

タンパク質の中性子解析における精密化手順

茶竹 俊行, 森本 幸生
(京都大学 原子炉実験所)

Protocol of Refinement in Neutron Crystallographic Analysis

Toshiyuki Chatake, Yukio Morimoto
(Research Reactor Institute, Kyoto University)

ABSTRACT

The protocol of refinement in neutron protein crystallographic analysis is described in this article. During the refinement, there are several features specialized for neutron experiment: (1) the mixed chemical structure of deuterium and hydrogen atoms, (2) solvent molecules around the protein, and the deuterium exchange in the polypeptide chain. We demonstrate the example of neutron refinement steps, using rubredoxin mutant, of which neutron diffraction collected using BIX-3 diffractometer at JRR-3 in Japan Atomic Energy Agency. The X-ray refinement program of CNS was used for the present refinement. Scattering lengths, stereochemical parameters, and command input files were modified for neutron experiment. These files would be the templates for the neutron protein crystallographic analysis.

Keywords: *Neutron crystallography, protein, refinement*

1. はじめに

中性子単結晶構造解析は蛋白質の水素原子を観測できるほぼ唯一の実験手法であり、生体内での酵素反応、遺伝情報の発現、機能制御を明らかにする上で非常に有用な情報を得ることが出来る [1-3]。また、水素原子を含んだ詳細な立体構造は、薬剤設計などの計算科学の初期モデルとして非常に期待されている [4-5]。現在、生体分子の中性子結晶解析は、JAEA の BIX-3 [6], BIX-4 [7], ILL の LADI [8], LANSCE の PCS [9] で進められており、年に数件の構造解析が報告されている。さらに、大強度陽子加速器計画 (J-PARC) で建設中の茨城県生命物質構造解析装置 (BIX-P1 型) により従来の数倍の解析が可能になると期待されている。

ところで、中性子結晶解析では、X 線解析用プログラムを中性子用に改変して構造精密化が行われているが、共通化した手順が報告されていないために、解析者が個々に手順を組む必要がある [10]。これは新規の研究者が中性子解析を始める上での大きな障壁となることが危惧される。今回の報告では、共通化手順を提案することを目的として、ルブレドキシン変異体の中性子解析を例にとり、生体分子の中性子結晶解析における構造精密化の具体的な手順を紹介する。

2. 解析の準備

2.1 システム要件

解析には以下のプログラムを使用する。

- (a) 精密化プログラム CNS [11]
- (b) モデリングプログラム MI-fit (Rigaku Inc.)

(c) スクリプト言語 Perl

精密化プログラム CNS は解析用の簡易言語を使用しているため、中性子解析用の改変が容易であるので、今回採用した。また、モデリングプログラムも、中性子解析で取り扱う軽水素原子と重水素原子の混合構造の設定が容易な MI-fit を採用した。これらは全てアカデミックフリーであり、Linux や Microsoft Windows 上で動作する。今回は、WindowsXP 上で CNS 1.1, MI-fit 2.8.0 と Perl 用に Cygwin をインストールして解析を行った。また Fedora core Linux 6 でも動作を確認している。

2.2. 中性子解析用のファイル

今回の精密化では、以下のファイルを準備した。

- (a) MI-fit 用の中性子用構造化学情報ファイル
 - (b) CNS 用中性子散乱長ファイル
 - (c) CNS 用中性子用構造化学情報ファイル
 - (d) CNS 用 コマンドファイル
 - (e) 中性子処理用 Perl スクリプト
- (a) は MI-fit のメニューから、【Dictionary】⇒【Load new dictionary】を行うことにより設定できる。(b)~(e) はディレクトリ内に階層構造で展開され、CNS による精密化の際に適宜読み出される。

2.3. 解析に必要な実験データ

解析に必要な実験データとして、

- (a) 中性子強度データ
 - (b) 格子定数
 - (c) 解析用の初期構造 (非水素構造)
- を準備する必要がある。(a) 中性子データは JAEA

BIX-3, BIX-4 を使用した場合には, 強度積算プログラム DENZO, SCALEPACK [12]から出力されたファイルを CNS 用にファイル変換して使用する. 今回は, BIX-3 で測定した 1.6 Å 分解能の中性子データをを用いる [13]. (b) 格子定数は中性子実験でも決定できるが, 波長分散や精度の問題があるので, 予め X 線により高精度で決定しておくことが望ましい. (c) 中性子解析では, 予め X 線結晶解析により水素原子以外の立体構造を決定して,これを初期構造として水素原子をモデルに組み込んでいく. この X 線実験の際には出来るだけ中性子実験を行った結晶と同じものを用いることが望ましい. 放射光実験で頻繁に行われる極低温 (~100 K) での測定では結晶格子が縮小してしまうため,これを初期構造に用いて精密化を行うと,収束に時間がかかる傾向がある. 中性子実験と同じ温度で X 線構造を使用することが望ましい. また, 中性子解析では重水中で作製した結晶 (完全重水素化の場合もある) を用いるが,最近の筆者の研究では, 重水中から得た蛋白質結晶は, 同一条件で軽水から作製した結晶とほぼ同一の立体構造をとることが明らかになった [14]. よって, 軽水結晶を用いて初期構造を求めても問題は無い.

3. 精密化の流れ

中性子解析では, X 線解析で作成した水素の無いモデルに水素原子を加えてその原子パラメーターを最適化することが, 構造決定の中心を占めている. 水素原子をモデルに加えていく順序を Fig. 1 に示す. ① 最初に, 軽水素 (H) 原子をモデルに追加する. 蛋白質中の軽水素は炭素と結合しており, 立体化学的に予測できる. CNS 上で自動発生させて追加する. ② 次に, 蛋白質中の重水素 (D) 原子の付加を行う. 蛋白質の中性子結晶解析では, 非干渉性散乱を抑えるために, 結晶を重水中で作成する. このため, 蛋白質のアミノ基や水酸基の水素原子は D 原子に置換される. これらの多くは位置の予測が難しいので, MI-fit 上で中性子フーリエ図を観察しながら, モデル上の正の中性子散乱密度が観測された所に D 原子の追加を行っていく. 水酸基やアミノ基の水素原子は構造生物学的に重要なものが多いので, 慎重に追加を行う. ③ この後, 溶媒領域にある重水 (D₂O) 分子を D 原子と同様に, MI-fit 上で確認しながら取り込んでいく. ④ 最終段階で, 主鎖のアミノ基の H/D 交換率を計算する. 分子内部のアミノ基の一部は H 原子のままであるため, 中性子フーリエ図では H 原子の負の散乱長に対応して, 負の核密度が観測される. そこで H 原子, D 原子の存在確率を最適化する.

これを実際の手順に当てはめたフローチャートを Fig. 2 に示す. 精密化では, Fig. 2 の各ステップを CNS, MI-fit, Perl を用いて行う.

4. 実際の解析処理

4.1. プログラムの実行方法

今回の精密化では, WindowsXP の Cygwin 用のユーザーディレクトリの下に解析用のディレクトリ

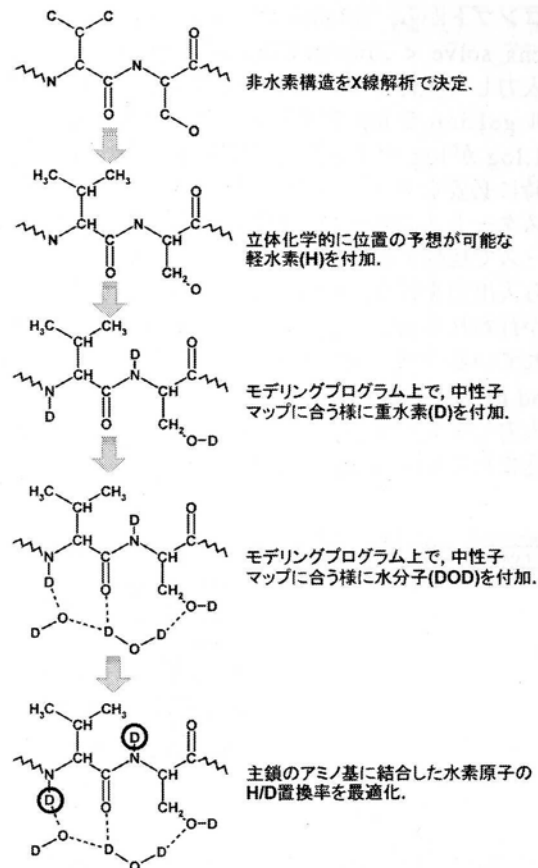


Fig.1 The scheme of building the hydrogen structure.

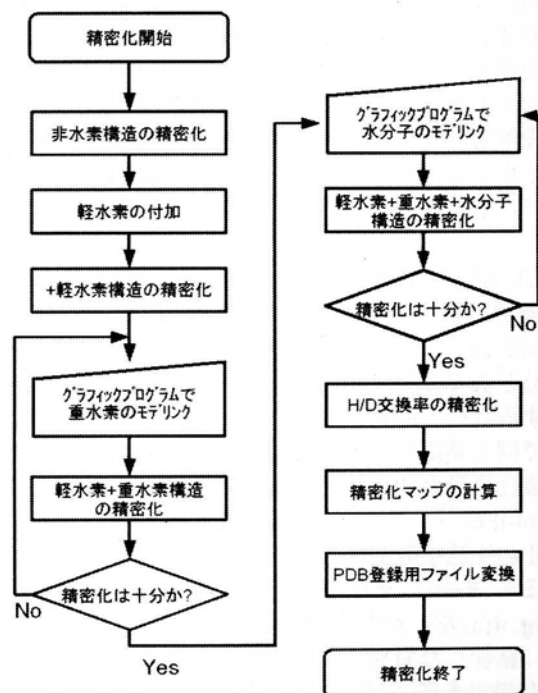


Fig.2 The flowchart of the refinement.

(rubmut) と六つのサブディレクトリを作成した (Fig. 3). CNS の実行は, WindowsXP のコマンド

プロンプトから, rubmut ディレクトリ上で
 cns_solve < ./inp/ge1.inp > ./log/ge1.log
 と入力して実行する. この場合には, コマンドファイル ge1.inp を inp ディレクトリから読み込まれて, ge1.log が log ディレクトリに書き込まれる. また, 同時に必要なファイルの入出力も行われる. MI-fit はスタートメニューから MI-fit を選択すると, GUI ベースで起動する. 必要なファイルは GUI メニューから入出力を行う. Perl スクリプトは, ファイル変換が行われる pdb ディレクトリに.pl の拡張子で保存されているので, pdb ディレクトリ内で
 hd.pl hd6.pdb hd6a.pdb
 と入力して実行する. この場合には, hd6.pdb が読み込まれて hd6a.pdb が出力される.

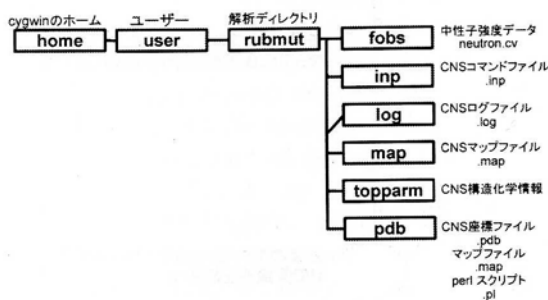


Fig.3 The directories for the refinement.

4.2. 精密化の実行手順

Fig. 2 のフローチャートに従って, Table 1 に示す実行手順を作成した. それぞれの実行内容に対する使用プログラム (及びコマンドファイル), 入出力ファイル, モデルに含まれている水素原子の種類, 各段階での R 値を表内に併記した. なお, 入出力ファイルについては, 重要なものだけを記述してある. 以下, 各手順について説明する.

4.2.1. 非水素構造の精密化 (Table 1 の 1-2)

解析の第一段階として, 非水素構造の精密化を行う (1-2). この精密化は特に, 中性子実験に用いた結晶と X 線実験の結晶に違いがある場合, それを補正するために重要である. 今回は, 中性子用と X 線用で同じ結晶を使用しているため, 精密化による構造変化は殆ど観察されない. Fig. 4(a)は精密化後の 2IFol-IFcl フーリエ図である. Trp36 のインドール環に正の中性子散乱密度 (Fig. 4(a)左) が観察されている. また, その周囲に, 負の中性子散乱密度 (Fig. 4(a)左) があるが, これらはインドール環の炭素に結合した H 原子の負の中性子散乱密度に対応するものである. さらに, インドール環の窒素原子 (黒丸) の上部に, 正の中性子散乱密度が突き出しているが, これは窒素原子に結合した D 原子の正の中性子散乱密度に対応している. これらは, 中性子データを用いて H, D 原子をモデルに導入できることを示している.

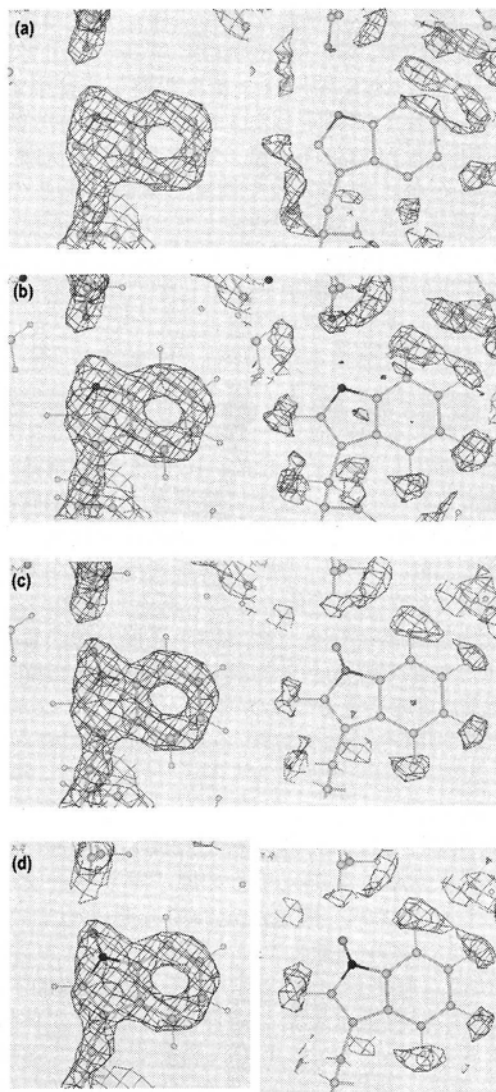


Fig.4 Neutron Fourier maps around Trp36 of (a) non-hydrogen, (b) hydrogen, (c) hydrogen and deuterium, and (d) final structures. The left and right pictures show positive and negative neutron scattering densities.

4.2.2. 軽水素構造の精密化 (3-4)

次に, H 原子をモデルに導入して精密化を行う. H 原子は CNS により自動発生させる. この結果は Fig. 4(b)であり, H 原子に対応した負の中性子散乱密度のマップが改善されていることが分かる. また, 窒素原子に結合した D 原子に対応する正の密度も前段階に続いて明瞭に観測することが出来る.

4.2.3. 軽重水素構造の精密化 (5-9)

D 原子の付加は, 中性子フーリエ図を確認しながら行う. まず, MI-fit で対応する正の中性子散乱密度が存在することを確認する. 長い側鎖を持つ Lys のアミノ基などは乱れにより観測できない場合があるので, これらをチェックしておく. また, 蛋白質内部のアミノ基は H 原子から D 原子に置き換わらず正の中性子散乱密度が観測できない時がある. これらもチェックする. Rubmut においては, Lys2, Lys6, Lys28, Lys50 の側鎖のアミノ基, Val4, Cys5,

処理内容	使用プログラム	入力ファイル	出力ファイル	水素原子	R 値
1 非水素構造の準備	CNS (ge1.inp)	xray.pdb (X 線構造)	ge1.pdb	—	
2 非水素構造の精密化	CNS (re1.inp)	ge1.pdb	re1.pdb	—	30.2
3 軽水素構造の準備	CNS (ge2.inp)	re1.pdb	ge2.pdb	H	
4 軽水素構造の精密化	CNS (re2.inp)	ge2.pdb	re2.pdb, m2.2fofc.map	↓	25.6
5 重水素原子のチェック	MI-fit	re2.pdb, m2.2fofc.map		↓	
6 重水素原の付加	CNS (ge3.inp)	re2.pdb	ge3.pdb	H,D	
7 重水素原子のモデリング	MI-fit	ge3.pdb, m2.2fofc.map	ge3_mifit.pdb	↓	
8	テキストエディタ	ge3.pdb, ge3_mifit.pdb	ge3.pdb	↓	
9 軽重水素構造の精密化	CNS (re3.inp)	ge3a.pdb	re3.pdb, m3.2fofc.map	↓	23.1
10 重水分子 (非水素) の追加	テキストエディタ	re3.pdb, xray.pdb	re3_plus_D2O.pdb	↓	
11 重水分子の準備 1	CNS (ge4.inp)	re3_plus_D2O.pdb	ge4.pdb	H,D, D ₂ O	
12 重水分子のモデリング	MI-fit	ge4.pdb, m3.2fofc.map	ge4_mifit.pdb	↓	
13	テキストエディタ	ge4.pdb, ge4_mifit.pdb	ge4a.pdb	↓	
14 重水分子の準備 2	CNS (ge5.inp)	ge4a.pdb	ge5.pdb	↓	
15 全体の精密化	CNS (re5.inp)	ge5.pdb	re5.pdb, m5.2fofc.map	↓	21.0
16 H/D 交換率の精密化準備	CNS (ge6.inp)	re5.pdb	ge6.pdb	H,D, D ₂ O	
17	CNS (al6.inp)	ge6.pdb	al6.pdb	H/D 交換率	
18 H/D 交換率の精密化	CNS (hd6.inp)	al6.pdb	hd6.pdb	↓	20.7
19 H/D 交換率の規格化	Perl (hd.pl)	hd6.pdb	hd6a.pdb	↓	
20 最終精密化	CNS (re6.inp)	hd6a.pdb	re6.pdb	↓	20.0
21 精密マップの計算	CNS (ma6.inp)	re6.pdb	m6.2fofc.map	↓	
22 PDB 登録用ファイル変換	Perl (pdb.pl)	re6.pdb	pdb.pdb	↓	
23 構造検討	MI-fit	re6.pdb, m6.2fofc.map		↓	

Tyr12, Cys38, Ala43 主鎖のアミノ基がこれにあたる。これらは次に行う重水素原子の付加において除外する。具体的には, ge3.inp の終わりのあたりに, 以下の命令を追加する。

```
delete selection = (name D and resname Val and resid 4) end
```

これは, Val4 の主鎖のアミノ基の D 原子をモデルから除外するという命令である。次に, MI-fit で付加された D 原子の位置の調整を行う。Fig. 5 に例を挙げる。Ser46 のトーション角を変更して, 中性子換わらず正の中性子散乱密度が観測できない時が密度にあわせる。MI-fit でセーブされる原子 pdb ファイルは CNS でそのまま使用することができない。セーブした PDB ファイルを基に ge4.pdb を変更する。この後, 重水素構造の精密化を行う。

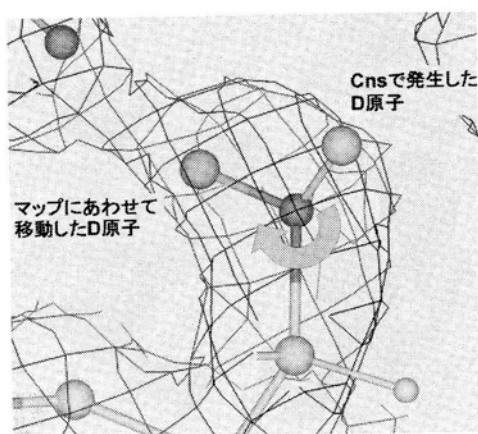


Fig.5 Modeling of D atom in Ser46.

4.2.4. 水分子の精密化 (10-15)

タンパク質中の D 原子を精密化した後は, 溶媒領域の重水 (D₂O) 分子の精密化を行う。この手順も, D 原子の精密化と同様に MI-fit 上でのモデリングと CNS での精密化の組み合わせにより行う。X 線構造で観測された D₂O の酸素原子を座標に加えて, CNS により D₂O の D 原子を発生させる。

比較的分解能の高い (~1.5 Å) の中性子解析では, 蛋白質を取り囲んでいる重水溶媒はバルク水ではなく, 配向した構造水として観測される。また, その配向の様子は水素結合や蛋白質表面の静電ポテンシャルにより, 3 次元的に多様である [15]。Fig. 6 に配向と D₂O 分子のダイナミクス関係を示す。左から, D-O-D の 3 原子全てが固定されているもの, D 原子が O-D 軸上を回転しているもの, O 原子が D...D 軸上を回転しているもの, D₂O 全体が, O 原子あるいは重心を中心として自由回転しているものである。これらの 4 種類それぞれに応じた精密化を行う必要がある。

これら 4 種類の水分子は, その形状に応じて観測された原子のみを分子モデルに導入する。しかしながら, 蛋白質本体と違い D₂O 分子はすべて二つの D 原子と O 原子から構成されている。このために, 観測された原子だけでフーリエ変換を行うと欠損した原子分の中性子散乱長が不足してしまい, 不正確な原子位置と温度因子を与える。これを避けるために, 欠損原子の中性子散乱長については近接原子にその不足分を与えて, D₂O 合計の中性子散乱長を等しくするようにして最適化を行う。各原子に割り当てる散乱長を Fig. 7 に示す。D 原子と O 原子の散乱長

は、それぞれ0.6671, 0.5803であり、総和が1.9145になる様に値を設定している。今回の最適化では、Fig. 6の(a)-(d)のタイプに対して、残基名をDOD, OD, DD, BALと設定して、中性子散乱密度の計算法を変えた。

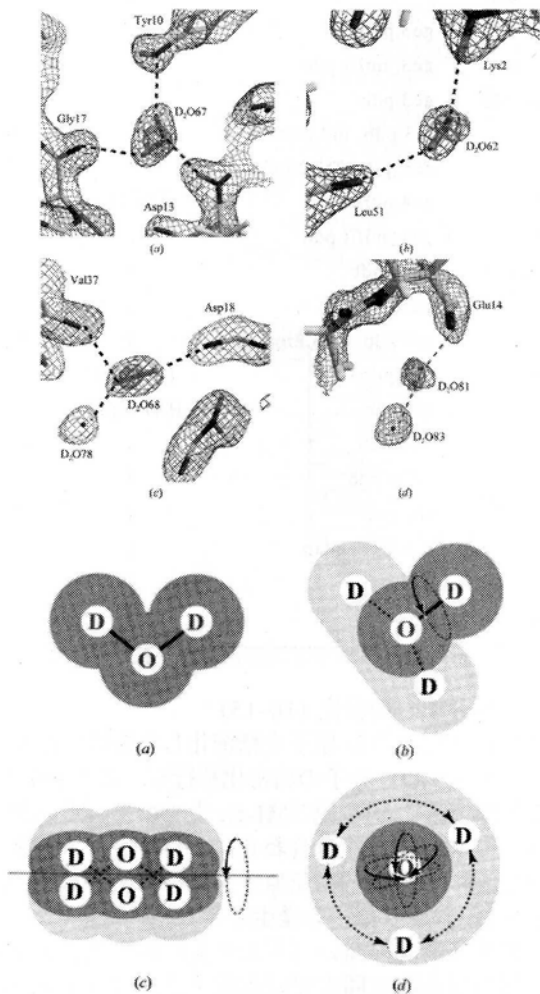


Fig.6 Dynamics of D₂O observed in neutron experiment.

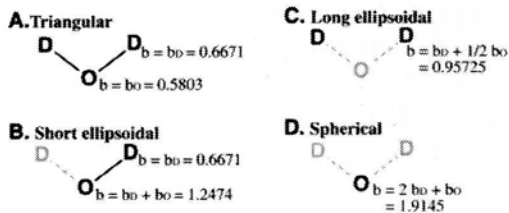


Fig.7 Scattering lengths for various shaped D₂O.

4.2.5. H/D 交換率の精密化 (16-20)

最後のステップとして、主鎖のH/D交換率の精密化を行う。分子内部のアミノ基の一部はH原子のままであるため、中性子フーリエ図ではH原子の負の散乱長に対応して、負の核密度が観測される。これに対応するために、初期値を50% H原子, 50% D原子として、それぞれの存在確率を最適化する。このH/D交換率の最適化は、理論上は極性残基の側

鎖のD原子についても行うことができるが、これらの側鎖を持つ極性アミノ酸は蛋白質表面で溶媒に接していることが多いため、100% D原子に変化しているとみなすことができる。実際の処理の流れとしては、これまで負の中性子散乱密度が観測されているために除外していた、Cys5, Tyr12, Cys38, Ala43の主鎖のアミノ基についてもD原子を発生させ、主鎖のD原子を分割する。今回は解析では、それぞれのsegidをAC1, AC2として、AC1とAC2に対してそれぞれ、H原子のD原子の中性子散乱長($b_H = -0.374$, $b_D = 0.667$)を与えて、占有率の精密化を行った。また、温度因子は結合しているN原子の値を代入した。その後、H原子とD原子の占有率の合計が1になる様にPerlスクリプトで規格化を行う。最後に、最適化した占有率のモデルで全体の精密化を行う。

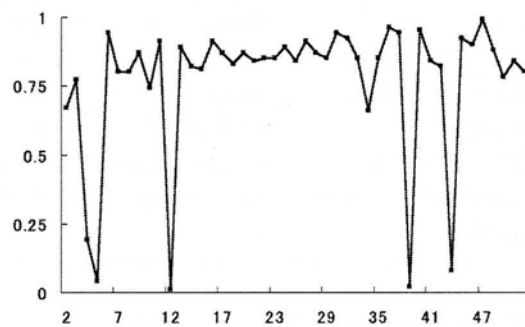


Fig.8 Occupancy of D atom of ND group in main-chain.

4.3. 精密化後の処理 (21-23)

精密化の後、精密なフーリエマップの計算とProtein Data Bankに登録するためのファイル形式の変換を行う。最終的な精密化マップを、Fig. 9に示す。現在のBIX-3,4で良好な中性子解析が行われた場合には、H原子D原子の90%以上を決定することができる。ところで、今回は手順の簡素化のために各手順を1回ずつしか行わなかったが、実際の解析では、各ステップを精密化が収束するまで数回繰

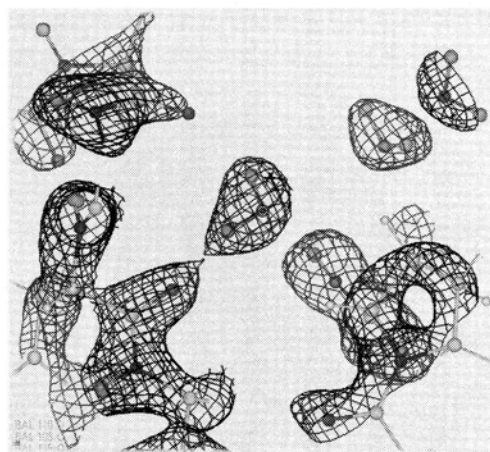


Fig.9 The final 2lFol-lFcl Fourier map in this refinement.

り返す必要がある。ただし、十分な分解能の中性子データ (1.5 ~ 1.8 Å程度) があれば、1 回ずつの精密化でもかなり良い解析結果をえることができる。今回の解析手順では $R = 20.0\%$ ($R_{\text{free}} = 24.2\%$) であり、精密化を数回行った場合と、ほぼ同じ結果を得ることができた。

今回の報告に使用したファイル、プロトコール等の記述、使用方法等の詳細については、著者 (chatake@rri.kyoto-u.ac.jp) まで問い合わせ下さい。

5. 謝辞

今回の解析に使用した CNS 用の構造化学情報ファイルはミュンヘン工科大学 Andreas Ostermann 博士と共に CNS の X 線用ファイルを変更したものである。中性子データの測定は、JAEA の栗原和男博士、茨城大学の田中伊知朗准教授、新村信雄教授、南カリフォルニア大の Robert Bau 教授と共に行った。

参考文献

- [1] N. Niimura, S. Arai, K. Kurihara, T. Chatake, I. Tanaka, R. Bau, *Cell Mol. Life Sci.* **63**, 285-300 (2006).
- [2] M. P. Blakeley, M. Cianci, J. R. Helliwell, P. J. Rizkallah, *Chem. Soc. Rev.* **33**, 548-557 (2004).
- [3] I. Tsyba, R. Bau, *Chemtracts.* **15**, 233-257 (2002).
- [4] 松崎尹雄, 蛋白質核酸酵素 **50**, 869-875 (2005).
- [5] K. Tanikawa, Y. Matsumoto, T. Matsuzaki, M. Shimizu, M. Matsumoto, M. Fukuoka, *Biol. Pharm. Bull.* **23**, 235-239 (2000).
- [6] I. Tanaka, K. Kurihara, T. Chatake, N. Niimura, *J. Appl. Cryst.* **35**, 34-40 (2002).
- [7] K. Kurihara, I. Tanaka, N. Niimura, M. Refai Muslih, A. Ostermann, *J. Synchrotron Radiat.* **11**, 68-71 (2004).
- [8] D. A. A. Myles, C. Bon, P. Langan, F. Cipriani, J. C. Castagna, M. S. Lehmann, *Physica B* **241&243**, 1122-1130 (1998).
- [9] P. Langan, G. Greene, B. P. Schoenborn, *J. Appl. Cryst.* **37**, 24-31 (2004).
- [10] 茶竹俊行, 森本幸生, 第 41 回京都大学原子炉実験所学術講演会報文集, 83-88 (2007).
- [11] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, *Acta Cryst. D* **54**, 905-921 (1998).
- [12] Z. Otwinowski, W. Minor, *Methods Enzymol.* **276**, 307-326 (1997).
- [13] T. Chatake, K. Kurihara, I. Tanaka, I. Tsyba, R. Bau, F. E. Jenney Jr, M. W. W. Adams, N. Niimura, *Acta Cryst. D* **60**, 1364-1373 (2004).
- [14] T. Chatake, K. Kurihara, R. Kuroki, Y. Morimoto, to be submitted.
- [15] T. Chatake, A. Ostermann, K. Kurihara, F. G. Parak, N. Niimura, *Proteins* **50**, 516-523 (2003).
- [16] K. Kurihara, I. Tanaka, T. Chatake, M. W. W. Adams, F. E. Jenney Jr., N. Moiseeva, R. Bau, N. Niimura, *Proc. Natl. Acad. Sci. USA* **101**, 11215-11220 (2004).

茶竹 俊行 (ちゃたけ としゆき)

略歴：1999 年 東京工業大学大学院生命理工学研究科博士後期課程修了。同年 日本原子力研究所 (現・日本原子力研究開発機構) 博士研究員, 2004 年 千葉科学大学 薬学部 助教授, 2006 年 8 月 京都大学原子炉実験所 助教授, 2007 年 4 月より現職。

所属：京都大学原子炉実験所 准教授
e-mail : chatake@rri.kyoto-u.ac.jp
専門：中性子構造生物学
趣味：スポーツ観戦



森本 幸生 (もりもと ゆきお)

略歴：1987 年 大阪大学大学院理学研究科博士課程修了。同年 姫路工業大学 助手, 1991 年 徳島大学工学部 助教授, 1996 年 姫路工業大学 (現 兵庫県立大学) 理学部 助教授, 2003 年より現職。

所属：京都大学原子炉実験所 粒子線基礎物性研究部門 教授
e-mail : morimoto@rri.kyoto-u.ac.jp
専門：タンパク質結晶学
趣味：バッハを聴くこと

