

Automatic Protonation state handling in eHiTS

Darryl Reid, Zsolt Zsoldos.
 Info@simbiosys.ca
 SimBioSys Inc., 135 Queen's Plate Drive, Suite 520
 Toronto, ON, M9W 6V1, Canada

The issue of protonation state is very important to the docking problem. Ligands and receptors with different protonation states can have dramatically different binding poses. However, it is common practice for many docking programs to ignore this issue and require that the user define a particular protonation state prior to running a docking experiment.

Protonation states of ligands and receptors are determined by the interaction between the two. Thus for any particular receptor-ligand pair there will generally be one correct protonation state. However for a different ligand, the protonation state of the receptor may be altered, to reflect the characteristics of the ligand. If a docking program were to pre-set the protonation state of the receptor then possible interactions with a ligand could be lost. A better solution, with a more appropriate score, can be found only if the program is run with different protonation states (not necessarily the neutral or the normally lowest energy form of the complex) or in solvent, but the form required to reach the lowest energy for the complex).

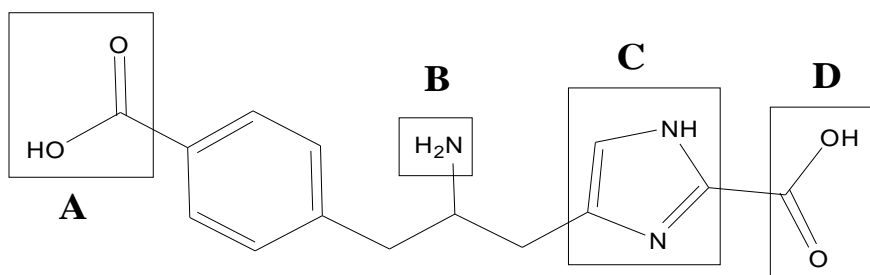


Figure 1: Sample ligand with 150 different possible protonation states. Functional groups A, B, C and D have various different forms important to docking.

Table 2: Possible protonation states for the functional groups identified in the molecule in Figure 4

A & D	
B	
C	

The molecule in Figure 4 has 150 possible protonation states. Table 2 shows the 5 possible protonation states for each of A and B, 2 for B and 3 for C, combined this leads to $5 \times 5 \times 2 \times 3 = 150$ different possible protonation states. Although, two pairs of states for A&D can be considered equivalent via rotations about the bond to R (swapping the roles of the 2 oxygen atoms), so a flexible docking program could work using only 3 protonation states for those fragments giving a total of $3 \times 3 \times 2 \times 3 = 54$ instead of 150. Most docking programs would need to dock all 150 (or at least 54) combinations separately to evaluate the different possibilities, not even considering different protonation states of the receptor.

eHiTS takes a unique approach to the protonation problem. eHiTS systematically evaluates all possible protonation states for the receptor and ligands, automatically for every receptor-ligand pair. It does this through the use of ambiguous chemical surface point types for positions that could be either protonated or deprotonated (i.e. have a lone pair). Then during the docking algorithm each state is evaluated and scored, selecting the best protonation state for each individual interaction without the combinatorial effect. For example, Figure 2 shows the interaction between a ligand atom (L) and a receptor atom (R). Here we can see two dummy surface point types placed on each L and R, the dummies can be either a hydrogen (H) or lone pair (LP). During the scoring stage of the eHiTS algorithm each of the possible four combinations are considered. Hydrogen versus Hydrogen and lone pair versus lone pair are ignored, since they will not give a good interaction, and a score is assigned to each of the other two possibilities. The interaction that gives the best score is the one that is selected for that point. By evaluating each interaction, one at a time, we can avoid the combinatorial explosion of all the combinations of possible protonation states. The result is a docking program that evaluates all possible protonation states for the receptor and ligand in a single run.

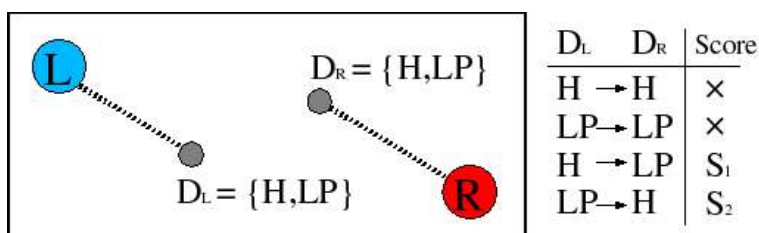


Figure 2: Interaction between a dummy surface point on a receptor atom and a dummy surface point on a ligand atom. The indicated dummy surface points can either represent a hydrogen or a lone pair. The possible interactions are listed in the table to the right, these are scored by eHiTS and the best score is assigned to the interaction.

Figure 3 shows an example of how the chemical properties (i.e. surface point types) are applied to the ring of a histidine residue. Both nitrogen atoms have the same surface point type towards the edge of the ring (Ambivalent H/LP) which indicates that it may be a donor H OR acceptor Lp depending on protonation state. This simultaneously handles all three protonation states of histidine.

It may be possible to generate unlikely or invalid protonation states for the entire ligand using this approach, however these rare cases are filtered out by eHiTS at later optimization and scoring phases.

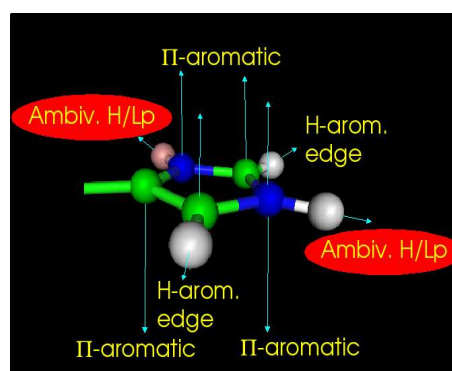


Figure 3: illustration of how surface point types are assigned to the ring portion of a histidine residue.