

学位論文 博士 (理学)

新規構造解析法開拓による
水溶液中蛋白質の立体構造研究

指導教員 中迫雅由 教授

令和元年度

慶應義塾大学大学院 理工学研究科

基礎理工学専攻 物理学専修

大出真央

論文要旨

蛋白質は細胞内部で他の生体高分子との相互作用や触媒作用、信号変換などを通して細胞の恒常性維持や細胞分裂といった生命現象を制御している。機能発現過程の蛋白質は熱揺らぎや水和水分子との相互作用による動的構造変化を生じているため、生命現象を素過程レベルで正しく理解するためには、溶液中での蛋白質の構造・動態を解析することが重要である。本研究では溶液中蛋白質の構造や動態を解析可能な手法として X 線小角散乱法(Small-angle X-ray scattering: SAXS)とクライオ電子顕微鏡法(Cryo-electron microscopy: cryoEM)に着目し、これらの手法による蛋白質構造・動態の新規解析法を構築し、様々な蛋白質に適用することで、その有効性を検証するとともに、各蛋白質の構造動態の解明に貢献した。

SAXS では、試料蛋白質溶液の散乱パターンから、慣性半径や会合状態、分子長などの試料構造情報が得られ、また非経験的分子形状推定法を適用することで分子概形が予測できる。SAXS での構造情報は結晶構造解析などと比べて低分解能ではあるが、高い測定環境自由度を活かして外部摂動による蛋白質構造変化を容易に測定できる。しかし、非経験的分子形状推定法で予測される分子形状には原理的な不確かさがある。本研究では、多変量解析を用いて多数の推定形状から尤もらしい形状を得る方法を開発し、先行研究よりも優れた成果を得た。この手法を、立体構造が未解明であった植物光受容蛋白質フォトトロピンとフィトクロム B の構造解析に適用した。光屈性等を制御するフォトトロピンでは、青色光受容前後での SAXS データから、全体形状と機能ドメイン配置を決定し、青色光吸収によって生じるリン酸化活性増大のメカニズムを低分解能で推定した。また、光形態形成を担うフィトクロム B については、クロマトグラフィーと SAXS を併用して得た SAXS データから、フィトクロム B の全体構造を提案し、赤色光励起されたフィトクロム B における溶液性状の変化を明らかにした。

これらの成果を踏まえ、さらに高い分解能で溶液中蛋白質の構造変化を探ることを目指し、cryoEM を用いた構造解析にも歩を進めた。cryoEM で取得できる氷包埋された蛋白質の観察像は極めて信号対雑音比が悪いが、構造揺らぎの中のあるコンフォメーションに固定されると考えられる。従って、膨大量の粒子像から蛋白質動態についての情報が得られると期待される。しかしながら、cryoEM による蛋白質動態解析手法は未成熟であり、数 MDa 程度の分子量の比較的巨大な試料についての解析が数例あるのみである。本研究では、cryoEM 像の持つ情報を分子動力学シミュレーションから得られる軌跡と組み合わせることで蛋白質の動態を解析する手法を新たに開発し、既存手法では解析が不可能な分子量 280 kDa のグルタミン酸脱水素酵素の溶液中での運動解析へと応用した。その結果、同酵素の機能ドメインに生じるナノメートルオーダーの運動を可視化するとともに、その運動の自由エネルギー地形を描くことに成功した。

目次

1章 序論

1.1 生命現象と蛋白質	1
1.2 蛋白質構造・動態解析の概略	2
1.3 本研究で用いる蛋白質構造・動態解析法	
1.3.1 X線小角散乱法	4
1.3.2 クライオ電子顕微鏡法	5
1.4 本研究で用いる解析法の問題点	
1.4.1 SAXSにおける問題点: 概形推定における不確かさ	8
1.4.2 cryoEMにおける問題点: 動態解析法の検討	8
1.5 本研究の概要と章立て	9

2章 X線小角散乱法

2.1 電磁波の放射	10
2.2 X線散乱	
2.2.1 Thomson 散乱	12
2.2.2 X線散乱と電子密度・構造因子	13
2.2.3 X線小角散乱	16
2.3 SAXSを用いた構造解析	
2.3.1 Guinier 近似	18
2.3.2 距離分布関数と分子最大長	20
2.3.3 Zernike-Prins の式と散乱プロファイルの濃度依存性	22
2.3.4 非経験的分子形状モデリング	23

3章 多変量解析を用いた蛋白質形状推定の高度化

3.1 蛋白質形状推定における多変量解析の導入	29
3.2 モデルの重ね合わせ	
3.2.1 慣性主軸を用いたモデル重ね合わせ	31
3.2.2 モデル一致度の評価	33
3.3 モデルのボクセルへの格納	34
3.4 主成分分析の基本原理	37
3.5 K 平均法によるクラスタリングと平均化処理	41
3.6 溶液中での形状既知蛋白質への本手法の適用	
3.6.1 好熱カビ由来 PDI への適用	
3.6.1.1 PDI 分子形状推定の詳細	42
3.6.1.2 提案手法による PDI 解析結果	45
3.6.2 シロイヌナズナ由来 PIL1 への適用	
3.6.2.1 PIL1 分子形状推定の詳細	47
3.6.2.2 提案手法による PIL1 解析結果	48
3.7 提案手法と既存手法との比較	50
3.8 議論	
3.8.1 提案手法の有用性	53
3.8.2 提案手法の改善可能性	54

4章 SAXS による植物蛋白質フォトトロピンの構造解析

4.1 青色光受容蛋白質フォトトロピン	
4.1.1 植物の生理現象と光信号伝達	55
4.1.2 シロイヌナズナ phot のドメイン構造と機能	

4.1.2.1 LOV ドメインの立体構造	58
4.1.2.2 LOV ドメインの光反応	58
4.1.2.3 STK ドメインの構造と機能	63
4.1.2.4 リンカー領域	63
4.1.2.5 LOV ドメインの役割	65
4.1.3 先行研究から予測される機能発現メカニズム	66
4.1.4 本研究の狙い	67
4.2 SPring-8 における X 線小角散乱実験	
4.2.1 測定試料	69
4.2.2 X 線小角散乱実験	70
4.2.3 測定手順	75
4.2.4 データ処理	75
4.3 phot1 LOV2-STK 断片の測定及び解析結果	
4.3.1 SAXS プロファイルの光依存的変化	76
4.3.2 慣性半径と原点散乱強度変化	76
4.3.3 距離分布関数の計算と最大分子長の比較	79
4.3.4 GASBOR と多変量解析による分子形状推定	80
4.3.5 変異体の測定による信号伝達機構の解析	85
4.4 全長 phot2 の測定及び解析結果	
4.4.1 SAXS プロファイルの光依存的変化	88
4.4.2 Guinier プロットによる慣性半径、原点散乱強度変化	88
4.4.3 距離分布関数の計算と最大分子長の比較	91
4.4.4 GASBOR と多変量解析による分子形状推定	91
4.4.5 変異体の測定による信号伝達機構の解析	97

4.5 議論	
4.5.1 phot1、phot2 機能ドメイン配置	98
4.5.2 phot1 における 475 番、636 番リジン残基の役割	101
4.5.3 phot2 における LOV1 の役割	103
4.5.4 phot 分子内信号伝達機構の普遍性	105
5章 SAXS による植物蛋白質フィトクロム B の構造解析	
5.1 赤色光・赤外光受容蛋白質フィトクロム B	
5.1.1 植物内におけるフィトクロムの役割	106
5.1.2 シロイヌナズナ由来 phyB のドメイン構造と機能	107
5.1.3 PSM 立体構造と機能	108
5.1.4 OPM 立体構造	110
5.1.5 先行研究で予測された機能発現メカニズムと本研究の狙い	110
5.2 SPring-8 における X 線小角散乱実験	
5.2.1 測定試料	112
5.2.2 X 線小角散乱実験	113
5.2.3 データ処理	115
5.3 Pr 型 phyB SAXS プロファイル	116
5.4 Pr 型 phyB 全体形状の推定	119
5.5 赤色光照射による phyB SAXS プロファイルの変化	121
5.6 Pfr 型 phyB による非特異的凝集の可逆性	123
5.7 Pfr 型 phyB 熱緩和の測定	125
5.8 議論	
5.8.1 Pr 型 phyB 四次構造の推定	128

5.8.2 赤色光受容に伴う非特異的凝集の意義	129
5.8.3 本研究の展望	131
6章 クライオ電子顕微鏡法	
6.1 電子顕微鏡における像形成理論	
6.1.1 位相コントラストと弱位相物体近似	133
6.1.2 位相コントラストとコントラスト伝達関数	134
6.2 cryoEM における観察試料	137
6.3 単粒子解析法	
6.3.1 概要	137
6.3.2 Beam Induced Motion とその補正	139
6.3.3 CTF の推定と補正	141
6.3.4 三次元再構成	143
6.3.5 高分解能構造の強調	145
6.4 連続的構造変化の解析方法	147
7章 cryoEM による GDH ドメイン運動解析	
7.1 超高度好熱菌由来グルタミン酸脱水素酵素	152
7.2 GDH の cryoEM 構造解析	
7.2.1 試料作成と cryoEM 観察	153
7.2.2 三次元再構成までの観察像処理	156
7.2.3 GDH 三次元再構成	156
7.2.4 Focused Classification による NAD ドメインコンフォメーションの分類	162
7.3 新規解析手法による GDH ドメイン運動解析	156

7.3.1 解析のための仮定と MD 計算の利用	163
7.3.2 MD 構造のマップへの変換	166
7.3.3 cryoEM 構造と MD 構造の比較	166
7.3.4 GDH ドメイン運動の自由エネルギー地形解析	167
7.4 cryoEM 構造に対する原子モデルの構築と比較	169
7.5 活性クレフト内部の構造	175
7.6 議論	
7.6.1 NAD ドメインの連続的構造変化	175
7.6.2 本研究の解析における仮定の妥当性	176
7.6.3 開発手法の改善点	177
7.6.4 GDH ドメイン運動解析における今後の展望	179
8 章 結語	180
付録 A アミノ酸の名称及び略記	182
付録 B Focused Classification による GDH 構造分類結果の詳細	183
B1. 各 Classification における分類構造の粒子増枚数と得られた構造	
B2 2nd - 10th Classification 分類構造と各構造の類似度分布、重み分布	
B3 各 Classification ごとに得られた GDH ドメイン運動の自由エネルギー地形	
B4 各 Classification 分類構造における 89 番トリプトファン残基ポテンシャルマップ	
付録 C cryoEM 観察像からの自由エネルギー地形解析方法	197
謝辞	200

略語一覧

BIM: beam induced motion

cryoEM: Cryo-electron Microscopy / クライオ電子顕微鏡法

CTF: Contrast Transfer Function / コントラスト伝達関数

DR: Dummy Residue / ダミー残基

DSA: Dummy Solvent Atom / ダミー溶媒原子

FMN: flavin mononucleotide

FSC: Fourier Shell Correlation

FTIR: Fourier Transform Infrared Spectroscopy / フーリエ変換赤外分光

GAF: cGMP phosphodiesterase/adenyl cyclase/FhlA

GDH: Glutamate Dehydrogenase / グルタミン酸脱水素酵素

GS-FSC: gold-standard Fourier Shell Correlation

HKRD: Histidine kinase-related domain

LOV: Light-Oxygen-Voltage Sensing

MD: Molecular Dynamics / 分子動力学

Mg-ATP: Adenosine TriPhosphate Magnesium

NMR: Nuclear Magnetic Resonance

NSD: Normalized Spatial Discrepancy

OPM: output module

PAS: Per-Arnt-Sim

PDB: Protein Data Bank / 蛋白質立体構造データベース

PDI: Protein Disulfide Isomerase

PCA: Principal Component Analysis / 主成分分析

phy: phytochrome

PKA: Protein Kinase A

PSM: photosensory module

P1L1 phototropin 1 LOV domain

phot: phototropin

PΦB: phytochromobilin

R/FR: Red/Far-Red

SAXS: Small-Angle X-ray Scattering / X線小角散乱法

STK: Serine/Threonine Kinase

S/N 比: Signal to Noise Ratio / 信号対雑音比

1 章 序論

1.1 生命現象と蛋白質

生命の基本単位は細胞である。脂質二重膜によって外界と隔てられた細胞内部では主に蛋白質や核酸、内部脂質、そして重量の 4-7 割を占める水で構成されている。さらにこれらの基本的な構成要素が細胞核やミトコンドリアといった細胞内小器官を形成しており、これによって細胞内部は nm から μm までの広範な空間スケールに跨った複雑かつ精緻な構造を有している。この環境内で蛋白質は他の生体高分子との相互作用や触媒作用、刺激応答としての信号変換などによって化学反応、即ち生命活動を制御する役割を担っている。従って生命現象を素過程レベルで正しく理解するためには蛋白質の構造・機能を理解することが不可欠となる。

蛋白質は 20 種類の疎水性あるいは親水性の側鎖を有するアミノ酸のペプチド結合による直鎖状高分子であり、数十から数千に及ぶその配列は核内の DNA に塩基配列として保存されている。「生命の設計図」とも呼ばれる DNA は言い換えれば「蛋白質の設計図」である。DNA 配列を基に生合成された蛋白質一次元鎖は、溶液中において静電相互作用や van der Waals 相互作用、水素結合、疎水性相互作用によって自由エネルギーが最小となるよう折り畳まれ、アミノ酸配列に特有な三次元立体構造を形成する(図 1.1a)。この折り畳みによって、蛋白質はその機能の発現に不可欠な相互作用サイトや触媒サイトを獲得する(図 1.1b)。また、蛋白質の立体構造は決して静的でなく、熱振動による揺らぎの中で複数の構造(コンフォメーション)間を遷移し、機能発現に伴って構造変化を生じる動的なものであり、構造揺らぎに内在する大規模運動と機能発現に伴う構造変化との間に強い相関があるという報告もなされている [Oroguchi *et al.*, 2009]。

蛋白質が水和環境下にあることは極めて重要である。そもそも、溶液中の蛋白質は溶媒和自由エネルギーに従って疎水性側鎖が表面に露出しないよう折り畳まれるため、その立体構造形成に周辺水分子との相互作用が不可欠であることは明らかである。また、蛋白質表面には親水性側鎖と水素結合によって特異的に相互作用する水和水分子の層が形成されるが(図 1.1a)、これら水和水との相互作用が蛋白質の機能発現を支配していることが現在までの研究から示唆されている[Bhat *et al.*, 1994; Luecke *et al.*, 1999; Nakasako *et al.*, 1999; Zhang *et al.*, 1999; Zhou *et al.*, 2001]。例えば超高度好熱菌由来グルタミン酸脱水素酵素(Glutamate Dehydrogenase: GDH)では、触媒作用に関わる活性クレフトの開閉運動と協奏してクレフト内部の局所的な水和構造変化が生じていることが複

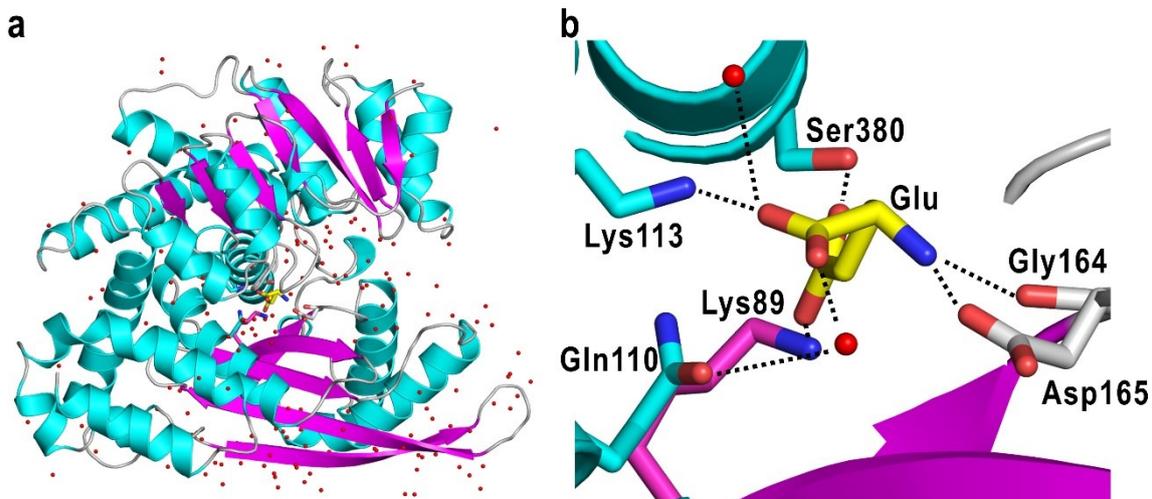


図 1.1 蛋白質立体構造と相互作用サイト。(a) *Clostridium symbiosum* 由来グルタミン酸脱水素酵素のグルタミン酸結合立体構造[Stillman *et al.*, 1993; PDB ID: 1BGV]。二次構造のうちらせん構造である α ヘリックスをシアン、直鎖構造である β ストランドをマゼンタ、その他のループ構造を白で着色している。蛋白質周囲の赤い球は水和水分子、蛋白質中心の黄色の分子は相互作用サイトと結合しているグルタミン酸である。(b) 相互作用サイト拡大図。基質であるグルタミン酸と相互作用している残基をスティック表示している。N 原子、O 原子はそれぞれ青、赤で着色しているが C 原子は各残基の形成する二次構造の色で着色している。また、水素結合を破線で表示している。本図は分子グラフィックスソフトウェア PyMOL によって作成した。

数手法によって確認されている[Nakasako *et al.*, 2001; Oroguchi & Nakasako, 2016]。このように蛋白質は水和環境なくして成り立たず、その機能発現メカニズムを明らかにするためには溶液環境下で測定・解析しなければならない。

1.2 蛋白質構造・動態解析の概略

蛋白質の立体構造解析において最も代表的な手法が X 線を用いた結晶構造解析法である[Drenth, 2007]。X 線構造解析においては、入射 X 線が試料中の電子によって弾性散乱されることで生じる X 線回折パターンを取得・解析する。蛋白質 1 分子からの X 線回折像の取得は散乱断面積が小さいために現状不可能であるが、多数の蛋白質分子が結

晶化による周期的な三次元配列を形成することで Bragg 回折が生じ、高い信号対雑音比 (Signal to Noise Ratio: S/N 比)での測定が可能となる。また、蛋白質結晶は van der Waals 相互作用や水素結合によって形成されており、結晶体積の 30-90%を溶媒が占めていることが大きな特徴である。このため、結晶中の蛋白質は多くの場合溶液中と同様の立体構造をとっていると考えられる。結晶構造解析法の最大の利点はその分解能の高さであり、標的蛋白質の良質な結晶さえあれば原子モデリングが可能な分解能の 3次元電子密度が取得できる。また、半世紀以上にわたって利用されてきた手法であるため、実験・解析環境が非常に洗練されていることも大きな長所である。蛋白質立体構造データバンク (Protein Data Bank: PDB)には現在(2019年10月)までに解析された 15 万以上の蛋白質や核酸、それら複合体の立体構造が登録されているが、そのうち 9割弱が結晶構造解析法による構造であり、蛋白質立体構造は基本的に結晶構造解析によって得られるものと言っても過言ではない。

結晶構造解析法によってこれまでに多くの蛋白質の立体構造が明らかにされてきたが、上でも述べた通り蛋白質は構造揺らぎを持った柔らかな存在であるため、結晶中の静的立体構造だけではなく、水溶液中の動態についても解析しなければならない。これに関して結晶構造解析では、例えばリガンド存在下、非存在下のように結晶化条件を変えた構造解析や、結晶学的非対称単位に複数分子が存在する場合の各分子間の構造比較 [Nakasako *et al.*, 2001; Stillman *et al.*, 1993]などが行われてきた。また、蛋白質結晶に溶媒が多量に含有されていることに着目し、湿度制御によって溶媒含有量を変化させることで、結晶中における蛋白質分子の充填形態制御とそれに伴うコンフォメーション変化の観測を試みた研究も存在する [Takayama & Nakasako, 2011]。しかしながら、そもそもの話ではあるが結晶構造解析を適用できる蛋白質は当然結晶が得られた蛋白質に限られるうえ、標的蛋白質に与えられる摂動はリガンド添加などに限定されてくる。また、生体内とはかけ離れた環境である結晶化溶媒中での構造のみが得られ、分子パッキングが本来の動態を阻害することも考えられる。もう一つの重要な問題は得られる各コンフォメーションについて単純な構造の比較しかできないという点である。例えば、各コンフォメーションの溶液中における出現頻度、即ち相対自由エネルギーや構造変化経路などに関する情報は基本的に得られない。これらの問題を踏まえると、蛋白質動態解析手法は以下の条件を満たすものが望ましい。

- ・非結晶条件下の試料を扱うことが可能
- ・測定系に対して外部摂動を容易に加えることが可能

・得られる動態について出現頻度などの詳細な解析が可能

本研究では、これらの要求を満たす蛋白質動態解析法として X 線小角散乱法、及びクライオ電子顕微鏡法に着目した。

1.3 本研究で用いる蛋白質構造・動態解析法

1.3.1 X 線小角散乱法

X 線小角散乱法(Small-Angle X-ray Scattering: SAXS)は溶液試料に X 線を照射し、その散乱パターンを測定する手法である[Guinier & Fournet, 1955; Feigin & Svergun, 1987] (図 1.2)。散乱パターンは試料構造からの Fraunhofer 回折に由来するため、得られたパターンを解析することで試料構造について情報を得ることができる。蛋白質の散乱断面積を考慮すると、高輝度放射光施設を利用した SAXS 測定であっても数 mg/mL の試料濃度が要求される。この時、照射 X 線のビームサイズを数百 μm 、試料厚さを数 mm とすると X 線照射領域中には 10^{14-15} 個程度の蛋白質が存在する。これらの蛋白質は照射 X 線に対して様々な配向を取るため、散乱パターンが蛋白質配向について平均化される。そのため、得られる散乱パターンは試料蛋白質について配向平均された球対称な構造因子に由来する等方的パターンとなり、蛋白質溶液の SAXS 測定から得られる構造情報は分子サイズや試料形状の異方性といった構造パラメータになる。これは結晶構造解析法のような他の蛋白質構造解析手法で得られる原子レベルの構造情報と比較して著しく低分解能である。しかしながら、SAXS では急速凍結などを行わずに溶液そのまま測定するため、光照射や温度変化などの外部摂動を容易に加えられる。また、試料セルとしてフローセルを用いることで、リガンドとの混合や pH 変化といった外部摂動に対する蛋白質応答を時分割測定することも可能である。SAXS はこのように、外部摂動による蛋白質機能発現の誘発とそれに伴う構造変化を散乱パターン変化として容易に測定できるといふ他に代わるもののない特徴がある。

先に述べた通り、蛋白質溶液試料について SAXS 測定から得られる散乱パターンは等方的である。実際の解析においては測定データの S/N 比を向上させるため円環平均を行うため、得られるデータは一次元の散乱プロファイルである。この散乱プロファイルからより高分解能の構造情報を抽出するために様々な試みが行われた。1970 年には散乱プロファイルを球面調和関数で展開し、この球面調和関数を Fourier 変換することで全体形状を推定する手法が提案された[Stuhrmann, 1970]。また、散乱プロファイルの形状から試料蛋白質の折れ畳み度を評価する Kratky プロットという手法も存在する

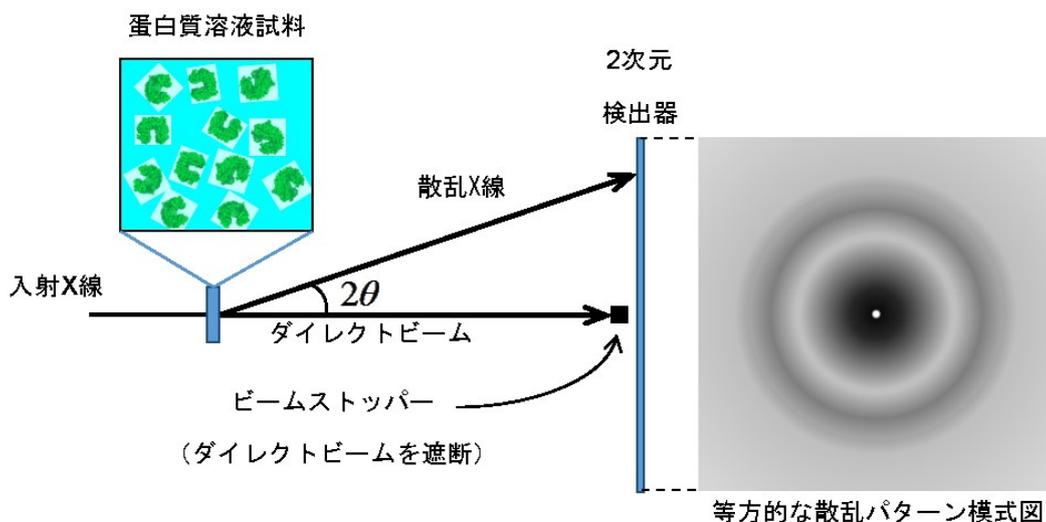


図 1.2 SAXS 模式図。入射 X 線の大部分は溶液試料を透過するため、透過 X 線は非常に高強度である。この強力な透過 X 線による検出器の損傷を避けるため、透過 X 線はビームストッパーによって遮断される。このため得られる散乱パターンからは小角領域が欠損している。

[Svergun *et al.*, 2013]。これら手法の中で現在主流となっている解析手法が 1990 年代後半から 2000 年代前半にかけて開発された非経験的分子形状モデリングと呼ばれる手法である[Svergun, 1999; Svergun *et al.*, 2001]。この手法では蛋白質を球の集合として粗視化し、実験で得られている SAXS プロファイルを説明できるように球の配置をランダムに変えつつ最適な球の配置を探っていく。このようにして蛋白質形状を球の集合によって表現することがこの手法の特徴である。また、この手法が普及した理由の一つに、解析が極めて容易であることが挙げられる。指定の必要なパラメータが極めて少なく、パラメータ入力から後は自動で推定形状が出力される。この扱いの容易さも相まって非経験的分子形状モデリングは SAXS による蛋白質構造解析においては標準的な解析ステップとなっている。

1.3.2 クライオ電子顕微鏡法

クライオ電子顕微鏡法(Cryo-electron Microscopy: cryoEM)は数百 kV の高電圧により加速された電子線を試料に照射し、透過投影像を実像で観察する手法である[Frank, 2006] (図 1.3)。電子-電子あるいは電子-原子核相互作用は極めて強く、蛋白質一分子の直接

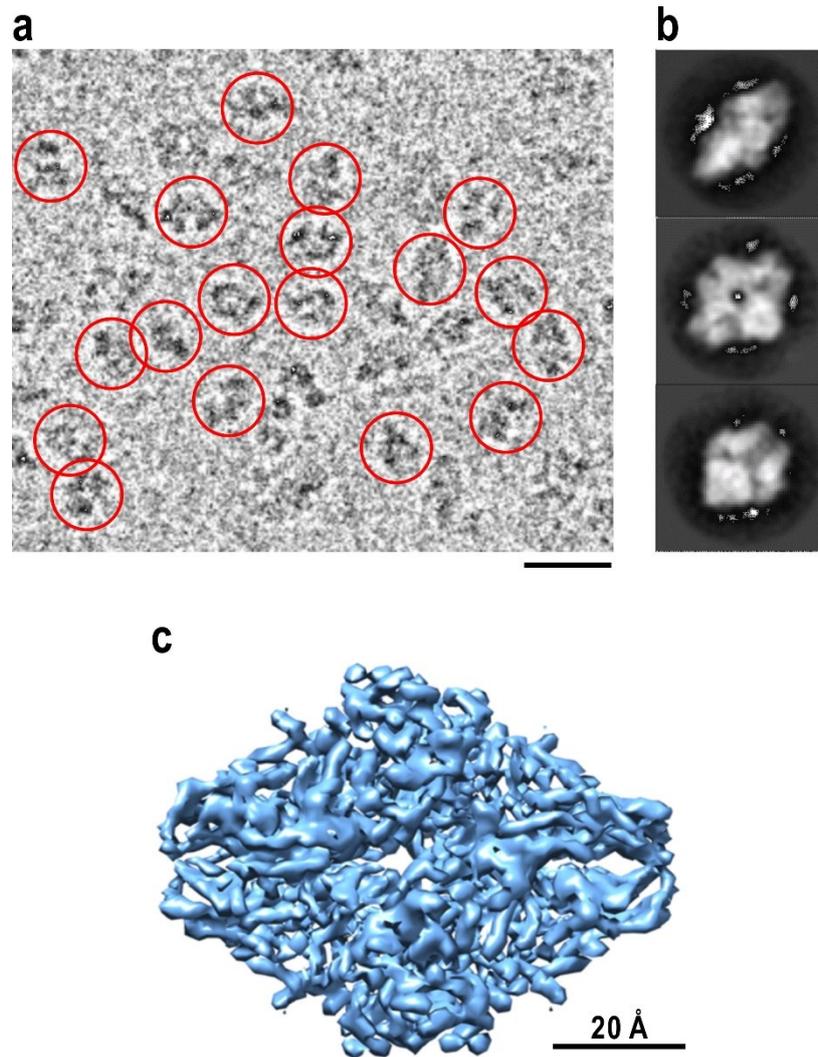


図 1.3 cryoEM による単粒子解析概略図。(a)大腸菌由来 β -galactosidase の cryoEM による観察画像。赤丸で囲まれた黒い像が蛋白質粒子像である。画像右下のスケールバーは約 100 Å。(b) 切り出された粒子像の投影方向別平均像。画像の横幅は約 100 Å である。(c)三次元再構成によって得られた β -galactosidase 3次元マップ。

観察を可能にする一方で様々な問題が生じる。まず電子線パスは空気による吸収・散乱を防ぐために試料周囲も含めて高真空である必要がある。試料厚みについても制限され、観察可能な厚みは 100 nm 程度が上限となる。また、溶液試料については電子線照射によって水分子から発生した酸素ラジカルが溶液内に拡散し、試料蛋白質に損傷を与えることが知られている。これらの問題を克服するために cryoEM では $\sim 1\mu\text{m}$ 径の孔に蛋白

質溶液を薄膜状に張り、急速凍結して観察する。液体窒素温度付近まで冷却した液体エタンを冷媒として用いるとその温度変化速度は 10^4 K/sec に到達するため、この速度で凍結することで蛋白質溶液は非晶質氷になり、溶液中での構造を保った氷包埋蛋白質の観察を可能にしている [Dubochet *et al.*, 1988]。

cryoEM で得られる構造情報は蛋白質一分子毎の投影像であるため、各粒子像は三次元構造についての情報を欠損している。また、低温環境下で電子線照射による酸素ラジカルの拡散は低減されるものの、試料損傷は皆無ではない。現在までの研究で明らかになっている蛋白質構造を致命的に破壊しない照射線量の限界はおよそ $20 \text{ e}/\text{\AA}^2$ 程度であり [Grant & Grigorieff, 2015]、この照射線量で得られる観察像は図 1.3a のように極めて S/N 比の悪い像である。このため、cryoEM では概して単粒子解析法と呼ばれる解析法が用いられる [Frank, 2006]。氷包埋された蛋白質分子は観察方向に対して各々自由な配向を取っていると考えられる。同一配向からの投影像を平均化することで S/N 比を向上させると共に、様々な配向からの投影像を基に三次元構造を再構成することが単粒子解析法の基本的な原理である (図 1.3b, c)。

2000 年代まで単粒子解析法による到達分解能はウイルスのような空間対称性の高い粒子を除いて 10 \AA 程度しかなかった。しかしながら、主に電子線検出器の劇的な性能向上により到達分解能が 2 \AA を超える構造解析の報告もなされるようになり [Merk *et al.*, 2016]、現在では cryoEM による単粒子解析法は結晶構造解析法などと並ぶ蛋白質構造決定法の一つとして位置付けられている。凍結時の溶液条件を変えることで構造変化を解析することも可能である。また、cryoEM の「蛋白質粒子像を一分子レベルで観察する」という特徴は蛋白質動態解析において他にないメリットである。氷包埋された蛋白質分子はその一つ一つが構造揺らぎの中のコンフォメーションに固定されると考えられるため、膨大量の粒子像を集めることで蛋白質の動態についての情報が得られる。従来の単粒子解析法において粒子像に含まれる構造揺らぎは分解能を妨げる要因でしかなかったが、到達分解能が向上し、解析アルゴリズムも洗練された昨今においてはむしろこの構造揺らぎを解析・可視化する試みが活発になされている [Bai *et al.*, 2015; Dashti *et al.*, 2013; Haselbach *et al.*, 2018; Nakane *et al.*, 2018]。

1.4 本研究で用いる解析法の問題点

1.3 節で説明した通り、結晶構造解析と比較して SAXS 及び cryoEM はより広範な試料に適用でき、より詳細に蛋白質動態を解析できる手法である (表 1.1)。しかしながら、

表 1.1 各解析手法の比較

	X 線結晶構造解析	SAXS	cryoEM
試料制限	結晶化の必要あり	特になし	特になし
試料の状態	結晶	溶液	溶液
到達分解能	原子レベル	分子概形～ ドメイン相対配置	原子レベル
構造変化の測定	リガンド結合 パッキング制御	光受容 リガンド結合 pH 変化への応答 温度変化	凍結時の条件変化 構造揺らぎ

蛋白質構造・動態解析に適用する上で両者ともに問題点を抱えている。

1.4.1 SAXS における問題点: 概形推定における不確かさ

SAXS による構造解析において、非経験的分子形状モデリングは極めて強力な手法であるが、概形推定に原理的な不確かさが生じるという大きな問題点が存在する。SAXS プロファイルは配向について平均化された形状に由来するため、蛋白質形状と SAXS プロファイルは 1 対 1 に対応せず、それぞれ異なった形状を持つ複数の粒子から同一形状の SAXS プロファイルが得られることがシミュレーション計算から報告されている [Petoukhov & Svergun, 2015]。従って、非経験的分子形状モデリングにおいて一度の解析から試料蛋白質形状を推定することは実質的に不可能である。この問題に対して、複数回のモデリングを行い、明らかに形状の異なる外れ値を除外した上で形状平均化を行うという手法が当初提案されたが [Volkhov & Svergun, 2003]、平均形状が尤もらしい形状とはなりえないため有効な手立てとはならなかった。階層的クラスタリングによって形状分類をする手法 [Petoukhov *et al.*, 2012] など提案されているものの、確立された手立ては現状存在しない。

1.4.2 cryoEM における問題点: 動態解析法の検討

cryoEM を利用した動態解析はここ数年で検討がなされるようになってきたトピックであり、手法としては未だ未成熟である。最も広く使われている手法が三次元構造を複

数構造に分類する手法である [Scheres, 2012; Bai *et al.*, 2015, Punjani *et al.*, 2017]。しかしながらこの手法は連続的な構造変化に弱く、また分類自体も不完全である場合が多い。連続的な構造変化を扱える手法も複数発表されているものの、これらの手法は散乱断面積が大きく像コントラストが良い、比較的巨大(分子量 ~数 MDa)な蛋白質や蛋白質複合体を対象としている [Schilbach *et al.*, 2018; Nakane *et al.*, 2018; Dashti *et al.*, 2014; Haselbach *et al.*, 2018]。これは主として、数百 kDa 程度の通常の蛋白質では像コントラストが悪く各粒子像のコンフォメーションの差異を比較することが困難なためである。従って、動態解析としての cryoEM の適用範囲を広げるためには、生体内での存在量も多く、生命現象の素過程を担っている数百 kDa 程度の蛋白質の連続的な構造変化を解析可能な手法の検討が必須である。

1.5 本研究の概要と章立て

本研究では蛋白質動態解析において SAXS 及び cryoEM の持つポテンシャルに注目し、多くの蛋白質に適用可能な動態解析手法を確立するべく 1.4 節に挙げた問題点の解決に取り組んだ。1.4.1 節で述べた課題に対しては多数の推定形状モデルから多変量解析によって形状の傾向を抽出・分類し、そこから尤もらしい形状を推定するという手法を開発した [Oide *et al.*, 2018a]。この際、電子顕微鏡との組み合わせによって形状推定の信頼度を高めるという試みも行った [Oide *et al.*, 2018b; Oide *et al.*, 2019]。また、1.4.2 節で述べた問題点については分子動力学(Molecular Dynamics: MD)シミュレーションとの組み合わせによって数十 kDa 程度の分子量での蛋白質動態を自由エネルギー地形として可視化する手法を開発した。

本論文では、まず SAXS の諸原理を 2 章で解説し、本論文の骨子となる多変量解析を用いた蛋白質形状推定手法の開発(3 章)と提案手法の応用である植物光受容蛋白質フォトトロピン(4 章)・フィトクロム B(5 章)の解析について記す。その後、6 章で cryoEM の諸原理を説明し、7 章で cryoEM と MD シミュレーションを組み合わせた蛋白質動態解析手法の開発と GDH ドメイン運動解析への適用を報告する。最後に、8 章にて研究の総括と今後の展望について述べる。

2章 X線小角散乱法

本章では X 線小角散乱法に関わる諸原理について説明する。初めに電磁波の放射現象に触れ、次いで電磁波の散乱現象及び X 線小角散乱法の原理、X 線小角散乱法から得られる実験データの解析法を述べる。

2.1 電磁波の放射 [砂川, 1999]

以下では真空中に局在して分布する電荷を考える。電荷位置が時間変化しないとき、電荷からは電気力線が伸びて静電場を形成している。電荷位置が激しく変動すると、電気力線が歪み、電場の時間変化として遠方へと伝播していく(図 2.1)。これが電磁波放射の直観的な説明である。運動する点電荷によって生じるポテンシャルは Liénard-Wiechert ポテンシャルとして定式化され、 \bar{x} に位置する点電荷から十分遠方の位置 \bar{r} に伝播する電場及び磁場は

$$\vec{E}(\bar{r}, t) = \frac{e}{4\pi\epsilon_0} \frac{\bar{n}(t'_0) \times \left\{ (\bar{n}(t'_0) - \vec{\beta}(t'_0)) \times \frac{d\vec{\beta}(t'_0)}{dt} \right\}}{c\alpha^3(t'_0)R(t'_0)} \quad (2.1)$$

$$\vec{B}(\bar{r}, t) = \frac{1}{c} \bar{n}(t'_0) \times \vec{E}(\bar{r}, t) \quad (2.2)$$

の形で与えられる。ただし、

$$R(t) = |\bar{r} - \bar{x}(t)| \quad (2.3)$$

$$\vec{\beta}(t'_0) = \frac{1}{c} \frac{d\bar{x}(t'_0)}{dt} \quad (2.4)$$

$$\alpha(t'_0) = 1 - \frac{1}{c} \bar{n}(t'_0) \cdot \frac{d\bar{x}(t'_0)}{dt} \quad (2.5)$$

である(図 2.2, 2.3)。また、 t'_0 は

$$t'_0 = t - \frac{|\bar{r} - \bar{x}(t'_0)|}{c} \quad (2.6)$$

の解として与えられる。式(2.2)より、 $\vec{E}(\bar{r}, t)$ と $\vec{B}(\bar{r}, t)$ は互いに直交しており、さらに $\bar{n}(t'_0) \cdot \vec{E}(\bar{r}, t) = 0$ である。従って、放射される電磁場は電磁波として振る舞うことが分かる。

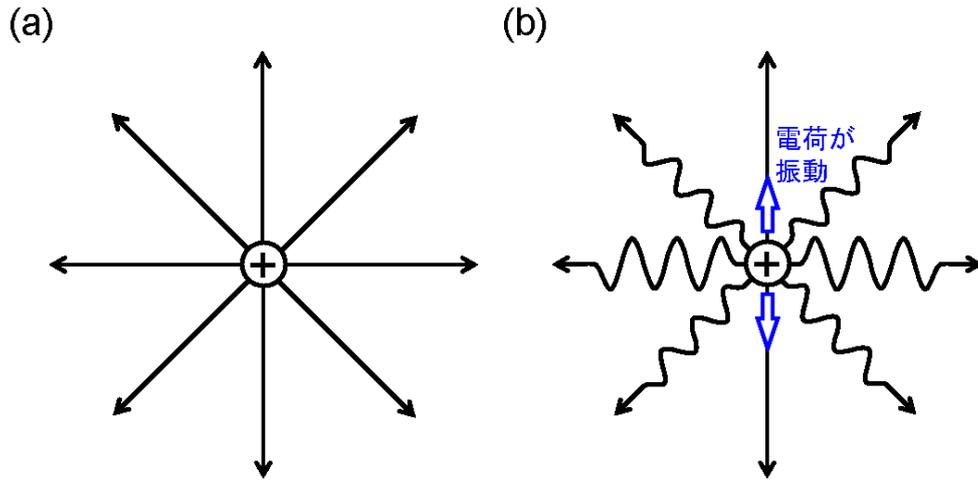


図 2.1 電荷の位置変動による電磁波の放射。静止した正電荷での電気力線(a)と振動する正電荷での電気力線(b)。図 2.1b では青矢印方向に正電荷が振動している。電荷の振動によって電気力線の始点位置が変化することで電場が振動し、遠方へ伝播していく。但し、振動方向へは電磁波が伝播せず、従って放射には角度依存性が存在することを記しておく。

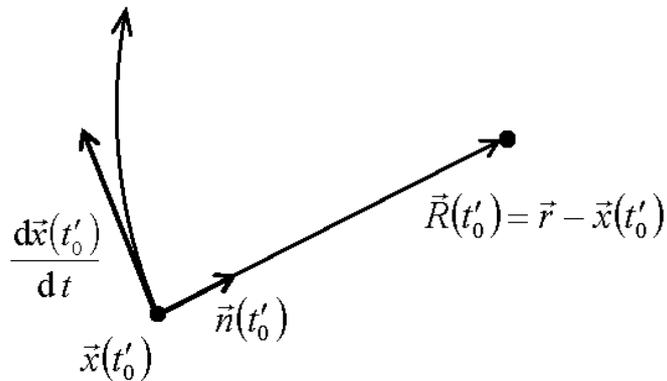


図 2.2 運動する点電荷と観測点との位置関係。時刻 t'_0 において点電荷は $\vec{x}(t'_0)$ に存在する。この時の観測点 \vec{r} との距離は $R(t'_0) = |\vec{r} - \vec{x}(t'_0)|$ となる。また、 $\vec{R}(t'_0)$ 方向の単位ベクトルが $\vec{n}(t'_0)$ である。 $\vec{\beta}(t'_0)$ は点電荷の速度ベクトルを光速で規格化したものであり、相対論的效果を取り入れる際に都合の良い表式である。因子 $\alpha(t'_0)$ については図 2.3 において詳しく説明する。

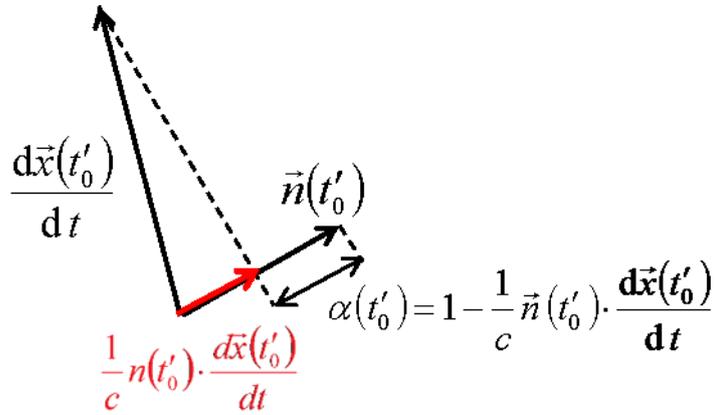


図 2.3 点電荷の運動による振幅の変化因子 $\alpha(t'_0)$ の模式図。この因子は点電荷の速度が観測点方向と直交している場合は常に 1 である。それ以外の場合、点電荷が静止しているときには 1 であるが、点電荷の速度が光速に近づくほど 0 に近づいてゆき、その傾向は点電荷の速度が観測点方向である時最も著しい。このように、 $\alpha(t'_0)$ は観測点での電磁波の振幅の、点電荷の運動方向に対する角度依存性を記述する因子である。

2.2 X 線散乱

2.2.1 Thomson 散乱 [Als-Nielsen & McMorrow, 2011]

2.1 節において、加速された点電荷から電磁波が放射されることを示した。これは、点電荷に外部から電磁波が入射すると、その電場によって点電荷が加速され、入射波を模した電磁波が放射されることを意味する。この素過程についての知見が無ければ点電荷によって電磁波が散乱される現象とみなすことができる。点電荷の微分散乱断面積は放射電磁波の Poynting ベクトルを用いた議論から

$$\frac{d\sigma}{d\Omega} = \frac{1}{(4\pi\epsilon_0)^2} \frac{e^2}{c^4 E_0^2} \left| \vec{n} \times \left(\vec{n} \times \frac{d^2 \vec{r}(t)}{dt^2} \right) \right|^2 \quad (2.7)$$

で与えられる。これを踏まえて最も単純な例である Thomson 散乱、即ち自由電子による弾性散乱について考える。後にもう一度触れるが、X 線散乱実験で観測しているのは測定試料を構成する原子が持つ電子による Thomson 散乱である。ここでは古典的な取

り扱いをしているため散乱は必ず弾性的であるが、量子力学的には散乱は必ずしも弾性的ではなく非弾性散乱をも考慮する必要がある。しかし、X線散乱実験に利用されるX線のエネルギーは~10 keV程度であり、このエネルギー領域では散乱は弾性的であるとして構わない。

1個の自由電子に平面波である電磁波が入射するときを考える。 \vec{n} を電子から観測点方向へ向く単位ベクトル、 \vec{j} を電磁波の偏り方向への単位ベクトル、入射波の波数ベクトルを \vec{k} とすると、これらのベクトルは図2.4のような位置関係になる。電磁波の振幅を E_0 、電気素量を e とすると、電子の運動方程式は

$$m \frac{d^2 \vec{r}(t)}{dt^2} = -e \vec{j} E_0 \exp[i(\vec{k} \cdot \vec{r}(t) - \omega t)] \quad (2.8)$$

と与えられる。ここから加速度を求め、(2.7)式に代入することで微分散乱断面積は

$$\frac{d\sigma}{d\Omega} = \frac{1}{(4\pi\epsilon_0)^2} \left(\frac{e^2}{mc^2} \right)^2 (\vec{n} \times (\vec{n} \times \vec{e}))^2 = \left(\frac{e^2}{4\pi\epsilon_0 mc^2} \right)^2 \sin^2 \Theta \quad (2.9)$$

という形となる。ただし Θ は図2.4の通り \vec{j} と \vec{n} とがなす角である。特に入射波が偏っていない場合には \vec{k} と \vec{n} がなす角 θ によって

$$\frac{d\sigma}{d\Omega} = \frac{1}{2} \left(\frac{e^2}{4\pi\epsilon_0 mc^2} \right)^2 (1 + \cos^2 \theta) = \frac{r_e}{2} (1 + \cos^2 \theta) \quad (2.10)$$

という形になる。式の通り、Thomson散乱の散乱断面積は θ に依存する。しかし、 θ が小さい領域では角度依存項は定数とみなせるため、Thomson散乱による散乱波は球面波であると近似してよい。この領域では $\Theta \cong \pi/2$ となるため、入射波が偏っている場合でも同様の近似が成り立つ。定数項 r_e は古典電子半径(Thomson散乱長)と呼ばれる量であり、およそ 2.82×10^{-15} mの値をとる。古典電子半径は電子による散乱の実際の強度を与えるものであるが、この値が非常に小さいためX線散乱においては高輝度のX線源が重要となる。X線散乱において散乱源となるのは物質中の電子であり、即ち物質によるX線の散乱はThomson散乱によって記述される。物質中の電子は原子核に束縛されており自由電子ではないが、X線のエネルギーはこの束縛エネルギーに比べて十分大きいいため、自由電子による散乱として扱うことができる。

2.2.2 X線散乱と電子密度・構造因子 [Svergun *et al.*, 2013]

図2.5のような形状の電子分布を持つ粒子によるX線の散乱を考える。この粒子の位置 \vec{r} における電子密度を $\rho(\vec{r})$ とする。この時、位置 \vec{r} での単位素片の散乱断面積は、こ

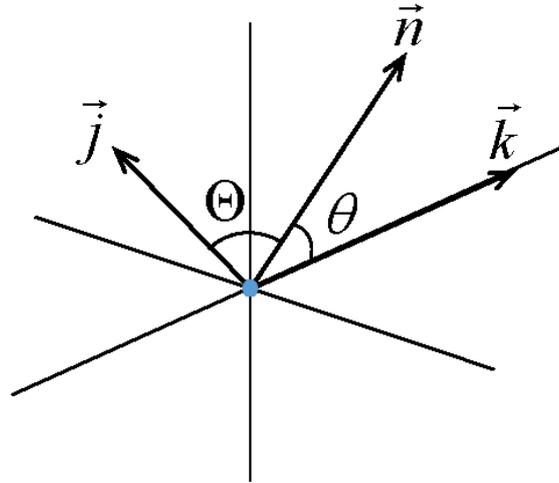


図 2.4 Thomson 散乱におけるベクトルの位置関係。図中においては位置関係がわかりやすくなるように \vec{k} を基準にした直交軸を設けている。

の $\rho(\vec{r})$ に比例すると考えられる。この粒子に波数ベクトル \vec{k} の波面の揃った X 線が入射し、 \vec{k}' の方向で散乱 X 線を観測する場合を考える。Maxwell 方程式の線形性より、粒子全体による散乱波は粒子内の各点による Thomson 散乱波の重ね合わせで表される。従って、粒子内に適当な原点 o をとり、 o に対する位置 \vec{r} の点の微小体積要素を $d^3\vec{r}$ とすると、粒子による散乱振幅に対する微小体積要素の寄与は $\rho(\vec{r})d^3\vec{r}$ と表すことができる。さらに、この体積要素の原点に対する位相のずれは

$$2\pi \vec{S} \cdot \vec{r} = (\vec{k}' - \vec{k}) \cdot \vec{r} \quad (2.11)$$

で表される。 $\vec{S} = (\vec{k}' - \vec{k})/2\pi$ で定義される \vec{S} を散乱ベクトルと呼ぶ。 \vec{k} と \vec{k}' がなす角を散乱角と呼び、これを通常 2θ で表す。この時、 $|\vec{S}| = 2 \sin \theta / \lambda$ となる。

ここから、粒子による \vec{S} に対する全散乱長 $\vec{F}(\vec{S})$ は

$$\vec{F}(\vec{S}) = \int \rho(\vec{r}) \exp(2\pi i \vec{S} \cdot \vec{r}) d^3\vec{r} \quad (2.12)$$

で与えられる。この $\vec{F}(\vec{S})$ を構造因子と呼ぶ。 $\vec{F}(\vec{S})$ は複素数であるため、散乱された X 線についてその振幅と位相を測定できれば試料粒子についてその電子分布、即ち構造を決定できる。しかしながら、X 線の周波数は 10^{18} Hz に達するため、現代の測定技術では位相を測定することは不可能である。また、X 線に対する物質の屈折率は 1 より極僅かに小さく、光学顕微鏡における凸レンズに対応する結像素子は存在しない。そのた

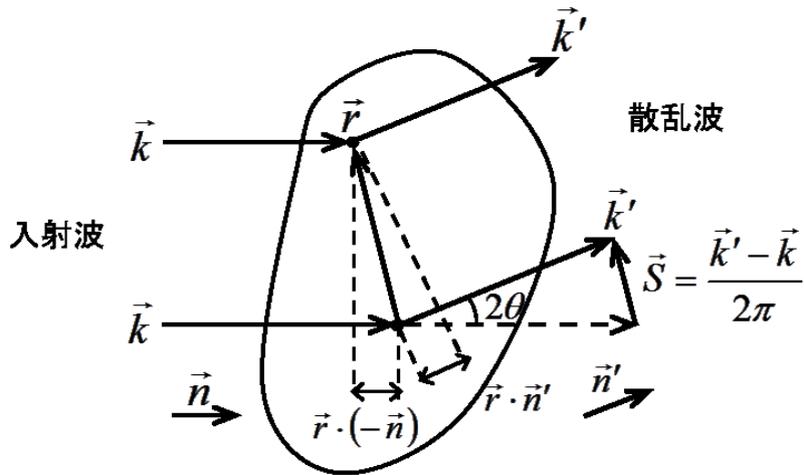


図 2.5 X 線散乱の模式図。2 点間での経路差は図に示されている通りである。入射波と散乱波の波数をそれぞれ $k = \vec{k} \cdot \vec{n}$, $k' = \vec{k}' \cdot \vec{n}'$ とすると、位相差は波数をかければよいため、最終的に式(2.11)の形で位相差が与えられる。

め、実際の実験で測定される物理量は散乱波の強度と見なせる。散乱強度 $I(\vec{S})$ は

$$\begin{aligned}
 I(\vec{S}) &= I_0 r_e^2 K \cdot \vec{F}(\vec{S}) \cdot \vec{F}^*(\vec{S}) \\
 &= I_0 r_e^2 K \cdot \int_{\vec{r}'} \int_{\vec{r}} \rho(\vec{r}) \rho(\vec{r}') \exp[2\pi i \vec{S} \cdot (\vec{r} - \vec{r}')] d^3 \vec{r} d^3 \vec{r}'
 \end{aligned}
 \tag{2.13}$$

として与えられる。ここで、 I_0 は入射 X 線強度、 K は I_0 、 r_e を除く種々の係数をまとめた係数である。従って、X 線散乱の散乱強度は試料粒子の電子密度分布を Fourier 変換し、複素自乗した形で与えられる。そのため、Thomson 散乱を考慮し、散乱強度について解析をすれば試料の電子密度分布、即ち形状や内部構造についての情報を得ることができる。

X 線散乱において散乱ベクトルの大きさと分解能との対応を与える式として X 線結晶構造解析で使用される Bragg の式が知られている。図 2.6 のような周期構造による回折を考えたとき、回折角 θ と距離 d に

$$2d \sin \theta = n\lambda \tag{2.14}$$

なる関係が成り立つ。 $n=1$ としてこの式を変形させると、

$$\frac{1}{d} = \frac{2 \sin \theta}{\lambda} = \frac{2k \sin \theta}{2\pi} = S \tag{2.15}$$

となる。従って、散乱ベクトルの逆数をとればある S における散乱強度はどの程度のス

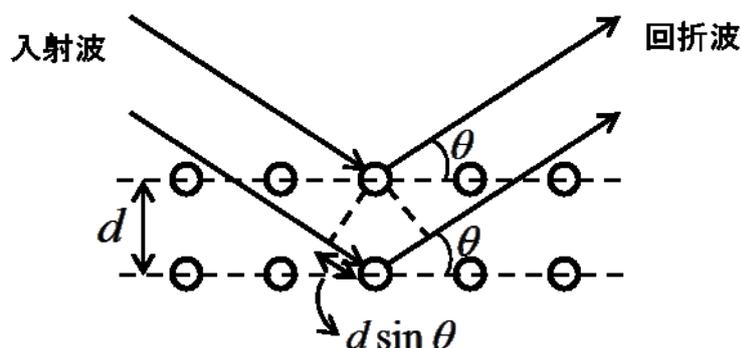


図 2.6 周期構造における入射-回折 X 線の経路図。図のように 2 つの入射波には $d \sin \theta$ の経路差が存在する。回折側においても同様の経路差が存在するため、2 つの波の経路差は $2d \sin \theta$ となる。この経路差が波長の整数倍となれば干渉による回折波が生じる。これを示したのが式(2.14)で与えられた Bragg の式である。

ケールの構造情報を反映しているのかを評価できる。

2.2.3 X 線小角散乱[Guinier & Fournet, 1955; Svergun *et al.*, 2013]

SAXS で測定する散乱角領域は光学系のセットアップや注目する空間スケールによるが、一般に $s = 0.02 \sim 0.05 \text{ \AA}^{-1}$ までを測定する。これは入射 X 線波長が 1 \AA の時、散乱角にして $2\theta = 15 \sim 30 \text{ mrad}$ 程度の領域である。Bragg の式から、この場合に SAXS で得られる構造情報は $> 20 \sim 50 \text{ \AA}$ の距離相関であり、会合状態や慣性半径といった概形情報やドメイン相対配置などが該当する。

蛋白質溶液から得られる散乱パターンは蛋白質粒子と溶媒分子の散乱強度となっており、蛋白質についての散乱プロファイルを取得するには円環平均の後、蛋白質溶液の散乱プロファイルから溶媒のみの散乱プロファイル (バックグラウンド) を差し引く操作が必要になる。また、蛋白質表面と相互作用し水和殻を形成する溶媒分子は、バルク溶媒とは異なった電子密度であり、バックグラウンドとは異なる散乱強度を持つ。従って、バックグラウンド減算で得られる散乱プロファイルは試料電子密度 ρ ではなく試料及び水和殻と溶媒との電子密度コントラスト $\Delta\rho$ 由来のものとなる(図 2.7)。散乱パターンに対して、円環平均、バックグラウンド減算を行い、実際に得られた散乱プロファイルを図 2.8 に示す。得られるデータが 1 次元のプロファイルであるため、SAXS 測

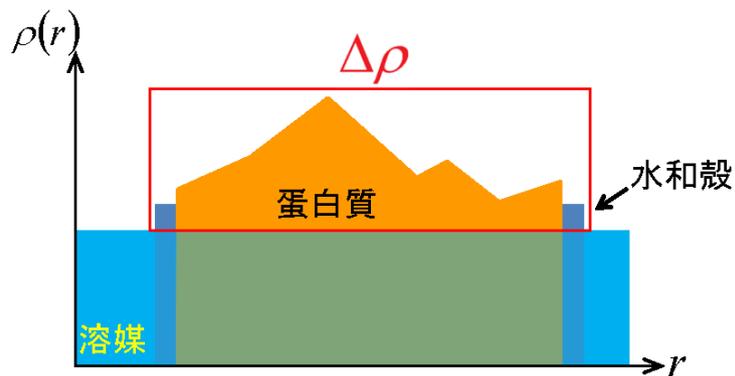


図 2.7 SAXS における蛋白質の電子密度コントラスト。SAXS 測定においては試料溶液散乱プロファイルから純粋溶媒散乱プロファイルを差し引くため、得られる散乱プロファイルは図の赤枠で囲まれた電子密度コントラストについての情報を含んでいる。蛋白質電子密度は純水と比較して大きく、通常の溶媒条件における SAXS 測定ではコントラストについて考慮しなくてよいが、高塩濃度溶媒など溶媒電子密度が高い場合にはコントラストが変化し(コントラスト変調)、得られる構造パラメータが変化することに注意が必要である。

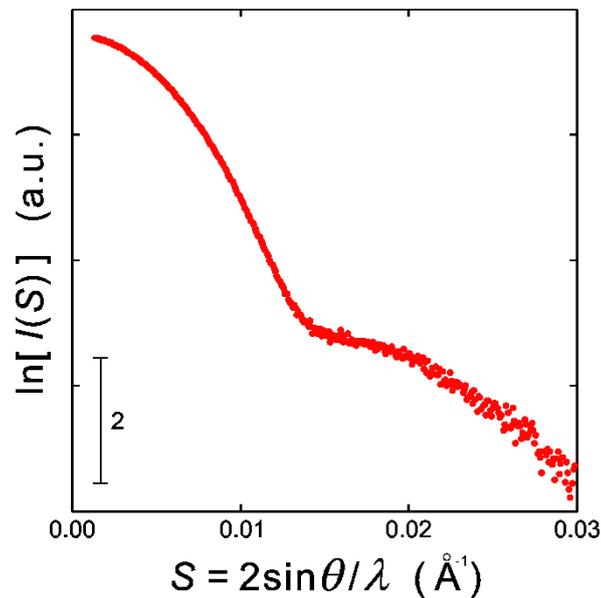


図 2.8 超高度好熱菌グルタミン酸脱水素酵素 1.0 mg/mL 溶液の SAXS プロファイル。 $S < 1.3 \times 10^{-3} \text{ \AA}^{-1}$ ($\sim 0.6 \text{ mrad}$) でのデータ欠損はビームストッパーによるものである。

定では蛋白質構造について深刻な情報欠損が生じているが、この散乱プロファイルから粒子慣性半径や最大分子長など蛋白質構造についての様々なパラメータを得ることができる。以降の節ではこれら SAXS における解析手法について説明する。

2.3 SAXS を用いた構造解析[Guinier & Fournet, 1955; Svergun *et al.*, 2013]

2.3.1 Guinier 近似

2.2.2 節および 2.2.3 節より、溶液散乱における蛋白質散乱強度は

$$I(S) = \frac{1}{4\pi} \int d\Omega \int_{\vec{r}} \int_{\vec{r}'} \rho(\vec{r}) \rho(\vec{r}') \exp[2\pi i \vec{S} \cdot (\vec{r} - \vec{r}')] d^3\vec{r} d^3\vec{r}' \quad (2.16)$$

で与えられる。積分の中の平面波成分について、位相部分 $2\pi i \vec{S} \cdot (\vec{r} - \vec{r}')$ が十分小さい(散乱角が十分小さい)として原点まわりで 2 次まで Taylor 展開し、角度積分を実行すると以下の近似式が得られる。

$$\ln I(S) = \ln I(0) - \frac{4\pi}{3} S^2 R_g^2 \quad \left(S \cdot R_g < \frac{1}{2\pi} \right) \quad (2.17)$$

$$I(0) = \int_{\vec{r}} \int_{\vec{r}'} \rho(\vec{r}) \rho(\vec{r}') d^3\vec{r} d^3\vec{r}' \quad (2.18)$$

$$R_g^2 = \frac{1}{2} \frac{\int_{\vec{r}} \int_{\vec{r}'} \rho(\vec{r}) \rho(\vec{r}') |\vec{r} - \vec{r}'|^2 d^3\vec{r} d^3\vec{r}'}{\int_{\vec{r}} \int_{\vec{r}'} \rho(\vec{r}) \rho(\vec{r}') d^3\vec{r} d^3\vec{r}'} \quad (2.19)$$

ただし、条件として以下の式を課す。

$$\int_{\vec{r}} \rho(\vec{r}) \vec{r} d^3\vec{r} = 0 \quad (2.20)$$

これは電子密度分布の重心を原点とする条件である。散乱強度は並進操作に対して対称であるため、この条件は自然に導入可能である。 $I(0)$ は原点散乱強度、 R_g は慣性半径と呼ばれる量である。この近似は、十分に小角の領域であれば散乱強度は Gauss 関数で近似可能であることを意味する。この近似を Guinier 近似と呼び、転じてこの近似の成り立つ小角領域を Guinier 領域と呼ぶ。(2.17)式に従って横軸に S^2 をとり、縦軸に $\ln I(S)$ をとれば散乱プロファイルは一次関数形となり、その切片および傾きから原点散乱強度、慣性半径を実験的に取得できる。図 2.9 は図 2.8 の散乱プロファイルに対する Guinier プロットであり、小角領域で散乱プロファイルが直線としてよく近似できることが確認できる。ただし、Guinier 近似が成り立つ条件として、溶液が単分散溶液である必要があり、多分散系ではこの近似は成り立たない。従って、試料蛋白質に非特異的

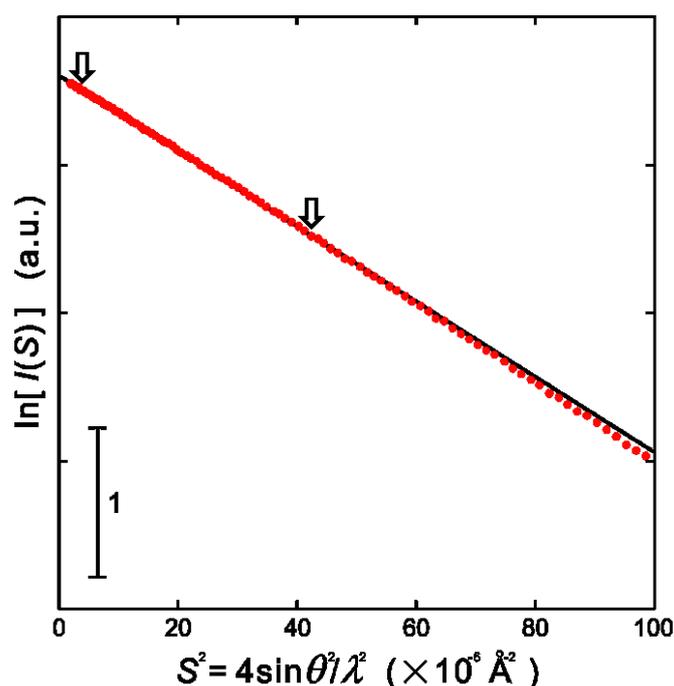


図 2.9 図 2.8 の散乱プロファイルに対する Guinier プロット。黒色の直線はグラフ中の矢印で挟まれた領域での Guinier 近似直線である。この直線の傾きから慣性半径はおよそ 44.3 Å と計算される。

凝集が生じている場合、粒子サイズについて多分散系となるため極小角において Guinier 近似が成り立たない。

原点散乱強度は、本来の測定ではビームストッパーの存在によって直接には得ることのできない値である。式の通り、物理的には総電子数の 2 乗にあたる値であるため、この値について分子量既知の試料と比較することで溶液中での会合状態を評価することができる。また、慣性半径は、電子密度の 2 次モーメントにあたる量である。この量を見ることで分子のサイズ・形状を見積もることができる。Guinier プロットでは、実験値は小角領域で直線近似可能だが、高角になると直線から外れる。これは (2.16) 式の平面波項を展開した際に無視された高次の項の寄与が大きくなるためである。例えば、棒状粒子はモーメントが大きくなるため散乱プロファイルは Guinier 近似による直線の上側に反る。一方、球状粒子では高次項の寄与が小さくなるため直線の下側に反っていく。この振る舞いを見ることによって慣性半径だけでなく試料蛋白質の異方性についても情報を得ることが可能である。さらに、慣性半径はその性質上同じ分子量であっても球状粒子より棒状粒子の方が大きくなるため、慣性半径と概形がわかれば分子量と合わせ

て粒子についての情報を包括的に得ることが可能である。

2.3.2 距離分布関数と分子最大長

X線散乱強度 $I(\vec{S})$ を逆 Fourier 変換すると、次の式が得られる。

$$\int_{\vec{S}} I(\vec{S}) \exp[-2\pi i \vec{S} \cdot \vec{u}] d^3 \vec{S} = \int_{\vec{r}} \rho(\vec{r}) \rho(\vec{r} - \vec{u}) d^3 \vec{r} = P(\vec{u}) \quad (2.21)$$

ここで、 $P(\vec{u})$ は電子密度の自己相関関数である。図 2.10 に電子密度分布と自己相関関数の関係性を模式的に示す。図 2.10 より、自己相関関数 $P(\vec{u})$ は電子密度 $\rho(\vec{r})$ と、 $\rho(\vec{r})$ を \vec{u} だけシフトさせた $\rho(\vec{r} - \vec{u})$ について互いに重なった領域の積分で表される。

溶液散乱実験から得られる散乱プロファイルは散乱ベクトルについて角度平均された一次元データであるため、散乱プロファイルを逆 Fourier 変換して得られる関数は厳密には三次元自己相関関数 $P(\vec{r})$ ではなく距離 r^2 が乗ぜられ配向平均された一次元の関数 $r^2 P(r)$ となる。これは距離分布関数と呼ばれる関数であり、一般的に図 2.11 のような形になる。ここで、距離の増加に応じて緩やかに減少していくのは、電子密度分布に明瞭なエッジが存在しないためである。ここで、距離分布関数がゼロになる $r(\neq 0)$ は、完全に重なっていた位置から全く重ならない位置まで移動した距離であり、これは即ち最大分子長 D_{\max} を与えるものである。従って、散乱プロファイルを逆 Fourier 変換することで最大分子長を得ることができる。しかし、実際には距離分布関数を散乱強度からの逆 Fourier 変換によって直接求めることは不可能である。これはビームストッパーによって小角領域が欠損しており、直接距離分布関数を求めることは非適切問題になるためである。この問題について、経験的パラメータから距離分布関数を推定し、さらに推定距離分布関数の Fourier 変換によって欠損のない散乱強度を計算する GNOM と呼ばれるソフトウェアが開発されている[Svergun, 1992]。このソフトウェアは最大分子長を入力するとその条件下で最適解となる距離分布関数、散乱プロファイルを出力する。ユーザーは、この最大分子長の値を様々に変えて試算し、距離分布関数は滑らかに減衰していくか、散乱強度は実験データとよく一致するか、といったことを手掛かりに最適な最大分子長と距離分布関数を決定する。

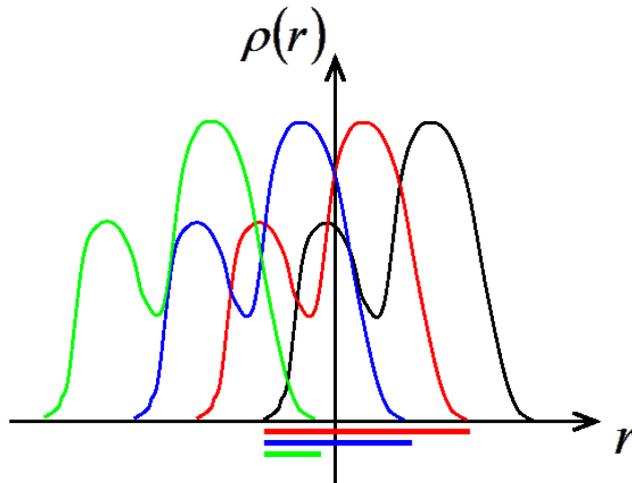


図 2.10 自己相関関数の直観的イメージ。一方の電子密度分布を固定し（黒）、もう一方の電子密度分布について \bar{u} を大きくして移動させていく（赤→青→緑）。この時各 \bar{u} での重なる領域を示したのがグラフ下部の線である。この領域について電子密度の積をとり、積分をすることで各 \bar{u} における自己相関関数の値 $P(\bar{u})$ が得られる。

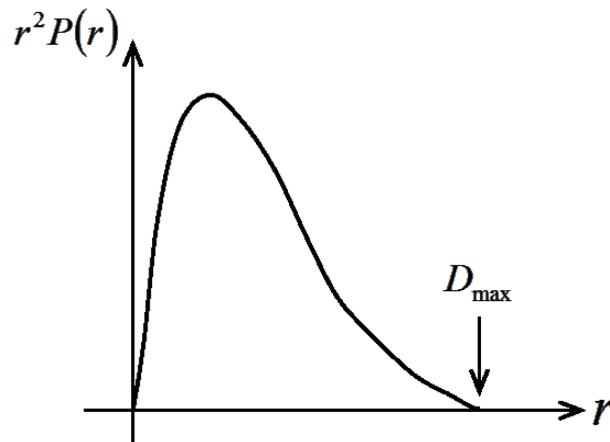


図 2.11 X 線溶液散乱実験における一般的な距離分布関数の形状。X 線溶液散乱実験の散乱プロファイルから得られる距離分布関数は一般にこのような形になる。図 2.10 より、完全に重なった状態から重なりが少なくなっていくにつれて $r^2P(r)$ の値は減少していく。これが 0 になった時電子密度の重なりが完全に消失する。この時の r が最大分子長 D_{\max} を与える。

2.3.3 Zernike-Prins の式と散乱プロファイルの濃度依存性

溶液散乱による散乱強度は今まで一粒子による散乱強度を角度平均したものを扱ってきた。しかしこれは、多数の蛋白質は完全に独立してランダムな配向を取っている、という仮定に基づいている。この仮定は粒子間の相互作用が完全に無視される希薄極限においては正しく成り立つが、有限の濃度を持つ場合、特に濃度が高い場合は粒子間相互作用が無視できず成り立たない。この粒子間の相互作用を盛り込んだのが以下の Zernike-Prins の式である。

$$I(S) = \bar{N} F^2(S) \left\{ 1 - \frac{1}{v_1} \int_0^\infty [1 - g(r)] \frac{\sin 2\pi Sr}{2\pi Sr} 4\pi r^2 dr \right\} \quad v_1 = \frac{V}{N} \quad (2.22)$$

ここで、 V は X 線の照射体積、 \bar{N} は照射体積中の平均粒子数、 v_1 は粒子の占める平均体積である。 $g(r)$ は動径分布関数と呼ばれ、 r 離れた距離における粒子数密度を表す。この関数が粒子間相互作用を表す関数である。この式は近似式であり、平均の照射体積を表す大きさ D_0 (例えば、 $D_0 = \sqrt[3]{V}$) について $S = 1/D_0$ より小角の領域では成り立たない。これは、この極小角領域では散乱強度に照射体積の Fourier 変換成分である強い散乱が入ってくるためである。しかし、この領域はビームストッパーによって遮られる領域であるため実際には問題にならない。例えば D_0 をビームサイズ程度の大きさである 0.2 mm とすると、この時の Zernike-Prins の式の小角限界は $5.0 \times 10^{-7} \text{ \AA}^{-1}$ であり、非常に小角の領域であることがわかる。

Zernike-Prins の式は散乱強度に粒子間相互作用の影響を考慮しているが、濃度の関数にはなっていない。これは、動径分布関数が粒子の幾何的配置のみに着目した関数だからである。この動径分布関数について、粒子間相互作用ポテンシャル $\Phi(r)$ を仮定し、これに対する Boltzmann 分布 $\exp(-\Phi(r)/k_B T)$ に置き換える。すると Zernike-Prins の式は

$$I(S) = \bar{N} F(S)^2 \left[1 + \frac{(2\pi)^{3/2} \beta(S)}{v_1} \right] \quad (2.23)$$

$$\beta(S) = \frac{1}{(2\pi)^{3/2}} \int_0^\infty \left[\exp\left(-\frac{\Phi(r)}{k_B T}\right) - 1 \right] \frac{\sin(2\pi Sr)}{2\pi Sr} 4\pi r^2 dr \quad (2.24)$$

となる。この式について Guinier 近似を適用することで、次の式が得られる。

$$\frac{KC}{I(S=0, C)} = \frac{1}{M_w} + 2A_2 C \quad (2.25)$$

$$R_g^2(C) = R_g^2(0) - B_{if} C \quad (2.26)$$

ここで、 C は蛋白質濃度、 K は定数、 A_2 は第二ビリアル係数である。 M_w は見かけの分子量と呼ばれるものであり、会合状態を反映する量である。 B_{ij} は粒子間相互作用を表す係数であり、 $B_{ij} < 0$ の場合粒子間には引力相互作用、 $B_{ij} > 0$ の場合斥力相互作用が働いていることになる。

2.3.4 非経験的分子形状モデリング

X線小角溶液散乱が実験手法として確立して以来、散乱プロファイルから試料について直接的に得ることのできる情報は、これまで説明してきた慣性半径、原点散乱強度、最大分子長といった構造についてのパラメータであり、実際の形状などを具体的に推し量ることは困難であった。しかし、溶液散乱による散乱プロファイルから三次元の分子概形を推定するアルゴリズムが考案され[Svergun, 1999; Svergun *et al.*, 2001]、分子の推定概形を簡便にモデリングすることが可能になった。ここでは、特に蛋白質を対象とした分子形状モデリングソフトウェア GASBOR[Svergun *et al.*, 2001]について説明する。

モデリングは何かしらの基本単位となるパーツを空間上に配置していくことにより行われる。蛋白質を構成する原子レベルまでの分解能は SAXS プロファイルでは到達し得ないため、蛋白質の最小の構造単位はアミノ酸残基とする。GASBOR では、このアミノ酸残基を平均化し、 $C\alpha$ 原子を中心とする球体と見なしたダミー残基 (Dummy Residue: DR) がモデリングの基本単位となる(図 2.12)。ダミー残基の中心を $C\alpha$ 原子とすることはモデリングをするうえで蛋白質の構造による束縛条件を満たすために重要である。蛋白質の主鎖を構成する $C\alpha$ 原子間の距離はおよそ 0.38 nm であるため、ダミー残基の直径を 0.38 nm とし、各ダミー残基が最低 1 つは他のダミー残基と接するように自由空間上に配置していけば「蛋白質はアミノ酸一次元鎖を折り畳んだものである」という条件を満たしながらモデリングをすることが可能となる。さらに最大分子長を入力パラメータとして束縛条件に加えることで、ダミー残基が配置される空間は分子最大長により制限された自由空間上となる。

GASBOR は構造モデルをある関数によって評価し、最適なモデルに近づけていくが、この際に使用されるアルゴリズムが焼きなまし法である[Kirkpatrick *et al.*, 1983]。焼きなまし法は、与えられた関数 $E(\vec{x})$ について最小値を与える \vec{x} を探索するためのアルゴリズムであり、次の 3 ステップから成る。

- ① ターゲット関数 $E(\vec{x})$ の解 \vec{x} について適当な初期値 \vec{x}_0 を用意し、初期値の近傍で新たな解の候補 \vec{x}' をランダムに生成する。

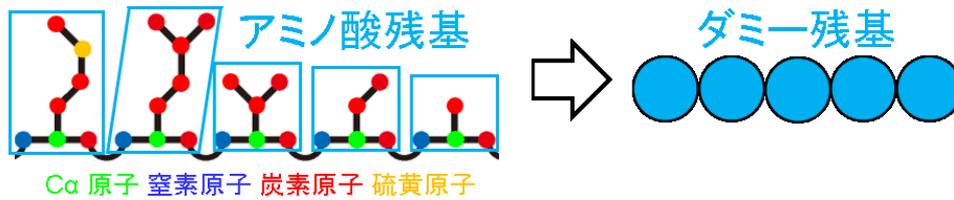


図 2.12 ダミー残基による蛋白質の粗視化。蛋白質のペプチド結合による鎖状構造(主鎖)中でペプチド結合に寄与しない炭素原子は $C\alpha$ 原子(図中の黄緑で表示された原子)と呼ばれ、蛋白質の主骨格をなす原子である。GASBOR はこの $C\alpha$ 原子を代表点としてダミー残基による蛋白質構造の粗視化を行っている。

② 以下で示す条件の下、 \vec{x}' を新たに解として採用するかどうかを決定する。

$$\Delta E = E(\vec{x}') - E(\vec{x}_0) < 0 \quad \dots \text{採用}$$

$$\Delta E = E(\vec{x}') - E(\vec{x}_0) > 0 \quad \dots \text{確率 } \exp(-\Delta E/T) \text{ で採用}$$

ここで T は任意のパラメータである。

③ \vec{x}_0 、あるいは \vec{x}' を新たな解として①に戻る。これらの操作をある T について適当な回数既に繰り返している場合、 T を小さくして①に戻る。

T は解の選択において重要な役割を果たすパラメータであり、実際の焼なましにおける温度に相当する。 $T=0$ とした場合、 $\Delta E < 0$ となる場合しか解が採用されない。この場合、解は極めて速く局所極小に収束することになる。 T について有限の値を与えればある確率で局所極小の谷から抜け出すことができる。従って、ある大きさの T から次第に小さくしていくことで解が局所極小に落ちることを避け、さらに T が十分に小さくなればその時点での近傍の極小に確実に近づけていくことができる。このようにして大域的極小に解を近づけていくのが焼なまし法の最大の特徴である。GASBOR においてはランダムに配置されたダミー残基鎖を初期解とし、ここからランダムにダミー残基を1つずつ移し替えていくことで新たな解を生成していく(図 2.13)。1度の操作では全体の構造は大きく変化しないため、解は十分に近傍同士にある。また、ダミー残基の移動先となる座標は、先に述べた次元鎖であるという条件を満たす座標に限定される。

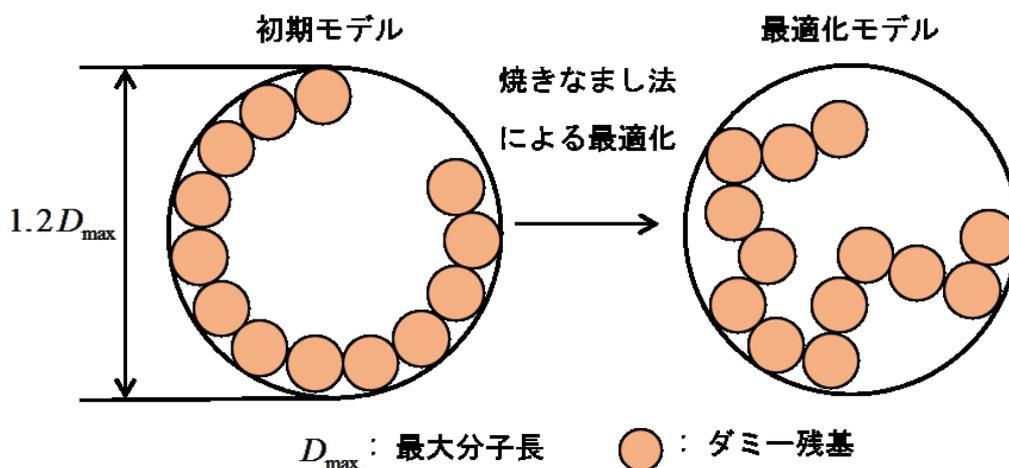


図 2.13 GASBOR 概略図

GASBOR の流れを簡単に説明すると次のようになる。

- ①最大分子長により領域の設定された自由空間内に入力残基数だけダミー残基を鎖状になるようランダムに配置する。(ダミー溶媒原子はダミー残基の位置に応じてその都度位置が決定される)
- ②焼きなまし法アルゴリズムに従い、モンテカルロ的にダミー残基モデルを操作し、ターゲット関数を計算、モデルを更新していく。
- ③パラメータ T が十分に小さくなり、かつ十分回数の計算がされている場合、最終的に得られたモデルをターゲット関数の大域解と判断し出力する。

実際に GASBOR で得られる蛋白質推定形状モデルが図 2.14 である。このように、ダミー残基集団の成す包絡面によって蛋白質形状を表現する。

次に GASBOR におけるターゲット関数について説明する。ターゲット関数は得られたモデルを逆空間、実空間双方の視点から評価するものであり、

$$E(\vec{r}) = \chi^2 + \alpha P(\vec{r}) \quad (2.27)$$

で与えられる。 χ^2 が逆空間に関する項、 $P(\vec{r})$ が実空間に関する項である。また α はターゲット関数への寄与が第一項と第二項で同程度となるようにするための重みである。第一項はモデルから得られた散乱プロファイル $I_{\text{DR}}(S)$ と実験による散乱プロファイル $I_{\text{exp}}(S)$ との一致性を評価する項である。溶液散乱プロファイルには蛋白質だけでなく水和殻による散乱も含まれているため、水和殻による散乱も計算しなければならない。その為、GASBOR 計算においてはダミー残基鎖の周囲にダミー溶媒原子

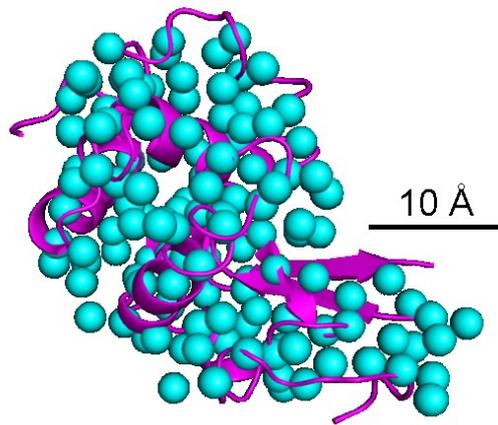


図 2.14 卵白リゾチーム結晶構造(Diamond, 1974; PDB ID: 6LYZ; マゼンタのリボンモデル)と卵白リゾチーム SAXS プロファイルから GASBOR によって計算された分子形状モデル(シアンのボールモデル)。ダミー残基の配置により蛋白質形状を表現している。ただし、ダミー残基の配置は実際のアミノ酸残基配置を表現するわけではない。GASBOR モデルは開発者である Svergun らが配布するデータ解析例より取得。

(Dummy Solvent Atom: DSA) を配置して、水和殻による散乱強度への寄与を考慮する。以上からモデルによる散乱プロファイル $I_{\text{DR}}(s)$ は Debye の式を用いて次のように与えられる。

$$I_{\text{DR}}(s) = \sum_{i=1}^{N+M} \sum_{j=1}^{N+M} g_i(s)g_j(s) \frac{\sin sr_{ij}}{sr_{ij}} \quad (2.28)$$

ここで、 N はダミー残基数、 M はダミー溶媒原子数である。 $g_i(s)$ はモデルの球体 1 つ 1 つの構造因子であり、次のように与えられる。

$$g_i(s) = \begin{cases} f(s) & \text{for a DR} \\ \frac{4\pi r_i^2}{M} \Delta r \cdot \delta\rho_b & \text{for a DSA} \end{cases} \quad (2.29)$$

ダミー残基について、 $f(s)$ はアミノ酸残基の構造因子を角度平均し、全残基について平均化した構造因子である。ダミー溶媒原子について、 Δr は水和殻の厚み、 $\delta\rho_b$ は水和溶媒分子のコントラスト電子密度である。従ってダミー溶媒原子の構造因子は水和殻の総電子数をダミー溶媒原子数で割ったものとなっている。ダミー溶媒原子はダミー残基とはやや扱いが違い、一様電子密度を持つ水和殻についてのサンプリング点として扱

っている。 $I_{\text{DR}}(s)$ によってターゲット関数の逆空間項 χ^2 は次のように定義される。

$$\chi^2 = \frac{1}{n-1} \sum_{j=1}^n \left[\frac{\langle c(s) \rangle I_{\text{DR}}(s_j) - c I_{\text{exp}}(s_j)}{\sigma(s_j)} \right]^2 \quad (2.30)$$

ここで、 $\sigma(s_j)$ は実験データの偏差である。 $\langle c(s) \rangle$ は $I_{\text{DR}}(s)$ を以下のように補正する関数である。 $I_{\text{DR}}(s)$ と $I_{\text{exp}}(s)$ のフィッティングは、小角領域でしかよく一致しない。これは、残基を球体として扱うことに起因する。そのため、これを高角側まで一致するように補正する関数 $\langle c(s) \rangle$ をかけてフィッティングを行う。 c はスケール因子であり、次のように定義される。

$$c = \left[\sum_{j=1}^n \frac{I_{\text{exp}}(s_j) I_{\text{DR}}(s_j)}{\sigma(s_j)^2} \right] / \left[\sum_{j=1}^n \frac{I_{\text{DR}}(s_j)^2}{\sigma(s_j)^2} \right] \quad (2.31)$$

第二項では、モデルが蛋白質の取りうる構造を満足しているかを評価する。この関数はエラー関数と呼ばれ、正確には次のように定義される。

$$P(\vec{r}) = \sum_k [W(R_k)(N_{\text{DR}}(R_k) - \langle N(R_k) \rangle)] + G(\vec{r}) + [\max\{0, (|\vec{r}_c| - r_0)\}]^2 \quad (2.32)$$

第一項は蛋白質の平均的残基分布による制限項である。ダミー残基モデルが図 2.15 のように分布しているとする。ある残基 k を中心として、厚さ 0.1 nm の球殻層を考え、距離 R_k における球殻内の残基分布数を $N(R_k)$ とする。これを複数の蛋白質の全残基について計算し、平均をとったのが平均残基分布関数 $\langle N(R_k) \rangle$ である。蛋白質の各残基に対する残基分布は平均的にこの関数の形に従うと考えられる。各残基における $\langle N(R_k) \rangle$ からのずれをエラーとして、モデルがこの残基分布に従っているかを評価するのが第一項である。蛋白質の構造は近距離秩序が支配的であるため、遠距離では偏差が大きくなる。その為、偏差に反比例する重み $W(R_k)$ をかけることで近距離での残差分布をより厳密に評価するようになっている。第二項はダミー残基の充填制限項である。この項によって蛋白質が一次元鎖の集合である条件、サブユニット同士が結合していない条件が満たされる。第三項は蛋白質重心による制限項である。GASBOR は入力された分子最大長 D_{max} に従って、半径 $1.2D_{\text{max}}$ の球内にダミー残基を配置していく。この項はこの時蛋白質の重心 \vec{r}_c が球の中心からドリフトしないように制限する項である。

GASBOR は計算に際して①散乱プロファイル、②最大分子長、③残基数の入力を必要とする。①、②については 2.3.2 節 で触れた距離分布関数計算ソフトウェア GNOM の出力ファイルに含まれるため、実際にはこの出力ファイルと残基数が実際に必要な入

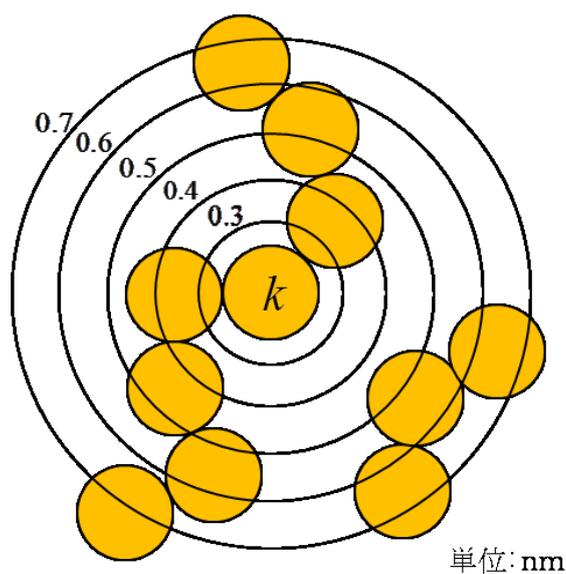


図 2.15 ダミー残基モデルにおける残基分布

力となる。残基数についてはシーケンスによる配列解析で正確な値を知ることが可能である。しかし、散乱プロファイルは水和殻を含めたコントラストによる散乱強度の配向平均であり、表面付近の細かい構造についての情報は失われている。その為、入力残基数は実際の値よりも多少低い値が妥当となる。どの程度の残基数が妥当であるかは厳密にはわからないため、実際の計算においては残基数を変えた複数のモデルを計算し、実験の散乱プロファイルと最もよく一致するものを妥当なモデルとして採用する。

3章 多変量解析を用いた蛋白質形状推定の高度化

1.4.1 節で述べたように SAXS における分子形状推定は不確かさを伴うものである。本研究では、この不確かさを解決するべく多変量解析を用いた分子形状モデルの分類法を開発した [Oide *et al.*, 2018]。本章では、提案手法の詳細を説明した後、実験データへの適用結果及び既存の分類手法との比較結果について述べる。

3.1 蛋白質形状推定における多変量解析の導入

1.4.1 節にて述べたように GASBOR を用いた蛋白質形状推定においては「推定形状が一意に決定できない」という問題が存在する。この問題を実際に示したのが図 3.1 であり、17 Å 分解能の SAXS プロファイルに対して複数回の GASBOR 計算を行った結果を示している。提示した問題が生じていなければ、一群の分子形状モデルの不確かさは図 3.1b のスケールバー(20 Å)よりも細かい構造に制限されるはずである。しかしながら、これらのモデルは概形からして大きく異なっている。この不確かさは「試料形状について配向平均された散乱プロファイルが得られる」という測定原理そのものに起因しており [Petoukhov & Svergun, 2015]、実験データのシグナルノイズ比や散乱プロファイルのサンプリング間隔、測定される回折角の範囲に関係なく生じる。例えば最大分子長 150 Å、残基数 500 の蛋白質について測定する場合、サンプリング定理より最低限必要な散乱プロファイルのサンプリング間隔は $1/(2 \times 150) \text{ \AA}^{-1} = 3.3 \times 10^{-3} \text{ \AA}^{-1}$ となる。これ以上細かいサンプリングは情報を増やさない冗長なものである。一方、GASBOR モデルがダミー残基の配置であることを考えると回折角は 4 Å 分解能、即ち 0.25 \AA^{-1} 程度までが必要な情報を持つ範囲であろう。この時変数量はおよそ 76 となるが、形状モデルの変数量は $500 \times 3 = 1500$ であるため、必要な変数量の 5% 程度にしかならない。従って、SAXS プロファイルからの形状推定は非適切問題となり不確かさが伴ってしまう。このように、本章で扱う問題は測定系の工夫や装置性能の改善といった実験面でのアプローチでなく、解析の高度化によって扱うべき問題である。

これに対して、本研究では多数の GASBOR モデルから推定形状の傾向を分布として抽出・分類（クラスタリング）し、そこから尤もらしい形状を推定するという形状推定スキームを提案する(図 3.2)。まず各分子形状モデルが互いによく重なるように並進・回転操作による重ね合わせを行い、各モデルを三次元のボクセルに格納することでダミー残基の三次元密度に変換する。このように処理すること

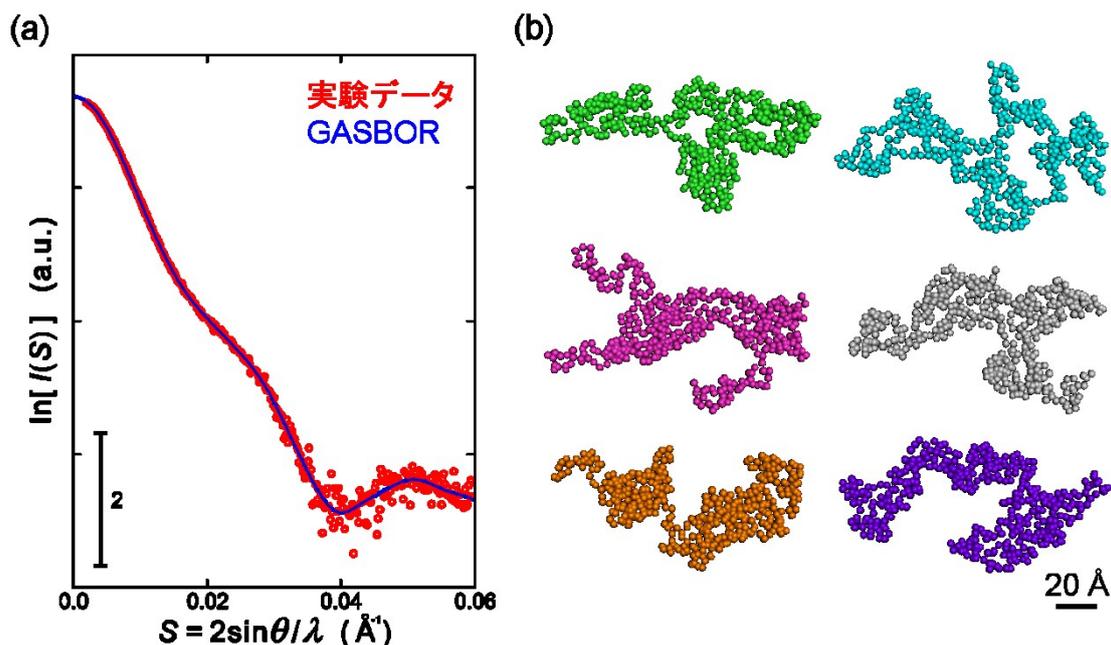


図 3.1 (a) 蛋白質 Protein Disulfide Isomerase の SAXS プロファイル。赤色の点が実験データ、青色の曲線が GASBOR 計算から得られた推定形状モデルのプロファイルである。両者は $S = 0.06 \text{ \AA}^{-1}$ 、即ち 17 \AA 分解能までよく一致している。(b)複数回の GASBOR 計算によって得られた蛋白質推定形状モデル。スケールバーは全モデルで共通である。

で、ある 1 つの分子形状モデルは各ボクセルの値によって張られる超多次元空間上の一点として表現される。多数の分子形状モデルが多次元空間上に形成する分布を主成分分析(Principal Component Analysis: PCA)で可視化し、似た形状毎に分類・平均化をすることでごく少数個の形状候補に削減する。提案手法はあくまで「形状の不確かさを分布として可視化することで形状推定の見通しをよくすること」が目的であり、提案手法が尤もらしい形状を抽出することはない。そのため、別の実験データや事前情報との組み合わせは必須である。次節以降で提案スキームの詳細を説明する。

3.2 モデルの重ね合わせ

重ね合わせのアルゴリズムは 2001 年に開発されたソフトウェア SUPCOMB [Kozin & Svergun, 2001]で使用されたアルゴリズムを基に作成した。SUPCOMB は

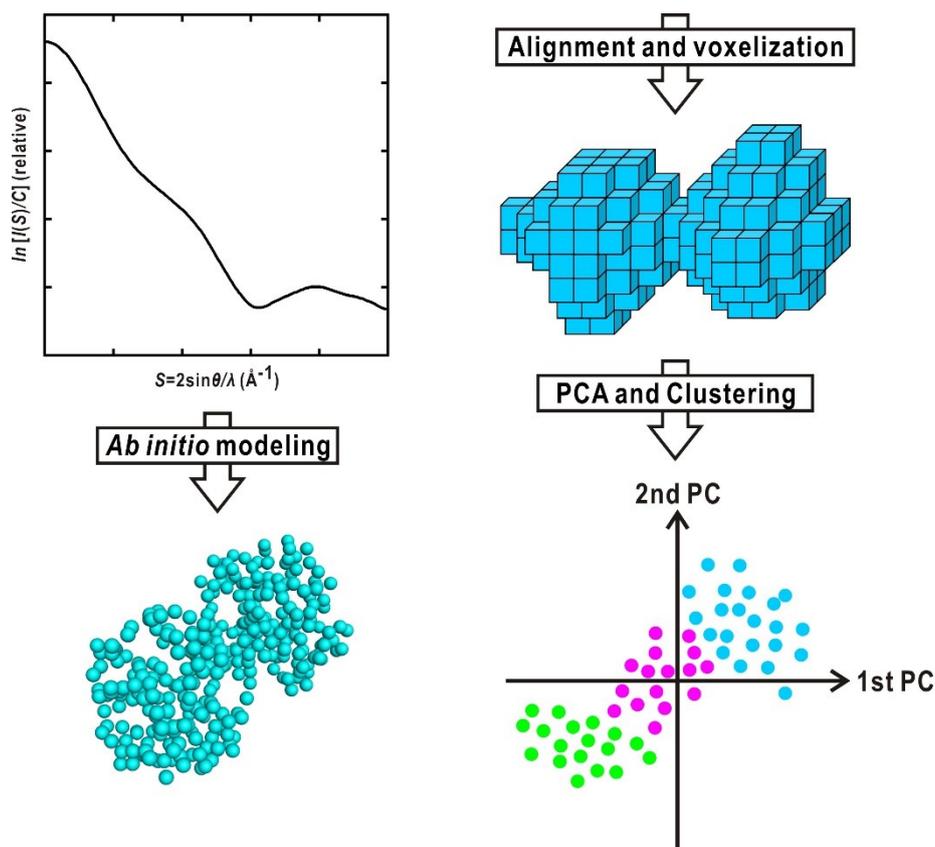


図 3.2 多変量解析による蛋白質形状推定スキーム

SAXS で得られる低分解能の形状モデルと X 線結晶構造解析などで得られる具体的な原子構造モデルとを重ね合わせることを目的としたソフトウェアである。重ね合わせのアルゴリズムを実装するにあたっては、重ね合わせたモデルの一致度を定義する必要がある。ここでは、具体的な重ね合わせの方法についてだけでなく、このモデル一致度の定義についても説明する。

3.2.1 慣性主軸を用いたモデル重ね合わせ

重ね合わせはモデルへの並進・回転操作によるが、参照モデルに対してモデルを逐一並進・回転させて最も一致度の高い相対位置・配向を探索する方法は計算量の観点から非現実的である。また、この方法では参照モデルに対して最も重なるように相対位置・配向が決定されるため、一致度が参照モデルに依存する、という問題が生じる。これらの問題を回避するため、慣性主軸による重ね合わせを導入する。 N 個のダミー残基から成るモデルに対し、 i 番目のダミー残基の座標を

$\vec{r}_i = (x_i, y_i, z_i)$ とすると、慣性テンソル I が次のように定義できる。

$$I = \begin{pmatrix} \sum_{i=1}^N (y_i^2 + z_i^2) & -\sum_{i=1}^N x_i y_i & -\sum_{i=1}^N z_i x_i \\ -\sum_{i=1}^N x_i y_i & \sum_{i=1}^N (z_i^2 + x_i^2) & -\sum_{i=1}^N y_i z_i \\ -\sum_{i=1}^N z_i x_i & -\sum_{i=1}^N y_i z_i & \sum_{i=1}^N (x_i^2 + y_i^2) \end{pmatrix} \quad (3.2)$$

ただし、モデルの重心（慣性中心）を原点に置いた。慣性テンソルは質点系の質量の分散を定義するテンソル量であるが、2 回対称テンソルであるため対角化可能である。対角化によって得られる固有値を主慣性モーメント、固有ベクトル軸を慣性主軸と呼ぶ。慣性主軸は質点系についてその分散を表現するように選ばれた直交基底であるため、2 つのモデルについて各主成分軸の単位固有ベクトル同士が重なるように並進・回転操作を行えば、最小限の計算量で尤もらしい重ね合わせが可能となる(図 3.3)。加えて、慣性主軸は他のモデルとは独立に決定されるため、参照モデルに依存した重ね合わせを防ぐことができる。

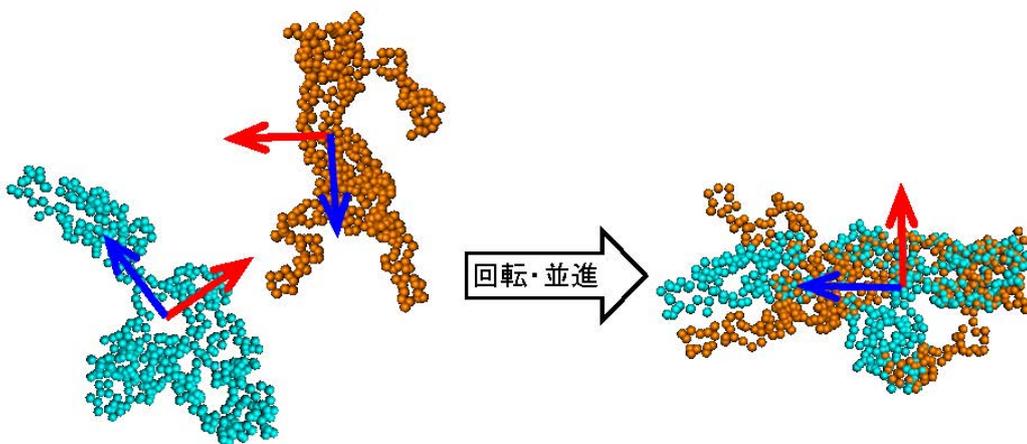


図 3.3 慣性主軸によるモデル重ね合わせの模式図。簡単のために 2 次元での図となっている。赤色、青色のベクトルはそれぞれのモデルでの慣性主軸の固有ベクトルであり、青色の軸は赤色の軸よりも主慣性モーメントの値が小さい。また、各モデルの固有ベクトルの原点はそのモデルの重心に置かれている。

3.2.2 モデル一致度の評価

前節で説明した方法では慣性主軸の単位固有ベクトルによって重ね合わせ方が一意に決まるように見える。しかし、実際には慣性テンソルの対角化において単位固有ベクトルの符号に任意性があるため、各軸での単位固有ベクトルの方向について $2^3=8$ 通りの組み合わせが存在する。3次元座標軸での固有ベクトルの取り方には右手系と左手系が存在し、これらは互いに重なり合わないため、結局固有ベクトルの取り方についての任意性は左手系を除外することで4通りとなる。ただし、単位固有ベクトルの組み合わせを考えるのは参照モデルに対して重ね合わせられるモデルについてだけであり、参照モデルの単位固有ベクトルは1通りに固定される。また、2.3.5節で触れた通り、GASBORによって推定される形状には鏡像の任意性が存在するため、モデルの鏡像についても一致度を評価する必要がある。従って、計8通りの重ね合わせパターンについてそれぞれの場合のモデル一致度を評価し、参照モデルに対して最も一致度の高い配向を決定しなければならない。以上の理由から、モデル一致度の指標を定義する。

GASBORモデルはダミー残基を要素とした集合と見なせるため、モデル一致度は集合間の距離によって定義できる。数学の分野では集合間の距離として Hausdorff 距離 h が用いられる [Huttenlocher *et al.*, 1993]。これは、モデル M_1 , M_2 及び各モデルに属するダミー残基 $m_1 \in M_1$, $m_2 \in M_2$ に対して

$$h(M_1, M_2) = \max_{m_1 \in M_1} \min_{m_2 \in M_2} \|m_1 - m_2\| \quad (3.2)$$

として定義され、直感的には「モデル M_1 に属する任意のダミー残基が、モデル M_2 に属するいずれかのダミー残基へ到達するために必要なユークリッド距離」と理解される(図 3.4)。この値が小さいほど2つのモデルは重なりが大きく、従って一致度が高いと言える。しかしながら、定義そのものの非対称性 ($h(M_1, M_2)$ と $h(M_2, M_1)$ が必ずしも一致しない) などからモデル一致度の指標とするには不適切であるため、本研究では SUPCOMB でモデル一致度の評価に用いられる Normalized Spatial Discrepancy (NSD) という指標を用いる。NSD は Hausdorff 距離を基本としており、次のようなモデルに依存した関数 ρ で定義される。

$$\text{NSD} : \rho(M_1, M_2) = \left\{ \frac{1}{2Nd^2} \left[\sum_{i=1}^N h^2(m_{1i}, M_2) + \sum_{j=1}^N h^2(m_{2j}, M_1) \right] \right\}^{\frac{1}{2}} \quad (3.3)$$

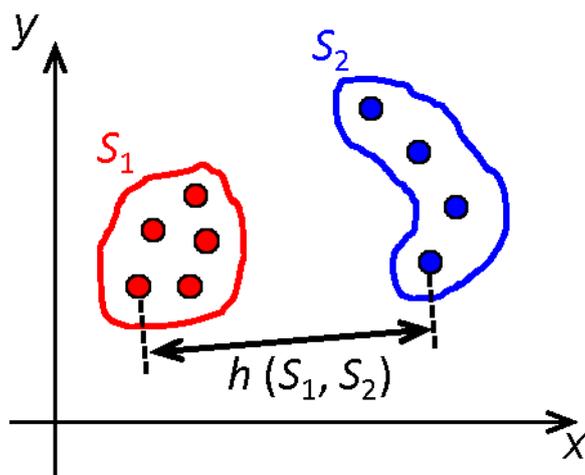


図 3.4 Hausdorff 距離の概念図。2 次元 $x-y$ 座標中のデータ点集合 S_1 に属する全ての点について、その点を中心にとった半径 $h(S_1, S_2)$ の円の中に、 S_2 に属する点が少なくとも 1 点存在することがわかる。注意点として、 $h(S_1, S_2)$ と $h(S_2, S_1)$ は必ずしも一致しない。これは、例えば S_1 が S_2 の部分集合である場合を考えれば明らかである。

ここで、 N はダミー残基数、 d は各モデルのダミー残基間距離である。モデル M_1 , M_2 それぞれを基準とした Hausdorff 距離を計算することでモデル一致度の指標としての信頼性を向上させているのが NSD の特徴である。この指標を用いて、8 通りの重ね合わせから最も参照モデルと一致度の高いモデルを決定する。

3.3 モデルのボクセルへの格納

主成分分析を用いてモデル形状を分類するためには、モデル形状の分散を記述できる多次元空間でモデルを取り扱わなければならない。ここでは、そのような多次元空間としてボクセルにより定義される空間を用いる。

まず、モデル形状を多次元空間で表現するため、得られた全ての GASBOR モデルを格納できる大きさを持つ箱を用意し、 J 個のボクセルに分割する。一つの GASBOR モデル M_1 をこの箱に格納した場合、各ボクセルに幾つのダミー残基が所属するかによって、このモデルを表現できるであろう(図 3.5)。今、各ボクセルの値によって定義される J 次元空間を考えれば、モデルはこの空間内の一点で表さ

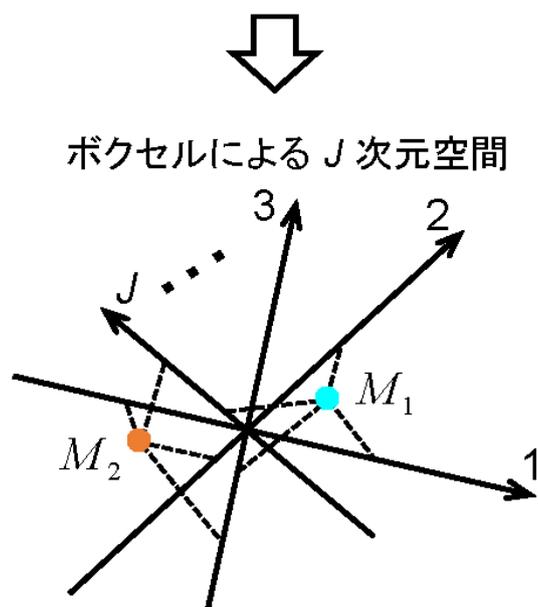
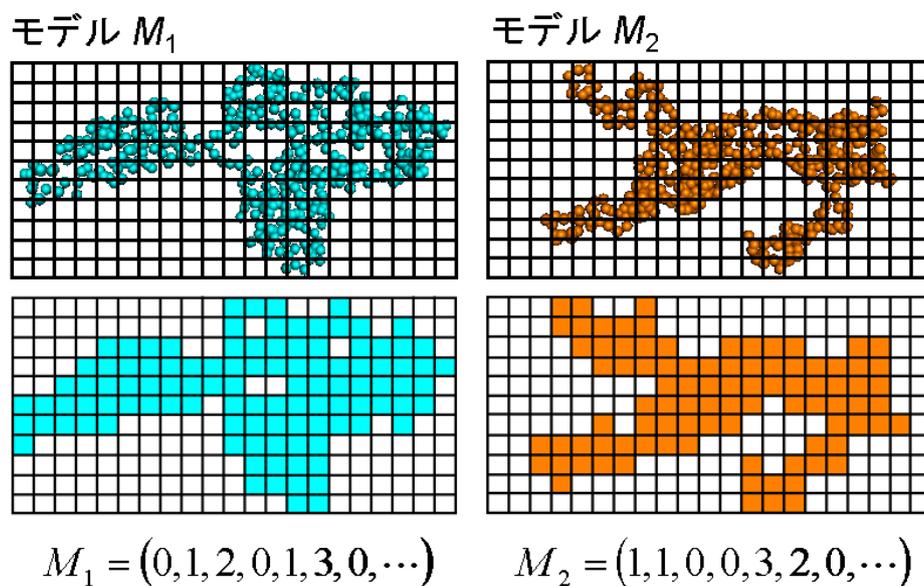


図 3.5 GASBOR モデルのボクセルへの格納。1 個の GASBOR モデルは各ボクセルに格納された値によって規定され、 J 次元空間上の 1 点で表現することが可能となる。その他の GASBOR モデルについても同様の操作を行うことで、形状の違いによる多次元空間上での分布が作られる。

れる。同様に、別のモデル M_2 についても箱への格納を行えば、 M_1 との形状の差異がダミー残基の所属するボクセルに反映され、 J 次元空間上で M_1 とは異なる一

点として表現される。このようにして得られた J 次元空間上での多数のモデル点の分布に対して主成分分析を行えば、モデル形状分布の可視化と形状分類が可能となるであろう。

モデルをボクセルへ格納するには、ボクセルサイズを最適化しなければならない。ボックスの大きさが一定であれば、ボクセルサイズとボクセル数 J は反比例関係にある。 J はモデル形状空間の次元でもあるため、ボクセルサイズの値に応じてモデル形状空間の次元数が増える。一般に、データ点間の距離をそのデータ点同士の類似度として捉える場合、データ点空間の次元の増加は類似度の低下を招く(図 3.6)。逆に、次元の過度な削減はデータ点間の差異を曖昧にする。従って、モデル形状の分類に適したボクセルサイズを見積もる必要が生じる。

直観的な定義になるが、適切なボクセルサイズを「1つのボクセルに1~2個のダミー残基が格納されるサイズ」と定義する。これより小さいボクセルサイズでは

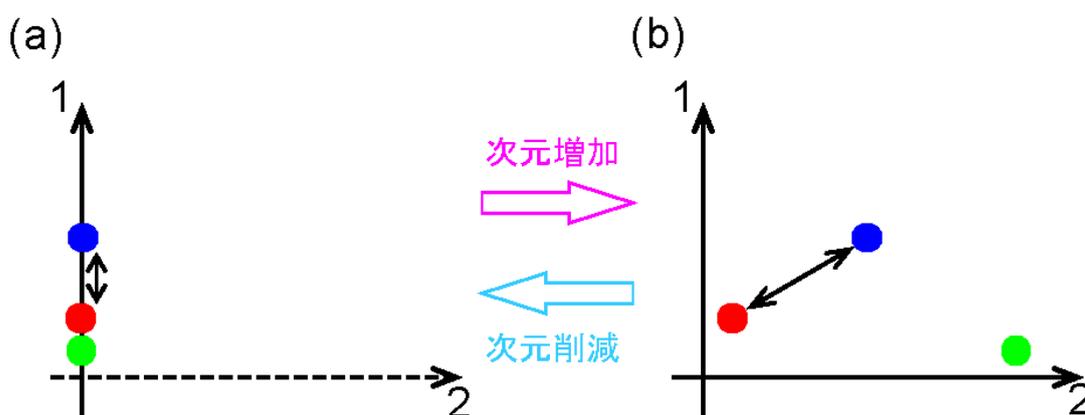


図 3.6 次元とデータ点類似度の相関。(a)から(b)へ次元を増加させると、図の青丸のように赤丸との距離が大きくなり類似度が低下する。この類似度の低下は次元の増加に伴って急速に甚だしくなる。このような空間の多次元性に起因する問題一般を「次元の呪い」と呼び後述するクラスタリングにおいても生じる問題である。また、逆に次元を削減する場合には赤丸と緑丸の類似度が不自然に増加する、ということが起こる。次元の増減に伴い生じる問題は相補的なものであるため、データ点の分類においては適切な次元を見極める必要がある。

GASBOR モデル形状内のボクセルに格納されるダミー残基の疎密による内部構造が生じ、類似度の低下を招くと考えられるからである。これを式で表せば、GASBOR モデルの体積を V 、ダミー残基数を N とすると、ボクセル一辺の長さ L について

$$L^3 \cong \frac{V}{N}$$

なる関係が成り立つため、GASBOR モデルの体積がわかれば適切なボクセルサイズを決定できる。GASBOR モデルの体積はモデル形状ファイル内で確認をすることができる。例えば 450 残基で計算をした場合 V は $7.65 \times 10^4 \text{ \AA}^3$ であり、ダミー残基 1 個当たりの占有体積は 170 \AA^3 となる。実空間拘束条件により GASBOR モデルの残基密度はあらゆる場合に一定であるため、ダミー残基 1 個当たりの占有体積はこの値で固定されたパラメータとして考えてよい。実際、ダミー残基数を変更して GASBOR 計算を行っても 1 個当たりの占有体積は変わらず 170 \AA^3 であった。従って、ボクセル一辺の長さは普遍的に $\sqrt[3]{170} \cong 5.5$ と定義できるが、ボクセルとしての扱いの容易さを考慮して適切な長さを 6 \AA と決定する。本節の最初に触れた適切なボクセルサイズの定義より、ここで決定されたボクセルサイズは適切なボクセルサイズの下限である。そのため、ボクセル格納後にローパスフィルター処理を施すことで意図しない類似度の低下を防いでいる。

3.4 主成分分析の基本原則 [Frank, 2006]

今、 J 個の独立な変数 $x^j (j = 1, 2, \dots, J)$ によって構成される N 個の従属変数 $\bar{x}_n = (x_n^1, x_n^2, \dots, x_n^J; n = 1, 2, \dots, N)$ を考える。これは J 次元空間中に分布する N 個のデータ点として解釈される(図 3.7a)。 $J \gg 1$ である場合、このデータ点の分布を直接可視化することは不可能である。そのため、主成分分析を用いてデータ点分布の傾向を抽出する。主成分分析とは「多次元空間中に分布するデータ点の集団について、データ点の分散を表現するような直交基底を張り直す操作」である。この定義に基づいて主成分分析の原理を説明する。

まず、データ点の分布を 1 次元の分布で表すことを考える。そのためには、データ点の分散をなるべく表現できる方向にデータ点集団を射影させるのが望ましい。これは、 \bar{x}_n を J 次元ベクトル $\bar{x}_n = (x_n^1, x_n^2, \dots, x_n^J)$ と捉えた場合に

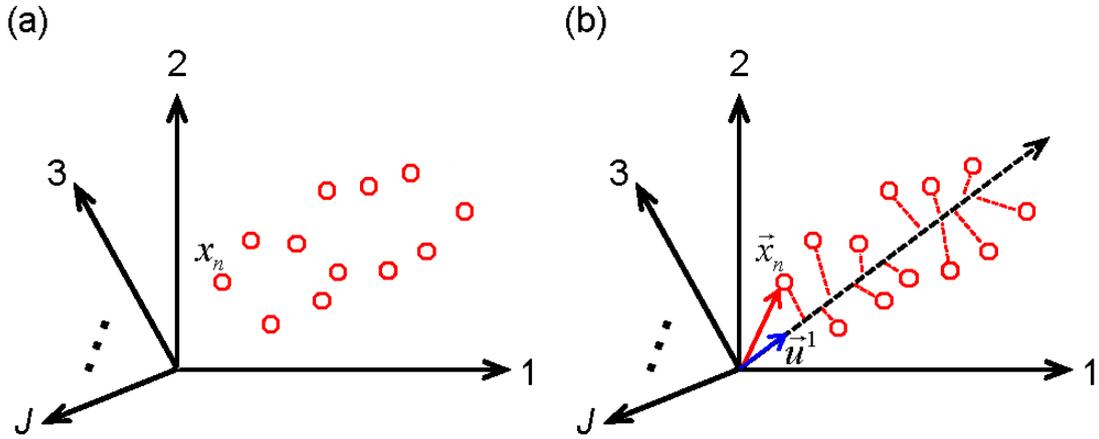


図3.7 J 次元空間上に分布するデータ点(a)とその分散を表現する方向への射影(b)。このようにして直交基底を選び直すのが主成分分析である。

$$\sum_{n=1}^N (\bar{x}_n \cdot \bar{u}^1)^2 = (\mathbf{X} \bar{u}^1)^T (\mathbf{X} \bar{u}^1) = \bar{u}^{1T} \mathbf{X}^T \mathbf{X} \bar{u}^1 = \lambda_1 \quad (3.4)$$

を最大にするような J 次元単位ベクトル \bar{u}^1 を求めることに等しい(図3.2 b)。式(3.4)の \mathbf{X} は

$$\mathbf{X} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_N \end{pmatrix} = \begin{pmatrix} x_1^1 & x_1^2 & \cdots & x_1^J \\ x_2^1 & x_2^2 & \cdots & x_2^J \\ \vdots & \vdots & \ddots & \vdots \\ x_N^1 & x_N^2 & \cdots & x_N^J \end{pmatrix} \quad (3.5)$$

と定義される行列である。また、 T は転置を意味する記号である。次に、データ点集団の \bar{u}^1 と直交する成分について、最も分散を表現できる J 次元単位列ベクトル \bar{u}^2 を求める。この時 \bar{u}^2 は

$$\sum_{n=1}^N (\bar{x}_n \cdot \bar{u}^2)^2 = (\mathbf{X} \bar{u}^2)^T (\mathbf{X} \bar{u}^2) = \lambda_2 < \lambda_1, \quad \bar{u}^1 \cdot \bar{u}^2 = 0 \quad (3.6)$$

なる条件を満たす。このようにして $\bar{u}^1, \bar{u}^2, \bar{u}^3, \dots, \bar{u}^j, \dots$ を求めていくと、これらは

$$\mathbf{D}^j \bar{u}^j = \mathbf{X}^T \mathbf{X} \bar{u}^j = \lambda_j \bar{u}^j \quad (3.7)$$

を満たす正規直交基底となっている。従って、主成分分析とは $J \times J$ 行列 \mathbf{D}^j の対角化に他ならない。 $\{\bar{x}^n\}$ の \bar{u}^j 方向への射影を第 j 主成分と呼ぶ。固有値の大きい

順、即ち \bar{u}^1 から順に少数個の主成分、多くの場合二、三個の主成分のみで $\{\bar{x}^n\}$ の分布を近似的に再現することで J 次元空間中での分布の傾向を抽出できる。ここまでの議論では $\{\bar{u}^j\}$ について原点の取り方に任意性が存在するため、原点としてデータ点集団 $\{\bar{x}^n\}$ の重心

$$\bar{\bar{x}} = (\bar{x}^1, \bar{x}^2, \dots, \bar{x}^J) = \left(\frac{1}{N} \sum_{n=1}^N x_n^1, \frac{1}{N} \sum_{n=1}^N x_n^2, \dots, \frac{1}{N} \sum_{n=1}^N x_n^J \right) \quad (3.8)$$

を選ぶ。行列 \mathbf{X} について、重心が原点となるよう並進操作を行うことで行列 \mathbf{D}^J は

$$\mathbf{D}^J = \mathbf{X}'^T \mathbf{X}' = (\mathbf{X} - \bar{\mathbf{X}})^T (\mathbf{X} - \bar{\mathbf{X}}) \quad (3.9)$$

と再定義できる。この時 \mathbf{D}^J の i 行 i' 列における行列要素 $d_i^{i'}$ は

$$d_i^{i'} = \sum_{n=1}^N (x_n^i - \bar{x}^i)(x_n^{i'} - \bar{x}^{i'}) \quad (3.10)$$

となり、 J 個の独立変数についての分散あるいは共分散である。そのため、 \mathbf{D}^J は分散共分散行列と呼ばれる。結局、主成分分析とは分散共分散行列 \mathbf{D}^J の対角化によって得られる直交基底への射影により、多次元空間での分布を低次元に落とし込む処理である。 J 次元空間全体での $\{\bar{x}^n\}$ の分散は \mathbf{D}^J の跡(トレース)として保存されているため、主成分分析によって J 次元空間中での分布の情報に欠損は生じない、ということをつけ加えておく。分散共分散行列はまた、

$$\mathbf{D}^N = (\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^T \quad (3.11)$$

として $N \times N$ 行列で定義することもできる。式(3.9)と式(3.11)で得られる固有値に違いはないが、一般に行列の対角化では行列の次元の 3 乗に比例して計算量が増加する。本研究では、 N はおよそ数百程度になる一方で、 J は試料蛋白質のサイズにもよるが数千~数万に及ぶ。その為、本研究では分散共分散行列を \mathbf{D}^N として定義する。ただし、対角化によって得られる固有値 $\lambda_i (i=1,2,\dots,q)$ は \mathbf{D}^N , \mathbf{D}^J で共通だが、固有ベクトルは \mathbf{D}^N , \mathbf{D}^J の次元が異なるため、

$$\begin{cases} \mathbf{D}^N \bar{v}_i = \lambda_i \bar{v}_i \\ \mathbf{D}^J \bar{u}_i = \lambda_i \bar{u}_i \end{cases} \quad i = 1, \dots, q \quad (3.12)$$

となる。 J 次元空間の低次元への射影に必要であるのは \mathbf{D}^J の固有ベクトル \bar{u}_i なので、以下の遷移公式を用いることで \mathbf{D}^N の固有ベクトル \bar{v} から \bar{u} を求める。

$$\begin{cases} \mathbf{U} = (\mathbf{X} - \bar{\mathbf{X}})^T \mathbf{V} \Lambda^{-\frac{1}{2}} \\ \mathbf{V} = (\mathbf{X} - \bar{\mathbf{X}}) \mathbf{U} \Lambda^{-\frac{1}{2}} \end{cases} \quad (3.13)$$

ただし、

$$\begin{cases} \mathbf{U} = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_q) \\ \mathbf{V} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_q) \end{cases} \quad (3.14)$$

$$\mathbf{\Lambda}^{-\frac{1}{2}} = \begin{pmatrix} \lambda_1^{-\frac{1}{2}} & & & 0 \\ & \lambda_2^{-\frac{1}{2}} & & \\ & & \ddots & \\ 0 & & & \lambda_q^{-\frac{1}{2}} \end{pmatrix} \quad (3.15)$$

である。ここから固有値の大きい数個の固有ベクトルを抽出して次元削減を行う。最適な削減後の次元数を検討するため、ある1つのGASBORモデルについて1~3次までのモデル近似で比較したところ、2次のモデル近似でおおよその概形を再現でき、3次のモデル近似は2次のモデル近似に比べて有意な違いが表れなかった(図3.8)。そのため、本研究では J 次元空間を2次元まで削減する。

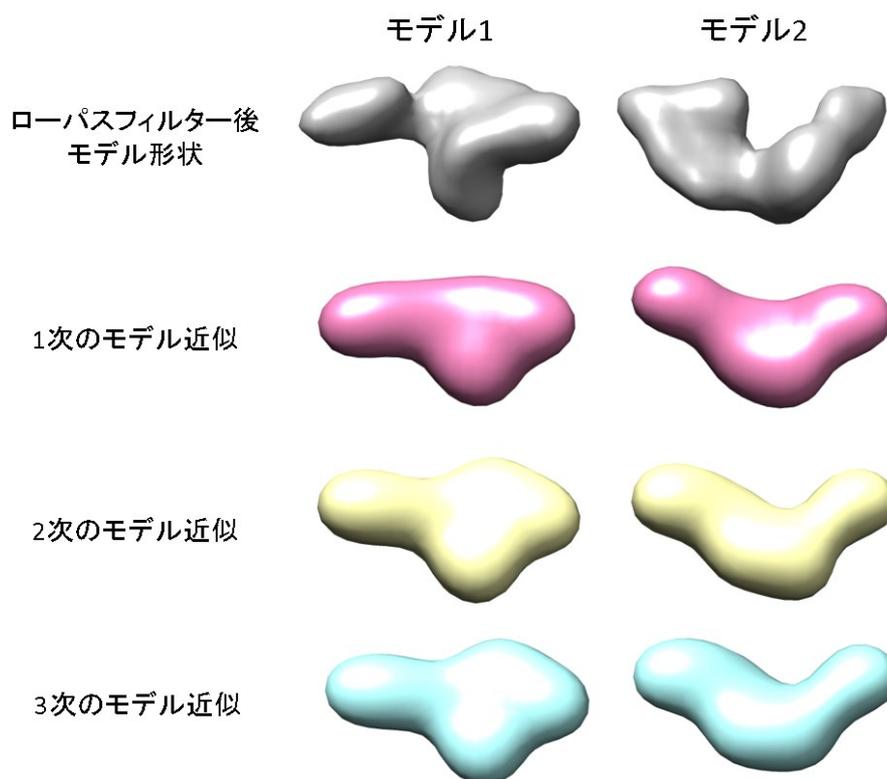


図 3.8 2 個の GASBOR モデルについて次数を変えて形状近似を行った結果。1 次近似と 2 次近似では 2 次近似がより良く形状を再現するが、3 次近似は 2 次近似と比較して有意な差が見られない。

3.5 K 平均法によるクラスタリングと平均化処理

主成分分析によって得られた 2 次元でのモデル形状分布について、クラスタリングにより形状を分類し、各クラスターで平均形状を作ることでモデル形状の分布を可視化する。クラスタリングには 1957 年に Steinhaus によって考案された K 平均法と呼ばれるアルゴリズムを用いる [Steinhaus, 1957]。K 平均法は直観的には以下のようにクラスタリングを行う [児玉渉、2011](図 3.9)。

- ① ランダムな初期位置に重心をセットする。
- ② 各データ点を最近接重心のクラスターに所属させる。
- ③ クラスターごとにデータ点の重心を計算し、重心位置を更新する。
- ④ 新たな重心について再度クラスタリングを行う。

上の処理を各重心の位置が収束するまで繰り返す。重心位置の収束は以下のように定義される値 T の収束度によって判断する。

$$T = \sum_{j=1}^K \sum_{x_n \in \text{cluster } j} d(x_n, c_j) \quad (3.16)$$

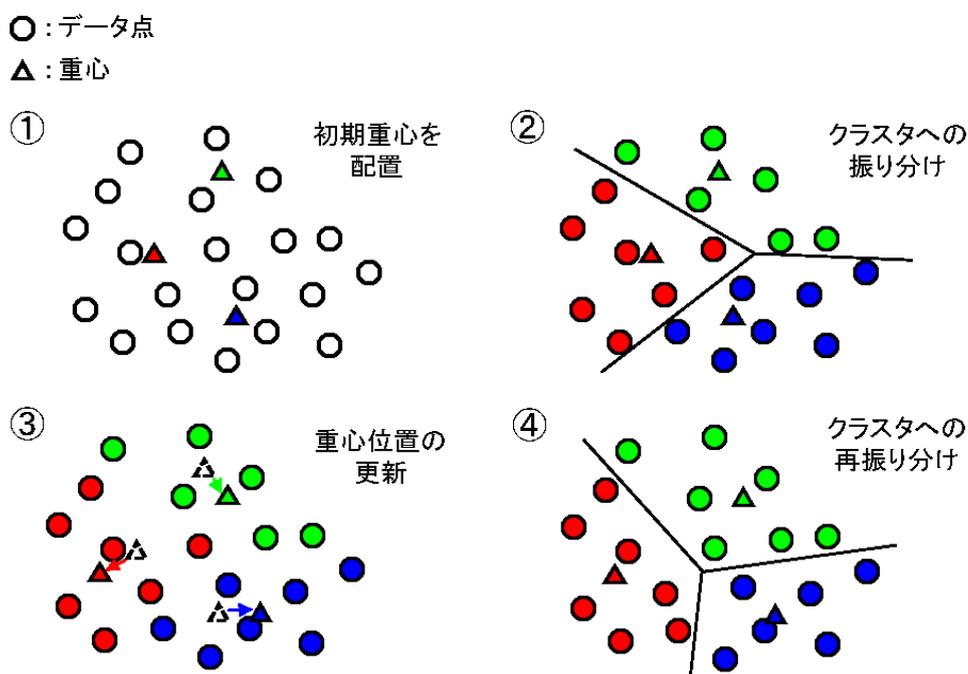


図 3.9 K 平均法によるクラスタリングのイメージ。③、④の操作を重心が収束するまで繰り返し行う。本研究では、式(3.16)で定義される T について、その変化量が 0.1% を切った場合に重心位置が収束したと判定している。

ただし、 K がクラスタ数 (重心数)、 x_n が 2 次元形状分布上でのモデル点位置、 c_j がクラスタ j の重心、 $d(x_n, c_j)$ が 2 点 x_n, c_j 間のユークリッド距離である。K 平均法では最終的な結果が初期重心位置に強く依存する。その為、本提案スキームでは異なる初期重心でクラスタリングを 100 回を行い、最も T が小さくなった結果を最終的なクラスタリング結果として採用する。また、K 平均法ではクラスタ数を指定する必要がある。本提案スキームでは、過不足なく分布の傾向を可視化できるだろうと考えクラスタ数を 10 に設定した。

それぞれのクラスタについてモデル形状の平均化処理を行うことで最終的なモデル形状の分布が可視化される。得られた 10 個の平均形状モデルを形状モデルの候補とし、この 10 個の中から尤もらしい形状を推定する。以上が提案スキームの一連の流れである。

3.6 溶液中での形状既知蛋白質への本手法の適用

提案手法を Protein Disulfide Isomerase(PDI)及び phototropin 1 LOV1 ドメイン (PIL1)の 2 種の蛋白質へ適用し、形状の分類が可能であるかどうかを確かめた。

3.6.1 好熱カビ由来 PDI への適用

3.6.1.1 PDI 分子形状推定の詳細

PDI は蛋白質内ジスルフィド結合形成の触媒としての機能を持ち、新生蛋白質の高次構造形成に関わる酵素である。酵母由来 PDI については結晶構造が解かれており、a、b、b'、a' の 4 つのドメインから成る J 字型の形状であることがわかっている(図 3.10)。好熱菌由来 PDI は酵母由来 PDI とアミノ酸配列の相同性が高いため、提案スキームによる分類が成功すれば 4 つのドメインが確認できる平均形状モデルが得られると考えられる。

本研究では 2010 年に測定された SAXS プロファイル [Nakasako *et al.*, 2010] (図 3.11)を用いた。GASBOR 計算では分子最大長とダミー残基数をパラメータとして与える必要があるため、まずこれらの値を決定した。分子最大長については距離分布関数計算ソフトウェア GNOM によって 128 Å に決定した(図 3.11)。ダミー残基数については 2.3.4 節末尾で述べたように、実際の残基数が最も実験の SAXS プロファイルを再現するとは限らない。実験の SAXS プロファイルをよく再現する

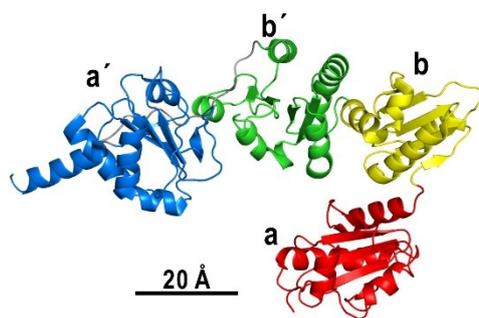


図 3.10 酵母由来 PDI 結晶構造[Tian *et al.*, 2006; PDB ID: 2B5E]

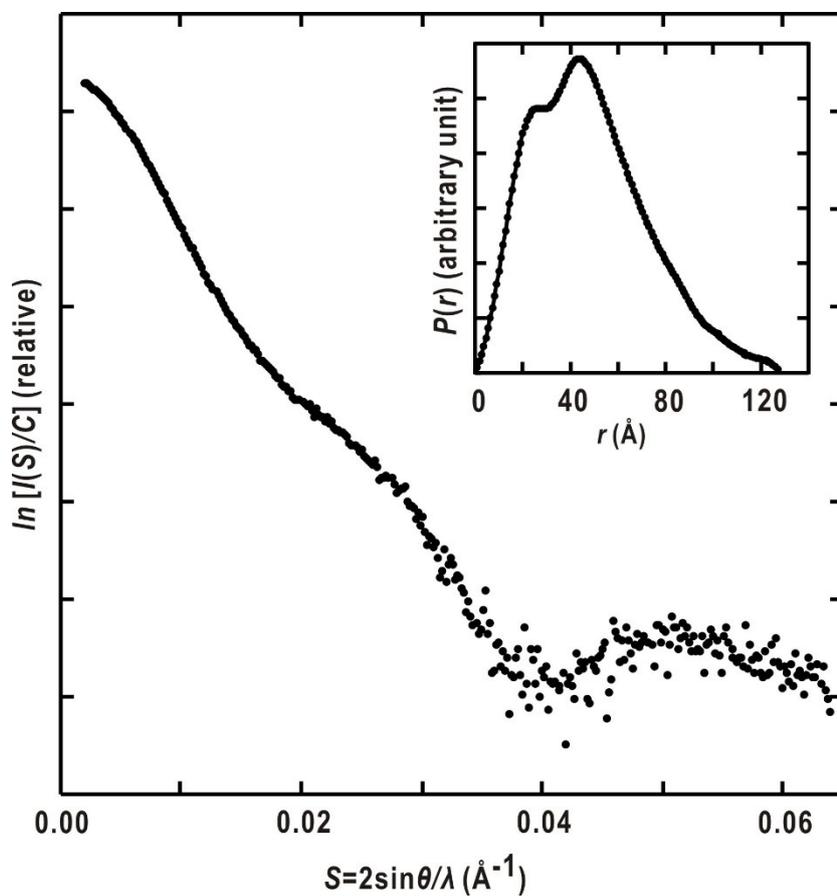


図 3.11 好熱カビ由来 PDI SAXS プロファイル。 $S = 0.025 \text{ \AA}^{-1}$ に見られる散乱プロファイルの肩はドメイン由来のものである。グラフ右上の挿入は GNOM によって計算された PDI 距離分布関数。

最適なダミー残基数を決定するため、実際の残基数 483 残基に対して 400 残基～500 残基まで 25 残基刻みでダミー残基数を変化させ、各ダミー残基数について 12 回の GASBOR 計算を行った。その結果、残基数ごとの χ^2 の平均値には変化がないものの、450 残基、500 残基での計算結果では χ^2 がより狭い範囲に収束する傾向にあった(図 3.12)。SAXS においては平均化や溶媒とのコントラストにより散乱プロファイルに寄与する残基が減少するため、実際の残基数よりも少ない 450 残基が適切であると判断された。このダミー残基数について SAXS プロファイルより GASBOR 計算を 576 回行い、得られた 576 個のモデルについて提案スキームを適用した。

提案手法による解析と併せてプログラム AMBIMETER [Petoukhov & Svergun, 2015]を用いて散乱プロファイルの不確かさを評価した。AMBIMETER は「入力された散乱プロファイルを与える形状がどの程度多く存在するか」を独自のスコア (ambiguity score)によって評価するプログラムである。開発者らの主張によると、スコアが 1.5 未満で「不確かさが少なく、経験的にはユニークな形状が約束される」、1.5 以上 2.5 未満で「モデリングによって複数形状が出力され得る」、2.5 以上で「対称性などの拘束を加えないと解析は難しい」とされている。

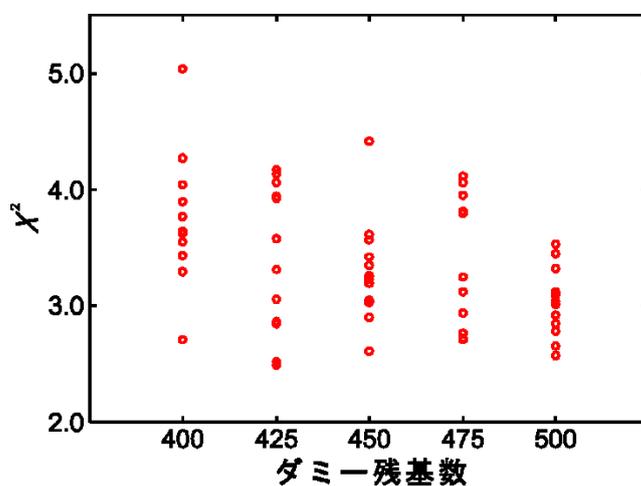


図 3.13 χ^2 の残基数依存性。ダミー残基数を変えて GASBOR 計算をした場合の χ^2 の残基数依存性をプロットしたグラフ。483 残基の PDI の場合、ダミー残基数 450、500 の場合に χ^2 が比較的低い値に収束する傾向にあった。

3.6.1.2 提案手法による PDI 解析結果

PDI SAXS プロファイルの ambiguity score は 2.45 であり、即ち PDI の形状推定が経験的には困難であることが示唆された。GASBOR モデル形状は平均形状を中心にほぼ均一な円盤状に分布しており(図 3.13)、ambiguity score による評価の通り様々な形状が出力された。平均形状はこれを反映して多様な形状を取っているため、酵母由来 PDI 結晶構造モデルを重ね合わせることで各クラスの形状の尤もらしさを評価した(図 3.14)。PDI の a' ドメインは b' ドメインとフレキシブルなループによって結合しており、基準振動解析の結果から a' ドメインに対して a、b、b' の 3 ドメイン(a-b-b' ドメイン)が剛体的に回転すると予測されている [Nakasako *et al.*, 2010]。そのため、結晶構造の重ね合わせにおいては剛体的であると考えられる a-b-b' ドメインの重なり具合を重視した。まず、全形状モデルを平均した場合、b' ドメインに対応する領域を欠損した不適切な形状が得られた。一方、各クラスでの平均形状を見ると、クラス I-V の 5 クラスにおいては a-b-b' ドメインが平均

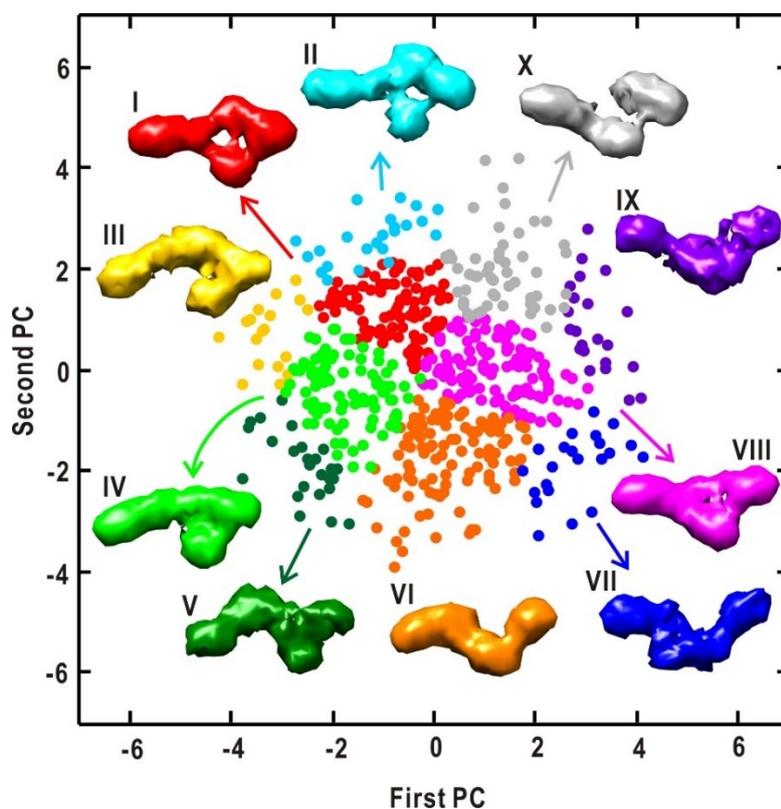


図 3.13 第 1 主成分(First PC)、第 2 主成分(Second PC)に射影された PDI モデル形状分布。

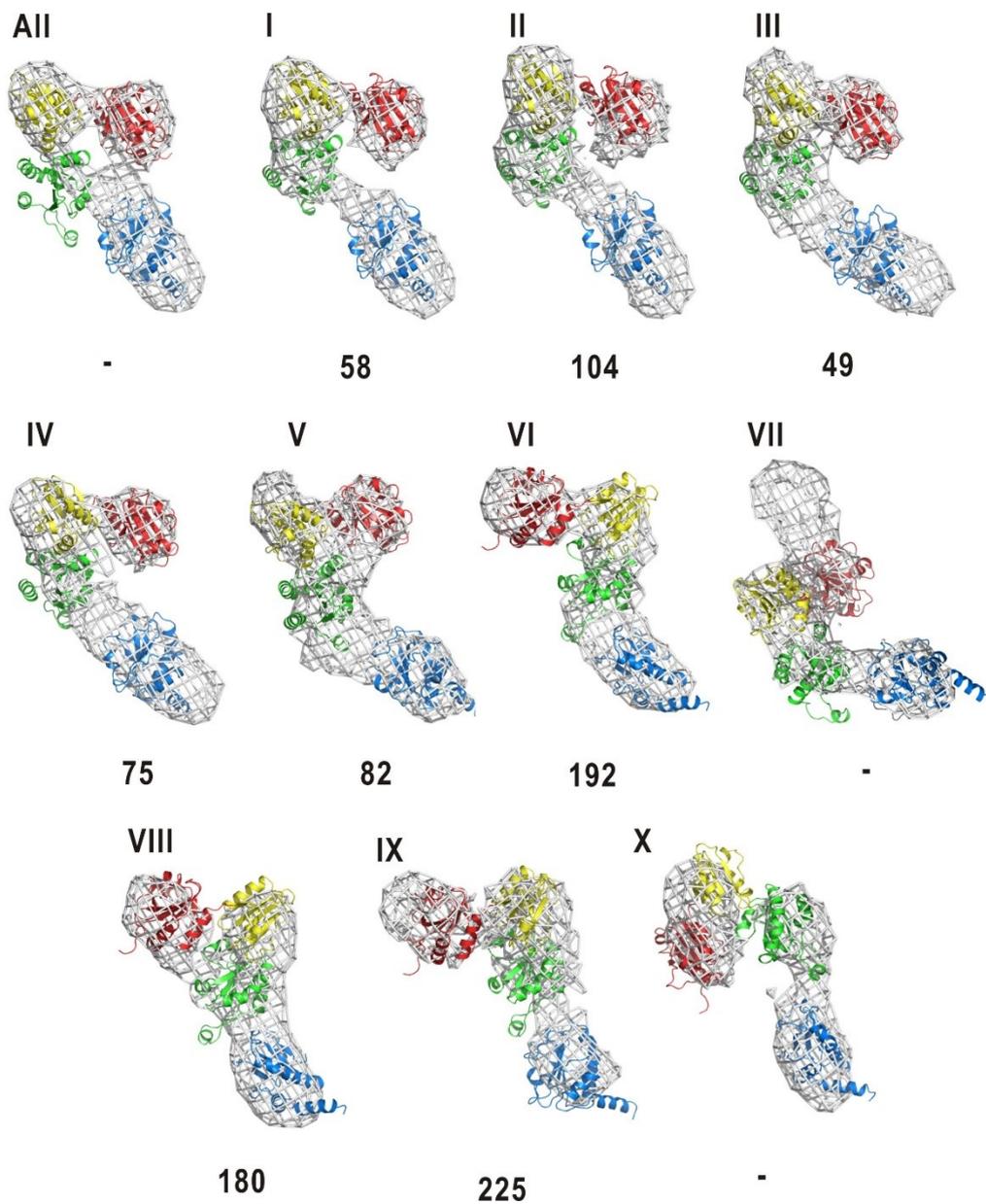


図 3.14 各クラスの平均形状と重ね合わせた PDI 結晶構造。各クラス下部の数字は各クラスに重ね合わせた結晶構造から CRY SOL によって計算されたプロファイルと実験データとの χ^2 である。なお、AII 及びクラス VII、X では χ^2 の値が「-」となっているが、これらのモデル平均形状は明らかにドメインの一部を欠損している(AII、クラス X)、あるいは余分な領域がある(クラス VII)ために CRY SOL による計算を行っていない。

形状とよく一致した。また、これらの 5 クラスを比較すると a-b-b' ドメインに対する a' ドメインの相対位置が異なっており、a' ドメインの柔軟性を反映しているものと考えられる。その他の 5 クラス(クラス VI-X)については a-b-b' ドメインの一致度が悪かった。実際、ソフトウェア CRYSOLOG [Svergun *et al.*, 1995]によって各クラスに重ね合わせた結晶構造モデルから SAXS プロファイルを計算し、実験データとの一致度を χ^2 によって比較したところ、クラス I-V と比較してクラス VI-X は χ^2 が大きい傾向にあった(図 3.14)。以上の結果より、クラス I-V の形状が尤もらしいと推定した。

3.6.2 シロイヌナズナ由来 P1L1 への適用

3.6.2.1 P1L1 分子形状推定の詳細

phototropin 1 は高等植物が有する青色光受容蛋白質であり、P1L1 は光受容・二量体形成ドメインとして機能する。phototropin は本研究における解析対象の 1 つであるため、詳細は後の章で述べることにし、ここでは先行研究 [Nakasako *et al.*, 2008]で得られている P1L1 結晶構造を示すのみに留める(図 3.15)。P1L1 はサブユニットとしても、二量体としても安定した構造をとっているが、結晶構造解析においては N-末端 6 残基と C-末端 25 残基は電子密度が消失しており、これらの領域は溶液中においてもフレキシブルであると考えられる [Nakasako *et al.*, 2008]。

P1L1 SAXS プロファイル [Nakasako *et al.*, 2004] (図 3.16)に対して、PDI と同様のプロセスによって分子最大長 100 Å、ダミー残基数 110(二量体全体で 220 残基)に決定した。決定されたパラメータに C2 対称性による拘束条件を加え、576 回の GASBOR 計算と提案手法による解析を行った。また、AMBIMETER による ambiguity score の評価も PDI と同様に行った。

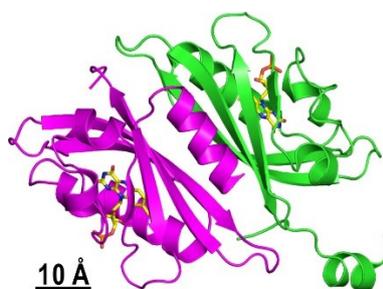


図 3.15 P1L1 結晶構造

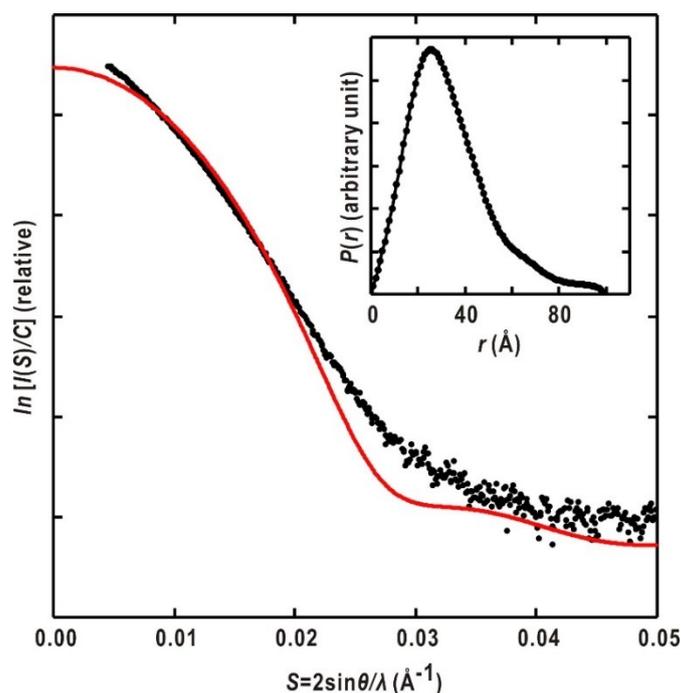


図 3.16 PIL1 SAXS プロファイル。赤い曲線は結晶構造から CRY SOL によって計算されたプロファイル。グラフ右上の挿入は GNOM によって計算された距離分布関数。

3.6.2.2 提案手法による PIL1 解析結果

PIL1 SAXS プロファイルの ambiguity score は 0.00 であり、不確かさのない極めてユニークなプロファイルであると評価された。GASBOR モデル形状はクラス I-III 付近を中心に分布しており、PDI の等方的かつ均一な分布とは異なる特徴を見せた (図 3.17)。また、各クラスの平均形状を比較すると平滑な楕円体型の I-VII と両端に突起が生えた形状の VIII-X の 2 種類の形状に大別されたが、そのうち I-VIII が全体の 80%以上を占めていた。PDI の多様な推定形状と比較して PIL1 推定形状はバリエーションに乏しく、これは散乱プロファイルの ambiguity score が低いことと相関していると考えられる。また、各クラス平均形状に対して PIL1 結晶構造を重ね合わせてみたところ、全てのクラスで平均形状の中心によく重ね合わせることができた(図 3.18)。結晶構造を重ね合わせても平均形状の両端に空白が生じたが、この空白は結晶構造解析で電子密度が現れなかった N-、C-末端と考えられる。実際、N・C-末端を欠いた結晶構造から CRY SOL によって計算された SAXS

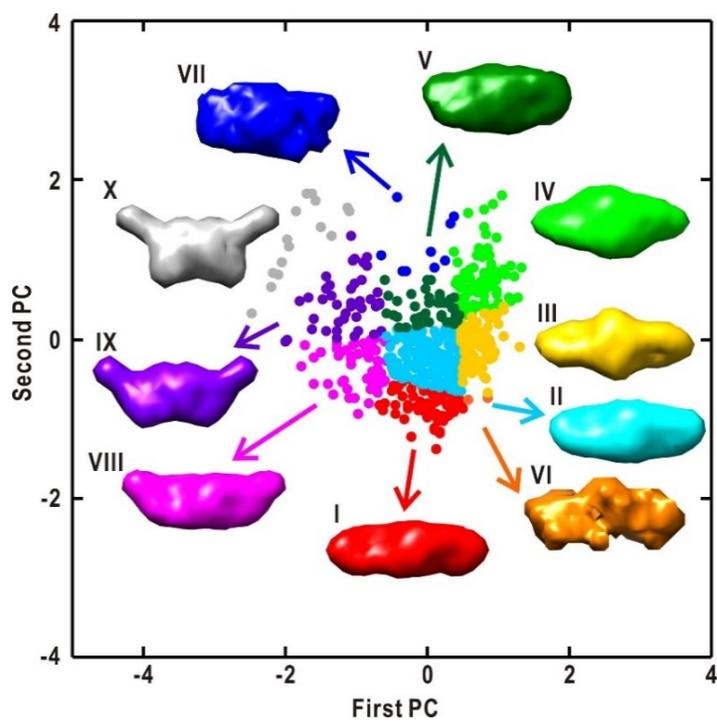


図 3.17 第 1, 第 2 主成分に射影された P1L1 モデル形状分布

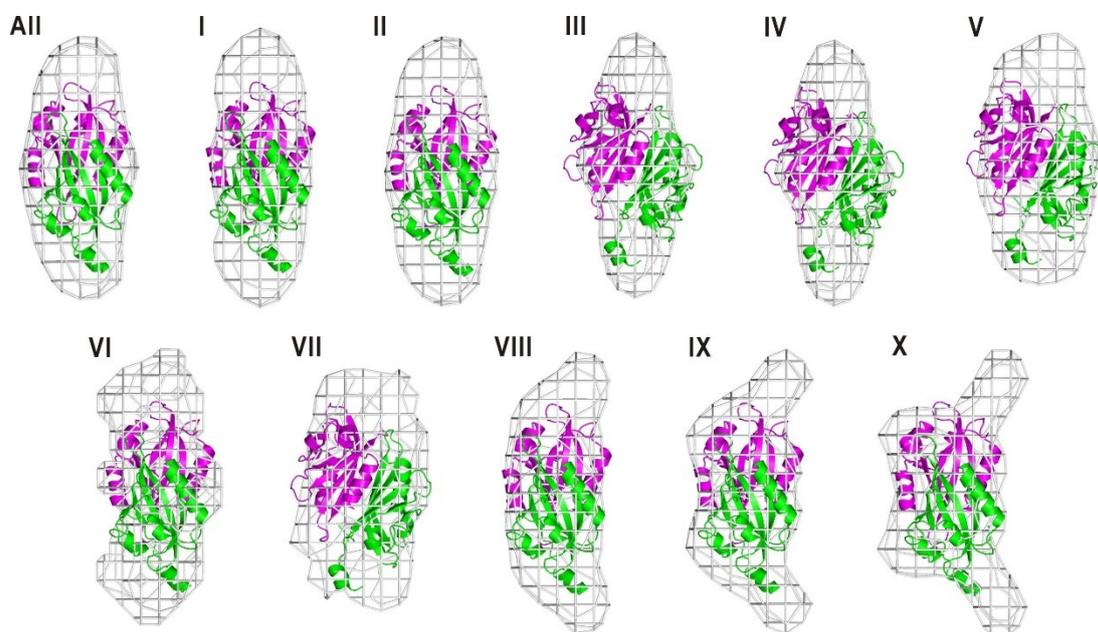


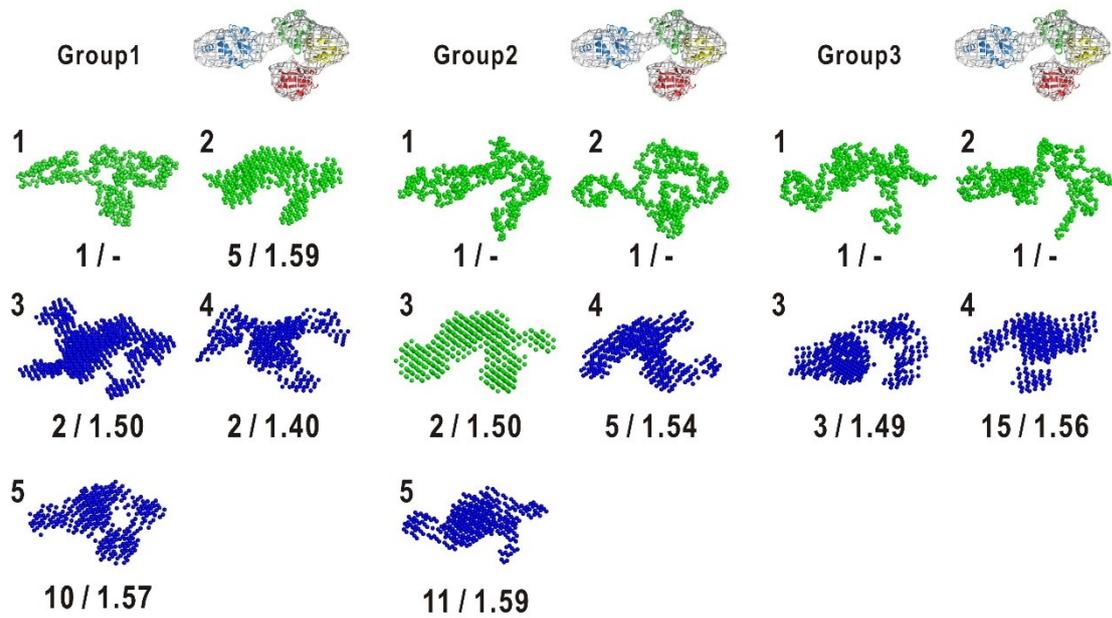
図 3.18 各クラスの平均形状と重ね合わせた P1L1 結晶構造。

プロファイルは実験データとは大きく異なるものであった。しかしながら、結晶構造にも現れなかった点を考慮すると N・C-末端領域は溶液中においても極めて揺らぎの大きい状態にあると考えられる。従って、クラス VIII-X のように空白領域が局在した形状は過剰に高分解能である可能性が高い。この点に加え、ambiguity score が低い中で 8 割以上が同様の形状に収束したことを考慮し、クラス I-VII の構造が尤もらしいと推定した。

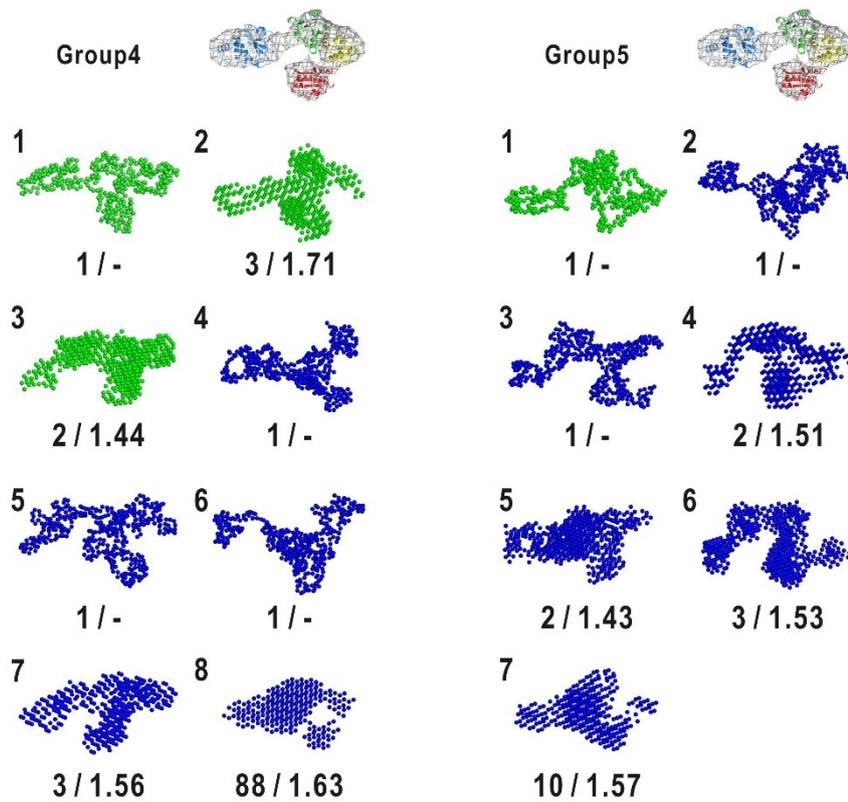
3.7 提案手法と既存手法との比較

3.6 節では数百個のモデルを提案手法によって分類することで蛋白質形状推定が可能であることを示した。本節では先行研究で提案されている手法との比較を通して提案手法の有用性を評価する。提案手法の特徴は①分類手法、②扱うモデル数の二点である。そのため、この二点に着目して比較していく。比較対象である既存手法には DAMCLUST を用いた [Petoukhov *et al.*, 2012]。DAMCLUST は現在までに提案されている手法の中では最も洗練された形状分類手法であり、階層的クラスタリングに基づいて NSD が小さいモデル同士を同じクラスと見なして同一クラスに統合していく手法である。クラス数を指定する必要がなく、大抵の場合 10 クラス未満に収束する。分子形状モデルは 3.6.1 節で解析した PDI のモデルを使用した。PDI は SAXS プロファイルの ambiguity score が大きく、空間対称性も有さない形状推定が困難な試料であるが、提案手法では尤もらしく形状分類ができた。そのため、このような試料に対して DAMCLUST がどの程度尤もらしく分類できるかによって提案手法の性能を測る。

まず、576 個の分子形状モデルから無作為に 20 個抽出したグループを 3 グループ作り (Group 1-3)、それぞれのグループについて DAMCLUST を適用した。20 個というモデル数は DAMCLUST が想定しているモデル数の上限である。20 個以上のモデルを処理することもできるが、その場合計算時間が非線形に増加する。図 3.19a は Group1-3 それぞれの分類結果であり、緑色が尤もらしい形状、青色が不適切な形状である。これを見ると、20 個のモデルは 5-4 クラスにしか分類されなかった。また、全モデルの 7 割以上が不適切な形状としてクラスタリングされ、残りのモデルも PDI の多様な形状の推定には不十分であった。この点の改善と推定形状の統計的有意性の保証には数百個のモデルを扱う提案手法が有効だと示唆された。



(a)



(b)

(c)

図 3.19 DAMCLUST によるモデル分類結果(次ページに続く)

(図 3.19 続き) (a)は 576 個のモデルからランダムに 20 個を選択した 3 グループ、(b)はランダムに 100 個を選択したグループ、(c)は図 3.13 の 10 クラスから各クラスの重心付近で 2 個ずつ選択したグループである。緑色が尤もらしい形状、青色が不適切な形状である。各グループ上部には比較のために提案手法で推定された PDI 分子形状モデルを載せている。各クラス下部の数字は左から所属モデル数、クラス内部での平均 NSD である。

次に、多数のモデル処理において DAMCLUST が提案手法と同等の分類が可能か評価するため、無作為に 100 個抽出したグループ(Group4)を作り、DAMCLUST を適用した(図 3.19b)。Group4 では Group1-3 よりもより提案手法に近い分類結果が得られると期待されたが、結果は Group1-3 と大差なく、さらに 9 割近くが 1 クラスに分類されるという非常に偏った分類結果となった。これは DAMCLUST のアルゴリズムそのものに起因する問題だと考えられる。また、計算時間についても比較を付け加えておく。提案手法が 576 個のモデルの分類に要した計算時間は 1 時間程度であったが、DAMCLUST では 100 個のモデルの分類に 20 時間程度を要した。従って提案手法と同程度のモデル数を DAMCLUST で扱うのは現実的でなく、また、無理に扱ったとしても Group4 の分類結果を見る限り適切に分類することは不可能と思われる。

最後に、DAMCLUST がモデル形状を適切に分類できるかという観点での評価を試みた。提案手法によって分類された 10 個のクラス(図 3.13)から各クラスの重心付近のモデルを 2 個ずつ抽出し、計 20 個のモデルから成るグループ(Group5)を作り DAMCLUST によって分類した(図 3.19c)。結果、Group5 は Group1-3 と異なり 7 クラスに分類された。この違いは形状の多様さが反映されるように恣意的なサンプリングを行ったことに起因すると考えられる。しかしながら、分離の結果自体は芳しくなく、ただ 1 つのクラスを除いて不適切な分類形状が得られた。

以上の結果から、提案手法は既存手法と比較して形状分類に優れており、特に PDI のようなプロファイルの ambiguity が大きい試料については強力な形状推定手法になりうると結論付けた。

3.8 議論

非経験的分子形状モデリングは蛋白質の溶液中での形状推定に極めて有用な手法であるが、推定形状の不確かさをどう扱うかは長年の課題であった。2012年に提案された DAMCLUST が一応の解決策となったものの、ambiguity の顕著な SAXS プロファイルから DAMCLUST によって適切な分子形状を推定することは依然として困難である。本研究では、数百個の分子形状モデルを扱うことで ambiguity の大きい SAXS プロファイルに対しても尤もらしい形状推定が可能であることを示した。本節では提案手法とその利用方法について議論する。

3.8.1 提案手法の有用性

提案手法による分類方法は既存手法である DAMCLUST と比較しても十分に有用な手法であり、特に計算速度の面で圧倒的に優れている。これは主にモデル重ね合わせのアルゴリズムに由来すると考えられる。DAMCLUST では NSD に基づいてクラスタリングを行うため、入力モデルによる全てのペアについて SUPCOMB による重ね合わせと NSD の計算を行う。従って、このプロセスでの計算量はモデル数 N に対して明らかに $O(N^2)$ で増加する。一方提案手法では、ある一つのモデルを参照として慣性主軸によって重ね合わせるため、計算量 $O(N)$ の高速な処理を可能にしている。

提案手法は分類性能においても DAMCLUST と遜色のない結果を示し、特に ambiguity の大きい PDI SAXS プロファイルについても適切な形状分類が可能であった。しかしながら、PCA によってモデル形状の分布を得るには少なくとも百個以上のモデルを扱う必要があり、PIL1 のような ambiguity のない試料について提案手法を適用するのはやや非効率的である。従って、ambiguity score を指標とした分類手法の使い分けが望ましい。3.6 節の結果から、推定形状の多様性は SAXS プロファイルの ambiguity score と明らかに相関している。そのため、形状分類手法の使い分けとしてプロファイルの ambiguity score を基準とすれば効率の良い形状推定が可能になると考えられる。例えば、PIL1 のように ambiguity score が小さい試料については 20 個程度のモデルを用意し DAMCLUST によって分類するのが効果的だと思われる。DAMCLUST は 20 個程度のモデルならば提案手法よりも計算時間が短く、さらにユーザーによるパラメータ入力を必要としないため極めて簡便に推定が可能である。一方、PDI のような ambiguity score が大きい試料につい

では提案手法を用いるのが最善である。3.7 節での比較から、PDI は数百個レベルで処理しなければその多様な形状を推定することができない。また、数百個のモデルを扱う場合、提案手法は分類性能、計算時間の両面で DAMCLUST と比較して圧倒的に性能が良い。以上の点から、ambiguity score を基準として、提案手法と既存手法それぞれを相補的に用いていくのが効果的であると考えられる。

3.8.2 提案手法の改善可能性

最後に、提案手法の改善可能性について触れておく。まずモデルの重ね合わせについてだが、現在の手法では一つのモデルに対してのみ重ね合わせを行うため分類結果は参照モデルに依存する。これについては 2-3 個のモデルを参照構造として重ね合わせ、その平均を利用するなどの処理で改善されることが考えられる。また、提案手法では次元削減に PCA を用いた。PCA はデータ点が分布する多次元空間上に直交基底を張り直す操作であるため、提案手法は多次元空間上の分布を二次元平面に切り出す線形な次元削減をしていることになる。超空間上での分布が線形な分布であるとは限らないため、非線形な次元削減手法[Coifman *et al.*, 2015; Bishop *et al.*, 1998; Maaten & Hinton, 2008]を用いることでより適切な分類が可能になると期待される。

4章 SAXS による植物蛋白質フォトトロピンの構造解析

前章で説明した蛋白質形状推定法を全体形状未知である蛋白質フォトトロピンに対して適用し、構造解析を行った [Oide *et al.*, 2016; Oide *et al.*, 2018]。本章でその詳細と結果を述べる。

4.1 青色光受容蛋白質フォトトロピン

4.1.1 植物の生理現象と光信号伝達

生物はその生命活動において光を様々な形で利用している。その中でも植物は光合成による炭水化物の産生で知られているように、とりわけ強く光に依存した生命活動を行っている。さらに、発芽や生長、花芽形成など、植物の生育に関わる重要な生理現象の多くは光信号を利用している[和田ら、2001]。これらの生理現象は光信号の検知、分子間相互作用、細胞内信号伝達など非常に洗練されたメカニズムによって制御されており、これら応答の初発過程を担っているのが光による物理刺激(信号)を検知し、他の蛋白質へ伝達することで光信号を生体内化学信号へと変換する光受容蛋白質である。

本章で解析対象とするフォトトロピン(phototropin: phot)は光受容蛋白質の一種であり、青色光のセンサーとして働く [Christie *et al.*, 1998; Christie, 2007]。phot が制御する生理現象は、光屈性[Sakai *et al.*, 2001] (図 4.1)、葉緑体光定位運動 [Kagawa *et al.*, 2001; Jarillo *et al.*, 2001] (図 4.2)、気孔開口 [Kinoshita *et al.*, 2001]など多岐にわたる。1880 年に Darwin 親子によって初めて植物の光屈性が報告され、その後 100 年以上後にシロイヌナズナ突然変異体の解析から光屈性を制御する蛋白質が同定され [Liscum *et al.*, 1995; Haula *et al.*, 1997]、光屈性(phototropism)にちなんでフォトトロピンと名付けられた。葉緑体光定位運動は植物細胞内の葉緑体が青色光照射に反応して移動する現象である。葉緑体の移動は細胞骨格を形成するアクチンフィラメントの伸縮によるものであること、F-アクチン結合蛋白質である THRUMIN 1 が phot との相互作用特性を持っており、phot の青色光受容に伴って THRUMIN 1 が活性化することなどが現在までに明らかにされている [Kong & Wada, 2011]。

気孔開口は phot が関わる現象の中で生化学的信号伝達経路の解明が最も進んでいる現象であり、phot の青色光受容が最終的に孔辺細胞中に存在する膜蛋白質

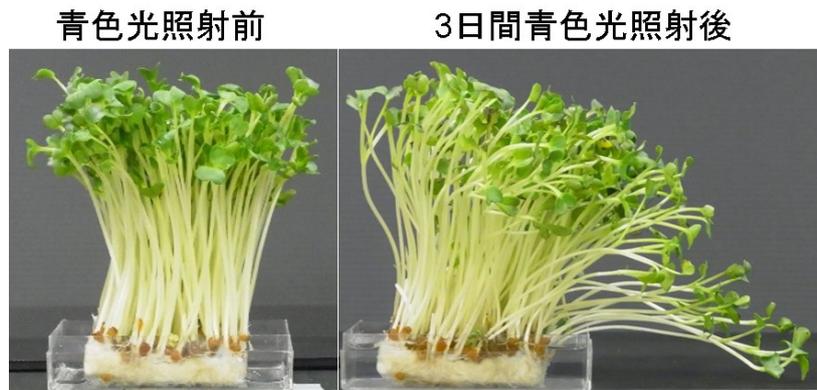


図 4.1 カイワレダイコンの光屈性。青色光照射前の画像(左)と向かって右側から青色光を3日間照射した後の画像(右)である。

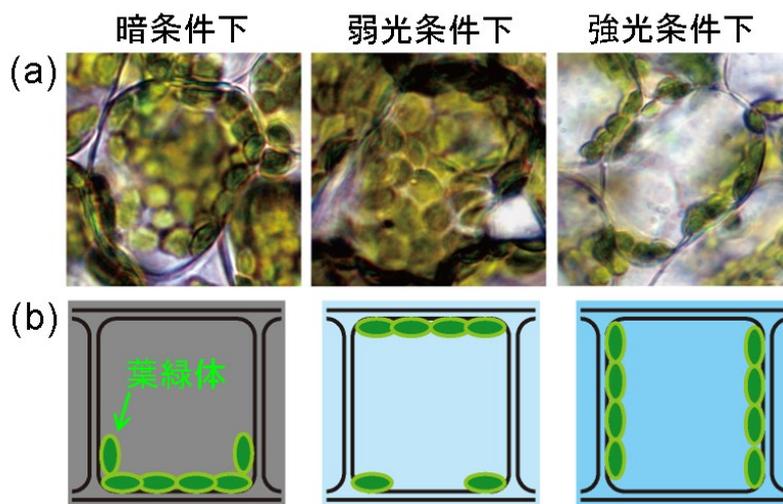


図 4.2 一個の植物細胞について葉緑体光定位運動を観察した光学顕微鏡画像(a)とその断面図(b)。暗条件下では細胞底部に貼りついているが(左図)、青色光を照射されると効率的に光合成を行えるよう細胞上部に集まる(中央図)。これを集合運動と呼ぶ。ここから青色光強度を上げると、集合反応とは対照的に青色光を避けるように細胞壁に貼りつく(右図)。これを逃避運動と呼ぶ。逃避運動は過剰な強度の光により光合成を行う光化学系が損傷することを避けるための生理現象と考えられている。

P-type H⁺-ATPase の活性化を引き起こすことが明らかにされている [Kinoshita & Shimazaki, 1999] (図 4.3)。さらに phot の青色光受容と H⁺-ATPase の活性化を繋ぐ信号伝達には BLUS1 と呼ばれる蛋白質が介在し、phot と BLUS1 が直接相互作用をすることも判明しているが [Takemiya *et al.*, 2013]、より詳細なメカニズム、例えば phot と BLUS1 の原子分解能での相互作用メカニズムや H⁺-ATPase 活性化へと至る分子間相互作用経路等の解明されていない。そのため、phot を起点とする青色光応答のメカニズムは依然として未解明な部分が多い。

シロイヌナズナを含むほぼ全ての高等植物はアミノ酸配列の異なる 2 種類の phot(phot1, phot2)を有している [Christie, 2007]。これら 2 つは同一の機能を有しているものの、青色光強度に対する感度が異なっている。phot1 はあらゆる強度の青色光に対して感受性を持つ一方で、phot2 は高強度の青色光下で機能する。そのため、葉緑体光定位運動の中でも強光に対して葉緑体が逃避する現象 (逃避運動; 図 4.2) は phot2 が特異的に制御している [Kagawa *et al.*, 2001]。

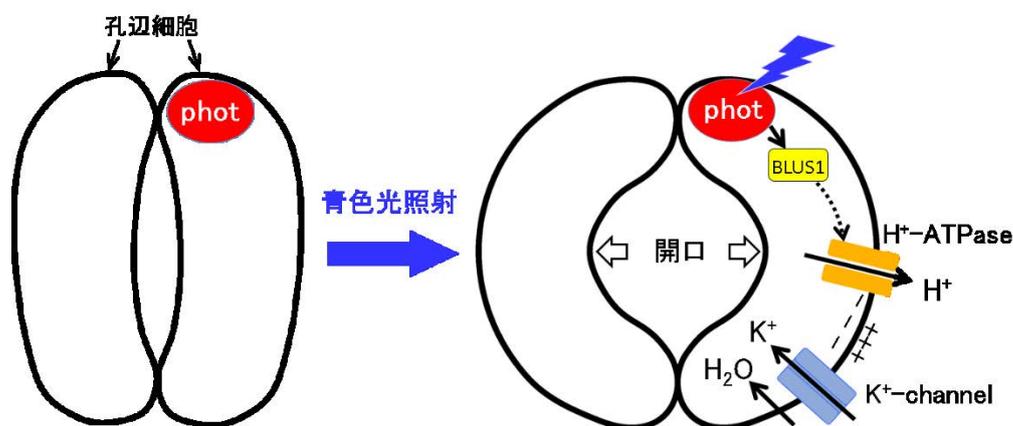


図 4.3 気孔開口模式図。孔辺細胞中に存在する phot が青色光を受容することで H⁺-ATPase が活性化されて細胞内のプロトンが細胞外へ輸送される。プロトン輸送に伴う膜電位の過分極はカリウム・チャンネルを介したカリウムイオンの取り込み、次いで浸透圧上昇による水の流入を招きこれによって孔辺細胞が膨張・変形する。phot の青色光受容から H⁺-ATPase の活性化までは複数の信号伝達過程を経由していると考えられているが、BLUS1 以外の詳細は不明である。

4.1.2 シロイヌナズナ phot のドメイン構造と機能 [Okajima, 2016]

本節以降特に断りのない限りシロイヌナズナ由来 phot についてのみ述べる。phot1, phot2 はそれぞれ 996 残基、915 残基のアミノ酸から成り (図 4.4)、共に 2 つの LOV(Light-Oxygen-Voltage Sensing: LOV)ドメイン(LOV1, LOV2)と 1 つのセリン・スレオニンキナーゼ(Serine/Threonine Kinase: STK)ドメインから構成されている(図 4.5)。

4.1.2.1 LOV ドメインの立体構造

LOV ドメインは、生体内信号伝達において分子間相互作用を担う一連の蛋白質群 Per-Arnt- Sim(PAS)スーパーファミリーの一種であり、その中でも特に光、酸素、電位の変化を検知するサブファミリーを構成する。また、phot1、phot2 の双方で LOV1、LOV2 の結晶構造が明らかにされており [Nakasako *et al.*, 2008; Christie *et al.*, 2012; Halavaty *et al.*, 2013] (図 4.6)、 phot1 LOV1、phot2 LOV1、及び phot1 LOV2 は二量体を形成することが明らかにされている。図 4.6 に示す様に LOV ドメインは 5 本の反平行 β シートと数本の α ヘリックスで構成されたポケット状の構造 (α/β -scaffold)を有し、このポケット内部にフラビンモノヌクレオチド(flavin mononucleotide: FMN)を非共有結合的に格納している(図 4.7)。FMN の発色団であるイソアロキサジン環 π -電子系の固有状態間遷移エネルギーが青色光エネルギーに対応するため、LOV ドメインは青色光受容ドメインとして機能する[Möglich *et al.*, 2009; Möglich *et al.*, 2010]。

4.1.2.2 LOV ドメインの光反応

暗中での LOV ドメインは 450 nm に極大を持つ吸収スペクトルを示し、D₄₅₀ と呼ばれる状態にある[Christie *et al.*, 1999]。FMN が青色光を吸収すると、ポケット内部のシステイン残基と共有結合を形成し(cysteinyll adduct)、吸収スペクトルが 390 nm に極大を持つ形状に変化する[Salomon *et al.*, 2000; Swartz *et al.*, 2001] (図 4.8a)。この状態の LOV ドメインを S₃₉₀ と呼ぶ。青色光受容による S₃₉₀ への遷移は青色光受容から数マイクロ秒以内に起きることが phot2 LOV2 のフーリエ変換赤外分光法(Fourier Transform Infrared Spectroscopy: FTIR)による測定から明らかにされている[Pfeifer *et al.*, 2009]。一方で、S₃₉₀ は数十秒から数分の時間スケールでシステイン - FMN 間の共有結合が切断され再び D₄₅₀ へ戻る[Kasahara *et al.*, 2002]。その

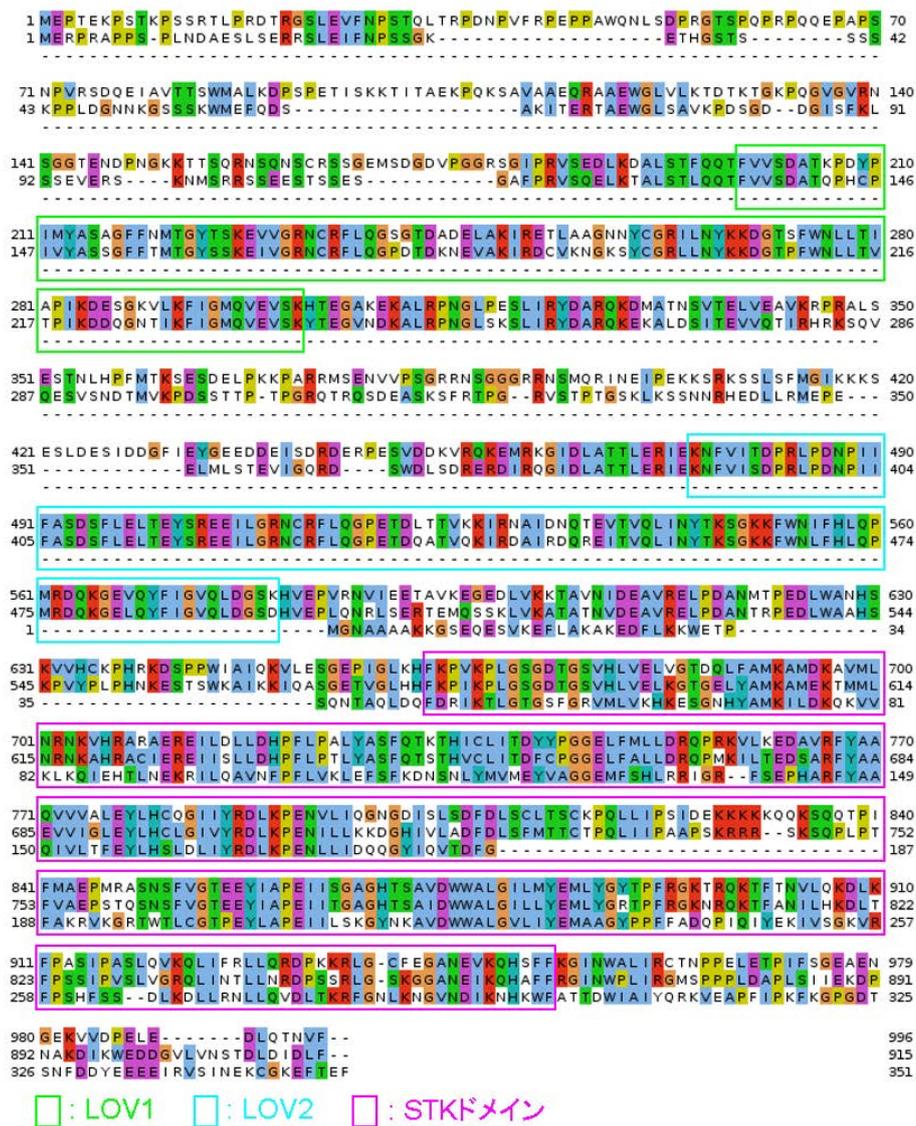


図 4.4 phot1 (上段)、phot2 (中段)、ハツカネズミ由来プロテインキナーゼ A (下段) のアミノ酸配列。類似した配列部分が重なるように並べており、着色された部分は類似性の高い領域である。類似配列の重ね合わせはソフトウェア ClustalW [Thompson *et al.*, 1994; Chenna *et al.*, 2003] によって計算し、ソフトウェア Jalview2 [Waterhouse *et al.*, 2009] にて表示した。アミノ酸の性質毎に着色される色が異なり、負電荷を帯びるアミノ酸はマゼンタ、正電荷を帯びるアミノ酸は赤、疎水性アミノ酸は青～緑系の色で着色されている。配列を囲む四角形は phot1 を基準としたドメインを構成する領域である(図 4.5 も参照)。

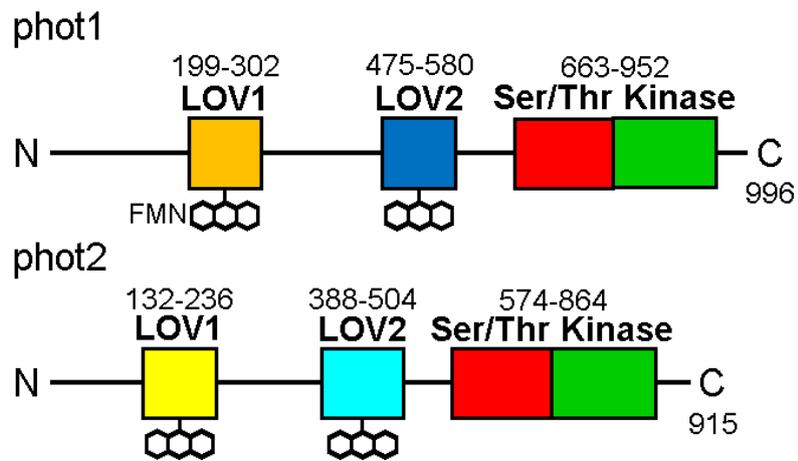


図 4.5 シロイヌナズナ由来 phot1(上)、phot2(下)の一次構造上の機能ドメイン配置。長さは各ポリペプチド鎖を構成する残基数に対応している。

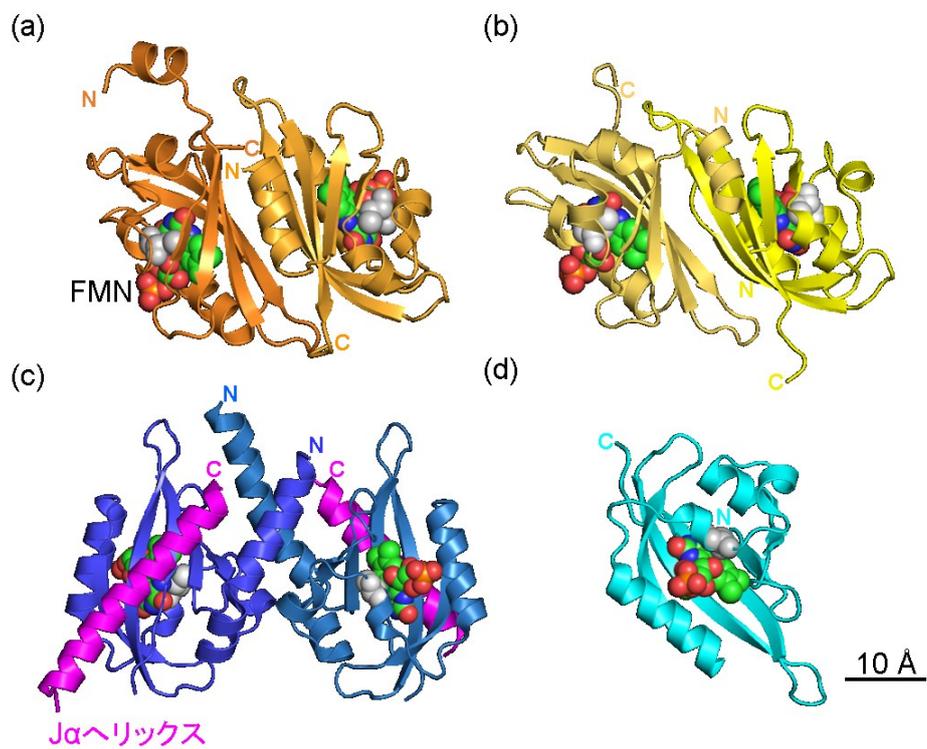


図 4.6 LOV ドメイン結晶構造(次ページに続く)

(図 4.6 続き) X 線結晶構造解析により得られた phot1 LOV1(PDB ID: 2Z6C)(a)、phot2 LOV1(PDB ID: 2Z6D)(b)、phot1 LOV2(PDB ID: 4HHD)(c)、phot2 LOV2(PDB ID: 4EEP)(d)の結晶構造。LOV ドメイン内の緑色(炭素)、赤色(酸素)、青色(窒素)の球で表示された分子が FMN である。また、LOV ドメイン骨格中の灰色の球は FMN と共有結合を形成するシステイン残基である。図 4.6c にてマゼンタで表示されたヘリックスは $J\alpha$ ヘリックスと呼ばれる LOV2 - STK ドメイン間に存在するヘリックスである。右下のスケールバーは全構造モデルに共通である。

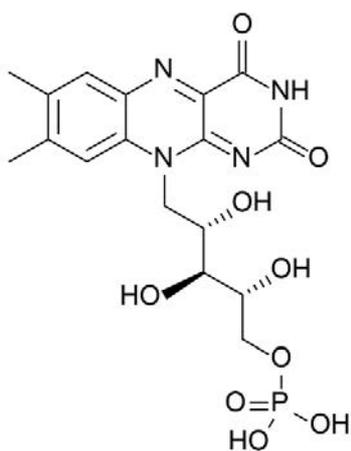


図 4.7 フラビンモノヌクレオチド構造式。炭素原子と結合した水素原子を省略して書かれている。図上部の 3つのベンゼン環が連なったような構造がイソアロキサジン環である。

ため、LOV ドメインは青色光照射に対して可逆的な光反応サイクルを示す(図 4.8b)。光反応の可逆性より、LOV ドメインは青色光照射によって D_{450} と S_{390} の光平衡状態に到達するが、平衡定数は青色光の強度に依存するため十分な強度の青色光を照射することで D_{450} の存在を無視できる程度まで S_{390} を蓄積させることが可能である。

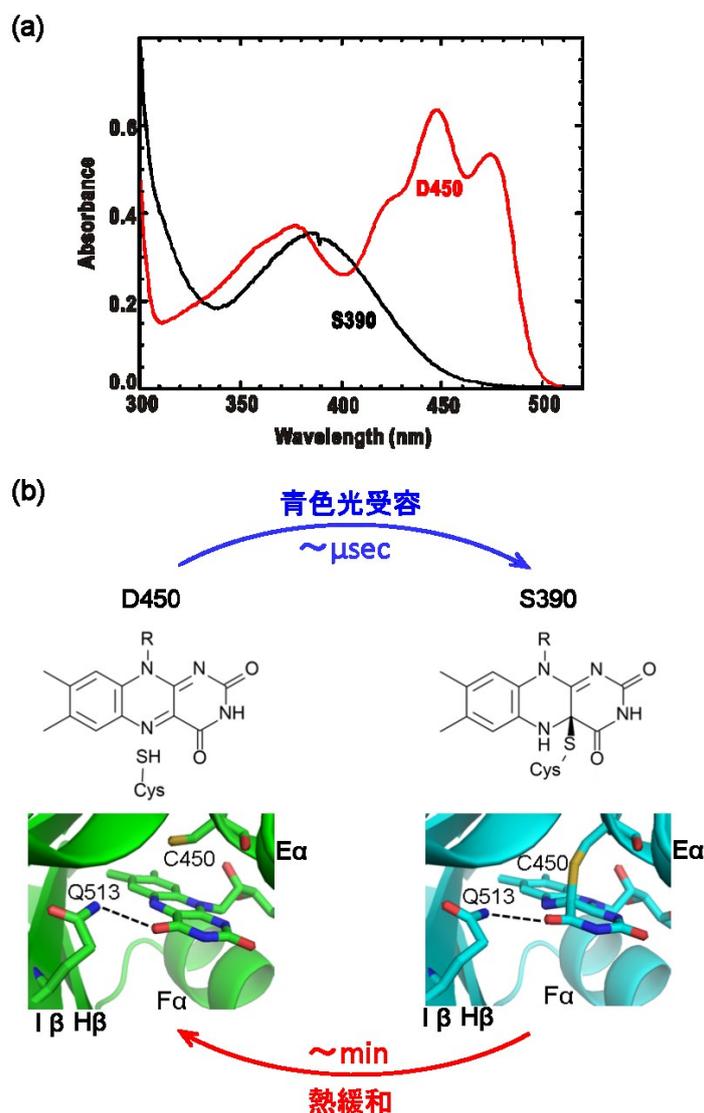


図 4.8 LOV ドメインの D₄₅₀、S₃₉₀ における吸収スペクトル(a)と光応答サイクルの模式図(b)。図 4.9a は phot1 LOV2 - STK 断片の吸収スペクトルであるが、LOV2 単独の吸収スペクトルとほぼ同一であることが確認されている。図 4.9b において、各状態の上段画像が FMN、LOV2 システイン残基の構造式、下段画像が対応するカラスムギ由来 phot1LOV2 結晶構造の拡大図である(D450: 2V0U、S390: 2V0W) [Halavaty & Moffat, 2007]。システイン基(C450)と FMN の共有結合により LOV ポケット内部の水素結合環境が変化しており、例えば FMN - グルタミン基(Q513)の相対位置が変化している。

4.1.2.3 STK ドメインの構造と機能

STK ドメインは蛋白質のセリン、スレオニン残基のヒドロキシ基をリン酸化する酵素（キナーゼ）として機能する。構造未知であるものの、アミノ酸配列が動物由来プロテインキナーゼ A(Protein Kinase A: PKA)に類似していることから、PKA に類似した構造を有するものと考えられている。また、PKA と STK ドメイン間のアミノ酸配列の相関性、相似性を利用した予測構造も得られている [Takayama *et al.*, 2011] (図 4.9)。PKA は N 末端側と C 末端側で 2 つのサブドメインに分かれており、それぞれ N ロープ、C ロープと呼ばれる。2 つのロープの間隙(クレフト)に活性中心が存在し、リン酸基の供給源となるアデノシン三リン酸マグネシウム (Adenosine TriPhosphate Magnesium: Mg-ATP) を結合する ATP 結合サイトを有している [Kornev *et al.*, 2006]。一般に蛋白質がリン酸化されると、リン酸基の持つ負電荷によって蛋白質の相互作用特性も変化することから(図 4.10)、STK ドメインはリン酸化による蛋白質間相互作用の制御、即ち蛋白質間での信号伝達のスイッチを担っていると考えられている。

STK ドメインには phot 自身をリン酸化する能力があること [Sakai *et al.*, 2001; Christie *et al.*, 1998; Christie *et al.*, 2002] が知られている。また、STK ドメインには activation loop と呼ばれるループ構造が存在し、phot1 では activation loop 中の 851 番セリン残基のリン酸化が機能発現に必要であることが *in vivo* で確認されている [Inoue *et al.*, 2008]。これらの生化学的知見から、phot の STK ドメインが STK ドメイン自身をリン酸化(自己リン酸化)すると考えられている。しかしながら、STK ドメインの自己リン酸化が分子内・分子間いずれの反応であるかは現在明らかになっていない。

4.1.2.4 リンカー領域

phot は LOV、STK ドメインとは別に、LOV2 と STK ドメインを繋ぐ特別なリンカー領域と呼ばれる領域を有している。リンカー領域のアミノ酸配列はドメイン領域の配列と比較して構成アミノ酸の種類が少なく、同一アミノ酸が連続して並ぶ頻度が高いという特徴があり、このような特徴を持つ配列は低複雑性配列と呼ばれている [西川, 2009]。低複雑性配列領域は構造が「本質的に不規則(intrinsic disorder)」であるとされており、実際にリンカー領域は不規則領域である可能性が高いことが配列解析から予測されている [Takayama *et al.*, 2011]。不規則領域は X

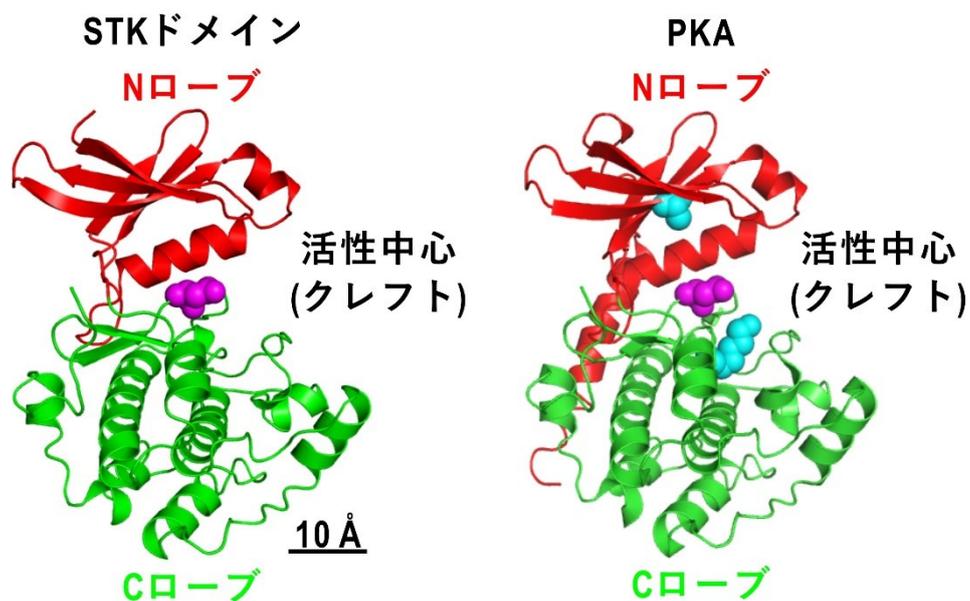


図 4.9 STK ドメイン予測構造モデル(左)。この構造は配列の類似した(図4.4)ハツカネズミ由来PKA結晶構造 [Akamine *et al.*, 2003; PDB ID: 1J3H] (右)を参照構造として予測されたホモロジーモデルである [Takayama *et al.*, 2011]。2つのドメインの間隙(クレフト)に活性中心が存在し、クレフトにリン酸化の対象となるアミノ酸残基と Mg - ATP を挟み込むことで酵素としての役割を果たす。PKA 結晶構造中に球で表示されている残基が Mg-ATP と相互作用し、ATP 結合サイトとして機能するアミノ酸残基である。マゼンタがアスパラギン酸、シアンがリジンである。STK ドメインについては 4.2.1 で述べる 720 番アスパラギン酸のみ表示している。

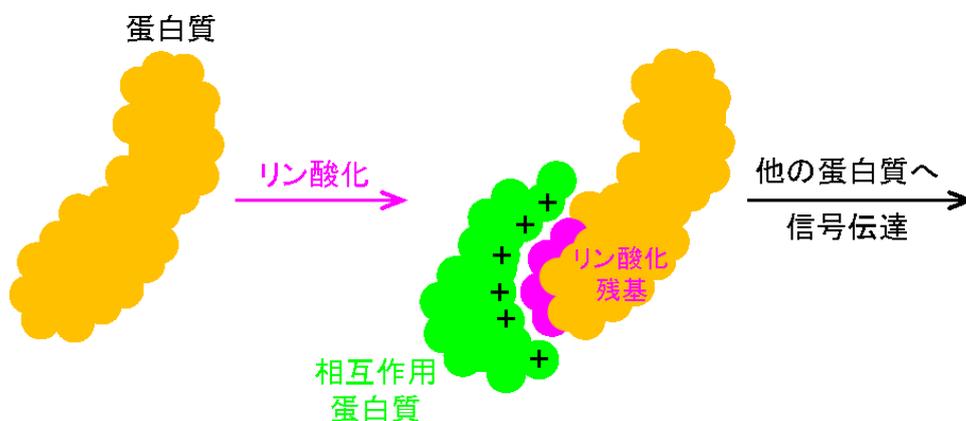


図 4.10 蛋白質のリン酸化と信号伝達(次ページに続く)

(図 4.10 続き)リン酸化された残基は負電荷を帯びるため、正電荷を持つ蛋白質と相互作用し複合体を形成する。この蛋白質複合体が他の受容体蛋白質に結合するなどして信号が伝達されていく。この他にもリン酸化は疎水性の残基を親水化させるため、蛋白質の構造変化をも起こしうる。蛋白質のリン酸化は生体内信号伝達において重要な役割を果たしうる化学反応である。

線結晶構造解析において電子密度が観察されないことが多く、そのため多型的な構造を持つか、あるいは定まった構造を持たないと考えられている。また、不規則領域が蛋白質間相互作用において重要な役割を持ちうることが核磁気共鳴法 (Nuclear Magnetic Resonance: NMR)を用いた研究で明らかにされており [Wright & Dyson, 1999]、phot についてもリンカー領域が機能発現する上で重要な役割を担っている可能性が高い。

4.1.2.5 LOV ドメインの役割

phot は 2 つの LOV ドメインを有しているが、機能・構造共に両者は大きく異なる。シロイヌナズナの phot1 と phot2 について、暗中では LOV2 によって STK ドメインの活性が抑制されており、青色光によって STK ドメインの活性が増加するということが分光測定とリン酸化活性測定から明らかにされている [Christie *et al.*, 2002]。この実験事実から、LOV2 - STK 断片が phot の機能発現に必要な最小単位であると思われている。一方、LOV1 は STK ドメインの活性制御に直接関わってはならず、むしろ phot の青色光センサーとしての感度を制御するドメインであることが示唆されている [Christie *et al.*, 2002; Cho *et al.*, 2007]。また、phot1 LOV1、phot2 LOV1 共に結晶構造が二量体であることからシロイヌナズナ由来 phot においては二量体サイトとしても機能していると考えられている。さらに、LOV1 と異なり、LOV2 は C 末端に連なる両親媒性ヘリックスを有している。J α ヘリックスと呼ばれるこのヘリックスは、phot1 LOV2 結晶構造(図 4.6c)にてその存在を確認できる。phot2 LOV2 についても、結晶構造で J α ヘリックスを有するカラスムギ由来 phot1 LOV2 との配列類似性より、ほぼ確実に J α ヘリックスを有すると予測されている [Takayama *et al.*, 2011]。J α ヘリックスはリンカー領域と同様に不規則領域と

予測分類されており、周辺環境変化の影響を受けやすいと考えられている。

4.1.3 先行研究から予測される機能発現メカニズム

4.1.2 節で説明したように、暗中には phot のリン酸化活性は LOV2 によって抑制されており、LOV2 の青色光受容がスイッチとなってリン酸化酵素として機能発現する。この青色光受容による制御メカニズムについて、構造生物学的観点から様々な研究がなされてきた。

S₃₉₀ ではシステイン残基との共有結合により LOV2 内部でのイソアロキサジン環の相対配向が変化し、FMN 周囲の水素結合環境が変化する [Crosson & Moffat, 2001; Crosson & Moffat, 2002] (図 4.8b)。このポケット内部で生じた摂動により、LOV2 ドメインの微小な全体構造変化が誘発されることが、カラスムギ由来 phot1 LOV2 の X 線結晶構造解析により確認されている [Halavaty & Moffat, 2007]。FTIR を用いた測定によると、phot2 LOV2 ドメインでは S₃₉₀ へ遷移した後、数百マイクロ秒で α ヘリックス及びループ構造の変化、さらに数ミリ秒後に β シート構造の変化が生じる [Pfeifer *et al.*, 2009]。S₃₉₀ への遷移により生じるこの構造変化は極めて微小であるものの、これに伴い LOV2 近傍の J α ヘリックスは部分的にアンフォールドすることが NMR 解析により示唆されている [Harper *et al.*, 2003]。この J α ヘリックスの大規模な構造変化はその不安定性に由来するものであろう。これらの事実から、青色光を受容した LOV2 の微小な構造変化が J α ヘリックスのアンフォールドといったより大規模な構造変化を誘発し、STK ドメインの活性に影響を与えているという機能発現メカニズムが想起される。

このメカニズムに基づくモデルとして、LOV2 が STK ドメインのクレフトを直接阻害しており、青色光受容による LOV2 ドメインの構造変化と J α ヘリックスのアンフォールドによって LOV2 の解離・STK ドメインの活性化が誘発されるというモデルが 2010 年に提案された [Pfeifer *et al.*, 2010] (図 4.11)。しかしながら、このモデルは FTIR 測定及び、LOV2 - STK ドメイン相互作用を提案した仮想的構造モデル [Tokutomi *et al.*, 2008] を下地としており、構造及び構造変化の直接観察に基づかない信憑性に欠けるモデルであった。

一方で、SAXS による phot2 LOV2-STK 断片の構造解析結果 [Takayama *et al.*, 2011] に基づき、LOV2 が STK ドメイン活性クレフトを直接阻害しないモデルが 2011 年に提案された (図 4.12)。このモデルでは、LOV2 は暗条件下で STK ドメ

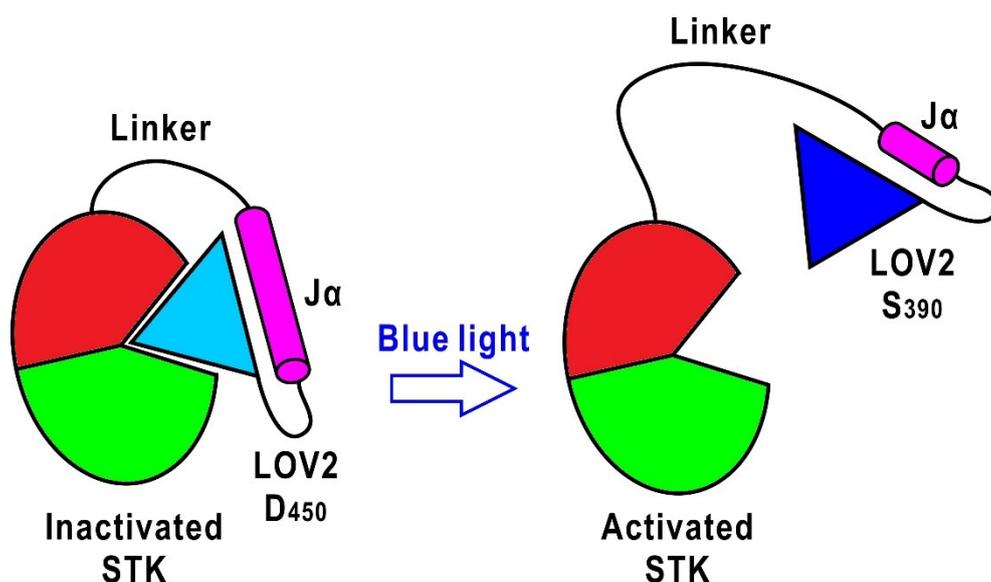


図 4.11 クレフト直接阻害による phot 機能制御モデル(Pfeifer *et al.*, 2010)。暗条件下では LOV2 が STK ドメインのクレフトと相互作用することでキナーゼ活性を阻害している。LOV2 が青色光を受容して S₃₉₀ へ遷移すると、LOV2 の構造変化により STK との相互作用が変化し、J α ヘリックスのアンフォールドと併せて LOV2 が STK ドメインから解離する。

ンの N ロープに接するように配置されており、リンカー領域は LOV2 - STK ドメインの境界面でコンパクトに折り畳まれていると推測している。LOV2 ドメインが S390 へ遷移すると、J α ヘリックスのアンフォールドのみならずリンカー領域の構造変化をも誘発し、これらの構造変化によって LOV2 ドメインが STK ドメインから引き離され、STK ドメインが活性化される。このモデルでは、リンカー領域が分子間シグナル伝達を担っていると解釈している。SAXS による分子形状の直接観察に基づくモデルであること、従来のクレフト阻害モデルでは SAXS プロファイルを説明できないこと [Okajima, 2016] 等から、現在ではこのモデルが分子内信号伝達メカニズムをよりよく説明するものと考えられている。

4.1.4 本研究の狙い

以上の背景より、phot については以下の事実が認められる。

(i) phot は青色光受容蛋白質であり、光屈性などの生理現象を制御するが、その信

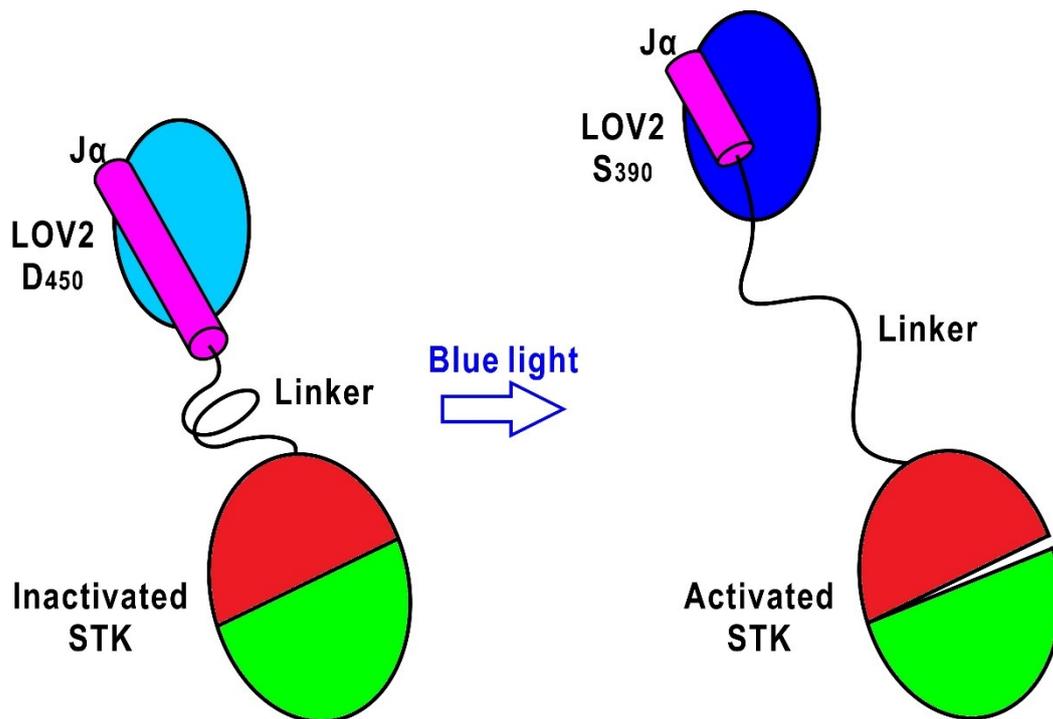


図4.12 LOV2ドメインがSTKドメインを間接的に制御するフォトトロピン機能発現モデル[Takayama *et al.*, 2011]。このモデルではLOV2が青色光によってS390に励起されるとJ α ヘリックスのアンフォールドやリンカー領域のコンフォメーション変化を引き起こし、これらの構造変化によりSTKドメインが活性化される。

号伝達メカニズム、特に初発過程である青色光信号を蛋白質相互作用に変換するメカニズムは明らかにされていない。

- (ii) phot1、phot2共にドメイン単位での構造が解明・予測されており、二量体であると予想されている(図4.13)。
- (iii) フォトトロピンの青色光受容による機能発現においてはLOV2、STKドメインがその中核を担っている。
- (iv) LOV2及びその周辺の青色光に依存した構造変化はX線結晶構造解析やNMRを用いて原子分解能で明らかにされており、これらの結果からphot分子内信号伝達メカニズムが推定されている。
- (v) 信号伝達メカニズムを説明するモデルとして、LOV2-STK間の相互作用がリンカー領域によって制御されている、とするモデルが提案されている。このモデルはSAXSで測定された分子形状を説明できるモデルであり、信憑性も高い。

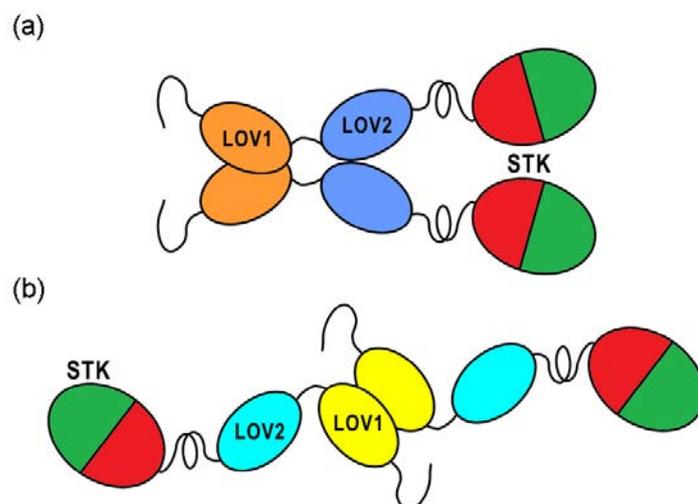


図 4.13 結晶構造から予測される phot1(a)、phot2(b)の全体形状予想図。phot1 では LOV1、LOV2 がそれぞれ二量体を形成するため、全体では図のような二量体を形成すると考えられる。phot2 では LOV1 のみが二量体を形成しているため、LOV1 のみによって二量体を形成していると考えられる。先行研究でも phot2 LOV2 - STK 断片では溶液中で単量体として存在していると報告されている [Takayama *et al.*, 2011]。

しかしながら、生体内での phot と同様の二量体、特に LOV1 - LOV2 - STK（全長）での全体形状解析、及び青色光照射に伴う構造変化を測定した例は未だ存在しない。そこで、本研究では phot1 の機能最小フラグメントである LOV2 - STK 断片及び全長 phot2 について X 線小角散乱法と 3 章で提案した分子形状推定法を用いて全体形状推定と青色光による構造変化を解析し、phot における分子内信号伝達機構を構造変化の観点から研究した。以下の節で測定、解析及び得られた結果について述べる。

4.2 SPring-8 における X 線小角散乱実験

4.2.1 測定試料

測定試料の発現・及び精製は大阪府立大学・理学系研究科・生物化学専攻・環境応答制御科学分野・生体光制御研究室(当時)の岡島公司博士からご提供いただいた。phot1 LOV2 - STK 断片は 20 mM Tris-HCl、100 mM NaCl、10% グリセロール、

1 mM EGTA を含む緩衝溶液(pH 7.4)に溶解させた。全長 phot2 は 20mM HEPES、200 mM NaCl、10% グリセロールを含む緩衝溶液(pH 7.5)に溶解させた。また、慣性半径及び原点散乱強度の濃度依存性を解析するため(2.節参照)、各試料について濃度を変えて複数回の SAXS 測定を行った(表 4.1)。

表 4.1 各試料の測定濃度

試料	濃度
phot1 LOV2 - STK 断片	0.50 mg/mL, 0.80 mg/mL, 1.60 mg/mL, 2.30 mg/mL
全長 phot2	0.50 mg/mL, 1.00 mg/mL, 1.25 mg/mL

4.1.2 節で触れた通りリン酸化は蛋白質に大きな構造変化を起こし得る。よって、発現後および精製過程での自己リン酸化による構造変化を避けるため、phot2 について STK ドメインの 720 番アスパラギン酸残基(Asp720)をアスパラギンに置換した点変異体を用いた。図 4.9 に示す様に Asp720 は ATP 結合サイトを構成する残基であり、この残基がアスパラギンに置換されると Mg-ATP との結合能を失い、キナーゼとして失活する[Takayama *et al.*, 2011]。一方、アスパラギン酸とアスパラギンは先端部の酸素原子一個が窒素原子に置き換わっているだけなので、Asp720 の置換が STK ドメインの構造に及ぼす影響は極めて小さいと考えられる。そのため、本研究ではこの Asp720 変異体を野生型とほぼ同一であると見なしている。

4.2.2 X 線小角散乱実験

SAXS 測定は、SPring-8 ビームライン BL45XU の小角散乱実験ハッチで実施した(図 4.14)。シンクロトロン蓄積リングで発生した放射光は、ダイヤモンド製 2 結晶モノクロメータによって単色化され、ビームラインに導かれる。Kirkpatrick-Beaze 配置の X 線ミラー (K-B ミラー) によって集光された X 線はビーム形状整形スリット、寄生散乱除去スリットを経て試料に入射する。このようにして得られた入射 X 線は、検出器位置でのビームサイズがおよそ縦 100 μm ×横 450 μm であった。また、試料位置の直前にはイオンチャンバーが挿入されており、入射強度を測定することで得られた SAXS プロファイルごとに強度の規格化が可能である。蛋白質溶液試料セルは容量 30 μL 、光路長 3 mm であり、窓材には 10 μm 厚の石英薄板

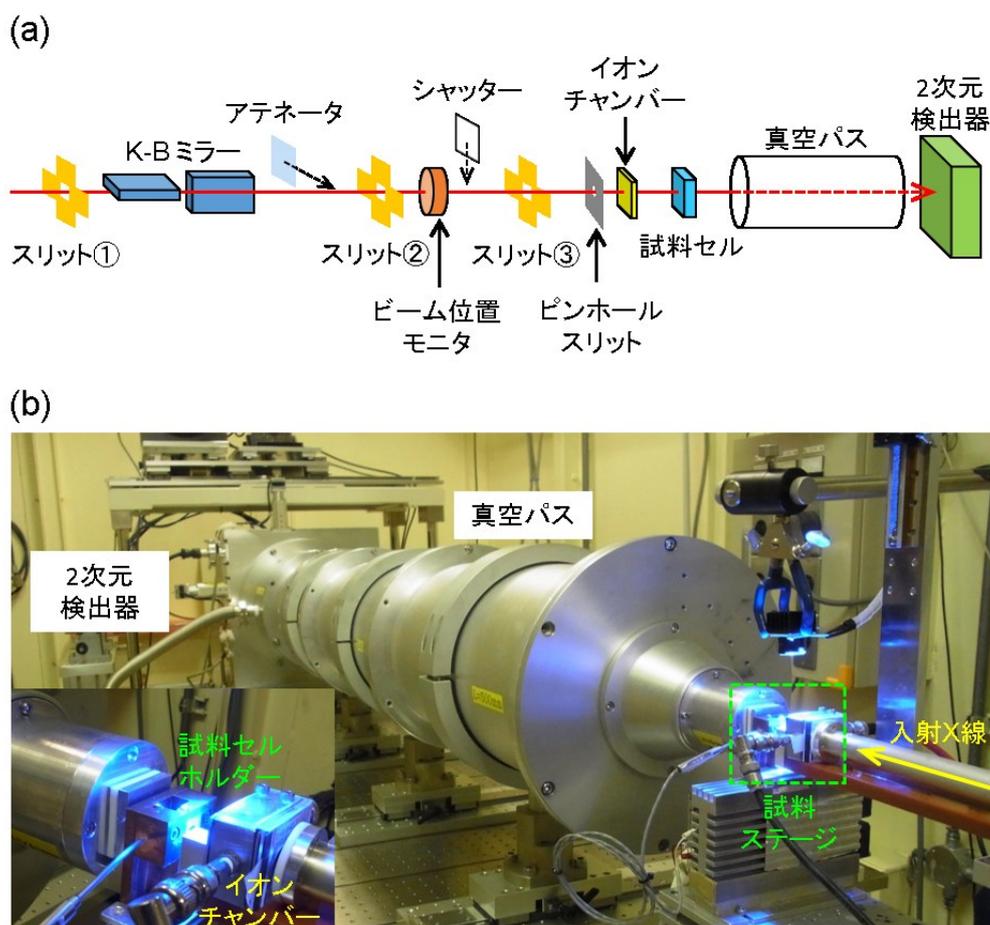


図 4.14 BL45XU の光学系模式図(a)と実際の実験ハッチ内の写真 (b)。4.14a においてはモノクロメータより下流のみの図となっている。4.14b 左下の写真は四角形で囲まれた部分の拡大写真である。スリット①の役割はビーム形状の整形、スリット②、③及びピンホールスリットは寄生散乱除去の役割を持つ。試料セル - 検出器間の真空パスは試料蛋白質溶液からの微弱な散乱を検出するため 10^{-2} Pa 程度の真空に保たれている。

を使用している(図 4.15)。phot1 LOV2 - STK 断片、全長 phot2 共に、波長 1.0000 Å、カメラ長 2.5 m で測定を実施した。

散乱 X 線検出器には 2 次元ピクセルアレイ検出器 Pilatus 300K-W (DECTRIS, Switzerland)を使用した(図 4.16)。この検出器の最大の特徴は、1 ピクセルごとに入射した光子数を数ミリ秒で高速読み出しできるという点にある。1 フォトンず

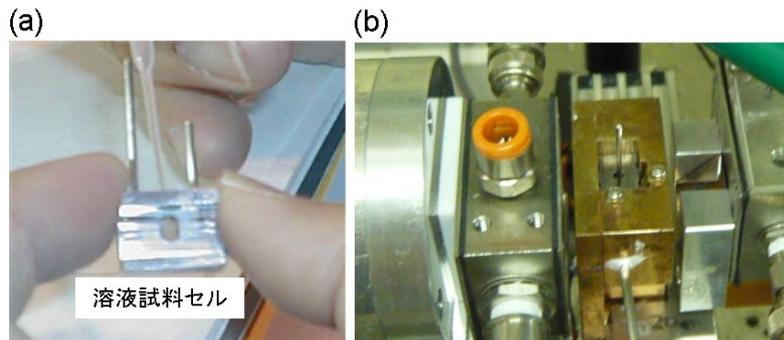


図 4.15 溶液試料セル(a)と試料セルホルダーにセットした状態(b)。試料セルへの注入部は試料溶液の蒸発を妨げるため非常に小さい。そのため試料注入には径の小さいゲルロード用のマイクロチップを使用する。

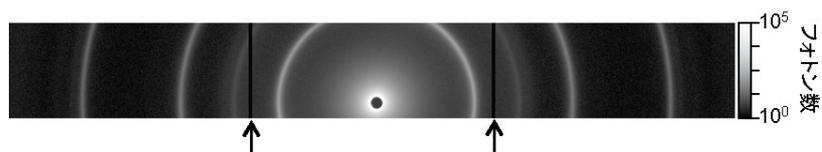
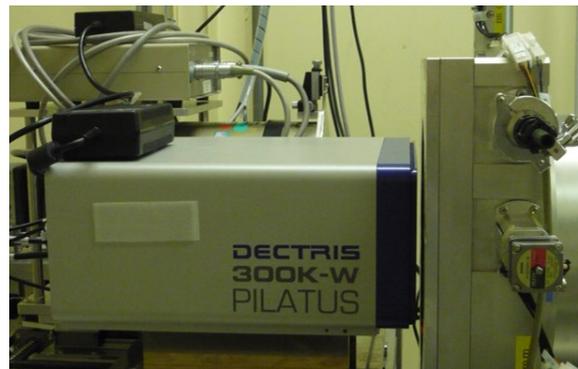


図 4.16 BL45XU 実験ハッチに設置された検出器 Pilatus300K-W(a)と標準試料であるベヘン酸銀粉末結晶からの回折パターン(b)。図 4.16a は検出器を横方向から撮影した写真であり、検出面は写真向かって右側である。図 4.16b の白い領域ほど光子が多く検出されており、中央部の黒丸はビームストッパーによる影である。また、矢印で示された黒帯は検出器の不感領域である。2 つの不感領域で散乱角が同一とならないようにビーム位置は検出器中心に対して数ピクセル横にシフトさせている。

つ数え上げるという特性から荷電結合素子型の検出器と異なり、ノイズの積算や検出飽和が起こらず、より質の高い実験データを取得できる。ただし一度に多数の光子が入射すると、光子の数え落としや最悪の場合過電流によりピクセルが破壊される恐れがあるため、入射 X 線のフラックスに注意する必要がある。今回の実験では、試料蛋白質溶液の散乱断面積が小さく、アテネータなしでも検出される光子数は毎秒 200 フォトンに達しないことから、アテネータは使用せずに測定を行った。なお、入射 X 線はそのほとんどが散乱されずに試料を透過する。この透過 X 線（ダイレクトビーム）は散乱 X 線と比較して桁違いに強度が高く、このダイレクトビームによる検出器の破壊を防ぐため、検出器直前にはダイレクトビームを吸収するビームストッパーが挿入されている。

実験では、試料溶液に青色光を照射して測定を行うため、試料セルに対して直接青色光を照射できるように青色光 LED(4515C, Nakagawa Chemicals, Japan)を試料セル直上に配置した(図 4.14b)。この青色光 LED は最大波長 480 nm、半値幅 40 nm である。4.1.3 節でも述べたが、LOV ドメインの S₃₉₀ は熱緩和によって D₄₅₀ へ戻るため、青色光を照射された phot 溶液は D₄₅₀ と S₃₉₀ の二成分系となる。SAXS では試料溶液に単分散性が求められるため、ほぼ全ての phot 分子が S₃₉₀ に励起されるよう照射量を検討しなければならない。SAXS 測定中の青色光照射条件を決定するため、phot1 LOV2-STK 断片および全長 phot2 双方について青色光照射条件を変えて複数の吸収スペクトルが測定された(図 4.17 上段)。赤色のスペクトル曲線は暗条件下で測定されたものであり、D₄₅₀ の吸光スペクトルに相当する。橙色、緑色、水色、紫色はそれぞれ 25-110 μmol/m²·s の粒子フルエンス率での青色光照射下での吸光スペクトルであり、一部の phot 分子の S₃₉₀ への移行によって吸光スペクトルが変化している。粒子フルエンス率は光の照射量の単位であり、単位面積・単位時間に何フォトン入射するか、という量である。エネルギーに換算すると、波長 λ、粒子フルエンス率 F の場合およそ

$$0.12 \times \frac{F}{\lambda} \text{ nJ}/\mu\text{m}^2 \cdot \text{s}$$

となる。ここで波長、粒子フルエンス率の単位はそれぞれ nm、μmol/m²·s である。前述の通り S₃₉₀ に相当する吸光スペクトルは熱緩和のために測定が困難であるが、phot1 LOV2 などの一部の緩和時間が長い LOV ドメインでは S₃₉₀ のスペクトルが測定されており、その結果から S₃₉₀ では波長 450 nm での吸収がほぼゼロとなるこ

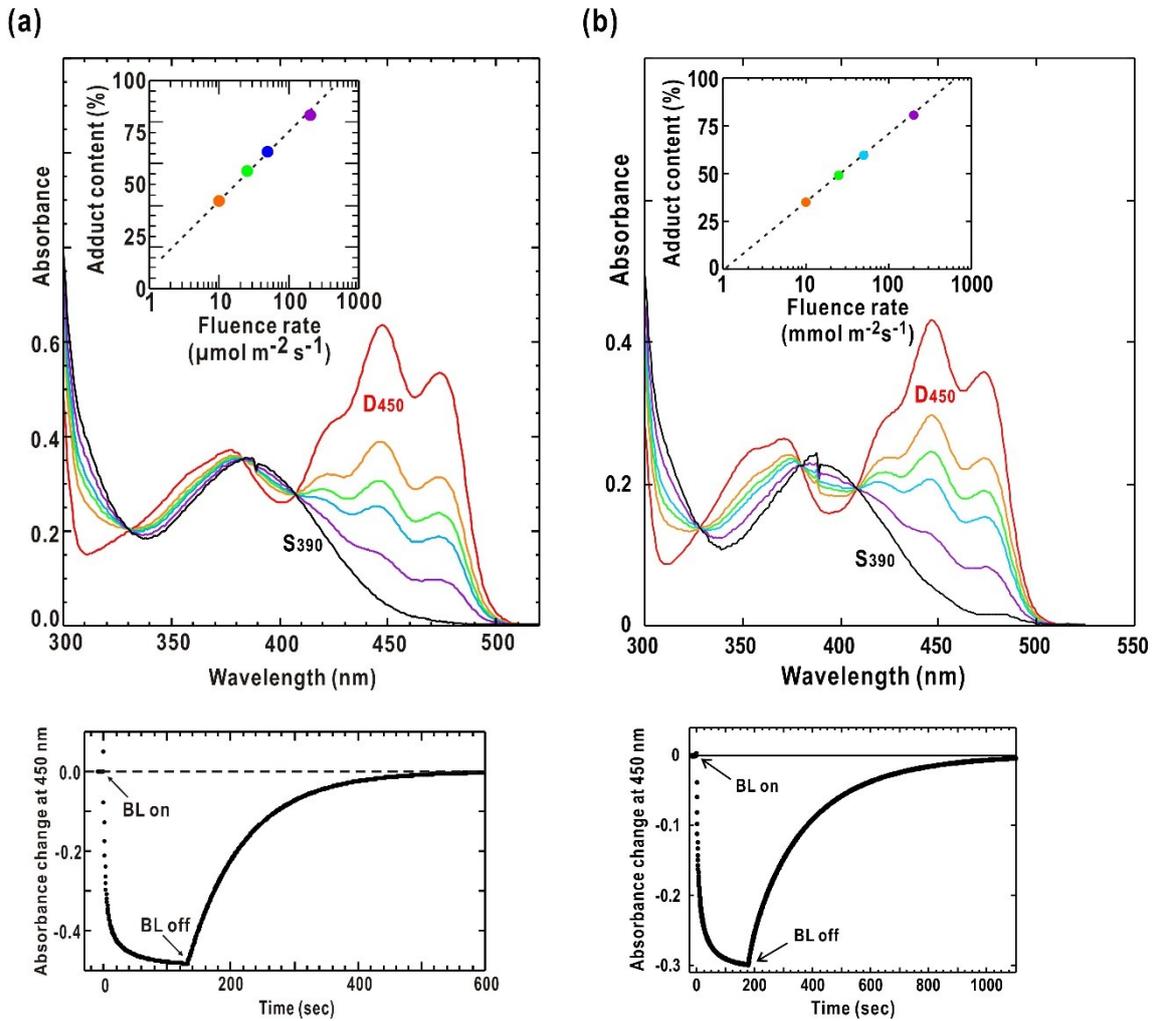


図 4.17 フォトリピンの青色光強度に依存した吸収スペクトル変化と青色光による波長 450 nm での吸光度の時間変化。図 4.17a が phot1 LOV2 - STK 断片、4.17b が全長 phot2 のデータである。吸収スペクトル曲線の色と粒子フルエンス率の色は対応している。なお、使用した分光計は波長 390 nm 付近で光源ランプが切り替わるため、これによる吸光度のずれが S₃₉₀ 吸収スペクトルで不連続点として表れている。

とが知られている [Okajima *et al.*, 2016]。この事実に基づき、図 4.17 における S₃₉₀ の吸収スペクトル(黒色)は青色光照射下でのスペクトル曲線から滑らかさを損なわない程度に D₄₅₀ のスペクトル曲線を差し引くことで得られている。また、450 nm での吸光度を S₃₉₀ での値と比較することで各照射量における S₃₉₀ へ遷移した phot

分子の割合(Adduct content)を見積もることが可能である(図 4.17 内挿)。この考えに基づいて Adduct content を粒子フルエンス率について外挿すると、粒子フルエンス率 $450 \mu\text{mol}/\text{m}^2\cdot\text{s}$ 以上の強度で青色光を照射すればほぼ全ての phot 分子が励起状態にあると見なせる。青色光 LED はこの閾値を超えた強度で照射できるように位置を調整した。併せて phot 全分子を励起させるために必要な照射時間の見積もりも行い(図 4.17 下段)、phot 分子を完全に S_{390} へ励起するためには 2 分程度の青色光照射が必要となることがわかった。さらに、 S_{390} へ励起された状態から全 phot 分子が熱緩和によって D_{450} へ戻るためには 15 分程度の時間を要することがわかった。本実験ではこの熱緩和後の phot に対しても暗条件下で測定を行い、構造変化の可逆性を評価した。

4.2.3 測定手順

露光時間は phot1 LOV2 - STK 断片が 60 秒、全長 phot2 が 20 秒であり、各試料ごとに以下の手順で測定した。また、各測定でのデータについて、光条件に応じて [] 内のように呼称する。

- ①緩衝液のみで測定する(バックグラウンド減算のための測定)。
 - ②緩衝液を試料セルから抽出後、セルを超純水で洗浄し、風乾する。
 - ③試料溶液を暗条件下で測定する。[Dark1]
 - ④試料溶液に対して青色光を 2 分間照射後、青色光照射条件下で測定する。[Light]
 - ⑤青色光照射を停止し、15 分後に試料溶液を再度暗条件下で測定する。[Dark2]
 - ⑥試料溶液をセルから抽出後、セルを緩衝液、超純水の順に洗浄し、風乾する。
- 測定を通じて試料セルは同一セルを使用し、入射 X 線に対して同じ向きで設置した。また、X 線照射によって石英窓、試料溶液共に放射線損傷が生じる。この影響を低減させるため、X 線照射中に試料セルを垂直方向に動作させ、照射位置を移動させながら測定を行った。ただし、石英窓は位置毎に吸収係数が微妙に異なるため、各測定で同一経路を移動させている。

4.2.4 データ処理

本実験に使用した検出器 Pilatus 300K-W は二次元検出器であるため、得られるデータは散乱パターンとなる。測定データの質を実験中直ちに確認するため、散乱プロファイルへと変換する処理を測定と並行して行った。SAXS プロファイルは

- ・ 散乱パターンの円環平均
- ・ 試料溶液散乱プロファイルからのバックグラウンド（溶媒散乱プロファイル）減算

によって得られる。散乱パターンの円環平均はソフトウェア FIT2D [Hammersley, *et al.*, 1996; Hammersley, 1997]を用いて行った。バックグラウンド減算は BL45XU のビームライン担当者である引間博士により開発されたソフトウェア DataView を用いて行った。DataView では FIT2D により得られた散乱プロファイルの確認、バックグラウンド減算、プロファイルの平均化処理などを行う。これら処理により得られた散乱プロファイルについて、試料蛋白質凝集の有無やバックグラウンド減算の妥当性などを評価しながら測定を行った。

4.3 phot1 LOV2 - STK 断片の測定及び解析結果

4.3.1 SAXS プロファイルの光依存的変化

phot1 LOV2 - STK 断片 0.8 mg/mL 溶液の SAXS プロファイルでは、 $s = 0.006 \text{ \AA}^{-1}$ 付近で青色光照射下の散乱プロファイルに変化が生じた (図 4.18)。変化が生じた散乱ベクトルの大きさはおよそ 150 \AA 分解能であるため、青色光受容によって分子全体のスケールで何らかの構造変化が生じていることが示唆された。また、差分プロファイルより、Dark1 から Light、Light から Dark2 にかけての散乱プロファイル変化がほぼ対称であった。そのため、phot1 LOV2 - STK 断片の青色光に依存した構造変化は可逆的な変化であったと言える。これは S_{390} から D_{450} への熱緩和と同様に LOV2 - STK 断片全体構造も暗条件下での構造に戻ったと解釈できる。

4.3.2 慣性半径と原点散乱強度変化

Guinier プロットを図 4.19a に示す。いずれのプロファイルも小角領域で直線近似でき、単分散溶液試料であった。慣性半径は Dark1 で $42.9 \pm 0.5 \text{ \AA}$ 、Light で $48.8 \pm 0.5 \text{ \AA}$ 、Dark2 で $43.5 \pm 0.5 \text{ \AA}$ であり、青色光照射に依存して $5\sim 6 \text{ \AA}$ 程度変化した。この結果より、phot1 LOV2 - STK 断片は青色光照射によって異方的な分子伸長あるいは膨張を起こし、熱緩和過程で可逆的な収縮を起こすと考えられる。また、図 4.19a に示すプロファイルの高角側では、Guinier 近似直線に対して上向きに反る振る舞いを見せており、比較的異方性の強い分子形状であることがわかる。一方、原点散乱強度は青色光受容でほとんど変化せず、分子量既知の卵白リゾチー

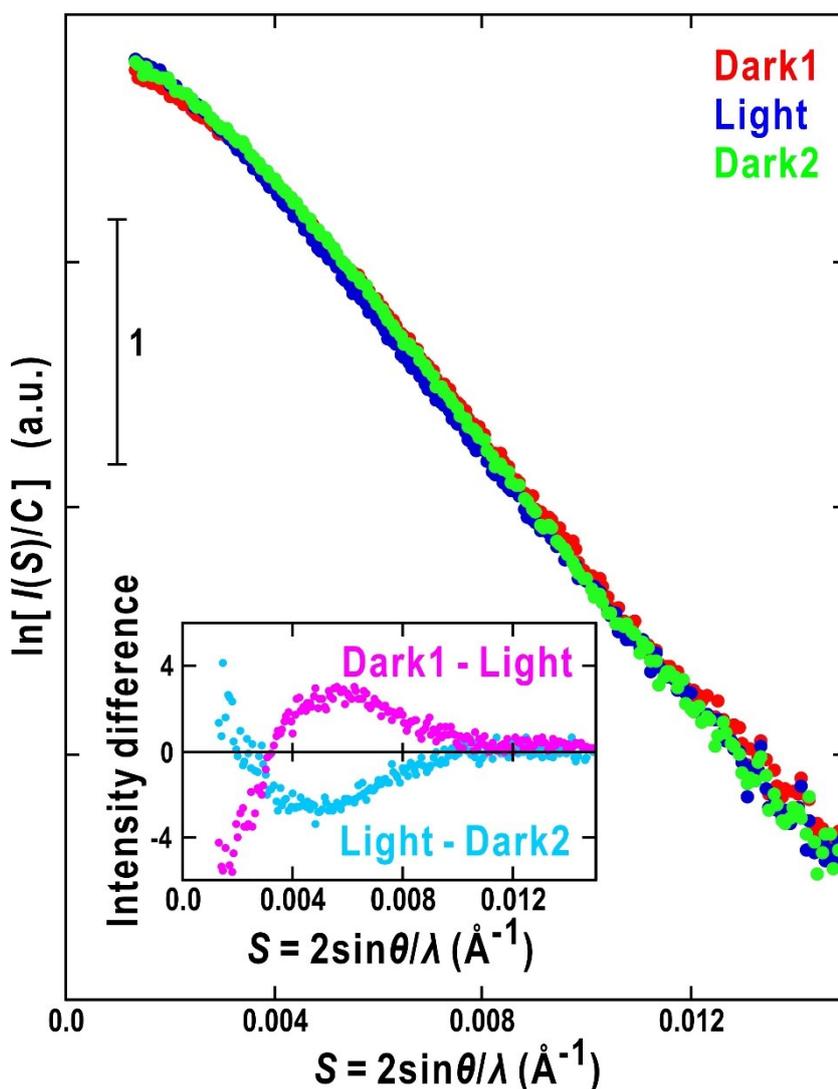


図 4.18 phot1 LOV2 - STK 断片の SAXS プロファイルと各プロファイルに対する差分プロファイル。赤色のプロファイルが Dark1、青色のプロファイルが Light、緑色のプロファイルが Dark2 である。差分プロファイルについてはマゼンタが Dark1 から Light を減算したプロファイル、シアンが Light から Dark2 を減算したプロファイルである。

ム原点散乱強度との比較によって二量体であることが示唆された。これは phot1 LOV2 ドメイン結晶構造が二量体であることと矛盾しない結果である。さらに複数の濃度での SAXS プロファイルから濃度依存性を解析したところ、慣性半径が濃

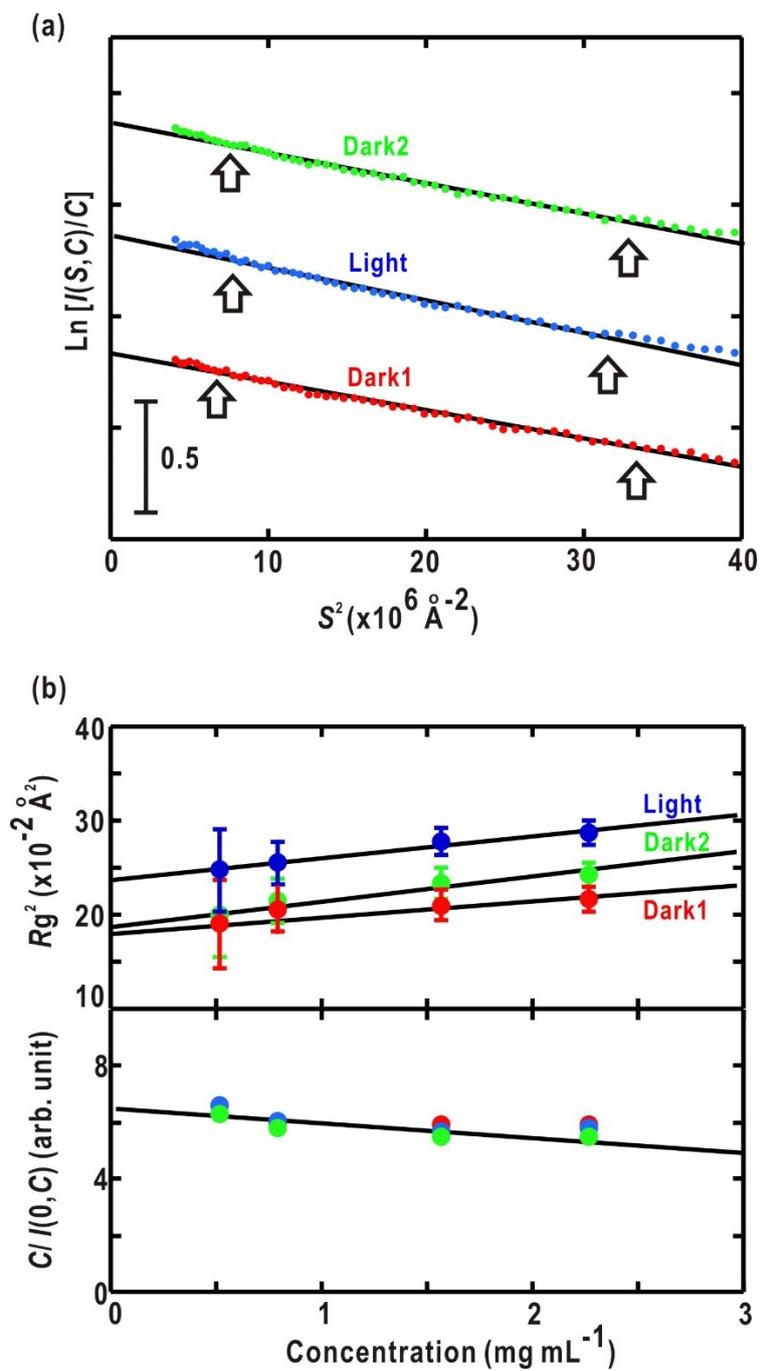


図 4.19 phot1 LOV2 - STK 断片 0.8 mg/mL 溶液 Guinier プロット(a) と各濃度での Guinier 解析結果から求められた原点散乱強度及び慣性半径の濃度依存性(b)。

度に依存して増加していた。式(2.26)を考慮すると、phot1 LOV2 - STK 分子は互いに引力相互作用を有していると考えられる。

4.3.3 距離分布関数の計算と最大分子長の比較

分子形状推定に先立ち、図 4.18 の SAXS プロファイルからソフトウェア GNOM[Svergun, 1992]によって距離分布関数・最大分子長を求めた(図 4.20)。各光条件における最大分子長は Dark1 で $146 \pm 2 \text{ \AA}$ 、Light で $162 \pm 2 \text{ \AA}$ 、Dark2 で $152 \pm 2 \text{ \AA}$ であった。青色光照射によって最大分子長は 15 \AA 程度増大し、その後暗条件下で熱緩和によって収縮するという振る舞いを見せており、Guinier 解析で求めた慣性半径の振る舞いと矛盾しない。本節で得られた最大分子長を用いて各光条件での散乱プロファイルごとに分子形状推定計算を行い、構造変化を検討する。

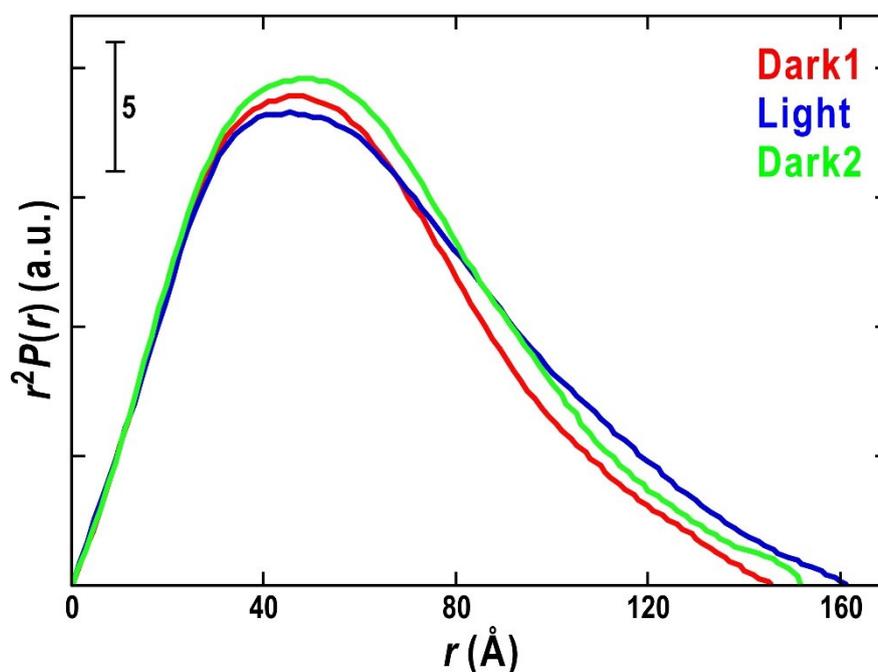


図 4.20 phot1 LOV2 - STK 断片 距離分布関数ソフトウェア GNOM により計算された各光条件における距離分布関数。グラフ中の矢印は各色の距離分布関数における収斂距離、即ち最大分子長を示している。

4.3.4 GASBOR と多変量解析による分子形状推定

まず GASBOR プロファイルが実験データに最もよく一致するダミー残基数を決定した。その結果、1000 残基(二量体であるため 1 サブユニットあたり 500 残基)を指定した場合に最もよく一致した(図 4.21)。この値は実際の残基数 547 の約 9 割の値である。SAXS での散乱強度は試料蛋白質の溶媒とのコントラストで決まるが、phot1 LOV2 - STK 断片は試料緩衝溶液に多量の NaCl が溶解しており、純水と比較して 1 割程度のコントラスト変動が生じている [Ibel & Stuhrman, 1975]。

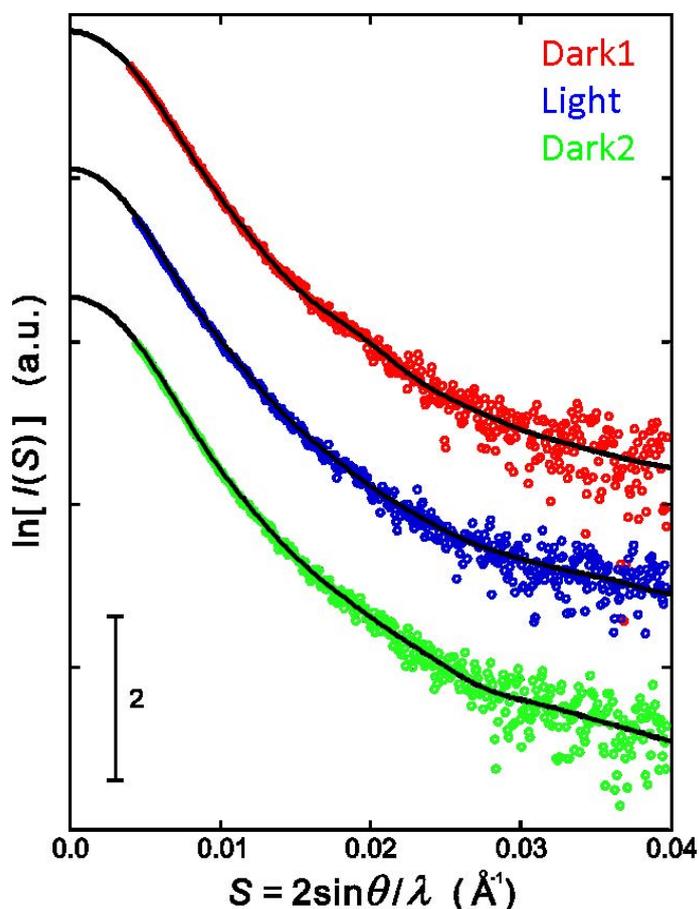


図 4.21 実験で取得された SAXS プロファイル(赤色、青色、緑色の点、図 4.18 と同一のプロファイル)及び実験データをもとに GASBOR から計算されたプロファイル(黒色の曲線)の比較。各散乱プロファイルは散乱強度についてオフセットした上でプロットしている。各光条件それぞれについて、GASBOR から計算されたプロファイルは実験データによく一致している。

この電子密度コントラストの減少や構造的に揺らぎの大きな領域が平均的に低コントラストになることを考慮すると実際の残基数の 9 割程度のダミー残基数は妥当であると考えられる。最大分子長とダミー残基数より Dark1、Light、Dark2 それぞれについて 576 回ずつ独立な GASBOR 計算を行った。なお、二量体と推定されたことからダミー残基の配置について C2 回転対称性の拘束条件を課した。

まず Dark1 推定形状モデルについて、多変量解析を用いてモデル形状分布の可視化を行った(図 4.22a)。phot1 LOV2 - STK 断片の推定形状は S 字型の形状と(クラス I-VIII)と V 字型(クラス XI、X)の 2 種類の形状に大別された。これら形状のどちらがより尤もらしいかを判断するため、phot1 LOV2 結晶構造[Halavaty & Moffat, 2013] (図 4.6c)との二回回転対象領域の一致度を指標とした。LOV2 は二量体形成サイトであるため、GASBOR モデル形状の回転対称軸原点付近に位置すると考えられ、配置の任意性が制限される。定量的に一致度を評価するため NSD に基づいた指標

$$\left\{ \frac{1}{N} \left[\sum_{i=1}^N h^2(m_i, M) \right] \right\}^{\frac{1}{2}}$$

を用いて各クラスの一貫性を検討した。ここで、 m_i は結晶構造中の i 番目の $C\alpha$ 原子、 M は GASBOR 平均形状モデルである。計算の結果、例えばクラス III では上の指標の値が 45.6 となった一方で、クラス IX では 52.3 となった。この結果より、phot1 LOV2 - STK 断片の二量体構造をクラス III のような S 字型であると推定した。Light、Dark2 GASBOR モデルに対して、Dark1 GASBOR モデルと同様の解析を行った結果を図 4.23、4.24 に示す。

Light で得られた各クラスを Dark1 推定形状と比較した結果、Dark1 と同様に S 字型(クラス I、II、V-VII)及び V 字型(クラス III、IV、VIII、IX)の形状が得られたものの、S 字型形状は Dark1 から伸長したような形状であった。4.3.2、4.3.3 節での SAXS プロファイル解析結果より、phot1 LOV2 - STK 断片は青色光照射によって分子伸長が生じることが明らかになっている。そのため、Light での形状は、クラス I のような伸長した S 字型であると推定した。Dark2 では S 字型(クラス I、VI)、V 字型(クラス VII-IX)と共に菱形の形状(クラス II-V)が表れた。SAXS プロファイルの可逆的変化を考慮し、Dark2 分子形状はクラス I に代表される Dark1 と同様の S 字型であると推定した。なお、Dark2 のモデル形状分布は菱形の形状が出現

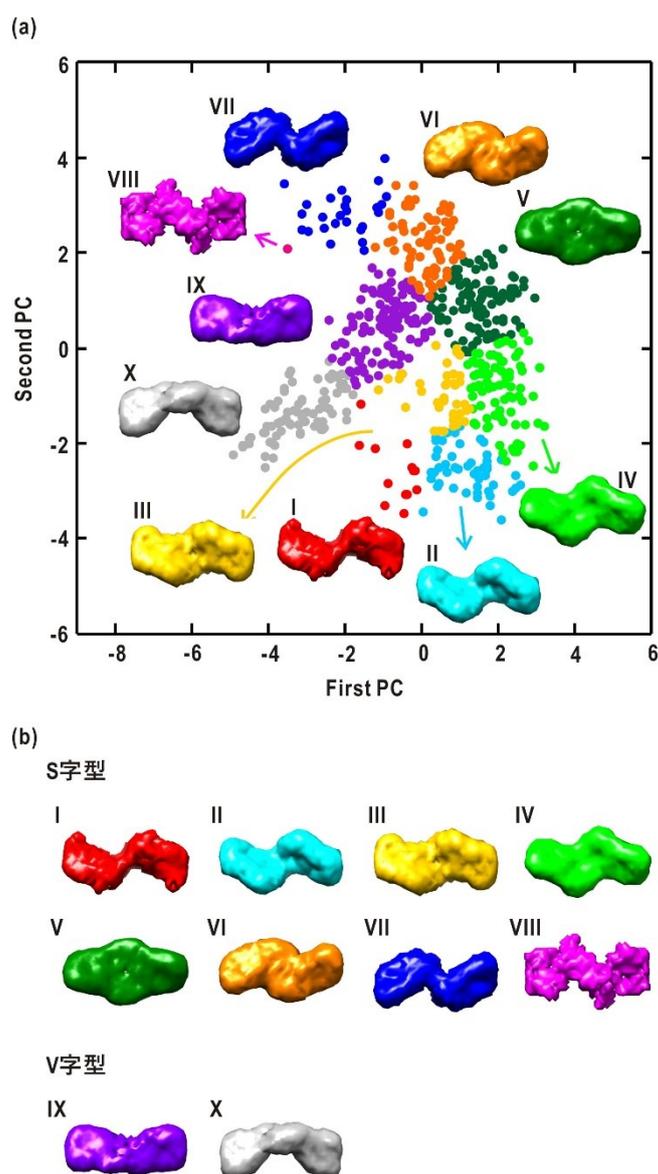


図 4.22 phot1 LOV2 - STK 断片の Dark1 状態における GASBOR モデル形状分布(a)と各クラスにおける平均形状の分類結果 (b)。モデル表示の閾値は各モデルで共通である。クラス 8 の形状が滑らかでないのは 1 個のモデルのみでクラスが形成されていることによる。

するなど Dark1 とはやや異なるものであった。図 4.18 の差分プロファイルが完全には対称でないことを考慮するとこれは Dark2 において熱緩和が不完全であったことに起因すると考えられる。Dark1、Light、Dark2 推定形状を比較すると、青色光に依存して伸長とわずかな屈曲を示すことが推測される(図 4.25)。

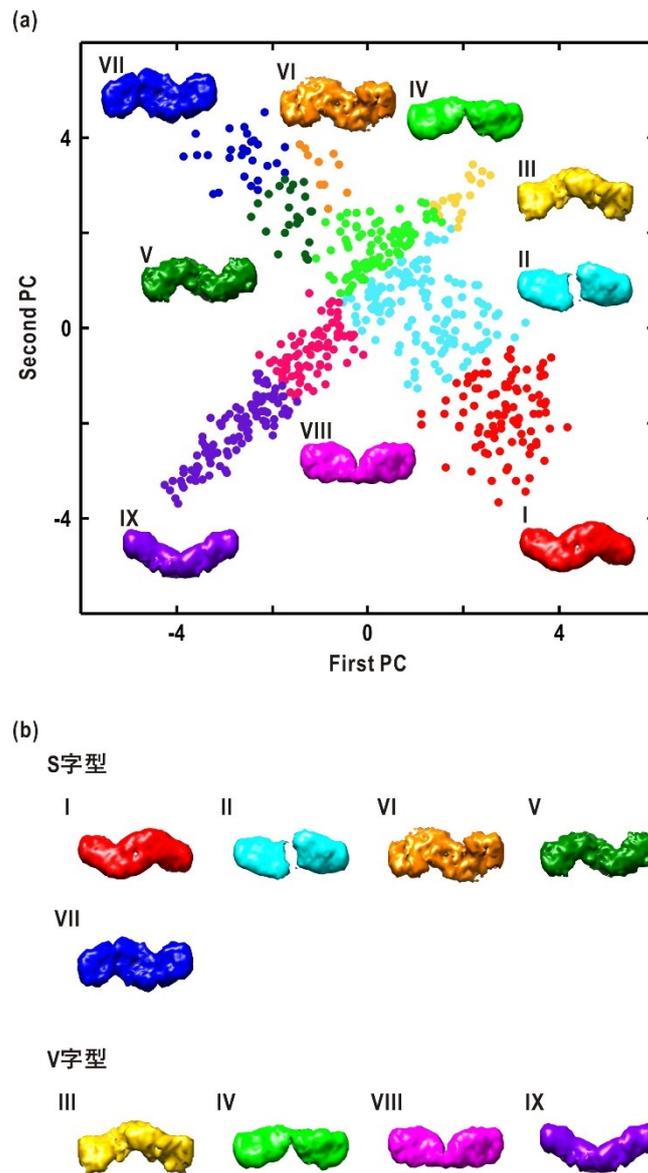


図 4.23 phot1 LOV2 - STK 断片の Light 状態における GASBOR モデル形状分布(a)と各クラスにおける平均形状の分類結果 (b)。クラスが9個になっているのはK平均法によるクラスタリングにおいて1つの重心がモデルの全く存在しない初期位置に置かれたことによる。

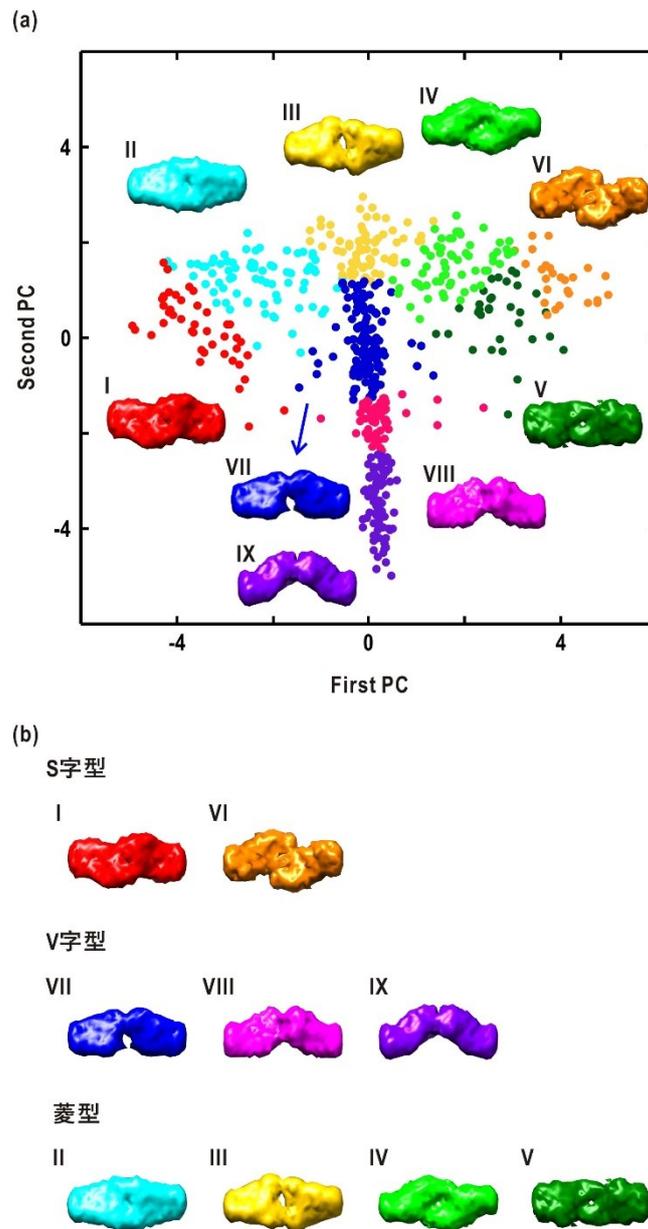


図 4.24 phot1 LOV2 - STK 断片の Dark2 状態における GASBOR モデル形状分布(a)と各クラスにおける平均形状の分類結果 (b)。クラスが 9 個になっているのは Light と同様の理由による。

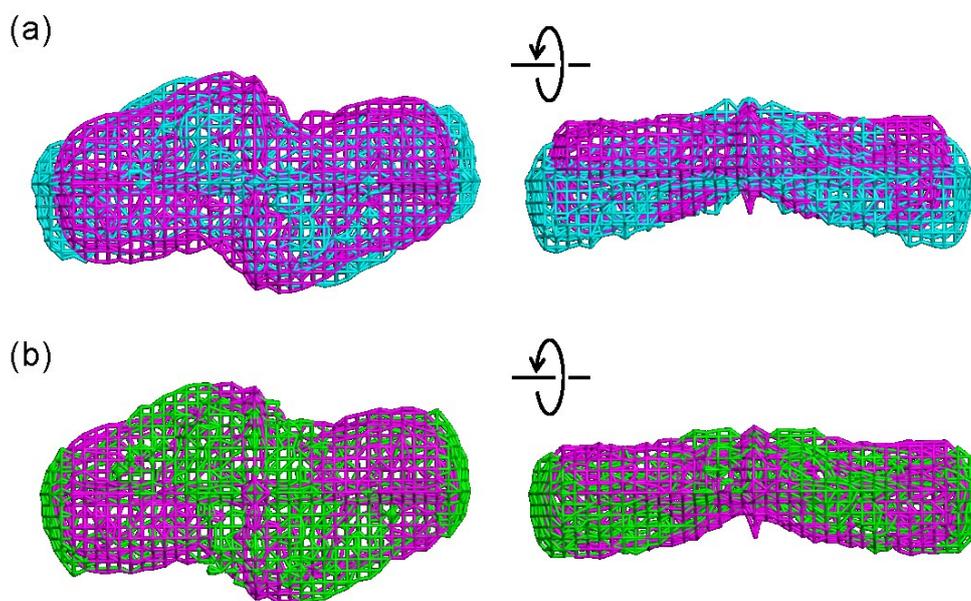


図 4.25 phot1 LOV2 - STK 断片の構造変化 phot1 LOV2 - STK 断片 Dark1、Light 形状比較図(a)と Dark1、Dark2 形状比較図(b)。マゼンタのモデルが Dark1、シアンのモデルが Light、黄緑色のモデルが Dark2 である。

4.3.5 変異体の測定による信号伝達機構の解析

分子内信号伝達機構についてさらに詳細な知見を得るため、変異体についても SAXS 測定を行った。phot の機能発現においては青色光受容に伴う LOV2 - STK ドメインの相対位置変化が重要であると推測されたが、phot1 では機能発現において重要な役割を担う残基が先行研究にて複数同定されている [Kashojiya *et al.*, 2015, Kashojiya *et al.*, 2016] (表 4.2; 図 4.26)。

表 4.2 phot1 機能阻害変異体

変異体名	位置	備考
D806N	STK	Mg-ATP 結合が阻害されリン酸化機能を喪失
C512A	LOV2	青色光受容による FMN との結合を阻害
K475A	LOV2(A' α /A β)	青色光に依存した STK の活性化が失われる
K636A	Linker region	青色光に依存した STK の活性化が失われる

D806N 変異体は STK ドメインの Mg-ATP 結合サイトである 806 番アスパラギン酸 (D806、付録 A) をアスパラギン(N)に置換した変異体である。全長 phot2 に対して導入した変異と同様に、構造変化を生じることなくリン酸化機能を欠損していると考えられる。C512 変異体は FMN と結合する LOV2 の 512 番システイン(C512) をアラニン(A)に置換した変異体であり、LOV2 の青色光受容機能を欠損している。K475A、K636A は上記 2 種類の変異体とは異なり、導入によって青色光受容やリン酸化機能とは独立して phot1 がリン酸化活性を失うことが明らかにされている。475 番リジン(K475)は LOV2 N 端側の A'α-Aβ 間ループ(A'α-Aβ ギャップ)に位置し、隣接する Jα と相互作用をしていると考えられている。また、636 番リジン(K636)は LOV2 - STK 間のリンカー領域に存在する。4.5.1 節で推定した四次構造によると K475、K636 は LOV2 - STK の界面に位置しており、従ってこれらの残基が LOV2 - STK 間の相対位置の制御に重要な役割を果たしていると予測される。本研究ではこれらの残基について点変異を導入した 4 つの変異体(D806N、C512A/D806N、K475A/D806N、K636A/D806N)について SAXS による構造変化解析を行った(図 4.27)。

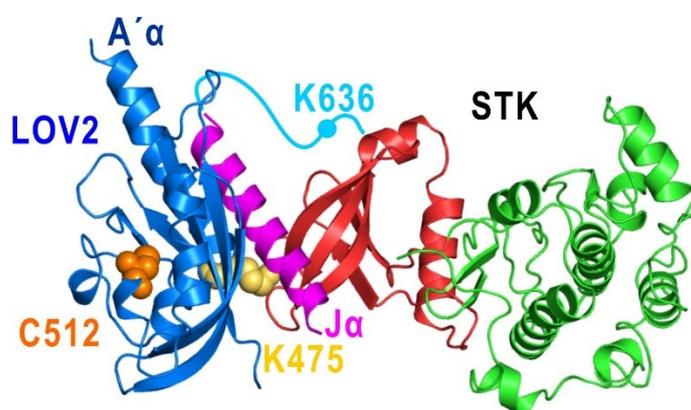


図 4.26 表 4.2 の変異導入位置。オレンジの残基が 512 番システイン残基、黄色の残基が 475 番リジン残基である。シアンで表示された 636 番リジン残基については構造の不明なリンカー領域に位置するためイラストで表示している。806 番アスパラギン酸残基については、STK ドメインがホモロジーモデルであるため省略しているが、PKA 結晶構造と類似した位置にあると予測される。

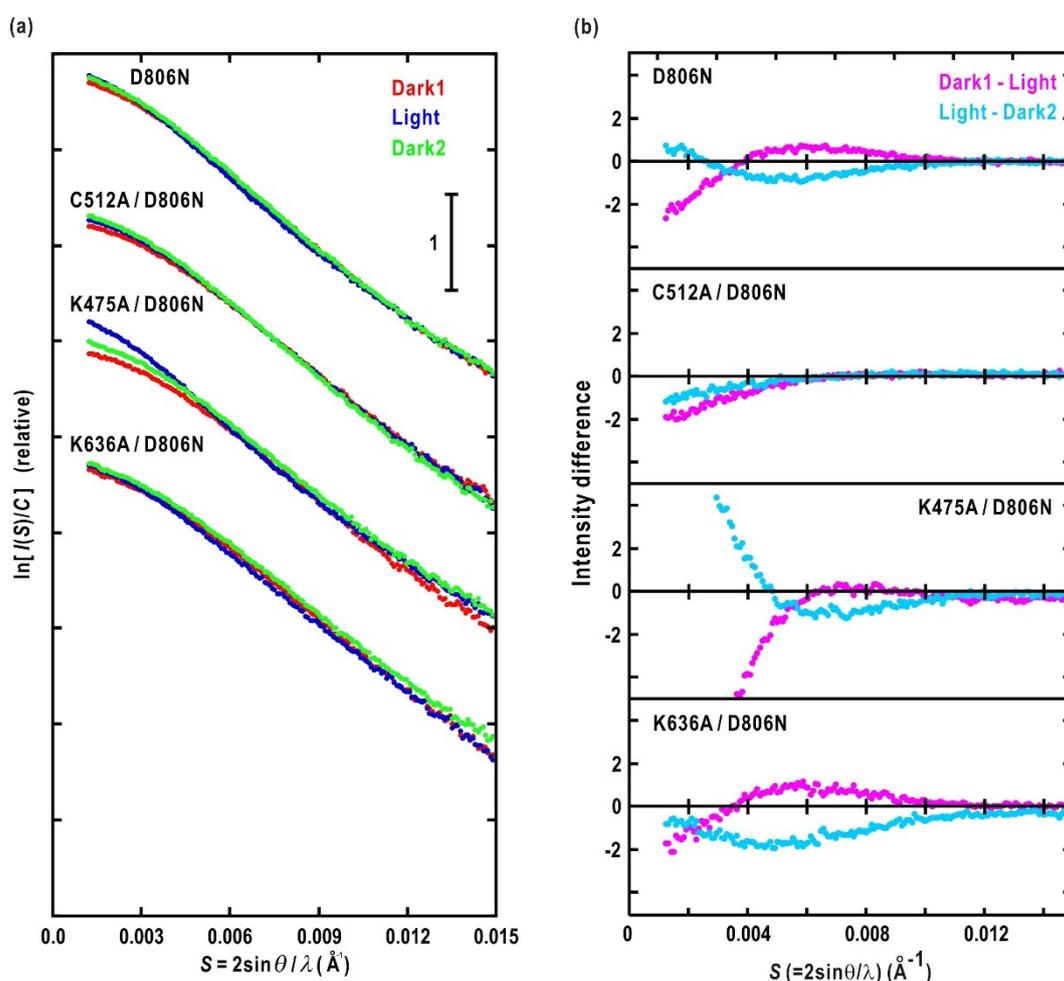


図 4.27 phot1 LOV2 - STK 断片変異体 SAXS プロファイル(a)とその差分プロファイル(b)。

D806N 変異体では、SAXS プロファイルは青色光照射による構造変化も含めて野生型の場合とよく一致する SAXS プロファイルが得られた。一方、C512A/D806N 変異体では青色光照射によって構造はほとんど変化しなかった。従って、野生型や D806N 変異体での青色光依存的な構造変化は S_{390} によるものであると確認できた。K475A/D806N 変異体では青色光照射によって $s < 0.004 \text{ \AA}^{-1}$ での散乱強度が大幅に上昇し、暗条件下で減少するという特異な挙動を見せた。式(2.18)、(2.22)を考慮すると原点近傍での散乱強度は総電子数(\cong 分子量)の二乗に比例する一方で散乱強度は濃度に線形に比例する。即ち、測定試料について会合が生じた場合小角の散乱強度は会合度に応じて増加する。従って、K475A/D806N 変異体は青色光

によって可逆的な凝集を生じ、熱緩和によって再び解離したと解釈される。凝集は蛋白質間相互作用や排除体積効果などによって引き起こされるものであり、いずれにしても青色光受容によって phot1 LOV2-STK 断片の分子表面性状の変化に起因すると考えるのが妥当であろう。K636 変異体については野生型と比較して特筆すべき散乱プロファイルの変化は見られなかった。

4.4 全長 phot2 の測定及び解析結果

4.4.1 SAXS プロファイルの光依存的変化

全長 phot2 1.0 mg/mL 溶液について、図 4.28 に示す SAXS プロファイルが得られた。ここで、各プロファイルの色と光照射条件の対応は phot1 LOV2 - STK 断片の SAXS プロファイル(図 4.18)と同様である。図 4.28 より、 $s = 0.006 \text{ \AA}^{-1}$ 付近で青色光照射と連動して散乱プロファイルの変化が生じている。変化が生じた散乱ベクトルの大きさはおよそ 150 \AA 分解能にあたるため、青色光依存的に分子全体に及ぶ何らかの構造変化が生じたものと考えられる。各光条件間での差分プロファイルを見ると、Dark1 から Light、Light から Dark2 への光条件変化に伴う散乱プロファイルの変化はほぼ対称的であることも確認でき、phot1 LOV2 - STK 断片と同様に青色光に依存した可逆的な構造変化が生じたと推測される。

4.4.2 Guinier プロットによる慣性半径、原点散乱強度変化

図 4.28 のプロファイルについて Guinier プロットを適用した結果が図 4.29a である。いずれのプロファイルも小角領域で直線によく一致する振る舞いを見せており、単分散性の良好な試料であったと言える。慣性半径の値は Dark1 で $44.3 \pm 1.4 \text{ \AA}$ 、Light で $43.7 \pm 1.4 \text{ \AA}$ 、Dark2 で $44.9 \pm 1.4 \text{ \AA}$ であり、phot1 LOV2 - STK 断片とは対照的に、青色光照射によってわずかに減少する挙動を見せた。原点散乱強度については、卵白リゾチームの原点散乱強度との比較から二量体を形成していることが示された。さらに、青色光照射によって若干の増加が認められたものの、増加量は会合状態の変化などの大規模な分子量変化によるものではない。従ってわずかな構造変化によって水和構造が変化し、水和水分子量が増加したことによると考えられる。また、図 4.29a で、高角側では Guinier 近似直線に対して上向きに反る振る舞いをすることから、phot1 LOV2-STK 断片同様に異方性の強い分子形状で

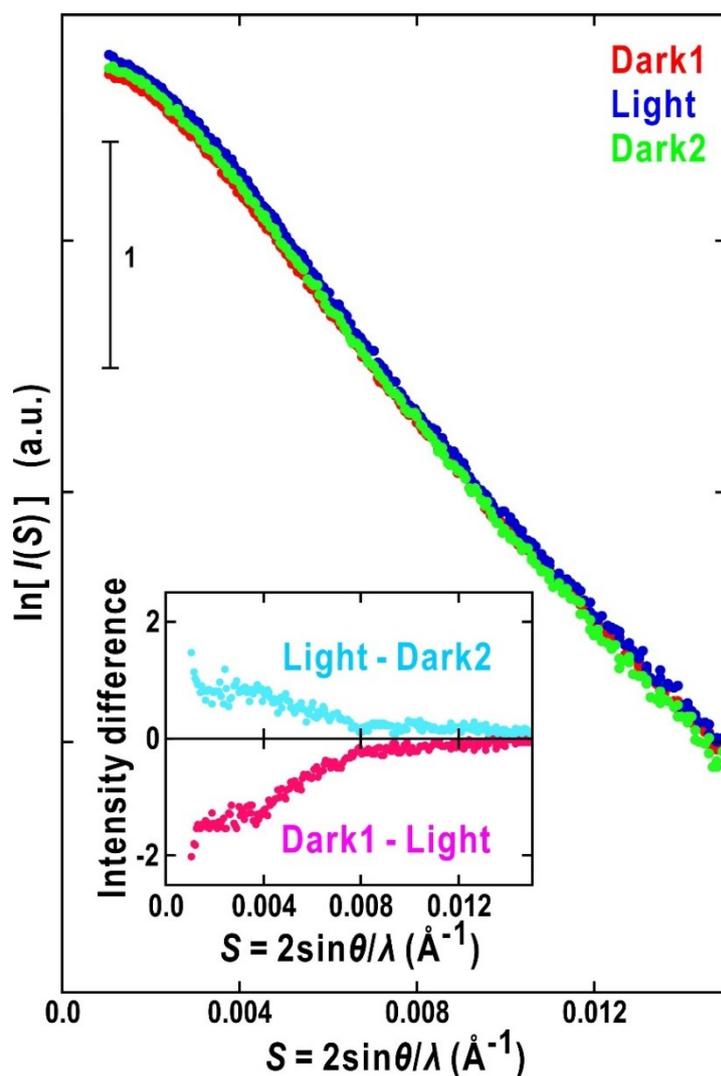


図 4.28 全長 phot2 の SAXS プロファイルと各プロファイルに対する差分プロファイル。赤色のプロファイルが Dark1、青色のプロファイルが Light、緑色のプロファイルが Dark2 である。差分プロファイルについてはマゼンタが Dark1 から Light を減算したプロファイル、シアンが Light から Dark2 を減算したプロファイルである。

あると予想された。他の濃度についても Guinier プロットを行い、式(2.25)及び式(2.26)を用いて慣性半径、原点散乱強度の濃度依存性を調べたところ、全長 phot2 の分子間相互作用は斥力相互作用であり、phot1 LOV2 - STK 断片とは異なる分子間相互作用を有していることが示された。

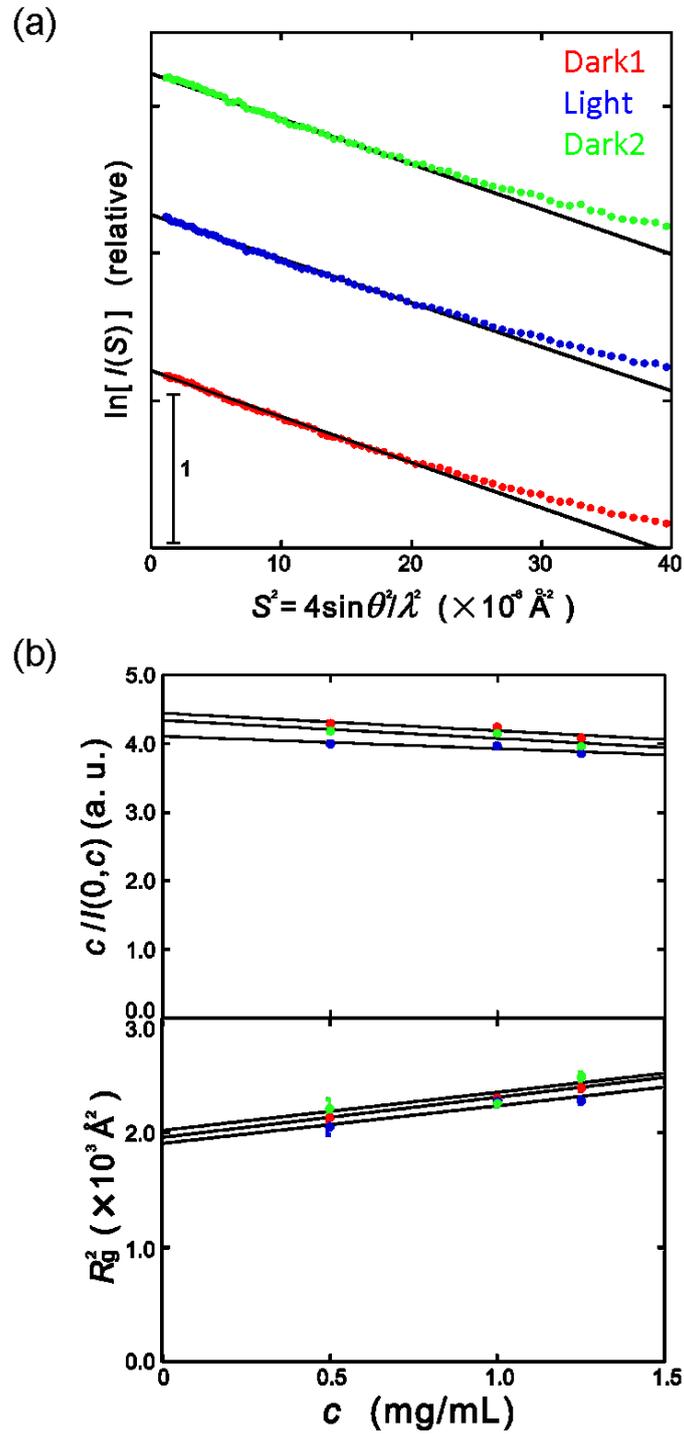


図 4.29 全長 phot2 慣性半径・原点散乱強度の濃度依存性全長 phot2 1.0 mg/mL 溶液 Guinier プロット(a)と各濃度での Guinier 解析結果から求められた原点散乱強度及び慣性半径の濃度依存性(b)。

4.4.3 距離分布関数の計算と最大分子長の比較

GASBOR による形状推定に先立ち、GNOM による距離分布関数計算を行い、最大分子長を推定した(図 4.30)。計算された距離分布関数より得られた各光条件における最大分子長は Dark1 で $188 \pm 2 \text{ \AA}$ 、Light で $186 \pm 2 \text{ \AA}$ 、 $187 \pm 2 \text{ \AA}$ であった。溶液中で単量体として存在する phot2 LOV2 - STK 断片の最大分子長 130 \AA [Takayama *et al.*, 2011]と比較すると全長 phot2 が溶液中で二量体を形成することが明らかである。青色光照射による最大分子長の変化は $1\text{-}2 \text{ \AA}$ 程度と分子サイズと比較して極めて小さな変化であったものの、その振る舞いは Guinier 解析で求めた慣性半径の変化と矛盾しないものであった。本節で得られた最大分子長を用いて各光条件での散乱プロファイルごとに GASBOR 計算を行い、青色光照射で生じた分子内構造変化を具体的に解析した。

4.4.4 GASBOR と多変量解析による分子形状推定

まず 3.8.2 節と同様にして GASBOR プロファイルが実験データに最もよく一致するダミー残基数を決定した。その結果 1200 残基(二量体であるため 1 サブユニットあたり 600 残基)を指定した場合に最もよく一致した(図 4.31)。これは実際の残基数である 915 残基のおよそ 7 割程度しかない値であった。これについて、以下

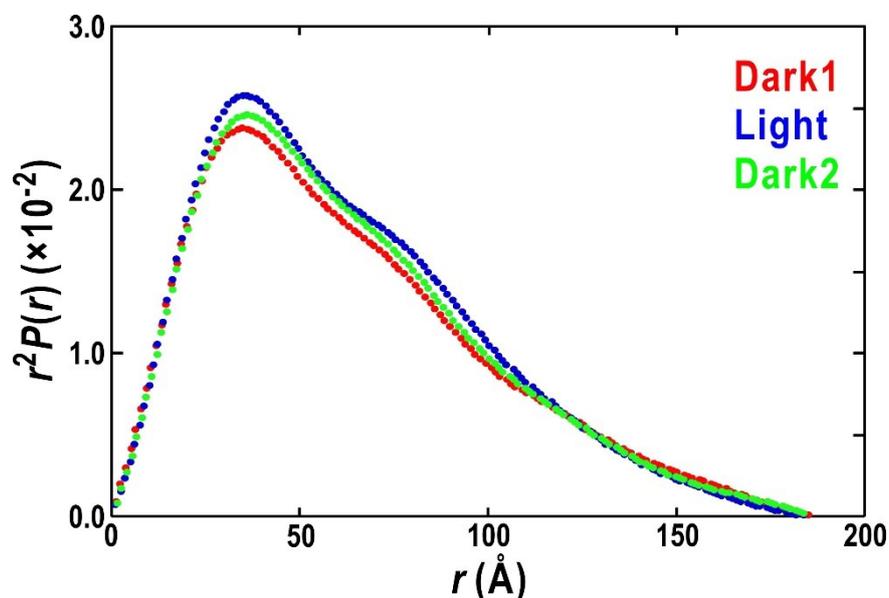


図 4.30 全長 phot2 距離分布関数ソフトウェア GNOM により計算された各光条件における距離分布関数。

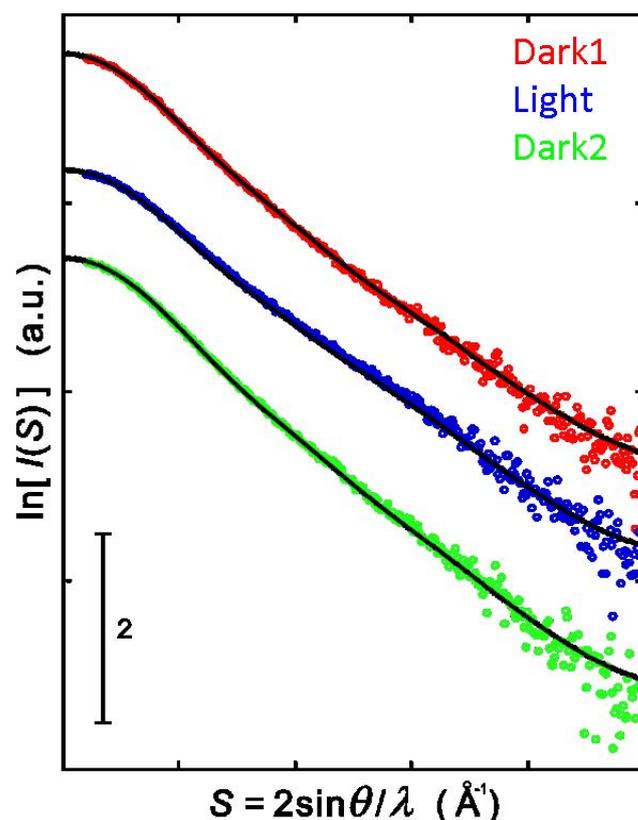


図 4.31 全長 phot2 SAXS プロファイルと GASBOR から計算されたプロファイルの重ね合わせ。各散乱プロファイルは散乱強度についてオフセットした上でプロットしている。

の理由が考えられる。

(i) ループ構造による散乱強度の低下

全長 phot2 では LOV1 N 端(116 残基)、LOV1 - LOV2 リンカー領域(98 残基)がループ構造であり、これらは構造の揺らぎにより SAXS での散乱強度に寄与しない可能性が高い。但し、LOV2 - STK のリンカー領域は散乱強度に寄与している可能性が高いと報告されている [Takayama et al., 2011]。

(ii) 溶媒条件によるコントラストの低下

全長 phot2 緩衝溶液には NaCl が多量に含まれているため、純水と比較して電子密度が高い。従って試料電子密度のコントラストは純水と比べて 88%程度まで低下する。GASBOR では溶媒として純水を仮定して散乱強度を計算するため、溶媒電子密度が純水よりも高い試料では見積もられる残基数が減少する。

以上二項目を勘案すると、想定される残基数はおよそ 600~650 程度となることから、今回予備的な計算から見積もられたダミー残基数は妥当であると言える。決定された最大分子長とダミー残基数より Dark1、Light、Dark2 それぞれについて 576 回ずつ独立な GASBOR 計算を行った。なお、二量体であることからダミー残基に対して C2 回転対称性の拘束条件を課すことでモデル形状の収束性の向上を図った。

Dark1 GASBOR モデルについて多変量解析を用いてモデル形状分布の可視化と分類を行った結果を図 4.32 に示す。全長 phot2 の推定形状は S 字型(クラス I-V)と V 字型(クラス VI-X)の 2 種類の形状に大別されたが、Dark1 の結果のみから 2 種類の形状の内どちらがより尤もらしいか判断することはできなかった。Light についても推定を行ったところ、Dark1 と同様に S 字型(クラス I-VII)および V 字型(クラス VIII-X)の形状が得られた(図 4.33)。前節までの SAXS プロファイル解析結果より、全長 phot2 は青色光照射によって何らかの構造変化を生じるものの、慣性半径及び最大分子長の変化は不確かさの域を超えない程度でしかなく、これらのパラメータから構造変化を推定することはできない。そこで、S 字型、V 字型それぞれの推定形状に対して LOV1、LOV2 結晶構造及び STK 予測構造を重ね合わせ、CRY SOL によって計算された SAXS プロファイルを実験データと比較することで推定を試みた(結晶構造の重ね合わせについては 4.5.1 節にて詳述)。Dark1 では S 字型構造との χ^2 が 1.2、V 字型構造との χ^2 が 17.7 であった。一方 Light では従って、S 字型構造との χ^2 が 18.7、V 字型構造との χ^2 が 2.4 であり Dark1 と対称な結果となった。以上より、Dark1 ではクラス III のような S 字型、Light ではクラス V のような V 字型であると推定された。Dark2 GASBOR モデル形状分布は概ね Dark1 と類似した形状であり、SAXS プロファイルもよく一致することから Dark1 と同様にクラス IV のような S 字型と推定された(図 4.34)。従って、青色光に依存した構造変化は可逆的な変化であったと考えられる。

Dark1、Light、Dark2 推定形状を比較すると(図 4.35)、全長 phot2 は青色光に依存してわずかな屈曲を示す。phot2 LOV2 - STK 断片に対する SAXS 測定では青色光に依存した分子伸長及び慣性半径の増加が観測されていたが[Takayama *et al.*, 2011]、今回の全長 phot2 において観測された慣性半径及び分子最大長の変化量は phot2 LOV2-STK 断片のものと比較して小さい。これは図 4.35a のように、伸長と屈曲が同時に生じるような構造変化であるためと考えられる。

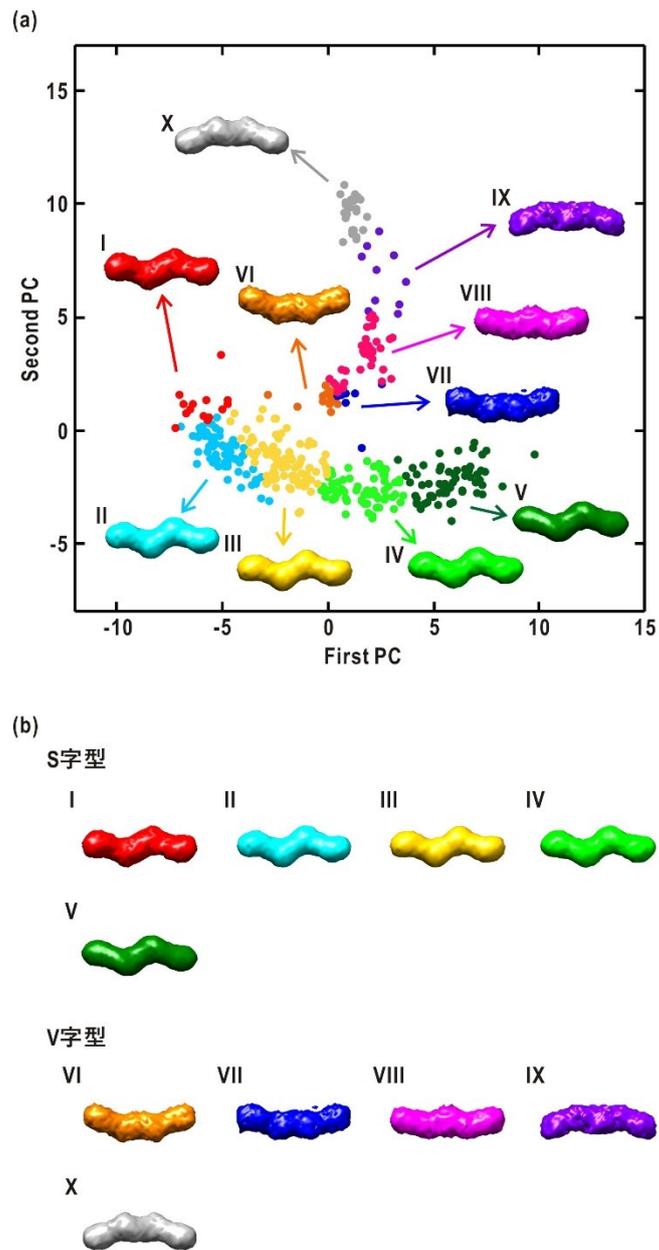


図 4.32 全長 phot2 Dark1 GASBOR モデル形状分布(a)と各クラスにおける平均形状の分類結果(b)。

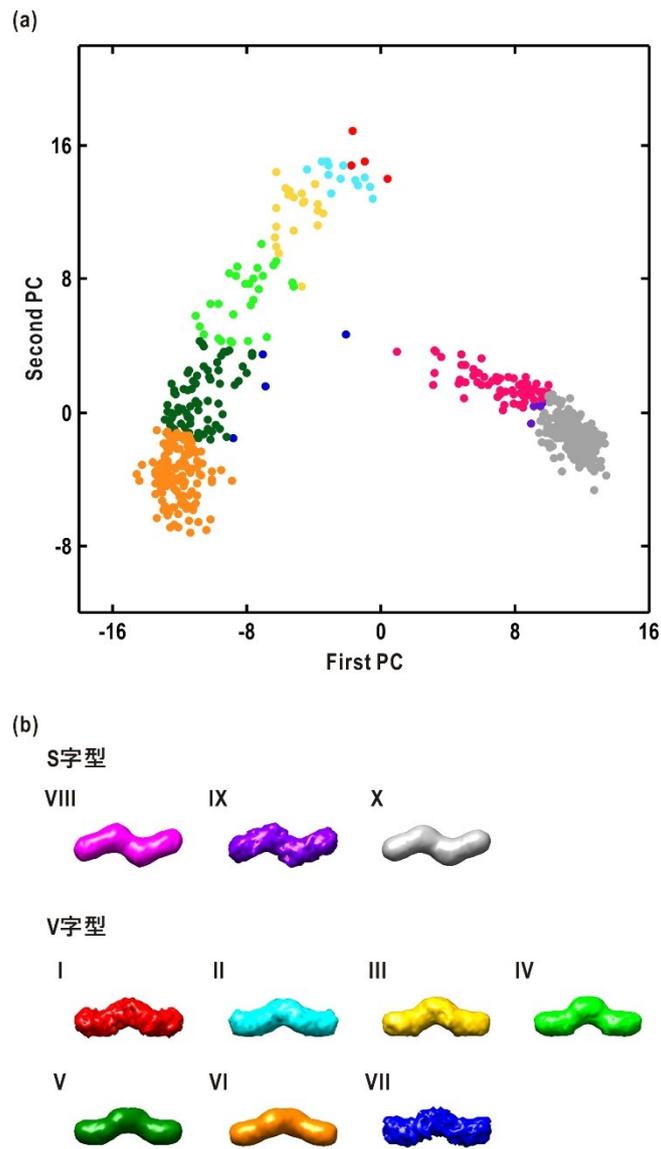


図 4.33 全長 phot2 Light GASBOR モデル形状分布(a)と各クラスにおける平均形状の分類結果(b)。

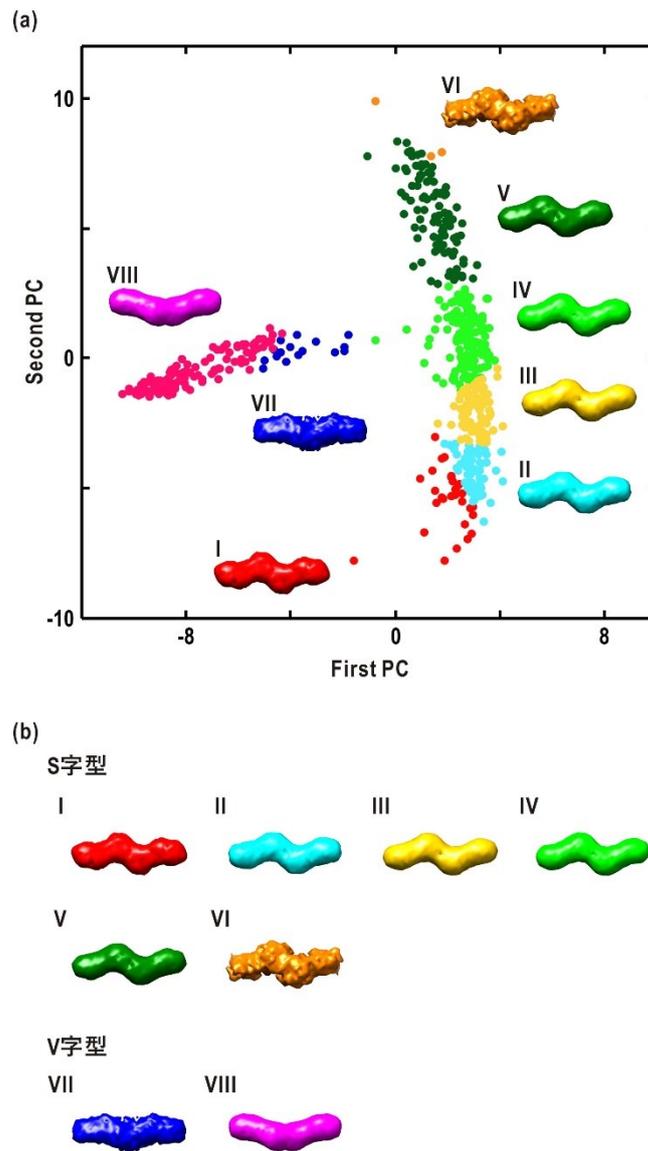


図 4.34 全長 phot2 Dark2 GASBOR モデル形状分布(a)と各クラスにおける平均形状の分類結果(b)。

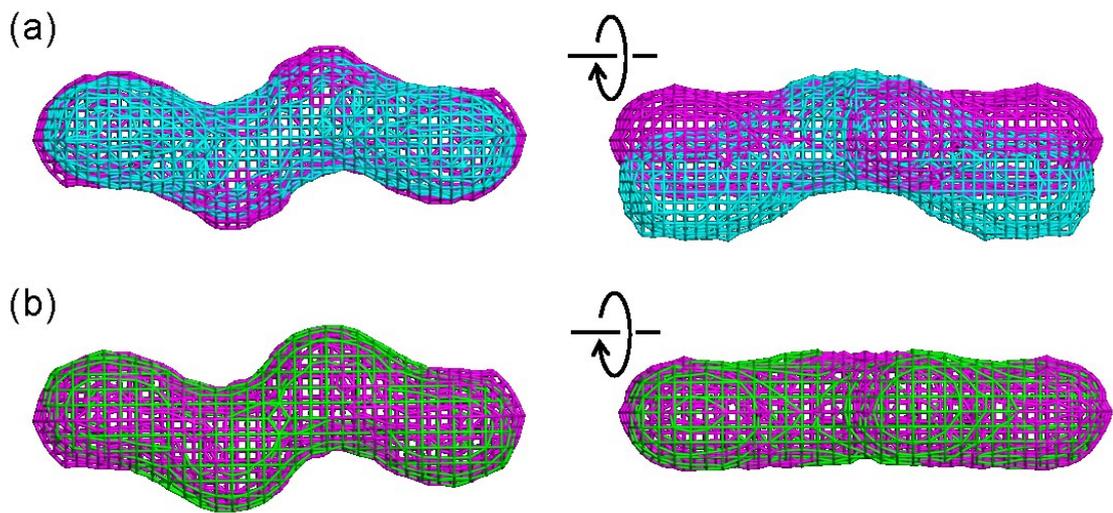


図 4.35 全長 phot2 Dark1、Light 形状比較図(a)と Dark1、Dark2 形状比較図(b)。マゼンタのモデルが Dark1、シアンのモデルが Light、黄緑色のモデルが Dark2 である。

4.4.5 変異体の測定による信号伝達機構の解析

全長 phot2 では LOV1、LOV2 両方を有しているが、先行研究によって機能の制御には LOV2 がより重要であると示唆されている [Okajima, 2016]。そこで青色光照射に伴う構造変化に対する LOV1、LOV2 それぞれの寄与を調べるため、変異体の測定を行った。4.1.2.2 節で述べた通り、 S_{390} に遷移した LOV ドメインでは FMN がポケット内部のシステイン基と共有結合を形成するため、このシステイン基に変異を導入することで LOV ドメインは失活する。本研究では LOV1 を失活させた C170A/D720N 変異体、LOV2 を失活させた C426A/D720N 変異体、双方を失活させた C170A/C426A/D720N 変異体の 3 種の変異体を用意し、D720N 変異体と同様の SAXS 測定を行った。

SAXS 測定の結果、C170A/D720N 変異体および C426A/D720N 変異体では、D720N 変異体と同様に青色光に依存した構造変化が観測されたものの、その度合いは D720N 変異体よりも小さかった(図 4.36)。一方で C170A/C426A/D720N 変異体は青色光照射でほとんど構造変化を示さなかった。従って、C170A/D720N 変異体では LOV2、C426A/D720N 変異体では LOV1 がそれぞれ青色光依存的な構造変化に寄

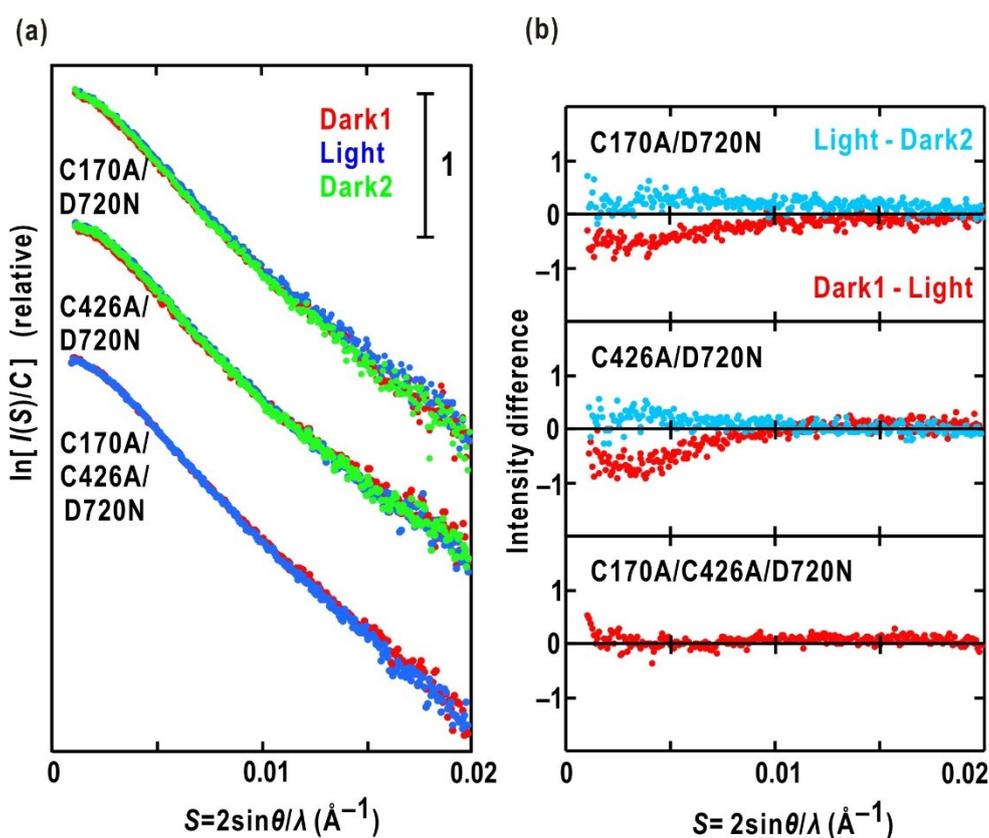


図 4.36 全長 phot2 変異体 SAXS プロファイル(a)とその差分プロファイル(b)。

与しており、野生型における構造変化は双方の寄与によるものと示唆された。

4.5 議論

4.5.1 phot1、phot2 機能ドメイン配置

4.3.4 節、4.4.4 節の結果より、暗条件下、青色光照射条件化それぞれの場合の phot1 LOV2 - STK 断片、全長 phot2 の形状が得られた。これら形状に対して LOV ドメイン結晶構造、STK ドメイン予測構造を重ね合わせることで 4 次構造、並びに青色光に伴う構造変化の実態を予想した。

phot1 LOV2 - STK 断片 Dark1 推定形状については、LOV2-J α 結晶構造(図 4.6c)が推定形状の中心とよく一致した。残りの領域は 2 つの STK 予測構造がよく当てはまり、STK の N 末端ドメインが J α ヘリックスと接するように配置された(図 4.37a)。一方、Light 推定形状においては LOV2-J α 二量体の相対配置を調整するこ

とで推定形状中心部によく一致した。この事実は先行研究で報告された青色光照射による LOV2-linker 二量体相対配置の変化と対応していると考えられる [Nakasako *et al.*, 2004]。また、STK ドメイン予測構造は Dark1 の場合と比較してより LOV2 ドメインから離れた配置になり、J α ヘリックスとのコンタクトも弱まった(図 4.37b)。リジン/アルギニンを特異的に加水分解する酵素を用いて各光状態の

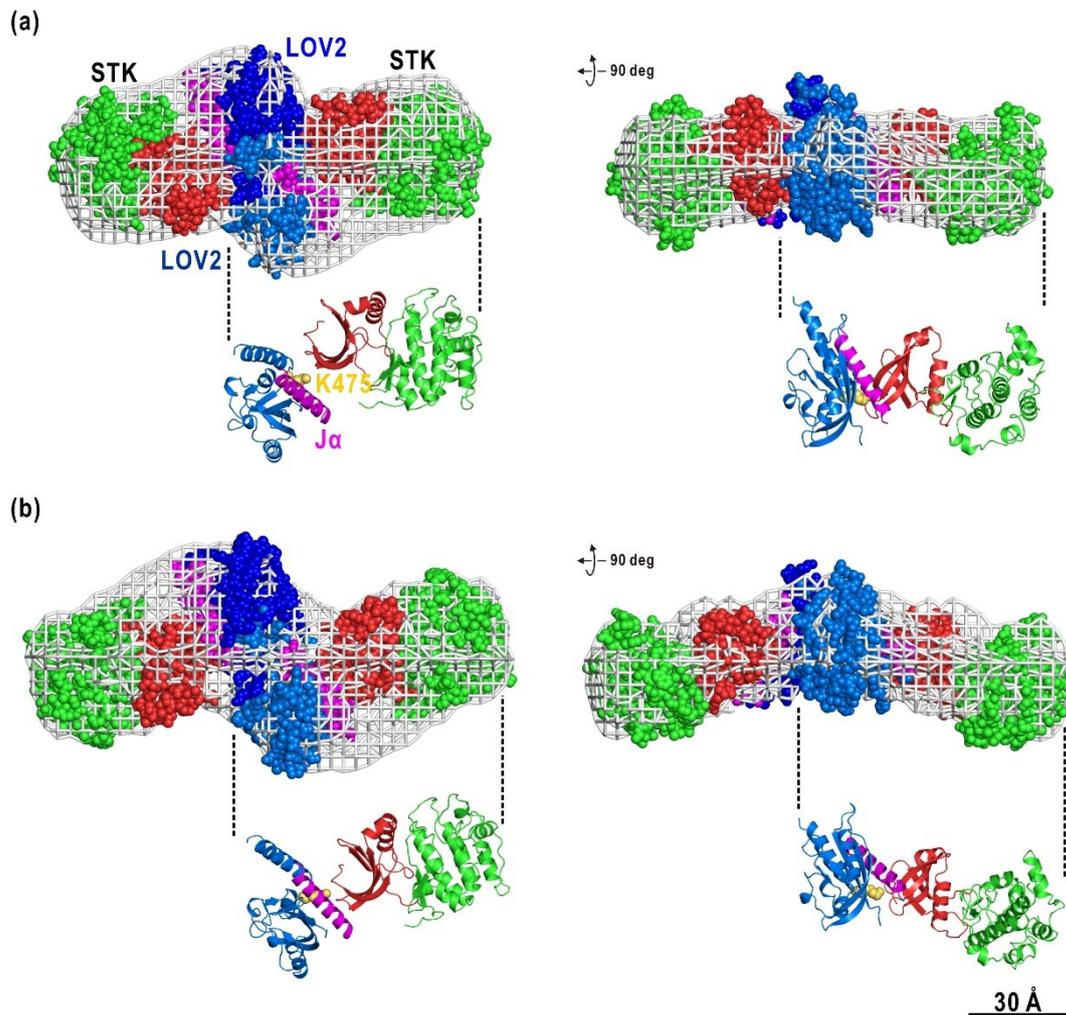


図 4.37 phot1 LOV2 - STK 断片 Dark1 推定四次構造(a)と Light 推定四次構造(b)。右図は左図の表示から 90 度回転させて表示している。また、475 番リジン残基周辺の構造変化を明らかに示すために、各推定四次構造下部にリボン表示で単一の LOV2-STK 断片を載せている。各ドメインの表示色は図 4.6、図 4.9 における表示色と対応している。

phot を処理した先行研究によると、青色光受容型 phot1 では J α ヘリックスの 603 番リジン残基が分解されたが、暗条件下では分解されなかったと報告されている [Kashojiya *et al.*, 2015]。以上より、phot1 LOV2-STK 断片においては青色光受容によって J α ヘリックス構造変化と連動した STK の相対配置変化が生じていると予測される。

本研究では phot1 について LOV2 - STK 断片で構造解析を行ったため、全長 phot1 の四次構造(図 4.13a)については未知である。しかしながら、phot1 LOV1 は分子間相互作用ではなく共有結合によって強固な二量体を形成している [Nakasako *et al.*, 2008] ため、LOV1 二量体も全長 phot1 において二量体構造にあると考えられる。従って、LOV1 と LOV2-STK 断片は回転対称軸を共有して二量体構造をとっていると考えるのが自然である。LOV2-STK 断片の対称点は S 字の中心にあるため、LOV1 の取り得る配置は図 4.37 右図(S 字型を横から見た図)において LOV2 の上または下に制限される。以上の議論に加えて LOV1 N 末端と LOV2 C 末端のつながりを考慮し、全長 phot1 の四次構造を推定した(図 4.38)。phot1 N 末端領域は 200 残基程の長さがあり、機能発現に伴い自己リン酸化される残基を多く含んでいる。このモデルが正しいならば N 末端領域と STK ドメインが相互作用可能なため (図 4.38)、分子間だけでなく分子内での自己リン酸化による機能発現の可能性も考えられる。しかしながら、LOV1 - LOV2 間のループ構造は約 170 残基であり、リンカー領域の 2 倍以上の長さである。この領域がどのような構造をとっているかは全く不明であるため、今回推定した四次構造はあくまで 1 つのモデルに過ぎない。

全長 phot2 については、LOV1 ドメインが溶液中・結晶中いずれの場合においても二量体を形成し [Nakasako *et al.*, 2004; Nakasako *et al.*, 2008]、LOV2-STK は溶液中で単量体として存在する [Takayama *et al.*, 2011]。従って、まず Dark1 推定形状について LOV1 二量体結晶構造を推定形状の中心に据えて重ね合わせた(図 4.39a)。Light 推定形状についても同様にして LOV1 による二量体形成を中心とした四次構造が推定された(図 4.39b)。Dark での構造と Light での構造を比較すると、LOV1 ドメインに対する LOV2 ドメイン、STK ドメインの相対配置が変化していた。先行研究で報告された LOV2-STK 断片の青色光依存的な構造変化と比較すると、LOV2-STK 断片で観測された分子伸長は全長 phot2 においては抑制され、代わりに屈曲するような相対位置変化が観測された。この違いは LOV1-LOV2 間の相互作用が影響していると考えられる。

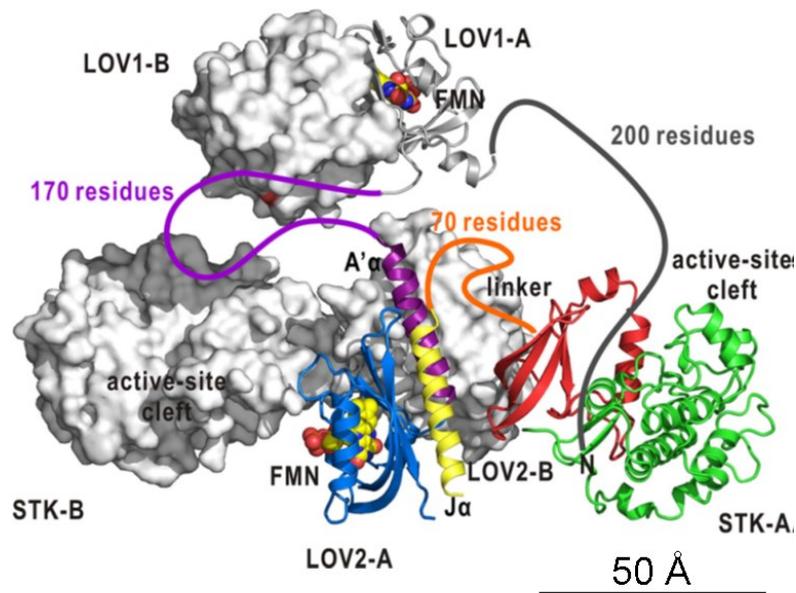


図 4.38 phot1 LOV2 - STK 断片 Dark1 推定四次構造(図 4.37a)に対して phot1 LOV1 結晶構造(図 4.6a)を配置した全長 phot1 推定四次構造。2 つのサブユニットについてそれぞれリボンモデル、空間充填モデルで表示している。配向は図 4.37 の右図と対応している。図中の黒色、紫色、橙色の曲線はそれぞれ結晶構造に現れなかった LOV1 N 末端、LOV1-LOV2 リンカー領域、LOV2-STK リンカー領域の推定配置である。

phot1、phot2 のいずれにおいても青色光受容によって LOV2 ドメインと STK ドメインの相対配置が変化するモデルが得られた。従って、phot の青色光に依存した分子内信号伝達メカニズムは伸長・屈曲運動による LOV2 - STK ドメインの相対位置変化であると示唆される。この伸長・屈曲運動を主として担うのは LOV2 - STK ドメインを物理的に繋ぐリンカー領域であると考えられる。これは先行研究 [Takayama *et al.*, 2011]で提案された分子内信号伝達モデルを支持する結果であると言える。

4.5.2 phot1 における 475 番、636 番リジン残基の役割

phot1 では野生型、D806N 変異体の両者において青色光依存的に $s = 0.005 \text{ \AA}^{-1}$ 付近での SAXS プロファイルの変化が観察された(図 4.18; 4.27)。一方、C512A/D806N

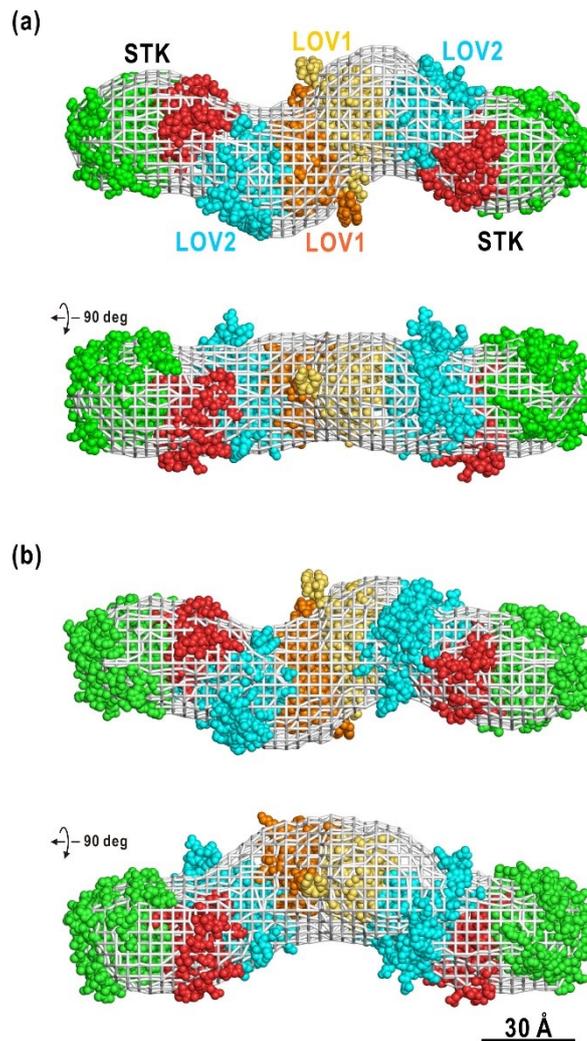


図 4.39 全長 phot2 Dark1 推定四次構造(a)と Light 推定四次構造(b)。各図下段は上段の表示から 90 度回転させて表示している。LOV1、LOV2 結晶構造及び STK ドメイン予測構造は空間充填モデルにて表示している。また各ドメインの表示色は図 4.6、図 4.9 における表示色と対応している。

変異体ではこの変化は観察されなかった。従って、この散乱角でのプロファイルの変化は青色光受容による構造変化であると考えられる。これを踏まえ、475 番、636 番リジン残基の役割について考察する。

K475A/D806N 変異体では青色光に依存して会合・解離が生じた。これは 4.3.5 節で述べた通り、青色光受容によって phot1 LOV2-STK 断片表面の性質(表面電荷

など)の変化に起因すると考えられる。この可逆的凝集は K475A 点変異に特有であり、また、先行研究より青色光照射は 475 番リジン残基と接する J α ヘリックスの構造に影響を与えることが明らかにされている [Harper *et al.*, 2003]。4.5.1 で予測された構造モデルでは 475 番リジン残基は LOV2-STK の界面に位置する残基であり、J α ヘリックスとも隣接している (図 4.26)。従って、青色光受容に伴い、475 番残基周辺は LOV2 ドメイン・J α ヘリックス双方の構造変化の影響を受けると共に、STK ドメインとのコンタクトが弱まることで溶媒への露出が大きくなると考えられる (図 4.37c)。この時 475 番残基近傍のコンフォメーションも変化しているならば、K475A の変異導入はこのコンフォメーション変化に影響を与え、結果的に Light 構造における表面の性質変化と凝集を引き起こすというモデルが推察される。凝集によってリン酸化活性が阻害されるのか、リン酸化活性を失うような構造変化 (あるいは構造変化の阻害)によって結果的に凝集を生じるのかは不明であるものの、J α ヘリックス及び 475 番リジン残基が青色光依存的な構造変化において重要な役割を果たしているということが強く示唆される。

K475A/D806N 変異体とは対照的に、K636A/D806N 変異体における光依存的な SAXS プロファイル変化は野生型・D806N 変異体とよく似ていた (図 4.27)。従って、K636A/D806N 変異体においては、青色光受容に伴って野生型と同様の構造変化が生じていたと考えられる。しかしながら、K636A の変異導入によってリン酸化活性は喪失していたため、構造変化に関わっているだろう 475 番リジン残基と対照的に 636 番リジン残基は LOV2 から STK へのシグナル伝達において重要な役割を果たしていると考えられる。

4.5.3 phot2 における LOV1 の役割

現在まで、青色光依存的な機能発現において本質的な役割を担っているのは LOV2 ドメインであり、LOV1 ドメインは青色光受容に伴う構造変化や STK の活性化にほとんど関わらないと考えられてきた。しかしながら、C170A/D720N 変異体、C426A/D720N 変異体の双方において、青色光照射による SAXS プロファイルの変化量は D720N 変異体と比較して小さかった。この結果は S₃₉₀ に遷移した LOV1・LOV2 ドメインが協同的に phot2 全体の構造変化に寄与している、ということを示唆している。生化学実験の結果もこのモデルを支持しており、例えばシロイヌナズナ由来 phot2 では C426A 変異体でもキナーゼ活性を持つことが *in vivo*、

in vitro の双方で確認されている [Christie *et al.*, 2002; Suetsugu *et al.*, 2013; Cho *et al.*, 2007]ほか、LOV1 が phot の青色光強度に対する感受性をコントロールしていることが報告されている。また、時間分解吸光度計測を用いた S₃₉₀ から D₄₅₀ への熱緩和測定によって、LOV1・LOV2 ドメインが互いの緩和時間に影響を与えていることが明らかにされている [Oide *et al.*, 2018] (図 4.40)。S₃₉₀-D₄₅₀ の存在比は光強度だけでなく緩和時間にも影響を受けると考えられるため、LOV1 ドメインは LOV2 ドメインの光反応サイクルを制御することで phot の機能を制御していると推測される。

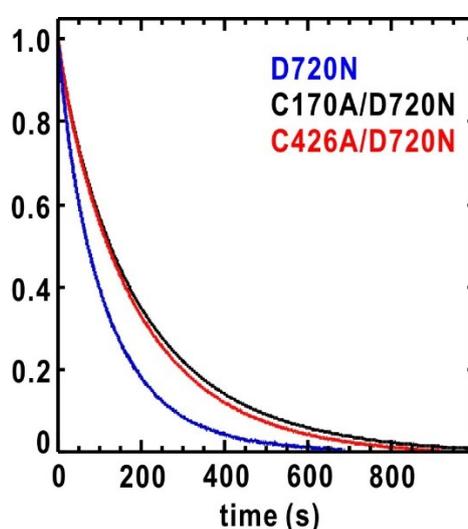


図 4.40 phot2 各種変異体の時間分解吸光度計測。測定波長は 450 nm であり、青色光照射停止からの経過時間でプロットされている。D720N 変異体は二成分の指数関数的減衰で近似でき、各成分の半減期はそれぞれ 54 秒、194 秒であった。前者が LOV2 の緩和、後者が LOV1 の緩和と推定される。一方、C170A/D720N、C426A/D720N 変異体はそれぞれ単一の指数関数的減衰で近似できた。C170A/D720N 変異体の半減期は 168 秒であり、D720N での LOV2 と比較して 3 倍の量であった。C426A/D720N 変異体の半減期は 68 秒であり、D720N での LOV1 と比較して半分以下の量であった。

4.5.4 phot 分子内信号伝達機構の普遍性

本研究では LOV2 - STK 相対配置の変化が phot 分子内信号伝達機構を担っているという、先行研究で提案されたモデルを支持する結果が phot1、phot2 双方で得られた。また、phot に関する先行研究としては 2014 年に単細胞真核生物である緑藻クラミドモナス由来 phot について SAXS による青色光に依存した構造変化の測定が行われている [Okajima *et al.*, 2014]。クラミドモナス由来 phot は本研究の解析対象である高等植物由来 phot と比較して、phot1、phot2 のようなアイソフォームを持たない、溶液中で単量体として存在する、などの相違点が存在する。また、生体内での役割も、走光性の制御など高等植物とは異なった機能を担っている [Huang & Beck, 2003; Im *et al.*, 2006; Trippens *et al.*, 2012]。このように高等植物由来 phot と機能面で大きく異なるクラミドモナス由来 phot であるが、シロイヌナズナ由来 phot1、phot2 と類似した機能発現メカニズムが推定されている [Okajima *et al.*, 2014]。従って、四次構造やアミノ酸配列の異なる 3 種類の phot (シロイヌナズナ由来 phot1、phot2、クラミドモナス由来 phot) で共通した分子内信号伝達メカニズムが推定されている。これは、「LOV2 - STK の相対配置変化が分子内信号伝達機構を担う」というモデルが phot における普遍的なメカニズムである可能性を示唆している。今後の phot 研究の進展によってこのメカニズムが普遍的なものであるか解明されることが期待される。

5章 SAXS による植物蛋白質フィトクロム B の構造解析

4章に続き、開発手法を全体形状未知である赤色光・遠赤色光受容蛋白質フィトクロム B に対して適用し、構造解析を行った。本章ではその詳細と結果を述べる。

5.1 赤色光・遠赤色光受容蛋白質フィトクロム B

5.1.1 植物内におけるフィトクロムの役割

フィトクロムは一部の緑藻及び確認された全ての陸生植物に存在する赤色光・遠赤色光受容蛋白質である [松下・長谷, 2004]。また、原核生物もフィトクロムと類似した蛋白質(バクテリオフィトクロム)を有することが報告されている [Yeh *et al.*, 1997]。これらの中には植物フィトクロムと同様の機能ドメイン構造を有する「カノニカルな」種の他に、基本的なドメインこそ共通するものの全体としては植物フィトクロムと大きく異なる機能ドメイン構成を有する種が存在する [Wiltbank & Kehoe, 2019]。バクテリオフィトクロムは基本的には異なる系統として扱われるが、カノニカルな種、あるいは断片については植物フィトクロムとの比較に用いられる。以下本章においては植物フィトクロムをフィトクロムと呼称し、原核生物由来のフィトクロムについてはバクテリオフィトクロムとして区別する。

フィトクロムが関わる生命現象は種子発芽や芽生えの緑化、花芽形成といった光形態形成に加えて避陰反応など多岐にわたるが、これらの反応は波長 600-700 nm の赤色光によって誘導され、波長 700-750 nm の遠赤色光によって打ち消される赤色光・遠赤色光(Red/Far-Red: R/FR)可逆反応であることが知られている [和田ら, 2001]。R/FR 可逆反応は最初にレタスの種子発芽において見いだされ [Borthwick *et al.*, 1952]、1959年にはこの現象の制御因子としてフィトクロムが発見・粗精製された [Butler *et al.*, 1959]。その後、1980年代末の遺伝子解析によってシロイヌナズナで 5 種類のフィトクロム遺伝子が発見され、対応する分子種がフィトクロム A-E (phytochrome A-E: phyA-phyE) と命名された [Sharrock & Quail, 1989]。最初に発見されたフィトクロムは phyA であり、phyA-E においてアミノ酸配列は互いに 50%程度共通している。また、現在では多くの陸生植物がシロイヌナズナと同様に phyA-E を有していることが明らかになっている [Mathews & Sharrock, 1997]。phyA-phyE のうち分子生物学レベルで研究が進展しているのは phyA、phyB の二種類のみである。また phyB-phyE はそれぞれ似通った振る舞いを示す。その

ため、以降は phyA、phyB のみに絞って記述する。

フィトクロムに典型的な R/FR 可逆反応は phyB によって制御されており、また phyB は低光量の赤色光に対する反応にも関わっている [Reed *et al.*, 1993; Shinomura *et al.*, 1996;]。一方で phyA は低光量反応のさらに 1/10000 程度の赤色光による超低光量反応と高照射量の遠赤色光に対する反応を制御しており、phyB とは大きく異なる振る舞いを示す [Nagatani *et al.*, 1993; Shinomura *et al.*, 1996]。両者は生体内における蛋白質量も大きく異なっており、phyA が暗所で生育された生体に多量に蓄積される一方で、phyB は前者の数十分の一程度の量であるものの光量によらず安定して存在している [Quail *et al.*, 1995]。また、phyA は光に対して不安定であり、光照射によって速やかに分解される [Furuya, 1993]。本研究では、R/FR 可逆反応を示す phyB についてその全体構造と機能発現メカニズムを明らかにすべく研究を行った。

5.1.2 シロイヌナズナ由来 phyB のドメイン構造と機能

以下ではシロイヌナズナ由来 phyB を単に phyB と記述する。phyB は 1172 残基のアミノ酸から成る分子量約 130 kDa の蛋白質である。また、phyA、phyB 共に溶液中では二量体を形成していることが明らかにされている [Tokutomi *et al.*, 1989; Nakasako *et al.*, 1990; Oide *et al.*, 2019]。ドメイン構造としては N-末端から PAS ドメイン [Taylor & Zhulin, 1999]、cGMP phosphodiesterase/adenyl cyclase/FhlA (GAF) ドメイン [Ho *et al.*, 2000]、phytochrome(PHY)ドメイン [Montgomery & Lagarias, 2002]、2 つの PAS ドメイン、ヒスチジンキナーゼと相同性の高い Histidine kinase-related domain (HKRD) [Yea & Lagarias, 1998]で構成されている(図 5.1) [和田ら、2001]。PAS ドメインは 4 章で触れた LOV ドメインと同様に PAS スーパーファミリーに属するものであるが、phot の LOV ドメインと異なり色素を持たない。

phyB においては GAF ドメインが光受容ドメインであり、色素であるフィトクロモビルン(phytochromobilin: PΦB)を共有結合によって格納している [Wu & Lagarias, 2000; Rockwell *et al.*, 2008]。また、phyB は PAS-GAF-PHY から成る N 末端側と PAS-PAS-HKRD から成る C-末端側の 2 領域がそれぞれ独立にモジュールとして立体構造を保持することが知られている。GAF ドメインを有する PAS-GAF-PHY が光受容モジュール(photosensory module: PSM)、PAS-PAS-HKRD がシグナル出力モジュール(output module: OPM)と呼ばれている。後で詳しく述べる

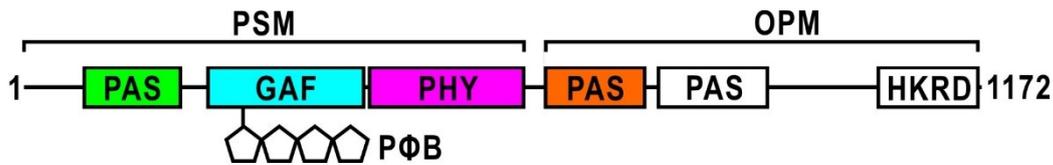


図 5.1 シロイヌナズナ由来 phyB 一次構造。GAF ドメインから伸びる五員環構造は色素分子である PΦB である。

が、phyB のシグナル伝達機構は不明であり、従って OPM が実際にシグナルの出力を担っている確証はない。しかしながら、PAS ドメインやキナーゼ様ドメインといったシグナル伝達に関与すると思われるドメインを有しており、またフィトクロム研究の歴史的な経緯より C-末端モジュールはシグナル伝達を担っていると考えられてきた。そのため、この部位はシグナル出力モジュールと呼ばれている。

5.1.3 PSM 立体構造と光反応

PSM の立体構造は 2014 年に X 線結晶構造解析によって得られ[Burgie *et al.*, 2014]、複数のヘリックスによって二量体を形成していることが明らかとなった(図 5.2)。PSM では GAF ドメインのみが光受容を担っているが、光受容による PΦB 周辺領域の微小な構造変化と協奏して PSM 全体で構造変化を示すことがバクテリオフィトクロム PSM の結晶構造解析の結果から報告されている[Yan *et al.*, 2009; Ulijasz *et al.*, 2010; Velazquez *et al.*, 2015; Burgie *et al.*, 2016]。従って、GAF ドメイン単体ではなく PSM 全体で光受容部位として機能していると考えられる。

5.1.1 節にて phyB は R/FR 可逆反応を制御していると述べた。これに対応して、phyB は赤色光受容型である Pr 型と遠赤色光受容型である Pfr 型の 2 種類の光受容型を取る [Quail, 2002](図 5.3)。Pr 型は波長 660 nm 付近に吸収ピークを持つ不活性な状態である。赤色光を受容した Pr 型 phyB は波長 720 nm 付近に吸収ピークを持つ Pfr 型 phyB に遷移し、活性を獲得する。この Pr 型から Pfr 型の遷移ではピコ秒からナノ秒スケールでの PΦB の光反応中間体遷移と、追従する秒スケールでの蛋白質構造変化が生じていることが共鳴ラマン分光と量子化学計算による先行研究から示されている [Mroginski *et al.*, 2006]。反対に Pfr 型は遠赤色光受容によって可逆的に Pr 型へと変換される。Pr 型は波長 760 nm 付近の光をほとんど吸収しないため、遠赤色光照射によって phyB はほぼ全てが Pr 型へと遷移する。一方で、

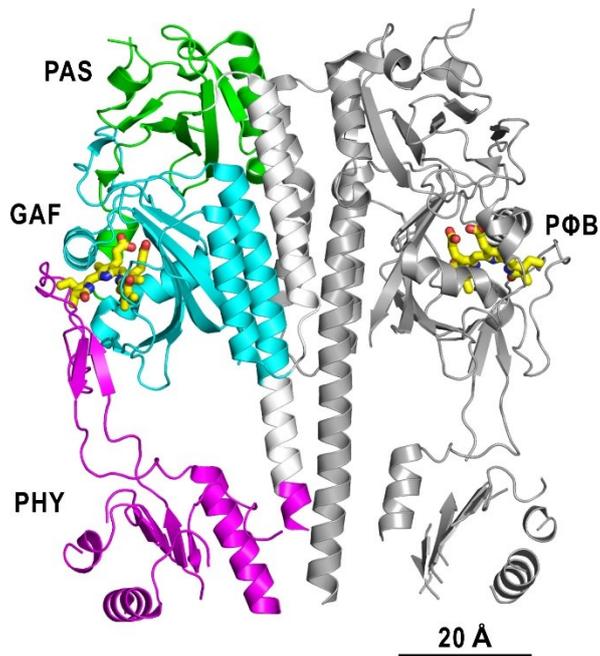


図 5.2 phyB PSM 二量体結晶構造(PDB ID: 4OUR)。GAF ドメイン内の黄色のスティックモデルが色素分子の PΦB である。

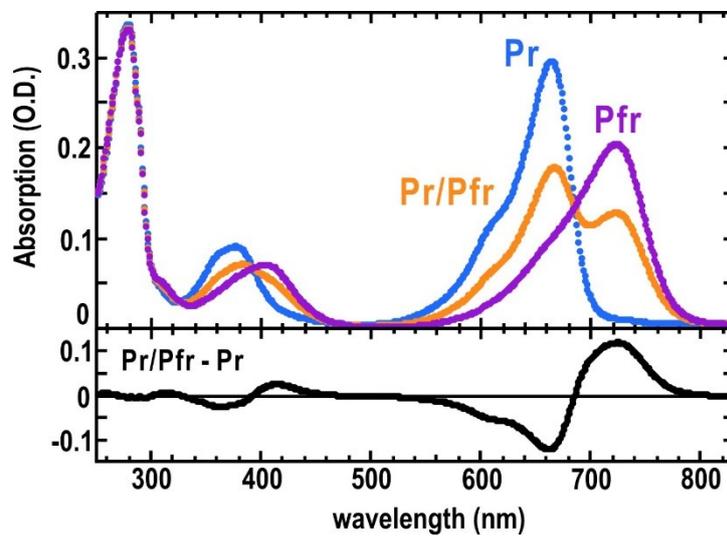


図 5.3 phyB 吸収スペクトル。下図は Pr/Pfr 混合状態と Pr 型との差スペクトルである。Pr/Pfr 混合状態は波長 660 nm、照射量 $1000 \mu\text{mol}/\text{m}^2 \cdot \text{s}^{-1}$ の赤色光照射による定常状態で測定している。Pfr 型スペクトルは Pr 型及び Pr/Pfr 混合状態スペクトルから計算されたものである。

Pfr 型は吸光ピークの裾野が波長 660 nm 付近まで広がっているため、一部の Pfr 型は再び Pr 型に遷移する。その結果、赤色光を照射された phyB は最終的に Pr/Pfr 型の混在する平衡状態に到達し、Pfr 型のみ phyB は実現不可能である。Pr 型-Pfr 型の光依存的遷移とは別に Pfr 型は Pr 型へ熱緩和によって遷移することが知られている [Remberg *et al.*, 1998]。この熱緩和は時定数が 20 分程度の(比較的)速い緩和と 90 分程度の遅い緩和の二成分で構成されており [Oide *et al.*, 2019](図 5.4)、先行研究によると、速い緩和が Pr-Pfr ヘテロ二量体での遷移、遅い緩和が Pfr-Pfr ホモ二量体での遷移に対応すると考えられている [Remberg *et al.*, 1998]。

5.1.4 OPM 立体構造

OPM 欠損 phyA が溶液中で単量体であったことから、OPM は phyA における二量体形成サイトとして機能すると考えられている [Nakasako *et al.*, 1990]。また、OPM は PSM と異なり全体形状が未だ不明であるものの、各ドメイン単位ではある程度の知見が得られている。まず、PAS ドメインは PAS スーパーファミリーに共通な α/β -scaffold 構造を有しているものと考えられる。実際、光合成細菌 *Rhodospseudomonas palustris* (*R. palustris*) 由来バクテリオフィトクロム PAS-GAF-PHY-PAS 断片結晶構造では C 末端由来 PAS ドメインが α/β -scaffold 構造を形成している [Bellini & Papiz, 2012] (図 5.5)。HKRD については、結晶構造既知の好熱菌 *Thermotoga maritima* 由来ヒスチジンキナーゼセンサードメインが *R. palustris* 由来バクテリオフィトクロム HKRD と配列相同性が高いと先行研究で報告されている [Evans *et al.*, 2004]。従って、phyB においても HKRD は類似した構造を取っていると予測される。しかしながら、実際のドメイン構造や OPM 内部でのドメイン配置などは現在まで明らかになっていない。

5.1.5 先行研究で予測された機能発現メカニズムと本研究の狙い

phyB、ひいてはフィトクロムの機能発現メカニズムに関する研究は長年にわたって多くのグループで行われてきたが、通説が二転三転とする極めて混沌としたものであった [松下、2008]。しかしながら、現在では OPM が核移行を担い、PSM の N 末端部位が核内での遺伝子発現調節に寄与するという説が受け入れられるようになった。まず OPM について、1996 年に phyB OPM 断片の核移行活性が報告され [Sakamoto & Nagatani, 1996]、その後全長 phyB と緑色蛍光蛋白質との融合蛋

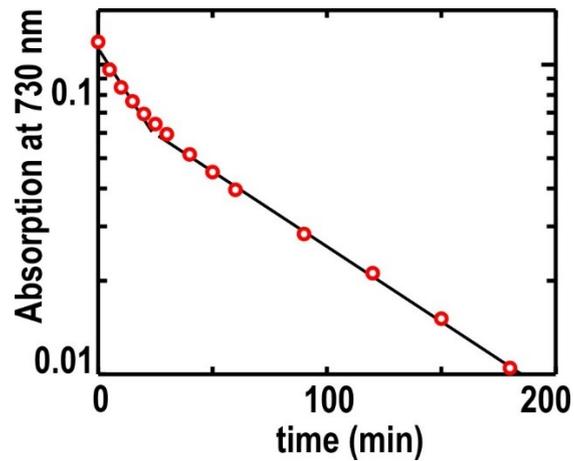


図 5.4 Pfr 型 phyB の熱緩和による吸光度変化。二成分の指数関数的減衰で近似可能である。

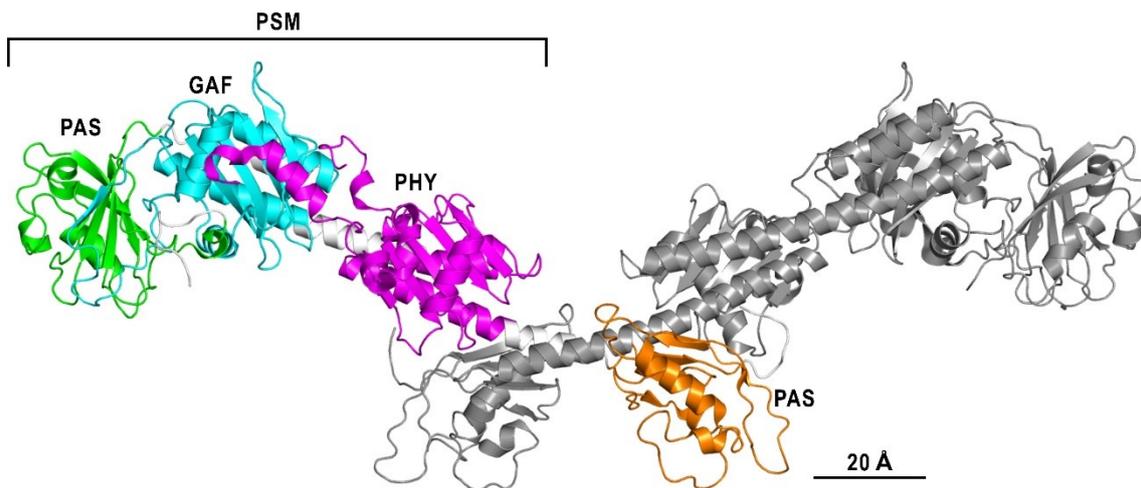


図 5.5 *R. palustris* 由来 PAS-GAF-PHY-PAS 断片結晶構造(PDB ID: 4GW9)。オレンジ色のドメインが OPM N 末端の PAS ドメインであり、反平行 β シートと α ヘリックスで構成された α/β -scaffold 構造を有していることが視認できる。

白質においても光依存的な細胞核移行が確かめられた [Yamaguchi *et al.*, 1999; Kircher *et al.*, 1999]。また、これらの結果と前後してシグナル伝達に重要な部位が OPM の特定領域に集中していることが変異株解析から明らかになった [Quail *et al.*, 1995]。このため、OPM が核移行とシグナル伝達を担っていると考えられてき

たが、2003年にPSM N-末端領域がシグナル伝達を本質的に担っており、OPMは光依存的な核移行だけを担っていることが報告された [Matsushita *et al.*, 2003]。また、網羅的遺伝子発現解析に基づき、シロイヌナズナゲノムの約一割の遺伝子がフィトクロムによって発現を制御されることが示されている [Tepperman *et al.*, 2006; Quail, 2007]。そのため、フィトクロムは遺伝子発現量を制御するという形でシグナルを伝達していると考えられている。これら先行研究に基づき、上述の機能発現メカニズムが有力な説となっている。

以上のようにフィトクロムの機能発現メカニズムについてはある程度の描像が得られているものの、その詳細、特に構造生物学的な知見はPSMの光依存的構造変化に留まっている。これは現在までに全長での構造解析がほとんどなされていないことが主な原因と考えられる。全機能ドメインが含まれたフィトクロムを扱った構造研究としてはPSM N末端を欠損したエンドウマメ由来 *phyA* についてSAXS解析及び電子顕微鏡観察を行った例がある [Nakasako *et al.*, 1990; Nakasako *et al.*, 2005]。1990年の電子顕微鏡観察に基づくモデルはOPMを二量体サイトとして扁平な楕円体及び円柱が平たく並んだような構造となっており(図 5.6)、*R. palustris* 由来バクテリオフィトクロムでも同様のモデルがSAXS解析から提案されている [Evans *et al.*, 2004]。その一方で *phyB* に関して全長での構造研究はなされていない。PSM二量体結晶構造は得られているものの、先行研究を考慮すると全長かつ溶液中では二量体を形成せず、サブユニットの相対配置が異なっている可能性が高い。事実、*R. palustris* 由来バクテリオフィトクロムではC末端にPASドメインが追加されたことで二量体形成機構が大きく異なっている(図 5.5)。従って、*phyB* について更なる知見を得るためには全長かつ溶液中での構造および光応答の解析を行わなければならない。そこで、本研究では全長 *phyB* についてX線小角散乱法と3章で提案した分子形状推定法を用いて全体形状を推定すると共に、種々の光照射条件でのSAXSプロファイルを比較することで *phyB* の光依存的な性状の変化を観測した。以下の節で測定、解析の詳細及び得られた結果について述べる。

5.2 SPring-8におけるX線小角散乱実験

5.2.1 測定試料

測定試料は本研究室の岡島公司特任助教(当時)によって確立されたプロトコルに基づいて発現・精製し、20 mM Tris-HCl、150 mM NaClを含む緩衝溶液(pH 7.5)

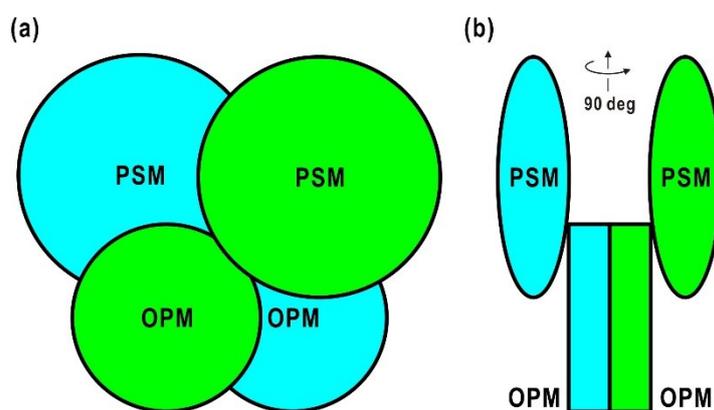


図 5.6 (a) 先行研究[Nakasako *et al.*, 1990]で提案されたエンドウマ
メ由来 phyA 推定形状。(b) (a)のモデルを 90 度回転させた図。

に溶解させた。phyB は溶液中で極めて不安定であり、時間経過によって非特異的凝集を生じやすい。そのため、本研究では試料をゲル濾過クロマトグラフィー(size exclusion chromatography: SEC)カラムに流し、SEC カラム通過後の溶液を直接フローセルに接続して SAXS 測定を行う SEC-SAXS 法 [Watanabe & Inoko, 2009]を用いた。SEC は試料溶液を多孔質ゲルが充填されたカラムに通過させることで試料を粒子径によって分離する手法である。ゲル中の多孔質を通過した場合粒子径が小さいほど経路長が長くなるため、遅く溶出する。このような特徴から、SEC-SAXS 法は凝集を生じた多分散試料でも単分散な SAXS プロファイルを測定できるという利点がある。SEC では試料溶液が数倍-十数倍に希釈されるため、カラムに注入する試料濃度を 7-8 mg/mL 程度の高濃度に調整した。

5.2.2 X 線小角散乱実験

phot と同様に、SAXS 測定は SPring-8 BL45XU で実施した。光学系セットアップは波長 1.0000 Å、カメラ長 2.0 m であった。本研究では理化学研究所の引間孝明研究員によって構築された SEC-SAXS 装置を用いて測定を行った。この装置では、固定の溶液試料セルに代わって内径 2 mm、厚さ 10 μm の石英キャピラリーを用いたフローセルを用いる(図 5.7)。これに送液ポンプ(UI22-110, FLOM, Japan)と SEC カラムを接続することで、SEC からの溶出試料溶液をそのままフローセルへ導いて SAXS 測定を行う。X 線照射野直前には UV 吸光計が設置されており、相対的な試料濃度を計測可能である。使用する SEC カラムは測定試料のサイズ(分子量)

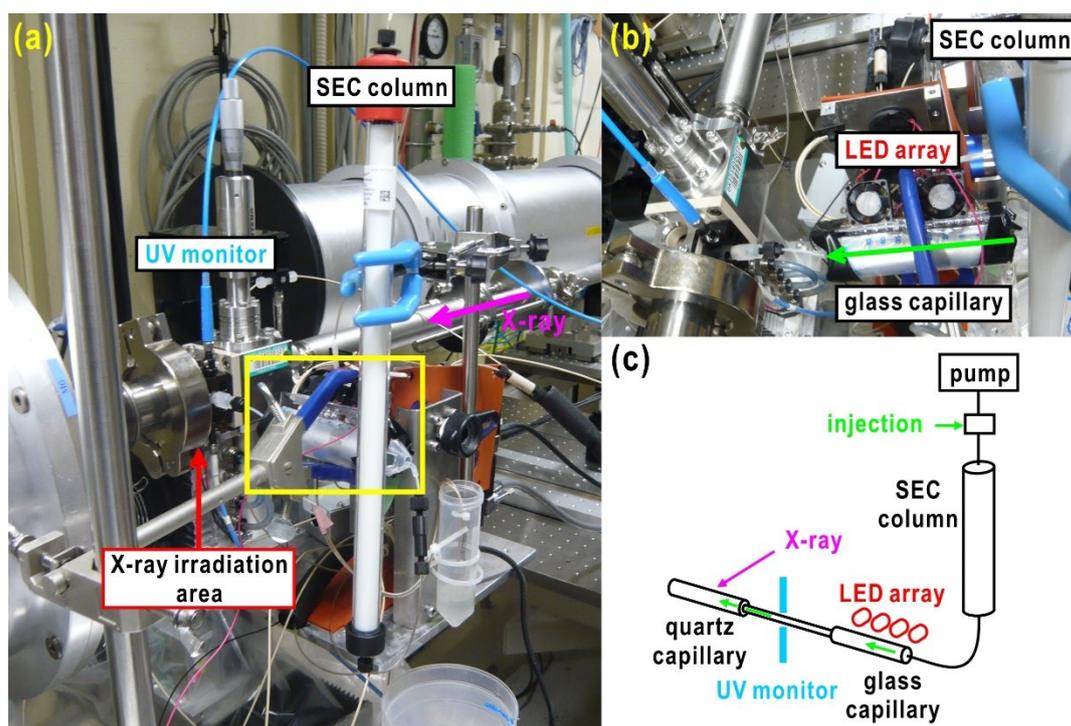


図 5.7 (a) SEC-SAXS 測定セットアップ。X 線照射野直上に伸びる水色のケーブルが UV 吸光度計である。(b) (a)の黄枠で囲まれた領域の別角度からの画像。黄緑は実際に SEC から溶出した試料溶液が通過する流路である。(c) 測定セットアップの模式図。(a)、(b)では確認できないが、X 線照射野には X 線光軸と垂直に石英キャピラリーが挿入されている。

によって決定されるが、本研究では phyB 二量体に相当する 200-300 kDa 程度を分離可能な Hiload 10/300 Superdex 200 Increase (GE Healthcare, Chicago, IL) を使用した。フローセルの流速は SEC カラムの耐圧限界を考慮して 0.75 mL/min に設定した。

今回の実験は SAXS による構造解析に加えて、phyB の赤色光に対する構造的な応答の観測が目的である。そのため、SEC カラムとフローセルの間に 100 mm 長のガラスキャピラリーを接続し、ガラスキャピラリー直上に複数の赤色光 LED(極大波長 660 nm)を設置した(図 5.7b, c)。今回のセットアップの場合、使用したガラスキャピラリーを通過するまでに $2000 \mu\text{mol}/\text{m}^2 \cdot \text{s}^{-1}$ の照射量で赤色光が約 10 秒間照射される。予備的に行われた実験では、Pr 型 phyB 溶液は $1000 \mu\text{mol}/\text{m}^2 \cdot \text{s}^{-1}$ の照

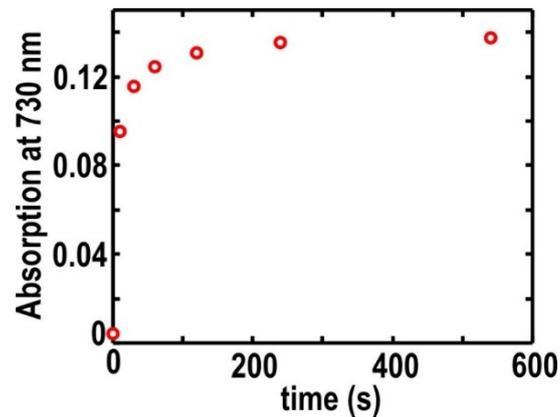


図 5.8 Pr 型から Pfr 型への遷移に伴う phyB の吸光度変化。Pfr 型は波長 730 nm に吸収極大を持つため遷移に伴って吸光度が増大している。

射量で 120 秒程度の時間をかけて光平衡状態へ到達した (図 5.8)。蛋白質濃度は芳香環側鎖やジスルフィド結合に由来する波長 280 nm の吸光度によって規格化可能であるため、規格化された吸光スペクトルにおける 730 nm での吸光度を Pr 型、Pfr 型スペクトルでの吸光度(図 5.3)と比較することで Pfr 型へ遷移した割合を計算可能である。図 5.8 の場合、SEC-SAXS 測定では光平衡状態には達していないものの、7 割程度の phyB が Pfr 型へ遷移していると考えられる。

5.2.3 データ処理

SEC-SAXS 法ではフローセルに対して連続的に X 線を照射し続け、適当なフレームレートに分割された一連の散乱パターンを取得する。本実験では露光時間 0.2 秒ごとに散乱パターンを記録した。記録された散乱パターンは測定装置と接続された計算機上で逐次 1 次元 SAXS プロファイルに処理された。また、入力された SAXS プロファイルを自動で Guinier 解析するプログラム AutoRg [Petoukhov *et al.*, 2007]を用いて慣性半径及び原点散乱強度を算出した。図 5.9 は標準試料であるグルタミン酸脱水素酵素(GDH)において SEC-SAXS 測定で得られた溶出プロファイルである。プロファイルの一点一点に対応して SAXS プロファイルが得られるが、各 SAXS プロファイルは露光時間が短く、S/N 比が悪い。その為、溶出ピークを中心に数十本のプロファイルを平均化することで S/N 比を向上させる。また、バックグラウンド散乱のプロファイルは試料溶出前の適当な領域を用いた。

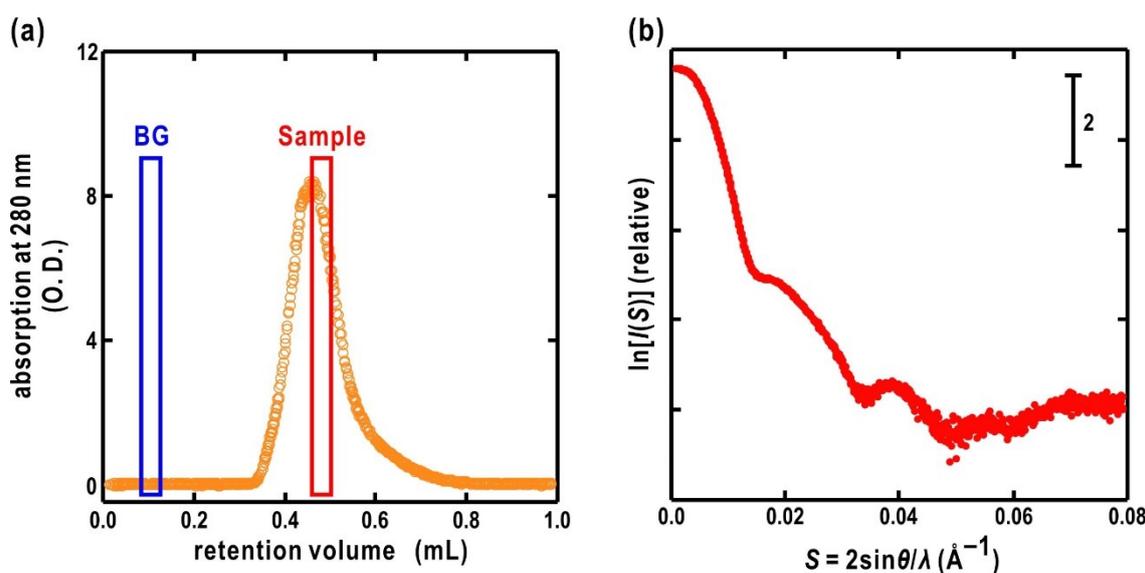


図 5.9 SEC-SAXS 測定による GDH SEC 溶出プロファイル(a)と SAXS プロファイル(b)。(b)のプロファイルは BG 領域の平均 SAXS プロファイルが差し引かれた Sample 領域の SAXS プロファイルを平均化することで得られている。

5.3 Pr 型 phyB SAXS プロファイル

まず、Pr 型 phyB について解析するべく、SEC カラムへ注入する直前に 2000 $\mu\text{mol}/\text{m}^2\cdot\text{s}^{-1}$ の照射量で波長 730 nm の遠赤色光を 60 秒間照射し SEC-SAXS 測定を行った(図 5.10a)。溶出プロファイルは複数のピークから成る多分散系の特徴を持ち、3 成分で近似された(図 5.10b; 表 5.1)。粒子系の大きい(溶出が速い)成分から順に α 、 β 、 γ と名付け、各成分について個別に平均化することでそれぞれの SAXS プロファイルを取得し、解析を行った(図 5.11a)。

γ 成分の Guinier プロットは直線で近似できる単分散を示し、慣性半径は 50.5 ± 0.2 Å と求められた(図 5.11b)。また、SAXS プロファイルから分子量を見積もるプログラム MoW [Fischer *et al.*, 2010]による解析では分子量約 218 kDa と見積もられた。これは phyB 単量体の分子量(130 kDa)のおよそ 2 倍程度の値であり、溶液中で二量体を形成するという先行研究による報告と矛盾しないものであった。距離分布関数から得られた最大分子長は 170 Å であった(図 5.11c)。また、 $S=0.015$ Å⁻¹ 付近になだらかなサブピークが見られた。式(2.15)よりこれは 60 Å 程度の距離の相関を反映しており、ドメインやサブユニットによる距離相関に由来すると考えられた。

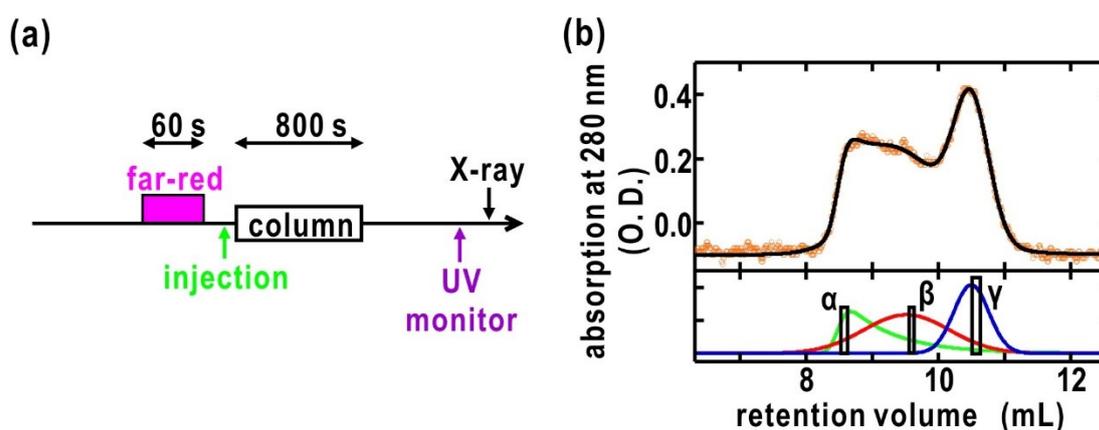


図 5.10 Pr 型 phyB SEC-SAXS 測定のための光照射条件(a)と実際に得られた SEC 溶出プロファイル(b)。(a)の矢印は時間の流れに対応している。(b)上段のオレンジの点の実験データ、黒い曲線が下段の3成分(α 、 β 、 γ)による近似曲線である。図 5.11 における各成分の SAXS プロファイルは(b)下段の黒枠領域の平均化による。本図(b)並びに図 5.13(b)、図 5.15(b)、図 5.17(b)における SEC 溶出プロファイルの曲線フィッティングは IGOR Pro を用いて計算した。

表 5.1 各測定条件と結果

光照射条件	SAXS 測定前の 光照射	各成分の割合 (%)			γ 成分のパラメータ	
		α	β	γ	R_g (Å)	M_w (kDa)
far-red (5.3 節)	none	28.1	41.7	30.2	50.5 ± 0.2	218
far-red (5.5 節)	red	35.1	31.3	33.6	---	---
far-red \rightarrow red \rightarrow far-red (5.6 節)	none	32.8	23.8	43.4	50.1 ± 0.3	220
far-red \rightarrow red (5.7 節)	none	75.9	6.8	17.3	51.0 ± 0.3	202

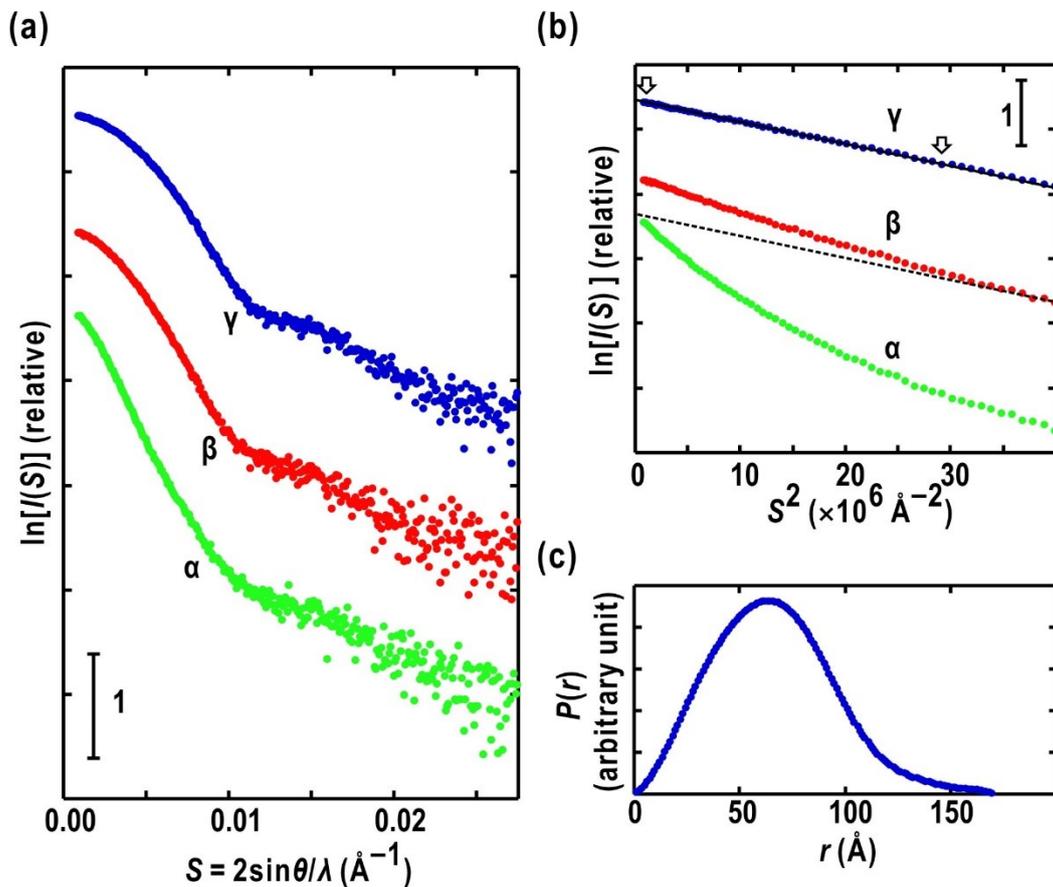


図 5.11 Pr 型 phyB SEC-SAXS 測定結果。(a)各成分の SAXS プロファイル。(b)各成分の Guinier プロット。(c) γ 成分の SAXS プロファイルから計算された距離分布関数。図 5.11 及び 5.14、5.16、5.18 の(a)、(b)では α 、 β 、 γ のプロファイルをオフセットさせてプロットしている。

一方、 α 、 β 成分では Guinier プロットが直線近似不可能であり、小角で上向きに反するという典型的な凝集体の SAXS プロファイルが得られた。 α 、 β 成分は γ 成分と比較して粒子径が大きいこと、小角領域を除いて α 、 β 成分と γ 成分の SAXS プロファイルがよく似ていることを考慮すると、 α 、 β 成分は凝集した Pr 型 phyB であると考えられる。仮に α 、 β 成分の凝集が、特定の表面で定まった配向による分子会合を示す特異的凝集であるとする、サブユニット配置に新たな距離相関が生じるため、 γ 成分と比較して $S=0.015 \text{\AA}^{-1}$ 近傍で散乱プロファイルに何らかの変

化が表れるはずである。ところが、 α 、 β 成分と γ 成分の SAXS プロファイルはこの散乱角でよく似ていることから、 α 、 β 成分は phyB 二量体の非特異的凝集の結果であると考えられた。各成分の積分から比率を計算すると、 γ に対応する単分散成分が3割、 α 、 β に対応する凝集成分が7割程度であった(表 5.1)。

5.4 Pr 型 phyB 全体形状の推定

5.3 節で得られた γ 成分の SAXS プロファイルについて3章の提案手法を適用することで Pr 型 phyB の全体形状推定を試みた。phot と同様に最適ダミー残基数を探索し、残基数 1075、最大分子長 170 Å、C2 回転対称性を有するという条件の下、560 回の独立した GASBOR 計算を行った。得られた 560 個の GASBOR モデルの形状分布は2つのグループに大別される非等方的な分布であった(図 5.12a)。クラス I-VI のグループは Twisted tube 型、クラス VII-X のグループは X 字型の形状であった。形状推定の判断材料として負染色電子顕微鏡観察を行ったところ、平均投影像は X 字型の形状を支持するものであった(図 5.12b)。従って、我々は Pr 型 phyB 全体形状をクラス IX に代表される X 字型であると推定した(図 5.12c)。推定形状は全長 120 Å の中心構造から 30-40 Å 程度の突起が X 字型に伸びた形状であった。

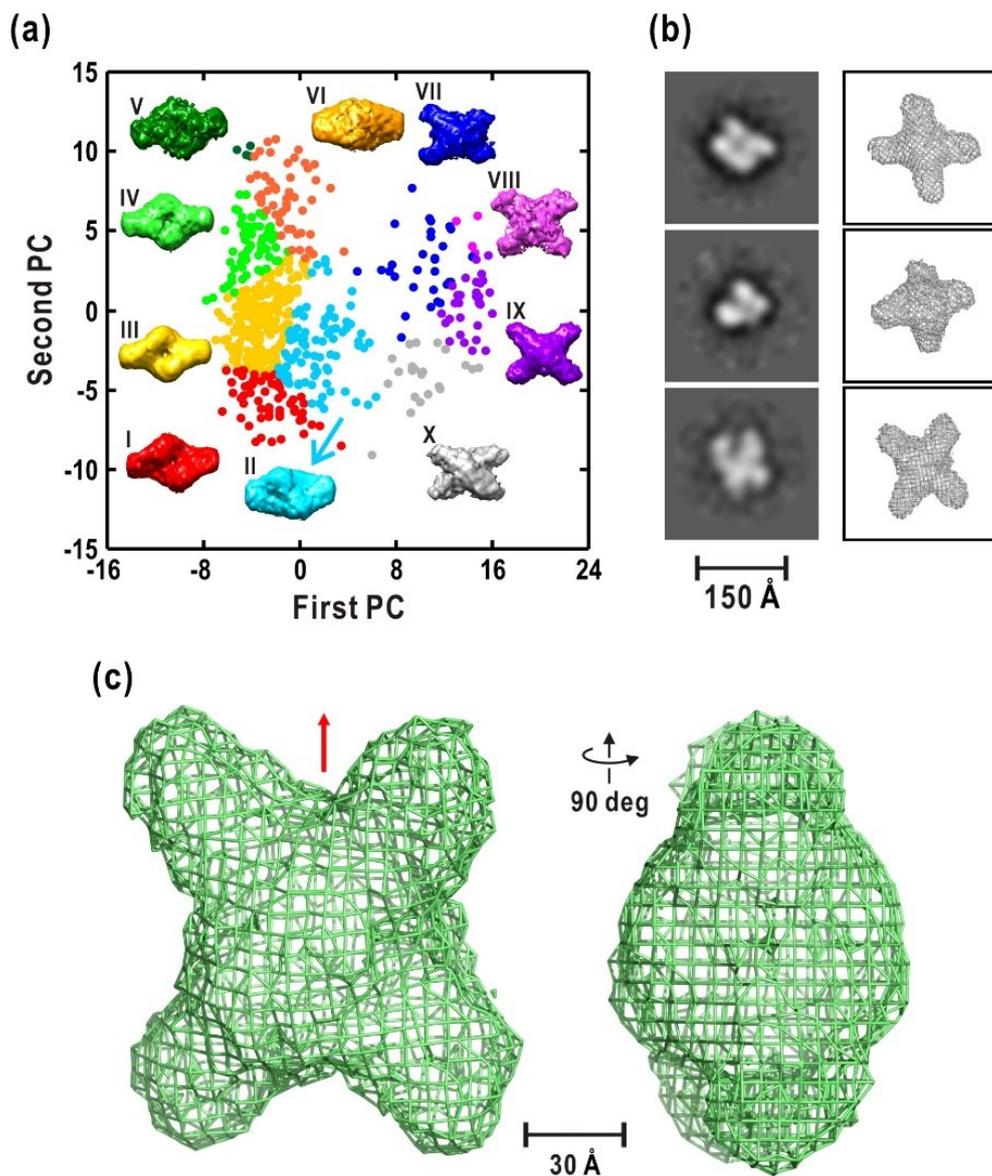


図 5.12 Pr 型 phyB 全体形状推定結果。(a)3 章の提案手法による Pr 型 phyB GASBOR モデル形状分布。(b)負染色観察で得られた Pr 型 phyB 投影像(左)と対応すると推定される配向からのクラス XI 平均形状(右)。電顕像解析プログラム RELION2.1[Kimanius et al., 2016]によって分類・平均化処理を行っている。(c)推定された Pr 型 phyB 推定全体形状。赤色の矢印は中心対称軸である。

5.5 赤色光照射による phyB SAXS プロファイルの変化

次に、赤色光照射による phyB の応答を観測する測定を行った(図 5.13a)。SEC カラムへの注入までは 5.3 節と同様であるが、SAXS 測定の直前にガラスキャピラリー中で赤色光を照射し、溶出プロファイル、SAXS プロファイルの比較を行った。得られた溶出プロファイルは図 5.10b とよく似た α 、 β 、 γ の 3 成分で近似されるものであった(図 5.13b; 表 5.1)。SEC からの溶出までは 5.3 節の測定と同等であるため、予想された結果であると言える。 α 、 β 成分についても Pr 型 phyB と同様に凝集由来の SAXS プロファイルであった(図 5.14a)。一方、 γ 成分は Pr 型 phyB と大きく異なり、Guinier プロットが直線近似不可能な SAXS プロファイルであった(図 5.14b)。Pr 型 phyB の SAXS プロファイルと差分によって比較すると、 $S < 0.003 \text{ \AA}^{-1}$ で散乱強度の上昇、 $0.003 \text{ \AA}^{-1} < S < 0.010 \text{ \AA}^{-1}$ で減少が見られた(図 5.14c)。小角領域の散乱強度上昇から、赤色光照射によって phyB が凝集を生じたことは明らかである。一方で $S = 0.015 \text{ \AA}^{-1}$ 近傍では散乱強度に変化がなかったため、Pfr 型への遷移に伴う凝集は非特異的なものであることが示された。なお、Pfr 型 phyB は Pr-Pfr ヘテロ二量体と Pfr-Pfr ホモ二量体の二種類が存在し得るが、本研究でこれらを区別することは不可能である。従って、これらを一括りにして Pfr 型 phyB とする。

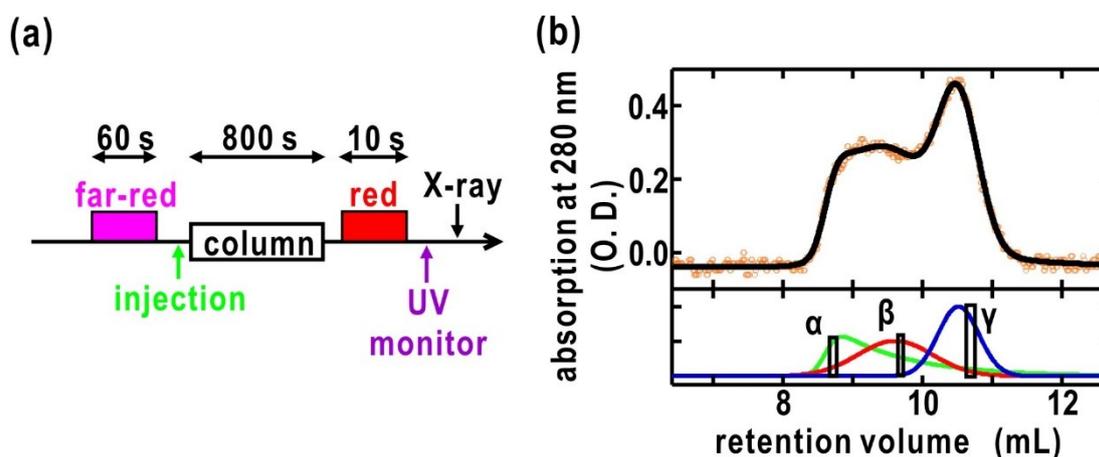


図 5.13 phyB 赤色光応答の SEC-SAXS 測定のための光照射条件(a) と実際に得られた SEC 溶出プロファイル(b)。5.3 節の条件と比較すると、SEC カラムからの溶出と X 線照射の間に 10 秒間の赤色光照射が挿入されている。

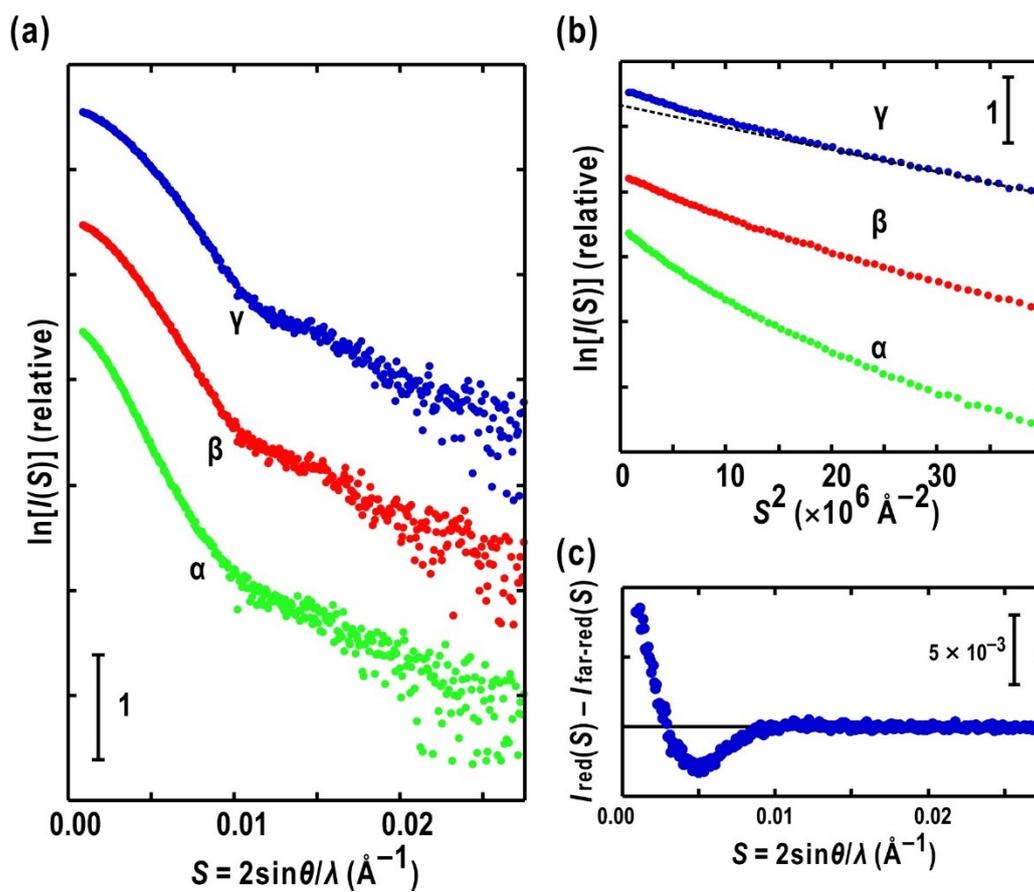


図 5.14 phyB 赤色光応答の SEC-SAXS 測定結果。(a)各成分の SAXS プロファイル。(b)各成分の Guinier プロット。(c) γ 成分の Pr 型 phyB SAXS プロファイル(図 5.11 の γ 成分)との差分プロファイル。

5.6 Pfr 型 phyB による非特異的凝集の可逆性

5.5 節では Pfr 型 phyB が非特異的業種を生じるという興味深い性質が明らかにされた。本節における測定では、この凝集が可逆的なものかどうかを評価する(図 5.15a)。この測定では 60 秒間の遠赤色光照射の後、30 秒間の赤色光照射を行った。これによって一部の Pr 型 phyB が Pfr 型へ遷移し、凝集を生じたと考えられる。次に遠赤色光を 30 秒間照射し、SEC カラムへ注入した。もしも凝集が非可逆的であるならば、凝集が増加することで α 、 β に対応する成分が増加し、 γ に対応する成分が溶出プロファイルの変化が期待される。

得られた溶出プロファイルはこれまでの測定と同様に 3 成分で近似可能であったが、 α 、 β 成分の減少と γ 成分の増加が見られた(図 5.15b; 表 5.1)。また、各成分の SAXS プロファイル、Guinier プロットは Pr 型 phyB とよく似たものであった(図 5.16)。 γ 成分について慣性半径、分子量を算出したところそれぞれ $50.1 \pm 0.3 \text{ \AA}$ 、220 kDa となり、Pr 型 phyB 二量体によく一致する値であった。溶出プロファイルの傾向より Pfr 型 phyB の非特異的凝集が可逆的であることは明らかであるが、本測定では Pr 型 phyB と比較して γ 成分が増加するという振る舞いが見られた。これは、5.3 節の測定では α 、 β 成分として分離された Pr 型凝集 phyB の一部が、赤色光・遠赤色光の逐次照射によって離合し単分散な Pr 型 phyB となったことを示唆するものである。

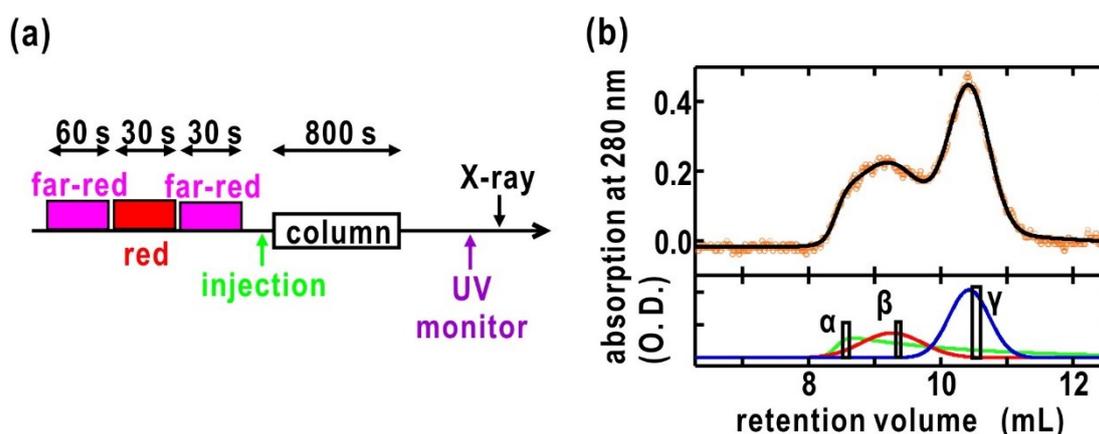


図 5.15 Pfr 型 phyB 非特異的凝集の可逆性測定のための光照射条件 (a)と実際に得られた SEC 溶出プロファイル(b)。5.3 節の条件と比較すると、SEC カラム注入前の遠赤色光照射に対して赤色光、遠赤色光の逐次照射を追加している。

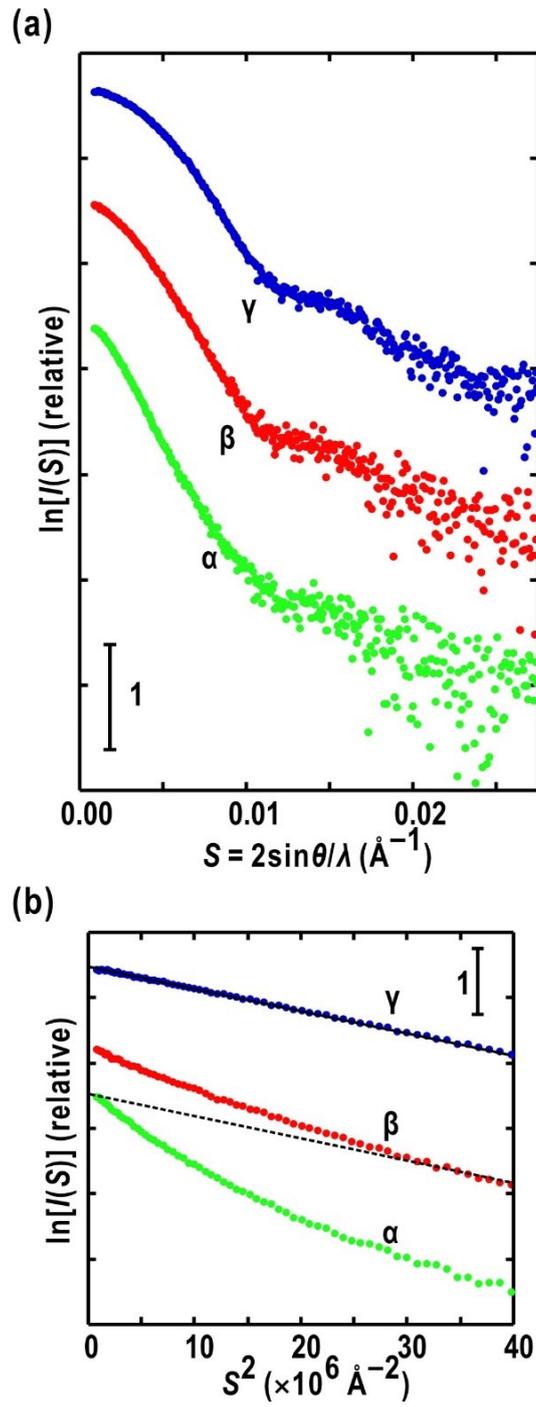


図 5.16 Pfr 型 phyB 非特異的凝集の可逆性測定結果。(a)各成分の SAXS プロファイル。(b)各成分の Guinier プロット。

5.7 Pfr 型 phyB 熱緩和の測定

5.1.3 節で触れた通り、時分割吸光度測定によって Pfr 型 phyB は熱緩和で Pr 型へ遷移することが明らかにされている。しかしながら、吸光度変化は発色団近傍の光化学的変化のみを反映するため、全体構造としても熱緩和が生じているかは不明である。そこで、構造的な熱緩和の有無を確認するべく、図 5.17a のように光照射と SEC-SAXS 測定を行った。この測定では、0 秒間の遠赤色光照射とそれに続く 30 秒間の赤色光照射を行った phyB 溶液を SEC カラムに注入した。SEC 開始から phyB の溶出までおよそ 800 秒を要するため、溶出時には一部の Pfr 型 phyB が Pr 型へ熱緩和していると考えられる。

溶出プロファイルはこれまでの測定と同様に α 、 β 、 γ の 3 成分で近似できたものの、これまでと異なり β 、 γ 成分が大きく減少し、 α 成分が支配的であった(図 5.17b; 表 5.1)。従って、事前の予想通りに単分散な Pr 型 phyB の一部が Pfr 型 phyB に変換され、凝集することで α 成分に移行したものと考えられる。 α 成分の SAXS プロファイルが Pr 型 phyB の場合とほぼ同様であったのに対し、 β 成分では小角領域での散乱強度上昇が抑制され、 γ 成分では Guinier プロットは直線近似が可能であったものの、慣性半径が 51.3 Å となり、若干の上昇が見られた(図 5.18)。 β 成分での SAXS プロファイル変化の理由は不明だが、 γ 成分に関しては増加した α 成分の溶出の裾野が小角の散乱強度に影響したものであり、 γ 成分自体は他の測定と同

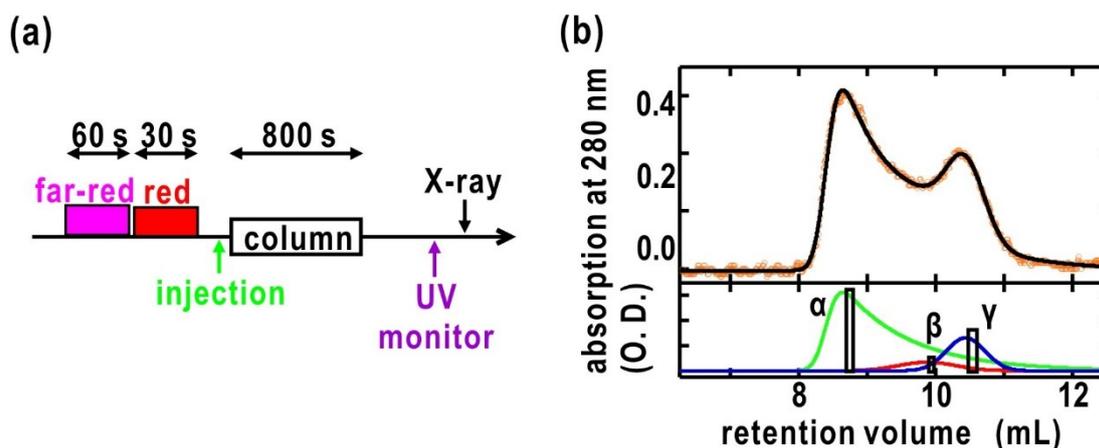


図 5.17 Pfr 型 phyB 熱緩和測定のための光照射条件(a)と実際に得られた SEC 溶出プロファイル(b)。5.3 節の条件と比較して、SEC カラム注入前の遠赤色光照射後に赤色光照射を追加している。

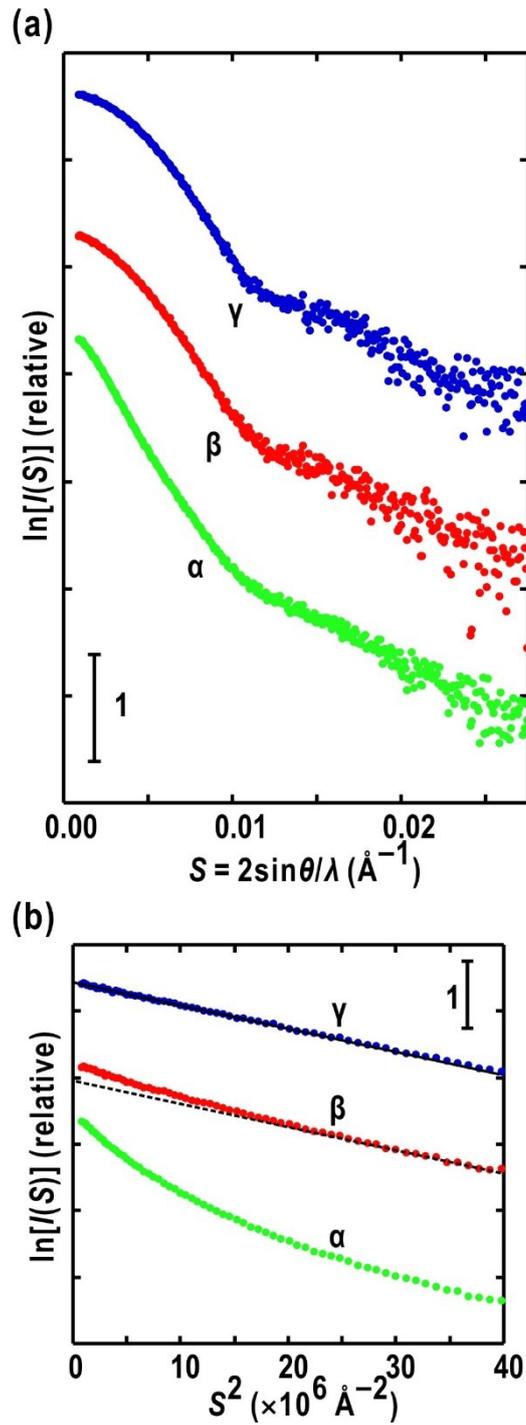


図 5.18 Pfr 型 phyB 熱緩和の SEC-SAXS 測定結果。(a)各成分の SAXS プロファイル。(b)各成分の Guinier プロット。

様に単分散な Pr 型 phyB であると判断した。

得られた溶出プロファイルから構造的熱緩和の有無を判断するため、これまでの測定結果との比較によって評価を試みた(図 5.19)。まず 5.3 節の結果より Pr 型 phyB では 68.1%が凝集状態、31.9%が単分散状態にある(図 5.11b; 表 5.1)。次に、ここに $1000 \mu\text{mol}/\text{m}^2 \cdot \text{s}^{-1}$ の照射量で赤色光を 30 秒間照射する Pr 型の 60%が Pfr 型に遷移すると考えられる(図 5.8)。Pfr 型 phyB は非特異的凝集を生じるため、この時点での単分散な phyB の割合は $0.4 \times 31.9 = 12.8\%$ と見積もられる。仮に構造的熱緩和が生じないならば、凝集した Pfr 型 phyB は光化学的に Pr 型へ遷移しても凝集したままであり、本節の測定における γ 成分の割合はこの見積もられた値になるはずである。しかしながら、実際の測定において γ 成分は見積もりよりも 1.4 倍大きい 17.3%となり、また慣性半径の値より単分散 phyB から成ることが確認された(図 5.17b、表 5.1)。この結果から、熱緩和によって Pfr 型から Pr 型へ光化学的に遷移した凝集 phyB は構造的にも Pr 型へ戻り、離合していることが示された。加えて、溶出プロファイルから見積もられる熱緩和した phyB の量は、分光測定で見積もられる量(図 5.4)よりも少ない。これは、構造的な熱緩和が光化学的な熱緩和に対し

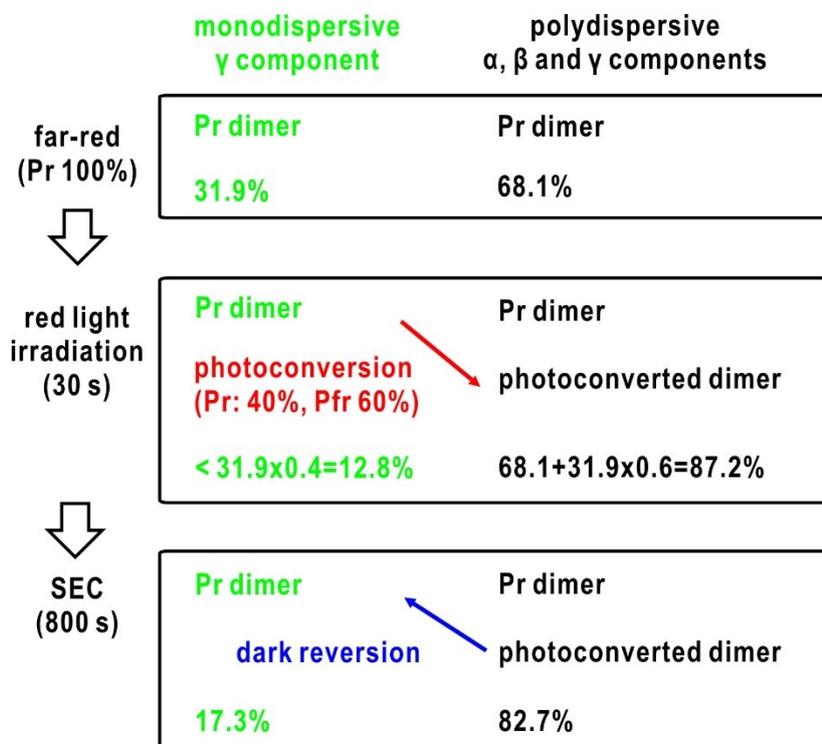


図 5.19 単分散成分の算出による Pfr 型 phyB 熱緩和の評価。

て遅延が生じていることを示唆している。

5.8 議論

5.8.1 Pr 型 phyB 四次構造の推定

5.4 節において Pr 型 phyB の全体形状を X 字型であると推定した(図 5.15)。本節では、先行研究で得られている PSM 結晶構造との比較を通して四次構造を推定すると共にその妥当性について議論する。phyB PSM(PAS-GAF-PHY)二量体結晶構造 [Burgie *et al.*, 2014]を 1 次構造におけるドメイン配列順序や中心対称軸を考慮して重ね合わせたところ、PHY ドメインを中心として開裂するように配置することで推定形状とよく一致した(図 5.20)。また、重ね合わせの結果と phyB ドメイン配列順序より OPM C 末端である HKRD ドメインにおいても開裂しており、二量体サイトが OPM PAS ドメインのいずれか、あるいは両方であると予測された。PSM 及び HKRD が開裂しているモデルは先行研究の結晶構造と相反するものであるが、OPM PAS ドメインを中心に二量体を形成しているバクテリオフィトクロム結晶構造 [Bellini & Papiz, 2012] (図 5.5)や PSM が開裂したバクテリオフィトクロム二量体クライオ電顕構造 [Li *et al.*, 2010]が報告されている他、OPM PAS ドメイン近傍に二量体サイトが存在すると予測した先行研究も存在する [Edgerton & Jones, 1992]。従って溶液中での構造が結晶構造から予測されるものと大きく異なる可能性は十分に考えられる。

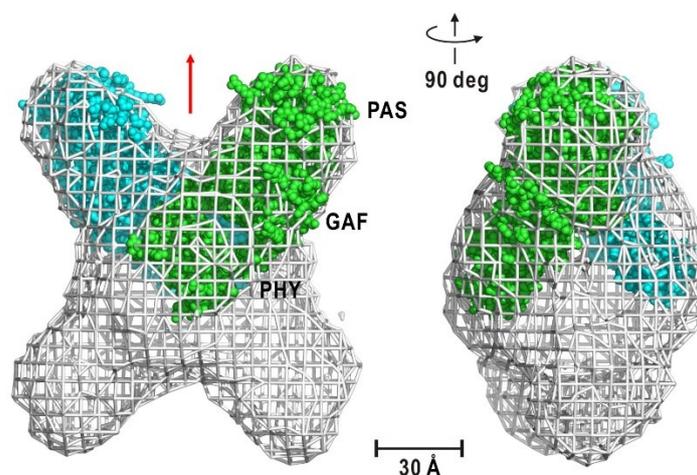


図 5.20 Pr 型 phyB 推定全体形状と PSM 結晶構造の重ね合わせ。

結晶構造との比較に加え、本研究室で過去に測定されたエンドウマメ由来 N 末端欠損 phyA(large phyA) の Pr 型における SAXS プロファイル [Nakasako *et al.*, 2005] (図 5.21a)に対して 3 章の提案手法を適用し、得られた Pr 型 large phyA 推定分子形状とも比較を行った。large phyA は N 末端 60 残基を欠損しているものの、機能ドメインは全て保持している。また、large phyA は phyB とのアミノ酸配列の相同性は高いが、phyB と比較して分子サイズがやや大きく先行研究では慣性半径が 57.1 Å と見積もられている。Pr 型 large phyA GASBOR モデル形状は等方的に分布したものの、phyB と同様に X 字型(クラス I-IV)と Twisted tube 型(クラス V-X)の 2 種類の形状に大別された(図 5.21b)。我々は phyB との配列相同性や先行研究での large phyA 電顕像[Nakasako *et al.*, 1990]を考慮し、クラス II に代表される X 字型の形状であると推定した(図 5.21c)。Pr 型 large phyA 推定形状に対して phyB PSM 結晶構造を重ね合わせたところ、phyB 推定形状と同様に PAS-GAF が開裂する配置でよく一致した。しかし、その開裂の度合いは phyB と比較して小さかった。5.1.1 節で述べたように、phyA と phyB は配列こそ相同性があるものの、生体内での振る舞いは大きく異なる。分子形状の比較から示唆された四次構造の差異が担う役割の違いに影響していると予測される。

5.8.2 赤色光受容に伴う非特異的凝集の意義

5.5、5.7 節における測定において Pfr 型 phyB の赤色光受容に伴う非特異的凝集が観測されたが、凝集に寄与している部位は全くもって不明である。しかしながら、藻類由来フィトクロムの OPM 欠損断片や PAS ドメインのみを OPM として有するバクテリオフィトクロムの SAXS 測定では Pr 型、Pfr 型のいずれにおいても凝集が確認されていない [Choudry *et al.*, 2018; Heyes *et al.*, 2019]。従って、phyB の非特異的凝集に寄与している部位は N 末端 PAS を除く OPM のどこかであると推測される。

Pfr 型 phyB の非特異的凝集は可逆的であったため、Pr 型 phyB の凝集とは異なるメカニズムであり、恐らくは構造変化に伴う分子表面の変化によるものであると考えられる。実際、Pfr 型への遷移に伴ってフィトクロム分子の表面が変化することは先行研究でも報告されている [Nakazawa *et al.*, 1991; von Horsten *et al.*, 2016]。Pr 型 phyB は生体内で生理活性を持たないため、この分子性状の変化はフィトクロムの機能発現メカニズムと密接に関連したものであると考えられる。仮説に過ぎ

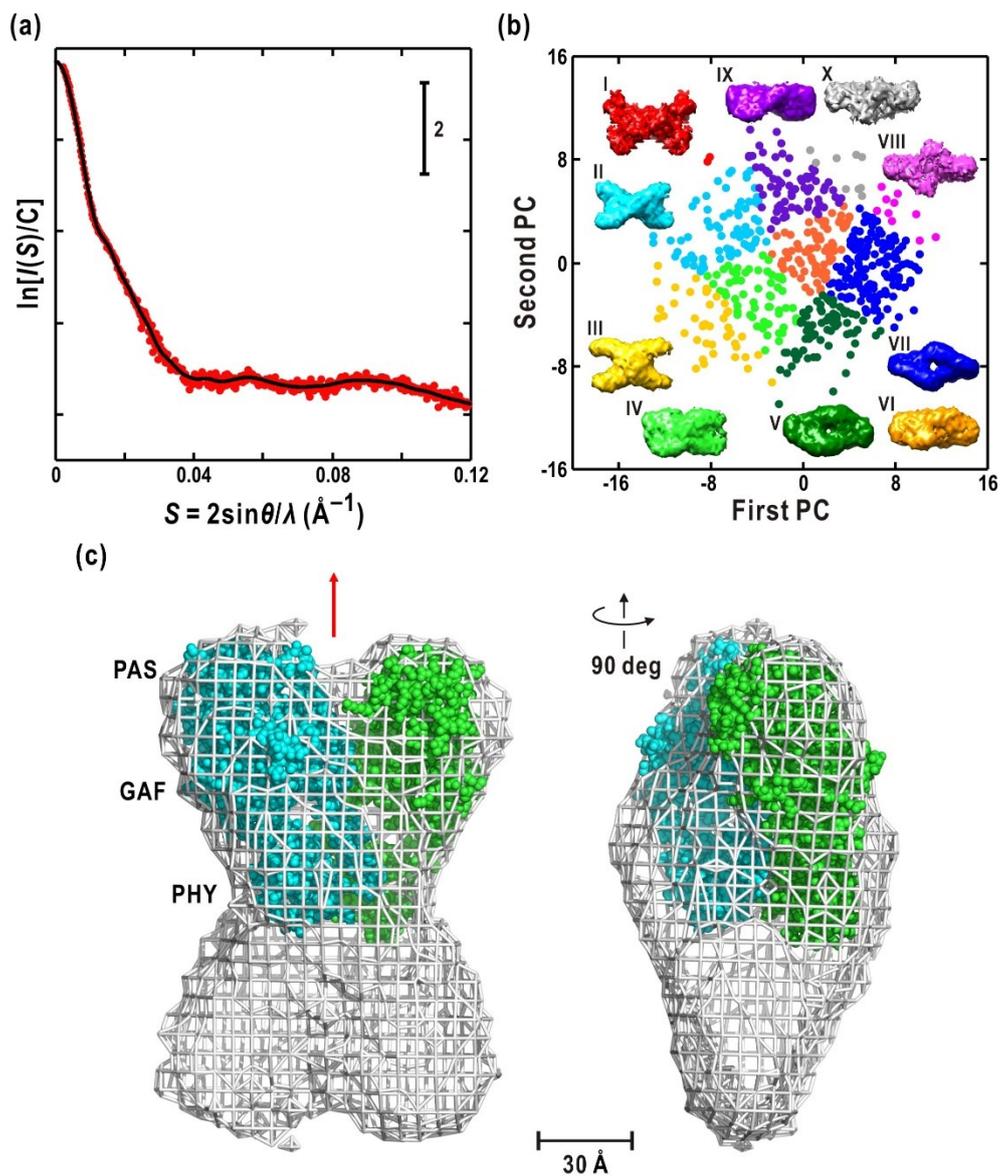


図 5.21 エンドウマメ由来 large phyA の全体形状推定。(a)Pr 型 large phyA SAXS プロファイル。黒の曲線は GASBOR モデルから計算されたプロファイルである。(b)3 章の提案手法によって得られた large phyA GASBOR モデル形状分布。(c)推定された Pr 型 large phyA 推定形状と phyB PSM 結晶構造の重ね合わせ。

ないが、我々はこの Pfr 型への遷移に伴う分子性情変化を細胞質内の核輸送受容体と相互作用するためのものであると推測している(図 5.22)。この仮説は現在唱えられている phyB の機能発現メカニズムにうまく合致するものである。第一に、光依

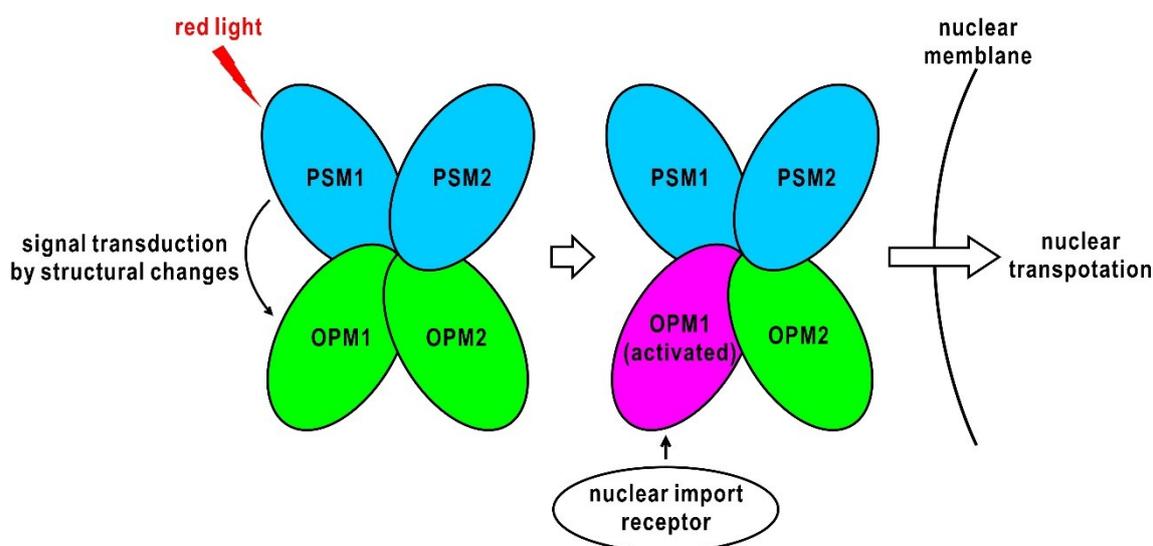


図 5.22 本研究で提案した赤色光受容による phyB 核移行メカニズムのモデル。本図においては簡単のために Pr-Pfr ヘテロ二量体で記述しているが、実際においても Pr-Pfr ヘテロ二量体で核移行するのか、あるいは Pfr-Pfr ホモ二量体のみが核移行するのかは現時点で不明である。

存在的な核移行[Sakamoto & Nagatani, 1996]では核輸送受容体との相互作用のスイッチングが必要になるが、光依存的な分子表面の変化はこの要求を満たすものである。本研究の測定では自己凝集を生じたが、これは *in vitro* であるために相互作用の相手が他の phyB 分子しかいなかったためだと考えられる。第二に、本研究で凝集が OPM 由来の性質であると示唆されたが、先行研究では核移行を OPM が独占的に担っていると報告されている [Matsushita *et al.*, 2003]。この2つの事実を単純に組み合わせると、OPM の核移行シグナルとしての分子表面変化が *in vitro* において自己凝集を引き起こしているというメカニズムが予想される。このように本研究で提案された仮説は現在の phyB シグナル伝達メカニズムにおける通説を補強するものとなっている。今後の研究によってこの仮説について検証がなされるものと期待される。

5.8.3 本研究の展望

本研究によって phyB の全体形状及び光依存的な分子性情変化が明らかになっ

た。これらは先行研究で報告されている事実と矛盾なく合致し得るものであり、phyB の機能発現メカニズムについて新たな知見を加えるものであった。今後の展望としては、やはり高分解能での全体形状解析が期待されるが、phyB の分子量を考慮すると X 線結晶構造解析あるいはクライオ電顕による解析が妥当と考えられる。X 線結晶構造解析をする上では結晶化の際に phyB の凝集しやすい不安定な性質が問題点となる。しかしながら、5.6 節の実験において赤色光、遠赤色光を照射して得られた Pr 型 phyB は他の測定条件と比較して単分散成分の割合が高かった。詳細なメカニズムは不明だが、赤色光・遠赤色光の逐次照射によって phyB の単分散性が改善されることが期待される。クライオ電顕を用いた構造解析に関しては、本研究室で既に着手しており、助走的な成果ではあるが、化学固定処理を施した Pr 型 phyB の 8 Å 分解能投影像が得られている(図 5.23)。今後の試料凍結条件の検討などでさらなる高分解能での投影像や三次元構造が得られるものと期待される。また、凍結時の光照射条件の変更などによって Pfr 型 phyB についても構造解析を行うことを現在検討中である。

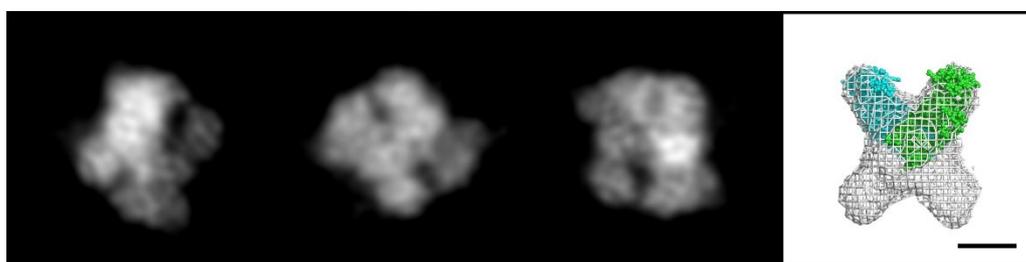


図 5.23 クライオ電顕観察像から得られた複数配向からの Pr 型 phyB 投影像。RELION 3.0 [Zivanov *et al.*, 2018]による分類・平均化処理を行っている。右図は同じスケールの phyB 推定全体形状であり、スケールバーは 50 Å である。

6章 クライオ電子顕微鏡法

3-5章までの研究により、X線小角散乱法を用いた分子形状推定について新規解析手法を考案すると共に、考案手法を種々の植物光受容蛋白質に適用することで各蛋白質の構造及び機能発現メカニズムについて新たな知見を与えることができた。これらの成果を踏まえ、さらに高分解能で溶液中蛋白質の構造変化を探るべく、本研究ではクライオ電子顕微鏡法を用いた構造解析を併せて行った。本章では像形成の基本的な原理を説明すると共に、観察像からの蛋白質構造解析法についても触れる。

6.1 電子顕微鏡における像形成理論

6.1.1 位相コントラストと弱位相物体近似 [堀内、1988]

自由空間を伝播してきた加速電圧 E の電子線が静電ポテンシャル V である物体を透過する状況を考える。入射電子線の波動関数を $\psi_0(\vec{r})$ とすると、透過直後の波動関数 $\psi(\vec{r})$ は物体との相互作用によって

$$\psi(\vec{r}) = q(\vec{r})\psi_0(\vec{r}) \quad (6.1)$$

のように変化する。 $q(\vec{r})$ は物体の透過関数と呼ばれる。この時、定常状態の Schrödinger 方程式を満たす電子線の波動関数 $\psi(\vec{r})$ は平面波で与えられ、その波数 k は

$$k = \frac{\sqrt{2me(E+V)}}{\hbar} \quad (6.2)$$

で与えられる。 $V=0$ である自由空間での波数との差を Δk とすると、 $V \ll E$ の条件下で

$$\Delta k = \frac{1}{\hbar} \left(\sqrt{2me(E+V)} - \sqrt{2meE} \right) \cong \frac{kV}{2E} \quad (6.3)$$

と近似的に表現できる。即ち、透過したポテンシャルに応じて位相が進む(図 6.1)。ここで、 m は電子質量、 e は電気素量である。観察されるポテンシャル場は数十 V だが、通常電子顕微鏡において電子線は数百 kV まで加速されるため、(6.3)式が成り立つ。物体の厚みを Δz とすると、入射電子線の物体透過で生じる全位相差は(6.3)式より

$$\int_0^{\Delta z} 2\pi \Delta k dz = \frac{\pi}{\lambda E} \int_0^{\Delta z} V dz = \sigma V_p, \quad \sigma = \frac{\pi}{\lambda E} \quad (6.4)$$

で与えられる。 σ は物体の相互作用係数、 V_p は物体の入射電子線方向への投影ポテンシャル、 λ は電子線の波長である。物体の厚みが薄ければ、入射電子線は振幅一定のまま位相のみ変化するため、透過関数を

$$q(\vec{r}) = \exp[i\sigma V_p] \quad (6.5)$$

と表せる。さらに、物体が非常に薄く、 $\sigma V_p \ll 1$ が成り立つ場合、透過関数は

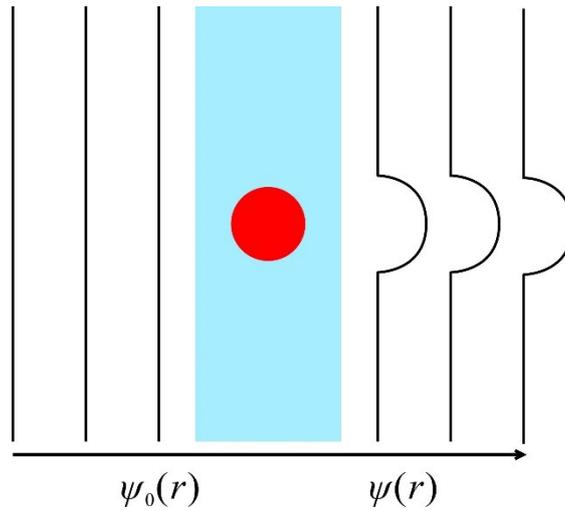


図 6.1 物体による電子線の散乱。黒線が平面波である電子線の波面である。赤い丸がポテンシャル V である物体である。クライオ電子顕微鏡法では観察試料は図の水色の四角のような薄い非晶質氷の膜である(6.2 節)。これはバックグラウンドとして一様に位相を進めるだけだが、物体による位相差はこのバックグラウンドとのコントラストとなることに注意する必要がある。一般に $V > 0$ であるため、物体との相互作用によって電子線の位相は進行方向に進む。

$$q(\vec{r}) \cong 1 + i\sigma V_p \quad (6.6)$$

と近似できる(弱位相物体近似)。この近似下では透過後の波動関数は入射波及び位相が $\pi/2$ ずれた散乱波の重ね合わせと見なせる。以降では基本的にこの近似が成り立っているものとして説明する。

6.1.2 位相コントラストとコントラスト伝達関数

前節にて、弱位相物体透過直後の電子線は入射波と位相シフトした散乱波の重ね合わせとして近似できることを示した。試料を透過した電子線は対物レンズによって後方焦点面で試料像の Fraunhofer 回折像を形成し、その後像面上で入射波と回折波の干渉光による像を形成する(図 6.2)。回折像の広がりに対して後方焦点面一像面間の距離が十分に大きければ、像面上の像は回折像を光源とした Fraunhofer 回折像に等しい。従って、電子顕微鏡における物体透過から像形成に至る過程は対物レンズ後方焦点面を仲立ちとして Fourier 変換を 2 回行うことに相当する。この時、像面上の位置 \vec{x} における強度

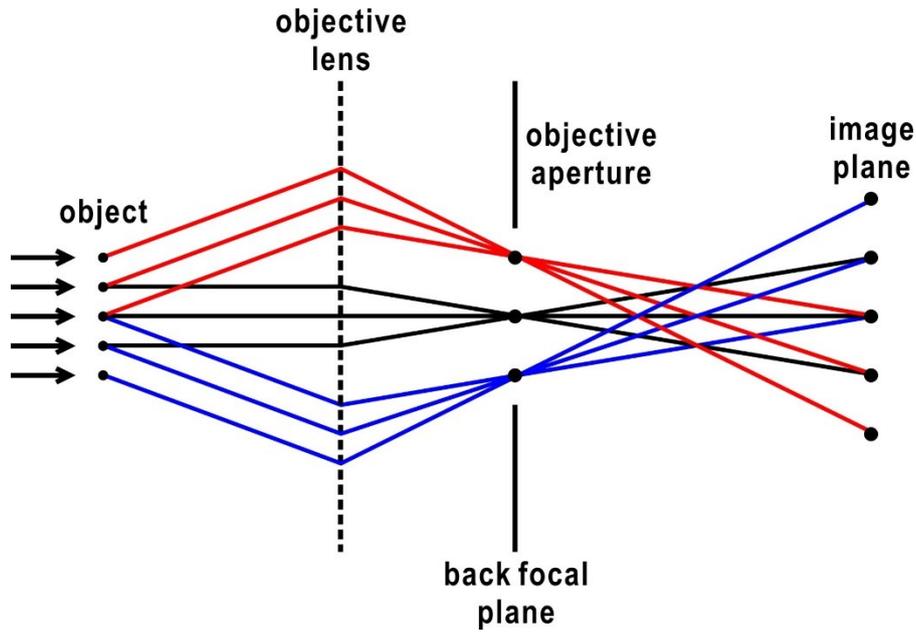


図 6.2 対物レンズによる結像の光路図。対物レンズ(objective lens)は簡略化して破線で示している。実際の測定では本図のように後方焦点面(back focal plane)に対物絞り(objective aperture)が挿入されている。後方焦点面では電子線の波動関数は物体の Fraunhofer 回折パターンを形成するため、対物絞りは高空間周波数でのカットオフとして作用する。

$I(\vec{x})$ は入射波 $\psi_0(\vec{x})$ と散乱波 $\psi_s(\vec{x})$ より

$$I(\vec{x}) = \psi(\vec{x})\psi^*(\vec{x}) = |\psi_0(\vec{x})|^2 + |\psi_s(\vec{x})|^2 + 2 \operatorname{Re} [\psi_0(\vec{x})\psi_s^*(\vec{x})] \quad (6.7)$$

となる。入射波の振幅を 1 とすると回折波の振幅は $\sigma^2 V_p^2$ であるため、第二項の寄与は極めて小さい。従って観察像のコントラストは第三項の干渉項由来である位相コントラストが支配する。

像面上の波動関数は物体の散乱に加えて対物レンズ収差由来の位相差による影響を受ける。これを踏まえ、位相コントラストを定式化するため像面上の振幅 $\psi(\vec{x})$ を後方焦点面での波動関数の逆 Fourier 変換として

$$\psi(\vec{x}) = \text{FT}^{-1} [Q(\vec{u}) \exp[2\pi i \chi(\vec{u})]] \quad (6.8)$$

で表す。ただし、 $Q(\vec{u})$ は透過関数による作用を含めた後方焦点面の位置 \vec{u} での波動関数、 $2\pi \chi(\vec{u})$ は対物レンズを通過した際に収差によって生じた位相変化である。(6.6)式より、弱位相物体では

$$\begin{aligned}\psi(\vec{x}) &= \text{FT}^{-1}\left[\left(\delta(\vec{u}) + i\sigma \text{FT}\left[V_p(\vec{r})\right]\right)\left(\cos 2\pi\chi(\vec{u}) + i \sin 2\pi\chi(\vec{u})\right)\right] \\ &= 1 - \sigma \text{FT}^{-1}\left[\text{FT}\left[V_p(\vec{r})\right]\left(\sin 2\pi\chi(\vec{u}) - i \cos 2\pi\chi(\vec{u})\right)\right]\end{aligned}\quad (6.9)$$

となる。(6.7)式を考慮すると、像コントラストは $\sin 2\pi\chi(\vec{u})$ によって決まる。この正弦型関数をコントラスト伝達関数(Contrast Transfer Function: CTF)と呼ぶ。高倍率観察では対物レンズ収差は球面収差係数 C_s が支配的であり、試料デフォーカス Δf による位相シフトと併せて CTF は

$$\sin 2\pi\chi(k) = \sin\left(\Delta f \pi \lambda k^2 - \frac{1}{2} C_s \pi \lambda^3 k^4\right)\quad (6.10)$$

の形で与えられる(図 6.3)。ただし、ここでは \vec{u} ではなく単なる空間周波数 k の関数とした。対物レンズが完全無収差の場合、試料像のコントラストは(6.7)式第二項のみから由来する極めて微弱なものとなる。そのため、電子顕微鏡観察では意図的に試料をデフォーカスさせて撮影する。デフォーカス量は撮影条件や加速電圧で変化するが、一般的な蛋白質構造解析では $0.5\text{-}3.0\ \mu\text{m}$ 程度である。

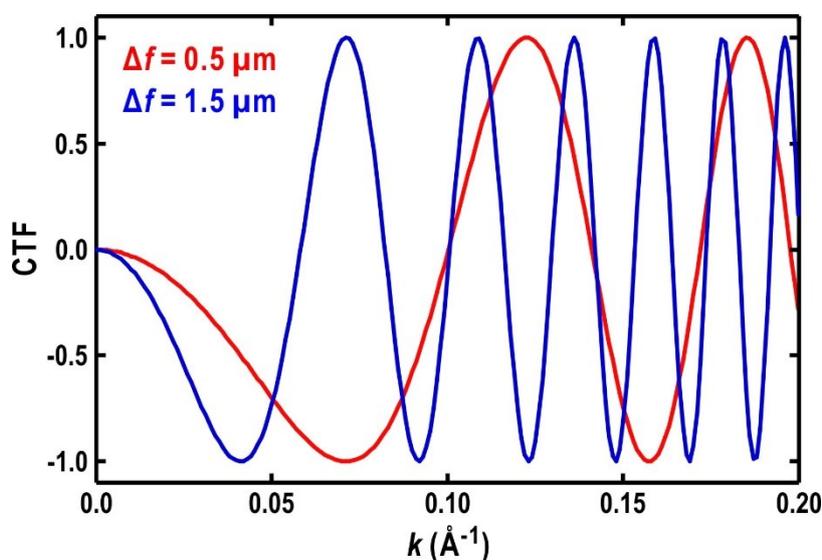


図 6.3 コントラスト伝達関数の理論曲線。球面収差係数は $2.7\ \text{mm}$ 、加速電圧は $300\ \text{kV}$ としている。図のように、デフォーカスを深くすると周期が短くなっていく。なお、このグラフは対物レンズに対して試料を離すアンダーフォーカスでの CTF 曲線である。オーバーフォーカスの場合アンダーフォーカスよりも低空間周波数にゼロ点が発生してしまい、コントラストが悪化する。そのため、電子顕微鏡観察は基本的にアンダーフォーカスで撮影する。

6.2 cryoEM における観察試料

電子顕微鏡の最大の特徴はプローブである電子線の相互作用の強さであるが、これによって様々な制限が生じる。まず、空気中では電子線が極めて短距離で減衰するため、電子線パスは試料周囲を含めて高真空に保たれる。当然このような環境では溶液試料は速やかに乾燥するため、電子顕微鏡が蛋白質構造解析に用いられ始めた当初は乾燥試料やグルコース溶液で包埋乾燥させた試料を観察していた[Henderson & Unwin, 1975]。これに対して、薄膜状溶液を急速凍結し、液体窒素などで冷却可能な低温ステージを搭載した電子顕微鏡で試料観察を行うクライオ電子顕微鏡法(cryoEM)が 1980 年代に開発・実用化された[Dubochet & McDowell, 1981; Adrian *et al.*, 1984; Dubochet *et al.*, 1988]。10⁵ K/sec 程度の速度での急速凍結することで溶液中の水分子は非晶質氷を形成する。これによって、乾燥や水分子の結晶成長による損傷のない溶液中構造の電子顕微鏡観察が可能となる。一般に cryoEM 観察においては試料溶液の厚みは数十 nm 程度であり、また蛋白質溶液は水素、炭素、窒素、酸素、硫黄を中心とした散乱能の小さい軽元素で構成されている。そのため、cryoEM 観察試料は弱位相物体として 6.1 節の像形成理論に従うと見なせる。薄膜状溶液の作成方法として現在主流となっているのは μm サイズの孔が並んだ親水化処理済カーボン膜に蛋白質溶液を滴下し余剰溶液を除去することで、試料溶液の表面張力を利用して成膜する方法である(図 6.4)。試料の急速凍結には液体窒素温度付近まで冷却した液体エタンを通常用いる。

一方で、cryoEM でも問題となるのが電子線照射による試料損傷である。溶液試料に電離放射線を照射すると非弾性散乱による共有結合の破壊や一部の水分子の酸素ラジカルへの転化により試料蛋白質が損傷する [中迫・山本、2015]。凍結試料では酸素ラジカルの拡散が抑制されるため損傷が低減されるものの、cryoEM 観察で試料構造を損なわないためには照射線量をおよそ $20 \text{ e}^-/\text{\AA}^2$ 以下に制限しなければならず[Grant & Grigorieff, 2015]、得られる像は S/N 比が極めて低い(図 6.5)。この低 S/N 比を克服するため、cryoEM による蛋白質構造解析で頻繁に用いられる手法が次節で説明する単粒子解析法である。

6.3 単粒子解析法 [Frank, 2006]

6.3.1 概要

6.2 節で説明した通り cryoEM で得られる蛋白質粒子像は極めて S/N 比が悪く、また

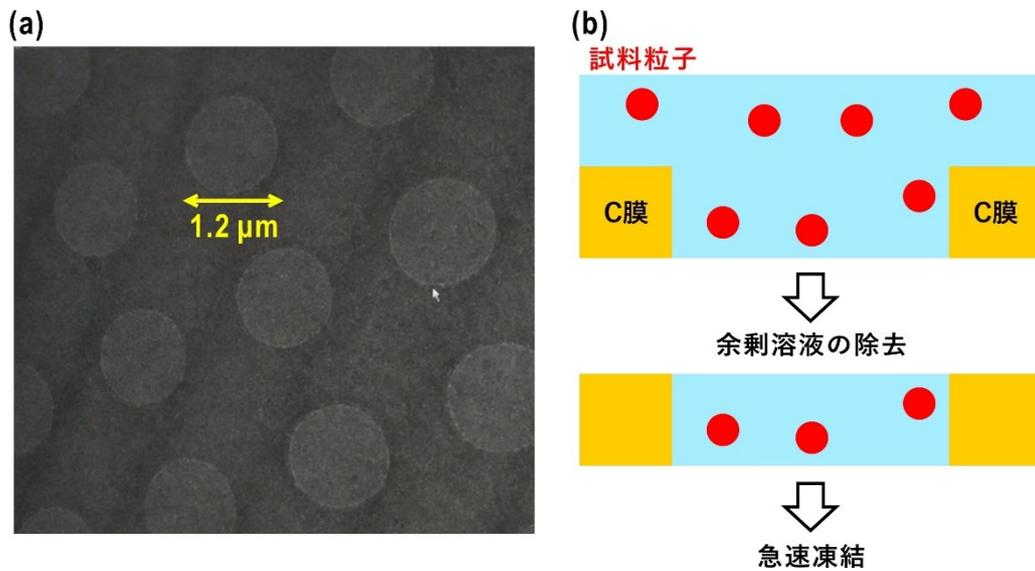


図 6.4 cryoEM 凍結試料の作成方法。(a) 試料支持膜であるカーボンフィルムの写真。円形の部分が 1.2 μm 径の孔である。(b) 薄膜状溶液作成の模式図。親水化処理されたカーボン膜に試料溶液を滴下すると、孔の部分に溶液が入り込む。この状態で余剰溶液を除去することで表面張力によって膜が張られる。

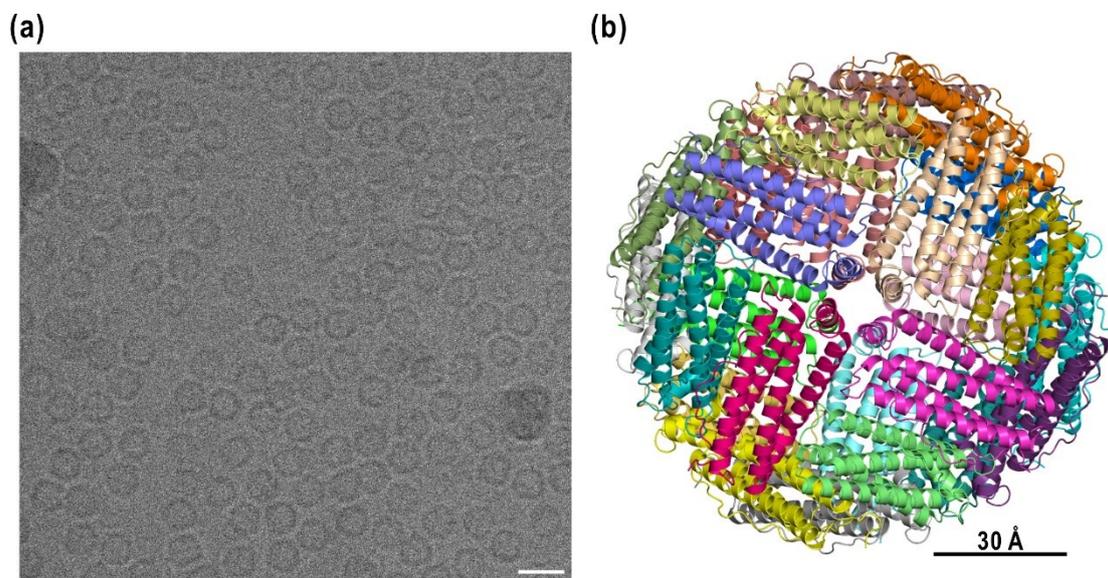


図 6.5 (a) ハツカネズミ由来 apo ferritin の照射線量 $20 \text{ e}^-/\text{\AA}^2$ での cryoEM 観察像。右下のスケールバーは 150 \AA 。(b) ハツカネズミ由来 apo ferritin 結晶構造(PDB ID: 6S61)。

得られる像は2次元投影像であるため、三次元構造はこれら粒子像から再構成しなければならない。蛋白質粒子はそのアミノ酸配列に固有な立体構造を取ると共に、薄膜状溶液中で光軸方向に対して自由な配向を取るはずである。これに基づいて単粒子解析法では「蛋白質粒子像は投影された配向のみが異なり、同一の構造(単粒子)由来である」と仮定し、同一配向からの投影像の平均化によるS/N比の向上と、異なる配向の投影像からの三次元再構成を行う。各粒子像の配向は未知であるため、実際にはプロジェクションマッチングという処理を交えて三次元構造を取得する。プロジェクションマッチングは参照構造から投影像を計算し、投影像と粒子像を比較することで配向を決定する処理である。三次元再構成は投影断面定理に基づいて行われる。この定理によると、「二次元投影像のFourier変換」と「三次元構造のFourier変換の、投影方向を法線として原点を通る断面」は等しい。従って、決定された配向に従って各粒子像のFourier変換をスライスとして挿入していくことで三次元構造のFourier変換が得られ、これを逆Fourier変換すると目的の構造となる。適当な初期構造から出発してプロジェクションマッチングと三次元再構成を反復的に繰り返すことで少しずつ配向の精度と3次元構造の分解能が改善されていき、最終的に試料蛋白質の三次元構造に収束する(図6.6)。適切な初期構造であれば正しい構造に収束することは数学的に確かめられている。経験則としては初期構造が概形レベルの分解能($\sim 50 \text{ \AA}$)で正しければ、粒子像に問題がない限り正しい構造に収束する。以降の小節ではcryoEM観察から三次元再構成に至るまでの重要な処理を簡単に説明する。

6.3.2 Beam Induced Motion とその補正

cryoEMでは試料への電子線照射に伴い、試料損傷だけではなく試料移動も生じる。これはBeam Induced Motion (BIM)と呼ばれる現象である。BIMは凍結試料を保持するカーボン膜の熱膨張による一方向への移動(global motion)と、凍結薄膜の膨張による非一様な移動(local motion)の2種類が存在する。BIMはかつてcryoEMにおいて到達分解能が $\sim 20 \text{ \AA}$ で頭打ちになっていた最大の要因である。現在では電子線検出器の大幅な性能向上によって観察像を短い露光時間で分割するフレーム撮影が可能になったため、各フレームの位置ずれ計算によるBIMの補正(motion correction)が標準的な解析の一部となっている(図6.7) [Brilot *et al.*, 2012; Campbell *et al.*, 2012; Li *et al.*, 2013]。motion correctionでは観察像全体のシフトによってglobal motionを補正し、その後観察像を 5×5 程度の領域に分割し、各領域をシフトさせることでlocal motionを補正する [Zheng *et al.*, 2018]。

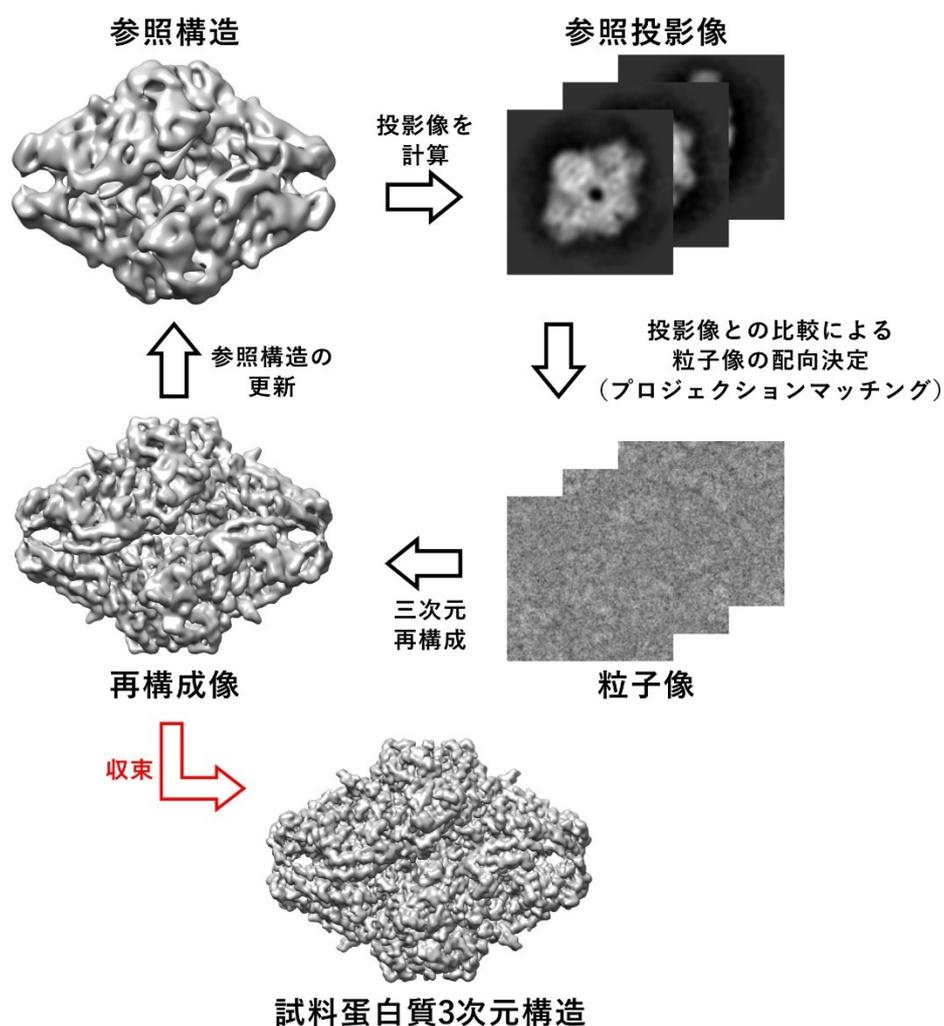


図 6.6 単粒子解析法の概略図

また、フレーム撮影にはもう一つの大きな利点が存在する。6.3 節で述べた通り試料損傷を避けるためには照射線量が $20 \text{ e}^-/\text{\AA}^2$ 以下に制限される。しかし、これは高分解能の構造情報を保つための基準であり、各分解能における許容照射線量は空間周波数に対して指数関数的に減少することが先行研究で報告されている [Unwin & Henderson, 1975; Baker *et al.*, 2010]。従って、各フレームについて、そのフレームの撮影時点での照射線量を許容する空間周波数の情報だけを取り出すことで構造情報の損失を回避しつつ像コントラストを改善することが可能である (Dose Filtration) [Grant & Grigorieff, 2015]。これはフレームごとの処理であるため、BIM の補正と並列して行われる。

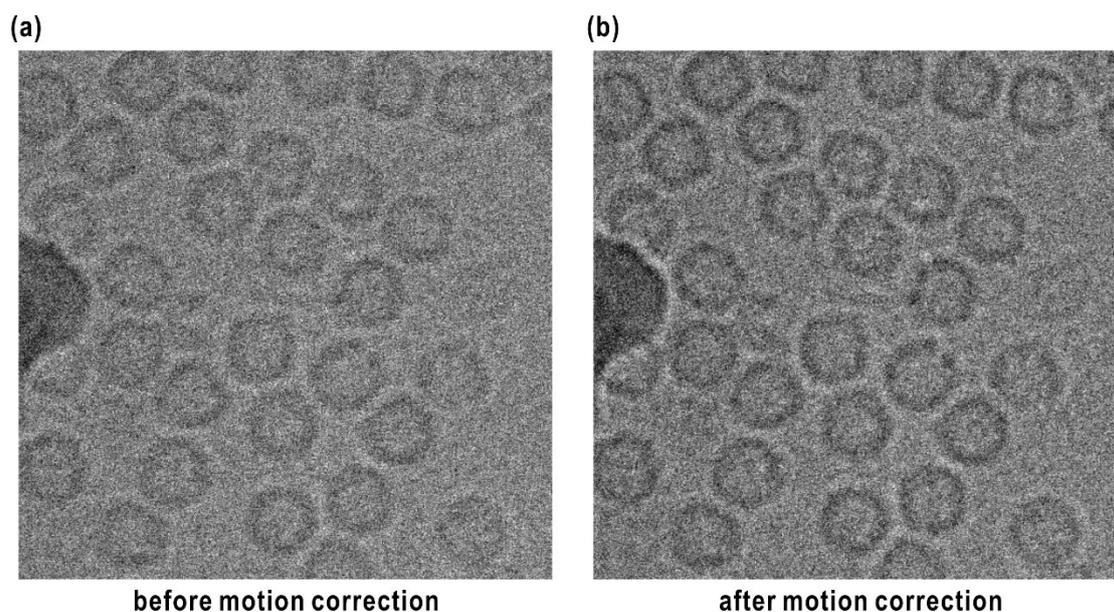


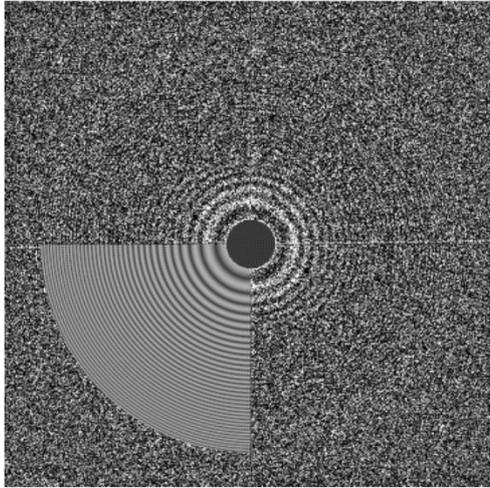
図 6.7 図 6.5 の cryoEM 観察像について motion correction 及び dose fractionation を行わずに積算した観察像(a)と行った観察像(b)。電子線照射線量は $100 \text{ e}^-/\text{\AA}^2$ である。補正によってコントラストが向上しており、例えば左のコンタミネーションの縁が鮮明化している。本図での補正には Unblur&Summovie [Brilot *et al.*, 2012; Grant & Griegorieff, 2015] を用いた。

6.3.3 CTF の推定と補正

cryoEM 観察像は投影ポテンシャルを反映した位相コントラスト像だが、CTF の Fourier 変換である点拡がり関数が畳み込まれ、構造情報に変調がかかっている。そのため正しい構造情報を得るためには CTF の補正とそのためのパラメータ推定を行わなければならない。変調は実空間では点拡がり関数の畳み込みだが、逆空間では CTF の乗算であり、取り扱いが単純である。そのため、推定・補正共に逆空間上で行う。

まず推定について簡単に説明する。CTF は複数のパラメータに依存するが、球面収差係数や電子線波長は装置によって決定される値であるため、CTF の推定とは即ちデフォーカス値の推定に等しい。cryoEM 観察像を Fourier 変換すると Thon ring と呼ばれる CTF 由来のリングパターンが表れる [Thon, 1971] (図 6.8)。デフォーカス値はこの Thon ring に対してフィッティングを行うことで推定する。なお、前節の BIM 補正を行うことで像コントラストが改善し、Thon ring の S/N 比も向上する。そのため、CTF の補正は BIM

(a)



(b)

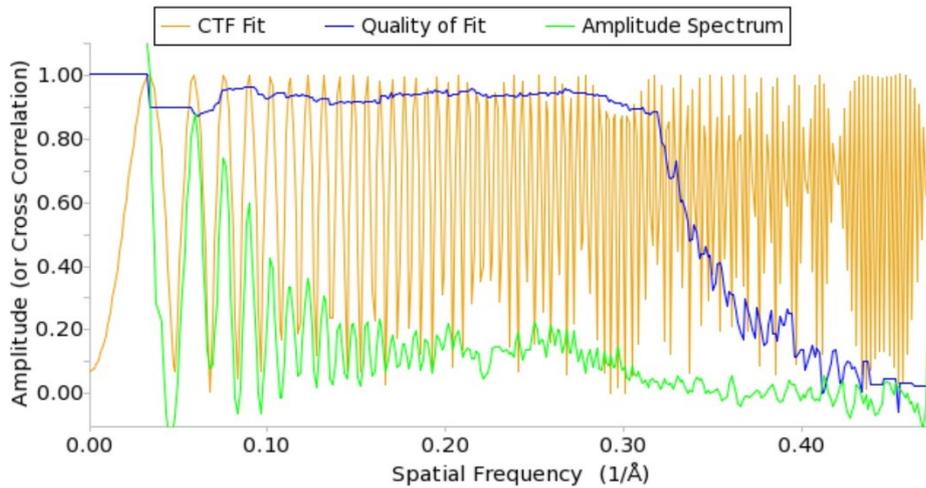


図 6.8 (a)cryoEM 観察像(図 6.7b の全体像)の Fourier 変換像。画像中心の同心円状パターンが Thon ring である。左下は推定パラメータから計算されたパターンであり、実験データの Thon ring と周期がよく一致していることがわかる。(b)一次元での CTF 曲線のフィッティング。緑色の曲線がバックグラウンド減算された実験データ、オレンジ色の曲線がフィッティング曲線、青色の曲線が実験データとフィッティング曲線の相関係数である。本図の CTF パラメータは CTFFIND 4.0 [Rohou & Grigorieff, 2015]で推定し、統合解析環境 cisTEM [Grant *et al.*, 2018]にて出力された図を用いている。

の補正後に行われる。

次に補正方法だが、CTF の補正は粒子像毎に Wiener フィルターによって行われる。観察された粒子像の Fourier 変換を \bar{X} 、正しい粒子像の Fourier 変換を \bar{F} とすると、 \bar{X} のピクセル j での値 X_j は

$$X_j = \text{CTF}_j F_j + N_j \quad (6.11)$$

と表せる。CTF_{*j*} はピクセル j での CTF の値、 N_j はピクセル j でのノイズの値である。ピクセル j に対する Wiener フィルター S_j は推定値 \hat{F}_j と真値 F_j の平均二乗誤差を最小化する関数、即ち

$$S_j = \arg \min_{S_j} \mathbb{E} \left[\left| \hat{F}_j - F_j \right|^2 \right], \quad \hat{F}_j = S_j X_j \quad (6.12)$$

として与えられる。ただし $\mathbb{E}[\cdot]$ はノイズのみ異なる複数の同様の粒子像のアンサンブル平均であり、ノイズが乱数であるためこのような定式化となっている。ノイズと信号の間に相関がない場合、Wiener フィルターは(6.12)式を偏微分することで

$$S_j = \frac{\text{CTF}_j^*}{\left| \text{CTF}_j \right|^2 + P_j^N / P_j^F} \quad (6.13)$$

となる。但し、 P_j^N 、 P_j^F はそれぞれノイズ、信号のパワースペクトルである。定性的には、分子部分で位相を打ち消しつつ、分母で振幅の補正を行うフィルターである。また、構造情報が失われている CTF のゼロ点近傍ではフィルターは単に S/N 比となって発散を防ぐ。このゼロ点近傍での情報欠損を補うため、単粒子解析では複数のデフォーカスの異なる画像を足し合わせることで欠損のない構造を得る。なお、(6.13)式では CTF を複素数として扱っているが(6.10)式では CTF は明らかに実数関数であり、実際においても電子線が光軸から傾斜しているなどの特殊な状況でのみ CTF を複素数として取り扱う必要が出てくる。そのため、Wiener フィルターにおいても実数として扱って実用上問題は無い。

6.3.4 三次元再構成

cryoEM の分野では多くの研究室によって開発された解析プログラムが公開されているが、その中でも三次元再構成関連のプログラムは他の処理部分と比較して圧倒的に数が多く、プロジェクションマッチングや再構成計算、収束判定など多くの点で流儀が異なる。ここでは、本研究で使用したプログラム RELION [Scheres, 2012a; Scheres 2012b] での定式化に従って説明する。

RELION では、三次元再構成は Wiener フィルターに基づいて行われる。反復計算の中で新たに再構成される三次元構造の Fourier 変換 \vec{v}^{new} について、ボクセル l での値 V_l^{new} は N 枚の粒子像 \vec{x}_i ($i = 1, 2, \dots, N$) から、以下の式によって計算される。

$$V_l^{\text{new}} = \frac{\sum_i^N \int_{\phi} \Gamma_{i\phi}^{\text{old}} \sum_{j=1}^J P_{ij}^{\phi T} \frac{\text{CTF}_{ij}}{\sigma_{ij}^{2\text{old}}} X_{ij} d\phi}{\sum_i^N \int_{\phi} \Gamma_{i\phi}^{\text{old}} \sum_{j=1}^J P_{ij}^{\phi T} \frac{\text{CTF}_{ij}^2}{\sigma_{ij}^{2\text{old}}} d\phi + \frac{1}{\tau_l^{2\text{old}}}} \quad (6.14)$$

ここで、 X_{ij} は粒子像 \vec{x}_i のピクセル j での値、 CTF_{ij} は粒子像 \vec{x}_i のピクセル j での CTF 値、 $\Gamma_{i\phi}^{\text{old}}$ は一回前の構造 \vec{v}^{old} とのプロジェクトンマッチングから決定された \vec{x}_i の配向が ϕ である確率密度分布である。RELION では重み $\Gamma_{i\phi}^{\text{old}}$ の下 ϕ についての積分を行うことで粒子像の配向を一意に決定されるパラメータではなく、広がりを持った分布として扱っている。 $P_{ij}^{\phi T}$ はスライスを三次元的に挿入する演算子であり、ピクセル j の値を与えられた配向 ϕ に従ってボクセル l へ挿入する $L \times J$ 行列として定義される。 $\sigma_{ij}^{2\text{old}}$ 、 $\tau_l^{2\text{old}}$ はそれぞれ \vec{x}_i のピクセル j におけるノイズの分散、参照構造 \vec{v}^{old} のボクセル l における信号の分散であり、(6.13)式における P_j^N 、 P_j^F と等価である。 $\sigma_{ij}^{2\text{old}}$ 、 $\tau_l^{2\text{old}}$ は参照構造及び決定された配向に依存するため \vec{v}^{new} と同様に以下の式によって更新される。

$$\sigma_{ij}^{2\text{new}} = \frac{1}{2} \int_{\phi} \Gamma_{i\phi}^{\text{old}} \left| X_{ij} - \text{CTF}_{ij} \sum_{l=1}^L P_{jl}^{\phi} V_l^{\text{old}} \right|^2 d\phi \quad (6.15)$$

$$\tau_l^{2\text{new}} = \frac{1}{2} \left| V_l^{\text{old}} \right|^2 \quad (6.16)$$

(6.15)式右辺は観測値と推定された真値(期待値)の分散であり、(6.11)式より明らかにノイズの分散に等しい。(6.16)式も信号の分散をほぼそのまま代入する式である。粒子像の配向、参照構造の収束によってこれらの値も正しい値へ収束する。従って、反復計算終了時に (6.14)式は理想的な Wiener フィルターとして振る舞い、尤もらしい三次元構造を与える。

上記の説明は単一構造への三次元再構成であるが、複数構造を仮定して構造を分類しながら三次元再構成を行うことも可能である。この場合でも初期構造は1つで十分であり、分割された粒子像のサブセット間での構造の偏りを起点として反復計算の中で自然と分類されていく。この方法は大きなサブユニットの有無を弁別できるほか[Scheres, 2012a]、誤って粒子像として切り出された夾雑粒子の弁別にも用いられる。また、この複数構造への分類は2次元投影像に対しても適用可能であり、夾雑粒子の除去に大きな

効果を発揮する(図 6.9)。そのため、切り出された粒子像は三次元再構成前に 2 次元投影像の状態で分類にかけられ、夾雑粒子が除去される。

本節では分解能の評価方法についても説明する。RELION では分解能の評価に Fourier Shell Correlation (FSC) [Saxton & Baumeister, 1982; van Heel & Stoffler-Meilicke, 1985]を用いる。FSC は 2 つの構造 \vec{v}_1 、 \vec{v}_2 がどの程度の分解能まで構造が一致しているかを評価する指標であり、空間周波数 k に対して

$$\text{FSC}(k) = \frac{\sum_{l \in k} V_{1l} \cdot V_{2l}}{\sqrt{\sum_{l \in k} |V_{1l}|^2 \cdot \sum_{l \in k} |V_{2l}|^2}} \quad (6.17)$$

で定義される、空間周波数に依存して減少する関数である(図 6.10)。ある cryoEM 構造の分解能を評価する場合、粒子像を 2 つのデータセットに分けて三次元再構成を行い、得られた 2 構造の FSC を計算することで分解能を評価する。現在の主流としては、FSC = 0.143 まで値が減衰した時の空間周波数を有効分解能として用いている [Rosenthal & Henderson, 2003]。なお、RELION では三次元再構成最初の段階でデータセットを 2 つに分け、それぞれ独立に三次元再構成を行った上で最終的な三次元構造を計算する。この方法を用いることで再構成においてオーバーフィッティングによるアーティファクトがほぼ出現しないことが確認されており [Chen *et al.*, 2013]、この方法で得られた二つの構造から計算される FSC を gold-standard FSC (GS-FSC) と呼ぶ [Scheres & Chen, 2012]。以降の本研究での cryoEM 構造に対する分解能評価は GS-FSC に基づくものである。

6.3.5 高分解能構造の強調

一般に、単粒子解析法で得られた 3 次元構造は評価された分解能と比して高分解能の特徴が見えない。これは電子線の不完全なコヒーレンスや検出器感度などによる高空間周波数でのコントラスト喪失が原因である。これは構造振幅の空間周波数的な減衰として表れる。そのため、高空間周波数の構造振幅を引き上げるように補正することで高分解能構造を強調する (post processing) [Rosenthal & Henderson, 2003]。先行研究では高空間周波数での減衰をガウス関数として近似している。コントラスト損失が構造全体で一定であると仮定すると、空間周波数 k での平均構造振幅 $V(k)$ は

$$V(k) \exp\left[-\frac{B}{4} k^2\right] \quad (6.18)$$

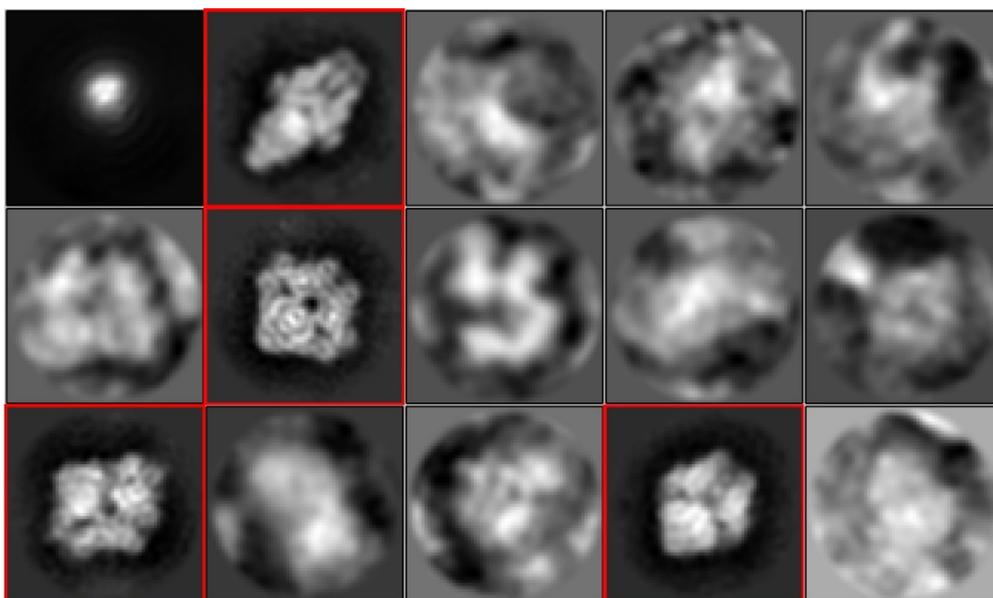


図 6.9 大腸菌由来 β -galactosidase 粒子像の 2 次元投影像分類結果。赤枠で囲まれた画像は β -galactosidase 投影像、それ以外は夾雑粒子か構造が壊れた β -galactosidase 粒子像であると考えられる。

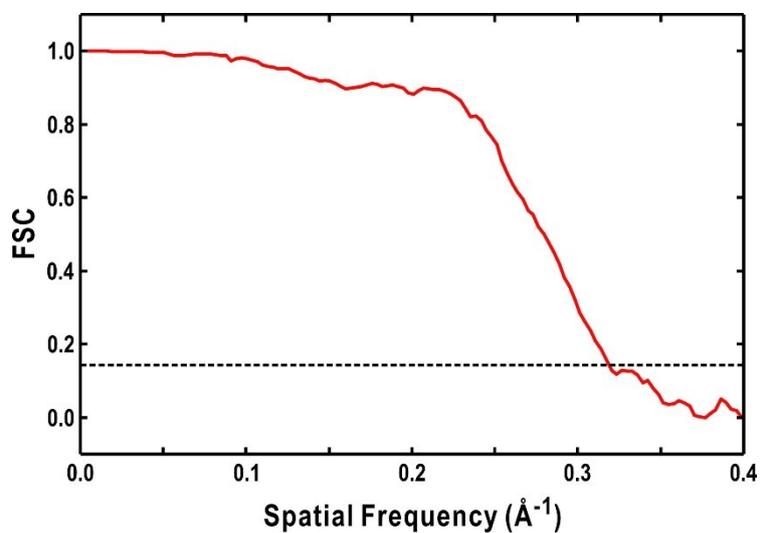


図 6.10 実際の 3 次元構造から得られた FSC 曲線。破線は $FSC=0.143$ であり、ここから分解能は 3.2 \AA と見積もられる。

と変調される。ここで、 B は損失の度合いを表すパラメータであり、結晶構造解析における温度因子に対応する。そのため、cryoEM においてもこのパラメータを結晶構造解析と同様に B-factor と呼称する。このガウス関数型の減衰を補正することで高分解能構造協調を行う。post processing 前後のポテンシャルマップを比較すると側鎖側鎖などの高分解能の構造が強調されることがわかる(図 6.11)。しかし、実際には構造揺らぎや非一様な試料損傷などによりコントラスト損失は構造全体で一定ではない。そのため、低分解能な領域は過剰な強調がなされることに注意する必要がある。

6.4 連続的構造変化の解析方法

6.3.1 節で述べた通り、単粒子解析法では試料蛋白質の構造が画一的であると仮定しているが、多くの蛋白質はその構造を連続的に変化させている。氷包埋された蛋白質分子はその一つ一つが構造変化の中のコンフォメーションに固定されると考えられるた

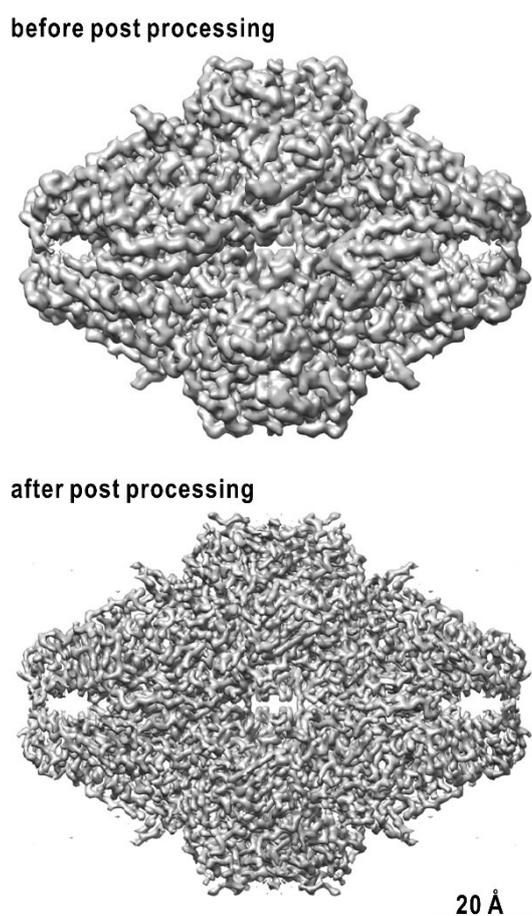


図 6.11 post processing による高分解能構造の強調。

め、cryoEM 観察像における蛋白質粒子像は構造多型を有している。2000 年代までこの構造多型は分解能向上を妨げる因子でしかなかったが、装置や解析手法の進歩によってより高分解能での構造解析が行われるようになった現在ではこの構造多型を可視化・解析する試みが活発になされている[Sorzano *et al.*, 2018]。本節ではこれらのうち代表的な手法と現状の問題点について簡単に説明する。

最初期の試みは 6.3.4 節で説明した構造分類に基づくアプローチである。蛋白質の構造多型はサブユニットの有無とは異なり、分子全体の中の一部の領域に生じる場合が多い。そのため通常の種類では分子全体の信号によって構造の差異が埋もれてしまい、弁別が困難であった。これを解決するべく 2015 年に提案されたのが Focused Classification と呼ばれる方法である[Bai *et al.*, 2015]。この方法では、まず通常通りに三次元再構成を行う。この時、構造多型が局所的であるため配向は正しく決定されていると仮定する。配向決定後、Particle Subtraction と呼ばれる処理によって興味ある部分(多型を示す領域)以外の構造を粒子像から差し引く(図 6.12)。この処理によって粒子像は興味ある部分のみになるため、これらに対して配向を固定して 3 次元構造分類を行うことで構造多型を可視化する。配向を固定して分類を行うのは画像全体に対して信号を持つ領域が僅かであり、プロジェクションマッチングがノイズに引きずられやすいためである。簡便かつ適用範囲の広い手法であるが、欠点は 3 次元構造分類そのものの不完全性である。人の目で見て明らかな程度に構造を分類できた例もあるが[Nakanishi *et al.*, 2018]、連続的に構造を変化させて作られたテスト粒子像について、適切に分類できなかったという結果も報告されている[Zhong *et al.*, 2019]。従って、ある程度の構造多型弁別は可能であるものの、完全な弁別は不可能な手法である。

Focused Classification をさらに一歩押し進めた方法が Multi-body Analysis と呼ばれる方法である[Schilbach *et al.*, 2017; Nakane *et al.*, 2018]。この方法では、蛋白質 3 次元構造を剛体として見なせる複数の領域に分割する。Particle Subtraction によって粒子像毎に各領域の投影像を取得し、各領域ごとに三次元再構成を行う。反復計算の最初に Particle Subtraction をやり直すことで、各領域の構造が改善されていく。また、1 枚の粒子像の中で領域ごとの配向が異なる場合、これは構造揺らぎを反映している。そのため、この配向のずれを解析することで分子の内部運動を可視化することが可能である[Nakane *et al.*, 2018]。しかしながら、Multi-body analysis は各領域の分子量が 200 kDa 以下になると配向決定が困難になることが報告されている [中根, 2018]。

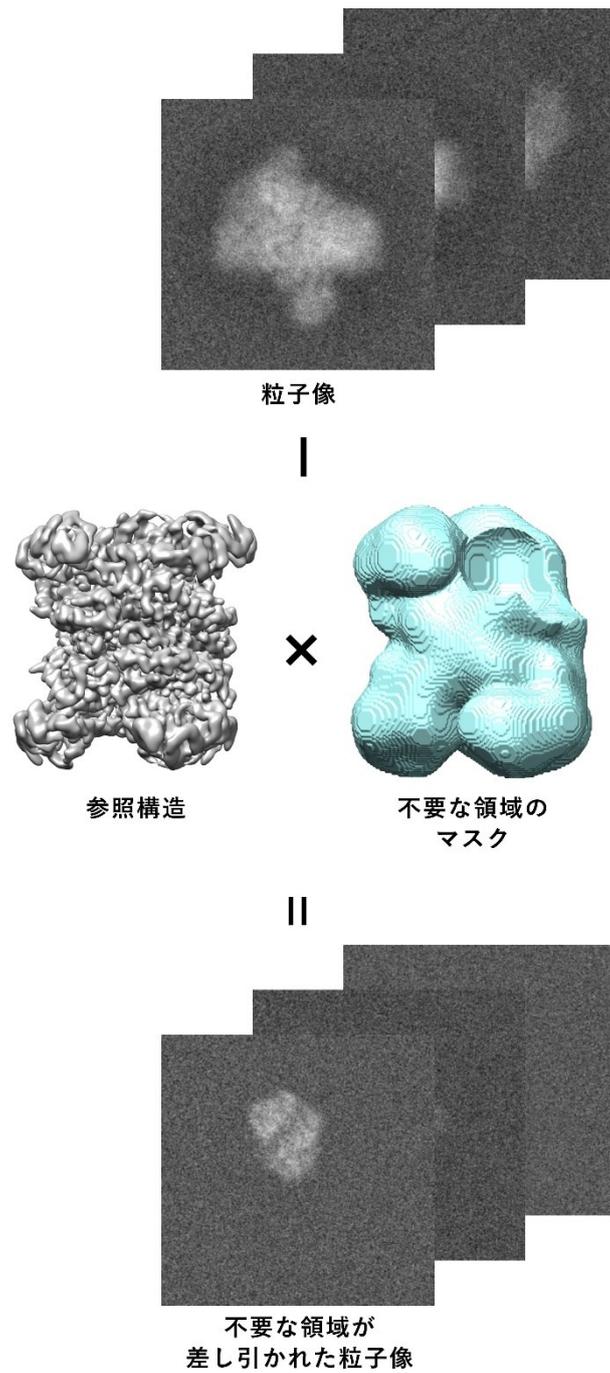


図 6.12 Particle Subtraction の概略図

上述の方法とは全く別の原理に基づくアプローチが多様体学習と呼ばれる方法である。cryoEM では蛋白質構造は 2 次元あるいは 3 次元の「画像」であり、ある 1 つの粒子像は各ボクセル、あるいはピクセルの値によって定義される。例えば $J \times J$ ピクセル

の2次元投影像は J^2 次元空間上の一点と見なせる。ある一つの配向について多数の粒子像を集めると、これらは各粒子像の構造の違いを反映して J^2 次元空間上で何らかの分布を形成する。一方、蛋白質の構造変化は一般に2-3の自由度で記述できるため、この粒子像の分布は J^2 次元上で2-3次元の超曲面(多様体)として分布していると考えられる。この J^2 次元空間上の多様体を低次元に射影し、構造の分布を可視化する方法が多様体学習である(図 6.13)。3節で紹介した主成分分析も多様体学習の一種である。多様体学習の大きな利点は試料蛋白質の構造変化について自由エネルギー地形を可視化できることである。熱平衡な溶液中の蛋白質1分子はカノニカルアンサンブルであると見なすことができる。この系における微視的状态とはコンフォメーションであり、あるコンフォメーション i を取る確率 ω_i は

$$\omega_i = \frac{1}{Z(\beta)} \exp[-\beta G_i], \quad \beta = \frac{1}{k_B T} \quad (6.19)$$

で与えられる。ここで、 k_B はボルツマン因子、 T は絶対温度、 $Z(\beta)$ は分配関数、 G_i はコンフォメーション i における自由エネルギーである。系の分配関数が不明な場合(6.18)式から自由エネルギーを直接求めることは不可能であるものの、自由エネルギー差を求めることは可能である。最頻コンフォメーションの出現確率を ω_0 とすると、コンフォメーション i の自由エネルギー差 ΔG_i は

$$\Delta G_i = G_i - G_0 = -k_B T \ln \frac{\omega_i}{\omega_0} \quad (6.20)$$

となる。この式について、コンフォメーションの出現確率 ω_i がcryoEM観察像における出現頻度と等価であるとする、多様体学習で得られた粒子像の分布を頻度分布に変換し、(6.19)式を適用することで自由エネルギー地形が得られる。

各構造の安定性や構造遷移における経路の定量的な議論を可能にするため、蛋白質動態解析において自由エネルギー地形が得られることは極めて重要である。現在提案されている方法としては配向ごとに自由エネルギー地形を計算し、1つの地形に統合する方法[Dashti et al., 2014]や、三次元構造分類で数百個の3次元構造を得たうえでこれらについて多様体学習を行う方法[Haselbach et al., 2018]が存在する。しかし、前者の方法はS/N比の低い生の投影像同士を比較するため、適用できる対象がS/N比が良いサイズの大きい試料に限定される。また、後者についても、三次元構造分類自体が不完全なため、計算結果が信頼できるのはサイズが大きくある程度分類がうまくいく試料に限られてくる。

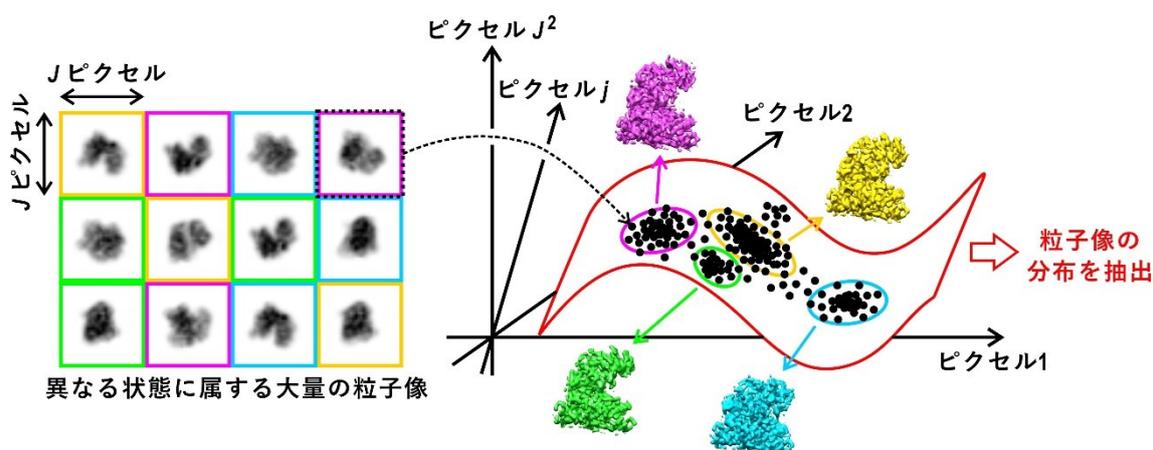


図 6.13 多様体学習の概略図

以上が現在提案されている方法の中でも比較的うまくいくことが確認されている方法であるが、問題点がいくつか存在する。まず Focused Classification においては上述した通り完全な分類を行うことは実質的に不可能である。また、単に構造を分類するため、構造揺らぎや自由エネルギー地形の可視化につなげることは困難である。Multi-body analysis や多様体学習は Focused Classification の欠点を克服しているものの、現在提案されている手法は全て分子量の大きい試料に適用が制限される。実際、これらの手法で解析ができた試料は小さい試料でも 1.4 MDa [Schilbach *et al.*, 2018] から 2.5 MDa [Nakane *et al.*, 2018]、大きい試料では 5 MDa を超える [Dashti *et al.*, 2014; Nakane *et al.*, 2018]。従って、本研究で扱ってきたような数百 kDa 程度の分子量の蛋白質について動態解析を行うためには新たな手法を開発しなければならない。そこで、我々は分子動力学(Molecular Dynamics: MD)計算を組み合わせた新規的な蛋白質動態解析手法を開発し、既存手法の適用が困難な分子量 280 kDa の試料蛋白質 GDH についてドメイン開閉運動の解析を行った。これらの詳細を次の 7 章で述べる。

7章 cryoEM による GDH ドメイン運動解析

前章末に述べた通り、現状の手法では巨大な超分子複合体でない限り、cryoEM 観察像から連続的な構造変化を解析することは困難である。本研究では超高度好熱菌由来グルタミン酸脱水素酵素の cryoEM 構造解析を通じて、MD 計算と組み合わせた新規的な蛋白質構造変化解析手法を開発し、既存手法では不可能な GDH ドメイン運動の自由エネルギー地形可視化に成功した。本章ではこれらの詳細について述べる。

7.1 超高度好熱菌由来グルタミン酸脱水素酵素

超高度好熱菌由来グルタミン酸脱水素酵素(GDH)は総分子量 280 kDa のホモ六量体蛋白質である(図 7.1a)。GDH は補酵素であるニコチンアミドアデニンジヌクレオチドリノ酸(NADP)と共にグルタミン酸を脱水素反応によって α -ケトグルタル酸へ変換する機能を持ち、尿素合成における必須酵素である。各サブユニットはヌクレオチド結合サイトを持つ NAD ドメインと六量体形成部位である core ドメインから構成され、ドメイン間はグルタミン酸結合サイトを有する活性クレフトとなっている(図 7.1b)。リガンド非結合結晶構造の各サブユニットを比較すると NAD ドメインの core ドメインに対する相対位置が異なっており(図 7.1b)、また SAXS プロファイルは結晶構造から計算されたプロファイルと異なるものであった [Nakasako *et al.*, 2001]。これらの結果から GDH は溶液中で自発的なドメイン開閉運動を行っていると考えられている。実際、GDH は結晶構造での構造変化と同様のドメイン開閉運動を行っており、またこの運動は活性クレフト内部の局所的な水和構造変化と協奏的に生じていることが MD 計算の結果から報告されている [Oroguchi & Nakasako, 2016]。

GDH は本研究における解析対象として複数の利点を備えている。まず、マルチドメイン蛋白質であり、構造変化が剛体的なドメイン間の相対変位として生じるため、解析が比較的容易である。3つ以上の剛体的領域から成る蛋白質では運動が複雑化するため、2つの機能ドメインのみから構成されている、という点も重要である。第二に GDH は先行研究が豊富であり、構造変化の様式が既知である他、結晶構造や SAXS プロファイル、MD 計算など cryoEM との相補的な比較が可能なデータが多く存在する。解析によって得られた動態をこれらのデータと比較することで妥当性などの評価が可能となる。

以上を踏まえ、本研究ではリガンド非結合 GDH の cryoEM 観察を行い、既存手法及び開発した新規手法を用いてドメイン開閉運動の解析を行った。

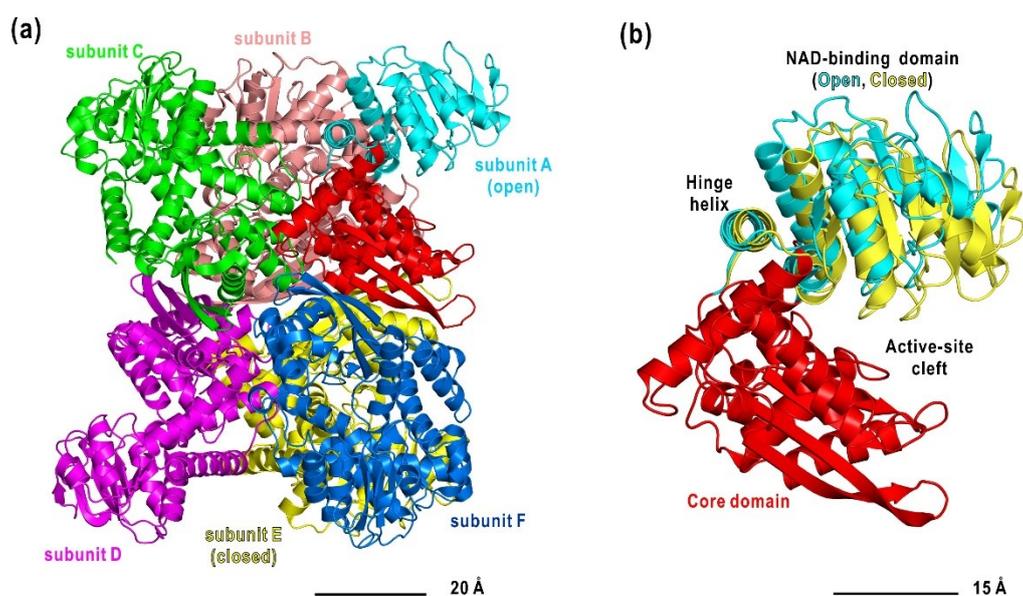


図 7.1 (a)GDH 六量体結晶構造(PDB ID: 1EUZ)。 (b) 開構造であるサブユニット A と閉構造であるサブユニット E の core ドメインによる重ね合わせ。

7.2 GDH の cryoEM 構造解析

7.2.1 試料作成と cryoEM 観察

GDH の発現・精製は先行研究で確立されているプロトコルに従って行った [Nakasako *et al.*, 2001; Oroguchi & Nakasako, 2016]。精製された GDH は 10 mg/mL 濃度で 5 mM Tris-HCl 緩衝液(pH 7.5)に溶解させ、試料作成に用いた。cryoEM 観察試料作成には通常 0.5-1.0 mg/mL 程度の試料溶液が用いられ、今回の GDH 溶液試料は極端に濃度が高い。これは、他の試料と比較して GDH の溶液濃度に対する cryoEM 観察時の視野内数密度が低いためである。この原因は恐らく試料作成時の余剰溶液除去において溶液を吸わせる濾紙と GDH の親和性が高いためと考えられている。この問題に対しては試料濃度を上げる程度しか対応策が存在しないため、上記濃度で調製している。

凍結試料の作成、及び cryoEM 観察は大阪大学大学院生命機能研究科難波研究室(当時)にて行われた。観察用グリッドには Mo R1.2/1.3 200mesh holey carbon grid (Quantifoil, Großlobichau, Germany)を使用した。これはモリブデン製グリッドに 1.2 μm 径の孔が 1.3 μm 間隔で空けられたアモルファスカーボン膜を載せたものである。凍結試料の作成には Vitrobot (Thermo Fisher, Waltham, MA) (図 7.2a)を使用した。Vitrobot は余剰溶液の除

去及び液体エタンへの突入による急速凍結を自動的に行う装置である。以下に凍結試料作成手順を簡単に記す。まず親水化処理済グリッドを保持したピンセットを Vitorobot 内部にセットする。Vitorobot 内部は温湿度が制御可能であり本実験では 4°C、湿度 98% に設定した。次に Vitorobot 側面に設けられた小窓から試料溶液 2-3 μL をグリッドに滴下する。滴下後、速やかに装置を操作することで濾紙が取り付けられたロボットアームによる余剰溶液除去と液体エタンへの突入による急速凍結が自動で行われる。濾紙をグリッドへ押し付ける時間は任意に設定でき、良好な膜厚の試料を得るためには試料溶液の粘性等を考慮してこの時間を調整しなければならない。本実験では 3 秒あるいは 5 秒間に設定し、その後の cryoEM 上でのスクリーニングを経て良好な試料について観察を行った。

クライオ電子顕微鏡は難波研究室の JEM-Z200FSC (CRYO ARM; JEOL, Tokyo, Japan) を使用した(図 7.2b)。また、電子顕微鏡の調整及び後述の自動測定は難波研究室の加藤貴之特任准教授(当時)の指導・支援の下で行った。CRYO ARM は加速電圧 200 kV、球面収差係数 1.4 mm の電子顕微鏡である。(6.4)式より、相互作用係数は加速電圧に反比例

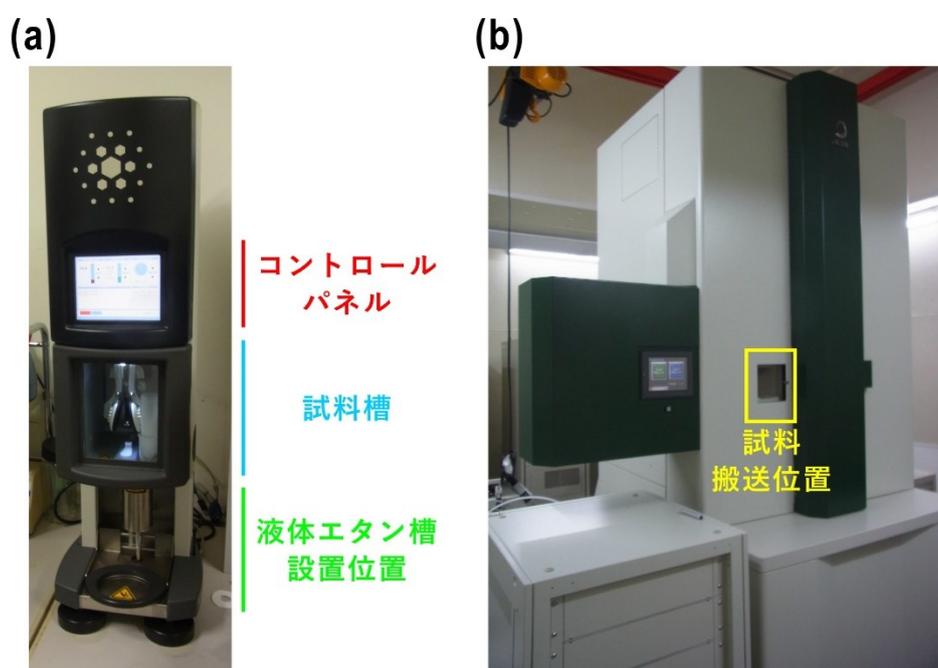


図 7.2 (a) Vitorobot 装置写真。温湿度などのパラメータは上部コントロールパネルで操作する。試料槽中心部の器具はロボットアームであり、ハの字状の部分に濾紙を取り付ける。(b) CRYO ARM 200 装置写真。黄枠の部分から凍結資料を搬送する。

するため、GDH 程度のサイズの分子では加速電圧が小さめの方が良好なコントラストが得られやすい。CRYO ARM には Ω -フィルターと呼ばれる電子線のエネルギーフィルターが存在する(図 7.3)。これは名前の通り Ω 型に電磁石が配置されており、このフィルターを通過した電子線は波長によって通過後の出射位置が変化するため、直後にスリットを設置することで波長依存的に電子線をフィルタリングできる。CRYO ARM 200 におけるスリット幅は 10 eV に相当する。これによって非弾性散乱電子線を除去することが可能であり、回折波の振幅が変化しコントラストが向上する。また、観察像の強度についてスリットの有無による比較を行うことで非弾性散乱電子の割合、即ち試料の厚みを評価することができる。経験則として、200 kV 電子線の場合良好な厚みの試料では 20-30% 程度の変化量に収まる。この指標によって凍結試料のスクリーニングを行った後、自動測定ソフトウェア JADAS(JEOL)によって自動観察を行った。JADAS は試料グリッドについて適当な領域を指定することで、自動的に指定領域内の孔を検出し撮影を行う。撮影条件は倍率 56,179 倍、照射電子線量 $8.15 \text{ e}^-/\text{\AA}^2/\text{sec}$ 、露光時間 10 秒であった。観察像は電子線検出器 K2-summit (Thermo Fisher, Waltham, MA)によって 50 フレームに分割して記録された。なお、上記倍率での観察像のピクセルサイズは 0.87 \AA であった。

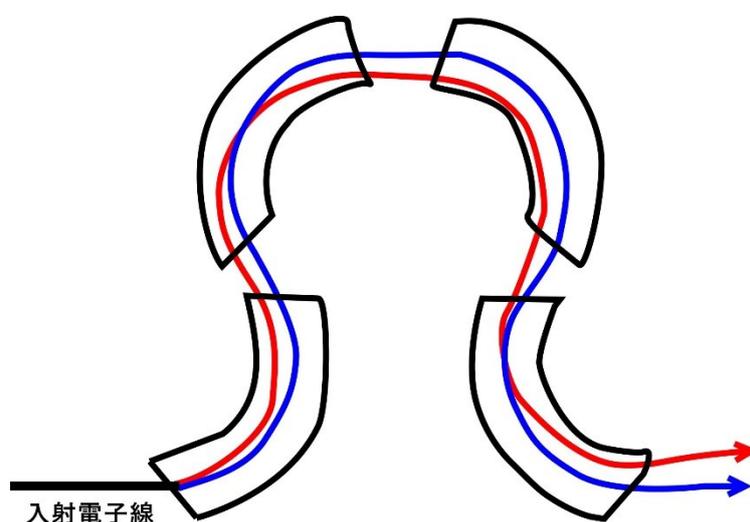


図 7.3 Ω -フィルター模式図。青色の経路の電子線と比較して赤色の経路の電子線は波長が短い。フィルター出射部にスリットを設けることで電子線の波長を選択できる。

7.2.2 三次元再構成までの観察像処理

JADAS によって自動測定された観察像から試料液膜が張っていない、非晶質氷が再結晶化している、エタン氷や霜の巨大な粒子が写り込んでいるなどの解析に適さない観察像を目視で除去し、最終的に 1784 枚の観察像を取得した。motion correction 及び dose fractionation を MotionCor2 [Zheng *et al.*, 2018]で処理し、Gctf[Zhang *et al.*, 2016]によって CTF パラメータの推定を行った (図 7.4)。この時点で各観察像の Thon ring を確認し、パラメータが誤っている観察像やデフォーカスが浅すぎる、あるいは深すぎる観察像を除外した。この処理後の観察像枚数は 1473 枚であり、デフォーカスの範囲はおよそ 0.5-2.3 μm に収まった。

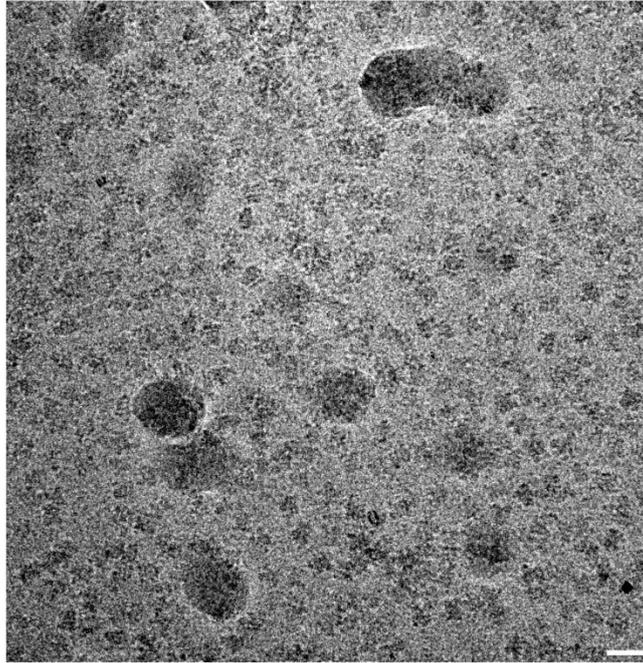
次に GDH 粒子像の切り出しであるが、これにはテンプレートマッチングと呼ばれる方法を使用した [Scheres, 2015]。テンプレートマッチングではまず目視で 1000 個程度の粒子像を選択し、2次元投影像分類によって S/N 比のよい平均投影像を得る。この中から異なる配向の粒子像を 3 個ほど選出し、これをテンプレートとして観察像内の相関の高い部分を粒子像として自動で切り出す。ただし、オーバーフィッティングを避けるためテンプレートに 20 \AA 程度のローパスフィルターをかける必要がある。テンプレートマッチングによって切り出された 339,009 個の粒子像を投影像分類にかけることで夾雑分子を排除した。分類後に残った粒子像は 207,707 個であり、平均投影像は 6 \AA 程度の分解能であった(図 7.5)。一般によりデータとされる条件は「平均投影像中に二次構造が視認できること」とされている。図 7.5 ではヘリックスと思しき構造などが見えており、比較的質の良好なデータセットであったと言える。この 207,707 個の粒子像から三次元再構成を行った。

7.2.3 GDH 三次元再構成

まず、7.2.2 節で得られた平均投影像から e2initialmodel によって三次元初期構造を作成した。e2initialmodel は単粒子解析用プログラムパッケージ EMAN2 [Tang *et al.*, 2007] に組み込まれているプログラムであり、複数の異なる配向からの平均投影像からコモンライン法[Crowther *et al.*, 1970]によって三次元再構成を行う。投影断面定理より、投影像の Fourier 変換は三次元 Fourier 変換の原点を通るスライスに等しいため、異なる配向である 1 組の投影像は逆空間上で一本の線分を共有する(コモンライン)。コモンライン法はこれに基づき、投影像間のコモンラインから各投影像の配向を決定する方法である。

得られた e2initialmodel で得られた構造を参照構造としてまず三次元構造分類を行い、

(a)



(b)

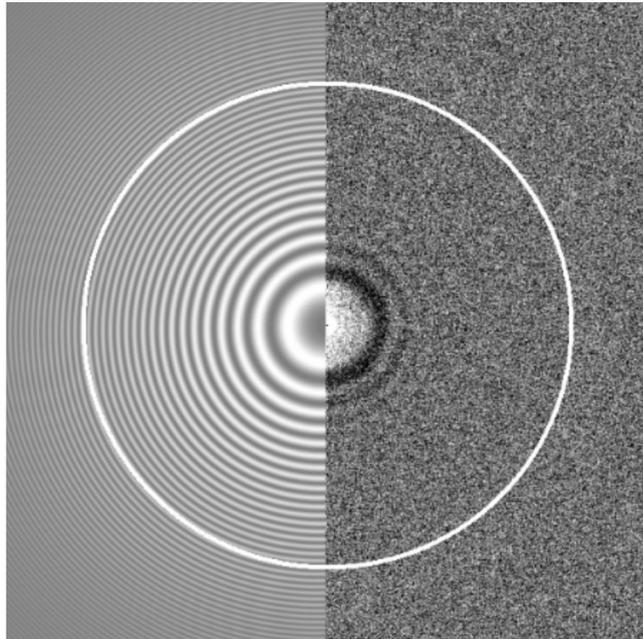


図 7.4 (a) motion correction、dose fractionation 処理を施した GDH cryoEM 観察像。デフォーカスは $1.3\ \mu\text{m}$ 、右下のスケールバーは $200\ \text{\AA}$ である。
(b) (a)の Thon ring (右)と推定されたパラメータから計算された CTF 曲線(右)。白い円は Gctf が評価したこの観察像の有効分解能であり、 $4.2\ \text{\AA}$ 相当の空間周波数である。

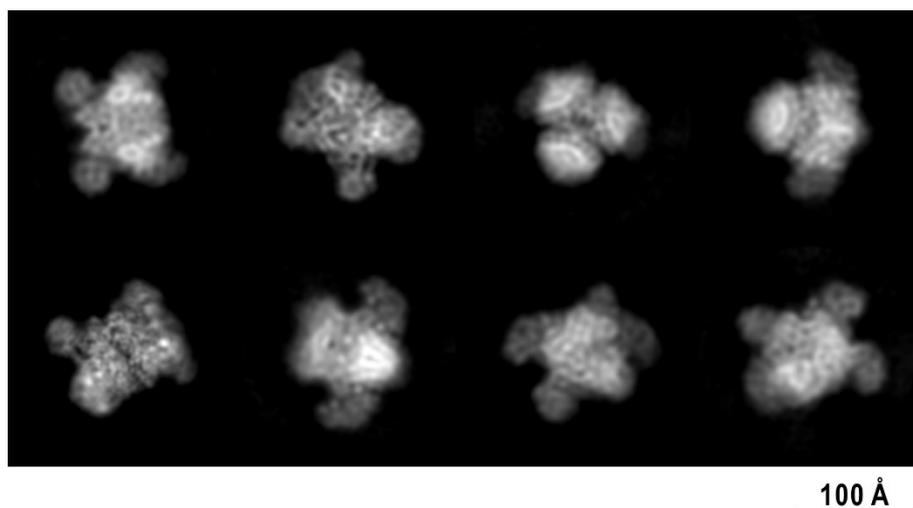


図 7.5 GDH 二次元平均等映像。例えば左下の投影像では紐のような構造が見えるがこれはヘリックス構造と推定できる。

周囲に粒子が存在するなど解析に適さない三次元構造を除去した。分類後に残った 169,240 個の粒子像を用いて三次元再構成を行った。なお、試料蛋白質に回転対称性が存在する場合、試料粒子像について探索すべき配向が制限され、また配向が制限されることで実質的に枚数が増え分解能が向上する。そのため、三次元再構成は GDH が持つ D3 回転対称性の拘束条件下で行った。得られた三次元構造は GS-FSC で 3.4 Å 分解能であった(図 7.6)。

粒子像の配向分布を確認したところ、三つの配向に粒子像が集中する非常に偏った分布となった(図 7.7a)。これは preferred orientation と呼ばれる現象であり、薄膜状溶液中の蛋白質分子が急速凍結されるまでのわずかな時間にブラウン運動によって気液界面へ到達することで生じると考えられている。また、多くの場合 preferred orientation にある粒子は気液界面に接触した部分、あるいは全体の構造がアンフォールドによって崩壊する [D'Imprima *et al.*, 2019]。GDH の場合、全体構造は崩壊していないものの、気液界面に接触したサブユニットは崩壊している可能性が高い。そこで、全粒子像のサブユニットのうちどれだけの割合が気液界面に接触しているか見積もりを行った。まず、粒子像が集中した各配向(Orientation 1-3)について割り当てられた粒子像の枚数を算出し全粒子像に対する割合を求めたところ、Orientation 1 から順に 2%、38.5%、10.6%であった(図 7.7a)。次に、各配向において GDH がどのように気液界面と接触し得るかを検討した(図 7.7b)。その結果、各配向で 2-4 個の NAD ドメインが接触し得ると判断した。こ

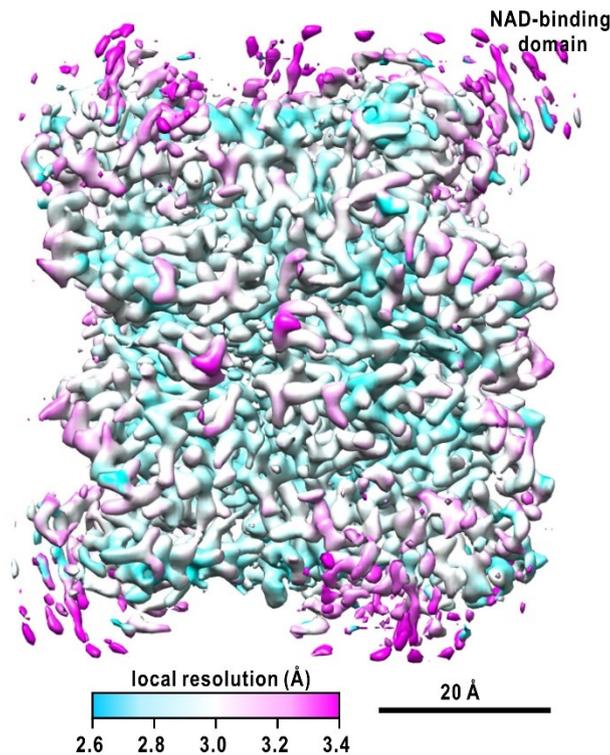


図 7.6 cryoEM 観察像から再構成された GDH 三次元構造。6.3.5 節で説明した処理により高分解能構造を強調して表示している。

こから、粒子像の中の全サブユニットの内最大 30%程度が NAD ドメインで気液界面に接触していると見積もられた。これら 30%の NAD ドメインは本来の構造を損なっており、一方で残りの 70%のサブユニットはバルク溶媒中でその構造を保っていたと考えられる。

局所的な分解能をプログラム blocres [Cardone et al., 2013]によって評価したところ、core ドメインでは 3.0-2.6 Å 程度であった。実際、ポテンシャルマップには二次構造やアミノ酸側鎖が良好に確認でき、結晶構造ともよく一致した(図 7.8a)。また、活性クレフト中には小分子程度のサイズの孤立したポテンシャルマップが確認された。これは試料溶液の組成を考慮すると tris 分子であると考えられる(図 7.9)。一方、NAD ドメインは局所分解能が 3.2 Å 以下であると見積もられ、ポテンシャルマップも結晶構造との重ね合わせが困難な特徴に乏しいものであった(図 7.8)。このドメイン間での分解能の差異はドメイン運動を反映したものであると考えられる。また、ポテンシャルマップから Fourier 変換によって SAXS プロファイルを計算し、Guinier 解析によって慣性半径を評

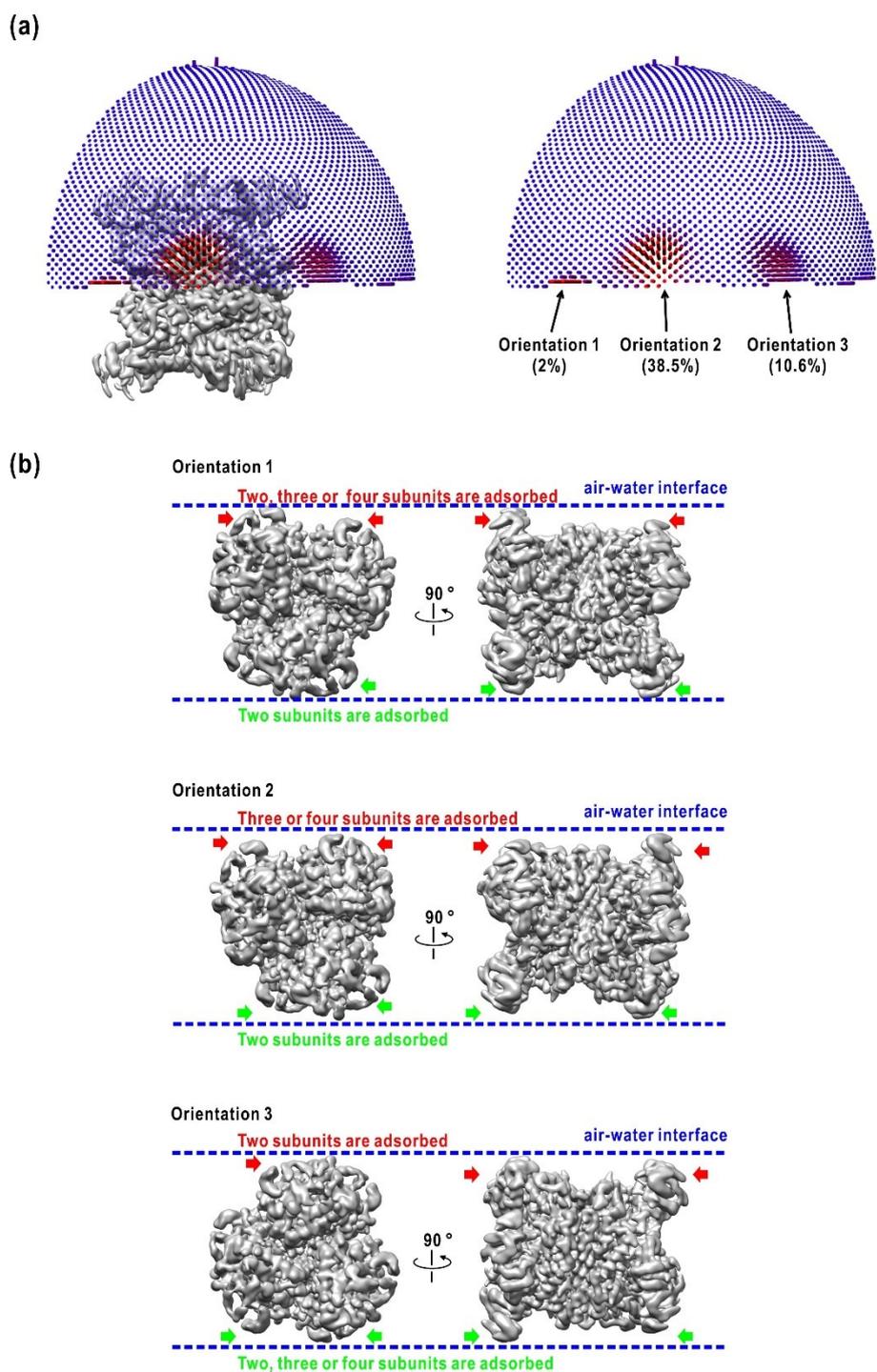


図 7.7 (a)GDH cryoEM 粒子像の配向分布。割り振られた粒子像の数が
多い配向ほど赤く着色されている。また、D3 回転対称性による拘束を
加えているため、立体角が制限されている。(b)各 preferred orientation
における GDH の気液界面との接触評価。青の破線が気液界面に対応
しており、矢印で示された部分が接触する領域である。

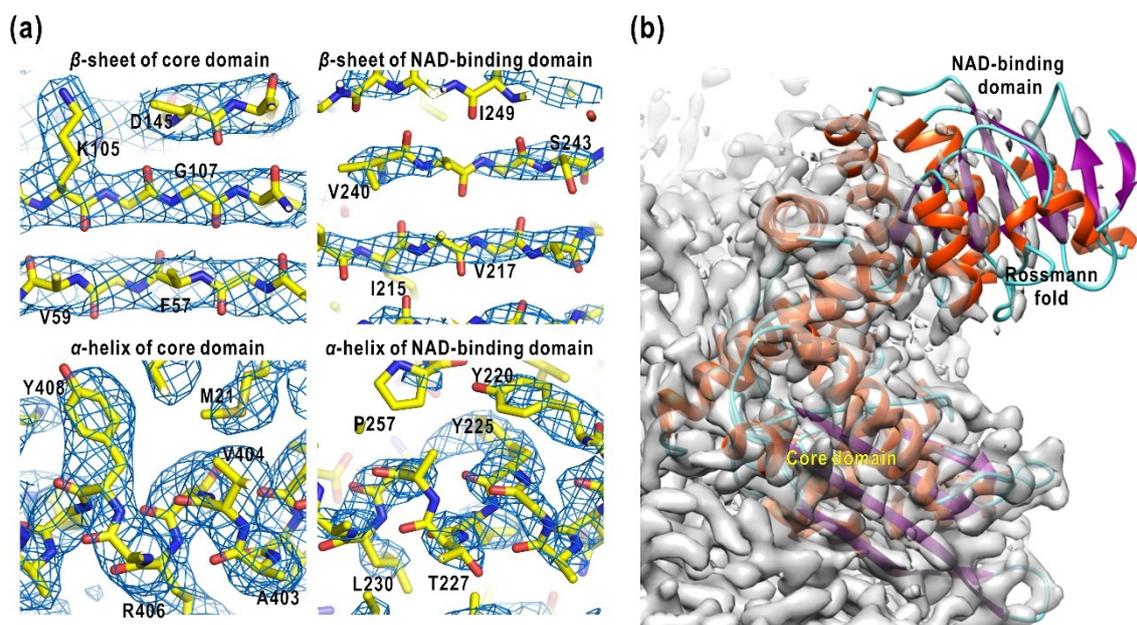


図 7.8 GDH cryoEM 構造と結晶構造との比較。(a) 結晶構造と各ドメインの二次構造の比較。左の列が core ドメイン、右の列が NAD ドメインの二次構造である。アミノ酸は一文字表記で記している。(b) サブユニット全体での結晶構造との比較。cryoEM 構造では NAD ドメインに相当する領域を欠損している。

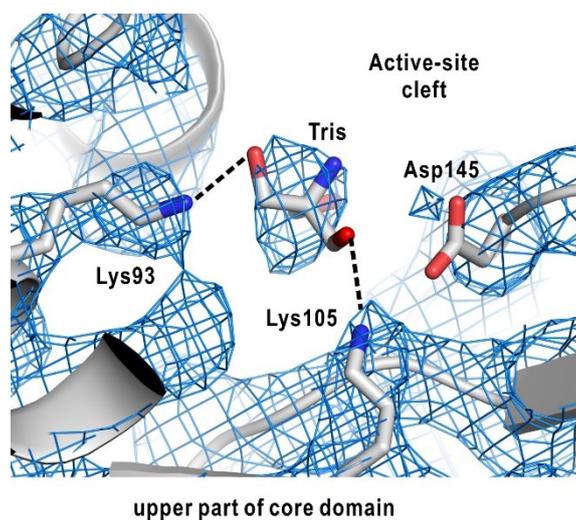


図 7.9 活性クレフト中の孤立したポテンシャルマップ。アミノ酸は 3 文字表記で記されている。水素結合が生じていると思われる極性原子間を破線で繋いでいる。

価したところ 42.6 Å となり、実験値である 43.9 Å [Nakasako *et al.*, 2001]に近い値が得られた。SAXS プロファイルは溶液中での構造揺らぎを反映するため、cryoEM から得られた GDH ポテンシャルマップは溶液中での構造揺らぎのアンサンブル平均と見なせる。

7.2.4 Focused Classification による NAD ドメインコンフォメーションの分類

7.2.3 節の結果より、cryoEM で得られた GDH 粒子像は NAD ドメインがドメイン開閉運動由来と思われる構造多型を示した。そこでこの構造多型を解析するべく、GDH 全体構造のうち、ある一つの NAD ドメインのみに着目した Focused Classification を行った(図 7.10)。また、Focused Classification を行うにあたって全粒子像に対して Symmetry Expansion と呼ばれる処理を行った[Scheres, 2016]。回転対称性による拘束条件下では一枚の粒子像に存在する全サブユニットを平均化しているため、C1(対称性なし)として配向を振り分け直さなければ分類ができない。一方で、GDH は MD 計算および酵素活性測定の結果から各サブユニットは独立に運動・機能発現することが報告されており [Oroguchi & Nakasako, 2016]、C1 として配向を振り分けて 1 つの NAD ドメインのみに着目すると残りの 5 つの NAD ドメインの構造揺らぎを捨てることになってしまう。Symmetry Expansion では分子の回転対称性に従って各粒子像の配向を振り分けなおすのではなく対応する配向に粒子像をコピーする。この操作を施すことで全ての NAD ドメインを母集団として分類することが可能となる。また、前章でも述べた分類の不確かさを評価するため、初期乱数のみを変えた Focused Classification を 10 回行い(1st-10th

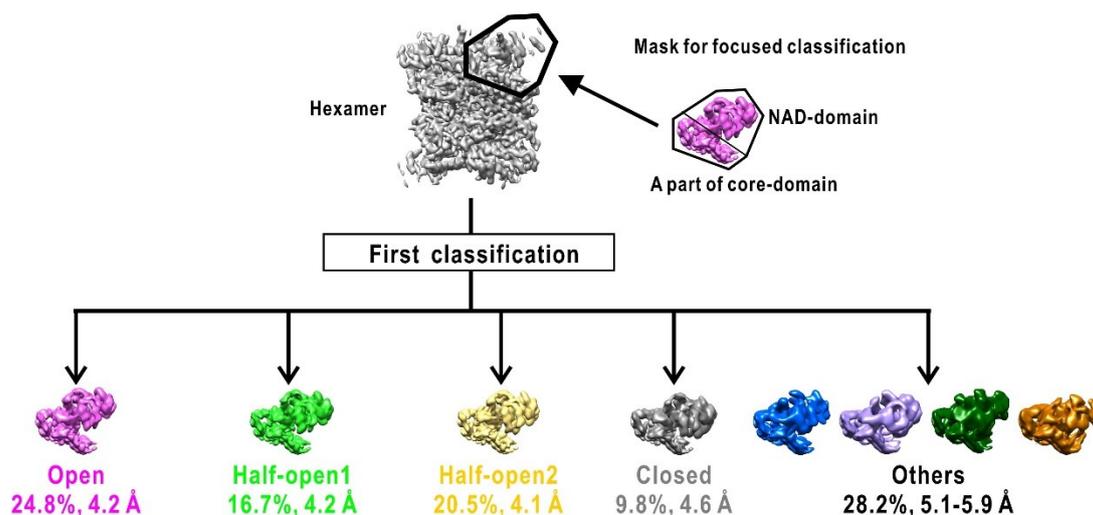


図 7.10 本研究で実施した Focused Classification の概略図。各クラスに記載された数字は粒子像の割合/分解能である。

Classification)、その結果の比較を試みた。

Symmetry Expansion 後の粒子像を Focused Classification によって 8 クラスに分類したところ、どの Classification においても 4-5 個のマップが NAD ドメインとして解釈可能であり(図 7.11、付録 B1, B2)、残りは NAD ドメインの形状が崩れた低分解能なマップであった。また、解釈可能なマップに分類された粒子像と低分解能マップに分類された粒子像の割合は全ての Classification で 7:3 程度であった(付録 B1)。低分解能マップの 3 割という数字は気液界面に接触しているであろうサブユニットの割合と一致しており、従ってこれらのマップは気液界面に接触して構造が壊れてしまったサブユニットであると解釈できる。以降は解釈可能なマップについてのみ議論する。

各 Classification 内でマップを比較すると、マップ間で局所分解能と共に活性クレフトの開き方に差異が見られた(図 7.11、付録 B2)。また、Classification 間でも得られた分類結果は互いに異なるものであった。例えば、1st Classification では開構造(open)、2種類の半開構造(half-open1, 2)、閉構造(closed)の 4 つの構造に分類された(図 7.11)。なお、half-open1 と half-open2 は図 7.11 では差異がほとんどないが、紙面に垂直な方向に NAD ドメイン位置が異なっている。一方で、他の Classification では open と half-open の中間構造(open/half-open)や half-open と closed の中間構造(half-open/closed)といった別の構造が現れた。これら結果の不一致は NAD ドメインの構造変化が連続的であることや、Focused Classification そのものが不完全であることが原因であると思われる。この結果を踏まえ、我々は Focused Classification で得られた構造が複数コンフォメーションの結合としてモデル化し、解析を試みた。次節でその詳細を述べる。

7.3 新規解析手法による GDH ドメイン運動解析

7.3.1 解析のための仮定と MD 計算の利用

前節の結果より Focused Classification を用いても、得られる構造は明らかに複数のコンフォメーションが混じったものであった。仮に一つの構造について構成する粒子像をコンフォメーションごとに選別できたならば、これらをコンフォメーションごとに三次元再構成した上で足し合わせた構造は選別前の構造と等しいと考えられるため、

$$V_l = \sum_{k=1}^K \omega_k V_{kl}, \quad \omega_k = \frac{N_k}{N_0} \quad (7.1)$$

なる式が近似的に成り立つと仮定する。即ち、cryoEM マップ \vec{v} のボクセル l での値 V_l が、それぞれが単一のコンフォメーション k である K 個のマップ \vec{v}_k のボクセル l での値

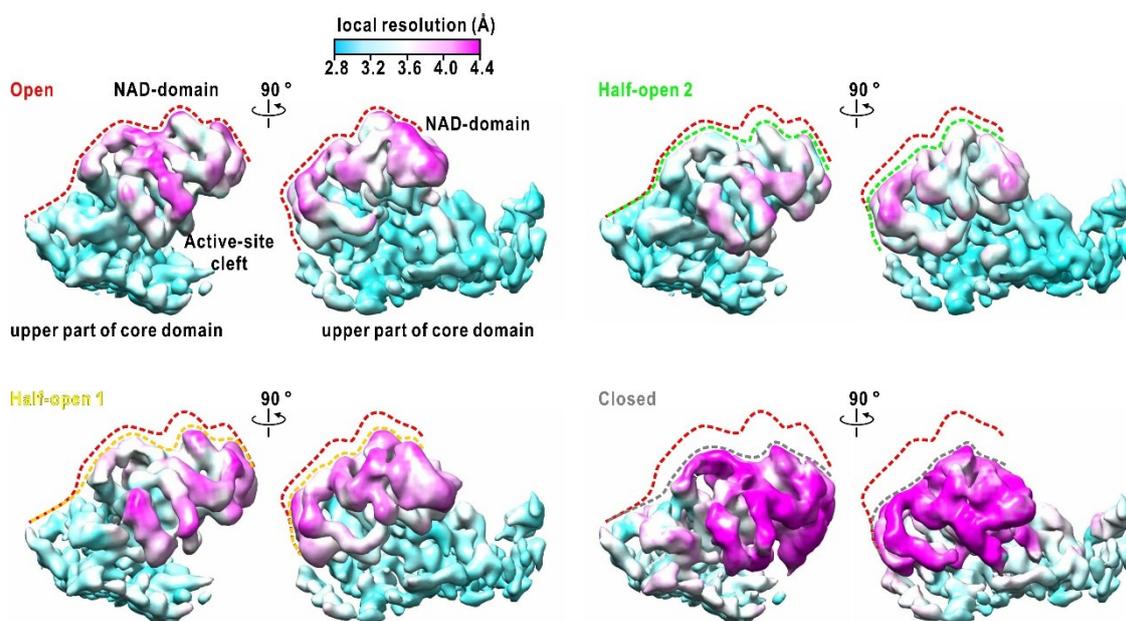


図 7.11 1st Classification で得られた NAD ドメインの複数のコンフォメーション。破線は各コンフォメーションの輪郭であり、色によって対応している。

V_{kl} ($k=1-K$)の線形結合で近似できるとする。 N_0 は \vec{v} を構成する粒子像の枚数、 N_k は \vec{v} を構成する粒子像のうちコンフォメーション k である粒子像の枚数である。この近似が実際に成立しうるかどうかは後ほど議論する。(7.1)式が成り立つとすれば、cryoEM 構造 \vec{v} から各コンフォメーションの重み ω_k が決定できる。この重みは各コンフォメーションの出現頻度に等しいため、さらにコンフォメーションの分布があれば両者を組み合わせて自由エネルギー地形を構築することができる。問題点は \vec{v}_k およびコンフォメーションの分布をどのように用意するかであるが、本研究では MD 計算を利用することで解決した。

7.1 節で述べた通り、GDH は MD 計算による先行研究が行われており、200 ns にわたるリガンド非結合状態での構造揺らぎのトラジェクトリーが得られている [Oroguchi & Nakasako, 2016]。結晶表面の溶媒露出 NAD ドメインの AFM による変位観察との比較より、この MD トラジェクトリーでは GDH の可能なコンフォメーションを十分にサンプリングしていることが確認されている [Oroguchi & Nakasako, 2016]。ただし、この報告はあくまで「GDH が構造変化しうる範囲をカバーしている」ことを保証するだけであり、実際の構造出現頻度を再現しているかどうかは不明である。この MD 構造のアンサ

ンブルについて PCA で構造揺らぎを解析したところ、2つの主成分で運動が説明できた(図 7.12a)。第一主成分はドメイン開閉運動であり、第二主成分はクレフトの開き方を保ったまま互いにねじれるように運動する剪断運動であった。また、第一、第二主成分への射影によって各構造をプロットすると、原点(平均構造)を最頻構造として広がるような分布であった(図 7.12b)。我々はこの MD 構造プロットから 28 個の参照構造を抽出した。これらの構造は GDH の取りうるコンフォメーションをスパースに尽くしているため、 \bar{v}_k に至適な構造である。また、図 7.12b のプロットはコンフォメーションの分布であり、サンプリングされた構造の点上に cryoEM 構造に対する重みをプロットすることで自由エネルギー地形を構築することが可能である。ただし、MD 構造は原子モデルであるため、マップに変換しなければならない。次節でマップへの変換方法について簡単に説明する。

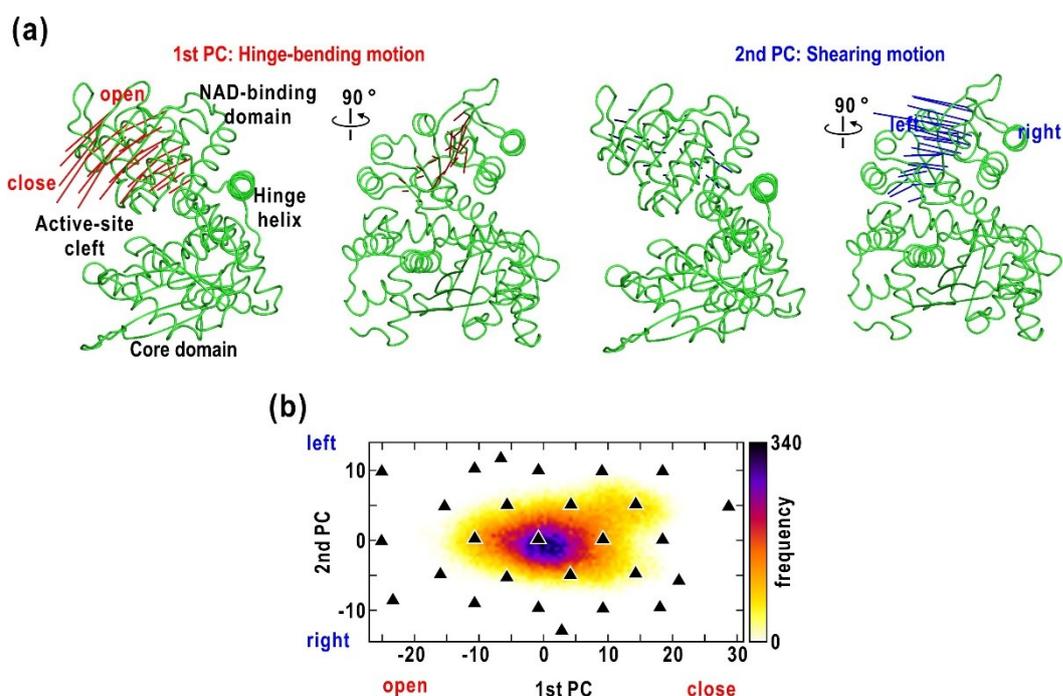


図 7.12 (a) PCA で得られた GDH MD トラジェクトリーにおける第一、第二主成分方向のドメイン運動。赤色、青色の線はそれぞれ第一、第二主成分の固有ベクトルである。(b) GDH MD トラジェクトリーの第一、第二主成分へのプロット。6 サブユニットすべてについてプロットしている。また、黒色の三角点は本研究で参照構造として用いた MD 構造である。

7.3.2 MD 構造のマップへの変換

サンプリングされた 28 個の MD 構造についてまず原子モデルを cryoEM 構造の core ドメインについて重ね合わせた。core ドメインは cryoEM 構造においても結晶構造とほぼ変わらない構造を有しており、またドメイン運動の基準となるドメインである。そのため、重ね合わせの基準として適している。MD 構造からマップへの変換は分子ビューアソフトである UCSF Chimera [Pettersen *et al.*, 2004] に内蔵されている molmap という機能を用いて行った。MD 構造から得られたマップを以下 MD マップと呼称する。

6.3 節で述べた通り、cryoEM 構造は理想的なポテンシャルマップと比較して高空間周波数側で構造振幅が減衰している。この影響を除くため、MD マップに減衰関数 $D(k)$ を逆空間上で乗じることで cryoEM 構造と同程度の減衰となるよう処理した。減衰関数 $D(k)$ は以下の式で決定した。

$$D(k) = \frac{\int |V_{EM}(k)| d\Omega_k}{\int |V_{Model}(k)| d\Omega_k} \quad (7.2)$$

ただし、 V_{EM} は GDH 六量体 cryoEM 構造の Fourier 変換、 V_{MD} は GDH 六量体結晶構造の Fourier 変換である。

7.3.3 cryoEM 構造と MD 構造の比較

まず、cryoEM 構造が MD 構造の重ね合わせで表現することができるかを検討するため、1st Classification の cryoEM 構造と MD マップの類似度を計算し、主成分空間上にプロットした。もしも cryoEM 構造が真に単一の構造であるならば、プロットされる類似度は単峰性の地形となるはずである。また、cryoEM 構造がノイズやコントラストの悪さで近似不可能であれば類似度の地形は起伏のない地形になると考えられる。類似度は

$$\exp \left[- \sum_i \frac{|V_i^{MD} - V_i^{EM}|^2}{2\sigma_i^{EM^2}} \right] \quad (7.3)$$

によって定義される量である。これは RELION でプロジェクションマッチングに使用されている指標を三次元に拡張したものである。cryoEM 構造に加わるノイズは理想的には検出器由来のガウス分布に従う白色雑音であるため、ガウス関数が用いられている。

各 cryoEM 構造について得られた主成分空間への類似度プロットは単峰性の分布でも平坦な分布でもなく、それぞれに特徴の異なるものであった(図 7.13)。open 構造は第一主成分、第二主成分が(-5~5, -10~0)である構造と類似度が高かった。一方、2つの half-

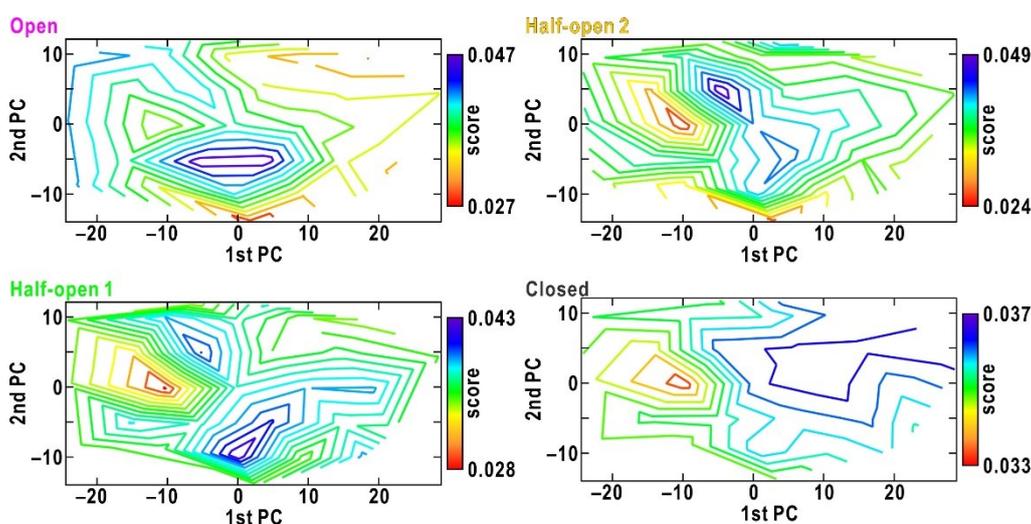


図 7.13 1st Classification における各 cryoEM 構造の MD 構造との類似度分布。本図及び図 7.14、図 7.15 の等高線プロットは IGOR Pro を用いて作成した。

open 構造では(-5 ~ 5, -10 ~ 0)と(-10 ~ 0, 5 ~ 10)の 2 地点にピークが生じた。また、closed 構造では主成分プロットの第一象限全体に広がる非常にブロードな類似度のピークが得られた。これは closed 構造 NAD ドメインの局所分解能の低さを反映しているものと考えられる。2nd – 10th Classification においても同様に類似度を計算したところ(付録 B2)、類似した cryoEM 構造間で類似度の分布も似通っていた。例えば、1st Classification の half-open1 構造と 7th Classification の #3 構造(half-open)などである。これらの結果は、ある cryoEM 構造についてその構造に特有の MD マップの組み合わせとして近似できることを示しており、従って(7.1)式の仮定を支持する結果であった。

7.3.4 GDH ドメイン運動の自由エネルギー地形解析

7.3.3 の結果を踏まえ、我々は cryoEM 構造 \vec{v}^m について

$$V_l^m = \sum_{k=1}^K \omega_k^m V_{kl}, \quad \sum_{k=1}^K \omega_k^m = 1 \quad (7.4)$$

が成り立つとし、重み ω_k^m を類似度から計算する方法を開発した(付録 C)。開発手法を各 cryoEM 構造に対して適用したところ、多くの cryoEM 構造で重みの分布は 1 つか 2 つのピークが支配的であり、重みの大きさこそ違うが、その位置は類似度分布においてピークが生じる位置と共通していた(図 7.14、付録 B2)。例えば、1st Classification の open

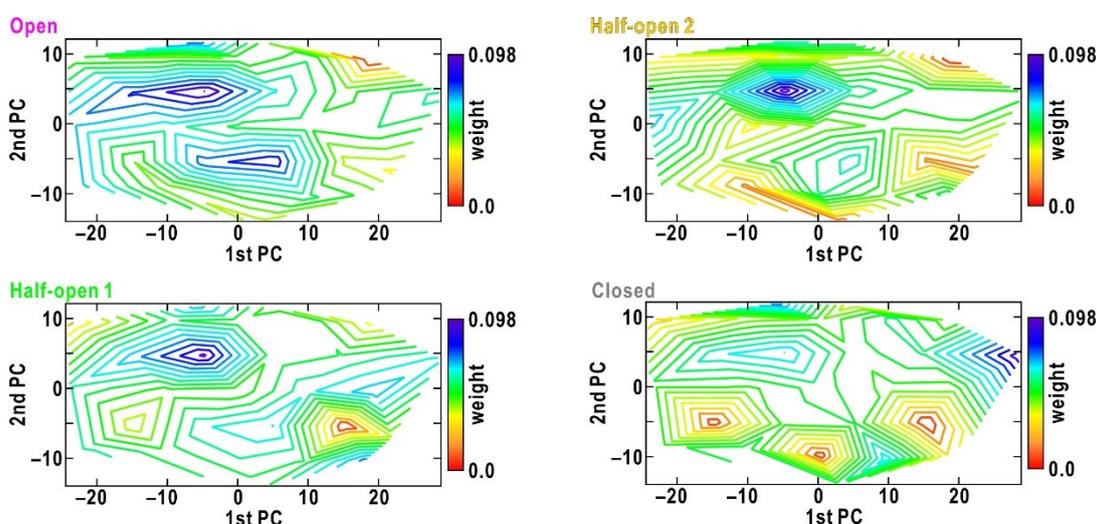


図 7.14 1st Classification の各 cryoEM 構造における MD 構造重み分布。

構造では類似度分布におけるピーク位置と同じ(-5 ~ 5, -10 ~ 0)のコンフォメーションに加えて、(-10 ~ 0, 0 ~ 5)のピークが出現した。また 2 つの half-open 構造では共に(-10 ~ 0, 0 ~ 10)、(0 ~ 10, -10 ~ 0)にピークが現れ、open 構造とはやや異なる分布であった。一方、closed 構造は他の 3 つの構造と分布が大きく異なり、(30, 5)が支配的であったが、(-10 ~ 0, 0 ~ 5)の open 側の conformation もわずかながら寄与していた。なお、これらの重みの計算はすべての cryoEM 構造において一部の重みが負の値をとった。しかしながら、重みが負の値となるコンフォメーションは共通しており、例えば(10 ~ 20, -10 ~ 0)のコンフォメーションは多くの cryoEM 構造で負の値をとった。そのため、不規則に負の値が現れるのではなく出現頻度の低いコンフォメーションがノイズなどによって負の値にずれ込んでいると考えられる。そこで cryoEM 構造ごとに最小(負の方向に最大)の重み ω_{\min}^m を用いて、 ω_k^m に

$$\omega_k'^m = \frac{\omega_k^m + |\omega_{\min}^m|}{1 + K|\omega_{\min}^m|} \quad (7.5)$$

という形でスケールリングを施した。本研究では $K=28$ である。これによって、重みの和を 1 に保ち、かつ ω_k^m の大小関係を保ったまま ω_{\min}^m が 0 となるようスケールリングされる。図 7.14 および付録 B2 における重みの分布はこのスケールリングが施されている。

7.2.3 節の結果より、cryoEM で観察された粒子像は急速凍結によって常温での構造揺らぎの中のあるコンフォメーションにトラップされていると考えられる。この時、ある

コンフォメーションのカノニカル集団としての出現頻度と cryoEM 粒子像における各コンフォメーションの出現頻度は等しいと見なせる。そこで、各 cryoEM 構造 m における MD マップ k の重み ω_k^m を以下のように足し合わせることで粒子像全体での MD マップ k の重み、すなわち MD マップ k のコンフォメーション出現頻度 ω_k を求めた。

$$\omega_k = \sum_m \frac{N_m}{N_0} \omega_k^m, \quad \sum_m N_m = N_0 \quad (7.6)$$

ただし、 N_0 が全粒子像の枚数、 N_m が cryoEM 構造 m に分類された粒子像の枚数である。また、(6.19)式より、 ω_k を用いて最安定構造(頻度 ω_{\max})との相対自由エネルギー ΔG_k は

$$\Delta G_k = -k_B T \ln \frac{\omega_k}{\omega_{\max}} \quad (7.7)$$

で求められる。ただし、 m についての和は同じ Classification における cryoEM 構造に対する和であり、従って自由エネルギー地形は Classification ごとに得られる。

得られた自由エネルギー地形は全 Classification でほぼ同様の形状となった(付録 B3)。これらの平均によって得られた最終的な自由エネルギー地形を見ると、広く浅い自由エネルギーの谷と 2 つの安定な窪みが現れ、MD 計算から得られた地形とは大きく異なっていた(図 7.15)。1 つは(-5 ~ 10, -10 ~ 0)に位置し、もう 1 つは谷の一部として(-10 ~ 0, 0 ~ 10)に形成されていた。この 2 つの窪みの深さの差はおよそ $0.2 k_B T$ であり、また両者の間には高さ $1.0 k_B T$ 程度のエネルギー障壁が存在した。谷の中での自由エネルギー変化量は $0.4 k_B T$ の範囲に収まっていた。これらのエネルギーの大きさは eV 単位で $\sim 10^{-2}$ 程度であり、光受容蛋白質の 1 光子吸収による励起が 1 eV 程度であることを考慮すると、室温での熱揺動として十分に起こりうるものである。谷は構造変化の経路としては第一主成分のドメイン開閉運動に沿ったものであるが、谷よりも自由エネルギーの低い窪みの存在を考慮すると NAD ドメインの運動経路は単なる開閉運動ではなくより複雑なものであると考えられる。

7.4 cryoEM 構造に対する原子モデルの構築と比較

7.3 節で説明した開発手法により GDH ドメイン運動の自由エネルギー地形の可視化に成功した。本節以降ではより高分解、あるいは部分構造に着目した解析について説明する。まず、1st Classification で得られた open、half-open1、half-open2 構造について分解能を上げるべく更なる分類を行った(図 7.16)。closed 構造に関しては元々の枚数が少なく、これ以上の分類では逆に分解能を悪化させると考えて行わなかった。得られた分類

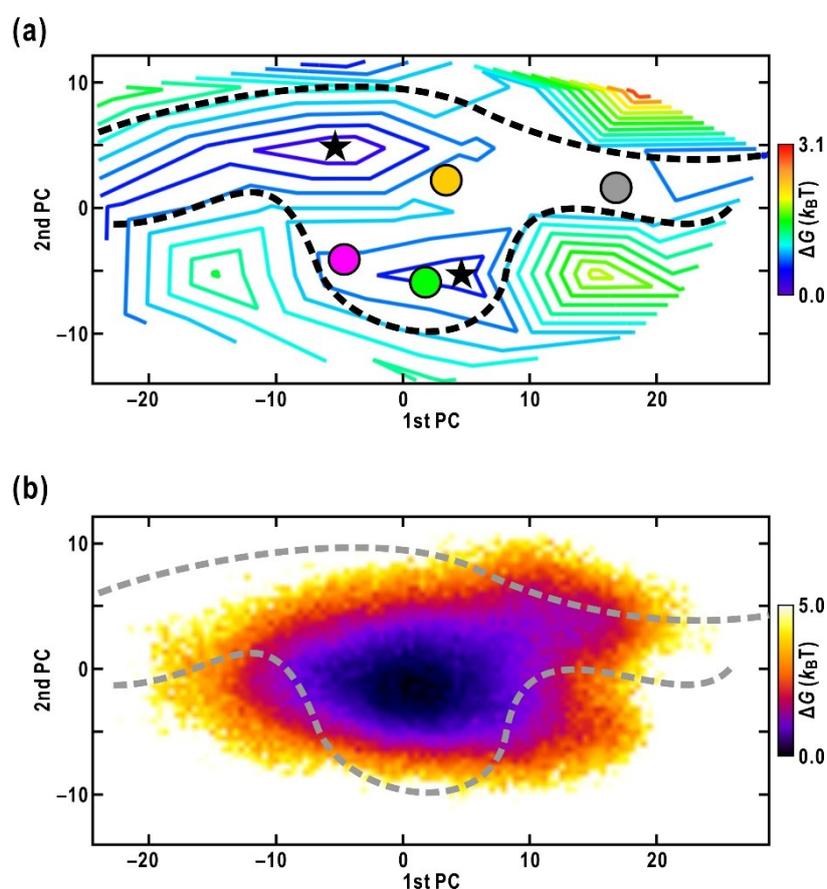


図 7.15 (a) 開発手法によって cryoEM 構造から解析された GDH ドメイン運動の自由エネルギー地形。丸印はそれぞれ 7.4 節の解析で得られた 1st Classification cryoEM 構造の原子モデル(マゼンタ: open、緑色: half-open1、黄色: half-open2、灰色: closed)を射影している。星印はエネルギー地形に形成された窪み(極小点)を示している。(b)(a)の自由エネルギー地形と MD 構造から得られた自由エネルギー地形の比較図。灰色の破線が(a)における谷の輪郭である。

結果に基づき、Particle Subtraction によるアーティファクトを避けるために減算前の粒子像で三次元再構成を行った。得られた三次元構造は図 7.8 と比較して分解能が非常に向上しており(図 7.17a、表 7.1)、二次構造やループの位置が判別可能であった(図 7.17b)。half-open2 は特にマップの質が良好であったため、まず half-open2 についてプログラム COOT[Emsley *et al.*, 2010]を用いて結晶構造を初期構造として原子モデルを作成した。その後、half-open2 の原子モデルについて、NAD ドメインを剛体的に移動させて他の 3

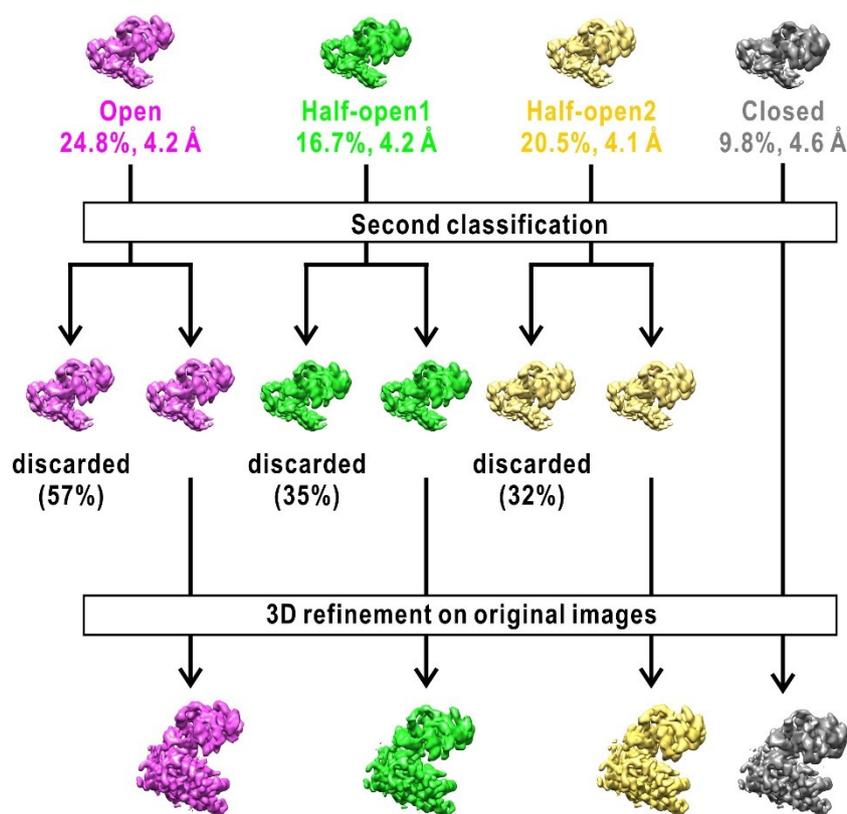


図 7.16 1st Classification cryoEM 構造の分類による精密化概略図

つの構造についてモデリングを行った。COOT によるモデル作成後、構造解析ソフトウェアスイート Phenix[Adam *et al.*, 2010]上で使用可能な Real-space refinement [Afonine *et al.*, 2018]によって各モデルを精密化した。Real-space refinement は立体化学的制限の下マップとモデルの一致度を向上させる処理である。

最終的に得られた原子モデル(図 7.17b)は主成分空間上に図 7.15a のように射影された。open 構造から half-open 構造にかけては $0.2 k_B T$ 程度のエネルギー差が存在し、また構造変化は 3 \AA 程度のドメイン開閉運動で与えられた(図 7.18a)。half-open1 構造から half-open2 構造にかけては 2.5 \AA 程度の剪断運動によって窪みから別の窪みへ遷移し(図 7.15a、7.18b)、half-open2 構造と closed 構造ではエネルギー差はほとんどないがドメインの開閉度合いが 3.5 \AA 異なった(図 7.18c)。これら原子モデルを結晶構造と比較したところ open、half-open2、closed はそれぞれサブユニット A、D、E とほぼ同様のコンフォメーションであった(図 7.19、表 7.1)。一方で half-open1 構造は結晶構造中には存在しなかった。これは結晶内部での相互作用によるものだと考えられる。

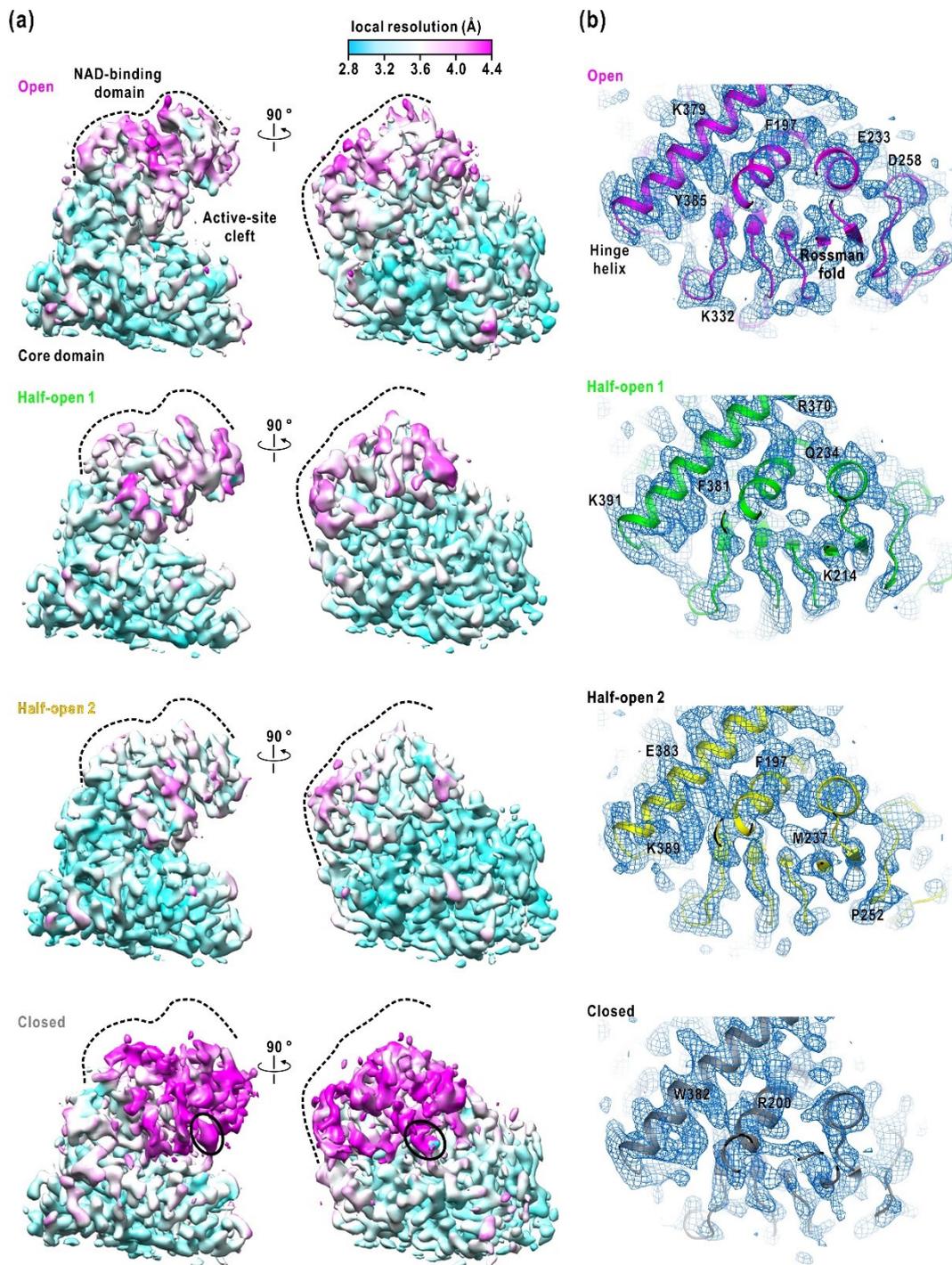


図 7.17 (a) 図 7.16 の解析で精密化された各コンフォメーションの cryoEM 構造。黒色の破線は open 構造の輪郭線である。(b) (a)のマップから作成された原子モデル。元のマップを青色で表示している。

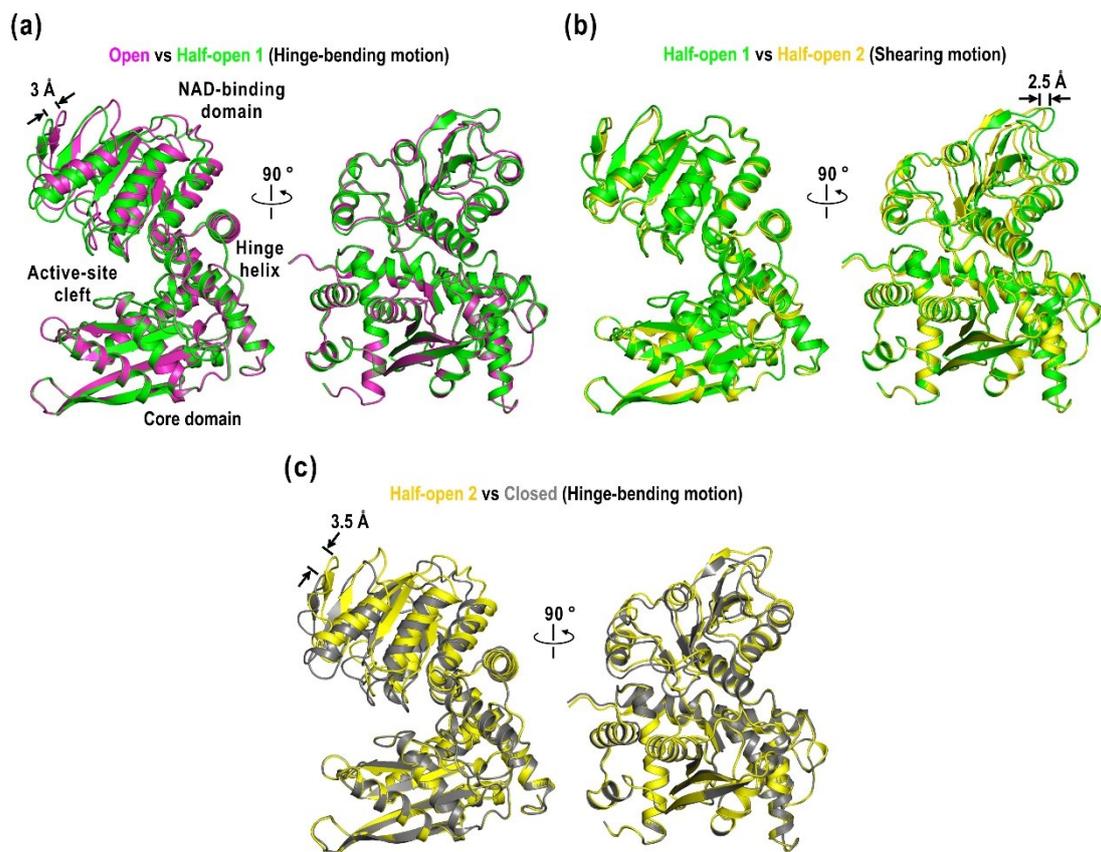


図 7.18 (a) open 構造と half-open1 構造、(b) half-open1 構造と half-open2 構造、(c) half-open2 構造と closed 構造の原子モデル比較図。それぞれ core ドメインについて重ね合わせている。

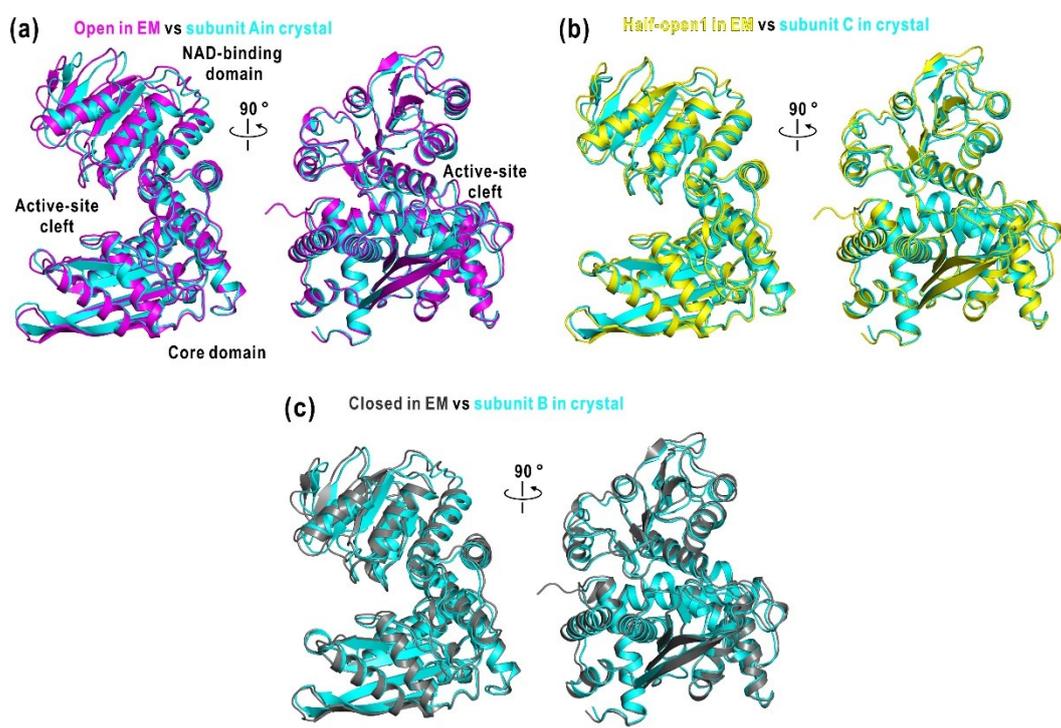


図 7.19 (a) open 構造、(b) half-open2 構造、closed 構造の表 7.1 で対応している結晶構造との重ね合わせ。結晶構造はすべてシアンで着色。

表 7.1 1st Classification cryoEM 構造の各種パラメータ

構造	粒子像枚数	分解能 (Å)	平均分解能 (Å)		一致する結晶構造	結晶構造との C _α -RMSD (Å) ^a
			NAD ドメイン / core ドメイン			
open	118,299	3.8	3.6 / 3.2		subunit A	1.32
half-open1	109,568	3.8	3.6 / 3.3		---	1.16 ^b
half-open2	146,327	3.7	3.4 / 3.1		subunits D (C)	0.90
closed	97,002	4.9	4.0 / 3.3		subunits E (B)	0.84
others	545,322	---	---		---	---

^a cryoEM 構造から得られた原子モデルと結晶構造を core ドメインに重ね合わせて計算した。

^b half-open1 は cryoEM でのみ得られたコンフォメーションであったが、比較のためサブユニット C と重ね合わせて計算した。

7.5 活性クレフト内部の構造

cryoEM 構造について、活性クレフト内部を見ると局所分解能が良好な core ドメインの一部である 89 番トリプトファン残基に明らかに構造多型が生じていた(図 7.20、付録 B4)。実際、core ドメイン内の他のトリプトファン残基は側鎖である五員環+六員環の構造(インドール環)が鮮明に見える一方で、89 番トリプトファンではインドール環の欠損や複数の突起の出現など、多様なマップが現れた。先行研究からこの 89 番トリプトファンは立体配座が変わりやすく、NAD ドメインの閉じ込みに対して立体障害となることでドメイン開閉運動制御機構の一端を担っていることが報告されている [Nakasako *et al.*, 2001; Oroguchi & Nakasako, 2016]。各 cryoEM 構造における 89 番トリプトファンの立体配座が一意でなく、複数の立体配座の足し合わせとして見えているのはこの 89 番トリプトファンの動きやすさによるものと思われる。

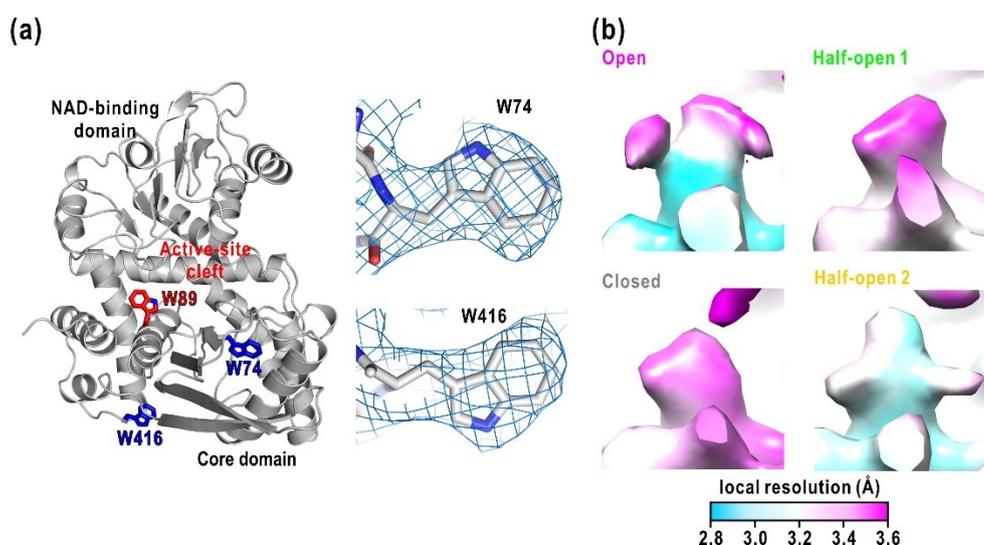


図 7.20(a) core ドメイン内の 3 つのトリプトファンの配置(左)と 74 番、416 番トリプトファンのマップおよび原子モデル(右)。(b) 各 cryoEM マップでの 89 番トリプトファンのポテンシャルマップ。局所分解能によって色付けされている。

7.6 議論

7.6.1 NAD ドメインの連続的構造変化

NAD ドメインは結晶構造で 3 種類のコンフォメーションをとっており、また MD 計

算においてもドメイン開閉運動が解析されていたものの、今までに解析に用いられた手法では各コンフォメーションの出現頻度や自由エネルギー地形を解析することは実質的に不可能であった。本研究では cryoEM を用いることで非結晶状態の GDH について構造解析をするとともに、溶液中での構造揺らぎについての一分子レベルでの情報を得ることができた。さらに、この情報を開発手法によって解析することで自由エネルギー地形の可視化に成功した。cryoEM 粒子像には Preferred Orientation が生じていたが、Focused Classification によって壊れた粒子像は十分に除去することができた。三次元構造分類には不完全性があるため完全な除去は不可能であるものの、各 Classification から得られた自由エネルギー地形はよく似通っていたため、実用上問ないレベルまで除去できていたといえる。また、得られた地形は 200 ナノ秒 MD 計算から得られた地形と大きく異なるものであった。従って、200 ナノ秒程度の MD 計算は取りうる構造揺らぎをスペースにサンプリングできるが、各コンフォメーションの出現頻度について現実の系を再現することは困難であると評価できる。MD 計算での構造変化の振る舞いは使用する力場のパラメータに依存するため[Beauchamp *et al.*, 2012]、現状では有限時間の MD 計算による現実のコンフォメーション出現頻度の再現は極めて困難であると考えられる。

7.6.2 本研究の解析における仮定の妥当性

本研究では新規解析手法の提案にあたって(7.1)式、あるいは(7.4)式が成り立つと仮定した。本節ではこの仮定の妥当性を検討する。

簡単のために $K = 2$ 、すなわち内在するコンフォメーションが構造 1、構造 2 の二つしかない系を考える。このとき、構造 1 と 2 を独立に三次元再構成して足し合わせた構造 \bar{v} のボクセル l における値 V_l は(6.14)式、(7.1)式より

$$V_l = \frac{N_1}{N} \frac{\sum_{i_1}^J \sum_j^J P_{ij}^{\phi_i T} \frac{\text{CTF}_{i_1 j}}{\sigma_{i_1 j}^2} X_{i_1 j}}{\sum_{i_1}^J \sum_j^J P_{ij}^{\phi_i T} \frac{\text{CTF}_{ij}^2}{\sigma_{ij}^2} + \frac{1}{\tau_{1l}^2}} + \frac{N_2}{N} \frac{\sum_{i_2}^J \sum_j^J P_{ij}^{\phi_i T} \frac{\text{CTF}_{i_2 j}}{\sigma_{i_2 j}^2} X_{i_2 j}}{\sum_{i_2}^J \sum_j^J P_{ij}^{\phi_i T} \frac{\text{CTF}_{i_2 j}^2}{\sigma_{i_2 j}^2} + \frac{1}{\tau_{2l}^2}} \quad (7.8)$$

で与えられる。ただし、配向 ϕ は各粒子像で ϕ_i に決定されているため、 ϕ についての積分は消去されている。分母第一項について、CTF はデフォーカスによって決まるがデフォーカスのバリエーションはさほど大きくない。また、測定中は電子線照射線量が一定であるため、ノイズの分散も似たような振る舞いになる。従って N 、 N_1 、 N_2 がそれぞれ十分に大きければ、足しこまれる CTF およびノイズの分散の値は構造 1 と構造 2 で同

じようにサンプリングされ、結果的に粒子像の枚数で互いにスケールするような値になっていると考えられる。以上の議論より、構造2について分母第一項を

$$\sum_{i_2}^{N_2} \sum_j^J P_{lj}^{\phi_i^T} \frac{\text{CTF}_{i_2j}^2}{\sigma_{i_2j}^2} = \frac{N_2}{N_1} \sum_{i_1}^{N_1} \sum_j^J P_{lj}^{\phi_i^T} \frac{\text{CTF}_{i_1j}^2}{\sigma_{i_1j}^2} \quad (7.9)$$

と近似する。すると、(7.8)式は

$$V_l = \frac{N_1}{N} \frac{\sum_{i_1}^{N_1} \sum_j^J P_{lj}^{\phi_i^T} \frac{\text{CTF}_{i_1j}^2}{\sigma_{i_1j}^2} X_{i_1j}}{\sum_{i_1}^{N_1} \sum_j^J P_{lj}^{\phi_i^T} \frac{\text{CTF}_{ij}^2}{\sigma_{i_1j}^2} + \frac{1}{\tau_{1l}^2}} + \frac{N_1}{N} \frac{\sum_{i_2}^{N_2} \sum_j^J P_{lj}^{\phi_i^T} \frac{\text{CTF}_{i_2j}^2}{\sigma_{i_2j}^2} X_{i_2j}}{\sum_{i_1}^{N_1} \sum_j^J P_{lj}^{\phi_i^T} \frac{\text{CTF}_{ij}^2}{\sigma_{i_1j}^2} + \frac{1}{\frac{N_2}{N_1} \tau_{2l}^2}} \quad (7.10)$$

と変形される。第二項の分母第二項(τ_{2l}^2 の項)を無視すれば、これは構造1と2の粒子像を混ぜた三次元再構成によって得られた構造と等価である。すなわち、構造を分けて三次元再構成した場合と、分けずに再構成した場合では τ_{2l}^2 が異なる。この τ_{2l}^2 の振る舞いを N_1 、 N_2 の比率で場合分けして定性的に考える。まず $N_1 > N_2$ の場合、 $\tau_{1l} > \tau_{2l}$ となる。これは、互いに構造が単一かつ粒子像の配向が正しく決定されているため信号の強さ、すなわちコントラストが枚数のみによって決まるためである。構造を分けずに再構成することは(7.10)式で $N_2 \tau_{2l}^2 / N_1$ を τ_{1l}^2 に置き換えることに等しいが、 N_1 と N_2 、 τ_{1l} と τ_{2l} の大小関係から明らかに $\tau_{1l}^2 > N_2 \tau_{2l}^2 / N_1$ である。従って、この場合構造2の構造振幅、即ち分解能が過大評価される。 $N_1 < N_2$ の場合、先ほどの議論における不等号が全て反転するためこの場合は構造2の分解能が過小評価される。以上より異なる構造をまとめて三次元再構成した場合、独立に再構成して足し合わせた場合と比べて分解能の過大評価、あるいは過小評価によって差異が生じている。この近似のずれは明らかに三次元再構成像内の構造多型が甚だしいほど大きくなる。今回の解析では分類された各構造に対して(7.1)式の近似を適用しており、最終的な自由エネルギー地形は各 Classification でほぼ同一の地形であった。そのため、ある程度分類されていれば十分に近似が成り立っていると考えられる。また、結果論的ではあるが本研究での開発手法を適用するならば、ある程度の分類を行ったうえで各構造ごとに適用することが望ましいという結論が得られる。

7.6.3 開発手法の改善点

開発手法は数十 kDa 程度の分子量の連続的構造変化を解析でき、この点で極めて有

用である。実際、GDH について Multi-body analysis を適用してみたところ計算は収束せず徐々に分解能が悪化していった。従って、このサイズの蛋白質の構造変化を解析するには開発手法が現状唯一の手法である。一方で、提案手法には様々な改善点が存在する。

一点目が重みの計算方法についてである。7.3.4 節で述べた通り、開発手法で計算された重みは負の値をとることがあった。これはおそらくノイズに由来するオフセットであるが、ノイズは避けようがないため解析面で工夫が必要となる。本研究では線形結合(重みの和が1であるため厳密にはアフィン結合)であると仮定したが、実際にはこれに加えて各重みがゼロ以上の値をとるという拘束を付加しなければならない。このとき、解くべき問題は線形代数的問題ではなく cryoEM 構造をよく説明する非負の重みを求める最適化問題となる。最適化問題には適切な初期値が必要であるため、今回の線形代数的解析で得られた重みを初期値として最適化問題を解く、という方法が効果的だと思う。

二点目は参照構造についてである。本研究では自由エネルギー地形を構築する参照構造として MD 計算で得られた構造を用いた。計算機性能の向上により数百 kDa 程度の蛋白質の MD 計算は容易になりつつあるため、開発手法は極めて有用であろう。一方で、それ以上のサイズの試料では計算コストの観点から MD 計算で参照構造を用意することが難しくなる。このような場合にはその他の手法で参照構造を得るのが望ましい。例えば、大きい分子であれば粗視化 MD でも十分にコンフォメーションのサンプリングは可能であるし、あるいは MDa サイズの粒子であれば Multi-body Analysis を適用して得られた構造をもとに本手法を用いて自由エネルギー地形を解析する、といった相補的利用が考えられる。MD 計算についてもいかに効率よく、コンフォメーションを尽くした参照構造が得られるか、という点は極めて重要であるため、位相空間の探索に優れた拡張アンサンブル法 [Sugita *et al.*, 2000; Harada & Kitao, 2011]などを積極的に利用していくなどの検討が今後求められる。

三点目は自由エネルギー地形を構築した座標系についてである。本研究では MD トラジェクトリーに対して PCA を行い、第一、第二固有ベクトルで張られた空間に構築した。しかし、主成分分析は構造変化を線形に近似するものであり、蛋白質構造変化は本来回転を主とする非線形なものである。そのため、原点近傍では精度が良いが、原点から離れた領域では歪みが大きくなる。そのため、この座標系についても検討の余地がある。例えば、蛋白質構造変化について多様体学習[Coifman *et al.*, 2015; Bishop *et al.*, 1998; Maaten & Hinton, 2008]を適用し、得られた非線形な固有ベクトルを座標系として用いる

といった改善案が考えられる。

7.6.4 GDH ドメイン運動解析における今後の展望

本研究で得られた自由エネルギー地形では機能発現と相関したドメイン開閉運動に対応する谷が存在したが、同時に開閉運動を阻害する 2 つの窪み(準安定状態)も見つかった。第二主成分である剪断運動はこの 2 つの準安定状態の間を遷移するような運動であるため、NAD ドメインの運動は単純な開閉運動よりも複雑であると考えられる。先行研究を踏まえると、GDH ドメイン運動におけるこれら 2 つの準安定状態の寄与を解析するためには、各コンフォメーションにおける水和構造変化の解析が必要である。しかしながら、cryoEM 構造では同程度の分解能の X 線結晶構造と比較して酸素原子、特に水分子が見えにくい[Zivanov *et al.*, 2018]。これはおそらく負電荷を帯びた酸素の原子散乱因子に由来するものと考えられる。従って、より詳細な解析のためには cryoEM 構造からの水和水分子予測 [Matsuoka & Nakasako, 2009; Matsuoka & Nakasako, 2010]のよ様な複合的なアプローチを考えていかなければならない。

ドメイン運動メカニズムという観点では点変異の導入や補酵素・基質との結合によってどのように自由エネルギー地形が変化するか、ということが興味の俎上に載る。本研究では GDH グルタミン酸結合サイトに tris らしき分子が結合していたものの、close 構造は安定な状態ではなかった。しかしながら、グルタミン酸結合状態の GDH 結晶構造が得られていることを考慮すると[Stillman *et al.*, 1993]、溶液にグルタミン酸を添加することで closed 構造が安定化するものと考えられる。開発手法を用いてこれを自由エネルギー地形変化の観点から論じることが可能になると期待される。

8章 結語

本研究は溶液中での蛋白質構造変化解析を軸に据え、SAXS 及び cryoEM を用いた解析手法の提案とその応用を行った。SAXS については、従来から用いられていた分子形状推定法の問題点を解決すべく、多変量解析による分子形状推定スキームを考案した。cryoEM については、現状では不可能であった数百 kDa の蛋白質構造変化解析を可能とすべく、MD 計算を援用した新規動態解析手法を開発した。また、各手法の応用として、光受容蛋白質フォトトロピン、フィトクロム B の SAXS による構造・機能発現メカニズム解析及び、cryoEM を用いた超高度好熱菌由来 GDH のドメイン運動解析を実施した。

SAXS における分子形状推定スキームは先行研究で提案された手法と比較しても性能に遜色はなく、テスト試料である PDI のような SAXS プロファイルの不確かさが甚だしい場合に特に有効であった。試料 SAXS プロファイルの不確かさに応じて既存手法と本研究の提案手法を使い分けることで効率的な分子形状推定が可能になると期待される。また、cryoEM では新規動態解析手法によって数百 kDa の蛋白質の、数十 kDa 程度の部分的な領域における構造揺らぎの自由エネルギー地形が明らかにされた。このスケールでの蛋白質動態の cryoEM による解析は現在までに例がなく、今後さらなる進展と応用が見込まれる。

応用研究においては、フォトトロピン及びフィトクロム B について現在まで不明であった全体形状を明らかにするとともに、光受容に伴う機能発現メカニズムの推定を行った。推定されたメカニズムは両者ともにあるドメインでの光受容による構造変化が離れたドメインの構造変化や活性化につながるというものであった。これらの蛋白質は共に植物光応答のシグナル伝達初発過程を担う蛋白質であり、両者で共通の機能発現メカニズムが推定されたことは興味深い結果である。今後の研究によって植物の光依存的生理現象の普遍的なメカニズムが明らかになるかもしれない。また、GDH に関しては、新規手法によって得られた自由エネルギー地形がドメイン開閉運動について単なる開閉運動ではなく剪断運動を交えた複雑な構造遷移を経由するものであるという可能性を示唆した。また、三次元構造分類によって得られた複数のコンフォメーションは共通して開閉運動に重要な 89 番トリプトファン残基の構造多型を示していた。今後は、コンフォメーションとこのような重要な局所領域の相関解析や、補酵素を添加した場合の自由エネルギー地形変化の解析などに歩を進めていくことが考えられる。

本研究で行われた解析や得られた結果は、それぞれの構造・動態解析手法の相補性を示すものでもあった。例えば、SAXS における分子形状推定においては結晶構造解析で得られた部分構造や cryoEM による低分解能投影像が推定の強力な補助となった。また、cryoEM においては SAXS プロファイルとの比較によって cryoEM 観察像から得られる構造アンサンブルが室温の溶液中での構造揺らぎを反映していることを確認できた。MD 計算は極めて強力な動態解析手法であるものの、MD 計算による自由エネルギー地形は cryoEM 観察像から得られた自由エネルギー地形とは大きく異なるものであった。これは限られた時間の MD 計算では溶液中での構造揺らぎのアンサンブルを再現できないことを示しており、現実の系における構造アンサンブルを解析可能な cryoEM の性質は動態解析手法として他にない長所である。一方で GDH cryoEM 構造において原子モデルが組めるほどに分類された構造であっても 89 番トリプトファン残基は構造多型を示した。アミノ酸一残基レベルでの構造分類はコントラストの観点から不可能であり、このような局所的な領域の構造多型解析には逆に MD 計算が有効な手法となる。今後はこのように複数手法を相補的に組み合わせるアプローチが極めて重要になっていくと予想される。

生命現象の理解という大きな枠組みの中で本研究は新規解析手法の提案と複数の実験手法の相補利用によって、蛋白質構造・動態解析をほんの少しだけ進めることができたものと考えている。今後の研究によって生命を理解するための構造生物学というツールがさらに発展していくことを期待し、結びとする。

付録 A

アミノ酸の名称及び略記

アミノ酸	三文字表記	一文字表記
アラニン	Ala	A
アルギニン	Arg	R
アスパラギン	Asn	N
アスパラギン酸	Asp	D
システイン	Cys	C
グルタミン	Gln	Q
グルタミン酸	Glu	E
グリシン	Gly	G
ヒスチジン	His	H
イソロイシン	Ile	I
ロイシン	Leu	L
リジン	Lys	K
メチオニン	Met	M
フェニルアラニン	Phe	F
プロリン	Pro	P
セリン	Ser	S
スレオニン	Thr	T
トリプトファン	Trp	W
チロシン	Tyr	Y
バリン	Val	V

付録 B

Focused Classification による GDH 構造分類結果の詳細

本研究では GDH cryoEM 観察像に対して Focused Classification を 10 回行い、それぞれの分類結果、および分類構造の解析結果を比較した。本付録にてこれらの詳細を掲載する。

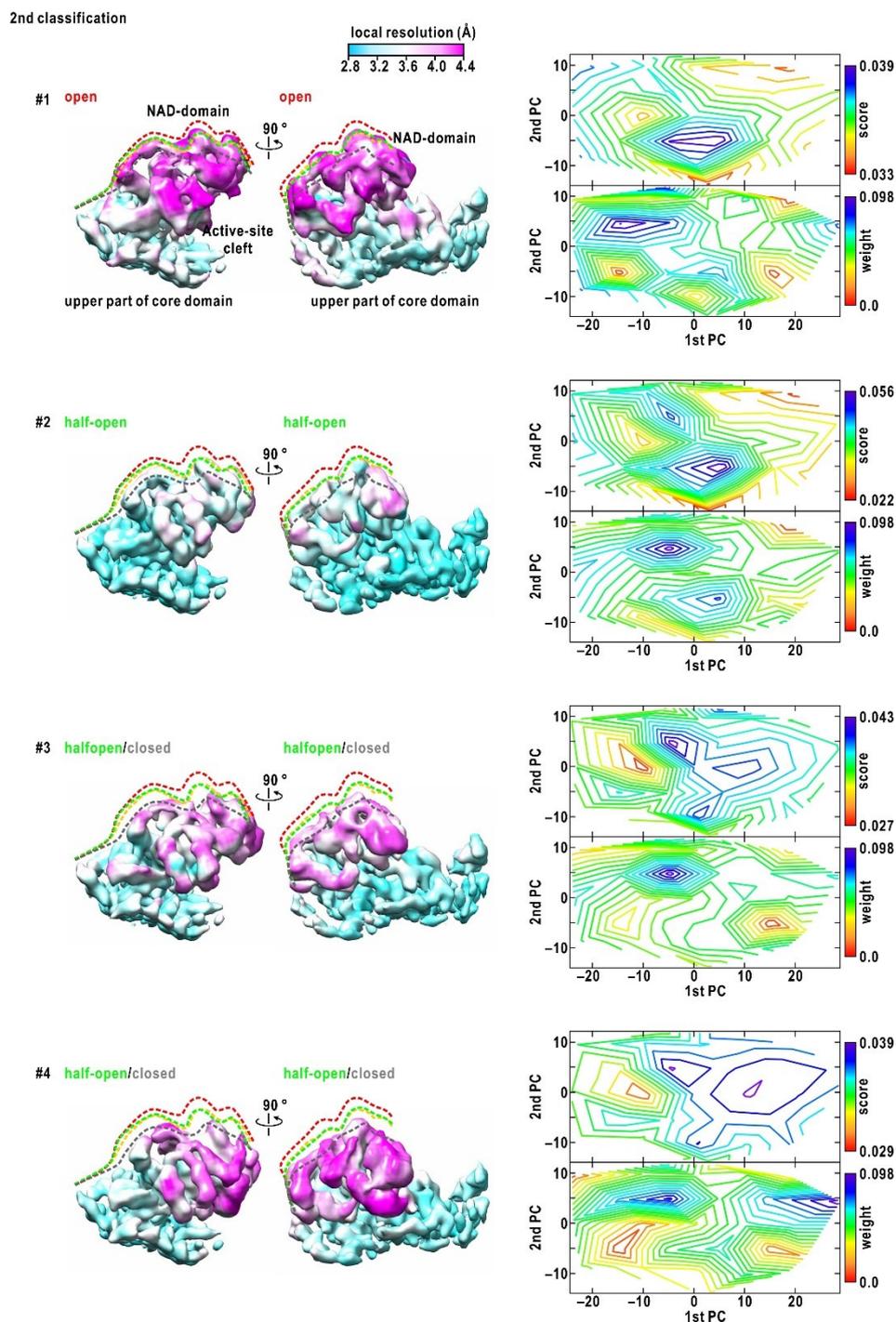
B1. 各 Classification における分類構造の粒子増枚数と得られた構造

Classification	粒子像の割合	分解能 (Å)	構造
1	0.248	4.06	open
	0.167	4.24	half-open
	0.205	4.06	half-open
	0.098	4.64	closed
	0.082/0.046/0.068/0.086	5.13/5.91/5.73/5.41	discarded
2	0.159	4.43	open
	0.275	3.90	half-open
	0.159	4.24	half-open
	0.122	4.33	closed
	0.063/0.055/0.067/0.101	5.91/5.91/5.41/5.13	discarded
3	0.188	4.43	open
	0.266	3.98	half-open
	0.191	4.06	half-open
	0.106	4.75	closed
	0.083/0.051/0.066/0.50	4.87/5.91/5.73/6.09	discarded
4	0.257	4.15	open
	0.158	4.24	half-open
	0.174	4.15	half-open
	0.093	4.53	closed
	0.074/0.052/0.090/0.101	5.41/6.09/4.75/4.75	discarded
5	0.210	4.15	open
	0.175	4.06	half-open
	0.080	4.87	half-open
	0.176	4.33	closed
	0.076/0.058/0.132/0.092	5.13/5.81/4.75/4.78	discarded

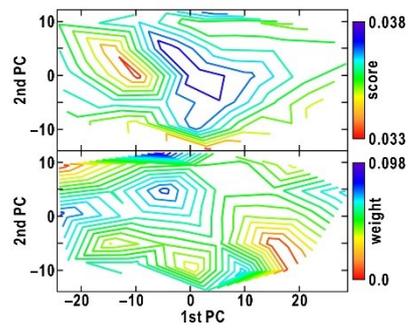
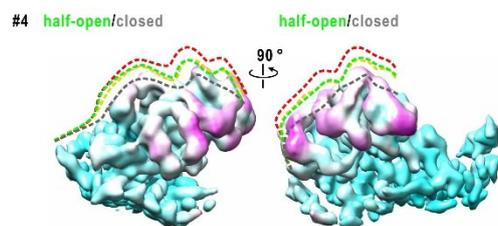
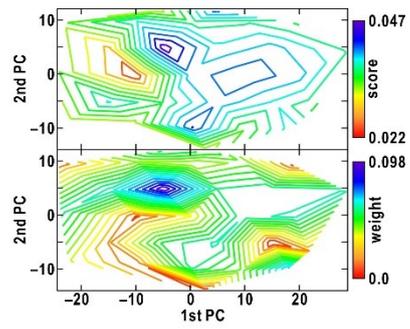
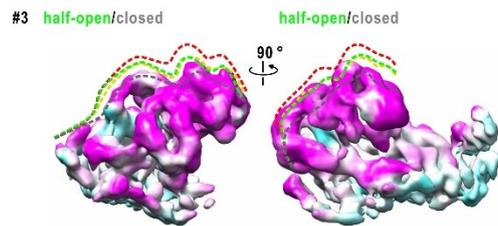
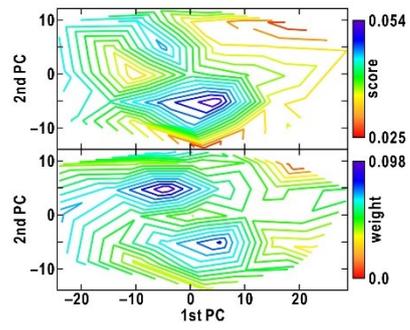
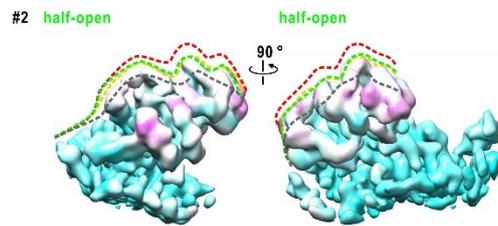
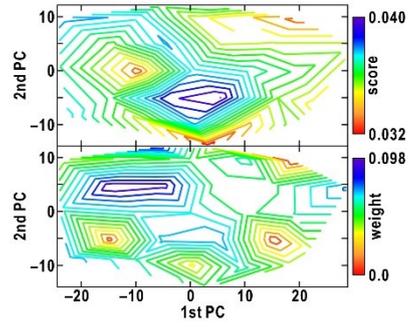
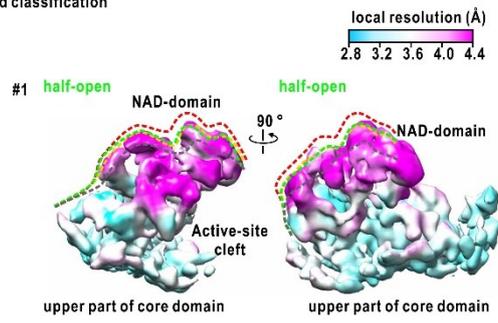
6	0.188	4.43	open
	0.202	4.06	half-open
	0.107	4.53	half-open
	0.124	4.33	half-open
	0.115	4.43	closed
	0.091/0.088/0.085	4.87/5.27/5.27	discarded
7	0.233	4.15	open
	0.133	4.24	half-open
	0.133	4.43	half-open
	0.123	4.33	half-open
	0.125	4.43	closed
	0.062/0.100/0.092	5.91/4.64/5.00	discarded
8	0.207	4.24	open
	0.209	3.98	half-open
	0.166	4.43	half-open
	0.110	4.43	closed
	0.093/0.094/0.053/0.069	4.75/5.13/6.09/5.57	discarded
	9	0.232	4.24
0.240		4.06	half-open
0.200		4.06	half-open
0.104		4.64	closed
0.055/0.057/0.062/0.050		5.91/5.73/5.73/5.91	discarded
10		0.238	4.15
	0.143	4.43	half-open
	0.237	3.90	half-open
	0.086	4.87	closed
	0.098/0.068/0.074/0.057	5.00/5.57/5.27/5.91	discarded

B2. 2nd - 10th Classification 分類構造と各構造の類似度分布、重み分布

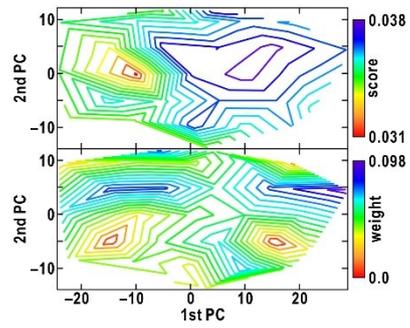
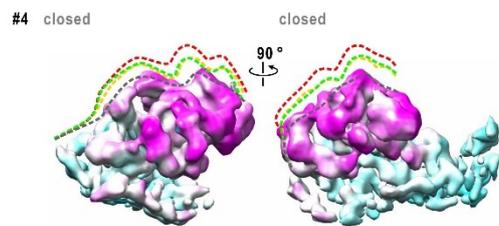
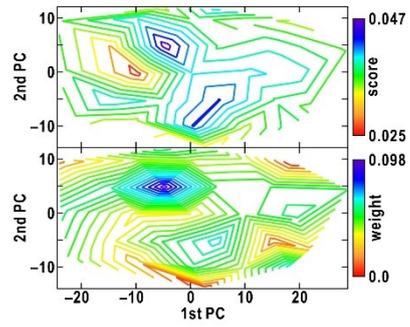
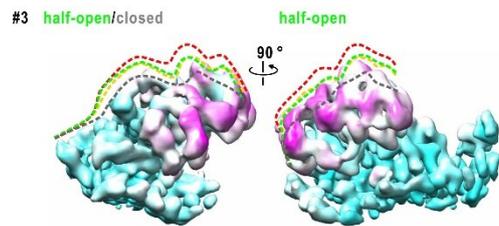
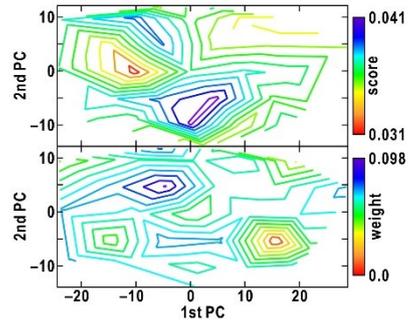
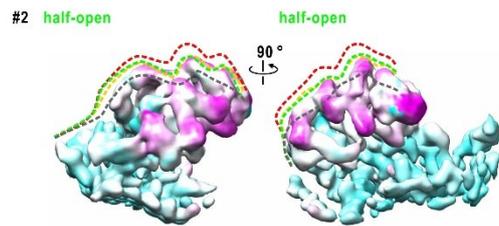
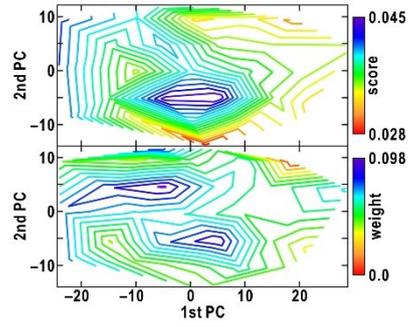
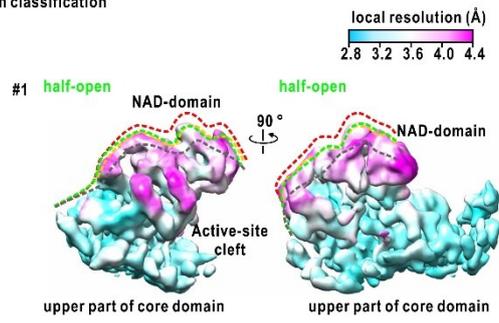
各分類構造の右に配置された等高線図は上が類似度分布、下が重み分布である。



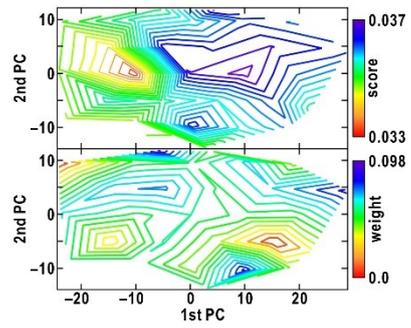
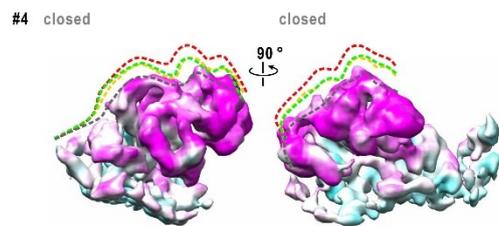
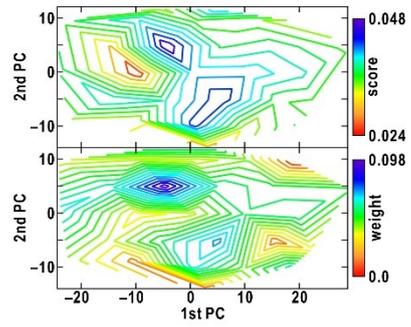
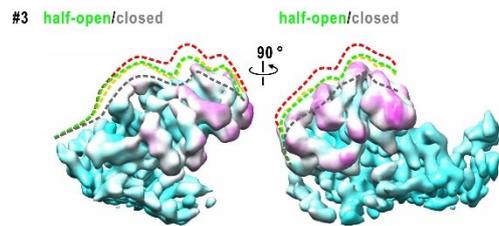
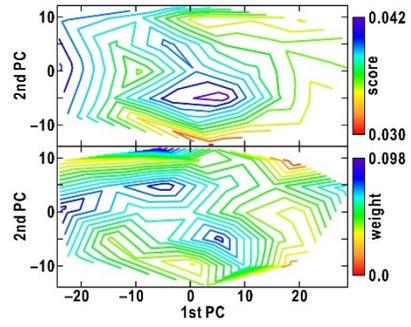
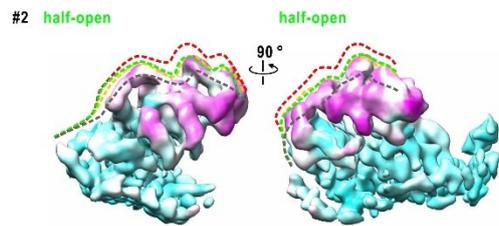
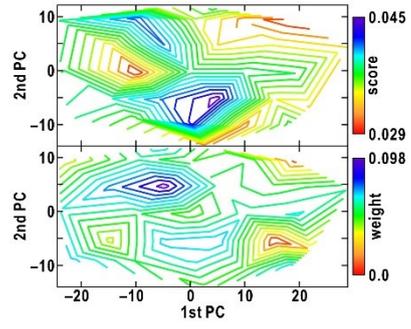
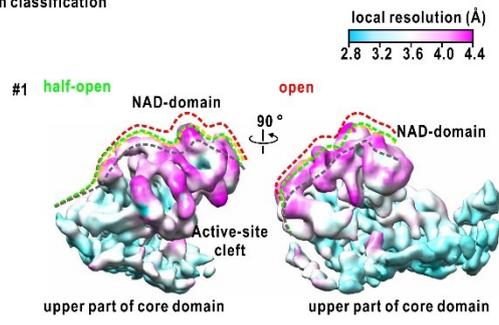
3rd classification



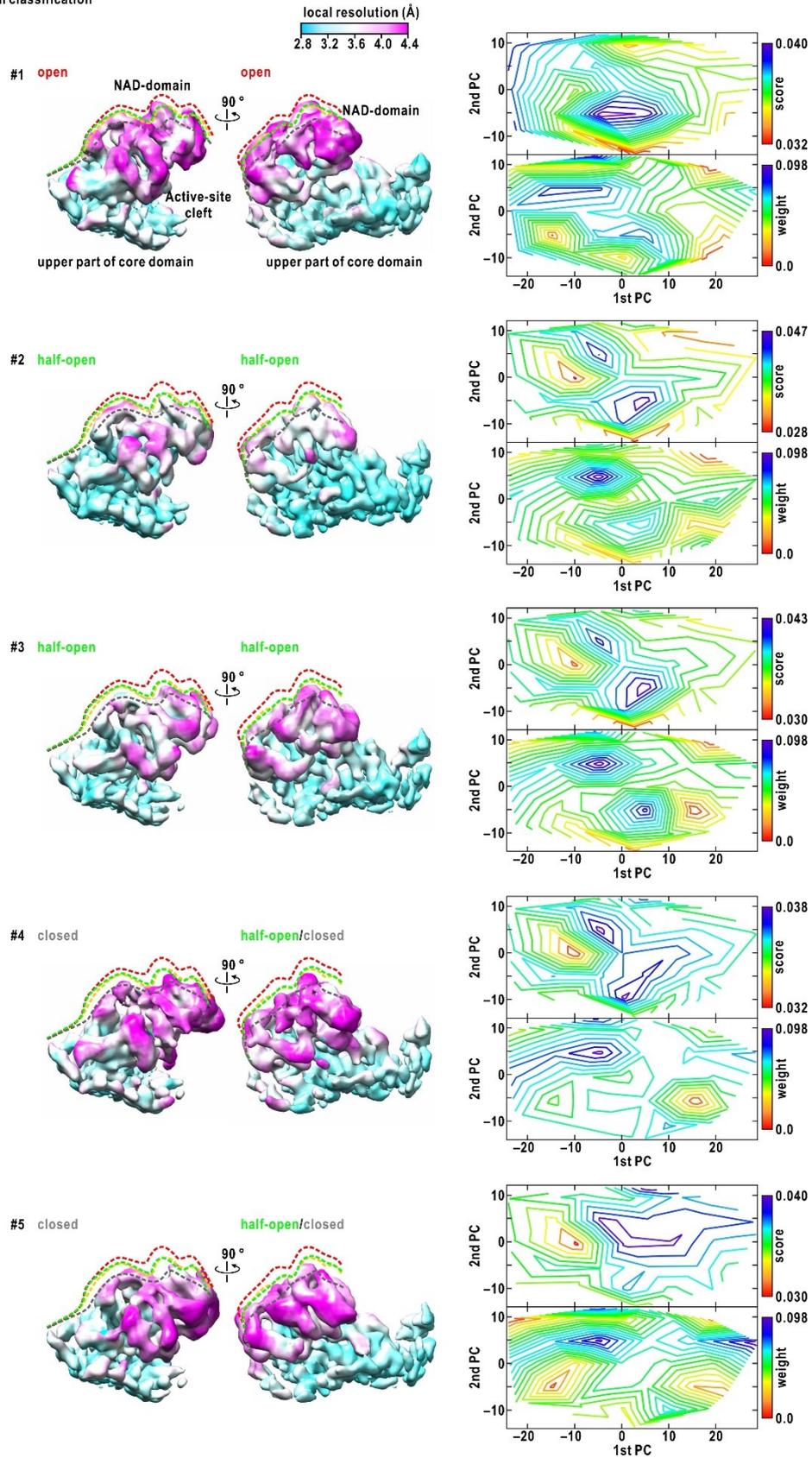
4th classification



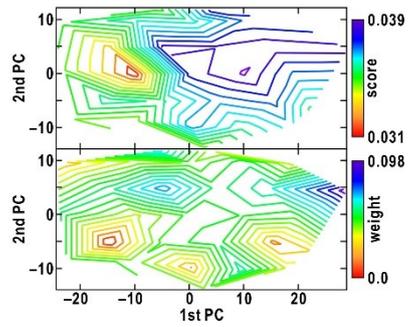
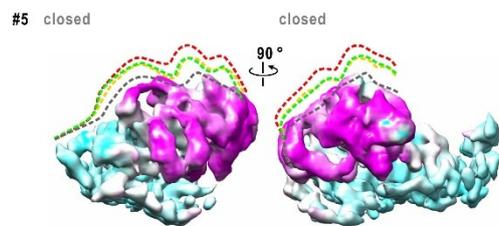
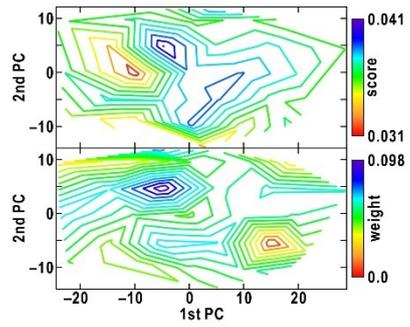
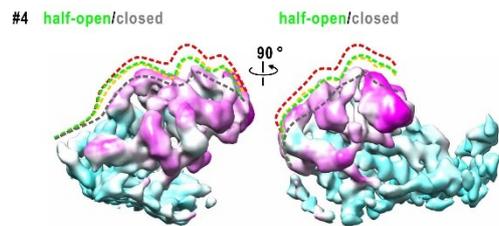
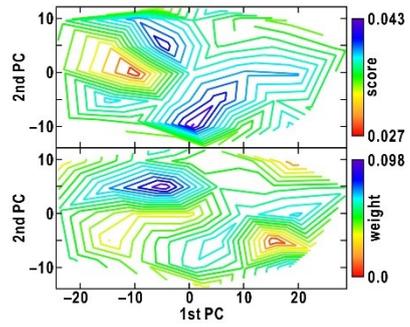
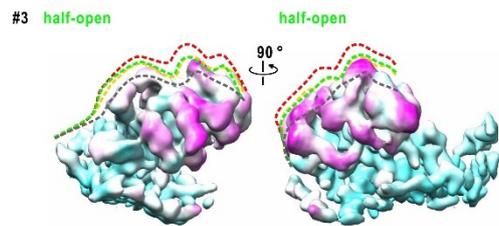
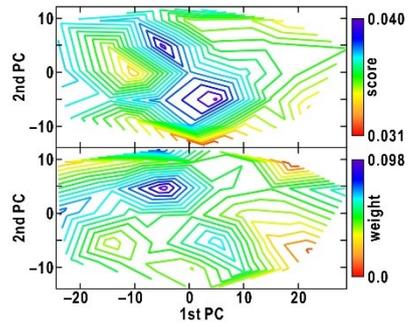
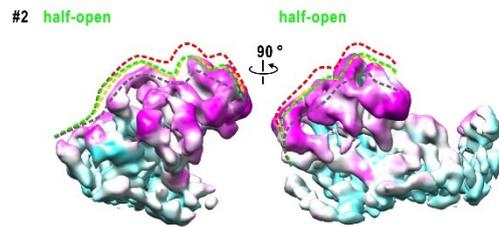
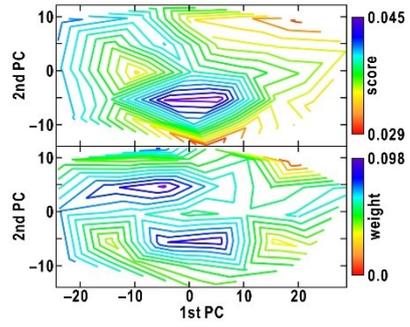
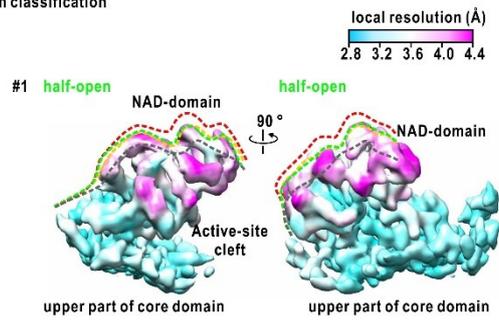
5th classification



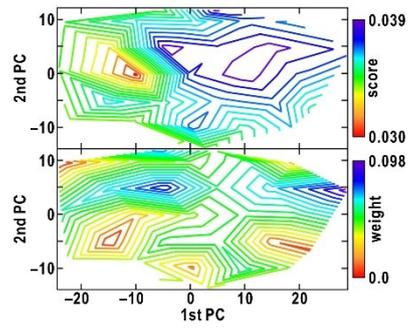
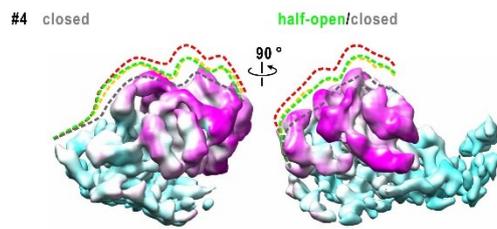
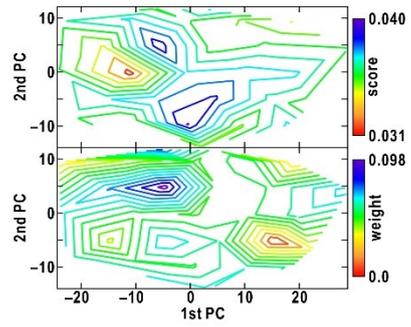
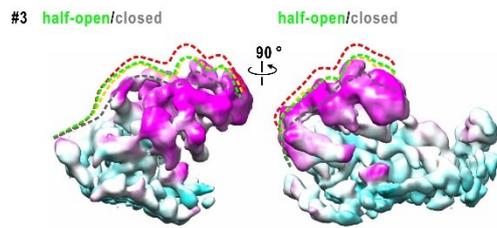
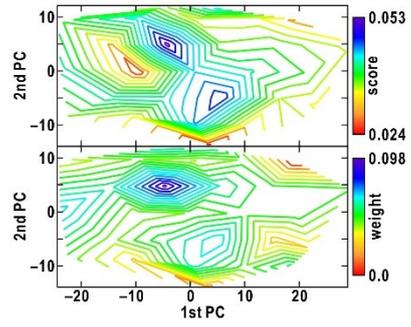
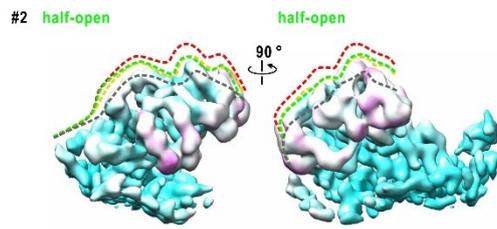
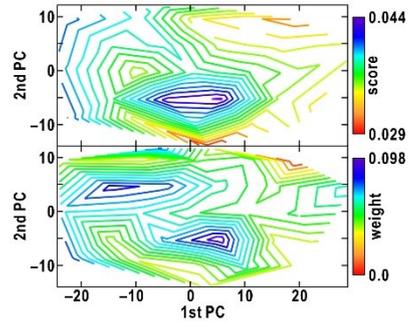
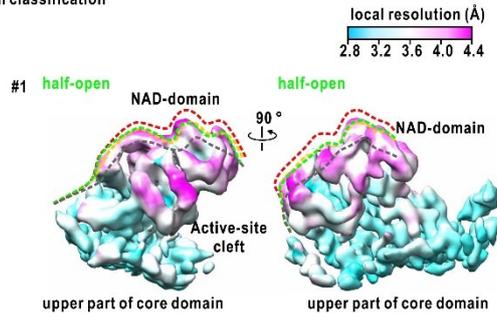
6th classification



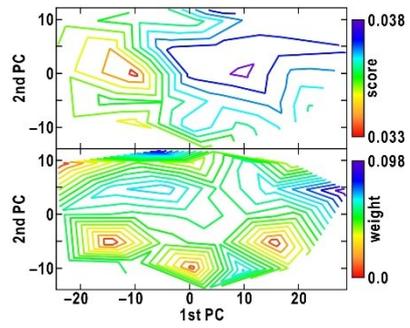
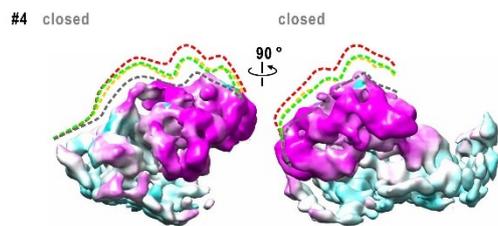
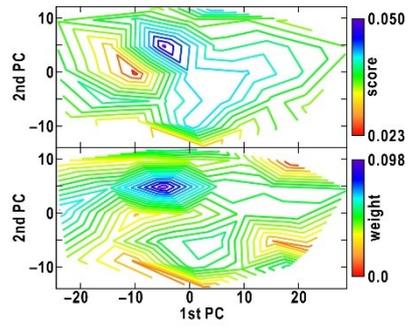
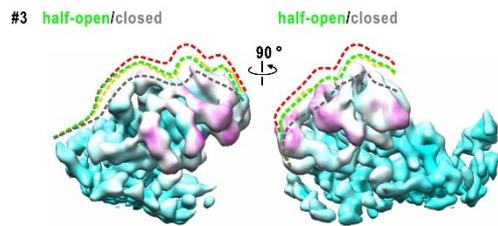
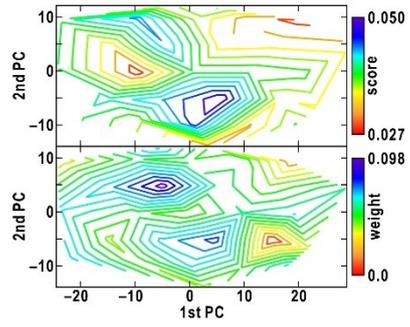
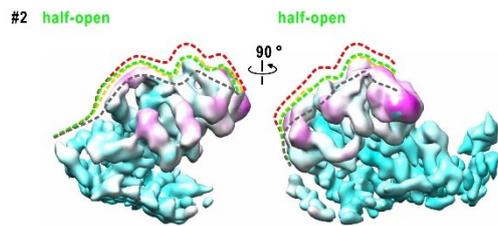
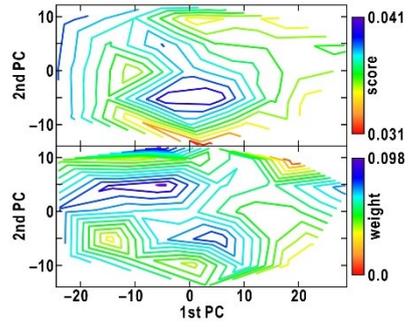
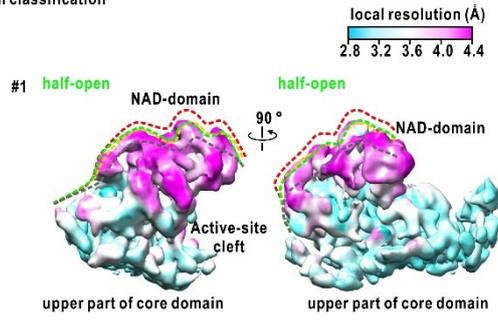
7th classification



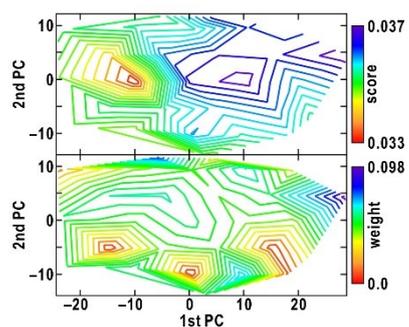
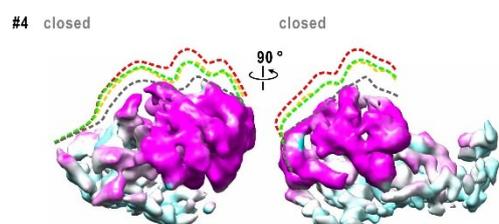
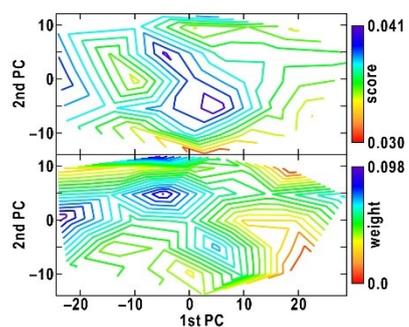
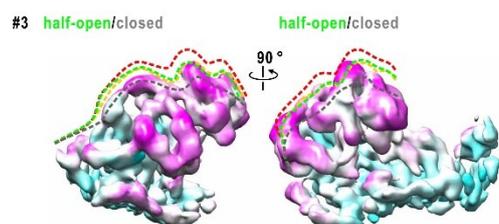
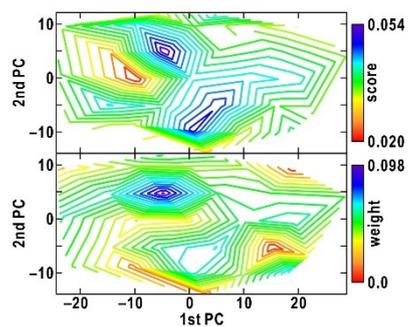
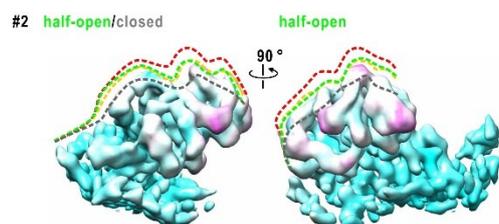
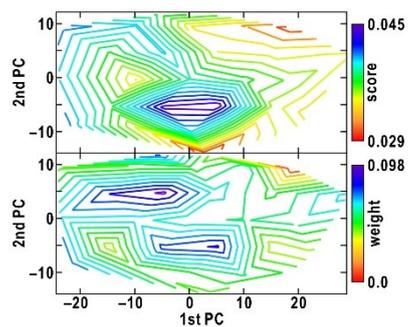
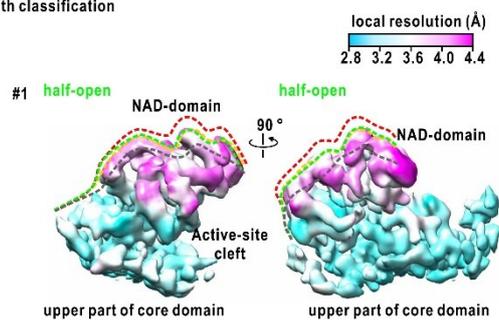
8th classification



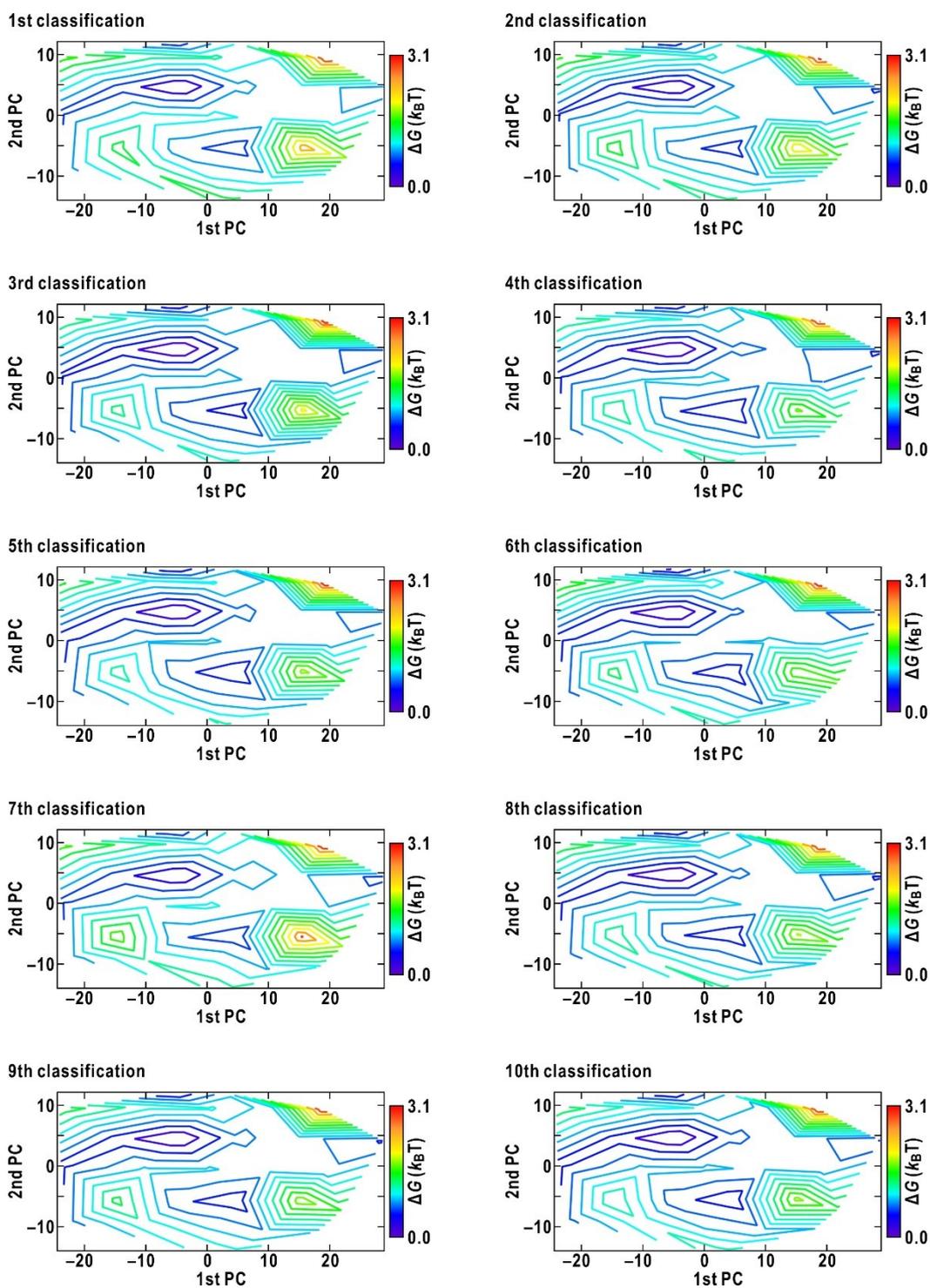
9th classification



10th classification

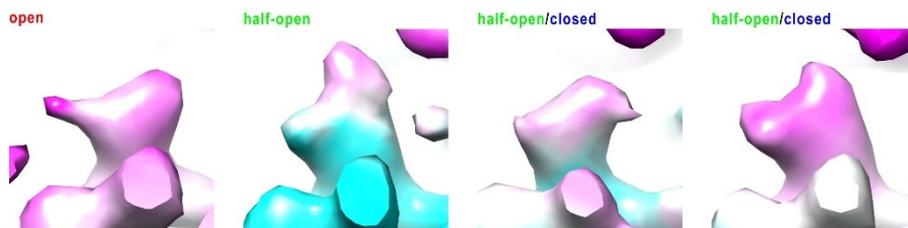


B3. 各 Classification ごとに得られた GDH ドメイン運動の自由エネルギー地形

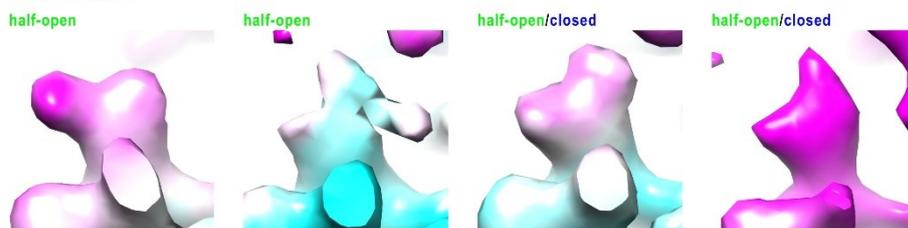


B4. 各 Classification 分類構造における 89 番トリプトファン残基ポテンシャルマップ

2nd classification



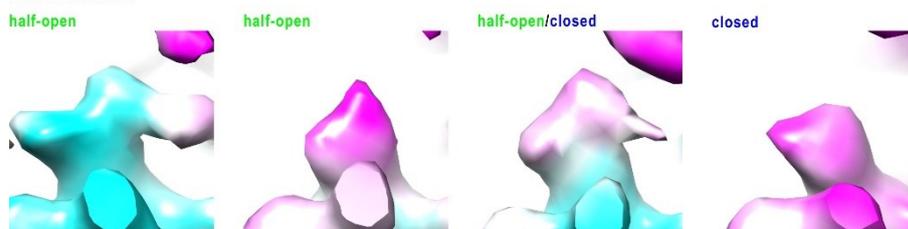
3rd classification



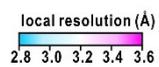
4th classification

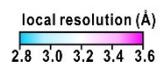
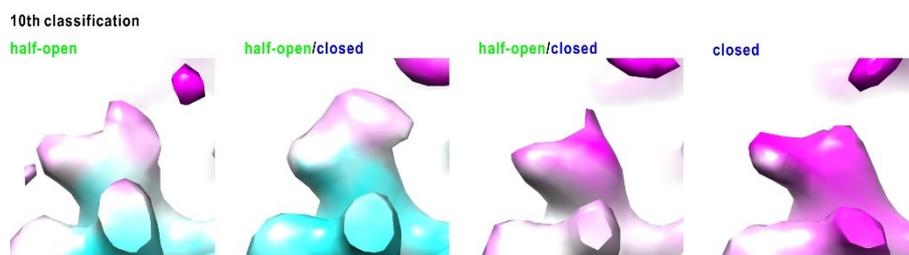
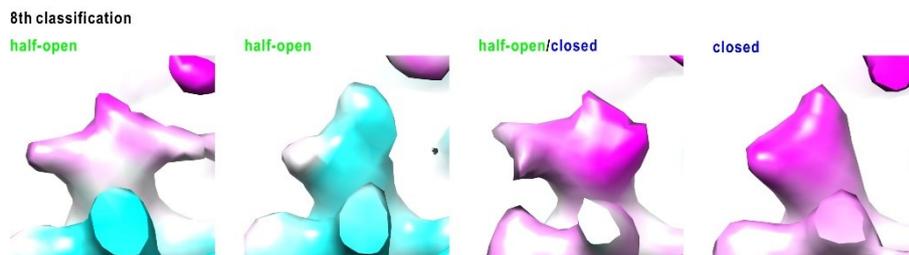
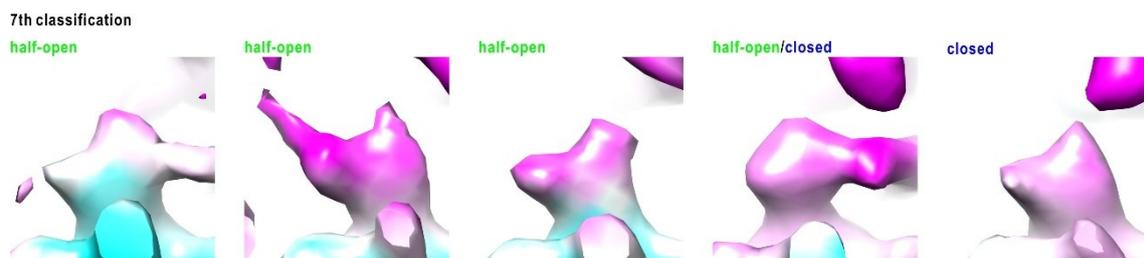


5th classification



6th classification





付録 C

cryoEM 観察像からの自由エネルギー地形解析方法

本研究では MD 計算で得られた構造を利用して cryoEM 観察像から自由エネルギー地形を推定する方法を開発し、GDH ドメイン運動の自由エネルギー地形の解析に適用した。本付録では cryoEM マップと MD マップの類似度から cryoEM マップにおける MD マップの重みを計算する方法の詳細を記す。

以下では逆空間上で議論する。本文 7.3.3 節の(7.3)式より、cryoEM マップ \bar{v} 及びあるコンフォメーション $i (i=1-K)$ の MD マップ \bar{v}^i の間の類似度を

$$\exp(-Z_i), \quad Z_i = \sum_l \frac{|V_l - V_l^i|^2}{2\sigma_l^2} \quad (\text{C1})$$

で定義する。ここで、 V_l 、 V_l^i はそれぞれ \bar{v} 、 \bar{v}^i のボクセル l における値、 σ_l^2 は \bar{v} のボクセル l におけるノイズの分散である。cryoEM マップを MD マップの線形結合で近似できると仮定すると、 V_l は

$$V_l = \sum_k \omega_k V_l^k, \quad \sum_k \omega_k = 1 \quad (\text{C2})$$

と書き換えられる。 V_l^k はコンフォメーション $k (k=1-K)$ である MD マップ \bar{v}^k のボクセル l における値である。(C2)式を(B1)式に代入すると

$$Z_i = \sum_l \frac{\left| \sum_k \omega_k V_l^k - V_l^i \right|^2}{2\sigma_l^2} \quad (\text{C3})$$

となる。 ω_k の規格化条件および半角公式より、(C3)式の分子部分は以下のように変形できる。

$$\begin{aligned} \left| \sum_k \omega_k V_l^k - V_l^i \right|^2 &= \left| \sum_k \omega_k \Delta V_l^{ki} \right|^2, \quad \Delta V_l^{ki} = V_l^k - V_l^i \\ &= \sum_{k_1}^K \sum_{k_2}^K \omega_{k_1} \omega_{k_2} \left| \Delta V_l^{k_1 i} \right| \left| \Delta V_l^{k_2 i} \right| \cos(\alpha_l^{k_1 i} - \alpha_l^{k_2 i}) \\ &= \sum_{k_1}^K \sum_{k_2}^K \omega_{k_1} \omega_{k_2} \frac{1}{2} \left\{ \left| \Delta V_l^{k_1 i} \right|^2 + \left| \Delta V_l^{k_2 i} \right|^2 - \left| \Delta V_l^{k_1 i} - \Delta V_l^{k_2 i} \right|^2 \right\} \\ &= \sum_{k_1}^K \sum_{k_2}^K \omega_{k_1} \omega_{k_2} \frac{1}{2} \left\{ \left| \Delta V_l^{k_1 i} \right|^2 + \left| \Delta V_l^{k_2 i} \right|^2 - \left| \Delta V_l^{k_1 k_2} \right|^2 \right\} \end{aligned} \quad (\text{C4})$$

ただし、 α_i^{ki} は ΔV_l^{ki} の位相である。(C4)式より、 Z_i は

$$Z_i = \sum_l \frac{1}{2\sigma_l^2} \sum_{k_1}^K \sum_{k_2}^K \omega_{k_1} \omega_{k_2} \frac{1}{2} \left\{ \left| \Delta V_l^{k_1 i} \right|^2 + \left| \Delta V_l^{k_2 i} \right|^2 \right\} - \sum_l \frac{1}{2\sigma_l^2} \sum_{k_1}^K \sum_{k_2}^K \omega_{k_1} \omega_{k_2} \frac{1}{2} \left| \Delta V_l^{k_1 k_2} \right|^2 \quad (C5)$$

と変形される。(C5)式の第二項は i について独立であり、すべての i に対して同じ値をとる。そのため、以下で定数 C として扱う。また、第一項の k_1 、 k_2 についての輪の部分は ω_k の規格化条件によって

$$\begin{aligned} \sum_{k_1}^K \sum_{k_2}^K \omega_{k_1} \omega_{k_2} \frac{1}{2} \left\{ \left| \Delta V_l^{k_1 i} \right|^2 + \left| \Delta V_l^{k_2 i} \right|^2 \right\} &= \sum_{k_1}^K \omega_{k_1} \frac{1}{2} \left| \Delta V_l^{k_1 i} \right|^2 \sum_{k_2}^K \omega_{k_2} \\ &\quad + \sum_{k_1}^K \omega_{k_1} \sum_{k_2}^K \omega_{k_2} \frac{1}{2} \left| \Delta V_l^{k_2 i} \right|^2 \\ &= \sum_k^K \omega_k \left| \Delta V_l^{ki} \right|^2 \end{aligned} \quad (C6)$$

の形にまとめられる。 l と k の和の位置を入れ替えると、(C5)式は

$$Z_i = \sum_k^K \omega_k a_i^k - C, \quad a_i^k = \sum_l \frac{\left| \Delta V_l^{ki} \right|^2}{2\sigma_l^2} \quad (C7)$$

となる。 C は未知の値であるが、上で述べた通り i に対して一定値をとる。そのため、 $1-K$ の中のあるコンフォメーション m と cryoEM との類似度 Z_m を Z_i から差し引いて

$$Y_i = Z_i - Z_m = \sum_k^K \omega_k (a_i^k - a_m^k) \quad (C8)$$

として新たな量 Y_i に変換することで除去できる。このようにして得られた Y_i ($i = 1-K$) は行列及び列ベクトルによって

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_m = 0 \\ \vdots \\ Y_K \end{pmatrix} = \begin{pmatrix} -a_m^1 & \cdots & a_1^m & \cdots & a_1^K - a_m^K \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_K^1 - a_m^1 & \cdots & a_K^m & \cdots & -a_m^K \end{pmatrix} \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_m \\ \vdots \\ \omega_K \end{pmatrix} \quad (C9)$$

と表現できる。 a^k の行列について逆行列が求められれば重みを求められるが、この行列は m 行目の成分がすべてゼロであり正則でなく、従って逆行列が存在しない。そこで、 ω_k の規格化条件に基づいて(C9)式の m 行目を以下のように置き換える。

$$\begin{pmatrix} Y_1 \\ \vdots \\ 1 \\ \vdots \\ Y_K \end{pmatrix} = \begin{pmatrix} -a_m^1 & \cdots & a_1^m & \cdots & a_1^K - a_m^K \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & \cdots & 1 & \cdots & 1 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_K^1 - a_m^1 & \cdots & a_K^m & \cdots & -a_m^K \end{pmatrix} \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_m \\ \vdots \\ \omega_K \end{pmatrix} \quad (\text{C10})$$

この置換によって a_i^k に関する行列が正則となり、逆行列及び重みが求められる。

謝辞

本研究は多くの方々のご指導・ご協力に支えられて行われました。

SPring-8 BL45XU での SAXS 測定は、次の方々と共同で実施されました。理化学研究所の山本雅貴利用システム開発研究部部長、引間孝明博士、兵庫県立大学の高山裕貴助教、大阪府立大学の徳富哲客員教授、嘉祥寺谷幸子氏(2014 年度博士卒)、慶應義塾大学の中迫雅由教授、荳口友隆専任講師、岡島公司特任助教、関口優希氏(2016 年度博士卒)、小林周氏(2017 年度博士)、中島真氏(2012 年度学士卒)、福田朝陽氏(2016 年度修士卒)らに深く感謝いたします。

cryoEM 観察は大阪大学大学院生命機能研究科難波研究室の設備を使用させていただきました。難波啓一教授(現特任教授)、加藤貴之特任准教授(現大阪大学蛋白質研究所教授)、並びに難波研究室の皆様にご深く感謝いたします。特に加藤教授には試料作成から測定、解析に至るまで多くの面でご指導、ご支援いただくと共に様々な議論をさせていただきました。

慶應義塾大学の岡島公司特任助教には大阪府立大学在籍時から測定試料の精製をしていただき、測定も共同で実施させていただきました。また、慶應義塾大学に移籍されてから試料の精製や測定だけでなく、解析や論文の議論、研究のご指導など多岐にわたってお世話になり、論文の執筆にあたって貴重な助言をくださいました。

加藤教授、岡島特任助教のご支援、ご指導なくしては本研究の遂行は不可能でした。お二人に心より感謝申し上げます。

慶應義塾大学物理学科の能崎幸雄教授、同学科の渡邊紳一教授、物理情報工学科の牧英之准教授にはご多忙の中、本博士論文の副査を引き受けていただきました。深く感謝いたします。

理化学研究所の山本雅貴部門長には SAXS 測定についてアドバイスをいただき、また私を理化学研究所研修生として受け入れていただくなど多岐にわたりお世話になりました。同じく、理化学研究所の引間孝明博士には、光学系の調整や試料窓の交換など SAXS 測定を支援していただくだけでなく、測定に際しても測定方法から解析方法まで丁寧に説明いただき、また様々なことにご対応いただきました。

筑波大学の嘉祥寺谷幸子博士には測定試料の精製をしていただき、また測定においても試料交換をしていただくなどお世話になりました。

慶應義塾大学の荳口友隆専任講師には GDH の MD 計算トラジェクトリーをご提供

いただただけでなく、様々な議論を通して研究に対する理解を深めることができました。

最後になりますが、指導教員の中迫雅由教授には、私を研究室へ迎えていただただけでなく、本研究を進めるにあたり様々な助言、熱心な指導を賜りました。本論文をはじめとした研究論文の執筆にあたって多くの叱咤激励と共に推敲をしていただき、書き上げることができました。箸にも棒にも掛からぬ若造であった私がここまで至ることができたのは偏に中迫教授のご指導のおかげでございます。

ご指導・ご鞭撻と共に私を支えてくださった皆様と、大学院への進学を支えてくださった家族に深く感謝いたします。

参考文献

- Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echol, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. and Zwart, P. H. (2010) “*PHENIX*: a comprehensive Python-based system for macromolecular structure solution”, *Acta Crystallogr.* **D66**, 213-221.
- Adrian, M., Dubochet, J., Lepault, J. and McDowell, A. W. (1984) “Cryo-electron microscopy of viruses”, *Nature* **308**, 32-36.
- Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A. and Adams, P. D. (2018) “Real-space refinement in *PHENIX* for cryo-EM and crystallography” *Acta Crystallogr.* **D74**, 531-544.
- Akamine, P., Madhusudan, Wu, J., Xuong, N. H., Ten Eyck, L. F. and Taylor, S. S. (2003) “Dynamic features of cAMP-dependent protein kinase revealed by apoenzyme crystal structure”, *J. Mol. Biol.* **327**, 159-171.
- Als-Nielsen, J. and McMorrow, D. (2011) *Elements of Modern X-ray Physics 2nd edition*, Wiley, New York.
- Bai, X. C., Rajendra, E., Yang, G., Shi, Y. and Scheres, S. H. W. (2015) “Sampling the conformational space of the catalytic subunit of human γ -secretase”, *eLife* **4**, 11182 (19 pages).
- Baker, L. A., Smith, E. A., Bueler, S. A. and Rubinstein, J. L. (2010) “The resolution dependence of optimal exposures in liquid nitrogen temperature electron cryomicroscopy of catalase crystals”, *J. Struct. Biol.* **169**, 431-437.
- Beauchamp, K. A., Lin, Y., Das, R. and Pande, V. S. (2012) “Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements” *J. Chem. Theory Comput.* **8**, 1409-1414.
- Bellini, D. and Papiz, M. Z. (2012) “Structure of a Bacteriophytochrome and Light-Stimulated Protomer Swapping with a Gene Repressor” *Structure* **20**, 1436-1446.
- Bhat, T. N., Bentley, G. A., Boulot, G., Greene, M. I., Tello, D., Dall’Acqua, W., Souchon, H., Schwarz, F. P., Mariuzza, R. A. and Poljak, R. J. (1994) “Bound water molecules and conformational stabilization help mediate an antigen-antibody association”, *Proc. Natl. Acad. Sci. USA* **91**, 1089-1093.
- Bishop, C. M., Svensén, M. and Williams, C. K. I. (1998) “GTM: The generative topographic mapping”, *Neural Comput.* **10**, 215-234.

- Borthwick, H. A., Hendricks, S. B., Parker, M. W., Toole, E. H. and Toole, V. K. (1952) “A Reversible Photoreaction Controlling Seed Germination”, *Proc. Natl. Acad. Sci. USA* **38**, 662-666.
- Brilot, A. F. , Chen, J. Z., Cheng, A. , Pan, J., Harrison, S. C., Potter, C. S., Carragher, B., Henderson, R. and Grigorieff, N. (2012) “Beam-induced motion of vitrified specimen on holey carbon film”, *J. Struct. Biol.* **177**, 630–637.
- Burgie, E. S., Bussell, A. N., Walker, J. M., Dubiel, K. and Vierstra, R. D. (2014a) “Crystal structure of the photosensing module from a red/far-red light-absorbing plant phytochrome”, *Proc. Natl. Acad. Sci. USA* **111**, 10179-10184.
- Burgie, E. S, Zhang, J. and Vierstra, R. D. (2016) “Crystal structure of *Deinococcus* phytochrome in the photoactivated state reveals a cascade of structural rearrangements during photoconversion”, *Structure* **24**, 448-457.
- Butler, W. L., Norris, K. H., Siegelman, H. W. and Hendricks, S. B. (1959) “Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants”, *Proc. Natl. Acad. Sci. USA* **45**, 1703-1708.
- Campbell, M. G., Cheng, A., Brilot, A. F., Moeller, A., Lyumkis, D., Veesler, D., Pan, J., Harrison, S. C., Potter, C. S., Carragher, B. and Grigorieff, N. (2012) “Movies of ice-embedded particles enhance resolution in electron cryo-microscopy”, *Structure* **20**, 1823-8.
- Cardone, G., Heymann, J. B. and Steven, A. C. (2013) “One number does not fit all: Mapping local variations in resolution in cryo-EM reconstructions”, *J. Struct. Biol.* **184**, 226-236.
- Chen, S., McMullan, G., Faruqi, A. R., Murshudov, G. N., Short, J. M., Scheres, S. H. W. and Henderson, R. (2013) “High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy”, *Ultramicroscopy* **135**, 24–35.
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003) “Multiple sequence alignment with the Clustal series of programs”, *Nucleic Acids Res.* **31**, 3497-3500.
- Cho, H. Y., Tseng, T. S., Kaiserli, E., Sullivan, S., Christie, J. M., and Briggs, W. R. (2007) “Physiological roles of the light, oxygen, or voltage domains of phototropin 1 and phototropin 2 in *Arabidopsis*”, *Plant Physiol.* **143**, 517-529.
- Choudry, U., Heyes, D. J., Hardman, S. J. O., Sakuma, M., Sazanovich, I. V., Woodhouse, J., De La Mora, E., Pedersen, M. N., Wulff, M., Weik, M., Schirò, G. and Scrutton, N. S. (2018) “Photochemical Mechanism of an Atypical Algal Phytochrome” *ChemBioChem* **19**, 1036-1043.

- Christie, J. M. (2007) "Phototropin blue-light receptors", *Annu. Rev. Plant Biol.* **58**, 21-45.
- Christie, J. M., Hitomi, K., Arvai, A. S., Hartfield, K. A., Mettlen, M., Pratt, A. J., Tainer, J. A. and Getzoff, E. D. (2012) "Structural tuning of the fluorescent protein iLOV for improved photostability", *J. Biol. Chem.* **287**, 22295-22304.
- Christie, J. M., Reymond, P., Powell, G. K., Bernasconi, P., Raibekas, A., Liscum, E. and Briggs, W. R. (1998) "*Arabidopsis* NPH1: A flavoprotein with the properties of a photoreceptor for phototropism", *Science* **282**, 1698-1701.
- Christie, J. M., Salomon, M., Nozue, K., Wada, M. and Briggs, W. R. (1999) "LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): binding sites for the chromophore flavin mononucleotide", *Proc. Natl. Acad. Sci. USA* **96**, 8779-8783.
- Christie, J. M., Swartz, T. E., Bogomolni, R. A. and Briggs, W. R. (2002) "Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function", *Plant J.* **32**, 205-219.
- Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F. and Zucker, S. W. (2005) "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps", *Proc. Natl. Acad. Sci. U. S. A.* **102**, 7426-7431.
- Crosson, S. and Moffat, K. (2002) "Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch", *Plant Cell.* **14**, 1067-1075.
- Crosson, S. and Moffat, K. (2001) "Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction", *Proc. Natl. Acad. Sci. USA* **98**, 2995-3000.
- Crowther, R. A., DeRosier, D. J. and Klug, A. (1970) "The reconstruction of a three-dimensional structure from projections and its application to electron microscopy", *Proc. R. Soc. Lond.* **317**, 319-340.
- Dashti, A., Schwander, P., Langlois, R., Fung, R., Li, W., Hosseinizadeh, A., Liao, H. Y., Pallesen, J., Sharma, G., Stupina, V. A., Simon, A. E., Dinman, J. D., Frank, J. and Ourmazd, A. (2014) "Trajectories of the ribosome as a Brownian nanomachine", *Proc. Natl. Acad. Sci. U. S. A.* **111**, 17492-17497.
- D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M. and Kühlbrandt, W. "Protein denaturation at the air-water interface and how to prevent it", *eLife* **1**, 8 (18 pages).
- Drenth, J. (2007) *Principles of Protein X-ray Crystallography Third Edition*, Springer, New York.
- Dubochet, J., Adrian, M., Chang, J. J., Homo, J. C., Lepault, J., McDowell, A. W. and Schultz, P. (1988) "Cryo-electron microscopy of vitrified specimens", *Q. Rev. Biophys.* **21**, 129-228.
- Dubochet, J. and McDowell, A. W. (1981) "Vitrification of pure water for electron microscopy", *J. Microsc.* **124**, 3-4.
- Edgerton, M. D. and Jones, A. M. (1992) "Localization of protein-protein interactions between

- subunits of phytochrome”, *Plant Cell* **4**, 161-171.
- Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010) “Features and Development of Coot”, *Acta Crystallogr.* **D66**, 486-501.
- Evans, K., Grossmann, J. G., Fordham-Skelton, A. P. and Papiz, M. Z. (2004) “Small-angle X-ray scattering reveals the solution structure of a bacteriophytochrome in the catalytically active Pr state”, *J. Mol. Biol.* **364**, 655-666.
- Feigin, L. A. and Svergun, D. I. (1987) *Structure Analysis by Small-Angle X-Ray and Neutron Scattering*, Plenum Press, New York.
- Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I. and Craievich, A. F. (2010) “Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale”, *J. Appl. Cryst.* **43**, 101-109.
- Frank, J. (2006) *Three-Dimensional Electron Microscopy of Macromolecular Assemblies*, Oxford University Press, U.K.
- Furuya, M. (1993) “Phytochromes: Their molecular species, gene families, and functions”, *Annu. Rev. Plant. Physiol. Plant Mol. Biol.* **44**, 617-645.
- Grant, T. and Grigorieff, N. (2015) “Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6”, *eLife* **4**, e06980 (19 pages).
- Guinier, A. and Fournet, G. (1955) *Small-Angle Scattering of X-rays*, Wiley, New York.
- Halavaty, A. S. and Moffat, K. (2007) “N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from *Avena sativa*”, *Biochemistry* **46**, 14001-14009.
- Halavaty, A. S. and Moffat, K. (2013) “Coiled-coil dimerization of the LOV2 domain of the blue-light photoreceptor phototropin 1 from *Arabidopsis thaliana*”, *Acta Cryst.* **F69**, 1316-1321.
- Hammersley, A. P. (1997) “FIT2D: An Introduction and Overview”, ESRF Internal Report ESRF97HA02T.
- Hammersley, A. P., Svensson, S. O., Hanfland, M., Fitch, A. N. and Häusermann, D. (1996) “Two-dimensional detector software: from real detector to idealized image or two-theta scan”, *High Press. Res.* **14**, 235-248.
- Harada, R. and Kitao, A. (2011) “Nontargeted parallel cascade selection molecular dynamics for enhancing the conformational sampling of proteins”, *J. Chem. Theory Comp.* **8**, 290-299.
- Harper, S. M., Neil, L. C. and Gardner, K. H. (2003) “Structural basis of a phototropin light switch”, *Science* **301**, 1541-1544.
- Haselbach, D., Komarov, I., Agafonov, D. E., Hartmuth, K., Graf, B., Dybkov, O., Urlaub, H., Kastner, B., Lührmann, R. and Stark, H. (2018) “Structure and conformational dynamics of

- the human spliceosomal B^{act} complex”, *Cell* **172**, 1-11.
- Henderson, R. and Unwin, P. N. (1975) “Three-dimensional model of purple membrane obtained by electron microscopy”, *Nature* **257**, 28-32.
- Heyes, D. J., Hardman, S. J. O., Pedersen, M. N., Woodhouse, J., De La Mora, E., Wulff, M., Weik, M., Cammarata, M., Scrutton, N. S. and Schirò, G. (2019) “Light-induced structural changes in a full-length cyanobacterial phytochrome probed by time-resolved X-ray scattering” *Commun. Biol.* **2**, 1.
- Ho, Y. S., Burden, L. M. and Hurley, J. H. (2000) “Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor” *EMBO J.* **19**, 5288-5299.
- Huala, E., Oeller, P. W., Liscum, E., Han, I. S., Larsen, E. and Briggs, W. R. “Arabidopsis NPH1: a protein kinase with a putative redox-sensing domain”, *Science* **278**, 2120-2123.
- Huang, K. and Beck, C. F. (2003) “Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga *Chlamydomonas reinhardtii*”, *Proc. Natl. Acad. Sci. USA* **100**, 6269-6274.
- Huttenlocher, D. P., Klanderma, G. A. and Rucklidge, W. J. (1993) “Comparing images using the Hausdorff distance”, *IEEE Trans. Pattern Anal. Mach. Intell.* **15**, 850-863.
- Ibel, K. and Stuhrman, H. B. (1975) “Comparison of neutron and X-ray scattering of dilute myoglobin solutions”, *J. Mol. Biol.* **93**, 255-265.
- Im, C. S., Eberhard, S., Huang, K., Beck, C. F. and Grossman, A. R. (2006) “Phototropin involvement in the expression of genes encoding chlorophyll and carotenoid biosynthesis enzymes and LHC apoproteins in *Chlamydomonas reinhardtii*”, *Plant J.* **48**, 1-16.
- Inoue, S., Kinoshita, T., Matsumoto, M., Nakayama, K. I., Doi, M. and Shimazaki, K. (2008) “Blue light-induced autophosphorylation of phototropin is a primary step for signaling”, *Proc. Natl. Acad. Sci. USA* **105**, 5626-5631.
- Jarillo, G. A., Gabrys, H., Capel, J., Alonso, J. M., Ecker, J. R. and Cashmore, A. R. (2001) “Phototropin-related NPL1 controls chloroplast relocation induced by blue light”, *Nature* **410**, 952-954.
- Kagawa, T., Sakai, T., Suetsugu, N., Oikawa, K., Ishiguro, S., Kato, T., Tabata, S., Okada, K. and Wada, M. (2001) “*Arabidopsis* NPL1: A phototropin homolog controlling the chloroplast high-light avoidance response”, *Science* **291**, 2138-2141.
- Kasahara, M., Swartz, T. E., Olney, M. A., Onodera, A., Mochizuki, N., Fukuzawa, H., Asamizu, E., Tabata, S., Kanegae, H., Takano, M., Christie, J. M., Nagatani, A. and Briggs, W. R. (2002) “Photochemical properties of the flavin mononucleotide-binding domains of the phototropins from *Arabidopsis*, rice, and *Chlamydomonas reinhardtii*”, *Plant Physiol.* **129**, 762-773.

- Kashojiya, S., Okajima, K., Shimada, T. and Tokutomi, S. (2015) “Essential role of the A α /A β gap in the N-terminal upstream of LOV2 for the blue light signaling from LOV2 to kinase in *Arabidopsis phototropin1*, a plant blue light receptor”, *PLoS One* **10**, e0124284.
- Kashojiya, S., Yoshihara, S., Okajima, K. and Tokutomi, S. (2016) “The linker between LOV2-J α and STK plays an essential role in the kinase activation by blue light in *Arabidopsis phototropin1*, a plant blue light receptor”, *FEBS Lett.* **590**, 139-147.
- Kimanius, D., Forsberg, B. O., Scheres, S. H. W. and Lindahl, E. (2016) “Accelerated cryo-EM structure determination with parallelization using GPUs in RELION-2”, *eLife* **5**, e18722 (21 pages).
- Kinoshita, T., Doi, M., Suetsugu, N., Kagawa, T., Wada, M. and Shimazaki, K. (2001) “Phot1 and phot2 mediate blue light regulation of stomatal opening”, *Nature* **414**, 656-660.
- Kinoshita, T. and Shimazaki, K. (1999) “Blue light activates the plasma membrane H⁺-ATPase by phosphorylation of the C-terminus in stomatal guard cells”, *EMBO J.* **18**, 5548-5558.
- Kircher, S., Kozma-Bognar, L., Kim, L., Adam, E., Harter, K., Schäfer, E. and Nagy, F. (1999) “Light quality-dependent nuclear import of the plant photoreceptors phytochrome A and B”, *Plant Cell* **11**, 1445–1456.
- Kong, S.G. and Wada, M. (2011) “New insights into dynamic actin-based chloroplast photorelocation movement”, *Mol. Plant* **4**, 771-781.
- Kornev, A. P., Haste, N. M., Taylor, S. S. and Eyck, L. F. (2006) “Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism”, *Proc. Natl. Acad. Sci. USA* **103**, 17783-17788.
- Kozin, M. and Svergun, D. I. (2001) “Automated matching of high- and low-resolution structural models”, *J. Appl. Cryst.* **34**, 33-41.
- Li, X., Mooney, P., Zheng, S., Booth, C. R., Braunfeld, M. B., Gubbens, S., Agard, D.A. and Cheng, Y. (2013) “Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM”, *Nat. Met.* **10**, 584–590.
- Li, H., Zhang, J., Vierstra, R. D. and Li, H. (2010) “Quaternary organization of a phytochrome dimer as revealed by cryoelectron microscopy”, *Proc. Natl. Acad. Sci. USA* **107**, 10872-10877.
- Liscum, E. and Briggs, W. R. “Mutations in the NPH1 locus of *Arabidopsis* disrupt the perception of phototropic stimuli”, *Plant Cell*. **7**, 473-85.
- Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P. and Lanyi, J. K. (1999) “Structural changes in bacteriorhodopsin during ion transport at 2 angstrom resolution”, *Science* **286**, 255-261.
- Mathews, S. and Sharrock, R. A. (1997) “Phytochrome gene diversity” *Plant Cell Environ* **20**,

666-671.

- Matsuoka, D. and Nakasako, M. (2009) "Probability distributions of hydration water molecules around polar protein atoms obtained by a database analysis", *J. Phys. Chem. B* **113**, 11274-11292.
- Matsuoka, D. and Nakasako, M. (2010) "Prediction of hydration structures around hydrophilic surfaces of proteins by using the empirical hydration distribution functions from a database analysis" *J. Phys. Chem. B* **114**, 4652-4663.
- Matsushita, T., Mochizuki, N. and Nagatani, A. (2003) "Dimers of the N-terminal domain of phytochrome B are functional in the nucleus", *Nature* **424**, 571-574.
- Merk, A., Bartesaghi, A., Banerjee, S., Falconieri, V., Rao, P., Davis, M. I., Pragani, R., Boxer, M. B., Earl, L. A., Milne, J. L. S. and Subramaniam S. (2016) "Breaking Cryo-EM Resolution Barriersto Facilitate Drug Discovery", *Cell* **165**, 1698-1707.
- Möglich, A., Ayers, R. A. and Moffat, K. (2009) "Structure and signaling mechanism of Per-ARNT-Sim domains", *Structure* **17**, 1282-1294.
- Möglich, A., Yang, X., Ayers, R. A. and Moffat, K. (2010) "Structure and function of plant photoreceptors", *Annu. Rev. Plant Biol.* **61**, 21-47.
- Montgomery, B. L. and Lagarias, J. C. (2002) "Phytochrome ancestry: sensors of bilins and light" *Trends Plant Sci* **7**, 357-366.
- Mroginski, M. A., Murgida, D. H. and Hildebrandt, P. (2007) "The chromophore structural changes during the photocycle of phytochrome: a combined resonance Raman and quantum chemical approach", *Acc. Chem. Res.* **40**, 258-266.
- Nagatani, A., Reed, J. W. and Chory, J. (1993) "Isolation and Initial Characterization of Arabidopsis Mutants That Are Deficient in Phytochrome A", *Plant Physiol.* **102**, 269-277.
- Nakane, T., Kimanius, D., Lindahl, E. and Scheres, S. H. W. (2018) "Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION", *eLife* **7**, 36861 (18 pages).
- Nakanishi, A., Kishikawa, J., Tamakoshi, M., Mitsuoka, K. and Yokoyama, Y. (2018) "Cryo EM structure of intact rotary H⁺-ATPase/synthase from *Thermus thermophilus*", *Nat. Comm.* **9**, 89 (10 pages).
- Nakasako, M., Fujisawa, T., Adachi, S., Kudo, T. and Higuchi, S. (2001) "Large-scale domain movements and hydration structure changes in the active-site cleft of unligated glutamate dehydrogenase from *Thermococcus profundus* studied by cryogenic X-ray crystal structure analysis and small-angle X-ray scattering", *Biochemistry* **40**, 3069-3079.
- Nakasako, M., Iwata, T., Matsuoka, D. and Tokutomi, S. (2004) "Light-induced structural

- changes of LOV domain-containing polypeptides from *Arabidopsis* phototropin 1 and 2 studied by small-angle X-ray scattering”, *Biochemistry* **43**, 14881-14890.
- Nakasako, M., Maeno, A., Kurimoto, E., Harada, T., Yamaguchi, Y., Oka, T., Takayama, Y., Iwata, A. and Kato, K. (2010) “Redox-dependent domain rearrangement of protein disulfide isomerase from a thermophilic fungus”, *Biochemistry* **49**, 6953-6962.
- Nakasako, M., Odaka, M., Yohda, M., Dohmae, N., Takio, K., Kamiya, N. and Endo, I. (1999) “Tertiary and quaternary structures of photoreactive Fe-type nitrile hydratase from *Rhodococcus* sp. N-771: roles of hydration water molecules in stabilizing the structures and the structural origin of the substrate specificity of the enzyme”, *Biochemistry* **38**, 9887-9898.
- Nakasako, M., Wada, M., Tokutomi, S., Yamamoto, K. T., Sakai, J., Kataoka, M., Tokunaga, F. and Furuya, M. (1990) “Quaternary structure of pea phytochrome I dimmer studied with small-angle X-ray scattering and rotary-shadowing electron microscopy” *Photochem Photobiol* **52**, 3-12.
- Nakasako, M., Zikihara, K., Matsuoka, D., Katsura, H. and Tokutomi, S. (2008) “Structural basis of the LOV1 dimerization of *Arabidopsis* phototropins 1 and 2”, *J. Mol. Biol.* **381**, 718-733.
- Nakazawa, M., Yoshida, Y. and Manabe, K. (1991) “Differences between the surface properties of the P_R and the P_{FR} forms of native pea phytochrome, and their application to a simplified procedure for purification of the phytochrome”, *Plant Cell Physiol.* **32**, 1187-1194.
- Oide, M., Hikima, T., Oroguchi, T., Kato, T., Yamaguchi, Y., Yoshihara, S., Yamamoto, M., Nakasako, M. and Okajima, K. (2019) “Molecular shape under far-red light and light-induced association of *Arabidopsis* phytochrome B”, *FEBS J.* 15095 (14 pages).
- Oide, M., Okajima, K., Kashojiya, S., Takayama, Y., Oroguchi, T., Hikima, T., Yamamoto, M. and Nakasako, N. (2016) “Blue light-excited Light-Oxygen-Voltage-sensing domain 2 (LOV2) triggers a rearrangement of the kinase domain to induce phosphorylation activity in *Arabidopsis* phototropin 1”, *J. Biol. Chem.* **291**, 19975-19984.
- Oide, M., Okajima, K., Nakagami, H., Kato, T., Sekiguchi, Y., Oroguchi, T., Hikima, T., Yamamoto, M. and Nakasako, M. (2018a) “Blue light-excited LOV1 and LOV2 domains cooperatively regulate the kinase activity of full-length phototropin 2 from *Arabidopsis*”, *J. Biol. Chem.* **293**, 963-972.
- Oide, M., Sekiguchi, Y., Fukuda, A., Okajima, K., Oroguchi, T. and Nakasako, M. (2018b) “Classification of *ab initio* models of proteins restored from small-angle X-ray scattering”, *J. Synchrotron Rad.* **25**, 1379-1388.
- Okajima, K. (2016) “Molecular mechanism of phototropin light signaling”, *J. Plant Res.* **129**, 149-157.

- Okajima, K., Aihara, Y., Takayama, Y., Nakajima, M., Kashojiya, S., Hikima, T., Oroguchi, T., Kobayashi, A., Sekiguchi, Y., Yamamoto, M., Suzuki, T., Nagatani, A., Nakasako, M. and Tokutomi, S. (2014) “Light-induced conformational changes of LOV1 (light oxygen voltage-sensing domain 1) and LOV2 relative to the kinase domain and regulation of kinase activity in *Chlamydomonas* phototropin”, *J. Biol. Chem.* **289**, 413-422.
- Oroguchi, T., Hashimoto, H., Shimizu, T., Sato, M. and Ikeguchi, M. (2009) “Intrinsic dynamics of restriction endonuclease *Eco*O109I studied by molecular dynamics simulations and X-ray scattering data analysis”, *Biophys. J.* **96**, 2808-2822.
- Oroguchi, T. and Nakasako, M. (2016) “Changes in hydration structure are necessary for collective motions of a multi-domain protein”, *Sci. Rep.* **6**, 26302 (14 pages).
- Petoukhov, M.V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D., Konarev, P. V. and Svergun, D. I. (2012) “New developments in the ATSAS program package for small-angle scattering data analysis”, *J. Appl. Crystallogr.* **45**, 342-350.
- Petoukhov, M. V., Konarev, P. V., Kikhney, A. G. and Svergun, D. I. (2007) “ATSAS 2.1 – towards automated and web-supported small-angle scattering data analysis”, *J. Appl. Crystallogr.* **40**, 223-228.
- Petoukhov, M. V. and Svergun, D. I. (2015) “Ambiguity assessment of small-angle scattering curves from monodisperse systems”, *Acta Cryst.* **D71**, 1051-1058.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. and Ferrin, T. E. (2004) “UCSF Chimera—a visualization system for exploratory research and analysis”, *J. Comput. Chem.* **25**, 1605–1612.
- Pfeifer, A., Majerus, T., Zikihara, K., Matsuoka, D., Tokutomi, S., Heberle, J. and Kottke, T. (2009) “Time-resolved Fourier transform infrared study on photoadduct formation and secondary structural changes within the phototropin LOV domain”, *Biophys. J.* **96**, 1462-1470.
- Pfeifer, A., Mathes, T., Lu, Y., Hegemann, P. and Kottke, T. (2010) “Blue light induces global and localized conformational changes in the kinase domain of full-length phototropin”, *Biochemistry* **49**, 1024-1032.
- Punjani, A., Rubinstein, J. L., Fleet, D. J. and Brubaker, M. A. (2017) “cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination” *Nat. Methods.* **14**, 290-296.
- Quail, P. H. (2002) “Phytochrome photosensory signalling networks”, *Nat. Rev. Mol. Cell. Biol.* **3**, 85-93.
- Quail, P. H. (2007) “Phytochrome-regulated Gene Expression”, *J. Integrative Plant Biol.* **49**, 11–20.
- Quail, P. H., Boylan, M. T., Parks, B. M., Short, T. W., Xu, Y. and Wagner, D. (1995) “Phytochromes: photosensory perception and signal transduction”, *Science* **268**, 675-680.

- Reed, J. W., Nagpal, P., Poole, D. S., Furuya, M. and Chory, J. (1993) “Mutations in the gene for the red/far-red light receptor phytochrome B alter cell elongation and physiological responses throughout Arabidopsis development”, *Plant Cell* **5**, 147-157.
- Remberg, A., Ruddat, A., Braslavsky, S. E., Gärtner, W. and Schaffner, K. (1998) “Chromophore incorporation, Pr to Pfr kinetics, and Pfr thermal reversion of recombinant N-terminal fragments of phytochrome A and B chromoproteins”, *Biochemistry* **37**, 9983-9990.
- Rockwell, N. C., Njuguna, S. L., Roberts, L., Castillo, E., Parson, V. L., Dwojak, S., Lagarias, J.C. and Spiller, S. C. (2008) “A second conserved GAF domain cysteine is required for the blue/green photoreversibility of cyanobacteriochrome Tlr0924 from *Thermosynechococcus elongatus*”, *Biochemistry* **47**, 7304-7316.
- Rohou, A. and Grigorieff, N. (2015) “CTFFIND4: Fast and accurate defocus estimation from electron micrographs”, *J. Struct. Biol.* **192**, 216-221.
- Rosenthal, P. B. and Henderson, R. (2003) “Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy”, *J. Mol. Biol.* **333**, 721–745.
- Sakai, T., Kagawa, T., Kasahara, M., Swartz, T., Christie, J. M., Briggs, W. R., Wada, M. and Okada, K. (2001) “Photochemical properties of the Flavin mononucleotide-binding domains of the phototropins from *Arabidopsis*, Rice, and *Chlamydomonas reinhardtii*”, *Proc. Natl. Acad. Sci. USA* **98**, 6969-6974.
- Sakamoto, K. and Nagatani, A. (1996) “Nuclear localization activity of phytochrome B”, *Plant J.* **10**, 859-868.
- Salomon, M., Christie, J. M., Knieb, E., Lempert, U. and Briggs, W. R. (2000) “Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor, phototropin”, *Biochemistry* **39**, 9401-9410.
- Saxton, W. O., and Baumeister, W. (1982) “The correlation averaging of a regularly arranged bacterial cell envelope protein”, *J. Microsc.* **127**, 127-138.
- Scheres, S. H. W. (2012a) “A Bayesian view on cryo-EM structure determination”, *J. Mol. Biol.* **415**, 406–418.
- Scheres, S. H. W. (2012b) “RELION: Implementation of a Bayesian approach to cryo-EM structure determination”, *J. Struct. Biol.* **180**, 519-530.
- Scheres, S. H. W. (2015) “Semi-automated selection of cryo-EM particles in RELION-1.3”, *J. Struct. Biol.* **189**, 114-122.
- Scheres, S. H. W. and Chen, S. (2012) “Prevention of overfitting in cryo-EM structure determination”, *Nat. Met.* **9**, 853–854.

- Schilbach, S., Hantsche, M., Tegunov, D., Dienemann, C., Wigge, C., Urlaub, H. and Cramer, P. (2017) “Structures of transcription pre-initiation complex with TFIID and Mediator”, *Nature* **551**, 204-209.
- Sharrock, R.A. and Quail, P.H. (1989) “Novel phytochrome sequences in *Arabidopsis thaliana*: structure, evolution, and differential expression of a plant regulatory photoreceptor family”, *Genes. Dev.* **3**, 1745-1757.
- Shinomura, T., Nagatani, A., Hanzawa, H., Kubota, M., Watanabe, M. and Furuya, M. (1996) “Action spectra for phytochrome A- and B-specific photoinduction of seed germination in *Arabidopsis thaliana*”, *Proc. Natl. Acad. Sci. USA* **23**, 8129-8133.
- Sorzano, C. O. S., Jiménez, A., Mota, J., Vilas, J. L., Maluenda, D., Martínez, M., Ramírez-Aportela, E., Majtner, T., Segura, J., Sánchez-Gracia, R., Rancel, Y., del Caño, L., Conesa, P., Melero, R., Jonic, S., Vargas, J., Cazals, F., Freyberg, Z., Krieger, J., Bahar, I., Marabini, R. and Carazo, J. M. (2019) “Survey of the analysis of continuous conformational variability of biological macromolecules by electron microscopy”, *Acta Crystallogr.* **F75** 19-32.
- Steinhaus, H. (1957) “Sur la division des corps matériels en parties”, *Bull. Acad. Polon. Sci.* **4**, 801-804.
- Stillman, T. J., Baker, P. J., Britton, K. L. and Rice, D. W. (1993) “Conformational flexibility in glutamate dehydrogenase: Role of water in substrate recognition and catalysis”, *J. Mol. Biol.* **234**, 1131-1139.
- Stuhrmann, H. B. (1970) “Interpretation of small-angle scattering functions of dilute solutions and gases. A representation of the structures related to a one-particle scattering function”, *Acta Cryst.* **A26**, 297-306.
- Suetsugu, N., Kong, S. G., Kasahara, M. and Wada, M. (2013) “Both LOV1 and LOV2 domains of phototropin2 function as the photosensory domain for hypocotyl phototropic responses in *Arabidopsis thaliana* (Brassicaceae)” *Am. J. Bot.* **100**, 60-69.
- Sugita, Y., Kitao, A. and Okamoto, Y. (2000) “Multidimensional replica-exchange method for free-energy calculations”, *J. Chem. Phys.* **113**, 6042-6051.
- Svergun, D. I. (1992) “Determination of the regularization parameter in indirect-transform methods using perceptual criteria”, *J. Appl. Cryst.* **25**, 495-503.
- Svergun, D. I. (1995) “CRYSOLO-a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates”, *J. Appl. Crystallogr.* **28**, 768-773.
- Svergun, D. I. (1999) “Restoring low resolution structure of biological macro-molecules from solution scattering using simulated annealing”, *Biophys. J.* **76**, 2879-2886.
- Svergun, D. I., Koch, M. H. J., Timmins, P. A. and May, R. P. (2013) *Small Angle X-ray and*

- Neutron Scattering from Solutions of Biological Macromolecules*, Oxford University Press, U.K.
- Svergun, D. I., Petoukhov, M. V. and Koch, M. H. (2001) "Determination of domain structure of proteins from X-ray solution scattering", *Biophys. J.* **80**, 2946-2953.
- Swartz, T. E., Corchnoy, S. B., Christie, J. M., Lewis, J. W., Szundi, I., Briggs, W. R. and Bogomolni, R. A. (2001) "The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin", *J. Biol. Chem.* **276**, 36493-36500.
- Takayama, Y. and Nakasako, M. (2012) "Humidity-controlled preparation of frozen-hydrated biological samples for cryogenic coherent x-ray diffraction microscopy", *Rev. Sci. Instrum.* **83**, 054301 (6 pages).
- Takayama, Y., Nakasako, M., Okajima, K., Iwata, A., Kashojiya, S., Matsui, Y. and Tokutomi, S. (2011) "Light-Induced Movement of the LOV2 Domain in an Asp720Asn Mutant LOV2-Kinase Fragment of *Arabidopsis* Phototropin 2", *Biochemistry* **50**, 1174-1183.
- Takemiya, A., Sugiyama, N., Fujimoto, H., Tsutsumi, T., Yamauchi, S., Hiyama, A., Tada, Y., Christie, J. M. and Shimazaki, K. (2013) "Phosphorylation of BLUS1 kinase by phototropins is a primary step in stomatal opening", *Nat. Commun.* **4**, 2094.
- Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I. and Ludtke, S. J. (2007) "EMAN2: an extensible image processing suite for electron microscopy", *J. Struct. Biol.* **157**, 38-46.
- Taylor, B. L. and Zhulin, I. B. (1999) "PAS domains: internal sensors of oxygen, redox potential, and light", *Microbiol Mol Biol Rev* **63**, 479-506.
- Tepperman, J. M., Hwang, Y.-S. and Quail, P. H. (2006) "phyA dominates in transduction of red-light signals to rapidly-responding genes at the initiation of *Arabidopsis* seedling deetiolation", *Plant J.* **48**, 728-742.
- Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice", *Nucleic Acids Res.* **22**, 4673-4680.
- Thon, F. (1971) "Phase contrast electron microscopy", In: *Electron Microscopy in Material Science* (U. Valdré, ed.), Academic Press, New York.
- Tian, G., Xiang, S., Noiva, R., Lennarz, W. J. and Schindelin, H. (2006) "The Crystal structure of yeast protein disulfide isomerase suggests cooperativity between its active sites", *Cell* **124**, 61-73.
- Tokutomi, S., Matsuoka, D. and Zikihara, K. (2008) "Molecular structure and regulation of phototropin kinase by blue light", *Biochim. Biophys. Acta* **1784**, 133-142.

- Tokutomi, S., Nakasako, M., Sakai, J., Kataoka, M., Yamamoto, K. T., Wada, M., Tokunaga, F. and Furuya, M. (1989) “A model for the dimeric molecular structure of phytochrome based on small-angle X-ray scattering”, *FEBS Lett.* **247**, 139-142.
- Trippens, J., Greiner, A., Schellwat, J., Neukam, M., Rottmann, T., Lu, Y., Kateriya, S., Hegemann, P. and Kreimer, G. (2012) “Phototropin influence on eyespot development and regulation of phototactic behavior in *Chlamydomonas reinhardtii*”, *Plant Cell* **24**, 4687-4702.
- Ulijasz, A. T., Cornilescu, G., Cornilescu, C. C., Zhang, J., Rivera, M., Markley, J. L. and Vierstra, R. D. (2010) “Structural basis for the photoconversion of a phytochrome to the activated Pfr form”, *Nature* **463**, 250-254.
- van der Maaten, L. and Hinton, G. (2008) “Visualizing data using t-SNE”, *J. Mach. Learn Res.* **9**, 2579-2605.
- van Heel, M. and Stöffler-Meilicke, M. (1985) “The characteristic views of *E. coli* and *B. stearothermophilus* 30S ribosomal subunits in the electron microscope”, *EMBO J.* **4**, 2389-2395.
- von Horsten, S., Straß, S., Hellwig, N., Gruth, V., Klasen, R., Mielcarek, A., Linne, U., Morgner, N. and Essen, L. O. (2016) “Mapping light-driven conformational changes within the photosensory module of plant phytochrome B”, *Sci. Rep.* **6**, 34366 (8 pages).
- Velazquez Escobar, F., Piwowarski, P., Salewski, J., Michael, N., Fernandez, Lopez M., Rupp, A., Qureshi, B. M., Scheerer, P., Bartl, F., Frankenberg-Dinkel, N., Siebert, F., Andrea Mroginski, M. and Hildebrandt, P. (2015) “A protonation-coupled feedback mechanism controls the signalling process in bathy phytochromes”, *Nat Chem* **7**, 423-430.
- Volkov, V. V. and Svergun, D. I. (2004) “Uniqueness of *ab initio* shape determination in small-angle scattering”, *J. Appl. Cryst.* **36**, 860-864.
- Watanabe, Y. and Inoko, Y. (2009) “Size-exclusion chromatography combined with small-angle X-ray scattering optics”, *J. Chromatogr. A* **1216**, 7461–7465.
- Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. and Barton, G. J. (2009) “Jalview Version 2-a multiple sequence alignment editor and analysis workbench”, *Bioinformatics* **25**, 1189-1191.
- Wiltbank L. B. and Kehoe, D. M. (2019) “Diverse light responses of cyanobacteria mediated by phytochrome superfamily photoreceptors”, *Nat. Rev. Microbiol.* **17**, 37-50.
- Wright, P. E. and Dyson, H. J. (1999) “Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm”, *J. Mol. Biol.* **293**, 321-331.
- Wu, S. H. and Lagarias, J. C. (2000) “Defining the bilin lyase domain: lessons from the extended phytochrome superfamily” *Biochemistry* **39**, 13487-13495.

- Yamaguchi, R., Nakamura, M., Mochizuki, N., Kay, S. A., and Nagatani, A. (1999) “Light-dependent translocation of a phytochrome B:GFP fusion protein to the nucleus in transgenic *Arabidopsis*”, *J. Cell Biol.* **145**, 437–445.
- Yang, X., Kuk, J. and Moffat, K. (2009) “Conformational differences between the Pfr and Pr states in *Pseudomonas aeruginosa* bacteriophytochrome”, *Proc. Natl. Acad. Sci. USA* **106**, 15639-15644.
- Yeh, K. C. and Lagarias, J. C. (1998) “Eukaryotic phytochromes: light-regulated serine/threonine protein kinases with histidine kinase ancestry” *Proc. Natl. Acad. Sci. USA* **95**, 13976-13981.
- Yeh, K. C., Wu, S. H., Murphy, J. T. and Lagarias, J. C. (1997) “A cyanobacterial phytochrome two-component light sensory system”, *Science* **277**, 1505-1508.
- Zhang, K. (2016) “Gctf: Real-time CTF determination and correction”, *J. Struct. Biol.* **193**, 1-12.
- Zhang, Z., Komives, E. A., Sugio, S., Blacklow, S. C., Narayana, N., Xuong, N. H., Stock, A. M., Petsko, G. A. and Ringe, D. (1999) “The role of water in the catalytic efficiency of triosephosphate isomerase”, *Biochemistry* **38**, 4389-4397.
- Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y. and Agard, D. A. (2017) “MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy”, *Nat. Met.* **14**, 331-332.
- Zhong, E. D., Bepler, T., Davis, J. H. and Berger, B. (2019) “Reconstructing continuously heterogeneous structures from single particle cryo-EM with deep generative models”, *arXiv* 1909.05215v1 (16 pages).
- Zhou, Y., Morais-Cabral, J. H., Kaufman, A. and MacKinnon, R. (2001) “Chemistry of ion coordination and hydration revealed by a K⁺ channel-Fab complex at 2.0 Å resolution”, *Nature* **414**, 43-48.
- Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E. and Scheres, S. H. W. (2018) “New tools for automated high-resolution cryo-EM structure determination inRELION-3”, *eLife* **7**, e42166 (22 pages)
- 砂川重信 (1999) “理論電磁気学 第3版”, 紀伊国屋書店, 東京.
- 中迫雅由, 山本雅貴 (2015) “我が国における生体分子・粒子の構造解析：過去・現在そして未来の展望”, *日本物理学会誌*, **70**, 702-712.
- 西川建 (2009) “天然変性蛋白質とは何か?”, *生物物理* **49**, 4-10.
- 堀内繁雄 (1988) “高分解能電子顕微鏡 —原理と利用法—”, 共立出版, 東京
- 松下智直 (2008) “植物の光情報受容体フィトクロムの細胞内シグナル伝達機構”, *光合*

成研究, **18**, 16-23.

松下智直, 長谷あきら (2004) “フィトクロムの細胞内シグナル伝達機構再考”, *蛋白質
核酸 酵素*, **49**, 749-757.

和田正三, 徳富哲, 長谷あきら, 長谷部光泰 (2001) “細胞工学別冊 植物細胞工学シリーズ 16 植物の光センシング - 光情報の受容とシグナル伝達 -”, 秀潤社, 東京.