

CScore™

RANK THE AFFINITY OF COMPOUNDS BOUND TO A TARGET
WITH CONSENSUS SCORING



CScore uses multiple types of scoring functions to rank the affinity of ligands bound to the active site of a receptor. The strengths of individual scoring functions combine to produce a consensus that is more robust and accurate than any single function for evaluating ligand-receptor interactions.

In virtual high throughput screening, compounds are "docked" into an active site and then "scored" to determine which ones are more likely to bind tightly to the target macromolecule. At present, no single scoring function can accurately predict the binding affinities for all types of ligands with all types of receptors. Consensus scoring is the simultaneous use of multiple different scoring functions in order to make virtual screening more predictive.

Applications

- Predict binding affinities of ligands against biological targets whose 3D coordinates are known
- Perform virtual screening to prioritize compounds for synthesis or screening
- Identify ligand-receptor binding modes

CScore is a consensus scoring program that integrates multiple, well known and extensively applied scoring functions from the scientific literature¹⁶: a GOLD™-like function, a DOC™-like function, ChemScore™, a PMF function, and FlexX™ (separately licensed).

Each individual scoring function is used to predict the affinity of ligands in candidate complexes. CScore creates columns in the SYBYL® Molecular Spreadsheet™ that report raw scores for each scoring function. CScore also creates a consensus column,

containing integers that range from 0 to the total number of scoring functions. Each complex whose score exceeds the threshold for a particular function adds one to the value of the consensus; configurations below the threshold contribute a zero. Consensus columns can also be calculated from any combination of numerical spreadsheet columns, so that key aspects of binding (e.g., the presence of a specific hydrogen bond) can be used to discriminate good configurations from bad ones.

D_score

The DOCK-like function is based on the work of Meng, Shoichet, and Kuntz.¹ It considers both electrostatic and hydrophobic contributions to the binding energy,² but does not include entropic terms. A distance-dependent dielectric attenuates charge-charge and other polar interactions. Nevertheless, Knegt et al.³ report that these terms tend to dominate the non-polar terms in this type of scoring function.

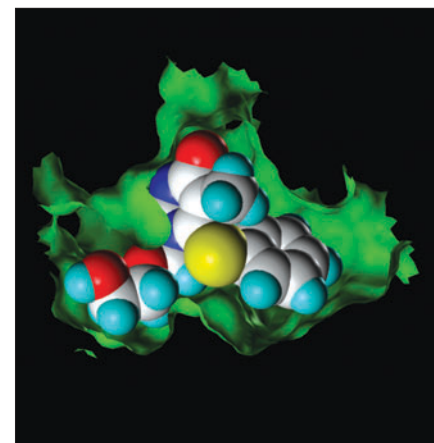
ChemScore

ChemScore⁴ is based on a diverse training set of 82 receptor-ligand complexes. It uses four terms that estimate contributions to binding energy from lipophilic interactions, metal-ligand binding, hydrogen bonding, and loss of ligand flexibility. A head-to-head evaluation of scoring functions demonstrated that ChemScore is more accurate and robust than many other widely used functions. The CScore implementation of ChemScore was validated in collaboration with the authors.

PMF_score

This function is based upon ligand-receptor atom-pair interaction potentials that are statistical in nature rather than empirical. Muegge and Martin⁵ derived these potentials from crystallographic data in the Protein Data Bank,⁶ drawing on hundreds

	2: Total_Score	14: G_SCORE	15: PMF_SCORE	16: D_SCORE	17: CHEMSCORE	18: CSCORE
1: HIV1_001	-15.86	-196.23	-44.58	-129.60	-30.32	5
4: HIV1_004	-12.61	-213.95	-46.13	-125.31	-30.13	5
6: HIV1_006	-12.47	-186.34	-48.97	-123.87	-28.39	5
8: HIV1_008	-12.06	-196.86	-50.60	-128.46	-29.54	4
11: HIV1_011	-10.99	-158.41	-73.84	-118.96	-25.02	4
14: HIV1_014	-9.96	-190.31	-68.71	-123.40	-24.64	4
20: HIV1_020	-8.70	-133.04	6.47	-94.17	-17.66	2



(Left) A Molecular Spreadsheet showing the results of a CScore analysis of inhibitors bound to HIV-1 Reverse Transcriptase. (Right) An inhibitor, 1-(2-hydroxyethyl-oxymethyl)-6-phenyl-thiothymine, bound to HIV-1 Reverse Transcriptase.

Advantages

- CScore offers multiple approaches for the evaluation of ligand-receptor interactions. The strengths of individual scoring functions combine to produce a consensus that is both more robust and more accurate than any single approach currently in use.
- CScore can be used to rank multiple configurations of the same ligand docked with a receptor, or to rank selected configurations of different ligands docked to the same receptor.
- The consensus column can be drawn from any combination of columns of a spreadsheet. This allows researchers to add functional data specific to a particular biological system to the CScore consensus.
- A subset extractor allows creation of selection criteria to identify the best docking results and then constructs a new database or Molecular Spreadsheet containing those results.

and sometimes thousands of examples of each interaction type. The resulting equation for all the atom-pair interactions in a protein-ligand complex can yield free energies directly, as solvation and entropic terms are treated implicitly. This type of scoring function tends to perform best at ranking different ligands bound to a single receptor.

G_score

The GOLD-like function is based on the work of Jones et al.^{7,8} and focuses on hydrogen-bonding interactions. A hydrogen-bonding term incorporates desolvation of donors and acceptors. A pair-wise dispersion term considers the contribution of non-polar or hydrophobic interactions, while a more typical molecular mechanics equation is used to calculate the internal strain energy of the ligand itself. This type of scoring function performs well in situations where there are significant polar interactions between the ligand and the receptor, but has difficulty with ligands that are primarily non-polar hydrophobic in nature.

FlexX

The FlexX scoring function^{9,10} (separately licensed) is based on an empirical function first reported by Böhm.¹¹ The equations divide hydrogen-bond, salt bridge, and non-polar interaction distances into a small number of discrete bins. An energy is assigned to each interaction based on which bin it occupies. Additional terms in the function are both entropic as well as enthalpic in nature, and consider the effect of buried surface area and the number of rotatable bonds in the ligand. The FlexX scoring function performs best in situations where the ligand has many site points (heteroatoms and ring centroids).

Features

- Integration of multiple scoring functions for evaluating binding affinity of ligands.
- Automated display of raw and consensus scores in the Molecular Spreadsheet.
- Quick comparison and analysis of different combinations of scoring methods.
- Numerical or functional data can also be used for consensus scoring.
- Relaxation of ligands in the receptor site prior to scoring.
- Tools for identifying the highest affinity ligands and extracting the results.

Hardware and Software Requirements

CScore is accessible through the SYBYL expert molecular modeling environment. It requires a license for SYBYL/Base in addition to the CScore license itself. FlexX scores require a FlexX license. SYBYL and CScore run on workstations operating under IRIX[®] (SGI[®]) or Linux[®] (x86).

Complementary Software

Integration of CScore in the SYBYL[®] expert molecular modeling environment allows users to access SYBYL's powerful molecular design and analysis tools.

- **CombiFlexX[®]** for rapidly docking combinatorial libraries into a receptor site.
- **Concord[®]** for generating accurate 3D coordinates.
- **Confort[™]** for generating sets of diverse, low energy conformers.
- **FlexX** for flexibly docking ligands into a binding site.
- **FlexX-Pharm[™]** for flexibly docking ligands under pharmacophore constraints.

- **FlexE[™]** for considering protein flexibility during docking calculations.
- **FlexS[™]** for performing shape-based screening of ligands in the absence of receptor structure.
- **MM3[™]/MM4[™]** for optimizing structures by molecular mechanics.
- **MOLCAD[™]** for visualizing molecular surfaces and molecular properties.
- **SiteID[™]** for finding and visualizing protein binding sites.
- **UNITY[®]** for locating compounds in databases that match a pharmacophore or fit a receptor site.

References

- Meng, C.; Shoichet, B. K.; Kuntz, I. D. *J. Comp. Chem.* **1992**, *13*, 505-524.
- Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comp. Chem.* **1986**, *7*, 230-252.
- Knegtel, R. M. A.; Bayada, D. M.; Engh, R. A.; v.d. Saal, W.; v. Geerestein, V. J.; Grootenhuis, P. D. J. *J. Comp.-Aid. Mol. Des.* **1999**, *13*, 167-183.
- Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolinine, G. V.; Mee, R. P. *J. Comp.-Aid. Mol. Des.* **1997**, *11*, 425-445.
- Muegge, Martin, Y. C. *J. Med. Chem.* **1999**, *42*, 791-804.
- Bernstein; Koetzle; Williams; Brice; Rodgers; Kennard; Shimanouchi; Tasumi. *Eur. J. Biochem.* **1977**, *80*, 319-324.
- Jones, G.; Willett, P.; Glen, R. C. *J. Mol. Biol.* **1995**, *245*, 43-53.
- Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727-748.
- Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, *261*, 470-489.
- Rarey, M.; Kramer, B.; Lengauer, T. *Bioinformatics* **1999**, *15*, 243-250.
- Böhm, H.-J. *J. Comp.-Aid. Mol. Des.* **1994**, *8*, 243-256.



WWW.TRIPOS.COM				CONTACT_US@TRIPOS.COM		
AUSTRALIA +61 (7) 5439 9775	CANADA +1 450 4334500	FRANCE +33 1 69 59 29 49	GERMANY +49 89 45 10 300	JAPAN +81 3 5166 1721	UNITED KINGDOM +44 1 908 650000	UNITED STATES 800 323 2960 +1 314 647 1099
©Tripos, Inc. 2005 0075 Printed in USA All trademarks are the property of their respective owners.						