

sievgeneNMR (DIRECTION 法) の改良

詳細設計書

2018 年 2 月 5 日

Copyright (C) 2006-2018 Next Generation Natural Product Chemistry (N²PC)

目次

1. 概要	3
2. 入力ファイルと計算式	4
3. ユーザインタフェース	6
4. 修正内容	9
4. 1. 修正内容一覧	9
4. 1. 1. NMR 種別の識別情報追加	9
4. 1. 2. NMR 種別での変数領域変更	11
4. 1. 3. タンパク側 NMR ファイル入力追加	12
4. 1. 4. タンパク側 NMR 計算処理追加	14
4. 1. 5. NMR 種別による呼び出し手続きの変更	17

1. 概要

NMR チームにおいて、開発された新しい蛋白質—リガンド複合体計測の手法を取り込む。この手法は、式は sievgenenMR (DIRECTION 法) と酷似している。異なるのは sievgenenMR (DIRECTION 法) では、蛋白質側をラジオ波照射し、リガンドのH原子のシグナル情報を得る、このリガンドのH原子のシグナル情報を再現するようにリガンドを動かす。これに対し、新しい手法では、リガンドの特定のH原子をラジオ波照射し、蛋白質のH原子のシグナル情報を得る、この蛋白質のH原子のシグナル情報を再現するようにリガンドを動かす。式はほぼ同じあるが、読み込む情報に違いがある。

sievgenenMR (DIRECTION 法) では、蛋白質側の全てのH原子（ないし特定種類の残基のH原子）のスピンを励起し、化合物側の全ての H 原子に磁気ダイポール相互作用で移行したスピンの緩和を計測データとする。蛋白質側のH原子から化合物側のH原子へのスピンの移行は $1/R^6$ で起こり、個々の化合物側Hについて $\Sigma 1/R^6$ （和は、蛋白質側の該当する全てのH原子についてとる）となる。そして、化合物側の全てのHのスピンのシグナルの実験値と、ドッキングにより得られた複合体構造から得られる理論値 ($\Sigma 1/R^6$) とを比較し、実験値と計算値の相関が高くなるように、化合物座標を選択し、動かす。

新しい手法では、励起させるHのスピンのと、観測するHのスピンの逆になる。化合物側の特定種類の（たとえばベンゼン環のH）H原子のスピンを励起し、蛋白質側の全てのH原子に磁気ダイポール相互作用で移行したスピンの緩和を計測データとする。化合物側のH原子から蛋白質側のH原子へのスピンの移行は $1/R^6$ で起こり、個々の蛋白質側Hについて $\Sigma 1/R^6$ （和は、化合物側の該当する全てのH原子についてとる）となる。そして、蛋白質側の全てのHのスピンのシグナルの実験値と、ドッキングにより得られた複合体構造から得られる理論値 ($\Sigma 1/R^6$) とを比較し、実験値と計算値の相関が高くなるように、「化合物座標」を選択し、動かす。

2. 入力ファイルと計算式

入力ファイルは、sievgeneNMR (DIRECTION 法)と同じくスピン緩和係数ファイルと、スピン緩和実験値ファイルの二種類がある。

個々の水素のスピン緩和の観測値はシグナルと呼ばれ、シグナルとスピン緩和の相関係数は下記の式で求められる。 S_b^i は、実験値のシグナルで、i-th の H 原子のシグナル、 s_a^i は理論値 ($\Sigma 1/R^6$) で、i-th の H 原子のシグナル、Nc は NMR シグナルを観測する H の総数とすると、相関係数 R は：

$$R_a^b = \frac{\sum_i (s_b^i - \frac{i}{Nc})(s_a^i - \frac{i}{Nc})}{\sqrt{\sum_i (s_b^i - \frac{i}{Nc})^2 \cdot \sum_i (s_a^i - \frac{i}{Nc})^2}} \dots (1)$$

$$Bi = s_b^i - \frac{i}{Nc} \dots (2) \quad Ai = s_a^i - \frac{i}{Nc} \dots (3)$$

$$s_a^i = \sum_j^{NP} \frac{1}{R_{ij}^6} \quad i \text{ は蛋白質の H, } j \text{ は化合物の H で、NP は該当する化合物の H の総数}$$

とおくと、(4)の式で表現される。

$$R_a^b = \frac{\sum_i Bi \cdot Ai}{\sqrt{\sum_i Bi^2 \cdot \sum_i Ai^2}} \dots (4)$$

実験的にスピン励起できるリガンドの H には複数の種類があるので、N 種類の実験を行った場合、各実験の信頼性や重要度が同じとして、それぞれの実験データと計算値の相関係数を R_x とし、これらの平均値を、全体の相関係数 R_{total} とする。

$$R_{total} = (R_{\text{ベンゼン環のH}} + R_{\text{脂肪族H}} + R_{\text{アミンH}} + \dots) / N$$

この R_{total} を最大化する力 F_{total} は、各相関係数の距離微分による力 F_x の平均値となる。

$$F_{total} = (F_{\text{ベンゼン環のH}} + F_{\text{脂肪族H}} + F_{\text{アミンH}} + \dots) / N$$

ここでは、各項 R_x , F_x について、計算式を導出することにする。

MD を行うため、リガンドの H 原子に働く力を求めるため、この x 、 y 、 z の微分を求める。

$$\begin{aligned} \frac{\partial R}{\partial x_j} &= \frac{1}{\sum_i B_i^2 \sum_i A_i^2} \left(\sqrt{\sum_i B_i^2 \sum_i A_i^2} \frac{\partial}{\partial x_j} \sum_i B_i A_i - \sum_i B_i A_i \frac{\partial}{\partial x_j} \sqrt{\sum_i B_i^2 \sum_i A_i^2} \right) \\ &= \frac{1}{\sum_i B_i^2 \sum_i A_i^2} \left(\sqrt{\sum_i B_i^2 \sum_i A_i^2} \sum_i B_i \left\{ \frac{\partial}{\partial x_j} A_i \right\} + \sum_i B_i A_i \sqrt{\sum_i B_i^2} \frac{1}{\sqrt{\sum_i A_i^2}} \sum_i A_i \left\{ \frac{\partial}{\partial x_j} A_i \right\} \right) \end{aligned} \quad (5)$$

{ } の部分は :

$$\begin{aligned} \frac{\partial}{\partial x_j} A_i &= \frac{\partial}{\partial x_j} \sum_i \frac{1}{R_{ij}^6} = -6 \frac{1}{R_{ij}^7} \frac{\partial}{\partial x_j} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \\ &= -6 \frac{1}{R_{ij}^7} \frac{(x_j - x_i)}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}} \end{aligned} \quad (6)$$

$s_a^i = \sum_j^{NP} \frac{1}{R_{ij}^6}$ は、リガンドの H が、蛋白質の H の重なって R=0 となり値が発散す

るのを押さえるため、van der Waals ポテンシャルの形

$$\begin{aligned} s_a^i &= \sum_j^{NP} E_{ij} \\ E_{ij} &= \begin{cases} -4 \left(\frac{A^{12}}{R_{ij}^{12}} - \frac{A^6}{R_{ij}^6} \right) & : \text{when } R \geq 2^{1/6} A \\ +1 & : \text{when } R < 2^{1/6} A \end{cases} \end{aligned}$$

で置き換える。係数 A の値は、デフォルトは、半径 A=1. 0 Å として、制御ファイルで変更できるようにするのが良い。

蛋白質の H では、CH₃, CH₂, NH₃, NH₂, フェニル基 (Phe, Tyr) の H (2 ペア x 2 個) 等が等価な水素原子である。これら等価な H のシグナル強度は同じである。

なお、本システムについて、使用法とテスト計算結果をプロジェクト研究員に説明すること。

3. ユーザインタフェース

ユーザインタフェースは既存の sievGeneNMR と共通化する方針とする。

従来のリガンド側シグナルを使用する方法と、タンパク側シグナルを使用する方法の識別には INPUT フェーズの新規オプションの NMRTYP で決定し、その他のパラメータは NMRTYP の設定値にしたがってタンパク側、リガンド側の種別を決定する。

NMRTYP のデフォルト値は、リガンド側シグナルを使用する NMRTYP = LIGAnd とする。

項番	フェーズ	項目	キーワード	値	内容
1	INPUT	NMR 計測指定	NMRTYPe	選択型	NMR 計測種別指定 (LIGAnd PROTein)
2		スピン緩和係数 ファイル(*1)	NMRCOE	選択型	スピン緩和係数ファイル 名(NORE READ)
3			NAMENM	文字列	スピン緩和係数ファイル 名(“”)
4		スピン緩和実験 値ファイル(*1)	NMREXP	選択型	スピン緩和実験値ファイル 名(NORE READ)
5			NAMENO	文字列	スピン緩和実験値ファイル 名(“”)
6		タンパク残基係 数ファイル	NMRPRO	選択型	タンパク残基係数ファイル 読み込み(NORE READ)
7			NAMENP	文字列	タンパク残基係数ファイル 名(“”)
8			原子間距離閾値	NMRRAD	実数型
9	DOCK	スコア係数	WETNMR	実数型	NMR のスコア係数(1.0)
10	MIN	スピン緩和計算 指定	USENMR	選択型	スピン緩和計算指定 (NO YES)

(*1)H 原子係数、NMR 実験値は NMRTYP=LIGA, PROT の指定で、それぞれリガンド側、タンパク側の情報として入力する。

(2-1) スピン緩和係数ファイル

化合物に含まれる水素でシグナルを計算する対象の水素原子を 1 行ごとに、原子 ID、原子名、力係数の順に空白区切りで記述する。

空行、または 1 カラム目が';'である行はコメント行とみなす。

力係数は前述の (5) 式の係数で、相関係数のみ計算し、力を加えない場合は 0.0 を指定する。

```
; NMR signal coefficient
;
2 H2 1.00000
11 H11 1.00000
12 H12 1.00000
```

(2-2) スピン緩和実験値ファイル

(2-2-1) リガンド側シグナル使用時のスピン緩和実験値ファイル

NMR 実験で得られたスピン緩和実験値を原子 ID、原子名、シグナル計算係数の順に空白区切りで記述する。

空行、または 1 カラム目が';'である行はコメント行とみなす。

```
;
; NMR test suits (laco, use excel sheet calculation)
;
2 H 0.012895381
11 H 0.009223408
12 H 0.004169454
```

(2-2-2) タンパク側シグナル使用時のスピン緩和実験値ファイル

リガンド側の対象水素とタンパクの各残基に対するスピン緩和実験値を指定する。残基への係数を残基 ID、残基名、実験値シグナルの順に空白区切りで記述する。

空行、または1カラム目が';'である行はコメント行とみなす。

“LIGAND”で開始されるブロックには化合物側の水素原子 ID を記述する。

“PROTEIN”で開始されるブロックには、1行に残基 ID、残基名、実験値シグナルを記述する。

```
; residue coefficient
; level 1 : H39 H41
LIGAND
39 41

PROTEIN
2 ALA 0.020645607
11 ALA 0.283182474
14 ASN 0.025803068
15 VAL 0.363536242

; level 2 : H40
LIGAND
40

PROTEIN
2 ALA 0.015803041
11 ALA 0.198817769
14 ASN 0.018111482
15 VAL 0.226836835
```

(2-3) タンパク残基係数ファイル

タンパクの各残基に対するスピン緩和計算の係数を指定する。残基への係数を残基 ID、残基名、係数の順に空白区切りで記述する。

空行、または1カラム目が';'である行はコメント行とみなす。

```
; residue coefficient
1 PRO 0.2
2 GLN 0.1
3 VAL 0.2
```

スピンを励起できるリガンド側の原子種は、ベンゼン環のHや、脂肪族のH, アミンのHなど複数種類ある。これらのHを別々に励起し、タンパク質側の緩和時間の変化を測定できるため、入力は、複数の原子種をそれぞれ励起したときのタンパク質のシグナルを入力できるように、複数のグループからなるようにする。

リガンド側のどの原子を励起した場合、タンパク質側のどのアミノ酸残基の緩和時間が変化するかをグループごとに入力できるようにする。

4. 修正内容

4. 1. 修正内容一覧

NMR 機能の実装のため、sievgeneMD に下記の修正を行う。

#	修正内容	対象ファイル	区分
1	NMR 種別の識別情報追加	Sievgene_Class. f90	修正
2		SievEnum_Type. f90	修正
3		SievInput_Option. f90	修正
4	NMR 種別での変数領域変更	Interact_Class. f90	修正
5	タンパク側 NMR ファイル入力追加	SievIOMethod. f90	修正
6	タンパク側 NMR 計算処理追加	Optional_Method. f90	修正
7	NMR 種別による呼び出し手続きの変更	Interact_Control. f90	修正
8		sievgene. f90	修正
9		Score_Method. f90	修正

4. 1. 1. NMR 種別の識別情報追加

タンパク側シグナルを利用する方法を識別する内部情報として、NMR の種別を示すフラグを追加する。

(1) テーブルの追加

Sievgene_Class. f90 に Main_t に NMR 種別を示す NMR_Type を追加する。

```
integer::NMRType          ! NMR type
                          ! ( NMR_LIGAND | NMR_PROTEIN)
```

(2) 定数定義の追加

SievEnum_Type. f90 に下記の NMR 種別で使用する定数を定義する。

```
!
! NMR type
!
integer, parameter::NMR_LIGAND = 1
integer, parameter::NMR_PROTEIN = 2
character(14), dimension(NMR_PROTEIN), parameter::
    NMR_TYPE = (' LIGAND-SIGNAL ',
                ' PROTEIN-SIGNAL' /) &
```

(3) 制御ファイルのオプション入力機能追加

SievInput_Option. f90 に下記の "NMR_TYP = LIGA | PROT" の選択肢を追加する。また、Input_InputOption 手続きに、オプション値の初期化手続き (Init_Value) と設定手続き

(Set_Value)の呼び出しを追加する。

```
Select_t('NMRTYP', 2, &! NMR type (LIGAnd | PROTein)
          ('LIGA', 'PROT', ' ', ' ', ' ', ' ', ' '), &
          ', ', ', ', ', ', ', ', ' /)) /)
:
call Init_Value(sopt, SelectOption, 'NMRTYP', 'LIGA')
:
call Set_Value(sopt, SelectOption, 'NMRTYP', main%NMRTyp)
```

4. 1. 2. NMR 種別での変数領域変更

NMR 計算に対し、リガンド側シグナルを利用する方法とタンパク側シグナルで、使用する変数領域を変更する。また、各手続きで読み込むデータをそれぞれの領域に設定する。

(1) テーブルの追加

Interact_Class.f90 の NMR 計算領域をリガンド側 (NMR%Ligand) とタンパク側 (NMR%Protein) の二種に変更する。既存の処理で NMR 用の変数をアクセスしていた部分は、全てリガンド側領域をアクセスするように変更する。

タンパク側シグナルを使用する場合、リガンドの H の種別毎にタンパク残基の係数を設定する配列 (LigandCoeff と ResidCoeff) を用意する。この配列のサイズで使用する MAX_HYDTYPE は 30 とする。

```
type NMRLigand_t
  integer::HydNum                ! number of ligand hydrogens
  integer, dimension(MAX_LIGAND)::HydList  ! ligand hydrogen ID list
  integer, dimension(MAX_LIGAND)::ID2List  ! hydrogen ID to list
  real*4, dimension(MAX_LIGAND)::Coefficient ! NMR coefficient
  real*4, dimension(MAX_LIGAND)::CalcValue ! calculate NMR value
  real*4, dimension(MAX_LIGAND)::ExprValue ! experiment NMR value
end type NMRLigand_t

type NMRProtein_t
  integer::HydNum                ! number of protein hydrogens
  integer, dimension(MAX_ATOM)::HydList  ! ligand protein ID list
  integer, dimension(MAX_ATOM)::ID2List  ! hydrogen ID to list
  real*4, dimension(MAX_ATOM)::Coefficient ! NMR coefficient
  real*4, dimension(MAX_ATOM)::CalcValue ! calculate NMR value
  real*4, dimension(MAX_ATOM)::ExprValue ! experiment NMR value

  integer::HydTypeNum            ! ligand hydrogen type
  real*4, dimension(MAX_LIGAND, MAX_HYDTYPE)::LigandCoeff ! ligand hydrogen
  real*4, dimension(MAX_RESID, MAX_HYDTYPE)::ResidCoeff ! protein residue
end type NMRProtein_t

type NMR_t
  type (NMRLigand_t)::Ligand
  type (NMRProtein_t)::Protein
  real*4, dimension(3, MAX_LIGAND)::Grad    ! NMR gradient
  real*4::ExprAverage                       ! average of exprValue
  real*4::Radius                            ! vdW radius
  integer::NMRType
end type NMR_t
#endif
```

4. 1. 3. タンパク側 NMR ファイル入力追加

(1) リガンド用 NMR ファイル入力手続き名の変更

下記の3種のリガンド用 NMR ファイル入力手続き名を変更する。

#	修正内容	対象ファイル	区分
1	スピン緩和係数、スピン緩和 実験値ファイル入力	Read_NMRSignal	Read_LigandNMRSignal
2	タンパク残基係数入力	Read_NMRData	Read_ProteinNMRData

(2) タンパク用スピン緩和係数ファイル、スピン緩和実験値ファイル入力の追加 呼び出し形式：

```
subroutine Read_ProteinNMRSignal(input, nmr, atomInfo)
```

引数：

```
type(SievInput_t), intent(in)::input          ! 入力ファイルクラス  
type(NMR_t), intent(out)::nmr                ! NMR 情報  
type(AtomInfo_t), intent(in)::atomInfo       ! タンパク原子情報
```

戻り値：

なし

機能：

下記のファイルを読み込み、情報を設定する。

(2-1) 入力：input%NMRCoefficient(スピン緩和係数ファイル入力用構造体)

設定内容：

```
タンパク側水素原子リスト：nmr%Protein%HydList  
タンパク側水素原子数      ：nmr%Protein%HydNum  
タンパク側原子の水素原子リストへのポインタ：nmr%Protein%Id2List  
タンパク側水素の力の係数：nmr%Protein%Coefficient
```

(2-2) 入力：input%NMRExpriment(スピン緩和実験値ファイル入力用構造体)

設定内容：

```
スピン緩和実験値：nmr%Protein%ExprValue
```

(3) タンパク用スピン緩和係数ファイル、スピン緩和実験値ファイル入力の追加

呼び出し形式：

```
subroutine Read_MultiProteinNMRData(  
    input, residName, residNum nmr, atomName, ligandAtomNum)
```

引数：

```
type(SievInput_t), intent(in)::input          ! 入力ファイルクラス  
character(3), dimension(MAX_RESID), intent(in)::residName ! タンパク残基名  
integer, intent(in)::residNum                ! タンパク残基数  
type(NMR_t), intent(out)::nmr               ! NMR 情報  
character*2, dimension(MAX_LIGAND), intent(in)::atomName ! リガンド原子名  
integer, intent(in)::ligandAtomNum          ! リガンド原子数
```

戻り値：

なし

機能：

リガンド側の水素情報を設定する。

- ・リガンド側水素数：nmr%Ligand%HydNum
- ・リガンド側水素リスト：nmr%Ligand%HydList
- ・リガンド側原子の水素原子リストへのポインタ：nmr%Ligand%ID2List

また、下記のファイルを読み込み、情報を設定する。ファイルを読み込まない場合はタンパク側残基の係数を 1.0 に設定する。

(2-1) 入力：input%NMRProtein(タンパク残基係数ファイル入力用構造体)

設定内容：

タンパク側残基係数：nmr%Protein%ResidCoeff(i, j)

※i, j は i 番目残基の j 番目の水素原子タイプを示す。

4. 1. 4. タンパク側 NMR 計算処理追加

NMR 計算処理 (Optional_Method.f90) に下記の手続きを追加する。

(1) タンパク側水素座標登録

呼び出し形式：

```
subroutine Set_ProteinCoordinate(obj, nmr, cord)
```

引数：

```
type(Object_t), intent(in)::obj      ! オブジェクト(原子、残基)情報
```

```
type(NMR_t), intent(in)::nmr        ! NMR 情報
```

```
real*4, dimension(MAX_ATOM, 3)::cord ! タンパク原子座標
```

戻り値：

なし

機能：

MINIMIZE フェーズでの atomInfo はリガンド側原子情報に更新されているため、タンパクの原子座標と残基 ID を保存する。

(2) タンパク側 NMR シグナル計算

呼び出し形式：

```
subroutine Calc_ProteinNMREffect(nmr, ligandCord, score)
```

引数：

```
type(NMR_t), intent(inout)::nmr      ! NMR 情報
```

```
real*8, dimension(3, MAX_ATOM), intent(inout)::ligandCord ! リガンド原子座標
```

```
real*4, intent(out)::score           ! NMR スコア
```

戻り値：

なし

機能：

蛋白質水素原子とリガンド側水素原子のスピンの緩和の計算値を求め、実験値との差分から得られた結果を用いて、リガンド側水素原子への力と NMR スコアを計算する。(詳細は「2. 入力ファイルと計算式」を参照)

(3) タンパク側の等価な水素の検出とスピン緩和実験値の補正

手続き名 :

```
subroutine Set_EquivalentProteinHydrogen(bondInfo, atomInfo, residInfo, nmr)
```

引数 :

```
type(BondInfo_t), intent(in)::bondInfo ! タンパク結合情報  
type(AtomInfo_t), intent(in)::atomInfo ! タンパク原子情報  
type(ResidInfo_t), intent(in)::residInfo ! タンパク残基情報  
type(NMR_t), intent(inout)::nmr ! NMR 情報
```

戻り値 :

なし

機能 :

タンパクの等価な水素原子を検出し、等価な水素同士でのスピン緩和実験値の値を平均値に変更する。

(4) タンパク側 NMR 計算値表示

手続き名 :

```
subroutine Display_ProteinNMRCalcValue(nmr, unit)
```

引数 :

```
type(NMR_t), intent(in)::nmr ! NMR 情報  
integer, intent(in)::unit ! 出力装置番号
```

戻り値 :

なし

機能 :

タンパクの等価な水素同士でのスピン緩和計算値を平均値に変更した値を表示する。

(5) タンパク側残基の対象性判定

手続き名 :

```
subroutine symmetry_Resid(numatom, mat, numbond, nbond, residInfo, no_identity)
```

引数 :

```
integer::numatom           ! 原子数  
integer::numbond           ! 結合数  
integer, dimension(MAX_ATOM, 2)::nbond ! 結合原子ペア  
integer, dimension(MAX_ATOM)::mat     ! 原子タイプ  
type(ResidInfo_t), intent(in)::residInfo ! 残基情報  
integer, dimension(MAX_ATOM)::no_identity ! 同一原子の添字
```

戻り値 :

なし

機能 :

水素原子に対し同一残基内にある水素原子との対称性を求め、同一のトポロジーの場合に相手原子の添字を設定する。

#identity(3) = 2 の場合、3 番目原子は 2 番目原子と対称であることを示す。この場合、identity(2) = 2 となる。

4. 1. 5. NMR 種別による呼び出し手続きの変更

NMR 計算種別に従い、タンパク側 NMR 処理またはリガンド側 NMR 処理を呼び出す。

(1) Interact_Control.f90

NMR 種別が NMR_LIGAND の場合 : Calc_NMREffect

NMR 種別が NMR_PROTEIN の場合 : Calc_ProteinNMREffect

(2) Score_Method.f90

NMR 種別が NMR_LIGAND の場合 : Calc_NMREffect, Display_NMRCalcValue

NMR 種別が NMR_PROTEIN の場合 : Calc_ProteinNMREffect, Display_ProteinNMRCalcValue

(3) sievgene.f90

NMR 種別が NMR_LIGAND の場合 :

Read_ProteinNMRData, Set_ProteinHydrogen, Read_LigandNMRSignal,
Set_EquivalentHydrogenParameter, Set_ProteinHydrogen

NMR 種別が NMR_PROTEIN の場合 :

Read_ProteinNMRSignal, Set_EquivalentProteinHydrogen,
Set_ProteinCoordinate, Read_MultiProteinNMRData

以上