QCalc 5.0 scoring manual β 1.0

Peter Hanspers

May 29, 2005

1 Introduction

Three scoring functions are now implemented in QCalc: X-Score [1], ChemScore [2] and PMF-Score [3]. X-Score and ChemScore are empiricial whilst PMF-Score is knowledge based.

2 Scoring

All scoring functions require the topology to be loaded and the correct mask be specified. The initial topology (with coordinates from the .top-file) can be scored to verify atom typing.

Both trajectory and restart files can be scored. The following options were added to QCalc when specifying trajectory or restart files:

- adding ",frames=every n" means calculations will only be performed on every n:th frame.
- adding ",frames=n-m" means calculations will only be performed on frames n to m.
- specifying "mean" instead of a file name calculates the mean of all frames processed since start or since the last time "mean" was specified.

The input requested is similar for all three functions. To avoid confusion, examples of typical inputs are now given.

2.1 X-Score

2.1.1 Input

Prompt	Input
Topology file:	c:/peter/data/P450/adm/adm.top
Qcalc>	xscore
Mask:	residue 1 407
Mask	.
Score initial topology? (yes/no)	yes
Q-atom (FEP) file:	c:/peter/data/P450/adm/lig.fep
Cofactor (. or EOL terminates):	restype=HEM
Cofactor (. or EOL terminates):	.
FF translation key:	qoplsaa
Scoring parameters:	xscore_default.input
Qcalc>	go
Enter names of coordinate or restart files	c:/peter/data/.../md.dcd,frames=every 5 mean

Cofactors X-Score uses different typing schemes to set atoms types for protein and ligands atoms. If needed, atoms in parts of the protein can be typed using the ligand atom-typing procedure by defining a cofactor. This is useful if the protein has some special residues, like HEM in P450, that are not defined in the X-Score residue library (file RESIDUE_DEF_XTOOL.dat). Ligand atoms are typed on the individual atom level, in contrast to the residue level for protein atoms, using data in file ATOM_DEF_XTOOL.dat. Cofactor definitions are made on separate lines and has the form: restype=RES, where RES is the residue name, e.g. HEM. All atoms, regardless of their proximity to each other, in residues named RES will be included in the cofactor RES and typed as if they were ligand atoms (though in every other respect they are considered as part of the protein). Any number of cofactors can be defined.

Force field A force field translation key has to be given to allow for the translation of atom types according to the Q convention to types according to the Sybyl convention. The translation tables are in the file ATOM_TYPE_CONVERSIONS.dat (shared by PMF-Score). A different translation file can be specified in the input file (see below).

Parameters Scoring parameters, output specifications and data files can be specified in an input file. Default parameters can be used by specifying "default" when asked for scoring parameters. In that case the following parameters and filenames are used:

SHOW_ABS	'NO'	! Show binding score for each atom?
SHOW_TOTAL	'YES'	! Show total binding score?
SHOW_LIGAND	'YES'	! Show ligand atoms?
SHOW_PROTEIN	'NO'	! Show protein atoms?
SHOW_COFACTOR	'YES'	! Show cofactor atoms?
APPLY_HPSCORE	'YES'	! Use hydrophobic contact algorithm?[1]
APPLY_HMSCORE	'YES'	! Use hydrophobic matching algorithm?
APPLY_HSSCORE	'YES'	! Use hydrophobic surface algorithm?
RESIDUE_DEFINITIONS	residue_def_xtool.dat	
ATOM_DEFINITIONS	atom_def_xtool.dat	
LOGP_DEFINITIONS	atom_def_xlogp.dat	
SURFACE_DEFINITIONS	surface_def_xtool.dat	
ATOM_TRANSLATIONS	atom_type_conversions.dat	

Default scoring coefficients are as reported in [1].

2.1.2 Output

When SHOW_LIGAND and/or SHOW_PROTEIN is specified, a list of ligand and/or protein atoms is displayed, showing the atom type, residue, atomic properties, neighbouring atoms and bonds for each atom. In addition, a list of bonds and aromatic rings detected are output.

When scoring, the contribution from each term (van der Waals (VDW), H-bonding (HB), hydrophobic contact (HP), matching (HM) and surface (HS) and rotational(RT)) is displayed along with the total score. If SHOW_ABS was specified the contributions for every ligand atom is displayed. Atomic binding score is always displayed when scoring the initial configuration.

X-Score results are in pKd units.

2.1.3 Data files

The format of the data files is further explained in the respective files.

2.2 ChemScore

2.2.1 Input

Prompt	Input
Topology file:	c:/peter/data/P450/adm/adm.top
Qcalc>	chemscore
Mask:	residue 1 407
Mask	.
Score initial topology? (yes/no)	yes
Q-atom (FEP) file:	c:/peter/data/P450/adm/lig.fep
Parameter file:	c:/peter/data/ff/chemscore_oplsaa.prm
Qcalc>	go
Enter names of coordinate or restart files	c:/peter/data/.../md.dcd,frames=every 5 mean

Parameter file ChemScore reads all atom parameters from a single parameter file, though there are different files for different force fields. The parameter file defines the atomic properties of all atom types.

2.2.2 Output

Prior to scoring, ChemScore outputs atom type and bond information for all ligand atoms, as well as information about rings detected. For every frame, the contribution from each term (H-bonds, metal contacts, lipophilic contacts and frozen rotatable bonds) along with the total score.

ChemScore results are in kJ/mol. A negative score means negative energy.

2.3 PMF-Score

2.3.1 Input

Prompt	Input
Topology file:	c:/peter/data/P450/adm/adm.top
Qcalc>	chemscore
Mask:	residue 1 407
Mask	.
Score initial topology? (yes/no)	yes
Q-atom (FEP) file:	c:/peter/data/P450/adm/lig.fep
Force field translation key:	qoplsaa
Scoring parameters:	pmfscore_default.input
Qcalc>	go
Enter names of coordinate or restart files	c:/peter/data/.../md.dcd,frames=every 5 mean

Presently, all atom defined as solvent atoms are ignored. The option of including critical water molecules defined as solvent will be added.

Force field As for X-Score, a force field translation key has to be given to allow for the translation of Q atom types to Sybyl atom types. The Sybyl type derived is only used to determine the hybridization of carbon and nitrogen atoms.

Scoring parameters Output options, data files and the maximum ring size considered are defined in an input file. The output options are similar to those in X-Score. The maximum ring size parameter determines the number of steps the ring finding algorithm will take in every search direction. Too low a setting will prevent the algorithm from finding all rings. Too high a setting will increase the time required for the ring search.

2.3.2 Output

When SHOW_LIGAND and/or SHOW_PROTEIN is specified, a list of ligand and/or protein atoms is displayed, showing the atom type, residue, atomic properties, neighbouring atoms and bonds for each atom. In addition, a list of rings detected are output. If specified, bonds are also output. Because of the nature

of the PMF-function, only the total score is displayed. Atomic binding score is always displayed when scoring the initial configuration.

It is safe to consider PMF-Score results as rankings. For details about converting the PMF-Score to free energy of binding, see [3]. A more negative score means better binding.

3 References

[1] Renxiao Wang, Luhua Lai and Shaomeng Wang, Further development and validation of empirical scoring functions for structure-based binding affinity prediction, Journal of Computer-Aided Molecular Design, 16, 11-26, 2002

[2] Matthew D. Eldridge, Christopher W. Murray, Timothy R. Auton, Gaia V. Paolini and Roger P. Mee, Empirical scoring functions: I. The deveoplment of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes, Journal of Computer-Aided Molecular Design, 11, 425-445, 1997

[3] Ingo Muegge and Yvonne C. Martin, A general and fast scoring function for protein–ligand interaction: A simplified approach, J. Med. Chem., 42, 701-894, 1999