User's Guide

UNIO'10

Automated NMR Data Analysis for Protein Structure Determination and More

Web : http://www.unio-nmr.eu
Email: torsten.herrmann@ens-lyon.fr

Chapter 1 Introduction

Before getting started with UNIO, there are a number of things you may want to know. This chapter introduces UNIO, tells you where to begin, how to install, register/obtain a license for academic or commercial use and how to use this documentation.

1.1 Welcome

UNIO enables you to perform automated NMR data analysis for protein 3D structure determination. UNIO represents the result of more than a decade of basic research performed in order to enable accurate, objective and highly automated protein structure determination by NMR. NMR data analysis interconnects the MATCH algorithm for sequence-specific backbone assignment, the ASCAN algorithm for side-chain assignment[2] and the CANDID algorithm for NOE assignment.[3] The ATNOS algorithm for robust signal identification in NMR spectra is employed to grant each of the three expert-systems direct access to the raw NMR data[4], thus enabling integrated and consistent data evaluation. The UNIO data analysis protocol requires only a minimal set of 3-6 NMR spectra, namely APSY and NOESY, and features high computational efficiency.

The experience gained so far (> 20 de novo structure determinations performed with the entire protocol, i.e., with successive use of the MATCH, ATNOS/ASCAN and ATNOS/CANDID modules) shows that the UNIO approach leads to accurate, objective and consistent NMR data interpretation. Individual UNIO modules, especially the CANDID module, have been used in several hundreds of de novo structure determination projects.

[1] Volk, J.; Herrmann, T.; Wüthrich, K. *J. Biomol.NMR.* **2008**, *41*, 127-138.

[2] Fiorito, F.; Damberger, F.F.; Herrmann, T.; Wüthrich, K. *J. Biomol. NMR* **2008**, *42*, 23-33.

[3] Herrmann, T.; Güntert, P.; Wüthrich, K. *J. Mol. Biol.* **2002**, *319*, 209-227.

[4] Herrmann, T.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **2002**, *24*, 171-189.

[5] Herrmann, T. Encycl. Magn. Res. 2010

The goal of this manual is to provide the most accessible, comprehensive and useful guide to UNIO possible. We aim to provide a level of depth that covers the advanced feature set, while remaining accessible to beginners. We hope you find this material useful and enjoy using UNIO.

1.2 Where to Begin

We recommend you to start by taking a look at the *Getting Started* section within the product documentation. You can access this by choosing **Getting Started** in the main window, then click on *Getting Started* for one of the three UNIO NMR data analysis tools, MATCH, ATNOS/ASCAN or ATNOS/CANDID.

1.3 System Requirements

1.3.1 All Operating Systems

Minimal requirements for best graphical user experience are:

- 1024x768 or larger monitor
- True color display (16-bit or 32-bit depth)
- ksh shell installed in the directory /bin, i.e., as shell path /bin/ksh

1.3.2 Requirements for Mac OS X Systems

You can perform UNIO NMR data analysis on any OS X system that includes:

- Any OS X-capable Macintosh
- Operating system: OS X 10.0 or later using Intel Processors. PPC Mac OSX systems are not supported.
- Memory: 256 MB free
- Disk space: at least 50 MB free

1.3.3 Requirements for Linux Systems

The minimal requirements to run UNIO are:

- Linux kernel 2.4 or above
- Processor type: i386, i686 or x86_64
- Memory: 256 MB free
- Disk space: at least 50 MB free

1.4 Installation Instructions

At present, UNIO is only available as a zip archive - there is currently no installer provided. You need to unzip the archive into a suitable folder, e.g., into your Desktop folder. From the Finder you can simply double click the zip-archive, then navigate to the created folder UNIO_10 and click on the executable file Unio10_Linux or Unio10_MacOSX to launch UNIO. From the command line you can do the following:

```
mv UNIO_10.zip ~/Desktop/.
cd ~/Desktop
gunzip UNIO_10-DistributionDate.tar.gz
tar xvf UNIO_10-DistributionDate.tar
cd UNIO_10
./Unio08_Linux or ./Unio08_MacOSX
```

1.5 Registration, Support and Updates

UNIO offers up and running scientific and technical support to all users to assist you in getting your NMR structure determination project started.

Information about registration, license for academic and commercial use, support and updates are available under the website
http://www.unio-nmr.eu

You can contact us by writing to the email address: torsten.herrmann@ens-lyon.fr

1.6 Using the Documentation

1.6.1 Documentation Conventions

Screenshots are used extensively throughout the documentation and are used to illustrate all features of UNIO (more than 100 figures are shown to get you acquainted with UNIO).

If icons, buttons, text field and labels or view areas of UNIO are mentioned in the text, then these items are highlighted in *italics* in the text.

Note: This style of text box provides some additional information that may be useful or emphasizes a key point.

Important: This style of text box reports about an important aspect of UNIO NMR data analysis.

Caution: This style of text box warns you about a potential problem and gives solutions to overcome this problem.

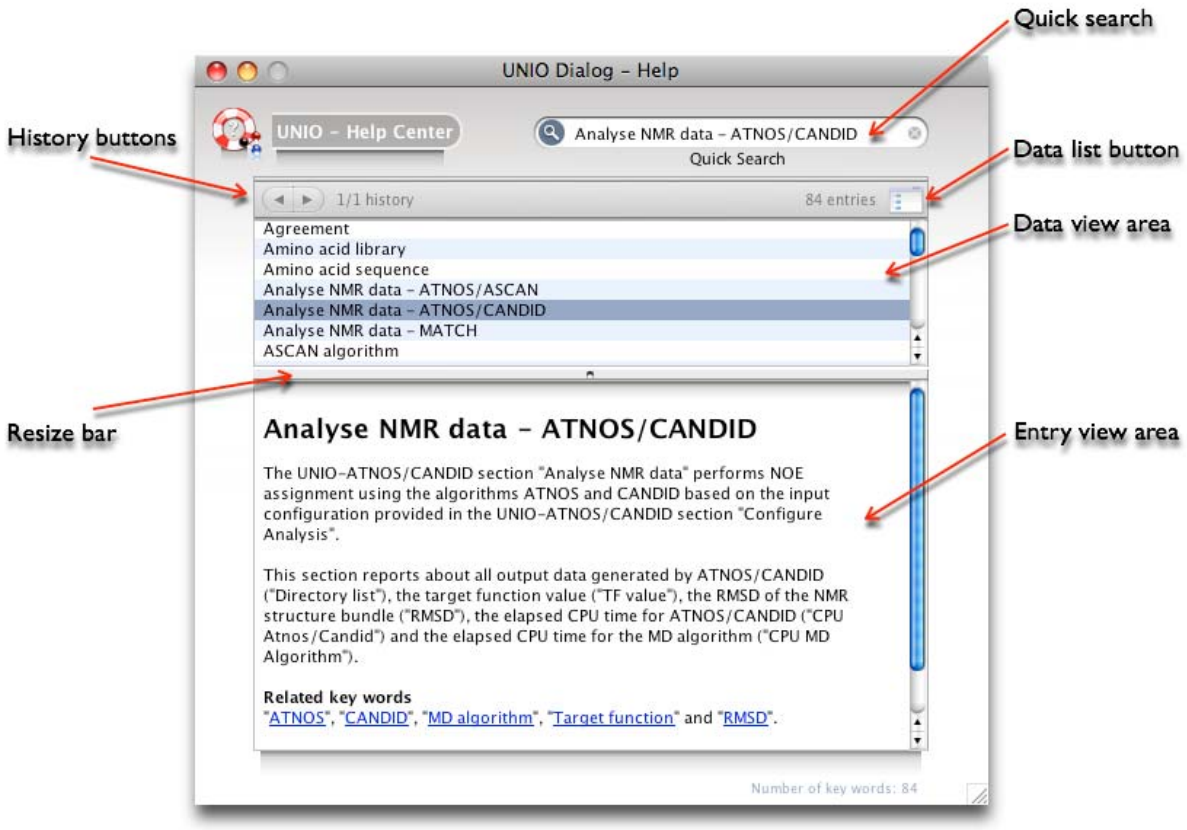
1.6.2 Navigating the Documentation

The UNIO documentation is spread across six components; the UNIO Help Center, the UNIO Syntax Dictionary, the Getting Started with MATCH, the Getting Started with ATNOS/ASCAN, the Getting Started with ATNOS/CANDID and this User's Guide.

The Getting Started documentations are aiming at introducing you to each of the three UNIO NMR data analysis tools and to provide you with enough knowledge to start performing your own NMR structure determination project. The UNIO Help Center will help you to find quick information while you performed a NMR data analysis task. The UNIO Syntax Dictionary will help you to write your own UNIO macros that can be executed in the UNIO Script Editor. The UNIO Script Editor along with the UNIO scripting language (ECHO) got designed to facilitate commonly required tasks, such as file format conversion, distance restraint or Ramachandran analysis, calculation of RMSD values and similar.

1.6.3 UNIO Help Center window

The UNIO Help Center contains complete information about all essential key words that will cross your path while performing a NMR structure determination. The UNIO Help Center can be searched for key words using the *Quick search* field in the top right of the help window.



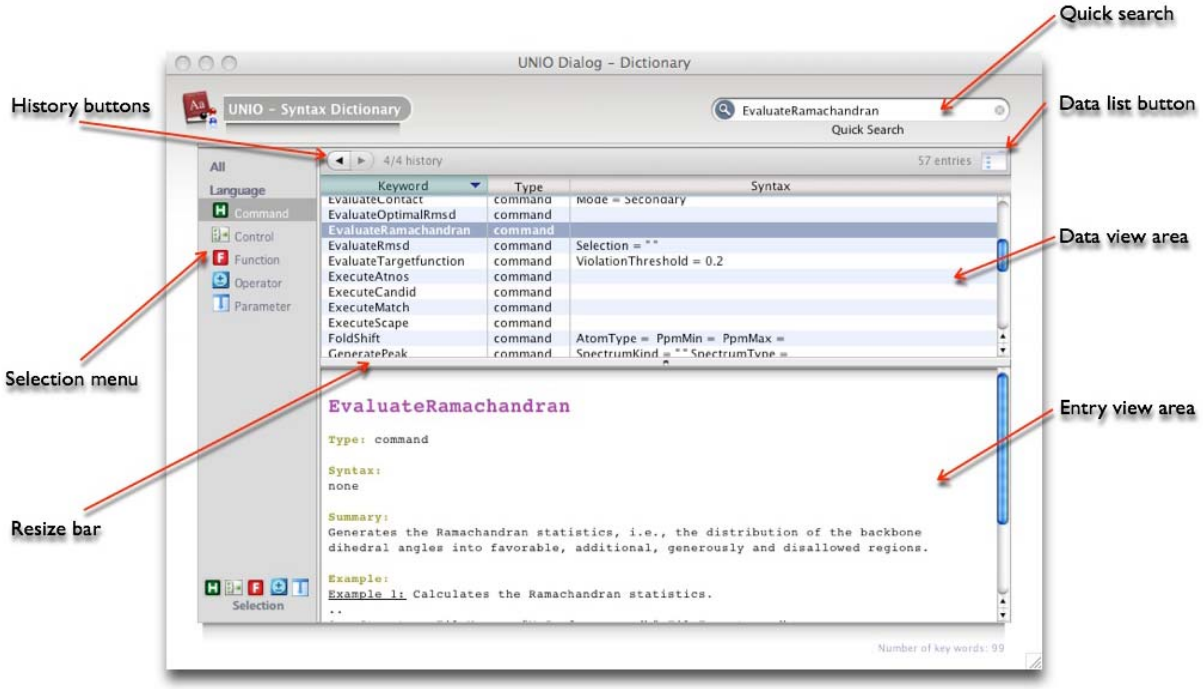
UNIO Help Center.

The documentation for a key word entry can be viewed by clicking on the entry header in the keyword index (*Data view area*). This will show the entry in the area underneath the *Data view area*. The height of the *Data view area* and *Entry view area* can be altered by dragging the *Resize bar*. Use the left and right *History buttons* to get forth and back to previous entries. Use the *Data list button* to see a listing of all keywords in the *Data view area*.

Tip: Use the listing of "Related key words" at the bottom of each entry to get a comprehensive coverage of a topic. Click on the hyperlink to jump to the given key word.

1.6.4 UNIO Syntax Dictionary window

The UNIO Syntax Dictionary provides complete information about all UNIO syntax words that can be used to write your own UNIO macro. The UNIO Syntax Dictionary can be searched for key words using the *Quick search* field in the top right of the help window.



UNIO Syntax Dictionary.

The documentation for a UNIO syntax word entry can be viewed by clicking on the entry header in the keyword index (*Data view area*). This will show the entry in the area underneath the *Data view area*. The height of the *Data view area* and *Entry view area* can be altered by dragging the *Resize bar*. Use the left and right *History buttons* to get forth and back to previous entries. Use the *Data list button* to see a listing of all syntax words in the *Data view area*. The *Data view area* can be sorted in ascending or descending order by mouse click on the Keyword column, the Type column or the Syntax column header. The *Selection menu* allows you to browse through the different UNIO syntax classes, i.e., UNIO syntax words are grouped into the Command, the Control, the Function, the Operator or the Parameter class according to their function in the UNIO scripting language (ECHO).

Tip: Use the listing of "Related key words" at the bottom of each entry to get a comprehensive coverage of a topic. Click on the hyperlink to jump to the given key word. The UNIO Script Editor (see below) got designed to support you in writing your own executable UNIO macros.

1.6.5 User's Guide

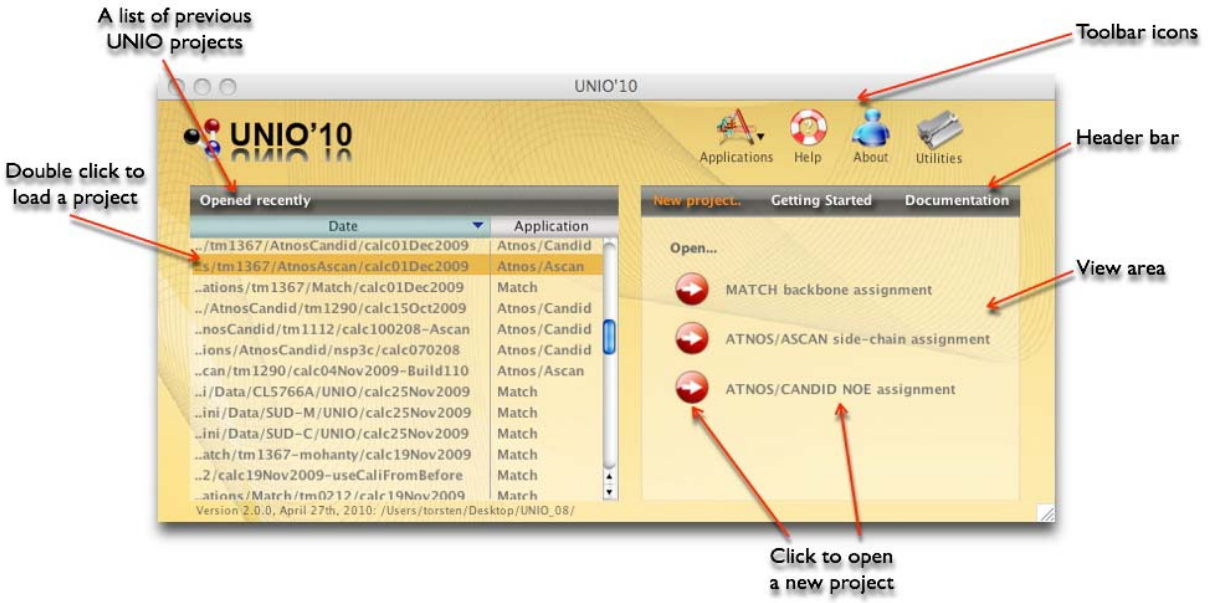
The User's Guide is a complete reference to UNIO. It details what the dialogs and windows do, and explains UNIO NMR data analysis in depth.

Chapter 2 The UNIO Environment

This section details the main components within UNIO NMR data analysis. The graphical user interface of UNIO contains all the features you need to quickly perform your own NMR data analysis and protein NMR structure determination. The main window allows you to open a new project, to load a previous project, to have access to Getting Started and User's Guide documentation and to open the UNIO Help Center, the UNIO Preference panel and the UNIO Notes window. The UNIO Application window (see below) allows you for each of the three UNIO data analysis tools - MATCH, ATNOS/ASCAN and ATNOS/CANDID - to configure your input, to analyze the NMR data and to evaluate the results obtained.

2.1 UNIO - Main window

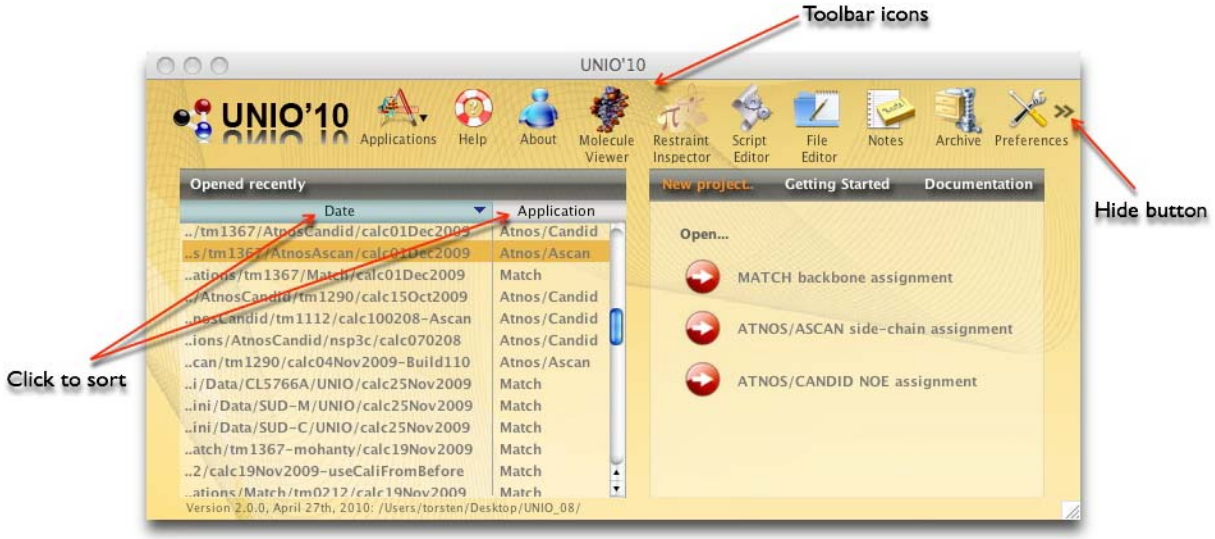
The UNIO main window consists of a main toolbar that provides easy access to commonly used functions. The left panel is used to load previous projects, the right view area allows you via the *Header bar* to start a new project ("New project.."), to browse the Getting Started documentation ("Getting Started") and to have access to the User's guide and the UNIO Help Center ("Documentation").



UNIO Main Window, section "New projects.."

Tip: The view area "*Open recently*" gives you fast access to previous projects. The list of previous projects can be sorted chronologically or Application-based by clicking on the corresponding column header.

A mouse click on the Toolbar icon "Utilities" reveals all UNIO utility programs: the UNIO Molecular Viewer, the UNIO Restraint Inspector, the UNIO Script Editor, the UNIO File Editor, the UNIO Notes, the UNIO Archive and the UNIO Preferences window. Clicking on the *Hide button* reduces the *Toolbar Icons* to its simplified view mode.



UNIO Main Window, section "New project.." Extended Toolbar icons view.

UNIO Main window, section "Getting Started".

UNIO Main window, section "Documentation".

Toolbar icons

The toolbar icons gives easy access to commonly used UNIO windows.

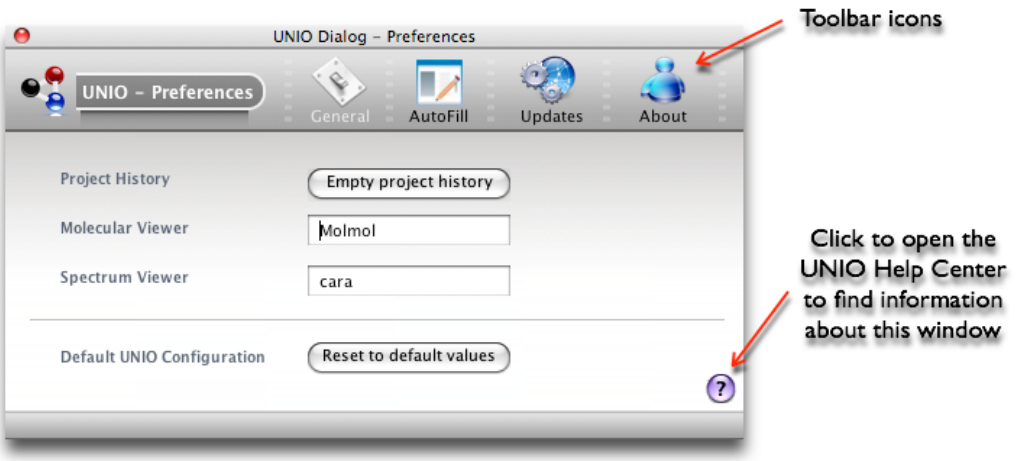
Applications	Opens an option menu to choose one of the three UNIO NMR data analysis tools, MATCH, ATNOS/ASCAN or ATNOS/CANDID.
Notes	Opens the UNIO Notes window.
About	Opens the UNIO About window.
Utilities	Enlarges the Toolbar to show all UNIO utility programs that are listed in the following rows.
Molecule Viewer	Opens the UNIO Molecule Viewer window.
Restraint Inspector	Opens the UNIO Restraint Inspector window.
Script Editor	Opens the UNIO Script Editor window.
File Editor	Opens the UNIO File Editor window.
Notes	Opens the UNIO Notes window.
Archive	Opens the UNIO Archive window.
Preferences	Opens the UNIO Preference window.

2.2 UNIO - Preferences window

The UNIO Preferences window is split into four sections "General", "AutoFill", "Updates" and "About". You can access the Preferences window from the *Toolbar icons* of the UNIO Main window (see Chapter 2.1).

2.2.1 Section "General"

The sub-window "General" allows you to empty the project history as listed in the UNIO Main window (see Chapter 2.1) and to specify external programs for displaying three-dimensional cartesian coordinates ("Molecular viewer"), and for displaying the NMR spectrum ("Spectrum viewer"). If you do not specify an external coordinate display progam, then the UNIO Molecular Viewer will be used to show the 3D protein structure. UNIO launches these external programs using the shell command given in the associated text fields. If you press the "*Reset to default values*" button, UNIO will be reset to its initial settings, however the history list will still be kept.



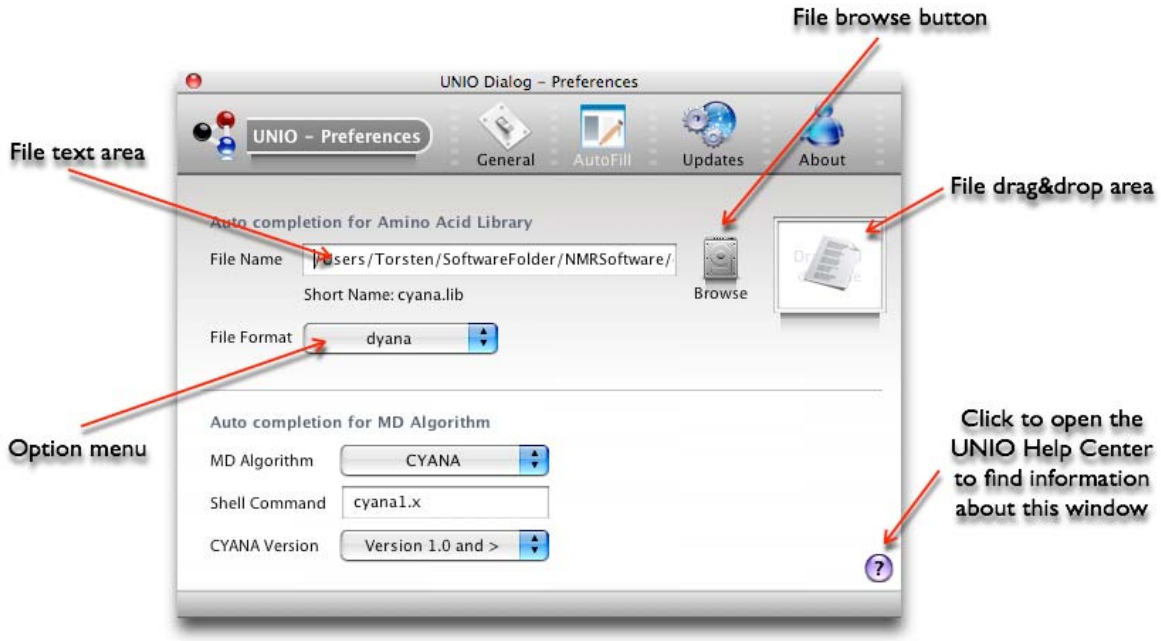
UNIO Preferences window, section General".

Tip: UNIO launches external programs using the shell commands provided. This assumes that the path for the shell commands to launch the programs are listed in your shell $PATH variable usually set in your .cshrc or .bashrc file in your home directory.

Tip: UNIO executes the shell command by adding the current file name of the NMR structure or NMR spectrum. This allows you to give optional parameters to your shell command. For example, you can type the command "Molmol -s" into the text field in order to automatically superimpose the NMR structure bundle.

2.2.2 Section "AutoFill"

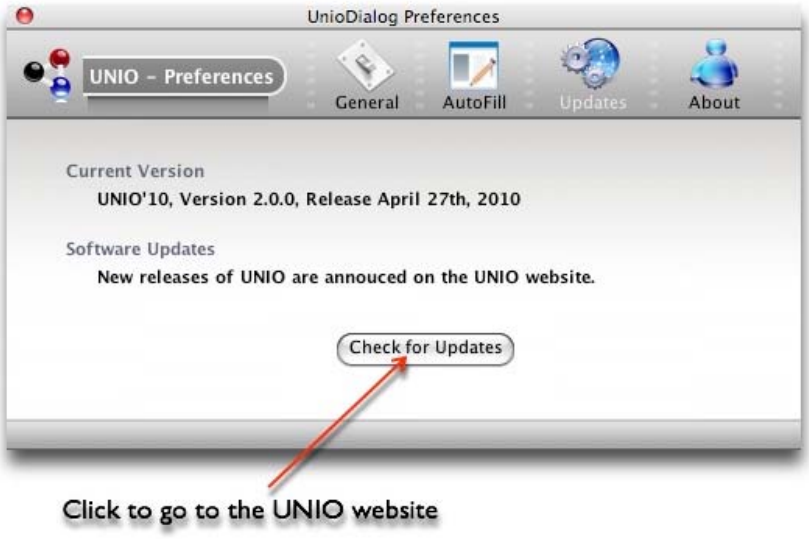
The sub-window "AutoFill" allows you to facilitate frequently performed tasks necessary for UNIO NMR data analysis. You can provide information about the amino acid library that is then automatically used in any new UNIO project. Similar you can use the AutoFill function for the MD program so that the NMR structures are always calculated with the given program for any new UNIO-ATNOS/CANDID project.



UNIO Preferences window, section "AutoFill".

2.2.3 Section "Updates"

The sub-window "Updates" allows you to check for new UNIO versions or updates. Pressing the *Check for Updates* button opens the UNIO website (http://www.unio-nmr.eu) in the system default internet browser.

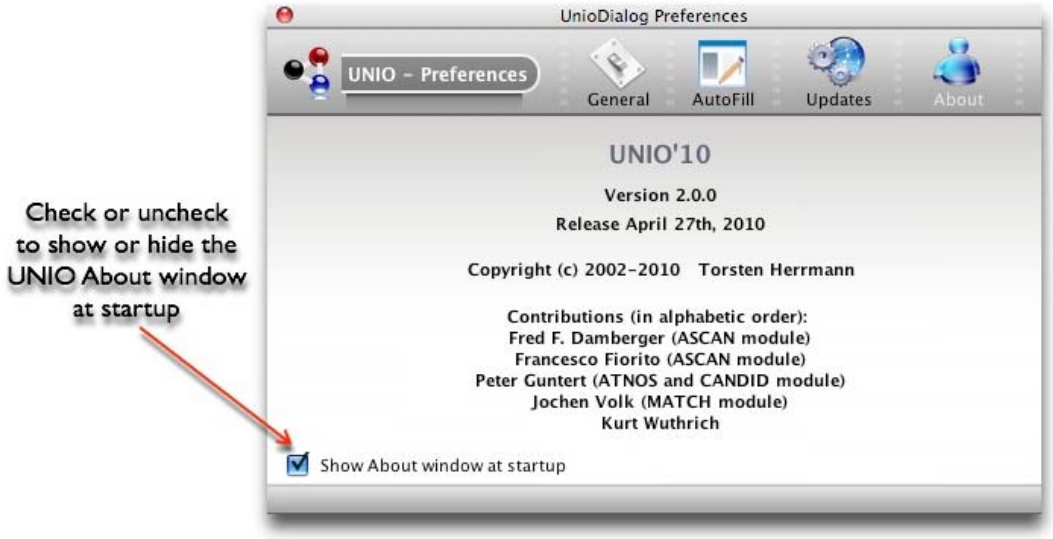


UNIO Preferences window, section "Updates".

2.2.4 Section "About"

The sub-window "About" provides information about the current software version, UNIO copyright and acknowledges people that contributed significantly to the UNIO NMR data analysis and protein NMR structure determination project.

Use the check box on the left bottom to show or hide the UNIO About window at startup.

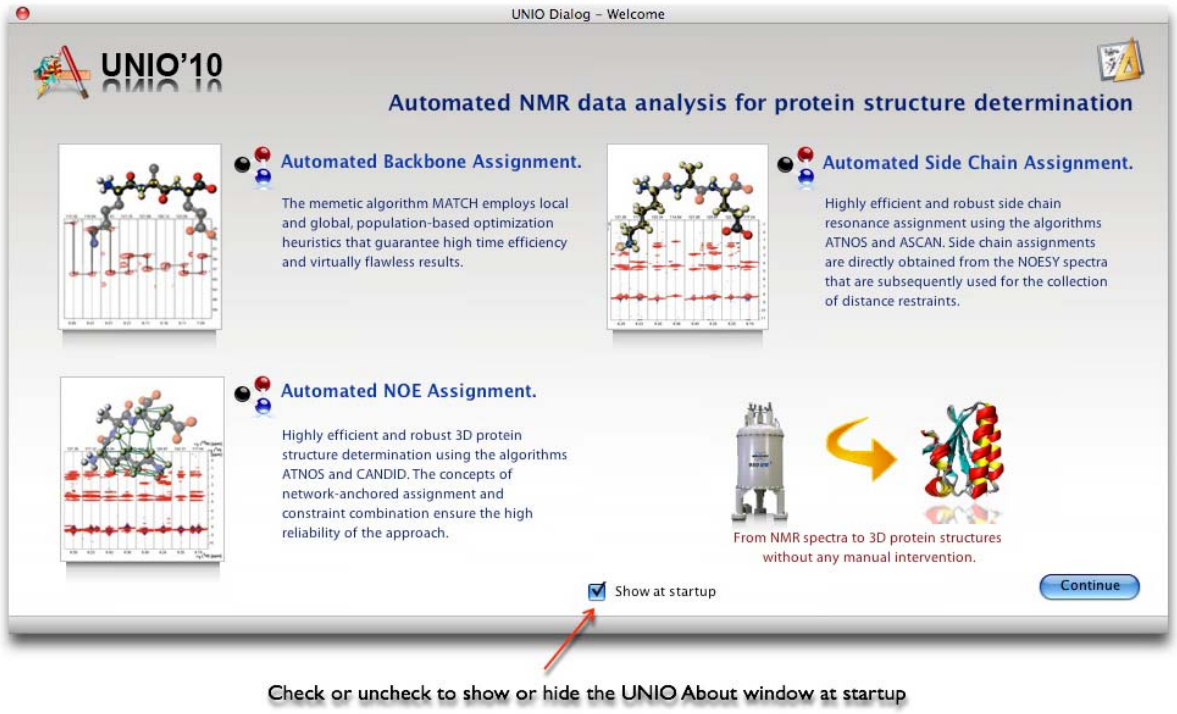


UNIO Preferences window, section "About".

2.3 UNIO - About window

The UNIO About window is shown at startup and provides information about the three UNIO NMR data analysis tools, MATCH for backbone assignment, ATNOS/ASCAN for side-chain assignment and ATNOS/CANDID for NOE assignment.

You can show or hide the UNIO About window at startup by checking or un-checking the option box at the bottom of the window.

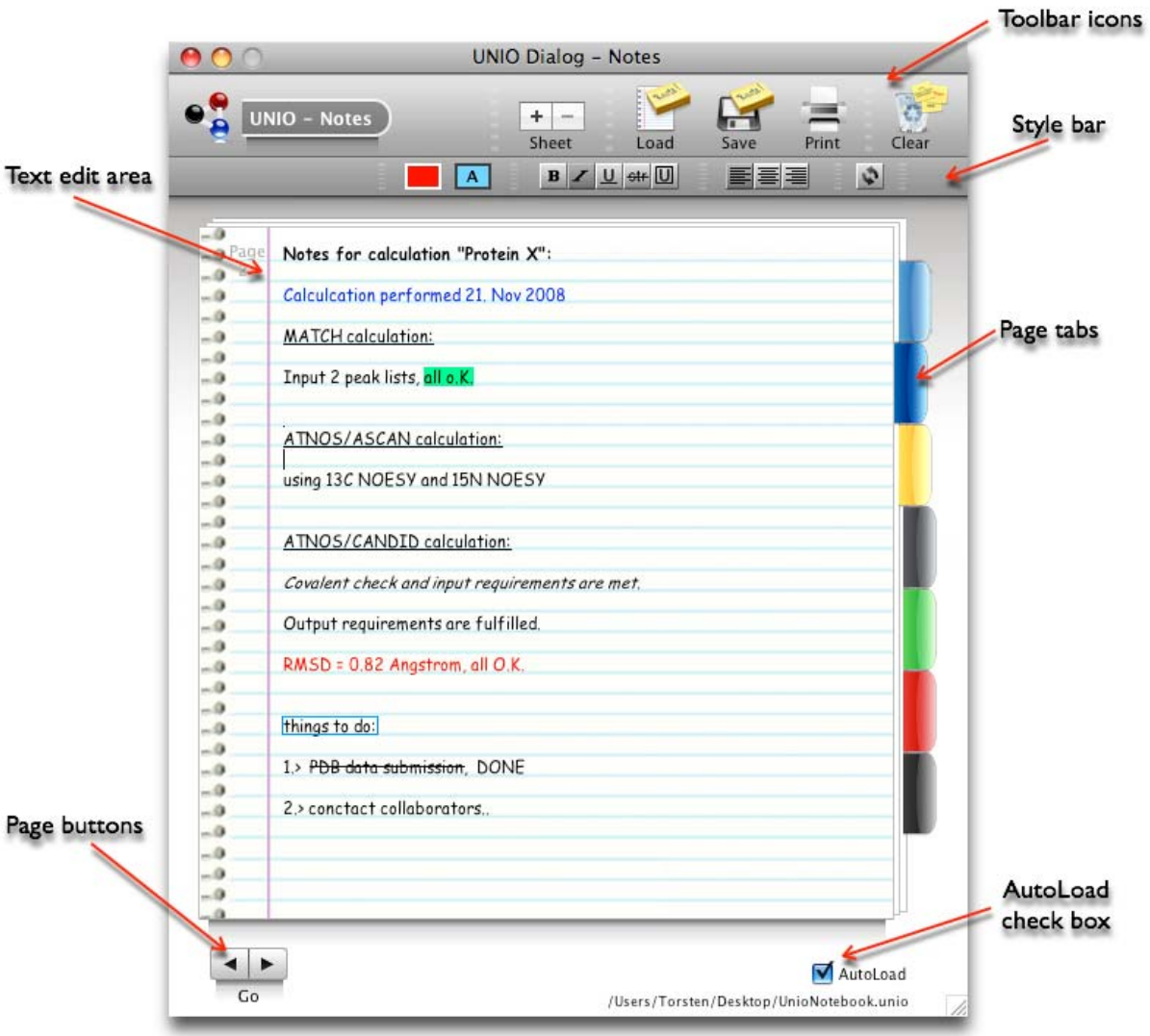


UNIO About window.

2.4 UNIO - Notes window

The UNIO Notes window is an advanced text editing and thoughts-storing tool and allows you to create and store notes within the UNIO environment.

The icons in the toolbar allow you to load an existing Note and to save or print a current Note. The *Clear button* in the toolbar deletes the text entries of all sheets. You can use the *Sheet buttons* in the toolbar to add and remove Note sheets. The *Page buttons* at the bottom of the window or the *Page tabs* on the left of the *Text edit area* allows you to quickly navigate between sheets. If the *AutoLoad button* is checked, then the UNIO Notes window will always open with the last Note that you have worked on. You can use the style bar to change the style of your text.



UNIO Notes window.

Style bar icons

The style bar icons allows you to change your text style. In the following the icons in the style bar are listed according to the appearance in the header bar from left to right.

Text color box	Opens the systems color palette that you can use to select the color of the text.
Text background color box	Opens the systems color palette that you can use to select the color of the text background.
Bold style icon	Change the text style to bold.
Italic text style icon	Change the text style to italics.
Underline text style icon	Underline text.
Strike out text style icon	Strike out text.
Box text style icon	Draw a box around the text.
Text align left icon	Align the text to the left.
Text align center icon	Center the text.
Text align right icon	Align the text to the right.
Style reset icon	Reset all selected styles and return to the default text style.

Tip: The *Style reset icon* at the most left of the style bar allows you to quickly return to the default text style that is plain style, black text color and no text background color.

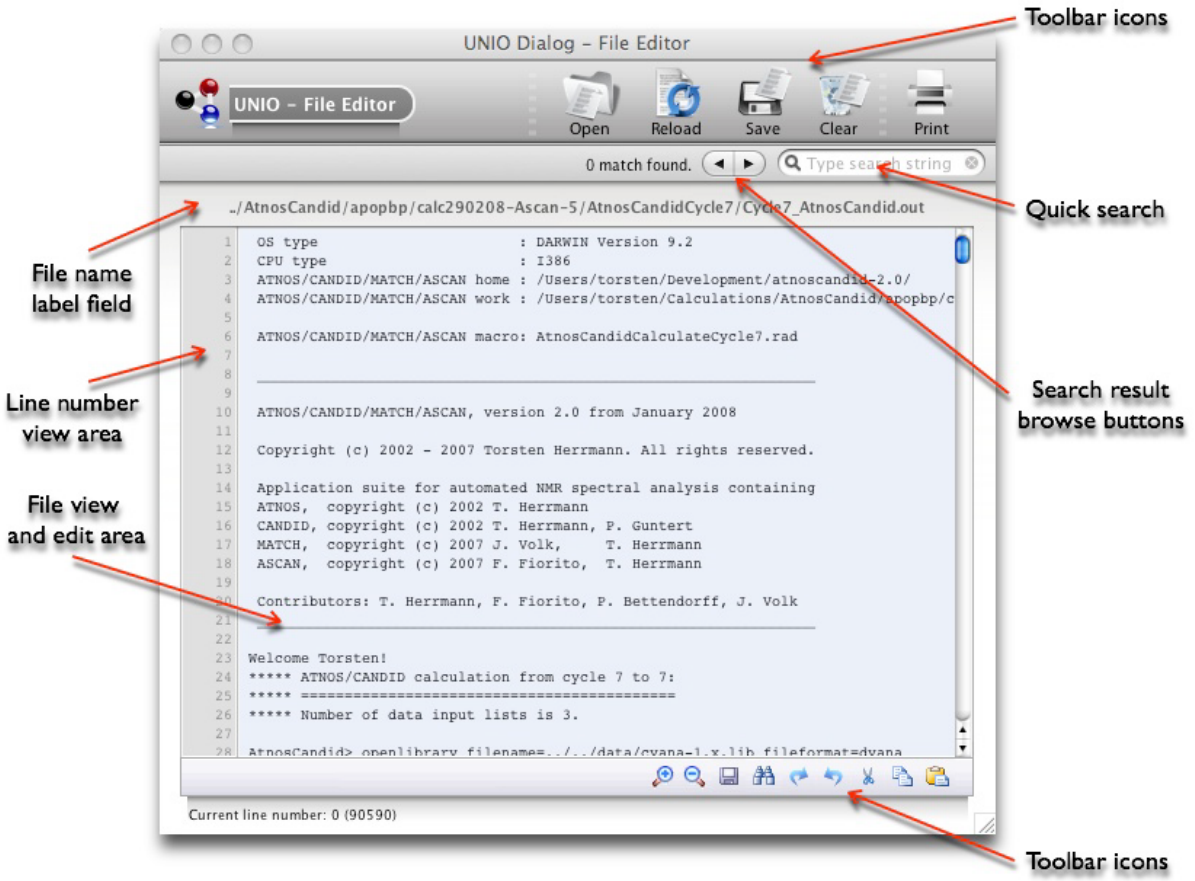
Tip: If you highlight text with the mouse and then click on a style icon, all text highlighted will change to the selected style.

2.5 UNIO - File Editor window

The UNIO File Editor allows you to open and edit files within the UNIO environment.

The top *Toolbar icons* allow you to open, reload, to save and to print the file. The Clear button resets the UNIO FIle Editor to its starting-up configuration. The path of the file is displayed in the *File name label field*. While browsing the *File view and edit area*, the current line of the file is displayed at the left bottom of the window. The *Line number view area* indicates the line number range of the currently displayed part of the file.

The bottom *Toolbar icons* allow you to increase and decrease the font size, to save the file, to find a particular search pattern, to redo and undo changes, to cut selected lines and to copy and paste selected text in the *File view and edit area* from and to the clipboard.



UNIO File Editor window.

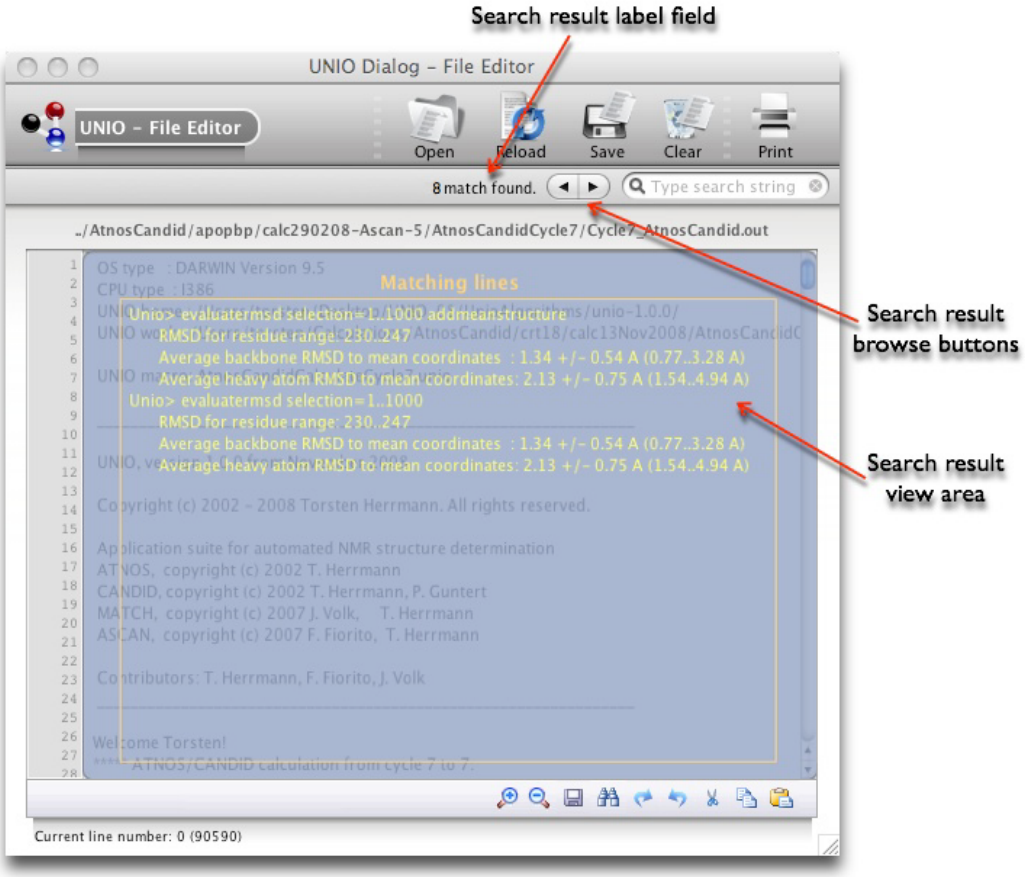
The UNIO File Editor offers also a *Quick search* field to find a string pattern in the given file. The table below gives an overview of allowed wild cards for the pattern filter.

Pattern filter

*	Matches zero or more of any character (including space character) Example: The filter pattern A*C matches ABC, AOPB or AS sdD.

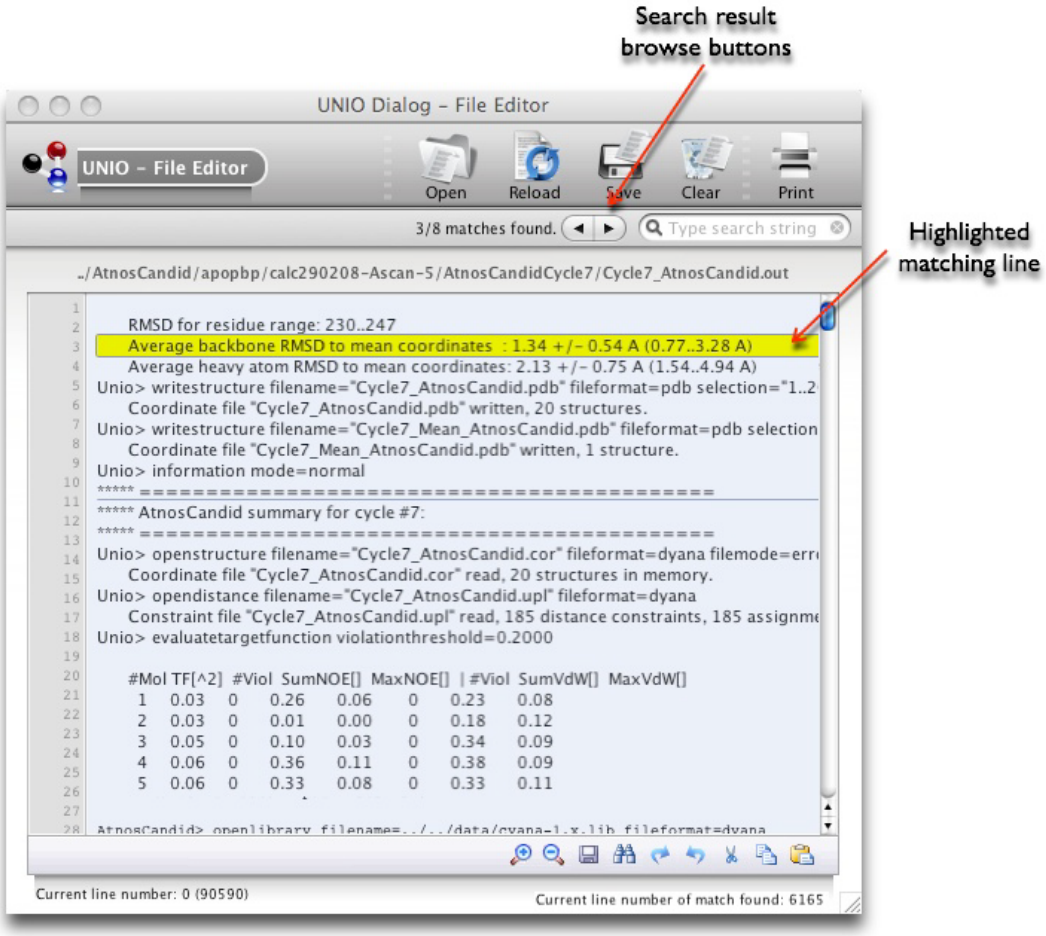
?	Matches exactly one character Example: The filter pattern A?C match ABC, but not AC or ADEC.
[chars]	Matches any of the characters in the brackets. Example: The filter pattern A[BC]D matches ABD, ACD, but not AD or ABCD.
[char-char]	Matches any character whose ASCII value is between the first and the second character. Example: The filter pattern A[B-K] matches ABD or AGD, but not ALD or ABKD
go to line N	Key phrase to fine line number N. Example: go to line 100; go to line -10.

While you are typing a search string in the *Quick search* field, matching lines are immediately displayed in the *Search result view area*. This feature allows you to narrow your search by refining the search pattern against the intermediate matching lines listed.



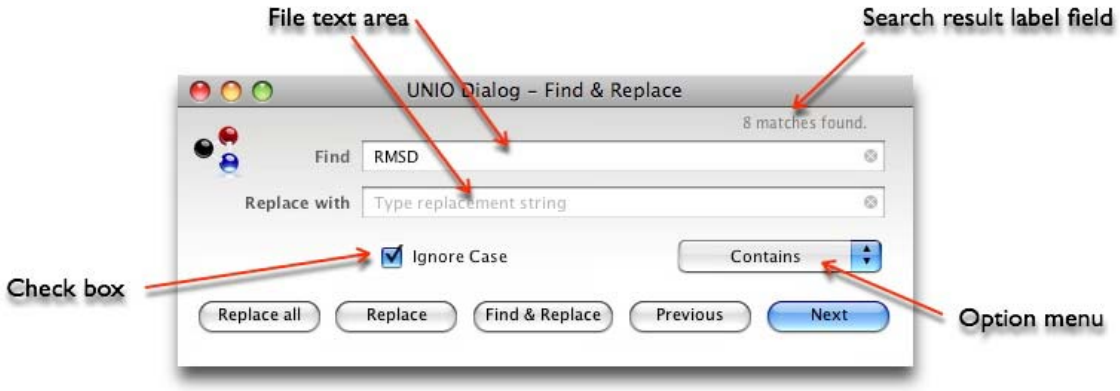
UNIO File Editor window while typing a search string in the Quick search field.

If you leave the *Quick search* field with your mouse pointer, the *Search result view area* disappears and the *File view and edit area* is active again. The *Search result browse buttons* allow you to navigate through the matching lines found. The matching lines are highlighted in yellow.



UNIO File Editor window while browsing through the pattern matches found.

A more advanced text search can be performed by using the UNIO Find&Replace window. The bottom *Toolbar icon* "Binocular" of the UNIO File Editor window launches this sub-window. Besides the above-mentioned wild cards for the pattern filter, the text search can be further controlled by the *Check box* to enable/disable case sensitive pattern matching, and by the Option menu which allows to choose between the following matching options: "Contains", "Beginning of the word" and "Full word". The UNIO Find&Replace window also allows you to replace matching pattern by a given string.



UNIO Find&Replace window.

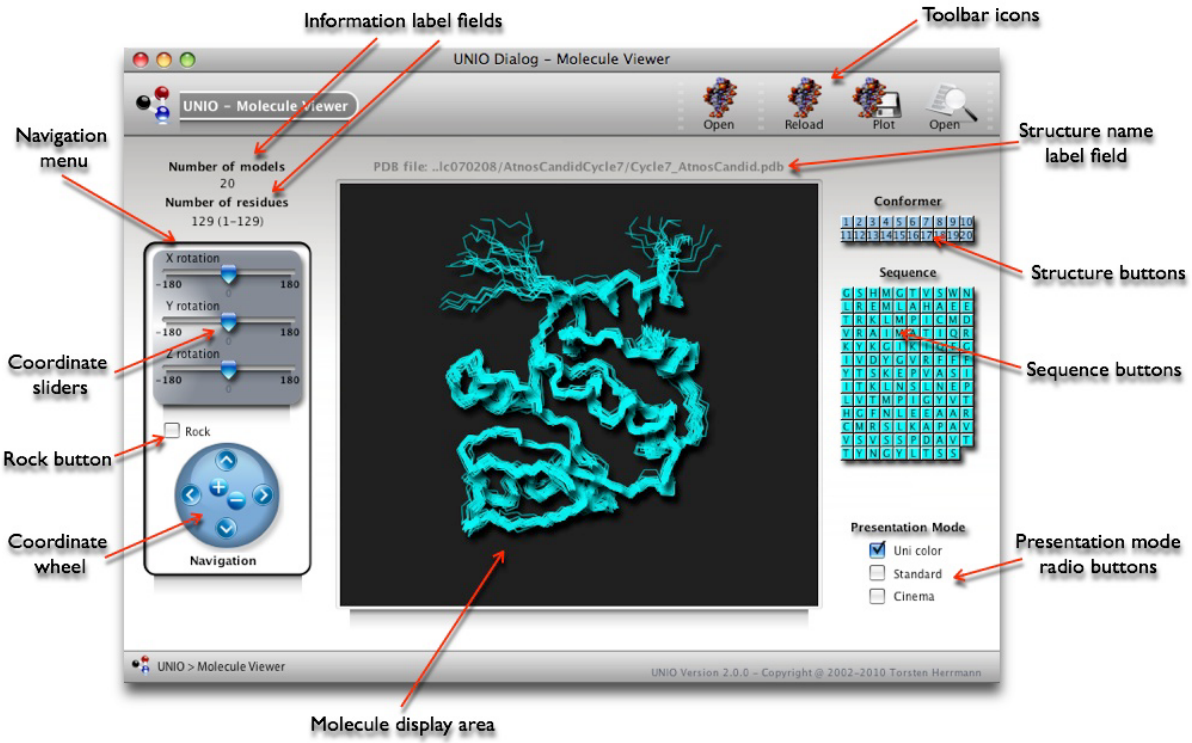
Similar to the *Quick search* field, while you are typing a search string in the *File text area* "Find" of the UNIO Find&Replace window, matching lines are immediately displayed in the *Search result view area* of the UNIO File Editor window.

Tip: If you are searching for individual restraints in a distance restraint file created as output from an ATNOS/CANDID calculation, then the UNIO Restraint Inspector offers you better search and display options than the UNIO File Editor.

2.6 UNIO - Molecule Viewer window

The UNIO Molecule Viewer allows you to display a protein structure within the UNIO environment. The coordinate file of the protein must be in Protein Data Bank (PDB) file format.

The top *Toolbar icons* allow you to open and reload a PDB file, to plot the protein structure as visualized in the *Molecule display area* in PNG file format with black-colored or transparent background, and to open the PDB file with the UNIO File Editor. The path of the PDB file is displayed in the *Structure name label field*. The number of models loaded and the number of the amino acids in the protein sequence are displayed in the *Information label fields*.

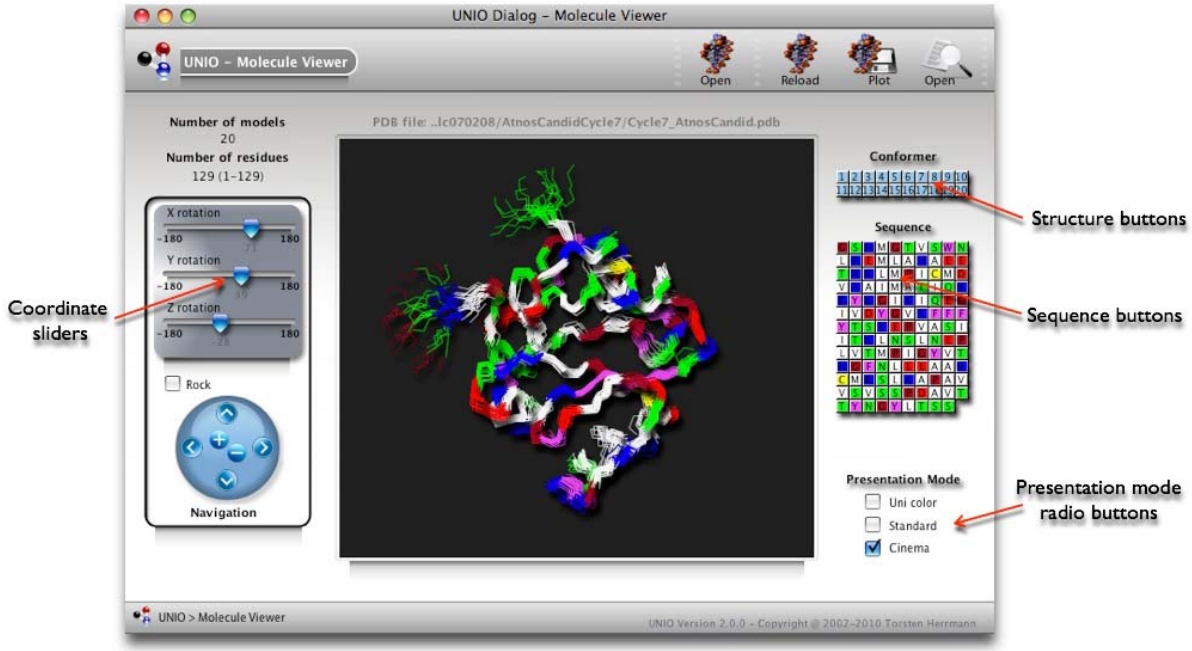


UNIO Molecule Viewer window after loading a protein structure.

For navigation of the protein structure, you can either use the *Navigation menu* or the left and the right mouse buttons. In the *Navigation menu*, you can either use the *Coordinate sliders* or the *Coordinate wheel* to change the perspective (view angle and zoom) of the protein structure. The *Rock button* allows you to animate the protein visualization. If you move your mouse pointer into the *Molecule display area*, then you can use the left mouse button (by keeping it press down) for rotation and the right mouse button for translation of the molecule. In addition, if you move with the mouse pointer over a residue in the structure, then the corresponding residue name and numbering is displayed at the bottom of the *Molecule display area*.

Individual models (conformers) of the protein can be selected and deselected by using the *Structure buttons*. Similar individual amino acid residues of the protein can be selected or deselected by using the *Sequence buttons*. A left mouse button click of a Sequence button triggers the selection/deselection mode. The right mouse button can be used to change the color of the respective residue.

The *Presentation mode radio buttons* allow you to choose between different representation modes for the protein structure. The radio button "Uni color" can be used to choose a uniform color of the structure bundle. The radio buttons "Standard" and "Cinema" performs an auto-coloring of the protein according to classification of amino acids (hydrophobic, polar, negatively charged, positively charged etc.) If you move with the mouse pointer over the word "Standard" or "Cinema" next to the *Presentation mode radio buttons*, then you can see the definition of the respective presentation mode displayed as tool tip.



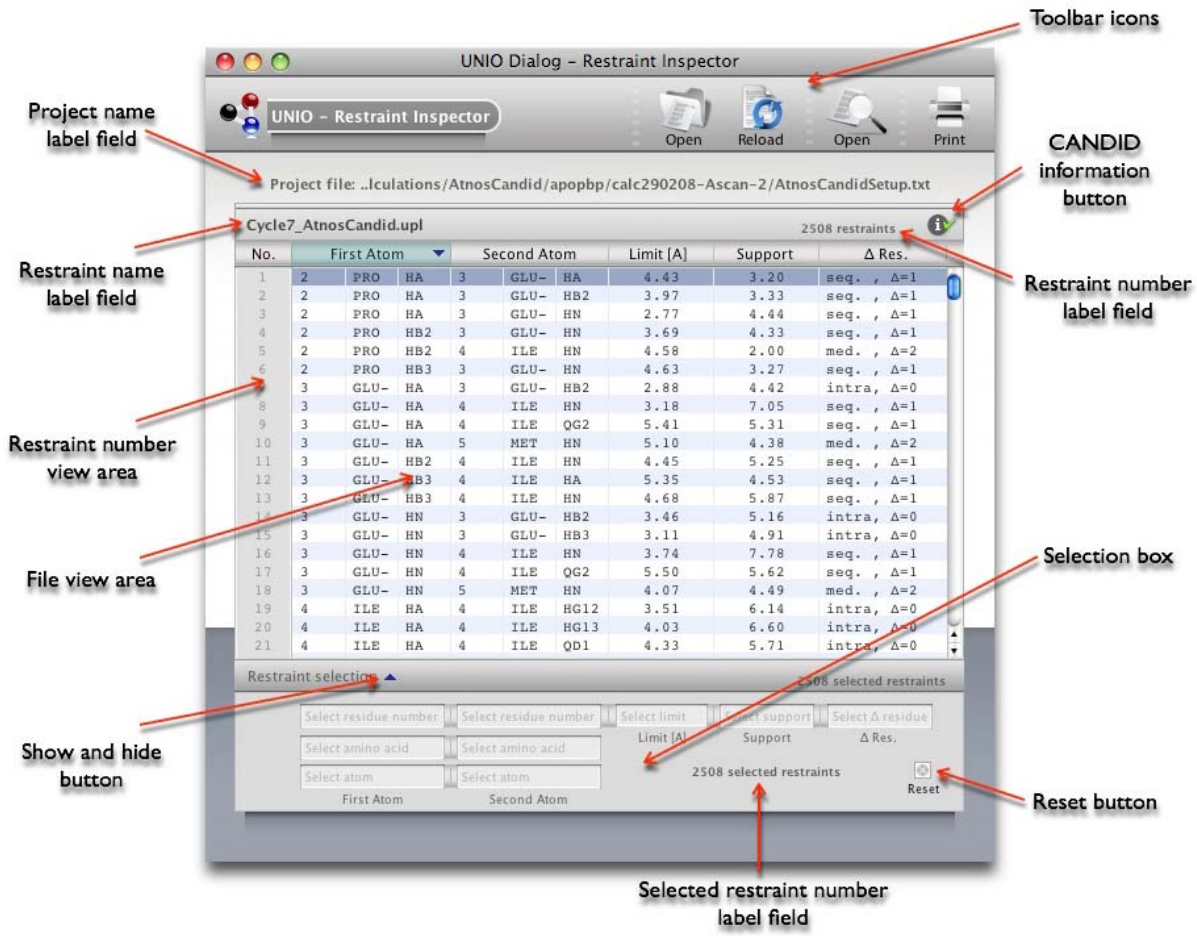
UNIO Molecule Viewer window. Protein presentation mode "Cinema" selected.

2.7 UNIO - Restraint Inspector window

The UNIO Restraint Inspector allows you to inspect the final distance restraint file from an ATNOS/CANDID calculation within the UNIO environment.

The top *Toolbar icons* allow you to open and reload the UNIO-ATNOS/CANDID configuration file, to load the distance restraint file into the UNIO File Editor and to print the restraints displayed in the *File view area*. The path of the ATNOS/CANDID configuration file is displayed in the *Project name label field*, and the name of the final ATNOS/CANDID distance restraint file is given in the *Restraint name label field*. The number of distance restraints loaded is displayed in the *Restraint number label field*. The *Restraint number view area* indicates the number range of the currently displayed part of the restraint file.

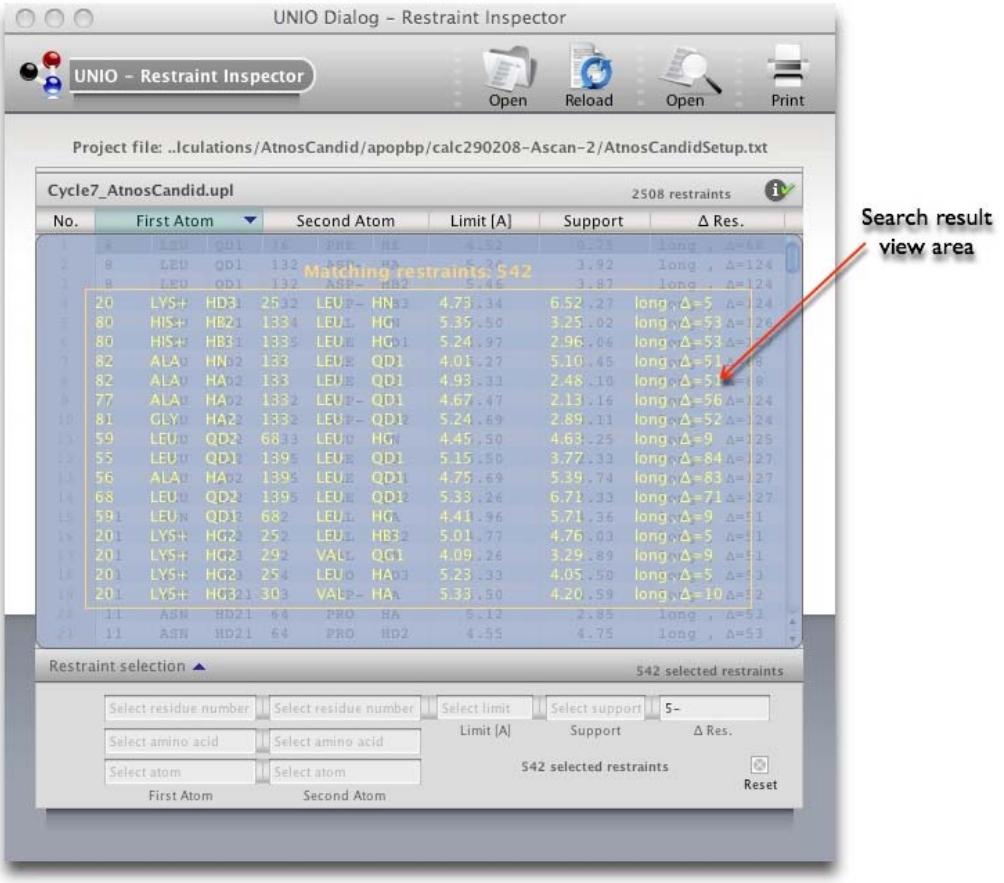
The restraints displayed in the File view area can be sorted in ascending or descending order by mouse click on the First Atom column, Second Atom column, Limit column, Support column or Δ Res. (residue difference) column header.



UNIO Restraint Inspector window.

The *Selection box* serves to perform a search for individual distance restraints that match given search criteria. While typing a search pattern, the matching restraints are

immediately displayed in the *Search result view area*. This feature allows you to narrow your search by refining the search pattern against the intermediate matching lines listed. If you leave the *Text field* with your mouse pointer, the *Search result view area* disappears and the *File view area* is active again. The number of selected distance restraints is shown in the *Selected restraint number label fiel*d. The *Reset button* can be used to reset all previous selections. The *Selection box* can be hidden and shown using the *Show and hide button*.



UNIO File Editor window while typing a search string in a field of the *Selection box*.

The UNIO Restraint Inspector can also be used to learn about the CANDID assignment process that led to a given distance restraint. Use a single mouse click to highlight a single restraint. Then, double mouse click on the highlighted restraint will launch the UNIO CANDID Assignment window (see below). The CANDID Assignment window lists all NOE cross peaks that have been assigned to a given distance restraint. For a given peak, CANDID characteristic parameters such as closeness of fit between the peak position and the chemical shift values of the assigned atom pair, average distance of the atom pair in the preceding ATNOS/CANDID iteration, network-anchored support, residual support etc.

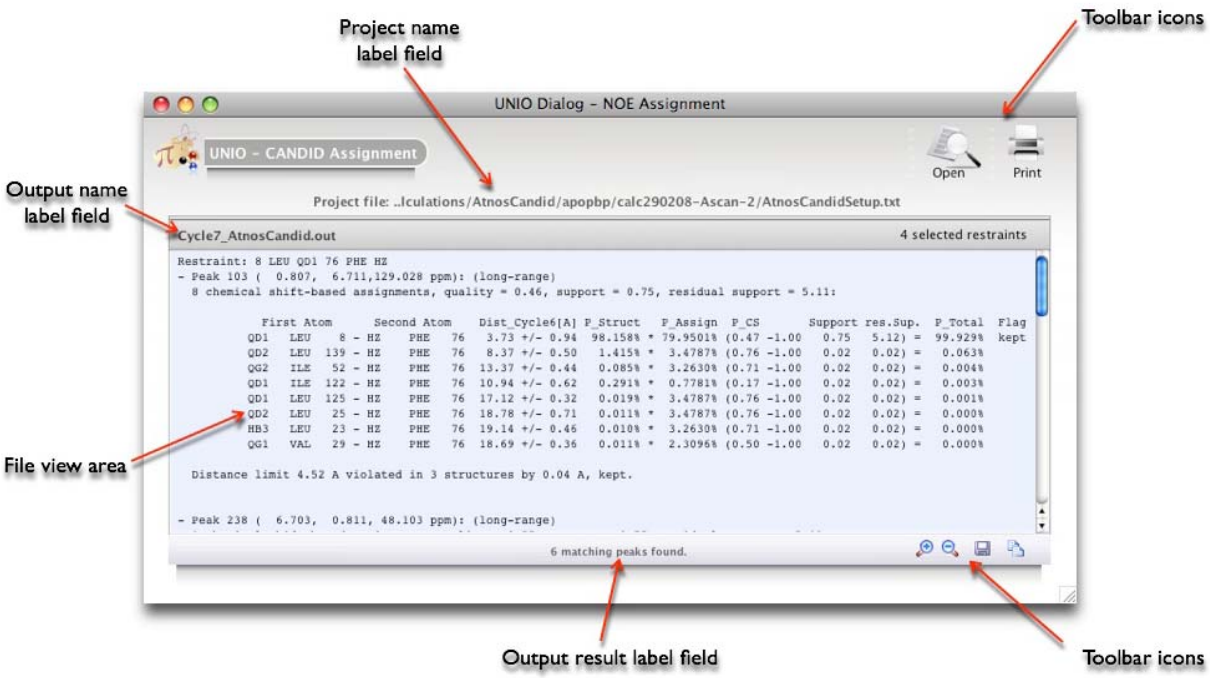
If you want to highlight several distance restraints at once, then use the command key plus mouse click to highlight multiple distance restraints. Use the *Information button* to see the corresponding CANDID output for all highlighted distance restraints.

2.8 UNIO - CANDID Assignment window

The UNIO CANDID Assignment window allows you to learn about the NOE assignment process performed by the UNIO-CANDID routine that led to a given distance restraint within the UNIO environment. The UNIO CANDID Assignment window can exclusively be launched out of the UNIO Restraint Inspector window (see above).

The top *Toolbar icons* allow you to open the ATNOS/CANDID cycle 7 output file with the UNIO File Editor and to print the text displayed in the *File view area*. The path of the ATNOS/CANDID configuration file is displayed in the *Project name label field*, and the name of the final ATNOS/CANDID output file is given in the *Output name label field*. The number of NOE cross peak that can be assigned to a given atom pair resp. distance restraint as selected in the UNIO Restraint Inspector window is shown in the *Output result label field*.

The bottom *Toolbar icons* allow you to increase and decrease the font size used in the *File view area*, to save the text displayed in the *File view area* into a file, and to copy the text shown in the *File view area* to the clipboard.



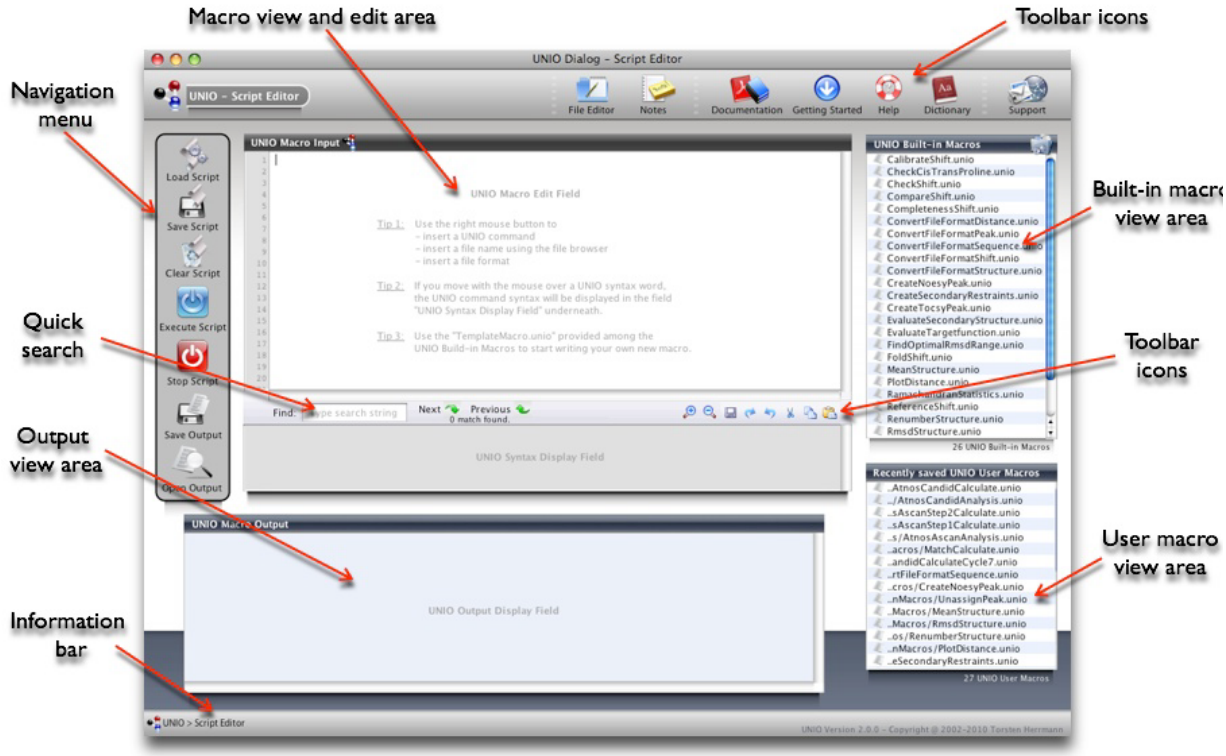
UNIO CANDID Assignment window.

The CANDID Assignment window lists for each matching peak found, CANDID characteristic parameters such as closeness of fit between the peak position and the chemical shift values of the assigned atom pair, average distance of the atom pair in the preceding ATNOS/CANDID iteration, network-anchored support, residual support etc. At the bottom of the text displayed, a CANDID NOE assignment legend is given that explains the meaning of the output shown in more detail.

2.9 UNIO - Script Editor window

The UNIO Script Editor allows you to write and execute your own UNIO macros within the UNIO environment. The *Navigation menu* allows to load a previous script or macro, to save a script, to clear the *Macro view and edit area*, to execute a script, to save the generated output displayed in the *Output view area* into a file and to open the output generated with the UNIO File Editor. The top *Toolbar icons* give you access to UNIO utility programs, such as the UNIO File Editor or the UNIO Notes window, or to open the UNIO documentations and the UNIO Syntax Dictionary window.

UNIO provides numerous built-in macros written for commonly performed tasks during protein NMR structure determination, such as file format conversion, distance restraint or Ramachandran analysis and similar. The *Built-in macro view area* provides a listing of all UNIO built-in macros or scripts. Similar, the *User macro view area* lists all user-created UNIO scripts and thus gives you fast access to previously written UNIO macros.

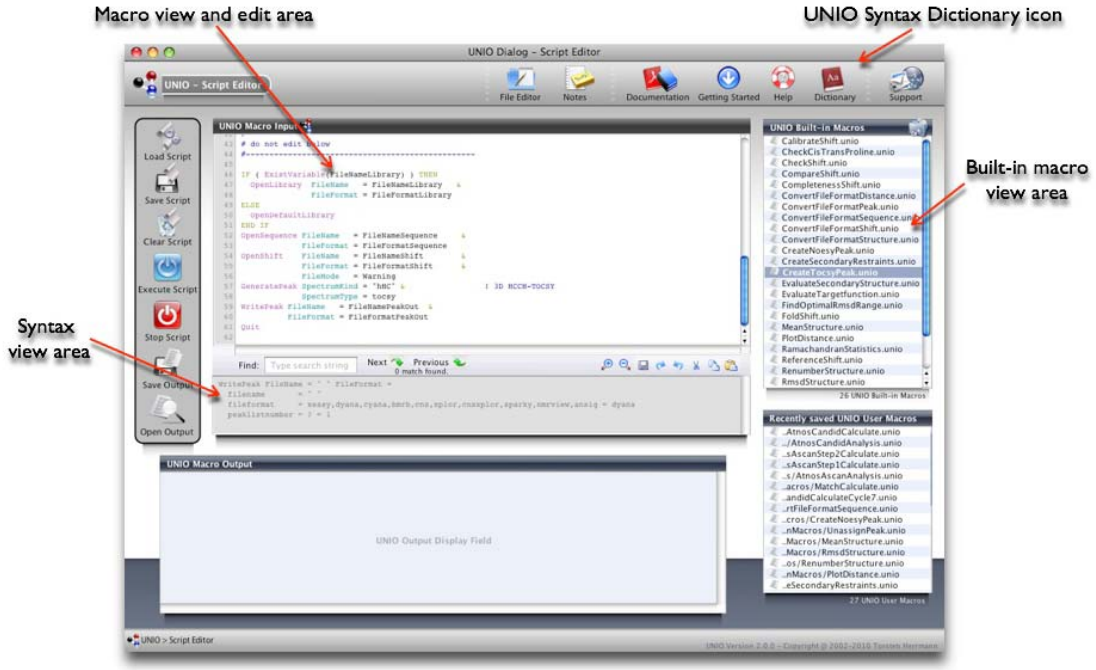


UNIO Script Editor window.

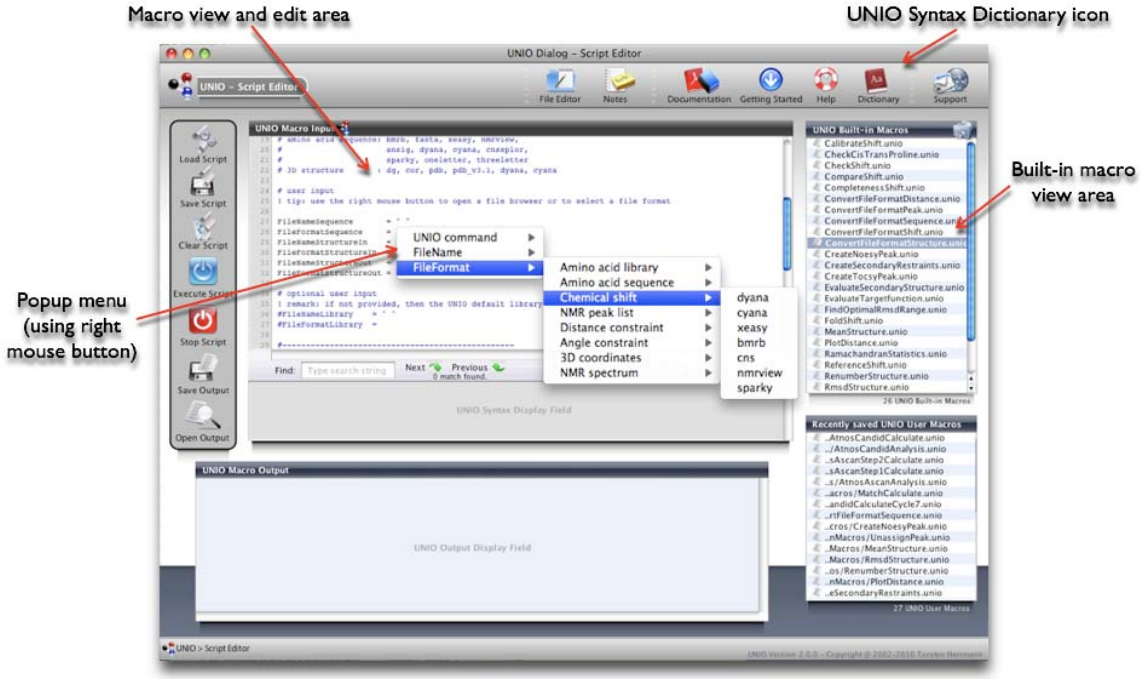
You can load a UNIO built-in script by using a mouse click on the corresponding macro name in the *Built-in macro view area*. By doing so, the highlighted script will be loaded and displayed in the Macro view and edit area (see next figure).

The header of each UNIO built-in macro contains a brief description of its task. Following below, the user needs to complete the input files that are required to execute the respective macro properly.

The *Macro view and edit area* got designed to display UNIO scripts in the most convenient way possible. To achieve this, the UNIO macro language is automatically color-coded according to the respective UNIO syntax type. You can use the *UNIO Syntax Dictionary icon* to learn all about the UNIO macro language and key words.



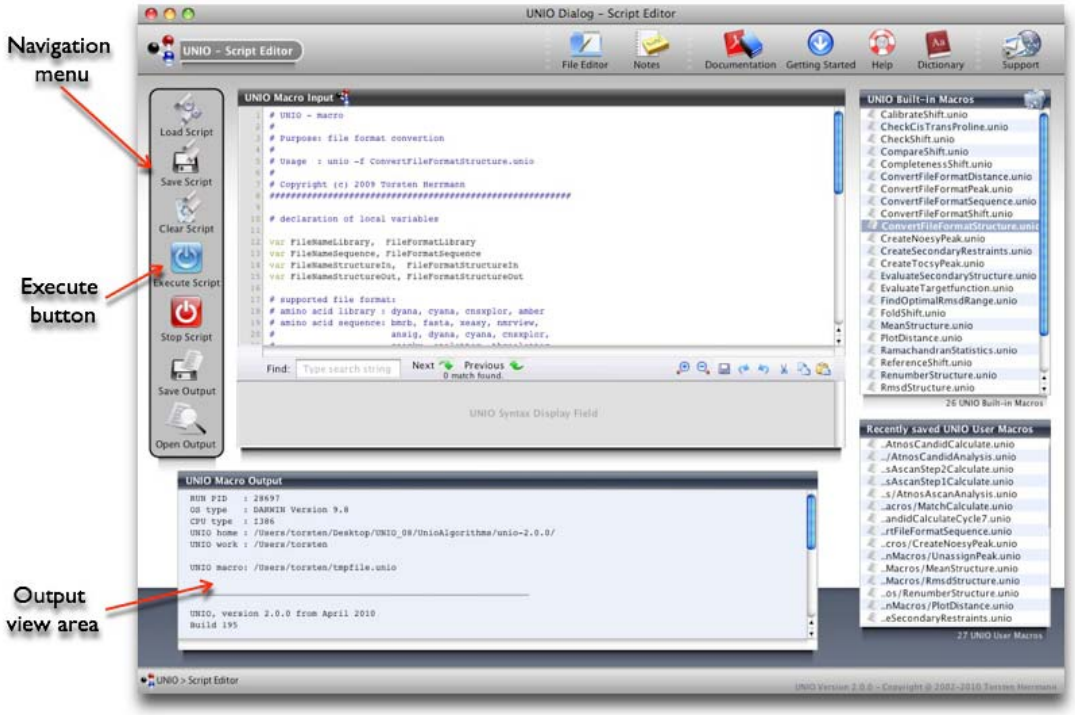
UNIO Script Editor window. After loading a UNIO built-in macro.



UNIO Script Editor window. After left mouse click in the *Macro view and edit area*.

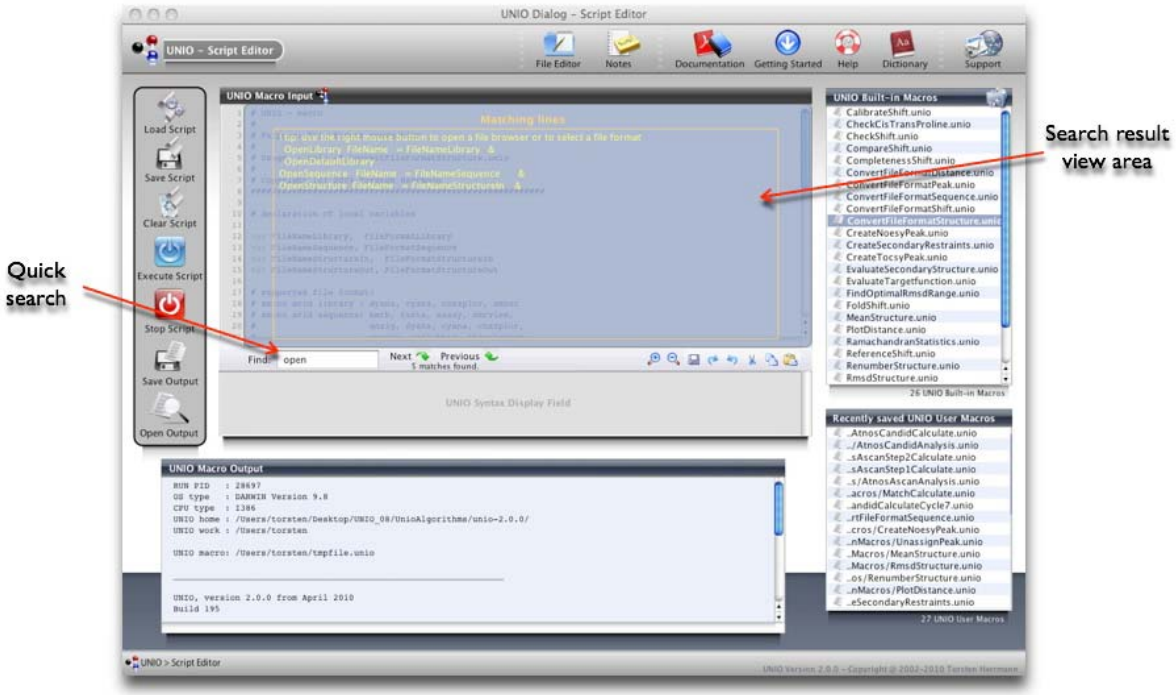
In order to facilitate the programming of your own UNIO scripts, two important features of the UNIO Script Editor have been designed. First, if you move with the mouse pointer over a word in the *Macro view and edit area*, the corresponding UNIO language syntax is displayed in the *Syntax view area*. Second, if you use the right mouse button in the Macro view and edit area, a popup menu is opened which allows you to insert a UNIO command or key word, to open the system file browser and to insert UNIO-supported file formats of files.

The *Execute button* in the *Navigation menu* allows you to evaluate your macro with UNIO. After UNIO has evaluated and executed your macro, the corresponding result is shown in the *Output view area*. You can then open the output with the UNIO File Editor or save it to a file by using the corresponding icons in the *Navigation menu*.



UNIO Script Editor window. After execution of a UNIO macro.

Similar to the *Quick search* field of the UNIO File Editor or the UNIO Restraint Inspector, the *Quick search* field of the UNIO Script Editor allows you to find string pattern present in the *Macro view and edit area*. While you are typing a search string in the *Quick Search* field, matching lines are immediately displayed in the *Search result view area*. This feature allows you to narrow your search by refining the search pattern against the intermediate matching lines listed.



UNIO Script Editor window while typing a search pattern in the *Quick search field*.

If you leave the *Quick search* field with your mouse pointer, the *Search result view area* disappears and the *Macro view and edit area* is active again. The *Search result browse buttons* next to the *Quick search* field allow you to navigate between the matching lines found. The matching lines are highlighted in yellow.

The *Toolbar icons* at the bottom of the *Macro view and edit area* allow you to increase and decrease the font size, to save the file, to redo and undo changes, to cut selected lines and to copy and paste selected text from and to the clipboard.

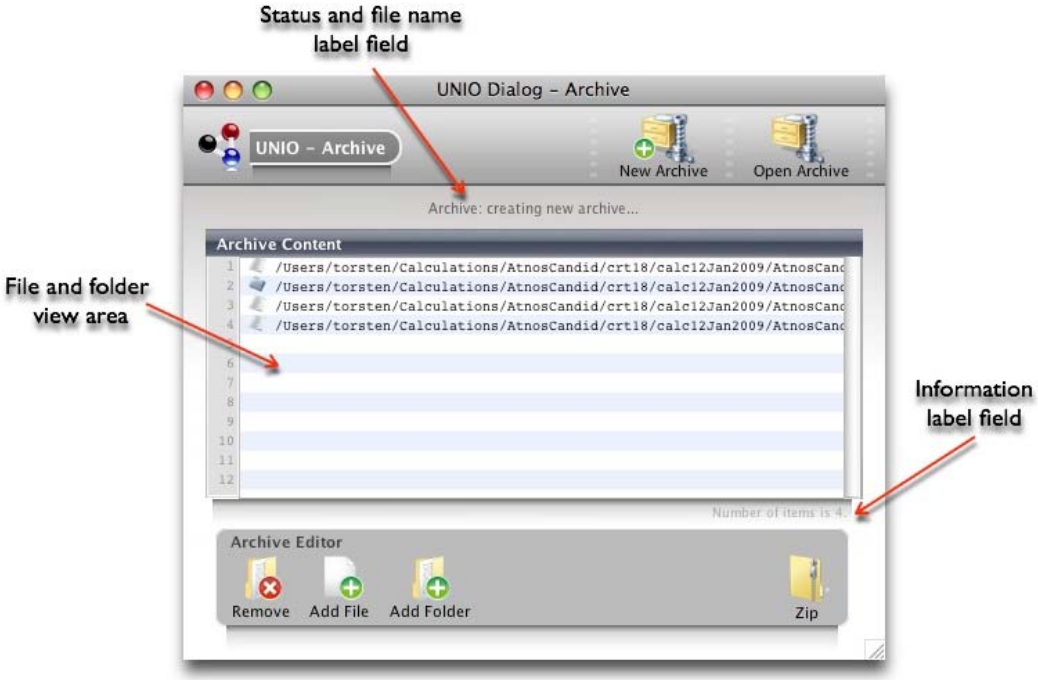
Tip: UNIO got designed to work with most commonly used programs for protein NMR structure determination. Therefore UNIO is highly compatible with a huge range of popular file formats, and consequently the UNIO Script Editor with its UNIO built-in macros represent an easily usable framework to convert between numerous file formats, such as protein sequence, chemical shifts, peak lists, distance restraints, 3D coordinates and similar.

2.10 UNIO - Archive window

The UNIO Archive window allows you to open and create ZIP-archives within the UNIO environment. The top *Toolbar icons* allow you create a new archive or to open an existing archive.

UNIO Archive window.

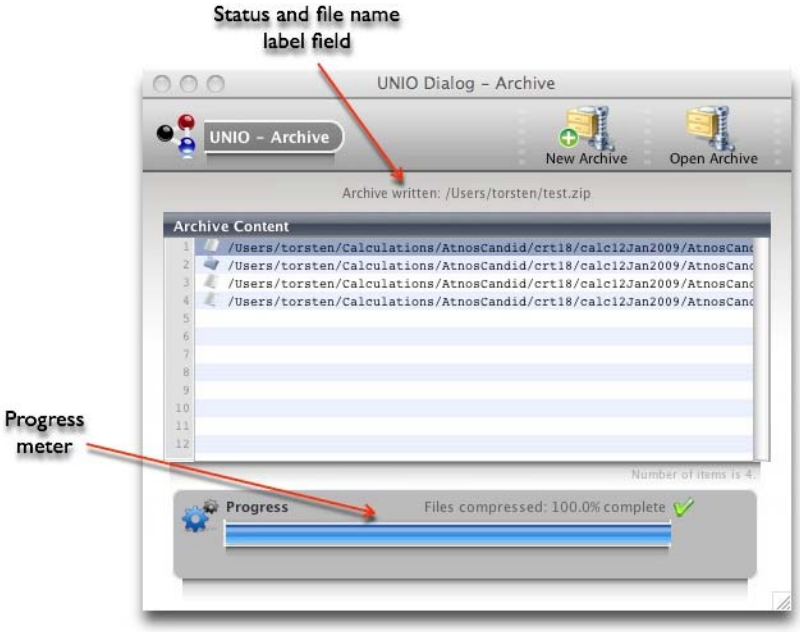
To create a new archive, use the the *New Archive* button in the *Toolbar icons* and then press the *Add File* or *Add Folder* button in the *Archive editor area*. Alternatively, you can



UNIO Archive window while creating a new archive.

also simply drag and drop single or multiple files in the *File and folder view area*. The *Status and file name label* field reports at the current status of the archiving process. The number of added files and folder are displayed in the *Information label field*.

To trigger the archiving process, use the *Zip* button in the *Archive editor area*. The path of the zip-file is then displayed in the *Status and file name label field*. The *Progress meter* reports about the percentage of compressed files and folders during archiving.



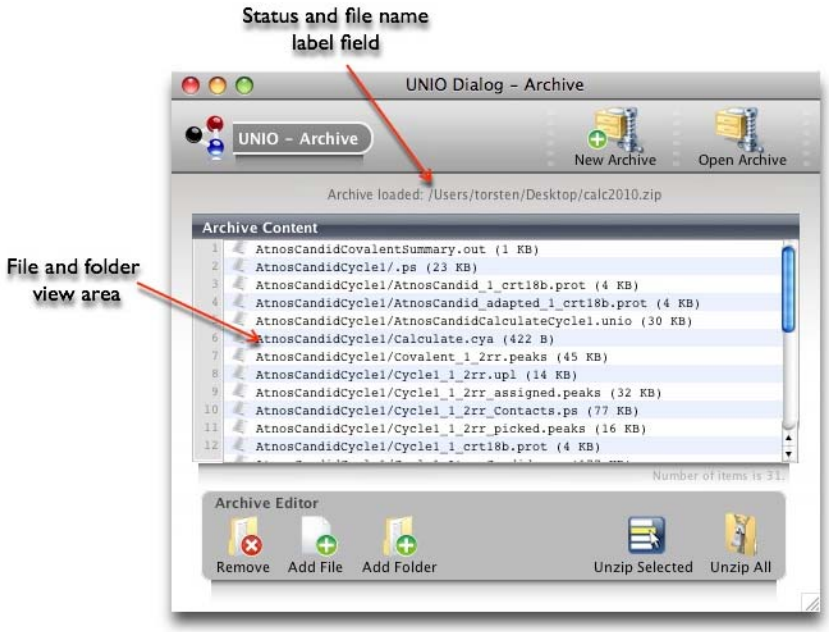
UNIO Archive window after successful archiving.

Tip: After opening an existing archive, you have several possibilities how to work with the archive (see next figure):

First, you can uncompress the entire archive using the *Unzip All* button in the Archive editor area.

Second, you can select single or multiple files in the File and folder view area (highlighting via mouse click) and then only uncompress these selected files and folders by using the *Unzip Selected* button in the Archive editor area.

Third, you can directly edit the archive, i.e. add and delete files and folders without the need to uncompress the entire archive first.



UNIO Archive window after opening an existing archiving.

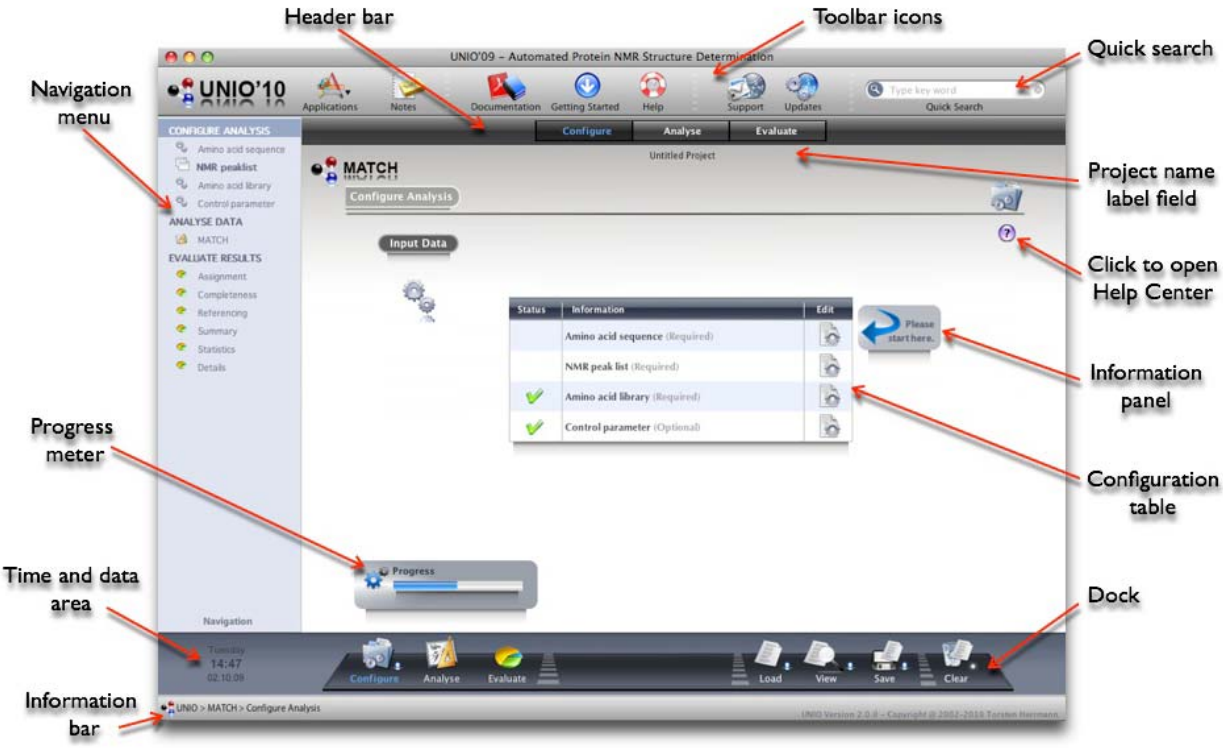
Caution: The UNIO Archive window requires operating system-supported compress and uncompress utilities. The UNIO Archive works in all cases under Mac OSX. However on many Linux systems the system-intern zip-utilities are not existing, and consequently the UNIO Archive system won't work.

2.11 UNIO - Application window for MATCH

The UNIO MATCH window allows you to perform automated sequence-specific backbone resonance assignment. The MATCH environment is split into three main sections "Configure Analysis", "Analyse NMR Data" and "Evaluate Results". The *Toolbar icons* give you fast access to frequently used functions of UNIO. The *Header bar* allows you to navigate between the three main sections of MATCH. The *Navigation menu* gives you fast access to all sub-sections of the three main MATCH sections. The *Dock* allows you to navigate between the three main sections and provides functionalities specific for one section.

2.11.1 Section "Configure Analysis"

The MATCH section "Configure Analysis" is meant to collect all input data required to perform MATCH backbone assignment. The *Configuration table* lists all four input components, launches the corresponding input windows and reports along with the *Progress meter* about the completeness of the input configuration. You can click on a row of the *Configuration table* to open the corresponding input window. A description of the individual windows and the required input data are given later in this User's Guide.



UNIO Application window for MATCH, section "Configure Analysis".

Toolbar icons

Applications	Opens an option menu for navigating between the UNIO NMR data analysis routines of MATCH, ATNOS/ASCAN or ATNOS/CANDID.
Notes	Opens the UNIO Notes window.
Documentation	Opens this User's Guide.
Getting Started	Opens the Getting Started documentation for MATCH.
Help	Opens the UNIO Help Center.
Support	Opens the system's default mail program.
Updates	Opens the UNIO website.
Quick Search	Opens the UNIO Help Center with the typed key word.

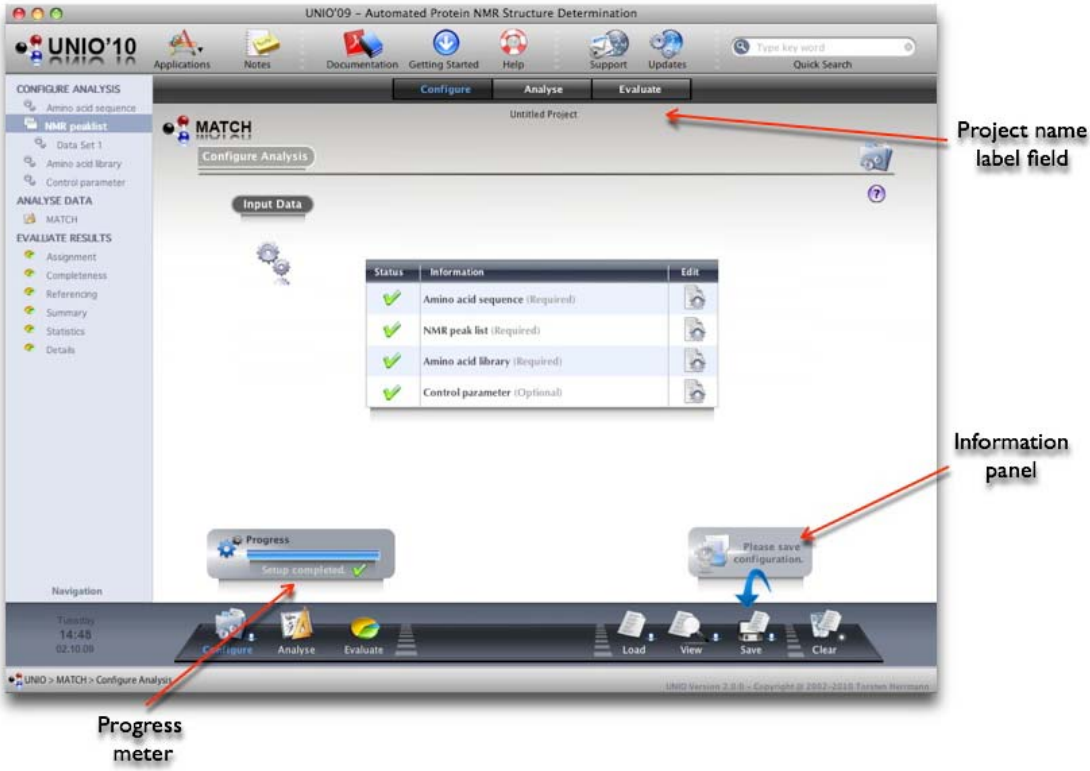
Header bar

Configure	Opens the MATCH main section "Configure Analysis".
Analyse	Opens the MATCH main section "Analyse NMR Data".
Evaluate	Opens the MATCH main section "Evaluate Results".

Dock

Configure	Opens the MATCH main section "Configure Analysis".
Analyse	Opens the MATCH main section "Analyse NMR Data".
Evaluate	Opens the MATCH main section "Evaluate Results".
Load	Opens the OS file dialog to load a MATCH configuration file.
View	Opens the UNIO Summary window that reports about all input data.
Save	Opens the OS file dialog to save a MATCH configuration file.
Clear	Clears all input configuration.

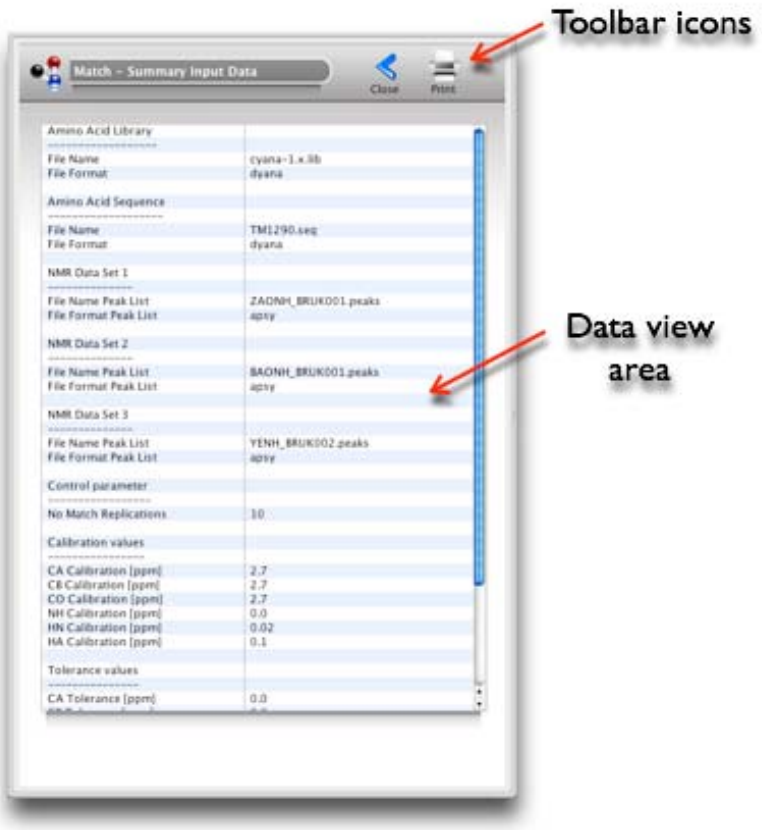
After you have provided all required input data, you first need to save your input data as MATCH configuration file, conventionally called "MatchSetup.txt". The directory containing your MATCH configuration file will be in the following considered as your project directory and all MATCH output data will be written in this directory. The name of this directory is displayed in the *Project name label field* underneath the *Header bar*.



UNIO Application window for MATCH, section "Configure Analysis".

Tip: The information about the required input data and file format is provided with the description of the respective input window (see later).

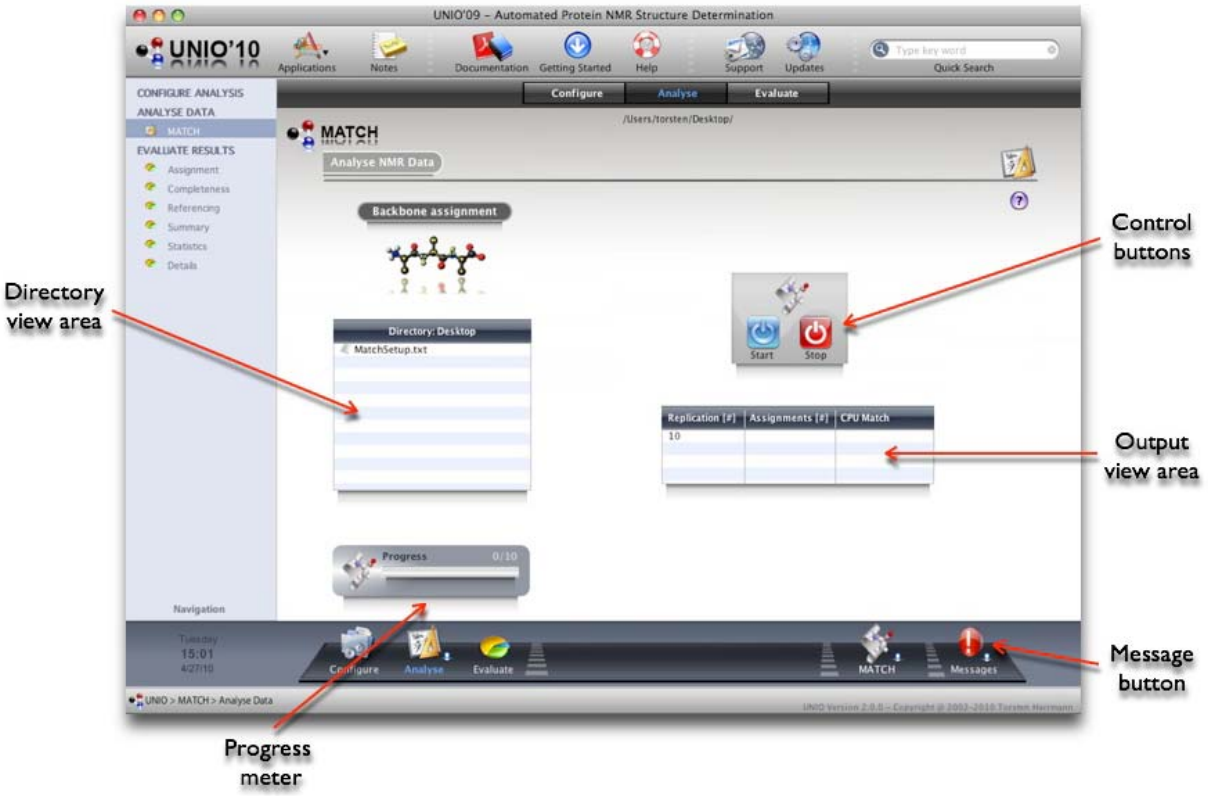
The *View button* in the *Dock* opens the UNIO Summary window which gives you an overview of all input data provided. The *Toolbar icons* allows you to close this window and to print the displayed text in the *Data view area* which corresponds to the MATCH configuration file.



UNIO Input Summary window for MATCH, section "Configure Analysis".

2.11.2 Section "Analyse NMR Data"

The MATCH section "Analyse NMR Data" represents the core of MATCH. By clicking on the *Control button* "Start", you initiate automated MATCH backbone assignment. During the NMR data analysis, the *Progress meter* reports about the state of the calculation, similar the *Directory view area* is getting updated frequently to report about MATCH output. At the end of MATCH data analysis, a brief summary about the assignment completeness and the CPU usage is given in the *Output view area*.

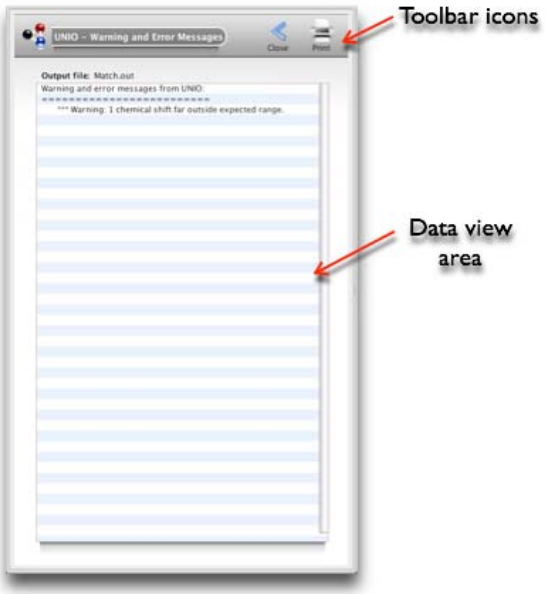


UNIO Application window for MATCH, section "Analyse NMR Data".

Tip: Single mouse click selects a file or lists a directory in the *Directory view area*. Double mouse click opens a file selected in the UNIO File Editor.

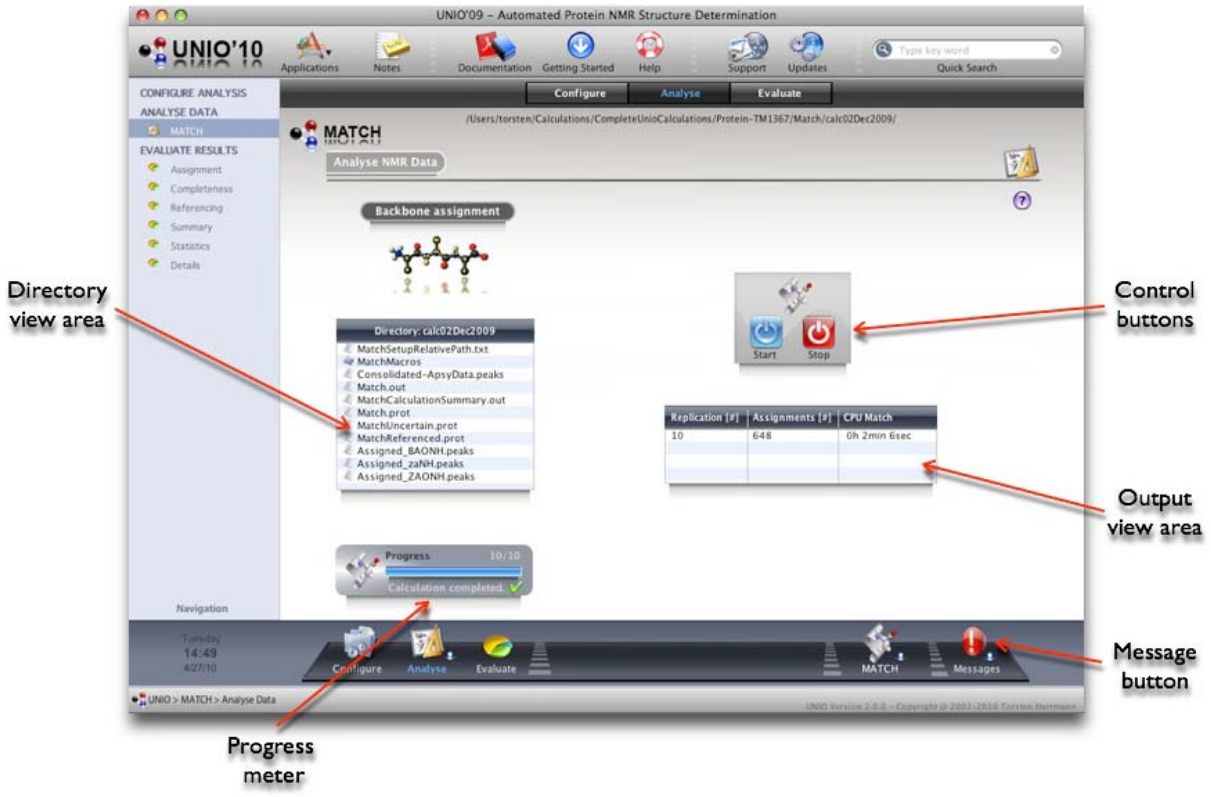
The *Message button* in the *Dock* opens the UNIO Messages window. This window displays all warning and error messages in all output files generated by UNIO.

Tip: The *Message button* might be useful to find potential problems in the input data and to correct your configuration file if necessary.



UNIO Warning and Error Messages for MATCH, section "Analyse NMR data".

After a MATCH calculation is completed, you can browse through all output files by mouse click in the *Directory view area*. The *Output view area* reports about the number of assigned backbone atoms and the CPU time spent. For detailed inspection of the results obtained, you can now proceed to the MATCH section "Evaluate Results".



UNIO Application window for MATCH, section "Analyse NMR Data". MATCH calculation completed.

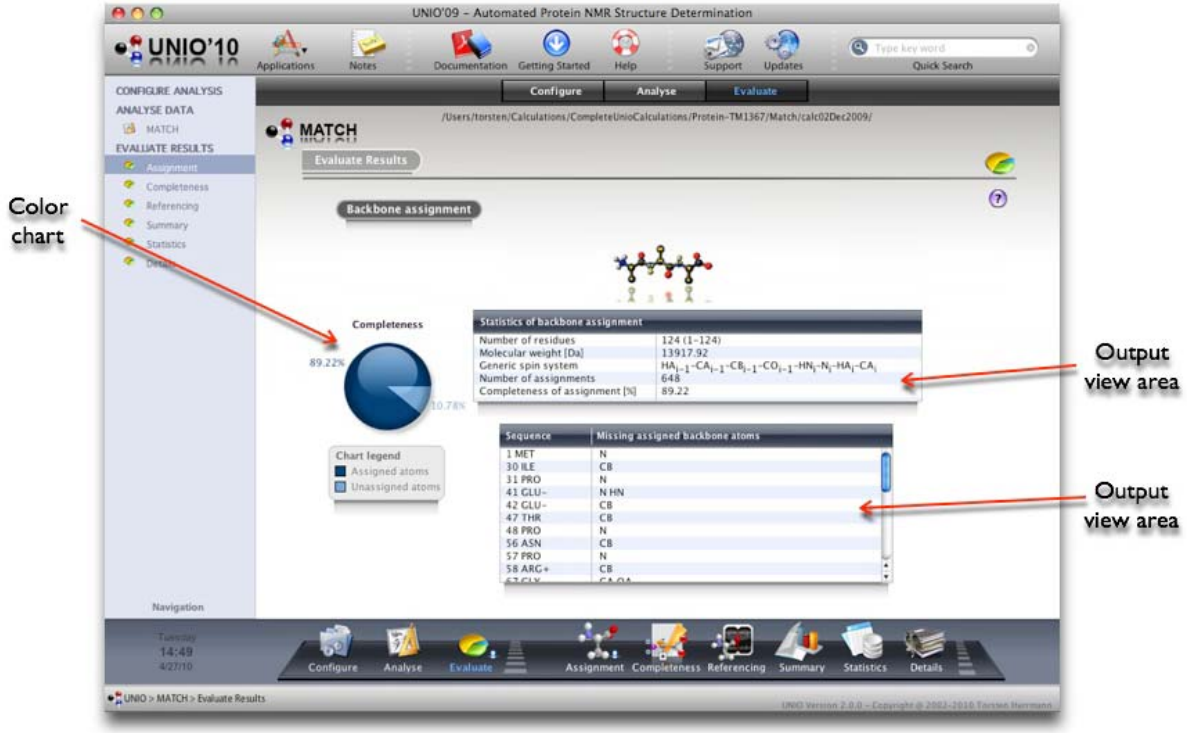
2.11.3 Section "Evaluate Results"

The MATCH section "Evaluate Results" consists of six sub-sections "Backbone Assignment", "Completeness", "Referencing", "Summary", "Statistics" and "Details". The *Navigation menu* or the *Dock* allows you to browse through these sections.

2.11.3.1 "Evaluate Results: Backbone Assignment"

The sub-section "Backbone Assignment" gives a brief overview of the results obtained.

The top *Output view area* provides a summary of characteristic MATCH parameters (see table below). The bottom *Output view area* lists all unassigned backbone atoms. The *Color chart* gives a quick report about the assignment completeness achieved.



UNIO Application window for MATCH, section "Evaluate Results: Backbone Assignment".

Table for sub-section "Backbone Assignment"

Number of residues	Gives the number of amino acid residues in the protein sequence.
Generic spin system	Shows the atom correlation within a tripeptide that got used for sequence-specific resonance assignment.
Number of assignments	Total number of assigned atoms in the protein sequence.
Completeness of assignments	Percentage of assigned atoms in the protein sequence.

2.11.3.2 "Evaluate Results: Completeness"

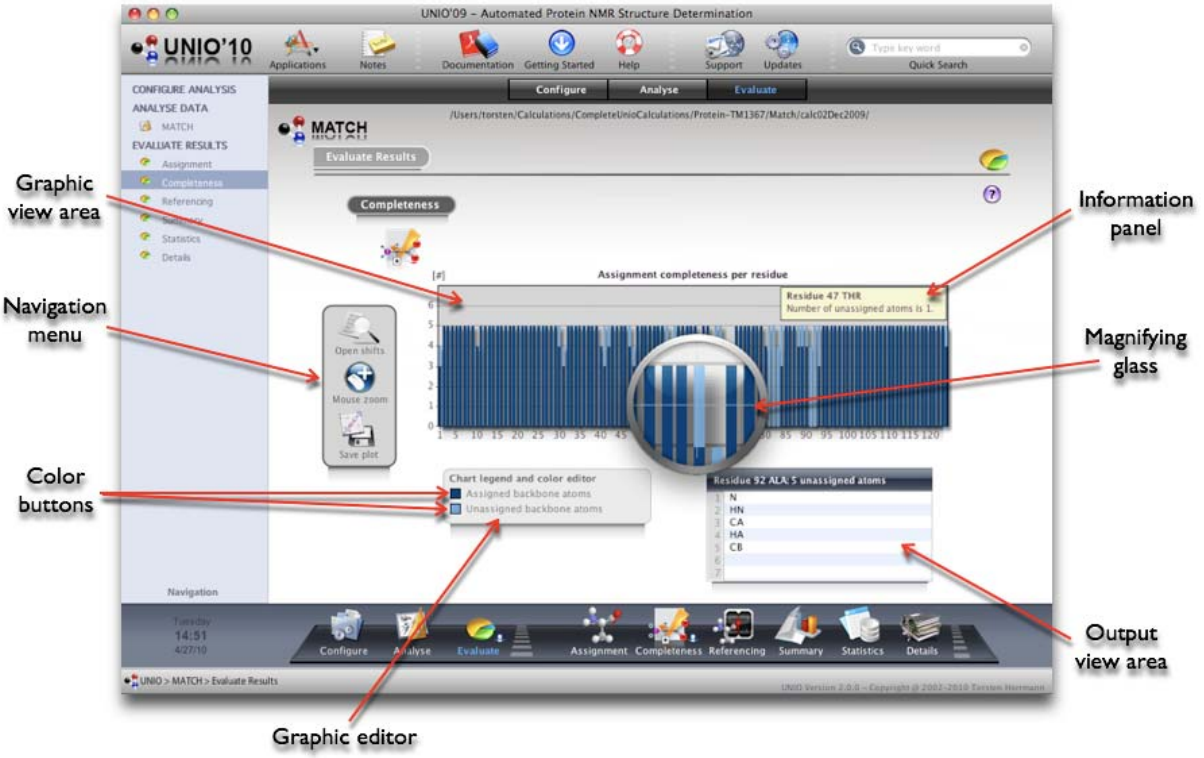
The sub-section "Completeness" informs about the assignment completeness as a function of the amino acid sequence. The *Graphic view area* allows the user to directly interact with the MATCH results.

To facilitate the navigation in the *Graphic view area*, a *Magnifying glass* is opened automatically as long the mouse pointer is within the *Graphic view area*. If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding residue name and the number of assigned and unassigned atoms are displayed in the *Information panel*.

If you click with the mouse on a chart bar in the *Graphic view area*, the assigned resp. unassigned atoms are listed in the *Output view area*.

The *Graphic editor* allows you to change the color-presentation of the chart. A mouse click on the *Color buttons* will launch the operating systems color palette window.

The *Navigation menu* serves three purposes: First, you can open the chemical shift list of the assigned backbone atoms with the UNIO File Editor; second, you can change the zoom factor used for the *Magnifying glass*; and third, you can save the residue chart in PNG file format with grey or transparent background.



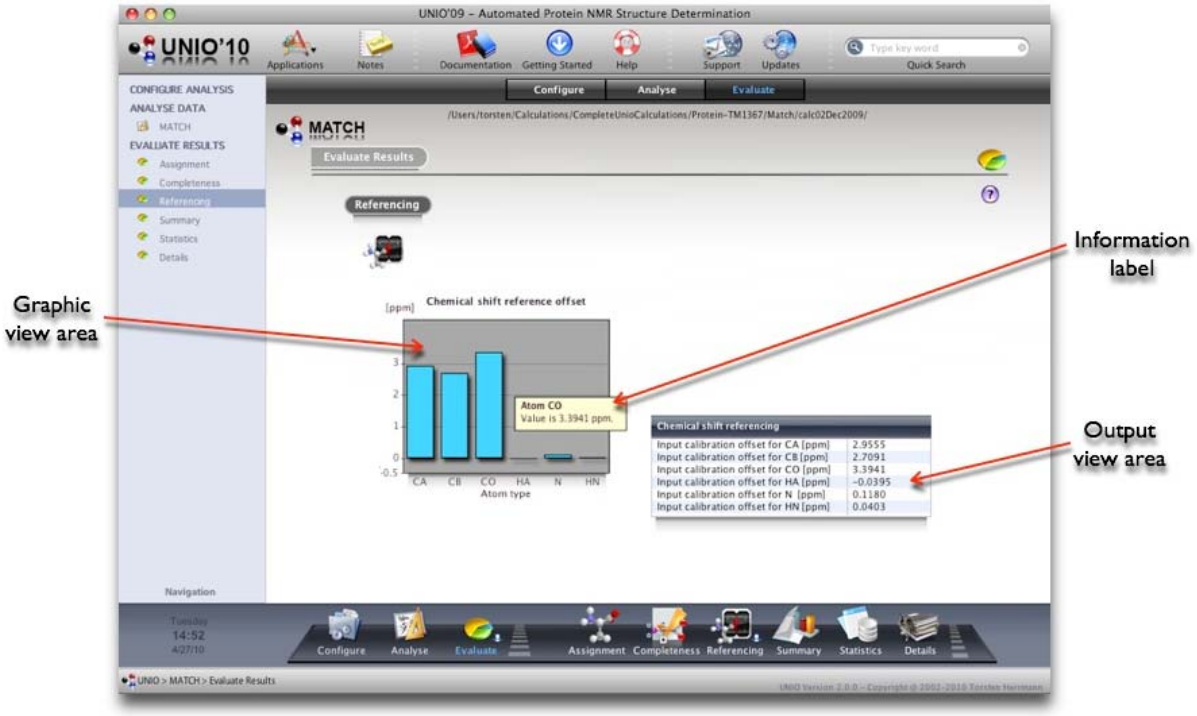
UNIO Application window for MATCH, section "Evaluate Results: Completeness".

2.11.3.3 "Evaluate Results: Referencing"

The sub-section "Referencing" informs about the chemical shift referencing offset of all backbone atoms for which correlations have been provided in the input configuration. The *Graphic view area* allows the user to directly visualize the respective referencing offset.

If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding atom name and the value of the reference offset in ppm are displayed in the *Information panel*.

The Output view area summarizes the chart information in table presentation.



UNIO Application window for MATCH, section "Evaluate Results: Referencing".

Important: The UNIO-MATCH algorithm performs an automated chemical shift referencing prior and at the end of the overall MATCH assignment process. Although this automated chemical shift referencing is extremely efficient and correct, it is nonetheless recommended to prepare the input for a MATCH calculation in such a way that the detected chemical shift reference offsets for the individual atoms are small. Please keep in mind that the sequence-specific assignment process is strongly guided by comparison between your input chemical shift values of the peak positions and the average chemical shift value of corresponding atoms found for NMR active nuclei in proteins taken from the BMRB data base.

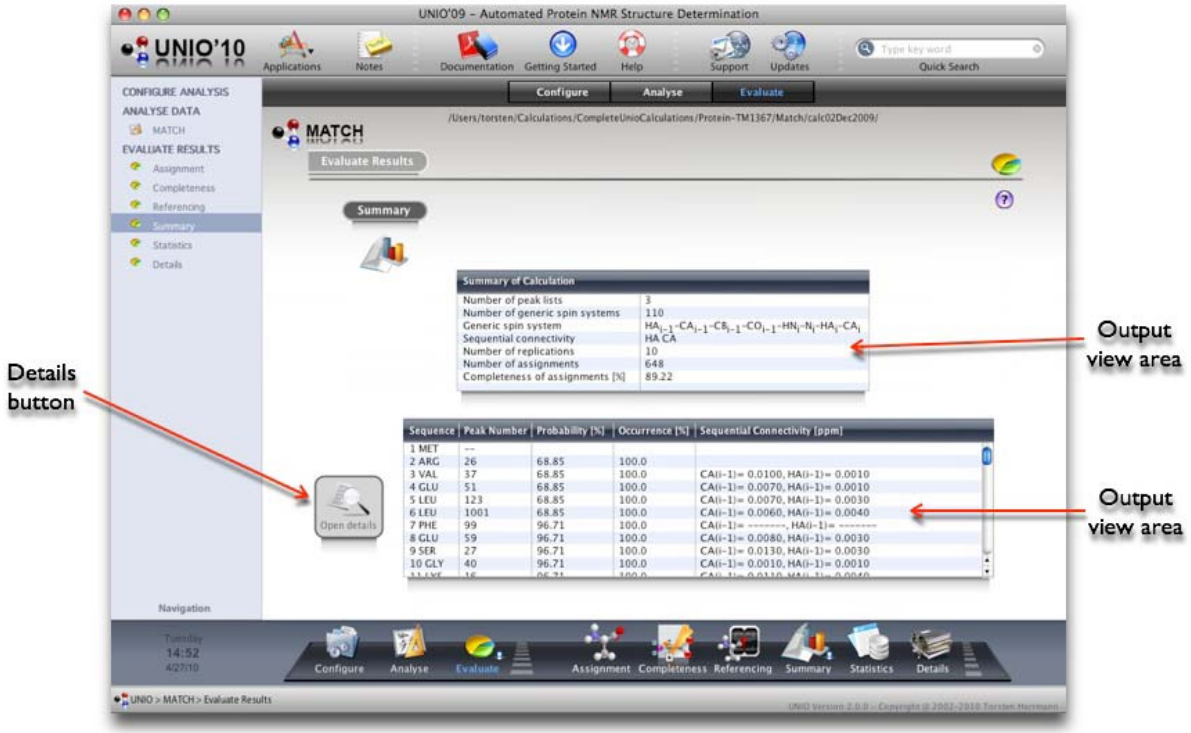
2.11.3.4 "Evaluate Results: Summary"

The sub-section "Summary" consists of two tables providing information about user-given input and MATCH output.

The upper table summarizes the main input and output information, such as the number of input peak lists used, the number of MATCH replications, the number of amino acid residues, the number of assignments and the assignment completeness.

The lower table reports about the resonance assignments made. For each assignment, its probability and the percentage of occurrence throughout the number of MATCH replications is given. The last column reports about the chemical shift agreement used for establishing sequential connectivities.

The *Details button* open the MATCH output with the UNIO File Editor and allows you to take a closer look at the MATCH assignment process.



UNIO Application window for MATCH, section "Evaluate Results: Summary".

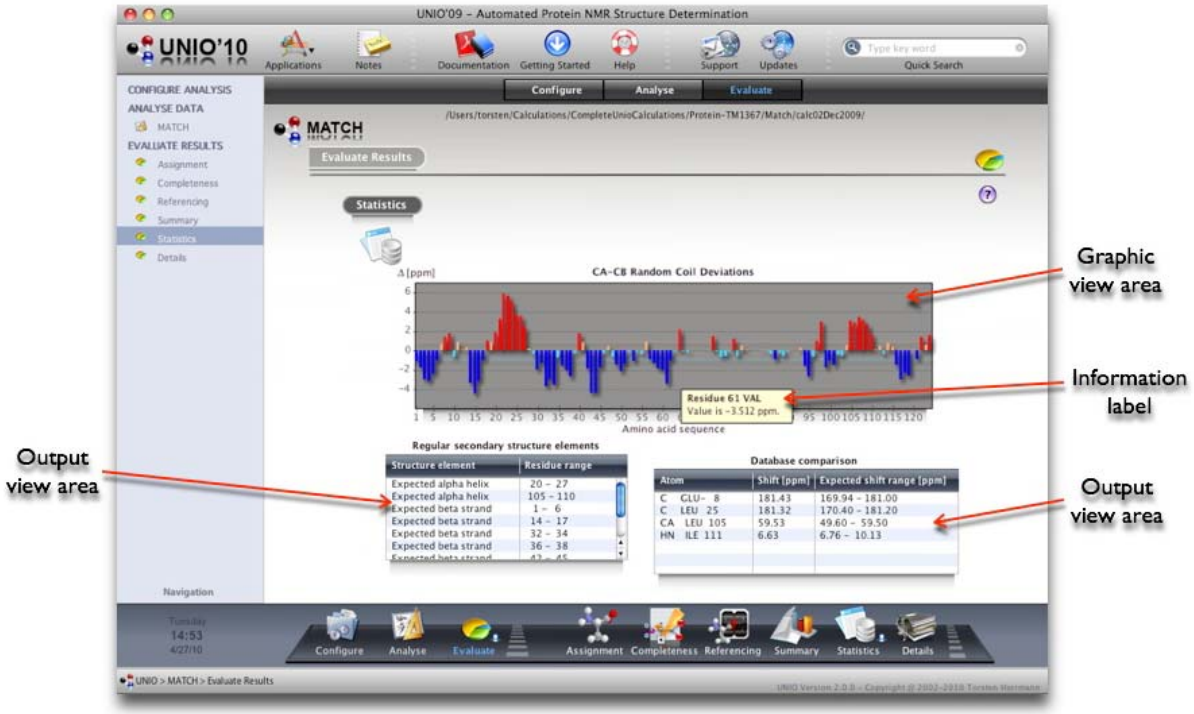
2.11.3.5 "Evaluate Results: Statistics"

The sub-section "Statistics" compares MATCH resonance assignments with average chemical shift values of proteins.

The *Graphic view area* plots the CA-CB random coil deviation along the protein sequence. If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding residue name and the value of the random coil deviation in ppm are displayed in the *Information panel*. Helical and strand secondary structure elements are color-coded in red and blue, respectively.

Consecutive amino acid residues with a chemical shift deviation above or below a certain threshold (orange and cyan lines in the *Graphic view area*) are listed as regular secondary structure elements in the table on the left side.

The table on the right side reports about chemical shift assignments that are unusual with respect to average chemical shift values found in proteins.



UNIO Application window for MATCH, section "Evaluate Results: Statistics".

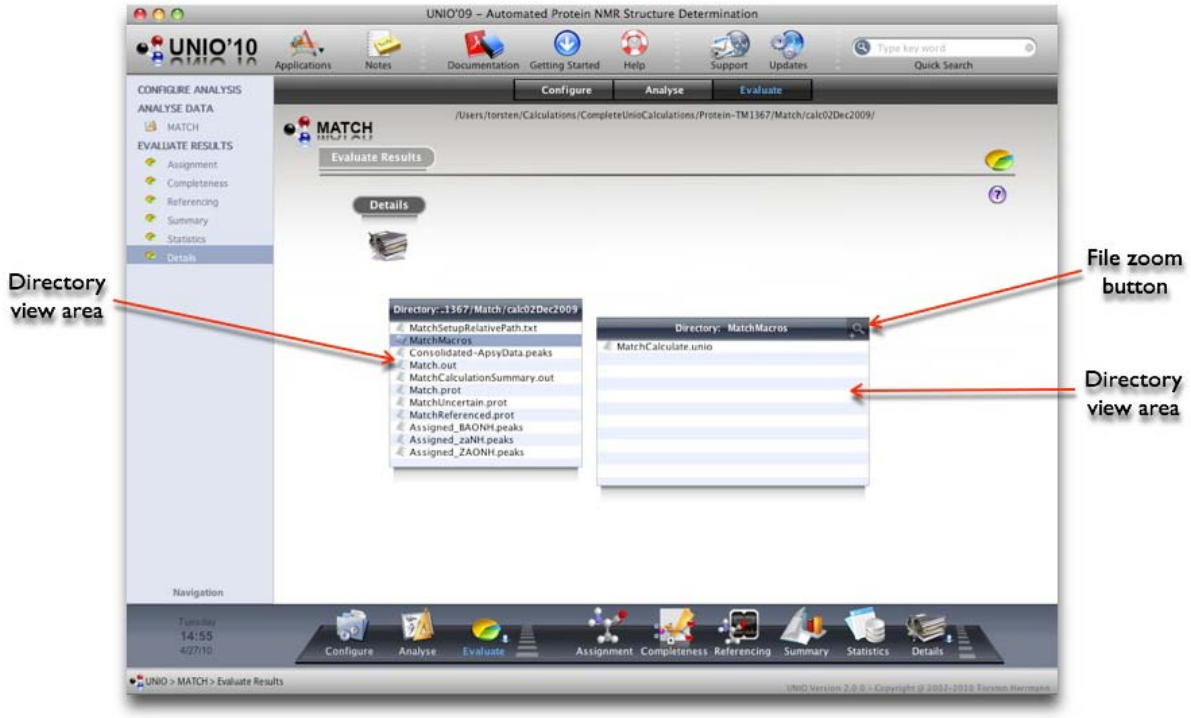
2.11.3.6 "Evaluate Results: Details"

The sub-section "Details" allows you to browse through all MATCH output directories and files.

A single mouse click displays the content of a directory in the right table, while double mouse click opens a file in the UNIO File Editor.

If your mouse pointer is over an entry of the left or right table and has been longer than 1 second in the *Output view area*, *UNIO File zoom* gets activated. The *UNIO file zoom* allows you to see the file content without explicitly opening the file. In addition, *UNIO File zoom* gives you at the bottom of the *File zoom area* information about the type of the file. You can use the *File zoom button* at the top of the left table to switch *UNIO File zoom* on and off.

Tip: Use the *UNIO File zoom* function to quickly see the content of a file and to learn about the type of file you are looking at.



UNIO Application window for MATCH, section "Evaluate Results: Details".

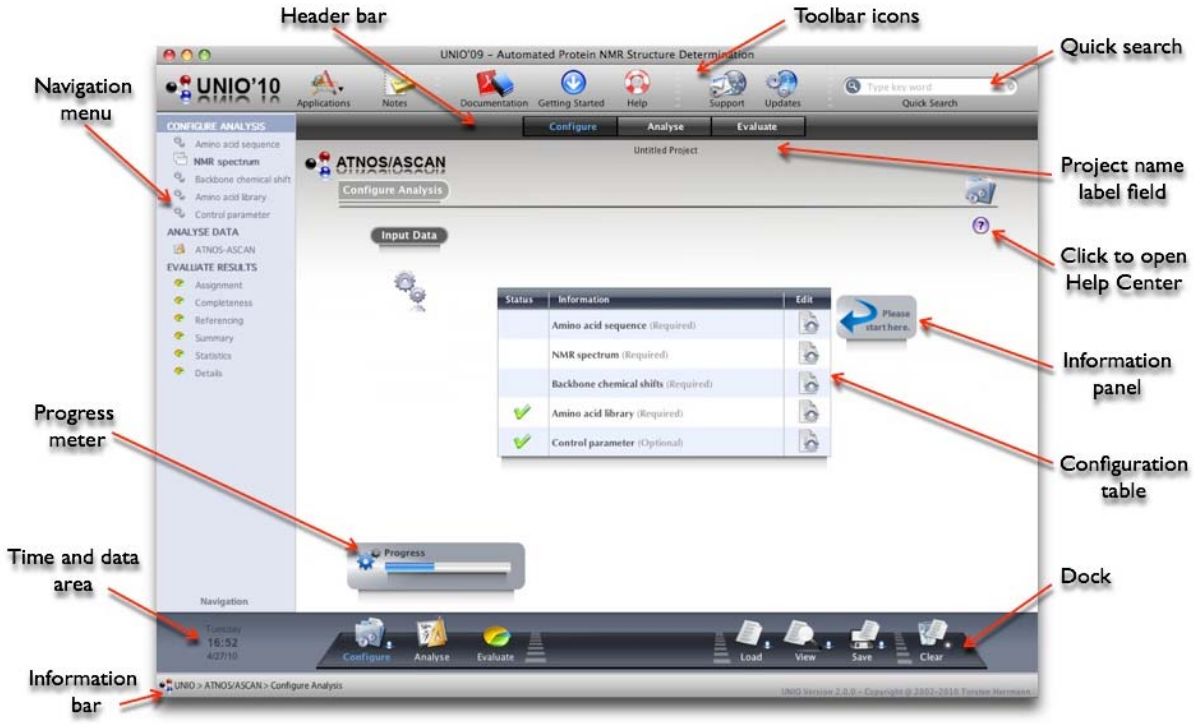
UNIO Application window for MATCH, section "Evaluate Results: Details". UNIO File zoom is active.

2.12 UNIO - Application window for ATNOS/ASCAN

The UNIO ATNOS/ASCAN window allows you to perform automated side-chain resonance assignment. The ATNOS/ASCAN environment is split into three main sections "Configure Analysis", "Analyse NMR Data" and "Evaluate Results". The *Toolbar icons* give you fast access to frequently used functions of UNIO. The *Header bar* allows you to navigate between the three main sections of ATNOS/ASCAN. The *Navigation menu* gives you also access to sub-sections of the three main ATNOS/ASCAN sections. The *Dock* allows you to navigate between the three main sections and provides functionalities specific for one section.

2.12.1 Section "Configure Analysis"

The ATNOS/ASCAN section "Configure Analysis" is meant to collect all input data required to perform ATNOS/ASCAN side-chain assignment. The *Configuration table* lists all five input components, launches the corresponding input windows and reports along with the *Progress meter* about the completeness of the input configuration. You can click on a row of the *Configuration table* to open the corresponding input window. A description of the individual input windows are given later in this User's Guide.



UNIO Application window for ATNOS/ASCAN, section "Configure Analysis".

Toolbar icons

Applications	Opens an option menu for navigating between the UNIO NMR data analysis routines of MATCH, ATNOS/ASCAN or ATNOS/CANDID.
Notes	Opens the UNIO Notes window.
Documentation	Opens this User's Guide.
Getting Started	Opens the Getting Started documentation for ATNOS/ASCAN.
Help	Opens the UNIO Help Center.
Support	Opens the system's default mail program.
Updates	Opens the UNIO website.
Quick Search	Opens the UNIO Help Center with the typed key word.

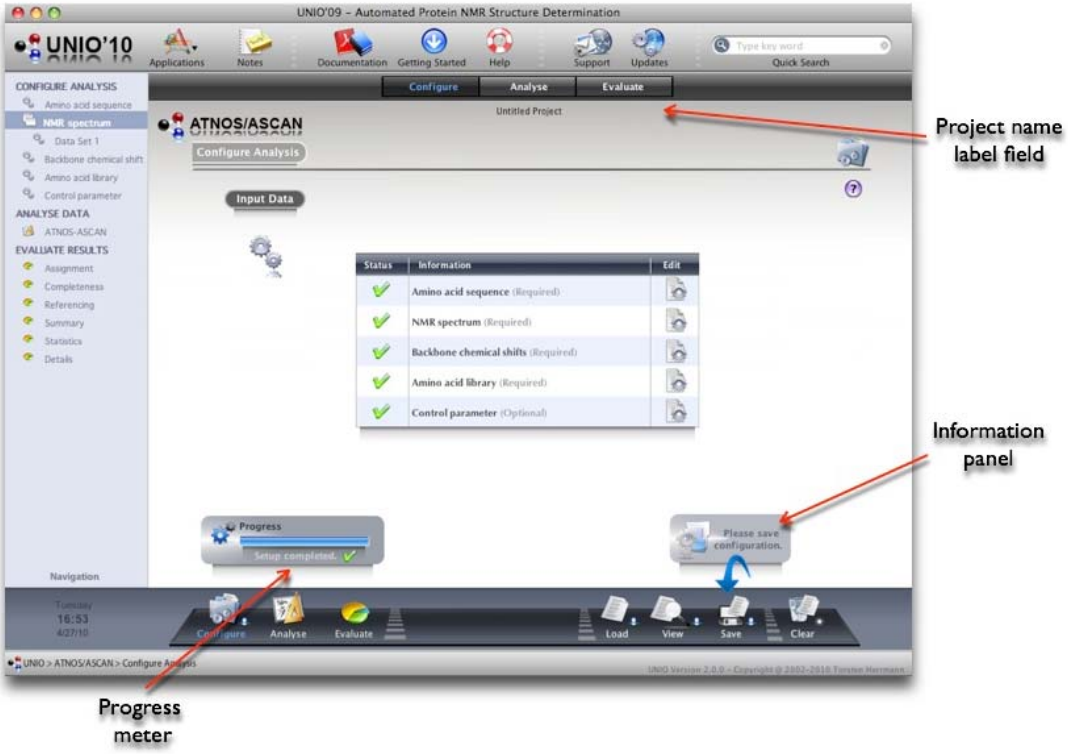
Header bar

Configure	Opens the ATNOS/ASCAN main section "Configure Analysis".
Analyse	Opens the ATNOS/ASCAN main section "Analyse NMR Data".
Evaluate	Opens the ATNOS/ASCAN main section "Evaluate Results".

Dock

Configure	Opens the ATNOS/ASCAN main section "Configure Analysis".
Analyse	Opens the ATNOS/ASCAN main section "Analyse NMR Data".
Evaluate	Opens the ATNOS/ASCAN main section "Evaluate Results".
Load	Opens the OS file dialog to load an ATNOS/ASCAN configuration file.
View	Opens the UNIO Summary window that reports about all input data.
Save	Opens the OS file dialog to save an ATNOS/ASCAN configuration file.
Clear	Clears all input configuration.

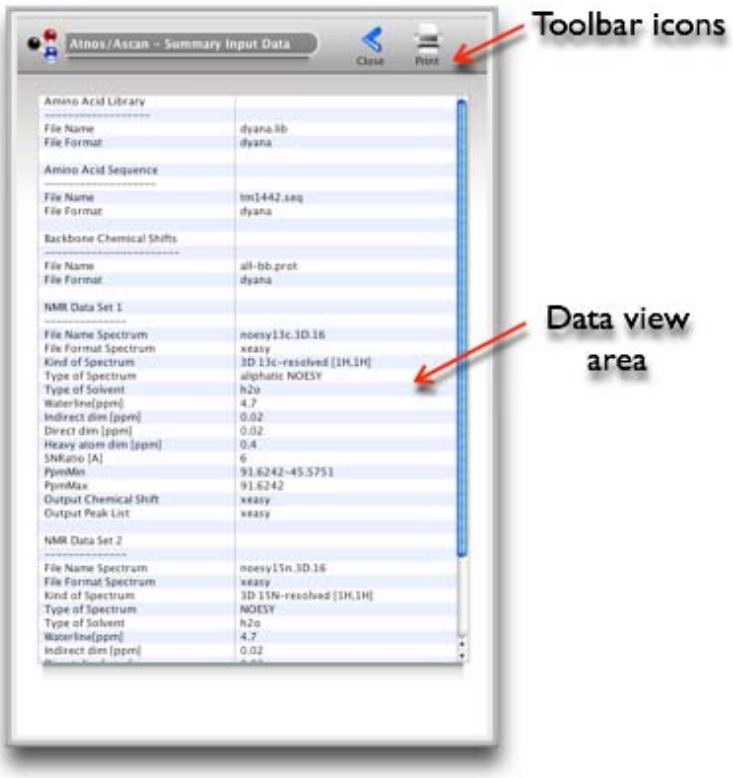
After you have provided all required input data, you first need to save your input data as ATNOS/ASCAN configuration file, conventionally called "AtnosAscanSetup.txt". The directory containing your ATNOS/ASCAN configuration file will be in the following considered as your project directory and all ATNOS/ASCAN output data will be written in this directory. The name of this directory is displayed in the *Project name label field* underneath the *Header bar*.



UNIO Application window for ATNOS/ASCAN, section "Configure Analysis".

Tip: The information about the required input data and file format is provided with the description of the respective input window (see later).

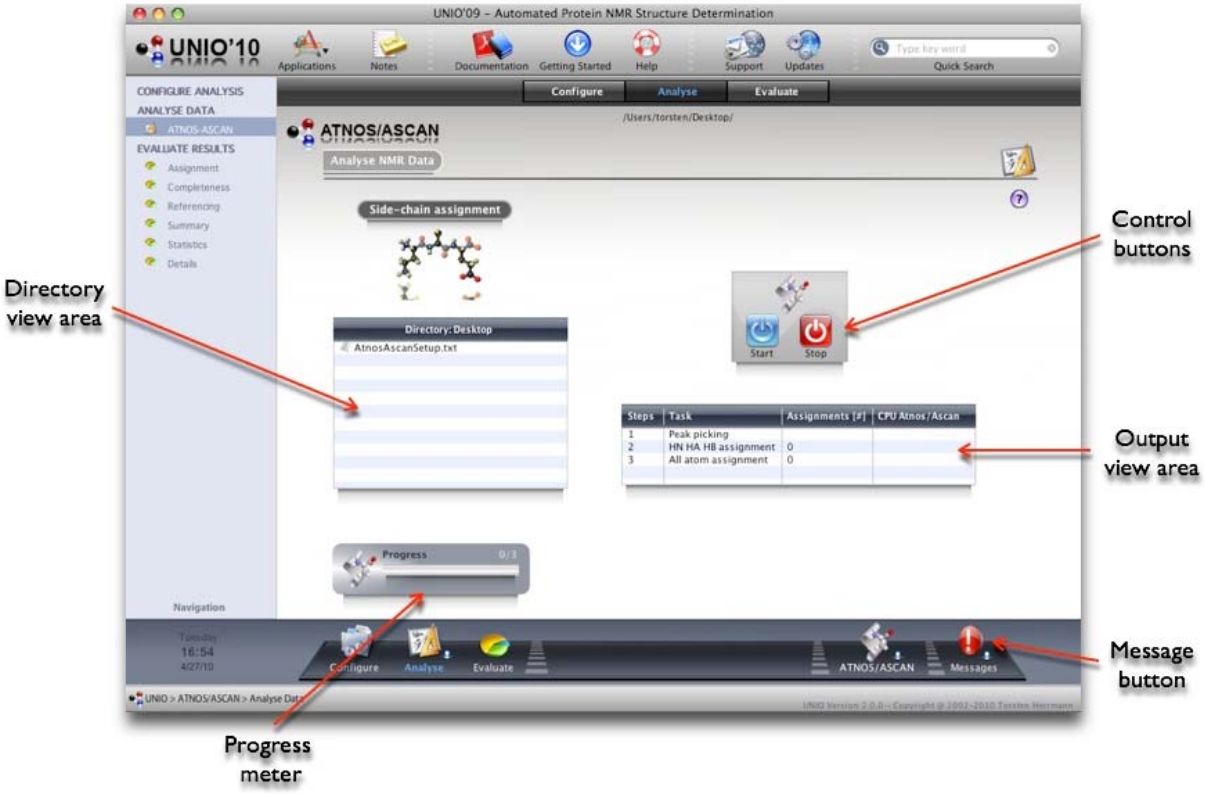
The *View button* in the *Dock* opens the UNIO Summary window which gives you an overview of all input data provided. The *Toolbar icons* allows you to close this window and to print the displayed text in the *Data view area* which corresponds to the ATNOS/ASCAN configuration file.



UNIO Input Summary window for ATNOS/ASCAN, section "Configure Analysis".

2.12.2 Section "Analyse NMR Data"

The ATNOS/ASCAN section "Analyse NMR Data" represents the core of ATNOS/ASCAN. By clicking on the *Control button* "Start", you initiate automated ATNOS/ASCAN side-chain assignment. During the NMR data analysis, the *Progress meter* reports about the state of the calculation, similar the *Directory view area* is getting updated frequently to report about ATNOS/ASCAN output. At the end of ATNOS/ASCAN data analysis, a brief summary about the assignment completeness and the CPU usage is given in the *Output view area*.

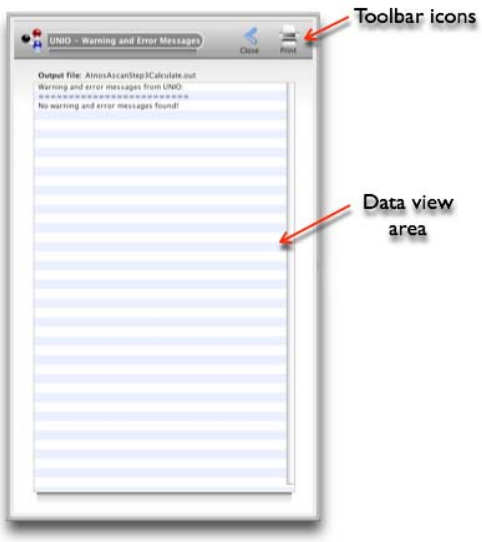


UNIO Application window for ATNOS/ASCAN, section "Analyse NMR Data".

Tip: Single mouse click selects a file or lists a directory in the *Directory view area*. Double mouse click opens a file selected in the UNIO File Editor.

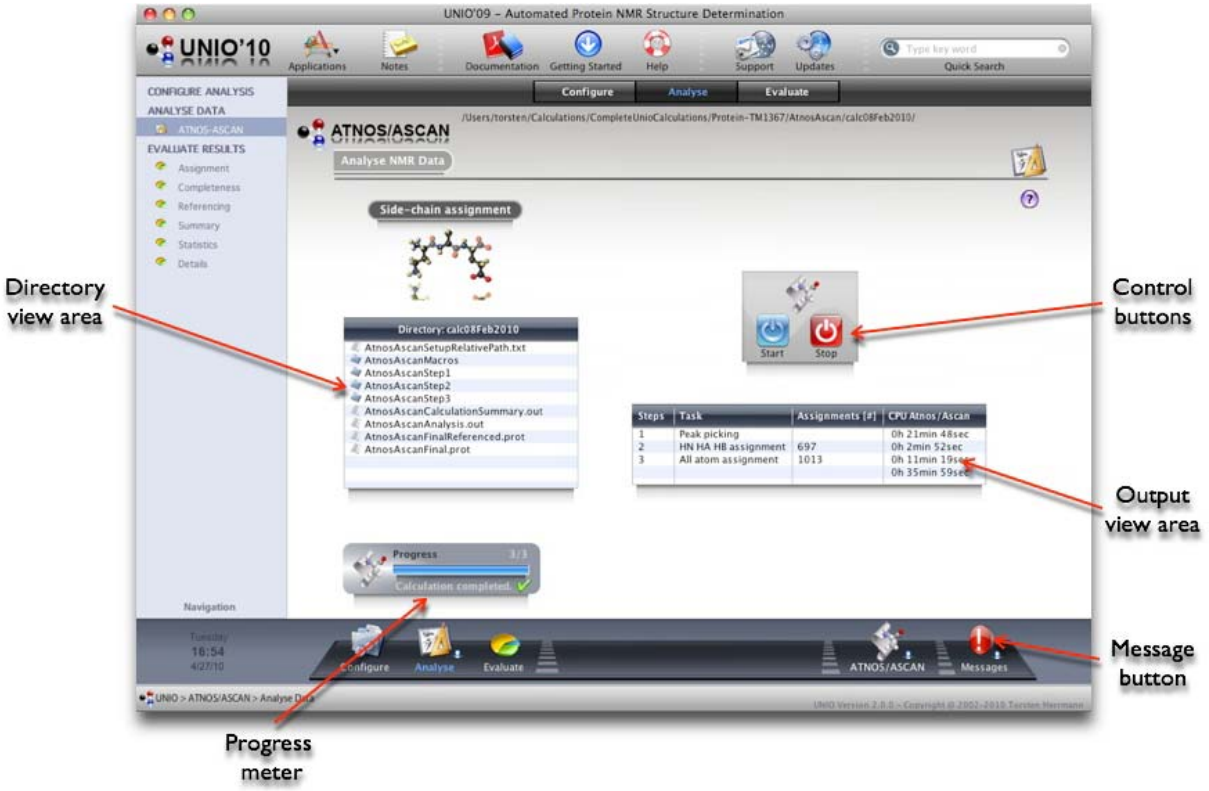
The *Message button* in the *Dock* opens the UNIO Messages window. This window displays all warning and error messages in all output files generated by UNIO.

Tip: The *Message button* might be useful to find potential problems in the input data and to correct your configuration file if necessary.



UNIO Warning and Error Messages for ATNOS/ASCAN, section "Analyse NMR data".

After a ATNOS/ASCAN calculation is completed, you can browse through all output files by mouse click in the *Directory view area*. The *Output view area* reports about the number of assigned atoms and the CPU time spent. For detailed inspection of the results obtained, you can now proceed to the ATNOS/ASCAN section "Evaluate Results".



UNIO Application window for ATNOS/ASCAN, section "Analyse NMR Data". Calculation completed.

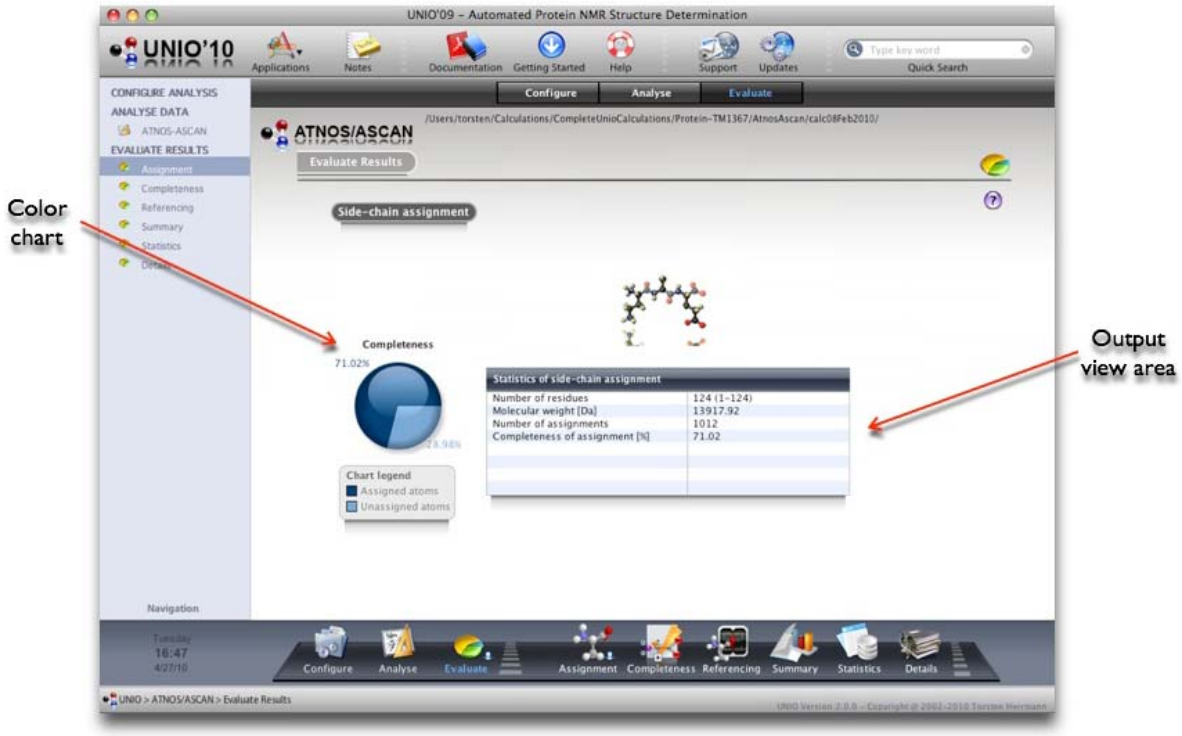
2.12.3 Section "Evaluate Results"

The ATNOS/ASCAN section "Evaluate Results" consists of six sub-sections "Side-chain Assignment", "Completeness", "Referencing", "Summary", "Statistics" and "Details". The *Navigation menu* or the *Dock* allows you to browse through these sections.

2.12.3.1 "Evaluate Results: Side-chain Assignment"

The sub-section "Side-chain Assignment" gives a brief overview of the results obtained.

The top *Output view area* provides a summary of the ATNOS/ASCAN results obtained (see table below). The *Color chart* gives a quick report about the assignment completeness achieved.



UNIO Application window for ATNOS/ASCAN, section "Evaluate Results: Side-chain Assignment".

Table for sub-section "Side-chain Assignment"

Number of residues	Gives the number of amino acid residues in the protein sequence.
Molecular weight	Gives the molecular weight of the protein in Dalton.
Number of assignments	Total number of assigned atoms in the protein sequence.
Completeness of assignments	Percentage of assigned atoms in the protein sequence.

2.12.3.2 "Evaluate Results: Completeness"

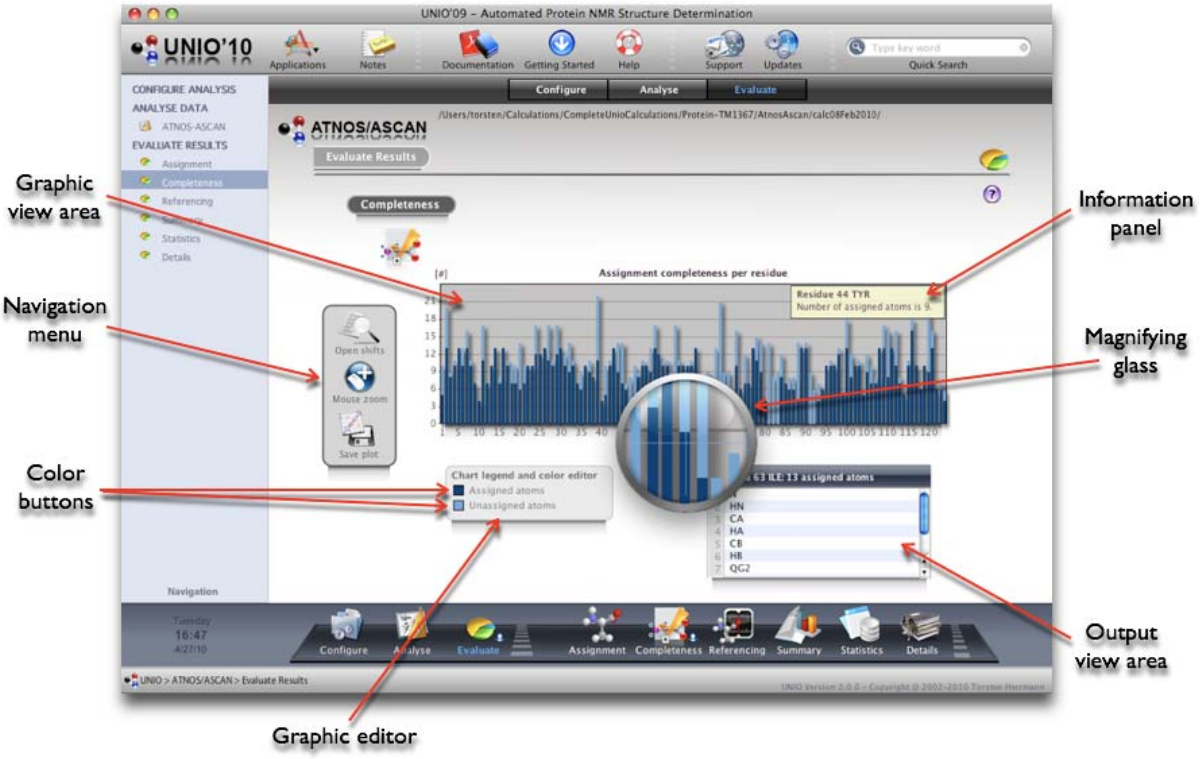
The sub-section "Completeness" informs about the assignment completeness as a function of the amino acid sequence. The *Graphic view area* allows the user to directly interact with the ATNOS/ASCAN results.

To facilitate the navigation in the *Graphic view area*, a *Magnifying glass* is opened automatically as long the mouse pointer is within the *Graphic view area*. If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding residue name and the number of assigned and unassigned atoms are displayed in the *Information panel*.

If you click with the mouse on a chart bar in the *Graphic view area*, the assigned resp. unassigned atoms are listed in the *Output view area*.

The *Graphic editor* allows you to change the color-presentation of the chart. A mouse click on the *Color buttons* will launch the operating systems color palette window.

The *Navigation menu* serves three purposes: First, you can open the chemical shift list of the assigned backbone atoms with the UNIO File Editor; second, you can change the zoom factor used for the *Magnifying glass*; and third, you can save the residue chart in PNG file format with grey or transparent background.



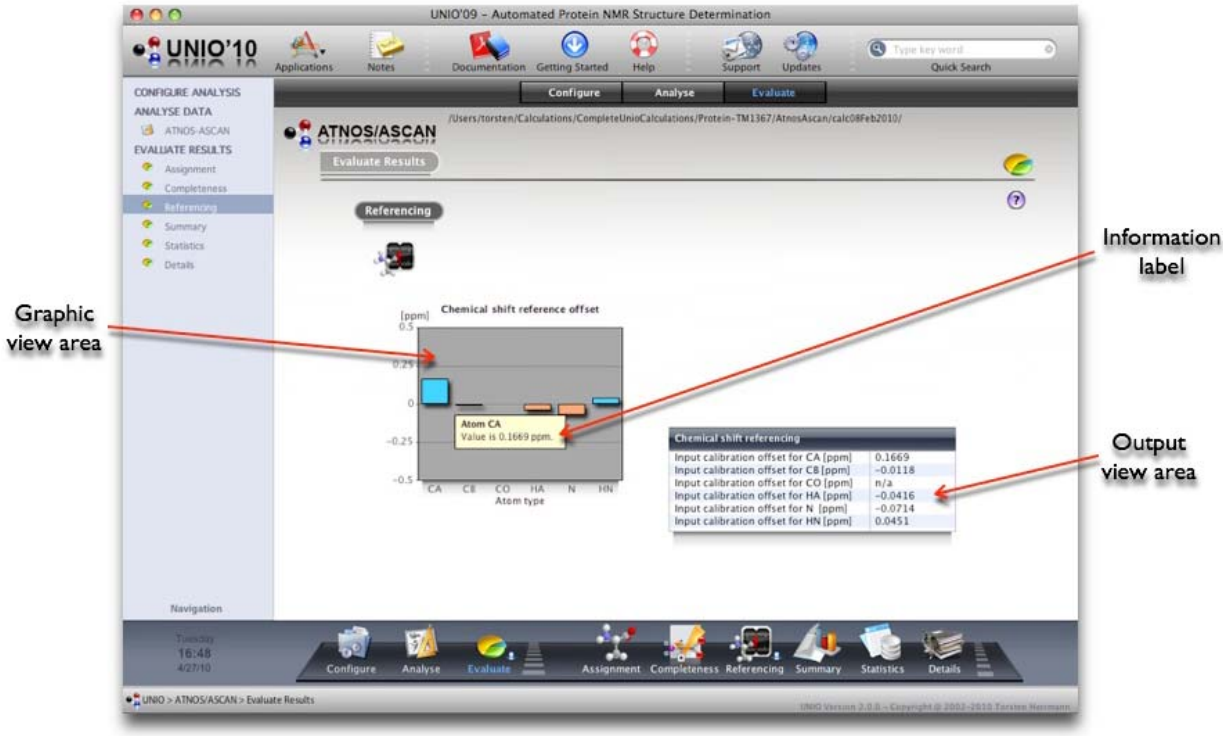
UNIO Application window for ATNOS/ASCAN, section "Evaluate Results: Completeness".

2.12.3.3 "Evaluate Results: Referencing"

The sub-section "Referencing" informs about the chemical shift referencing offset of all backbone atoms. The *Graphic view area* allows the user to directly visualize the respective referencing offset.

If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding atom name and the value of the reference offset in ppm are displayed in the *Information panel*.

The Output view area summarizes the chart information in table presentation.



UNIO Application window for ATNOS/ASCAN, section "Evaluate Results: Referencing".

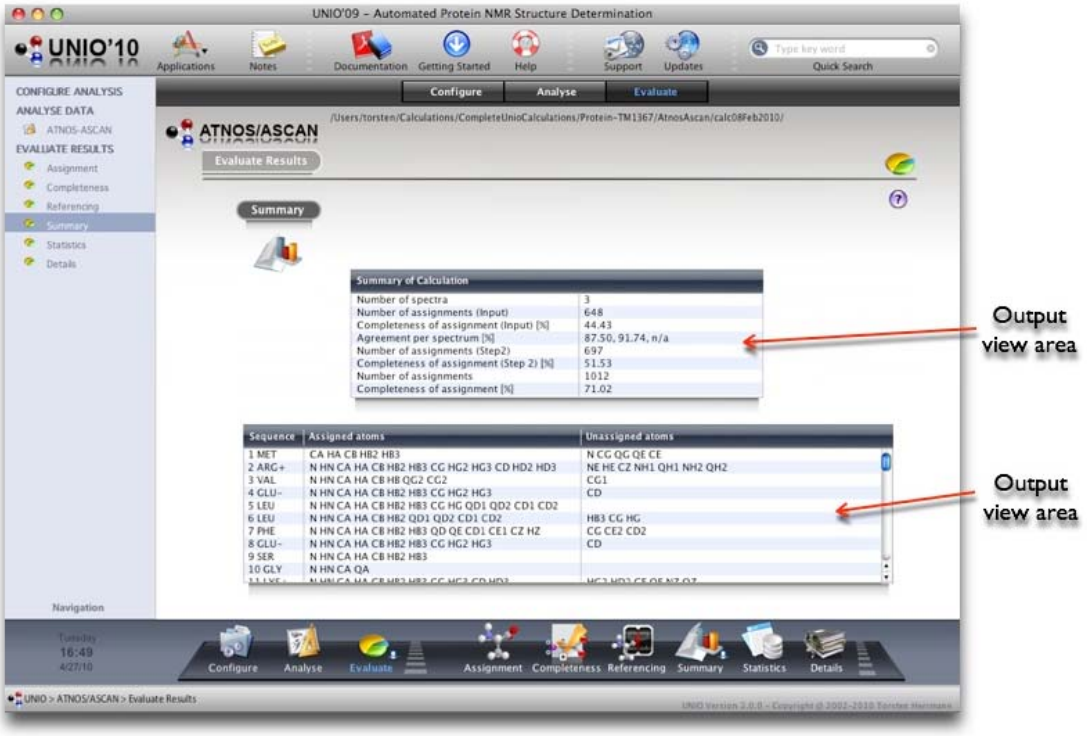
Important: The UNIO-ATNOS/ASCAN algorithm performs an automated chemical shift referencing prior and at the end of the overall ATNOS/ASCAN assignment process. Although this automated chemical shift referencing is extremely efficient and correct, it is nonetheless recommended to prepare the input for a ATNOS/ASCAN calculation in such a way that the detected chemical shift reference offsets for the individual atoms are small. Please keep in mind that the sequence-specific assignment process is strongly guided by comparison between your input chemical shift values of the peak positions and the average chemical shift value of corresponding atoms found for NMR active nuclei in proteins taken from the BMRB data base.

2.12.3.4 "Evaluate Results: Summary"

The sub-section "Summary" consists of two tables providing information about user-given input and ATNOS/ASCAN output.

The upper table summarizes the main input and output information, such as the number of input NMR spectra used, the number of assigned input atoms, the assignment completeness of the input assignment, the number of assignments and the assignment completeness after ATNOS/ASCAN step 2 and 3.

The lower table reports for each amino acid residue in the protein sequence about missing resonance assignments.



UNIO Application window for ATNOS/ASCAN, section "Evaluate Results: Summary".

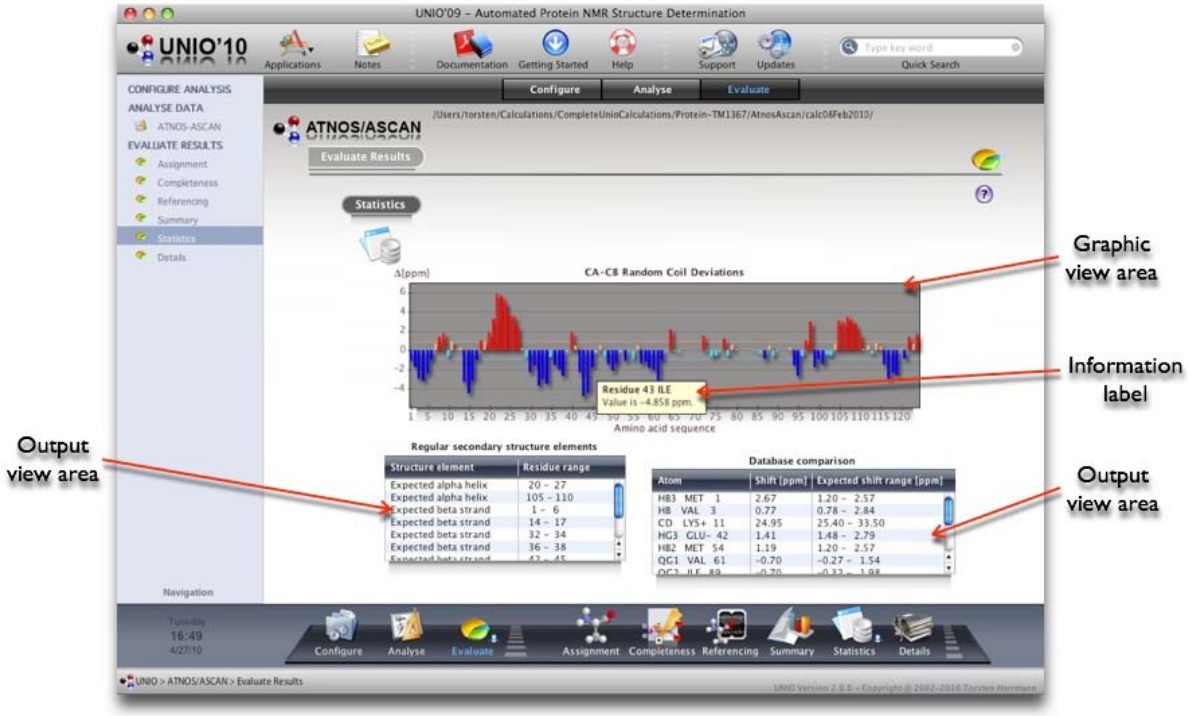
2.12.3.5 "Evaluate Results: Statistics"

The sub-section "Statistics" compares ATNOS/ASCAN resonance assignments with average chemical shift values of proteins.

The *Graphic view area* plots the CA-CB random coil deviation along the protein sequence. If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding residue name and the value of the random coil deviation in ppm are displayed in the *Information panel*. Helical and strand secondary structure elements are color-coded in red and blue respectively.

Consecutive amino acid residues with a chemical shift deviation above or below a certain threshold (orange and cyan lines in the *Graphic view area*) are listed as regular secondary structure elements in the table on the left side.

The table on the right side reports about chemical shift assignments that are unusual with respect to average chemical shift values found in proteins.



UNIO Application window for ATNOS/ASCAN, section "Evaluate Results: Statistics".

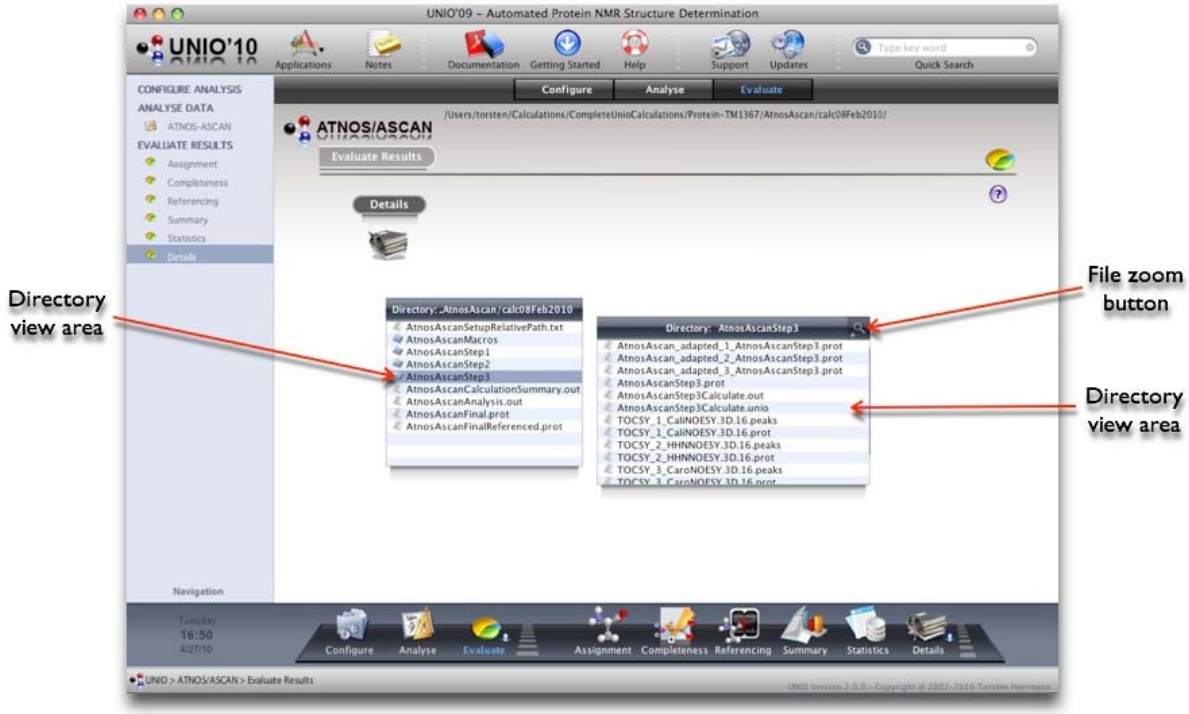
2.12.3.6 "Evaluate Results: Details"

The sub-section "Details" allows you to browse through all ATNOS/ASCAN output directories and files.

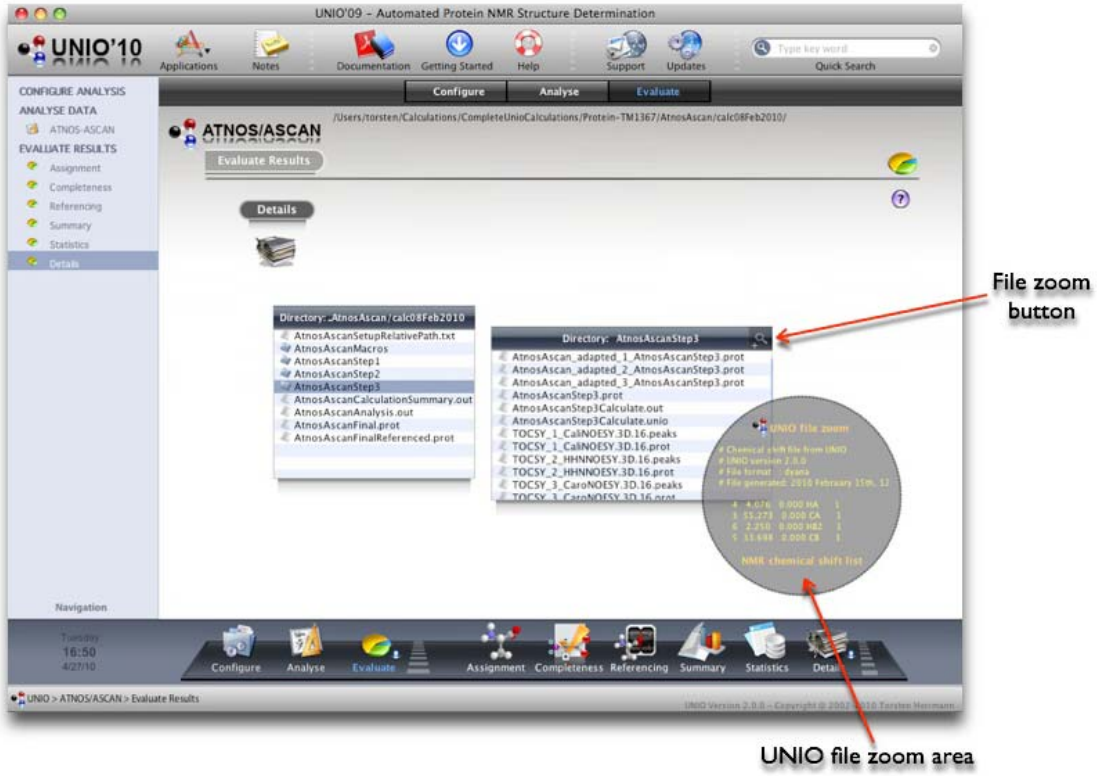
A single mouse click displays the content of a directory in the right table, while double mouse click opens a file in the UNIO File Editor.

If your mouse pointer is over an entry of the left or right table and has been longer than 1 second in the *Output view area*, *UNIO File zoom* gets activated. The *UNIO file zoom* allows you to see the file content without explicitly opening the file. In addition, *UNIO File zoom* gives you at the bottom of the *File zoom area* information about the type of the file. You can use the *File zoom button* at the top of the left table to switch *UNIO File zoom* on and off.

Tip: Use the *UNIO File zoom* function to quickly see the content of a file and to learn about the type of file you are looking at.



UNIO Application window for ATNOS/ASCAN, section "Evaluate Results: Details".



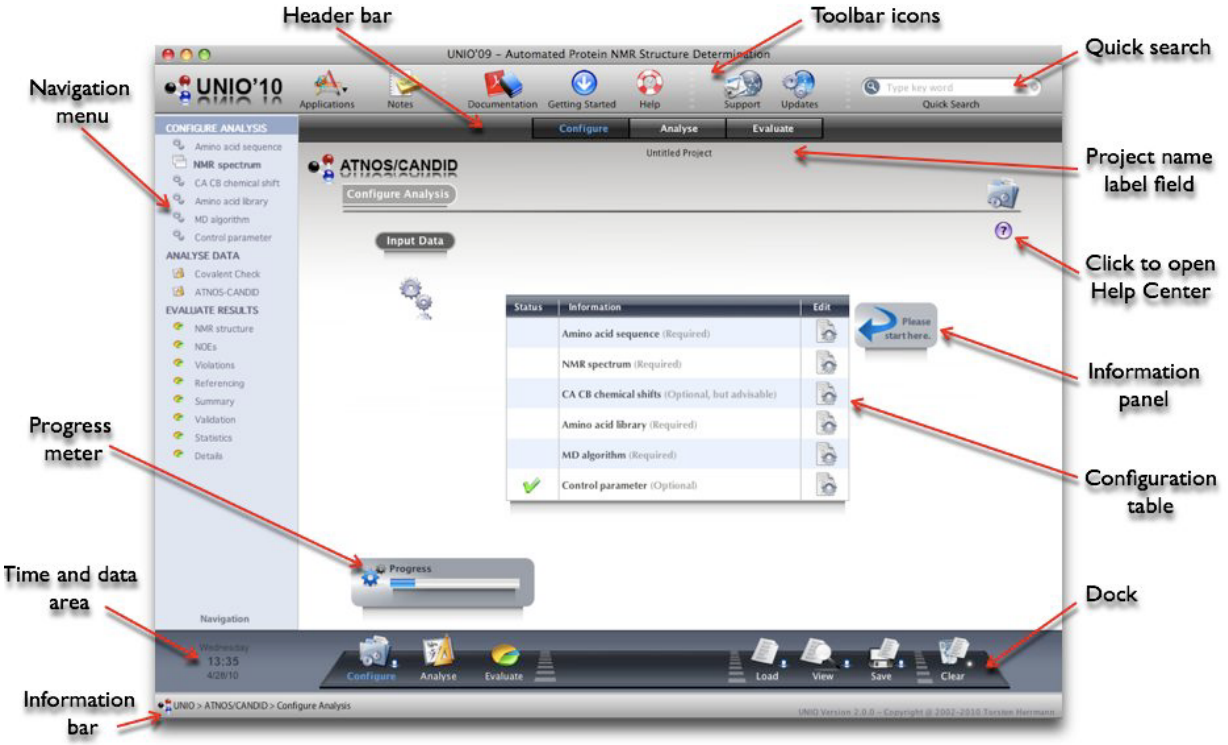
UNIO Application window for ATNOS/ASCAN, section "Evaluate Results: Details". UNIO File zoom active.

2.13 UNIO - Application window for ATNOS/CANDID

The UNIO ATNOS/CANDID window allows you to perform automated NOE cross peak assignment. The ATNOS/CANDID environment is split into three main sections "Configure Analysis", "Analyse NMR Data" and "Evaluate Results". The *Toolbar icons* give you fast access to frequently used functions of UNIO. The *Header bar* allows you to navigate between the three main sections of ATNOS/CANDID. The *Navigation menu* gives you also access to sub-sections of the three main ATNOS/CANDID sections. The *Dock* allows you to navigate between the three main sections and provides functionalities specific for one section.

2.13.1 Section "Configure Analysis"

The ATNOS/CANDID section "Configure Analysis" is meant to collect all input data required to perform ATNOS/CANDID NOE assignment. The *Configuration table* lists all five input components, launches the corresponding input windows and reports along with the *Progress meter* about the completeness of the input configuration. You can click on a row of the *Configuration table* to open the corresponding input window. A description of the individual input windows are given later in this User's Guide.



UNIO Application window for ATNOS/CANDID, section "Configure Analysis".

Toolbar icons

Applications	Opens an option menu to navigate to the NMR data analysis routines of MATCH, ATNOS/ASCAN or ATNOS/CANDID.
Notes	Opens the UNIO Notes window.
Documentation	Opens this User's Guide.
Getting Started	Opens the Getting Started documentation for ATNOS/CANDID.
Help	Opens the UNIO Help Center.
Support	Opens the system's default mail program.
Updates	Opens the UNIO website.
Quick Search	Opens the UNIO Help Center with the typed key word.

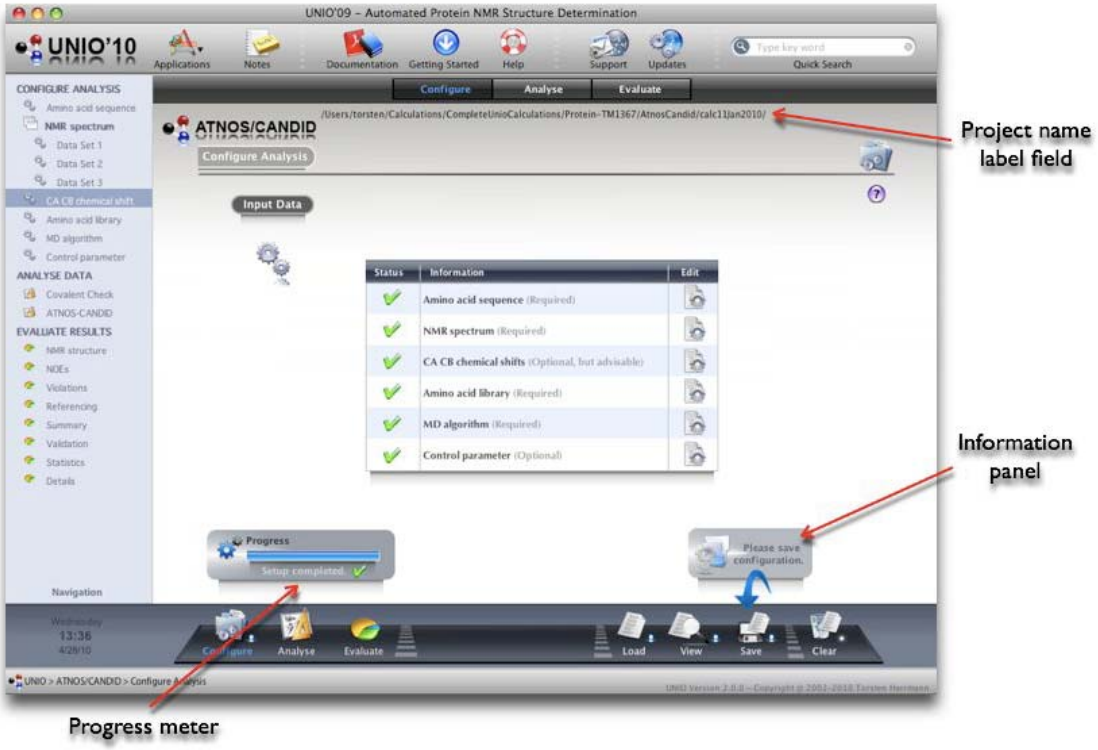
Header bar

Configure	Opens the ATNOS/CANDID main section "Configure Analysis".
Analyse	Opens the ATNOS/CANDID main section "Analyse NMR Data".
Evaluate	Opens the ATNOS/CANDID main section "Evaluate Results".

Dock

Configure	Opens the ATNOS/CANDID main section "Configure Analysis".
Analyse	Opens the ATNOS/CANDID main section "Analyse NMR Data".
Evaluate	Opens the ATNOS/CANDID main section "Evaluate Results".
Load	Opens a file dialog to load an ATNOS/CANDID configuration file.
View	Opens the UNIO Summary window that reports about all input data.
Save	Opens a file dialog to save an ATNOS/CANDID configuration file.
Clear	Clears all input configuration.

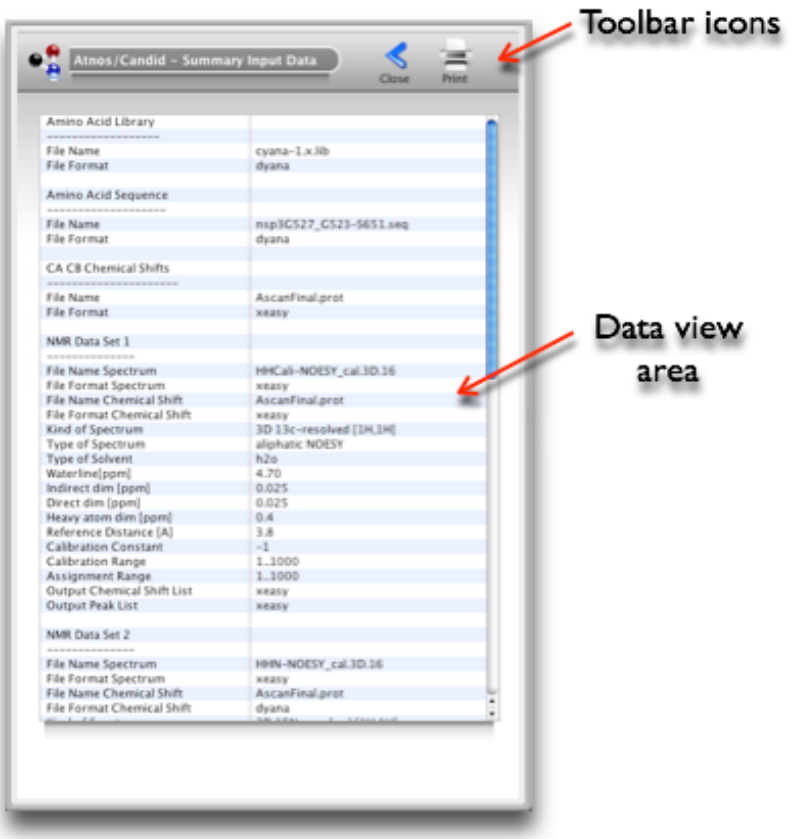
After you have provided all required input data, you first need to save your input data as ATNOS/CANDID configuration file, conventionally called "AtnosCandidSetup.txt". The directory containing your ATNOS/CANDID configuration file will be in the following considered as your project directory and all ATNOS/CANDID output data will be written in this directory. The name of this directory is displayed in the *Project name label field* underneath the *Header bar*.



UNIO Application window for ATNOS/CANDID, section "Configure Analysis".

Tip: The information about the required input data and file format is provided with the description of the respective input window (see later).

The *View button* in the *Dock* opens the UNIO Summary window which gives you an overview of all input data provided. The *Toolbar icons* allows you to close this window and to print the displayed text in the *Data view area* which corresponds to the ATNOS/CANDID configuration file.



UNIO Input Summary window for ATNOS/CANDID, section "Configure Analysis".

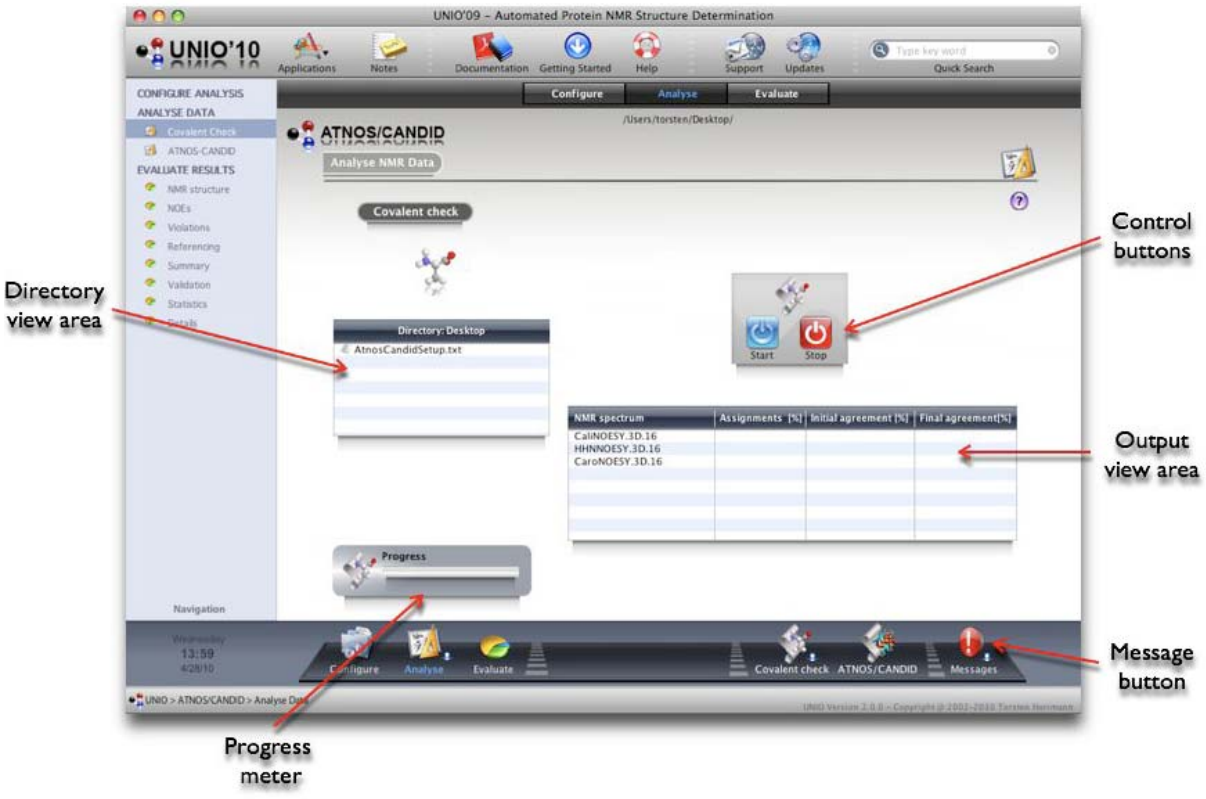
2.13.2 Section "Analyse NMR Data"

The ATNOS/CANDID section "Analyse NMR Data" represents the core of ATNOS/CANDID. This section is split into two section, the "Covalent check" and "ATNOS/CANDID" structure determination.

2.13.2.1 "Analyse NMR Data: Covalent Check"

The sub-section covalent check verifies the fitting of the chemical shift values of the assigned atoms with the NMR spectra used. Resonance assignment have usually not been obtained using NOESY spectra. Except for the cases that you have used the ATNOS/ASCAN side-chain assignment module. In this case you have always have a perfect agreement between the chemical shift values of the assigned atoms and the NOESY spectra. In all other cases, in order to verify that the spectrum calibration or referencing has been done correctly, the "Covalent check" has been introduced. The covalent check tells about the calibration agreement between chemical shifts of assigned atoms and the NOESY spectra.

The covalent check reports the percentage of validated covalent contacts defined as number of observed contacts divided by the number of covalent structure-imposed contacts. To this end, the ATNOS/CANDID algorithm generates a set of pairwise combinations of protons i and j (i is not equal j) for which sequence-specific resonance assignments are available, and which have covalent structure-imposed upper distance limits shorter than a user-given threshold (For example, the maximal distance between the intraresidual atoms HA and HN is always smaller than 5 Angstrom, and therefore must lead to an observable signal in the NOESY spectrum).

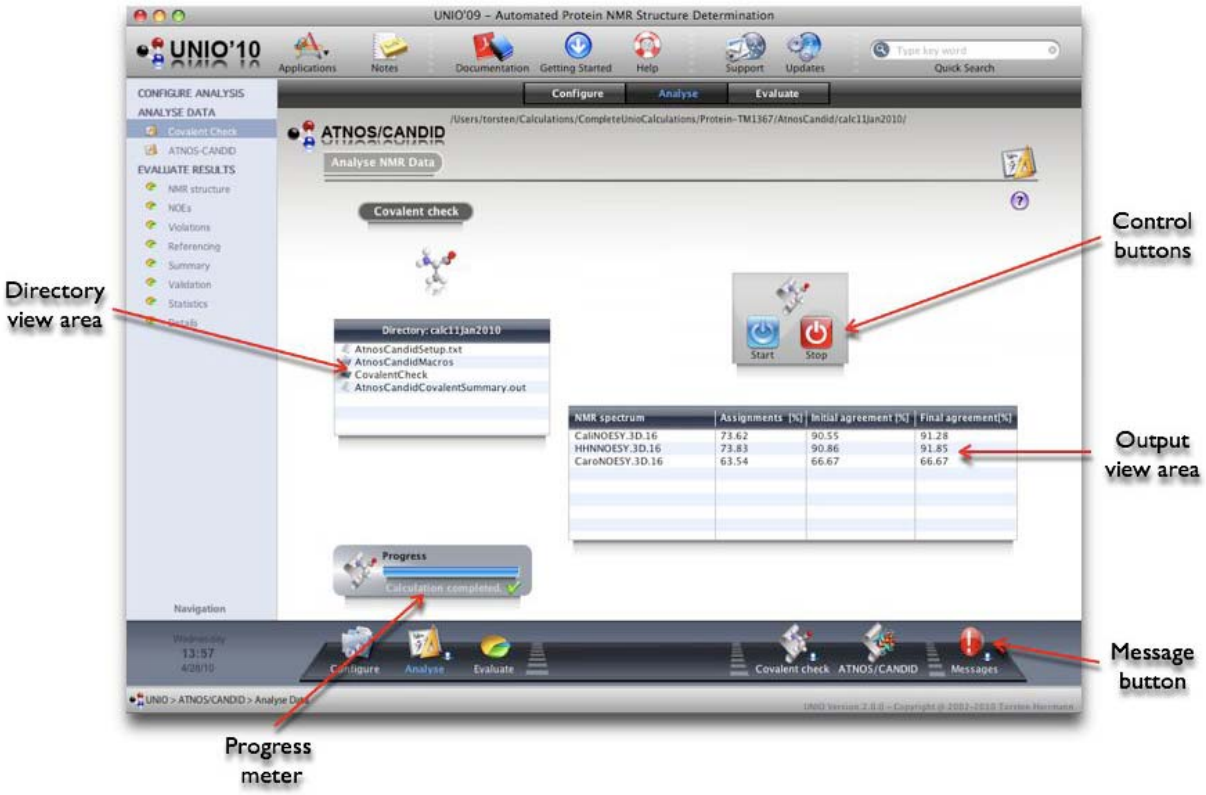


UNIO Application window for ATNOS/CANDID, section "Analyse NMR Data: Covalent check".

A covalent contacts between two atoms is validated if a local extremum is found in the NMR spectrum whose frequency coordinates agree with the chemical shift assignment of the corresponding atom pair within a user-given tolerance window. A low percentage of confirmed covalent contacts is indicative for an insufficient calibration between resonance assignments and NOESY spectra.

By clicking on the *Control button* "Start", you initiate the ATNOS/CANDID covalent check. During the NMR data analysis, the *Progress meter* reports about the state of the covalent check, similar the *Directory view area* is getting updated frequently to report about ATNOS/CANDID output.

At the end of the ATNOS/CANDID covalent check, a brief summary about the assignment completeness, the initial chemical shift agreement and the final chemical shift agreement is given. The initial chemical shift agreements reports about the percentage of covalent contacts found based on the input chemical shifts. ATNOS/CANDID has a built-in calibration routine that is applied during the covalent check. Thus, the final agreement reports about the chemical shift agreement after automated calibration of the resonance assignments to the NOESY spectra has been applied.



UNIO Application window for ATNOS/CANDID, section "Analyse NMR Data: Covalent check".

2.13.2.2 "Analyse NMR Data: ATNOS/CANDID"

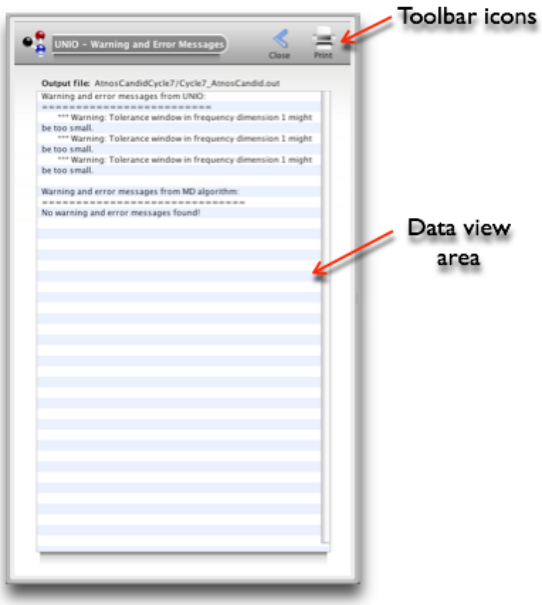
The sub-section "ATNOS/CANDID" represents the automated NOE assignment and structure determination part of UNIO. By clicking on the *Control button* "Start", you initiate automated ATNOS/CANDID NOE assignment. During the NMR data analysis, the *Progress meter* reports about the state of the calculation, similar the *Directory view area* is getting updated frequently to report about ATNOS/CANDID output. At the end of ATNOS/CANDID data analysis, a brief summary about the evolution of characteristic ATNOS/CANDID parameters and the CPU usage is given in the *Output view area*.

UNIO Application window for ATNOS/CANDID, section "Analyse NMR Data: ATNOS/CANDID".

Tip: Single mouse click selects a file or lists a directory in the *Directory view area*. Double mouse click opens a file selected in the UNIO File Editor.

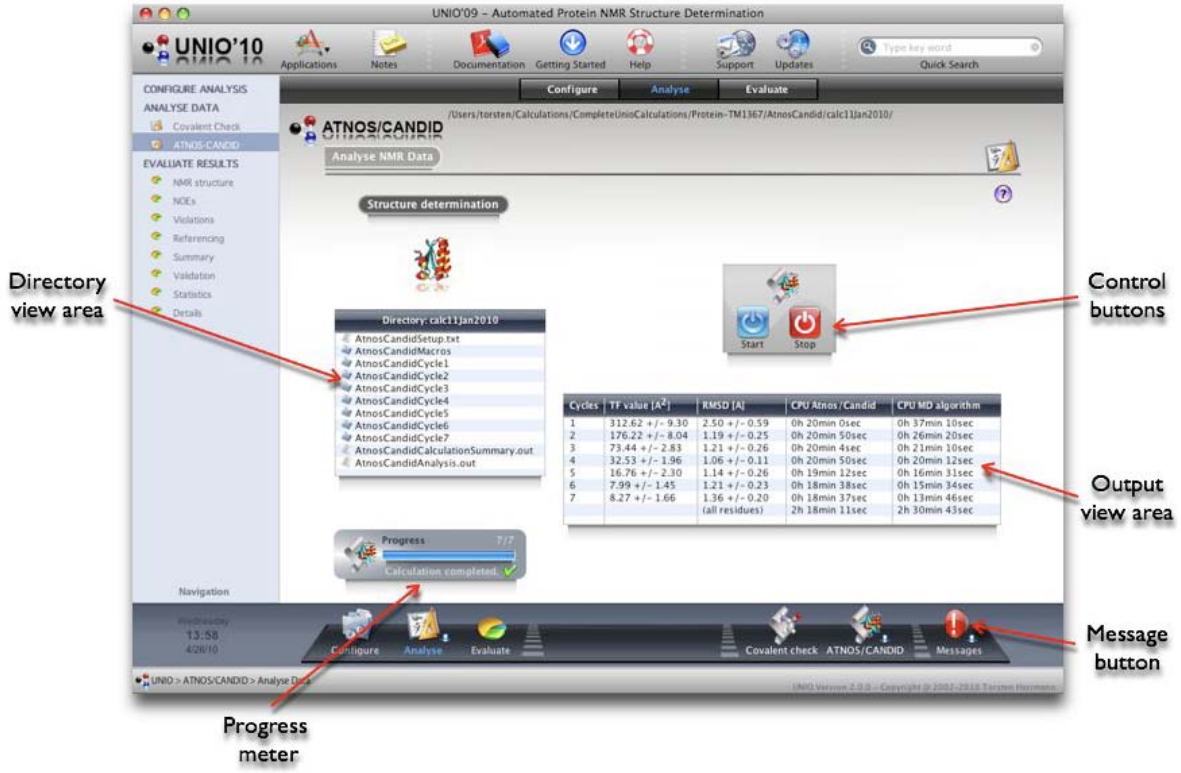
The *Message button* in the *Dock* opens the UNIO Messages window. This window displays all warning and error messages in all output files generated by UNIO.

Tip: The *Message button* might be useful to find potential problems in the input data and to correct your configuration file if necessary.



UNIO Warning and Error Messages for ATNOS/CANDID, section "Analyse NMR data".

After a ATNOS/CANDID calculation is completed, you can browse through all output files by mouse click in the *Directory view area*. The *Output view area* reports about the evolution of characteristic ATNOS/CANDID parameters and the CPU time spent. For detailed inspection of the results obtained, you can now proceed to the ATNOS/CANDID section "Evaluate Results".



UNIO Application window for ATNOS/CANDID, section "Analyse NMR Data: ATNOS/CANDID".

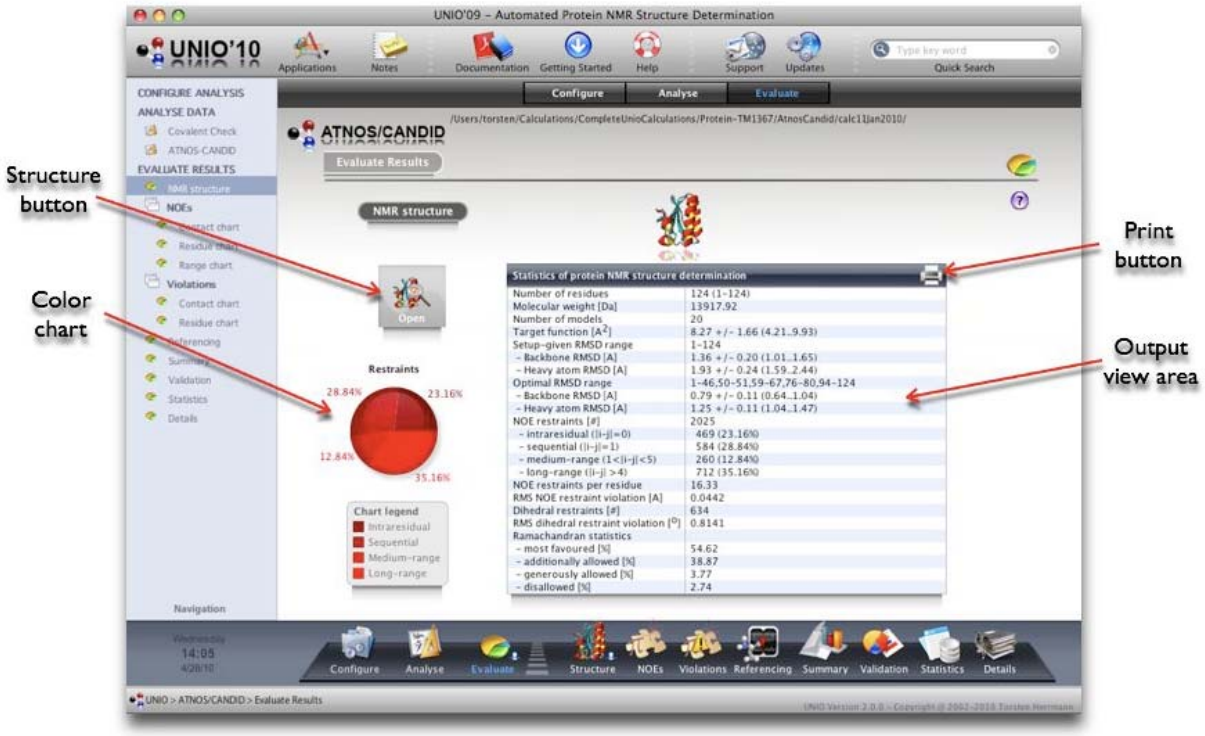
2.13.3 Section "Evaluate Results"

The ATNOS/CANDID section "Evaluate Results" consists of eight sub-sections "Structure", "NOEs", "Violations", "Referencing", "Summary", "Validation", "Statistics" and "Details". The *Navigation menu* or the *Dock* allows you to browse between these sections.

2.13.3.1 "Evaluate Results: Structure"

The sub-section "Structure" gives a brief overview of the results obtained.

The *Output view area* provides a summary of characteristic parameters for the NMR structure obtained by the ATNOS/CANDID approach (see table below). The *Print button* allows you to make a printout of the *Output view area*. The *Structure button* allows you to display the 3D NMR structure using the UNIO Molecule Viewer or an external molecular viewer that you have specified in the UNIO Preferences window. The *Color chart* gives a quick report about the distance restraints used to calculate the final NMR structure bundle.



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: Structure".

Table for sub-section "Structure"

Number of residues	Gives the number of amino acid residues in the protein sequence.
Molecular Weight	Gives the molecular weight of the protein in Dalton.
Number of models	Gives the number of conformers used to represent the final protein structure.

Target function	Reports quantitatively about the agreement between 3D structure and experimental restraints.
Setup-given Backbone RMSD	Calculates the RMSD by superimposing the bundle of conformers for best fit using all N, CA and CO atoms using the residue range specified by the user.
Setup-given Heavy atom RMSD	Calculates the RMSD by superimposing the bundle of conformers for best fit using all heavy atoms using the residue range specified by the user.
Optimal RMSD range Backbone RMSD	Calculates the RMSD by superimposing the bundle of conformers for best fit using all N, CA and CO atoms using the residue range automatically determined by UNIO for best superposition.
Optimal RMSD range Heavy atom RMSD	Calculates the RMSD by superimposing the bundle of conformers for best fit using all heavy atoms using the residue range automatically determined by UNIO for best superposition.
NOE restraints	Lists the number of upper distance restraints used to calculate the 3D structure.
NOE restraints pre residue	Gives the number of distance restraints per residue.
RMS NOE restraint violations	Gives the value for the root mean square violation of the distance restraints.
Dihedral angle restraint	Lists the number of dihedral angle restraints used to calculate the 3D structure.
RMS dihedral restraint violations	Gives the value for the root mean square violation of the dihedral angle restraints.
Ramachadran plot	Lists the Ramachandran statistics.

Important: The reported target function value is calculated by UNIO using a harmonic potential and includes upper distance restraints, dihedral angle restraints and vdW interactions. If you want to see the target function value as calculated by the MD algorithm used, then you need to have a look into the corresponding MD output file in the directory AtnosCandidCycle7. Conventionally this file is called Cycle7_MD.out.

2.13.3.2 "Evaluate Results: NOEs"

The sub-section "NOEs" informs about the NOE restraints used to calculate the final NMR structure bundle. It consists of three further sub-sections "Contact plot", "Residue plot" and "Range plot". The main UNIO light-blue colored *Navigation menu* allows you to browse between these sub-sections of "NOEs".

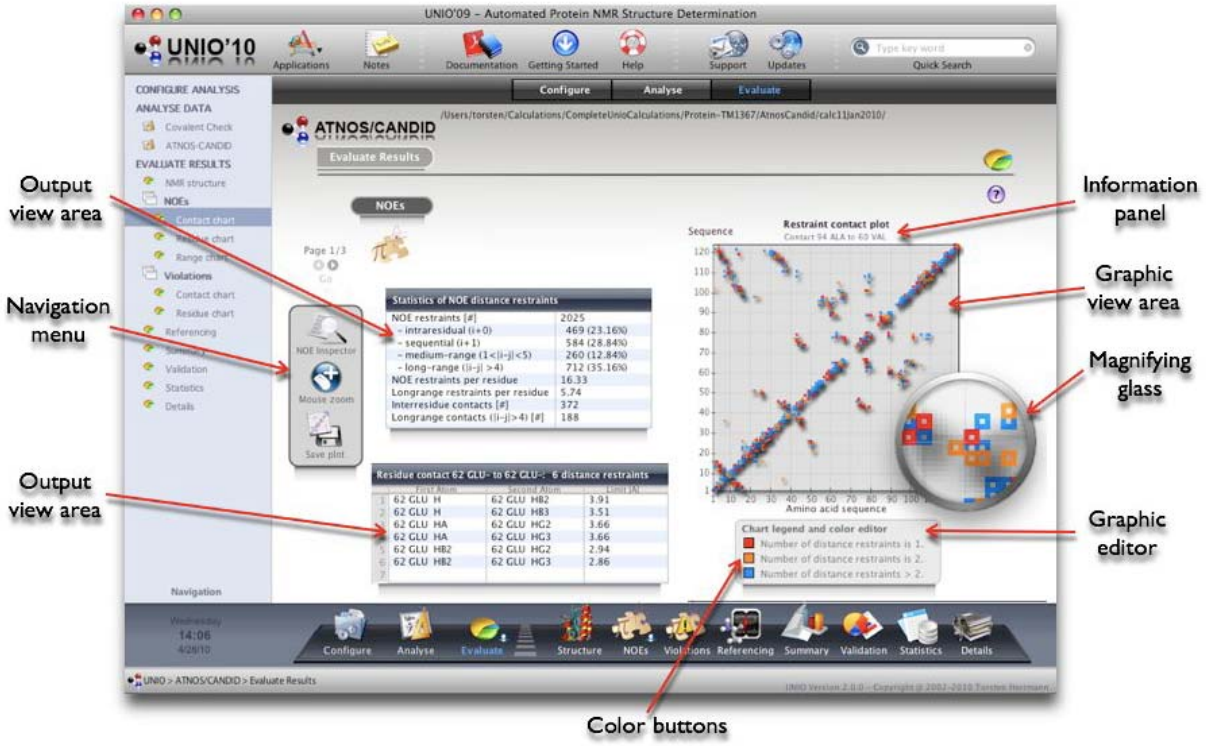
2.13.3.2.1 "Evaluate Results: NOEs: Contact plot"

The sub-section "NOEs: Contact plot" displays the set of distance restraints that was used to calculate the final NMR structure in form of a residue-residue contact plot. The *Graphic view area* allows the user to directly interact with the ATNOS/CANDID results.

The top *Output view area* provides a statistical summary for the distance restraints. To facilitate the navigation in the *Graphic view area*, a *Magnifying glass* is opened automatically as long the mouse pointer is within the *Graphic view area*. If the mouse pointer is over a residue-residue contact of the *Graphic view area*, the corresponding residue names are displayed in the *Information panel*.

If you click with the mouse on a residue-residue contact in the *Graphic view area*, the corresponding distance restraints are listed in the bottom *Output view area*.

A single mouse click in the bottom *Output view area* allows you to highlight an individual distance restraint, and a double mouse click will launch the UNIO Restraint Inspector that will provide you with all information about the respective distance restraint.



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: NOEs: Contact plot".

The *Graphic editor* allows you to change the color-presentation of the chart. A mouse click on the *Color buttons* will launch the operating systems color palette window.

The *Navigation menu* serves three purposes: First, you can open the complete set of distance restraints with the UNIO Restraint Inspector; second, you can change the zoom factor used for the *Magnifying glass*; and third, you can save the residue-residue contact chart in PNG file format with grey or transparent background.

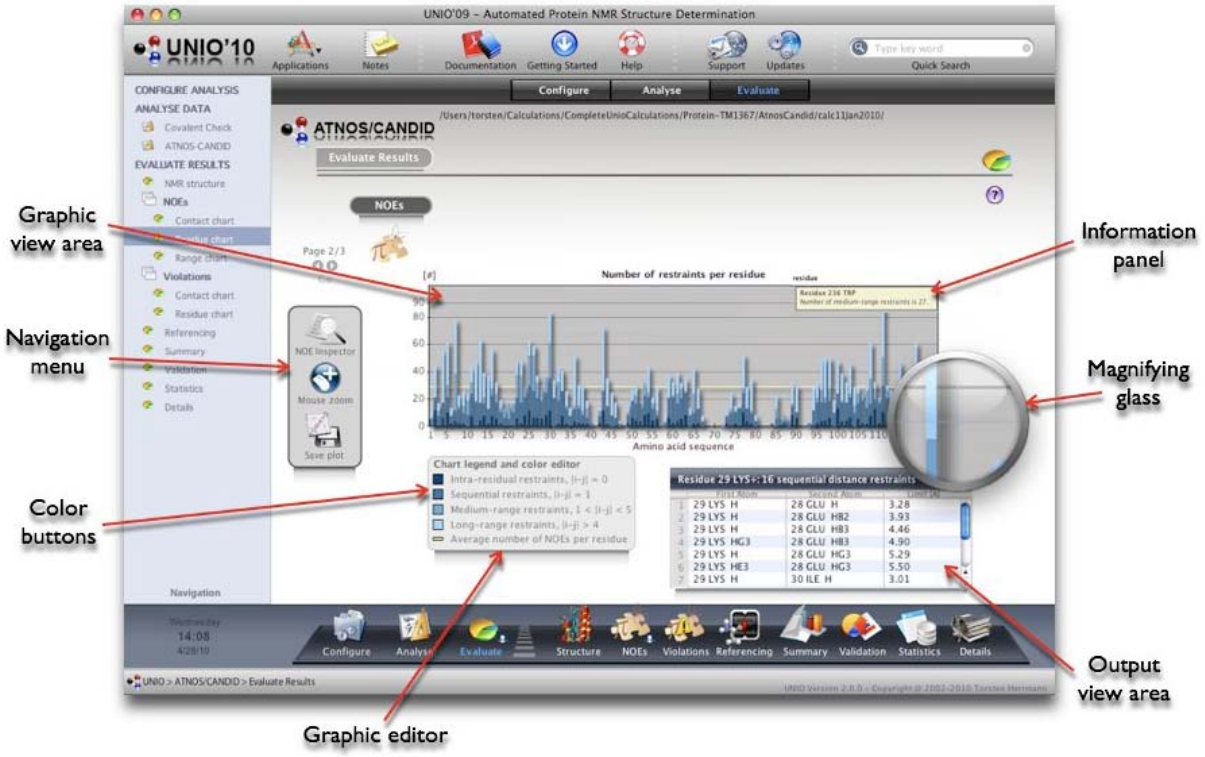
2.13.3.2.2 "Evaluate Results: NOEs: Residue plot"

The sub-section "NOEs: Residue plot" displays the set of distance restraints that was used to calculate the final NMR structure as a function of the amino acid sequence and color-coded according to the class of distance restraint. The *Graphic view area* allows the user to directly interact with the ATNOS/CANDID results.

To facilitate the navigation in the *Graphic view area*, a *Magnifying glass* is opened automatically as long the mouse pointer is within the *Graphic view area*. If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding residue name, the number and class of distance restraints are displayed in the *Information panel*.

If you click with the mouse on a chart bar in the *Graphic view area*, the corresponding distance restraints are listed in the *Output view area*.

A single mouse click in the *Output view area* allows you to highlight an individual distance restraint, and a double mouse click will launch the UNIO Restraint Inspector that will provide you with all information about the respective distance restraint.



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: NOEs: Residue plot".

The *Graphic editor* allows you to change the color-presentation of the chart. A mouse click on the *Color buttons* will launch the operating systems color palette window.

The *Navigation menu* serves three purposes: First, you can open the complete set of distance restraints with the UNIO Restraint Inspector; second, you can change the zoom factor used for the *Magnifying glass*; and third, you can save the residue chart in PNG file format with grey or transparent background.

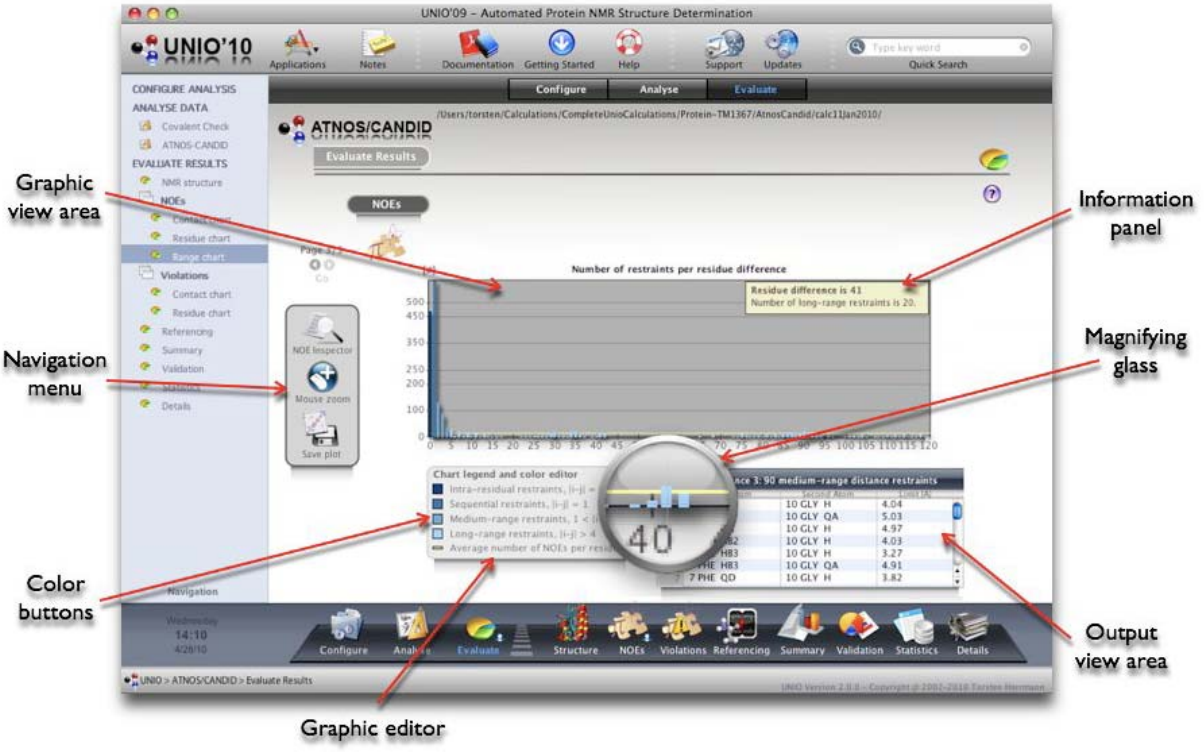
2.13.3.2.3 "Evaluate Results: NOEs: Range plot"

The sub-section "NOEs: Residue plot" displays the set of distance restraints that was used to calculate the final NMR structure as a function of the residue difference. The *Graphic view area* allows the user to directly interact with the ATNOS/CANDID results.

To facilitate the navigation in the *Graphic view area*, a *Magnifying glass* is opened automatically as long the mouse pointer is within the *Graphic view area*. If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding residue difference, the number and class of distance restraints are displayed in the *Information panel*.

If you click with the mouse on a chart bar in the *Graphic view area*, the corresponding distance restraints are listed in the *Output view area*.

A single mouse click in the bottom *Output view area* allows you to highlight an individual distance restraint, and a double mouse click will launch the UNIO Restraint Inspector that will provide you with all information about the respective distance restraint.



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: NOEs: Range plot".

The *Graphic editor* allows you to change the color-presentation of the chart. A mouse click on the *Color buttons* will launch the operating systems color palette window.

The *Navigation menu* serves three purposes: First, you can open the complete set of distance restraints with the UNIO Restraint Inspector; second, you can change the zoom factor used for the *Magnifying glass*; and third, you can save the residue chart in PNG file format with grey or transparent background.

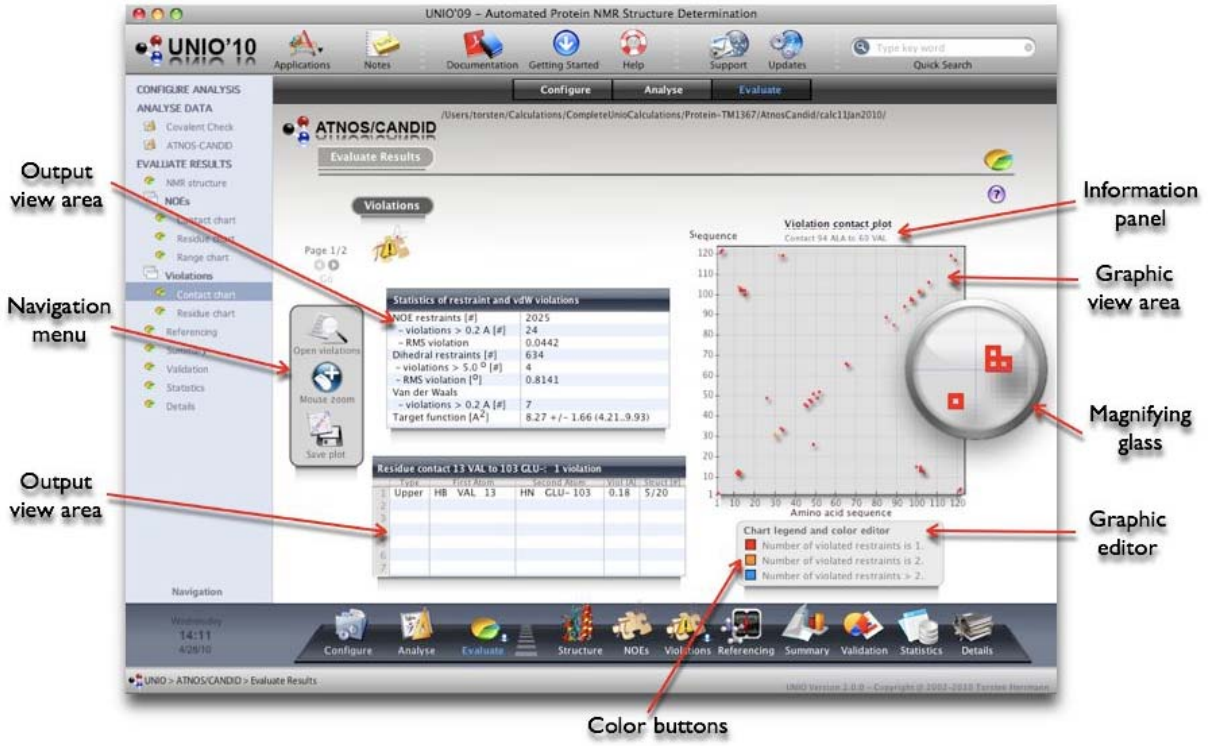
2.13.3.3 "Evaluate Results: Violations"

The sub-section "Violations" informs about the conformational restraints that are violated in the final NMR structure bundle. It consists of two further sub-sections "Contact plot" and "Residue plot". The main UNIO light-blue colored *Navigation menu* allows you to browse between these sub-sections of "Violations".

2.13.3.3.1 "Evaluate Results: Violations: Contact plot"

The sub-section "Violations: Contact plot" displays the set of distance restraints that are violated in the final NMR structure in form of a residue-residue contact plot. The *Graphic view area* allows the user to directly interact with the ATNOS/CANDID results.

The top *Output view area* provides a statistical summary for the violated conformational restraints. To facilitate the navigation in the *Graphic view area*, a *Magnifying glass* is opened automatically as long the mouse pointer is within the *Graphic view area*. If the mouse pointer is over a residue-residue contact of the *Graphic view area*, the corresponding residue names are displayed in the *Information panel*.



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: Violations: Contact plot".

If you click with the mouse on a residue-residue contact in the *Graphic view area*, the corresponding violated conformational restraints are listed in the bottom *Output view area*.

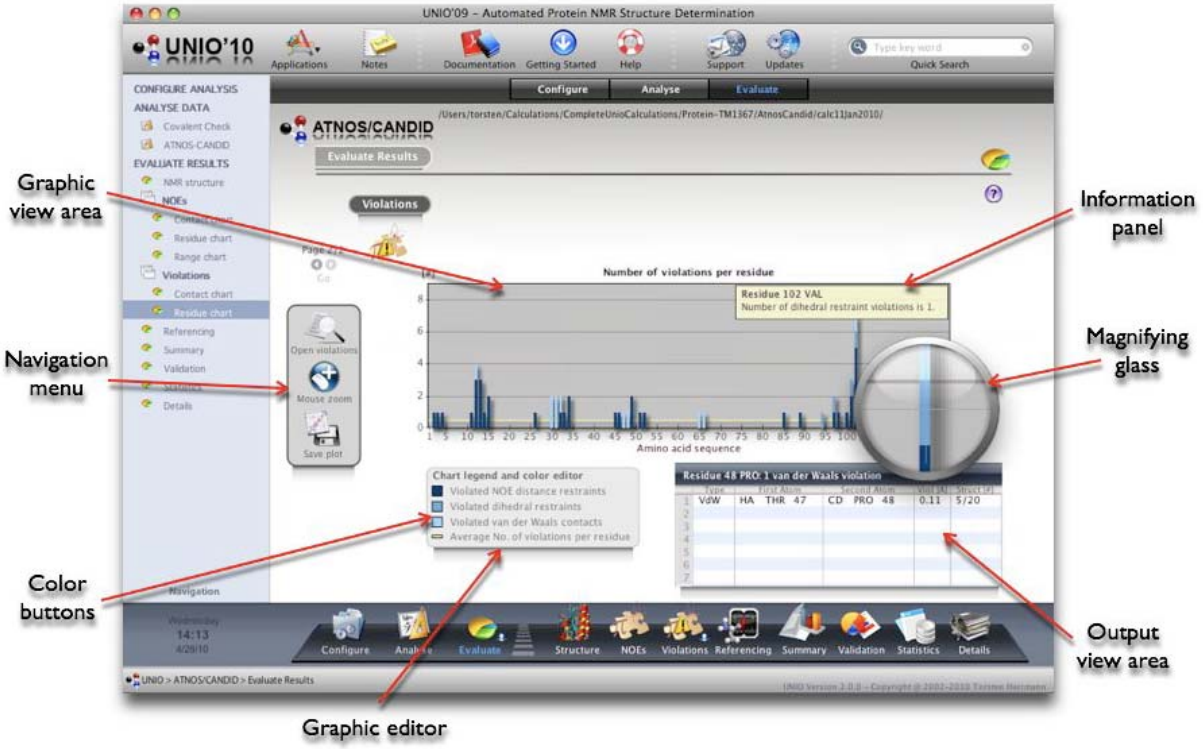
The *Graphic editor* allows you to change the color-presentation of the chart. A mouse click on the *Color buttons* will launch the operating systems color palette window.

The *Navigation menu* serves three purposes: First, you can open the complete listing of violated conformational restraints in the UNIO File Editor; second, you can change the zoom factor used for the *Magnifying glass*; and third, you can save the residue-residue contact chart in PNG file format with grey or transparent background.

2.13.3.3.2 "Evaluate Results: Violations: Residue plot"

The sub-section "Violations: Residue plot" displays the set of conformational restraints that are violated in the final NMR structure as a function of the amino acid sequence and color-coded according to the class of conformational restraint. The *Graphic view area* allows the user to directly interact with the ATNOS/CANDID results.

To facilitate the navigation in the *Graphic view area*, a *Magnifying glass* is opened automatically as long the mouse pointer is within the *Graphic view area*. If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding residue name, the number and class of violated conformational restraint are displayed in the *Information panel*.



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: Violations: Residue plot".

If you click with the mouse on a chart bar in the *Graphic view area*, the corresponding violated conformational restraints are listed in the *Output view area*.

A single mouse click in the *Output view area* allows you to highlight an individual distance restraint, and a double mouse click will launch the UNIO Restraint Inspector that will provide you with all information about the respective distance restraint.

The *Graphic editor* allows you to change the color-presentation of the chart. A mouse click on the *Color buttons* will launch the operating systems color palette window.

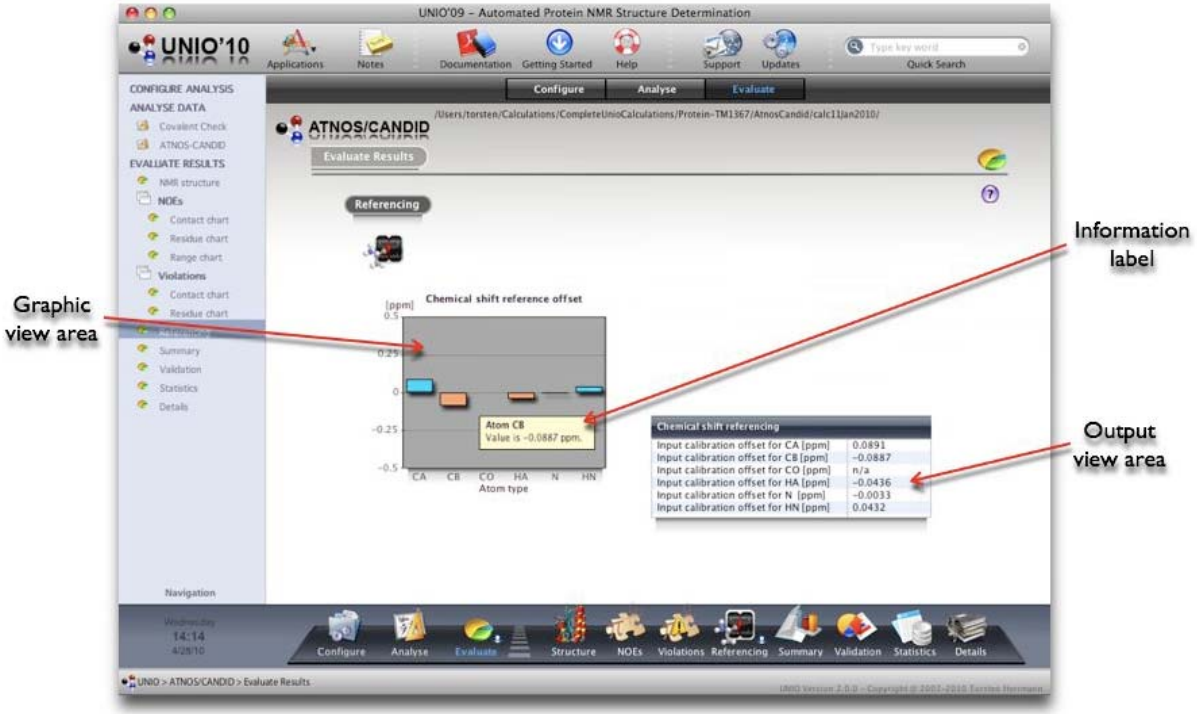
The *Navigation menu* serves three purposes: First, you can open the complete listing of violated conformational restraints with the UNIO File Editor; second, you can change the zoom factor used for the *Magnifying glass*; and third, you can save the residue chart in PNG file format with grey or transparent background.

2.13.3.4 "Evaluate Results: Referencing"

The sub-section "Referencing" informs about the chemical shift referencing offset of all backbone atoms. The *Graphic view area* allows the user to directly visualize the respective referencing offset.

If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding atom name and the value of the reference offset in ppm are displayed in the *Information panel*.

The Output view area summarizes the chart information in table presentation.



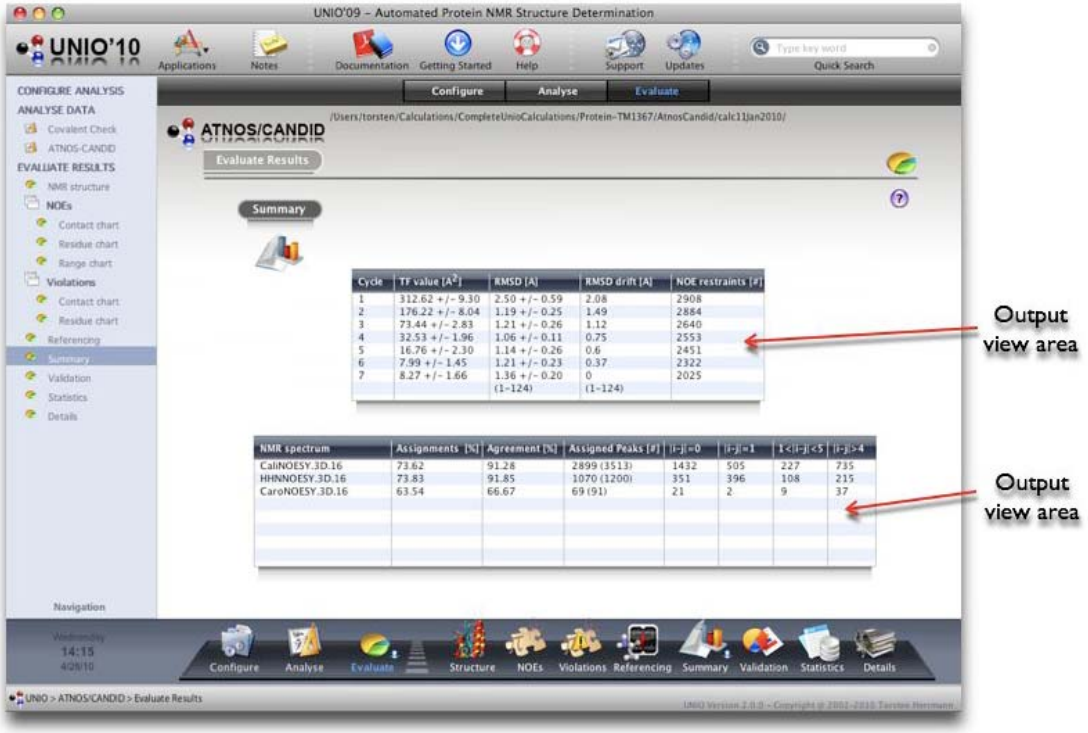
UNIO Application window for ATNOS/CANDID, section "Evaluate Results: Referencing".

2.13.3.5 "Evaluate Results: Summary"

The sub-section "Summary" consists of two tables providing information about input and ATNOS/CANDID output.

The upper table summarizes the evolution of characteristic parameters from ATNOS/CANDID cycle 1 to cycle 7, such as the target function value, the RMSD and the RMSD-Drift.

The lower table reports for each input spectrum about the assignment completeness, the chemical shift agreement and the number of assigned cross peaks. The number of assigned cross peaks is further split into four classes, intraresidual, sequential, short-range and long-range cross peak assignments.

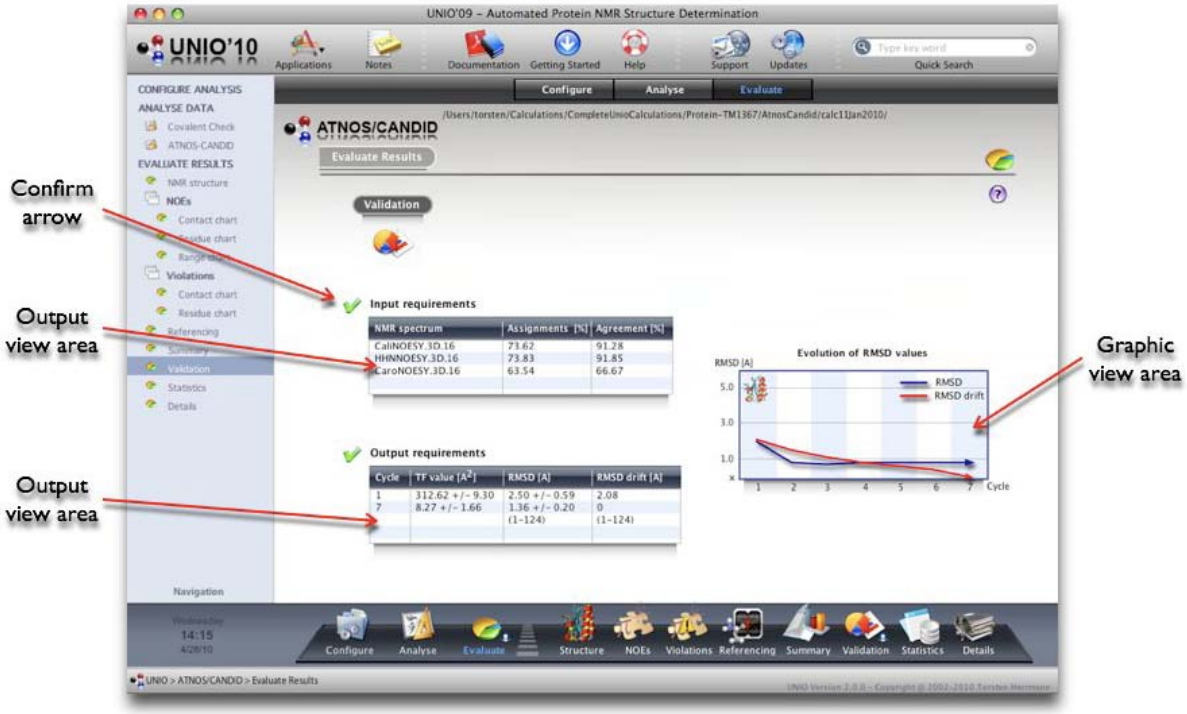


UNIO Application window for ATNOS/CANDID, section "Evaluate Results: Summary".

2.13.3.6 "Evaluate Results: Validation"

The sub-section "Validation" reports about characteristic ATNOS/CANDID parameters that are used to assess the correctness of the resulting 3D structure.

The validation is split into "Input requirements" and "Output criteria" as listed in the upper and lower table. If the input requirements and output criteria are met, *Confirm arrows* are shown to signal acceptance of the result. The evolution of the RMSD and the RMSD-Drift is depicted in the *Graphic view area*.



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: Summary".

Input requirement 1: The chemical shift list must contain more than 90% of the non-labile and backbone amide ^{1}H chemical shifts. If 3D heteronuclear-resolved [^{1}H,^{1}H]-NOESY spectra are used, more than 90% of the ^{15}N and/or ^{13}C chemical shifts must also be available. The chemical shift values must be correct.

Important: The criterion 1 is more relaxed if the resonance assignments have been obtained with ATNOS/ASCAN. ATNOS/ASCAN is only assigning conformationally meaningful side-chain atoms that are involved in many NOE signals. For example, side-chain atoms close to the protein's surface are hardly assigned. This intrinsic feature of ATNOS/ASCAN leads to a lower completeness of the resonance assignments. However, since only assignments for side-chain atoms are missing that do not contribute to meaningful NOE signals, the criterion 1 can be relaxed to 70-75% of atoms that must be assigned.

Input requirement 2: ATNOS/CANDID must validate NOE signals for at least 85% of all pairwise combinations of protons i and j (i is not equal j) for which sequence-specific resonance assignments are available, and which have covalent structure-imposed upper distance limits shorter than a user-given threshold. For example, the maximal distance between the intraresidual atoms Hα and HN is always smaller than 5 Å, and therefore must lead to an observable signal in the NOESY spectrum.

Please keep in mind that the second requirement only makes a statement about the calibration of the chemical shift list to the NOESY spectrum. A low percentage of validated covalent NOE contacts is usually indicative of insufficient chemical shift agreement between the chemical shift list and the NOESY spectrum. Whenever the difference between the NOE cross peak positions and the chemical shift of a given atom pair is larger than the predetermined tolerance range, then ATNOS/CANDID will fail to identify the corresponding NOE cross peak. In such a case, a higher percentage can usually be obtained by re-calibrating the chemical shift list using a constant value for all hydrogen atoms.

Output criterion 1: The average final ATNOS/CANDID target function value for the bundle of conformers used to represent the structure from the first ATNOS/CANDID cycle should be below 150 $Å^2$, and the corresponding value for the last ATNOS/CANDID cycle should be below 2-3 $Å^2$.

Output criterion 2: The average backbone RMSD to the mean coordinates for the structured parts of the polypeptide chain should be below 3.0 Å for the structure bundle of ATNOS/CANDID cycle 1.

Output criterion 3: The RMSD-Drift between the mean coordinates of the k^{th} and the last ATNOS/CANDID cycles calculated for the backbone atoms of the structured parts of the polypeptide chain should be in the order of the RMSD value of the k^{th} cycle.

Definition RMSD-Drift: RMSD-Drift is calculated as the RMSD between the mean coordinates of the bundles of conformers obtained after the k^{th} and the seventh ATNOS/CANDID cycles.

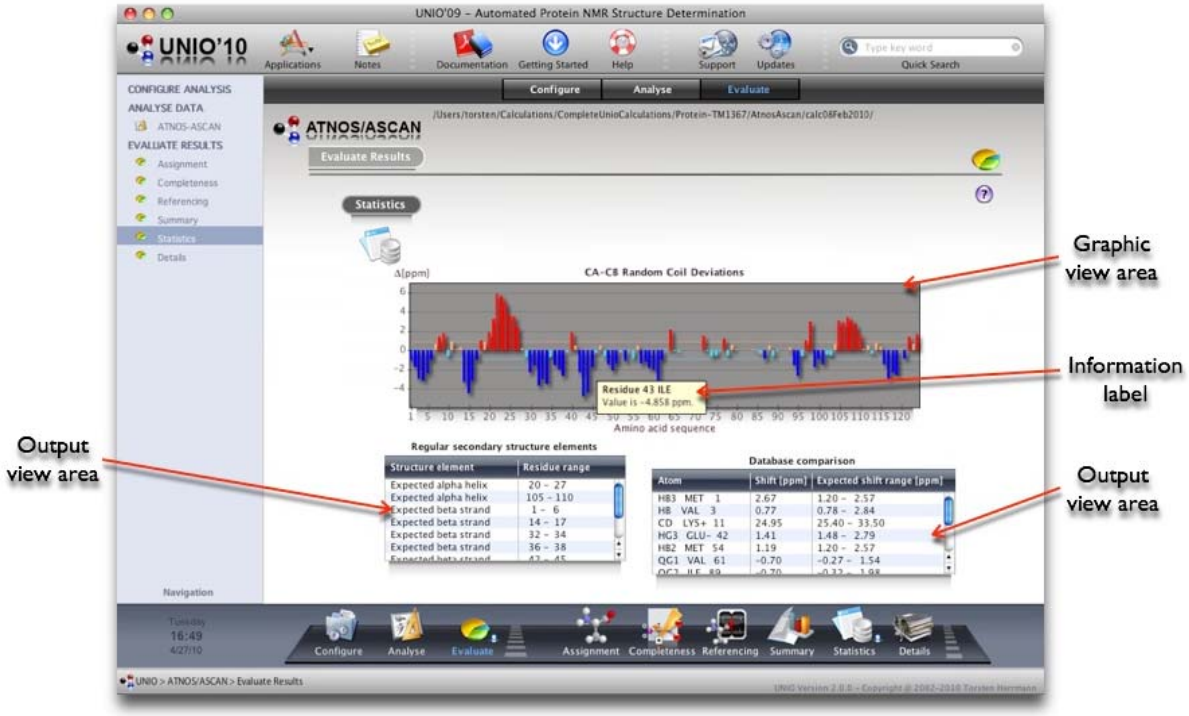
2.13.3.7 "Evaluate Results: Statistics"

The sub-section "Statistics" compares the input resonance assignments with average chemical shift values of proteins.

The *Graphic view area* plots the CA-CB random coil deviation along the protein sequence. If the mouse pointer is over a residue chart bar of the *Graphic view area*, the corresponding residue name and the value of the random coil deviation in ppm are displayed in the *Information panel*. Helical and strand secondary structure elements are color-coded in red and blue respectively.

Consecutive amino acid residues with a chemical shift deviation above or below a certain threshold (orange and cyan lines in the *Graphic view area*) are listed as regular secondary structure elements in the table on the left side.

The table on the right side reports about chemical shift assignments that are unusual with respect to average chemical shift values found in proteins.



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: Statistics".

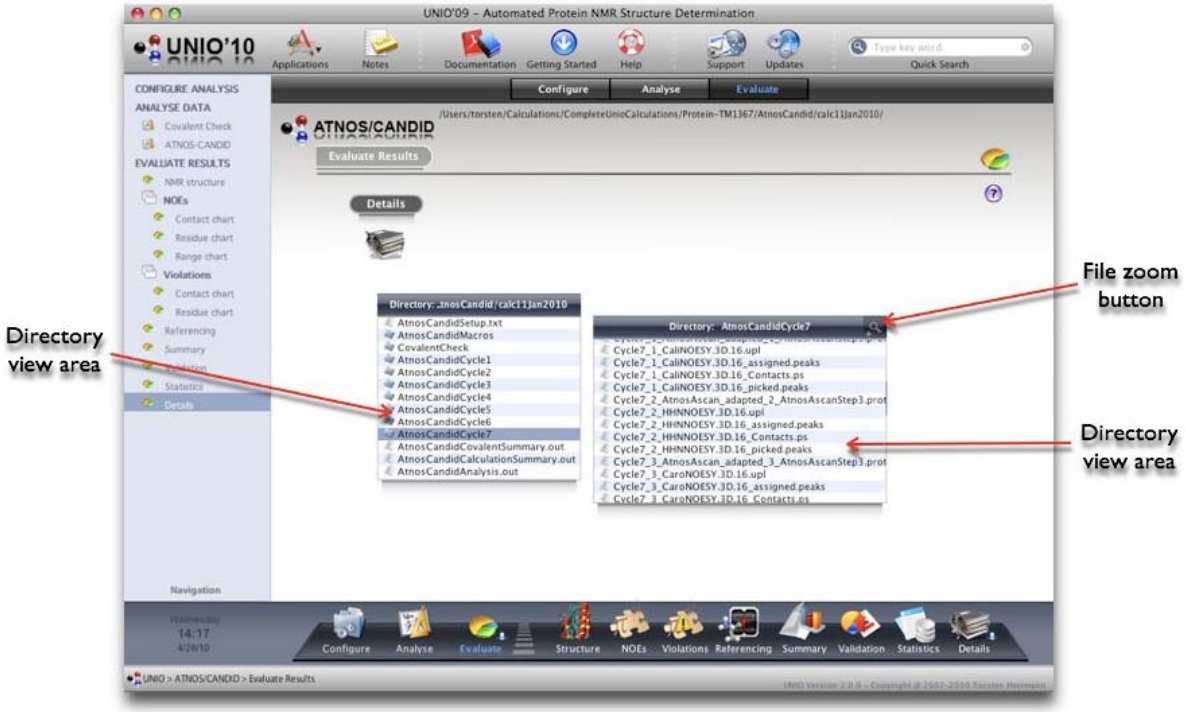
2.13.3.8 "Evaluate Results: Details"

The sub-section "Details" allows you to browse through all ATNOS/CANDID output directories and files.

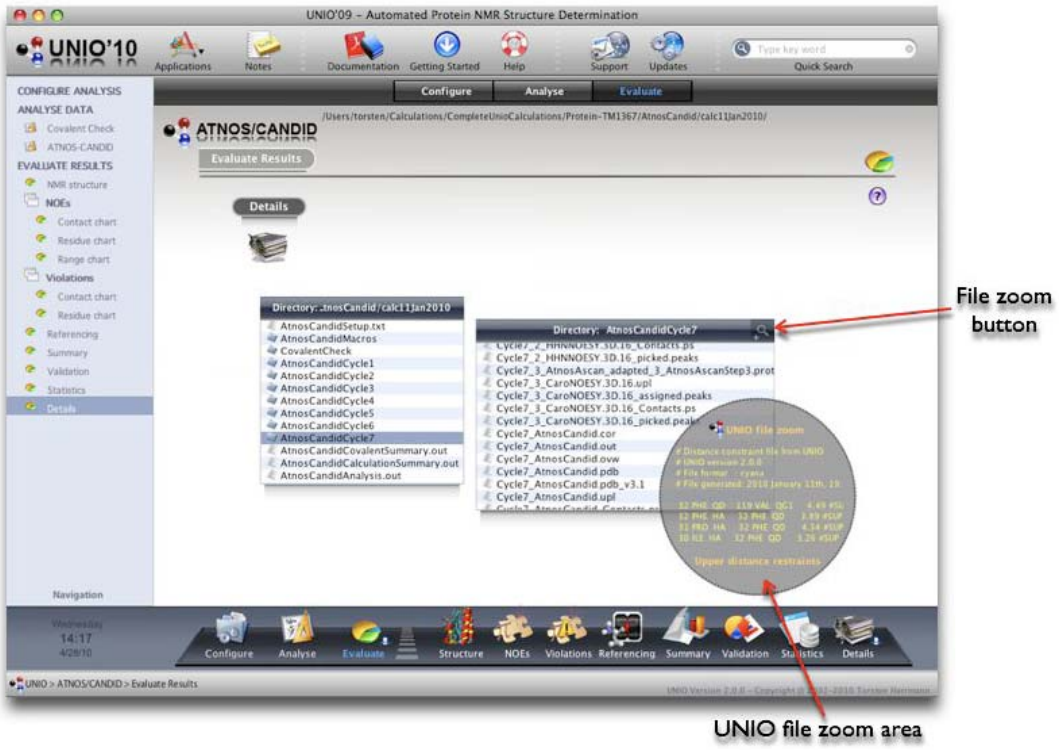
A single mouse click displays the content of a directory in the right table, while double mouse click opens a file in the UNIO File Editor.

If your mouse pointer is over an entry of the left or right table and has been longer than 1 second in the *Output view area*, *UNIO File zoom* gets activated. The *UNIO file zoom* allows you to see the file content without explicitly opening the file. In addition, *UNIO File zoom* gives you at the bottom of the *File zoom area* information about the type of the file. You can use the *File zoom button* at the top of the left table to switch *UNIO File zoom* on and off.

Tip: Use the *UNIO File zoom* function to quickly see the content of a file and to learn about the type of file you are looking at.



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: Details".



UNIO Application window for ATNOS/CANDID, section "Evaluate Results: Details". File zoom active.

2.14 UNIO - Configuration windows

The UNIO Configuration windows allows you to provide all required input data to perform UNIO NMR data analysis either with MATCH or ATNOS/ASCAN or ATNOS/CANDID.

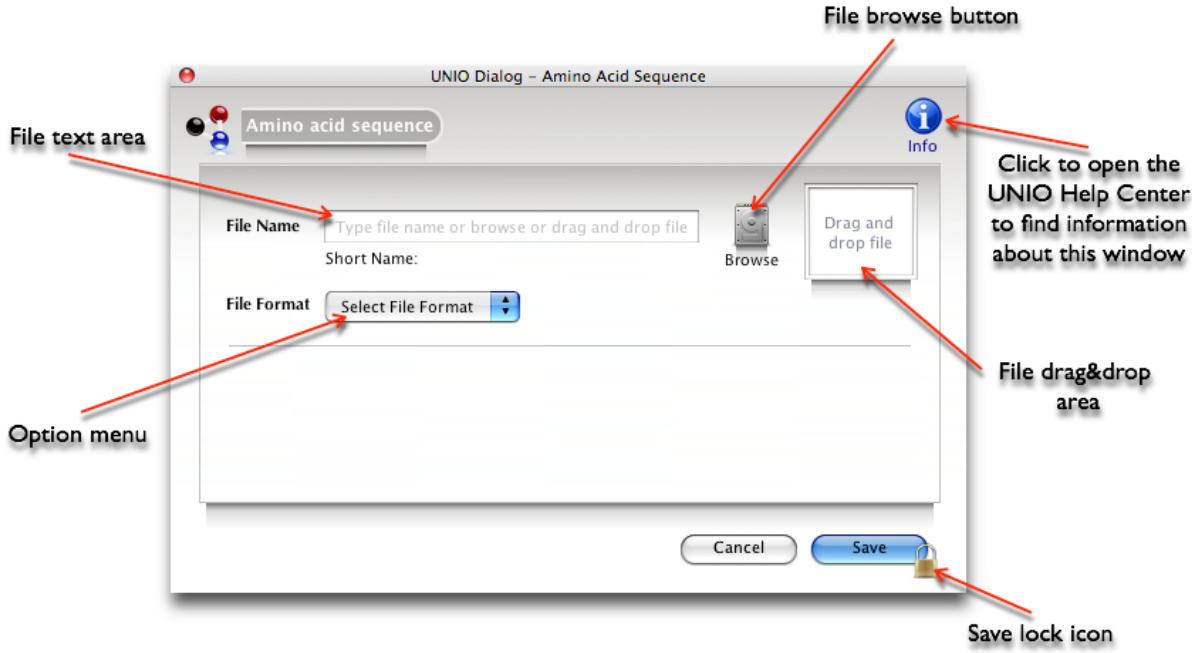
2.14.1 Configuration window "Amino Acid Sequence"

The "Amino Acid Sequence" configuration window allows you to provide the file name of your protein's sequence and the corresponding file format.

The amino acid sequence is the order in which amino acid residues, connected by peptide bonds, lie in the polypeptide chain in peptides and proteins.

There are three ways to provide your amino acid sequence file, either by typing the complete path of the file in the *File text area*, or by using the system's file browser that is opened by clicking on the *File browse button*, or by dropping the file into the *File drag&drop area*. You can select the corresponding file format using the *Option menu* button.

As long as the input is incomplete, the *Save lock icon* is shown at the right bottom of the window. After completion of the input, you can save your configuration by pressing the *Save* button.



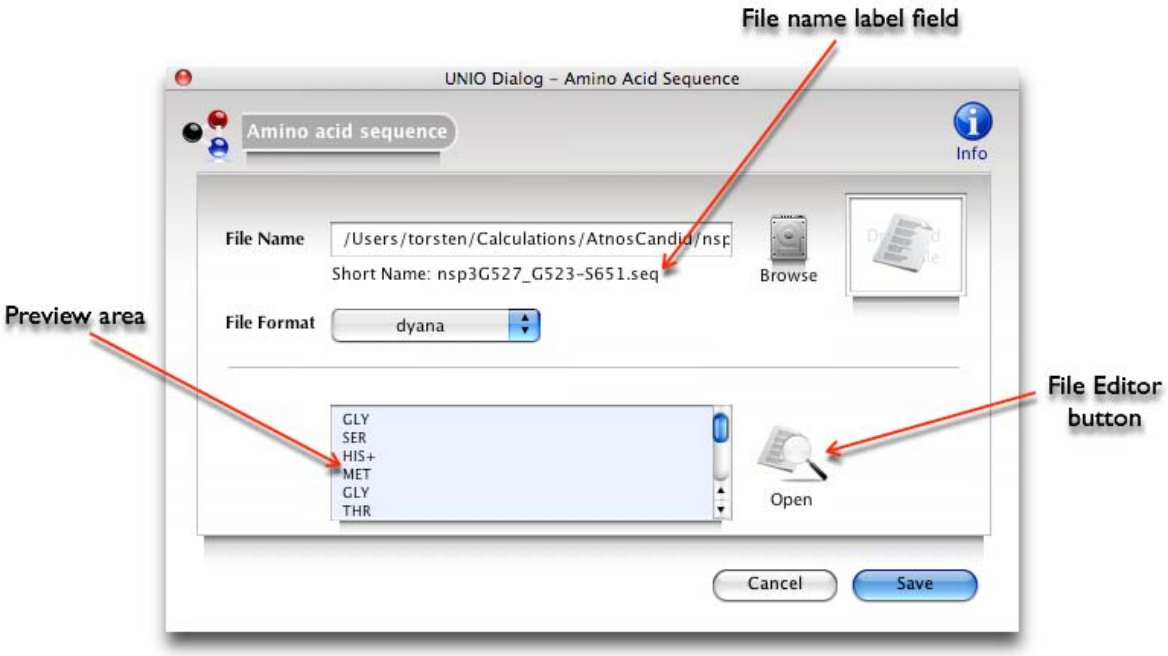
UNIO Configuration window "Amino Acid Sequence".

Supported file formats are BMRB, CNS, CYANA, DYANA, XEASY, XPLOR, one letter code, three letter code, FASTA, SPARKY, NMRVIEW, ANSIG.

Caution: If you want to apply ATNOS/CANDID data analysis in combination with a MD algorithm, then the file format of the amino acid sequence must be compatible with the MD algorithm selected.

After you have provided the amino acid sequence file, the *Preview area* and the *File Editor button* will be shown. The *Preview area* shows the file content and might be helpful to verify the correctness of the file format selected.

If you want to open and edit the sequence file, then you can press the *File Editor button* that launches the UNIO File Editor.



UNIO Configuration window "Amino Acid Sequence".

Caution: For many MD algorithms, such as CYANA or CNS the maximal allowed length of any input file name including the file path is 80 characters.

Tip: At the moment when you save your input configuration, you can choose between the save options "Absolute file path" or "Relative file path". The option "Relative file path" often overcomes the above mentioned problem of too long file names for MD algorithms. Additionally the option "Relative file path" is more convenient, then you want to transfer UNIO projects between different users or between multiple computers with different directory structures.

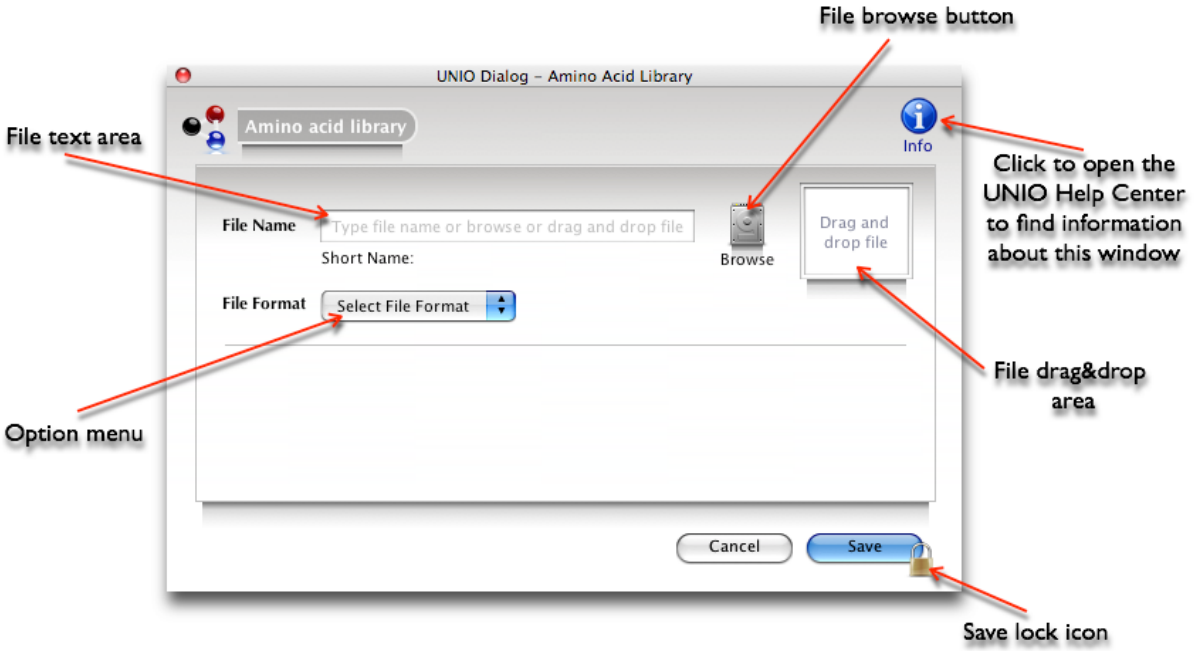
2.14.2 Configuration window "Amino Acid Library"

The "Amino Acid Library" configuration window allows you to provide the file name of your amino acid library and the corresponding file format.

The amino acid library declares the type of atoms, the atom nomenclature, the dihedral angle definition, the covalent connectivities and the standard covalent polypeptide geometry. Protein structure determination using Nuclear Magnetic Resonance (NMR) requires the use of molecular dynamics programs that incorporate both NMR experimental data and a priori knowledge about the covalent polypeptide structure of the protein. Atomic parameters for each amino acid type are encoded in libraries ("Amino acid library") used by structure calculation programs such as CYANA, CNS, XPLOR etc.

There are three ways to provide your amino acid library file, either by typing the complete path of the file in the *File text area*, or by using the system's file browser that is opened by clicking on the *File browse button*, or by dropping the file into the *File drag&drop area*. You can select the corresponding file format using the *Option menu* button.

As long as the input is incomplete, the *Save lock icon* is shown at the right bottom of the window. After completion of the input, you can save your configuration by pressing the *Save* button.



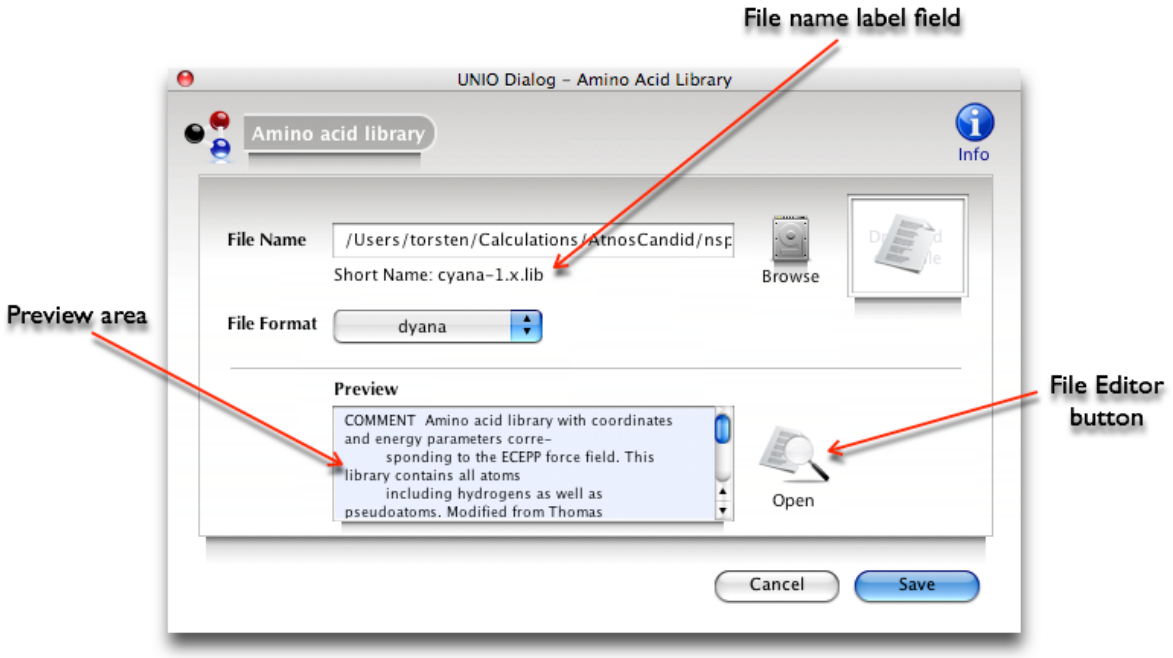
UNIO Configuration window "Amino Acid Library".

Supported file formats are AMBER, CNS, CYANA, DYANA, XPLOR, XPLOR-NIH.

Caution: If you want to apply ATNOS/CANDID data analysis in combination with a MD algorithm, then the file format of the amino acid library must be compatible with the MD algorithm selected.

After you have provided the amino acid library file, the *Preview area* and the *File Editor button* will be shown. The *Preview area* shows the file content and might be helpful to verify the correctness of the file format selected.

If you want to open and edit the amino acid library file, then you can press the *File Editor button* that launches the UNIO File Editor.



UNIO Configuration window "Amino Acid Library".

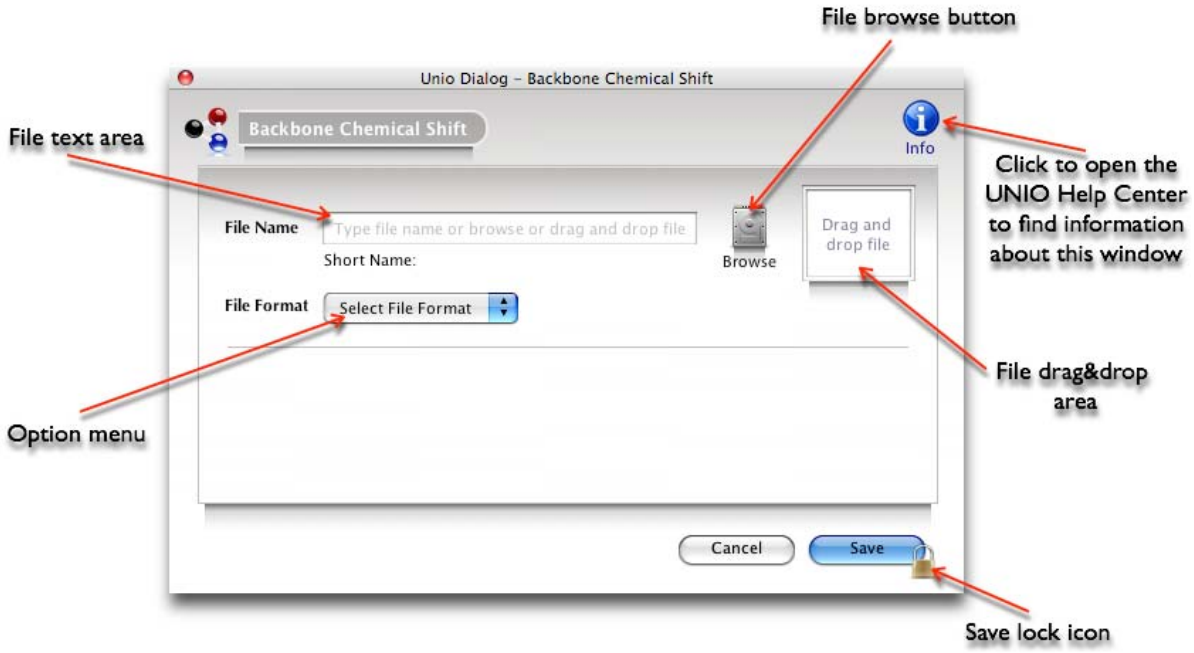
Tip: You can specify an amino acid library in the UNIO Preferences window. If you do so, then the information provided will be automatically loaded into any new UNIO project. This AutoFill feature of UNIO facilitates commonly performed configuration tasks.

Tip: For the UNIO modules MATCH and ATNOS/ASCAN, the UNIO default amino acid library "amber94.lib" is automatically chosen as amino acid library. There is no need to choose your own amino acid library for these two data analysis modules.

2.14.3 Configuration window "Backbone Chemical Shift"

The "Backbone Chemical Shift" configuration window allows you to provide the file name of backbone resonance assignments and the corresponding file format. The backbone chemical shifts represent the starting point for ATNOS/ASCAN side-chain assignment.

There are three ways to provide your backbone chemical shift file, either by typing the complete path of the file in the *File text area*, or by using the system's file browser that is opened by clicking on the *File browse button*, or by dropping the file into the *File drag&drop area*. You can select the corresponding file format using the *Option menu* button. As long as the input is incomplete, the *Save lock icon* is shown at the right bottom of the window. After completion of the input, you can save your configuration by pressing the *Save* button.

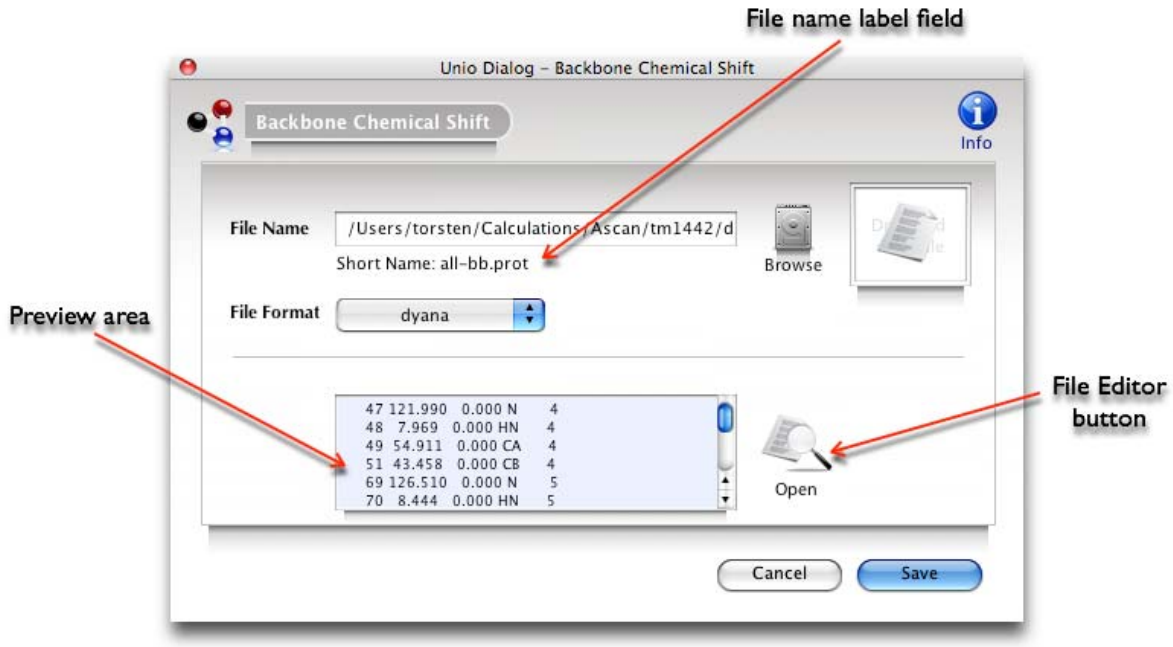


UNIO Configuration window "Backbone Chemical Shift".

Supported file formats are BMRB, CYANA, DYANA, NMRVIEW, SPARKY, XEASY.

After you have provided the amino acid sequence file, the *Preview area* and the *File Editor button* will be shown. The *Preview area* shows the file content and might be helpful to verify the correctness of the file format selected.

If you want to open and edit the backbone chemical shift file, then you can press the *File Editor button* that launches the UNIO File Editor.



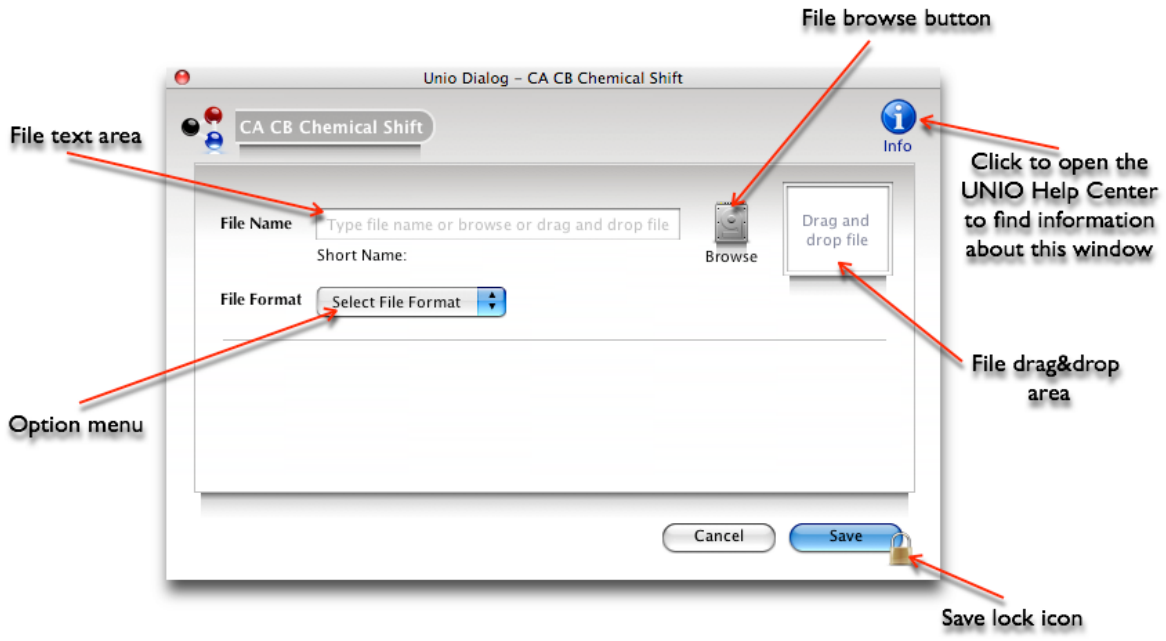
UNIO Configuration window "Backbone Chemical Shift".

2.14.4 Configuration window "CA CB Chemical Shift"

The "CA CB Chemical Shift" configuration window allows you to provide the file name of your - assumed - correctly calibrated CA and CB chemical shift file and the corresponding file format.

The CA and CB chemical shift values are evaluated to define regular secondary structure elements by comparing the difference of CA-CB chemical shifts for three consecutive residues with the corresponding random coil chemical shift values. The information about regular secondary structure elements are used for both, ATNOS signal identification and CANDID NOE assignment. Additionally, upper distance restraints and torsion angle restraints are automatically derived from this input and added as conformational restraints to the input for the MD algorithm, conventionally called "RegularSecondaryContacts.upl" and "RegularSecondaryContacts.aco", respectively.

There are three ways to provide your CA CB chemical shift file, either by typing the complete path of the file in the *File text area*, or by using the system's file browser that is opened by clicking on the *File browse button*, or by dropping the file into the *File drag&drop area*. You can select the corresponding file format using the *Option menu* button. As long as the input is incomplete, the *Save lock icon* is shown at the right bottom of the window. After completion of the input, you can save your configuration by pressing the *Save* button.



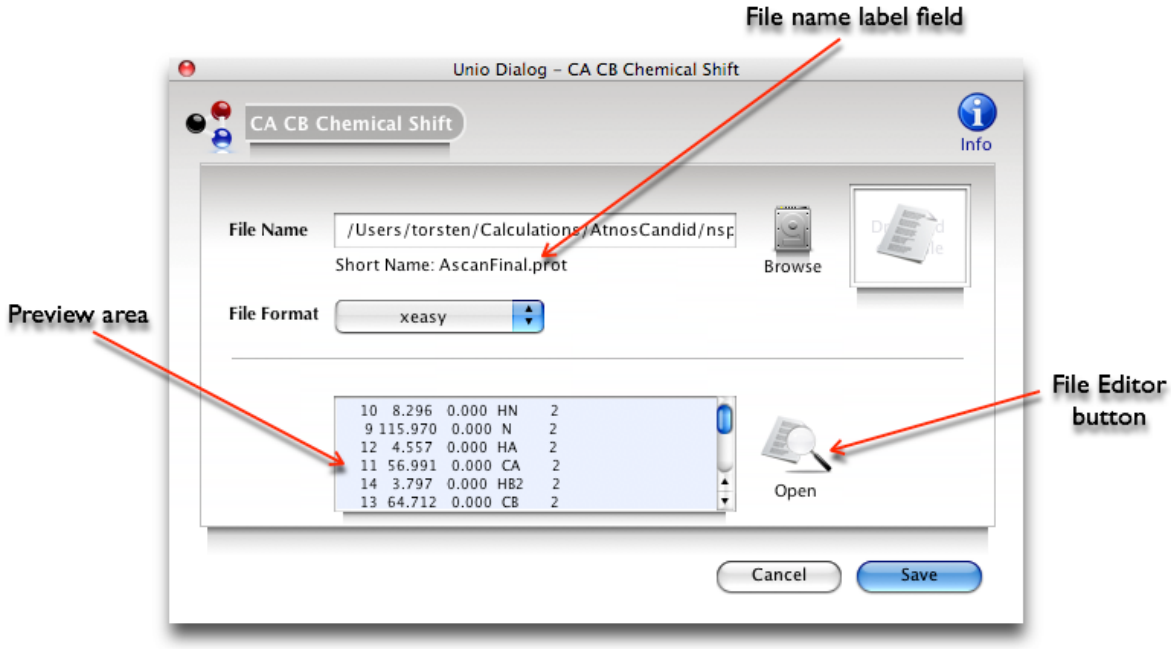
UNIO Configuration window "CA CB Chemical Shift".

Supported file formats are BMRB, CYANA, DYANA, NMRVIEW, SPARKY, XEASY.

Caution: The chemical shift values for the CA and CB atoms must be correctly calibrated. Otherwise the applied comparison using chemical shift statistics of atoms in proteins may lead to incorrect results for the reported regular secondary structure elements and automatically derived conformational restraints.

After you have provided the CA CB chemical shift file, the *Preview area* and the *File Editor button* will be shown. The *Preview area* shows the file content and might be helpful to verify the correctness of the file format selected.

If you want to open and edit the CA CB chemical shift file, then you can press the *File Editor button* that launches the UNIO File Editor.

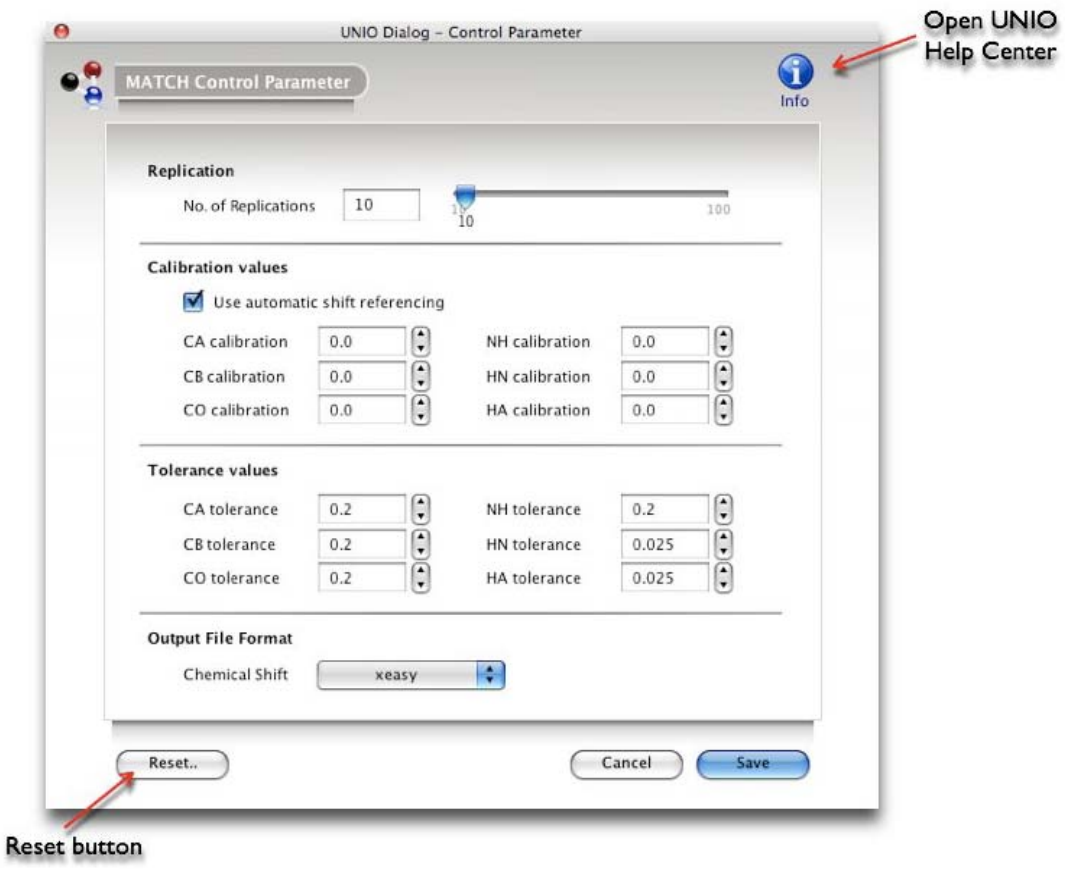


UNIO Configuration window "CA CB Chemical Shift".

2.14.5 Configuration window "MATCH Control Parameter"

The "MATCH Control Parameter" configuration window allows you to change parameter values used for MATCH backbone resonance assignment.

You can specify the number of independent MATCH replication calculations, the option of automated chemical shift referencing and/or the correct calibration offsets for individual atoms, give tolerance values used for establishing sequential connectivities of individual atoms and choose the output file format of the chemical shift list written out by UNIO-MATCH. You can use the Reset button to return to the default parameter values.



UNIO Configuration window "MATCH Control Parameter".

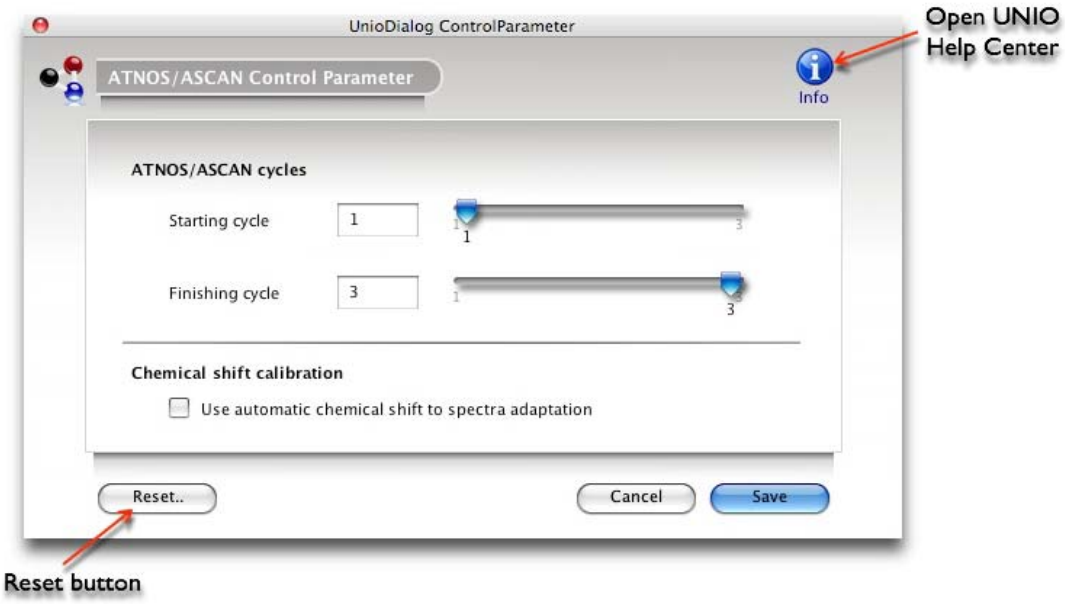
Important: It is recommend to use the default MATCH control parameters. The option "Use automatic shift referencing" is recommend to be used. MATCH is able to perform automated input peak list referencing prior to any resonance assignments made, and at the end of MATCH reference assignment to perform an exact chemical shift referencing for all backbone atoms provided in the input peak lists.

Tip: It is recommended to keep the MATCH default parameter configuration and only change values if you know exactly what you are doing!

2.14.6 Configuration window "ATNOS/ASCAN Control Parameter"

The "ATNOS/ASCAN Control Parameter" configuration window allows you to change parameter values used for ATNOS/ASCAN side-chain resonance assignment.

You can specify the starting and finishing cycle of ATNOS/ASCAN and the option for automated input backbone chemical shift to spectra adaptation. You can use the Reset button to return to the default parameter values.



UNIO Configuration window "ATNOS/ASCAN Control Parameter".

Important: It is strongly advised to correctly calibrate/reference all spectra prior to the use of ATNOS/ASCAN.

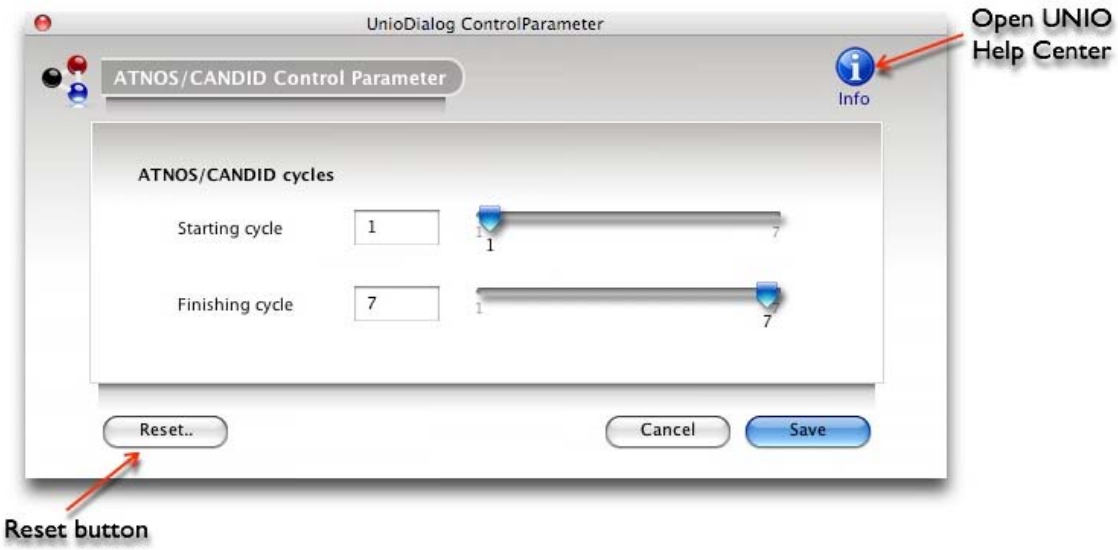
Important: It is recommend to use the default ATNOS/ASCAN control parameters. The option "Use automatic chemical shift to spectra adaption" is not necessarily recommend to be used. ATNOS/ASCAN is able to perform automated input chemical shift to spectra referencing based on expected peak pattern in NOESY spectra. However, since only the backbone assignments can be used for this purpose, the accuracy will only be 90-95%. If you use this option, then it is advised to carefully verify the automatically referenced chemical shift list which can be found in the directory "AtnosAscanStep1" along with the corresponding covalent peak lists.

Tip: It is recommended to keep the ATNOS/ASCAN default parameter configuration and only change values if you know exactly what you are doing!

2.14.7 Configuration window "ATNOS/CANDID Control Parameter"

The "ATNOS/CANDID Control Parameter" configuration window allows you to change parameter values used for ATNOS/CANDID NOE assignment.

You can specify the starting and finishing cycle of ATNOS/CANDID. You can use the Reset button to return to the default parameter values.

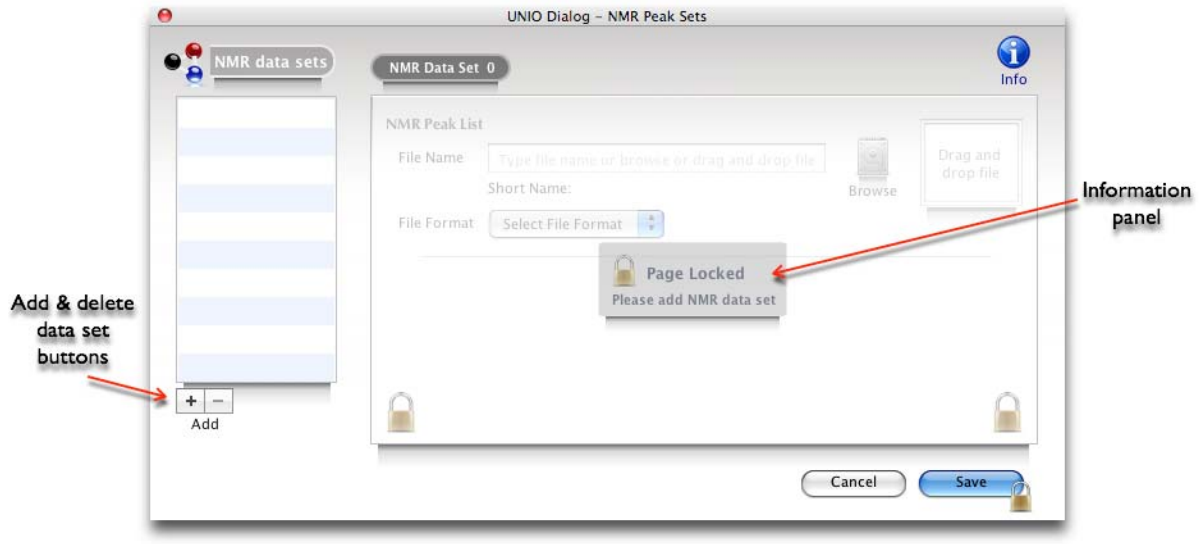


UNIO Configuration window "ATNOS/CANDID Control Parameter".

2.14.8 Configuration window "Peak Data Sets" for MATCH

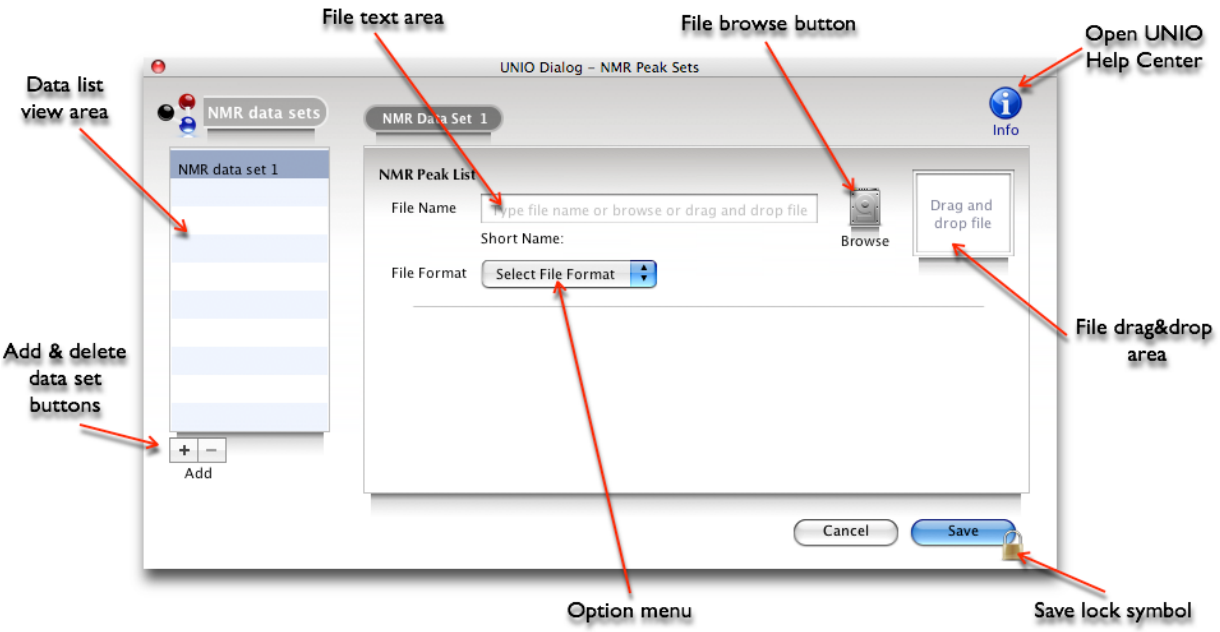
The "Peak Data Sets" configuration window allows you to provide your NMR data sets for MATCH backbone resonance assignment.

You can use the *Add&delete data set buttons* to add a new data set or to remove an existing data set. If no data set is yet provided, an *Information panel* indicates that the input data area is locked and first a data set needs to be added.



UNIO Configuration window "Peak Data Sets".

After you have added a new data set by pressing the *Add data set button*, you can continue to provide the file name of your NMR peak list file and the corresponding file format.

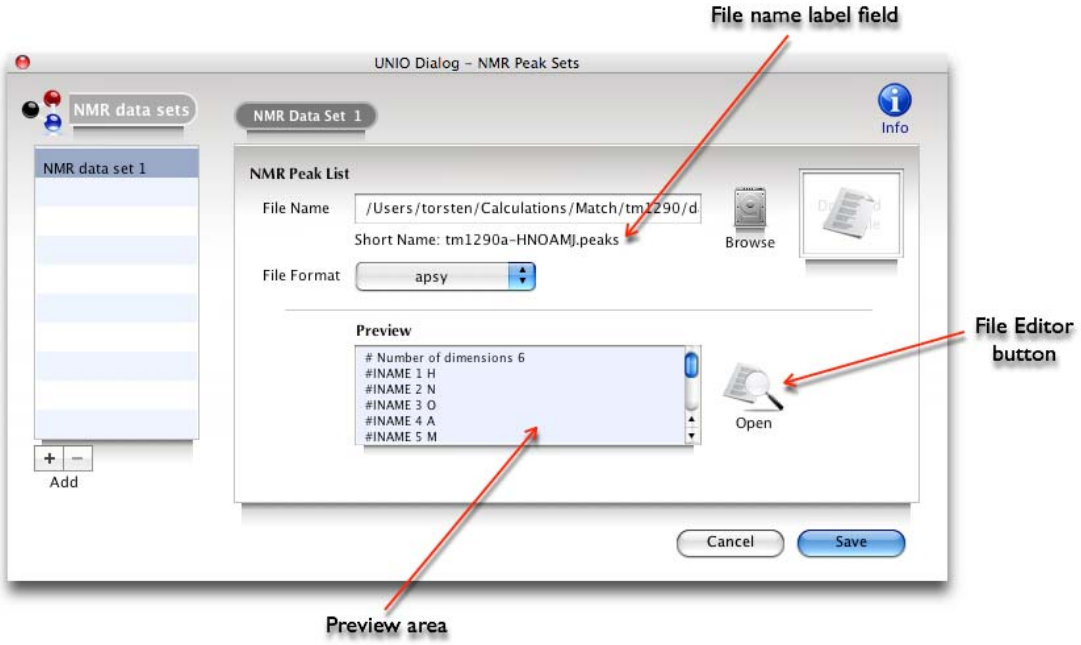


UNIO Configuration window "Peak Data Sets".

There are three ways to provide your NMR peak list file, either by typing the complete path of the file in the *File text area*, or by using the system's file browser that is opened by clicking on the *File browse button*, or by dropping the file into the *File drag&drop area*. You can select the corresponding file format using the *Option menu* button. As long as the input is incomplete, the *Save lock icon* is shown at the right bottom of the window. After completion of the input, you can save your configuration by pressing the *Save* button.

After you have provided the NMR peak list file, the *Preview area* and the *File Editor button* will be shown. The *Preview area* shows the file content and might be helpful to verify the correctness of the file format selected.

If you want to open and edit the NMR peak list file, then you can press the *File Editor button* that launches the UNIO File Editor.



UNIO Configuration window "Peak Data Sets".

Tip: The *Data list view area* at the right of the window allows you to navigate between different NMR peak lists. The currently active NMR data set is highlighted in the *Data list view area*. The selection can be changed by a single mouse click on the corresponding row.

MATCH backbone assignment can be performed starting from three types of information provided by the user: first from one or several APSY peak lists ("APSY"), second from a single consolidated generic spin-system peak list containing information about user-given extensive spin-system editing ("GSS"), and third from a partially assigned user-given single consolidated triple-resonance assignment list ("CARA_SS").

In the following, the three different types of MATCH input peak lists will be introduced.

2.14.8.1 Configuration window "Peak Data Sets: APSY Input Data"

The header of an APSY input peak list specifies the dimensionality and the type of the NMR experiment used. The experiment type is described with a 1-letter code for each dimension using the definition given in the table below.

The body of an APSY peak list is made up by the correlation peaks found in the NMR experiment and corresponds largely to the XEASY file format for NMR peak lists. The first column represents the peak number. The following m columns represent the frequency position of the signal in the m-dimensional NMR experiment.

Number of dimensions 5
#INAME 1 B
#INAME 2 A
#INAME 3 O
#INAME 4 N
#INAME 5 H
1 16.6040 49.0772 174.5171 106.8183 7.6631 1 U 0.000E+00 0.000E+00 e 0 0 0 0 0 0 0
2 28.3436 54.1580 172.9739 105.0573 6.1699 1 U 0.000E+00 0.000E+00 e 0 0 0 0 0 0 0

Example for the file format of an APSY peak list.

The APSY input peak lists must correspond to NMR APSY experiments that encode sufficient information in order to perform sequence-specific backbone resonance assignment. The above shown APSY peak list would be insufficient for the MATCH assignment process, since no information about sequential connectivities to a preceding spin system is given.

Tip: A typical set of APSY peak lists for MATCH backbone assignment consists of a 5-dimensional BAONH, a 5-dimensional ZAONH and a 4-dimensional zaNH APSY experiments.

1-letter code	Nucleus	Atom nomenclature	Relative residue position
H	^{1}H	amide proton	i
N	^{15}N	amide nitrogen	i
A	^{13}C	alpha carbon	i-1
B	^{13}C	beta carbon	i-1
O	^{13}C	carbonyl carbon	i-1
M	^{15}N	amide nitrogen	i-1
J	^{1}H	amide proton	i-1
Z	^{1}H	alpha proton	i-1
Y	^{1}H	alpha proton	i
z	^{1}H	alpha proton	i and i-1
a	^{13}C	alpha carbon	i and i-1
b	^{13}C	beta carbon	i and i-1
E	^{13}C	alpha carbon	i
D	^{13}C	beta carbon	i
I	^{1}H	aliphatic proton	i-1
K	^{1}H	aliphatic proton	i
C	^{13}C	aliphatic carbon	i-1
G	^{13}C	aliphatic carbon	i

Table for APSY experiment identifier.

Examples: The 4D-APSY-HNCOCA experiment gets the indentifier NOAH, or AONH, depending on the order the evolution periods appear in the pulse program; the 3D-APSY-HNCA is identified by NaH or aNH; the 3D-APSY-HNCO has the identifier ONH or NOH and 5D-APSY-HACACONH has ZAONH

Tip: The consolidation of the individual input peak lists into a listing of generic spin systems is done automatically by MATCH. Thus in contrast to many other backbone assignment process no manual spin system editing is required.

Tip: Instead of APSY experiments, you can also use conventional triple-resonance peak lists as input for MATCH under the assumption that you prepare the triple-resonance peak lists in analogy to the file format of APSY peak lists.

2.14.8.2 Configuration window "Peak Data Sets: GSS"

Although MATCH is designed to perform automatically the consolidation of the input peak lists into a listing of generic spin systems ("GSS"), you can also run MATCH with your own spin system edited peak list ("Consolidated peak list").

The header of a consolidated peak lists contains the necessary information about the atom names and the relative residue position for the manually edited spin systems.

The body of a consolidated peak list is made up by the correlation peaks found in the NMR experiment and corresponds largely to the XEASY file format for NMR peak lists. The first column represents the peak number. The second column represents a flag for the spin system ambiguity and should always be set to the value 1. The following m columns represent the frequency position of the signal in the m-dimensional NMR experiment.

```
HA  CA  CB  CO  HN  N  HA  CA
left  left  left  left  intra  intra  intra  intra
0.0  0.0  0.0  0.0  0.0  0.0  0.0  0.0
  1  1    4.040   61.353   37.643  175.785    7.976  127.458    4.264   52.654
  2  1    5.045   60.566   35.280  175.474    9.047  127.574    4.663   54.580
```

Example for the file format of a consolidated peak list ("GSS").

2.14.8.3 Configuration window "Peak Data Sets: CARA_SS"

MATCH also provides support for CARA edited partially assigned spin systems ("CARA_SS").

The accepted file format for a CARA edited consolidated peak lists looks as shown below.

```
# Number of dimensions 3
#INAME 1 HN
#INAME 2 C13
#INAME 3 N15
   1   8.674  45.497 121.055 0 U        0.000e+000 0.00e+000 -   0  212  982  213 0
# CB-1 32
   2   8.674  57.235 121.055 0 U        0.000e+000 0.00e+000 -   0  212  214  213 0
# CA 32
   3   8.674  29.683 121.055 0 U        0.000e+000 0.00e+000 -   0  212  216  213 0
# CB 32
```

```
   4   8.674  56.980 121.055 0 U        0.000e+000 0.00e+000 -   0  212  215  213 0
# CA-1 32
```

Example for the file format of a CARA partially assigned spin system peak list ("CARA_SS").

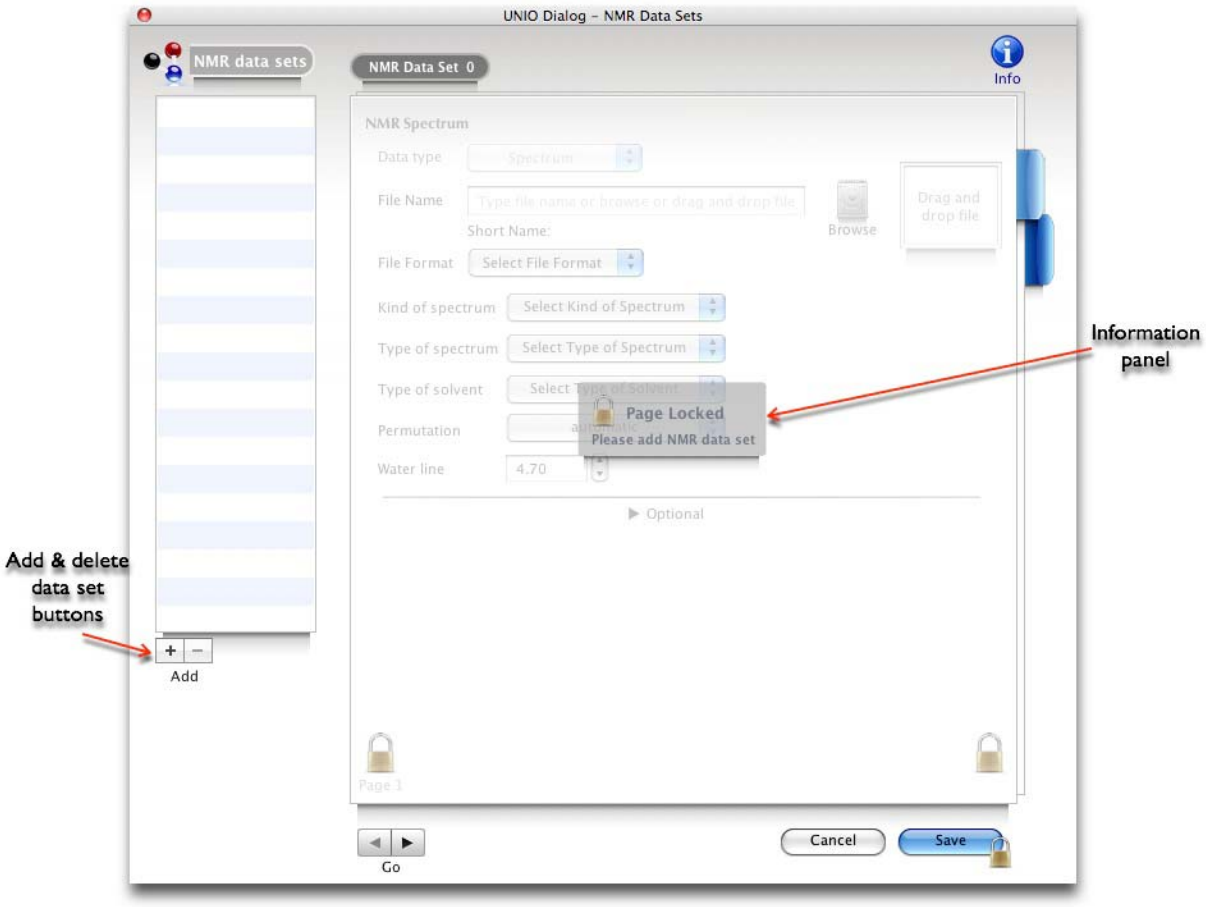
Tip: The default parameters of a MATCH backbone assignment calculation are chosen for APSY-type input data. If you are using a CARA edited spin system list, then the tolerance values used for establishing sequential shift matching may need to be increased.

2.14.9 Configuration window "NMR Data Sets" for ATNOS/ASCAN

The "NMR Data Sets" configuration window allows you to provide your set of NMR spectra as input for ATNOS/ASCAN side-chain resonance assignment.

You can use the *Add&delete data set buttons* to add a new NMR data set or to remove an existing NMR data set. If no data set is yet provided, an *Information panel* indicates that the input data area is locked and first a data set needs to be added.

This configuration window consists of two pages or sheets. The first page collects all required input data, while the second page offers advanced and/or optional input features. The *Page buttons* at the bottom of the window or the *Page tabs* on the left side of this window allows you to quickly navigate between these two sheets.



UNIO Configuration window "NMR Data Sets" for ATNOS/ASCAN.

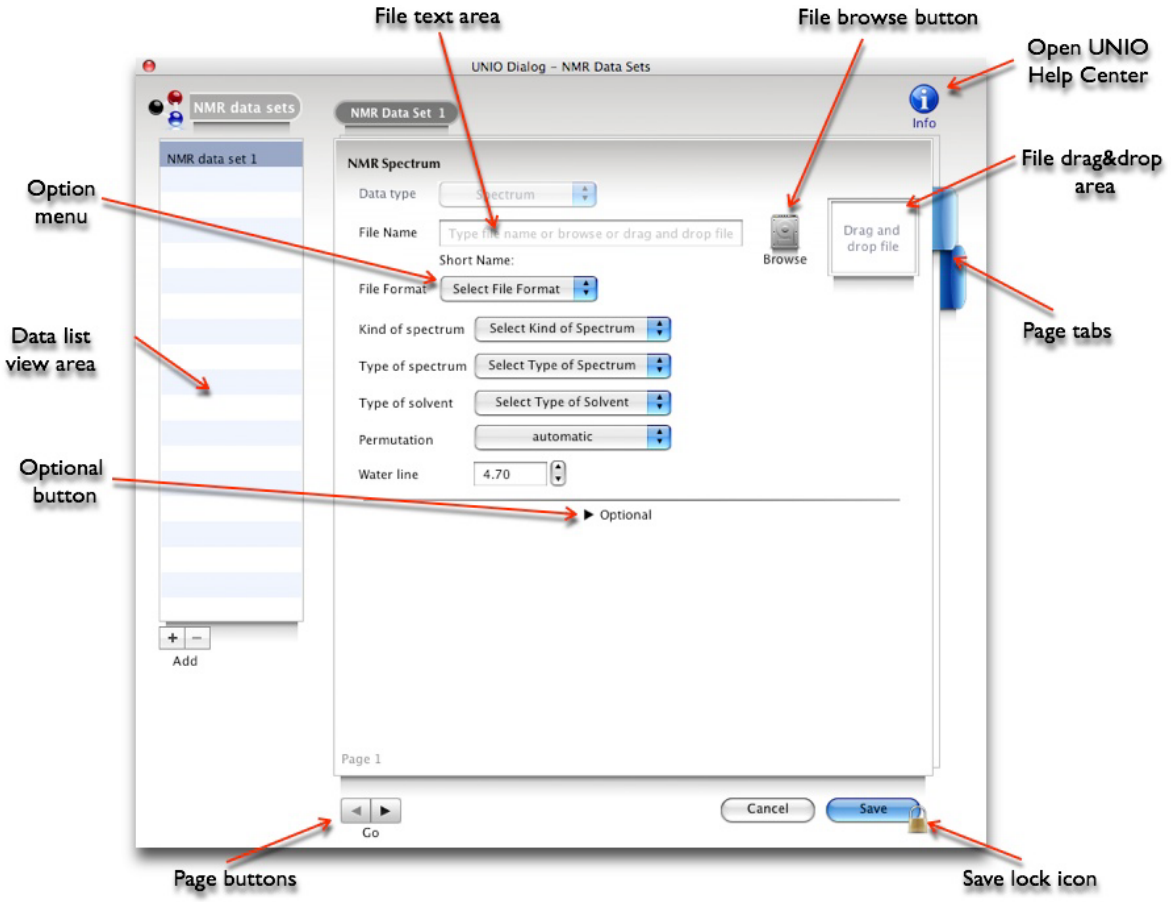
After you have added a new data set by pressing the *Add data set button*, you can continue to provide the file name of your NMR spectrum file the corresponding file format of the spectrum and additional information about the spectrum as the kind of spectrum, the type of spectrum, the type of solvent used and the value of the waterline.

There are three ways to provide your NMR spectrum file, either by typing the complete path of the file in the *File text area*, or by using the system's file browser that is opened by clicking on the *File browse button*, or by dropping the file into the *File drag&drop area*.

You can use the corresponding *Option menus* to choose between different possibilities for the above-mentioned spectrum-specific information.

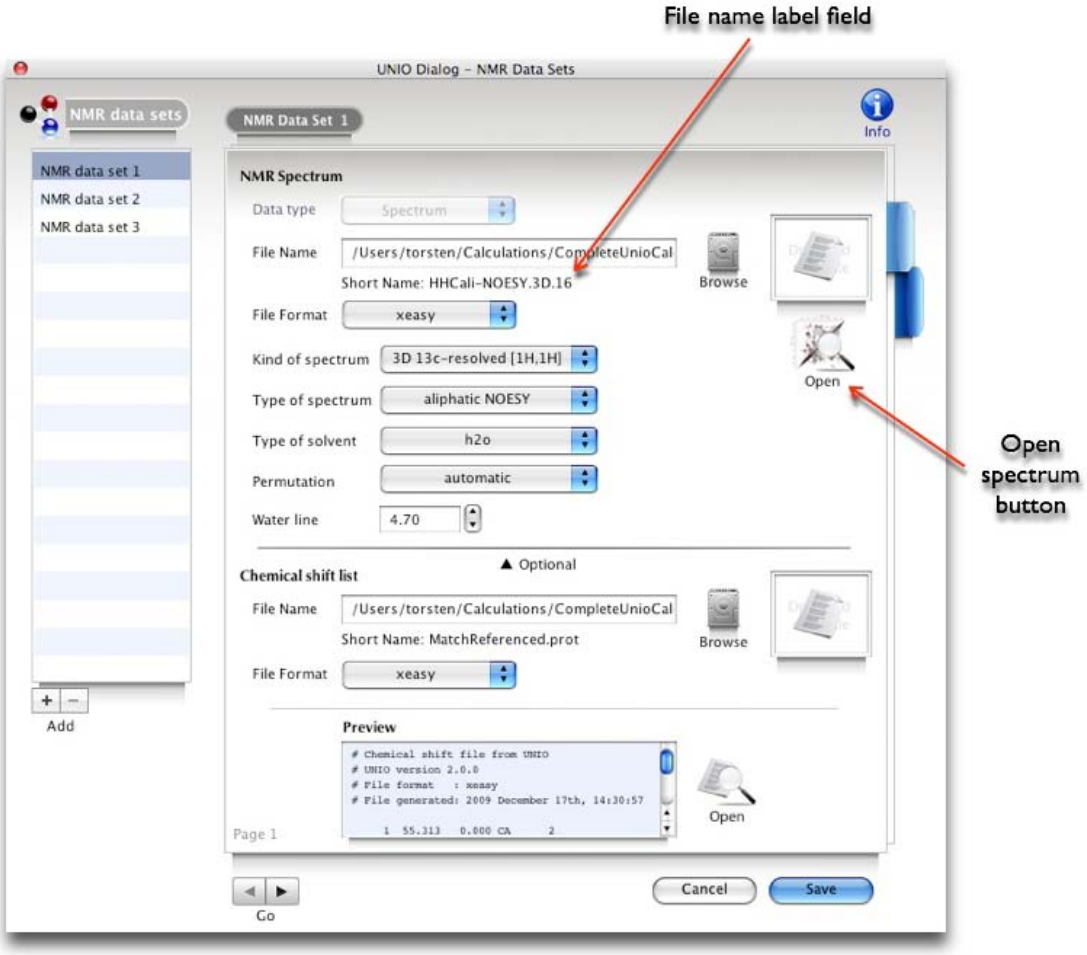
Optionally, you can provide for each input NMR spectrum an individual input chemical shift list. If you click on the *Optional button*, then a submenu will be shown that allows you to provide a chemical shift list. By default the chemical shift list that you have provided as "Backbone chemical shift" list in the input configuration will be taken as chemical shift list for all input NMR spectra.

As long as the input is incomplete, the *Save lock icon* is shown at the right bottom of the window. After completion of the input, you can save your configuration by pressing the *Save* button.



UNIO Configuration window "NMR Data Sets" for ATNOS/ASCAN.

After you have provided the NMR data set file, the *Open Spectrum button* will be shown. If you want to open the current NMR spectrum, then you can press this button that launches the spectrum view program as specified in the UNIO Preferences window.

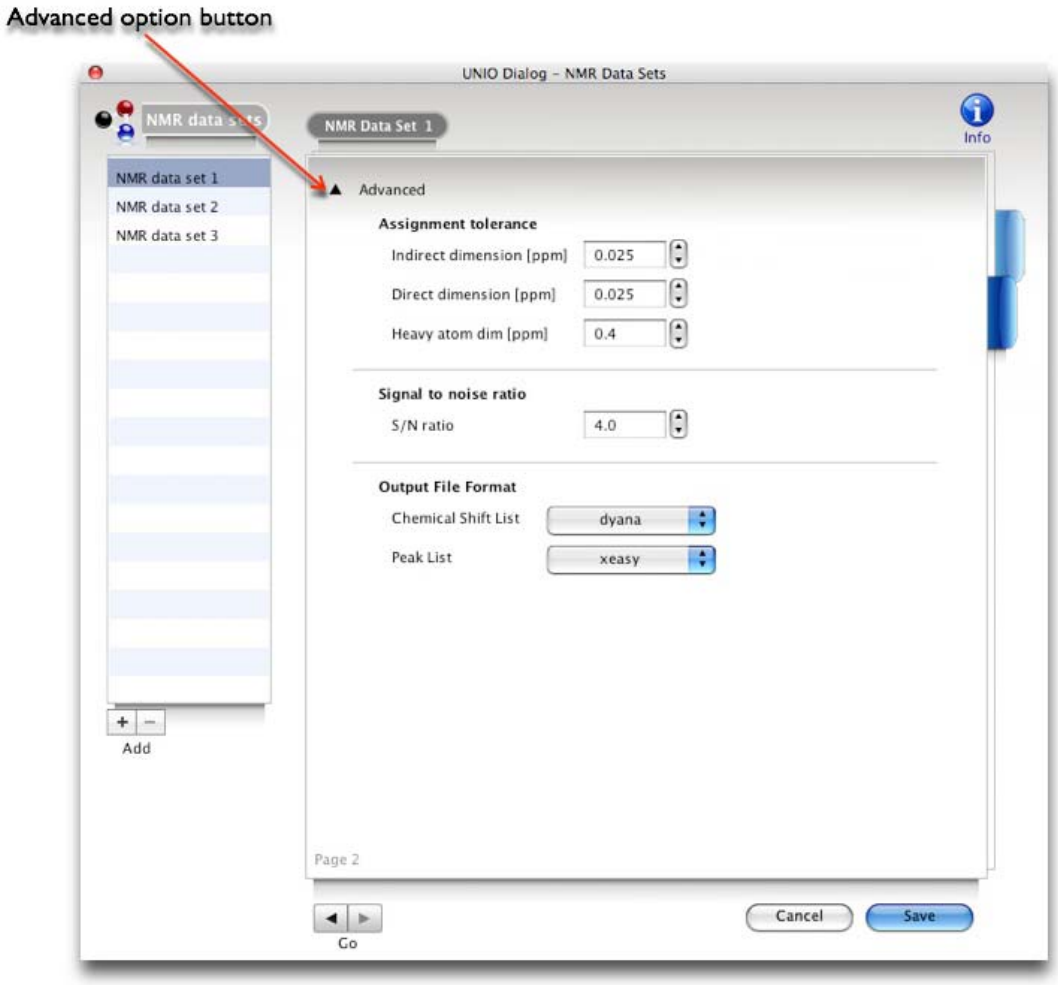


UNIO Configuration window "NMR Data Sets" for ATNOS/ASCAN.

You can use the *Page buttons* at the bottom of the window or the *Page tabs* on the left side of this window to navigate to the advanced, optional input features of this configuration window.

Tip: Currently only XEASY and BRUKER spectrum file format are supported by UNIO. However you can easily convert other spectrum file formats into XEASY format using the program CARA that you can download free-of-charge from the website www.nmr.ch.
After you have loaded your spectrum into CARA, go to Tools -> Export EASY format to generate an XEASY spectrum file format.

If you press the *Advanced option button*, the optional data entry area will get visible. The second page of this configuration window allows you to change parameter values used for ATNOS/ASCAN side-chain resonance assignment. You can specify the tolerance for chemical shift assignment for all three spectrum dimensions, the initial signal-to-noise ratio for signal identification and choose the output file format for the chemical shift and peak list written out by UNIO-ATNOS/ASCAN.



UNIO Configuration window "NMR Data Sets" for ATNOS/ASCAN.

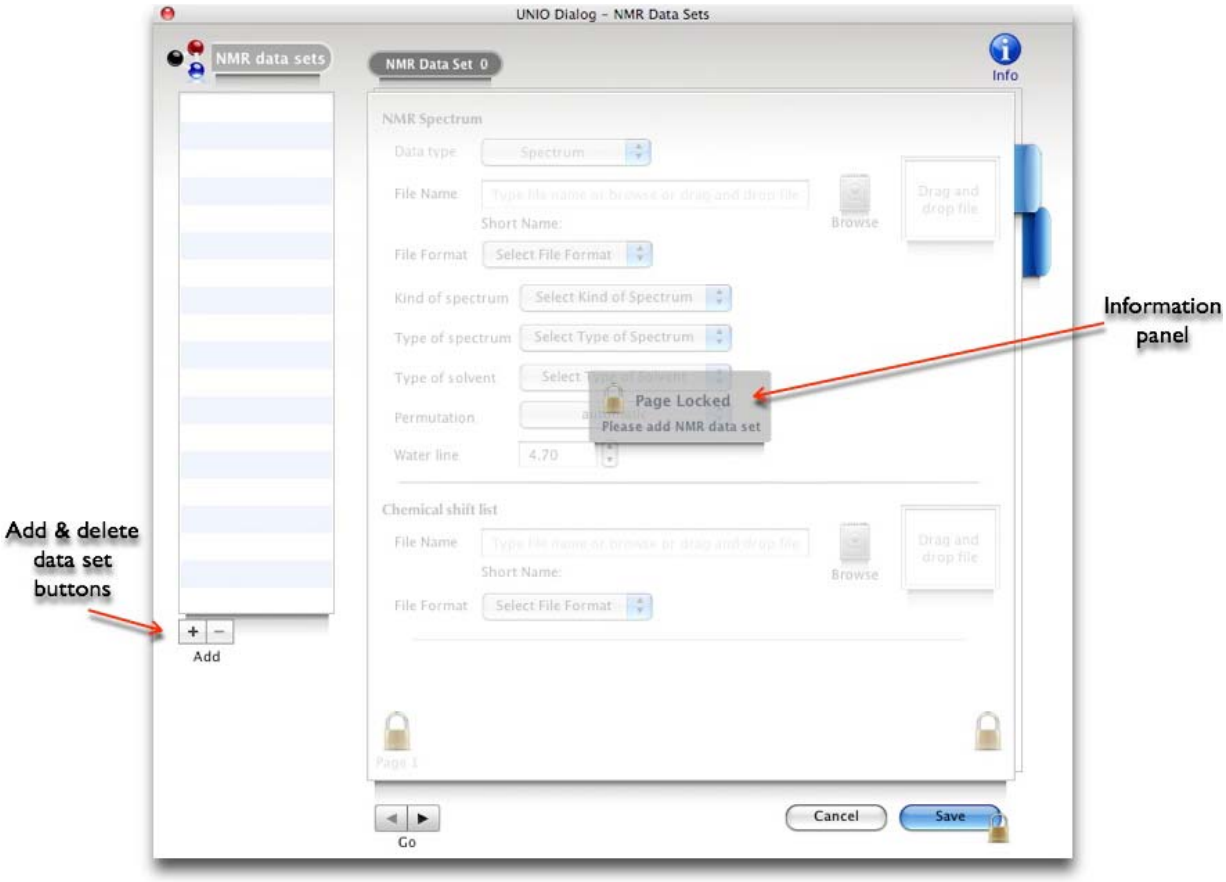
Tip: It is recommended to keep the ATNOS/ASCAN default parameter configuration and only change values if you know exactly what you are doing!

2.14.10 Configuration window "NMR Data Sets" for ATNOS/CANDID

The "NMR Data Sets" configuration window allows you to provide your set of NMR spectra as input for ATNOS/CANDID NOE cross peak assignment.

You can use the *Add&delete data set buttons* to add a new NMR data set or to remove an existing NMR data set. If no data set is yet provided, an *Information panel* indicates that the input data area is locked and first a data set needs to be added.

This configuration window consists of two pages or sheets. The first page collects all required input data, while the second page offers advanced and/or optional input features. The *Page buttons* at the bottom of the window or the *Page tabs* on the left side of this window allows you to quickly navigate between these two sheets.



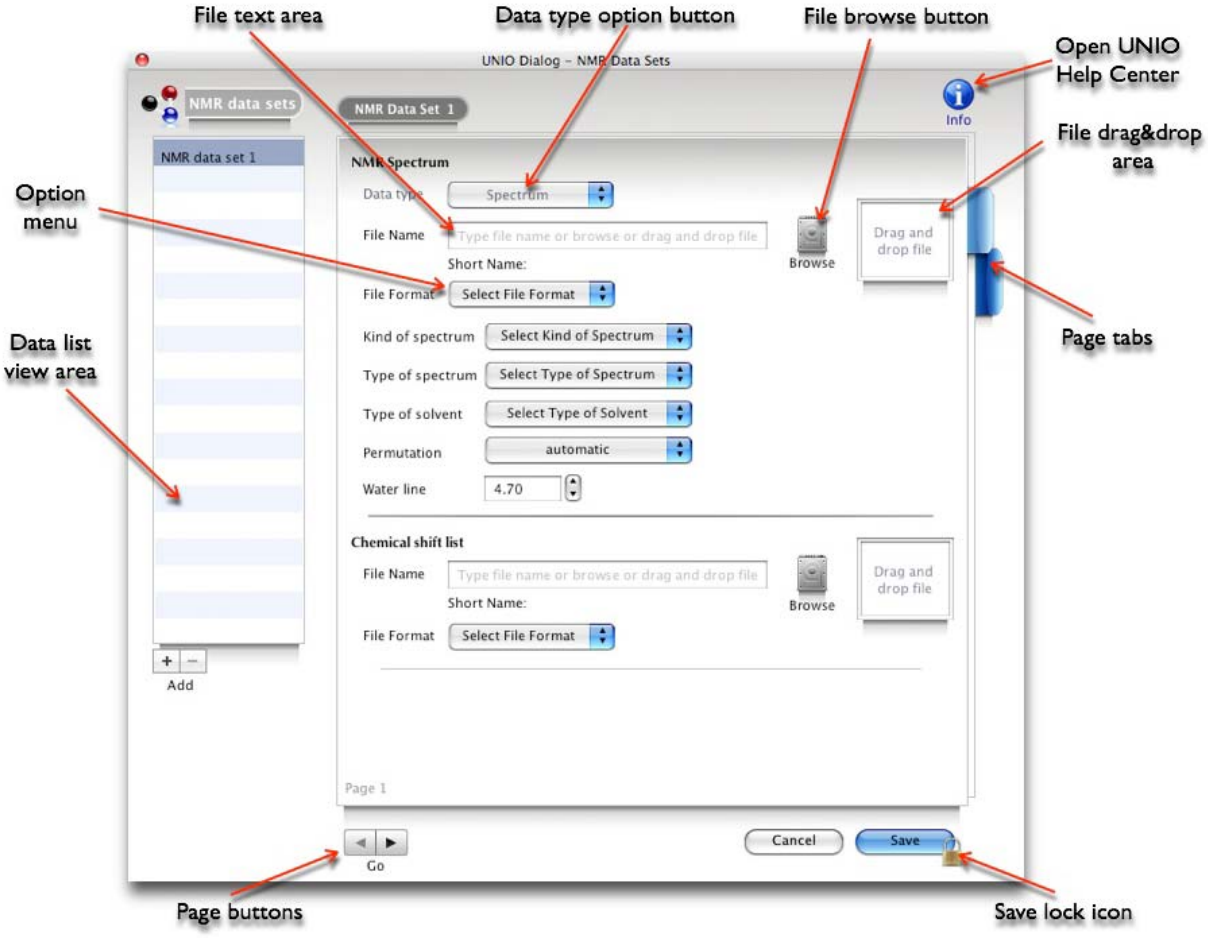
UNIO Configuration window "NMR Data Sets" for ATNOS/CANDID.

After you have added a new data set by pressing the *Add data set button*, you can choose between two kind of input data for ATNOS/CANDID NOE assignment: "Spectrum" or "Peak list" using the *Data type option button*.

After you have chosen the input data type, you can continue to provide the file name of your NMR spectrum/peak list file, the corresponding file format of the spectrum/peak list and additional information about the spectrum/peak list as the kind of and the type of spectrum/peak list, the type of solvent used and the value of the waterline. For each NMR spectrum/peak list, you need also to provide the file name of a chemical shift file and the corresponding file format.

There are three ways to provide your NMR spectrum/peak list file and your chemical shift file, either by typing the complete path of the file in the *File text area*, or by using the system's file browser that is opened by clicking on the *File browse button*, or by dropping the file into the *File drag&drop area*. You can use the corresponding *Option menus* to choose between different possibilities for the above-mentioned spectrum-specific information and the file formats.

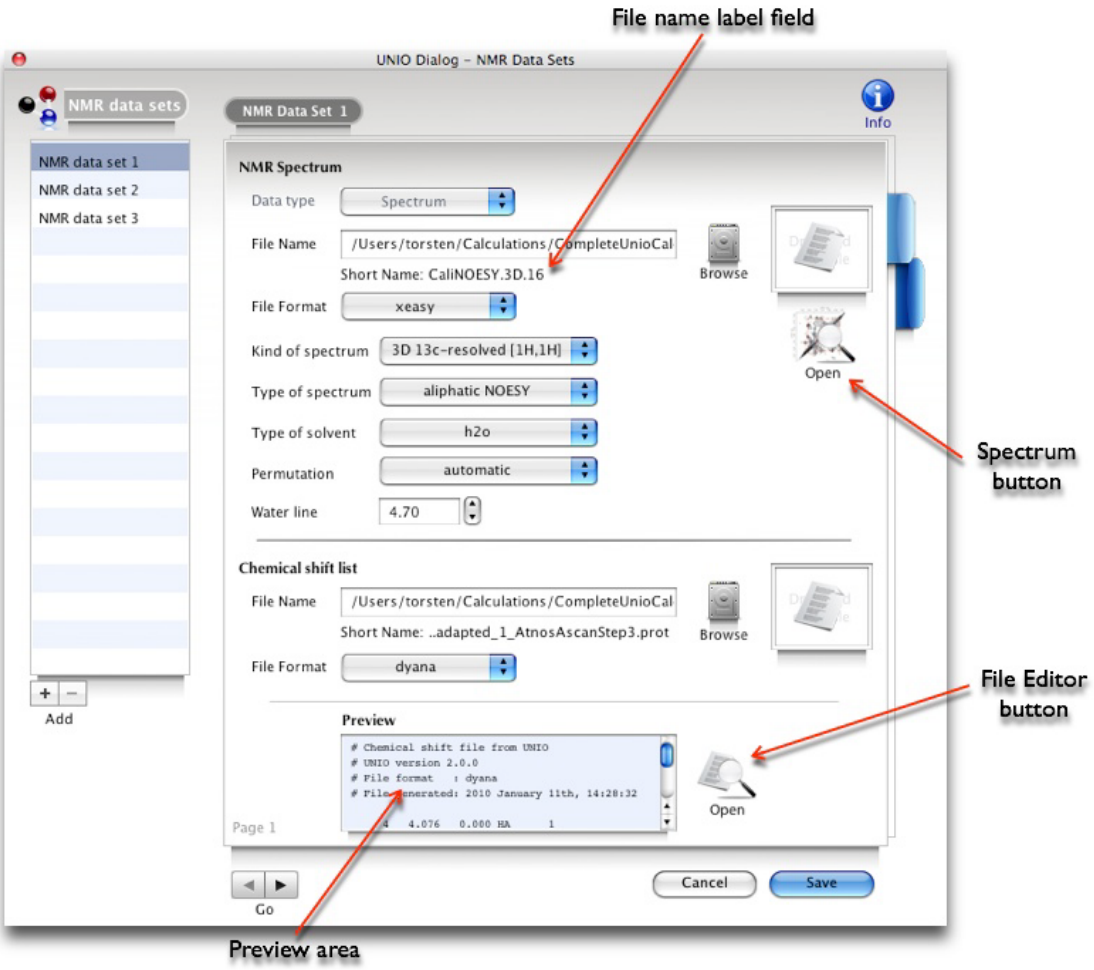
As long as the input is incomplete, the *Save lock icon* is shown at the right bottom of the window. After completion of the input, you can save your configuration by pressing the *Save* button.



UNIO Configuration window "NMR Data Sets" for ATNOS/CANDID.

After you have provided the NMR data set file, the *Open Spectrum button* will be shown. If you want to open the current NMR spectrum, then you can press this button that launches the spectrum view program as specified in the UNIO Preferences window.

Similar, the *Preview area* and the *File Editor button* will be shown as soon as the chemical shift file is loaded. The *Preview area* shows the file content and might be helpful to verify the correctness of the file format selected. If you want to open and edit the chemical shift file, then you can press the *File Editor button* that launches the UNIO File Editor.



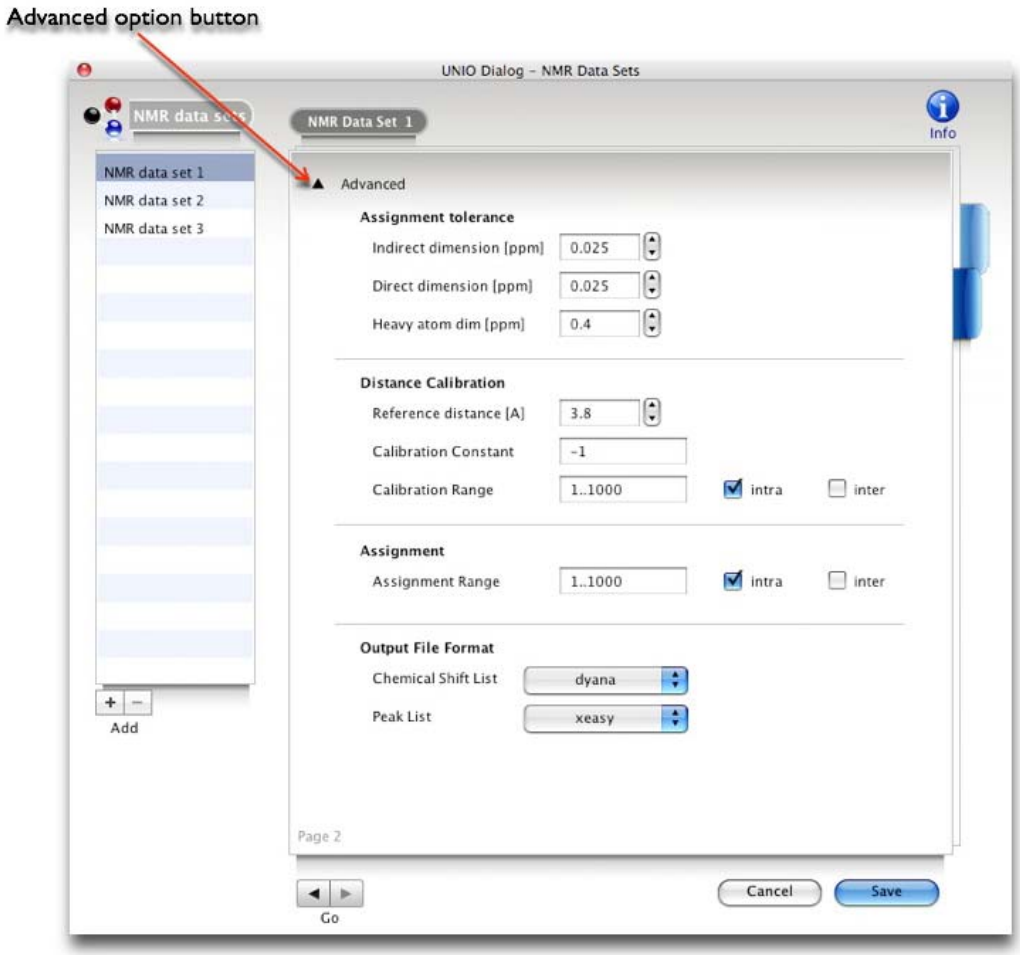
UNIO Configuration window "NMR Data Sets" for ATNOS/CANDID.

Tip: Currently only XEASY and BRUKER spectrum file format are supported by UNIO. However you can easily convert other spectrum file formats into XEASY format using the program CARA that you can download from the website www.nmr.ch.
After you have loaded your spectrum into CARa, go to Tools -> Export EASY format to generate an XEASY spectrum file format.

You can use the *Page buttons* at the bottom of the window or the *Page tabs* on the left side of this window to navigate to the advanced, optional input features of this configuration window.

If you press the *Advanced option button*, the optional data entry area will get visible. The second page of this configuration window allows you to change parameter values used for ATNOS/CANDID NOE cross peak assignment.

You can specify the tolerance for chemical shift assignment for all three spectrum dimensions, give information about the applied distance calibration, specify the range of atoms in the protein sequence that can be assigned within the given spectrum and choose the output file format for the chemical shift and peak list written out by UNIO-ATNOS/CANDID.



UNIO Configuration window "NMR Data Sets" for ATNOS/CANDID.

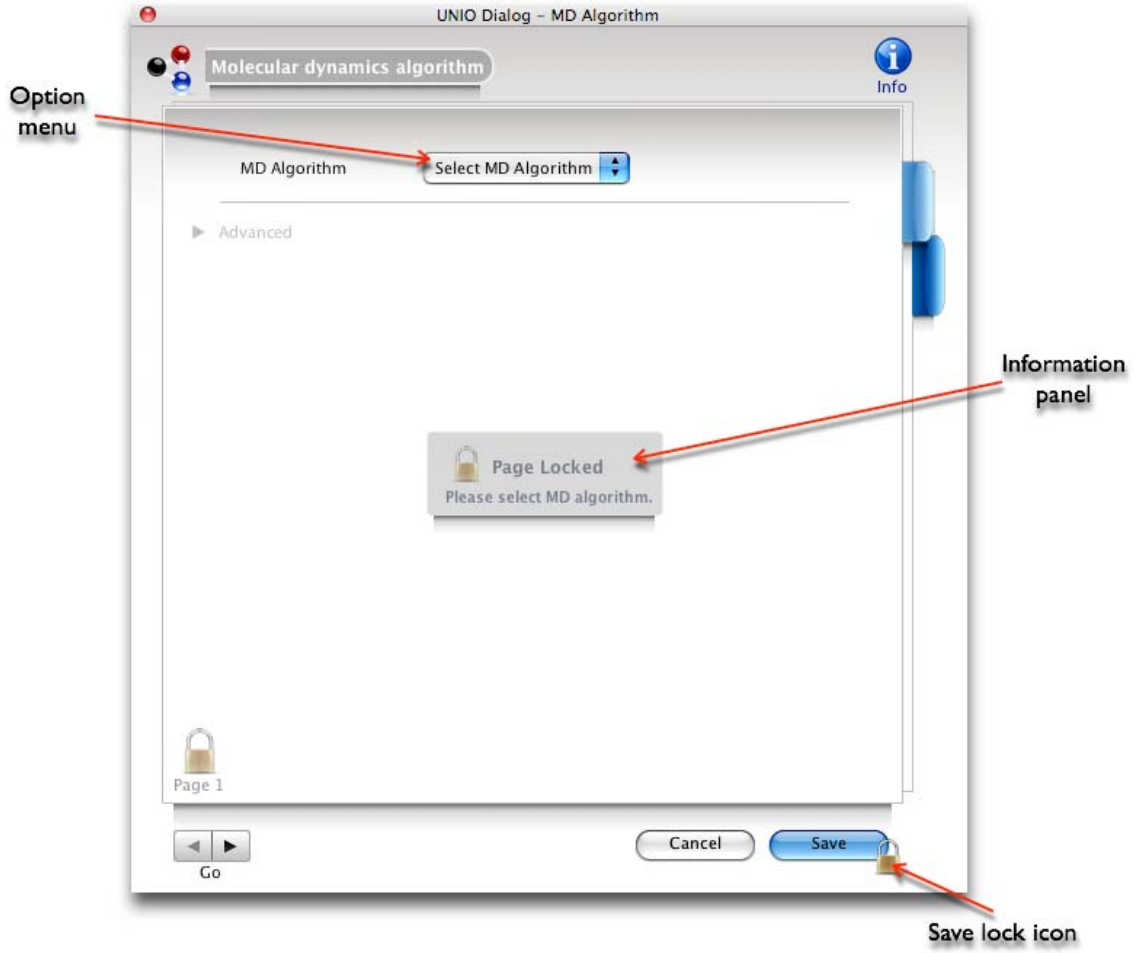
Tip: It is recommended to keep the ATNOS/CANDID default parameter configuration and only change values if you know exactly what you are doing!

2.14.11 Configuration window "MD Algorithm"

The "MD Algorithm" configuration window allows you to select and configure the molecular dynamics algorithm that you want to use in combination with ATNOS/CANDID NOE cross peak assignment.

ATNOS/CANDID generates in each iteration step (ATNOS/CANDID cycles 1..7) a list of meaningful upper distance restraints that is passed on to the selected MD algorithm for calculating the 3D protein structure. UNIO supports commonly used structure calculation engines such as CYANA, CNS, XPLOR and XPLOR-NIH.

You can use the *Option menu button* to choose a MD algorithm. If no MD program is yet selected, an *Information panel* indicates that the input data area is locked and first a choice need to be made to open the input data area.



UNIO Configuration window "MD Algorithm".

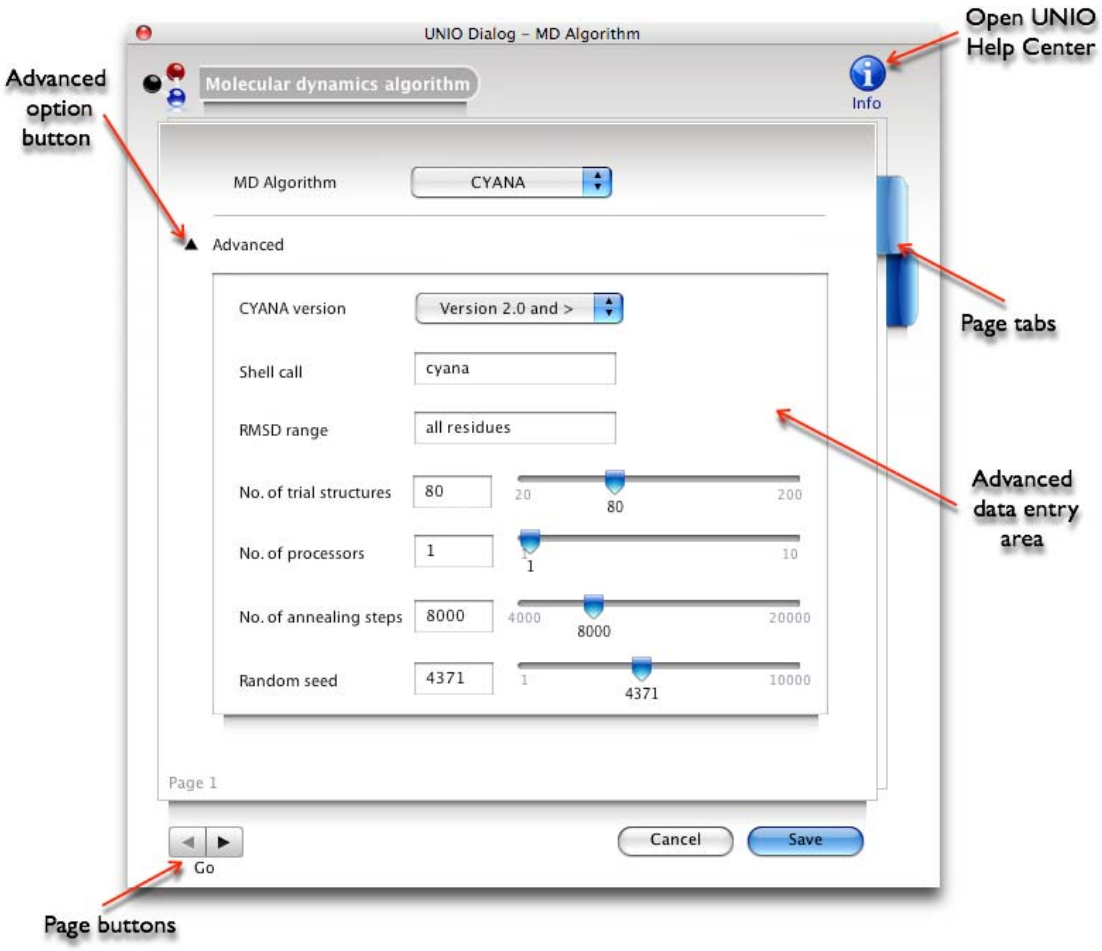
The "MD algorithm" configuration window consists of two pages or sheets. The first page allows to modified default parameters used in combination with the MD program, while the second page offers advanced and/or optional input features.

The *Page buttons* at the bottom of the window or the *Page tabs* on the left side of this window allows you to quickly navigate between these two sheets.

2.14.11.1 Configuration window "MD Algorithm" for CYANA

If you press the *Advanced option button*, the optional data entry area will get visible. This input area allows you to change parameter values used for ATNOS/CANDID NOE cross peak assignment in combination with CYANA torsion angle dynamics.

You can specify the CYANA version, the MD shell command to call the program, the RMSD range, the number of trial structures, the number of processors, the number of annealing steps and the starting random seed.



UNIO Configuration window "MD Algorithm" for CYANA.

Advanced CYANA configuration options

CYANA version	UNIO supports CYANA version 1.0, 2.0 and 3.0 and higher.
Shell call	Specifies the executable shell call of the program as it is found in your shell $PATH. Default value: Same name as the MD algorithm

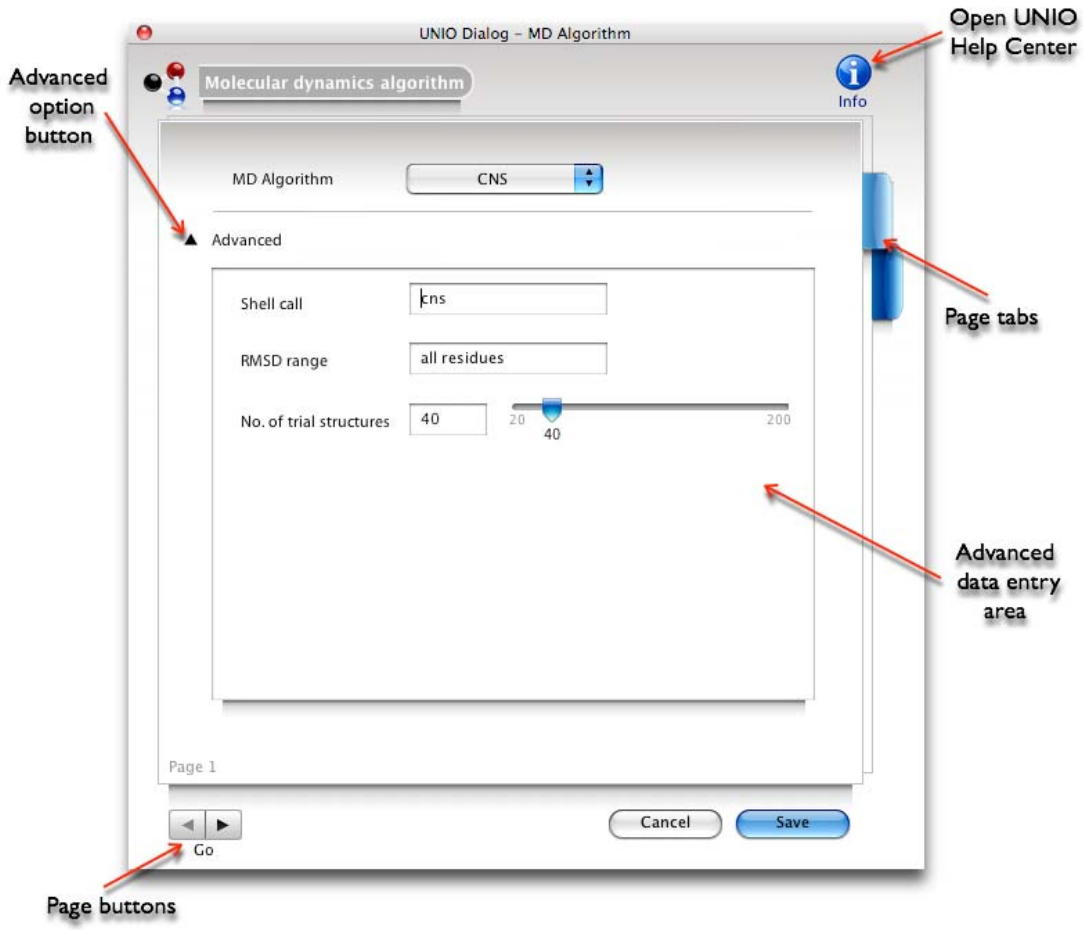
RMSD range	Specifies the residue range for the superposition used to calculate the pair-wise root-mean square deviation (RMSD) between all pairs of conformers in the NMR structure bundle for the backbone atoms and for all heavy atoms. Default value: "all residues" Syntax examples: 2-5,10-35,57,62,100-135
No. of trial structures	CYANA calculates N trial structures from which a subset n<N is selected to represent the NMR structure. Default value: 80
No. of processors	Number of processors that can be used by CYANA. Default value: 1
No. of annealing steps	CYANA performs simulated annealing with a total of N MD steps used to go from the starting to the final temperature. Default value: 8000
Random seed	Random number generator seed used to generate the starting conformers for the following simulated annealing schedule. Default value: 4371

Tip: It is recommended to keep the CYANA default parameter configuration and only change values if you know exactly what you are doing!

2.14.11.2 Configuration window "MD Algorithm" for CNS and XPLOR

If you press the *Advanced option button*, the optional data entry area will get visible. This input area allows you to change parameter values used for ATNOS/CANDID NOE cross peak assignment in combination with CNS, XPLOR or XPLOR-NIH structure calculation.

You can specify the MD shell command to call the program, the RMSD range and the number of trial structures.



UNIO Configuration window "MD Algorithm" for CNS, XPLOR, XPLOR-NIH.

Advanced CNS, XPLOR, XPLOR-NIH configuration options

Shell call	Specifies the executable shell call of the program as it is found in your shell $PATH. Default value: Same name as the MD algorithm
RMSD range	Specifies the residue range for the superposition used to calculate the pair-wise root-mean square deviation (RMSD) between all pairs of conformers in the NMR structure bundle for the backbone atoms and for all heavy atoms. Default value: "all residues" Syntax examples: 2-5,10-35,57,62,100-135

No. of trial structures	CNS, XPLOR or XPLOR-NIH calculate N trial structures from which a subset n<N is selected to represent the NMR structure. Default value: 40

Tip: It is recommended to keep the CNS, XPLOR or XPLOR-NIH default parameter configuration and only change values if you know exactly what you are doing!

UNIO-ATNOS/CANDID in combination with CNS

The following description is only valid for ATNOS/CANDID NOE assignment in combination with CNS structure calculation.

At the moment, you have to copy by hand the following three files into your project directory: "generate_seq.inp", "generate_extended.inp" and "anneal.inp". To create these files, go to the CNS website and follow the menu "Input files -> General" (for generate_seq.inp) and "Input files -> NMR" (for generate_extended.inp and anneal.inp). The number of trial structures for "anneal.inp" must correspond to the number of trial structures you have specified in the "MD algorithm" configuration window.

Next, open the file "anneal.inp" with a text editor and ensure that the following lines look exactly like this:

```
{===>} nmr.noe.file.1="Cycle?_AtnosCandid.upl";
{===>} nmr.noe.file.2="";
{===>} nmr.noe.ave.mode.1="sum";
{===>} nmr.noe.ave.mode.2="sum";
{===>} nmr.cdih.file="";
{===>} pdb.out.name="Cycle?_MD";
```

Tip: Templates can be found in "$DIR_INSTALL/UnioAlgorithms/unio-1.0.0/example/cns/". The provided "anneal.inp"-file has been tested successfully and proven to provide suitable CNS - MD parameters in combination with ATNOS/CANDID.

UNIO-ATNOS/CANDID in combination with XPLOR-NIH

At the moment, you have to copy by hand the following three files into your project directory: ”generate.inp”, “generate_template.inp” and “sa.inp”. To create these files, go to the website of XPLOR-NIH. The number of trial structures for "sa.inp" must correspond to the number of trial structures you have specified in the “MD algorithm” configuration window.

Next open the file "sa.inp" with a text editor and ensure that the following lines look exactly like this:

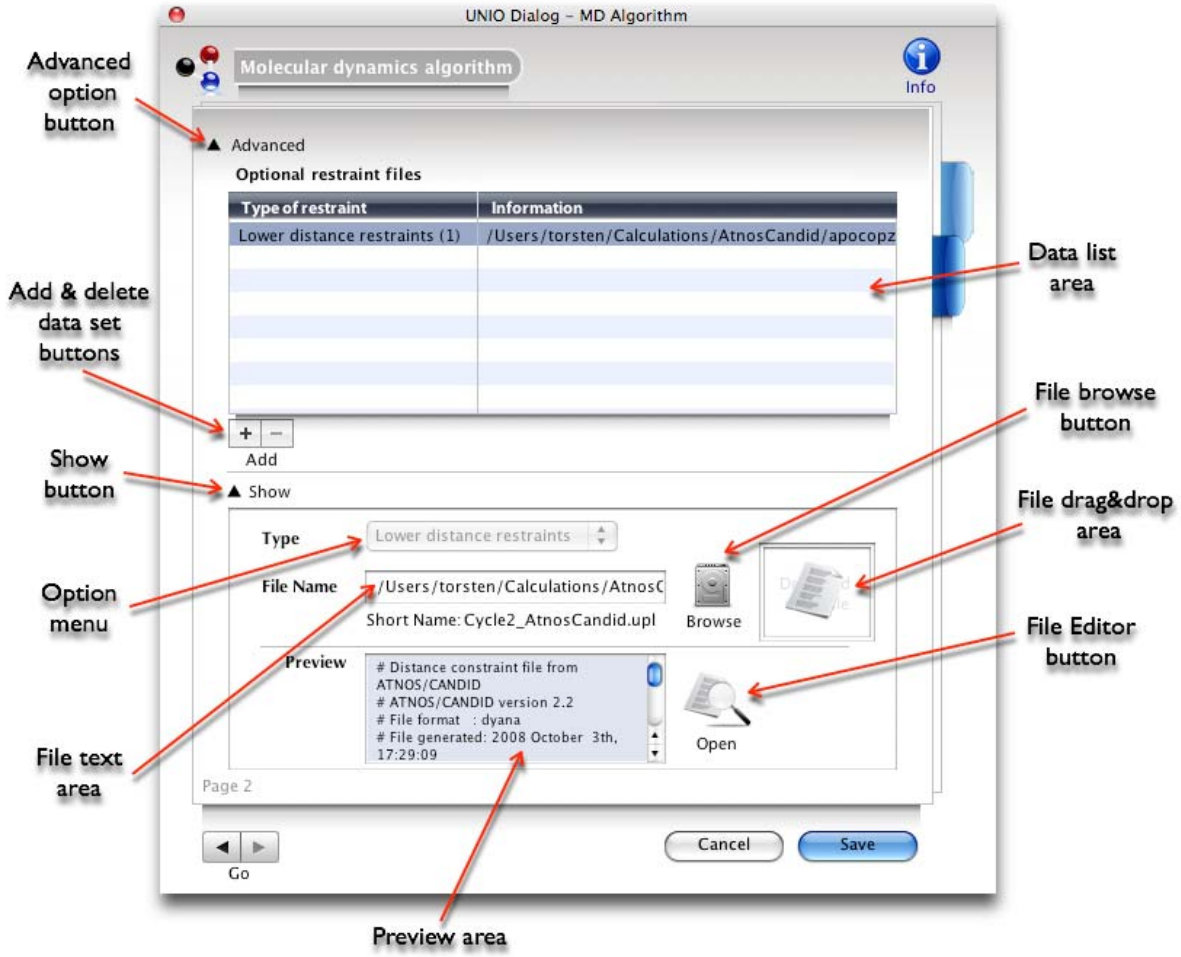
```
noe
nres = 75000
class all
@Cycle?_AtnosCandid.upl
!@RegularSecondaryContacts.upl
end

noe
ceiling 1000
averaging all sum
potential all square
scale all $knoe
sqconstant all 1.0
sqexponent all 2 end

restraints dihed
reset
scale $kcdi
nass = 5000
!@RegularSecondaryContacts.aco
end

{====>} {*Name(s) of the family of final structures.*}
evaluate ($file = "Cycle?_MD_" + encode($count) + ".pdb")
write coor output= $file end
```

Tip: Templates can be found in "$DIR_INSTALL/UnioAlgorithms/unio-1.0.0/example/xplor-nih/". The provided "sa.inp"-file has been tested successfully and proven to provide suitable XPLOR-NIH - MD parameters in combination with ATNOS/CANDID.

You can use the *Page buttons* at the bottom of the window or the *Page tabs* on the left side of this window to navigate to the advanced, optional input features of the second of this configuration window.

If you press the *Advanced option button*, the optional data entry area will get visible. The second page of this configuration window allows you to provide supplementary conformational restraints, such as lower and upper distance restraints, dihedral angle restraints, residual dipolar couplings, J-couplings, hydrogen bonds and disulfide bonds.

All given information will only be used by the MD algorithm selected. ATNOS/CANDID is not making any use of these restraint files.



UNIO Configuration window "MD Algorithm".

After you have added a new restraint file by pressing the *Add data set button*, you can continue to provide the file name of your restraint file. There are three ways to provide your restraint file, either by typing the complete path of the file in the *File text area*, or by using the system's file browser that is opened by clicking on the *File browse button*, or by dropping the file into the *File drag&drop area*.

As long as the input is incomplete, the *Save lock icon* is shown at the right bottom of the window. After completion of the input, you can save your configuration by pressing the *Save* button.

The *Preview area* and the *File Editor button* will be shown as soon as the restraint file is loaded. The *Preview area* shows the file content and might be helpful to verify the correctness of the file format that must be compatible with the selected MD algorithm. If you want to open and edit the restraint file, then you can press the *File Editor button* that launches the UNIO File Editor.

To delete a restraint file, first you need to select the corresponding restraint is the *Data list area* and then press the *Delete Data set button*.

Important: Currently, optional restraints are only automatically passed on for ATNOS/CANDID in combination with CYANA. If you are using a different structure calculation engine, then you need to add this restraints by hand in the corresponding annealing protocols.

Important: The file format of the provided restraint files must be compatible with the MD algorithm used.

Chapter 3 UNIO Tips and Tricks

This section describes some tips and tricks for working with UNIO, mainly in the context of highly automated/fully automated complete protein structure determination with UNIO. A special emphasis is given on aspects to be considered for the input preparation and on best practice for moving from one NMR data analysis module to the next.

3.1 UNIO - MATCH

The MATCH backbone assignment module got designed to work highly efficient and accurate with APSY-type experiments that encode sufficient information in order to perform sequence-specific backbone resonance assignment.

Minimal requirement for the input data is the presence of sequential correlations, i.e., the two conventional triple-resonance 3D experiments HNCA and HN(CO)CA would be sufficient as input for MATCH. However, since MATCH as any backbone assignment program relies strongly on data base comparison between the input chemical shift values of the peak positions and the average chemical shift value of corresponding atoms found for NMR active nuclei in proteins, it is recommended that the input peak lists contain information about both the CA and the CB resonance frequencies.

Tip: A typical set of APSY input peak lists for MATCH backbone assignment consists of a 5-dimensional BAONH, a 5-dimensional ZAONH and a 4-dimensional zaNH APSY experiments.

MATCH is able to perform automated chemical shift referencing for all backbone atoms, i.e. for HA, HN, N, CA, CB, CO. Nonetheless, it is advised to accurately calibrate your input peak lists prior to the use of MATCH.

Tip: MATCH provides two listings of assigned backbone atoms called "Match.prot" and "MatchReferenced.prot". The first list "Match.prot" gives the resonance frequencies of the assigned atoms as derived from the frequency coordinates of the input signals. The second list "MatchReferenced.prot" gives the chemical shift values of the assigned backbone atoms after automated chemical shift referencing. If your input peak lists were correctly calibrated, then these two output chemical shift lists will be almost identical.

3.2 UNIO - ATNOS/ASCAN

The ATNOS/ASCAN side-chain assignment module works best if the input backbone chemical shift list contains information about the HN, N, HA, CA and CB atoms. The above mentioned typical set of APSY experiments are exactly providing this information. In addition, it is advised that the input NOESY spectra have been correctly calibrated prior to the use of ATNOS/ASCAN.

Tip: A typical set of input NOESY spectra for ATNOS/ASCAN side-chain assignment consists of a 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY with the carrier frequency in the aliphatic region, a 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY with the carrier frequency in the aromatic region and a 3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY.

If all your input NOESY spectra are correctly calibrated, then you can use the MATCH output chemical shift list "MatchReferenced.prot" as input backbone chemical shift list for ATNOS/ASCAN.

For proper performance of the side-chain assignment process, ATNOS/ASCAN must validate NOE signals for at least 85% of all pairwise combinations of protons i and j (i is not equal j) for which sequence-specific resonance assignments are available, and which have covalent structure-imposed upper distance limits shorter than a user-given threshold. For example, the maximal distance between the intraresidual atoms Hα and HN is always smaller than 5 Å, and therefore must lead to an observable signal in the NOESY spectrum.

In the directory "AtnosAscanStep1", you can find a peak list of expected covalent peak pattern for each input NOESY spectrum. The percentage of agreement between input chemical shift list and NOESY spectra are reported in the section "ATNOS/ASCAN: Evaluate Results: Summary". If the reported percentage of agreement is low, then it is recommended to use the above-mentioned peak lists to detect possible disagreements between the calibration of a given NOESY spectrum and the input backbone chemical shift list.

Tip: ATNOS/ASCAN provides similar to MATCH two listings of assigned atoms called "AtnosAscanFinal.prot" and "AtnosAscanFinalReferenced.prot". The first list "AtnosAscanFinal.prot" gives the resonance frequencies of the assigned atoms as derived from the frequency coordinates of the NOE signals. The second list "AtnosAscanFinalReferenced.prot" gives the chemical shift values of the assigned atoms after automated chemical shift referencing. If your input peak lists were correctly calibrated, then these two output chemical shift lists will be almost identical.

Tip: If you want to use the ATNOS/ASCAN chemical shifts for the following step of ATNOS/CANDID NOE assignment, then it is recommended to use the individually adapted chemical shift lists for each NOESY spectra that are provided in the directory "AtnosAscanStep3".
The ATNOS/ASCAN chemical shift lists for best use with ATNOS/CANDID are called "AtnosAscan_adapted_1_AtnosAscanStep3.prot and correspondingly.

Tip: In the directory AtnosAscanStep3, a TOCSY peak list is automatically generated for each input NOESY spectrum. Use these peak lists to quickly verify the resonance frequencies of the ATNOS/ASCAN assigned atoms.

3.3 UNIO - ATNOS/CANDID

For proper performance, the ATNOS/CANDID module must met certain input and output criteria that are listed below.

Input requirement 1: The chemical shift list must contain more than 90% of the non-labile and backbone amide ^{1}H chemical shifts. If 3D heteronuclear-resolved [^{1}H,^{1}H]-NOESY spectra are used, more than 90% of the ^{15}N and/or ^{13}C chemical shifts must also be available. The chemical shift values must be correct.

Important: The criterion 1 is more relaxed if the resonance assignments have been obtained with ATNOS/ASCAN. ATNOS/ASCAN is only assigning conformationally meaningful side-chain atoms that are involved in many NOE signals. For example, side-chain atoms close to the protein's surface are hardly assigned. This intrinsic feature of ATNOS/ASCAN leads to a lower completeness of the resonance assignments. However, since only assignments for side-chain atoms are missing that do not contribute to meaningful NOE signals, the criterion 1 can be relaxed to 70-75% of atoms that must be assigned.

Input requirement 2: ATNOS/CANDID must validate NOE signals for at least 85% of all pairwise combinations of protons i and j (i is not equal j) for which sequence-specific resonance assignments are available, and which have covalent structure-imposed upper distance limits shorter than a user-given threshold. For example, the maximal distance between the intraresidual atoms Hα and HN is always smaller than 5 Å, and therefore must lead to an observable signal in the NOESY spectrum.

Please keep in mind that the second requirement only makes a statement about the calibration of the chemical shift list to the NOESY spectrum. A low percentage of validated covalent NOE contacts is usually indicative of insufficient chemical shift agreement between the chemical shift list and the NOESY spectrum.

Whenever the difference between the NOE cross peak positions and the chemical shift of a given atom pair is larger than the predetermined tolerance range, then ATNOS/CANDID will fail to identify the corresponding NOE cross peak. In such a case, a higher percentage can usually be obtained by re-calibrating the chemical shift list using a constant value for all hydrogen atoms.

Output criterion 1: The average final ATNOS/CANDID target function value for the bundle of conformers used to represent the structure from the first ATNOS/CANDID cycle should be below 150 $Å^2$, and the corresponding value for the last ATNOS/CANDID cycle should be below 2-3 $Å^2$.

Output criterion 2: The average backbone RMSD to the mean coordinates for the structured parts of the polypeptide chain should be below 3.0 Å for the structure bundle of ATNOS/CANDID cycle 1.

Output criterion 3: The RMSD-Drift between the mean coordinates of the k^{th} and the last ATNOS/CANDID cycles calculated for the backbone atoms of the structured parts of the polypeptide chain should be in the order of the RMSD value of the k^{th} cycle.

Definition RMSD-Drift: RMSD-Drift is calculated as the RMSD between the mean coordinates of the bundles of conformers obtained after the k^{th} and the seventh ATNOS/CANDID cycles.