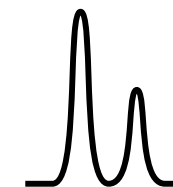


Taxonomic distance

A score, inversely proportional to the taxonomic distance between the biological source of the annotated compound and the biological source of the candidate structures.

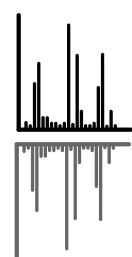


Physico-chemical consistency

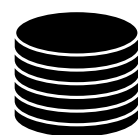
A score, proportional to the closeness of physico-chemical properties measured or calculated between the annotated compound and the candidate structures. Typically: retention time, retention order, collisional cross section.

$$S_{\text{tot}} = w_1 S_1 + w_2 S_2 + w_3 S_3 + w_4 S_4$$

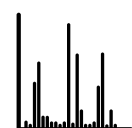
Spectral similarity score



exp
e.g. GNPS or ISDB-DNP approach
exp. or in-silico



Fingerprint similarity score



exp
e.g. Sirius approach
Structural DB

1000111010101010
1100101011101110

Structural consistency

When annotating spectrally organised data (i.e a molecular network), a structural consistency is expected between nodes of a cluster. A structural similarity distance is calculated among all annotations of the cluster. If a validated annotation is present within the cluster, remaining annotations are reranked so that the distance between this validated annotation and the remaining annotations is kept minimal. If no validated annotations are present, the annotations are reranked in order to obtain maximal structural consistency at rank 1.

