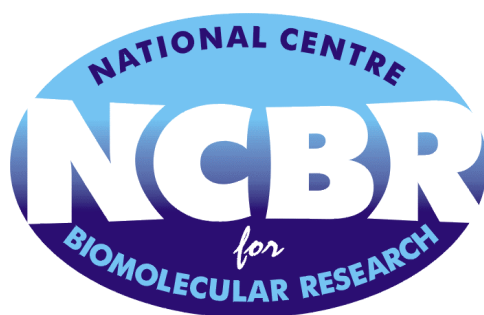


TRITON

Manual and Tutorials

version 3.0



March 2003

Program:

Martin Prokop

Concept:

Jiří Damborský, Michal Kutý, Jaroslav Koča

Manual and tutorials:

Michal Boháč, Jiří Damborský, Martin Prokop

Official web page:

<http://ncbr.chemi.muni.cz/triton/triton.html>

Contact e-mail address:

triton@chemi.muni.cz

References to be cited if you use:**TRITON**

Prokop, M., Damborský, J., Koča, J., [TRITON: In Silico Construction of Protein Mutants and Prediction of Their Activities](#). *Bioinformatics*, 2000. **16**: 845-846.

Damborský, J., Prokop, M., Koča, J., [TRITON: Graphic Software for Rational Engineering of Enzymes](#). *Trends in Biochemical Sciences*, 2001. **26**: 71-73.

DRIVER

Černohorský, M., Kutý M., and Koča J., [A Multidimensional Driver for Quantum Chemistry Program MOPAC](#). *Computers & Chemistry*, 1996. **21**: 35-44.

© Martin Prokop, Jiří Damborský, Michal Boháč, Michal Kutý and Jaroslav Koča

National Centre for Biomolecular Research, Faculty of Science, Masaryk University
Kotlářská 2, 611 37 Brno, Czech Republic

Contents

1	Introduction	4
2	New in version 3.0	5
3	Methodology	6
4	Reliability	7
5	Limitations	7
6	Graphic Interface	8
7	Installation	9
8	Setup	10
9	Basic Operations Tools	12
10	Projects Management Tools	16
11	Data Analysis Tools	19
12	Wizard Mutagenesis	21
13	Wizard Reaction	27
14	Wizard Optimization	36
15	Tutorials	42
15.1	Tutorial 1: Construction of single-point mutant proteins	43
15.2	Tutorial 2: Modelling of enzymatic reaction using a minimal model	44
15.3	Tutorial 3: Modelling of enzymatic reaction using a medium model	46
15.4	Tutorial 4: Modelling of enzymatic reaction using a large model	49
15.5	Tutorial 5: Assignment of atomic charges to a protein	53
16	Acknowledgements	55
17	References	55

1 Introduction

The program TRITON [1] is a graphic software package for *in silico* construction of protein mutants, for modelling enzymatic reactions and for the analysis of essential interactions between the enzyme and its substrate. The protein mutants are constructed by homology modelling using the external program MODELLER [2] while the reactions are modelled using semi-empirical quantum-mechanic method of external program MOPAC [3]. The program TRITON provides graphic tools for the preparation of the input data files, for running calculation and for the analysis of the output data. Implementation ensures the overall integrity of consecutive steps of the modelling of mutants and evaluation of their catalytic properties. TRITON is suitable for molecular biologists and biochemists, non-specialists in computer modelling.

The program TRITON is being developed in the National Centre for Biomolecular Research (formerly Laboratory of Biomolecular Structure and Dynamics), Faculty of Science, Masaryk University, Brno, Czech Republic since 1996. The program is freely provided to academic users via the official web site <http://ncbr.chemi.muni.cz/triton/triton.html>.

2 New in Version 3.0

- New **Optimization wizard** was added for geometry optimization, single point calculation and optionally also the transition state search calculation using the program MOPAC.
- New version of **MODELLER 6v2** was implemented with many bugs fixed and *.top* file improved.
- New version of **MOPAC2002** with **MOZYME** and **PM5** algorithms was ported.
- New possibility for setting and running **consecutive reactions** was added (only with MOPAC2000 or MOPAC2002).
- **Export** of each MOPAC calculated geometry as separate *.pdb* file was made possible.
- **Parallel driving** of two distances at the same time was implemented. This procedure enables double driving of two pairs of atoms, where the distance of each pair is changed by different value within the same driving step.
- **Select neighbourhood** option was added. The Wizard Reaction and Wizard Optimization contain this option for simple selection of atoms or whole residues, which are within specified distance to selected atom/residue too.
- Improvements in visualization of output data were made:
 - New **Electrostatics Progress Graph** allows user to visualize: (i) progress of partial charge for selected atom or total residual charge for selected residue during the modeled reaction or (ii) progress of electrostatic interaction energy between selected atoms/residues during the reaction.
 - Improved **Residue Interactions Graph** allows user to visualize the electrostatic interaction energies of selected residues with respect to specified atom. The exact values of interactions are shown upon moving the mouse on the specified column. The graph can be easily modified by exclusion/addition of specific residues from their list.
- Many **bugs** from previous version 2.0beta were **fixed**.

3 Methodology

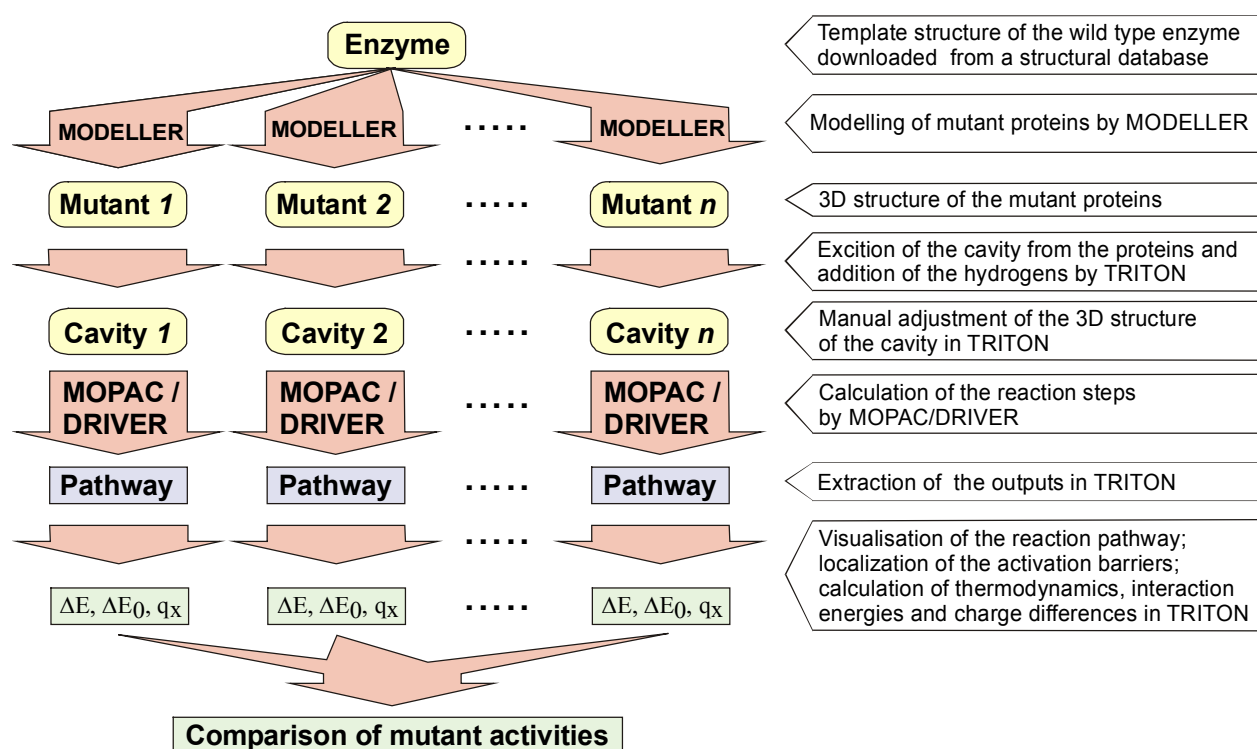


Fig 3.1 Methodology of in silico site-directed mutagenesis implemented in TRITON

1. Modelling of protein mutants

The 3D structures of mutants are generated (**Fig 3.1**) by the program MODELLER [2]. This program uses the method of the satisfaction of spatial restrains for model building with the structure from the structural database as a template and the amino acid sequence of the studied protein with the desired substitution as the target sequence.

2. Preparation of the input data for reaction pathway calculation

Only the active site residues, substrate, co-substrate and co-factors are excised from the protein structure and included in the calculation. Hydrogens are added and the structure is manually adjusted to resemble enzyme-substrate complex before the reaction. The positions of all backbone atoms are fixed.

3. Calculation of the reaction pathway

The reaction pathway is calculated by the semi-empirical quantum chemistry program MOPAC [3]. The subroutine DRIVER [4] is used for reaction pathway mapping by step by step changing and fixing of proposed reaction co-ordinate while optimizing remaining internal co-ordinates of the system.

4. Analysis of the output data from the reaction pathway calculation

The changes in energy and geometry of the system during the reaction can be visualized and animated. The activation barriers of the reaction, changes in partial atomic charges of active site residues and interaction energies between the residues and substrate in ground and transition state can be calculated and plotted.

4 Reliability

Reliability of the calculations, i.e. the ability to reproduce experimental data, is crucial to every modelling technique. We are still long way from quantitative predictions of the effect of particular mutation on the protein activity. In many instances, however, it is valuable to provide qualitative or semi-quantitative information about crucial amino acids for catalysis. Such predictions can guide experiments or can assist during interpretation of the experimental data. Systematic computational site-directed mutagenesis of the haloalkane dehalogenase in position 172 [5] was conducted in parallel with the experiment [6]. Experimental activities of 16 mutant proteins were compared with theoretically derived parameter, the charge difference of reactant and product on the halide-stabilizing hydrogen atom of residue 172. This parameter could safely distinguish protein mutants with a high activity from mutants displaying a low activity (**Fig 4.1**).

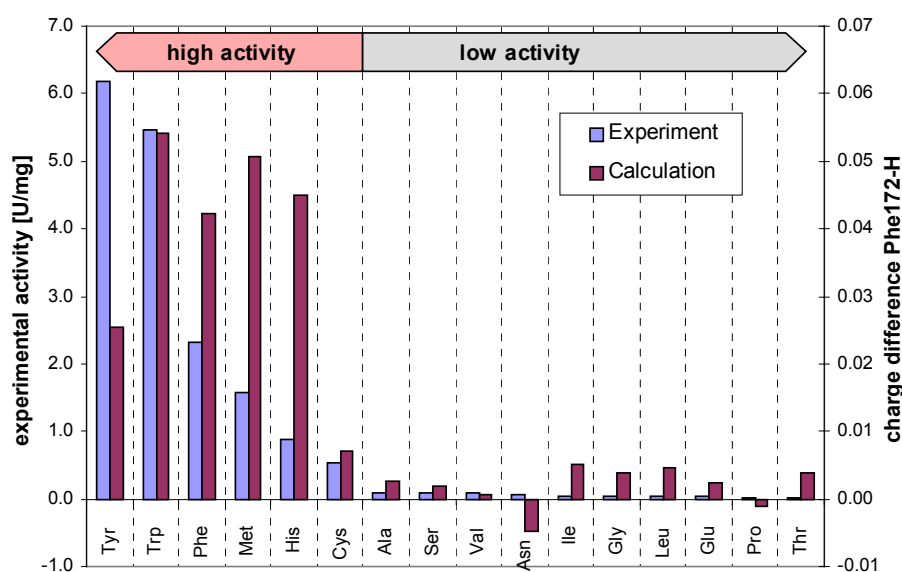


Fig 4.1 Comparison of calculation with experiment for haloalkane dehalogenase mutants

5 Limitations

There is no computational method available to-date, which would predict the effect of mutation on the protein structure with absolute certainty. User must be aware that even single-point mutation can result in a conformational change in the protein structure, which may not be modelled properly by the MODELLER. Furthermore, modelling of biochemical reactions using the MOPAC is restricted to selected active site residues neglecting the rest of the protein. This limitation was recently challenged by the MOZYME algorithm of the MOPAC2000 package. MOZYME algorithm is a SCF algorithm of Localized Molecular Orbitals, which can effectively speed-up the calculation. Application of MOZYME allows inclusion of systems having about 1000 atoms (~30 residues) into a reaction path modelling with reasonable length of calculation. However the applicability of a semi-empirical method for modelling of a particular reaction should be validated against the experimental data or against the results from the *ab initio* calculations performed with the selected parts of modelled system.

6 Graphic Interface

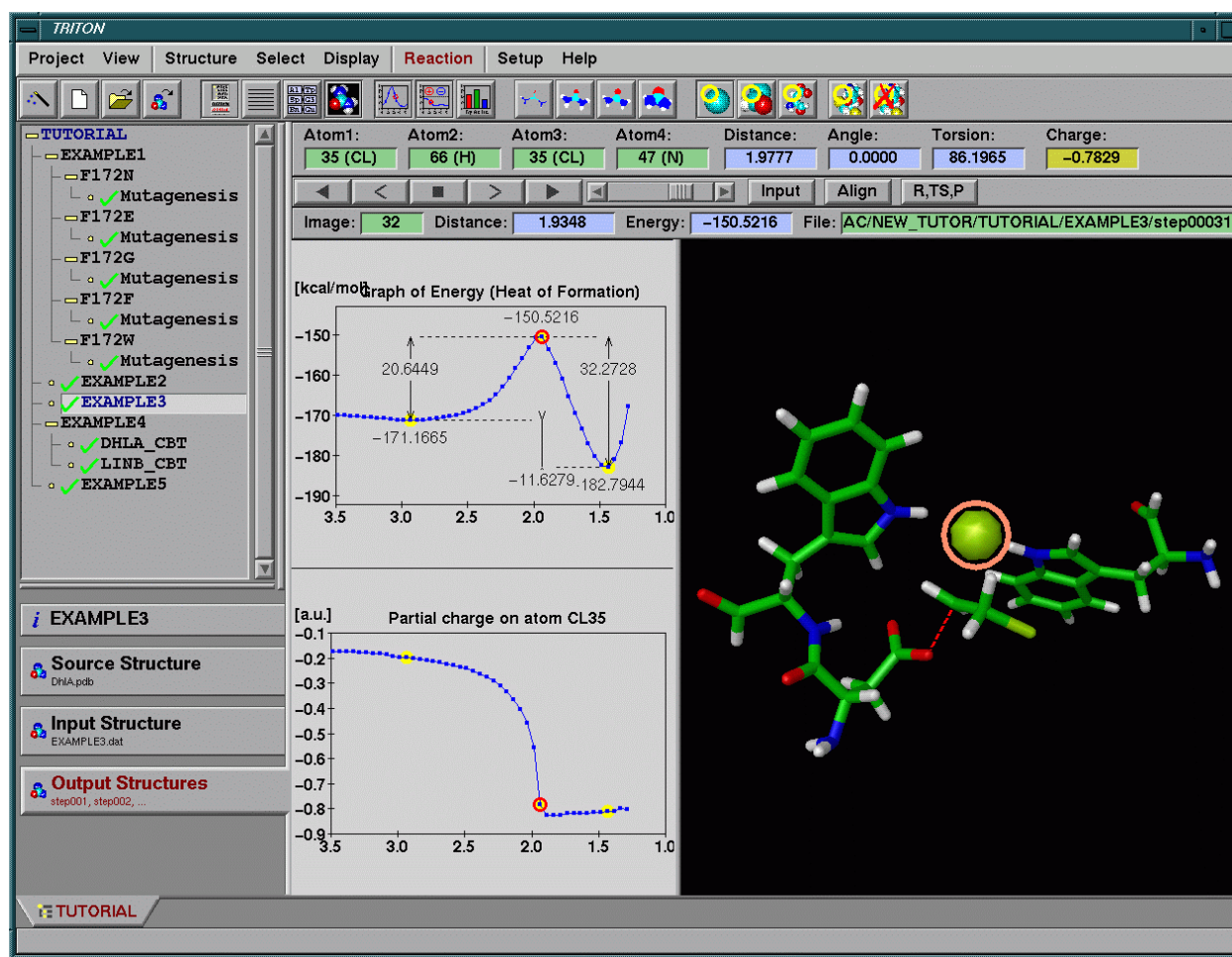


Fig 6.1 Graphic interface of the program TRITON

The graphic interface of the program TRITON is presented in **Fig 6.1**. Several different projects can be opened in the TRITON at the same time. Switching among them is enabled by the **Project Folders**, which are localized on the bottom of the screen. The window for every project can be divided into several parts. The **Project Tree Window** needed for the management of the project hierarchy and the **Bookmarks** for easy access to the data of specific subprojects are situated on the left side of the window. Bookmarks serve for switching among Source, Input and Output Structures. The middle part of main window can serve for the display of the graphs. The **Energy Graph** is a plot of the energy (heat of formation) against the reaction co-ordinate (middle-top window). The **Electrostatics Progress Graph** is a plot of the partial atomic charge on a selected atom, sum of partial charges on a group of selected atoms or electrostatic interaction energy of selected atoms against the reaction co-ordinate (middle-bottom window). The **Residue Interactions Graph** is plotting the values of residual interactions (not shown on this figure; see **Fig 15.6**). A **3D Model Window** enables visualization and manipulation of the structure (right window). The **Animation Bar** enables control of the animation of pre-calculated reaction (above the Energy Graph window). The **Measure Box** serves for measuring of distances, angles and torsion angles (above the 3D Model window).

7 Installation

For installation of TRITON follow these instructions:

1. Download TRITON distribution package from the official web page <http://ncbr.chemi.muni.cz/triton/triton.html>. Program TRITON is available in versions running under IRIX, Linux and NetBSD operating systems. Program is distributed as compressed archive file *triton-irix-3.0.tar.Z*, *triton-linux-3.0.tar.gz* or *triton-netbsd-3.0.tar.gz*.

2. Unpack files under IRIX using the command

```
uncompress -c triton-irix-3.0.tar.Z | tar xvf -
```

under Linux use

```
tar xzvf triton-linux-3.0.tar.gz
```

or under NetBSD use

```
tar xzvf triton-netbsd-3.0.tar.gz
```

All files will be unpacked into *triton-3.0* directory.

3. Log in as root. You will be able to install package programs into system directories. By installing the package as non-root, script will install the files into user directories. In latter case you cannot delete previously created *triton-3.0* directory after the installation.
4. Run the installation script ***install.sh*** placed in *triton-3.0* directory for automatic installation of TRITON package components. Script will ask which components should be installed (TRITON 3.0, MODELLER 6v2, DRIVER 1.2 and MOPAC 7.01).
5. Installation script will ask for MODELLER activation key if you choose to install it from this package. You can receive this key *via* e-mail after signing the license agreement on <http://salilab.org/modeller/registration.shtml>.
6. Installation script will automatically set up the paths in TRITON configuration file for all MODELLER, MOPAC and DRIVER executables installed from distribution package.
7. In the case of problems with installation, please read the INSTALL file in *triton-3.0* directory.
8. Please note that for calculations of bigger systems the newer version of MOPAC is needed. You can receive the MOPAC2002 binaries and necessary license key by e-mail request at ccs@fqspl.com.pl. The license keys for 30-days evaluation copy or full version of MOPAC2002 must be specified if you want to use it with TRITON package.

8 Setup

The program TRITON requires the external program MODELLER for construction of protein mutants, and the programs MOPAC and DRIVER for modelling of chemical and biochemical reactions. Executable files of these programs are provided in the distribution package. For more information about their versions look into README files in *triton-3.0* directory. In the case that you will use only programs distributed within TRITON package, they will be set up automatically during the installation. In the case you will be using your own versions of MOPAC or MODELLER, please set up the right locations of their executables using following instructions.

1. Roll down the *Menu: Setup* on the top of the TRITON screen (**Fig 8.1**) and select any of *Setup MODELLER*, *Setup MOPAC* or *Setup DRIVER* options.



Fig 8.1 Menu of program TRITON with Setup menu indicated

2. Configure MODELLER in *Setup MODELLER* window (**Fig 8.2**). Choose the version of your MODELLER program. Next, specify the location of the executable file on your computer. Special configurations for various versions of MODELLER and the position of its executable file on various machines can be saved and restored by user. Also the initial specification of the license key for running MODELLER can be done using this wizard after the successful registration.

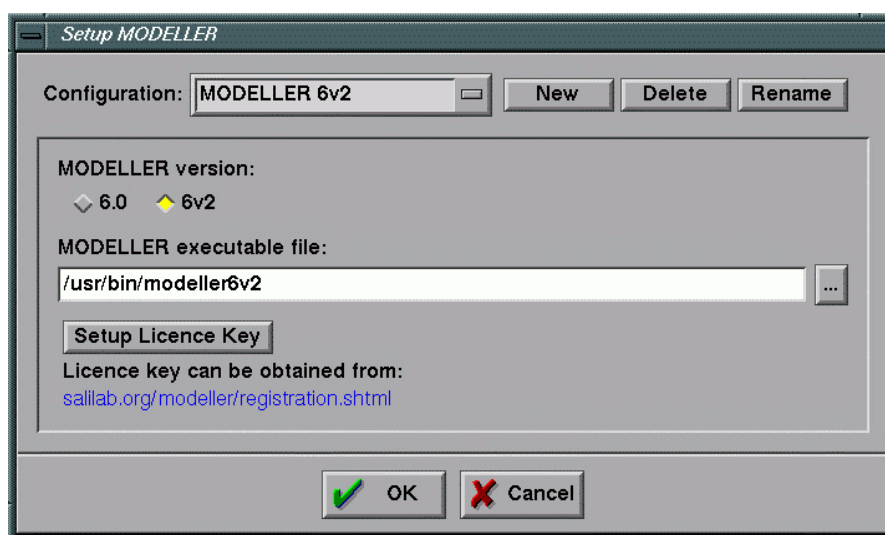


Fig 8.2 Window for setup of MODELLER

3. Configure MOPAC in *Setup MOPAC* window (**Fig 8.3**). Choose the version of your MOPAC program. Next, specify the location of the executable files of MOPAC on your system. Special configurations for various versions of MOPAC and the position of its executable file on various machines can be saved and restored by user.

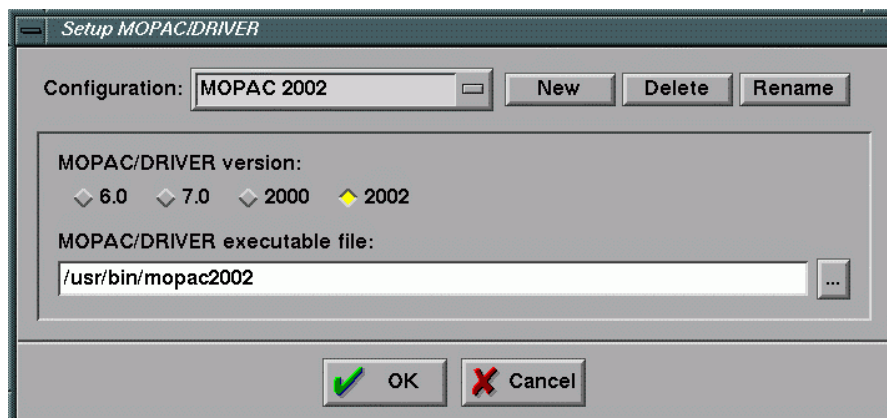


Fig 8.3 Window for setup of MOPAC

4. Configure DRIVER in *Setup DRIVER* window (**Fig 8.4**). Choose the version of your DRIVER program. Next, specify the location of the executable files of DRIVER on your system. Special configurations for various versions of DRIVER and the position of its executable file on various machines can be saved and restored by user.

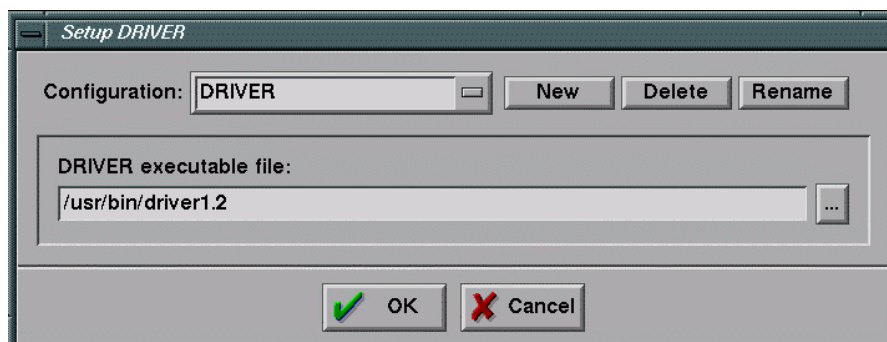


Fig 8.4 Window for setup of DRIVER

9 Basic Operations Tools

The program TRITON enables common work with molecular structures. The molecules which are stored in different formats can be read into the program, visualized using various molecular models, rendered, coloured, labelled, superimposed, etc.

Reading a molecular structure

Structure can be loaded from a file by *Menu: Structure/Load*. Name of the file with structure must be specified in a dialog box. Format of the input file can be specified in *Format* field (or [autodetect] can be left). Only PDB, MOPAC input (internal co-ordinates) and MOPAC output (internal co-ordinates) formats are currently fully supported. The formats MOL2 and TINKER are experimental. Loaded structure is shown as 3D-model. Other structures can be opened in a new folder, added to opened structure or added to project. The folders are shown on the bottom of the window. If you add the structure to already opened structure you can manipulate both structures independently using *Menu: Structure/Manipulate*.

Viewing of a structure as 3D-model, text, list of residues

Structures can be represented as 3D-model (*Menu: View/3D-model*), as the text of source file from which structure was read (*Menu: View/Flat File*) or as a list of amino acid residues and ligands (*Menu: View/Residues*). Note, that residue list can be visualized if the structure is read from the file in PDB format only. All three representations can be displayed at the same time and can be used for selecting of atoms, showing/hiding of atoms, measurement, etc. The changes made in one view are immediately made also in two other views.

Showing/hiding atoms

Atoms can be shown/hidden by items in *Menu: Display (Hide Selected Atoms, Hide Unselected Atoms, Show All Atoms, Hide All Atoms)*. You can also use middle-mouse button in *Flat File* view or *Residues* view to show/hide atoms.

Working with atom subsets

Subsets of atoms are handled using the dialog window *Menu: Structure/Subset*. Subsets can be created or deleted. If the new subset is created, selected atoms are added to the subset. Selected atoms can be added to subset by *Add Atoms* button. *Remove Atoms* button is used for removing all atoms from the subset. Atoms of subset are selected and unselected by clicking on a subset item in a list. Clicking on a subset item by middle-mouse button hides/shows atoms of the subset.

Selection of atoms

Atoms can be selected and unselected by <Ctrl><Left Mouse Button>. This works in 3D-model, flat file view, in residue list and residue interaction graph. In the subsets lists and the residue lists in wizards, clicking the mouse without <Ctrl> can be used. It can be specified whether atoms, residues, molecules or structures will be selected in *Menu: Select*. Here are also other options for selecting of atoms under this menu (*Select All*, *Unselect All*, *Select Backbone*, *Inverse Selection*). The *Select neighbourhood* is very useful option allowing to select atoms or the whole residues, which are within the specified distance from selected atom or residue.

Measurement of distances and angles

Distances, angles and torsion angles of atoms can be visualized using the measure box (*Menu: View/Measure Box*). Values are displayed upon clicking on atoms. Partial charge value can also be displayed in this box if it is included in a source file, e.g. originating from reaction calculation.

Rendering of 3D-models

Rendering of atoms in 3D-model can be achieved using *Menu: Structure/Render*. Available models are *Wire*, *Sticks*, *Balls and Sticks*, and *CPK*. The change in model rendering is applied on entire structure unless particular atoms are selected.

Coloring of atoms

Color of atom can be set in the dialog box *Menu: Display/Atom Colors*. Atoms can be colored by element, by specified color or by partial charge. The last option is enabled only if the partial charge values are included in source file, e.g. originating from reaction calculation.

Manipulation of multiple structures

Multiple structures can be conveniently manipulated using *Menu: Structure/Manipulate*. Program shows the dialog box with the list of structures read in a single folder. Every item in the list includes following fields: *Show* (show/hide structure), *Dial* (enable/disable manipulation with the structure), *Sel* (enable/disable selection of atoms in a structure), specification of color (choose *By Atoms* or any of several colors). Setting of fields is changed upon clicking on it.

Superposition of structures

Superposition of structures can be conducted in a dialog box *Menu: Structure/Superimpose*. This box is enabled only if more than one structure is read in a folder. It is required to assign first and second structure in the fields *Structure 1* and *Structure 2*. The algorithm to be used for superposition must be specified (note: quaternion superposition and RMS are not implemented in current version). If *Specify three atoms* algorithm is used, three atoms of the first structure must be specified by clicking on them by mouse or directly by writing the numbers of atoms in the field *1, 2 and 3*. If the option *Second structure* is checked, three atoms of second structure must be specified otherwise the atoms with the same numbers as in first structure are used. *Apply* button will start transformation of the second structure in following way: position of atom *1* in both structures will be aligned, lines connecting atoms *1* and *2* in both structures will be fitted and finally the planes defined by atoms *1, 2 and 3* in both structures will be fitted.

Labelling of atoms and residues

Atom labels can be shown in *Menu: Display/Atom Labels*. It is possible to show the element (*by Element*), the number of atom (*by Number*) and the location of atom in a residue read from PDB files (*by Location*). Also the name of the residues can be shown (*Menu: Display/Residue Labels*).

Special effects in 3D-model

Some options of 3D-model are listed in *Menu: Display/Options*. Parallel or perspective projection can be set on and off. *Quick rotations* specifies whether 3D-model will be changed to wired model during translation and rotation of the structure.

Usage of mouse in 3D-model view

<Left Mouse Button> Choose atom.

<Ctrl><Left Mouse Button> Select/unselect atom (residue/molecule/structure).

<Middle Mouse Button> Translate the structure in XY plane (your screen).

<Ctrl><Middle Mouse Button> Translate the structure in Z axis (to foreground or to background).

<Middle Mouse Button><Right Mouse Button> Scale the model.

<Ctrl><Middle Mouse Button><Right Mouse Button> Change size of balls and diameter of tubes in the model.

<Right Mouse Button> Rotate structure around X and Y axes.

<Ctrl><Right Mouse Button> Rotate structure around Z axis.

<Left Double Click> outside of any atom or residue Unselect all atoms

Usage of mouse in flat file view

<Left Mouse Button> Choose atom, which was read from this line.

<Ctrl><Left Mouse Button> Select/unselect atom (residue/molecule/structure), which was read from this line (respectively residue/molecule/structure which includes this atom).

<Middle Mouse Button> Show/hide atom, which was read from this line.

<Left Double Click> Unselect all atoms.

<Ctrl><Left Double Click> Select all atoms.

Usage of mouse in residues view

<Left Mouse Button> Choose first atom of a residue.

<Ctrl><Left Mouse Button> Select/unselect atoms of residue.

<Shift><Ctrl><Left Mouse Button> Select/unselect all residues between previously chosen residue and this residue.

<Middle Mouse Button> Show/hide atoms of a residue.

<Left Double Click> outside of any residue Unselect all atoms.

<Ctrl><Left Double Click> on first or last residue of a chain Select all residues of a chain.

Usage of mouse in graphs

<Left Mouse Button> Choose step on reaction pathway (Energy Graph, Electrostatics Progress Graph).

<Ctrl><Left Mouse Button> Select atoms of a residue (only in Residue Interactions Graph).

<Midle Mouse Button> Change scaling of the Y axis (when mouse is vertically moved).

<Ctrl><Midle Mouse Button> Move up or down along the Y axis (when mouse is vertically moved).

<Right Mouse Button> Show context menu.

10 Project Management Tools

The data within TRITON are organized into so-called projects. The projects are important for effective work with molecular structures, for setting up the calculations, and for data visualization and analysis. Each project can include one or more subprojects enabling construction of a hierarchical system to store data. Every newly created project is associated with a unique directory containing a file with *.prj* extension. This file includes all information about the project. The project file is in ASCII format and can be manually edited. Other files used in the project, i.e. co-ordinates, input and output files, are also placed in the project directory.

Creating a new project

A new project can be created using *Menu: Project/New*. Project type must be specified: *General Project*, *Mutagenesis*, *Reaction*. Meaning of these projects is explained in following paragraphs.

Opening an existing project

Formerly created project can be opened using *Menu: Project/Open*. Then the name of the project file (with extension *.prj*) must be specified.

Adding an existing project

Existing subproject can be added to an opened project using *Add Existing Subproject* in the context menu of the project (it is shown after pressing right mouse button on the project item in the project list). Then the project file (with extension *.prj*) must be specified. Opened project is added as a subproject under the actual project on which the context menu was opened). By this way it is possible to create a hierarchical system of projects.

Saving a project

Project file (with extension *.prj*) is saved automatically after some changes are made to the project, e.g., addition/deletion of subprojects. Projects can also be saved explicitly by user from *Menu: Project/Save*.

Deleting a project

Projects can be deleted from *Menu: Project/Delete*. This can also be done from the context menu of a subproject in the subprojects list. Project which should be deleted must be specified in a dialog box. Notice: In this version of TRITON, the data files are actually not deleted; the project is only removed from parent project file (with extension *.prj*). Removed subprojects can be added to another project as described in the paragraph *Adding an existing subproject*.

Closing a project

Open projects can be closed using *Menu: Project/Close*. It closes all projects in the actual folder. The folders are displayed on the bottom of the program window. Separately opened structures have their own folders and can be closed in the same way.

Adding a structure into a project

A structure can be added into the project using *Menu: Structure/Load*. File with structure must be specified in a dialog box. Then the *Project* item must be checked in the right side of the dialog (in the section *Add To*). The structure will be read and added as a separate bookmark into the project.

10.1 General project

General project is typically used when the user initiates study of a new biomolecule. General project serves as a “parent project” under which a number of subprojects (e.g., *Reaction project*, *Optimization project* or *Mutagenesis project*) are created. General project is created by *Menu: Project/New* and pressing of *General Project* button in a dialog box. Then the project name must be entered. After this user must specify if the project will be opened in the TRITON screen separately, i.e. in its own folder, or if it will be added to one of already opened projects. General project by default does not contain any data, but user can add the structures into it (see *Adding structure into a project* in paragraph above). Then the structure is quickly available for opening and for using in consequently created (computational) subprojects (see *Wizard Mutagenesis* and *Wizard Reaction* or *Optimization* chapter).

10.2 Mutagenesis project

Mutagenesis project is used for *in silico* construction of protein mutants using the external program MODELLER. Before the very first Mutagenesis project is created, MODELLER program must be installed, activated and configured (see section Setup on page 10). The only input requested for this project is a file with three-dimensional structure of a protein molecule. This structure will be used as a template structure for homology modelling. This file must be in PDB format (*.pdb* extension). It is highly recommended that the file already contains the hydrogen atoms and substrate molecule in case that the mutant proteins will be further utilized for modelling of the reaction. All input data are specified using the wizard (see Wizard Mutagenesis on page 21). After finishing this wizard, the program will generate the input files for MODELLER and run the calculation. A separate subproject is created for every mutant. Each mutant project contains the *Template structure*, i.e. a wild type structure, and the *Output Structure*, i.e. a mutant structure. You can switch between them using the bookmarks on the left side of TRITON screen (see *Fig 6.1*).

10.3 Reaction project

Reaction project is used to calculate reaction pathway using the external program MOPAC. Before the very first Reaction project is created, MOPAC and DRIVER programs must be installed and configured (see section Setup on page 10). The only input required for this project is a file with the three-dimensional structure of the system (usually protein plus substrate). It is highly recommended that this file is in PDB format (*.pdb* extension). Calculation can be done on one structure or on a set of mutant structures generated using Mutagenesis project. All input data are specified using the wizard (see Wizard Reaction on page 27). Upon finishing this wizard, the program will generate the input files for MOPAC/DRIVER and run the calculation. If reactions for the set of mutants are calculated, a separate Reaction project is created for each mutant. Reaction project contains *Source structure*, *Input structure* and *Output Data*. You can switch between them using bookmarks on the left side of TRITON screen (see **Fig 6.1**). *Source structure* is a structure used as input file specified in wizard. *Input structure* is the input structure read by MOPAC/DRIVER (*.dat* extension). Input structure is generated by wizard using the information provided by the user through the wizard, e.g. atoms included in the calculation, fixed atoms, reaction co-ordinate, etc. *Output Data* consist of output structures produced during MOPAC/DRIVER calculation, i.e. the structures along the reaction pathway.

10.4 Optimization project

Optimization project is used to perform various MOPAC calculations, such as simple single point calculation, geometry optimization or their combination as a consecutive reaction. Before the very first MOPAC calculation project is created, MOPAC program must be installed and configured (see section Setup on page 10). The only input required for this project is a file with the three-dimensional structure of the system in PDB format. All input data are specified using the wizard (see Wizard Optimization on page 36). Upon finishing this wizard, the program will generate the input files for MOPAC and run the calculation. Optimized geometry or electronic distribution represented by partial charges can be then visualized using TRITON after the calculation is done. Optimized geometry can be then saved as PDB file including the important information in the file. Optimization project contains *Source structure*, *Input structure* and *Output Data*. You can switch between them using bookmarks on the left side of TRITON screen (see **Fig 6.1**). *Source structure* is a structure used as input file specified in wizard. *Input structure* is the input structure read by MOPAC (*.dat* extension). Input structure is generated by wizard using the information provided by the user through the wizard, e.g. atoms included in the calculation, fixed atoms, etc. *Output Data* consist of output structure produced by MOPAC calculation.

11 Data Analysis Tools

Calculations performed using TRITON often result in a large amount of data, which need to be conveniently visualized and interpreted. Data for analysis are extracted from the output files generated by MOPAC or MODELLER, plotted using interactive graphs and saved in ASCII files for later use. Interactive graphs include the plot for animation of the reaction, energy plot and partial charges and non-bonded interactions analysis.

Manipulating animations

The Animation Bar (**Fig 11.1**) serves for animations based on data obtained from the reaction mapping calculation. The buttons with symbols situated on the top left corner serve for running such animations. *Input* button shows the input structure used for the calculation. *Align* button can be used for manual selection of three atoms which will be fixed during the animation of the reaction. *R, TS, P* button can be used for manual setting of the reactants, transition states and products of the calculated reaction. Bottom part provides the information about the number of images, current value of driven distance, total energy (heat of formation in kcal/mol) of the visualized structure, and its location.

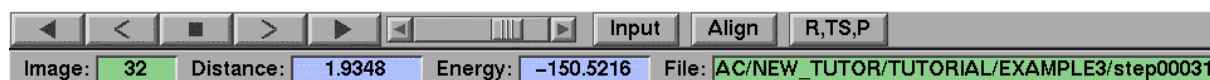


Fig 11.1 Animation Bar for manipulation with reaction mapping outputs

Energy visualization

The changes of energy during the reaction can be visualized using *Menu: Reaction/Energy Graph (Fig. 6.1)*. Energy (heat of formation) of every structure along the reaction pathway can be obtained from the graph. Geometry of every structure on the reaction pathway can be displayed in the 3D Model window by picking on the corresponding point in the graph.

Visualization of the partial charge changes

Changes of partial charges on selected atoms during the reaction can be analyzed using *Menu: Reaction/Electrostatics Progress Graph (Fig. 6.1)*. Geometry of every structure in the Electrostatics Progress Graph can be displayed in the 3D Model window by picking on the corresponding point in the graph. Progress of the partial charge of one selected atom, sum of partial charges of group of selected atoms or non-bonded electrostatic interaction energies between selected atoms during the reaction can be visualized by choosing between them in *Graph Settings...* window which appears after clicking <right mouse button>.

Visualization of the residue interactions

Non-bonded interactions between selected part of the molecule (substrate) and every amino acid in a protein can be calculated. Interacting atoms can be selected in the 3D Model Window using <Shift> button. The graph is created by *Menu: Reaction/Residue Interactions Graph*. Graph settings can be changed using *Graph Settings...* window (**Fig 11.2**). This window can be opened by <right mouse button> and choosing *Graph Settings...* option. Two different modes can be used. In *mode 1*, several parameters can be simultaneously visualized for particular reaction state (reactant - R, transition state - TS, product - P). In *mode 2*, the values of selected parameter for different reaction states can be simultaneously visualized.

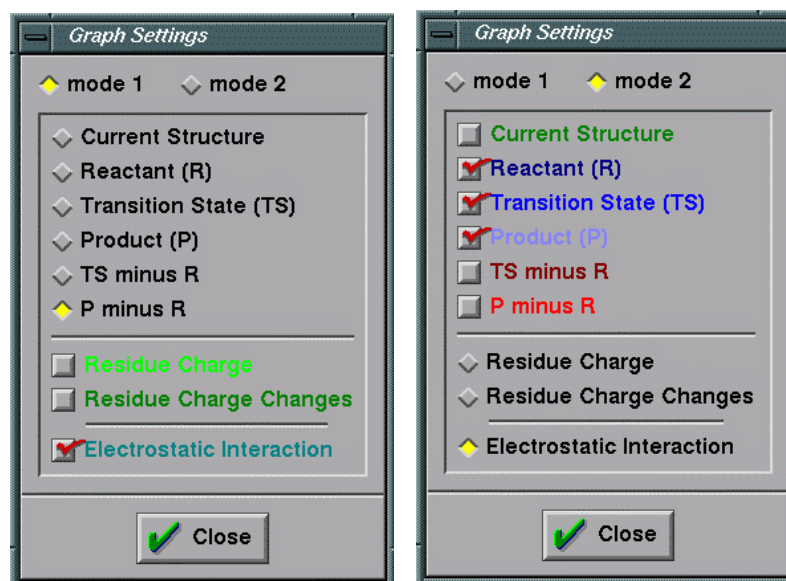


Fig 11.2 *Graph Settings windows for mode 1 and mode 2*

Extraction and saving of the data

The values can be extracted and stored from each graph by pressing <right mouse button> with mouse placed on the graph and by selecting *Write Energy File...*, *Write Charge File...* or *Write Residue Interactions...*. The values will be saved into *.txt file with the specified name.

Extraction of structures to PDB files

Output geometry from any of the steps from reaction path mapping or from simple MOPAC calculation (e.g., optimized geometry) can be saved to PDB file format by choosing *Menu: Structure/Write...* and specifying the name of PDB file.

12 Wizard Mutagenesis

IMPORTANT: Make sure that the external program MODELLER is properly installed on your system and configured using the *Setup MODELLER* window (see page 10) before running any calculation.

- 1) Roll down the *Menu: Project* and choose *New Project* option or press the button on the top left of the TRITON screen (**Fig 12.1**) to initiate the wizard.



Fig 12.1 Icon (first left) to initialize wizards

- 2) Select the *New Project* button to create a new project (**Fig 12.2**).

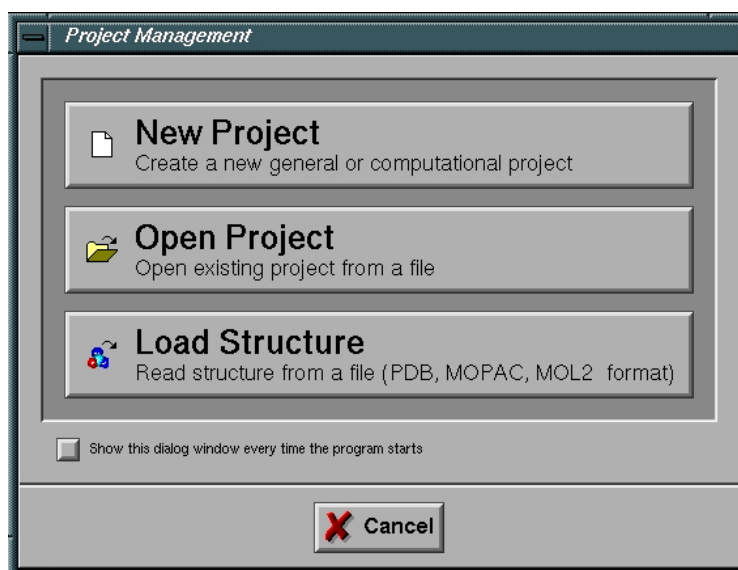


Fig 12.2 Window for project management

- 3) Press the *Mutagenesis* button in *Project Specification* window (**Fig 12.3**). This will open the Mutagenesis wizard and the introduction to this type of project will appear (**Fig 12.4**).

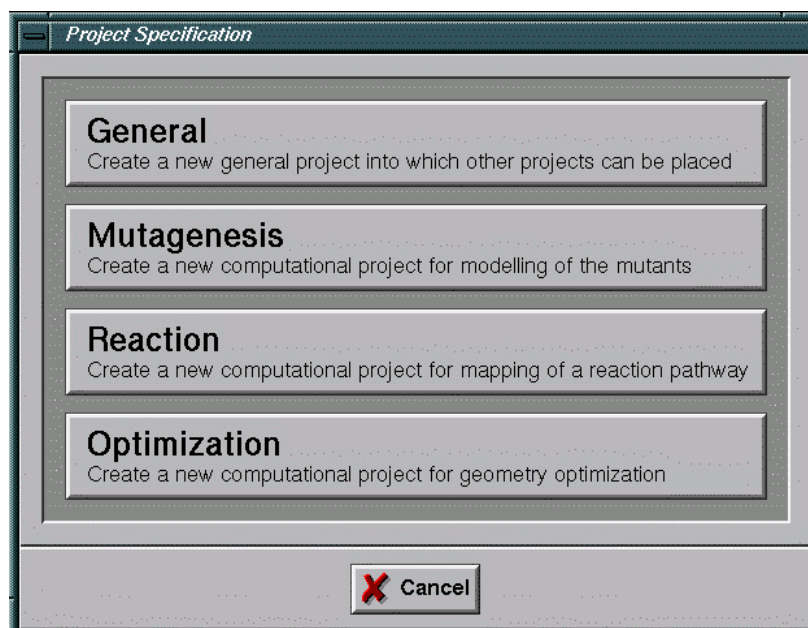


Fig 12.3 Window for specification of projects

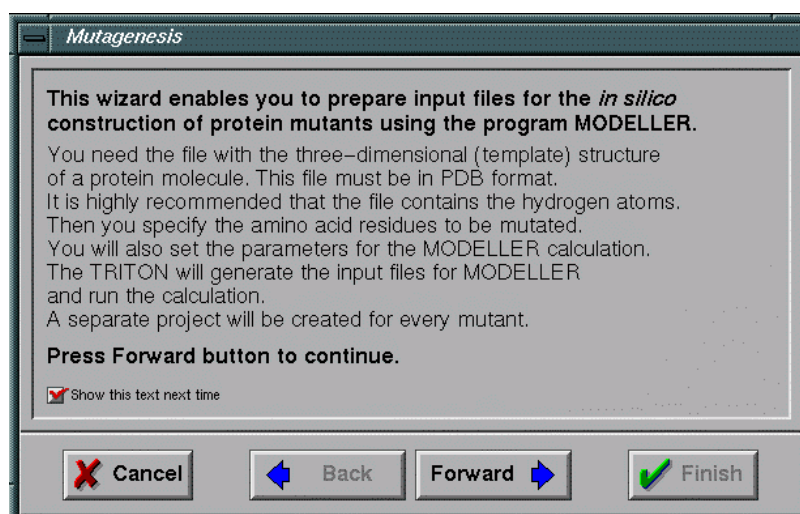


Fig 12.4 Window with introduction to mutagenesis project

- 4) Input the project name into *Project name* field. This project can be created as a independent project and placed into specified directory (**Fig 12.5**), or can be added as a subproject into opened project (**Fig 12.6**). In this particular case the project with name EXAMPLE2 is created and it is a part of general project called TUTORIAL located in *~/tmp* directory.

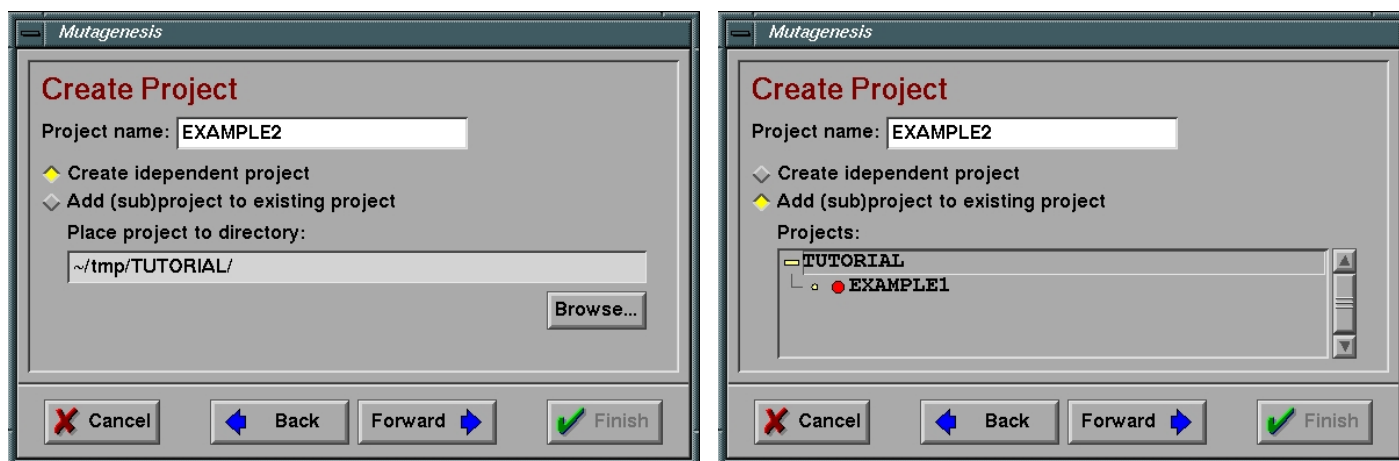


Fig 12.5 and Fig 12.6 Windows for specification of project name and its location

- 5) Specify the input structure for mutagenesis. You can read this structure from a file in PDB format (**Fig 12.7**) or take the structure from previous project (**Fig 12.8**). The source file of selected structure can be then copied into a new directory or directly linked from a source directory. Latter possibility will save space on a hard disc, but you must make sure not to delete or move the source structure file.

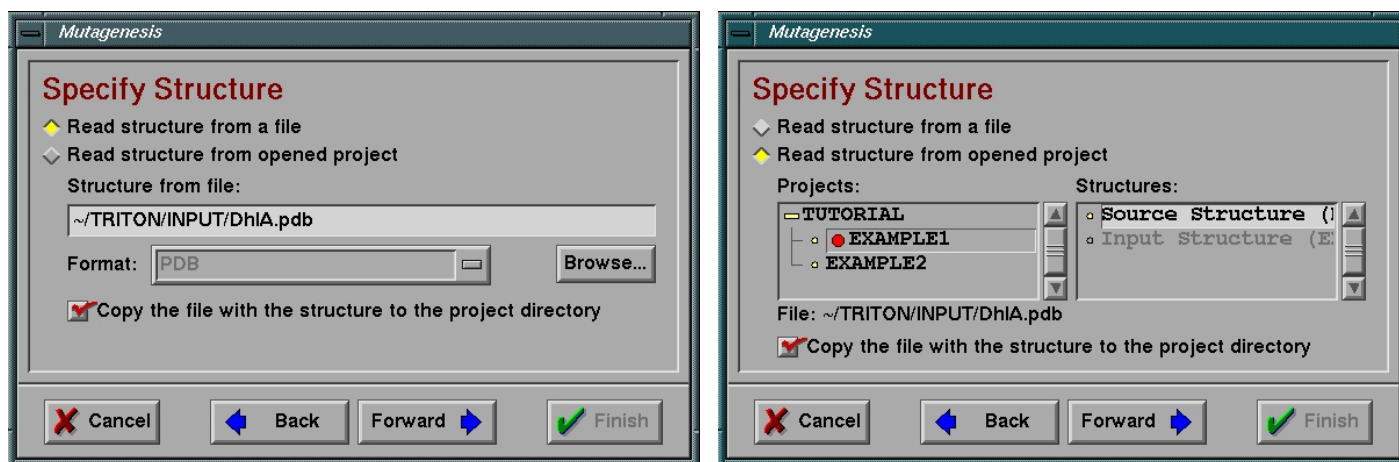


Fig 12.7 and Fig 12.8 Windows for specification of input structure for mutagenesis

- 6) Specify the residues to be mutated (**Fig 12.9**). From the window showing all available residues choose those you want to mutate and press the button with $\rightarrow$. Selected residues will move into *Selected residues* window. If you select wrong residue, you can remove it from this list by selecting it and pressing the button with $\leftarrow$.

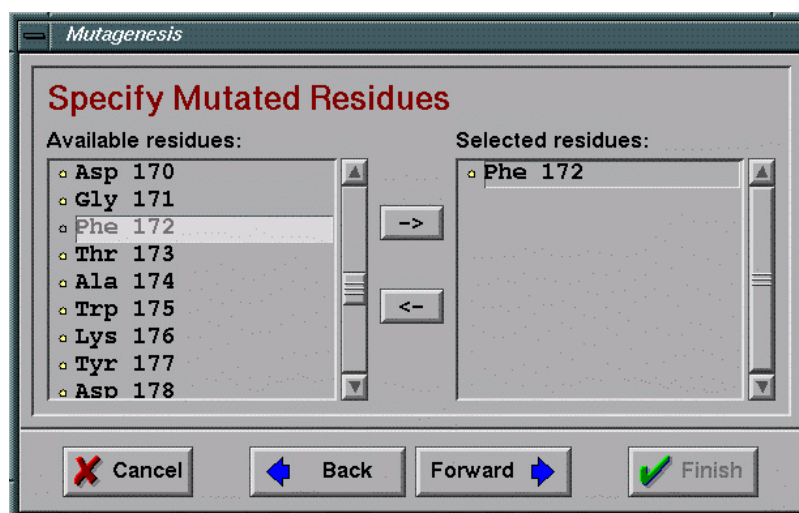


Fig 12.9 Window for specification of residues for mutagenesis

- 7) Specify the substitutions you want to construct (**Fig 12.10**). You can choose any of 20 amino acids by checking their buttons on the right side of the window. Requested substitutions will be automatically introduced in the structure of the protein. Separate subproject will be created for every mutation and process of homology modelling using MODELLER will automatically proceed.

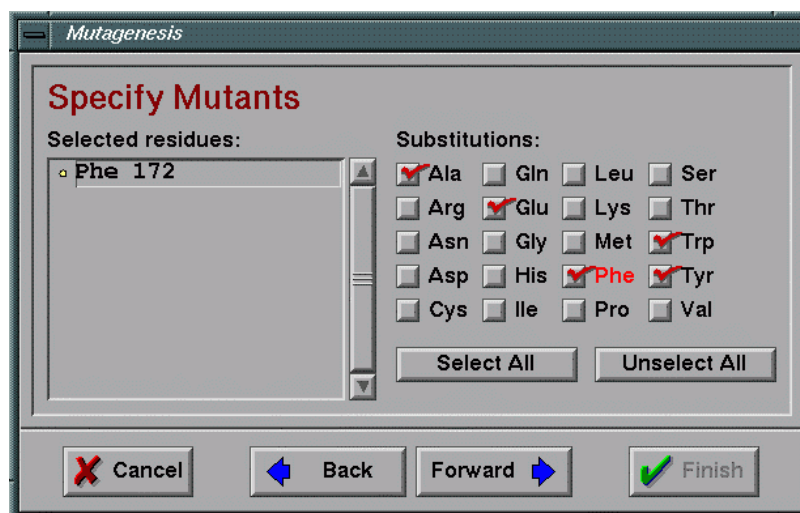


Fig 12.10 Window for specification of requested substitutions

- 8) Specify the parameters for MODELLER calculation (**Fig 12.11**). You can specify the level of the molecular dynamics refinement, allowed deviation (in Å) of the mutant structure from the template structure and the request for transfer of water molecules, heteroatoms and hydrogens from template to model. By choosing *Read (add) hydrogens* option while template has no hydrogens, MODELLER will automatically add them. For more details look into a theory section of the MODELLER manual. You have also a last chance to open *MODELLER Setup* window and configure the program (see page 10).

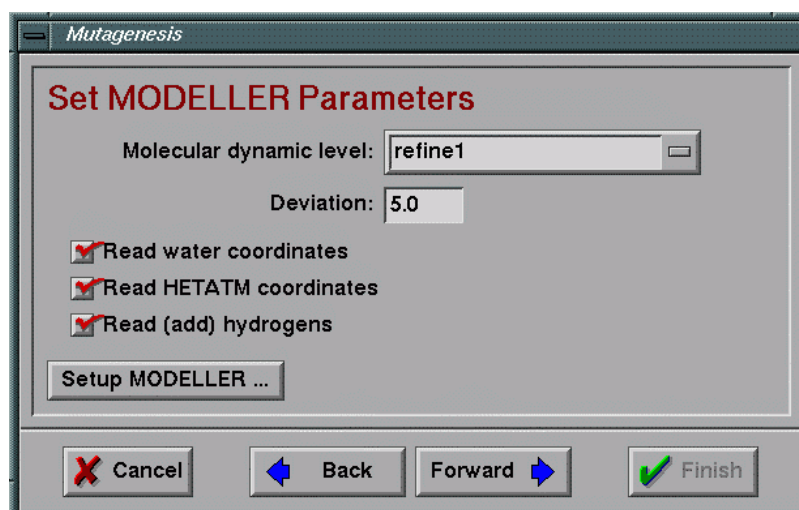


Fig 12.11 Window for specification of MODELLER parameters

- 9) Modify the MODELLER parameters file (.top) directly in TRITON (**Fig 12.12**). This option is implemented in TRITON for more experienced users who want to modify the parameters for calculation. Default parameters are already pre-adjusted for smooth operation of MODELLER program.

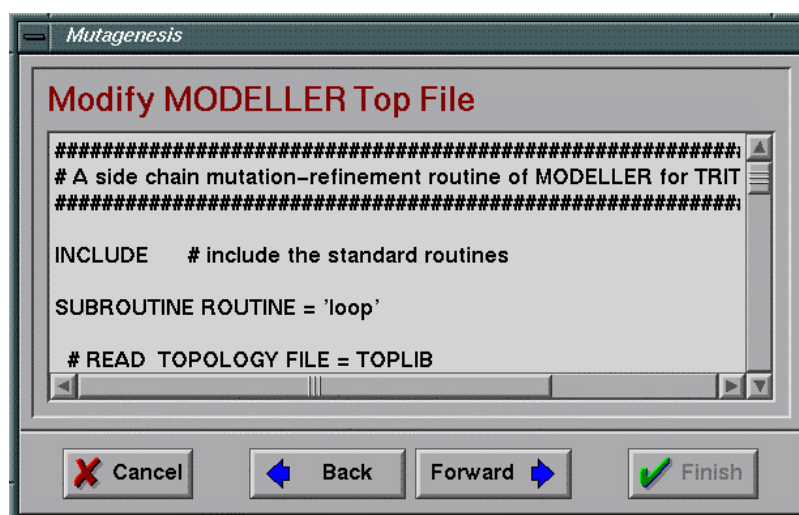


Fig 12.12 Window for specification of parameters for MODELLER

- 10) Finalize the mutagenesis project (**Fig 12.13**). All parameters necessary for calculation were entered and the user can either start the calculation on current computer or generate the input files, move them to another computer and run the calculation there. Press the *Finish* button to complete the wizard.

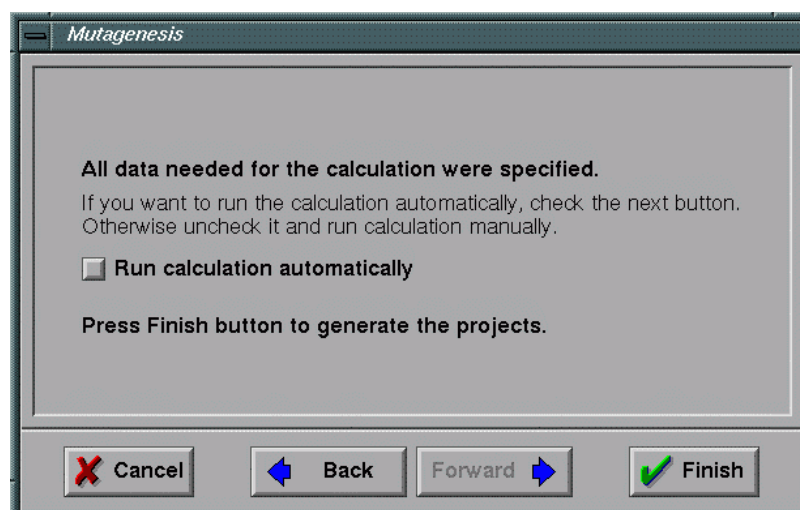


Fig 12.13 Window to finalize wizard mutagenesis

13 Wizard Reaction

IMPORTANT: Make sure that the external programs MOPAC and DRIVER are properly installed on your system and configured using the *Setup MOPAC* and *Setup DRIVER* window (see page 11) before running any calculation.

- 1) Roll down the *Menu: Project* and choose *New Project* option or press the button on the top left of the TRITON screen (**Fig 13.1**) to initiate the wizard.



Fig 13.1 Icon (first left) to initialize wizards

- 2) Select the *New Project* button to create a new project (**Fig 13.2**).

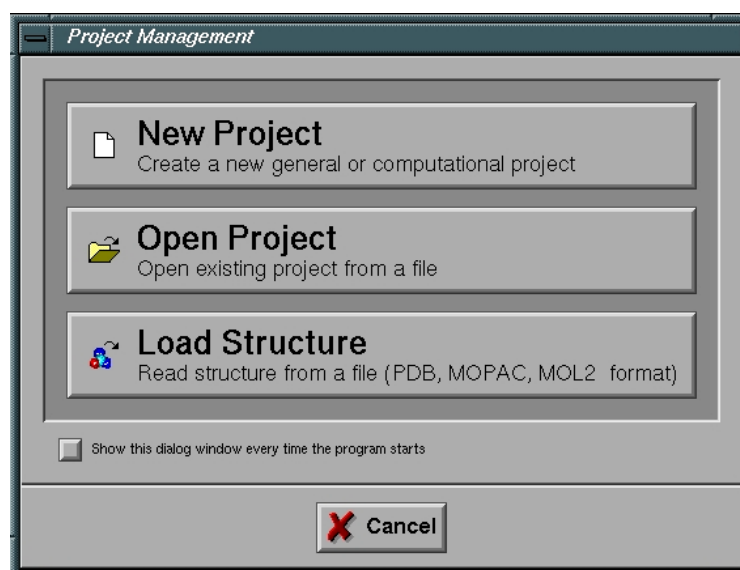


Fig 13.2 Window for project management.

- 3) Press the *Reaction* button in *Project Specification* window (**Fig 13.3**). This will open the Reaction wizard and the introduction for this type of project will appear (**Fig 13.4**).

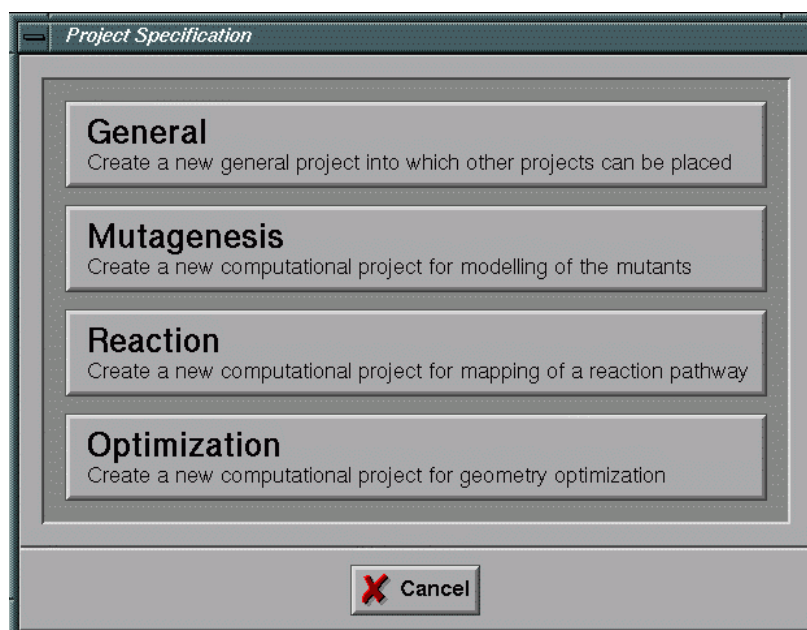


Fig 13.3 Window for specification of projects

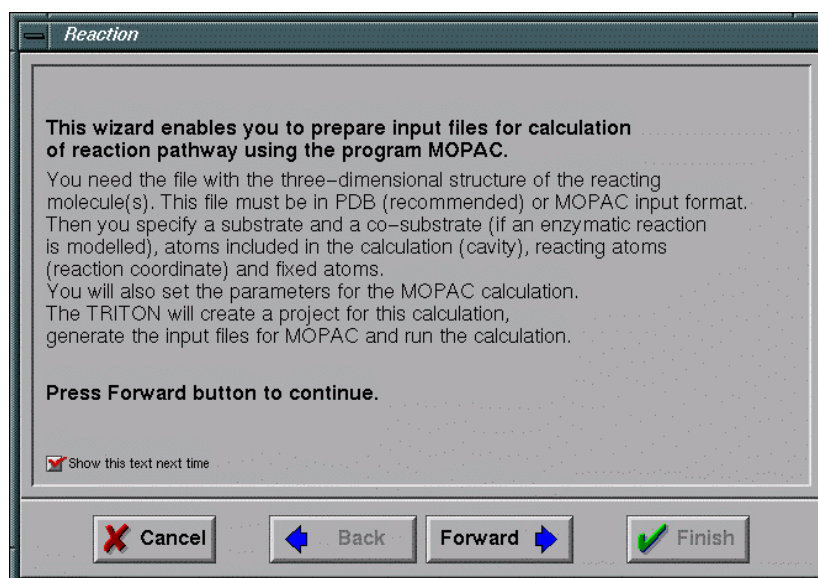


Fig 13.4 Window with introduction to reaction project

- 4) Specify the type of reaction you wish to model using MOPAC/DRIVER (**Fig 13.5**). You can choose between enzymatic and chemical reaction. *Enzymatic reaction* assumes that you are dealing with protein structure consisting of amino acids and ligands, while *Chemical reaction* assumes the study of small organic molecules.

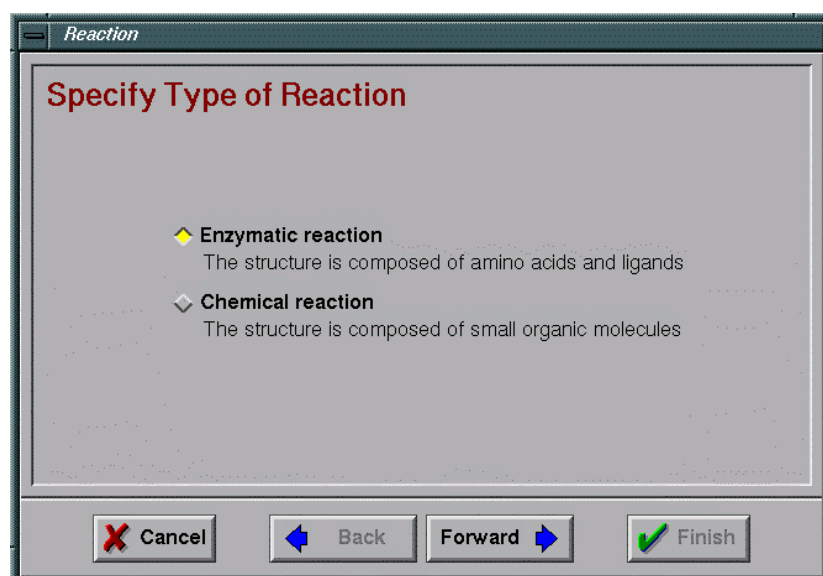


Fig 13.5 Window for specification of reaction type

- 5) Specify the number of protein structures you will use for reaction mapping (**Fig 13.6**). *Single protein* option allows you to perform calculation with only one protein, while the *Set of Proteins* option allows you to perform calculations on entire set of proteins originating from previous mutagenesis project. *Set of Proteins* option is unavailable if no previous mutagenesis project is present in any of TRITON opened projects.

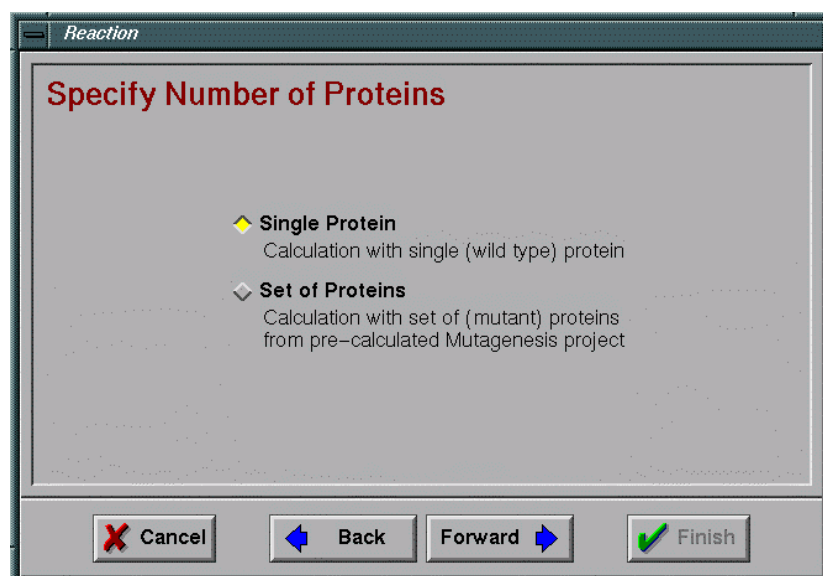


Fig 13.6 Window for specification of the number of protein structures

- 6) Input the project name into *Project name* field. You can create a new independent project (**Fig 13.7**), add a new project into an existing project or create a new multi-project within existing mutagenesis project (**Fig 13.8**). Using multi-project option will automatically add reaction subproject with corresponding input data for the calculation into each of the mutagenesis projects. The name of the reaction subproject should be specified in following window (**Fig 13.9**).

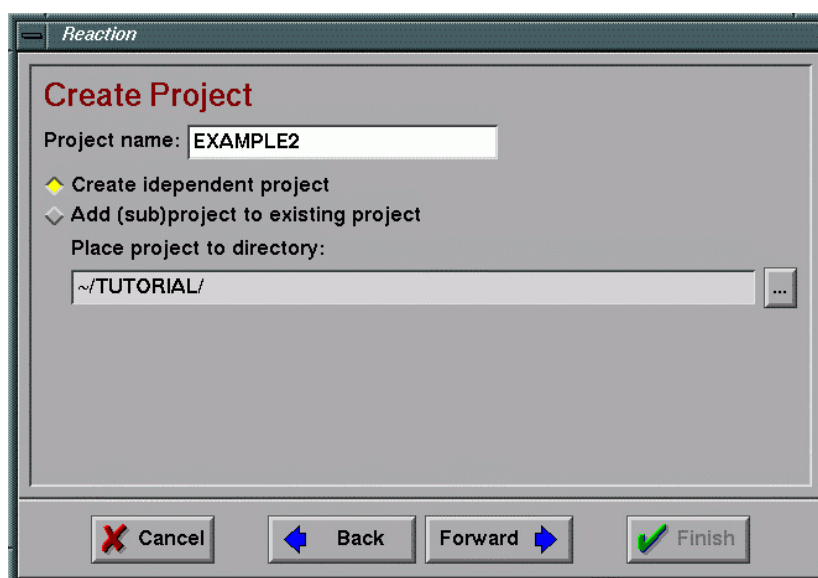


Fig 13.7 Window for specification of project name and its location

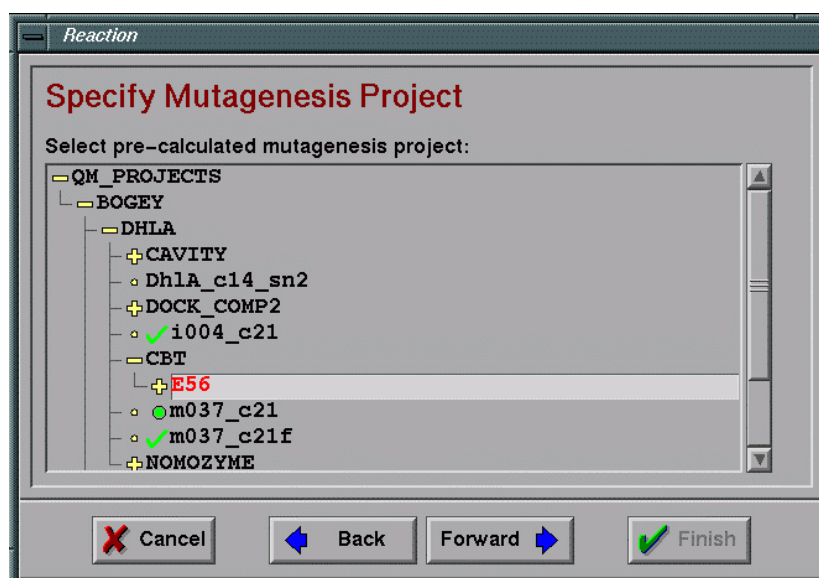


Fig 13.8 Window for specification of source mutagenesis project used for preparation of reaction project.

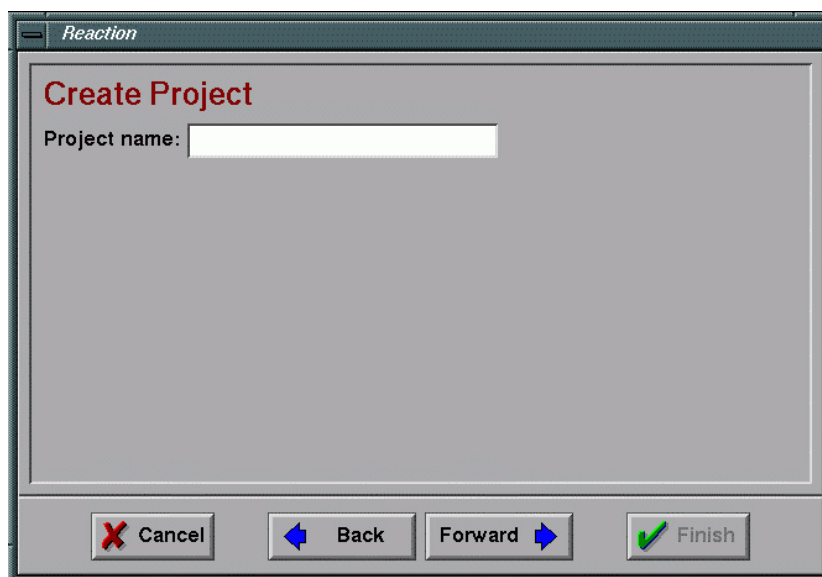


Fig 13.9 Window for specification of reaction project name

- 7) Specify the file name and location of input structure for reaction mapping. You can read this structure from a file (**Fig 13.10**) or use the structure from opened project (**Fig 13.11**). Supported file format is PDB containing information about residues.

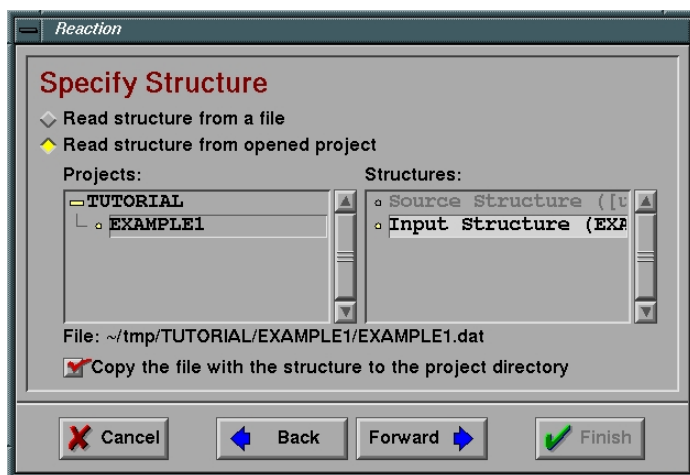
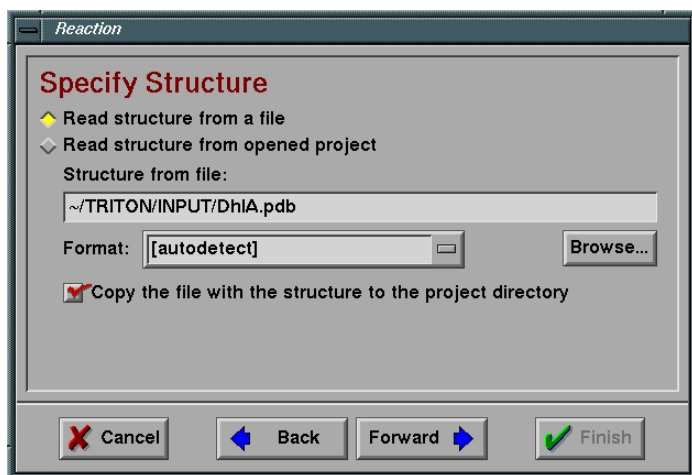


Fig 13.10 and Fig 13.11 Windows for specification of input structure for reaction mapping

- 8) Select the substrates or co-substrates for the reaction (**Fig 13.12**). They will become a part of the active site model and will be included in the calculation. Selection can be achieved in several different ways. The molecule can be selected from *Ligands* list containing recognized “hetero-residues”, i.e. listed in PDB file as HETATM. You can also open the list of residues by checking the button *Show Residue List* and choose any other hetero-residue by clicking on it while holding <Ctrl>.

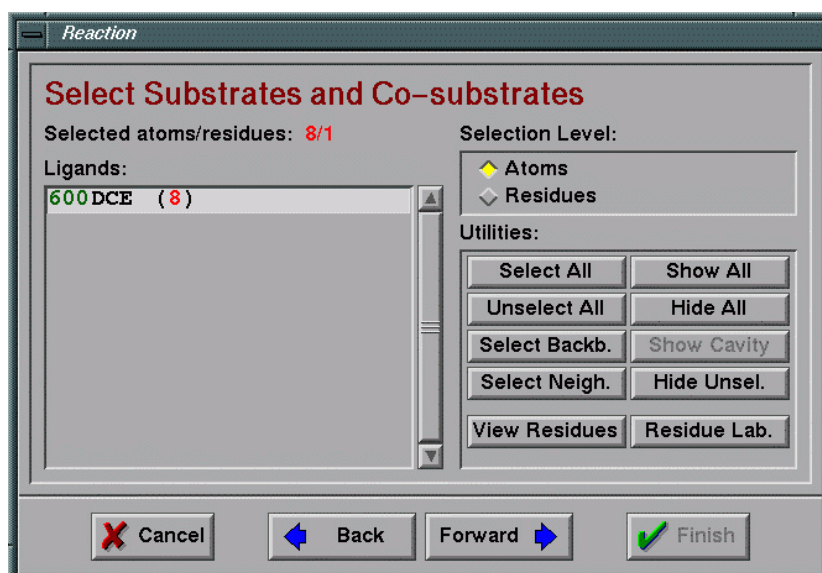


Fig 13.12 Window for selection of substrate and co-substrate

- 9) Select the amino acid residues for the reaction (**Fig 13.13**). These residues will become a part of the active site model. Selection can be achieved in several different ways. You can choose the residues by clicking on their names in the *Residues* list. The number in brackets behind the residue code shows the number of atoms forming this residue. You can also open the list of residues by checking the button *View Residues* and choose residue by clicking on it while holding <Ctrl>. Another option for selecting the model is to use *Select Neigh.* option, which allows you to select the residues within specified radius from defined atom/residue. The number *Selected atoms/residues* shows the total number of atoms/residues in selected model. All selected residues are also shown in 3D model window.

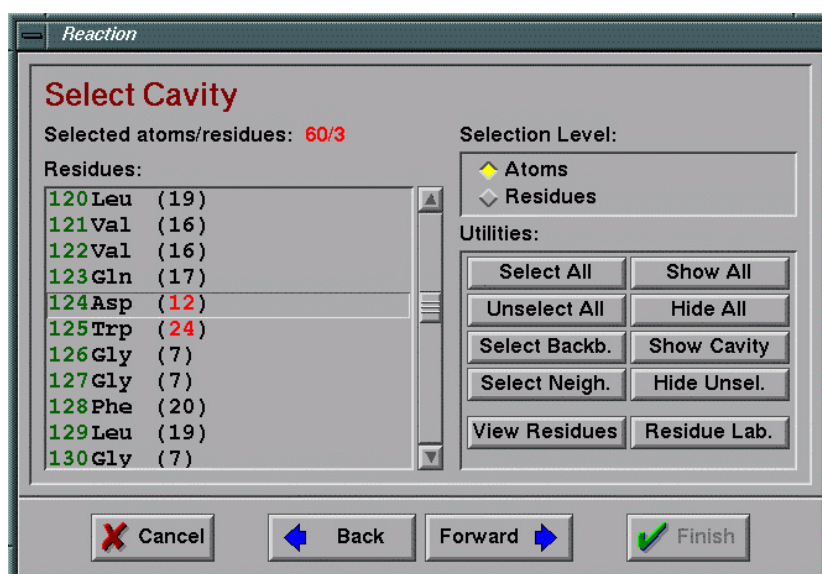


Fig 13.13 Window for selection of residues for cavity

- 10) Specify two atoms, which will interact and create a new bond. Distance of these two atoms will be considered as reaction co-ordinate (**Fig 13.14**). The distance will be reduced/increased step by step simulating the chemical reaction. You can specify the two reacting atoms by clicking on them in 3D-model window. In the case shown below, the atom 2 will interact with the atom 1985. You must also specify the size of driving step, i.e. the change in distance between two reacting atoms. Default value is 0.05 Å. The value must be negative if the atoms should be approaching each other. The value of *Dist.* defines the final distance between the two approaching atoms and it is automatically set to be smaller than the optimal distance of the formed bond.

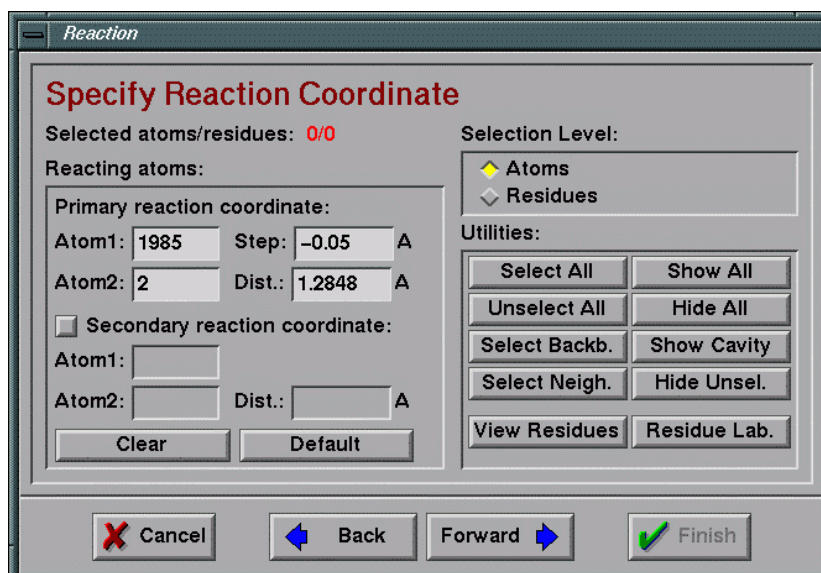


Fig 13.14 Window for reaction co-ordinate specification

- 11) Select atoms to be fixed during the reaction mapping (**Fig 13.15**). Some atoms of the cavity must be fixed to keep the overall geometry of the enzyme cavity intact. It is recommended to fix the heavy atoms of the backbone and let the side chains move during geometry optimization using *Select Backb.* option. Such a model provided satisfactory results for simulating conditions in the enzyme active site during the catalysis.

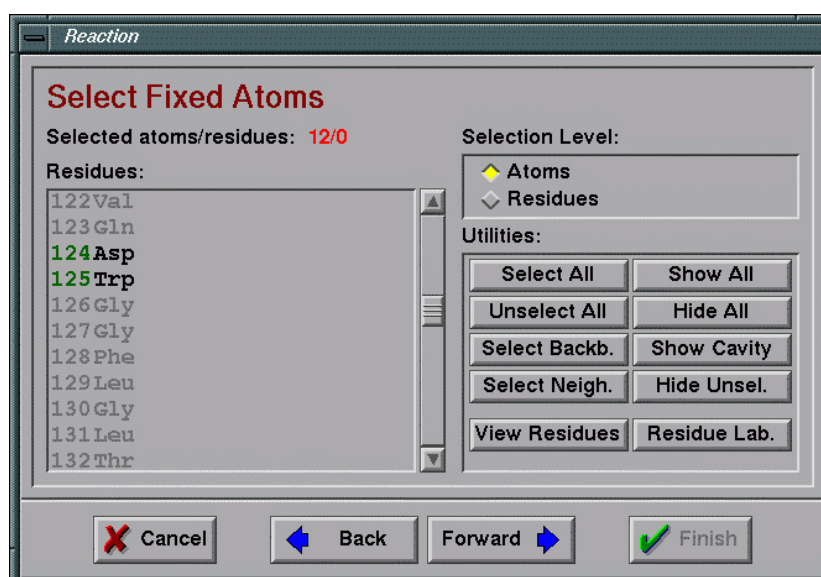


Fig 13.15 Window for selection of fixed atoms

- 12) Specify the parameters for MOPAC calculation (**Fig 13.16**). *Total charge* defines the total charge of studied model. It is dependent on the amino acids in the model. *Method* defines the type of semi-empirical Hamiltonian to be used for the calculation (AM1, PM3 or PM5). For *SCF* procedure conventional or much faster MOZYME option can be chosen (see *Introduction* section for more details). BFGS is suggested *Optimizer* for this type of calculations but the EF can be also chosen here. Next, some other KEYWORDS for MOPAC calculation must be specified. Optimal set of KEYWORDS is checked by default. For more details see MOPAC manual. You are also given a last chance to open *MOPAC Setup* window and *DRIVER Setup* window and configure the programs (see page 10).

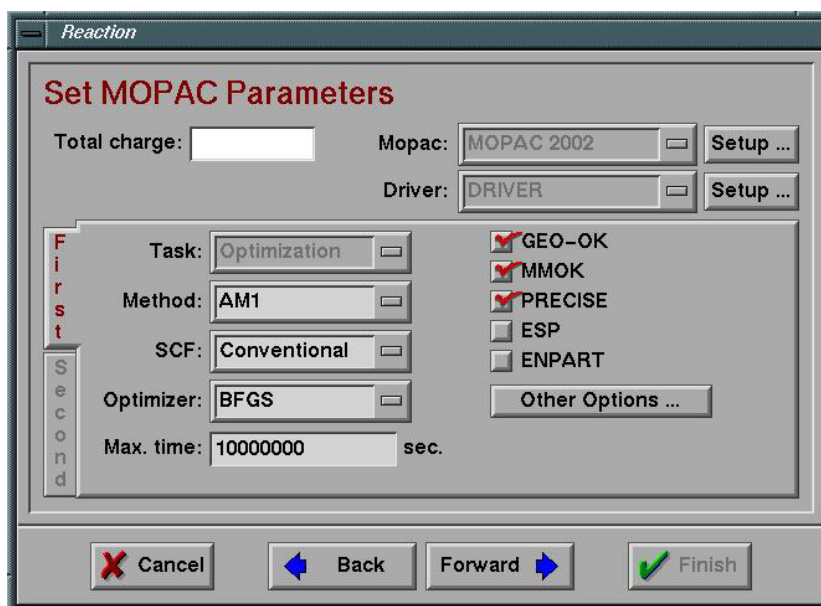


Fig 13.16 Window for specification of MOPAC keywords

- 13) Input additional information consisting commentary or additional keywords for MOPAC (**Fig 13.17**). This input is optional.

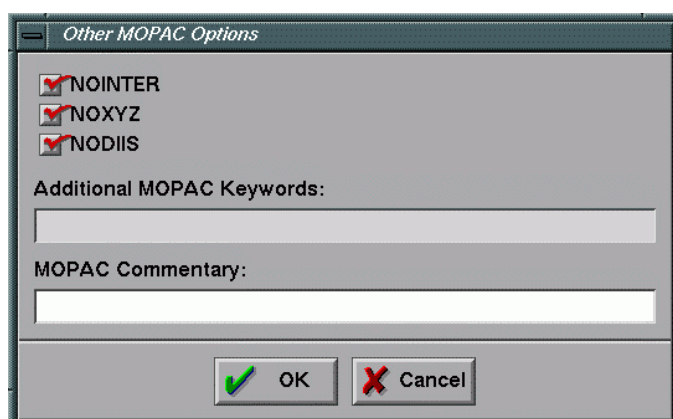


Fig 13.17 Window for additional information

- 14) Finalize the reaction project (-**Fig 13.18**). All parameters necessary for calculation were entered and the user can either start the calculation on current computer or generate the input files, move them to another computer and run the calculation there. Press the *Finish* button to complete the wizard.

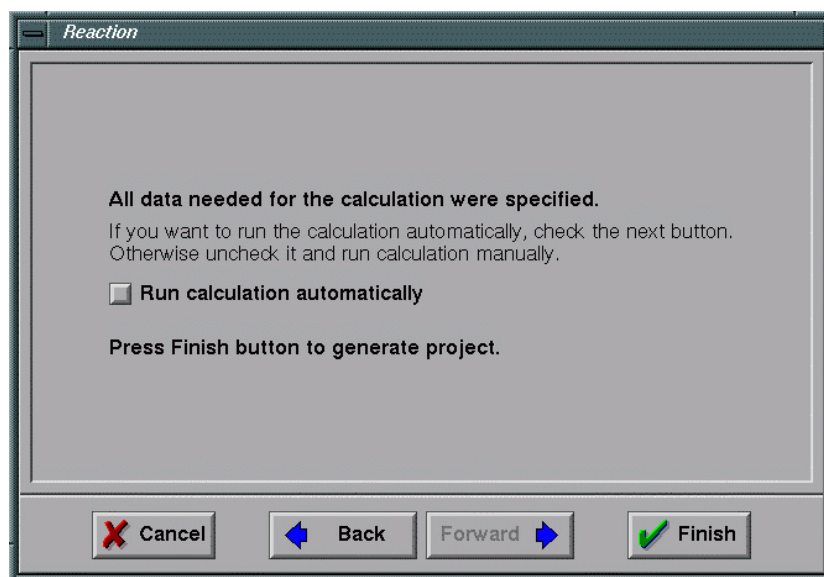


Fig 13.18 Window for finalizing of reaction wizard

14 Wizard Optimization

IMPORTANT: Make sure that the external program MOPAC is properly installed on your system and configured using the *Setup MOPAC* window (see page 11) before running any calculation.

- 1) Roll down the *Menu: Project* and choose *New Project* option or press the button on the top left of the TRITON screen (**Fig 14.1**) to initiate the wizard.



Fig 14.1 Icon (first left) to initialize wizards

- 2) Select the *New Project* button to create a new project (**Fig 14.2**).

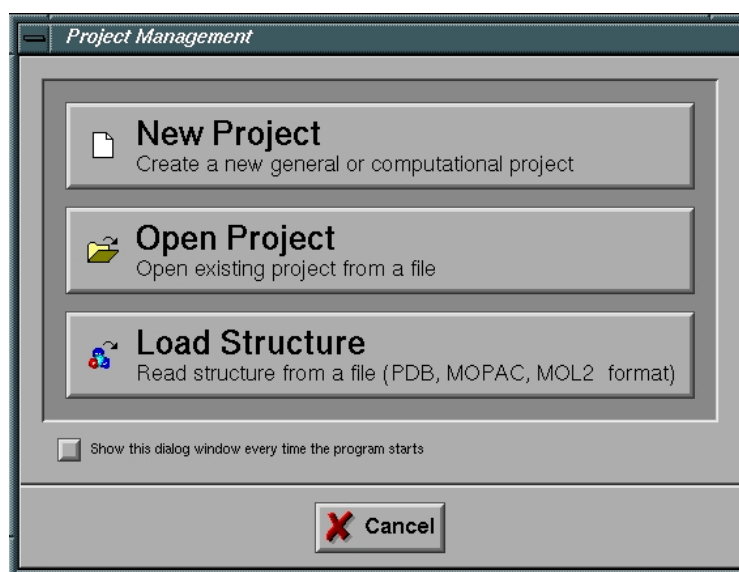


Fig 14.2 Window for project management

- 3) Press the *Optimization* button in *Project Specification* window (**Fig 14.3**). This will open the Optimization wizard and the introduction for this type of project will appear (**Fig 14.4**).

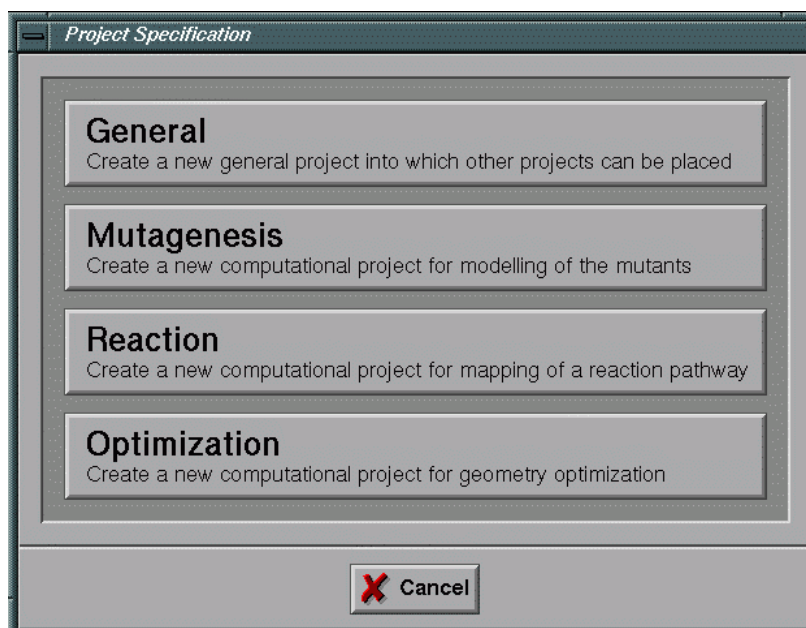


Fig 14.3 Window for specification of projects

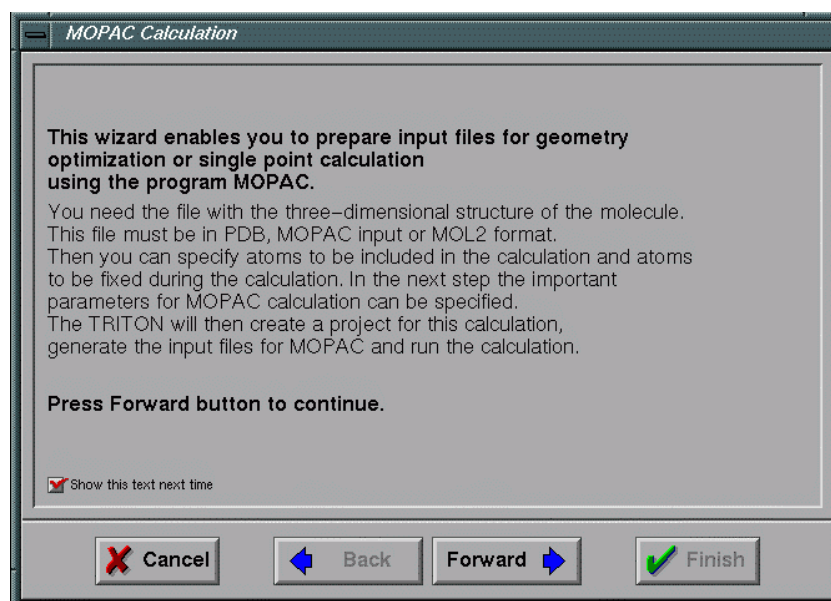


Fig 14.4 Window with introduction to optimization project

- 4) Input the project name into *Project name* field (**Fig 14.5**). This will name the project and all input files for the calculation. Then the project should be placed into the specified directory.

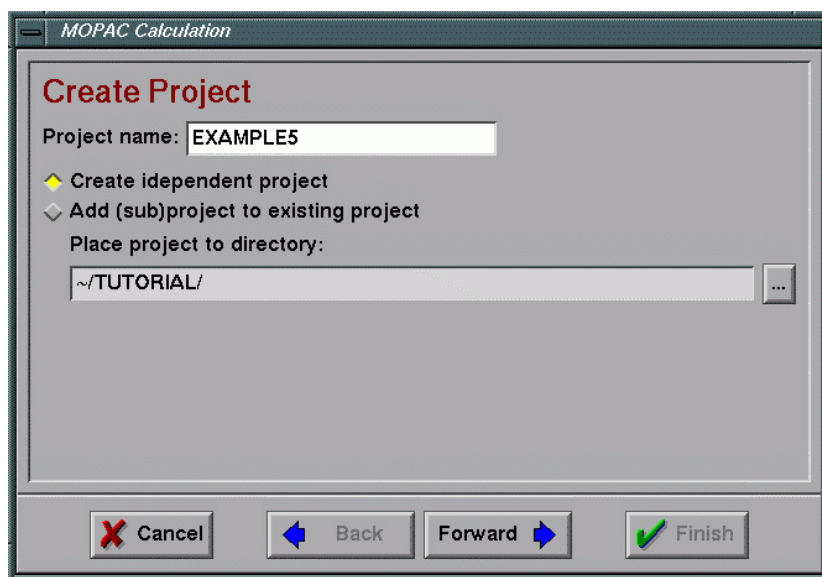


Fig 14.5 Window for specification of project name and location

- 5) Specify the file name and location of input structure for MOPAC calculation (**Fig 14.6**). You can read this structure from a file in PDB format. The structure must be properly protonated before reading it into TRITON!

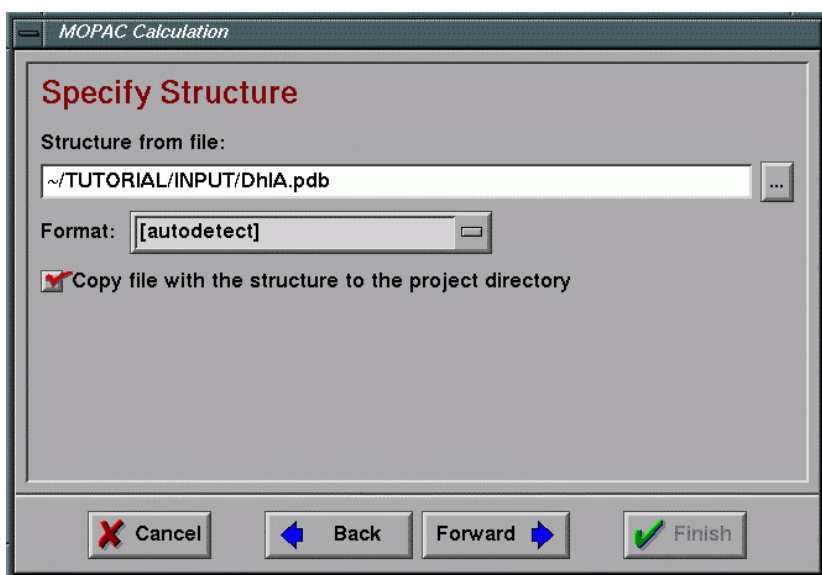


Fig 14.6 Window for specification of input structure for optimization project

- 6) Select the residues to be included into the calculation (**Fig 14.7**). These residues will become a part of the model and will be included into the calculation. Selection can be achieved in several different ways. The residue can be selected from *Residues* list containing all residues listed in PDB file. The number in brackets behind the residue code

in *Residue* list shows the number of atoms forming this residue. You can also open the list of residues by checking the button *View Residues* and choose residue by clicking on it while holding <Ctrl>. Another option for selecting the model is to use *Select Neigh.* option, which allows you to select the residues within specified radius from defined atom/residue. The number behind *Selected atoms/residues* shows the total number of atoms/residues in selected cavity. All selected residues are also shown in 3D model window.

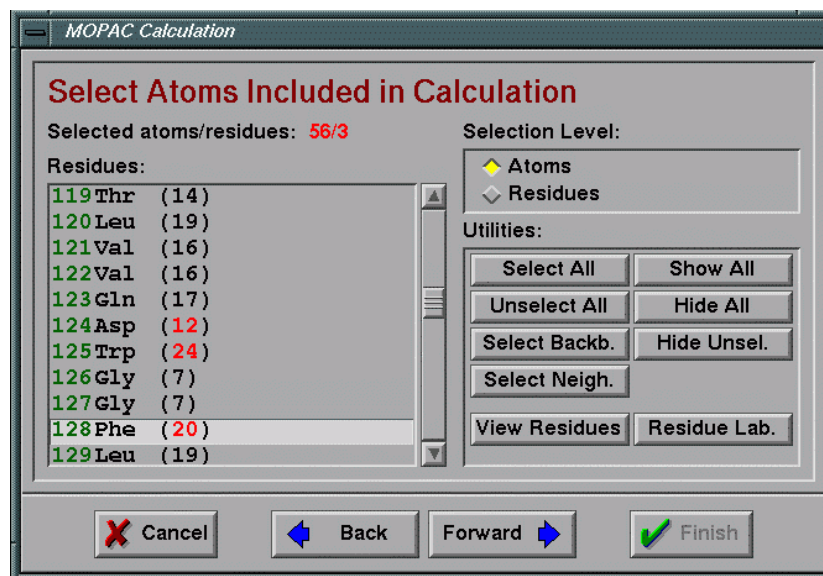


Fig 14.7 Window for selection of residues used for calculation

- 7) Select the atoms to be fixed during the calculation (**Fig 14.8**). Some atoms of the model may need to be fixed in order to simulate specific structures such as protein backbone. It is recommended to fix the heavy atoms of the backbone and let the side chains move during geometry optimization. Such a model provided satisfactory results for simulating conditions in the enzyme active site during the catalysis.

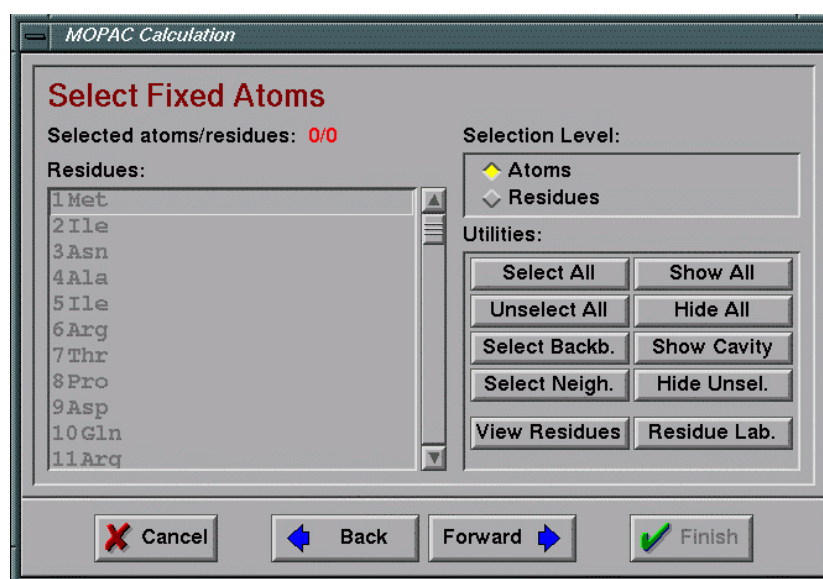


Fig 14.8 Window for selection of fixed atoms

- 15) Specify the parameters for MOPAC calculation. The consecutive reaction can be performed here, for example an initial optimization using PM3 method (**Fig 14.9**) followed by a single point calculation using AM1 method (**Fig 14.10**). *Total charge* defines the total charge of studied system. It is dependent on the amino acids in the model. *Method* defines the type of semi-empirical Hamiltonian to be used for the calculation (AM1, PM3 or PM5). For *SCF* procedure the Conventional or much faster MOZYME option can be chosen (see *Introduction* section for more details). BFGS is suggested *Optimizer* for this type of calculations but the EF can also be chosen instead. Next, some other KEYWORDS for MOPAC calculation must be specified. Optimal set of KEYWORDS is checked by default. For more details see MOPAC manual. You are also given a last chance to open *MOPAC Setup* window and *DRIVER Setup* window and configure the programs (see page 11).

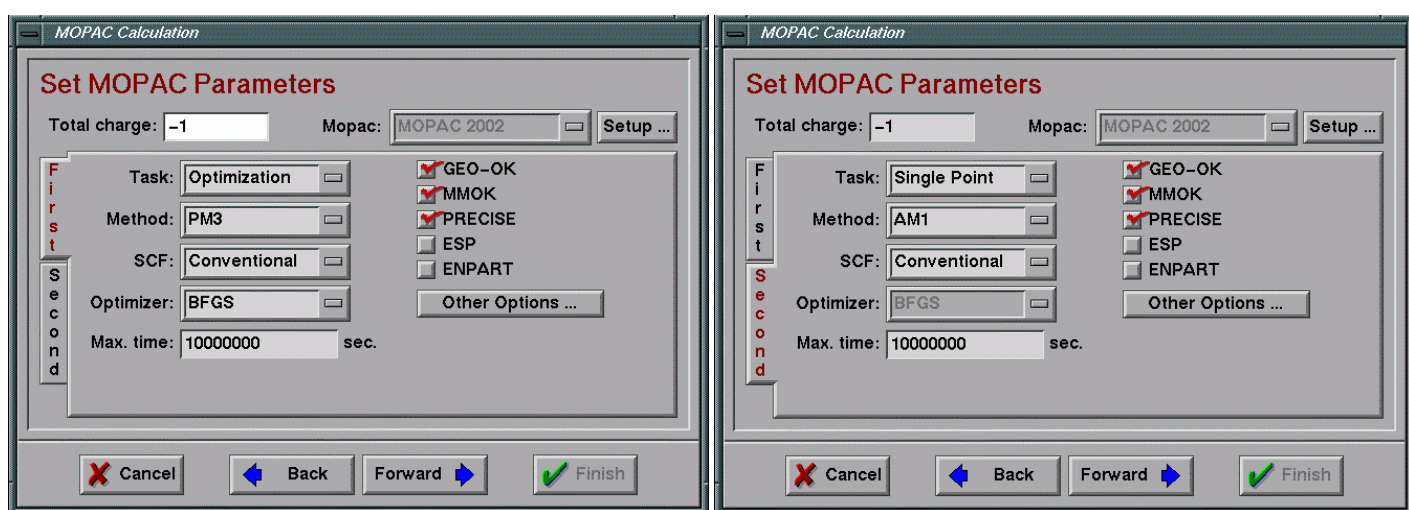


Fig 14.9 and Fig 14.10 Windows for specification of MOPAC parameters

- 8) Input additional information consisting commentary or additional keywords for MOPAC (**Fig 14.11**).

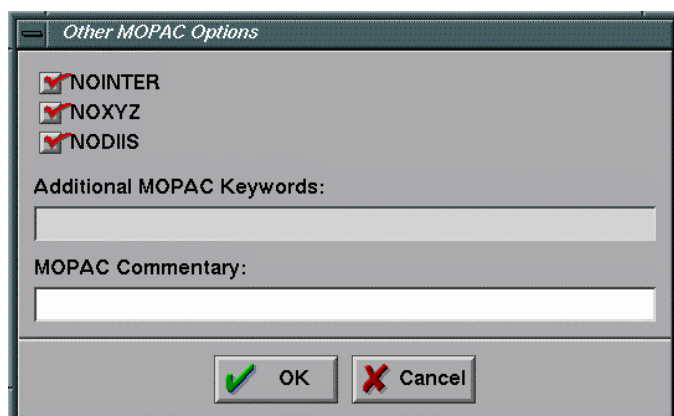


Fig 14.11 Window for additional information

- 9) Finalize the optimization project (**Fig 14.12**). All parameters necessary for calculation were entered and the user can either start the calculation on current computer or generate the input files, move them to another computer and run the calculation there. Press the *Finish* button to complete the wizard.

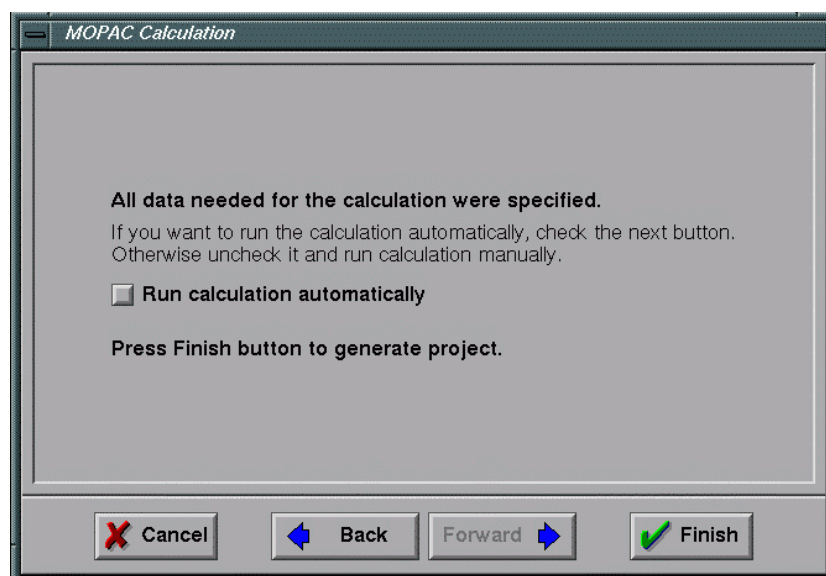


Fig 14.12 Window for finalizing of the optimization wizard

15 Tutorials

The purpose of tutorials is to introduce to user the program TRITON using the examples from “real” studies published in peer-reviewed literature. All files needed to run tutorials are provided in the TRITON distribution package. The subdirectory *tutorial* will be created in the directory */usr/doc/triton-3.0* during the installation. Tutorial contains subdirectories INPUT and EXAMPLE1 to EXAMPLE5. Subdirectory INPUT contains five PDB files: (i) *2dhc.pdb* as the original PDB file of the protein haloalkane dehalogenase from the soil bacterium *Xanthobacter autotrophicus* GJ10 (DhlA) complexed with the substrate 1,2-dichloroethane (DCE) directly downloaded from the Protein Databank <http://www.rcsb.org/pdb>, (ii) *DhlA.pdb* representing the same protein structure with added hydrogen atoms and removed water molecules, (iii) *DhlA_CBT.pdb* representing protein DhlA including docked substrate 1-chlorobutane (CBT) and two catalytic waters from the active site and added hydrogen atoms (iv) *LinB_CBT.pdb* representing the haloalkane dehalogenase from soil bacterium *Sphingomonas paucimobilis* UT26 (LinB) taken from high resolution X-ray structure (Oakley, A., et al., unpublished yet) including two catalytic waters, later docked substrate CBT and added hydrogen atoms and (v) *LinB_Cl.pdb* representing the same high-resolution structure of LinB with experimentally found chloride and water molecules with hydrogen atoms added to the whole system. All polar hydrogens were added to the protein using the WHATIF software package [7] and non-polar hydrogens using the program XLEAP [8]. The hydrogens can be also added directly within TRITON using the MODELLER, but this routine has not been extensively validated yet. Thus it is strongly recommended to protonate the structure before reading it into TRITON package. The subdirectory *tutorial* further contains the file *TUTORIAL.prj* which is a general project consisting of all five pre-calculated tutorial examples (*EXAMPLE1.prj*, *EXAMPLE2.prj*, *EXAMPLE3.prj*, *EXAMPLE4.prj* and *EXAMPLE5.prj*). These examples can be used to verify correctness of your tutorial calculations. *TUTORIAL.prj* can be opened directly from TRITON using *Menu: Help/Tutorial*.

15.1 Tutorial 1

Construction of single-point mutant proteins

Background

This tutorial will show how to construct single-point mutants of an enzyme. Current example is based on the study of Damborský *et al.* [5] who conducted saturated mutagenesis in the position 172 of the haloalkane dehalogenase from *Xanthobacter autotrophicus* GJ10. The purpose of the study was to investigate the role of the active site residue Phe172 for the dehalogenation reaction and to validate modelling procedure against the experimental data (see page 7).

Setup

- Change to the directory you want to work in (e.g., *~/tmp*).
- Start the program TRITON by typing ***triton &***.
- Create a new general project called ***MY_TUTORIAL*** (**Fig 12.1**, **Fig 12.2** and **Fig 12.3**).
- Create a new mutagenesis project, give it a name ***EXAMPLE1*** (**Fig 12.1**, **Fig 12.2** and **Fig 12.3**) and place it in ***MY_TUTORIAL*** project as its new subproject (**Fig 12.6**).
- Specify the ***Dh1A.pdb*** file from the ***INPUT*** directory as source structure (**Fig 12.7**).
- Choose the residue Phe172 as a residue to be mutated in this example (**Fig 12.9**).
- Choose the amino acids for substitutions: Asn, Gly, Glu, Trp and Phe^{wt} (**Fig 12.10**).
- Set up the MODELLER run using following parameters: **refine1** and **deviation 5.0**. Uncheck the *Read water co-ordinates* button since waters are not included in ***Dh1A.pdb*** and should not be included in the MODELLER calculation. Check the *Read HETATM co-ordinates* and *Read (add) hydrogens* buttons since heteroatoms and hydrogens are present in ***Dh1A.pdb*** (**Fig 12.11**).
- Finish the preparation of calculation without making any changes in a ***.top*** file (**Fig 12.12**) and run the calculation automatically (**Fig 12.13**).
- If you want to skip the calculations for this tutorial, go to subdirectory ***/usr/doc/triton-3.0/tutorial/EXAMPLE1*** and read ***EXAMPLE1.prj*** into TRITON.

Timing

The calculation for a single mutant takes around one minute on SGI Octane workstation with 1360MHz R12000 processor and 512 MB RAM. Output data for a single mutant take 1 to 2 MB of space on the hard drive.

Analysis

The 3D structures of single-point mutants of the haloalkane dehalogenase in PDB format represent the direct output of this calculation. The structures can be manipulated, visualized and aligned within TRITON. Most importantly, the structures can be further used for reaction mapping and prediction of their relative activities as presented on page 6.

15.2 Tutorial 2

Modelling of enzymatic reaction using a minimal model

Background

This tutorial will show how to model a biochemical reaction. The microbial enzymes haloalkane dehalogenases catalyze dehalogenation reaction by S_N2 reaction mechanism (**Fig 15.1**). The active-site aspartic acid (Asp124) serves as a nucleophile during the reaction and attacks the carbon atom attached to the leaving halogen atom. The simplest model which is essential for mapping of this reaction is composed of the nucleophile and the substrate (1,2-dichloroethane; DCE) neglecting the rest of the enzyme [9].

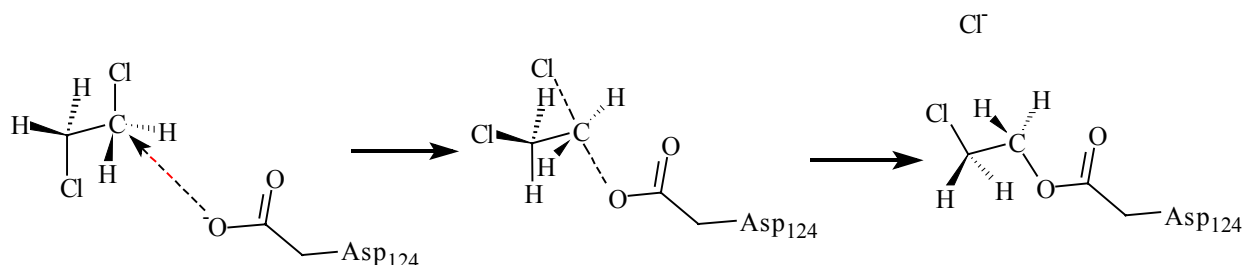


Fig 15.1 Scheme of the dehalogenation reaction using a minimal model

Setup

- Start TRITON by typing **triton &** in your MY_TUTORIAL directory.
- Open general project MY_TUTORIAL (**Fig 13.1** and **Fig 13.2**).
- Start a new reaction project (**Fig 13.2** and **Fig 13.3**) and specify its type as *Enzymatic reaction* (**Fig 13.5**) and a number of proteins as *Single protein* (**Fig 13.6**).
- Name the project **EXAMPLE2** and place it into MY_TUTORIAL project as its subproject (**Fig 13.7**).
- Take the *Dh1A.pdb* file from INPUT directory as source structure (**Fig 13.10**).
- Check the *DCE* molecule and set it as the substrate molecule in *List of Ligands* (**Fig 13.12**).
- Check the Asp124 and set it as the residue, forming cavity in *List of Residues* (**Fig 13.13**).
- Specify the atoms forming the reaction co-ordinate (**Fig 13.14**). Choose the carbon atom on DCE, which is closer to Asp124 and binds Cl (C2) and the oxygen on Asp124 (O1985) by clicking on them in 3D structure window. The original distance between them is 3.48 Å (see in **Fig 11.1**) and final distance after driving is 1.28 Å (see “Dist.” in **Fig 13.14**). Set the DRIVER size as **-0.05 Å** (atoms are approaching and their distance shortens).
- You do not need to select any atoms to be fixed during the calculation. Your model has only one amino acid residue and you do not need to simulate mutual position of amino acids in active site of enzyme using fixation (**Fig 13.15**).
- Setup the MOPAC/DRIVER calculation parameters (**Fig 13.16**). Set the total charge of the system to **-1** (Asp124 makes the system negatively charged), *SCF* to **Conventional** and *Optimizer* to **BFGS**. Select the **AM1 Hamiltonian** and increase the maximal

calculation time up to 10^8 seconds. Check the KEYWORDS: **GEO-OK**, **MMOK**, **PRECISE**, **NOINTER**, **NOXYZ** and **NODIIS** (*Fig 13.17*).

- Input optional commentary and finish the reaction wizard (*Fig 13.17* and *Fig 13.18*). If you are working on the machine on which you want to perform the calculations and you have MOPAC and DRIVER properly configured, let the calculation run automatically. If you want to run the calculation on different machine, copy the whole directory EXAMPLE2 onto this machine. Note that MOPAC and DRIVER must be installed there too! Then start the calculation by executing the *run.sh &*.
- Follow the calculation by examining the *run_output* file.
- Once the calculation is successfully finished, copy the whole directory EXAMPLE2 back onto the machine with TRITON installed, put it into your tutorial directory (MY_TUTORIAL) and analyze the results graphically in TRITON.
- If you want to skip the calculations for this tutorial, go to subdirectory */usr/doc/triton-3.0/tutorial/EXAMPLE2* and read *EXAMPLE2.prj* into TRITON.

Timing

The calculation of the whole reaction path takes 128 seconds on Origin200 SGI computer with 180 MHz R10000 processor and 512 MB of RAM. Output data take 1 MB of space on the hard drive.

Analysis

The TRITON will enable you to read the outputs automatically if you finished the calculation properly and if you obtained the output files in your *EXAMPLE2* directory. Press the *Output Structures* bookmark on the left part of the TRITON screen to read MOPAC/DRIVER outputs. Visualize the changes in geometry of the system during the reaction using animation tools. The animation can be started using *Menu: Reaction/Run Forward*, stopped using *Menu: Reaction/Stop Animation* or simply managed using Animation Bar (*Fig 11.1*). Note the change in conformation of DCE from *trans* to *gauche* during the reaction. Display the Energy Graph using *Menu: Reaction/Energy Graph*. Examine the transition state structure by clicking on the point in the energy plot corresponding to the structure with the highest energy (step 33). Inspect the reactant and product structures in a similar manner clicking on the points corresponding to the structures with lowest energy before (step 9) and after (step 41) the reaction. Compare the value of the activation energy barrier (31.3 kcal/mol) with barrier obtained in Tutorial 3 where two stabilizing residues are present in the model. You can observe positive ΔH indicating unfavourable energetics of the products due to missing stabilization of Cl^- .

15.3 Tutorial 3

Modelling of enzymatic reaction using a medium model

Background

This tutorial will show how to model biochemical reaction and how to study stabilization effects of the active site residues. The two active site residues, Trp125 and Trp175, were proposed to play very important role during the dehalogenation reaction. The stabilizing role of these residues of the haloalkane dehalogenase was implicated from crystallographic [10] and site-directed mutagenesis experiments [11]. Quantum chemical calculation of the dehalogenation reaction in the model system composed of Asp124, Trp125, Trp175 and DCE (**Fig 15.2**) enabled to “see” stabilizing role of the tryptophan residues in the catalysis [9].

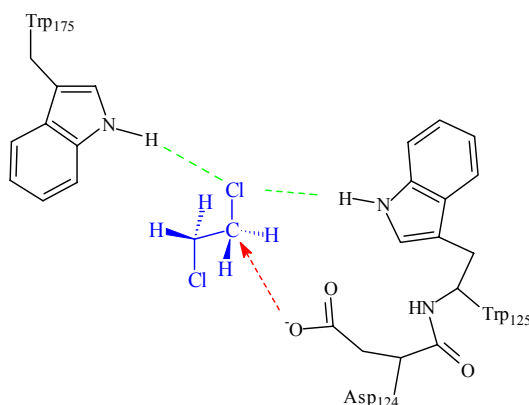


Fig 15.2 Scheme of the dehalogenation reaction using a medium model

Setup

- Start TRITON by typing **triton &** in MY_TUTORIAL directory.
- Open general project MY_TUTORIAL (**Fig 13.1** and **Fig 13.2**).
- Start a new reaction project (**Fig 13.1**, **Fig 13.2** and **Fig 13.3**) and specify its type as *Enzymatic reaction* (**Fig 13.5**) and a number of proteins as *Single protein* (**Fig 13.6**).
- Name the project **EXAMPLE3** and place it into MY_TUTORIAL project as its subproject (**Fig 13.7**).
- Take the *DhlA.pdb* file from the INPUT directory as source structure (**Fig 13.10**).
- Check the *DCE* molecule and set it as the substrate molecule in *List of Ligands* (**Fig 13.12**).
- Check the Asp124, Trp125 and Trp175 residues and set them as residues forming the cavity model in *List of Residues* (**Fig 13.13**).
- Specify the atoms forming the reaction co-ordinate (**Fig 13.14**). Choose the carbon atom on DCE, which is closer to Asp124 and binds Cl (C2) and the oxygen on Asp124 (O1985) by clicking on them. The original distance between them is 3.48 Å (see in **Fig 11.1**) and final driven distance is 1.28 Å (see “Dist.” in **Fig 13.14**). Set the DRIVER step size as – 0.05 Å.

- Fix the heavy atoms of amino acids, which form the backbone of enzyme using *Select backbone* utility (**Fig 13.15**). You have to fix these atoms during the optimization procedure to preserve mutual positions of amino acids in the active site of the enzyme.
- Setup the MOPAC and DRIVER calculation parameters (**Fig 13.16**). Set the total charge of the system to **-1**, *SCF* to **Conventional** and *Optimizer* to **BFGS**. Select the **AM1 Hamiltonian** and increase the maximal calculation time up to 10^8 seconds. Check the KEYWORDS: **GEO-OK, MMOK, PRECISE, NOINTER, NOXYZ and NODIIS** (**Fig 13.17**).
- Input optional commentary and finish the reaction wizard (**Fig 13.17** and **Fig 13.18**). If you are working on the machine on which you want to perform the calculations and you have MOPAC/DRIVER properly configured, let the calculation run automatically. If you want to run the calculation on a different machine, copy the whole directory **EXAMPLE3** onto this machine. Note that MOPAC and DRIVER must be installed there too! Then start the calculation just by typing the **run.sh &**.
- Follow the calculation by examining the *run_output* file.
- Once the calculation is successfully finished, copy the whole directory **EXAMPLE3** back onto the machine with TRITON installed, put it into your tutorial directory and analyze the results graphically in TRITON.
- If you want to skip the calculations for this tutorial, go to subdirectory */usr/doc/triton-3.0/tutorial/EXAMPLE3* and read *EXAMPLE3.prj* into TRITON.

Timing

The calculation of the whole reaction path takes about 30 minutes on Origin200 SGI computer with 180 MHz R10000 processor and 512 MB RAM. Output data take about 2.5 MB space on the hard drive.

Analysis

The TRITON will enable you to read outputs automatically if you finished the calculation properly and if you obtained the output files in your *EXAMPLE3* directory. Press the *Output Structures* bookmark on the left part of the TRITON screen to read MOPAC/DRIVER outputs. Visualize the changes in geometry of the system during the reaction using animation tools. Note the displacement of the side-chains of Trp125 and Trp175 immediately after the halide ion is released from DCE (step 32) indicating that these residues are involved in halide stabilization. Stabilization of Cl^- by nitrogen-bound hydrogens (atoms H24 and H66) of tryptophan residues can be confirmed by investigating the development of charges on these residues during the reaction using *Menu: Reaction/Electrostatics Progress Graph* and selection of the specified atom. Display both the Energy Graph (**Fig. 15.3**) and the Electrostatics Progress Graph (**Fig. 15.4**) in parallel and note the jump in charge after the transition state structure (TS = middle yellow point). Compare the changes in charge on nitrogen-bound hydrogens with some phenyl-ring hydrogens (e.g., atom H64 on Trp175). Note the difference in scaling in the graphs for H66 and H64!!! Note that scale of each graph may be changed by moving the mouse while the <middle mouse button> is pressed. The scale of the graph may be also translated along the y axis (partial charge) by moving the mouse while the <CTRL> and <middle mouse button> is pressed. Compare the value of the activation energy barrier ($E^a = 20.6$ kcal/mol) with barriers obtained in Tutorial 2. The

activation barrier is lowered by 10.5 kcal/mol due to the contribution of stabilizing tryptophans.

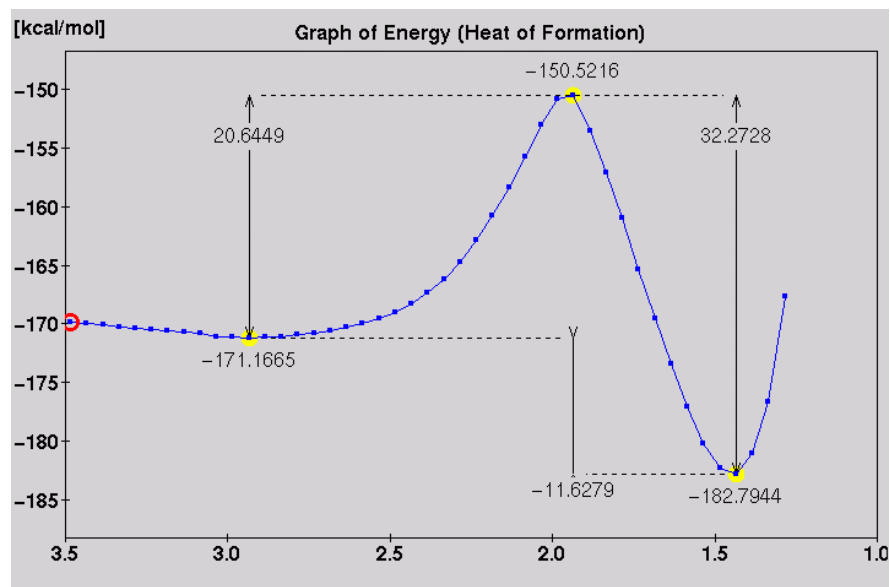


Fig 15.3 Energy Graph showing the changes in energy during the reaction

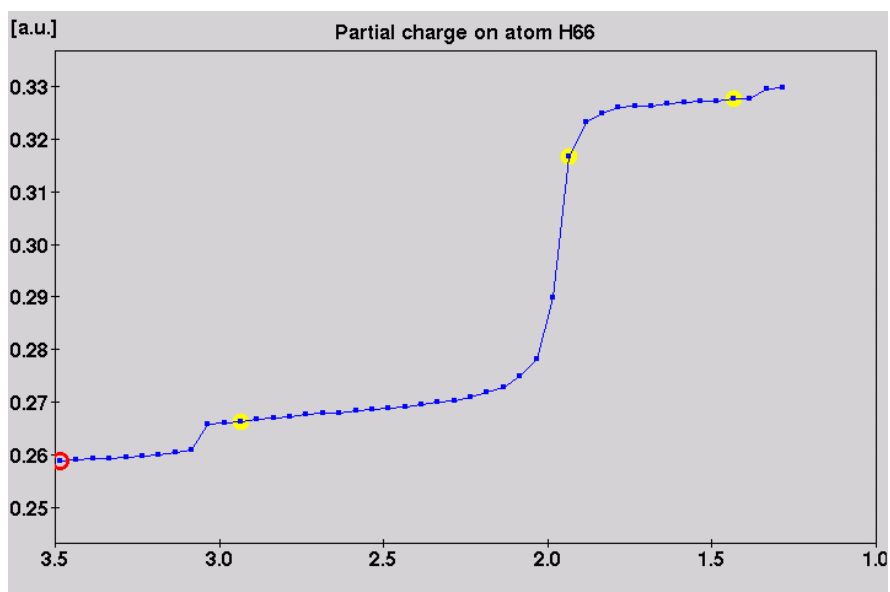


Fig. 15.4 Electrostatics Progress Graph showing the changes in charge during the reaction

15.4 Tutorial 4

Modelling of enzymatic reaction in two different enzymes using a large model

Background

This tutorial will show how to model biochemical reaction and investigate the relative contribution of the active site residues for the catalysis using two different haloalkane dehalogenases as model systems. Analogous comparison of the catalytic efficiency for the enzymes DhIA from *Xanthobacter autotrophicus* GJ10, LinB from *Sphingomonas paucimobilis* UT26 and DhaA from *Rhodococcus rhodochrous* NCIMB13064 was recently published [12] and will be partly reproduced within this tutorial. Calculation of the S_N2 reaction step conducted with larger models of the enzyme active sites is essential for getting an information about the potential stabilizing role of all surrounding amino acid residues. The active sites of DhIA and LinB are composed of 21 amino acids and the substrate molecule 1-chlorobutane (CBT) in docked orientation. **Fig 15.5** shows the products of modeled S_N2 reaction with alkyl-enzyme intermediate and halide ion present in the active site. Relative electrostatic contributions of the residues to the stabilization of released halide in DhIA and LinB and their comparison among the studied enzymes will be investigated. Note that this calculation using 21 amino acids model can be effectively performed only by using MOZYME option, which is included in MOPAC2000 or MOPAC2002 programs (see p. 9 how to obtain 30-days evaluation or full version of MOPAC2002).

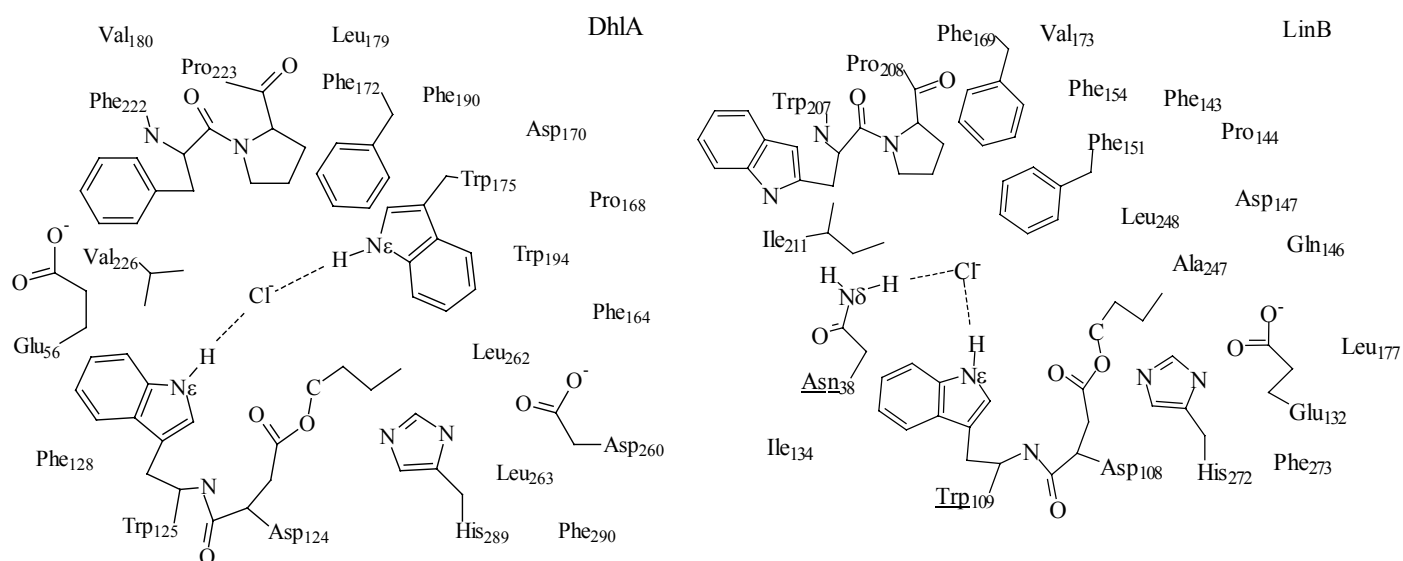


Fig 15.5 Schemes of reaction products for S_N2 step of the dehalogenation reaction in DhIA and LinB

Setup

- If you are not already running TRITON, start it by typing *triton &* in the MY_TUTORIAL directory.
- Open general project MY_TUTORIAL (**Fig 13.1** and **Fig 13.2**).
- Start a new general project and name it EXAMPLE4 (**Fig 13.1**, **Fig 13.2** and **Fig 13.3**).
- Start a new reaction project (**Fig 13.1**, **Fig 13.2** and **Fig 13.3**), specify its type as *Enzymatic reaction* (**Fig 13.5**) and a number of proteins as *Single protein* (**Fig 13.6**).
- Name the project DHLA_CBT and place it into MY_TUTORIAL project as its subproject (**Fig 13.7**).
- Take the *DhlA_CBT.pdb* file from the INPUT directory as a source structure (**Fig 13.10**).
- Check the CBT molecule and set it as the substrate molecule. Check also the catalytic water molecule WAT312 for the calculation in *List of Ligands* (**Fig 13.12**).
- Check 21 residues for DhlA enzyme: Glu56, Asp124, Trp125, Phe128, Phe164, Pro168, Asp170, Phe172, Trp175, Leu179, Val180, Phe190, Trp194, Phe222, Pro223, Val226, Asp260, Leu262, Leu263, His289 and Phe290 and set them as residues forming the cavity model in *List of Residues* (**Fig 13.13**).
- Specify the atoms forming the reaction co-ordinate (**Fig 13.14**). Choose the carbon atom on CBT (C4866), which binds Cl (CL4865) and the oxygen on Asp124 (O1977) by clicking on them. The original distance between them is 2.97 Å (see in **Fig 11.1**) and final driven distance is 1.27 Å. Set the DRIVER step size as **-0.05 Å**.
- Fix the heavy atoms of amino acids, which form the backbone of enzyme using *Select backbone* utility (**Fig 13.15**). You have to set these atoms to be fixed during the optimization procedure to preserve mutual positions of amino acids in the active site of enzyme.
- Setup the MOPAC/DRIVER calculation parameters (**Fig 13.16**). Set the total charge of the system to **-4** (Asp124, Glu56, Asp170 and Asp260 are negatively charged), *SCF* to **MOZYME** and *Optimizer* to **BFGS**. Select the **AM1** Hamiltonian and increase the maximal calculation time up to 10^8 seconds. Check the KEYWORDS: **GEO-OK**, **MMOK**, **NOINTER**, **NOXYZ** and **NODIIS** (**Fig 13.17**). By using keyword **PRECISE** the problems with geometry and energy converging may occur for such a big system and thus this keyword should be omitted.
- Input optional commentary and finish the reaction wizard (**Fig 13.17** and **Fig 13.18**). If you are working on the machine on which you want to perform the calculations and you have MOPAC and DRIVER properly configured, let the calculation run automatically. If you want to run the calculation on different machine, copy the whole directory EXAMPLE4 including DHLA_CBT subdirectory onto this machine. Note that MOPAC and DRIVER must be installed there too! Then start the calculation just by executing the *run.sh &*.
- Start second reaction project using the same procedure as for previous DHLA_CBT.prj.
- Name the project LINB_CBT and place it to the EXAMPLE4 project too.
- Take the *LinB_CBT.pdb* file from the INPUT directory and check the CBT and WAT298 from the *List of Ligands*.
- Specify following 21 residues forming the cavity model of LinB enzyme: Asn38, Asp108, Trp109, Glu132, Ile134, Phe143, Pro144, Gln146, Asp147, Phe151, Phe154, Phe169, Val173, Leu177, Trp207, Pro208, Ile211, Ala247, Leu248, His272 and Phe273.
- Specify the atoms forming the reaction co-ordinate. Choose the carbon atom (C4621) which binds Cl and oxygen on Asp108 (O1671). The original distance is 2.87 Å and final driven distance 1.27 Å.

- Set up all MOPAC parameters to be similar with previous *DHLA_CBT* project. Only the charge must be set to **-3** (Asp108, Glu132 and Asp147 are negative).
- Run the calculation automatically on the local machine or copy the whole directory LINB_CBT into the EXAMPLE4 directory on the remote machine and run the calculation there.
- Follow the calculations by examining their *run_output* files.
- Once the calculations are successfully finished, copy the whole directory EXAMPLE4 with both subdirectories back onto the machine with TRITON installed, put it into your tutorial directory and analyze the results graphically in TRITON.
- If you want to skip the calculations for this tutorial, go to subdirectory */usr/doc/triton-3.0/tutorial/EXAMPLE4* and read *EXAMPLE4.prj* into TRITON.

Timing

The calculation of the whole reaction path using MOZYME option takes about 34 hours for DhIA and about 33 hours for LinB. They were both performed on Origin200 SGI server with 180 MHz R10000 processor and 512 MB RAM. Output data take 14.3 MB space on the hard drive.

Analysis

Read the output files from EXAMPLE4 subdirectory into TRITON after successfully finished calculation. Visualize the changes in geometry of the system during the reaction using animation tool and the changes in energy during the reaction by the Energy Graph. Estimate the relative contribution of the active site residues to halide stabilization by investigating the electrostatic interactions between each residue and Cl⁻ released during the reaction. These interactions calculated as the difference between reactant and product state interactions are showed in Residue Interactions Graph (**Fig 15.6** and **Fig 15.7**). This graph shows the interaction for DhIA and LinB where the residues are structurally superimposed for better comparison. Select the Cl⁻ atom while you have <Shift> pressed. It is apparent that the largest stabilization in DhIA enzyme is provided by Trp125 and Trp175 forming the hydrogen bonds, followed by interactions with Pro223, Val226 and Phe172. Glu56 is the most repulsive residue due its negative charge and small distance to the halide. Similarly Asp170 and Asp260 show strong repulsion with Cl⁻ due their negative charge. The most stabilizing residues in LinB enzyme are Asn38 and Trp109 forming the hydrogen bonds with the halide. Pro208, Ile211, Phe151 and Phe169 also show significant attraction towards halide. The residues Glu132 and Asp147 are representatives of negatively charged residues with strong repulsion. The rest of amino acids are much less contributing to stabilization of the transition state. As it can be seen from **Fig 15.6** and **Fig 15.7** the two compared enzymes differ in contributions of several amino acids. Repulsive Glu56 in DhIA is substituted by attractive Asn38 in LinB, attractive Trp175 in DhIA is substituted by less efficiently stabilizing Phe151, etc. Note that scale of each graph may be changed by moving the mouse while the <middle mouse button> is pressed. The scale of the graph may be also translated along the y axis by moving the mouse while the <CTRL> and <middle mouse button> is pressed. To highlight the examined residue in 3D Model window simply click on its bar in graph. Changes in electrostatic interactions during the reaction, i.e. in the reactant, the transition state and the product structures, can also be investigated by changing the *Graph Settings* (**Fig 13.2**). The Van der Waals interactions are not supported in this version. The values of partial charges and the

interaction energies can be stored in a file by right clicking on the graph and pressing the *Write Residue Interactions...*

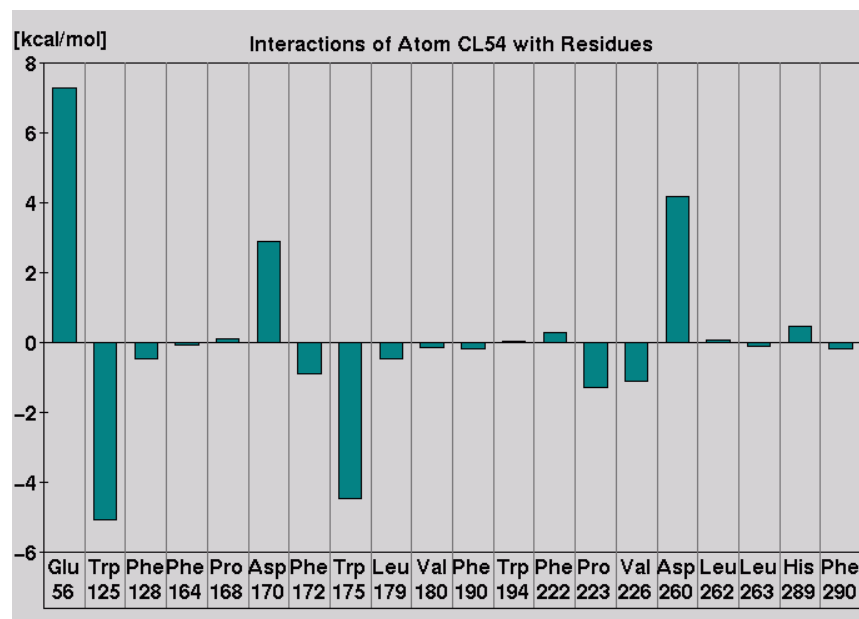


Fig 15.6 Graph of the differences of electrostatic interactions between product and reactant for the active site residues of Dh1A

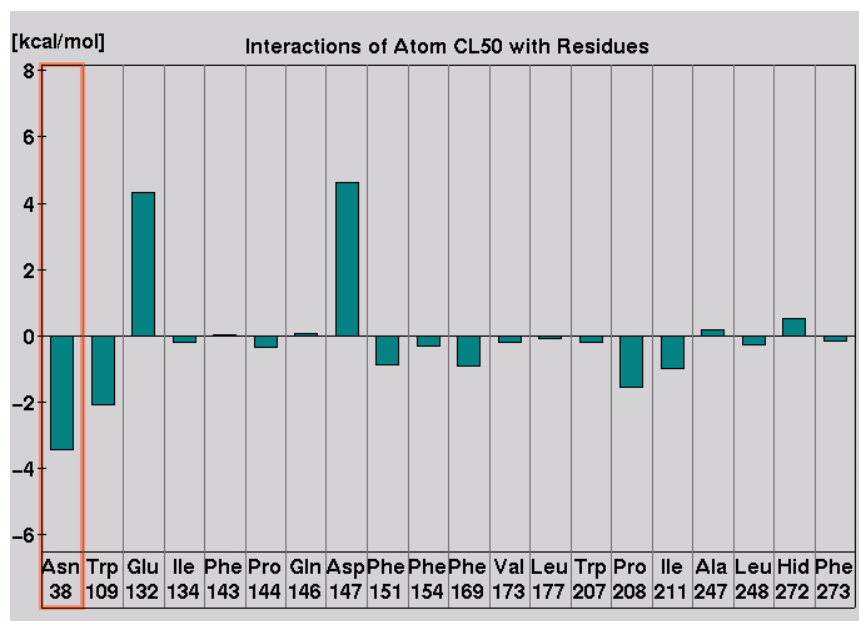


Fig 15.7 Graph of the differences of electrostatic interactions between product and reactant for the active site residues of LinB

15.5 Tutorial 5

Assignment of atomic charges to a protein

Background

This tutorial will show how to calculate the charge distribution on the whole enzyme Dhla using a single point semi-empirical QM calculation and how to investigate the relative contribution of all the residues for the catalysis. Calculation of charge distribution on the whole enzyme can serve for finding of interactions with bound substrate or with a product (here halide ion) released during the reaction. This tutorial will show how to calculate the interaction energies between all enzyme amino acids and chloride ion released during the S_N2 reaction. Input structure is the high-resolution structure of LinB *LinB_Cl.pdb* with chloride ion and water oxygens present in the active site. The single point calculation of the whole enzyme can be performed only by using MOZYME option, which is included in MOPAC2000 or MOPAC2002 programs (see p. 9 how to obtain 30-days evaluation or full version of MOPAC2002).

Setup

- If you are not already running TRITON, start it by typing *triton &* in the MY_TUTORIAL directory.
- Open general project MY_TUTORIAL (**Fig 14.1** and **Fig 14.2**).
- Start a new optimization project (**Fig 14.1**, **14.2** and **Fig 14.3**).
- Name the project EXAMPLE5 and place it into MY_TUTORIAL project as its subproject (**Fig 14.5**).
- Take the *LinB_Cl.pdb* file from the INPUT directory as a source structure (**Fig 14.6**).
- Check all residues and set them as residues forming the calculation model in *List of Residues* (**Fig 14.7**).
- You don not need to fix the heavy atoms of amino acids forming the backbone of enzyme using *Select Backb.* utility (**Fig 14.8**) because this is only the single point calculation.
- Setup the MOPAC calculation parameters (**Fig 14.9**). Set the total charge of the system to **-13** (this value corresponds with sum of charged residues), *Task to* **Single Point**, *SCF to* **MOZYME**, use **AM1 Hamiltonian** and then increase the maximal calculation time up to 10^8 seconds. Check the keywords **MMOK**, **NOXYZ** and add the **PL** and **MEMORY=900** keywords in window for additional keywords (**Fig 14.11**). The keyword **PRECISE** must be unchecked due to possible problems with energy convergence!
- Input optional commentary and finish the reaction wizard (**Fig 14.11** and **Fig 14.12**). If you are working on the machine on which you want to perform the calculations and you have MOPAC properly configured, let the calculation run automatically. If you want to run the calculation on different machine, copy the whole directory EXAMPLE5 onto this machine. Note that MOPAC2000 or MOPAC2002 must be installed there too! Then start the calculation just by executing the *run.sh &*.
- Follow the calculation by examining the *run_output* file.
- Once the calculation is successfully finished, copy the whole directory EXAMPLE5 back onto the machine with TRITON installed, put it into your tutorial directory and analyze the results graphically in TRITON.

- If you want to skip the calculations for this tutorial, go to subdirectory */usr/doc/triton-3.0/tutorial/EXAMPLE5* and read *EXAMPLE5.prj* into TRITON.

Timing

The calculation of the whole reaction path takes about 3.5 hours on SGI Octane computer with 1360 MHz R12000 processor and 512 MB RAM. Output data take 3.5 MB space on the hard drive.

Analysis

Read the output files from EXAMPLE5 subdirectory into TRITON after successfully finished calculation. Estimate the electrostatic interactions of all residues of LinB with released halide CL1001. These interactions can be visualized using Residue Interactions Graph (example is in **Fig 15.8**). Select the CL1001 molecule while you have <Shift> pressed and after you checked the *Menu: Select/Residues* option. It may be difficult to find the halide within the whole enzyme molecule, so user can use *Menu: Display/Hide All Atoms* and then show only the halide by clicking on it in the *List of residues (Menu: View/Residues)* using <middle mouse button>. It is apparent that the largest interactions are provided by charged residues. User can select which of the residues should be shown in the graph by specifying them in the *Displayed Residues...* window. This window will appear after clicking the <right mouse button> on the graph and choosing the *Displayed Residues...* option. To highlight the residue in 3D Model window click on bar in a graph. The graph may be non-comprehensive thus export of the values into the text file (*Write Residue Interactions...* option) and their analysis using some data management programs is advisable. **Fig 15.8** serving as example shows the graph of electrostatic interactions of selected residues from the LinB enzyme. See the attraction between doubly protonated His172 (Hip272) which corresponds to the state of His after the Ad_N catalytic step.

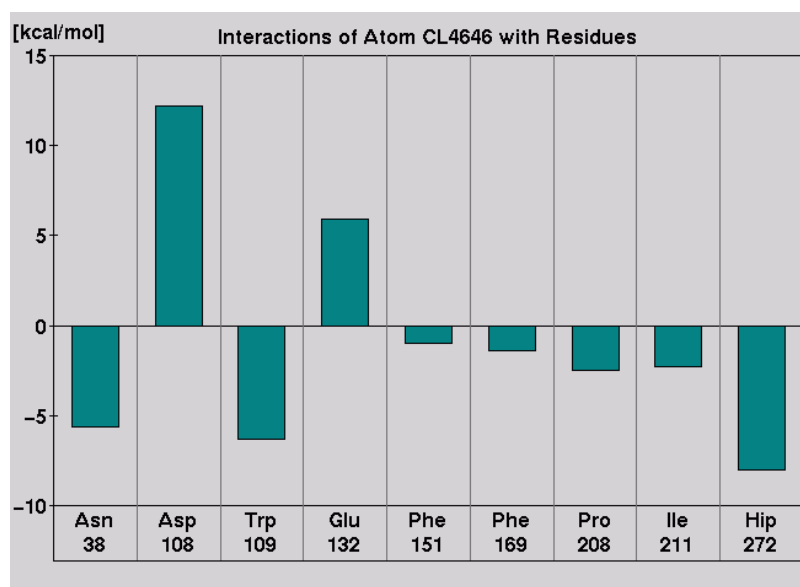


Fig 15.8 Graph of electrostatic interactions for selected residues in the LinB

16 Acknowledgements

We thank to Dr. Martin Černohorský (Czech Supercomputing Centre, Czech Republic) for providing the program DRIVER; Drs. Andrej Sali, Andras Fiser and Bozidar Yerkovich (The Rockefeller University, USA) for kind help with preparation of top files and for providing the program MODELLER; and the Czech Supercomputing Centre for providing computational facilities for testing TRITON. We also thank to Luboš Vrbka (NCBR, Czech Republic) for his help with development of TRITON import routine.

17 References

1. Prokop, M., Damborsky, J., and Koca, J., [TRITON: in silico construction of protein mutants and prediction of their activities](#). *Bioinformatics*, 2000. **16**: 845-846.
2. Sali, A. and Blundell, T.L., Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 1993. **234**: 779-815.
3. Stewart, J.J.P., MOPAC - A semiepirical molecular-orbital program. *Journal of Computer-Aided Molecular Design*, 1990. **4**: 1-45.
4. Cernohorsky, M., Kutý, M., and Koca, J., A multidimensional driver for quantum chemistry program MOPAC. *Computers & Chemistry*, 1996. **21**: 35-44.
5. Damborsky, J., Bohac, M., Prokop, M., Kutý, M., and Koca, J., [Computational site-directed mutagenesis of haloalkane dehalogenase in position 172](#). *Protein Engineering*, 1998. **11**: 901-907.
6. Schanstra, J.P., Ridder, I.S., Heimeriks, G.J., Rink, R., Poelarends, G.J., Kalk, K.H., Dijkstra, B.W., and Janssen, D.B., Kinetic characterization and X-ray structure of a mutant of haloalkane dehalogenase with higher catalytic activity and modified substrate range., *Biochemistry*, 1996. **35**: 13186-13195.
7. Vriend, G., WHAT IF: A molecular modelling and drug design program. *Journal of Molecular Graphics*, 1990. **8**: 52-56.
8. Schafmeister, C.E.A.F., Ross, W.S., and Romanovski, V., LEAP. 1995, *University of California San Francisco*: San Francisco, USA.
9. Damborsky, J., Kutý, M., Nemec, M., and Koca, J., [A molecular modelling study of the catalytic mechanism of haloalkane dehalogenase: 1. quantum chemical study of the first reaction step](#). *Journal of Chemical Information and Computer Sciences*, 1997. **37**: 562-568.
10. Verschueren, K.H.G., Seljee, F., Rozeboom, H.J., Kalk, K.H., and Dijkstra, B.W., Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. *Nature*, 1993. **363**: 693-698.
11. Kennes, C., Pries, F., Krooshof, G.H., Bokma, E., Kingma, J., and Janssen, D.B., Replacement of tryptophan residues in haloalkane dehalogenase reduces halide binding and catalytic activity. *European Journal of Biochemistry*, 1995. **228**: 403-407.
12. Bohac, M., Nagata, Y., Prokop, Z., Prokop, M., Monincova, M., Koca, J., Tsuda, M., Damborsky, J., [Halide-stabilizing residues of haloalkane dehalogenases studied by quantum mechanic calculations and site-directed mutagenesis](#). *Biochemistry*, 2002. **41**: 14272-14280.