

# Techniques For Drug-Design

**Jas Bhachoo and Daniel Robinson**



**SCHRÖDINGER.**

---

# INTRODUCTION

# Introduction

---

- Ligand based drug-design (LBDD)
  - As its name suggests LBDD focuses solely on the structure of the ligands.
    - Often LBDD is the only possibility, many important target classes have little or no structural data available.
    - Nonetheless, with care, powerful and useful models can be developed which can provide guidelines to synthetic-chemists.
- Structure based drug-design (SBDD)
  - In SBDD we also get to consider the structure and influence of the target protein.
    - We have to be very careful in ensuring that the protein structure we're using is reasonable for the job in hand.
    - Of course if this condition is met, the protein structure provides us with a host of valuable information.
- Joint approaches
  - Naturally there is no reason to favour one approach over the other.
    - If structural data is available we should use it, but the techniques of LBDD still remain useful.

---

# LIGAND PREPARATION

# The Importance of Ligand Preparation

---

- Ligands can come from a variety of sources, each of which have a variety of different issues:
  - Compounds from a third party database.
    - Molecules are entered by a variety of people, often in completely distinct locations, with little or no reference to each other.
      - Frequently there are issues with salt information, ionisation, tautomerisation and particularly chirality/E/Z definitions.
  - In-house databases.
    - Usually more consistent than third party databases as the chemists should be speaking to one another.
      - Nonetheless, chemists and modellers frequently have different ideas about the functionality needed to bind.
      - Issues still can exist with salt information, ionisation, and tautomerisation.
  - Hand-drawn structures.
    - Compounds entered by the modeller should be the most consistent and relevant for binding.
      - However ambiguities can often persist with regards to ionisation states and occasionally tautomers.

# The Importance of Ligand Preparation

---

- Ligand preparation helps to circumvent these issues.
  - Salt stripping.
    - Simple removal of counterions and neutralisation of any remaining charge gets all ligands into a common state for further analysis.
  - Ionisation/Tautomerisation.
    - It is possible to generate quite accurate models for the pKa of most ionisable centres.
      - These models enable us to generate a set of possible ionisation states for every ligand.
  - Chirality/E/Z.
    - If chirality is not specified then it is reasonable to assume that a mixture of stereocentres are present.
      - Software can be used to enumerate all of these centres to ensure that all cases are sampled.

## Additional Information: pKa Models

---

- The Hammett and Taft equations form the basis of many pKa estimation programs.

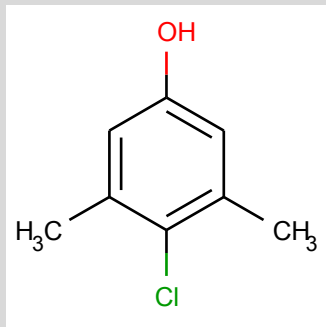
- The form of this equation is as follows:

$$pKa = pKa^0 - \rho \sum_i \sigma_i$$

- In this equation each ionisable group has a baseline pKa denoted  $pKa^0$ .
    - The sensitivity of this group to perturbations by neighbouring groups is then given by  $\rho$ .
    - The actual intrinsic effect of the neighbouring groups is summarised by the various  $\sigma$ .

## Additional Information: pKa Models

- Let us apply the equation to the pKa of the following phenol:



pKa <sup>0</sup> phenol	9.92
ρ phenol	2.23
σ meta-CH <sub>3</sub>	-0.06
σ para-Cl	0.24

- The pKa would thus be estimated as:

$$\text{pKa} = 9.92 - 2.23 * ( 0.24 - 0.06 - 0.06 ) = 9.70$$

- The experimentally determined value is 9.71.



# Automated Ligand Preparation in the Schrödinger Suite (LigPrep)

- The LigPrep module enables a large number of ligands to be processed quickly
- C:\Schrodinger2009\SchrodingerShell  
Access Command Line in Windows
  - Input: Maestro, SD, or SMILES format
  - Two choices are available: OPLS\_2005, MMFFs
  - Generate ionization states that are significantly populated in the pH range specified using ‘template based’ or ‘Hammett and Taft based’ method
  - Epik protocol for ligands bound to a metal
  - Tautomers are thermally accessible and interchangeable isomers of molecules in which the non-hydrogen topology of the molecule remains unmodified while hydrogen atoms are relocated and bond orders change
  - Chirality information includes parities and bond directions from SD files, and the chirality property from Maestro files
  - Rings template (ring\_conf adding more rings)

The screenshot shows the LigPrep application window with the following settings:

- Use structures from:** File
- File name:** (empty text box) **Browse...**
- Filter criteria file:** (empty text box) **Create...** **Browse...**
- Force field:** OPLS\_2005
- Ionization:**
  - ☐ Retain original state
  - ☐ Neutralize (best for QikProp)
  - ☒ Generate possible states at target pH: 7.0 +/- 2.0
- Using:**
  - ☐ Ionizer
  - ☒ Epik
  - ☐ Add metal binding states
- ☒ Desalt ☒ Generate tautomers
- Stereoisomers**
- Computation:**
  - ☐ Retain specified chiralities (vary other chiral centers)
  - ☒ Determine chiralities from 3D structure
  - ☐ Generate all combinations
- Generate at most:** 32 per ligand
- Generate low energy ring conformations:** 1 per ligand
- Output format:** ☒ Maestro ☐ SDF
- Buttons:** Start... Close Help

```
SchrodingerShell
WKS-SCHROD-JAS# ligprep -h
LIGPREP_VERSION 23108
LigPrep. Copyright (c) 2009 Schrodinger, LLC.

Usage: ligprep [options] ... < -imae ! -isd ! -ismi > < input filename > \
      < -omae ! -osd > < output filename >

-ismi : input file in SMILES format
-isd  : input file in SDfile format
-imae  : input file in Maestro format
-osd   : output file in SDfile format
-omae  : output file in Maestro format

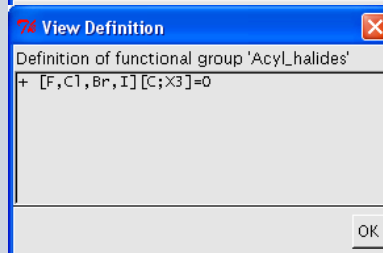
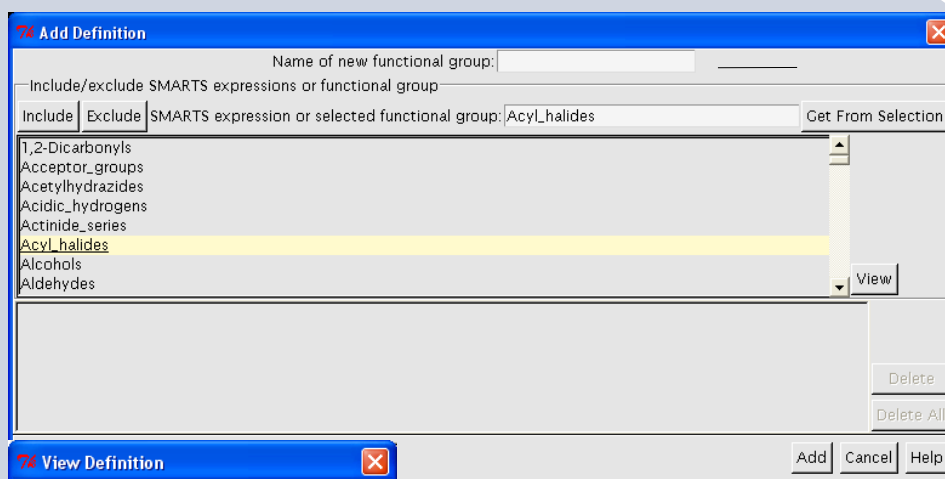
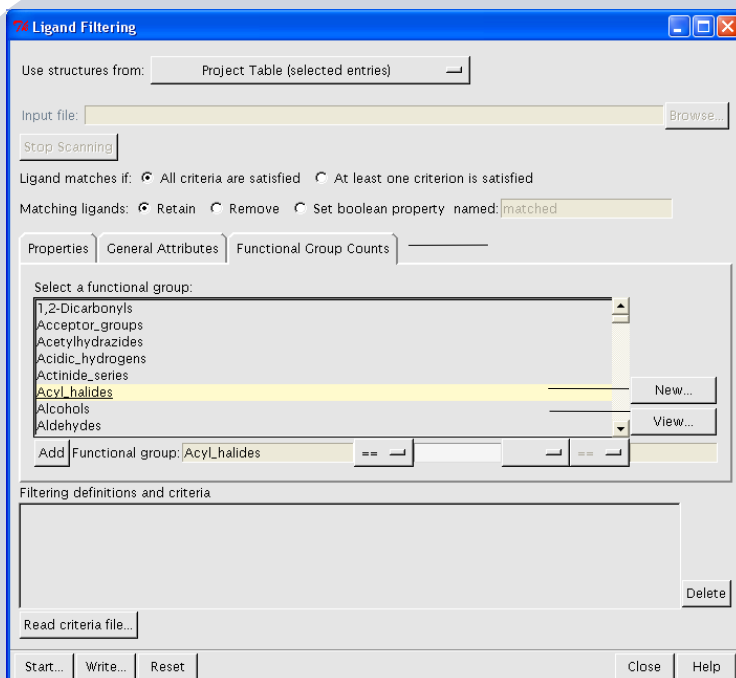
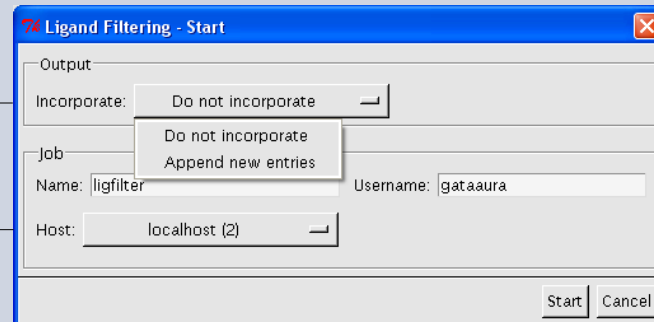
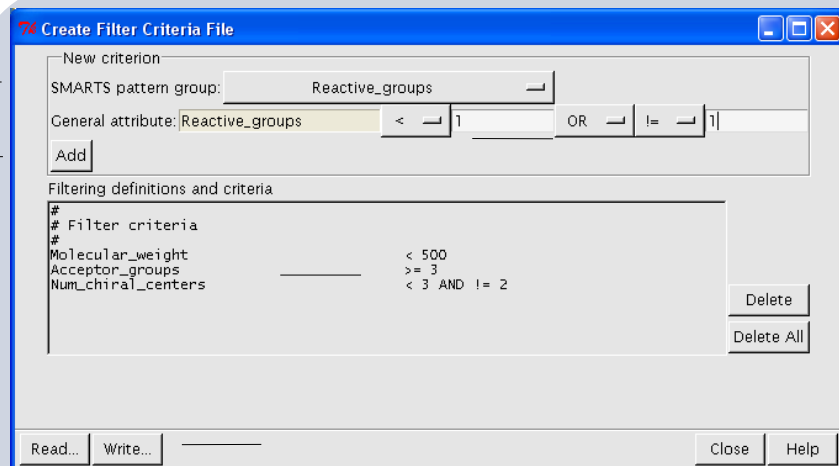
Typical Usage:
Process a 1D SMILES file to produce a 3D Maestro file
  ligprep -ismi <input.smi> -omae <output.mae>
Process a 2D SD file to produce a 3D SD file
  ligprep -isd <input.sd> -osd <output.sd>
Process a 2D SD file to produce a Maestro file
  ligprep -isd <input.sd> -omae <output.mae>
Process a 2D SD file to produce a Maestro file suitable for QikProp
  ligprep -qik -isd <input.sd> -omae <output.mae>
Process a 2D SD file to produce a Maestro file with various ionization
states, tautomers, and stereoisomers.
  ligprep -expand_its -isd <input.sd> -omae <output.mae>
Process a 2D SD file to produce a Maestro file with various ionization
states, tautomers, and stereoisomers with up to 8 ring conformations
per ligand.
  ligprep -r 8 -expand_its -isd <input.sd> -omae <output.mae>
Process a 3D Maestro file using chiralities from input geometry
  ligprep -g -imae <input.mae> -omae <output.mae>
Process a 3D Maestro file generating up to 64 stereoisomers
and ignoring input chiralities
  ligprep -ac -s 64 -imae <input.mae> -omae <output.mae>

Meta Options: (these turn on a number of actual options)
Specific
-qik          : set defaults appropriate for qikprop
               This option turns on the desalter, neutralizer,
               and the tautomerizer.
-unt          : Only run premin and bmin to untangle structures.
General
-adjust_<its> : adjust to a suitable state.
-vary_<its>   : generate different states.
-expand_<its> : aggressively generate different states.
-retain_<its> : Retain characteristics of the input structures.
               specifying 'i' also turns on retention of tautomers.
<its> means specify one or more of:
```

# LigPrep Command Line

```
SchrodingerShell
Options:
-a          : append structures to the output file
             default: overwrite the output file if it
             already exists.
-ac         : Do not respect existing chirality properties or
             use chiralities from the input geometry. Generate
             stereoisomers for all chiral centers up to the
             number in permitted (see -s option).
-bff #      : instruct bmin to use the forcefield specified.
             only 10 (MMFFs), 11 (OPLS2001), or
             14 (OPLS2005, default) are supported.
-bvac       : bmin minimization done in a vacuum.
-br         : instruct bmin to retain structures with
             incorrect chiralities
-bns        : instruct bmin to perform minimizations
             only
-btc #      : torsional constraints to use in bmin.
             # = -1 do not use torsional constraints.
             # = 0 torsional constraints for C=C,
             #       carboxylic acids, esters and amides.
             # = 2 torsional constraints for C=C. (Default)
-bts #      : torsional sampling to use in bmin.
             (overrides -btc option and sets it to 1.)
             # = 0 Restricted
             # = 1 Intermediate
             # = 2 Enhanced
             # = 3 Extended
-d          : use the desalter
             default: use the desalter
-DEBUG      : Show details of operation of the toplevel
             scripts.
-epik       : use Epik for ionization and tautomerization
-epik_metal_binding!-emb : run Epik with the metal_binding option
             so that states appropriate for
             interacting with metal ions in protein
             binding pockets are also generated.
-es <specfile> : Use non-standard epik pKa file.
-f <filename> : filter structures using specifications
             from <filename>
             default: do not filter.
             This uses ligfilter by default.
-fc <filename> : The name of the custom file listing
             composite descriptors
             requires -f
             This is passed to ligparse.
-fs <filename> : The name of the custom file listing SMARTS
             patterns to be used in examining the dataset
```

# Ligand Filtering in Maestro



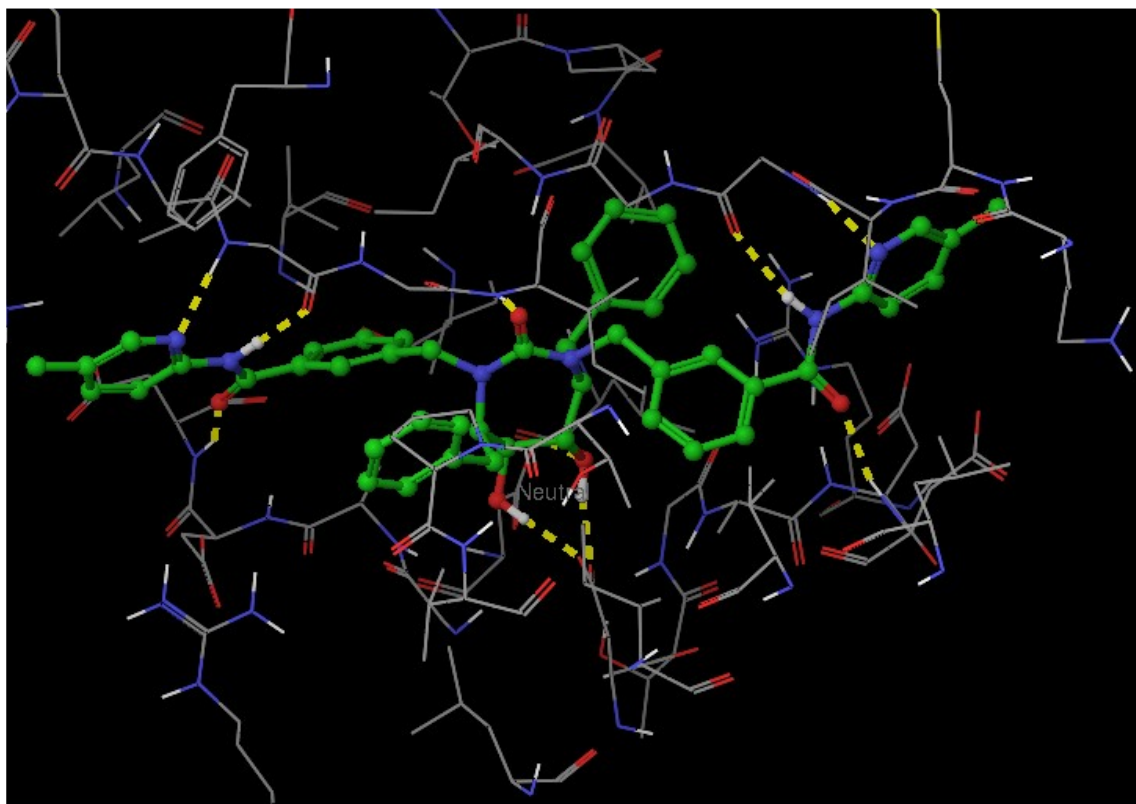
# What is Pharmacophore Modelling About? – Aims!

---

- When I don't have receptor information or I want to augment/corroborate my SB studies in an independent manner...
- What is common amongst my ligands, that seem to show experimental activity?
- If I can understand what is common amongst particularly active compounds and quantify it, I can start to make things that are more active
- If I can understand what is common amongst less active or inactive compounds, then I know what NOT to do!

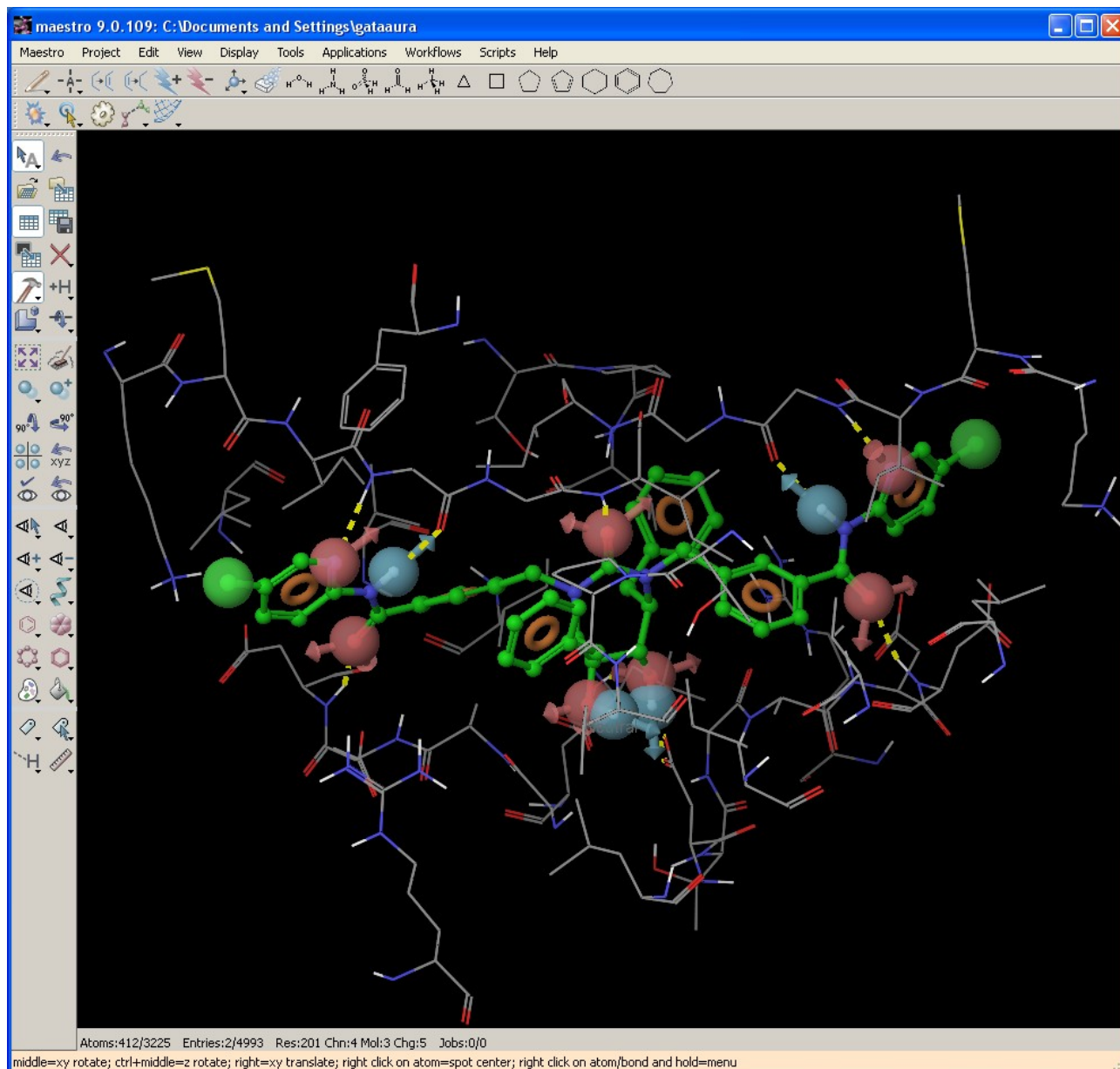
# Basic Concepts of Pharmacophore Models

---



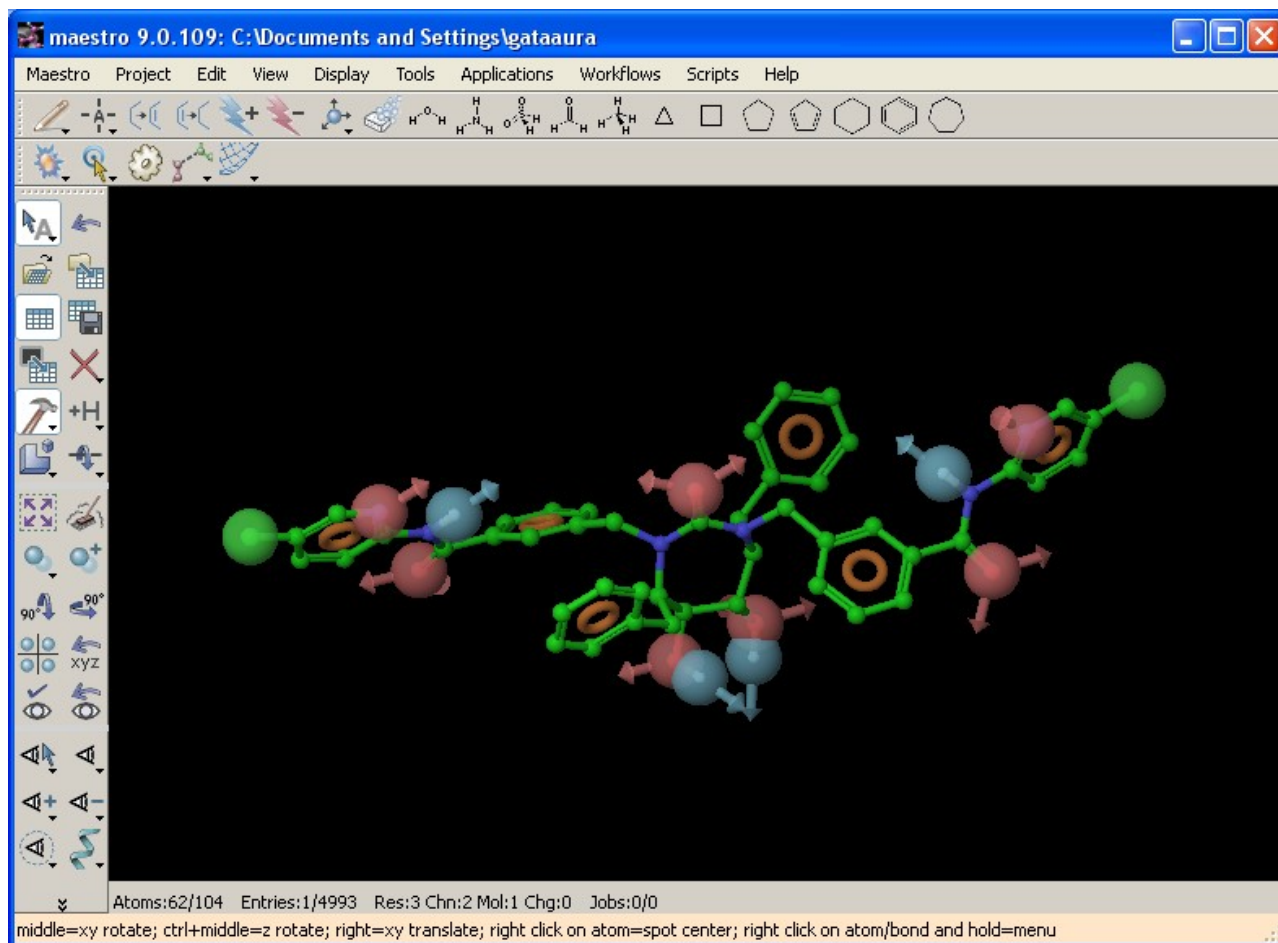
**The Protein  
Shows The  
Precise  
Nature Of  
The  
Interactions**

# Basic Concepts of Pharmacophore Models



**Translate  
these  
Interactions  
into ligand  
features  
A D N P R H**

# Starting Point for Pharmacophore Development

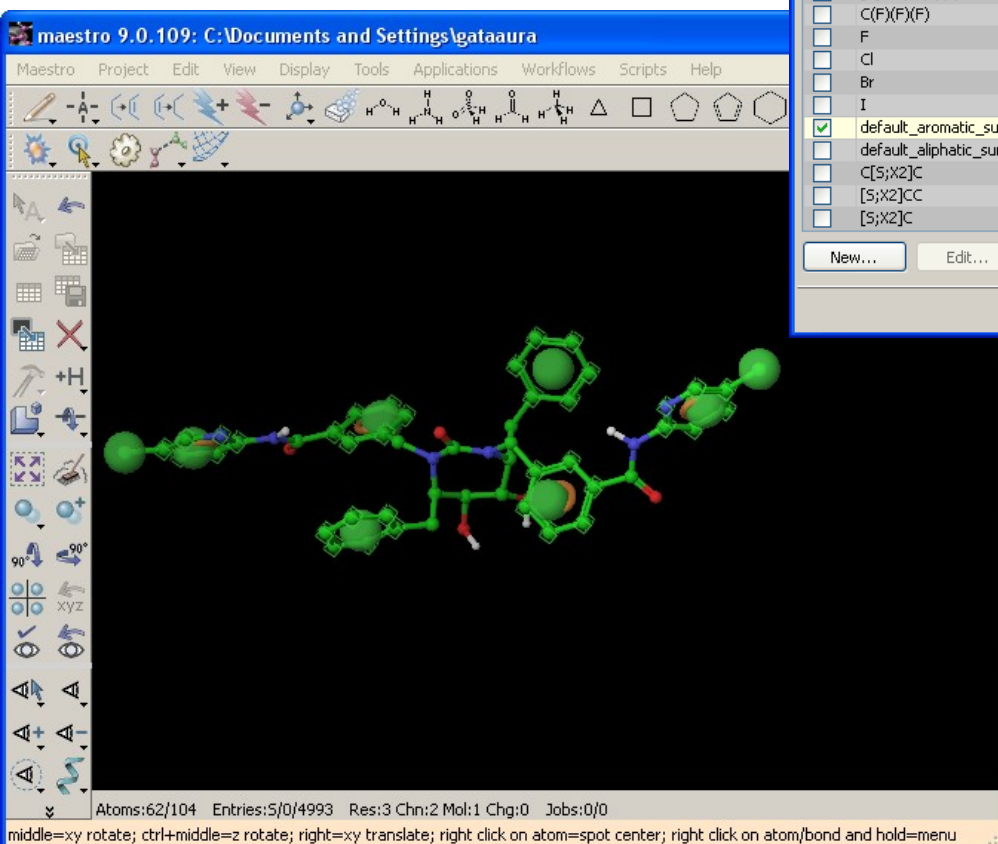


**The Structure Of  
The Ligands Alone  
Determine These  
interactions...  
Measure of  
similarity to  
understand what  
infers experimental  
activity**

**All done in  
absence of  
protein!**



# Image of Editing Features – Rings seen as Hydrophobes



**Edit Features**

Feature set: Import From Run... Import From File... Export... Reset All

Definitions

Feature: Hydrophobic (H) Add Custom Feature... Delete Custom Feature

☐ Projected points only Distance: 1.80 Å Set Distances SAS Options...

Pattern list:

Mark	Pattern	Geometry	Projected Point Type	Atom Numbers	Distance	Exclude	Ignore
<input type="checkbox"/>	[a]Cl	group	none	2	1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[a]Br	group	none	2	1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[a]I	group	none	2	1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[a]C(F)(F)(F)	group	none	2,3,4,5	1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[a][CH2]C(F)(F)(F)	group	none	2,3,4,5,6	1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[a]O[CH3]	group	none	2,3	1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[a]S[CH3]	group	none	2,3	1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[a]OC(F)(F)(F)	group	none	2,3,4,5,6	1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	C(F)(F)(F)	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	F	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	Cl	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	Br	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	I	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input checked="" type="checkbox"/>	default_aromatic_surface	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	default_aliphatic_surface	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	C[S;X2]C	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[S;X2]CC	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[S;X2]C	group	none		1.800	<input type="checkbox"/>	<input type="checkbox"/>

New... Edit... Delete

☒ Apply marker offset

OK Apply Cancel Help



# Features are defined by SMARTS patterns in Phase

**Edit Features**

Feature set: Import From Run... Import From File... Export... Reset All

Definitions

Feature: Acceptor (A)

☐ Project Distance: 1.80 Å  

Pattern list

Mark	Pattern	Geometry	Projected Point Type	Atom Numbers
<input type="checkbox"/>	Acceptor (A)			
<input type="checkbox"/>	Donor (D)			
<input type="checkbox"/>	Hydrophobic (H)			
<input type="checkbox"/>	Negative (N)			
<input type="checkbox"/>	Positive (P)			
<input type="checkbox"/>	Aromatic Rings (R)			
<input type="checkbox"/>	Custom (X)	vector	acceptor, sp, 1 lp	1
<input type="checkbox"/>	Custom (Y)	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	Custom (Z)	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	[N;X2](=N-O)[a]	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	[n;X2]1ccccc1	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	[n;X2]([a])([a])	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	[N;X2](=C~[C,c])(~[*])	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	[N;X3](C)(C)[N;X3]C	vector	acceptor, sp3, 1 lp	1
<input type="checkbox"/>	[N;X2](=C)(~[*])	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	[N;X2](~[C,c])=[N;X2]	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	[n;X2]1c[nH]cc1	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	O=[S;X4](=O)([!#8])([!#8])	vector	acceptor, sp, 3 lp	1
<input type="checkbox"/>	[O;X2]C	vector	acceptor, sp3, 2 lp	1
<input type="checkbox"/>	[O;X2]N	vector	acceptor, sp3, 2 lp	1
<input type="checkbox"/>	[O;X1]=[C,c]	vector	acceptor, sp2, 2 lp	1
<input type="checkbox"/>	o	vector	acceptor, sp2, 1 lp	1
<input type="checkbox"/>	[O;X2](C)C	vector	acceptor, sp3, 2 lp	1

☒ Apply marker offset

# Typical Steps in Pharmacophore Modeling

- Select data set (need a good range of activity)
- Ensure structures are represented in 3D
- Generate conformations for each ligand
- Generate sites
  - Pre-defined SMARTS patterns map features from ligand chemistry
- How many of your ligands do you want to use to build your model?
- What is common amongst your ligands?

# PHASE 3.0 - Basic Overview of Features



- 
- Phase Workflow Wizard
    - Creating Pharmacophore Models in Phase wizard
    - Creating QSAR Models in Phase wizard
  - Receptor Based Excluded Volume GUI
    - Creating Excluded Volumes
  - Search a Database
    - Database Searches with default options

---

# **Pharmacophore Models In Phase Wizard**

# Import Ligands

Develop Pharmacophore Model --- run1

File Display Step

**Prepare Ligands** Add ligands and ligand conformer sets to use as input. Clean up ligand structures, and generate ligand conformations if needed.

Add ligands: From File... From Run... From Project... ☒ Separate stereoisomers

Edit selected ligands: Clean Structures... Generate Conformers...

Ligands: (36 total)

In	Name	Activity	Pharm Set	# Conformations
◇	Molecule-1	5.509		1
◇	Molecule-2	5.456		1
◇	Molecule-3	5.469		1
◇	Molecule-4	6.824		1
◇	Molecule-5	7.824	active	1
◇	Molecule-6	7.060		1
◇	Molecule-7	7.167		1
◇	Molecule-8	8.398	active	1
◇	Molecule-9	8.301	active	1
◇	Molecule-10	5.620		1
◇	Molecule-11	6.076		1
◇	Molecule-12	5.886		1
◇	Molecule-13	5.839		1
◇	Molecule-14	5.252		1
◇	Molecule-15	6.060		1
◇	Molecule-16	7.824	active	1
◇	Molecule-17	7.796	active	1
◇	Molecule-18	6.097		1
◇	Molecule-19	5.301		1

Activity Thresholds...

< Back Next >

Prepare Ligands Create Sites Find Common Pharmacophores Score Hypotheses Build QSAR Model

Close Help

Add From Project

Choose one or more entries:

- endo-impl.1
- endo-impl.2
- endo-impl.3
- endo-impl.4
- endo-impl.5
- endo-impl.6
- endo-impl.7
- endo-impl.8
- endo-impl.9
- endo-impl.10
- endo-impl.11
- endo-impl.12
- endo-impl.13

Choose an activity property:

<none>  
Activity

Note: Activity values must be in units of  $-\log[\text{concentration}]$

☒ Convert property values:  
Activity =  $-\log [ 1.00 * \text{value} ]$

OK Cancel

Activity Thresholds

Active if activity above: 7.5

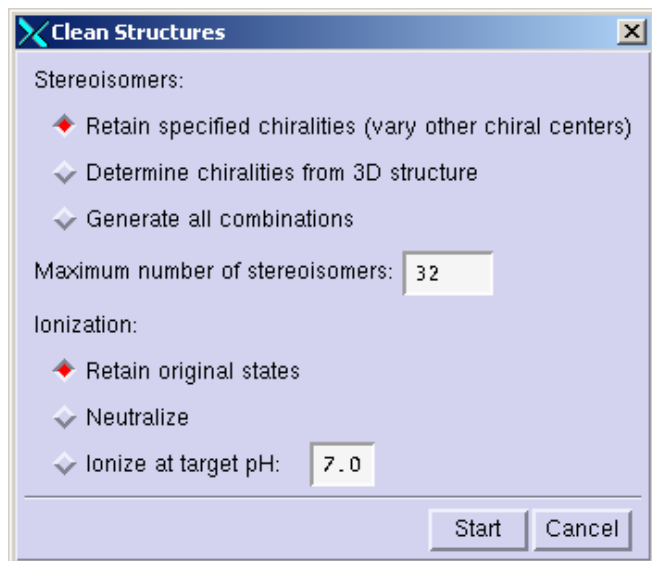
Inactive if activity below: 5.0

Maximum activity in table = 8.398  
Minimum activity in table = 4.076

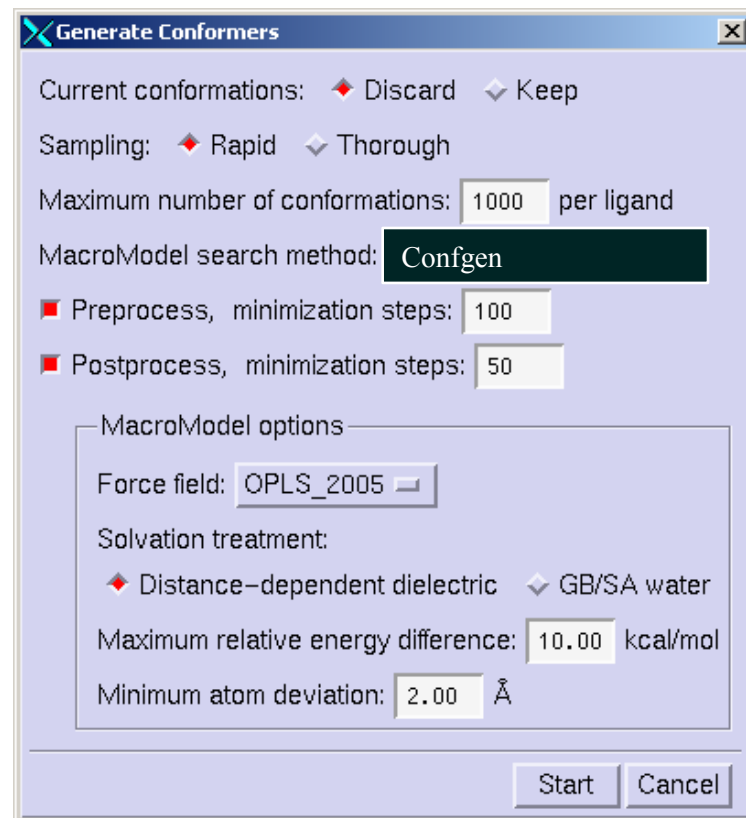
OK Cancel

- Add Structures from Project or File
- Choose Activity Property
- Assign “Actives Set” from Activity

# Prepare Ligands



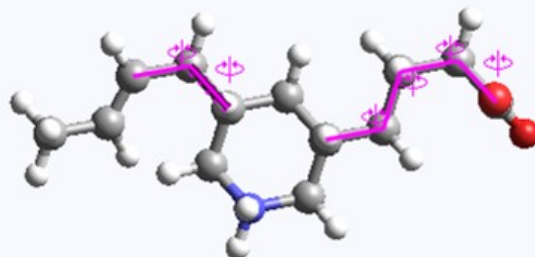
- Clean Structures
  - 2D → 3D
  - Expand stereoisomers
  - Ionize/neutralize



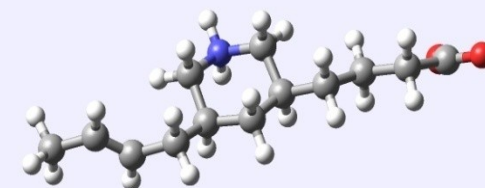
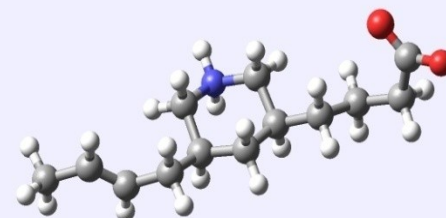
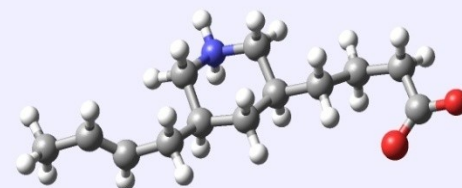
- Generate Conformers
  - New ConfGen Panel
  - Undergone extensive testing
  - Like charges, atoms close proximity etc

# ConfGen

- Confgen
  - Minima of core rotatable bonds systematically identified and sampled
  - Terminal rotamer groups then sampled
  - No electrostatics
  - Ring template library from MMod
  - 1-4 interactions to define potential



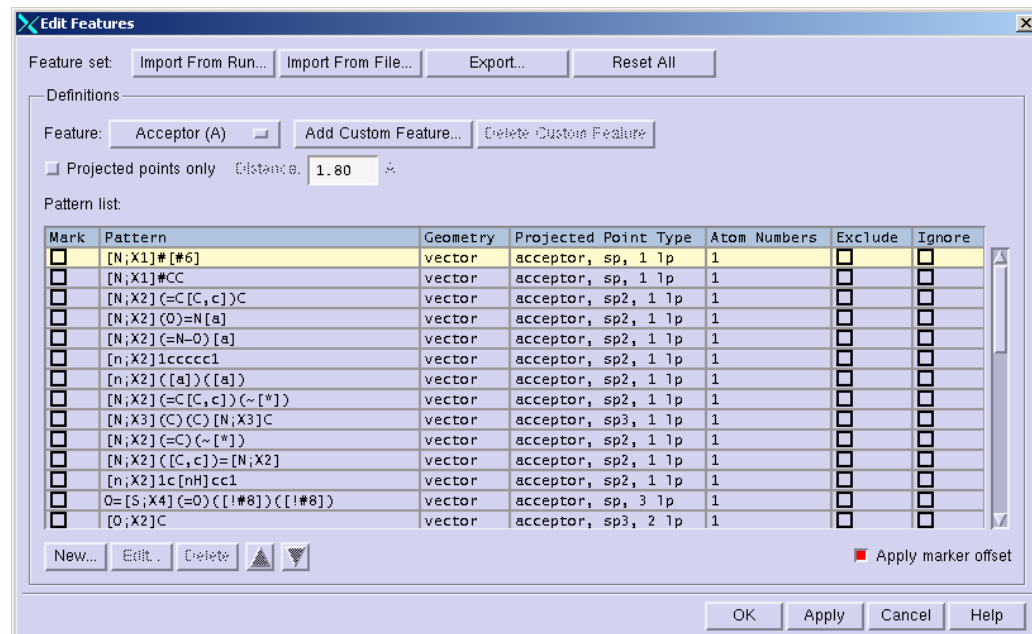
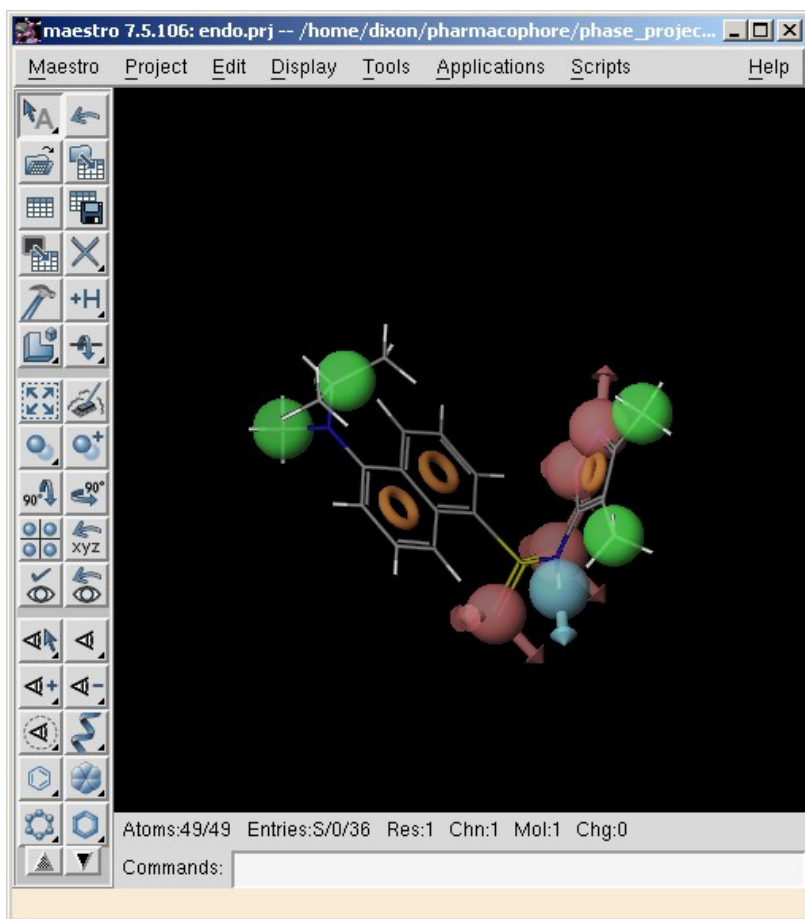
1. Core rotatable groups



2. Terminal rotamers



# Create Pharmacophore Sites



Edit Features

Feature set: Import From Run... Import From File... Export... Reset All

Definitions

Feature: Acceptor (A) Add Custom Feature... Delete Custom Feature

☒ Projected points only Distance: 1.80

Pattern list:

Mark	Pattern	Geometry	Projected Point Type	Atom Numbers	Exclude	Ignore
<input type="checkbox"/>	[N;X1]*[#6]	vector	acceptor, sp, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[N;X1]*CC	vector	acceptor, sp, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[N;X2](=C[C,c])C	vector	acceptor, sp2, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[N;X2](O)=N[a]	vector	acceptor, sp2, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[N;X2](=N-O)[a]	vector	acceptor, sp2, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[n;X2]1ccccc1	vector	acceptor, sp2, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[n;X2]([a])([a])	vector	acceptor, sp2, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[N;X2](=C[C,c])(~[*])	vector	acceptor, sp2, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[N;X3](C)(C)[N;X3]C	vector	acceptor, sp3, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[N;X2](=C)(~[*])	vector	acceptor, sp2, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[N;X2]([C,c])=[N;X2]	vector	acceptor, sp2, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[n;X2]1c[nH]cc1	vector	acceptor, sp2, 1 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	O=[S;X4](=O)([!#8])([!#8])	vector	acceptor, sp, 3 lp	1	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	[O;X2]C	vector	acceptor, sp3, 2 lp	1	<input type="checkbox"/>	<input type="checkbox"/>

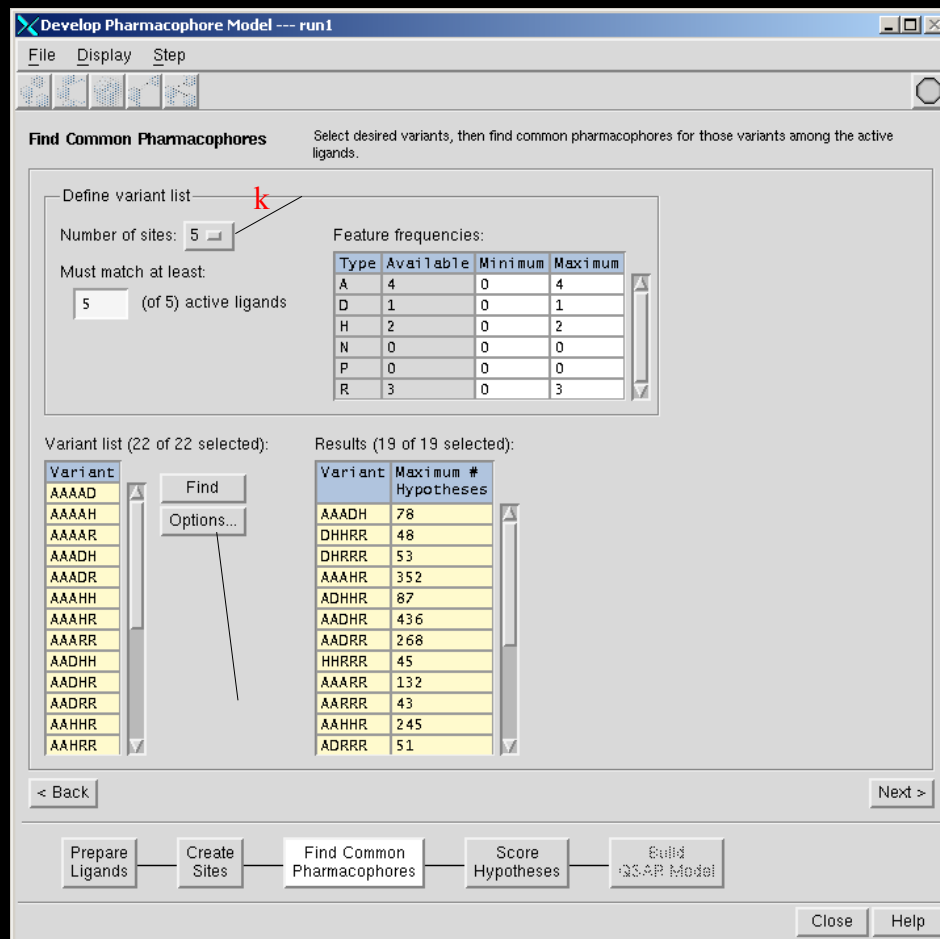
New... Edit... Delete

☒ Apply marker offset

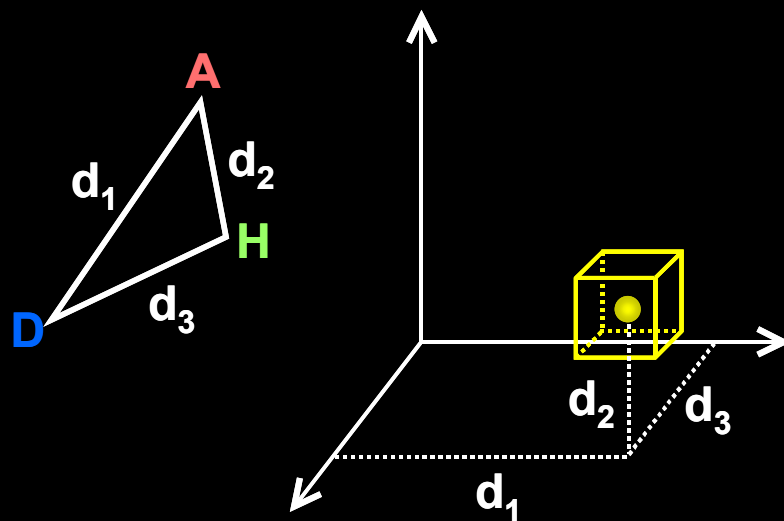
OK Apply Cancel Help

- Map/Visualize Pharmacophore Features
- Customize Definitions
  - Expand built-in feature types (A, D, H, N, P, R)
  - Create new feature types (X, Y, Z)
  - Choose vector features or “projected points”

# Find Common Pharmacophores

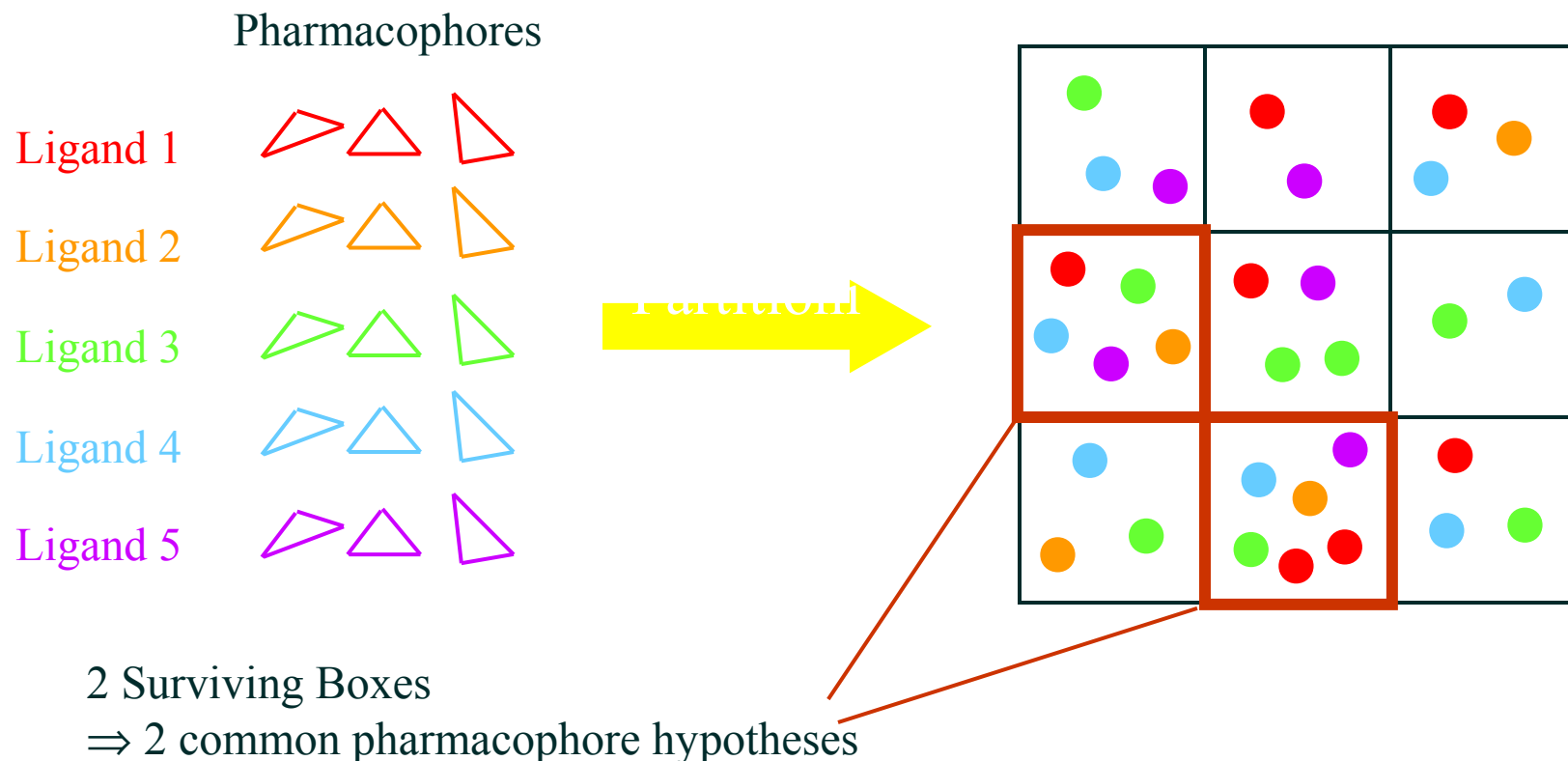


- Set Number of Sites
- Prohibit or Allow “Misses”
- Restrict Feature Frequencies
- Define Options for Partitioning Pharmacophores into High-Dimensional “Boxes”



Each  $k$ -point  $P$  represented by a vector of  $n$  distances, where  $n = k*(k-1)/2$ . Each I.S. dist  $d$  is filtered through a binary decision tree,

# Find Common Pharmacophores - Details



i.e. A Pharmacophore is mapped into a box of finite size according to its I.S. distances. If at-least 1 P from each of the 5 actives (as in QSG eg) is mapped into same  $1A^3$  box, this facilitates identification of a CP

# Score Hypotheses

**Score Actives**

Vector and site filtering

Keep those with RMSD below:  Å

Keep those with vector scores above:

Keep the top:  %

Keep at least:  and at most:

☒ Use feature matching tolerances:

Type	Tolerance
A	1.00
D	1.00
H	1.50
N	0.75

Survival score formula

\* vector score

+  \* site score

☒ +  \* volume score

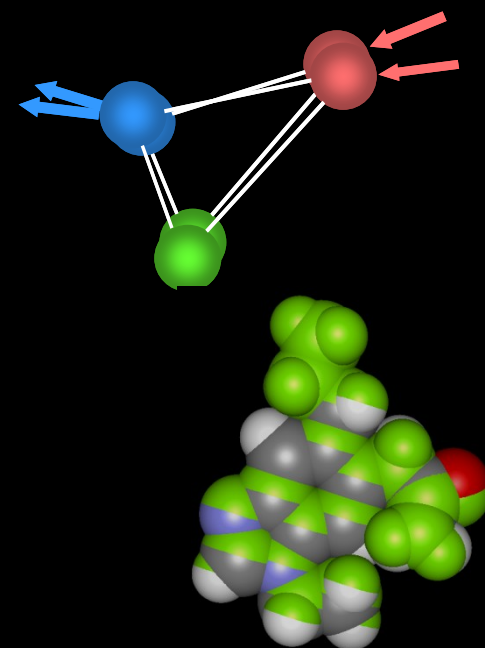
☒ +  \* selectivity score

+  ^ (number of matches - 1)

-  \* reference ligand relative conformational energy

+  \* reference ligand activity [min=4.076, max=8.398]

Start Cancel Help



**Score Inactives**

Adjusted survival score =

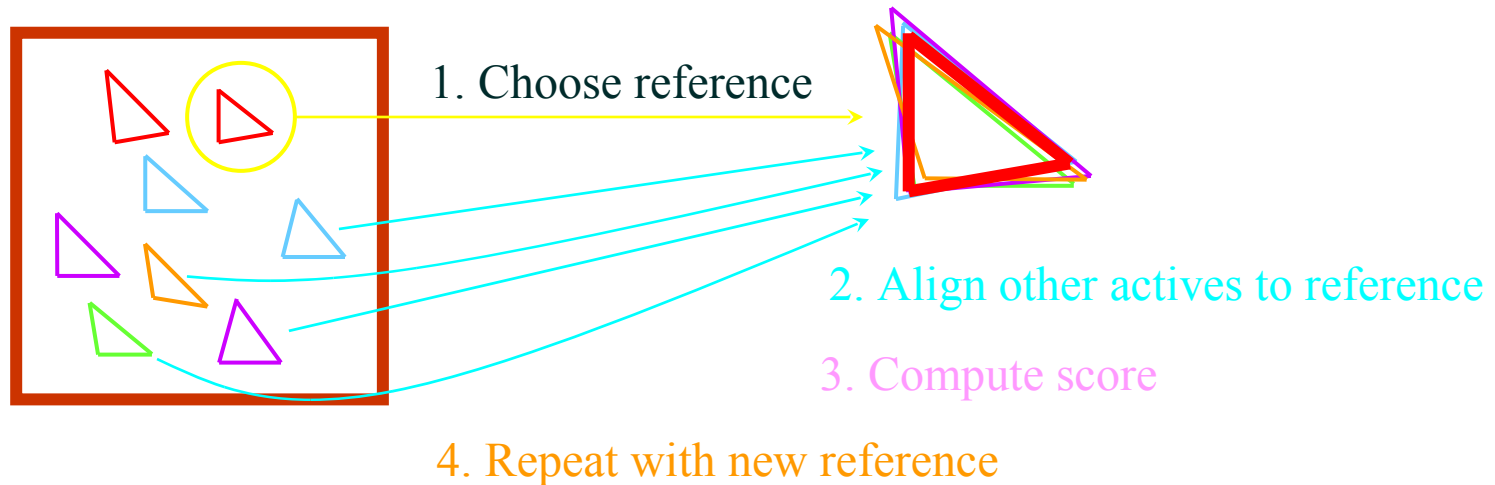
Survival score - (  \* inactive match score )

Start Cancel Help

- Score with Respect to Actives
  - Rank by alignment quality (site, vector, volume)
  - Favor hypotheses from high-activity ligands
  - Favor hypothesis from low-energy conformers
- Score with Respect to Inactives
  - Penalize hypotheses that match inactives

# Score Hypotheses - Details

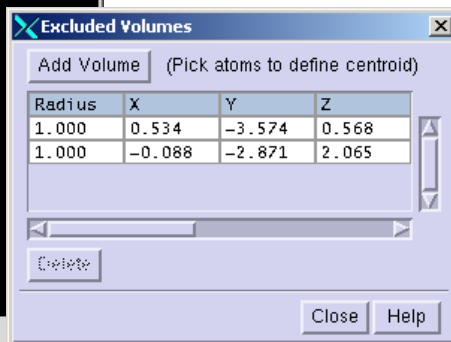
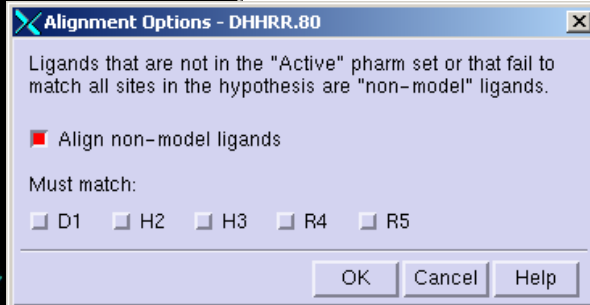
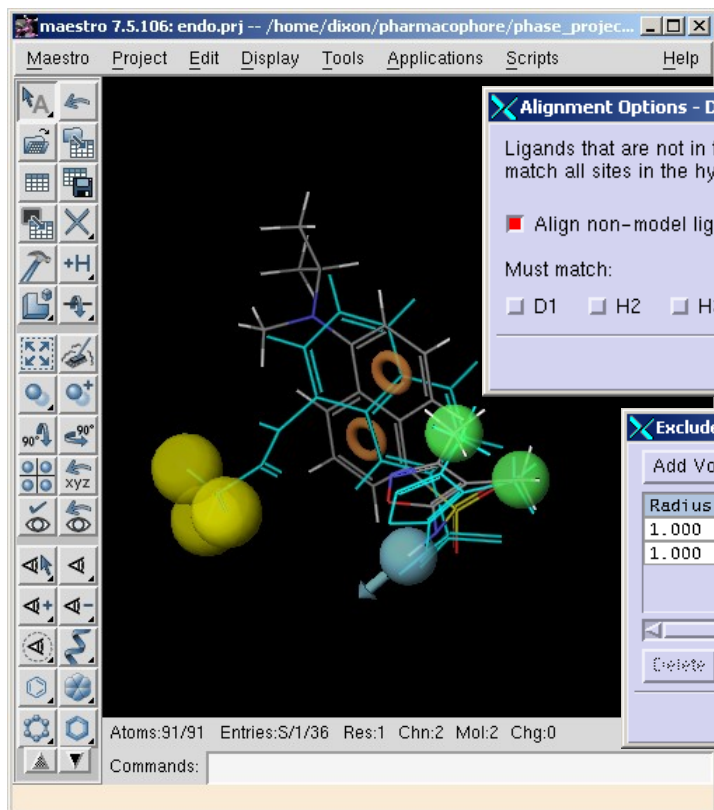
## Surviving Box



- Highest Scoring Reference → Hypothesis
- Use Inactives to Adjust Score
  - Penalise (re-rank) hypotheses if they match inactives

i.e. At-least 1 P from each of the 5 actives is mapped into same finite box ->LS alignment-> best alignment score (RMSD, Vector, Vol) -> which best P?

# Visualize Hypotheses



Develop Pharmacophore Model --- run1

File Display Step

Score Hypotheses Score and examine hypotheses, define excluded volumes, and select a hypothesis for the creation of a QSAR model.

Score Actives... Score Inactives... Rescore...

Hypotheses:

In	ID	Survival	Survival -inactive	Post-hoc	Site	Vector	Volume	Selectivity	# Matches
◇	DHHR. 61	14.335	12.894		0.98	0.999	0.844	2.118	5
◇	DHHR. 57	14.335	12.894		0.98	0.999	0.844	2.118	5
◇	DHHR. 80	14.312	13.408		0.98	1.000	0.870	2.068	5
◇	DHHR. 82	14.312	13.408		0.98	1.000	0.870	2.068	5
◇	DHHR. 79	14.279	13.370		0.97	1.000	0.842	2.068	5

Export... Delete Excluded Volumes...

Alignments for hypothesis DHHR.80:

In	Ligand Name	Activity	Pharm Set	Fitness	# Sites Matched	Relative Energy
◇	Molecule-4	6.824				
◇	Molecule-5	7.824	active	2.90	5	4.616
◇	Molecule-6	7.060				
◇	Molecule-7	7.167				
◇	Molecule-8	8.398	active	3.00	5	5.117
◇	Molecule-9	8.301	active	2.85	5	4.424
◇	Molecule-10	5.620				
◇	Molecule-11	6.076				

Alignment Options Model ligands only

< Back Search for Matches Next >

Prepare Ligands Create Sites Find Common Pharmacophores Score Hypotheses Build QSAR Model

Close Help

- View Multi-ligand Alignments
- Create Excluded Volumes based on inactives (drag on screen)

# Geometric Clustering of Hypotheses in GUI/CL

- Hierarchical, agglomerative clustering
- Similarity based on site and vector scores of aligned hypotheses

HypoID	Survival	Cluster	Size	AvgSim
-----				
AHHRRR_109	14.7308	1	6	0.98398
AHHRRR_113	14.7308	1	6	0.98398
AHHRRR_116	14.7054	1	6	0.98812
AHHRRR_112	14.7054	1	6	0.98812
AHHRRR_111	14.6287	1	6	0.99073
AHHRRR_115	14.6287	1	6	0.99073
-----				
AHHRRR_136	14.7314	2	3	0.98584
AHHRRR_135	14.6595	2	3	0.98895
AHHRRR_133	14.7592	2	3	0.98091
-----				
.				
.				
.				
-----				
ADHRRR_187	14.6018	69	1	0.00000
-----				
AADHRR_157	14.0773	70	1	0.00000

$$Sim(i, j) = \frac{\langle i | j \rangle}{\sqrt{\langle i | i \rangle \langle j | j \rangle}}$$

Essentially equivalent

Singleton (unique)

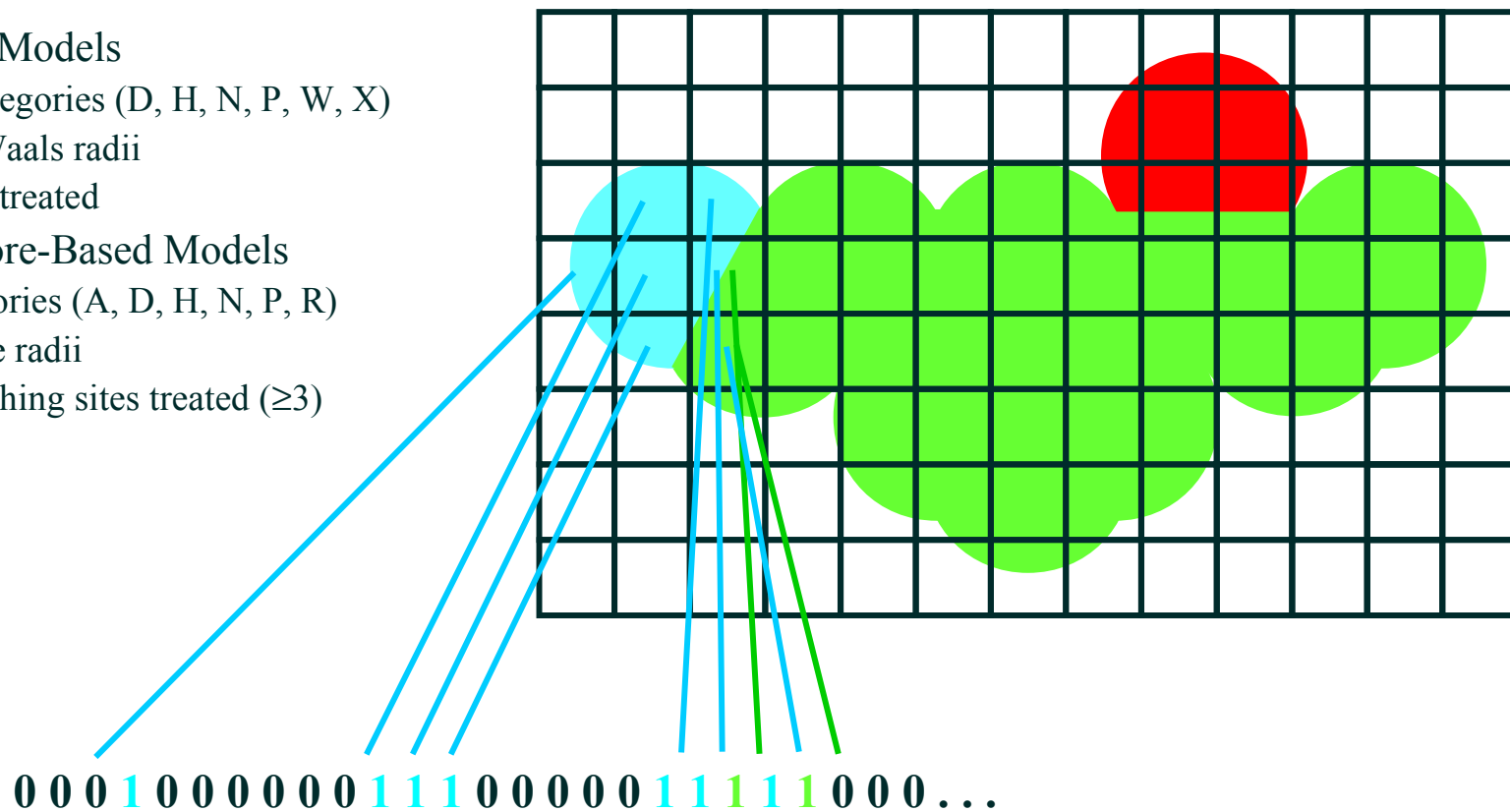
---

# **QSAR Models In Phase Wizard**



# Build QSAR Models

- Atom-Based Models
  - 6 atom categories (D, H, N, P, W, X)
  - Van der Waals radii
  - All atoms treated
- Pharmacophore-Based Models
  - Site categories (A, D, H, N, P, R)
  - Adjustable radii
  - Only matching sites treated ( $\geq 3$ )

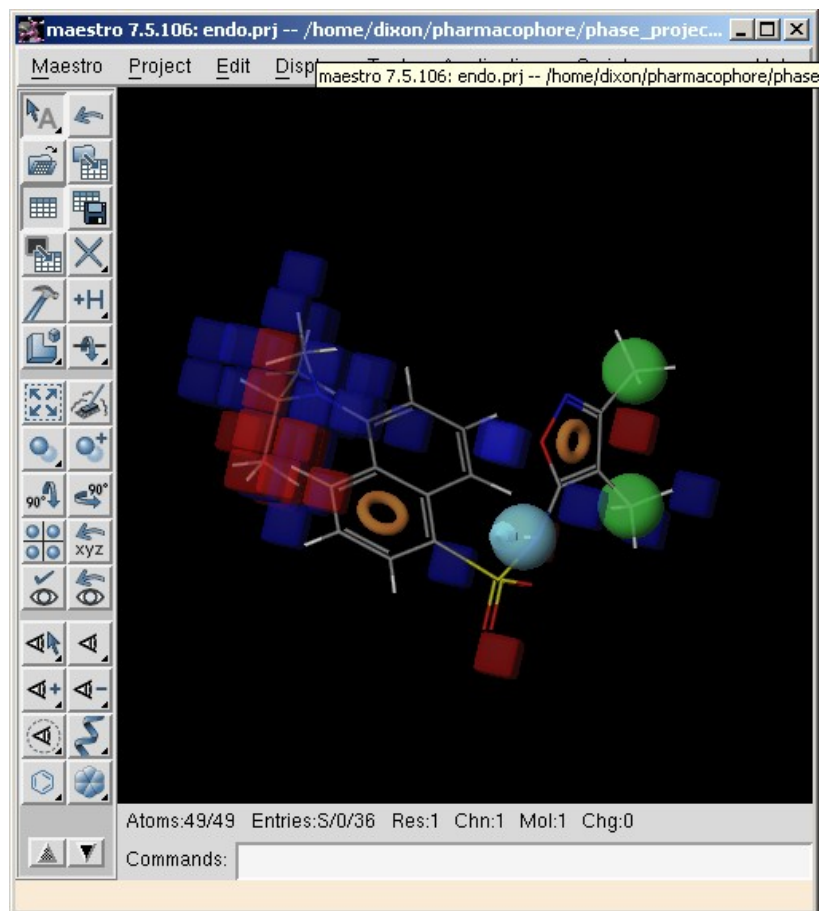


Ligand bit string – independent variables for PLS

Ligand activity – dependent variable for PLS

QSAR models based on PLS, applied to set of bits that encode if ligand atoms or ligand features occupy various cube-shaped elements of space (space that's occupied by aligned training set ligands)

# Analyze/Visualize QSAR Models



Develop Pharmacophore Model --- run1

File Display Step

**Build QSAR Model** Build QSAR models and examine the predicted structure-activity relations for the matching ligands.

Build Models Options... QSAR Visualization...

Hypothesis Scores QSAR Results

In	ID	# Factors	SD	R-squared	F	P	RMSE	Q-squ
◇	DHRRR_58	1	0.8271	0.4885	23.9	5.001e-05	0.8571	0.4495
		2	0.6857	0.6625	23.6	2.184e-06	0.7364	0.5936
		3	0.4402	0.8667	49.8	3.191e-10	0.5274	0.7916
◇	DHRRR_61	1	0.8309	0.4837	23.4	5.641e-05	0.9301	0.3515
		2	0.6942	0.6541	22.7	2.934e-06	0.7853	0.5375
		3	0.4352	0.8607	51.2	2.458e-10	0.5847	0.7438

Export... Delete Excluded Volumes...

Alignments for hypothesis DHRRR\_58:

In	Ligand Name	QSAR Set	Activity	# Factors	Predicted Activity	Pharm Set	Fitness
◇	Molecule-7	test	7.167	1	6.87		2.82
				2	6.93		
				3	6.85		
◇	Molecule-8	training	8.398	1	7.11	active	3.00
				2	7.43		
				3	7.91		
^	Molecule-9	test	8.301	1	7.09	active	2.85
				2	7.36		

Random training set: 75 % Apply

< Back Search for Matches

Prepare Ligands Create Sites Find Common Pharmacophores Score Hypotheses Build QSAR Model

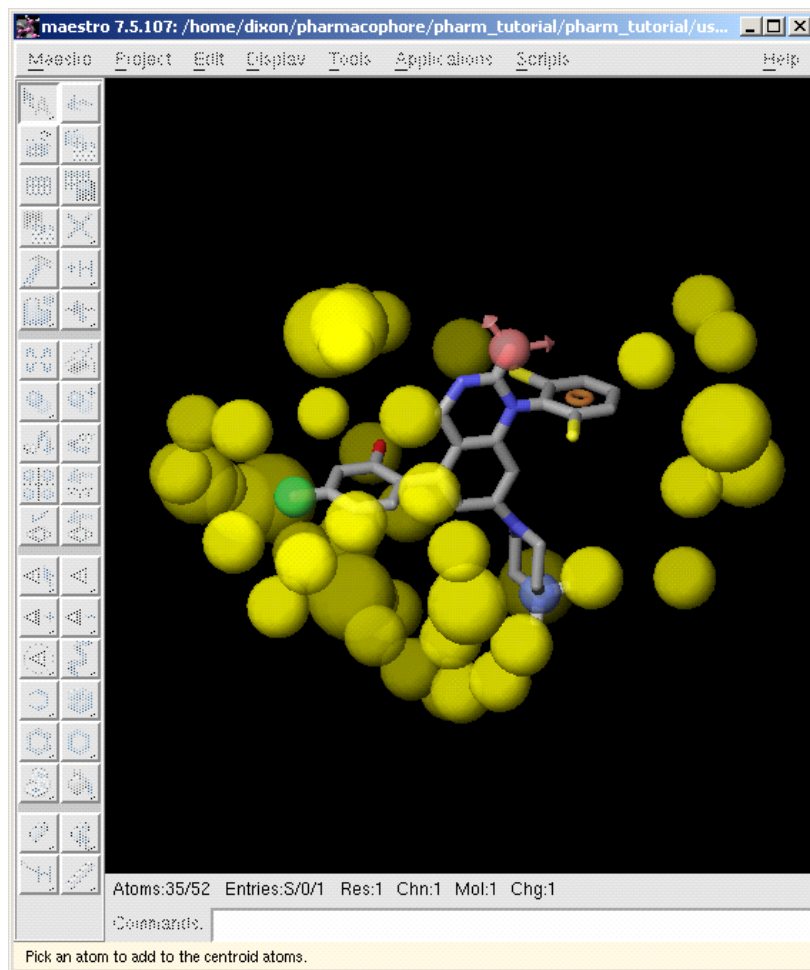
Close Help

- Compare Statistics
- Visualize Models
  - Good vs. bad regions
  - Broken down by ligand and occupation type

---

## Receptor-Based Excluded Volumes

# Receptor-Based Excluded Volumes



Phase Excluded Volume Receptor GUI

Choose Phase Hypothesis

Browse...

Select Atoms from Workspace Receptor

☒ Panel(ASL) Select  Clear

☐ Workspace Finish  Clear

Excluded Volume Spheres

Radii Sizes

☒ Use vdw radii of receptor atoms

☐ Use a fixed radius

☐ Use atom-level Maestro property

Radii Scaling Factors

☒ Use a fixed scaling factor

☐ Use atom-level Maestro property

Sphere Filters

☐ Ignore receptor atoms whose surfaces are within  Å of ligand surface

☐ Limit excluded volume shell thickness to  Å

Excluded Volume Output Option

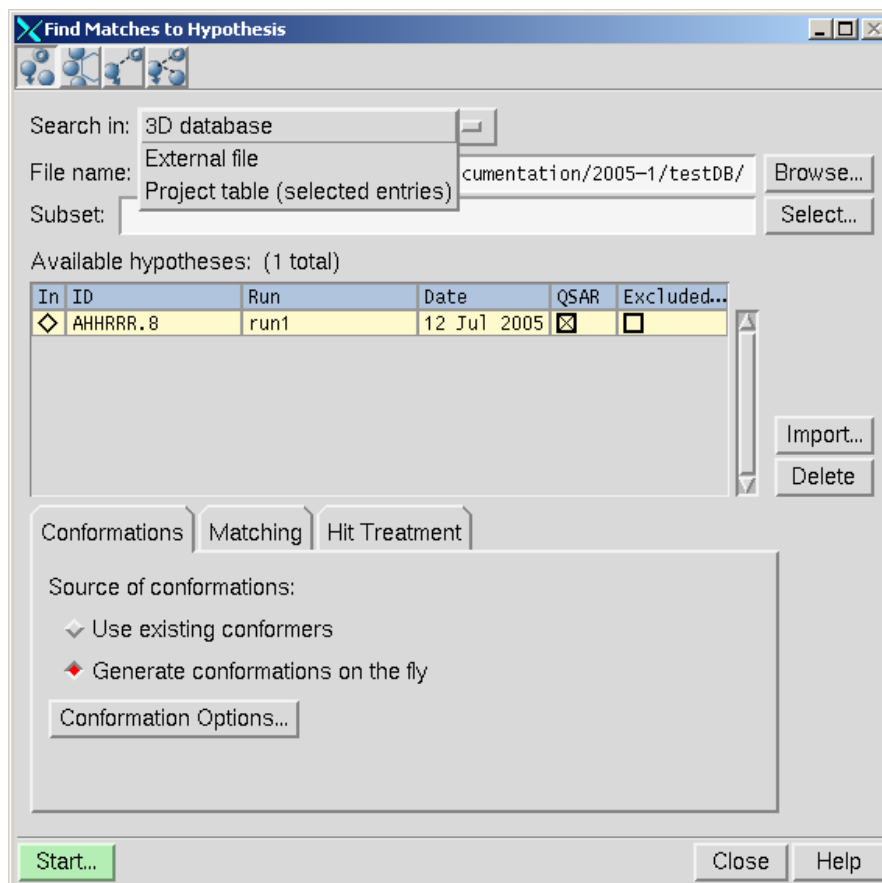
☐ Append to existing excluded volumes

Create Excluded Volumes Reset Close Help

---

# Find Matches to Hypothesis

# Searching Domains



- 3D Database (Binary SQLite format)
- External File
- Project Table Entries
- Search Existing Conformations
- Generate Conformations On-the-Fly

# Matching Options

**Find Matches to Hypothesis**

Search in: Project table (selected entries) ☐ Remote database

File name:  Browse...

Subset:  Browse...

Hypothesis: 257:AAHNR.225 Choose... Import...

Conformers Matching Hit Treatment

Mode: ☒ Find new matches ☐ Use saved matches

☐ Generate sites during search

☐ Limit CPU time for matching to 0.1 seconds / molecule

☒ Save all matches to disk

☐ Perform pre-screen using database keys

Distance matching tolerance: 2.0 Å Advanced...

Must match on at least: 6 site points

☒ Prefer partial matches involving more sites

☐ Consider atom types when computing volume scores

Start... Write... Close Help

Disable to force the search to consider 4-point and 3-point matches even if 5-point matches are found

Compute volume using only the overlap between atoms of the same MacroModel type

Advanced Matching Options

?

X

Permitted matches

Prohibited matches

Row	Site	Tolerance	Must match	A	D	H	N	P	R	X	Y	Z	A	D	H	N	P	R	X	Y	Z
1	A4	1.00	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	A5	1.00	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	H7	1.50	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	N9	0.75	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	R12	1.50	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	R13	1.50	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Apply:

☐ Tolerances

☐ Must match

☐ Permitted / prohibited matches

Save to Hypothesis

OK

Cancel

Help



# Hit Filters

**Find Matches to Hypothesis**

Search in: 3D database

File name: /scr/dixon/db\_benchmark/2006-1/GUI/testDB Browse...

Subset: Browse...

Hypothesis: 98 Select... Import...

**Conformers** **Matching** **Hit Treatment**

☐ Use QSAR model ☐ Apply excluded volumes

Return at most: 1 hits per molecule, 1000 hits total

Fitness = 1.00 \* (1.0 - align score / 1.20 )  
+ 1.00 \* vector score + 1.00 \* volume score

Reject hits with: ☐ align score > 1.20 ☒ vector score < -1.00 ☒ volume score < 0.00

Start... Close Help

- Enable/Disable Excluded Volumes
- Control Hits/Molecule
- Reject hits based on cutoffs

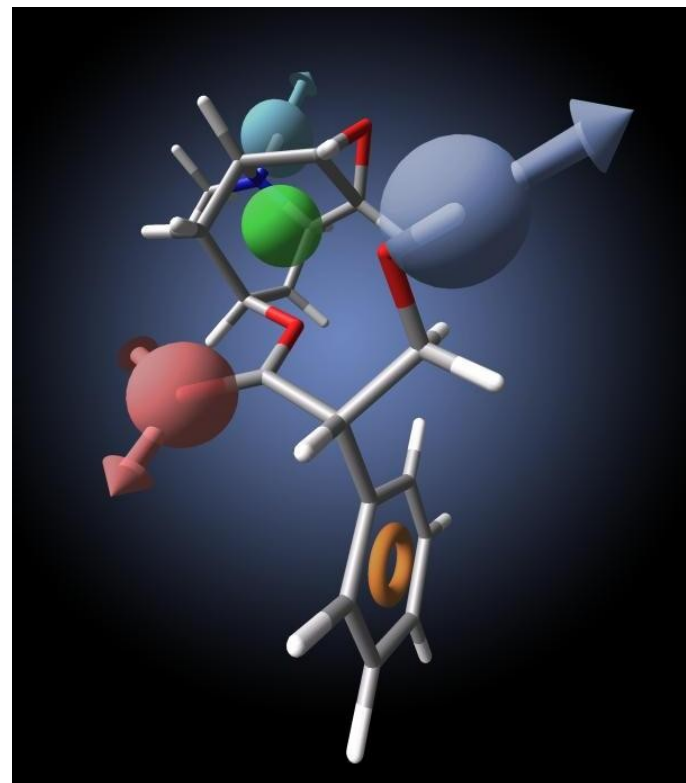
# Contact Information

**Steve Dixon**

**Product Manager, Ligand-Based Drug Design**

**Software Inquiries**

**[help@schrodinger.com](mailto:help@schrodinger.com)**



[www.schrodinger.com](http://www.schrodinger.com)

**SCHRÖDINGER.**