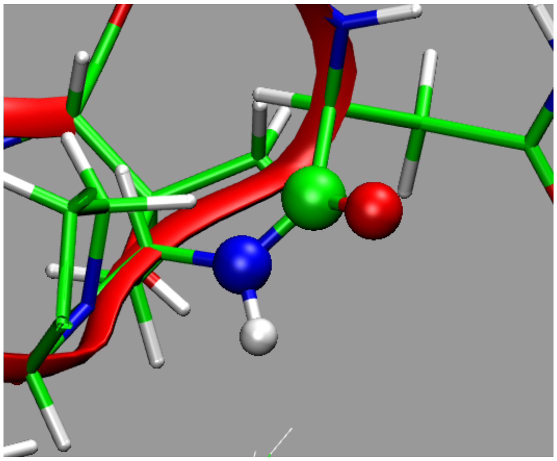
University of Illinois at Urbana-Champaign
NIH Resource for Macromolecular Modeling and Bioinformatics
Beckman Institute
Computational Biophysics Workshop

Structure Check



Authors:
Eduard Schreiner
Leonardo G. Trabuco

October 7, 2016

A current version of this tutorial is available at
http://www.ks.uiuc.edu/Training/Tutorials/

Contents

1 Chirality in proteins and nucleic acids 5

1.1 Introduction 5

1.2 Checking stereochemistry of the structure 5

1.3 Correcting chirality 7

1.4 Preserving stereochemistry during MD simulation 9

2 *cis* peptide bonds in proteins 11

2.1 Introduction 11

2.2 Checking configuration of peptide bonds 11

2.3 Converting *cis* peptides into *trans* configuration 12

2.4 Preserving configuration of peptide bonds during MD simulation 16

Introduction

The purpose of this tutorial is to demonstrate how to use the VMD plugins `cispeptide` and `chirality` together with `AutoIMD` to ensure proper structure of peptide bonds and correct stereochemistry of proteins and nucleic acids. The described plugins are designed to work with proteins and nucleic acids and understand atom names as used by the force fields CHARMM and AMBER as well as the standard PDB naming convention for non-hydrogen atoms.

WARNING: The current implementation of the `chirality` plugin does not support DNA yet.

The following points aim to motivate the necessity of the of the presented plugins:

- Structures from PDB may contain errors in stereochemistry and peptide bond configuration;
- Errors in peptide bond configuration can be introduced by modeling;
- Poor starting conditions, i.e. atoms being too close to each other, can generate huge forces during the initial optimization of the structure leading to errors in structure;
- Molecular dynamics (MD) protocols which apply forces on the molecule may generate forces large enough to introduce errors.

Since MD force fields are defined to support either type of chirality or peptide bond configuration, optimization of the structure will not correct introduced errors.

At his stage, we would like to give a brief outline of the strategy behind the described plugins. The general strategy to ensure correct peptide bond and stereochemical configuration consists of four steps.

1 **Identify unusual configuration at chiral centers and peptide bonds** All *cis* peptide bonds are regarded being unusual. In case of chiral centers, those are regarded unusual which occur less often in nature, e.g. D-form of amino acids.

2 **Inspect the found irregularities** Although it is possible to convert all peptide bonds and chirality centers into their “usual” configuration (this feature will be implemented in the future), it is safer to decide for each irregularity individually. First, different chiral forms of biomolecules and true *cis* peptide bonds are used in biology. Second, it is not always unambiguous how to correct an unusual configuration; this is especially the case for peptide bonds.

3 **Correct the structure if necessary using structure optimization (and equilibration)** The correcting happens in two steps. First one atom is moved to “preset” a proper configuration. In a second step an the structure is relaxed using an MD force field plus restraints enforcing the proper configuration.

4 **If proceeding with further MD simulations, apply restraints to preserve chirality and peptide bond configuration** This step may be necessary, if the molecular system is used in a simulation in which large forces may be applied, e. g., MDFF, TMD etc.

1 Chirality in proteins and nucleic acids

1.1 Introduction

All amino acids but glycine have at least one chiral center at C_α (see Fig. 1). Threonine and isoleucine have an additional chiral center at C_β. According to the D-/L- naming convention, naturally occurring amino acids are found in the L-configuration. Note, however, that D-amino acids do occur in biology, e. g., in cell walls of bacteria. Nucleic acids also have chiral centers. For example, in DNA the atoms C1', C3', and C4' are chiral, while RNA has an additional chiral center at C2'. Chirality is central to all molecular interactions in biological systems. A simple experiment demonstrates the principle: try to shake someone's left hand with your right.

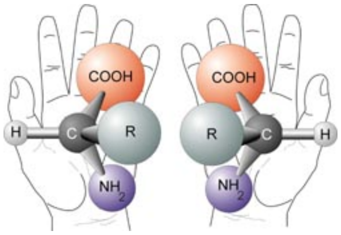


Figure 1: Chirality at the C_α carbon atom of amino acids. The two enantiomers are mirror images of each other, like the left and the right hand. Source: http://web99.arc.nasa.gov/ãstrochm/chirality.jpg

1.2 Checking stereochemistry of the structure

To demonstrate the usage of the plugin we will start by checking a structure consisting of the protein EF-Tu in complex with a tRNA$^{\mathrm{Phe}}$. The used structure is based on the the PDB structure 1OB2 in which errors have been introduced manually.

Currently, the `chirality` plugin can be used from the text console and through a GUI. The available commands provided by the pugin can be obtained by typing `chirality` on the Tk Console or the VMD console. The following information should be printed in the console:

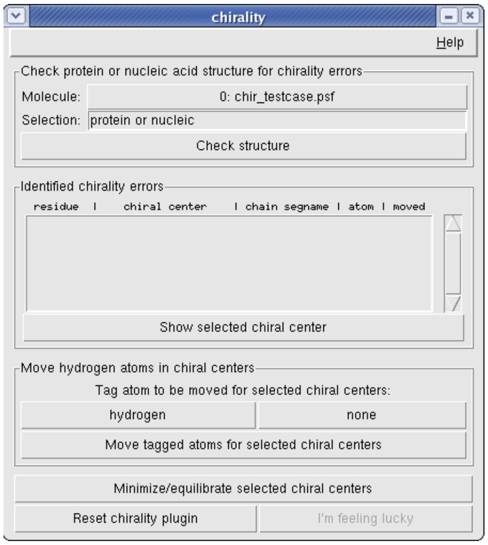
```
Commands:
  check     -- identify chirality errors
  list      -- list identified chirality errors
  minimize -- fix chirality errors using energy minimization
  move      -- move hydrogen atom to fix chirality
  reset     -- reinitialize plugin state
  restrain -- generate NAMD extrabonds file to prevent chirality changes
  show      -- visualize identified chirality errors
```

In the same way, the usage information for any of the provided commands can be obtained. For example, typing `chirality restrain` will show the syntax of the command that generates restraints for the impropers at chiral centers which can be used in NAMD.

In the following, we will describe how to use the `chirality` plugin through its GUI, but all of the tasks can also be accomplished from the console.

1 Load the files `chir_testcase.psf` and `chir_testcase.pdb` into a new session of VMD.

2 Open the `chirality` GUI by selecting Extensions → Modeling → Fix Chirality Errors in the VMD Main menu. In the upper part of the GUI (see Fig. 2) the user can specify the molecule and the atom selection to be tested.


chirality
Help
Check protein or nucleic acid structure for chirality errors
Check structure
Identified chirality errors
residue | chiral center | chain segname | atom | moved
Show selected chiral center
Move hydrogen atoms in chiral centers
Tag atom to be moved for selected chiral centers:
hydrogen
none
Move tagged atoms for selected chiral centers
Minimize/equilibrate selected chiral centers
Reset chirality plugin
I'm feeling lucky

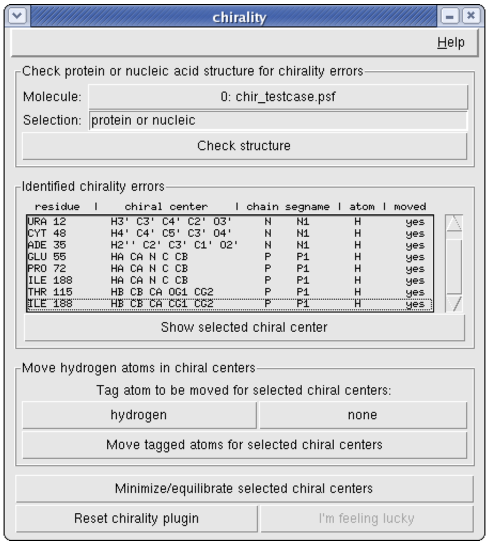
Molecule: 0: chir_testcase.psf

Selection: protein or nucleic

Figure 2: The chirality GUI window.

3 The button Check structure initializes the test. Identified unusual chiral center configurations are displayed in the Identified chirality errors form (see Fig. 3). In the present case there are 9 chiral errors identified in both the protein and RNA parts of the loaded structure.

4 By selecting an entry in the Identified chirality errors form and hitting the button Show selected chiral centers it is possible to visually inspect all the individual identified unusual chiral centers. In the created representation (see Fig. 4) the chiral center is highlighted by a transparent purple sphere, while the four bonded atoms are shown in CPK.


chirality
Help
Check protein or nucleic acid structure for chirality errors
Check structure
Identified chirality errors
Show selected chiral center
Move hydrogen atoms in chiral centers
Tag atom to be moved for selected chiral centers:
hydrogen
none
Move tagged atoms for selected chiral centers
Minimize/equilibrate selected chiral centers
Reset chirality plugin
I'm feeling lucky

Molecule: 0: chir_testcase.psf

Selection: protein or nucleic

```
 residue  |    chiral center     | chain segname | atom | moved
URA 12       H3' C3' C4' C2' O3'     N     N1       H        yes
CYT 48       H4' C4' C5' C3' O4'     N     N1       H        yes
ADE 35       H2'' C2' C3' C1' O2'    N     N1       H        yes
GLU 55       HA CA N C CB            P     P1       H        yes
PRO 72       HA CA N C CB            P     P1       H        yes
ILE 188      HA CA N C CB            P     P1       H        yes
THR 115      HB CB CA OG1 CG2        P     P1       H        yes
ILE 188      HB CB CA CG1 CG2        P     P1       H        yes
```

Figure 3: Identified unusual chiral center configurations in the chirality GUI window.

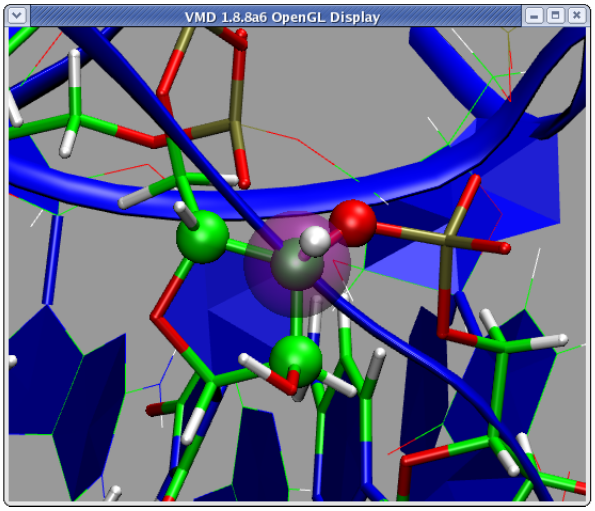


Figure 4: Representation of the chiral error in U12 of the tRNA

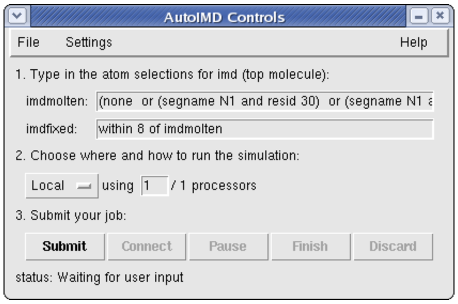
1.3 Correcting chirality

1 If the shown chiral center has the wrong chirality, the user can tag an atom to be moved to flip the chirality at the selected center. This is

done by hitting the button hydrogen. Note, currently only the moving of a hydrogen atom at a chiral center is supported. If the chirality at the displayed center is correct, nothing needs to be done. It is also possible to untag a tagged atom by selecting it and hitting none.

2 Once a chiral center has been inspected, and the atom to be moved has been tagged, the actual moving of the atom can be executed by hitting the button Move tagged atoms for selected chiral centers.

3 Since the described procedure (simple moving of the hydrogen atom) generates an unphysical geometry of the molecule, it is necessary to optimize the structure using an MD force field. The `chirality` plugin uses the `AutoIMD` plugin for this purpose. This final step is accomplished by selecting the chiral centers which should be relaxed and hitting the button Minimize/equilibrate selected chiral centers. This will open the AutoIMD Controls window (see Fig. 5).



1. Type in the atom selections for imd (top molecule):

imdmolten: (none or (segname N1 and resid 30) or (segname N1 a

imdfixed: within 8 of imdmolten

2. Choose where and how to run the simulation:

Local using 1 / 1 processors

3. Submit your job:

status: Waiting for user input

Figure 5: AutoIMD window called by `chirality` plugin.

4 Select Settings → Minimization Mode in the AutoIMD Controls window and hit then the Submit button. Once the `AutoIMD` session started, hit the Connect button and a minimization will start. Finish the simulation when the structure relaxed. Usually, it takes up to a few thousand steps until atoms do not move anymore. Is the structure minimized, hit the Finish in the AutoIMD Controls window. For more details about using the `AutoIMD` plugin, the user is referred to the AutoIMD user's guide
http://www.ks.uiuc.edu/Research/vmd/plugins/autoimd/ .

5 **WARNING:** Depending on the quality of the initial structure, it can happen that not all errors can be fixed in one run. Thus, it is important to check the final result again and correct the remaining errors, if necessary. In order to check the structure after the performed minimization/equilibration using AutoIMD, the `chirality` plugin should be reset. This can be done by hitting the Reset chirality plugin button at the bottom of the chirality window. Note, in the current implementation it will be necessary to select the molecule again in the top part of the chirality window. Note also that if the molecule to correct is solvated, it may be necessary to equilibrate the minimized parts of the structure. This can be done by selecting Settings → Equilibration Mode and submitting a new `AutoIMD` run.

6 Finally, it is a good advice to minimize the structure in the very last step. This should be done by selecting Settings → Minimization Mode in the AutoIMD Controls window, which is still open from the last `AutoIMD` session, and hit then the Submit button. As mentioned in the introduction, additional restraints during the correction of chirality errors. The restraints will be used always when `AutoIMD` is started from the chirality window. The equilibrium values for these restraints are very approximate and bring the configuration in the correct region. Thus, in order to obtain a optimized structure, it is necessary NOT to use the restraints defined by `chirality`. This is achieved by reusing the AutoIMD window from the previous minimization/equilibration; the restraints will not be taken into account.

7 Once all minimization/equilibration steps are performed, the corrected structure should be saved. This can be done in the following way. Select the molecule in the VMD main window. Create a representation containing all atoms in the Graphical Representations window. Select File → Save Coordinates in the VMD main window. Select `all` in the Selected atoms and hit Save. Finally, specify the file name in the Chose filename to save trajectory and hit OK.

1.4 Preserving stereochemistry during MD simulation

As mentioned in the introduction, structure optimizations and MD protocols applying external forces can introduce chirality errors into the simulated structure. To preserve stereochemistry, the `chirality` plugin offers the possibility to generate restraints in form of `extrabonds` which can be used in a NAMD simulations. This feature currently is available only from the console.

1 Assuming the structure of interest is loaded in the top molecule in VMD and one wishes to obtain restraints for all chiral centers in proteins and nucleic acids, the command reads

```
chirality restrain -o chirality-extrabonds.txt
```

The file `chirality-extrabonds.txt` contains now two improper restraints for each chiral center. One includes the chiral center and its non-hydrogen binding partners. In the second restraint, the chiral center is replaced by the hydrogen bound to it. The equilibrium values of the two impropers are set to the values present in the used structure.

WARNING: Note, the specified restraints modify the potential energy function of the system, and thus should not be used in equilibrium production runs.

Rather, these restraints are supposed to be used only in simulations in which large forces occur due to poor initial geometry or due to externally applied forces, e.g., MDFF or TMD. Note also that although the final structure resulting from such simulations will be correct, the behavior of the system during the simulation is artificial.

2 *cis* peptide bonds in proteins

2.1 Introduction

In naturally occurring proteins most peptide bonds are in the *trans* configuration (see Fig. 6). However, sometimes *cis* peptide bonds do occur. The vast majority of *cis* peptides is observed at a proline, Xaa-Pro, Xaa being any amino acid. But non-proline Xaa-non Pro *cis* bonds are also found in proteins, although they occur much less frequently than Xaa-Pro (see A. Jabs *et al.*, JMB, **286**, 291–304 (1999)).

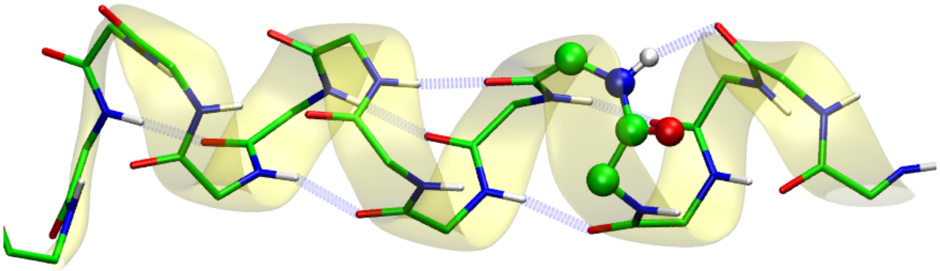


Figure 6: Structurally optimized α-helix containing a *cis*peptide bond. The *cis* peptide bond is shown in CPK, while hydrogen bonds within the helix are shown as thick dashed blue lines. The *cis* configuration of the peptide bond disrupts the hydrogen bond network stabilizing the helix. Note, the hydrogen bond network is broken not only locally.

The configuration of the peptide bond is central to the sort of secondary structure the protein backbone can adopt. It is easy to understand by imagining a normal α helix and converting a *trans* peptide bond into its *cis* form. The result is that the hydrogen bond network stabilising the helix is broken (see Fig. 6) and the helix will be unstable in long simulations.

2.2 Checking configuration of peptide bonds

To demonstrate the usage of the plugin we will start by checking a structure consisting of the protein EF-Tu in complex with a tRNA-Phe, although it is clear that in this example only the protein part will contain peptide bonds. The used structure is based on the the PDB structure 1OB2 in which errors have been introduced manually.

Currently, the `cispeptide` plugin can be used from the text console and through a GUI. The available commands provided by the pugin can be obtained by typing Type `cispeptide`. The following information should be printed in the console:

```
Usage: cispeptide <command> [args...]
Commands:
  check    -- identify cis peptide bonds
  list     -- list identified cis peptide bonds
  minimize -- fix cis peptide bonds using energy minimization
  move     -- move specified atom to convert to trans
  reset    -- reinitialize plugin state
  restrain -- generate NAMD extrabonds file to restrain peptide bonds
  set      -- define how to modify a given cis peptide bond
  show     -- visualize identified cis peptide bonds
```

In the same way the usage information for any of the provided commands can be obtained. For example, typing `cispeptide restrain` will show the syntax of the command that generates restraints for the dihedral angle at the peptide bond which can be used in NAMD.

In the following, we will describe how to use the `cispeptide` plugin through its GUI, but all of the tasks can also be accomplished from the console.

1 Load the files `pept_testcase.psf` and `pept_testcase.pdb` into a new session of VMD.

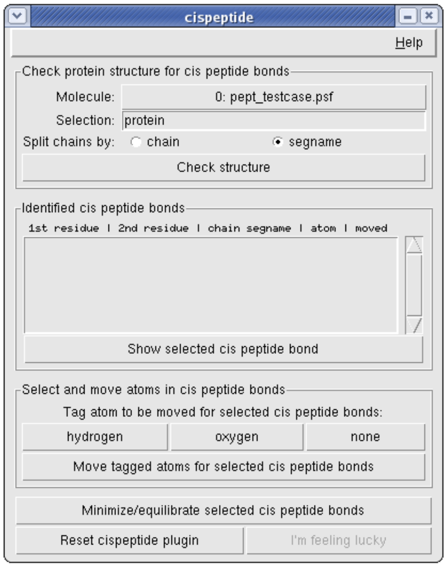
2 Open the `cispeptide` GUI by selecting Extensions → Modeling → Fix Cis Peptide Bonds in the VMD Main menu. In the upper part of the GUI (see Fig. 7) the user can specify the molecule and the atom selection to be tested.

3 The button Check structure initializes the test. Peptides identified to be in *cis* configuration are displayed in the Identified cis peptide bonds form (see Fig. 8). In the present case there are 4 *cis* peptide bonds identified in EF-Tu.

4 By selecting an entry in the Identified cis peptide bonds form and hitting the button Show selected cis peptide bond it is possible to visually inspect all the individual identified *cis* peptides. In the created representation (see Fig. 9) the *cis* peptide bond is highlighted in CPK representation.

2.3 Converting *cis* peptides into *trans* configuration

WARNING: The `cispeptide` plugin identifies **all** *cis* peptide bonds. As noted in the introduction, there are cases in which *cis* peptides do occur in proteins.


cispeptide
Help
Check protein structure for cis peptide bonds
Check structure
Identified cis peptide bonds
Show selected cis peptide bond
Select and move atoms in cis peptide bonds
Tag atom to be moved for selected cis peptide bonds:
hydrogen
oxygen
none
Move tagged atoms for selected cis peptide bonds
Minimize/equilibrate selected cis peptide bonds
Reset cispeptide plugin
I'm feeling lucky

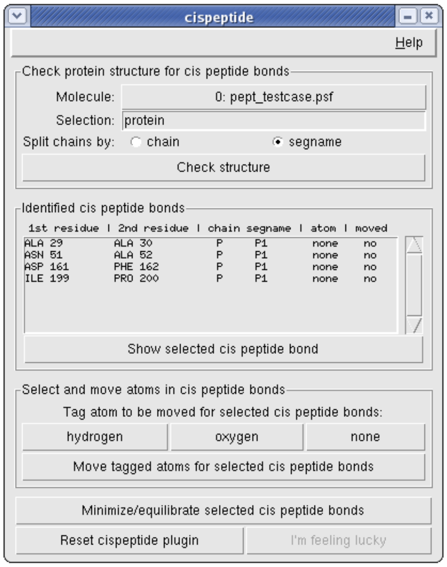
Molecule: 0: pept_testcase.psf

Selection: protein

Split chains by: ○ chain ● segname

```
 1st residue | 2nd residue | chain segname | atom | moved
```

Figure 7: The cispeptide GUI window.


cispeptide
Help
Check protein structure for cis peptide bonds
Check structure
Identified cis peptide bonds
Show selected cis peptide bond
Select and move atoms in cis peptide bonds
Tag atom to be moved for selected cis peptide bonds:
hydrogen
oxygen
none
Move tagged atoms for selected cis peptide bonds
Minimize/equilibrate selected cis peptide bonds
Reset cispeptide plugin
I'm feeling lucky

Molecule: 0: pept_testcase.psf

Selection: protein

Split chains by: ○ chain ● segname

```
 1st residue | 2nd residue | chain segname | atom | moved
ALA 29        ALA 30         P      P1        none   no
ASN 51        ALA 52         P      P1        none   no
ASP 161       PHE 162        P      P1        none   no
ILE 199       PRO 200        P      P1        none   no
```

Figure 8: Identified *cis* peptide bonds in the cispeptide GUI window.

Thus, it should be checked if a given *cis* peptide is real or is a mistake. This can be done by literature research, by comparing high resolution crystallographic

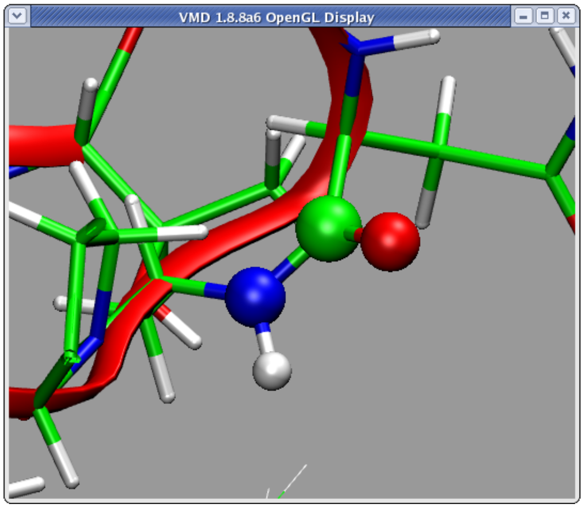


Figure 9: Representation of the *cis* peptide bond between Asn 51 and Ala 52 within EF-Tu

models, or by comparing crystallographic models of similar proteins.

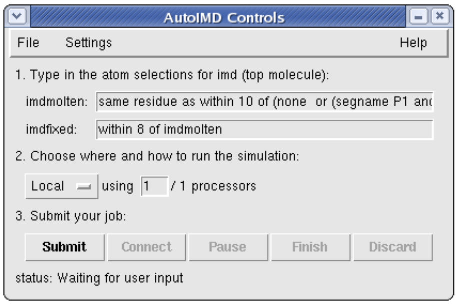
1 If the shown peptide bond is supposed to have a *trans* configuration, the user can tag the atom to be moved to flip the configuration from *cis* to *trans*. This is done by hitting the button hydrogen or oxygen.

WARNING: The choice which atom to move is crucial and has to be made based on the local structure around the peptide under consideration. The decision is particularly critical helices and β-sheets: the relaxed backbone should integrate in the existing hydrogen bond network characteristic for the specific secondary structure motif (see Fig. 6).

If the identified *cis* peptide bond is supposed to be there, nothing needs to be done. It is also possible to untag a tagged atom by selecting the corresponding *cis* peptide and hitting none.

2 Once a *cis* peptide bond has been inspected, and the atom to be moved has been tagged, the actual moving of the atom can be executed by hitting the button Move tagged atoms for selected cis peptide bonds.

3 Since the described procedure (simple moving of one atom) generates an unphysical geometry of the molecule, it is necessary to optimize the structure using an MD force field. The `cispeptide` plugin uses the `AutoIMD` plugin for this purpose. This final step is accomplished by selecting the *cis* peptides which should be relaxed and hitting the button Minimize/equilibrate selected cis peptide bonds. This will open the AutoIMD Controls window (see Fig. 10).



1. Type in the atom selections for imd (top molecule):

imdmolten: same residue as within 10 of (none or (segname P1 anc

imdfixed: within 8 of imdmolten

2. Choose where and how to run the simulation:

Local using 1 / 1 processors

3. Submit your job:

status: Waiting for user input

Figure 10: AutoIMD window called by `cispeptide` plugin.

4 Select Settings → Minimization Mode in the AutoIMD Controls window and hit then the Submit button. Once the `AutoIMD` session started, hit the Connect button and a minimization will start. Finish the simulation when the structure relaxed. Usually, it takes up to a few thousand steps until atoms do not move anymore. Is the structure minimized, hit the Finish in the AutoIMD Controls window. For more details about using the `AutoIMD` plugin, the user is referred to the AutoIMD user's guide http://www.ks.uiuc.edu/Research/vmd/plugins/autoimd/.

5 **WARNING:** Depending on the quality of the initial structure, it can happen that not all errors can be fixed in one run. Thus, it is important to check the final result again and correct the remaining errors, if necessary. In order to check the structure after the performed minimization/equilibration using IMD, the `cispeptide` plugin should be reset. This can be done by hitting the Reset cispeptide plugin button at the bottom of the cispeptide window. Note, in the current implementation it will be necessary to select the molecule again in the top part of the cispeptide window. Note also that if the molecule to correct is solvated, it may be necessary to equilibrate the minimized parts of the structure. This can be done by selecting Settings → Equilibration Mode and submitting a new AutoIMD run.

6 Finally, it is a good advice to minimize the structure in the very last step. This should be done by selecting Settings → Minimization Mode in the AutoIMD Controls window, which is still open from the last `AutoIMD` session, and hit then the Submit button. As mentioned in the introduction, additional restraints during the correction of *cis* peptides. The restraints will be used always when `AutoIMD` is started from the cispeptide window. The equilibrium values for these restraints are very approximate and bring the configuration in the correct region. Thus, in order to obtain a optimized structure, it is necessary NOT to use the restraints defined by `cispeptide`. This is achieved by reusing the AutoIMD window from the previous minimization/equilibration; the restraints will not be taken into account.

7 Once all minimization/equilibration steps are performed, the corrected structure should be saved. This can be done in the following way. Select the molecule in the VMD main window. Create a representation containing all atoms in the Graphical Representations window. Select File → Save Coordinates in the VMD main window. Select `all` in the Selected atoms and hit Save. Finally, specify the file name in the Chose filename to save trajectory and hit OK.

2.4 Preserving configuration of peptide bonds during MD simulation

As mentioned in the introduction, structure optimizations and MD protocols applying external forces can flip the configuration of a peptide bond in the the simulated structure. To preserve peptide configuration, the `cispeptide` plugin offers the possibility to generate restraints in form of `extrabonds` which can be used in a NAMD simulations. This feature currently is available only from the console.

1 Assuming the structure of interest is loaded in the top molecule in VMD and one wishes to obtain restraints for all peptide bonds in the system, the command reads

```
cispeptide restrain -o peptide-extrabonds.txt
```

The file `peptide-extrabonds.txt` contains now one dihedral restraint for each peptide bond. The equilibrium value of this dihedral is set to the value present in the used structure.

WARNING: Note, the specified restraints modify the potential energy function of the system, and thus should not be used in equilibrium production runs. Rather, these restraints are supposed to be used only in simulations in which large forces occur due to poor initial geometry or due to externally applied forces, e.g., MDFF or TMD. Note also that although the final structure resulting from such simulations will be correct, the behavior of the system during the simulation is artificial.

Acknowledgements

Development of this tutorial was supported by the National Institutes of Health (NIH 9P41GM104601 - Resource for Macromolecular Modeling and Bioinformatics).