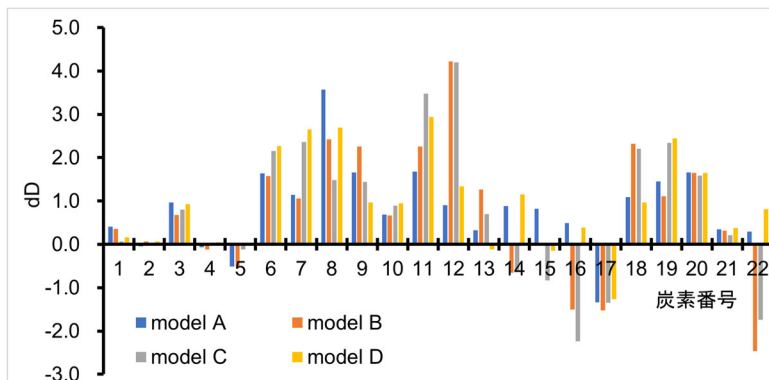


3-3. 計算結果の評価について

計算結果を客観的にどのように評価するかが問題になります。その方法について同じデータを用いて柵瀬押ししてみました。

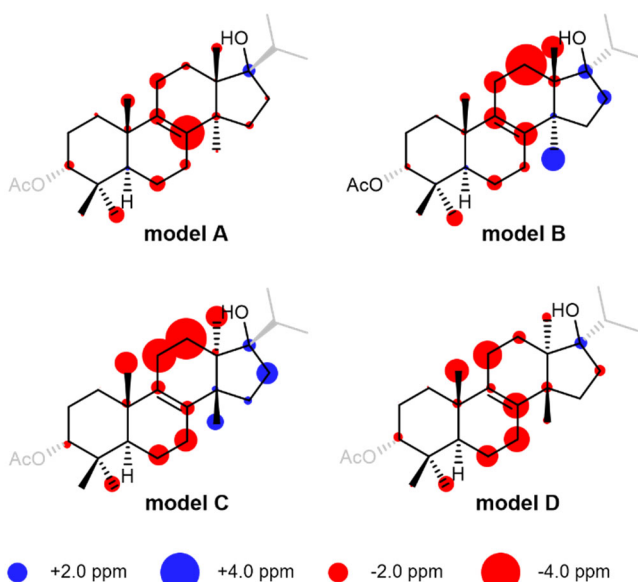
① 化学シフト差をグラフで示す。

作成は最も簡単です。ただ、判断が難しい場合が多いと思います。



② マッピングする。

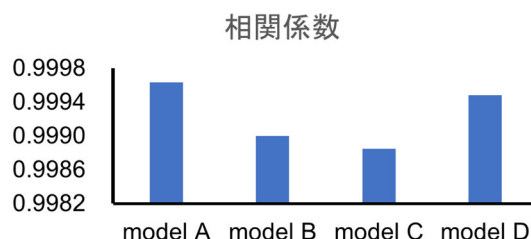
原理的には①と同じですが、手間はかかりますが、視覚的に判断が容易になります。また、差異の大きい部分が集中する場合が多く、多くの場合、その周辺の構造がおかしいと推定でき、より深い議論が可能になります。



③ 相関係数を求める。

実験値と、どれだけよく符合するかを求めるもので、エクセルでは CORREL 関数を使用します。単位無しです。化学シフト計算の精度は高いため、異性体を含めどれも 1.0 に近い数値となってしまいます。表では差異の判断はしづらいですが、グラフにすると多少見やすくなります。

model	相関係数
model A	0.99963
model B	0.99900
model C	0.99885
model D	0.99948



④ 最大誤差 (MAX deviation) を求める。

単位は ppm です。単独で用いられることは無く、他の評価と合わせて使用されます。数値が小さい方が一致度は高くなりますが、最小の値を示す、異性体が正しい構造とは限らないと思います。エクセルでは、一度実験値との差異の一覧を作成したのち、MAX 関数を用います。

model	Max dev (ppm)
model A	3.57
model B	4.22
model C	4.20
model D	2.94

⑤ 差異の絶対平均 (MAE) を求める。

差異の絶対値の平均を求めます。単位は ppm です。計算は簡単ですが、後述の RMSD と大きく結果が異なることはまれです。エクセルでは、一度実験値との差異の絶対値一覧を作成したのち、AVERAGE 関数を用います。

model	MAE (ppm)
model A	1.00
model B	1.32
model C	1.40
model D	1.11

⑥ 差異の偏差二乗平均平方根 (RMSD) を求める。

統計的に一致度を求める最も一般的な方法です。単位は ppm です。エクセルでは、一度実験値との差異の絶対値一覧を作成したのち、SQRT 関数と AVERAGE 関数を用います。

model	MAE (ppm)
model A	1.26
model B	1.66
model C	1.78
model D	1.45

⑦ DP4 期待値を求める。

2010 年に Goodman らが発表した異性体間における期待値の比を百分率で求める方法で、最近元も使用される評価法です。

詳細は後述しますが候補異性体間での期待値の比較ですので、**候補構造に真の構造が含まれていなくても、その中の一つの異性体 100% としてしまう場合もあります**。従って、RMSD 値、や Max deviation 値を併記する方が適切です。

model	DP4
model A	79.2%
model B	1.8%
model C	0.6%
model D	18.5%

なお、オリジナル論文の計算条件では MMFF で配座解析後、最安定配座から 10 kJ/mol 以内の配座について、その構造を維持したまま B3LYP/6-31G(d,p) で化学シフトを計算、ボルツマン分布で化学シフトを平均化して求めると決められています。従って、Spartan のプロトコルで計算した場合、正確には DP4 では無いこととなります。

現在の Spartan での標準偏差 1.77ppm を代入すると、正しい異性体(model A)の期待値は高くなります。ただ、論文等で標準偏差の

model	DP4 (Goodman のパラメータ)	DP4 (Spartan のパラメータ)
model A	79.2%	88.9%
model B	1.8%	0.4%
model C	0.6%	0.1%
model D	18.5%	10.6%

根拠を説明することは大変です。また Goodman らのパラメータを使用した場合、正しい異性体のスコアは低くなる傾向にあるため、結果として判断基準としてはより厳しくなり、問題は無いはずですが。