PNNL-14890

HDViewer

User's Manual

Version 1.0

September 2004

Dr. Timothy K. Firman,[†] Dr. Benjamin P. Hay,[†]
and Dr. Kevin E. Gilbert[§]

[†] *Molecular Interactions and Transformations Group, Chemical Sciences Division, Pacific Northwest National Laboratory, PO Box 999, Richland, WA 99352*

[§]*Serena Software, Box 3076, Bloomington, IN 47402*

Parts of this document were updated December 2014 by code administrator Benjamin Hay, now at the Supramolecular Design Institute, Oak Ridge, TN 37830.

Email: hayben@comcast.net

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

HDViewer is provided by the authors "as is" and any express or implied warranties of merchantability and fitness for a particular purpose are disclaimed. In no event shall the author be liable for any direct, indirect, incidental, special, exemplary, or consequential damages (including, but not limited to, procurement of substitute goods or services; loss of use; data or profits; or business interruption) however caused and on any theory of liability, whether in contract, strict liability, or tort (including negligence or otherwise) arising in any way out of the use of this software, even if advised of the possibility of such damage.

ACKNOWLEDGMENT

The concept of creating HDViewer arose during a collaboration with PCModel® author Kevin Gilbert to add an automated molecular mechanics post-processing capability to HostDesigner (a work still in progress). The HDViewer code, which is adapted from the graphic user interface of PCModel®, was written by Kevin. Beta versions were tested by TKF and BPH during the course of their host design research funded through the Environmental Management Science Program and the Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Sciences, Office of Science, U.S. DOE. Beta testing was performed at the Pacific Northwest National Laboratory (PNNL), operated by Battelle for the DOE under contract DE-AC06-76RLO 1830.

TABLE OF CONTENTS

1.0 INTRODUCTION

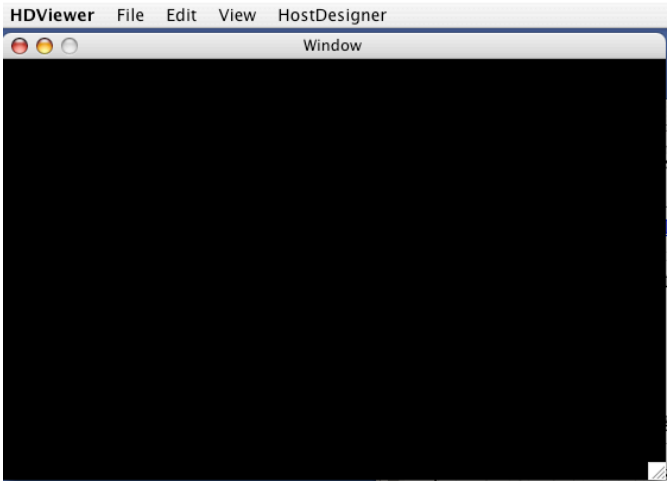
This manual describes the features of HDViewer, a graphic user interface for HostDesigner. HDViewer can import structures from several molecule file formats, generate the host fragment input files and control files needed to run HostDesigner, view HostDesigner output files, and export structures in several molecule file formats.

The HostDesigner package can be downloaded at no cost from the http://sourceforge.net/ website. It consists of the HostDesigner source code, HDViewer executables (MacOSX or Windows), a linkage structure library (LIBRARY), and one additional data file needed to operate the system (CONSTANTS). In addition, the download includes several example input files and the user manuals.

2.0 INSTALLATION

HDViewer is a stand alone executable. Simply drag and drop the HDViewer icon into the applications folder.

3.0 THE HDVIEWER INTERFACE



After startup HDViewer presents a menu bar and a window in which a structure will be displayed.

The 'File' menu contains standard I/O operations 'Open', 'Save', 'Page Setup', and 'Print'. The 'Open' and 'Save' options are discussed in more detail in Section 4. The 'Print' option may be used to print or save to a .pdf file the structure depicted on the main screen.

The 'Edit' menu options are 'Delete', 'Query', and 'Set Title'. 'Delete' removes the current structure. 'Query' is used to obtain distance, angle, and dihedral angle information (see Section 3.3 below). 'Set Title' allows the user to give a name to the structure that will be displayed in the title bar above the structure window (currently showing 'Window'). This name is also used in some file formats. If present, it will be read in and displayed on the structure window title bar.

The 'View' menu options control the appearance of the molecule. Choices include 'Stick', 'Tubes', 'Ball and Stick', and 'CPK'. By default, the 'Stick' view is selected. There is also a 'Labels' sub menu to allow the display of textual information. Options are 'Atom Symbols', 'Bonds Only', 'Atom Numbers' (serial numbers), 'Atom Type' (MM3 atom type numbers), and 'Atomic Charge'. By default, 'Labels' is set to 'Bonds Only'.

The ‘HostDesigner’ menu options ‘Control File’, ‘Linker Host File’, ‘Overlay Host File’, and ‘Run HostDesigner’ are discussed in detail in Sections 4, 5, and 6.

3.1 Selecting Atoms

There are several instances where the user is required to select atoms in a structure. These occur when creating HostDesigner fragment files (see Section 5) and when in Query mode (see Section 3.3). When selecting an atom it is recommended to use the ‘Stick’ option under View in order to clearly see the atom center. To do so, move the cursor on in the main window over an atom. A small square appears to highlight the atom currently being pointed at. Click (left-click in Windows systems) to select the highlighted atom. The atom will now show as a small ball, indicating that it has been selected. If you wish to select multiple atoms, continue clicking on new atoms. Clicking an already selected atom toggles it off. Clicking anywhere in the window not on an atom clears all selected atoms, unless HDViewer is in Query mode.

3.2 Rotating and Translating Structures

A structure can be rotated by holding down the mouse button (right button in Windows systems) with the cursor on the background and moving the mouse. The same operation will translate the structure when the appropriate key is depressed (option key for MacOSX, Ctrl key for Windows).

3.3 Query mode

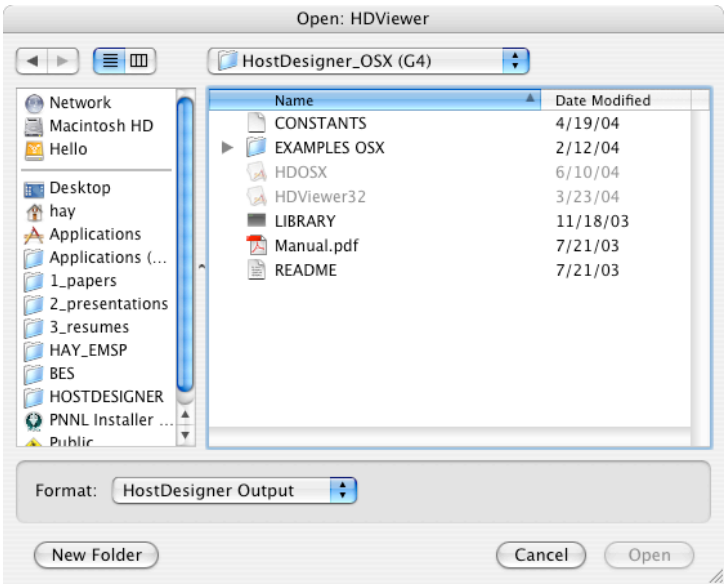
Query mode is used to obtain geometric information about the structure. By default ‘Query’ mode is off. When the option is selected from the ‘Edit’ menu a check appears next to ‘Query’ to indicate that it is on. There are four types of queries, which are apply depending on how many atoms are selected:

- Selecting one atom, then clicking any location not an atom, places a label of the selected atom’s xyz coordinates at the final location (and a partial charge label)
- Selecting two atoms, then clicking another location gives the interatomic distance
- Selecting three atoms then clicking another location gives the angle formed by those atoms
- Selecting four atoms then clicking another location gives the dihedral angle formed by those atoms

Unselecting the Query option will clear all labels. If the Query mode is on and you enter another mode in which you need to select atoms, either of the Link Host File (Section 5.1) or the Overlay Host File (Section 5.2) options, then Query mode will become inactive and no new labels may be assigned, though the existing ones will remain.

4.0 SUPPORTED INPUT AND OUTPUT FILE TYPES

HDViewer can read molecular structures in several file formats including HostDesigner output, HostDesigner fragment, PCModel®, Chem3D® Cartesian Coordinates 1, Protein Data Bank, SDF MOL, and XYZ. HDViewer can write structures in several formats including HostDesigner fragment, PCModel®, Chem3D® Cartesian Coordinates 1, SDF MOL, and PDB. Examples for each of these formats are given later in this section.



A structure is imported by selecting 'File' from the menu bar followed by 'Open' in the drop down menu. This brings up a standard file open window shown above. Choose the desired file format from a drop down menu toward the bottom of this window, currently showing "HostDesigner Output".

For each atom in a structure, the minimum information required to create structure input files for HostDesigner is the atom label, Cartesian coordinates in angstroms, atom type number (either MMX, MM2, or MM3 atom types are acceptable), and a list of connected atoms. Some input formats provide all this information, but others do not. The information included in the different input format types is summarized below.

Input file format	Cartesian Coordinates	Atomic Labels	Atom Type Numbers	Connectivity Table
HD Output	Yes	Yes	Optional	Optional
HD Fragment	Yes	Yes	Yes	Yes
PCModel	Yes	No	Yes	Yes
Chem3D.CC1	Yes	Yes	Yes	Yes
PDB	Yes	Yes	No	Yes
SDF	Yes	Yes	No	Yes
XYZ	Yes	Yes	No	No

HDViewer will use whatever information a file provides. HDViewer will attempt to generate missing information. The editing of structure information is not a feature of this viewer. If HDViewer generates incorrect missing information, the structure should be saved in to a format that contains all the needed information (such as the Chem3D.cc1) and then edited as a *text* file by opening it with, for example, TextEdit (Macintosh) or Notepad (Windows). After making any specific changes to the text and saving the corrected file, HDViewer can then read the structure back in.

Here is a brief description of the seven types of input geometry files allowed:

4.1 HostDesigner Output file

This is the file format used by HostDesigner to output lists of structures. This file format is described in detail in Section 3.6 of the HostDesigner User's Manual.

4.2 HostDesigner Fragment file

This is the file format used by HostDesigner to input molecular fragments. This file format is described in detail in Section 3.3 and 3.4 of the HostDesigner User's Manual.

4.3 PCModel® file

This is the default file format of Serena Software's molecular modeling software PCModel®. Here is an example of a file in PCModel® format (for methanol):

```
{PCM Methanol
NA 6
FL EINT4 UV1 PIPL1
SSNAME
AT 1 1  -0.7205   0.0111   0.0192 B 2 1 3 1 4 1 5 1 S 0  C 0.28
AT 2 6   0.6952   0.0095   0.0164 B 1 1 6 1 S 0  C -0.68
AT 3 5  -1.0847   1.0413   0.0125 B 1 1 S 0
AT 4 5  -1.0847  -0.5096   0.9083 B 1 1 S 0
AT 5 5  -1.0782  -0.5053  -0.8747 B 1 1 S 0
AT 6 21   0.9830   0.4740   0.8207 B 2 1 S 0  C 0.4
}
```

This format includes atom type numbers, Cartesian coordinates, connectivity list with bond orders, and atom charges, but is missing atom labels. HDViewer will automatically derive atomic labels from the atom type numbers in this format (assuming them to be MMX atom types). When an atom type is not recognized, then the label for that atom will be left blank. It is strongly recommended that the user check that the labels derived by HDViewer are as desired.

4.4 Chem3D® 'Cartesian Coordinates 1' File

This is an output based on a CambridgeSoft's Chem3D® 'Cartesian Coordinates 1' file format. When generating this file with Chem3D®, there are a number of options available and it is important that the correct options are specified. In this case, Connection Table should be 'By Serial Number' and the boxes for 'Include Serial Numbers' and 'Include Atom Type Text Numbers' should be checked.

Here is an example of a file in this format (for methanol):

```
6
  C     1   -0.720548    0.011053    0.019205    1     2     3     4     5
  O     2    0.695182    0.009480    0.016444    6     1     6
  H     3   -1.084684    1.041335    0.012469    5     1
  H     4   -1.084661   -0.509614    0.908277    5     1
  H     5   -1.078228   -0.505271   -0.874717    5     1
  H     6    0.983028    0.474010    0.820695   21     2
```

For each atom, the atom label is first listed, then the serial number and the three Cartesian Coordinates. The next column is the atom type (MM2 types if created with Chem3D®) and finally the serial number of each of the atoms attached is listed. This format includes all the information HDViewer needs explicitly.

4.5 PDB File

This is an output based on the Protein Data Bank file format. Here is an example of a file in this format (for methanol):

```
HEADER    MOLECULE
COMPND    Untitled
AUTHOR    GENERATED BY PCMODEL V8.5
HETATM    1  C   UNK   1         -0.721   0.011   0.019  1.00  0.00
HETATM    2  O   UNK   1          0.695   0.009   0.016  1.00  0.00
HETATM    3  H   UNK   1         -1.085   1.041   0.012  1.00  0.00
HETATM    4  H   UNK   1         -1.085  -0.510   0.908  1.00  0.00
HETATM    5  H   UNK   1         -1.078  -0.505  -0.875  1.00  0.00
HETATM    6  H   UNK   1          0.983   0.474   0.821  1.00  0.00
CONECT    1    2    3    4    5
CONECT    2    1    6
CONECT    3    1
CONECT    4    1
CONECT    5    1
CONECT    6    2
MASTER        0    0    0    0    0    0    0    0    6    0    6    0
END
```

This format includes Cartesian coordinates, a connectivity list, and atomic labels but does not include atom type numbers. HDViewer will automatically derive MMX atom type numbers from the atom labels and connectivity data. In cases where HDViewer is not able to recognize an atom type, then it is assigned a value of 300. It is strongly recommended that the user check that the atom types derived by HDViewer are as desired.

4.6 SDF Mol File

This is an output based on the .sdf, or .mol file format. Here is an example of a file in this format (for methanol):

```
SDF  Methanol
PCMODEL  v8.5   1.00000     0.00000

  6  5  0  0  0  0              1 V2000
   -0.7205    0.0111    0.0192 C   0  0  0  0  0  0
    0.6952    0.0095    0.0164 O   0  0  0  0  0  0
   -1.1025   -0.5102   -0.8832 H   0  0  0  0  0  0
   -1.0990   -0.5098    0.9233 H   0  0  0  0  0  0
   -1.0990    1.0544    0.0198 H   0  0  0  0  0  0
    0.9830    0.4740    0.8207 H   0  0  0  0  0  0
  1  2  1  0
  1  3  1  0
  1  4  1  0
  1  5  1  0
  2  6  1  0
M END
Untitled

$$$$
```

This format includes Cartesian coordinates, a connectivity table with bond orders, and atomic labels but does not include atom type numbers. HDViewer will automatically derive MMX atom type numbers from the atom labels and connectivity data. In cases where HDViewer is not able to recognize an atom type, then it is assigned a value of 300. It is strongly recommended that the user check that the atom types derived by HDViewer are as desired.

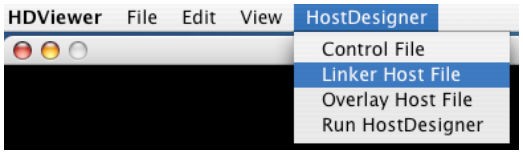
4.7 XYZ File

This is the simplest input format. Here's a sample input for methanol:

```
6
Methanol
  C   -0.720548    0.011053    0.019205
  O    0.695182    0.009480    0.016444
  H   -1.084684    1.041335    0.012469
  H   -1.084661   -0.509614    0.908277
  H   -1.078228   -0.505271   -0.874717
  H    0.983028    0.474010    0.820695
```

This format includes only Cartesian coordinates and atomic labels. Thus, HDViewer must assign both the atom types and the connectivity lists. This should work acceptably for many simple organic molecules, but may cause trouble with host – guest systems. As with other cases where HDViewer generates missing information (Section 4.3, 4.5, and 4.6), it is strongly recommended that the user check that the connectivity and the atom types derived by HDViewer are as desired.

5.0 GENERATING HOSTDESIGNER COMPLEX FRAGMENT INPUT FILES



The 'HostDesigner' menu includes options for generating the two types of structure input files used by HostDesigner. These options, 'Linker Host File' and 'Overlay Host File', are described in this section. To use either option, an input structure must already have been imported (see Section 3) and four characteristics for each atom (Cartesian coordinates, atomic labels, atom type numbers, and connectivity list) should be correct.

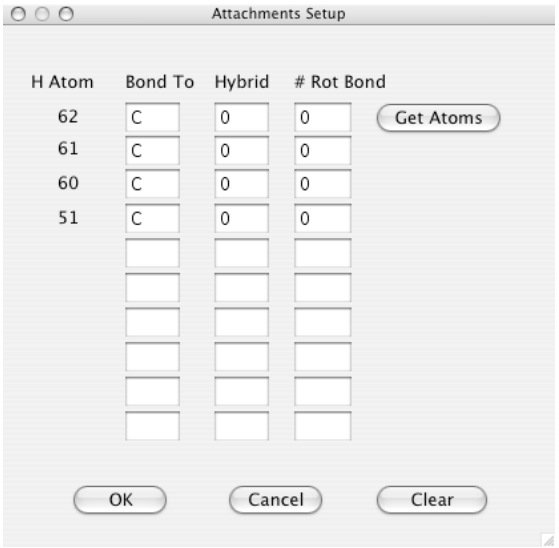
5.1 Linker Input Complex Fragments

This section describes how to use HDViewer to generate complex fragment input files required by the LINKER mode of HostDesigner. The format of these files is described in detail in Section 3.3 of the HostDesigner User's Manual.

The 'Linker Host File' option under the 'HostDesigner' menu brings up the 'Link File Setup' window pictured to the right. There are four buttons here to use to assign settings.

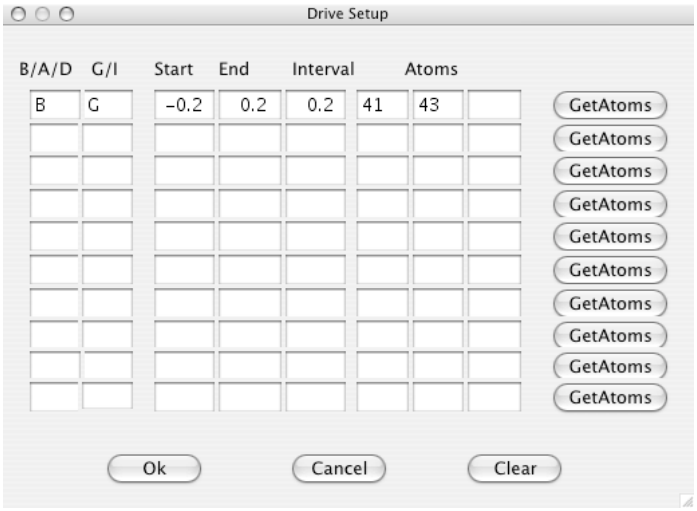
Every complex fragment must have at least one atom that is identified as the guest in order for the LINKER algorithm to work. The 'Guest Atom(s)' button is used for this purpose. Select all of the atoms that belong to the guest and then click the 'Guest Atom(s)' button. The number now displayed next to the 'Guest Atom(s)' button is the number of atoms that have been selected for the guest.

The 'Macro Atom' button is an option to specify one terminal methyl carbon that may be used to close a ring to form a macrocycle. Specifying a macro atom is strictly *optional* and is only needed when the **macro** keyword will be used in the CONTROL file. The number displayed in next to the 'Macro Atom' button is the serial number of the single atom selected.



The "Attachment Points" button brings up the 'Attachment Setup' window. This is used to assign the positions on the structure where HostDesigner will form bonds with links, and to assign characteristics and restrictions on which links are used at each location. Once 'Attachments Setup' window is open, select one to ten hydrogen atoms and click the 'Get Atoms' button. The serial numbers of the atoms selected will appear on the left, and default values for the other three options for each of these attachment points appear in the next three columns. A more detailed description of what values should be used for the following three columns appears in Section 3.3 of the HostDesigner

User's Manual, where the four columns correspond to the information in Lines 4+n to 3+n+na in that section. Clicking 'OK' will save the displayed settings, clicking 'Cancel' will leave the 'Attachment Setup' window as it was before it was last invoked, and clicking 'Clear' clears all settings in this window.



The 'Drive Setup' button is optional. It is used to specify geometry drives as described in Section 3.5 of the HostDesigner User's Manual. To add a drive, select two to four atoms and click a "Get Atoms" button. This will assign the drive type to be a B (bond or distance), A (angle), or D (dihedral) if 2, 3, or 4 atoms are chosen, respectively. Only the central two atoms are needed for HostDesigner to do a dihedral drive, so only those two are shown in that case. The G/I column defaults to I (internal) division. The user must add the start, end, and interval for each drive by clicking in the box and typing in the values. A total of 10 drives can be specified. Clicking 'OK' will save the displayed settings, clicking 'Cancel' will leave the 'Drive Setup' window as it was before it was last invoked, and clicking 'Clear' clears all settings in this window.

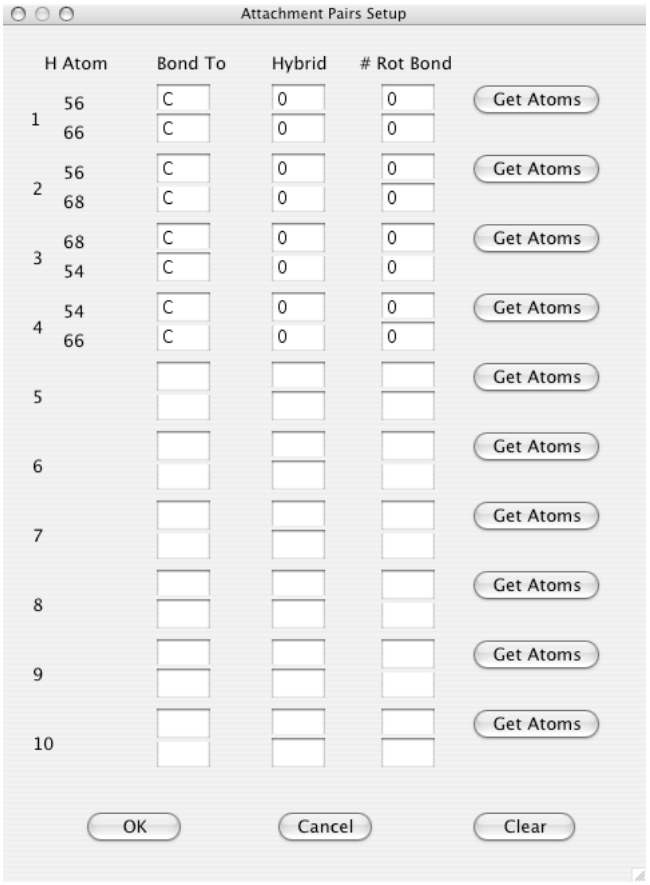
Once all settings have been made, a complex fragment input file for LINKER can be saved by clicking 'Ok' on the 'Link File Setup' window. This brings up a 'Save Host File' window that allows you to type in a file name and to specify a destination for the file. Clicking the 'Save' button writes the file.

The 'Cancel' button on the 'Link File Setup' window allows an exit without saving a complex fragment file.

5.2 Overlay Input Host Fragments

This section describes how to use HDViewer to generate complex fragment input files required by the OVERLAY mode of HostDesigner. The format of these files is described in detail in Section 3.4 of the HostDesigner User's Manual. As the formats are very similar, the options for OVERLAY are very similar to those used in LINKER, with the primary differences being that the macrocycle option is not applicable to OVERLAY runs and that attachment points are specified in a pairwise fashion.

The "Guest Atom(s)" selection works as described in Section 5.1.



The 'Attachment Pairs' menu button brings up the 'Attachment Pairs Setup' window. This is used to assign pairs of positions on the structure where HostDesigner will form bonds with links, and to assign characteristics and restrictions on which links are used at each location. For each pair, select two hydrogen atoms and click a "Get Atoms" button. The serial numbers of the two atoms will appear on the left and the defaults for the other attributes will appear in the three right columns. Each pair should be chosen once (order does not matter). As with previous menus, clicking 'Ok' button saves the displayed settings, clicking 'Cancel' will leave the 'Attachment Setup' window as it was before it was last invoked, and clicking 'Clear' clears all settings in this window.

The optional 'Driver Setup' button brings up the same 'Drive Setup' window described above in Section 5.1.

Once desired settings have been made, a complex fragment input file for use with OVERLAY can be saved by clicking 'Ok' on the 'Overlay File Setup' window. This brings up a 'Save Host File' window that allows you to type in a file name and to specify a destination for the file. Clicking the 'Save' button writes the file.

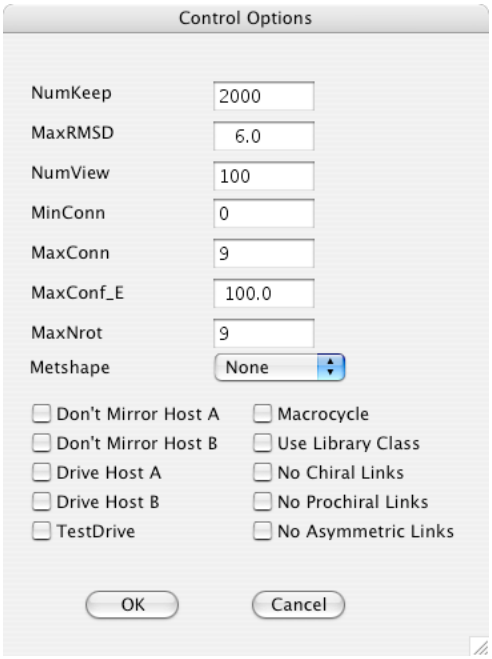
The 'Cancel' button on the 'Overlay File Setup' window allows an exit without saving a complex fragment file.

6.0 GENERATING HOSTDESIGNER CONTROL FILES

This section describes how to use HDViewer to generate 'control' files required by both the LINKER and OVERLAY mode of HostDesigner. The 'control' file contains instructions for how to execute the current job. The keywords and ramifications of using each keyword are described in detail in the in Section 3.2 of the HostDesigner User's Manual, so the descriptions given here will be brief and with the intention of clarifying how to set the various options in the interface rather than explaining them in any detail.

Selecting 'Control File' under the 'HostDesigner' pull down menu opens the 'Control Setup' window. At the top are two buttons for the mandatory selection of the type of run to do, LINKER or OVERLAY. The next three entries specify input file names. HostA is used to specify the first complex fragment file to be used (or the only one, in the case of OVERLAY). Clicking on the 'HostA' button (Keyword: **hosta=**) brings up a file selection window in which you should click on a fragment file, presumably created using the instructions in sections 5.1 or 5.2. In an identical fashion, the 'HostB' button (Keyword: **hostb=**) is used to specify the name of the second complex fragment file to be used (not needed for OVERLAY). The 'Link Library' button (Keyword: **linklib=**) is used to specify an alternate link library location. The default setting specifies the standard LIBRARY supplied with the distribution. The 'Output Name' field (Keyword: **out=**) specifies the prefix of the output files that will be saved.

The Options button invokes the 'Control Options' window in which many other settings may be adjusted.



The seven input windows and the 'Metshape' pulldown menu all correspond to their respective keywords in the HostDesigner menu. The ten clickable options represent the following keywords as described in the HostDesigner User's Manual, Section 3.2:

mirroraoff	**macro**
mirrorboff	**useclass**
drivea	**nochiral**
driveb	**noprochiral**
testdrive	**noasym**

The 'OK' button saves the set options, and the 'Cancel' button reverts to the last saved version.

The ‘Read Control’ button on the ‘Control Setup’ window will bring up a browse menu to find and read in an existing HostDesigner control file, which will replace all the options currently set in the window.

Once desired settings have been made, a control file can be saved by clicking ‘Ok’ in the ‘Control Setup’ window. This brings up a ‘Save Control File’ window that allows you to type in a file name and to specify a destination for the file. Clicking the ‘Save’ button writes the file.

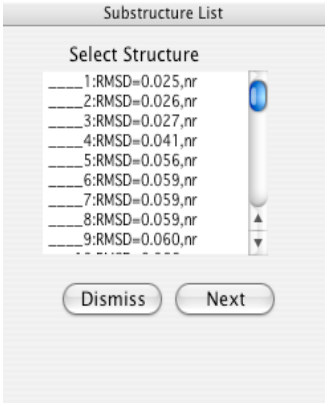
The ‘Cancel’ button on the ‘Control Setup’ window exits without saving a control file.

7.0 LAUNCHING HOSTDESIGNER

This feature no longer works.

8.0 VIEWING HOSTDESIGNER OUTPUT FILES

When HostDesigner executes it writes two output files that contain lists of Cartesian coordinates for the host structures that were generated. These files will be named ‘prefix_1.hdo’ and ‘prefix_2.hdo’ using the prefix specified by the **out=** keyword (see Section 6 above). The format of the .hdo files is described in detail in Section 3.6 of the HostDesigner User’s Manual.



When HDViewer opens a HostDesigner output file (see Section 4.0 above), the first structure in the list and a ‘Substructure List’ window appear in the main window. The structures may be stepped through using the ‘Next’ button or scrolled through and selected directly by positioning the cursor over the desired structure in the list and clicking the mouse button. The current selected structure can be written to a file by using the ‘Save’ option under the ‘File’ drop down menu. Clicking the ‘Dismiss’ button of the ‘Substructure List’ window will close the window, but leave the selected structure displayed in the main window.

As each structure is selected, the title bar of the structure window is changed to reflect the additional information about the structure (see below for an example). After the HDViewer v1.0 identifier, the rest of this title contains information relevant to how the structure was scored and the identity of the link used to build it. The information consists of the position in the list, the RMSD, the number of bonds that would have restricted rotation on guest complexation, a conformational energy estimate for the host, the formal name of the hydrocarbon from which the link was derived, and the serial number of the link. The first 19 characters of this information appear in the ‘Substructure List’ window for each structure.

9.0 HOW TO CITE HDVIEWER IN THE LITERATURE

In publishing results obtained either in part of in full from use of HDViewer, the user should use the following citation:

Citation:

(a) Firman, T. K.; Hay, B. P.; Gilbert, K. E., *HDViewer 1.0 User's Manual*, 2004, Pacific Northwest National Laboratory, Richland, WA 99352. (b) *HDViewer* is part of the HostDesigner download package, available at no cost from the following website: http://hostdesigner20.sourceforge.net.