

MDシミュレーションによる XFEL溶液散乱の実現の検討

横浜市立大学大学院
生体ナノシステム科学研究科
荳口友隆

Overview

コヒーレント光(XFEL@SPring8)を用いた
新しいタンパク質溶液散乱法の提案



異方的な散乱パターンから、
高次のタンパク質構造情報が引き出せることが期待できる。

(研究①)

フェムト秒パルスレーザーによるタンパク質1分子の溶液散乱

(研究②)

少数タンパク質分子による溶液散乱法の提案

横浜市立大学・佐藤研究室と北海道大学・西野吉則研究室の共同研究

研究 ①

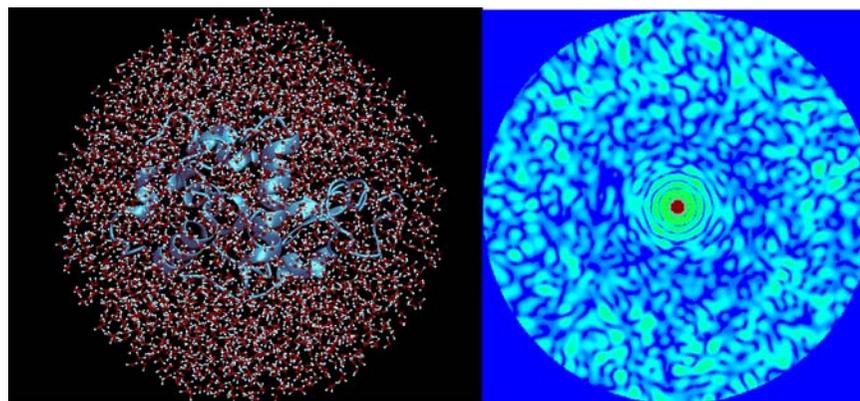
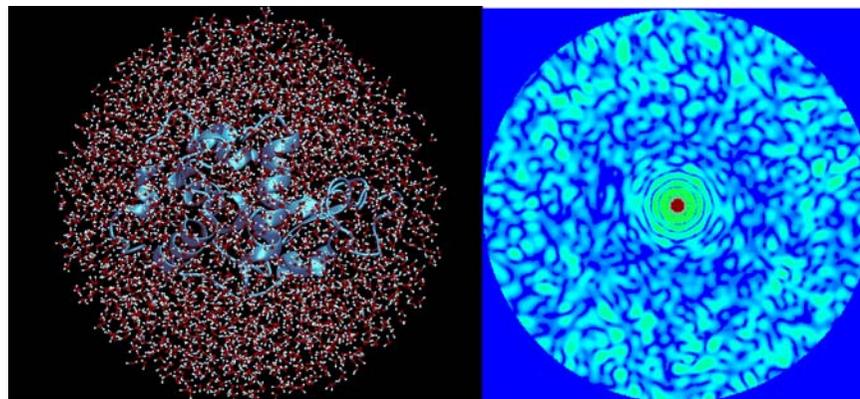
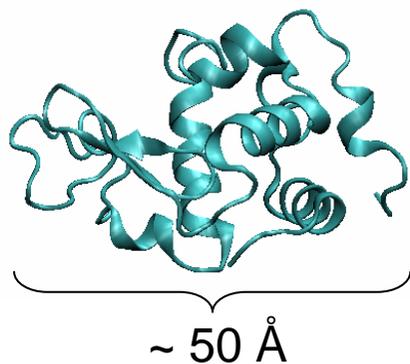
フェムト秒パルスレーザーによる
タンパク質1分子の溶液散乱

タンパク質1分子の溶液散乱： 溶液中の生体分子のフェムト秒スナップショット

何故、XFELで溶液散乱を行うか？

溶液中のリゾチーム

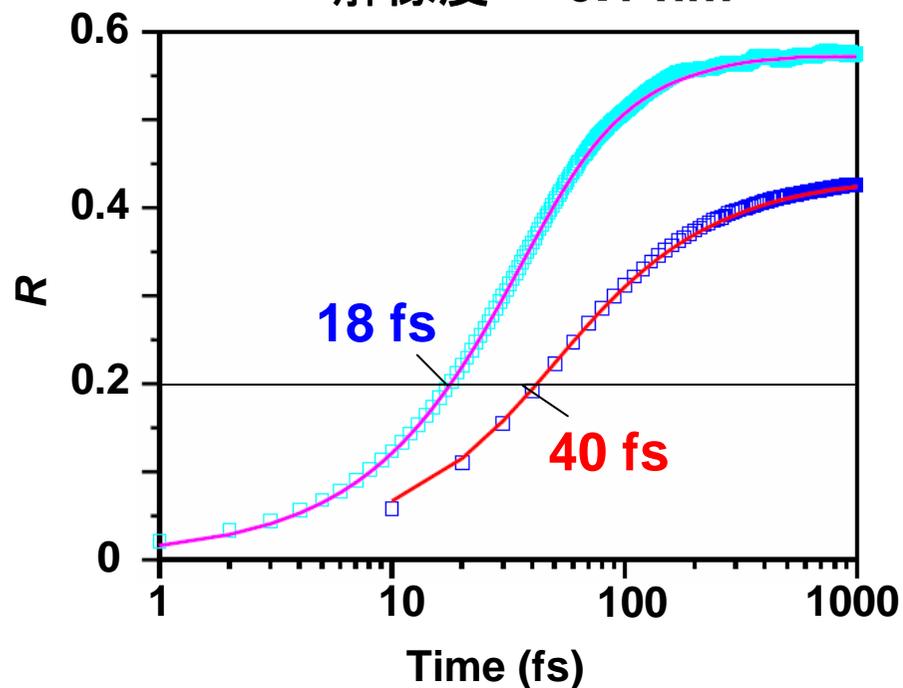
分子動力学計算



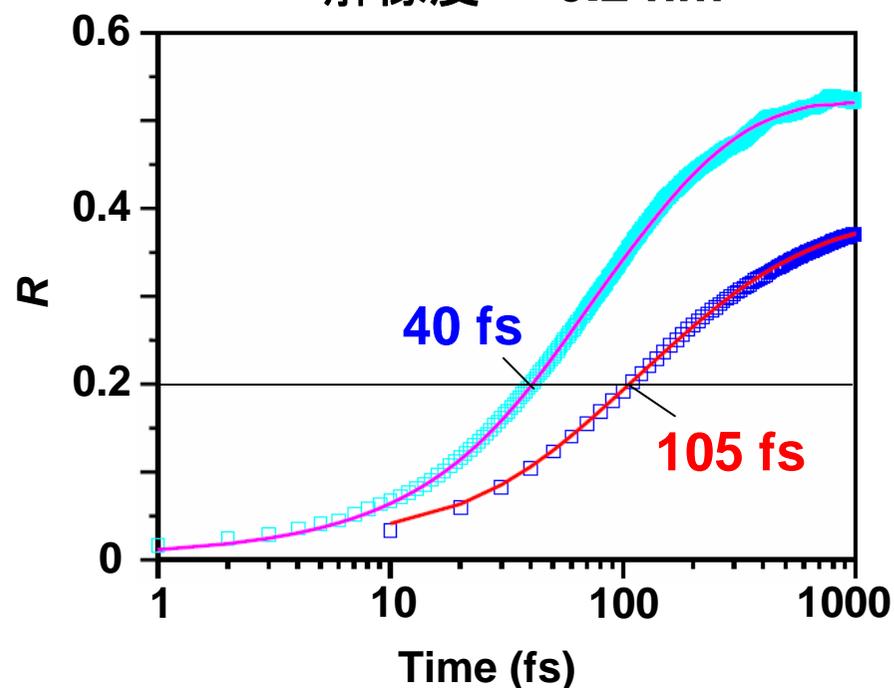
XFELの時間幅(~15フェムト秒)の時間スケールでは
溶液中の分子構造はフリーズしている

散乱スペックルのパターンが維持される時間の見積もり

解像度 = 0.1 nm



解像度 = 0.2 nm



R-factorの時間変化

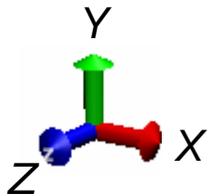
$$R(0, \langle \tau \rangle) = \frac{\sum_{i,j} \left| |F(k_{ij}, 0)| - \sqrt{\langle S(k_{ij}, t) \rangle_{0 \leq t \leq \tau}} \right|}{\sum_{i,j} |F(k_{ij}, 0)|}$$

- 各時間における散乱スペックルの初期構造からのずれ
- 平均散乱スペックルの初期構造からのずれ

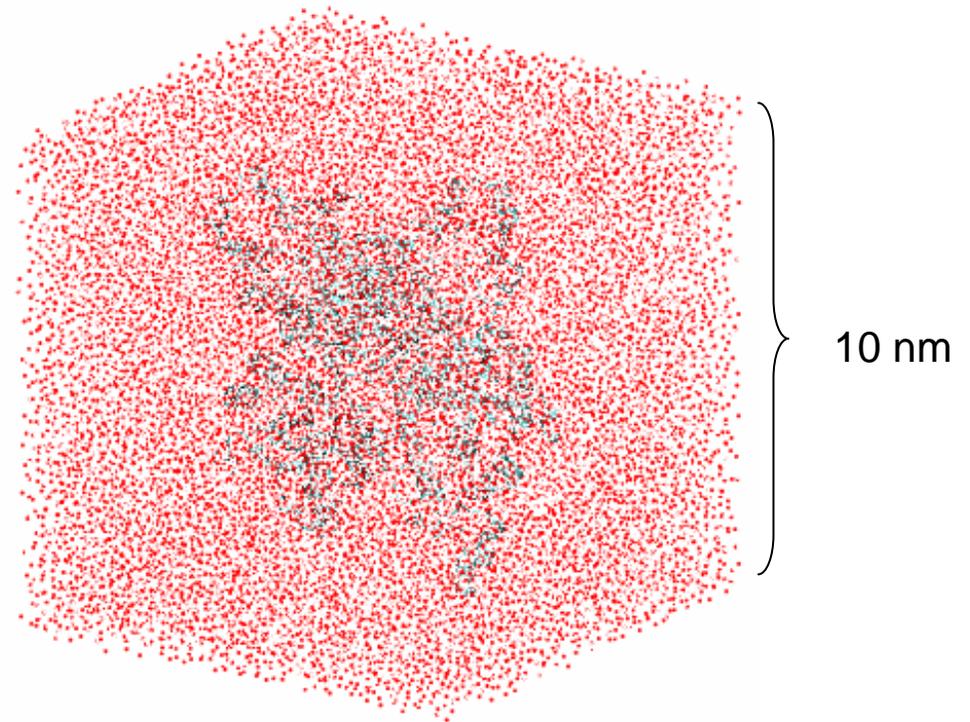
MDを用いたタンパク質分子1個の15 fsパルス溶液散乱の計算



Catalaseの構造



溶液中

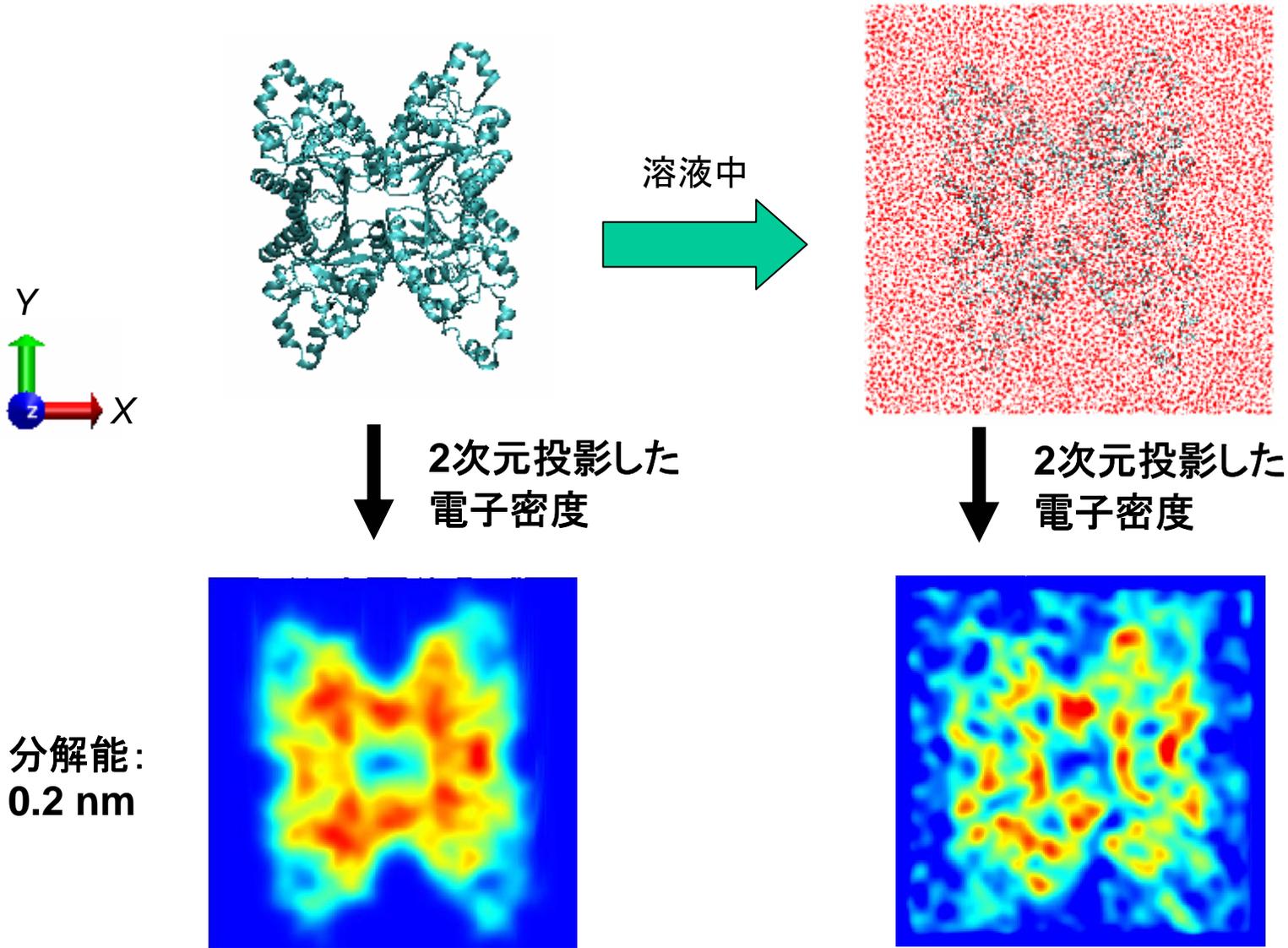


溶液中のCatalase

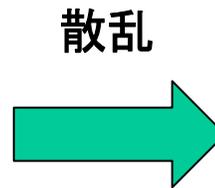
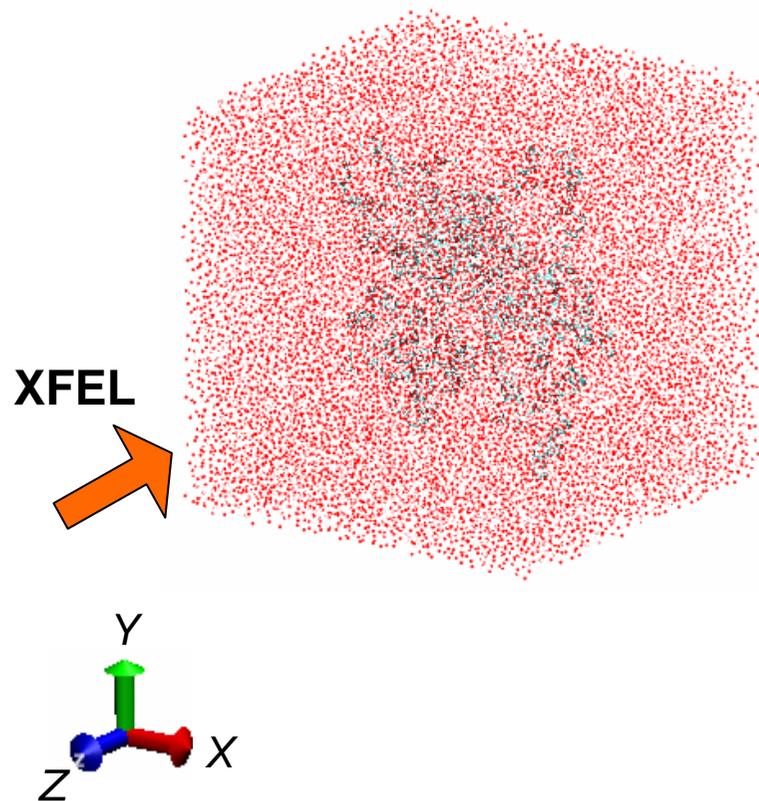
MDの計算条件:

- Simulation time: 5,000 fs
- MD box size = 12.8 nm

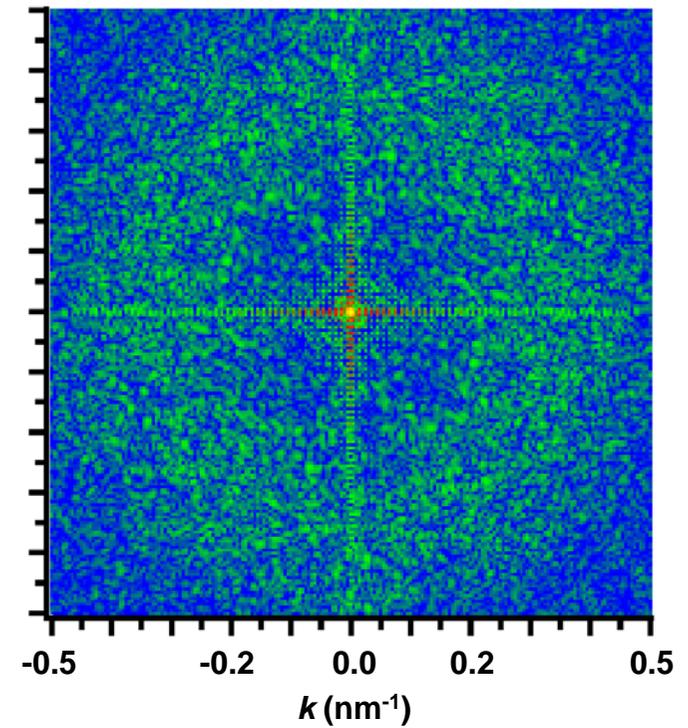
MDを用いたタンパク質分子1個の15 fsパルス溶液散乱



15 fsパルスにおける1分子溶液の散乱スペックル



15 fsの散乱スペックル



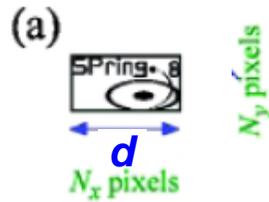
分解能:
0.2 nm

Oversamplingによる像の回復

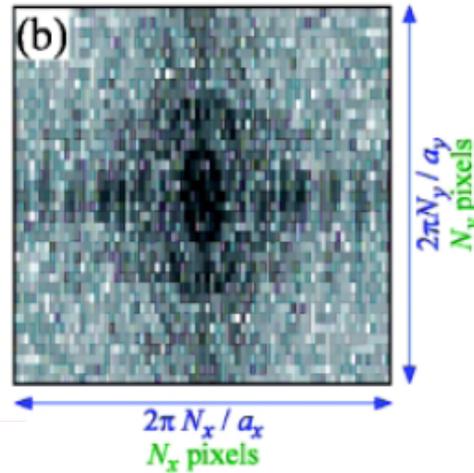
実空間での
パラメーター数

$$N_x \cdot N_y$$

Sample Image



Diffraction Pattern



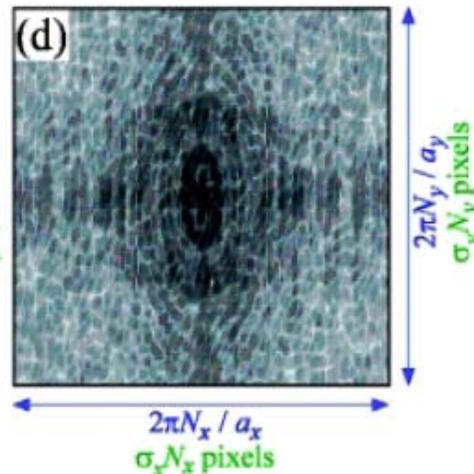
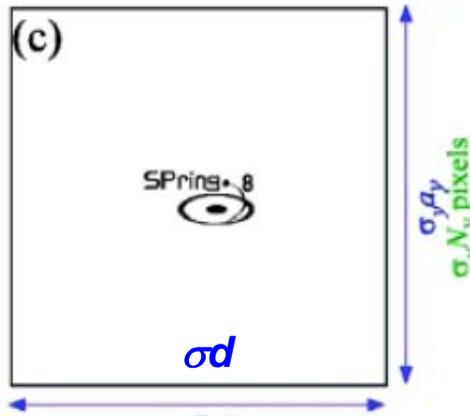
逆空間での
パラメーター数

$$N_x \cdot N_y / 2$$

サンプリング間隔 = $1/d$

Oversampling

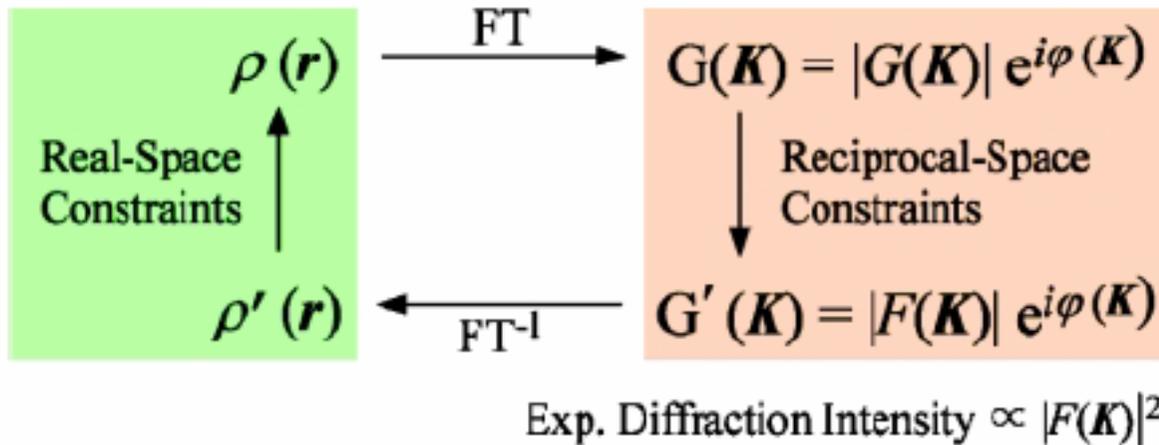
$$N_x \cdot N_y$$



$$\sigma^2 N_x \cdot N_y / 2$$

サンプリング間隔 = $1/\sigma d$

Oversamplingによる像の回復: Hybrid input and output method (HIO)



$|F(\mathbf{k})|$: 実験散乱強度

$G(\mathbf{k})$: 計算散乱強度

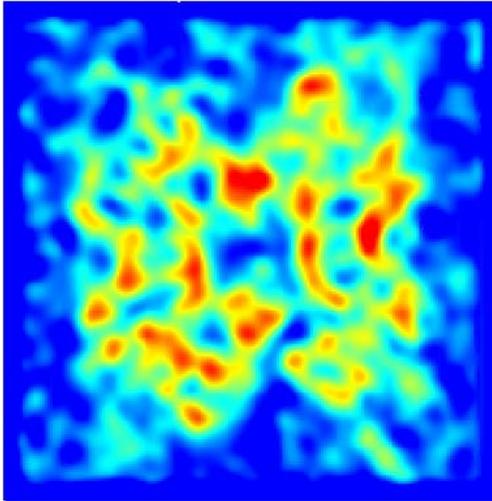
$\rho(\mathbf{r})$: 入力電子密度分布

$\rho'(\mathbf{r})$: 出力電子密度分布

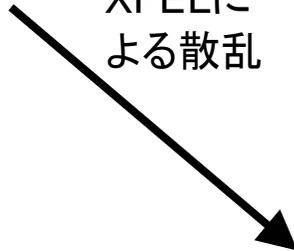
$$\rho_{n+1}(\mathbf{r}) = \begin{cases} \rho'_n(\mathbf{r}), & (\text{サポート内}) \\ \rho_n(\mathbf{r}) - \beta \rho'_n(\mathbf{r}), & (\text{サポート外}) \end{cases}$$

Oversamplingによる像の回復

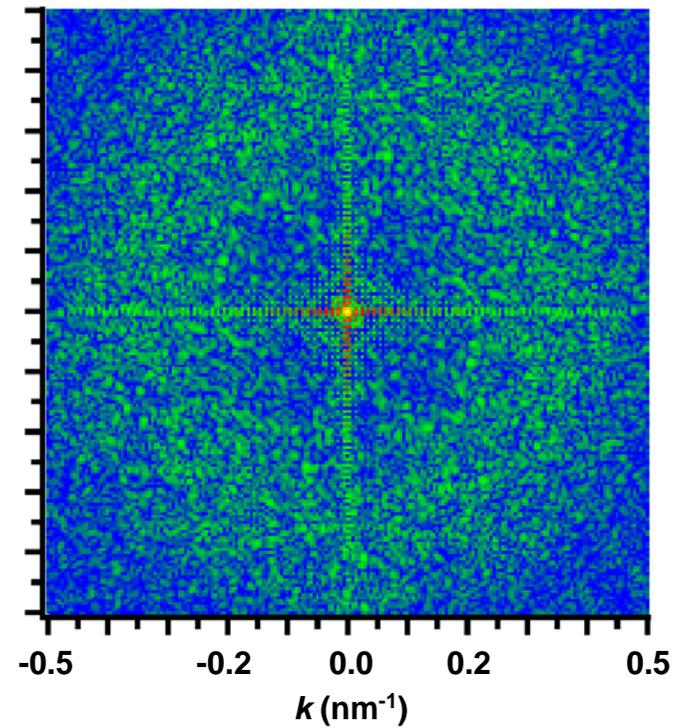
実際の電子密度分布



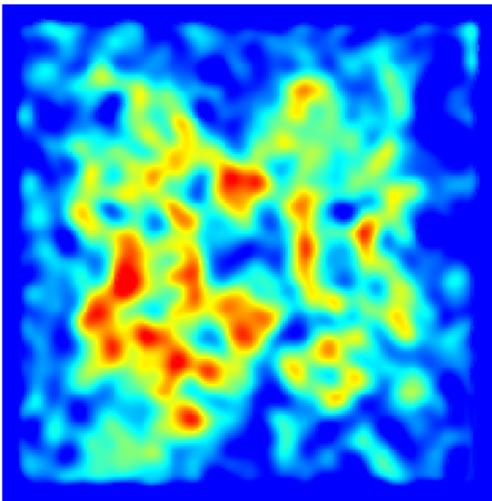
XFELによる
散乱



計算散乱強度



復元した電子密度分布



HIOによる
復元



分解能:
0.2 nm

まとめ

(まとめ)

- 15 fs パルス・X線レーザーによる、1分子溶液散乱の検証。
- 15 fs パルス・X線レーザーであれば、溶液構造も含めてフリーズした電子密度像が得られる。

(今後の課題)

- 15 fs パルス以上での水分子の動きはどれほど像回復を邪魔するかの検討。
- ビームストップ、ポアソン・ノイズ、エwald球の球面性がどれほど像回復を邪魔するかの検討

研究 ②

少数タンパク質分子による
溶液散乱法の提案

より実現が容易な測定系： 少数タンパク質分子による溶液散乱の提案

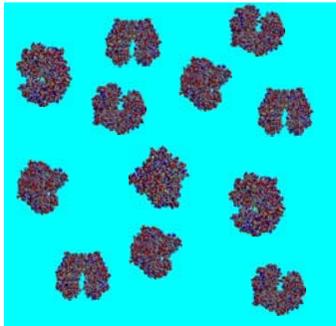
Motivation:

1分子の溶液散乱は非常に有用と期待されるが、実現は困難が予想される。



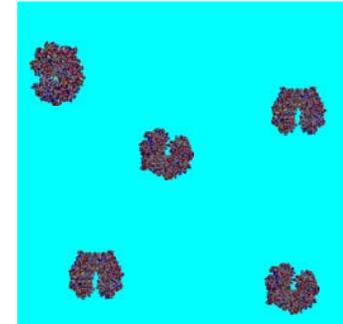
XFELによる少数タンパク質分子による溶液散乱ならば比較的容易。

(通常の溶液散乱)



- 照射体積に多数の分子(10^{12} 個)
- 分子の配向がランダム

(少数タンパク質分子の溶液散乱)



- 照射体積に多数の分子($10^{1\sim3}$ 個)
- 分子の配向が完全にランダムにはならない。

少数のタンパク質分子の散乱から、 どのように構造情報を引き出すか？

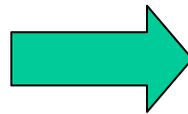
複数のタンパク質分子の散乱：

1分子の三次元散乱強度：

$$I_0(\mathbf{k}) = \sum_l \sum_{l=-m}^m I_{lm}(k) \underline{Y_{lm}(\omega)}$$

球面調和関数展開

複数分子



測定される散乱強度：

$$I(\mathbf{k}) = \int I_0(\omega, \mathbf{k}) \cdot \underline{N_t(\omega)} d\omega$$

溶液中にあるタンパク質分子の回転分布

少数のタンパク質分子の散乱から、 どのように構造情報を引き出すか？

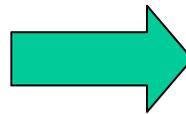
複数のタンパク質分子の散乱：

1分子の三次元散乱強度：

$$I_0(\mathbf{k}) = \sum_l \sum_{l=-m}^m I_{lm}(k) \underline{Y_{lm}(\omega)}$$

球面調和関数展開

複数分子



測定される散乱強度：

$$I(\mathbf{k}) = \int I_0(\omega, \mathbf{k}) \cdot \underline{N_t(\omega)} d\omega$$

溶液中にあるタンパク質分子の回転分布

溶液中において、タンパク質分子が少数で、かつそれが全てフリーズしている (10 fsパルス)

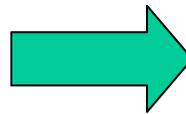
少数のタンパク質分子の散乱から、 どのように構造情報を引き出すか？

複数のタンパク質分子の散乱：

1分子の三次元散乱強度：

$$I_0(\mathbf{k}) = \sum_l \sum_{l=-m}^m I_{lm}(k) Y_{lm}(\omega)$$

複数分子



測定される散乱強度：

$$I(\mathbf{k}) = \int I_0(\omega, \mathbf{k}) \cdot \underline{N_t(\omega)} d\omega$$

溶液中にあるタンパク質分子の回転分布

溶液中において、タンパク質分子が少数で、かつそれが全てフリーズしている (10 fsパルス)



タンパク質分子の回転分布が完全にランダムでなくなる：

回転分布密度関数：

$$N_t(\omega) = \frac{N}{8\pi^2} + \frac{\alpha_t(\omega)}{\uparrow}$$

↑ Random (通常の溶液散乱) ↑ Fluctuation

少数のタンパク質分子の散乱から、 どのように構造情報を引き出すか？

通常の溶液散乱：

1分子の三次元散乱強度：

$$I_0(\mathbf{k}) = \sum_l \sum_{l=-m}^m I_{lm}(k) Y_{lm}(\omega)$$

Random



測定される散乱強度：

$$I(k) = N \cdot I_{00}(k)$$

0次項

少数分子の溶液散乱：

1分子の三次元散乱強度：

$$I_0(\mathbf{k}) = \sum_l \sum_{l=-m}^m I_{lm}(k) Y_{lm}(\omega)$$

Not
random



測定する量： 散乱強度の相関をとる

$$\begin{aligned} & \langle I(\mathbf{k}_1) \cdot I(\mathbf{k}_2) \rangle - I(k_1) \cdot I(k_2) \\ &= \sum_l P_l(\cos \phi) \sum_{l=-m}^m I_{lm}(\mathbf{k}_1) I_{lm}^*(\mathbf{k}_2) \end{aligned}$$

高次項の連立方程式が得られる。

少数のタンパク質分子の散乱の検討

溶液中に複数のタンパク質分子(数十～数百個)がある場合の散乱スペクトルを原子レベルで計算する。

(原子 + 一様電子密度モデルの組み合わせ)

MD: タンパク質+タンパク質周りの溶媒分子(水和構造)

一様電子密度: バルク領域 (タンパク質分子によって摂動を受けてない溶媒領域)

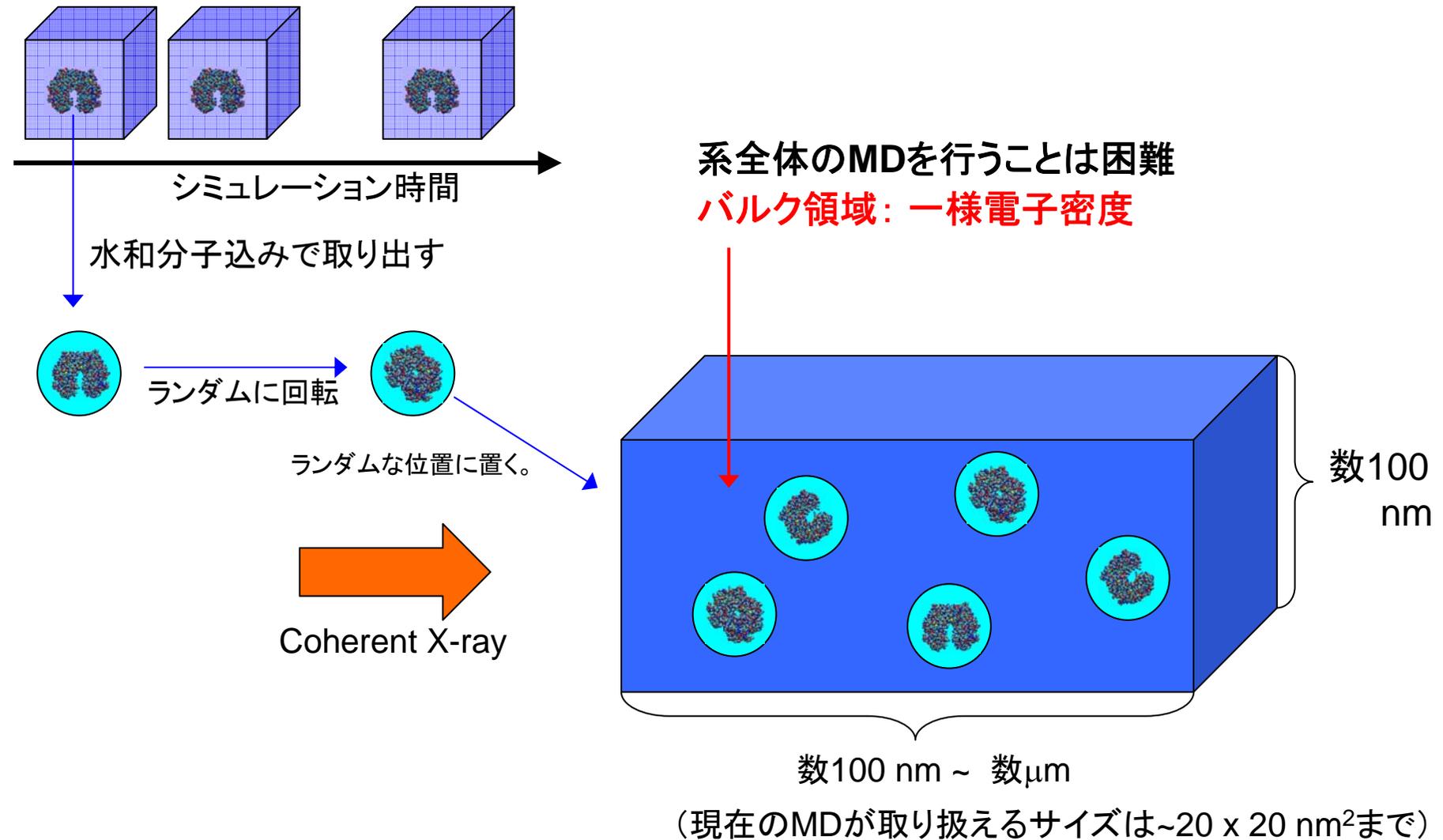


(実験で取り扱うような溶液サイズのMDは
現段階では不可能であるため)

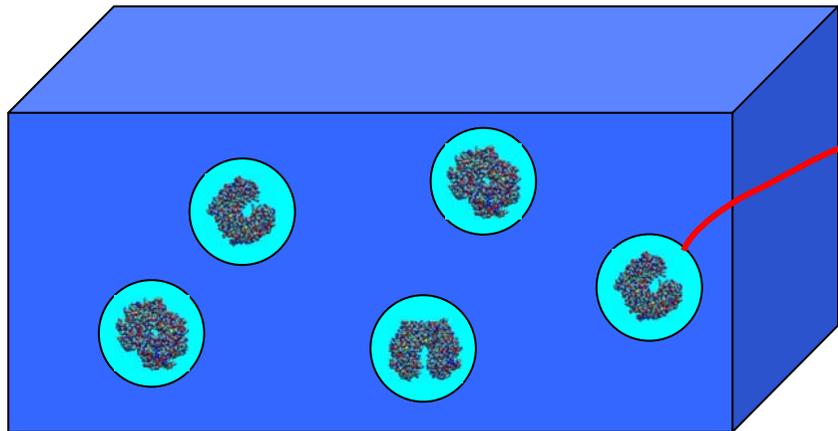
- どのような実験パラメーターであれば、タンパク質の散乱が十分な強度で得られるか？
- 溶液中に複数のタンパク質分子があっても、異方的な散乱スペクトルは得られるのか？

原子 + 一様電子密度モデルの組み合わせ

タンパク質溶液のMD



原子 + 一様電子密度モデルの組み合わせ



領域 Σ_i : 位置は $\mathbf{r}_{\Sigma i}$

i 番目のタンパク質 + 水和構造
の散乱長は、

$$F_{\Sigma i}(\mathbf{k}) = \sum_{j=1}^{N_{\Sigma i}} \bar{f}_j(\mathbf{k}) e^{-2\pi i \mathbf{k} \cdot \mathbf{r}_j}$$

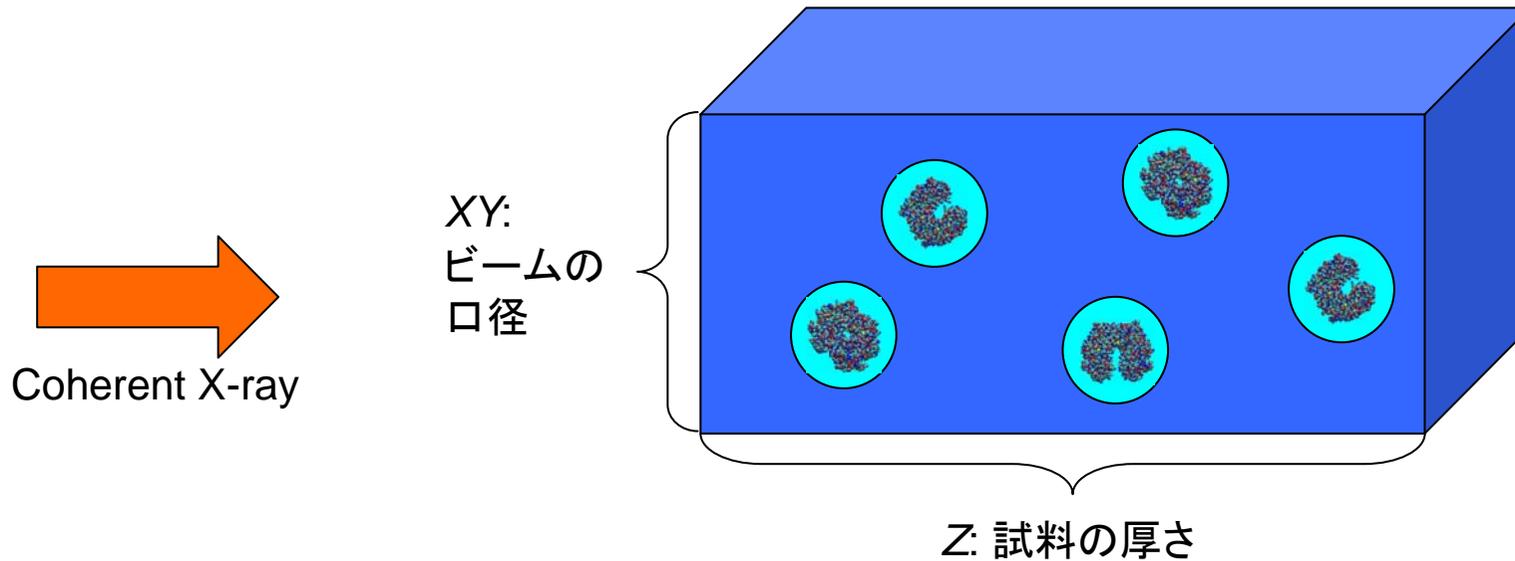
系全体の散乱長は、 $F(\mathbf{k}) = S(\mathbf{k}) + \sum_{i=1}^{N_{\text{protein}}} F_{\Sigma i}(\mathbf{k}) e^{-2\pi i \mathbf{k} \cdot \mathbf{r}_{\Sigma i}} = S(\mathbf{k}) + F_p(\mathbf{k})$

系全体の散乱強度は、 $I(\mathbf{k}) = |S(\mathbf{k})|^2 + 2S(\mathbf{k})\text{Re}(F_p(\mathbf{k})) + |F_p(\mathbf{k})|^2$

$$= \underbrace{I_B(\mathbf{k})}_{\text{バルクの散乱}} + \underbrace{I_{BP}(\mathbf{k})}_{\text{Cross term}} + \underbrace{I_P(\mathbf{k})}_{\text{タンパク質からの散乱}}$$

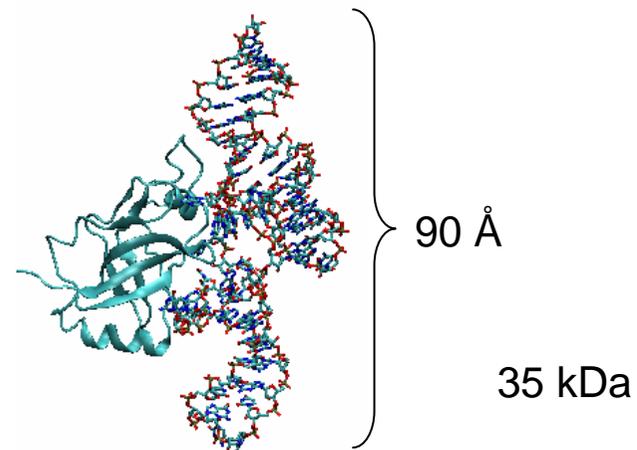
分子間距離情報のないタンパク質からの散乱 $I_P^0(\mathbf{k}) = \sum_{i=1}^{N_{\text{protein}}} |F_{\Sigma i}(\mathbf{k})|^2$ も計算しておく

計算で検討する実験パラメーター



- 入射光子数 = 5×10^{11} photon
- 波長: $\lambda = 0.25$ nm
- 最大解像度: $1/k_{\max} = 1$ nm
- 口径: $X = Y = 50, 100$ nm
- 試料の厚さ: $Z \leq 10$ μm
- Sampling間隔: $D = 1/2X$
- タンパク質の濃度: 0.5 g/L 以上
(通常の溶液散乱でのタンパク質濃度)

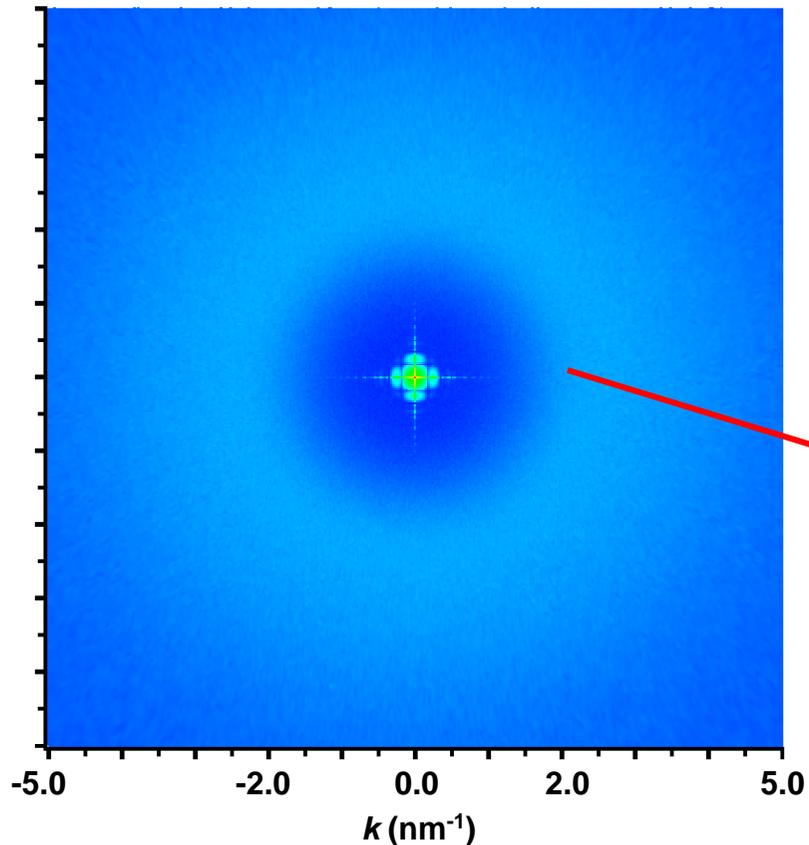
RNA binding protein/RNA (tmRNA)



計算条件の検討: 最大解像度

分子動力学シミュレーションから計算した純水の散乱

100 snapshotの平均



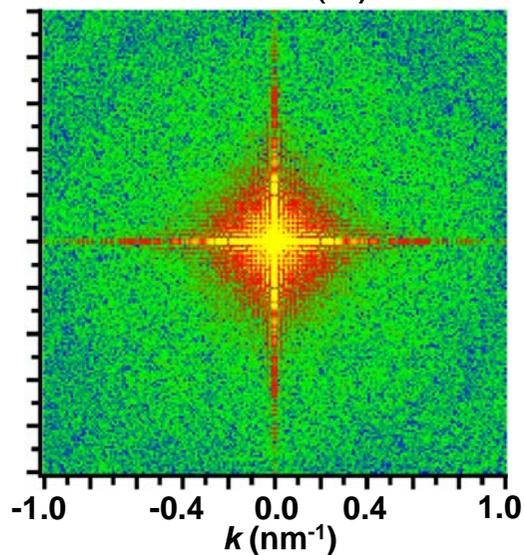
- (計算条件)
- MD: 20 nmの立方体
 - 集光面積: $X = 100$ nm, $Y = 100$ nm
 - $\lambda = 0.2$ [nm]
 - Max resolution = 0.2 [nm]
 - サンプルング間隔: $D = 40$ nm

0.5 nm程度のところに、水分子間動径分布のFirst peakに対応する環を持つ。

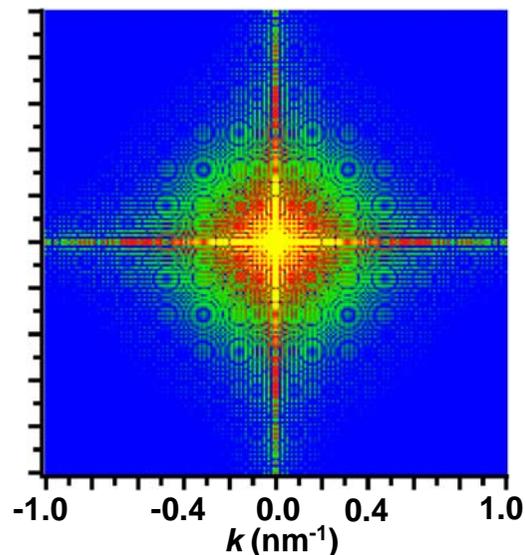
最大解像度は1 nmに設定。

tmRNA: $X=Y=50\text{nm}$, $Z=2,500\text{nm}$, $\sim 0.6\text{ g/L}$ ($n=64$ 個)

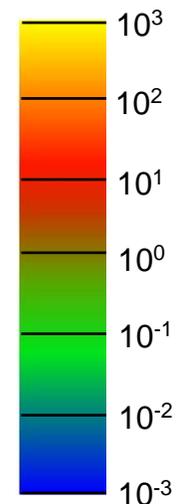
トータル $I(\mathbf{k})$



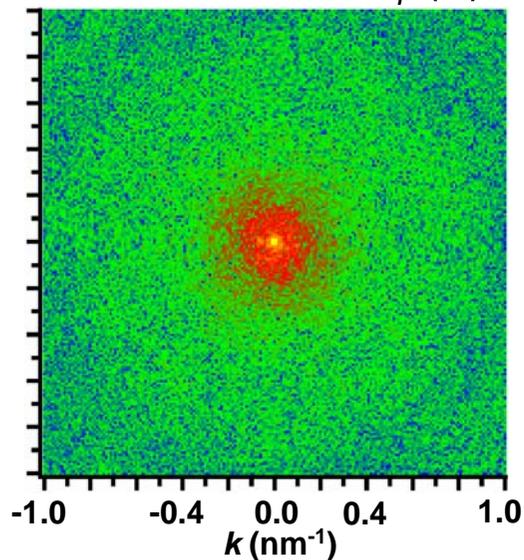
バルク $I_B(\mathbf{k})$



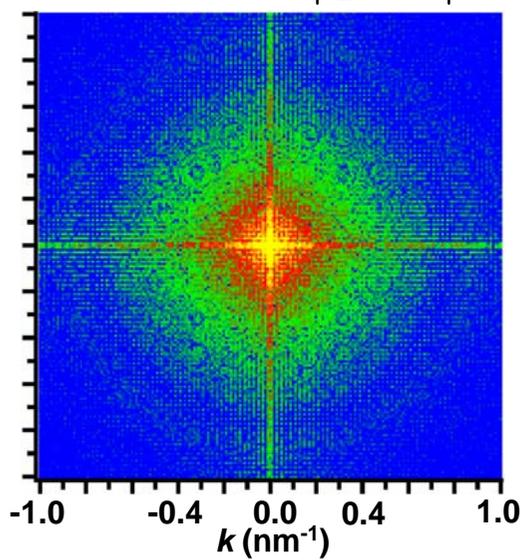
[photon]



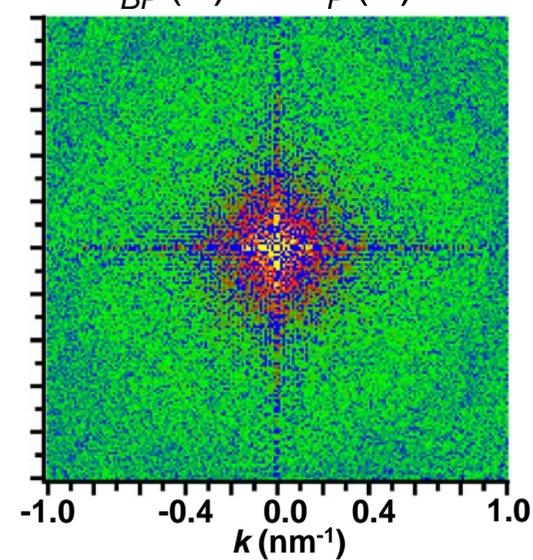
タンパク質+水和構造 $I_P(\mathbf{k})$



クロスターム $|I_{BP}(\mathbf{k})|$



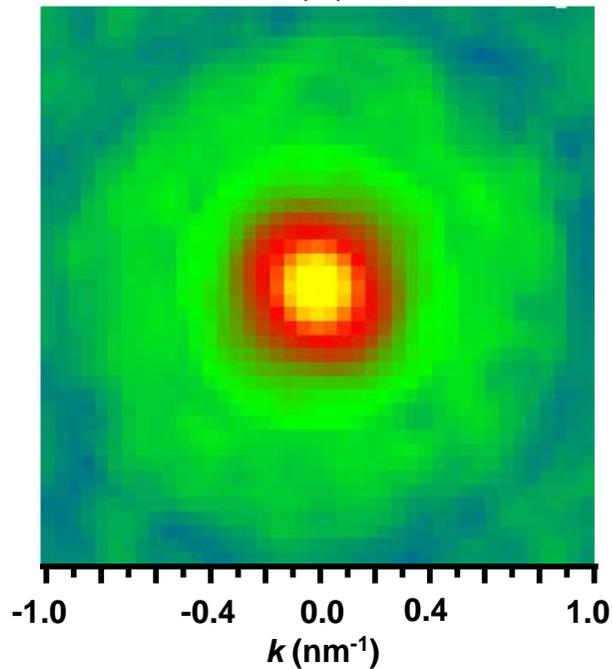
$I_{BP}(\mathbf{k}) + I_P(\mathbf{k})$



tmRNA: $X=Y=50\text{nm}$, $Z=2,500\text{nm}$, $\sim 0.6\text{ g/L}$ ($n=64$ 個)

仮に、クロスタームを処理できたとして、タンパク質サイズ(10nm x2)に Sampling 間隔を広げると、はっきりと異方性を観察することができる。

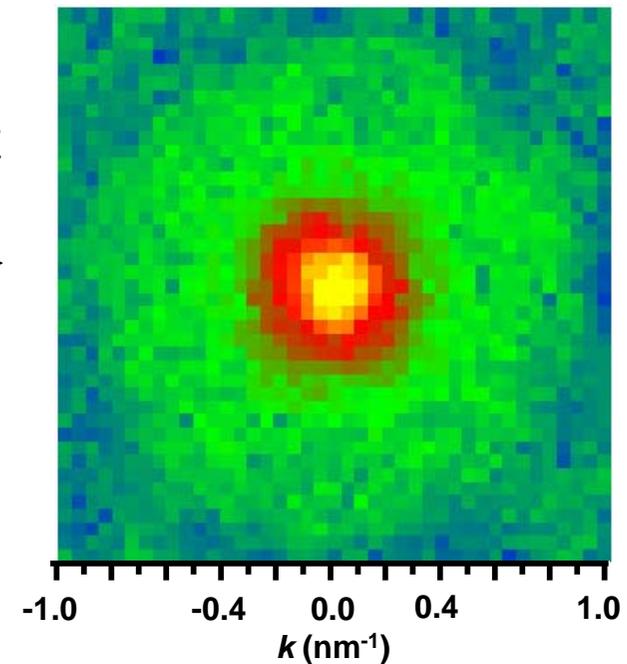
$$I_P^0(\mathbf{k}) = \sum_{i=1}^{N_{\text{protein}}} |F_{\Sigma i}(\mathbf{k})|^2$$



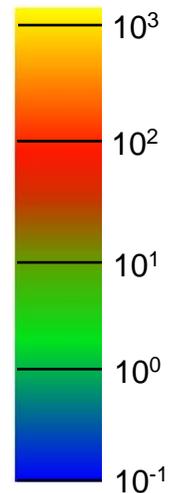
タンパク質間距離
の情報有り。



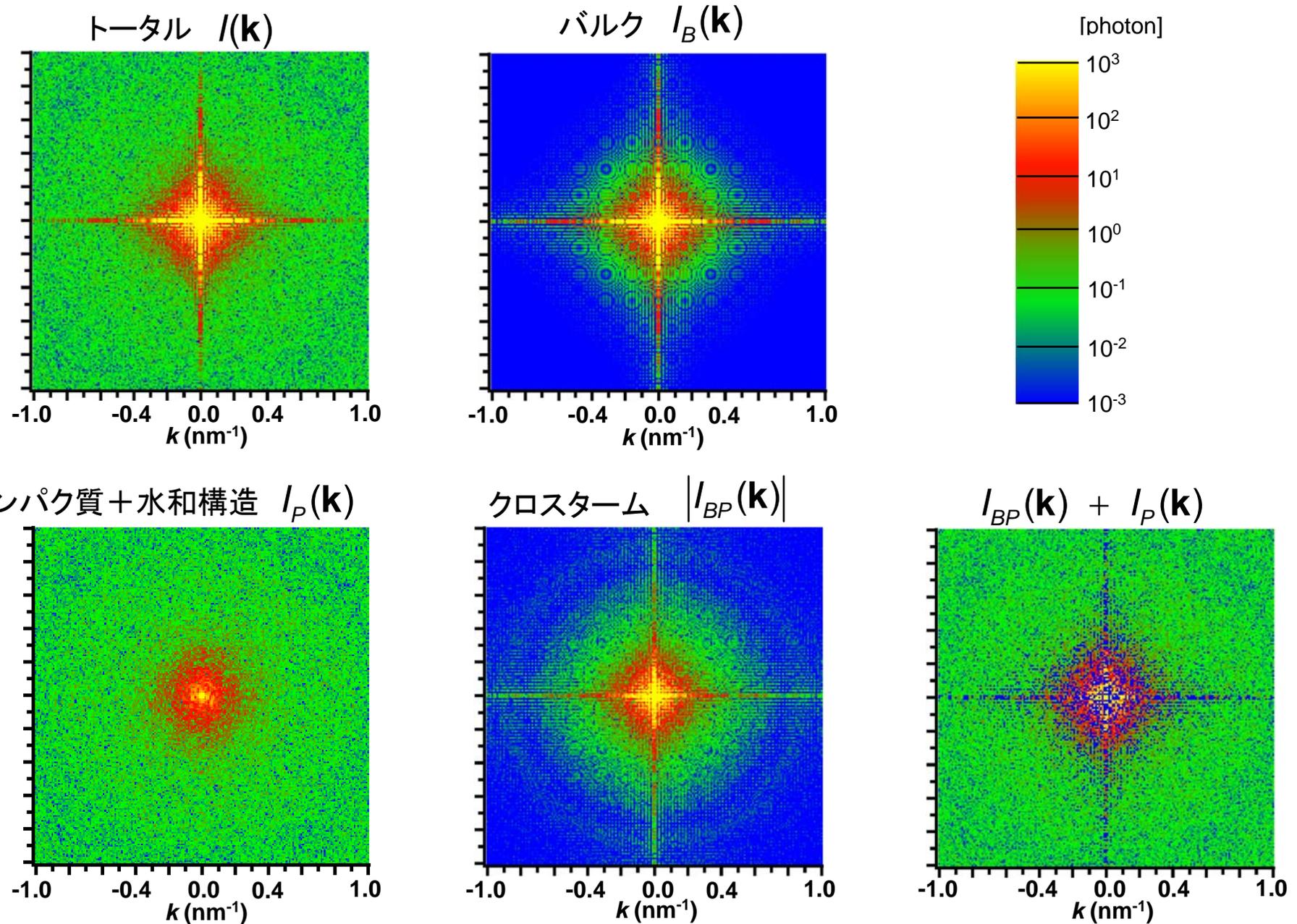
$I_P(\mathbf{k})$



[photon]



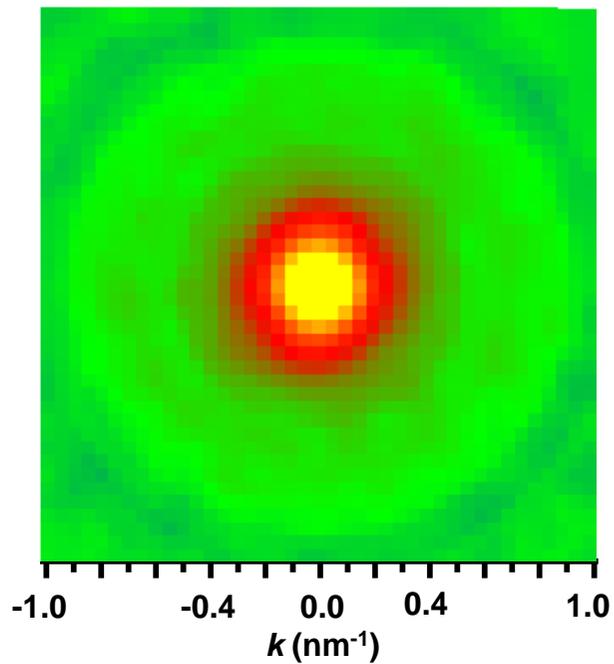
tmRNA: $X=Y=50\text{nm}$, $Z=2,500\text{nm}$, $\sim 1.2\text{ g/L}$ ($n=128$ 個)



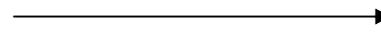
tmRNA: $X=Y=50\text{nm}$, $Z=2,500\text{nm}$, $\sim 1.2\text{ g/L}$ ($n=128$ 個)

同様に異方性を観察することができる。

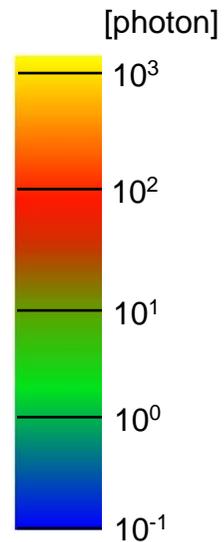
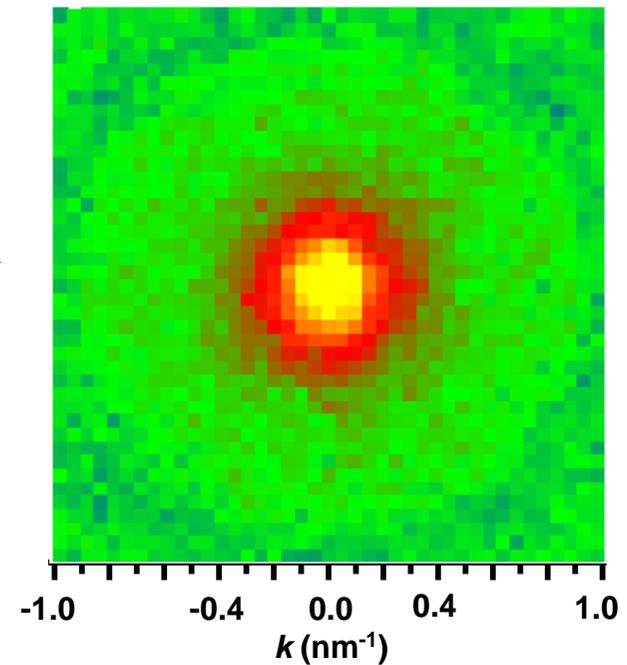
$$I_P^0(\mathbf{k}) = \sum_{i=1}^{N_{\text{protein}}} |F_{\Sigma i}(\mathbf{k})|^2$$



タンパク質間距離
の情報有り。



$I_P(\mathbf{k})$



まとめ

(まとめ)

- 少数タンパク質分子による新しい溶液散乱測定法の提案
- tmRNAの場合、タンパク質の個数が~200個程度あれば、異方的な散乱スペクトルが得られる。
- 1 nm以上、2 nm以下の広角の情報は完全に得られる。

(今後の課題)

- ただし、小角はクロスタームによる邪魔が大きい。
 - ・取り除く方法を考える(溶液の形状から、バルクの散乱が強く出るピクセルを推定し、解析からあらかじめ省いておく)
 - ・もっと大きい口径を検討する。

謝辞

(横浜市立大学)

佐藤衛 先生

池口満徳 先生

(北海道大学)

西野吉則 先生

(Spring8)

城地保昌 先生

別所義隆 先生