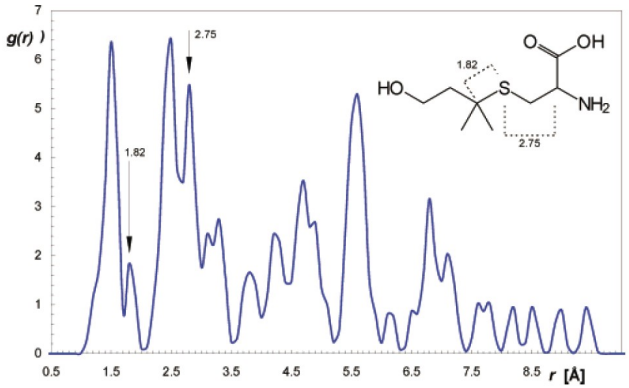
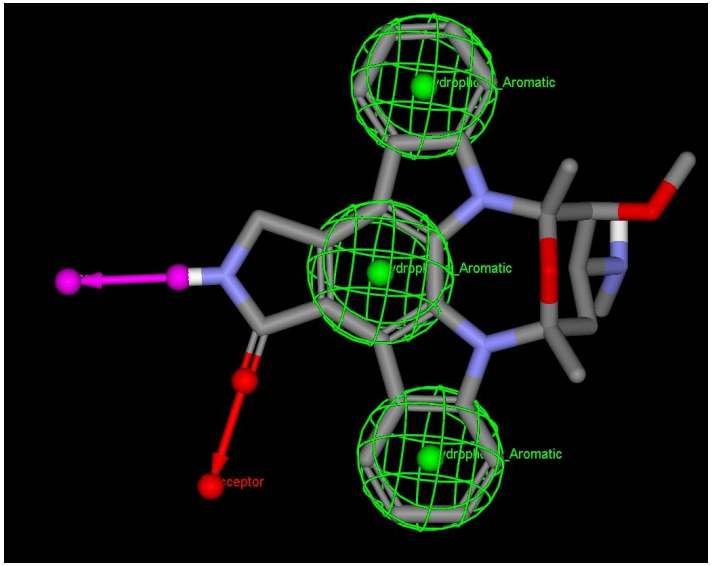
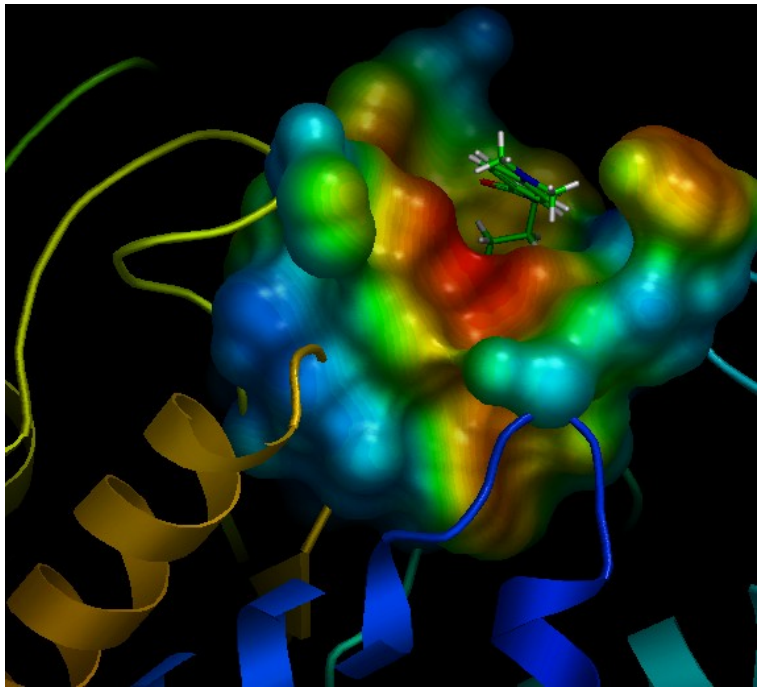
Virtual Screening

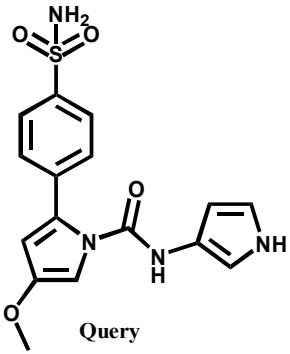
Anders Poulsen, S*BIO Pte Ltd

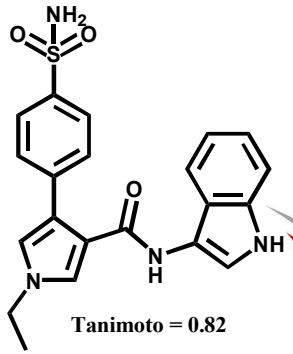


0100100110001101010011…









S*BIO™

Introduction

Understanding the relationship between a molecule and its biological activity are the central theme in medicinal chemistry.

But the biological activity of a molecule cannot be calculated from its chemical structure by first principles.

Problem:
Given a structure of the target or a ligand of the target find new ligands

S*BIO™

Introduction

When is Virtual Screening Relevant?

	Ligand Structure Known	Ligand Structure Not Available
Target Structure Known	Docking	Docking
Target Structure Not Available	Pharmacophore Modeling, Fingerprint Search, Autocorrelation Methods.	HTS

Appropriate for Computational Chemistry

Not Appropriate for Computational Chemistry

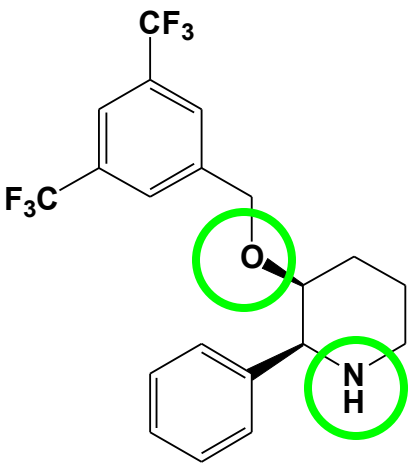
S*BIO™

Pharmacophore Concept

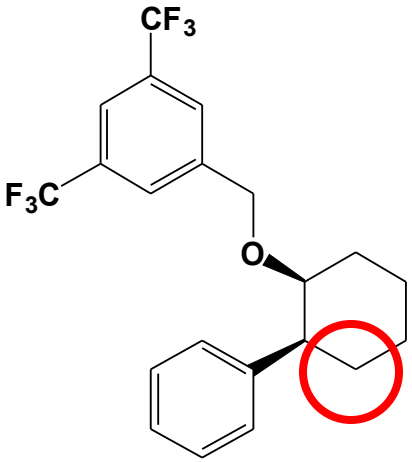
Observations that modification of some parts of a ligand results in minor changes of activity, whereas modifications of other parts of the ligand results in large change of activity.

Pharmacophore element: Atom or functional group essential for biological activity

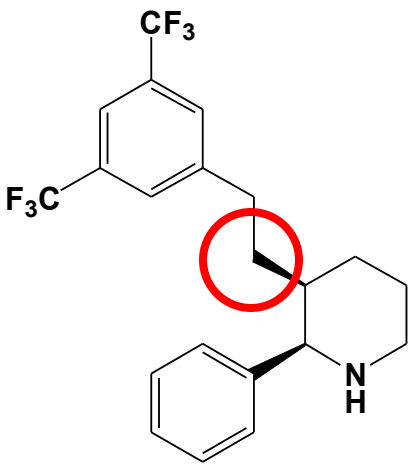
3D Pharmacophore mode: Collection of pharmacophore elements including their relative position in space



NK1 IC50 = 1.4 nM



NK1 IC50 = 2600 nM



NK1 IC50 = 1500 nM

S*BIO™

Concepts

Lock & Key Concept

The structure(s) of a molecule determines the biological activity of that molecule.

There must be a structural and electronic complementarity between the receptor and the ligand.

S*BIO™

Fingerprints

Describe the molecule as a bit-string of features.
For each feature: If feature present write 1
If feature absent write 0

...0100100110001101010011...

Fingerprints are almost unique.
Reduced representation that are easy and fast to calculate.
Excellent for searching large databases (>10 million compounds).

Answers to simple questions you can ask about molecules

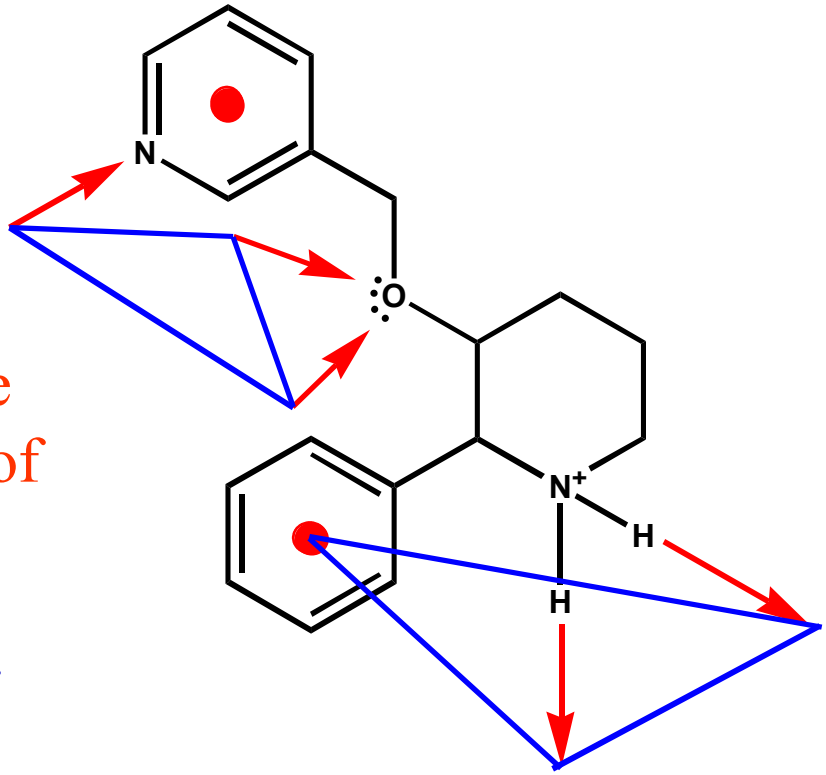
S*BIO™

Fingerprints

Example of Pharmacophore fingerprint

Each bit in the fingerprint represents a triangle with a particular geometry

Purely abstract representation based on the concepts of Hydrophobe, HB-donor and HB-acceptor.

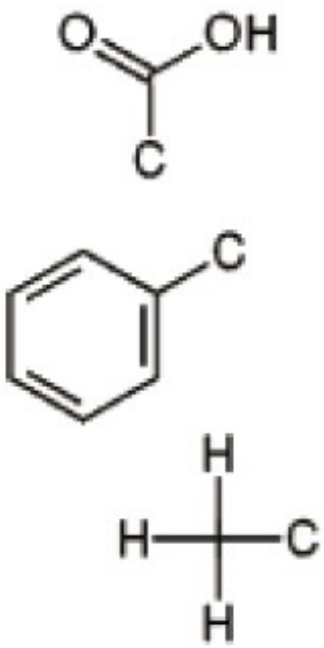


Red arrows: The representation (of the molecule)
Blue triangles: The features (of the fingerprint)

S*BIO™

Fragments

Describe molecule in terms of its buildingblocks and substituents
Connectivity between fragments are not encoded

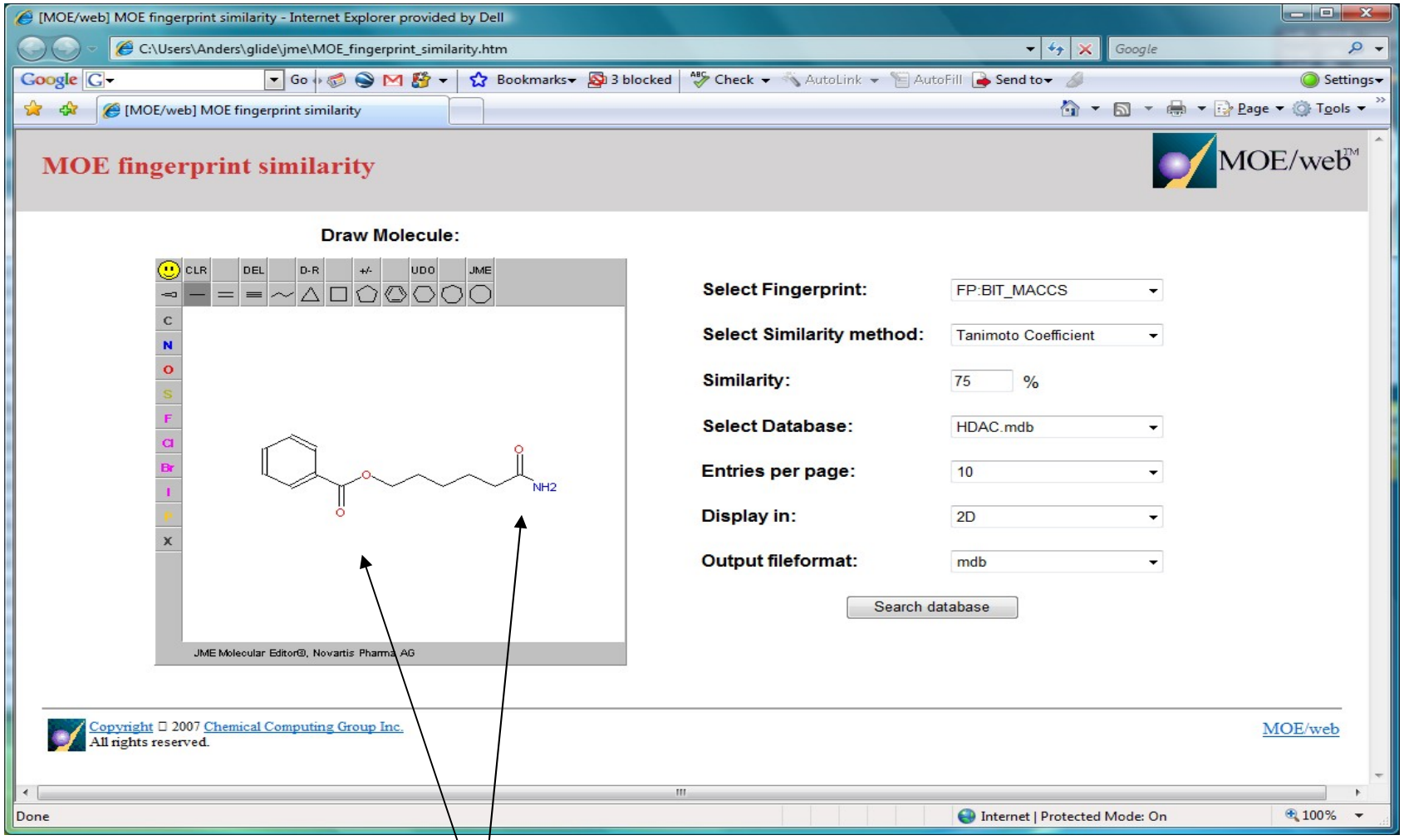


Closer to the traditional way that chemists think about molecules.
A fussy way of doing substructure search.

Link to Fingerprint Search

S*BIO™

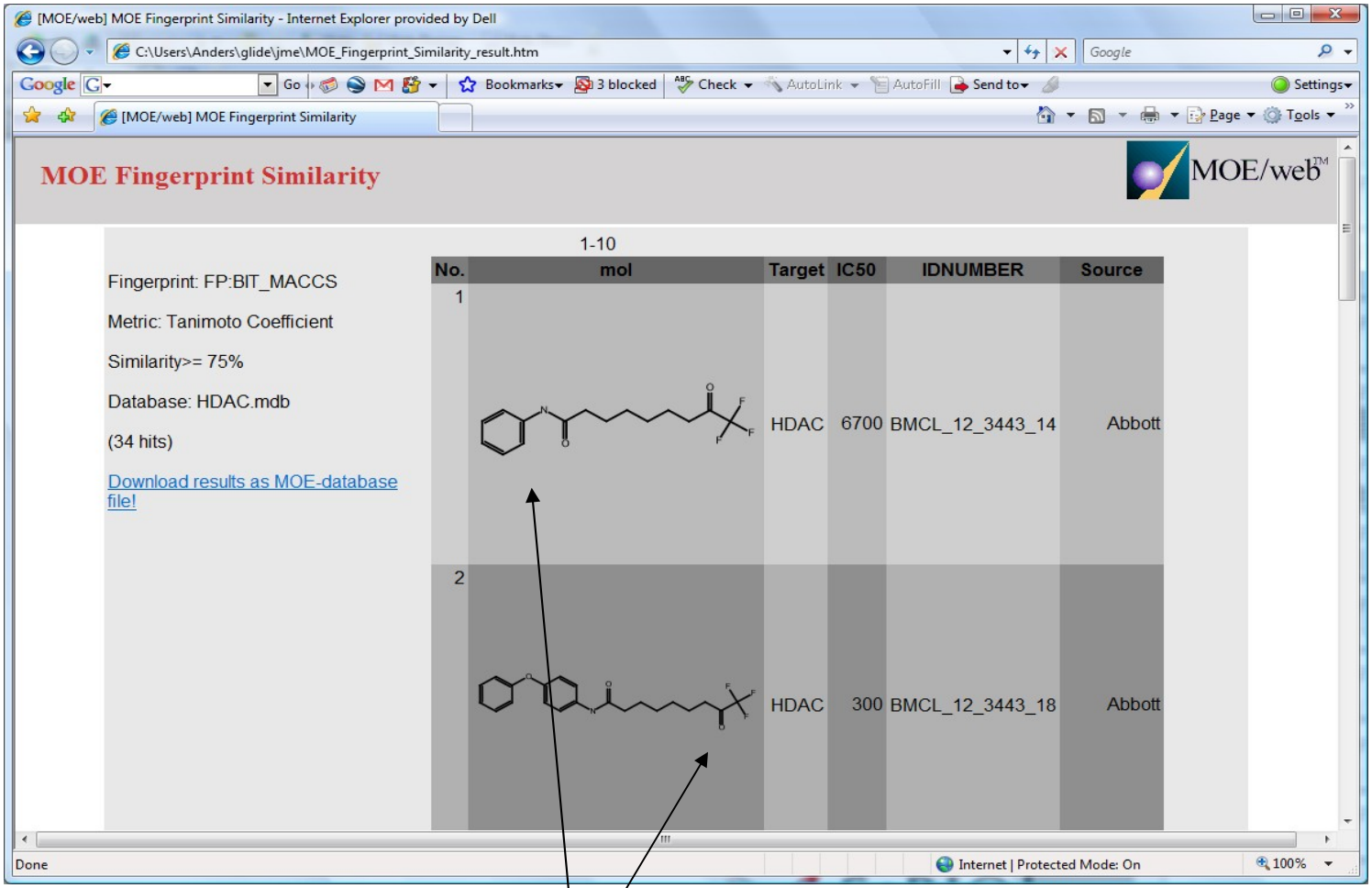
Fragments



Notice shape & functional groups

S*BIO™

Fragments



Notice shape & functional groups

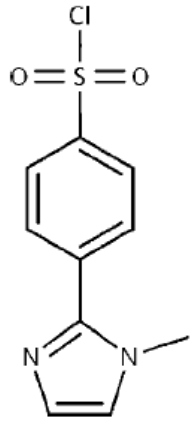
S*BIO™

Fragments

$$I = \frac{n_C}{n_A + n_B + n_C}$$

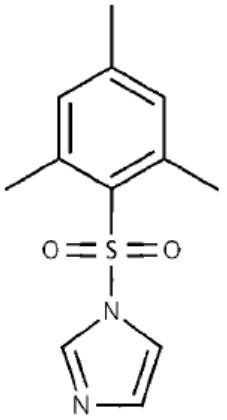
*n*A: Features in A not in B
*n*B: Features in B not in A
*n*C: Features in both A and B

A:



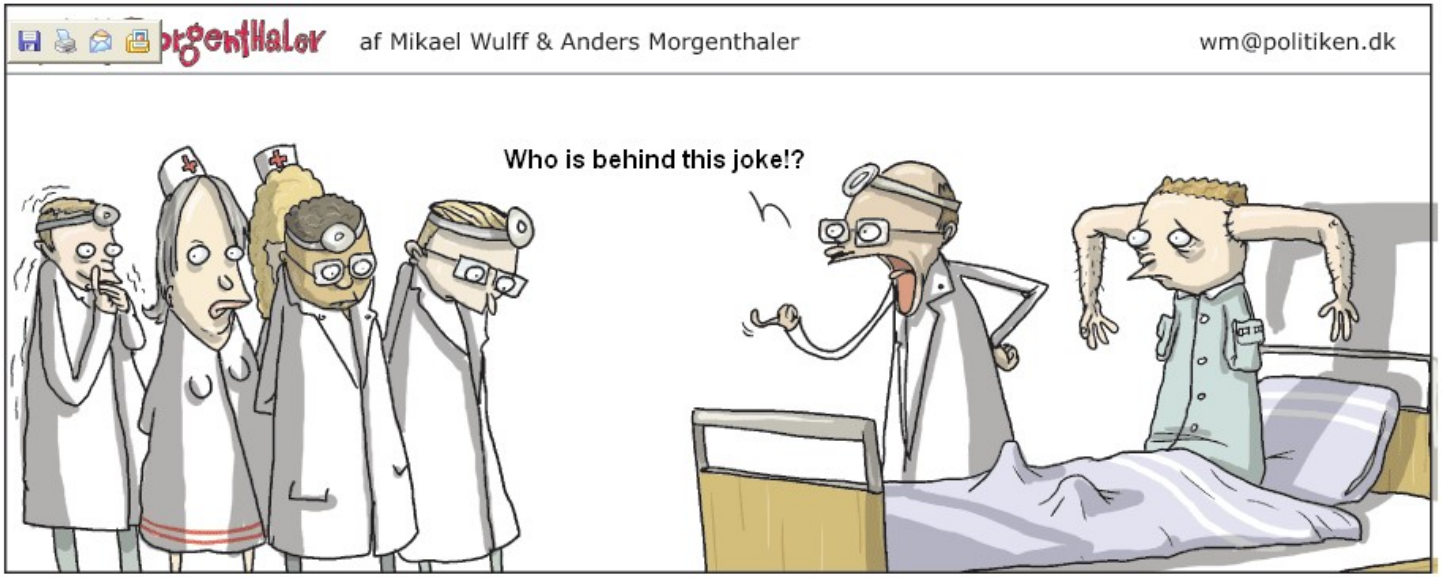
≡1.00

B:



0.91

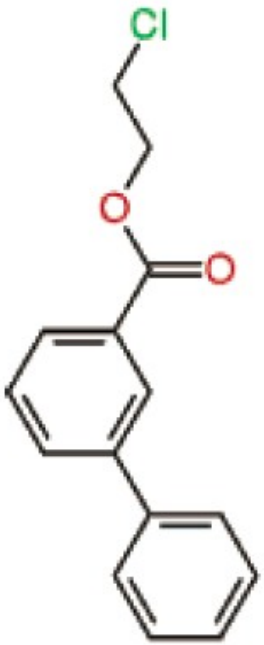
Does not encode information on the global topology



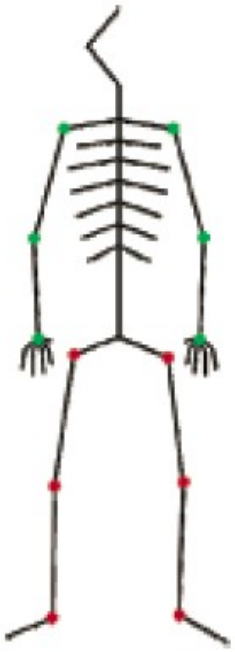
S*BIO™

Topology

How the chemist think of a molecule
Consider the explicit bonding and atom types.
Usually distances are measured in bonds between atoms.
Each atom encodes a property e.g. partial charge or polarizability.



Bonds

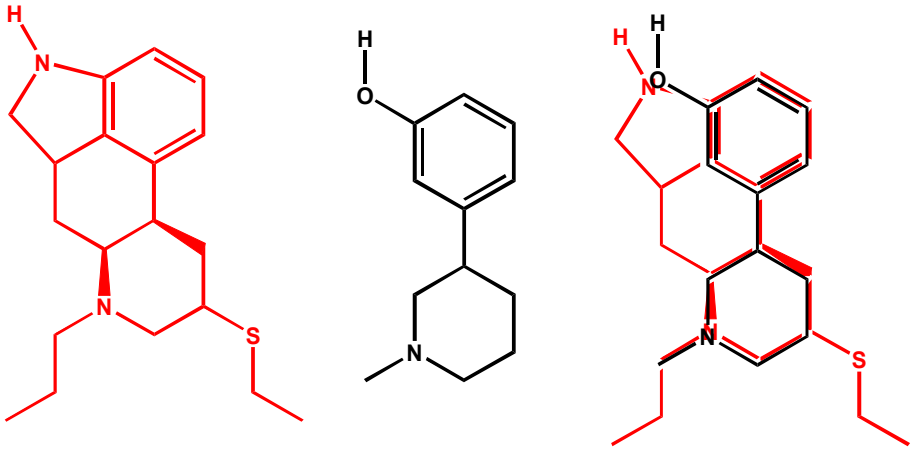


Joints

Traditional substructure searching (>10 million compounds)

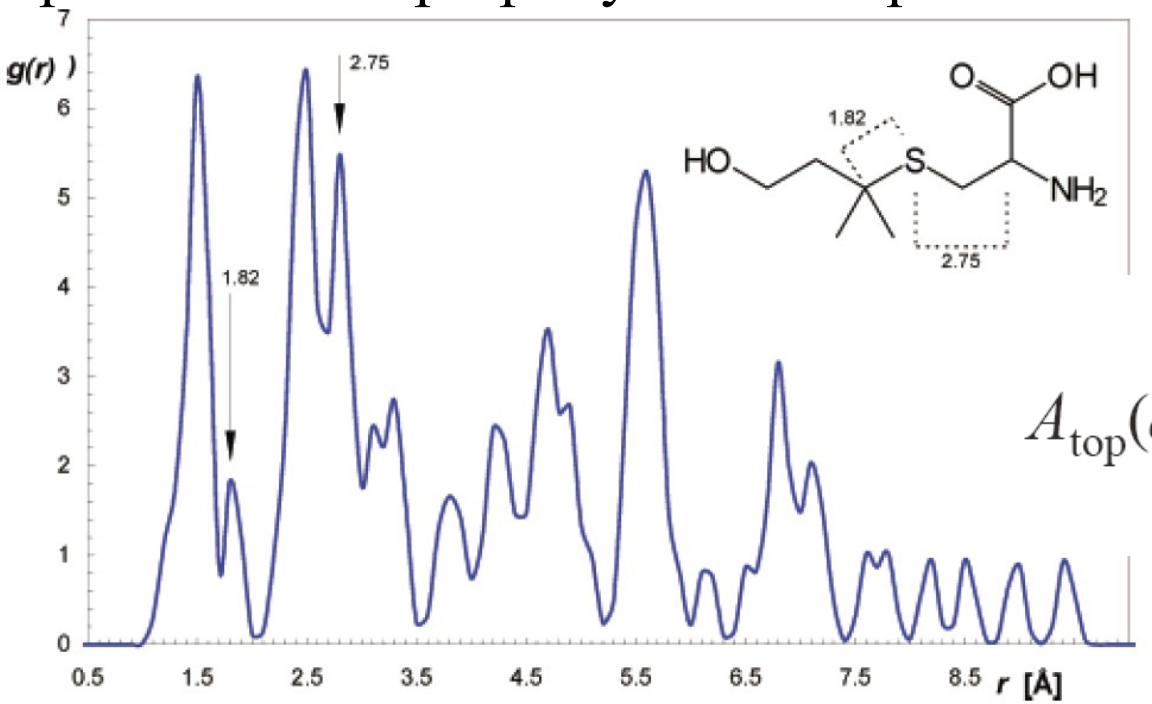
S*BIO™

Autocorrelation



Construct a graph from each atom pair distance. The Y-axis will be the sum over distance and the product of the property the atom pairs encode.

Excellent for searching large databases (>1 million compounds).



$$A_{\text{top}}(d) = \sum_{j=i+1}^{N} \sum^{N-1} a_i a_j \, \delta(d - d_{ij})$$

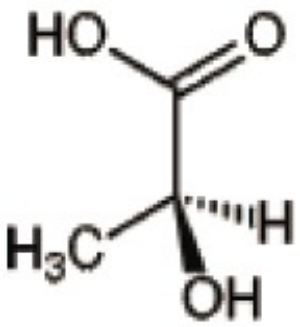
S*BIO™

Chirality

All proteins are chiral, so protein ligands usually have different affinities depending on their chirality.

Distance-dependent descriptors cannot encode chirality.

Chirality is a 3D property.

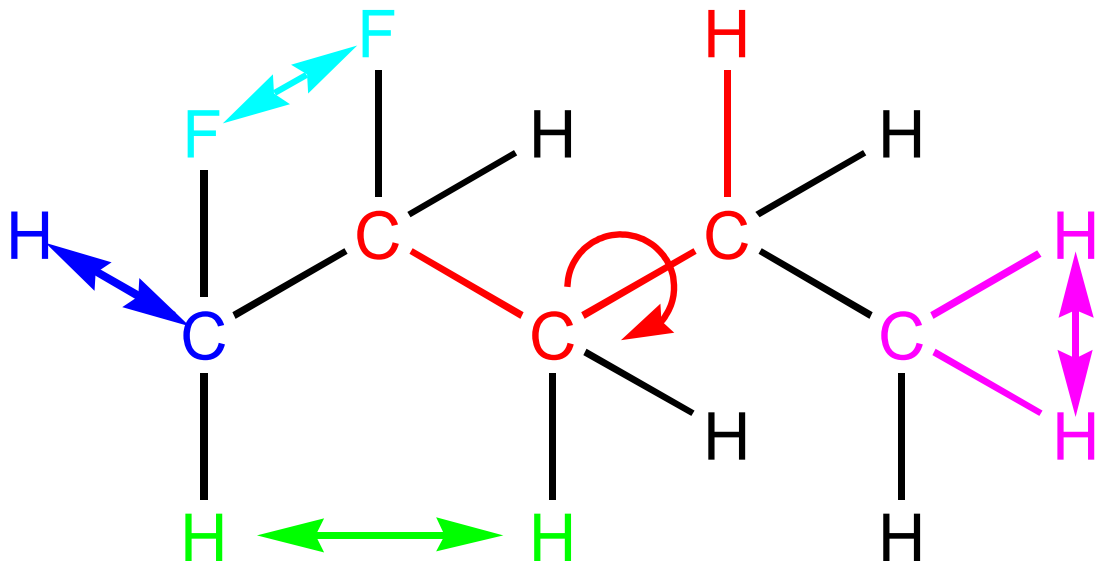


S*BIO™

3D

More than 30 million compounds are known.
Only 250,000 compounds (<1%) are crystalised.

X-ray bond distances, angles and torsion angles provides information for the construction of force-fields in which molecules can be represented as equations.



$$V_r = \sum_{1}^{Nbonds} \frac{k_r}{2}(r_i - r_0)^2$$

$$V_\theta = \sum_{1}^{Nangles} \frac{k_\theta}{2}(\theta_i - \theta_0)^2$$

$$V_\phi = \sum_{1}^{Ntorsions} \frac{1}{2} V_1 (1 + \cos\phi_i)$$

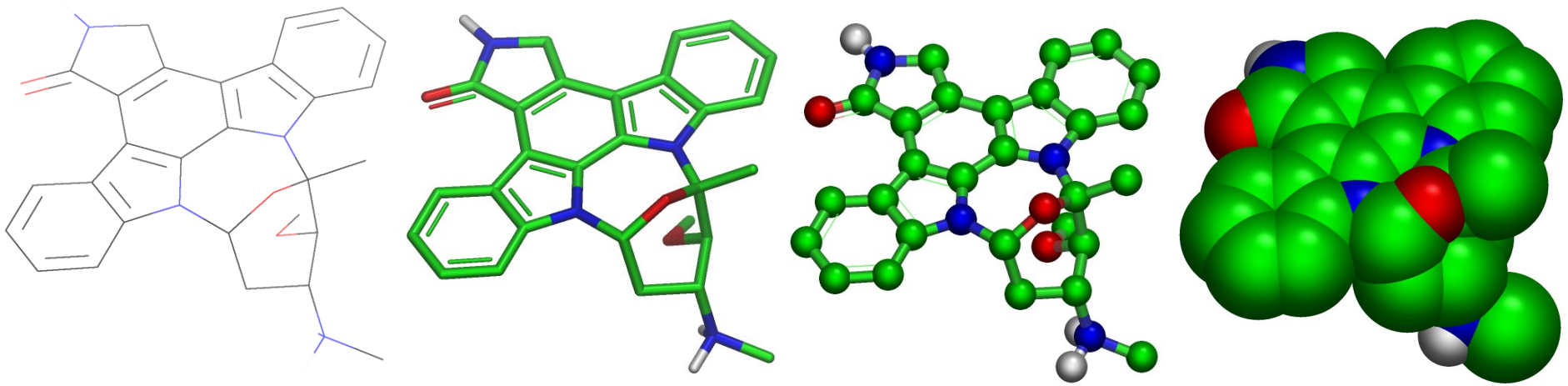
$$V_{vdw} = \sum_{1}^{NBpairs} \left(\frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^{6}}\right)$$

$$V_{el} = \sum_{1}^{all\ pairs} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$

S*BIO™

Docking

Structure-based methods like docking
Excellent for searching databases up to ~100,000 compounds).

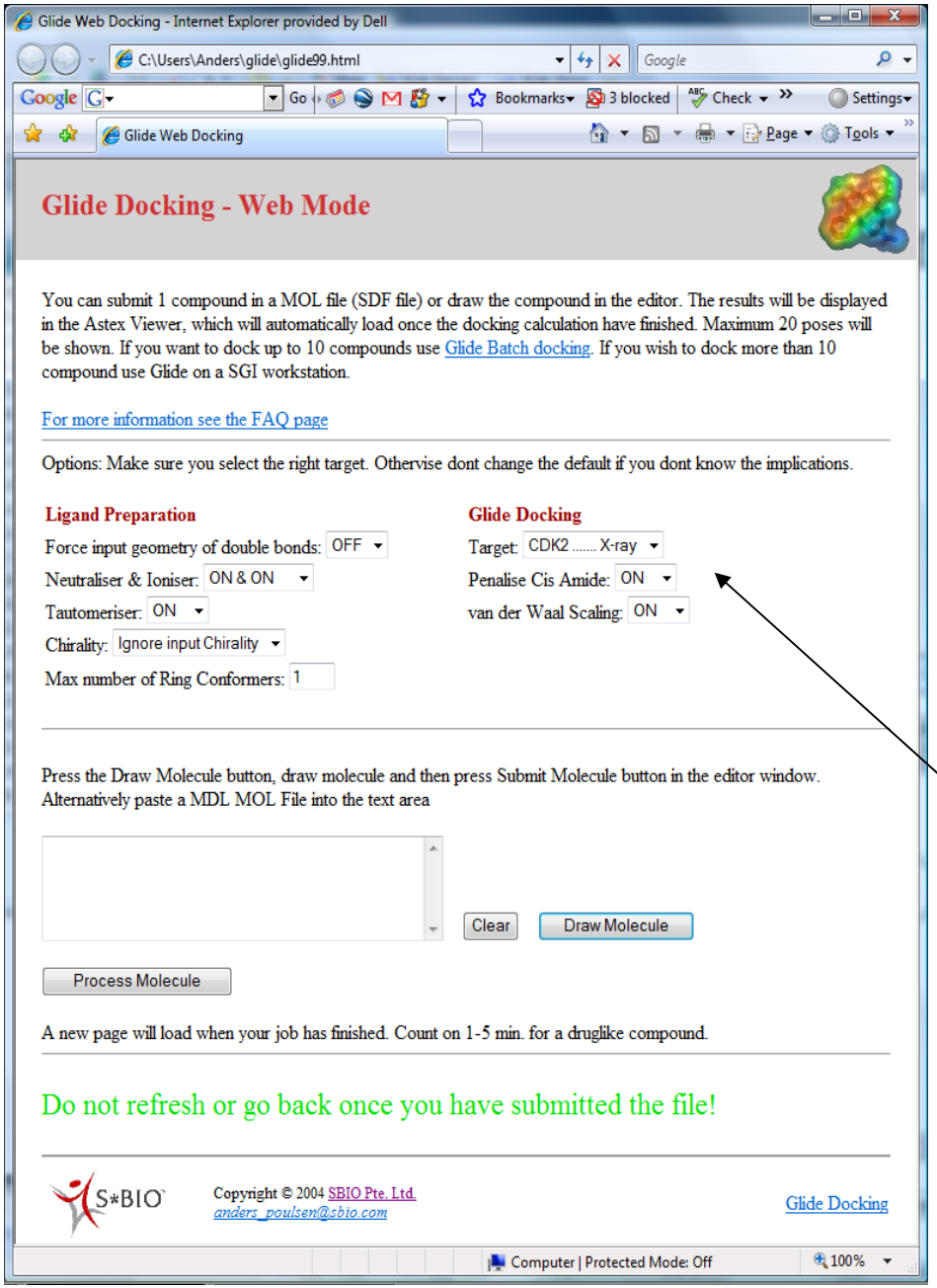


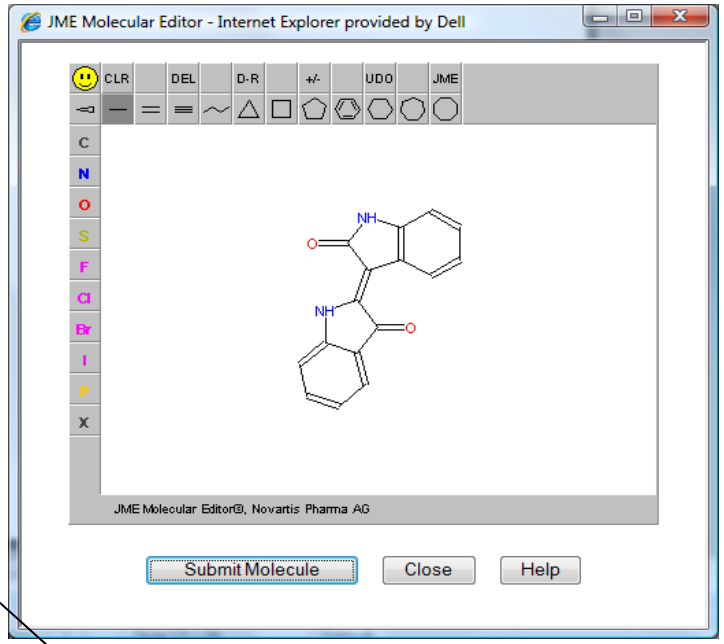
Computer graphics visualisations of forcefield koordinates

Link to Docking

S*BIO™

Docking

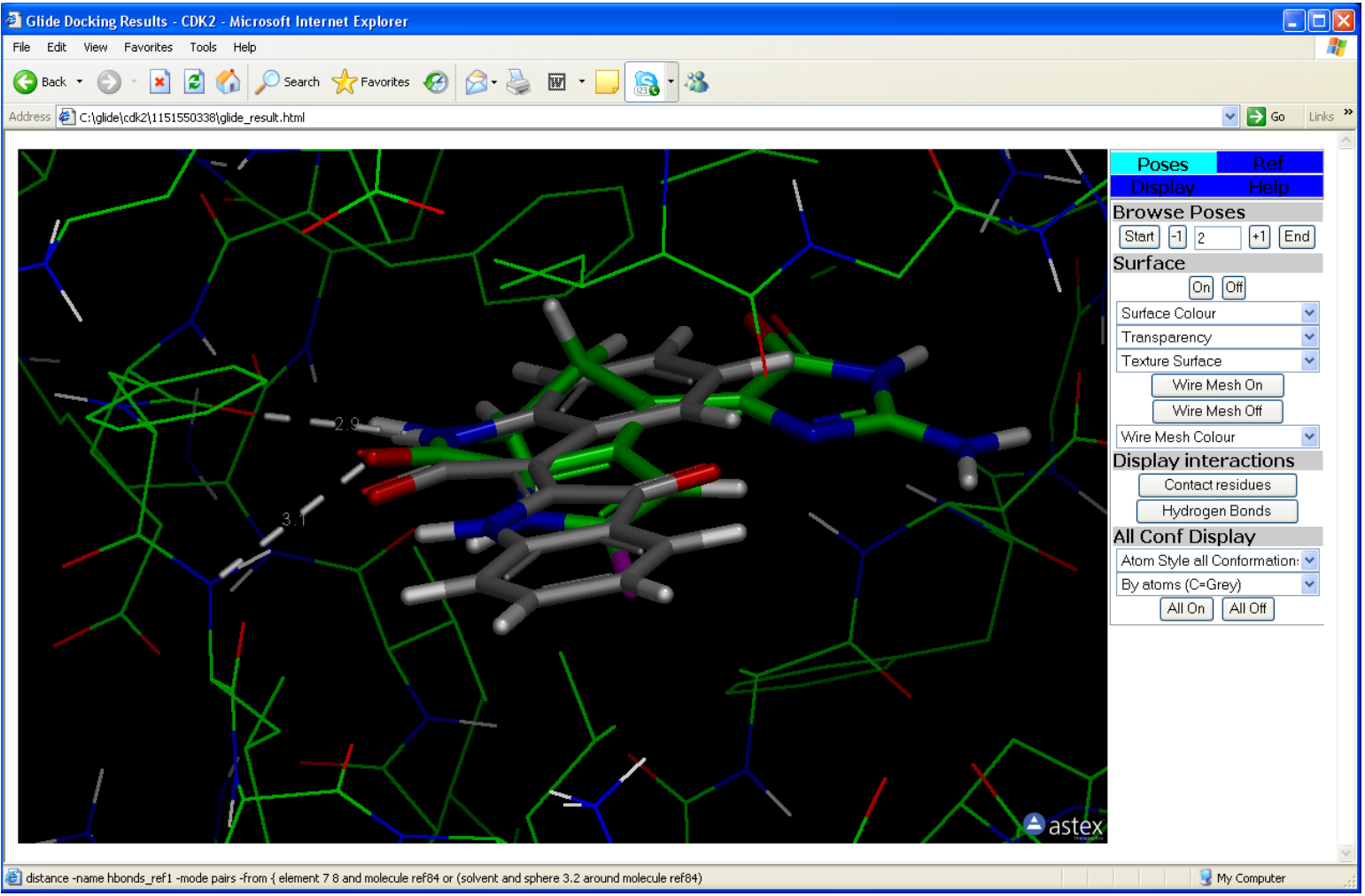




Draw Molecule & Select Target

S*BIO™

Docking



- Highest ranked pose not experimental bindingmode.
- Use available ligand data to evaluate Docking
- Select compounds by visual inspection, Dont rely on docking score alone

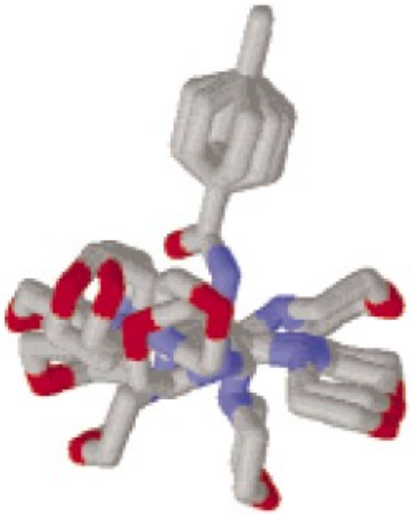
S*BIO™

Flexibility (4D)

Single bonds are rotatable, so most molecules have more than one conformation (shape).

In most 3D methods finding the bioactive conformation is essential and this remains one of the main challenges of medicinal chemistry today

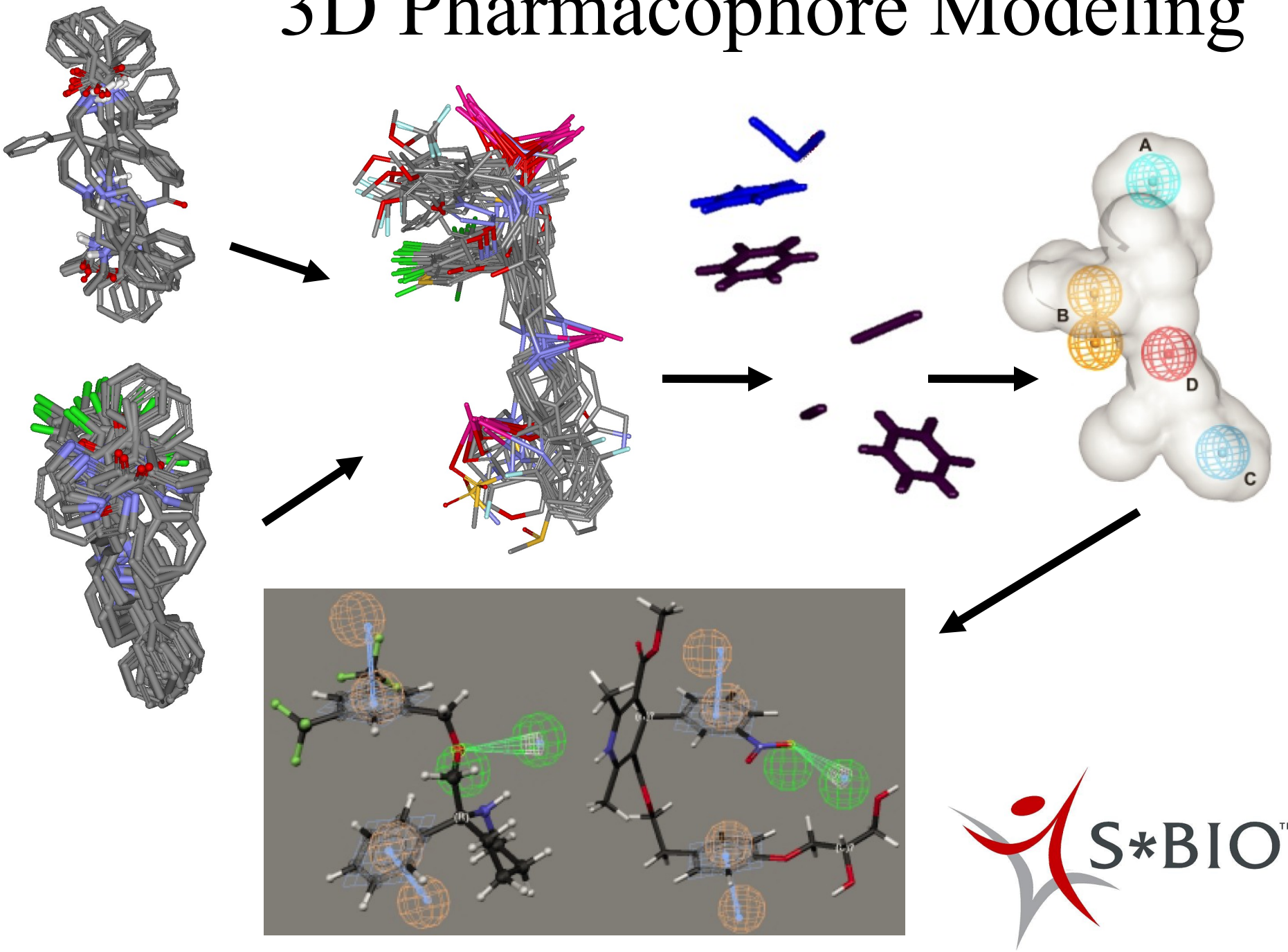
$$E = \sum_{1}^{N\text{angles}} \frac{k_\theta}{2}(\theta_i - \theta_0)^2 + \sum_{1}^{NB\text{pairs}} \left(\frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^{6}}\right) + \sum_{1}^{\text{all pairs}} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} + \sum_{1}^{N\text{bonds}} \frac{k_r}{2}(r_i - r_0)^2 + \sum_{1}^{N\text{torsions}} \frac{1}{2} V_1(1 + \cos\phi_i)$$



Find the minima of the forcefield equation

S*BIO™

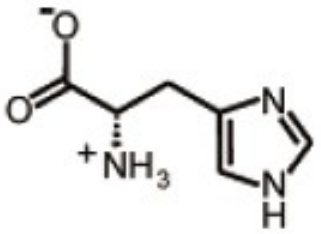
3D Pharmacophore Modeling

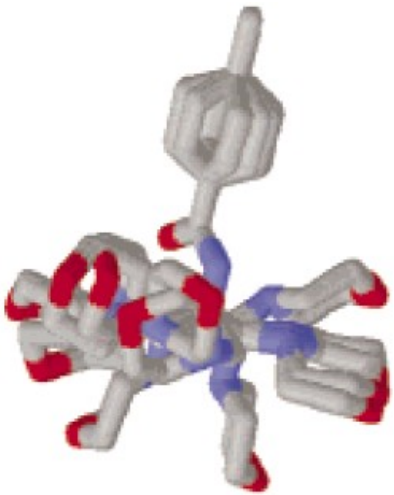


S*BIO™

Conclusion

...0100100110001101010011...





Chemo-Informatics

1D

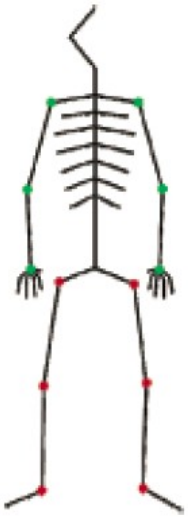
2D

Molecular Modeling

3D

4D

Increased complexity
Increased computational expense



S*BIO™