



OpenEye
Scientific Software

EON

Release 2.2.0.5

OpenEye Scientific Software, Inc.

June 03, 2013

CONTENTS

1	Front Matter	1
2	Installation and Platform Notes	3
2.1	Licenses	3
2.2	General Installation	3
2.3	MPI	5
2.4	Uninstallation	7
3	Usage	9
3.1	Introduction	9
3.2	Command Line Interface	9
3.3	Report File	14
4	Theory	15
5	Release Notes	17
5.1	EON v2.2.0 (<i>June 2013</i>)	17
5.2	EON v2.1.0	17
5.3	EON v2.0.0	18
6	List of Selected EON Publications	21
7	Citation	23
	Index	25

FRONT MATTER

Copyright 1997-2013 OpenEye Scientific Software, Santa Fe, New Mexico. All rights reserved.

All rights reserved. This material contains proprietary information of OpenEye Scientific Software. Use of copyright notice is precautionary only and does not imply publication or disclosure.

The information supplied in this document is believed to be true but no liability is assumed for its use or the infringement of the rights of others resulting from its use. Information in this document is subject to change without notice and does not represent a commitment on the part of OpenEye Scientific Software.

This package is sold/licensed/distributed subject to the condition that it shall not, by way of trade or otherwise, be lent, re-sold, hired out or otherwise circulated without OpenEye Scientific Software's prior consent, in any form of packaging or cover other than that in which it was produced. No part of this manual or accompanying documentation, may be reproduced, stored in a retrieval system on optical or magnetic disk, tape, CD, DVD or other medium, or transmitted in any form or by any means, electronic, mechanical, photocopying recording or otherwise for any purpose other than for the purchaser's personal use without a legal agreement or other written permission granted by OpenEye.

This product should not be used in the planning, construction, maintenance, operation or use of any nuclear facility nor the flight, navigation or communication of aircraft or ground support equipment. OpenEye Scientific Software, shall not be liable, in whole or in part, for any claims arising from such use, including death, bankruptcy or outbreak of war.

Windows is a registered trademark of Microsoft Corporation. Apple, OS X, and Macintosh are registered trademarks of Apple Inc. UNIX is a registered trademark of the Open Group. Linux is a registered trademark of Linus Torvalds. Red Hat is a registered trademark of Red Hat, Inc. SUSE, SLED and SLES are registered trademarks of Novell, Inc. Ubuntu is a registered trademark of Canonical Ltd.

SYBYL is a registered trademark of Tripos, L.P. MDL is a registered trademark and ISIS is a trademark of Accelrys Software Inc. SMILES is a trademark, SMARTS and SMIRKS are registered trademarks of Daylight Chemical Information Systems, Inc. MacroModel is a trademark of Schrodinger, LLC.

Python is a registered trademark of the Python Software Foundation. Django is a registered trademark of the Django Software Foundation. Java is a registered trademark of Oracle and/or its affiliates.

Other products and software packages referenced in this document are trademarks and registered trademarks of their respective vendors or manufacturers.

INSTALLATION AND PLATFORM NOTES

2.1 Licenses

To run EON you will need to obtain a license file for EON from OpenEye Scientific Software (business@eyesopen.com). The license file should be pointed to by the environment variable **OE_LICENSE**.

If you intend on running multi-processor EON via Open MPI, only the master machine needs access to the license file.

On Windows, the environment variables can be set under the system Control Panel.

2.2 General Installation

2.2.1 Linux

Linux distributions are provided as a gzipped tarball of the distribution tree described below. Installation is performed by simply untarring the file in the desired location. The top-level directory in the tarball is named `openeye`. Distributions for different Linux variants can be installed into the same location, allowing multiple Linux versions to be run from a single shared directory.

To ensure that the installed applications can be called from the command line, be sure to add the full path of the `openeye/bin` subdirectory to the **PATH** environment variable. For instance, if the distribution was installed into `/usr/local/openeye`, the **PATH** environment variable should contain: `/usr/local/openeye/bin`.

Under the top-level `openeye` directory are the following subdirectories:

- arch** This directory contains the collection of platform specific subdirectories. Each subdirectory contains the actual installed executables and support libraries for the associated platform. In the platform specific subdirectory there will be a subdirectory for each application. Within that will be another subdirectory for each version of that application.
- bin** This directory contains a startup script for each application that has been installed. This script determines, at run-time, what the current platform is and then calls the appropriate executable in the `arch`. This script enables the easy co-existence of multiple platforms and versions of any OpenEye application in the same distribution tree.
- data** This directory contains all of the associated data for the installed applications. There will be a subdirectory for each installed application and within that subdirectory there will be another subdirectory for each specific version of that application.

docs This directory contains all of the documentation associated with the installed applications. There will be a subdirectory for each installed application and within that subdirectory there will be another subdirectory for each specific version of that application.

examples This directory contains all of the examples associated with the installed applications. There will be a subdirectory for each installed application and within that subdirectory there will be another subdirectory for each specific version of that application.

The startup script discussed in the section on the `bin` directory above will have the same name as the installed executable with which it is associated. When the script is called, it will attempt to determine the current platform and run the appropriate executable if installed. If an appropriate executable cannot be found, the script will report that information, as well as a list of the currently installed platforms. The auto-detection can be overridden by setting one of two environment variables:

- **OE_ARCH** can be used to specify a colon separated list of compatible distributions for the current platform such as:

```
redhat-RHEL5-x64:redhat-RHEL4-x64
```

Specification of this environment variable overrides the auto-detection process, if it is present. If none of the compatible distributions listed are found, the script will fall back to the auto-detection process.

- **APPNAME_OE_ARCH** can be used to specify a colon separated list of compatible distributions for a specific application (as specified by changing the **APPNAME** text in the environment variable name) just like **OE_ARCH** as detailed above.

Specification of this environment variable overrides the **OE_ARCH** environment variable as well as the auto-detection process. If none of the compatible distributions listed are found, the script will fall back to the **OE_ARCH** list first and then to the auto-detection process.

Specifying this variable provides a simple way to customize the behavior for individual applications on non-standard platforms.

The startup script also supports a few commandline arguments including:

- | | |
|---------------------|--|
| -path | Specifying this argument will output the full path of the executable to be run. The executable will not be started if this argument is present. |
| -print_arch | Specifying this argument will output the details of the current platform as detected by the script as well as which platform-version of the executable is being run. The executable will be started if this argument is present. |
| -use_version | Specifying this argument followed by a specific version number allows the user to control which released version of the executable to run. |

2.2.2 Windows

Windows distributions are provided as a standard EXE installer. For installation double click the executable and follow the installation instructions. By default, OpenEye applications will install into the `C:\Program Files (x86)` directory (for 32-bit applications) or `C:\Program Files` (for 64-bit applications).

Under the application directory (`C:\Program Files\Application Name`) there are subdirectories for:

bin This directory contains the application executable.

data This directory contains all of the associated data for the installed applications.

docs This directory contains all of the documentation associated with the installed applications.

examples This directory contains all of the examples associated with the installed applications.

An OpenEye group with an application specific subgroup will be added to the *Start* menu. The application specific subgroup will contain links to the documentation, the uninstaller, and, for some applications, a Windows command shell with **PATH** settings already defined to allow the user to simply type the executable name at the prompt without concern for where the executable is actually installed. Links are also included to add and remove the installed location to the user's default path.

For graphical applications, a link to the application will be created on the desktop as well as in the application specific subgroup of the *Start* menu.

2.2.3 Mac OS X

Mac OS X distributions are provided as a *dmg* disk image. For installation, double click the *.dmg* file to open it, and drag the application to the Applications folder.

A folder containing documentation and example data is included in the disk image (Right click, Show Package Contents). Under the top level *Contents* folder there are subdirectories for:

data This directory contains all of the associated data for the installed applications.

docs This directory contains all of the documentation associated with the installed applications.

MacOS/bin This directory contains the application executable.

The documentation and example data can be copied to any convenient location. Graphical applications have built-in documentation, which is available from the application's Help menu.

For command-line only tools, an application named "Install Command Line Support" can be run from the *.dmg* file, and this will allow the user to add the application's location to the user's **PATH** environment variable. Command-line applications can also be run from the Applications folder, in which case they will open a terminal window with a properly configured environment.

2.3 MPI

2.3.1 Open MPI

This application uses the Open MPI implementation of MPI, found at <http://www.open-mpi.org>. This version includes a full Open MPI install, so no additional software is needed.

There are two requirements to run under Open MPI.

- Every machine in the cluster must have the same version of the application installed.
- The path to the application's *bin* directory must be in the **PATH** environment variable on all machines, and before any other locations that may contain MPI executables (*orted*, *mpirun*, etc).

2.3.2 Using Open MPI

To run under Open MPI on a single machine:

```
prompt> [application executable] -mpi_np [number of processes] [other application options]
```

To run Open MPI on multiple machines, begin by generating a text file that will include the MPI hosts you plan to use and the number of processes on each (for this example, we'll call this file *hosts*). The file should contain a line for each machine with the name of the machine, a space, then *slots=N*, where N is the number of processors for your run. For this example, you would want the file to look like:

```
c1 slots=4
c2 slots=4
c3 slots=4
c4 slots=4
c5 slots=4
```

Now the following command-line will start a job with 1 master and 19 slaves.

```
prompt> [application executable] -mpi_hostfile hosts [other application options]
```

The master will be on c1, where the job is being started. All the i/o for the run will be from the master machine, and all results and logging information will be combined.

2.3.3 Open MPI License

Most files in this release are marked with the copyrights of the organizations who have edited them. The copyrights below are in no particular order and generally reflect members of the Open MPI core team who have contributed code to this release. The copyrights for code used under license from other parties are included in the corresponding files.

```
Copyright (c) 2004-2010 The Trustees of Indiana University and Indiana
                        University Research and Technology
                        Corporation. All rights reserved.
Copyright (c) 2004-2010 The University of Tennessee and The University
                        of Tennessee Research Foundation. All rights
                        reserved.
Copyright (c) 2004-2010 High Performance Computing Center Stuttgart,
                        University of Stuttgart. All rights reserved.
Copyright (c) 2004-2008 The Regents of the University of California.
                        All rights reserved.
Copyright (c) 2006-2010 Los Alamos National Security, LLC. All rights
                        reserved.
Copyright (c) 2006-2010 Cisco Systems, Inc. All rights reserved.
Copyright (c) 2006-2010 Voltaire, Inc. All rights reserved.
Copyright (c) 2006-2011 Sandia National Laboratories. All rights reserved.
Copyright (c) 2006-2010 Sun Microsystems, Inc. All rights reserved.
                        Use is subject to license terms.
Copyright (c) 2006-2010 The University of Houston. All rights reserved.
Copyright (c) 2006-2009 Myricom, Inc. All rights reserved.
Copyright (c) 2007-2008 UT-Battelle, LLC. All rights reserved.
Copyright (c) 2007-2010 IBM Corporation. All rights reserved.
Copyright (c) 1998-2005 Forschungszentrum Juelich, Juelich Supercomputing
                        Centre, Federal Republic of Germany
Copyright (c) 2005-2008 ZIH, TU Dresden, Federal Republic of Germany
Copyright (c) 2007      Evergrid, Inc. All rights reserved.
Copyright (c) 2008      Chelsio, Inc. All rights reserved.
Copyright (c) 2008-2009 Institut National de Recherche en
                        Informatique. All rights reserved.
Copyright (c) 2007      Lawrence Livermore National Security, LLC.
                        All rights reserved.
Copyright (c) 2007-2009 Mellanox Technologies. All rights reserved.
Copyright (c) 2006-2010 QLogic Corporation. All rights reserved.
Copyright (c) 2008-2010 Oak Ridge National Labs. All rights reserved.
```

Copyright (c) 2006-2010 Oracle and/or its affiliates. All rights reserved.
Copyright (c) 2009 Bull SAS. All rights reserved.
Copyright (c) 2010 ARM ltd. All rights reserved.
Copyright (c) 2010-2011 Alex Brick . All rights reserved.
Copyright (c) 2012 The University of Wisconsin-La Crosse. All rights reserved.

Additional copyrights may follow

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer listed in this license in the documentation and/or other materials provided with the distribution.
- Neither the name of the copyright holders nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

The copyright holders provide no reassurances that the source code provided does not infringe any patent, copyright, or any other intellectual property rights of third parties. The copyright holders disclaim any liability to any recipient for claims brought against recipient by any third party for infringement of that parties intellectual property rights.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS 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.

2.4 Uninstallation

2.4.1 Linux

To uninstall a single distribution of a product the relevant subdirectories for that product and version simply need to be deleted from within the following directories:

arch In the `openeye/arch` directory is a platform specific subdirectory. Within this are directories for each installed product and within those are subdirectories for each version of the product. Delete the subdirectory for the version which is to be uninstalled. For example, to delete or uninstall v1.0.0 of a product, delete the folder "`<product_name>/1.0.0`".

data In the `openeye/data` directory is a subdirectory for each installed product and within those are subdirectories for each version of the product. Delete the subdirectory for the version which is to be uninstalled.

docs In the `openeye/docs` directory is a subdirectory for each installed product and within those are subdirectories for each version of the product. Delete the subdirectory for the version which is to be uninstalled.

examples In the `openeye/examples` directory is a subdirectory for each installed product and within those are subdirectories for each version of the product. Delete the subdirectory for the version which is to be uninstalled.

2.4.2 Windows

Installation of an OpenEye product on Windows causes an OpenEye group with an application specific subgroup to be added to the *Start* menu. One of the items in the application specific subgroup is a link to the uninstaller. Clicking on the uninstaller initiates a wizard which guides the user through uninstallation.

For graphical applications, uninstallation also removes the application's link from the desktop and *Start* menu.

2.4.3 Mac OS X

To uninstall a single distribution of an application simply drag the application from the Application folder to the Trash.

Some applications may, at the user's request, install symbolic links in the `/usr/local/bin` directory, and modify the `PATH` variable in the `.openeyerc.*` file(s) in the user's home directory. The symbolic links may be safely deleted after uninstallation, and the `.openeyerc.*` file(s) can be edited, if desired, to remove obsolete entries.

USAGE

3.1 Introduction

EON calculates the electrostatic similarity between two small molecules in the form of an Electrostatic Tanimoto (ET) score. Given a query molecule and a set of interesting molecules (ROCS overlay hits, for example), EON will calculate the Electrostatic Tanimoto between each database molecule and the query. Note that EON does not perform any overlay or alter the input orientation of the structures. They must be pre-aligned to the query on input. Also, since electrostatics calculations require high quality partial charges, EON will calculate new partial charges for the input structures using MMFF94. If the user provides an input file that contains structures with higher-quality partial charges, EON can use them as well.

EON is also dependent on pKa state and formal charges as these have a significant impact on electrostatics. EON now has the ability to adjust both the query and database molecule to a neutral pH model. This feature is on by default, but can be turned off by using appropriate command line flags.

Since electrostatics overlays are very dependent on alignment and require a good quality alignment between query and database molecule, ROCS provides the best input to EON. However, electrostatic complementarity is more dependent on subtle conformational changes than shape is, so there are several steps that can be taken to ensure the best possible success with EON.

Firstly, one can ensure that ROCS outputs multiple interesting conformers per molecule. ROCS includes a flag `-eon_input` that allows generation of a multi-conformer set of ROCS-aligned output specifically for input into EON. This file can be generated in parallel with a ROCS hit list so that in a single ROCS run you can find ROCS hits and prepare EON input. Please see the ROCS documentation for more detail on these flags.

Secondly, EON reads one or more conformers from the input file and uses technology from OMEGA to expand terminal torsions to search for subtle changes in conformation that might increase the score without changing the overall shape overlap with the query. To score just the input conformers and not search for alternate terminal conformations, the `-scoreonly` flag is provided.

Part of understanding EON results is visualization of the electrostatic grids used in the calculation. Although off by default, when writing EON results to a binary (OEB) file, ET grids can be attached to each molecule and visualized using the EON View mode in VIDA.

Since EON calculations can be time-consuming (approximately 1 molecule per second per CPU), EON can use the same distributed computing technology, Open MPI, that ROCS uses to help distribute the workload across a cluster of machines.

3.2 Command Line Interface

A description of the commandline interface can be obtained by executing EON with the `--help` option.

```
prompt> eon --help
```

will generate the following output:

Help functions:

```
eon --help simple      : Get a list of simple parameters
eon --help all         : Get a complete list of parameters
eon --help <parameter> : Get detailed help on a parameter
eon --help html        : Create an html help file for this program
```

3.2.1 Required Parameters

-dbase <filename>

File containing one or more 3D molecules to score against a reference or query molecule. If only this flag is given, EON will use the first molecule in the dbase file as the query and score all the remaining molecules against it. This is most useful when scoring ROCS results when ROCS was run with `-eon_input` equal to **true** since the ROCS query and therefore EON query will be the first in the input file. The query or reference molecule can also be specified separately using the `-query` flag below.

File format for `-dbase` can be one of:

File type	Extension
OEBinary	.oeb .oeb.gz
SDF	.sdf .mol .sdf.gz .mol.gz
MOL2	.mol2 .mol2.gz
PDB	.pdb .ent .pdb.gz .ent.gz
MacroModel	.mmod .mmod.gz

3.2.2 Optional Parameters

Execute Options

-param

The argument for this flag is the name of a file containing control parameters. The control parameter file acts to either replace or augment the command line interface. All parameters necessary for program execution may be provided in the control parameter file, although any command given explicitly on the command line will supersede options found in the parameter file. The application generates a new parameter file containing the full set of execution parameters upon every execution. The name of the parameter file is created by combining the prefix base name with the `‘.parm’` extension.

-mpi_np <n>

Specifies the number of processors *n* when the application is run in MPI mode.

-mpi_hostfile <filename>

Specifies the name of the file containing processors configuration. For every host this file should contain a line `host_name slots=n` where *n* is the number of processors on the host.

Input Options

-query <filename>

File containing one 3D molecule to use as a query. File format can be any of the formats given in the table for `-dbase` above.

-charges

Specifies charges to be used on the query. Default is to calculate them internally with `mmff`. The option `-charges existing` will use precalculated charges that must be set in the input files.

[default = mmff]

-scdbase

Since EON reads multi-conformer input molecules, when reading from non-binary files, EON will compare consecutive molecules and if they are determined to be the same structure, they will be concatenated into a single, multi-conformer molecule. If the user desires to score each input conformation independently, then using this flag will turn off the conformer comparison step.

[default = false]

Output Options**-offormat <extension>**

Format of output structure file(s). The default is `oeb` so that ET score can be included as tag data. The format for the file is determined by giving the extension. Valid values include all formats listed above for `-query`

[default = oeb]

-prefix <prefix>

Defines a prefix used to name output files. Using `-prefix FOO` will create a hits structure file named like `FOO_hits.oeb` and a report file, `FOO.rpt`

[default = EON]

-besthits <N>

Process entire dbase file but only keep `N` best scores sorted by property specified by `-rankby`. Using a value of `0` implies no hitlist will be maintained and structures will all be scored and output in input order.

[default = 500]

-maxhits <N>

Stop after finding first `N` hits. This option overrides any setting for `-besthits`

[default=0]

-rankby <score>

Property to use to sort hitlist. Values include `ET_combo`, `ET_pb` and `ET_coul`.

[default = ET_combo]

-cutoff <score>

Minimum score to keep as a hit.

[default = -1.0]

-outputquery

Write the query to the top of the hits file. This make visualization of results much easier inside VIDA.

-scoreonly

Turn off the terminal torsion conformer search and just score each input conformation as-is.

-hitsfile <filename>

Explicit filename for writing hits. Overrides the default filename created from `-prefix`.

-reportfile <filename>

Explicit filename for writing hits. Overrides the default filename created from `-prefix`.

-sdTags

This parameter controls whether to attach score information to output molecules as SD data.

Log Output Options

-logfile <filename>

Filename for log file. Overrides log filename created from *-prefix*.

-progress

Method for showing job progress on the command line. Choices include:

- percent - show a percent complete progress bar (DEFAULT)
- log - echo the log message for each molecule
- dots - show dots as in EON 1.1
- none - print nothing to console

[default = percent]

-statusfile

Write status info to this file. Use “none” for no status file.

-verbose

Give verbose output to console instead of simple progress.

[default = false]

ZAP/PB Options

-fixpka_query

Apply a neutral pH model to the query molecule.

[default = true]

-fixpka_dbase

Apply a neutral pH model to the database molecules.

[default = true]

-salt

Add salt to the Zap calculation. To aid in moderating large, local charges, salt is added into the calculation. Legal values are between 0.0 and 0.1 (mM).

[default = 0.04]

-spacing

Sets the grid spacing for the internal Zap calculations.

[default 0.5]

-writegrid

Write ET grid to output attached to each molecule. Useful for visualization in VIDA but this only works when writing hits to an OEB (.oeb or .oeb.gz) file. Note that while this feature is quite useful, grids do take a large amount of memory so care should be taken when using this feature for hit lists of more than 500 molecules.

[default = false]

Omega Options

-ewindow

Omega energy window used for conformer selection.

[default = 10.0]

-rms

Omega RMS threshold used to determine duplicate conformations

[default = 0.3]

-sampleHydrogens

Sets whether hydrogens will be sampled. This option enables sampling of hydrogen locations for -OH, -SH, and amines.

[default = false]

3.2.3 Example Commands

The simplest way to run EON is to use the `-eon_input` flag in ROCS to create a set of ROCS-aligned structures with the ROCS query at the top of the file. By default, if EON is only provided a dbase file, it will assume the first molecule is the query.

```
prompt> eon -dbase rocs_eon_input.oeb.gz
```

Or you can provide a single molecule in a query file and a set of molecules in a dbase file. A common example would be to use a ROCS query as the EON query and the ROCS hits file as the dbase file for EON.

So for example:

```
prompt> eon -dbase rocs_hits_1.oeb -query rocsquery.sdf
```

will score all the structures in `rocs_hits_1.oeb` against the molecule in `rocsquery.sdf` and place the structures in `EON_hits.oeb` (with ET scores in SD tag data) and a table of results in `EON.rpt`.

To keep only the best 100 hits, use:

```
prompt> eon -dbase rocs_hits_1.oeb -query rocsquery.sdf -besthits 100
```

Note that the use of an output structure file is mostly for generation of a single file containing structures and tag data such that loading this single file into VIDA provides easy analysis of the results. If however, the only real desire is for the numerical scores, use `-nostructs` to suppress the creation of an output structure file.

```
prompt> eon -dbase rocs_hits_1.oeb -query rocsquery.sdf -nostructs
```

By default, EON calculates partial charges using MMFF94. However, if you have structure files that already contain good partial charges in both your dbase file and query file, you can tell EON to use them instead:

```
prompt> eon -dbase rocs_hits_chgs.mol2
-query rocsquery_chg.mol2 -charges existing
```

To prevent continually over-writing output files, the `-prefix` flag allows you to give unique names to these files.

```
prompt> eon -dbase rocs_hits_1.oeb -query rocsquery.sdf -prefix FOO
```

will write the hit structures into a file named `FOO_hits.oeb` and the numerical values will be in `FOO.rpt`. The parameter file for this run will likewise be named `FOO.parm`.

To prevent EON for searching alternate terminal torsions, use the `-scoreonly` flag.

```
prompt> eon -dbase rocs_eon_input.oeb.gz -scoreonly
```

Finally, to create ET grids as used in the calculation and attach them to each output molecule. Note, this only works for OEB output and can be very memory intensive.

```
prompt> eon -dbase rocs_eon_input.oeb.gz -oformat .oeb.gz -writegrid
```

3.3 Report File

The EON report file format appears as a tab-delimited file with the following fields. Since the names of the query and the hits are of indeterminate length, fixed size fields for these names could result in loss of information. Unfortunately this gives a file that is hard to read in a terminal session, but it can easily be read into a spreadsheet program or into the data manager in VIDA.

Name This is the name of the database molecule.

EONQuery This is the name of the query molecule.

Rank The numerical ranking in the hitlist, based on the chosen score to rank by. Using the defaults, this is **ET_combo**. Can be altered by using *-rankby* command line switches. If no hitlist was used in the calculation, this field will be 0 (zero).

ET_pb This is the value of electrostatic Tanimoto, using full Poisson-Boltzmann (PB) electrostatics.

ET_coul This is the value of electrostatic Tanimoto using only the coulombic part of PB electrostatics.

ET_combo Sum of **ET_pb** and **EON_shape_tani**. This is a useful score that takes into account both shape match and ET match.

EON_shape_tani This the shape Tanimoto between the given molecule and the query. For calculations that use *-shapeonly*, this will be the same as the output Tanimoto from ROCS. When EON is allowed to alter terminal torsion, this will give the final shape Tanimoto.

THEORY

EON uses a field-based measure of Tanimoto to compare the electrostatic potential of two small molecules. This electrostatic potential is calculated internally using Zap, OpenEye's Poisson-Boltzman (PB) electrostatics toolkit.

The basic equation for a field Tanimoto is:

$$Tanimoto_{A,B} = \frac{\int A(\vec{r}) * B(\vec{r})}{\int A(\vec{r}) * A(\vec{r}) + \int B(\vec{r}) * B(\vec{r}) - \int A(\vec{r}) * B(\vec{r})}$$

The two boundary cases for Electrostatic Tanimoto occur when B = A:

$$\begin{aligned} Tanimoto &= \frac{\int A(\vec{r}) * A(\vec{r})}{\int A(\vec{r}) * A(\vec{r}) + \int A(\vec{r}) * A(\vec{r}) - \int A(\vec{r}) * A(\vec{r})} \\ &= 1 \end{aligned}$$

and the opposite case, when B = -A:

$$\begin{aligned} Tanimoto &= \frac{\int A(\vec{r}) * -A(\vec{r})}{\int A(\vec{r}) * A(\vec{r}) + \int -A(\vec{r}) * -A(\vec{r}) - \int A(\vec{r}) * -A(\vec{r})} \\ Tanimoto &= \frac{-\int A(\vec{r}) * A(\vec{r})}{\int A(\vec{r}) * A(\vec{r}) + \int A(\vec{r}) * A(\vec{r}) + \int A(\vec{r}) * A(\vec{r})} \\ &= -\frac{1}{3} \end{aligned}$$

In EON, we report two different Electrostatic Tanimoto(ET) measures, based on the outer dielectric used in the PB calculation. ET_pb uses an outer dielectric of 80, while ET_coul uses a value of 2.0. The rationale for using a PB electrostatic field is that the external potential is dampened by orientation of the aqueous solvent. It is a common observation that proteins essentially act to reproduce the aqueous desolvation of well-bound ligands. As a result a PB electrostatic field is more likely to correctly capture the essential elements of binding than that from the Coulombic field. However, this would still seem to be a point to be proven. As such we provide both Tanimotos. They typically track each other very closely.

For hit list ranking, we also report a score (ET_combo) that is the sum of the Shape Tanimoto (ST) and the PB Electrostatic Tanimoto (ET_pb).

RELEASE NOTES

5.1 EON v2.2.0 (June 2013)

5.1.1 New features

- The default hitlist format has been changed from sdf to oeb for increased functionality and decreased filesize. The output format is adjustable with the *-offormat* parameter.
- The *-spacing* parameter has been added. This parameter can be used to adjust the grid spacing of the Zap/PB calculations. [Case 16576]
- An option has been added to allow hydrogen atoms in -OH, -SH, and amines to take part in conformational sampling. This new option can be enabled via the *-sampleHydrogens* parameter. By default, hydrogen atoms are not sampled.
- Now SD tags are prefixed with *EON_*. The tags are optional with the *-sdTags* parameter. Additionally, any existing ROCS tags will not be removed because ROCS and EON tags no longer conflict. [Case 10096]

5.1.2 Bug fixes

- A bug has been fixed where the multiconformer check was being applied to OEB format files. Additionally, the default for *-scdbase* has been changed from *false* to *true* to prevent inadvertent merging of molecules with other file formats. [Case 16393]
- Warnings have been improved for charging failures. [Case 14433]
- *-writegrid* will now write out a grid with the same spacing as the actual Zap/PB calculation. Previously the attached grid had tighter spacing which could lead to confusion. [Case 13651]

5.1.3 Other changes

- PVM (parallel virtual machine) is no longer supported. OpenMPI version 1.6 is supported on all platforms. The *-mpi_np* and *-mpi_hostfile* flags are now used to run EON in MPI mode. These new flags replace the *oempirun* script.
- This will be the last release to support SuSe 10.

5.2 EON v2.1.0

June 2011

5.2.1 Enhancements

- The command line flags for choosing which charges to use have been simplified. There is a single flag `-charges`. The default is still to calculate MMFF charges. To use charges you have precalculated outside EON, first make sure you have charges for both the query and the dbase molecules, then use `-charges existing`.

Note: To use `existing` charges you must provide input files in a format that can store charges. This means input should be in either OEB or MOL2.

- This release adds EON to the set of applications that now use a common script in `openeye/bin` to determine the appropriate architecture and run the appropriate binary.
- On Windows, there is now an installer that installs the documentation and sets up a command prompt to facilitate running the EON command line.

5.2.2 Bug Fixes

- Fixed a bug to ensure that SD data present in the database is passed on to the EON hitlist. Note that this is specific to SD data other than ROCS scores. Since EON can modify the conformation, any ROCS scores would no longer be valid so they are removed from the output. Note that the score reported as `EON_ShapeTanimoto` is the final Shape Tanimoto calculated from the best conformation found inside EON.
- Fixed a bug that caused a crash when trying to write an empty hitlist.
- Fixed a bug that prevented `-salt` from working.
- Adding better error checking and messages for input. All input must be molecules and cannot be shape queries.

5.3 EON v2.0.0

August 2007

5.3.1 Enhancements

This is a significant new release of EON. This release incorporates many internal changes in how Electrostatic Tanimoto is calculated compared to EON 1.1.

- Salt in the PB calculation is now on by default to help moderate large, local charges. This should make EON 2.0 perform much better than 1.1 in the case of molecules with formal charges.
- OpenEye's neutral pH model has been incorporated allowing the setting of a pH model on both the query and database molecules. Since a significant use of EON is similarity searching, having a consistent model between the query and database molecules is very important. However, for the case where the states are well defined, this feature can be turned off.
- Instead of the slow and buggy `-spin` model for searching terminal torsions, EON 2.0 uses technology from OMEGA to only search reasonable torsions and to prevent creation of bad structures while searching for better electrostatic overlap.
- The hitlist is now sorted by default by a new score, `ET_combo`, which is the sum of the Shape Tanimoto (ST) and the Electrostatic Tanimoto (ET_pb).
- To make visualization of results easier, EON can write out the actual ET grids to the hits file, attached to each hit. This file can be loaded into VIDA 3.0 or later and the grids can be visually compared.

- There are also several changes to the mechanics of running the calculation including a new *-progress* feature and a more detailed PVM log output for PVM jobs.

LIST OF SELECTED EON PUBLICATIONS

- S.W. Muchmore, A.J. Souers and I. Akritopoulou-Zanze, **The Use of Three-Dimensional Shape and Electrostatic Similarity Searching in the Identification of a Melanin-Concentrating Hormone Receptor 1 Antagonist**, *Chemical Biology & Drug Design*, Vol. 67(2), pp 174-176 **2006**
- Markt, P., Petersen, R.K., Flindt, E.N., Kristiansen, K., Kirchmair, J., Spitzer, G., Distino, S., Schuster, D., Wolber, G., Laggner, C. & Langer, T., **Discovery of Novel PPAR Ligands by a Virtual Screening Approach Based on Pharmacophore Modeling, 3D Shape and Electrostatic Similarity Screening**, *Journal of Medicinal Chemistry*, Vol. 51(20), pp. 6303-6317, **2008**.
- Naylor, E., Arredouani, A., Vasudevan, S.R., Lewis, A.M., Parkesh, R., Mizote, A., Rosen, D., Thomas, J.M., Izumi, M., Ganesan, A., Galione, A. & Churchill, G.C., **Identification of a Chemical Probe for NAADP by Virtual Screening**, *Nature Chemical Biology*, Vol. 5, pp. 220-226, **2009**.
- Tresadern, G., Bemporad, D. & Howe, T., **A Comparison of Ligand Based Virtual Screening Methods and Application to Corticotropin Releasing Factor 1 Receptor**, *Journal of Molecular Graphics and Modelling*, Vol. 27(8), pp. 860-870, **2009**.
- Zavodszky, M.I., Rohatgim, A., Van Voorst, J.R., Yan, H. & Kuhn, L.A., **Scoring Ligand Similarity in Structure-Based Virtual Screening**, *Journal of Molecular Recognition*, Vol. 22(4), pp. 280-292, **2009**.
- Lopez-Ramos, M. & Perruccio, F., **HPPD: Ligand- and Target-Based Virtual Screening on a Herbicide Target**, *Journal of Chemical Information and Modeling*, Vol. 50(5), pp 801-814, **2010**.

CITATION

Note: To cite EON please use the following:

EON 2.2.0.5: OpenEye Scientific Software, Santa Fe, NM. <http://www.eyesopen.com>.

INDEX

Symbols

-besthits <N>
 eon command line option, 11

-charges
 eon command line option, 10

-cutoff <score>
 eon command line option, 11

-dbase <filename>
 eon command line option, 10

-ewindow
 eon command line option, 12

-fixpka_dbase
 eon command line option, 12

-fixpka_query
 eon command line option, 12

-hitsfile <filename>
 eon command line option, 11

-logfile <filename>
 eon command line option, 12

-maxhits <N>
 eon command line option, 11

-mpi_hostfile <filename>
 eon command line option, 10

-mpi_np <n>
 eon command line option, 10

-offormat <extension>
 eon command line option, 11

-outputquery
 eon command line option, 11

-param
 eon command line option, 10

-prefix <prefix>
 eon command line option, 11

-progress
 eon command line option, 12

-query <filename>
 eon command line option, 10

-rankby <score>
 eon command line option, 11

-reportfile <filename>
 eon command line option, 11

-rms
 eon command line option, 12

-salt
 eon command line option, 12

-sampleHydrogens
 eon command line option, 13

-scdbase
 eon command line option, 11

-scoreonly
 eon command line option, 11

-sdTags
 eon command line option, 11

-spacing
 eon command line option, 12

-statusfile
 eon command line option, 12

-verbose
 eon command line option, 12

-writegrid
 eon command line option, 12

A

APPNAME_OE_ARCH, 4

E

environment variable

 APPNAME_OE_ARCH, 4
 OE_ARCH, 4
 OE_LICENSE, 3
 PATH, 3, 5

eon command line option

 -besthits <N>, 11
 -charges, 10
 -cutoff <score>, 11
 -dbase <filename>, 10
 -ewindow, 12
 -fixpka_dbase, 12
 -fixpka_query, 12
 -hitsfile <filename>, 11
 -logfile <filename>, 12
 -maxhits <N>, 11
 -mpi_hostfile <filename>, 10

- mpi_np <n>, 10
- offormat <extension>, 11
- outputquery, 11
- param, 10
- prefix <prefix>, 11
- progress, 12
- query <filename>, 10
- rankby <score>, 11
- reportfile <filename>, 11
- rms, 12
- salt, 12
- sampleHydrogens, 13
- scdbase, 11
- scoreonly, 11
- sdTags, 11
- spacing, 12
- statusfile, 12
- verbose, 12
- writegrid, 12

O

- OE_ARCH, 4
- OE_LICENSE, 3

P

- PATH, 3, 5