Chapter 12
Medicinal Chemistry

*This chapter illustrates applications of **Spartan** to problems of relevance to medicinal chemists.*

Medicinal chemistry is a diverse field, the common thread being its concerns with the design, properties and behavior of molecules that are important in biological systems. The tutorials in this chapter illustrate some of the roles that calculations might play.

The first addresses simple models to estimate the rate at which molecules are transported between blood (hydrophilic media) and the tissue in the brain (hydrophobic media); this is not something that can be directly calculated, and instead we will rely first on the so-called polar surface area (PSA) and later refine this using electrostatic potential maps.

The second and third tutorials use similarity analysis to identify "new" molecules (drugs) that are structurally or functionally similar to molecules with known drug action. In the second, involving the antihistamine terfenadine, similarity analysis is not carried out between molecules but between a molecule and an environment. In the third, similarity analysis is used to suggest compounds that "look like" morphine and hence might exhibit similar analgesic action.

The last tutorial accesses the on-line Protein Data Bank (PDB). While calculations on biopolymers (proteins and RNA/DNA strands) are limited to molecular mechanics models, quantum chemical calculations can be routinely applied to molecules bound to biopolymers. The focus is not the molecule itself, but rather the binding environment. This is turned into a so-called pharmacophore (a simplified representation of the environment); a search for other molecules that may also be consistent with this environment is illustrated.

no calculations

Anticipating Blood-Brain Transport

To be effective, a drug must be capable of transportation to its target. For orally administered drugs, this includes transfer through cell membranes in the intestine. In the case of most neural drugs, this also involves traversing the blood-brain barrier. In this tutorial, you will first examine how well the polar surface area (the area of space-filling model due to nitrogen and oxygen atoms together with any attached hydrogens) correlates with log (C_{brain}/C_{blood})*.

1. Open ***blood brain transport*** from the ***medicinal chemistry*** sub-directory (***Tutorials*** directory).** This comprises a list of drugs/drug candidates for which experimental data relating to the ratio of concentrations in the brain and in the blood (C_{brain}/C_{blood}) are available. These span a range of nearly 5 log units.

2. Select **Spreadsheet** from the **Display** menu () and size the spreadsheet to show all compounds, several extra rows and three columns, in addition to the column of experimental data **log(brain/blood)** that is already displayed. Select **Properties** from the **Display** menu () and *click* on the **QSAR** tab in the dialog that results. *Click* on  (Post) to the left of **PSA**. Polar surface area values will fill one column in the spreadsheet.

3. *Click* on the **Add** button at the bottom of the spreadsheet and select (*click* on) the **Linear Regression** tab at the top of the dialog that appears. Select **log(brain/blood)** from the **Fit** menu and *click* on **PSA** in the box under **Using**. *Click* on **Apply**. Select **Properties** from the **Display** menu () and then *click* on the cell labeled **Fit1** at the bottom of the spreadsheet under the **Label** column. This brings up the **Regression Properties** dialog in which RMSD and R^2 values are reported. The smaller

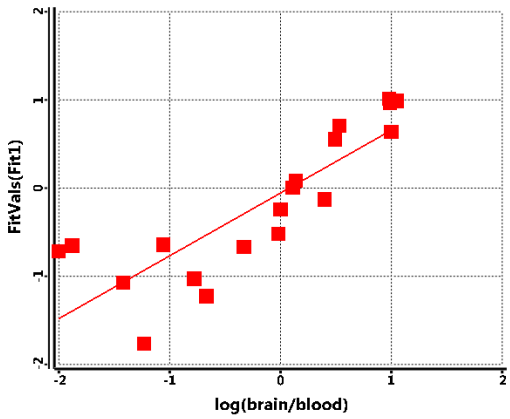
* J. Kelder *et al*, Pharmaceutical Res., **16**, 1514 (1999). All 19 named compounds considered in this paper have been included.

** For Windows, this directory is found in ***Program Files/Wavefunction/Spartan20***. It must be copied to another location available to the user prior to opening it in ***Spartan***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located at the top level of the ***Spartan'20*** disc image.

RMSD and the closer to unity R^2, the better the fit. Close the **Regression Properties** dialog.

4. Make a plot of log(brain/blood) vs. the regression fit. Select **Plots** from the **Display** menu ( ). *Click* on in the bar at the top of the plots pane to bring up the **Plots** dialog. Select **log (brain/blood)** from the **X Axis** menu, **Fit Vals (Fit 1)** from the **Y Axes** list and *click* on **Create**. *Click* on and check the box to the left of **Curve** and the radio button next to **Least Squares** and finally the **Done** button.

A least-squares line is drawn through the points.



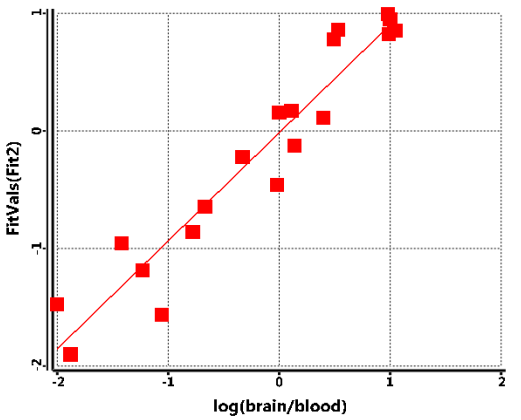
Next, you will consider an alternative definition of polar surface area based on electrostatic potential maps.

5. Select **Surfaces** from the **Display** menu ( ) and select ***electrostatic potential map*** in the dialog that results. An electrostatic potential map for the selected drug appears (maps for the remaining drugs may be seen by stepping through the list using the step buttons in the bottom left of the screen). The overall size and shape is that of the electron density and corresponds roughly to a conventional space-filling model. The colors indicate the value of the electrostatic potential. Colors toward red designate areas of negative potential (where a positive charge is most likely to be attracted), while colors toward blue designate areas of positive potential (where a positive charge is least likely to be attracted). To see the molecular skeleton underneath the electrostatic potential map,

change to a transparent or mesh display. *Click* on the map to select it and select **Transparent** or **Mesh** from the **Style** menu that appears at the bottom right of the screen.

6. Select **Properties** from the **Display** menu ( ). *Click* on the map. *Click* on the button to the left of **P-Area** the **Surfaces Properties** dialog that results. Polar areas, defined as the area for which the absolute value of the electrostatic potential is > 100 kJ/mol*, will be added to the spreadsheet.

7. *Click* on the **Add** button at the bottom of the spreadsheet and *click* on the **Linear Regression** tab at the top of the dialog that results. Select **log(brain/blood)** from the **Fit** menu and *click* on **Polar Area** from the quantities in the box under **Using**. *Click* on **Apply**. Bring up the **Regression Properties** dialog by *clicking* on the cell labeled **Fit2** under **Label** in the spreadsheet. Note that the R^2 value is better (closer to unity) for this fit than for the previous fit (to PSA defined in the usual manner). Make a plot of log (brain/blood) vs. the new regression fit. Select **Plots** from the **Display** menu ( ) and *click* on in the bar at the top of the plots pane. Choose **log (brain/blood)** from the **X-Axis** menu and **FitVals (Fit 2)** from the **Y Axes** list and *click* on **Create**. *Click* on and again specify **Least Squares** for the **Curve** in the **Edit Plot** dialog and *click* **Done**. Which property provides better correlation with transport across the blood/brain barrier, PSA or polar area?



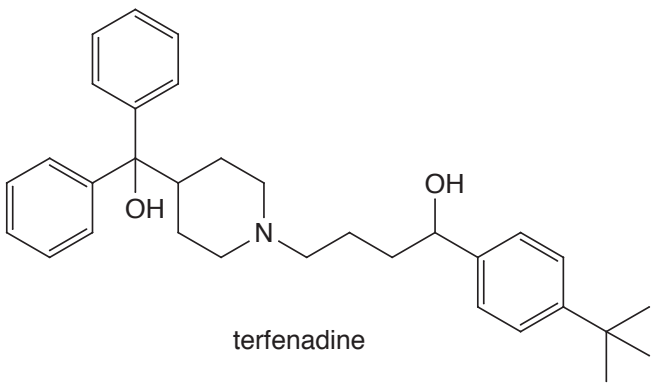
* This value may be changed in the **Settings Preferences** dialog (**Preferences** under the **Options** menu; **Chapter 25**).

8. Close ***blood brain transport*** and any open dialogs.

5 mins

Terfenadine. A Potassium Channel Blocker?

Cardiovascular toxicity due to blocking of potassium ion channels is commonly screened early in the drug development process. One way this is done is to compare drug candidates to a pharmacophore deduced from 3D QSAR studies. This tutorial uses a published pharmacophore* to see whether the antihistamine, terfenadine should be considered a potential channel blocker.



terfenadine

1. Open ***terfenadine similarity library*** from the ***medicinal chemistry*** sub-directory (***Tutorials*** directory).** This file contains the coordinates for a diverse selection of several hundred conformers of terfenadine (the actual number is reported in the **Molecule Properties** dialog) obtained from a similarity library calculation. It has been supplied in order to save computer time. Close ***terfenadine similarity library***.

2. Open ***potassium channel blocker pharmacophore*** from the ***medicinal chemistry*** sub-directory (***Tutorials*** directory). This is a five-point pharmacophore comprising four aromatic hydrophobe centers (purple spheres) and one positive ionizable

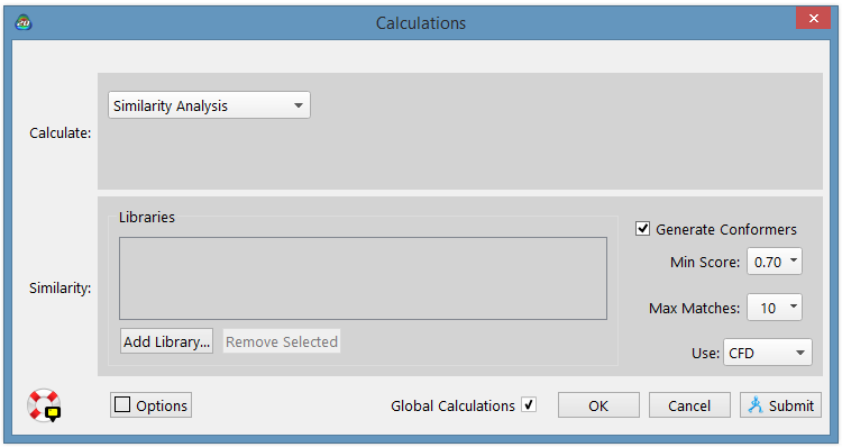
* S. Ekins *et al*, J. Pharmacology and Experimental Therapeutics, **301**, 427 (2002).

** For Windows, this directory is found in ***Program Files/Wavefunction/Spartan20***. It needs to be copied to another location available to the user prior to opening it in ***Spartan***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located at the top level of the ***Spartan'20*** disc image.

center (red sphere).

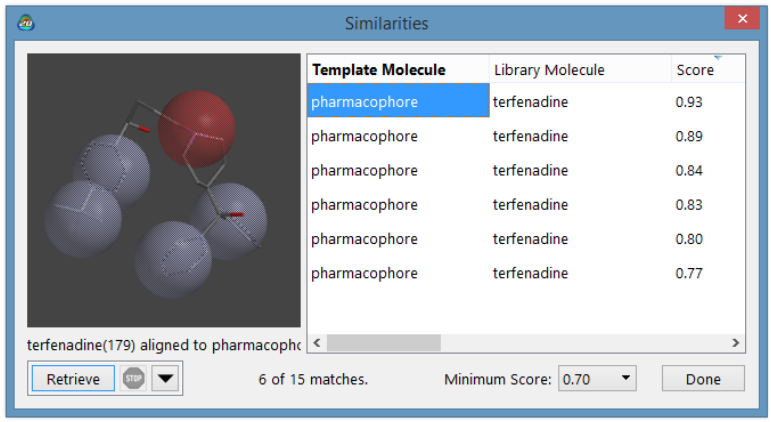
3. Select **Set Similarity Centers** from the **Geometry** menu ( ) and then select **CFD** from the menu at the bottom right of the screen next to **Similarity by:**. In turn, *click* on each of the (five) pharmacophore elements. In response, each will be surrounded by a violet circle indicating that it is to be used in the similarity analysis calculation.

4. Select **Calculations...** from the **Setup** menu ( ). Select **Similarity Analysis** from the top left menu to the right of **Calculate** inside the **Calculations** dialog. This leads to a new configuration of the **Calculations** dialog.



Click on **Add Library** to bring up a file browser. Locate ***terfenadine similarity library*** in the ***medicinal chemistry*** sub-directory (***Tutorials*** directory) and *click* on **Open**. This indicates that the library of terfenadine conformers will be searched for a match to the pharmacophore. *Click* on **Submit**.

5. When the analysis has completed, select **Similarities...** from the **Display** menu ( ) to bring up the **Similarities** dialog.



Multiple hits will appear at the right of the dialog, together with a similarity score (limiting on 1.0 which means a perfect match). *Click* on the hit with the highest score. A composite graphic of terfenadine and the pharmacophore will appear in a window at the left of the dialog. You can manipulate the two (as a single object) in the usual way (you need to position the cursor inside the window). Note that the positive ionizable center is positioned above nitrogen, while the four aromatic hydrophobes are above phenyl rings and the *tert*-butyl group.

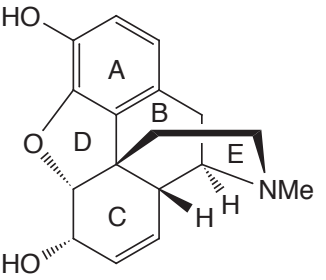
> The scoring algorithm not only accounts for position but also for direction of any nitrogen or oxygen centers that overlap with (non-hydrophobe) pharmacophore elements.

6. You can, retrieve a selected hit (without retrieving the pharmacophore), by *clicking* on the **Retrieve** button at the bottom left of the **Similarities** dialog. This can be used for further analysis.

9. Close all open documents and dialogs.

15 mins

Morphine. Structure vs. Pharmacophore



morphine

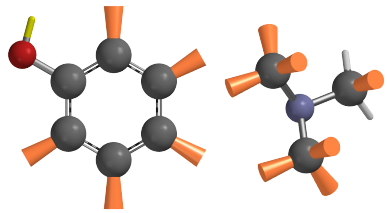
Three "chemical aspects" of morphine appear to be required in order for the molecule to act as an analgesic. These are: i) the nitrogen center (assumed to be protonated in the protein-bound complex), ii) the aromatic ring and iii) the hydroxyl group attached to the aromatic ring. The loss of any of these results in significant reduction of activity. This knowledge may either be used directly to identify other likely analgesics based on structure, or indirectly to construct a simple 3-point pharmacophore that in turn may be employed to find compounds with potentially comparable analgesic activity.

This tutorial comprises three parts. In the first part, you will identify molecules in the Spartan Spectra and Properties Database (SSPD) that incorporate the three required elements. In practice, you will look for molecules that incorporate a nitrogen center substituted by a methyl group and two other (sp^3) carbons and a substituted phenol. You will then select a single hit and generate a conformer library for similarity analysis simply in order to save computer time. This is a superficial investigation. To be comprehensive, one would need to use several (or all) hits.

1. Select **New Build** from the **File** menu ( ) and bring up the organic model kit. Select **Benzene** from the **Rings** menu and *double click* anywhere on screen. Select sp^3 oxygen ( ) and *click* on one of the free valences of benzene. You have made phenol. Select sp^3 nitrogen ( ), *double click* on a blank region on screen to insert a non-bonded sp^3 nitrogen. Two

molecules (phenol and ammonia) appear on screen. *Click* on sp^3 carbon ( ) and *click* on all three free valences on the nitrogen. Phenol and trimethylamine now appear on screen. *Click* on .

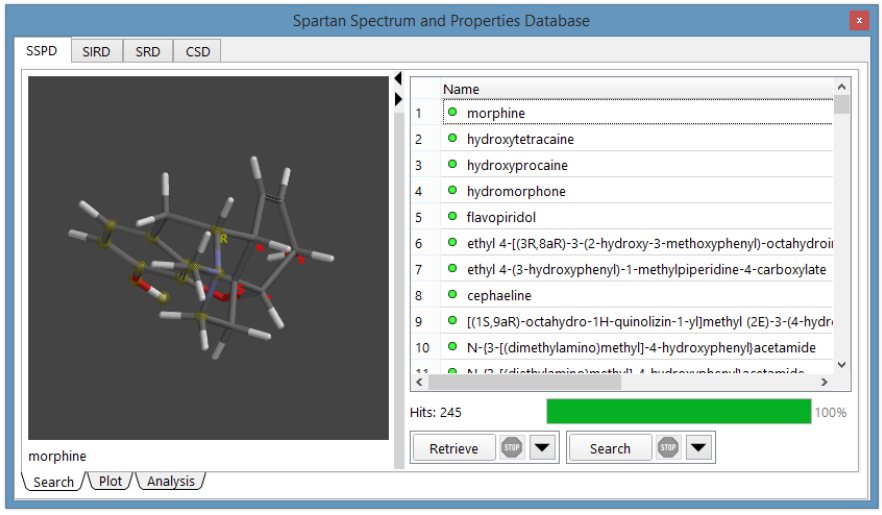
2. Select **Structure Query** from the **Search** menu ( ). *Click* on all five free valences on the carbons in the phenol fragment and on all three free valences on two of the three methyl groups and on one free valence for the remaining methyl group in the trimethylamine fragment. Orange cones will appear at all (twelve) selected positions.



This defines a search in which anything can be "grown" from the selected positions (including hydrogen) but a hit must contain phenol and a nitrogen bonded to three sp^3 carbon centers.

3. Select **Databases** from the **Search** menu ( ). *Click* on the **SSPD** tab in the dialog that results. *Click* on **Search** at the bottom of the **SSPD** dialog.

4. To sort in alphabetical order, *click* on **Name** at the top of the list inside in the scroll box at the right of the **SSPD** dialog. Find and *click* on ***cephaeline***.*



* The search above was done on the full ≈300,000 molecule ***SSPD***. The smaller subset bundled with the ***Spartan'20*** install will find significantly fewer "hits".

Its structure will be displayed in the window at the left of the dialog. You can manipulate the model in the usual way by positioning the cursor inside this window. *Click* on  to the right of the **Retrieve** button at the bottom of the **SSPD** dialog to bring up the **Retrieve Options** dialog. *Check* **New Document** under **Retrieve Options** and *click* on **OK**. Finally, *click* on the **Retrieve** button. Dismiss the **SSPD** dialog either by again selecting **Databases** from the **Search** menu () or by *clicking* on  at the top right-hand corner of the **SSPD** dialog.

Save the molecule you retrieved under the name ***cephaeline***. Close the file used for the search (phenol and trimethylamine) without saving.

5. For similarity library generation, it is advised to use a more restrictive conformational analysis routine. For this example, the molecule retrieved from the database has been set to do so. To confirm this, select **Set Torsions** from the **Geometry** menu (). Your molecule should be augmented with three yellow cylinders (only). If this is *not* the case, select **Preferences** from the **Options** menu. In the **Settings** tab, switch the **Conformer Rules** to **skeletal**, and *click* **OK**. You will still be in the **Set Torsions** mode, in the lower right of the interface, click on the **Reset** button to apply the changes to your molecule.

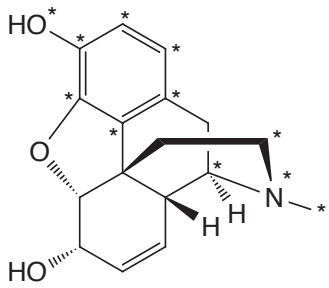
6. Select **Calculations...** from the **Setup** menu (). Select **Similarity Library** from the top left menu to the right of **Calculate**. No further information is required. *Click* on **Submit** at the bottom of the dialog. Name it ***cephaeline library***. You can continue with the tutorial while you are waiting for the job to complete. Note, however, that you need to close ***cephaeline library*** before it is used in **Step 9**.

In the second part of this tutorial, you will specify the key structural components in morphine and perform a similarity analysis based on structure using the conformer library for cephaeline.

7. You can either sketch morphine (much easier than building it in 3D) or to save time, open it: ***morphine*** from the ***medicinal***

chemistry sub-directory (***Tutorials*** directory).* If you sketch it, *click* on when you are done. The name should appear at the bottom of the screen as morphine is in SSPD. If it does not, you have made a mistake.

8. Select **Set Similarity Centers** from the **Geometry** menu ( ) and select **structure** from the menu at the bottom right of the screen. *Click* on the eleven *'ed atoms in the figure below. These atom centers define the phenol and the primary amine discussed earlier. A violet circle will be drawn around each. If you accidentally select the wrong atom, *click* on the "offending" circle to deselect it.

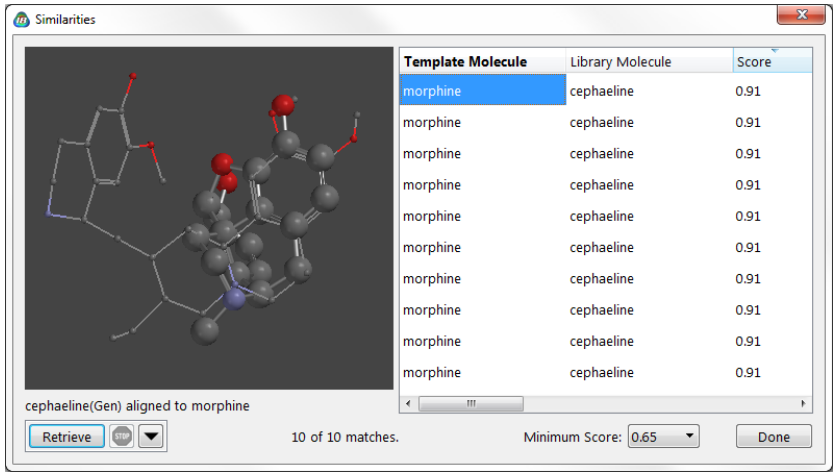


9. Select **Calculations...** from the **Setup** menu ( ). Select **Similarity Analysis** from the top left menu to the right of **Calculate** inside the **Calculations** dialog. *Click* on **Add Library** to bring up a file browser, locate and select ***cephaeline library***, and *click* on **Open**. Make certain that **Structure** is selected from the **Use** menu at the lower right of the dialog. You have requested that the library of cephaeline conformers be searched for a match to the key structural elements of morphine. *Click* on **Submit**. Supply the name ***morphine***.

10. The similarity analysis will require a few seconds. When it completes, select **Similarities...** from the **Display** menu ( ).

---

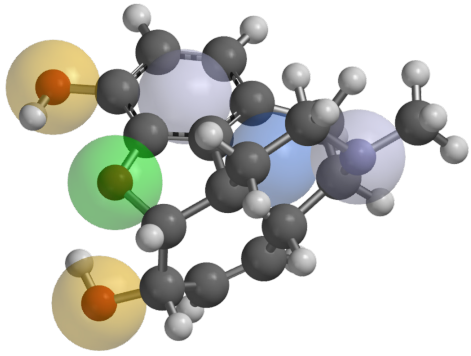
* For Windows, this directory is found in ***Program Files/Wavefunction/Spartan20***. It must be copied to another location available to the user prior to opening it in ***Spartan***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located at the top level of the ***Spartan'20*** disc image.



11. Hits appear in the box at the right of the dialog together with a score. Sort according to score (best on top) by *clicking* on **Score** at the top of the box. Select (*click* on) one of the best matches. A composite graphic consisting of morphine and the matching cephaeline conformer from the library will appear in a window at the left. You can manipulate the two structures (as a single object) in the usual way.

Repeat the similarity analysis using CFD's rather than structure.

12. Select **Set Similarity Centers** from the **Geometry** menu ( ), but this time select **CFD** from the menu at the bottom right of the screen. CFD's will augment the morphine structure.



13. Designate three of the CFD's as similarity centers. Select (*click* on) the CFD over the phenolic oxygen. A violet circle will surround this center indicating that it is to be used in the similarity analysis. (If you accidentally select the wrong CFD,

click on the circle and it will disappear.) Select (*click* on) the CFD at the middle of the phenol ring and the CFD at nitrogen.

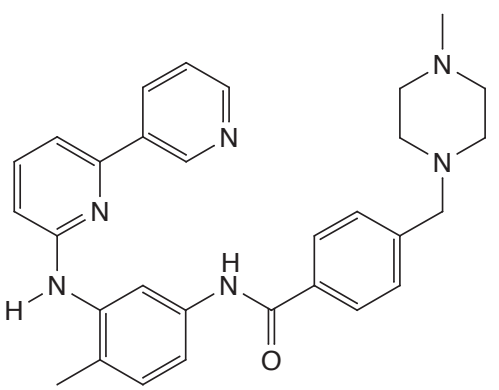
14. Select **Calculations...** from the **Setup** menu ( ). This should already designate **Similarity Analysis** and the appropriate library (***cephaeline***). Change the entry in the **Use** menu to **CFD**. *Click* on **Submit**.

15. When the analysis is complete, select **Similarities...** from the **Display** menu ( ). One or more hits should appear in the box at the right of the dialog. Select a hit to get a composite graphic (CFD and structure from the library). You can easily see the extent to which the two are matched.

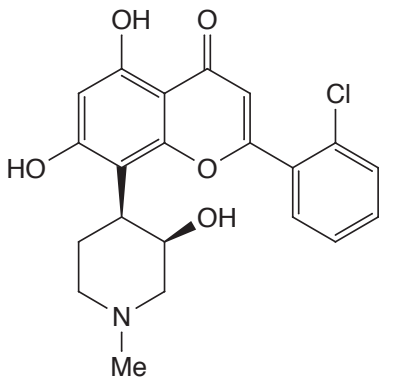
16. Close any open documents and dialogs.

5 mins

Gleevec. Making a Pharmacophore from PDB



gleevec



flavopiridol

In this tutorial, you will start with a protein structure from the PDB in which the so-called protein kinase inhibitor, gleevec, is found*. Instead of abstracting the molecular structure of gleevec, you will abstract its "binding environment", that is, the locations from which gleevec interacts with the protein host in terms of hydrogen-bond or charge-charge interactions (non-steric contacts). This information,

* B. Nager *et. al*., Cell, **112**, 859 (2003): PDB identification 1opj.

together with knowledge of either the steric requirements of the guest or the space occupied by the protein (excluded volumes), constitutes a structural pharmacophore, that is, a "skeleton" into which other possible "guests" might be accommodated. You will then use a subset of the non-steric contacts together with excluded volume elements to see to what extent this pharmacophore fits another protein kinase inhibitor, flavopiridol.

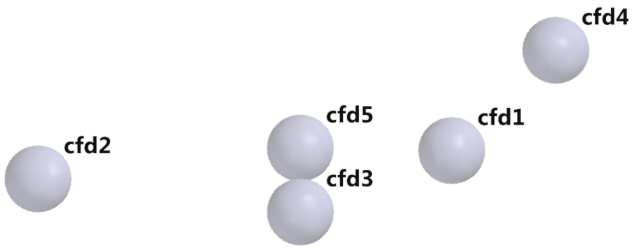
1. Select **Access PDB Online** from the **File** menu () to obtain the PDB file ***1opj*** from the PDB or open ***1opj*** from the ***medicinal chemistry*** sub-directory (***tutorials*** directory)*. Select **Extract Ligands** from the **File** menu (), *click* on the model for gleevec (it will be designated as Ligand STI 3 or 4) inside the protein structure (displayed as a set of transparent spheres similar to a space-filling model) to select it, and *click* on the **Exact Ligands** button at the lower right of the screen. Inside the **Extract Ligands** dialog, *check* both **HBA/HBD** and **+/- Centers** and **Excluded Volume Centers**. If any other items are selected, make certain to deselect them before you *click* on **OK**.

2. You are finished with ***1opj*** and may close it. Also, simplify the display of the extracted information by removing the excluded volume elements from view. Select **Configure...** from the **Model** menu () and *click* on the **CFDs** tab. *Uncheck* **Excluded Volumes** and *click* on **Apply** (this will hide the excluded volumes and leave the **Configure** dialog on screen).

3. The model that now appears comprises five purple spheres representing non-steric contacts between gleevec and its protein host. Turn on labels. In the **Configure** dialog, *click* on the **Labels** tab, *check* **CFD Labels** and *click* on **OK**.

---

* For Windows, this directory is found in ***Program Files/Wavefunction/Spartan20***. It must be copied to another location available to the user prior to opening it in ***Spartan***. For Linux, this is found in the directory where ***Spartan*** was installed. For Macintosh, this is located at the top level of the ***Spartan'20*** disc image.



These have been assigned based on the observation that each is within a predefined distance of (one or more) complementary groups from residues in the protein host. Because hydrogen positions are not assigned in the experimental X-ray structure it may not be possible to say whether a particular site on the guest is acting as a hydrogen-bond donor or acceptor. It may not even be possible to say whether a particular site is protonated or deprotonated (although pH may often dictate this). To be "safe", all non-steric contacts have initially been given four CFD definitions: hydrogen-bond acceptor (**HBA**), hydrogen-bond donor (**HBD**), positive ionizable (**+**) and negative ionizable (**–**). Examination of the structure of gleevec suggests more focused assignments: **1** (**HBA**, +), **2** (**HBA**), **3** (**HBD**), **4** (**HBA**) and **5** (**HBA**). (See previous figure for numbering.)

4. You will use only three elements of this pharmacophore (together with the excluded volume elements) to assess whether or not flavopiridol is likely to fit into the same host environment as gleevec. Select **Properties** from the **Display** menu ( ). *Click* on **CFD1**, remove the checks next to **HBD** and **-Ionizable** (leaving **HBA** and **+Ionizable**). Repeat the process, designating **CFD3** as **HBD** only, and **CFD5** as **HBD** only. Next, select **Set Similarity Centers** from the **Geometry** menu ( ). Ensure that the **CFD** (not **Structure**) is selected in the lower right corner of the screen. *Click* **CFDs 1**, **3**, and **5**. A violet circle will appear on each, indicating that it is to be used in the similarity analysis calculation. *Click* on .

5. Select **Calculations...** from the **Setup** menu ( ) and then **Similarity Analysis** from the top left menu to the right of **Calculate**. *Click* on **Add Library...** and then locate and select

flavopiridol in the ***medicinal chemistry*** sub-directory (***Tutorials*** directory) and *click* on **Open**. Make certain that **CFD** is selected next to **Use:** in the lower right of the **Calculations** dialog. *Click* on **Submit**. Name the document ***flavopiridol fit to gleevec pharmacophore***.

6. When the similarity analysis has completed, select **Similarities...** from the **Display** menu (). Sort the hits according to similarity score (*click* on **Score** at the top of the dialog). One after the other, *click* on the top scoring entries in the box to the right of the dialog and examine the fits in the window at the left.

7. Close any open documents and dialogs.