calculated using an analytical function that sums the individual energy terms.

Full flexibility of protein composed by thousands atoms demands the exploration of 3N dimensional space where N is number of atoms in a molecule. One possibility to reduce the number of degrees of freedom is based on the fact that under normal conditions the molecule cannot receive a sufficient amount of energy to deform bonds and bond angles substantially. These structural elements of molecule are considered as rigid and molecular motions are represented as rotations around joints. But number of joints in the protein model is also very high. For example for protein contained ~ 4500 atoms the number of rotational joints corresponding to fully flexible protein is ~1500. Therefore this model affords to regard only partial flexibility of protein and molecular motions have been described as loop fluctuations [1]. The molecular model including the atom-atom interactions makes possible to describe the protein motions by one or several degrees of freedom. In such “collective degree” of freedom all atoms displace simultaneously. It corresponds to an evident fact that an isolated motion of one atom is impossible in a molecule. Collective degrees of freedom may be extracted from Molecular Dynamics trajectory of atoms or may be represented by normal modes of molecule.

1.2 Structure of report

The report consists of four main parts. The first part is the introduction. The second part of report gives the brief review of the computational methods for representation of protein flexibility in the docking simulations. The molecular dynamics and normal mode analysis are two widely used methods, which apply the atom-atom interactions to calculate the protein flexibility. The comparison of these methods explains our choice of Normal Mode Analysis method to model the protein flexibility. This part of report contains the description of simplified models, which are used for the calculation of normal modes of large molecules.

The third part describes the methods used in our investigations: Motion-planning approach, Normal Mode Analysis, Elastic Network Model and Rotation Translation Block approximation. It also represents the combined Motion-planning and Normal Mode Analysis method developed during our investigations. It explains the use of the normal modes as collective degrees of
freedom to represent the protein flexibility. The definition of amplitudes of normal vibrations and iterative normal mode calculations are discussed.

The fourth part of report contains the description of the ElNemo package implementing the normal modes calculations for large molecules using the Tirion potential and building block approximation. It gives the main details of the Move3d-AMBER interface composed of first part related to LEAP module of AMBER and second part related to SANDER module. It describes the ambmov program created for preparation of input pdb files. This part of report also explains the specification of force field and the choice of the parameters determining the minimization of energy and potential energy function.
Chapter 2

Protein flexibility in the docking simulations

Historically the theories of protein-ligand docking have passed from “lock-and-key” theory to “induced fit” [4, 5]. According to the first theory the docking is simply the rigid body geometric problem of fitting together two complementary shapes. According to the second interactions between protein and ligand convert the initial complex of two molecules to a more stable conformation during the binding process. At present the explanation of docking process termed “preexisting equilibrium” is used [6]. According to this theory it is very important to examine different forms of protein because a ligand can bind to a less populated conformation. Therefore current methods for computational modeling of protein-ligand binding treat the receptor as a flexible object [4, 5, 7, 8].

2.1 Experimental determination of protein structures and dynamics

The most popular experimental methods for determination of protein structure are X-ray crystallographic analysis and nuclear magnetic resonance (NMR) [9]. However using traditional X-ray and NMR methods it is possible to determine only the average structure of protein. The crystallographic value, B-factor corresponds to the thermal oscillations of atoms in the crystal [10]. Some insight into the time scale of protein motions has been obtained from Raman spectroscopy [11], Mossbauer absorption spectroscopy and incoherent neutron scattering [12]. The most detailed information about dynamics properties of proteins has been determined by the complex methods combining experimental NMR techniques [13], circular dichroism and fluorescence measurements [14] with molecular dynamics simulations.

2.2 Classification of protein motions

It was established that the native state of proteins must be considered as a heterogeneous ensemble of conformations that interconvert on the picosecond to nanosecond timescale [13]. The side chains occupy multiple rotametric states and can characterized as having liquid-like characteristics. However, the backbone can be considered as having the solid-like rigidity that makes possible to determine certain motions of proteins. The systematization of protein motions is represented in the Database of Macromolecular Movements created by Gerstein group [15, 16]. The database contains detailed descriptions of a few hundred motions. According the data given in the paper [16] the classification of protein motions is follow:

- Hinge 45%, shear 14%, partial refolding 4%, allosteric 7%, other/non-allosteric 7%, unclassifiable 20%, and notably motionless (on the basis of packing).
- Domain 62%, fragment 22 %, subunit 11%, and complex 5% (on the basis of size);

More than half of motions are the hinges and shears (Fig.2.1) that are caused by the tight packing of atoms inside of protein. During such movement the atoms can displace without colliding with neighboring atoms. The protein structure can formed of relatively large domains, which change only little during the conformational transitions. A number of shear motions have to be
concatenated to produce a large overall motion because the individual shear motion is very small whereas the hinge motion of protein is less restricted and more localized.

![Figure 2.1 Schematic representation of shears and hinge motions](image)

### 2.3 Computational methods for presentation protein flexibility

The large number of computational methods used to describe the protein flexibility is caused by complexity of a problem [4, 5, 7, 8]. Protein flexibility is a result of high dimensionality of the system. The simplest methods represent the flexible molecule using an ensemble of conformational states, which may be received by the experimental methods or be generated computationally. Unfortunately, the opportunities of experimental methods are limited. The inaccuracies of computational methods for generation of multiple protein structures may be compensated by a huge number of protein forms in the ensemble [17]. The earliest attempts to accommodate small changes in form of protein were through the use of implicit methods with a “soft” scoring function [18]. More complex methods treat the protein molecule as mechanical system composed by rigid parts (domains) connected by flexible joints [19] or consider the protein flexibility as loop fluctuations [1]. They predict successfully the large-scale motions in proteins. However, an artificial division of the molecule into rigid and flexible parts moving independently may result in the protein motions very far from real because the movement of each atom in a molecule is determined by the complex system of intramolecular interactions. The most rigorous computational methods to reproduce the molecular motions are molecular dynamics (MD) and normal mode analysis (NMA) that calculate the trajectories of atoms using Newton's Second Law and potential energy functions of atom-atom interactions.

### 2.4 Molecular Dynamics
The MD method is widely used to simulate the molecular motions [20]. The potential energy function of the method can be composed for several molecules that makes possible to investigate not only one molecule but also the complex of molecules. The limitations of MD method are imposed firstly by the approximate nature of the force field that can produce a wrong random fluctuations; secondly by the relatively short time scale (of the order of a nanosecond) computationally accessible whereas small proteins have been shown to fold completely in tens of microseconds [21]. Several methods have been developed to follow the transition between the open and closed structures on the nanosecond time scale. One of these, targeted MD, was used to calculate the transition pathway of the bacterial chaperonin GroEL and to obtain the detailed information about the changes of molecular structure which has been impossible to determine by experimental techniques [22].

The MD trajectories of atoms are very complicated therefore the special methods for their analysis are used in the studies of the protein motions. They include principal component analysis [23-26], singular value decomposition [2, 3], and Fourier transforms techniques [27].

The collective degrees of freedom corresponding to the principal components are very convenient way to describe the protein flexibility. For example, they are used to reduce the dimension of the conformational space of proteins defined as (3N-6), where N is a number of atoms in the molecule [2,3]. Because the fist principal components of 2ns MD trajectory of HIV-1 protease include the main part of total motions it reduces the number of degrees of freedom from thousands to less than 10 [3]. However, the first principal components determine only the directions of the largest fluctuations of atomic positions in the MD trajectory and using them as collective degrees of freedom it is impossible to reproduce a quite realistic protein motions received by time consumed MD method.

2.5 Normal mode analysis

The potential energy function of NMA method has a harmonic form with one minimum and the normal modes can be calculated only for an isolated molecule.

2.5.1 Comparison of molecular dynamic and normal mode analysis methods

The more complicate MD method has the following advantages in comparison with NMA:

- The MD collective trajectory of atoms is more plausible and it includes the conformational transitions between different local minima whereas the protein motions received by NMA are the oscillations of atoms around one minimum.
- In NMA the solvent effects cannot be take into account explicitly whereas MD is known as one of the best methods to determine the influence of solvent on the internal motions of protein [20].

In the ideal case of harmonic form of potential energy function with one minimum the trajectories received by MD and NMA coincide. Therefore the detailed investigation of potential energy landscape of proteins has allowed to compare two methods and to reveal harmonic and anharmonic aspect in protein dynamics [23-25]. Thus as a results of principal component analysis and jumping-among-minima model analysis of 50 ps MD trajectory of human lysosyme, protein motions are shown to consist of three types of collective modes, multiply hierarchical modes, singly hierarchical modes, and harmonic modes (Fig. 2.2) [23]. The total number of calculated modes is 6117. Multiply hierarchical modes, the number of which accounts only for
0.5% of all modes, dominantly determine the total atomic fluctuations reproducing the hinge-bending motions of human lysosyme.

![Diagram](image)

**Figure 2.2** Free energy surface is schematically shown for (a) multiply hierarchical modes, (b) singly hierarchical modes, and (c) harmonic modes

Dauber-Osguthorpe et al. have been obtained surprisingly high agreement between the resulting motions received by MD and NMA methods for chymotrypsin-like serine protease [27]. M. Zacharias has carried out the comparison of two methods for RNA molecule [26]. Although quantitative differences of results were found, qualitative pattern of atom position and helical descriptor fluctuations along the sequence was similar in both methods.

### 2.5.2 Representations of protein motions by normal modes calculated using all-atom model

The normal modes calculated using atom-atom potentials were applied to reproduce the protein motions in the crystals were the thermal atomic fluctuations related to a crystallographic structure are observed [27].

Zacharias and Sklenar were used the harmonic modes corresponding to eigenvectors with lowest eigenvalues of the Hessian matrix calculated using GROMOS87 force field as independent variables to describe the receptor flexibility in the simulation of the binding of a ligand to DNA [28]. It was indicated that the application of harmonic mode relaxation approach was suitable due to relatively small (~ 1 Å) RMSD between the complexed and free form for investigated system. Kolossvary and Guida also applied the low-frequencies normal modes to obtain flexible protein binding sites in the docking simulation to study a ligand binding to PNP [29]. They used AMBER force field and the special conformational search procedure LMOD.

Cui et al. have been applied normal modes calculated with all-atom model to explore the inherent flexibility of the α, β and γ subunits of F1-ATPase in isolation and as part of αβγ complex [30]. The structurally relatively rigid and flexible regions of subunits and their conformational changes determining the functional cycle of F1-ATPase were revealed.
2.5.3 Simplified force field

The NMA with all-atom model is applied successfully to study the high-frequency localized motions of proteins and the low-frequency motions with relatively small amplitudes, which correspond to the dynamics around an energy minimum [31]. However, this method is not suitable to represent global movements of large domains.

- “First, global domains motions haven no (or very little) energy contribution from internal degree of freedom of the domains because there is no deformation.
- Second, the long-range interactions between domains are weaker than the short-range interactions between neighboring atoms” [32].

Therefore, the simplified force field or Elastic Network Model should be used for representation of large amplitude motions of protein. Within the frame of the approach proposed by Tirion [34], the standard detailed potential energy functions is replaced by

\[ E_p = \sum_{d_{ij} < R_c} c(d_{ij} - d_{ij}^0)^2 \]  

(2.5.3.1)

where \( d_{ij} \) is the distance between atoms \( i \) and \( j \), \( d_{ij}^0 \) being the distance between these two atoms in the given studied crystallographic structure. The sum is restricted to atom pairs separated by less than \( R_c \), which is an arbitrary cut-off parameter; only the Cα atoms have been taken into account. Such model can be interpreted as representing the potential as a function of the residue positions (Fig 2.3). Schematically the potential energy function of simplified force field can be interpreted as the smoothed form of the all-atom energy function (Fig. 2.4). Experimental values of generic force constants \( c \) in equation (2.5.3.1) may become available soon from new neutron scattering experiments [34].

The calculations of normal modes of large molecules as proteins using standard methods are problematic. The huge sum of atom-atom potentials of a protein yields a function with a large number of narrow minima. The normal modes must be calculated in the global minimum because the equilibrium positions of the atoms in the normal vibrations must be in the condition of rest determined by the lowest energy corresponding to the zero absolute temperature. Therefore, using NMA with a potential energy function including atom-atom interactions of a protein it is necessary to find the global minimum in which the second derivative of potential energy (Hessian) is calculated. The diagonalization of a Hessian matrix also is a very difficult task caused by high rank of Hessian matrix (3Nx3N). The use of memory of NMA is defined as \( O(N^2) \) and CPU time as \( O(N^3) \).

The use of simplified model of protein resolves two significant problems of MNA calculations. First, this is the problem of global minima. The energy function given by the equation is in a minimum for any chosen configuration of any system. Second, this is a memory requirement for storing the Hessian matrix and a time for its diagonalization. The rank of Hessian calculated with simplified model is determined by number of residues instead of number of atoms.
Figure 2.3 Tirion or Elastic Network Model of protein.

Figure 2.4 A schematic one-dimensional view of the potential energy surface of a protein showing two kinds of harmonic approximations: an approximation to a local minimum, and an approximation to the smoothed-out potential well [32].

2.5.4 Representations of protein motions by normal modes calculated with simplified models

In the seminal work of Tirion [33] in which simplified potential was introduced, the tests performed on a periplasmic maltodextrin binding protein were demonstrated that a single-parameter model can reproduce complex vibrational properties of macromolecular systems.
Hissen [31] compared the results obtained by simplified force field and CHARMM19, AMBER94 force fields for three well-studied proteins, crambin, lysosyme, and aspartate transcarbamylase. In further work, Tama and Sanejouand [36] used the conformational transitions of proteins determined as differences between closed and open crystallographic structures to validate the simplified model. It was shown that the simplified potentials reproduce the experimental data for five proteins of various sizes better than the standard detailed potentials. Alexandrov et al. [37] carried out the same test for 377 proteins for which the conformational changes are represented in the Database of Macromolecular Movements. Their results suggest that the lowest-frequency normal mode can identify the most mobile parts of molecules and predict the direction of conformational transitions occurring in proteins. A Web tool NMA movie generator or MolMovDB (available from http://molmovdb.org/nma) was developed for motions prediction by NMA of proteins with simplified force field. The Web server ElNemo accessed at http://igs-server.cnrs-mrs.fr/elnemo also was developed for protein movement analysis and the generation of templates for molecular replacement using simplified model [38]. The Web servers WEBnm@ http://www.bioinfo.no/tools/normalmodes [39], MoVies http://ang.cz3.nus.edu.sg/cgi-bin/prog/norm.pl [40], and Web server created by Delarue group http://lorentz.immstr.pasteur.fr/nma also contain the possibility to perform normal mode calculations of proteins.

In their review Karplus and McCammon [20] have been indicated that more global MD simulations are likely to be initiated with less detailed models, which was used already for normal mode calculations. The comparison of the major modes of MD trajectory of HIV-1 protease obtained with all-atoms and simplified models with Go potential was shown very small differences between them [3].

At present the applications of NMA with simplified potentials cover wide areas of structural biology:

• The study of protein conformational changes upon ligand binding [35, 36, 41].
• The study of membrane channel opening and closure [42].
• The analysis of structural movement of the ribosome [43].
• Viral capsid maturation [44].
• Transconformations of the SERCA1 Ca-ATPase [45, 46].
• The analysis of domain motions in large proteins in general [31, 47]
• The study of the global motions of subunits in $\alpha_3\beta_1\gamma$ complex of F$_1$-ATPase whereas the detailed force field was used to study more local motions of subunits [30].

2.7 Conclusions

The internal motions of proteins are strongly influenced by anharmonic effects and solvent damping [32]. Nevertheless, the harmonic aspect in protein dynamics is sufficiently large and NMA is widely used to study the protein motions. The examples of usage of normal modes for representation of protein dynamics are summarized in Table (see Appendix A). The all-atom model reproduces mainly small-scale local motions in proteins whereas the simplified model reproduces the backbone motions. Evidently, the simulation of some molecular processes requires an atomistic approach. But it is possible to predict that the “dynamic selectivity” [20] of proteins is related to the backbone motions rather than to the side chains motions characterized by high uncertainty [13]. Therefore the NMA method with simplified model was chosen for our investigations.
Chapter 3

Normal modes as collective degrees of freedom for Motion Planning

3.1 Normal Mode calculations

3.1.1 Principle

The classical Lagrangian for the vibrations of a protein with \( N \) atoms is given by

\[
L = T - V,
\]

where \( V \) is the potential energy describing interactions among atoms, and \( T \) is the kinetic energy

\[
T = \frac{1}{2} \sum_{i=1}^{N} m_i \left( \dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2 \right),
\]

where the dot notation has been used for derivatives with respect to time. The above expression for \( T \) can be rewritten by introducing mass-weighted Cartesian displacement coordinates. Let

\[
q_i = \sqrt{m_i} \cdot (x_i - x_{eq}), \ldots, q_{3N} = \sqrt{m_N} \cdot (z_N - z_{eq}),
\]

in which the \( q_i \) coordinate is proportional to the displacement from the equilibrium value \( q_{eq} \). Expanding potential energy in Taylor series and neglecting all terms with powers greater than two (harmonic approximation), potential energy will assume the form

\[
V = \frac{1}{2} \sum_{i,j=1}^{3N} f_{ij} q_i q_j,
\]

where \( f_{ij} \), the force constants, are defined as the second derivatives of the potential energy function:

\[
f_{ij} = \frac{\partial^2 V}{\partial q_i \partial q_j}.
\]

Substituting \( T = \frac{1}{2} \sum_{i=1}^{3N} \dot{q}_i^2 \) and \( V = \frac{1}{2} \sum_{i,j=1}^{3N} f_{ij} q_i q_j \) in the Lagrange’s equation of motion

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0,
\]

we obtain

\[
\ddot{q}_i + \sum_{j=1}^{3N} f_{ij} q_j = 0, \quad i=1..3N,
\]

or in matrix form

\[
\ddot{\mathbf{q}} + \mathbf{Fq} = 0.
\]

This is the set of \( 3N \) coupled 2nd order differential equations with constant coefficients. It can be solved by assuming a solution of the form

\[
q_i = A_i \cos(\omega t + \varphi).
\]

This substitution converts the set of differential equations into a set of \( 3N \) homogeneous linear equations:

\[
\sum_{j=1}^{3N} f_{ij} A_j - \omega^2 A_i = 0,
\]

or

\[
\mathbf{F} \ddot{\mathbf{A}} - \omega^2 \mathbf{A} = 0.
\]
This problem may now be solved with any eigenvector/eigenvalue solution method. One of the simplest is to attempt to diagonalize the matrix $F$ and extract the eigenvalues from the diagonal. It turns out that six eigenvalues of $F$ are zero for a non-linear molecule. This result can be expected from the fact that there are three degrees of freedom associated with the translation of the center of mass, and three with rotational motion of the molecule as a whole. Since, there is no restoring force acting on these degrees of freedom, their eigenvalues are zero.

Associated with each eigenvalue is a coordinate, called normal mode coordinate $Q_i$. The normal modes represent a set of coordinates related to the old one by an orthogonal linear transformation $U$:

$$Q = Uq,$$  \hspace{1cm} (3.1.1.10)

Such that the transformation matrix $U$ diagonalizes $F$:

$$UFU^T = \Lambda \text{ (diagonal)}.$$  \hspace{1cm} (3.1.1.11)

This transformation has a deep impact on the resulting form of the differential equations: (7b) transforms to

$$\ddot{Q} + \Lambda Q = 0,$$  \hspace{1cm} (3.1.1.12)

but since $\Lambda$ is diagonal, the equations (12) are effectively decoupled:

$$\ddot{Q}_i + \lambda_i Q_i = 0, \ldots, \ddot{Q}_{3N} + \lambda_{3N} Q_{3N} = 0,$$  \hspace{1cm} (3.1.1.13)

and the system therefore behaves like a set of $3N$ independent harmonic oscillators, each oscillating without interaction with the others.

It is of considerable importance to examine the nature of the above solutions. It is evident from Eq. (3.1.1.8) that each atom is oscillating about its equilibrium position with the same frequency and phase for a given solution $\omega_k$. In other words, each atom reaches its position of maximum displacement at the same time, and each atom passes through its equilibrium position at the same time. A mode of vibration having all these characteristics is called a normal mode of vibration, and its frequency is known as a normal mode frequency.

### 3.1.2 Rotation Translation Block (RTB) method

The sum of the potential energy of protein (Eq. 2.5.3.1) can include all atoms of a molecule. The building block approximation proposed by Tama [35] is applied for diagonalization of Hessian matrix received by all-atom model. The block is one or more residues which are regarded as rigid body (Fig. 3.1).

- block = 1 or several residues
- rotation + translation of block => new basis
- expression of Hessian in this new basis
- diagonalization of a matrix $6nB*6nB$

![Figure 3.1 Schematic representation of Rotation Translation Blocks (RTB) method [35]](image-url)
In standard approaches, the normal modes of a system are obtained through the diagonalization of the Hessian matrix $H$, that is, the $3N \times 3N$ matrix of the second derivatives of the potential energy with respect to the mass-weighted coordinates, where $N$ is the number of atoms of the system. In the RTB approach, $H$ is first expressed in a basis defined by the rotations and translations of $n_b$ blocks, $H_b$, the projected Hessian, being given by:

$$H_b = P^T HP$$  \hspace{1cm} (3.1.2.1)$$

where $P$ is an orthogonal $3N \times 6n_b$ matrix built with the vectors associated with the local rotations and translations of each block. Approximate low-frequency normal modes of the system, thus, are obtained by diagonalizing $H_b$, a matrix of size $6n_b \times 6n_b$, the corresponding $(3N)$ atomic displacements being obtained as:

$$A_p = PA_b$$  \hspace{1cm} (3.1.2.2)$$

where $A_b$ is the matrix of the eigenvectors of $H_b$.

The final objective is the treatment of huge macromolecules. In the first step, the blocks of residues are defined and the six rotation and translation modes of each block, $U_i$, are determined and stored. These $6n_b$ vectors form the new basis of small dimension that corresponds to the projector $P$. In the second step, the Hessian matrix is expressed in this new basis, separately for each coupling or diagonal block, $H_{ij}$

$$H_{ij}^b = U_i^T H_{ij} U_j$$  \hspace{1cm} (3.1.2.3)$$

The set of $n_i^2 H_{ij}^b$ block-matrices forms the matrix $H_b$. Thus, only the small dimension vectors $U_i$ and the small $6n_b \times 6n_b H_b$ matrix have to be stored. In the last step, $H_b$ is diagonalized with standard methods.

### 3.2 Robot motion planning

Motion planning is a classical problem in robotics [54]. It consists in computing feasible motions for articulated robots in workspaces cluttered with obstacles. In recent years, motion planning techniques have undergone considerable development and have been successfully applied to challenging problems in diverse application domains, including computational biology.

#### 3.2.1 Sampling-Based Motion Planning Algorithms

Sampling-based motion planners have been designed for exploring constrained high-dimensional spaces. Most of them are variants of the Probabilistic RoadMap (PRM) approach [ ]. The basic principle of PRM is to compute a connectivity graph (the roadmap) encoding representative feasible paths in the search space (e.g. the molecular conformational space). Nodes correspond to randomly sampled points that satisfy feasibility requirements (e.g. collision freeness) and edges represent feasible subpaths computed between neighbouring samples. Once
the roadmap has been constructed, it is subsequently used to process multiple motion queries or to determine ensemble properties of mobility. Variants of the PRM framework have been designed for solving single-query problems without preprocessing the complete roadmap. For example, the *Rapidly-exploring Random Trees* (RRT) algorithm \[56\] expands random trees rooted at the query positions and advancing toward each other through the use of a greedy heuristic. Such variants are well suited to highly constrained problems for which the solution space has the shape of a long thin tube. Whereas constructing a roadmap within the tube would require a high density of samples, the random tree variant benefits from the shape of the tube to naturally steer the expansion.

### 3.2.2 Applications to Computational Biology

Recently, PRM-based algorithms have been successfully applied to study molecular motions involved in biological processes such as protein-ligand interactions \[57, 58\], protein folding \[59, 60\] and also RNA folding \[61\]. The main difference in the molecular adaptation of the PRM framework is that the binary collision detection, used in robotic applications, is replaced by a molecular force field. Sampled conformations are accepted on the basis of their potential energy and roadmap edges are weighted according to their energetic cost. Although the framework is general enough to use any molecular force field, the techniques above generally consider simple potentials (including van der Waals and electrostatic terms) for sake of efficiency. The major strength of these sampling-based techniques is their ability to circumvent the energy trap problem encountered by classical simulation techniques which waste a lot of time trying to escape from the local minima of the molecular energy landscape. Singh *et al.* \[57\] and Apaydın *et al.* \[58\] showed promising results from the study of binding sites for flexible ligands, assuming a rigid model of the protein to limit the dimension of the conformational space. Protein flexibility, which plays an important role in protein-ligand interactions however, is considered for protein folding applications using simplified models such as articulated backbone with bounding spheres for the side-chains \[59\] or a vector-based approximation of secondary structure elements \[60\]. J. Cortes *et al.* \[1\] have shown the efficiency of a geometric treatment of molecular constraints for modelling large-amplitude motions and flexible molecular models.

### 3.2.3 Geometric View of Molecular Constraints

**Molecular degrees of freedom.** Molecular mechanics force fields consider bonded and non-bonded atomic interactions separately. Bonded interactions concern the variation in the relative position of bonded atoms which is usually given in internal coordinates: bond lengths (stretching), bond angles (bending) and dihedral angles (torsion). Slight variations in bond lengths and bond angles from their ideal values produce a large increase in energy. Due to the stiffness of these two terms, both parameters are generally kept constant and the molecular chain is considered as an articulated mechanism with revolute joints modeling bond torsions.

**Loop closure constraints.** In many studies, the global molecular architecture is known and only segments of the molecular chain (loops and unstructured segments not involved in the secondary elements) are possible flexible elements. The first and last atoms of these flexible segments must remain bonded to the fixed neighboring atoms in the chain. Thus, kinematic loop
closure constraints are introduced in the molecular chain. They reduce the subset of feasible conformations of the articulated molecular model. Similar constraints also appear in cyclic molecules and in the presence of disulphide bonds.

**Main repulsive constraints.** For non-bonded interactions, the repulsive part of the van der Waals term is the most important contributor. A large amount of energy is required to get two non-bonded atoms significantly closer than the van der Waals equilibrium distance. Thus, acceptable conformations must respect geometric constraints for steric clash avoidance.

**Main attractive constraints.** Conversely, two important attractive interactions are responsible for the globular shape of macromolecules and strongly participate to molecular docking: the hydrophobic interactions and the hydrogen bonds. They restrain the relative locations of the involved atoms and therefore imply geometric (distance/orientation) constraints on the molecular model.

**Geometric Molecular Model.** The algorithms deal with all-atom models of molecules, which are considered as articulated mechanisms with atoms represented by spheres. Groups of rigidly bonded atoms form the bodies and the articulations between bodies correspond to bond torsions. A Cartesian coordinate frame is attached to each group. The relative location of consecutive frames is defined by a homogeneous transformation matrix which is a function of the rotation angle between them. We follow a similar method to that of Zhang and Kavraki [62] to define such frames and matrices between rigid groups. Figure 3.2 shows the mechanical model for an amino acid residue. The rigid secondary structure elements (alpha helices and beta sheets) are modeled as rigid groups of backbone atoms with articulated side-chains. Since secondary structure elements are fixed in the model, loop closure constraints are introduced in the in-between segment. Similar closure constraints can also be introduced in the model of a molecule to consider non-bonded interactions, such as hydrogen bonds, that impose the spatial proximity between some atoms of the protein.

![Figure 3.2 Mechanical model of an amino acid residue (phenylalanin).](image)

It is composed of five rigid bodies, classified in:

- backbone rigids: $R_{b1} = \{N\}, R_{b2} = \{C_\alpha, \ C_\beta\}, R_{b3} = \{C, 0\}$
- side-chain rigids: $R_{s1} = \{C_\gamma\}, R_{s2} = \{C_\delta1, \ C_\delta2, \ C_\epsilon1, \ C_\epsilon2, \ C_\zeta\}$

The rotations between rigid atom groups are $\phi$ and $\psi$ for the backbone, and $\gamma_1$ and $\gamma_2$ for the side-chain.
3.3 Combined Motion Planning and Normal Mode Analysis approach

3.3.1 Main idea of the approach

Within the frame of the harmonic approach the molecular motions at physiological temperatures can be decomposed on normal modes with low frequencies (under 30 cm\(^{-1}\)). The normal vibrations of protein molecules with such frequencies are collective movements in which almost all atoms of molecule participate. The eigenvectors received by Normal Mode Analysis contain the relative amplitudes and directions of motions for each atom. In normal mode corresponding eigenvalue \( \lambda_i \) all atoms oscillate at the same frequency, \( \nu_i \) (\( \nu_i = \sqrt{\lambda_i} / 2\pi \)). In other words, each atom reaches its position of maximum displacement at the same time, and each atom passes through its equilibrium position at the same time. Therefore the normal mode with low frequency can be regarded as collective degree of freedom in which all atoms in protein displace simultaneously (Fig. 3.3). Such collective degrees of freedom reduces the dimension of space explored by Motion Planning techniques applicable to molecular modeling form thousands to units.

Figure 3.3 Collective degree of freedom corresponding to low frequency normal mode of aspartate transcarbamylase [48]
3.3.2 Collective degrees of freedom in Internal coordinates corresponding to torsion angles

Usually the main part of calculations using force fields are carried out in Cartesian coordinates and eigenvectors corresponding to normal modes and principle components extracted from Molecular Dynamics trajectory represent the atom displacements in Cartesian coordinates. The using collective degree of freedom in these type of coordinates is problematic for large amplitude conformational changes of proteins, because it produces no real distortions of molecular structure [29].

Normal modes can be calculated in any set of coordinates, not only in the commonly used Cartesian coordinates [31]. To obtain normal modes in a set of coordinates \( q_i, i = 1 \ldots 3N \), the transformation matrix \( C \) between the differentials of these coordinates and those of mass of the mass-weighted Cartesian coordinates \( x_i^*, i = 1 \ldots 3N \), must be calculated.

\[
x_i^* = m_i^{1/2} (x_i - x_i^0)
\]  
(3.3.2.1)

where \( m_i \) is mass of atom, \( x_i \) and \( x_i^0 \) are Cartesian coordinates determining displaced and equilibrium positions of atoms. The transformation matrix is defined by

\[
C_{ij} = \frac{\partial q_i}{\partial x_j^*}
\]  
(3.3.2.2)

Any complete set of coordinates will yield the same normal modes. However, it is possible to leave out some of coordinates \( q_i \), which is physically equivalent to keeping the corresponding degrees of freedom fixed. After a transformation to internal coordinates it is possible to eliminate bond and bond angle coordinates, leaving only the dihedrals. When the atomic displacements are sufficiently small, the dependency between Cartesian coordinates and torsion angles is linear. The eigenvectors in coordinates corresponding to the torsion angles may be defined as the difference between torsion angles calculated for equilibrium and displaced positions of atoms.
3.3.3 Representation of protein motions using one collective degree of freedom

It has been shown that main part of the known protein movements can be modelled by displacing the studied structure using at most two low-frequency normal modes [52]. For example, the 9th normal mode of lipase can reproduce the transition from open to closed form (Fig. 3.4). This structural change correlate with the conformational transition received using time expensive molecular dynamics method [63].

![Figure 3.4](image)

*Figure 3.4* Transition from open (i) to closed (ii) form of lipase corresponding to the displacement along 9th normal mode.

3.3.4 Exploration of multidimensional space of several collective degrees of freedom

Exploration of multidimensional space of several collective degrees of freedom makes possible to describe protein movements corresponding to linear combination of several normal modes with low frequencies.

An important mathematical property of normal modes: the normal mode vectors $q_i$, being the eigenvectors of a matrix, form a basis of the 3N-dimensional configuration space of protein [32]. Therefore, any vector $d$ in configuration space and thus any type of motion can be written as a superposition of normal mode vectors with suitable prefactors $p_i$ which are the projections of $d$ onto mode $i$. Mathematically, the projection are defined by

$$p_i = d \cdot q_i$$  \hspace{1cm} (3.3.4.1)
The structural change of protease corresponding to the movement in the 3-dimentional space of 7, 8, and 9th normal modes is represented in Fig. 3.5

Figure 3.5 The 3-dimentional space composed of three collective degrees of freedom corresponding to 7 (\( \vec{Q}_1 \)), 8 (\( \vec{Q}_2 \)), and 9th (\( \vec{Q}_3 \)) normal modes of protease. The conformation transition of the molecule is represented as \( 0.5 \vec{Q}_1 + 0.5 \vec{Q}_2 - 0.5 \vec{Q}_3 \). The scale of the space is determined approximately.
3.3.5 Definition of amplitudes of normal vibrations

The eigenvectors represent only the relative displacements of atoms. However, the exploration of normal modes as collective degree of freedom in Motion-planning modeling requires the correct definitions of the limits of varying for each degree of freedom. Values of maximal atom displacements corresponding to normal mode can be determined exactly. The amplitudes of atoms movement corresponding to normal modes, \( a_i \), is proportional to the temperature, \( T \),

\[
a_i = \frac{(2k_B T)^{1/2}}{\nu_i} \quad (3.3.5.1)
\]

where \( k_B \) is Boltzmann’s constant and \( \nu_i = \omega_i / 2\pi \) is the frequency of \( i \)th normal mode (eigenvalue \( \lambda_i = \omega_i^2 \)).

Since absolute frequencies in the simplified model of protein are not determined accurately, the amplitudes of atoms movements corresponding to normal modes also cannot be defined exactly. Approximately, for example, these values it is possible to establish using the crystallographic data containing the atomic displacements (RMSD) of open–closed transition. The linear relationship between approximate frequencies calculated with simplified model and the frequencies calculated with standard approach was revealed [35]. It makes possible to determine the amplitudes of normal vibrations of protein using equation (3.3.5.1) and the coefficient defining the transition between two sets of frequencies. When frequencies are not used to define the amplitudes of normal vibrations, only small number of normal modes with frequencies, which are assumed equal, should be used.

3.3.6 Different ways to represent backbone and sidechains motions of proteins

Experimental and theoretical studies were shown that protein motions can be decomposed into two components, backbone and sidechains motions [13, 64]. They are characterized by different properties [13]: sidechains can characterized as having liquid-like characteristics and backbone as having solid-like rigidity. Hinsen et al. [64] have decomposed a molecular dynamics trajectory of 1.5 ns for C-phycocyanin dimer surrounded by a layer of water into three contribution: the global motion of the residues, the rigid-body of the sidechains relative to the backbone, and the internal deformations of the sidechains. It was shown that they are almost independent by verifying the factorization of the incoherent intermediate scattering function. Therefore, backbone and sidechains motions of a protein can be represented by different ways in the docking simulations.

Normal modes calculated with simplified force field are used to represent the backbone motions of proteins. The representation of side chains as rigid ones in this model can produce the resulting trajectory containing unphysical configurations in which some interatomic interactions are very small. In such case the minimization of a protein structure can be used. It was shown that a relatively small number of steps of steepest descent minimization are sufficient to avoid high energy contributions [29]. Motion-planning methods using the collision detector are more efficient way to resolved this problem.

3.3.7 Nonlinearity of the protein conformational changes by an iterative usage of normal mode calculations
Under 200 K the motions of various proteins may be regarded as harmonic, whereas at physiological temperatures, protein flexibility arises from fluctuations between different states [34]. A “conformational substates” [65-67] or “jumping through the minima” [23] models describe such type of motions.

Relatively short time of normal modes calculations using simplified protein model and type of potential energy function of this model makes possible to receive the normal modes for different protein conformations which are considered as minima on the potential energy surface. The corresponding normal coordinates may be used to describe complex rearrangement of a protein structure during the docking process. For example, Miyashita et al. have represented the nonlinearity of the conformational changes of adenylate kinase by an iterative usage of normal modes calculations [49].

3.8 Conclusions

The new combined method for presentation of large amplitude global protein motions based on Motion Planning approach and Normal Mode Analysis with simplified molecular model has been developed. The efficiency of the method is caused firstly, by possibilities of Motion Planning techniques to calculate in reasonable computational time long trajectories of molecules and large amplitude conformational changes of protein and secondly, by rapidity of Normal Mode Analysis with simplified model giving quite real trajectories of atoms. The exploration of the method requires the correct definition of amplitudes of normal vibrations. The inaccuracy of harmonic approximation may be compensated by an iterative usage of normal mode calculations.
Chapter 4

Molecular Mechanics block of BioMove3D package

Molecular Mechanics block of BioMove3D package represents the interfaces to two molecular modeling packages applying the atom-atom interactions: ElNemo and AMBER. The interface of Move3D to ElNemo is used to implement the normal mode calculations of protein using simplified model. The interface of Move3D to AMBER makes possible to implement different calculations applying atom-atom interactions: calculations of potential energy of a molecule, optimization of a molecular structure by minimization of an energy function and molecular dynamics calculations.

4.1 Structure of ElNemo package


The source code of ElNemo is written in FORTRAN language. The ElNemo package is composed of two main parts, PDBMAT and DIAGRTB.

4.1.1 Calculations of potential energy of protein molecule using Elastic Network Model

PDBMAT is a fortran program in which Tirion's model, also called the Elastic Network Model, is implemented. Within the frame of this model, atoms less than cutoff away from each other’s are linked by springs of same strength. Pdbmat is the "Tirion" matrix builder: from a pdb file, yields a "Tirion" matrix, in the following format: i, j, non-zero-ij-element. Output matrix is saved in file with extension sdijf. The CUTOFF parameter defines which atom-atom interactions should be kept in the elastic network model. CUTOFF equals 8 Å by default. The parameter NRBL defines how many residues should be treated as a rigid body. This parameter is defined automatically as a function of protein size by default. The parameter KNONB determines the strength of springs between the atoms in Tirion's model. The value 8 kcal/mol·Å² should be chosen when all hard atoms of molecule are taken in account. When only Cα atoms are regarded, KNONB should be 10-13 kcal/mol·Å². If the molecular model with all atoms including the hydrogens is explored the value of the parameter KNONB may be less (6 kcal/mol·Å²). Masses of atoms are all set to one by default. In origin version of ElNemo the atomic coordinates found in pdb file were used. In our version the coordinates corresponding to the positions of the geometric primitives (spheres) of current robot are taken. The coordinates and masses of atoms are saved in the file ended by xyzm.

4.1.2 Transition from Cartesian coordinates to Rotation Translation coordinates and diagonalization of Hessian matrix
**DIAGRTB** is a fortran program with a standard diagonalization routine to obtain the corresponding normal modes of the system. For the large system as protein molecule the RTB approximation is used. DIAGRTB is composed of the sequence of following FORTRAN subroutines.

BLOCPDB determines block lengths, as a consequence of the number of residues per block. Pdb file is used to receive the information about amino acids in the protein. This subroutine saves the coordinates and masses of atoms in the temporary file rtb.pdb.

PREPMAT reads the sdijf file where all non-zero elements of Tirion’s matrix are written and splits them into blocks. The temporary file rtb.matbloc is created. It contains the elements corresponding to blocks and sizes of blocks.

RTB carries out the transition from Cartesian coordinates to Rotation-Translation coordinates. It reads the elements of matrix placed in the rtb.matbloc file and creates two temporary files: rtb.sdijb containing the non-zero elements of "Tirion" matrix in Rotation-Translation coordinates and rtb.block in which these elements are grouped per blocks.

DIAGSTD performs a diagonalization of a real symmetric matrix based on the QL algorithm. The original EISPACK library is used. The matrix written in rtb.sdijf is used as input. The calculated eigenvalues and eigenvectors are placed in a file with eigenfacs extension.

RTBTOMODES performs the passage from the eigenvectors of "Tirion" matrix in Rotation-Translation coordinates to eigenvectors in Cartesian coordinates and writes them in a file having eigenrtb extension.

### 4.1.3 Additional programs of ElNemo package

There are additional subroutines to analyse calculated normal modes of an investigated molecule.

CHECK_MODES calculates DEGREE OF COLLECTIVITY. This value indicates the fraction of residues that are significantly affected by a given mode. For maximal collective movements the degree of collectivity tends to be a value of one, whereas for localized motions, where the normal mode movement only involves few atoms, the degree of collectivity approaches zero. While low-frequency normal modes are expected to have collective characters, especially those related to functional conformational changes of proteins [36], computed ones sometimes happen to be localized. In such cases, they correspond to motions of some extended parts of the system, as often observed in crystallographic protein structures for N- or C-termini.

The OVERLAP (PROJ_MODES programme) measures the degree of similarity between the direction of a chosen conformational change and the direction of a given normal mode. The difference vector between the reference structure and a second conformation of the same protein or that of a close homologue here defines a conformational change. ElNemo reports cumulative values for the square of the overlap, starting with the lowest-frequency nontrivial normal mode. Note that, because the normal modes form a basis, this cumulative sum reaches a value of one when it is computed over all modes. If the considered conformational change has a collective character, the cumulative sum usually reaches a value of 0.7–0.8 already within the 20–50 lowest-frequency modes [41]. What makes NMA useful for predicting protein movements is the fact that in a large number of cases, one or two low-frequency normal modes, i.e. those with the highest overlap are enough for providing a fair description of the conformational change [36].
B-FACTORS (BFACFTORS programme) are computed from the mean square displacement $<R^2>$ of the first 100 lowest-frequency normal modes using the relationship $B = (8\pi^2/3) <R^2>$ and linear scaling to the observed B-factors in the reference structure as described in [36]. Correlations between NMA and crystallographic B-factors are usually found to be >0.5–0.6 [44], while values >0.8 have been reported [36]. Adjusting force constant parameter KNONB can slightly improve such correlations. This probably reflects the fact that modifying KNONB affects low-frequency densities [33]. The comparison between computed and observed crystallographic B-factors provides a measure of how well the protein’s flexibility in its crystal environment is described by the normal modes.

ROOT MEAN SQUARE DISTANCES (RMS_MODE programme) between the normal mode perturbed models and a second (not necessarily sequence-identical) structure are computed by a rigid body superposition using the lsqman software [68]. Reported are the RMSD between all C\alpha atoms of the two protein conformations, the number C\alpha atoms that are closer than 3 Å in the rigid body superposition and the RMSD between those atoms only. These numbers can be used as a proxy for the overlap in the case of not 100% sequence-identical proteins.

4.2 Interface of Move3d to AMBER

4.2.1 Input files

Widely used practices in molecular simulations to take the initial positions of atoms of molecule from experimental crystallographic data found in a pdb file of Protein Data Bank [http://www.rcsb.org/pdb/]. The hydrogen positions cannot be determined by crystallographic techniques and standard pdb files do not contain the hydrogen atoms. The AMBER has the possibility to compose complete pdb files including the hydrogens. For adding the hydrogen atoms three internal coordinates (bond length, bond angles and torsion angle) from the prep files (in directory $AMBERHOME/dat/leap/prep) are used. The positions determined by such method should be precise by energy minimization to avoid high energetic close contacts of atoms. The small number of minimization cycles generally is sufficient to displace light hydrogen atoms. But a large number of iterations should be used to precise the positions of hard atoms if it is necessary.

The ambmov program has been created for the preparation of input pdb file for Move3d-AMBER simulations. There are not the standard atom names for hydrogen atoms and AMBER may not accept the pdb file containing them. Therefore the first stage of the preparation of the input file is the deleting all hydrogens in the original pdb file. Then the AMBER adds the hydrogen atoms using its libraries. The minimization of energy is used to define more exactly the molecular structure and the new pdb file containing the minimized positions of all atoms is produced. The restraining with a harmonic potential is used to keep a ligand in the initial position during minimization.

4.2.2 Preparatory part for calculations of atom-atom interactions
The preparatory functions for calculations of atom-atom interactions are related to LeaP part of AMBER. The name LEaP is an acronym constructed from the names of older AMBER software modules, it replaces: link, edit, and parm.

At the beginning of a work with the interface Move3d-AMBER a pdb file is read and the following AMBER structures: OBJEKTs, UNITs, RESIDUEs, ATOMs, and PARMSETs (Parameter Sets) are created.

Information that used to generate the AMBER structures:

1) **Topology**: connectivity, atom names, atom types, residue names, and charges. This information comes from database, which is found in the $AMBERHOME/dat/prep directory. It contains topology for the standard amino acids as well as N- and C-terminal charged amino acids, DNA, RNA, and common sugars. Topology information for other molecules (not found in the standard database) is kept in a special file with extension .prep which is generated using the AMBER program antechamber ($AMBERHOME/exe directory).

2) **Force field**: Parameters for all bonds, angles, dihedrals, and atom types in the system. The standard parameters for several force fields are found in the $AMBERHOME/dat/parm directory. AMBER proposes very good defined force fields for amino acids, DNA, RNA, and sugars (Table). The special force field GAFF has been developed for other molecules. Atom types in GAFF are more general and cover most of the organic chemical space. When some parameters for a molecule can not be founded in the file gaff.dat, the special program parmchk ($AMBERHOME/exe directory) is used. It defines missing force field parameters and write them in a file ended by .parm.

The topology and force field parameters defined for a protein or a protein-ligand system are written in a special AMBER file ended by .parm. The format of this file is founded in Appendix C of “Amber 8 Users’ Manual”.

### 4.2.3 Specification of force field

The files founded in directory $AMBERHOME/dat/leap/cmd are used to specify the AMBER force field. It is necessary to use link for file chosen. For example:

```
   cd $AMBERHOME/dat/leap/cmd
   ln –s leaprc.ff03 leaprc
```

In such case force field ff03 (“ff” is force filed) will be used.
Table 4.1 Specification of force fields

<table>
<thead>
<tr>
<th>Filename</th>
<th>Topology</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>leaprc.ff86</td>
<td>Weiner et al. 1986</td>
<td>parm91X.dat</td>
</tr>
<tr>
<td>leaprc.ff94</td>
<td>Cornell et al. 1994</td>
<td>parm94.dat</td>
</tr>
<tr>
<td>leaprc.ff96</td>
<td></td>
<td>parm96.dat</td>
</tr>
<tr>
<td>leaprc.ff98</td>
<td></td>
<td>parm98.dat</td>
</tr>
<tr>
<td>leaprc.ff99</td>
<td></td>
<td>parm99.dat</td>
</tr>
<tr>
<td>leaprc.ff03</td>
<td></td>
<td>parm99.dat+frcmod.ff03</td>
</tr>
<tr>
<td>leaprc.ff02</td>
<td>Duan et al. 2003 reduced (polarizable) charges</td>
<td>parm99.dat</td>
</tr>
<tr>
<td>leaprc.ff02EP</td>
<td>“ + extra points</td>
<td>parm99EP.dat</td>
</tr>
<tr>
<td>leaprc.gaff</td>
<td>none</td>
<td>gaff.dat</td>
</tr>
<tr>
<td>leaprc.glycam04</td>
<td>Woods et al.</td>
<td>glycam04.dat</td>
</tr>
<tr>
<td>leaprc.glycam04EP</td>
<td></td>
<td>glycam04EP.dat</td>
</tr>
</tbody>
</table>

4.2.4 Running the energy calculations, energy minimization and molecular dynamics calculations

The functions of running the energy calculations, energy minimization and molecular dynamics calculations relate to Sander part of AMBER. The acronym stands for Simulated Annealing with NMR-Derived Energy Restraints. The source code of Sander is written in FORTRAN language.

The reading of the first lines of parm file gives the memory size required for calculations of atom-atom interactions of an investigated molecular system. Then the memory for the variables identifying the molecular topology and the force field parameters is allocated, and these variables are defined by the reading of the rest part of parm file.

The same function is used to run energy calculations, energy minimization and molecular dynamics. The choice between minimization of energy and molecular dynamics calculations is defined by IMIN flag of AMBER. When IMIN = 1, the minimization of energy is implemented and IMIN = 0 determines molecular dynamics calculations. Energy minimization with zero number of iterations is used to calculate the potential energy of the investigated molecules.

4.2.5 Parameters of energy minimization

Sander implements two methods of minimization: Steepest Decent and Conjugate Gradient Appendix A). The NTMIN flag of AMBER is used to select the method of minimization. When a molecular structure is far from minimum the SD method is more efficient (NTMIN = 2). When a molecular structure is close to minimum the CG method is preferable.
(NTMIN = 0), NTMIN = 1 determines the mixed SD + CG minimization. In such case the method of minimization will be switched from SD to CG after NCYC cycles.

The special function is used to update the conformation of an investigated molecule after the minimization. But the updated form does not correspond exactly to the molecular conformation defined by minimized coordinates because little changes of bond and bond angle take place during minimization whereas the articulated model of Move3d affords to manipulate only with torsions angles.

4.2.6 Parameters of the molecular potential energy function

AMBER has eight ways to change the potential energy function of a molecule using the NTF flag:

- NTF = 1 complete interaction is calculated (default);
- NTF = 2 bond interactions involving H-atoms omitted, use with NTC=2 — bonds involving hydrogen are constrained;
- NTF = 3 all the bond interactions are omitted, use with NTC=3 — all bonds are constrained;
- NTF = 4 angle involving H-atoms and all bonds are omitted;
- NTF = 5 all bond and angle interactions are omitted;
- NTF = 6 dihedrals involving H-atoms and all bonds and all angle interactions are omitted;
- NTF = 7 all bond, angle and dihedral interactions are omitted;
- NTF = 8 all bond, angle, dihedral and non-bonded interactions are omitted

4.3 The main capacities of Molecular Mechanics block of BioMove3d

1. Calculations of atom-atom interactions for different biological molecules
2. Minimization of the potential energy function of an investigated molecule using three methods: Steepest Decent, Conjugate Gradient and mixing Steepest Decent and Conjugate Gradient method
3. Molecular Dynamics simulations
4. Possibility to choose used force filed from the force fields available in AMBER
5. Possibility to change the parameters of minimization of energy: maximal number of interactions, number of Steepest Decent interactions in case of mixing Steepest Decent and Conjugate Gradient method
6. Possibility to change the parameters of the potential energy function of an investigated molecule in case when it is not necessary to calculate all terms
7. Energy calculations, minimization of energy and molecular dynamics simulations may be carried out not only for one molecule but for molecular system composed of protein and ligand
8. Normal mode analysis of protein molecule using simplified model
9. Exploration of collective degrees of freedom corresponding to normal modes of protein separately for each normal mode
10. Exploration of 5-dimention space of five collective degree of freedom corresponding to normal modes with lowest frequencies
References


### Appendix A

**Table** Studies of the dynamical features of macromolecules by normal mode analysis

<table>
<thead>
<tr>
<th>Article</th>
<th>Molecules</th>
<th>Methods</th>
<th>Structural changes</th>
<th>Normal modes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delarue and Sanejouand 2002 [41]</td>
<td>DNA-dependent polymerases, from 326 to 3500 residues</td>
<td>Elastic Network Model (ENM) including only $C_\alpha$ atoms</td>
<td>Open-closed transitions</td>
<td>100 lowest frequency Normal Mode (NM), the lowest and/or the second lowest frequency (more than 50% of motions)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Domain motions: translations, rotations up to $45^\circ$; rearrangement with rmsd 6.5 Å</td>
<td>one to four of the lowest frequency</td>
</tr>
<tr>
<td>Thomas et al., 1999 [48]</td>
<td>Aspartate Transcarbamylase: 6 catalic (C) chains of 310 res. and 6 regulatory (R) chains of 153 residues</td>
<td>All-atom model CHARMM-22; Diagonalisation In a Mixed Basis set (DIMB) approach for NM calculations; investigation the structural deformation by Vector analysis of NM</td>
<td>Transition from low affinity tensed (T) state to high affinity relaxed (R) state involving large tertiary and quaternary changes</td>
<td>53 lowest frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Global dynamics</td>
<td>with frequencies 0.5 cm$^{-1}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Specific segment motions: 80s and 240s loops (5 and 10 Å, respectively), the segment [50-55] (2 Å), helix H2</td>
<td>with frequencies below 5 cm$^{-1}$</td>
</tr>
<tr>
<td>Valadie et al., 2003 [42]</td>
<td>Integral membrane protein Mscl</td>
<td>ENM including only $C_\alpha$ atoms</td>
<td>Gating process in which molecule passes through closed form (C), first open state (CE) and open form (O). Gating mechanism is described as a tilt of transmembrane helices to its principal axis.</td>
<td>100 lowest frequency</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Protein</td>
<td>Model Details</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>---------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Miyashita et al., 2005, 2003 [49]</td>
<td>Adenylate kinase</td>
<td>All-atom ENM; Rotation Translation Block (RTB) approach for NM calculations; Nonlinear Elastic Model representing system nonlinearity through an iterative usage of NM calculations and Cracking (or partial unfolding) Model</td>
<td>C → O transition: number 11, 31, 64 recover 65% of motion; O → C transition: 5, 17, 20, 29, 37 — 65%; C → CE transition: 8, 11, 29 — 50%; C E → C transition: 3, 27, 32 — 60%; C E → O transition: 10 lowest frequency — 50%; O → CE transition: 7 lowest frequency — 50%</td>
<td></td>
</tr>
<tr>
<td>Hinsen et al., 1999 [47]</td>
<td>Citrate synthase (C), HIV-1 reverse transcriptase (H), aspartate transcarbamylase (A)</td>
<td>ENM including only $C_\alpha$ atoms; Domains definition using NM</td>
<td>Domain motions: Four lowest frequency (C); Seven lowest frequency (H); 20 lowest frequency (A)</td>
<td></td>
</tr>
<tr>
<td>Li and Cui, 2002 [50]</td>
<td>Ca-ATPase</td>
<td>All-atom model CHARMM; RTB approach for NM calculations</td>
<td>Structural flexibility of cytoplasmic domains (especially N), indentifies hinge regions (A/M and A/P), correlated motions between cytoplasmic and transmembrane domains; mainly the lowest frequency and also several low frequency with number 2-6,15,17 with frequencies between 1-4 cm$^{-1}$</td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>System/Proteins</td>
<td>ENM Description</td>
<td>Conformational Change</td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Reuter et al., 2003 [46]</td>
<td>Ca-ATPase</td>
<td>ENM including only Cα atoms; DomainFinder program; Steepest descent minimization of the side chains after displacement along NM</td>
<td>Transition from open (E1Ca) to closed (E2TG) form</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>number 7,8,15 — 55.9% 7,8 (large movement of N domain) 15 (movement of luminal loops)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Domain motions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>number 7-16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Two types of motions for the transmembrane helix: i), correlated twist of the 10 M helix; ii), independent rotations and translations of each M helix;</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>number 9 — 30%</td>
<td></td>
</tr>
<tr>
<td>Tama and Sanejouand 2001 [36]</td>
<td>20 proteins</td>
<td>ENM including only Cα atoms</td>
<td>Open-closed transitions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A lot of information on the nature of the conformational change is often found in a single low-frequency NM</td>
<td></td>
</tr>
<tr>
<td>Tama and Brooks, 2005 [51]</td>
<td>Icosahedral Viral Capsids (viruses with high symmetry)</td>
<td>ENM including only Cα atoms; RTB approach for NM calculations (block is a protein)</td>
<td>Transition between two conformations , rmsd 7.2 (NöV molecule), 9.7(CCMV), and 15.6 Å (HK97)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>500 lowest frequency — 95% first non-degenerate — 90%(NöV, CCMV), 65%(HK97), fist and second non-degenerate 90%(HK97)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Predictions of conformation changes when only one state is known</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>few lowest frequency</td>
<td></td>
</tr>
<tr>
<td>Tama et al., 2000 [35]</td>
<td>12 proteins of various sizes, ranging from 46 to 858 residues</td>
<td>All-atom model CHARMM; Standard and RTB approach for NM calculations (number of residues per block 1-5 or more)</td>
<td>Root Mean Square Fluctuations (RMSF) of Cα atoms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30 lowest frequency</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Open-closed transitions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>one low frequency</td>
<td></td>
</tr>
<tr>
<td>Suhre and Sanejouand 2004 [52]</td>
<td>Maltodextrin-binding protein, HIV-1 protease, Glutamine-binding protein</td>
<td>All-atom ENM; RTB approach for NM calculations</td>
<td>Crystallographic R-factor</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>One or two low-frequency NM</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>System</td>
<td>Model/Approach</td>
<td>Summary</td>
<td>Motions, Frequency</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Wang et al., 2005 [53]</td>
<td>S-adenosyl-L-homocysteine hydrolase</td>
<td>All-atom model CHARMM; Block approach for NM calculations</td>
<td>Open → closed transition by a hinge bending motion of two large domains within each subunit</td>
<td>four lowest frequency</td>
</tr>
<tr>
<td>Cui et al., 2004 [30]</td>
<td>F1-ATPase 30,856 atoms</td>
<td>All-atom model CHARMM and standard approach to calculate NM of individual subunits in αβγ complex; RTB approach for calculations of NM of all molecule</td>
<td>Various global and local motions of subunits, RMSF, more local motions of each subunit</td>
<td>Three lowest frequency, 30 lowest frequency, Four lowest frequency</td>
</tr>
<tr>
<td>Zacharias and Sklenar, 1999 [28]</td>
<td>DNA</td>
<td>All-atom model GROMOS87; Minimization of energy using eigenvectors as independent variables</td>
<td>Conformational changes during ligand binding, rmsd ~ 1 Å</td>
<td>5-40 lowest frequency</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Overall bending and minor groove closing and opening deformation</td>
<td>number 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Overall bending and some minor groove motion</td>
<td>number 2 and 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Major groove opening and closing deformation</td>
<td>number 4</td>
</tr>
<tr>
<td>Kolossvary and Guida, 1999 [29]</td>
<td>PNP</td>
<td>All-atom model AMBER; Inverse power method for NM calculations; conformational search procedure LMOD</td>
<td>Conformational changes during ligand binding</td>
<td>Low frequency</td>
</tr>
<tr>
<td>Kim et al., 2003 [44]</td>
<td>Caspsid bacteriophage HK97</td>
<td>Elastic Network Interpolation (ENI), only Cα atoms are included; symmetry constraints for NM calculations</td>
<td>Maturation from Prohead II to Head II: rigid body motion of two compact domain (A and P) with two refolding extensions (N-arm and L-loop) and approach of two particular residues associated with isopeptide bonds that make hexagonal cross-links in Head II.</td>
<td>Four lowest frequency</td>
</tr>
</tbody>
</table>
Appendix B

Example: Calculations of energy of atom-atom interactions and energy minimization for molecular system composed of two proteins (~10 000 atoms) using BioMove3d and AMBER

**Parameters of minimization:**

NTMIN = 1, MAXCYC = 100, NCYC = 10, DRMS = 0.0001

There is not differences between energy terms calculated using BioMove3d and Amber because the same atomic coordinates corresponding to coordinates of pdb file are used in two cases. However, there are differences for energy minimization of molecular system after structural changes related to 0.15 amplitude of 7th normal mode of Protein. It is explained by differences in the coordinates used in BioMove3d and the coordinates taken from the file created to pass them to AMBER.

**Calculation of atom-atom interactions using BioMove3d**

<table>
<thead>
<tr>
<th>NSTEP</th>
<th>ENERGY</th>
<th>RMS</th>
<th>GMAX</th>
<th>NAME</th>
<th>NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.3327E+04</td>
<td>4.2718E-01</td>
<td>4.7481E+00</td>
<td>CD</td>
<td>7150</td>
</tr>
<tr>
<td></td>
<td>BOND = 291.9135</td>
<td>ANGLE = 1252.0575</td>
<td>DIHED = 5314.4706</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VDWAALS = -4877.3943</td>
<td>EEL = -42372.0735</td>
<td>HBOND = 0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-4 VDW = 1884.9254</td>
<td>1-4 EEL = 25179.4094</td>
<td>RESTRAINT = 0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Minimization (SD+CG) after displacement corresponding to 0.15 of 7th normal mode of Protein using BioMove3d**

<table>
<thead>
<tr>
<th>NSTEP</th>
<th>ENERGY</th>
<th>RMS</th>
<th>GMAX</th>
<th>NAME</th>
<th>NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.4677E+04</td>
<td>5.6022E-01</td>
<td>1.1334E+01</td>
<td>OE1</td>
<td>2633</td>
</tr>
<tr>
<td></td>
<td>BOND = 291.9135</td>
<td>ANGLE = 1252.0529</td>
<td>DIHED = 5314.4255</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VDWAALS = -4867.7394</td>
<td>EEL = -42372.0725</td>
<td>HBOND = 0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-4 VDW = 1884.9254</td>
<td>1-4 EEL = 25179.4247</td>
<td>RESTRAINT = 0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NSTEP</th>
<th>ENERGY</th>
<th>RMS</th>
<th>GMAX</th>
<th>NAME</th>
<th>NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>-1.4688E+04</td>
<td>1.7920E-01</td>
<td>2.8675E+00</td>
<td>OE1</td>
<td>2633</td>
</tr>
<tr>
<td></td>
<td>BOND = 291.2525</td>
<td>ANGLE = 1256.9781</td>
<td>DIHED = 5314.5692</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VDWAALS = -4873.6500</td>
<td>EEL = -43737.1734</td>
<td>HBOND = 0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-4 VDW = 1884.9340</td>
<td>1-4 EEL = 25179.0981</td>
<td>RESTRAINT = 0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NSTEP</th>
<th>ENERGY</th>
<th>RMS</th>
<th>GMAX</th>
<th>NAME</th>
<th>NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.4659E+04</td>
<td>2.5234E-01</td>
<td>3.7678E+00</td>
<td>OE1</td>
<td>2633</td>
</tr>
<tr>
<td></td>
<td>BOND = 292.4153</td>
<td>ANGLE = 1256.6760</td>
<td>DIHED = 5317.3215</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VDWAALS = -4875.9247</td>
<td>EEL = -43707.9889</td>
<td>HBOND = 0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-4 VDW = 1882.4649</td>
<td>1-4 EEL = 25175.7199</td>
<td>RESTRAINT = 0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Calculation of atom-atom interactions using AMBER**

<table>
<thead>
<tr>
<th>NSTEP</th>
<th>ENERGY</th>
<th>RMS</th>
<th>GMAX</th>
<th>NAME</th>
<th>NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.3327E+04</td>
<td>4.2718E-01</td>
<td>4.7481E+00</td>
<td>CD</td>
<td>7150</td>
</tr>
<tr>
<td></td>
<td>BOND = 291.9135</td>
<td>ANGLE = 1252.0575</td>
<td>DIHED = 5314.4706</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VDWAALS = -4877.3943</td>
<td>EEL = -42372.0735</td>
<td>HBOND = 0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-4 VDW = 1884.9254</td>
<td>1-4 EEL = 25179.4094</td>
<td>RESTRAINT = 0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Minimization (SD+CG) after displacement corresponding to 0.15 of 7th normal mode of Protein using AMBER**

<table>
<thead>
<tr>
<th>NSTEP</th>
<th>ENERGY</th>
<th>RMS</th>
<th>GMAX</th>
<th>NAME</th>
<th>NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1.4694E+04</td>
<td>5.5939E-01</td>
<td>1.1325E+01</td>
<td>OE1</td>
<td>2633</td>
</tr>
<tr>
<td></td>
<td>BOND = 291.9135</td>
<td>ANGLE = 1252.0529</td>
<td>DIHED = 5314.4257</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VDWAALS = -4867.6033</td>
<td>EEL = -43749.6624</td>
<td>HBOND = 0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-4 VDW = 1885.0025</td>
<td>1-4 EEL = 25179.4244</td>
<td>RESTRAINT = 0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NSTEP</th>
<th>ENERGY</th>
<th>RMS</th>
<th>GMAX</th>
<th>NAME</th>
<th>NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>-1.4705E+04</td>
<td>1.7884E-01</td>
<td>2.8292E+00</td>
<td>OE1</td>
<td>2633</td>
</tr>
<tr>
<td></td>
<td>BOND = 291.2428</td>
<td>ANGLE = 1252.9524</td>
<td>DIHED = 5314.5500</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VDWAALS = -4873.7225</td>
<td>EEL = -43754.0772</td>
<td>HBOND = 0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-4 VDW = 1884.9361</td>
<td>1-4 EEL = 25179.1289</td>
<td>RESTRAINT = 0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NSTEP</th>
<th>ENERGY</th>
<th>RMS</th>
<th>GMAX</th>
<th>NAME</th>
<th>NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>-1.4803E+04</td>
<td>2.6375E-01</td>
<td>3.8679E+00</td>
<td>OE1</td>
<td>2633</td>
</tr>
<tr>
<td></td>
<td>BOND = 292.3456</td>
<td>ANGLE = 1256.8268</td>
<td>DIHED = 5317.3233</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VDWAALS = -4876.1960</td>
<td>EEL = -43851.5958</td>
<td>HBOND = 0.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-4 VDW = 1882.3800</td>
<td>1-4 EEL = 25175.4574</td>
<td>RESTRAINT = 0.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>