

## シンクロトロン放射光を使った 原子分解能の蛋白質結晶構造解析

野中 孝昌

長岡技術科学大学生物系\*および理化学研究所播磨研究所

### Protein Crystal Structure Analysis using Synchrotron Radiation at Atomic Resolution

Takamasa NONAKA

Department of BioEngineering, Nagaoka University of Technology  
And The Institute of Physical and Chemical Research (RIKEN), RIKEN Harima Institute

We can now obtain a detailed picture of protein, allowing the identification of individual atoms, by interpreting the diffraction of X-rays from a protein crystal at atomic resolution, 1.2 Å or better. As of this writing, about 45 unique protein structures beyond 1.2 Å resolution have been deposited in the Protein Data Bank. This review provides a simplified overview of how protein crystallographers use such diffraction data to solve, refine, and validate protein structures.

#### 1. はじめに

1978年, Sakabeらは分子量約5.8 kDaのインスリンの4°Cで測定した1.2 Å分解能のX線回折強度データに対して, ブロック対角法による最小二乗法で結晶構造精密化を行った<sup>1,2)</sup>。等方性温度因子が30 Å<sup>2</sup>以下の原子の異方性温度因子を精密化した結果, 差フーリエ図上で蛋白質のX線結晶構造解析では初めて水素原子を確認することに成功した。

1997年, 塩野らは分子量14 kDaのプロテアーゼ ScNP (*S. caespitosus*) の1.05 Å分解能のX線回折強度データを用いて, *ab initio*に構造を決定することに成功した<sup>3)</sup>。パターンサーチで見出した亜鉛原子による位相から出発して, 独自に開発したプログラム LODEM<sup>4)</sup>によりほとんどの原子位置を決定しうる明瞭な電子密度分布図を得た。

1998年, Harataらは分子量16 kDaの七面鳥リゾチーム (1.12 Å) と17 kDaのヒトリゾチーム (1.15 Å) の12°Cで測定したデータを用いて結晶構造の精密化を行った<sup>5)</sup>。分子量10 kDaを越える蛋白質としては初めての完全行列最小二乗法による精密化であり, 水素原子を除く全原子の異方性温度因子が決定された。α炭素の異方性温度因子を用いて剛体振動解析を行い, 両リゾチームの内部運動を

ORTEP 図上に見事に示した。

これら日本における atomic resolution (一般的な化合物の最短の原子間距離1.2 Åの分離が可能という意味で Sheldrick が真の意味の atomic resolution と定義している<sup>6)</sup>。未だ定まった訳語はないと思われるが, 本稿では原子分解能という訳語を充てることにする。) でのX線結晶構造解析に関連する報告には, Dauterらの総説“Proteins at atomic resolution”<sup>7)</sup>から始まる一連の Current opinion in structural biology 誌上に見られる蛋白質の原子分解能の構造解析に関する総説<sup>8,9)</sup>に述べられているエッセンスがほぼ全て述べられているといってもよい。すなわち, 蛋白質結晶の原子分解能のX線回折強度データを得ることができれば次のようなことが可能となる。

(1) 非水素原子の異方性温度因子と“riding”水素原子(直接共有結合している非水素原子の位置から立体化学的な理想的位置に配置された水素原子)を加味して結晶構造の精密化を行えば, 最終的な R-factor が10%を切ることもあり得る。等方的に精密化した場合と比べて, 高角での R-factor が小さいので原子位置の誤差が非常に小さくなる。実際, Harataらのリゾチームの座標の誤差は約0.03 Åであり<sup>5)</sup>, 2 Å分解能程度で精密化された構造の Luzz-

\* 長岡技術科学大学生物系 〒940-2188 長岡市上富岡町1603-1  
TEL 0258-47-9430 FAX 0258-47-9400 e-mail nonaka@vos.nagaokaut.ac.jp

tiの方法<sup>10)</sup>で見積もった誤差が0.20–0.25 Å程度であるのと比較して極めて正確であるといえることができる。

(2) 差フーリエ図上に溶媒分子や乱れた構造のみならず、水素原子をも容易に見出すことができるようになる。Rubredoxinでは2/3以上<sup>11)</sup>、Sakabeらのインスリンでは実に90%の水素原子の同定が可能となっている<sup>1,2)</sup>。蛋白質と直接水素結合で結びつけられていない多くの水分子を見出すことも可能で、rubredoxinでは結晶中に存在する水分子の50%以上が同定されている。側鎖のみならず、場合によっては主鎖までも静and/or動的に乱れている構造に至る所に見出すことができる。乱れたアミノ酸残基の割合が優に10%を超えることもしばしばある。また、精密化の最終段階の差フーリエ図では $\pm 0.4 \text{ e}\text{\AA}^{-3}$ を超える残余のピークがなくなる。Harataらのリゾチームの場合は、 $-0.21 \text{ e}\text{\AA}^{-3} \sim +0.26 \text{ e}\text{\AA}^{-3}$ である<sup>5)</sup>。

(3) 温度因子があまり大きくない場合には、炭素、窒素、酸素原子を明瞭に区別することができる。すなわち、電子一個分の違いを判別することが可能となる。したがって、通常の分解能では、困難であったグルタミン、アスパラギン、およびヒスチジンの側鎖の表裏を同定することができる。

(4) 結晶中に金属原子などの重い原子を含む場合は、パターンサーチなどによりそれらの原子位置を決定し、得られた初期位相をSHELXS<sup>12)</sup>やLODEM<sup>4)</sup>などのプログラムで拡張/改良することで、*ab initio*に構造を決定することができる。

筆者らも遅時きながら、2種類の蛋白質の原子分解能のX線結晶構造解析を行っているので、これまでの経験を踏まえつつ、結晶化から構造の評価までの概説を行うことにする。1つはアミノ酸残基数111のウシガエル卵由来のレクチン(RCELと呼ぶことにする)で、もう一つは419残基の細菌由来のキチン分解酵素(CA6と呼ぶ)である。

## 2. 結晶化

原子分解能の結晶構造解析を行っていて、もっとも多く受ける質問の一つは、“どうすれば原子分解能の回折を与える蛋白質の結晶を得ることができるのか?”ということである。ところが、私どもはこの質問には明快に答えることができない。実際、RCELもCA6も結晶化条件の初期探索はともに広く一般に行われているランダムスクリーニング法を採用したにすぎないし、多少の条件の最適化を行った後は特別な工夫はなにも行っていない。数多くの原子分解能の構造解析に関する研究を行っているWilsonらの総説<sup>7,8)</sup>にも、結晶化に関する記述はほとんど見られない。いくつかの国々で第3世代の放射光施設が共同で利用され、受光面の大きな検出器がルーチンに使われている以上、原子分解能の回折強度データ収集の正否はいかにいい結晶を得るか、という一点にかかっているといってもよい。

純度の高い試料を使うということは当然としても、クライオ条件を確立することは、望ましいことではあっても必須ではない。EU 3-D Validation Networkの総説<sup>13)</sup>に記述されている原子分解能の解析が行われた8つの例のうち、実に7つは室温でデータの収集が行われたのである。「1.はじめに」で紹介した3種類の例とCA6も室温かまたはそれに近い温度で、データ収集が行われている。すなわち、放射線損傷が少なければ室温でも大きな問題はないのである。CA6はSPring-8のBL44B2で測定を行ったが、1個の結晶から1.13 Å分解能のデータを収集することができた。Harataらはより生理条件に近い状態での構造を得るために、意図的に室温での測定を行っている<sup>5)</sup>。

単位格子内の分子量が同じなら、単位格子が小さい方がより強い反射強度を与える。しかしながら、リゾチームのように様々な種類の結晶が得られるならともかく、結晶化条件を変化させることで単位格子の大きさを制御することは実際上不可能である。残された方法は結晶自体を大きくすることである。最近の論文には結晶の大きさが記載されていないことが多いので不完全な調査ではあるが、もっとも小さいものでも $0.5 \times 0.5 \times 0.2 \text{ mm}^3$ である。それでは、結晶を大きくするにはどうすればよいか。これには従来から様々な方法が試みられているが、原子分解能の解析でもっとも多いのはシーディング法である。しかも、マイクロシーディングとマクロシーディングを組み合わせている例が多い。筆者らはシーディング法での結晶化を行ったことはないが、測定の時には大きいものから選んで使っている。

## 3. 測定

原子分解能のX線回折強度データの収集のためには、高角の弱い反射を精密に測定するため検出器の受光面はできるだけ大きく、また低角が飽和せぬようダイナミックレンジもできるだけ広い方がよい。さらに、高角領域の斜入射による積分計算の困難を排するためには、湾曲した検出器か $2\theta$ が可変の検出器であることが望ましい。湾曲タイプとしてはPFのBL6Cのイメージングプレートを搭載した円筒カメラが理想的であるし、 $2\theta$ を変化させることができればダイナミックレンジが狭いにしろ読みとり時間の短いCCDカメラが極めて効率的である。平板でかつ $2\theta$ を変化させることのできない場合には、低角、中角、および高角用に結晶と検出器の距離を設定し直す必要があるだろう。

一定面積の受光面でより多くの回折強度データを収集するためと、吸収の影響を無視し得るほど小さくするために、波長は1 Å以下で測定されることが多い。波長の選択は、X線の強度、検出器の大きさや位置分解能、結晶—検出器間距離、必要とする分解能、および結晶の大きさを勘案して決めるべきであるが、筆者らは十分な強度が得られるもっとも短い波長で測定を行っている。

測定法は振動法とワイセンベルグ法に限定されている

が、いずれにしろ S/N 比をよくするためにはできるだけ振動角を小さくし露出時間を長くするのがよい。しかしながら、マシンタイムに限りがある現状では S/N 比を犠牲にしても振動角を大きくしなければならない場合がほとんどである。

確かに、室温でも原子分解能の X 線回折強度データを収集することは可能な場合もあるが、たとえ液体窒素温度にしても放射線損傷は起こるようである。できるだけ精密なデータ収集、特に超高分解能 (Kuhn らの定義<sup>14)</sup>によれば、0.9 Å を超える分解能) のデータを得ようとするなら液体ヘリウムによる冷却も必要であろう。

さて、原子分解能に限られたことではないが、得られたデータの分解能がどれだけであると表現するのはなかなか難しい問題である。Acta Crystallographica Section D の notes for authors では、 $I \geq 3\sigma(I)$  の反射の最外角シェルにおけるコンプリートネスが70%以上、かつ全体のコンプリートネスが93%以上、さらに全体の  $R_{\text{merge}}$  が20%以下と非常に厳しい要求を行っている。これに対し、原子分解能の場合には Sheldrick がもう少し現実的な定義<sup>6)</sup>を行っており、広く受け入れられているようである。すなわち、 $I \geq 2\sigma(I)$  の反射の最外角シェルにおけるコンプリートネスが50%以上で、 $R_{\text{merge}}$  がおおむね25%以下であることを満たす分解能である。(それにしても、いったいいくつかのシェルに区切ればいいのだろうか?) 参考のため Table 1 に筆者らが SPring-8 の BL44B2 で測定した RCEL のデータについてまとめてみた。

#### 4. 位相決定

Frazaõ らのシトクローム c6 の *ab initio* 構造決定<sup>15)</sup>を除けば、最初から新規蛋白質の構造が原子分解能で決定された例は今のところほとんどない。いずれの例も、すでに構造が決定されている蛋白質の分解能を向上させた精密化である。しかしながら、原子分解能のデータが位相決定の上で極めて有効に利用できることは、いくつかの論文が示している<sup>16-18)</sup>とおりである。

RCEL は当初、回転対陰極型 X 線発生装置上に設置されたイメージングプレート X 線回折装置でデータの収集を行った。ターゲットは銅であり、検出器の大きさが  $190 \times 190 \text{ mm}^2$  で、結晶-検出器間距離は55 mm 以下にすることができなかったので、1.54 Å が分解能の限界であった。RCEL はすい臓由来リボスクレアーゼスーパーファミリーに属しているので、牛すい臓由来リボスクレアーゼ A (RNase A) の構造を使って分子置換法で最終的には構造を決定することができた。ところが、RCEL には異なる位置に架かったジスルフィド結合があり、その周囲のモデルを構築するには延べで千にも上るオミットマップを繰り返し作製し、全構造を構築するには数ヶ月を要した。その後、PF で1.14 Å 分解能、SPring-8 で1.06 Å 分解能のデータを得ることができたので、試みに部分構造か

Table 1. Summary of data collection and processing statistics

No. of crystals	1
Beamline	44B2
Temperature (K)	100
Diffractometer	Rigaku R-Axis IV
Data collection method	Oscillation
IP size (mm <sup>2</sup> )	300 × 300
Detector distance (mm)	196
Spindle axis	Parallel to [1 1 0]
Wavelength (Å)	0.7
Space group	$P3_121$
Unit-cell dimensions (Å)	$a=b=41.653, c=118.628$
Resolution range (Å)	34.5–1.06
Processing software	
Data collection	DENZO <sup>30)</sup>
Scaling	SCALA <sup>31)</sup>
$R_{\text{merge}}(I)$	0.042 (0.250) <sup>†</sup>
Completeness	0.971 (0.924) <sup>†</sup>
$\langle I/\sigma \rangle$	7.9 (2.9) <sup>†</sup>
Multiplicity	3.9 (2.9) <sup>†</sup>
No. of unique reflections	53,295 (3,672) <sup>†</sup>

<sup>†</sup>Values in parentheses are for the highest resolution shell (1.09–1.06 Å).

ら出発して徐々に位相を改良していくプログラム ARP<sup>19,20)</sup>を適用してみた。RCEL の111アミノ酸残基のうち、8 残基を占める最大の  $\alpha$ -ヘリックスは RNase A の対応する部分と側鎖を含めてほぼ同一の構造であり、分子の中心に近いところに位置している。そこで、この部分構造が正しいと仮定して、ARP で位相の改良/拡張を行った。SGI 社の Origin2000による丸一日程度の計算で Fig. 1 に

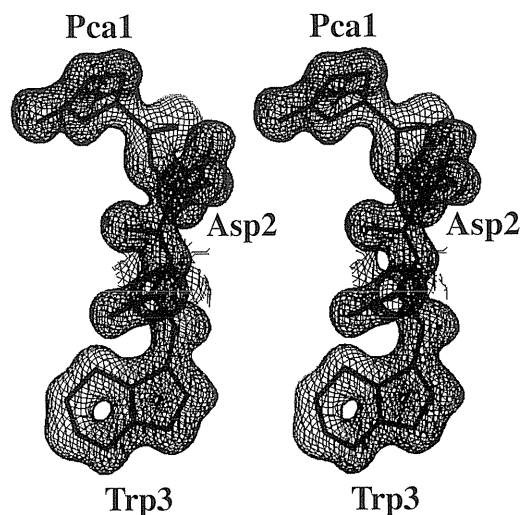


Figure 1. Stereo drawing of a fragment of the ARP ( $3F_0-2F_0, \alpha_c$ ) map for bullfrog egg lectin (RCEL) at 1.14 Å resolution with atoms connected on the basis of interatomic distances. These atoms were automatically placed without stereochemical restraints and thus they deviate from the final, correct positions. "Pca" is the three-letter code for "pyroglutamic acid".

示すように、原子一つ一つを明瞭に区別することのできる質の高い電子密度分布図を得ることができた。原子分解能のデータがあれば数ヶ月の作業を1日に短縮することも可能である。

## 5. 精密化

原子分解能のX線回折強度データが集まれば、パラメータの数に対して十分データの数が多くなるので、通常の分解能では不可能な異方性温度因子の精密化も可能となる。蛋白質の異方性の温度因子を決定できる広く普及したプログラムとしては、最近まで事実上SHELXL-97<sup>21)</sup>およびそれ以前のバージョンしか存在しなかったが、CCP4<sup>22)</sup> ver.3.5のREFMAC<sup>23)</sup>も個々の原子の異方性温度因子を精密化できるようになったようである。RCELの場合、データ/パラメータ比が4を超えるのでSHELXL-97を用いて全ての非水素原子の異方性温度因子 (Fig. 2) を決定することができたが、データ/パラメータ比が3程度でも十分異方性温度因子の精密化が可能である。

また、解析例が多いとは言えないので、原子分解能の結晶構造精密化のプロトコルは確立までには至っていない。それぞれの研究者が試行錯誤で精密化を行っている段階であるが、(1)等方性温度因子を用いた計算で十分収斂した後、異方性温度因子を導入する。(2)乱れと水分子を追加しつつ、精密化を続ける。(3)“riding”水素原子を導入する。(4)ブロック対角法または完全行列最小二乗法による精密化、というステップを踏むのが一般的なようである。さらに付け加えるならば、(5)立体化学的制限を課さない精密化、となるであろう。RCELの(1)~(3)のステップにおけるR-factorの変化をプロットしたものがFig. 3である。異方性温度因子を導入すると4%程度、“riding”水素原子を導入すると1%程度R-factorが下がっていることが分かる。

データ/パラメータ比が4に満たないリゾチームのケースに対して完全行列最小二乗法による精密化を成功させたHarataらのプロトコル<sup>9)</sup>は、現在のところ、もっとも優れているように思える。収斂の状況を監視しつつ、最初はブロック対角化法から始め、徐々にブロックを大きくしていくやり方は大いに参考になる。一方、Kuhnらは原子位置の誤差、あるいは結合長や結合角の推定標準偏差を算出するために、精密化の最後に1サイクルのみ完全行列最小二乗法の計算を行っている<sup>14)</sup>。論文に記述がないので十分に収斂しているか否かは定かではないが、立体化学的制限を課さずに精密化を行っている点は高く評価できる。

## 6. 構造の評価

ブロック対角法あるいは完全行列最小二乗法により精密化を行えば、個々のパラメータに対して推定標準偏差が与えられる。したがって、通常の分解能の解析のようにLuzzatiの方法<sup>10)</sup>によらなくとも、原子位置の誤差を求め

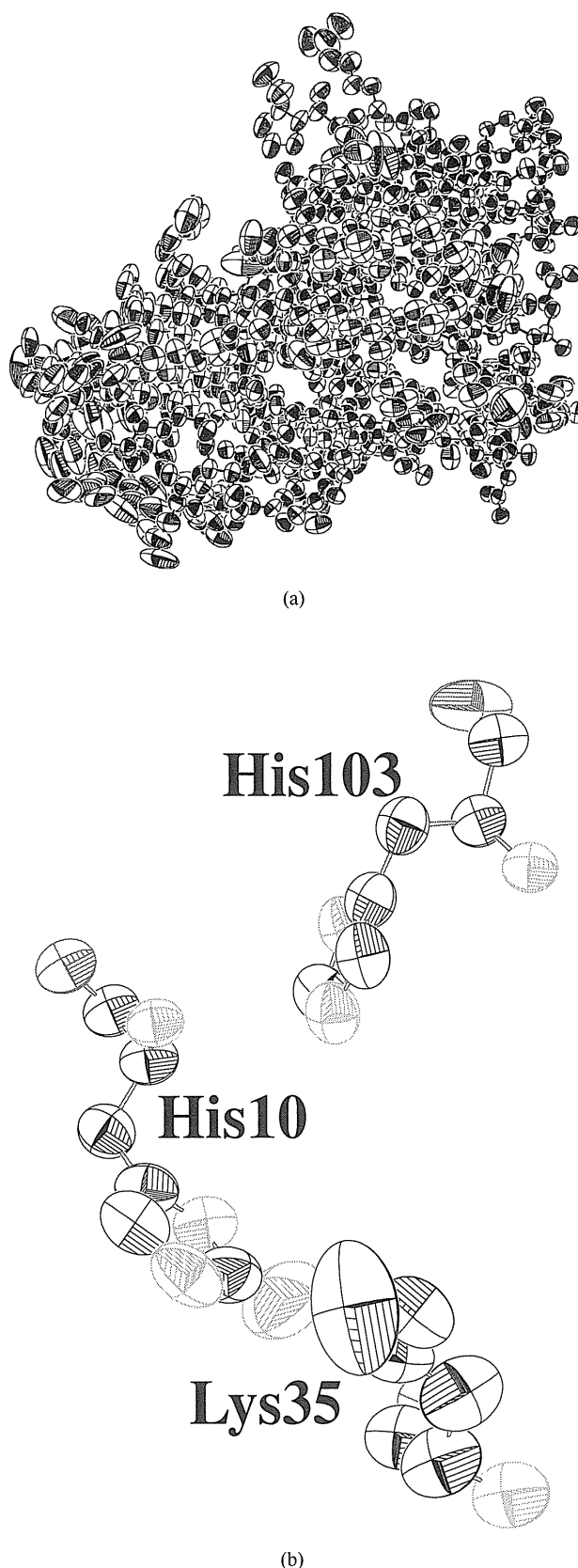


Figure 2. Anisotropic displacement ellipsoids drawn at 50% probability level for the all nonhydrogen atoms of bullfrog egg lectin (RCEL). (a) Overall structure of RCEL. (b) Catalytic center of RCEL.

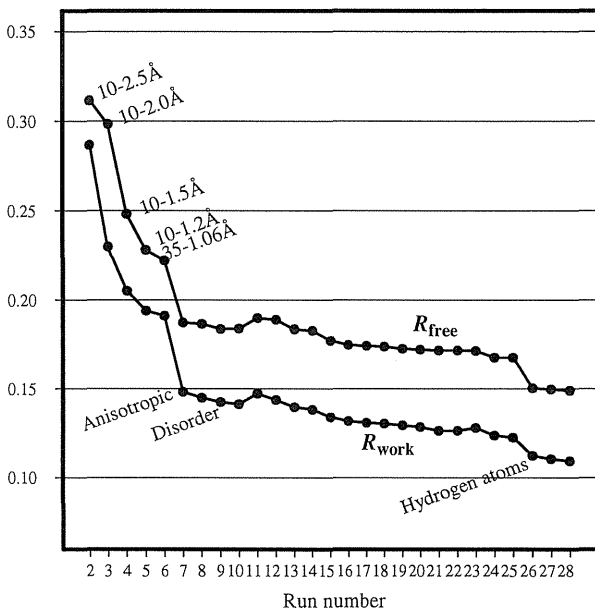


Figure 3. Course of the R factors during the refinement of bullfrog egg lectin (RCEL).

ることができる。RCELの場合、ブロック対角法による最小二乗法で精密化した後の個々の原子位置の推定標準偏差を平均した値は $0.046 \text{ \AA}$ であった。実験室系の $1.54 \text{ \AA}$ 分解能のデータを用いたX-PLOR<sup>24)</sup>による精密化では、Luzzatiプロットより平均誤差を求めると $0.20 \sim 0.25 \text{ \AA}$ であった。これに比べると原子分解能では、精密化の度合いが非常に高まっていることが分かる。

温度因子がそれほど大きくない場合には、電子密度分布図より炭素、窒素、および酸素の区別ができることはすでに述べた。実際にRCELではどのように見えるのかをFig. 4に示した。さらに、より定量的に示すため、B-factorと原子中心の電子密度の関係をFig. 5に示した。B-factorが $20 \text{ \AA}^2$ 以下の場合にはほぼ完全に元素の判別を付けることができる。

蛋白質の立体構造の評価法としてはRamachandranプロット<sup>25)</sup>が視覚的に捉えることができてもっともわかりやすい。ところが、原子分解能の場合には、プロットするまでもなくほとんど全てのアミノ酸残基がエネルギー的に許容される範囲に収まってしまうので、もっと厳密な評価法を適用せねばならない。Wilsonらを中心とするEU 3-D validation networkでは、命名法、立体化学的パラメータ、水素結合、溶媒、温度因子、充填密度、あるいは格子定数にいたるさまざまな評価の対象について“評価”を行っている<sup>13)</sup>。8つの原子分解能の構造に対して、PROCHECK<sup>26)</sup>、PROVE<sup>27)</sup>、SQUID<sup>28)</sup>、およびWHATCHECK<sup>29)</sup>の4つのプログラムを適用したうえで、原子分解能の構造、評価プログラム、および原子分解能の構造解析の将来について述べている。詳しくは原著を参照

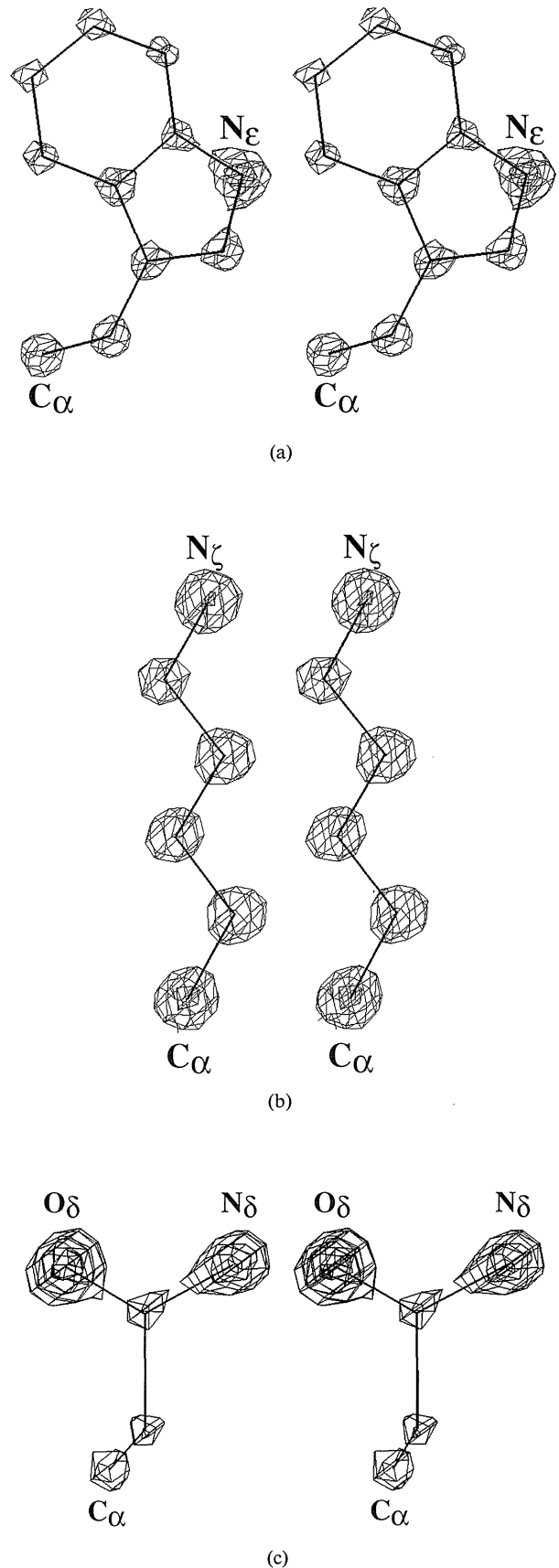


Figure 4. Stereo drawings of the omit electron density maps for the three residues, (a) Trp3, (b) Lys9, and (c) Asn38, of bullfrog egg lectin (RCEL).

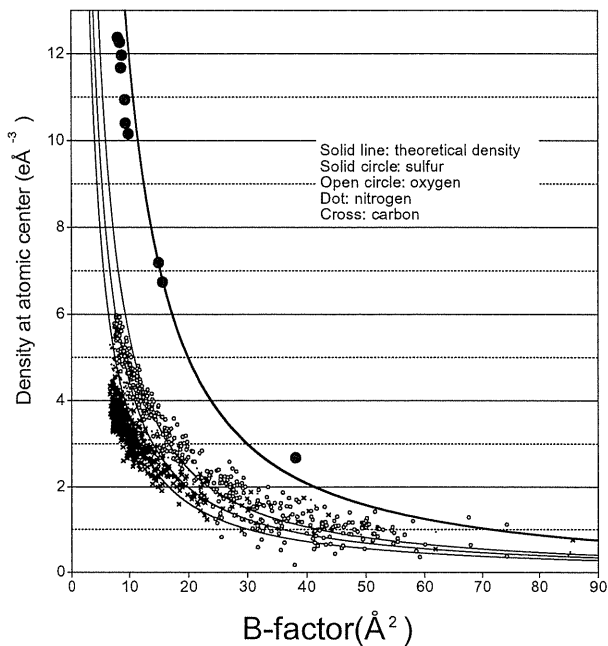


Figure 5. Observed ( $F_o$ ,  $\alpha_c$ ) and theoretical electron densities at atomic centers as a function of temperature factor.

していただくとして、注目すべき点を2つだけ簡単に述べる。(1)原子分解能の構造のポストリファインメントの結果、格子定数に0.5%もの誤差があることが分かった。ハンブルグのEMBLでは格子定数を実験的に決定する方法が改善され、誤差は0.1%以下になった。(2)ペプチド結合の2面角 $\omega$ の理想値からのずれは、原子分解能の構造の方がむしろ大きい。つまり、従来の精密化における制限がきつすぎたのであり、2面角には制限を課さない方がよい。

## 7. 終わりに

蛋白質結晶から原子分解能、さらには超高分解能のX線回折強度データがぞくぞくと得られるようになってきた。どうやら、そこには宝が眠っているらしい。動的情報や水素原子はすでに見えた。結合電子も見えつつある。*Ab initio* 構造決定も増えるだろう。さてその次は？

## 謝辞

PFのBL18BでのRCELの最初の原子分解能のデータ収集では、PFの渡邊信久博士より、SPring-8のBL44B2でのRCELとCA6のデータ収集では理研播磨研究所の足立伸一博士より協力をいただいた。両博士に心より謝意を表す。

## 参考文献

- 1) N. Sakabe et al.: Excerpta Medica: International Congress series; no.468 (Excerpta Medica, Amsterdam, 1978) pp. :73-80.
- 2) 坂部知平他: 蛋白質核酸酵素 **24**, 582 (1979).
- 3) 塩野正明他: 日本結晶学会平成9年度年会(つくば) 講演要旨集 P-039 (1997).
- 4) M. Shiono and M. M. Woolfson: Acta Cryst. **A48**, 451 (1992).
- 5) K. Harata et al.: Proteins **30**, 232 (1998).
- 6) G. M. Sheldrick: Acta Cryst. **A46**, 467 (1990).
- 7) Z. Dauter et al.: Curr. Opin. Struct. Biol. **5**, 784 (1995).
- 8) Z. Dauter et al.: Curr. Opin. Struct. Biol. **7**, 681 (1997).
- 9) S. Longhi et al.: Curr. Opin. Struct. Biol. **8**, 730 (1998).
- 10) V. Luzzati: Acta Cryst. **5**, 802 (1952).
- 11) Z. Dauter et al.: Proc. Natl. Acad. Sci. USA **93**, 8836 (1996).
- 12) G. M. Sheldrick: Acta Cryst. **D49**, 18 (1993).
- 13) EU 3-D Validation Network: J. Mol. Biol. **276**, 417 (1998).
- 14) P. Kuhn et al.: Biochemistry **37**, 13446 (1998).
- 15) C. Frazão et al.: Structure **3**, 1159 (1995).
- 16) C. W. Carter: Structure **3**, 147 (1995).
- 17) M. M. Woolfson and J-X. Yao: Acta Cryst. **A46**, 409 (1990).
- 18) C. M. Weeks: Acta Cryst. **D51**, 33 (1995).
- 19) A. Perrakis: Acta Cryst. **D53**, 448 (1997).
- 20) Perrakis: Nat. Struct. Biol. **6**, 458 (1999).
- 21) G. M. Sheldrick and T. R. Schneider: Methods in Enzymology **277**, 319 (1997).
- 22) Collaborative Computational Project, Number 4: Acta Cryst. **D50**, 760 (1994).
- 23) G. N. Murshudov: Acta Cryst. **D55**, 247 (1999).
- 24) A. T. Brünger and L. M. Rice: Methods in Enzymology **277**, 243 (1997).
- 25) C. Ramakrishnan and G. N. Ramachandran: Biophys. J. **5**, 909 (1965).
- 26) R. A. Laskowski: J. Appl. Cryst. **26**, 283 (1993).
- 27) J. Pontius: J. Mol. Biol. **264**, 121 (1996).
- 28) T. J. Oldfield: J. Mol. Graphics **10**, 247 (1992).
- 29) R. W. W. Hoof: Nature **381**, 272 (1996).
- 30) Z. Otwinowski and W. Minor: Methods in Enzymology **276**, 244 (1997).
- 31) P. R. Evans: Proceedings of CCP4 Study Weekend, pages 114-122 (1993).