Prime 4.0

Quick Start Guide

Schrödinger Press

May 2015

Contents

Document Conventions v

Chapter 1: Introduction 1

1.1 **Preparing for the Exercises** 1

1.2 **Creating a Maestro Project** 3

1.3 **Setting Maestro Preferences** 3

Chapter 2: Homology Modeling of a Single Chain 5

2.1 **Importing the Query Sequence** 5

2.2 **Finding Sequence Homologs** 7

2.3 **Editing the Alignment** 9

2.4 **Building a Model Structure** 14

2.5 **Refining Target Regions of the Structure** 16

2.5.1 Refining Loops 16

2.5.2 Minimizing Target Regions 18

Chapter 3: Homology Modeling of a Homomultimer 21

3.1 **Preparing the Template** 21

3.2 **Importing the Sequence** 24

3.3 **Importing the Template** 24

3.4 **Building the Model** 25

Chapter 4: Homology Modeling of a Heteromultimer 27

4.1 **Preparing the CDK4 Chain** 27

4.2 **Preparing the Cyclin-E Chain** 28

4.3 **Building the Model** 30

Getting Help 31

Glossary 35

Document Conventions

In addition to the use of italics for names of documents, the font conventions that are used in this document are summarized in the table below.

Font	Example	Use
Sans serif	Project Table	Names of GUI features, such as panels, menus, menu items, buttons, and labels
Monospace	`$SCHRODINGER/maestro`	File names, directory names, commands, environment variables, command input and output
Italic	*filename*	Text that the user must replace with a value
Sans serif uppercase	CTRL+H	Keyboard keys

Links to other locations in the current document or to other PDF documents are colored like this: Document Conventions.

In descriptions of command syntax, the following UNIX conventions are used: braces { } enclose a choice of required items, square brackets [ ] enclose optional items, and the bar symbol | separates items in a list from which one item must be chosen. Lines of command syntax that wrap should be interpreted as a single command.

File name, path, and environment variable syntax is generally given with the UNIX conventions. To obtain the Windows conventions, replace the forward slash / with the backslash \ in path or directory names, and replace the $ at the beginning of an environment variable with a % at each end. For example, `$SCHRODINGER/maestro` becomes `%SCHRODINGER%\maestro`.

Keyboard references are given in the Windows convention by default, with Mac equivalents in parentheses, for example CTRL+H (⌘H). Where Mac equivalents are not given, COMMAND should be read in place of CTRL. The convention CTRL-H is not used.

In this document, to *type* text means to type the required text in the specified location, and to *enter* text means to type the required text, then press the ENTER key.

References to literature sources are given in square brackets, like this: [10].

Introduction

This manual provides a tutorial introduction to using Prime for protein structure prediction by homology modeling.

- Chapter 2 provides an introduction to homology modeling for a single chain that covers most of the tasks that are performed in homology modeling, including running a BLAST search, selecting a template, editing the alignment, building the model, and refining the model.
- Chapter 3 provides an exercise on building a homology model for a homomultimer with four identical chains. This exercise includes the preparation of a tetrameric template in the Protein Preparation Wizard panel from a template that includes only the asymmetric unit.
- Chapter 4 provides an exercise on building a homology model for a heteromultimer with two different chains.

It is assumed that you have access to the third-party programs and databases (PDB, BLAST, HMMER/Pfam). If you are running Prime on Linux, it is assumed that you have downloaded and installed the optional (but highly recommended) third-party secondary structure prediction program PSIPRED. To find out how to obtain third-party programs, go to the Third Party Programs page of our website.

The tutorial is designed to run from a local copy of the PDB and BLAST databases. If you run the tutorial with web download, you might notice some differences.

For further information on using Maestro, see the Maestro online help or the *Maestro User Manual*. For more information about Prime features, see the *Prime User Manual*.

1.1 Preparing for the Exercises

To run the exercises, you need a working directory in which to store the input and output, and you need to copy the input files from the installation into your working directory. This is done automatically in the Tutorials panel, as described below. To copy the input files manually, just unzip the `prime` zip file from the `tutorials` directory of your installation into your working directory.

On Linux, you should first set the SCHRODINGER environment variable to the Schrödinger software installation directory, if it is not already set:

csh/tcsh: `setenv SCHRODINGER` *installation-path*

sh/bash/ksh: `export SCHRODINGER=`*installation-path*

If Maestro is not running, start it as follows:

- **Linux:** Enter the following command:

  ```
  $SCHRODINGER/maestro -profile Maestro &
  ```

- **Windows:** Double-click the Maestro icon on the desktop.

  You can also use Start → All Programs → Schrodinger-2015-2 → Maestro.

- **Mac:** Click the Maestro icon on the dock.

  If it is not on the dock, drag it there from the SchrodingerSuites2015-2 folder in your Applications folder, or start Maestro from that folder.

Now that Maestro is running, you can start the setup.

1. Choose Help → Tutorials.

   The Tutorials panel opens.

2. Ensure that the Show tutorials by option menu is set to Product, and the option menu below is labeled Product and set to All.

3. Select Prime Quick Start Guide in the table.

4. Enter the directory that you want to use for the tutorial in the Copy to text box, or click Browse and navigate to the directory.

   If the directory does not exist, it will be created for you, on confirmation. The default is your current working directory.

5. Click Copy.

   The tutorial files are copied to the specified directory, and a progress dialog box is displayed briefly.

If you used the default directory, the files are now in your current working directory, and you can skip the next two steps. Otherwise, you should set the working directory to the place that your tutorial files were copied to.

6. Choose Project → Change Directory.

7. Navigate to the directory you specified for the tutorial files, and click OK.

You can close the Tutorials panel now, and proceed with the exercises.

1.2 Creating a Maestro Project

You should create a new named Maestro project to save your work, in case you want to complete the exercises at a later time. If you are using an existing Maestro session, it is advisable to create a new project to keep the tutorial separate from your other work. When you start Maestro, a scratch project is created, which must be named in order to keep it for later use.

1. Choose Project → New.

   The New Project dialog box is displayed. The Look in option menu should contain the current Maestro working directory.

2. In the File name text box, type `PrimeTutorial`.

3. Click Save.

This procedure creates a project named `PrimeTutorial`. The work that you do during the exercises that follow is automatically saved in this project for later use.

1.3 Setting Maestro Preferences

Maestro has many options for displaying structures in the Workspace. Here you will set some preferences that are needed for the first exercise. You will also need to display the Saved Views toolbar.

1. Choose Maestro → Preferences.

   The Preferences panel opens.

2. On the left, under Molecular representation, click Ribbons.

   The ribbons preferences are displayed on the right.

3. For Helix interior, select Same as exterior.

4. From the Atoms to hide when ribbons are created option menu, choose All associated atoms.

5. Close the Preferences panel.

6. If the Saved Views toolbar is not displayed, display it:

- From Maestro, click Saved Views on the Manager toolbar, or choose Window → Toolbars → Saved Views.
- From BioLuminate, choose Edit → Preferences → Toolbars → Select Toolbars Displayed → Saved Views.

Homology Modeling of a Single Chain

This chapter provides an exercise for building and refining a model of a query sequence for which a sequence homolog can be identified using BLAST. While the tutorial is self-contained, you may find it useful to refer to the *Prime User Manual* or the online help for more detailed information on the panels or procedures.

If you have not already done so, complete the setup sections in Chapter 1.

2.1 Importing the Query Sequence

The query sequence that will be used is closely related to that of phosphoglycerate kinase from *Pyrococcus furiosus*, but has been modified slightly to provide a case that best demonstrates various features of Prime's Homology Modeling workflow:

```
>Query
YNRTVFLRVDLNSPMSNGKVQSDARFRAVLPTIKYLIESGAKVVVGTHQGKEYSTTEEHARILSELLNMH
VEYVEDYAIFGISKARERAAMKPGEVIVLENLRFSAEEFVRKLSQVIDLVVNDAFAAAHRSQPSLVGFAR
IKPMIMGFL
```

In this section, you will import the query sequence from the tutorial directory into Prime as the first step of the Structure Prediction workflow.

1. Open the Structure Prediction panel.

   From Maestro:

   - Choose Tasks → Homology Modeling in the main window, then click the Structure Prediction Wizard icon in the Homology Modeling panel.

     

   From BioLuminate:

   - Choose Tasks → Homology Modeling → Advanced Homology Modeling in the main window.

   The Structure Prediction panel opens at the first step, Input Sequence. By default the Guide is not displayed. If you want to display it, choose Step → Guide.

2. Click File and select `PrimeTutorial1.fasta`, then click Open.

   The sequence is displayed in the Prime sequence viewer (Figure 2.1). At this stage, there is no structure to display in the Workspace.

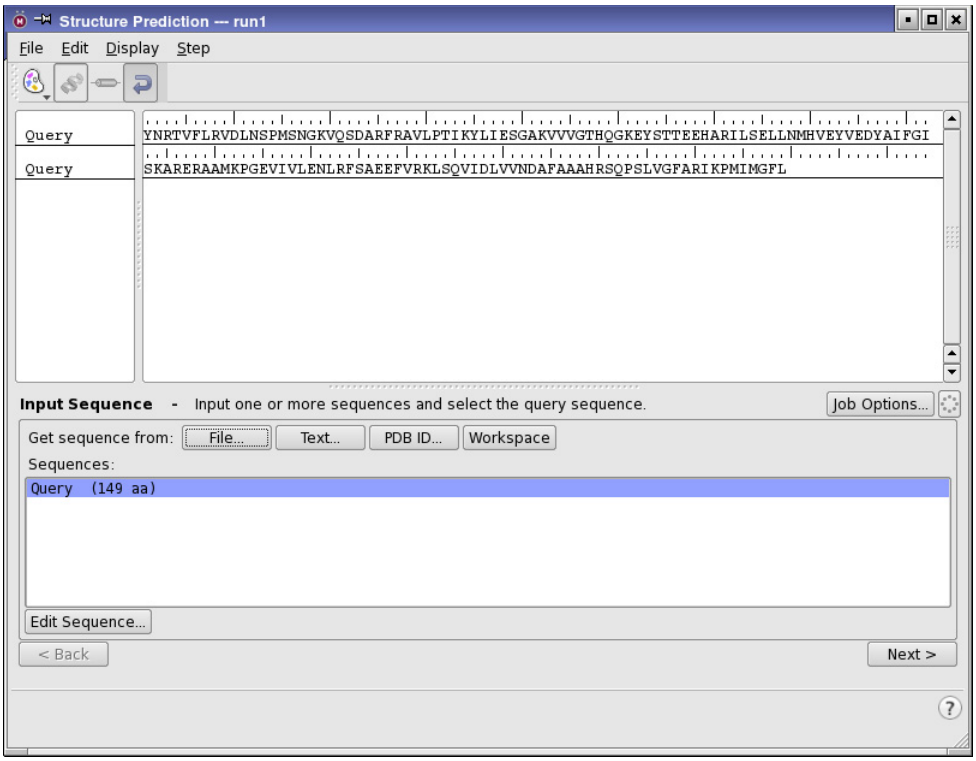

Figure 2.1. The Input Sequence step after import.

   Unlike the Prime sequence viewer, the Workspace sequence viewer in the lower part of the Maestro main panel displays sequences only for named entries in a project. Until the end of this tutorial, when the finished structure is added to the Project Table, the Workspace sequence viewer remains empty.

3. Click Next to proceed to the next step, Find Homologs.

2.2 Finding Sequence Homologs

In this step, you will search for homologous proteins with known structure using BLAST, then select one homolog as a template.

1. Click BLAST Homology Search.

   The search job is started. This search usually takes less than 1 minute on a 1-GHz processor. When the job finishes, a list of potential templates is displayed in the Homologs table. The highest-scoring template is selected by default, as shown in Figure 2.2.

The PDB and BLAST databases provided are continually being updated. Therefore, the rank order and scores of the homologs found might differ slightly from that shown.

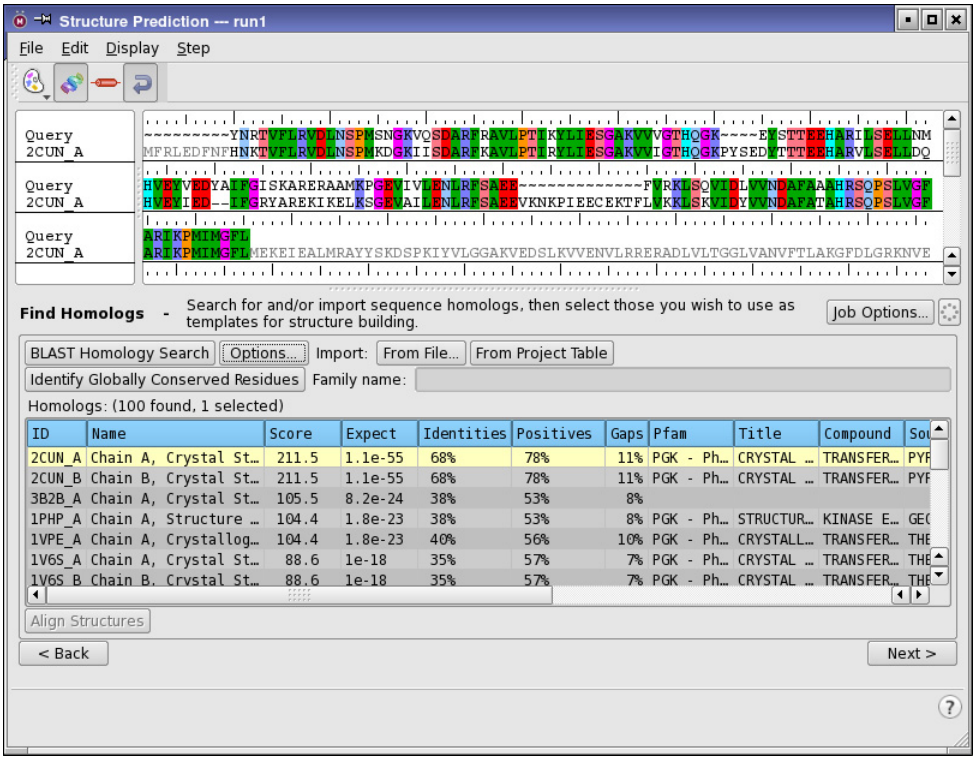

Figure 2.2. The Find Homologs step after searching for homologs.

2. If the SSA is not visible in the sequence viewer, click the View SSA button on the toolbar.

This button displays the secondary structure assigment in the sequence viewer. If it was not selected, when you select it the assignment for the homolog is displayed in the sequence viewer, with `_ssa` added to the homolog name.

3. Select the 1VPE_A template (by clicking its row).

   This template should be near the top of the Homologs table.

   The BLAST alignment between the template and query sequences is displayed in the Prime sequence viewer, along with the secondary structure assignment of the template. In addition, the selected template is displayed in the Workspace.

4. Zoom in on the region of the template that is aligned to the query (the colored region of the ribbon representation.) and manipulate the view to resemble Figure 2.3. You might want to use the SHIFT key to restrict rotations to the X or Y axis.

Figure 2.3. The 1VPE_A template, showing the region aligned with the query.

5. Save the view so you can easily return to it: click the Save View button on the Saved Views toolbar, and use the default name.

The next step, to obtain HMMER/Pfam family and sequence data, is optional. If you do not want to do this step, skip to Step 7.

6. (Optional) Click Identify Globally Conserved Residues.

   Note: You must have a local installation of the Pfam database to complete this step.

   This job should take 2 to 3 minutes to complete. A Hidden Markov Model (HMM) is generated from a multiple sequence alignment and used to identify the query family and provide information about which residues are conserved in the consensus sequence.

   When the job finishes, the family appears in the Family name text box, and the sequence is displayed in the sequence viewer, labeled `Query_pfam`. Only the conserved residues are displayed, and colored according to the color of the residue in the template.

   A minus sign appears beside the query name in the sequence viewer: this is a collapse/expand "button". Clicking on the minus sign hides the Pfam sequence, and the minus sign becomes a plus sign; clicking on the plus sign displays the Pfam sequence again.

You can now continue to the next step:

7. Ensure that 1VPE_A is still selected.
8. Click the Next button.

   The next step is Edit Alignment. The template is again automatically fit to fill the Workspace.

9. To return the view to the one you saved in the previous step, click the View1 button on the Saved Views toolbar.

2.3 Editing the Alignment

Because the alignment provided by the Find Homologs step is based only on sequence information, there is room for improvement. For example, the default alignment has placed a gap at query residue His59, which corresponds to the middle of a helix in the template (Figure 2.4). Therefore, it is unlikely that the alignment returned by BLAST is correct in this region. This can be rectified either by hand-editing the alignment or by using the Prime Align program, which takes secondary structure into account.

Before making changes to the BLAST alignment, save the current run:

1. From the File menu, choose Rename.
2. Enter `Blast_Alignment` in the dialog box, then click OK.

Name the new run you will be working in:

3. From the File menu, choose Save As.
4. Type `New_Alignment` in the text box and click OK.

   The `New_Alignment` run is the one that is now open. The `Blast_Alignment` run has been closed, but can be reopened at any time.

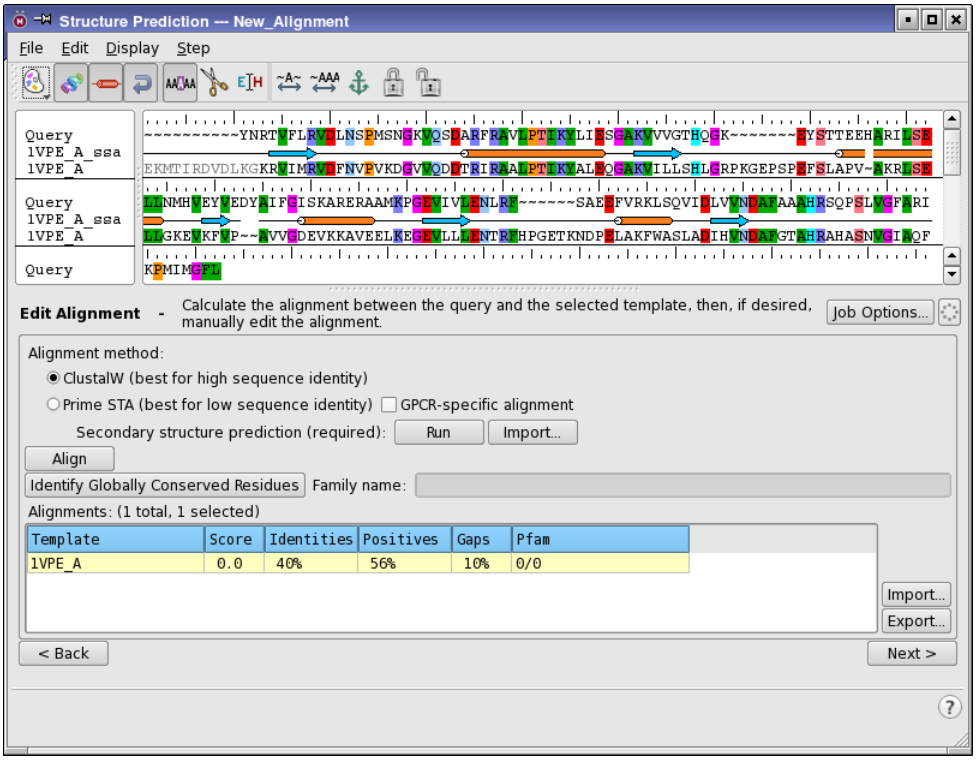

Figure 2.4. Initial view of the Edit Alignment step.

There are two alignment methods available. One uses ClustalW, and the other uses Prime's alignment program, STA. The latter will be used in this exercise.

5. In the Alignment method section, select Prime STA.

In order to deal with the fact that secondary structure prediction is only about 75% accurate,[1] Prime supports running two distinct secondary structure prediction programs. One of these, SSpro, is bundled with Prime. However, the other, PSIPRED, is not (and is not available on Windows). See Section 2.4 of the *Installation Guide* for information on third-party programs.

1. For example, visit the EVA site at http://cubic.bioc.columbia.edu/eva/sec/res_sec.html for more details

Now generate secondary structure predictions for the query to use in the alignment program:

6. Click Run.

   If the optional SSP program PSIPRED was installed (strongly recommended), this job should take about 5 minutes. When the SSP job finishes, the secondary structure predictions of the query are displayed in the sequence viewer, as in Figure 2.5.

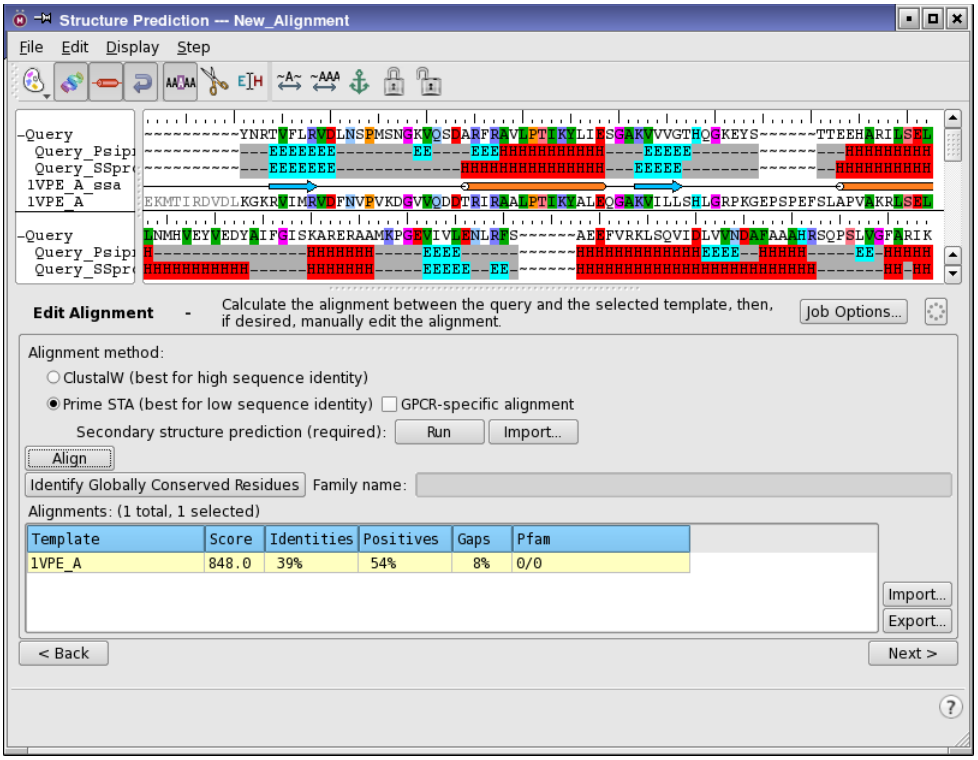

Figure 2.5. The Edit Alignment step after running SSP and Align.

This and subsequent operations may produce different views of the structure in the Workspace. Click the View1 button on the Saved Views toolbar as needed.

7. Click Align.

   The alignment program starts running. This job may take 20 minutes to complete.

   Once the alignment job finishes, the new alignment is displayed in the sequence viewer and the values in the Alignments table are updated. The template's Score, which was 0.0

prior to running the Align job, is now a non-zero number. In addition to some other minor changes in the alignment, the gap at His59 has been moved to an adjacent loop. This makes more physical sense and is likely to result in a more accurate homology model.

8. (Optional) The true secondary structure of the template is shown graphically in the sequence viewer above the template sequence. To get an indication of the accuracy of the SSP programs in particular regions of the sequence, run the SSP programs on the template: right-click on the template in the sequence viewer and choose Run SSP from the menu.

While the Align program took secondary structure into account in producing the alignment between query and template, it did not explicitly consider tertiary structure. You will next perform some manual editing of the alignment that accounts for tertiary structure.

Residues that are not being used in the current alignment are undisplayed, revealing where gaps exist in the alignment.

9. From the Color Property button menu, choose Residue Property.

   Aligned residues in the template in the Workspace are now colored according to the query's Residue Property (Figure 2.6). That is, they are colored according to the residue type to which they will be converted once the model is built.

Figure 2.6. The template colored by the query's residue property.

10. Examine the structure to confirm that hydrophobic residues (green) are directed toward the interior of the protein and charged residues (negative: red, positive: blue) are directed toward solvent (polar uncharged residues are colored cyan).

    The only exception is Helix 97-104 (template numbering), shown in Figure 2.6. To find this helix in the Workspace:

    a. Scroll the Prime sequence viewer to the second row.

    b. Locate the residue labeled (98) Asp97 by moving the pointer over the residues in the sequence viewer for the template.

    c. Drag to select the residues in the template from (98) Asp97 to (105) Glu104.

    The selected residues are highlighted in the Workspace with yellow markers.

11. Remove the markers by clicking in a blank area in the sequence viewer or the Workspace.

Several charged residues appear to be directed towards the interior of the protein, which is likely to result in buried charges in the model once built. This problem can be rectified by manually editing the alignment in this region. Fortunately, there is a two-residue gap near the helix that allows for some flexibility in the local alignment.

12. Change to Slide Freely mode by clicking Slide Freely on the Prime toolbar:

    

13. Drag residue Leu106 (of the template) to the left by two positions.

    The original gap is closed, a new C-terminal gap is created, and the Workspace is updated. The problematic charged residues are now mapped to residues directed outward, which is more physically reasonable (see Figure 2.7).

Now that an optimal alignment between query and template has been generated, you can proceed to the next step.

14. Click Next to proceed to the Build Structure step.

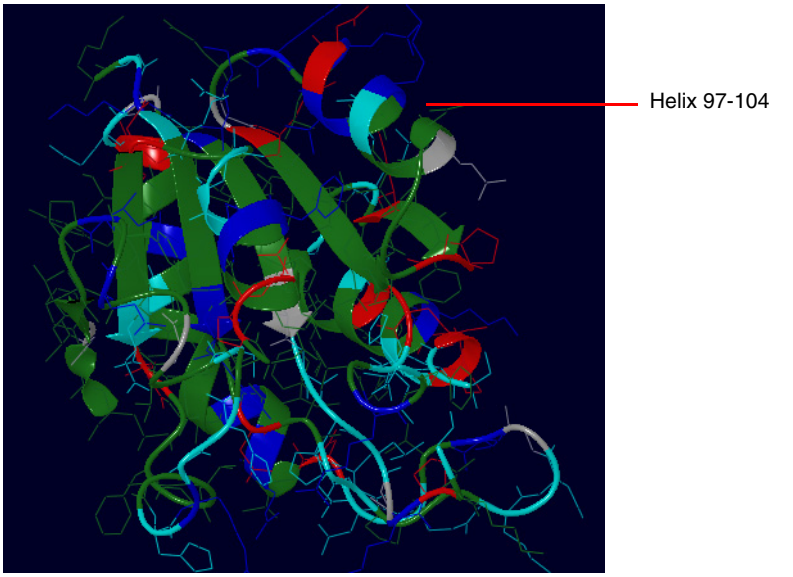

Figure 2.7. Helix from Figure 2.6 after hand editing.

2.4 Building a Model Structure

The Build Structure program builds insertions, closes gaps, and predicts side-chain conformations of non-conserved residues to produce a model with no unphysical clashes. However, it does this efficiently, without extensive conformational sampling. The structure produced in the Build Structure step is likely to represent only a local energy minimum and not the global minimum. Therefore, regions with gaps in the alignment are likely to require refinement in the refinement step.

In this part of the exercise, you will construct a homology model that is based on the alignment produced in the previous step and that includes the template ligand 3PG.

1. Select the ligand 3PG from the Include ligand and cofactors list .

   The selected ligand is highlighted in the Workspace.

2. Under Model building method, select Energy-based.

   This method is slower, but more accurate.

3. Click Build Model.

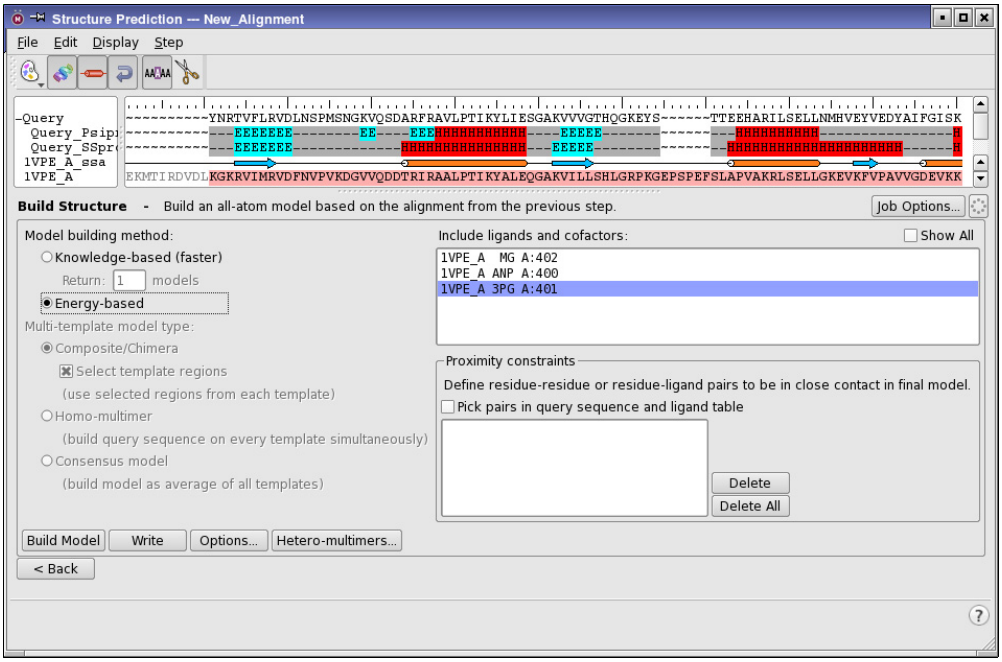

Figure 2.8. The Build Structure step.

This job takes about 5 minutes on a 2 GHz processor. The progress of the job is displayed in the Log file text area. When the job finishes, the structure is exported to the project as a new entry, and is displayed in the Workspace, colored according to the conservation of the template coordinates (Figure 2.9).

A dialog box opens, asking if you want to open the Refinement panel to refine the structure.

4. Click Yes.

   The Structure Prediction panel closes, and the Refine Loops panel opens.

Structures visible in the Workspace while working in the Structure Prediction panel are scratch entries (not yet part of the Project Table.) The Workspace sequence viewer does not display scratch entries. Now that this structure is a Project Table entry, its sequence and SSA are displayed in the Workspace sequence viewer. The Workspace sequence viewer is not displayed by default. To display it, choose Window → Sequence Viewer (Maestro) or Edit → Settings → Show Sequence Viewer (BioLuminate) in the main window.

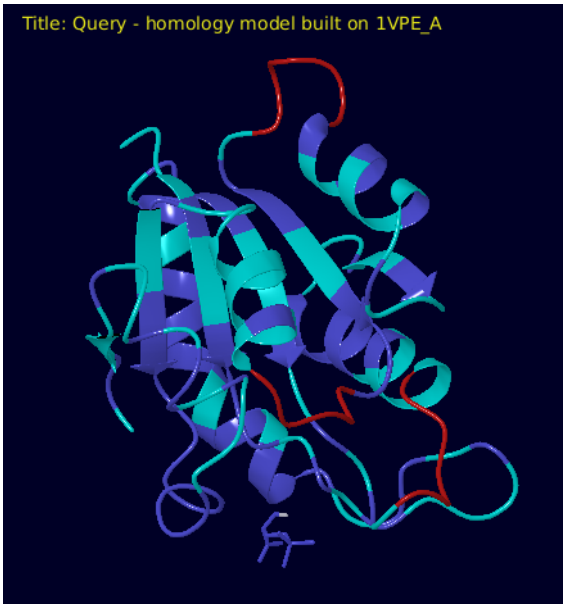

Figure 2.9. Workspace after building structures.

2.5 Refining Target Regions of the Structure

To improve the structure most efficiently, you should focus refinement efforts on areas of the structure that are likely to be problematic. In general terms, this means refining loops (particularly where insertions have been made or gaps closed) and re-predicting side-chain conformations. A particular structure may also have atom position clashes, non-ideal bond lengths and angles, and residues with unfavorable energies.

2.5.1 Refining Loops

In the first refinement exercise, you will refine one of the non-templated loops using the default sampling method.

If the Refine Loops panel is not open, open it:

- From Maestro, choose Tasks → Protein Refinement → Refine Loops.
- From BioLuminate, choose Tasks → Loop + Sidechain Prediction → Refine Loops.

  The Refine Loops panel opens.

1. For Find loops in Workspace structure, click Non-Template.

   The Loops table is populated with loops that did not originate from the template, of which there are three.

2. Click the row for loop1 (but not in the Run column).

   Loop 1, which includes residues 52 through 56, is selected. Markers appear in the Workspace to indicate the location of this loop in the structure.

   Refinement of loops of six or more residues should be performed using extended, not default, sampling. You can change the sampling method in the Structure Refinement Options dialog box.

3. Click the check box for loop1 in the Run column.

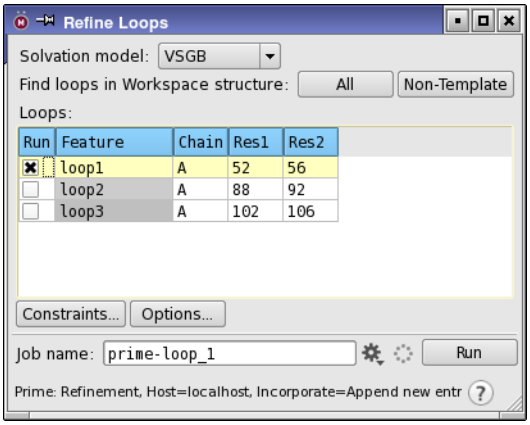

Figure 2.10. The Refinement panel with loop1 selected for refinement.

4. Click the Settings button.

   

   The Job Settings dialog box opens.

5. Choose Append new entries as a new group from the Incorporate option menu.

6. Enter `LoopRefinement` in the Name text box.

7. Click Run to launch the job.

   The refinement calculation is started. This job may take several minutes on a 2 GHz processor. When the job finishes, the predicted structure is incorporated into the Project Table and is displayed in the Workspace.

While we have been referring to the calculation that was just performed as a *refinement*, it is more accurately described as a *prediction*. The so-called refinement of loop 52-56 was in fact an ab initio loop prediction, because the program initially deleted the loop, reconstructed it in a particular way, and then exhaustively sampled it to identify the lowest energy conformation.

Refinement of loops that are less than 9 residues long yield excellent results in a large majority of cases. Loops 10 to 12 residues long yield very good results in a majority of cases. Loops 13 to 15 residues long produce a low energy conformation most of the time, but probably not the global minimum. Loops 16 to 20 residues long produce a low energy conformation, but refinement of loops this long will take on the order of 1-2 days. Loops longer than 20 residues long should not be attempted, partly because of the sampling problem, but also because the run times will be unreasonably long.

2.5.2 Minimizing Target Regions

Since only side chains (not the backbone) of residues within 7.5 Å were sampled during the previous loop refinement, it is not unreasonable to minimize the local environment of the loop before considering refinement complete.

1. Open the Minimize panel:
   - **Maestro:** choose Tasks → Protein Refinement → Minimize.
   - **BioLuminate:** choose Tasks → Implicit Solvent Refinement + Analysis → Minimize.
2. Click the Atom Selections button and choose Select.

   

   The Atom Selection dialog box opens with the Residue tab displayed and Residue number selected.
3. In the Residue Number text box, enter `52-56` and click Add.
4. Click Proximity.

   The Proximity dialog box is displayed.
5. Type `8.5` in the text box, select Residues, and click OK.
6. Click OK in the Atom Selection dialog box.

   Loop 52-56 and all residues within 8.5 Å are now selected.

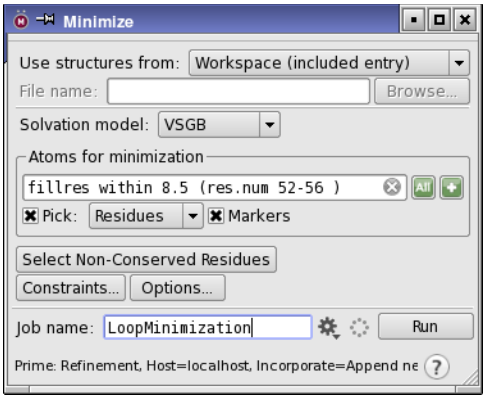

Figure 2.11. The Refinement panel with residues selected for minimization.

7. Enter `LoopMinimization` in the Job name text box.
8. Click Run to launch the job.

   The job is run with the default options. When the job finishes, the minimized structure is automatically incorporated into the project. Once a homology model has been refined, it can be used as input to other Schrödinger programs. Here, only one loop has been refined, so refinement of the other loops should be done before using the structure.

Homology Modeling of a Homomultimer

This chapter provides an exercise for building a homomultimer for hERG, which is a tetramer. The template that will be used is 1orq. This is not the highest scoring template that can be found from a BLAST search, but it is chosen to illustrate the use of a template in a realistic situation, where the similarity may not be high.

PDB structures often only include the asymmetric unit, so it may be necessary to build the multimeric template structure by applying the crystallographic symmetry operations. In the first part of the exercise, you will produce the desired tetrameric template from the PDB structure 1orq. Following that, you will create the homology model for hERG.

3.1 Preparing the Template

The template can be prepared by using the Protein Preparation Wizard. In addition to generating the tetramer, the parts of the structure that are not needed are deleted, and the chains are renamed so they can be distinguished when building the homology model.

1. Click the Prep Wiz button to open the Protein Preparation Wizard panel, or choose Tasks → Protein Preparation.

2. Enter 1orq in the PDB text box.
3. Select Biological unit.

   Selecting this option applies the symmetry operations to generate the tetramer, rather than importing the asymmetric unit.
4. Click Import.

   The tetramer is displayed in the Workspace after some processing.
5. In the main window, choose Select from the Delete button menu, or choose Edit → Delete → Select.

   The Atom Selection dialog box opens.

6. In the Atom tab, choose Element.
7. Select Cd from the list, click Add, then click OK.

   The cadmium atoms are removed from the structure. Cadmium or other heavy metals are often added to aid the determination of the structure from the X-ray data.
8. In the Review and Modify tab of the Protein Preparation Wizard panel, click Analyze Workspace.

   The structure is analyzed and the tables are populated.
9. Deselect Fit on select.
10. Select chains A and B, then click Delete.
11. In the waters table, press CTRL+A, then click Delete.
12. In the hets table, select hets 7 - 24, then click Delete.

    The hets table contains four copies of the potassium ions in the ion channel, which were generated when the biological unit was imported. Only one copy is needed.

The next stage is to relabel the chains.

13. In the main window, choose Edit → Build (or 3D Builder) → Residue Properties.
14. Choose Chain Name from the Property option menu.
15. Enter E in the Chain name text box, and click All.

    All atoms in the structure are now labeled as chain E. The four chains making up the tetramer will now be relabeled, leaving only the potassium ions in chain E.
16. Choose Molecules from the Pick option menu.

    The Pick option is also turned on when you do this.
17. Enter A in the Chain name text box, and pick one of the chains.

    See Figure 3.1 for an example of picking the chains.
18. Repeat the previous step with chain names B, C, and D for the other three chains.

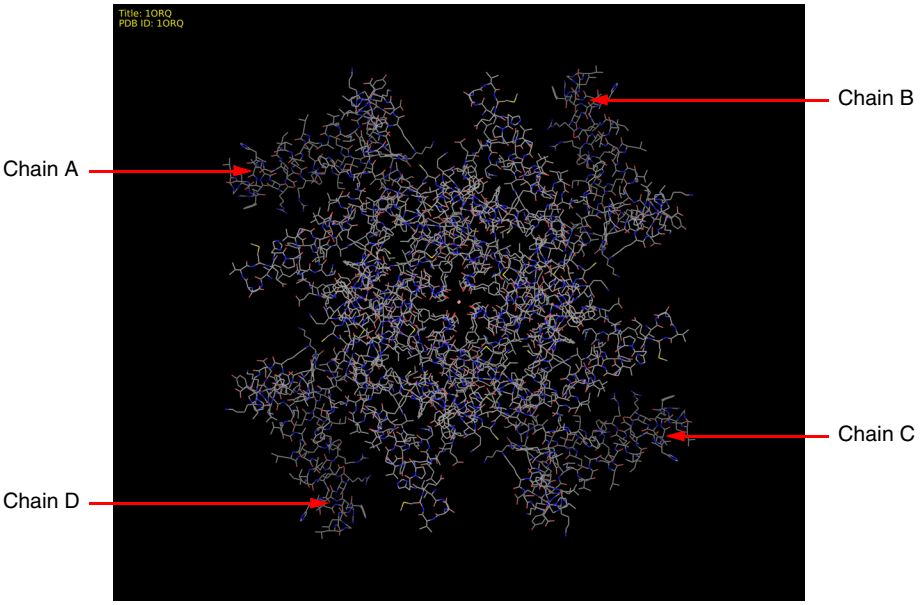

Figure 3.1. Chain labeling in the tetrameric template.

19. In the main window, choose Select from the Delete button menu, or choose Edit → Delete → Select.

    

    The Atom Selection dialog box opens.
20. In the Residues tab, choose Residue number from the list.
21. Enter 1-104 in the Residue number text box and click Add.
22. In the Chain tab, select chain E.
23. Click Subtract, then click OK.

The structure is now ready for use as a template. If you examine the status bar below the Workspace, it should report that there are 4178 atoms, 550 residues, 5 chains, and 10 molecules.

3.2 Importing the Sequence

If the Structure Prediction panel is not already open, open it now.

From Maestro:

- Choose Tasks → Homology Modeling in the main window, then click the Structure Prediction Wizard icon in the Homology Modeling panel.

From BioLuminate:

- Choose Tasks → Homology Modeling → Advanced Homology Modeling in the main window.

The Structure Prediction panel opens at the first step, Input Sequence. By default the Guide is not displayed. If you want to display it, choose Step → Guide.

1. Choose File → New, and name the run `hERG`.
2. For Get sequence from, click File.
3. Select `hERG.fasta`, then click Open.

   The sequence is displayed in the Prime sequence viewer.
4. Click Next to proceed to the next step, Find Homologs.

3.3 Importing the Template

1. Open the Project Table panel by pressing CTRL+T in the Workspace, or clicking the Table toolbar button.

2. Select the 1ORQ entry.

   If this is a new project, it should be the only entry in the table. It should be the one that is marked with a red square in the In column, as it is displayed in the Workspace.
3. In the Structure Prediction panel, click From Project Table.
4. If a dialog box prompts you about overwriting the existing `1orq.pdb` file, click Yes.

   Four sequences are displayed in the Homologs table. The first of these is selected.

5. Shift-click the last sequence in the table to select all four sequences.
6. Click Next to proceed to the next step, Edit Alignment.
7. In the warning dialog box that is displayed, click Continue.

   This warning is displayed because you selected multiple templates, and there are several ways of using multiple templates for homology modeling. In this case, the structures are properly aligned for building a multimer, so there is no need to perform any actions.

   The Edit Alignment step is displayed, showing the query and the four template chains.

3.4 Building the Model

The initial alignment of the query is satisfactory for model building, so you do not need to perform any alignment.

8. Click Next to proceed to the Build Structure step.
9. Click Continue in the warning dialog box that is displayed.

   The warnings are largely related to the missing loops in the template, which will not be built into the model. You can run a loop refinement to fill in the missing loop later if you wish.
10. Under Model-building method, choose Energy-based.
11. Under Multi-template model type, select Homo-multimer.
12. Select all of the cofactors (potassium ions) in the Include ligands and cofactors list.
13. Click Build Model.

The job is started. It takes about 20 minutes. When it finishes, a dialog box opens, asking if you want to open the Refinement panel to refine the structure. Normally, you should refine the structure before using it for other modeling studies. For this structure, since the template was missing a loop in each chain, you would probably want to run a loop prediction on these loops. This exercise is finished with the building of the model, so you can dismiss this dialog box.

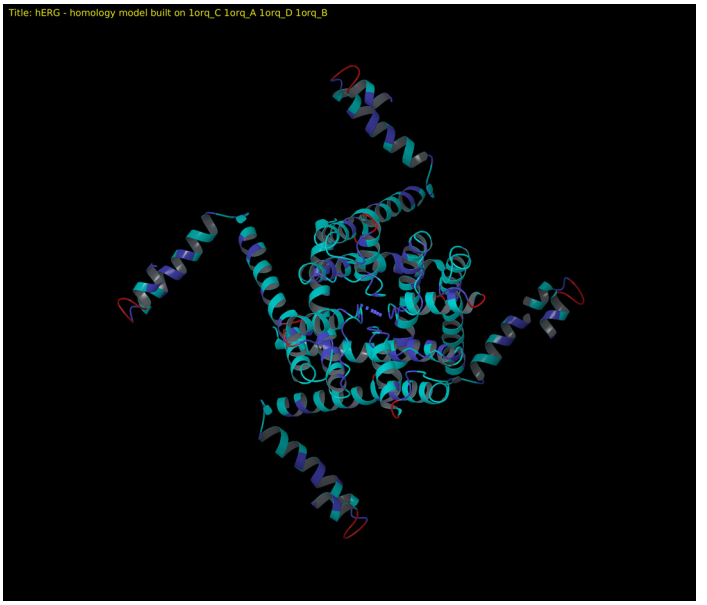

Figure 3.2. Final model of hERG.

Homology Modeling of a Heteromultimer

This chapter provides an exercise for building a heteromultimer, consisting of CDK4 and Cyclin-E.

If the Structure Prediction panel is not already open, open it now.

From Maestro:

- Choose Tasks → Homology Modeling in the main window, then click the Structure Prediction Wizard icon in the Homology Modeling panel.

From BioLuminate:

- Choose Tasks → Homology Modeling → Advanced Homology Modeling in the main window.

The Structure Prediction panel opens at the first step, Input Sequence. By default the Guide is not displayed. If you want to display it, choose Step → Guide.

4.1 Preparing the CDK4 Chain

1. Choose File → New.
2. Name the run `CDK4`.
3. For Get sequence from, click File.
4. Select `cdk4.fasta` and click Open.

   The sequence is displayed in the Prime sequence viewer.
5. Click Next to proceed to the next step, Find Homologs.
6. For Import, click From File.
7. Choose `3f5x.pdb` and click Open.

   Four sequences are displayed in the Homologs table. The first of these is selected.
8. Select the 3f5x_A sequence in the table.

9. Click Next to proceed to the next step, Edit Alignment.

   The Edit Alignment step is displayed, showing the query and the selected template chain.

The alignment of the query is satisfactory, so you do not need to perform any alignment.

10. Click Next to proceed to the Build Structure step.
11. Select 3f5x_A EZV A:300 in the Include ligands and cofactors list.

This finishes the preparation of the first chain of the heteromultimer.

4.2 Preparing the Cyclin-E Chain

1. Choose File → New.
2. Name the run `CyclinE`.
3. For Get sequence from, click File.
4. Select `cyclinE.fasta` and click Open.

   The sequence is displayed in the Prime sequence viewer.
5. Click Next to proceed to the next step, Find Homologs.
6. For Import, click From File.
7. Choose `3f5x.pdb` and click Open.

   Four sequences are displayed in the Homologs table. The first of these is selected.
8. Select the 3f5x_B sequence in the table.
9. Click Next to proceed to the next step, Edit Alignment.

   The Edit Alignment step is displayed, showing the query and the selected template chain.
10. Click the View SSA button.

    

    The secondary structure assignment of the template is shown.
11. Click the Add/Remove Anchors button.

    
12. Click on residue Tyr175 (Y) in the query to place an anchor at this point.

This residue is at the beginning of a gap in the template. An anchor symbol is placed in the ruler above the residue.

13. Click the Slide Freely button.

14. Slide template residue Leu341 (L) left to close the template gap.
15. This residue is the sixth residue after the gap.
16. Click the Add/Remove Anchors button.

17. Click on residue Asp195 (D) in the query to place an anchor at this point.

    This residue is a little after the gap in the template. An anchor symbol is placed in the ruler above the residue.

18. Click the Slide Freely button.

19. Slide query residue Pro200 (P) left to close the query gap.

    The alignment of the query is now satisfactory.

20. Click Next to proceed to the Build Structure step.
21. Click Continue in the warning dialog box that is displayed.

    The warnings are related to the missing loop in the template, which will not be built into the model. You can run a loop refinement to fill in the missing loop later if you wish.

This finishes the preparation of the second chain of the heteromultimer.

4.3 Building the Model

Now that both chains of the heteromultimer are prepared, you can build the model.

1. Under Model-building method, choose Energy-based.
2. Click Hetero-multimers.

   The Hetero-multimer Modeling dialog box opens. The instructions at the top describe the process you have just completed.
3. Select the CDK4 and CyclinE runs from the list.

   These are the only runs on the list if you ran this exercise with a new project. Otherwise, you might have runs from the previous tutorials (or other modeling runs).
4. Click Start.

   The dialog box closes, and the model-building job is started. The job takes about 25 minutes to run.

Figure 4.1. Final model of the heteromultimer.

Getting Help

Information about Schrödinger software is available in two main places:

- The `docs` folder (directory) of your software installation, which contains HTML and PDF documentation. Index pages are available in this folder.
- The Schrödinger web site, http://www.schrodinger.com/, In particular, you can use the Knowledge Base, http://www.schrodinger.com/kb, to find current information on a range of topics, and the Known Issues page, http://www.schrodinger.com/knownissues, to find information on software issues.

Finding Information in Maestro

Maestro provides access to nearly all the information available on Schrödinger software.

To get information:

- Pause the pointer over a GUI feature (button, menu item, menu, ...). In the main window, information is displayed in the Auto-Help text box, which is located at the foot of the main window, or in a tooltip. In other panels, information is displayed in a tooltip.

  If the tooltip does not appear within a second, check that Show tooltips is selected under General → Appearance in the Preferences panel, which you can open with CTRL+, (⌘,). Not all features have tooltips.

- Click the Help button in the lower right corner of a panel or press F1, for information about a panel or the tab that is displayed in a panel. The help topic is displayed in the Help panel. The button may have text or an icon:

  

- Choose Help → Online Help or press CTRL+H (⌘H) to open the default help topic.
- When help is displayed in the Help panel, use the navigation links in the help topic or search the help.
- Choose Help → Documentation Index, to open a page that has links to all the documents. Click a link to open the document.

- Choose Help → Search Manuals to search the manuals. The search tab in Adobe Reader opens, and you can search across all the PDF documents. You must have Adobe Reader installed to use this feature.

For information on:

- Problems and solutions: choose Help → Knowledge Base or Help → Known Issues → *product*.
- New software features: choose Help → New Features.
- Python scripting: choose Help → Python Module Overview.
- Utility programs: choose Help → About Utilities.
- Keyboard shortcuts: choose Help → Keyboard Shortcuts.
- Installation and licensing: see the *Installation Guide*.
- Running and managing jobs: see the *Job Control Guide*.
- Using Maestro: see the *Maestro User Manual*.
- Maestro commands: see the *Maestro Command Reference Manual*.

Contacting Technical Support

If you have questions that are not answered from any of the above sources, contact Schrödinger using the information below.

Web:	http://www.schrodinger.com/supportcenter
E-mail:	help@schrodinger.com
Mail:	Schrödinger, 101 SW Main Street, Suite 1300, Portland, OR 97204
Phone:	+1 888 891-4701 (USA, 8am – 8pm Eastern Time)
	+49 621 438-55173 (Europe, 9am – 5pm Central European Time)
Fax:	+1 503 299-4532 (USA, Portland office)
FTP:	ftp://ftp.schrodinger.com

Generally, using the web form is best because you can add machine output and upload files, if necessary. You will need to include the following information:

- All relevant user input and machine output
- Prime purchaser (company, research institution, or individual)
- Primary Prime user
- Installation, licensing, and machine information as described below.

Gathering Information for Technical Support

The instructions below describe how to gather the required machine, licensing, and installation information, and any other job-related or failure-related information, to send to technical support. Where the instructions depend on the profile used for Maestro, the profile is indicated.

For general enquiries or problems:

1. Open the Diagnostics panel.
   - **Maestro:** Help → Diagnostics
   - **Windows:** Start → All Programs → Schrodinger-2015-2 → Diagnostics
   - **Mac:** Applications → Schrodinger2015-2 → Diagnostics
   - **Command line:** `$SCHRODINGER/diagnostics`
2. When the diagnostics have run, click Technical Support.

   A dialog box opens, with instructions. You can highlight and copy the name of the file.
3. Upload the file specified in the dialog box to the support web form.

   If you have already submitted a support request, use the upload link in the email response from Schrödinger to upload the file. If you need to submit a new request, you can upload the file when you fill in the form.

If your job failed:

1. Open the Monitor panel, using the instructions for your profile as given below:
   - **Maestro/Jaguar/Elements:** Tasks → Monitor Jobs
   - **BioLuminate/MaterialsScience:** Tasks → Job Monitor
2. Select the failed job in the table, and click Postmortem.

   The Postmortem panel opens.
3. If your data is not sensitive and you can send it, select Include structures and deselect Automatically obfuscate path names.
4. Click Create.

   An archive file is created, and an information dialog box with the name and location of the file opens. You can highlight and copy the name of the file.
5. Upload the file specified in the dialog box to the support web form.

   If you have already submitted a support request, use the upload link in the email response from Schrödinger to upload the file. If you need to submit a new request, you can upload the file when you fill in the form.

6. Copy and paste any log messages from the window used to start the interface or the job into the web form (or an e-mail message), or attach them as a file.
   - **Windows:** Right-click in the window and choose Select All, then press ENTER to copy the text.
   - **Mac:** Start the Console application (Applications → Utilities), filter on the application that you used to start the job (Maestro, BioLuminate, Elements), copy the text.

If Maestro failed:

1. Open the Diagnostics panel.
   - **Windows:** Start → All Programs → Schrodinger-2015-2 → Diagnostics
   - **Mac:** Applications → SchrodingerSuite2015-2 → Diagnostics
   - **Linux/command line:** `$SCHRODINGER/diagnostics`
2. When the diagnostics have run, click Technical Support.

   A dialog box opens, with instructions. You can highlight and copy the name of the file.
3. Upload the file specified in the dialog box to the support web form.

   If you have already submitted a support request, use the upload link in the email response from Schrödinger to upload the file. If you need to submit a new request, you can upload the file when you fill in the form.
4. Upload the error files to the support web form.

   The files should be in the following location:

- **Windows:** `%LOCALAPPDATA%\Schrodinger\appcrash`
  (Choose Start → Run and paste this location into the Open text box.)
  Attach `maestro_error_`*`pid`*`.txt` and `maestro.exe_`*`pid_timestamp`*`.dmp`.
- **Mac:** `$HOME/Library/Logs/CrashReporter`
  (Go → Home → Library → Logs → CrashReporter)
  Attach `maestro_error_`*`pid`*`.txt` and `maestro_`*`timestamp_machinename`*`.crash`.
- **Linux:** `$HOME/.schrodinger/appcrash`
  Attach `maestro_error_`*`pid`*`.txt` and `crash_report_`*`timestamp_pid`*`.txt`.

If a Maestro panel failed to open:

1. Copy the text in the dialog box that opens.
2. Paste the text into the support web form.

Glossary

alignment—The optimal matching of residue positions between sequences, typically a query sequence and one or more template sequences.

anchor—A constraint on alignment set at a given residue position. Alignment changes must preserve the query-template pairing at that residue until the anchor is removed.

ASD—Atom Selection dialog box.

ASL—Atom Specification Language.

button menu—The menu available from a toolbar menu button, which you open by holding down the left mouse button.

Comparative Modeling—Protein structure modeling based on a query-template match with a substantial percentage of identical residues (usually 50% or greater sequence identity).

composite template—A type of template used in the Threading Path, produced from the core (invariable) and variable regions of a family of structurally similar proteins.

constraints—Tools to keep regions of a sequence (alignment constraints) or structure (during minimization) in a particular configuration.

deletions—The residues missing from a query sequence that are present in a template sequence.

entry—A structure or set of structures and associated properties. Entries are represented as a row in the project table, and can be used as input for jobs.

Fold Recognition—The use of secondary structure matching and profiles generated from multiple sequence/structure alignments to find templates when sequence methods are unsuccessful.

gaps—The spaces in an alignment resulting from insertions and deletions.

HETATMs—The atoms of residues, including amino acids, that are not one of the standard 20 amino acids. In PDB files, HETATM.

homolog—A sequence/structure related to the query sequence; i.e., a sequence with many of the same residues in the same patterns as the query sequence. Usually these sequences are derived from the same family and may have similar function.

insertions—The extra residues found in a query sequence that are not found in a template sequence.

loop—A region of undefined secondary structure.

Maestro toolbar—The array of icon buttons which provides tools for common Maestro tasks, located by default along the left side of the main window. There are buttons for operations such as moving structures in the Workspace, changing what is displayed, opening a project, or undoing the most recent Maestro operation.

Main menu bar—The menu bar at the top of the main Maestro window below the Auto-Help window. The main menu bar contains menu titles (Maestro, Project, Edit, etc.) that, when clicked, display menus from which selections can be made.

menu button—A toolbar button that has a menu, which you open by holding down the left mouse button. The button has a black triangle in the lower right corner.

Prime toolbar—The row of icon buttons which provides tools for common Prime tasks, located near the top of the Prime-SP panel.

project—A collection of related data, such as structures with their associated properties. In Prime a project comprises one or more *runs* (executions of the Prime workflow). The project may include data that does not appear in the *project table*.

project table—The Maestro panel associated with a project, featuring a table with rows of entries and columns of properties.

query sequence—A sequence of unknown structure or fold.

Ranking Score—The score used to rank composite templates derived from different seed templates. Generated by the Global Scoring Function in the Threading Path.

refinement—An improvement of a model structure through energy-based optimization of selected regions.

run—A single execution of the Prime workflow using a particular set of choices (of templates, of Paths, and of settings). Each run belongs to a *project*. Runs cannot be saved without saving the project to which they belong.

SSA—Secondary structure assignment.

SSP—Secondary structure prediction.

sequence viewer—An area in which protein sequences are displayed. Right-clicking a sequence opens an *option menu*.There are sequence viewers in the Prime–SP panel and in the Maestro main window. The Prime sequence viewer displays query and template sequences,

including family and conservation data in sequence format, *SSA*s, and *SSP*s. The Workspace sequence viewer displays the sequence and (by default) the SSA for the structures included in the Workspace, provided that they are entries in a named Maestro project.

template sequence—A sequence of known structure and fold used as a basis for building a model of the query.

Threading—A structure prediction process in which *Fold Recognition* is used to define templates, then backbone models are built via alignment to composite templates and refined. May be used when query-template sequence identity is low.

Workspace—The open area in the center of the Maestro main window in which structures are displayed.

Z-Score—Measures the compatibility of the query sequence with the model structure, relative to the compatibility of randomly shuffled sequences of the same composition.

120 West 45th Street
17th Floor
New York, NY 10036

101 SW Main Street
Suite 1300
Portland, OR 97204

245 First Street
Riverview II, 18th Floor
Cambridge, MA 02142

8910 University Center Lane
Suite 270
San Diego, CA 92122

155 Gibbs St
Suite 430
Rockville, MD 20850-0353

Dynamostraße 13
D-68165 Mannheim
Germany

Zeppelinstraße 73
D-81669 München
Germany

Potsdamer Platz 11
D-10785 Berlin
Germany

Quatro House
Frimley Road
Camberley GU16 7ER
United Kingdom

8F Pacific Century Place
1-11-1 Marunouchi
Chiyoda-ku, Tokyo 100-6208
Japan

No. 102, 4th Block
3rd Main Road, 3rd Stage
Sharada Colony
Basaveshwaranagar
Bangalore 560079, India

SCHRÖDINGER®