Subject: Display Full Similarity Matrix
Posted by caius on Thu, 07 Jan 2016 18:20:31 GMT
View Forum Message <> Reply to Message

Is it possible to display the full similarity matrix, either as a separate document or within the program for export? The manual states that the matrix is calculated completely anyway when conducting similarity analysis, but the "structure pair document" only displays information on compounds with similarity > 0.8

Subject: Re: Display Full Similarity Matrix
Posted by thomas on Sun, 10 Jan 2016 22:45:32 GMT
View Forum Message <> Reply to Message

No, currently there is no way to write the entire similarity matrix into a file, although it is created internally for various purposes. It never seemed obvious to me that there would be a need. If you are able to program in Java, you may download the source code and write a small program to calculate the matrix. This would only need a few lines of code. Let me know, if I can be of help with that.

Thomas

Subject: Re: Display Full Similarity Matrix
Posted by caius on Mon, 11 Jan 2016 21:01:19 GMT
View Forum Message <> Reply to Message

I was able to approximate matrix output using macros and the calculated column generation, but the results from chemsim(Flexophore_of_Structure, idcode) don't match the Structure Similarity values available in the similarity chart or the Structure[Structure List, Similarity] filter. For instance, use of chemsim does not yield a self similarity score of 1. Is there any way to get the outputs of "Structure Similarity" or "Structure (is similar to Flexophore)" into the table?

Subject: Re: Display Full Similarity Matrix
Posted by thomas on Fri, 15 Jan 2016 21:45:21 GMT
View Forum Message <> Reply to Message

The Flexophore is the only descriptor that give self-similarity values slightly below 1.0,
because under the hood the Flexophore is generated from a fresh set of representative conformers,
which every time look a little different. Nevertheless, similarity values from two different Flexophore descriptors should usually be above 0.99 if both were generated from the same molecule.
Structure filters, chemsim() function or graphical view axis all use the same flexophore similarity calculation, but have the same issue that there is a slight variation in similarity values, if the

flexophore is calculated multiple times.