

CP3 の計算について

CP3 の計算には、異性体セットについての実測値と計算値の両方の化学シフトが必要です。

CP3 の Java アプレットは Goodman グループのホームページから入手可能です (<https://www-jmg.ch.cam.ac.uk/tools/nmr/CP3.html>)。このアプレットでは実測及び計算化学シフトを入力するのみで CP3 スコアとベイズ確率理論に基づくそれぞれの異性体の確率を計算します。しかし、Java アプレットは Windows10 以降のデフォルト設定ではサポートされておらず、少々面倒なセッティングが必要です。

(Java を起動する場合以下を参考にしてください。 https://www.java.com/ja/download/help/win10_faq_ja.html)

オリジナルの CP3 の Java アプレットにはシグナルの帰属という概念はなく、アプレットにある “Assign resonances” というボタンを押すと、実測値、計算値とも化学シフトは大きい順にそれぞれ並べ替えられます。ですが、十分に帰属した場合はこの機能はむしろ余計です。しかし、どうも “Assign resonances” を押さない、即ちこの機能が無効にしても自動的に化学シフトは並び替えられる仕様になっているようです。

CP3 Java アプレットで CP3 確率を計算した方は経験されたと思いますが、非常に高い頻度で一方の異性体の確率を 100% として示します。ほとんど化学シフト差がない場合でも、100% を与えてしまう場合があります。CP3 スコアを確率に変換する際、Supporting Information に記載されていた説明内容で、パラメータの一つ解釈できないものがありました。このパラメータが CP3 確率を極端に変動させるすることが判りました。

CP3 は極めて敏感に異性体を区別します。異性体間で化学シフトの差が著しく小さい場合は、実験値計算値とも化学シフトを小数点第 2 位まで考慮すると良いかもしれません。また、プロキラルなメチレンプロトンの様に、帰属に曖昧さが残る場合は計算から除いた方が賢明かもしれません。

以下、CP3の詳細です。

まずCP3スコアを求めます。

① CP3スコアを求める手順（異性体AとBの比較）

Δ_{exp}^i 及び Δ_{calc}^i を以下に定義

$$\Delta_{exp}^i = \delta_{exp\ A}^i - \delta_{exp\ B}^i \quad (i \text{ 番目の実測化学シフトの差})$$

$$\Delta_{calc}^i = \delta_{calc\ a}^i - \delta_{calc\ b}^i \quad (i \text{ 番目の計算化学シフトの差})$$

$f_3^i(\Delta_{exp}^i, \Delta_{calc}^i)$ 及び $f_3^i(\Delta_{exp}^i, -\Delta_{calc}^i)$ を以下に定義

$$\Delta_{calc}^i / \Delta_{exp}^i > 1 \text{ の場合: } f_3^i(\Delta_{exp}^i, \Delta_{calc}^i) = (\Delta_{exp}^i)^3 / \Delta_{calc}^i$$

$$\text{それ以外の場合: } f_3^i(\Delta_{exp}^i, \Delta_{calc}^i) = \Delta_{exp}^i \Delta_{calc}^i$$

$$-\Delta_{calc}^i / \Delta_{exp}^i > 1 \text{ の場合: } f_3^i(\Delta_{exp}^i, -\Delta_{calc}^i) = (\Delta_{exp}^i)^3 / (-\Delta_{calc}^i)$$

$$\text{それ以外の場合: } f_3^i(\Delta_{exp}^i, -\Delta_{calc}^i) = \Delta_{exp}^i (-\Delta_{calc}^i)$$

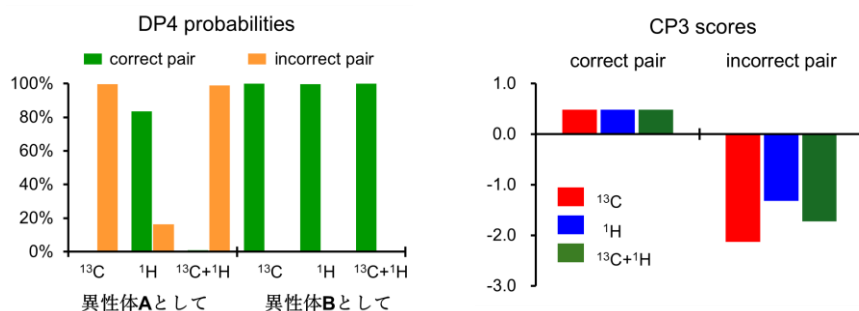
分母となる Δ_{exp}^i 或いは Δ_{calc}^i が 0 の場合、エラーを避ける目的でそれらに 0.0001 を代入する様になっています。

正しい組み合わせ、及び正しくない組み合わせのCP3スコア（それぞれ $CP3_{correct}$, $CP3_{incorrect}$ ）は、以下の式で定義されています。実際は炭素核、水素核について別個に求めます。 $^{13}C+^1H$ CP3スコアは、 ^{13}C CP3スコアと 1H CP3スコアの平均と定義されています。

$$CP3_{correct} = \sum_i f_3^i(\Delta_{exp}^i, \Delta_{calc}^i) / \sum_i (\Delta_{exp}^i)^2$$

$$CP3_{incorrect} = \sum_i f_3^i(\Delta_{exp}^i, -\Delta_{calc}^i) / \sum_i (\Delta_{exp}^i)^2$$

これらの値をグラフ化すると以下のようになります。原理的にCP3スコアの最大値は1です。最小値は特定されません。CP3スコアが大きい時、「もっともらしい」ことを示し、負の値が大きくなるほど「疑わしい」ことを示します。このグラフで、正しいペアとして帰属するには十分な情報となります。ここまでは、「化学シフトの順に並べ替える」という作業が無いことを除き、オリジナルのCP3と全く同じです。以下の例では、DP4での予想はいずれの異性体を計算しても同じ異性体を高い確率で支持してしまうことを意味しています。同じデータを使用したCP3解析では正しい組み合わせを明確に支持しています。



② CP3 スコアを確率へ変換

ベイズ推定理論に基づき、別途 Goodman らが求めた、正しい組み合わせ、及び正しくない組み合わせの CP3 期待値、及び標準偏差を利用して確率に変換します。期待値及び標準偏差は確率に変換する際の敏感さを決定するパラメータになります。Goodman のオリジナルでは、以下の式で二行目の定義の解釈が出来ませんでした。これが Goodman の CP3 Java アプレットで求めた場合、確率が異常に敏感になる原因であることも判りました（実はメールで問い合わせましたが、返答がありませんでした。）この変換は確率計算であることから、一つのパラメータを特殊化して重みをかけるするものではなく、公平に計算されるはずと考え、以下の対照的な定義と解釈し直しました。Goodman のアプレットでは頻繁に確率 100%を与えますが、下の式を用いた場合は 100%になることはまずありません。Goodman のアプレットで求めたものとは異なった値を与えますが、結論が逆転することはありません。それは CP3 スコアで既に結論が出ているからです。DP4 でも同様ですが、CP3 確率が多少変化しても議論に大きく影響することはありません。

以下は、エクセルでの計算を前提に NORM.DIST 関数を用いて示しました。

$$\begin{aligned}P(R_1|AC_1) &= 1 - \text{NORM.DIST}(\text{CP3}_{\text{correct}}, \text{expectation}_{\text{correct}}, \sigma_{\text{correct}}, \text{TRUE}) \\P(R_2|AC_1) &= 1 - \text{NORM.DIST}(\text{CP3}_{\text{incorrect}}, \text{expectation}_{\text{incorrect}}, \sigma_{\text{incorrect}}, \text{TRUE}) \\P(R_1|AC_2) &= 1 - \text{NORM.DIST}(\text{CP3}_{\text{correct}}, \text{expectation}_{\text{incorrect}}, \sigma_{\text{incorrect}}, \text{TRUE}) \\P(R_2|AC_2) &= 1 - \text{NORM.DIST}(\text{CP3}_{\text{incorrect}}, \text{expectation}_{\text{correct}}, \sigma_{\text{correct}}, \text{TRUE})\end{aligned}$$

	correct pair	incorrect pair
¹³ C expectation	0.547	-0.487
¹³ C standard deviation (σ)	0.253	0.533
¹ H expectation	0.478	-0.786
¹ H standard deviation (σ)	0.305	0.835
¹³ C+ ¹ H expectation	0.512	-0.637
¹³ C+ ¹ H standard deviation (σ)	0.209	0.499

その後、以下の式により確率に変換します。

$$\text{CP3 probability for correct pair (\%)} = \frac{P(R_1|AC_1)P(R_2|AC_1)}{P(R_1|AC_1)P(R_2|AC_1) + P(R_1|AC_2)P(R_2|AC_2)} \times 100$$

$$\text{CP3 probability for incorrect pair (\%)} = \frac{P(R_1|AC_2)P(R_2|AC_2)}{P(R_1|AC_1)P(R_2|AC_1) + P(R_1|AC_2)P(R_2|AC_2)} \times 100$$

エクセルの雛型を作成しました。

<https://view.officeapps.live.com/op/view.aspx?src=http%3A%2F%2Fmolecules.a.la9.jp%2Fpersonal%2FCP3excel.xlsx&wdOrigin=BROWSELINK>

1. Smith, S. G.; Goodman, J. M., *J. Org. Chem.* **2009**, *74*, 4597-4607.