

PF研究会「ERLサイエンスワークショップⅡ」  
2011年4月27日－28日、KEK-PF

# 天然変性タンパク質をターゲットとした 新しい構造生物学

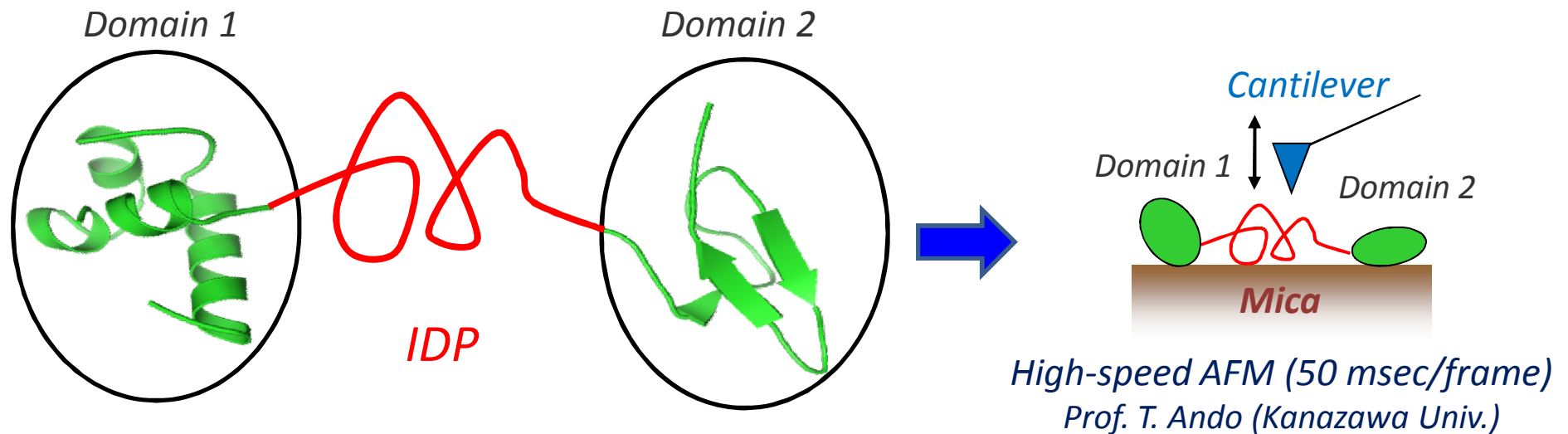
—将来光源を用いた1分子解析を目指して—

佐藤 衛

横浜市立大学 大学院 生命ナノシステム科学研究科

# 天然変性タンパク質

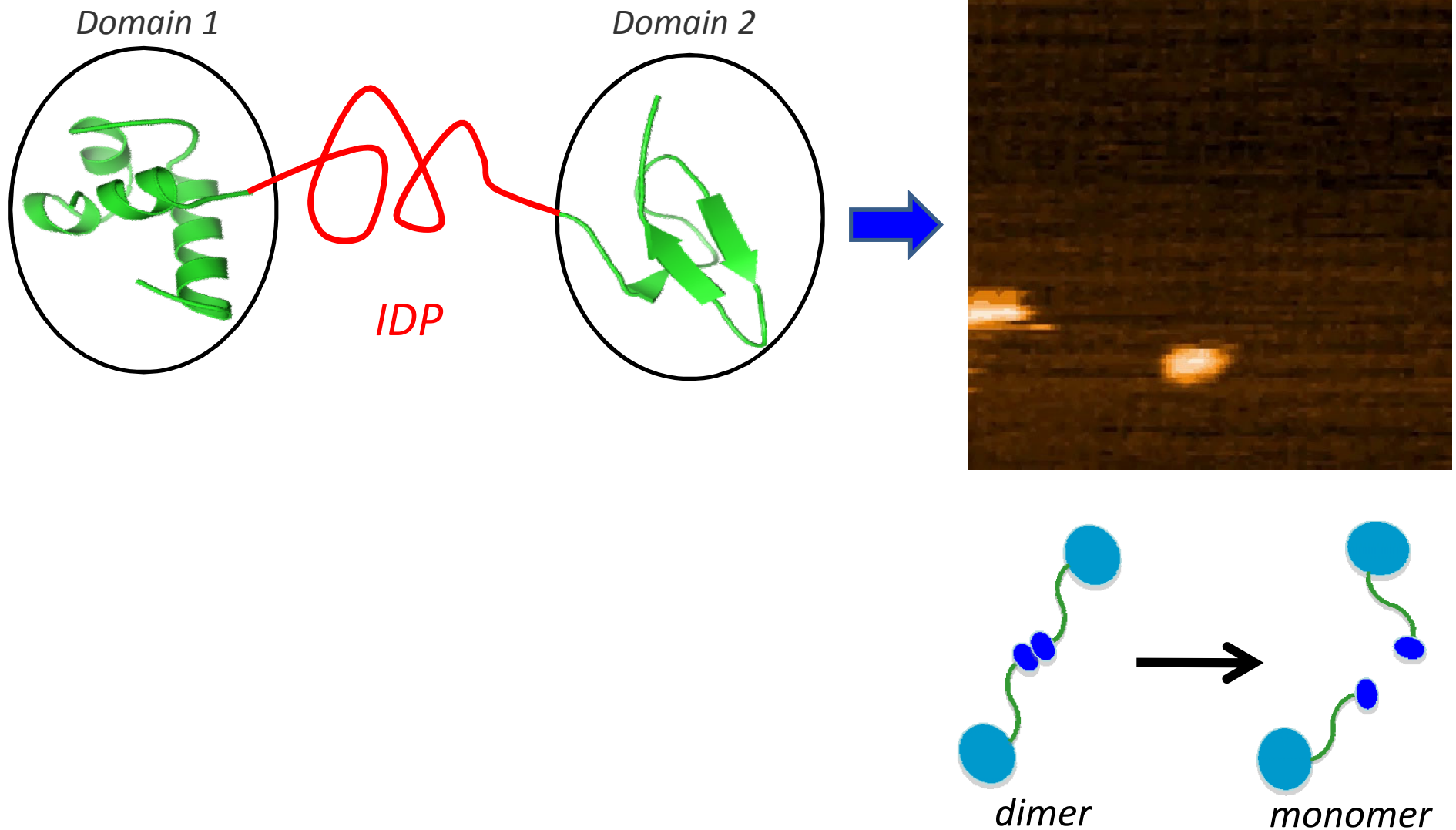
## *Intrinsically Disordered Protein: IDP*



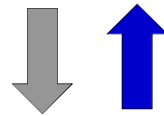
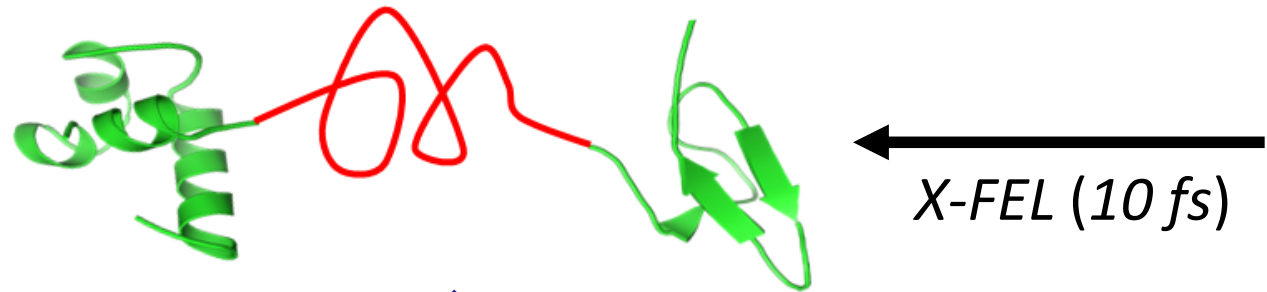
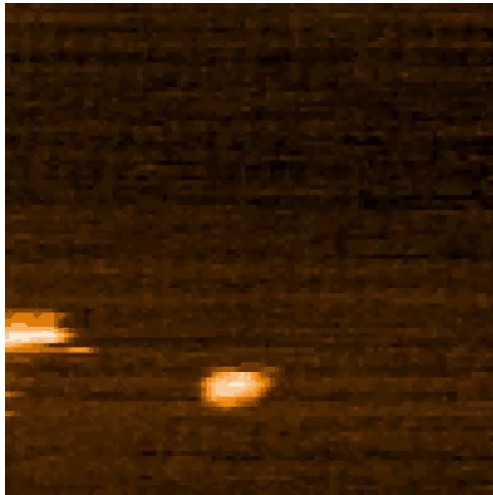
*IDPs are characterized by a lack of stable folded structure along their entire lengths or in localized regions between domains, when they exist as isolated polypeptide chains under physiological conditions*

# 天然変性タンパク質

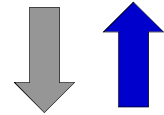
*Intrinsically Disordered Protein: IDP*



# X-ray Scattering from Single *IDP*

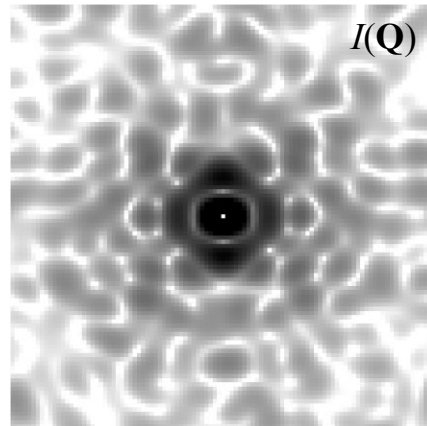


散乱因子：  $F(\mathbf{Q}) = \int \rho(\mathbf{r}) e^{-i\mathbf{Q}\cdot\mathbf{r}} d^3\mathbf{r}$      $|\mathbf{Q}| = (4\pi/\lambda) \sin \theta$



*Oversampling*による像の回復

散乱強度：  $I(\mathbf{Q}) = |F(\mathbf{Q})|^2 = \int d^3\mathbf{r}' \int d^3\mathbf{r} \rho(\mathbf{r}') \rho(\mathbf{r}) e^{-i\mathbf{Q}\cdot(\mathbf{r}-\mathbf{r}')}$



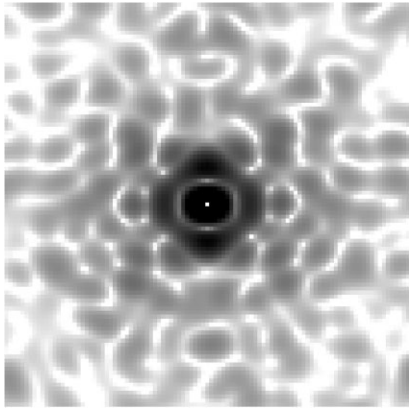
二次元検出器



# 天然変性タンパク質をターゲットとした新しい構造生物学



コヒーレント光を用いた  
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  
タンパク質構造情報が引き出せるか



## 研究 1

フェムト秒パルスレーザーによるタンパク質 1 分子からの溶液散乱

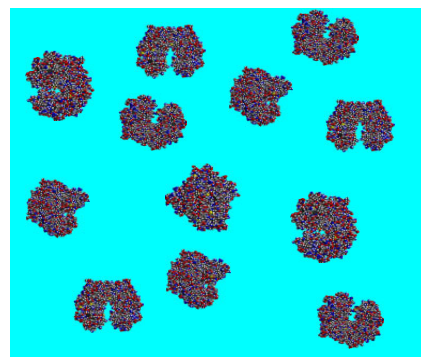
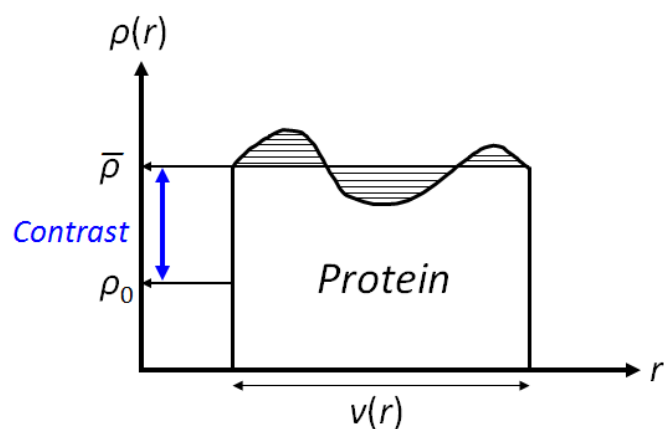
## 研究 2

少数タンパク質分子からの溶液散乱

北大・西野吉則研究室との共同研究

# Small-Angle X-ray Scattering (SAXS)

## Multi-domain proteins in solution



X-rays

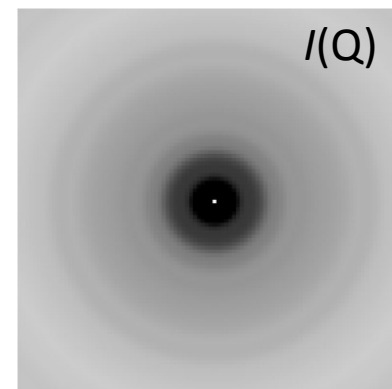
溶媒の電子密度  $\rho_0$

タンパク質  $10^{12}$  分子

散乱因子:  $F(\mathbf{Q}) = \int (\rho(\mathbf{r}) - \rho_0) e^{-i\mathbf{Q}\cdot\mathbf{r}} d^3\mathbf{r}$  (溶媒からの Contrast)

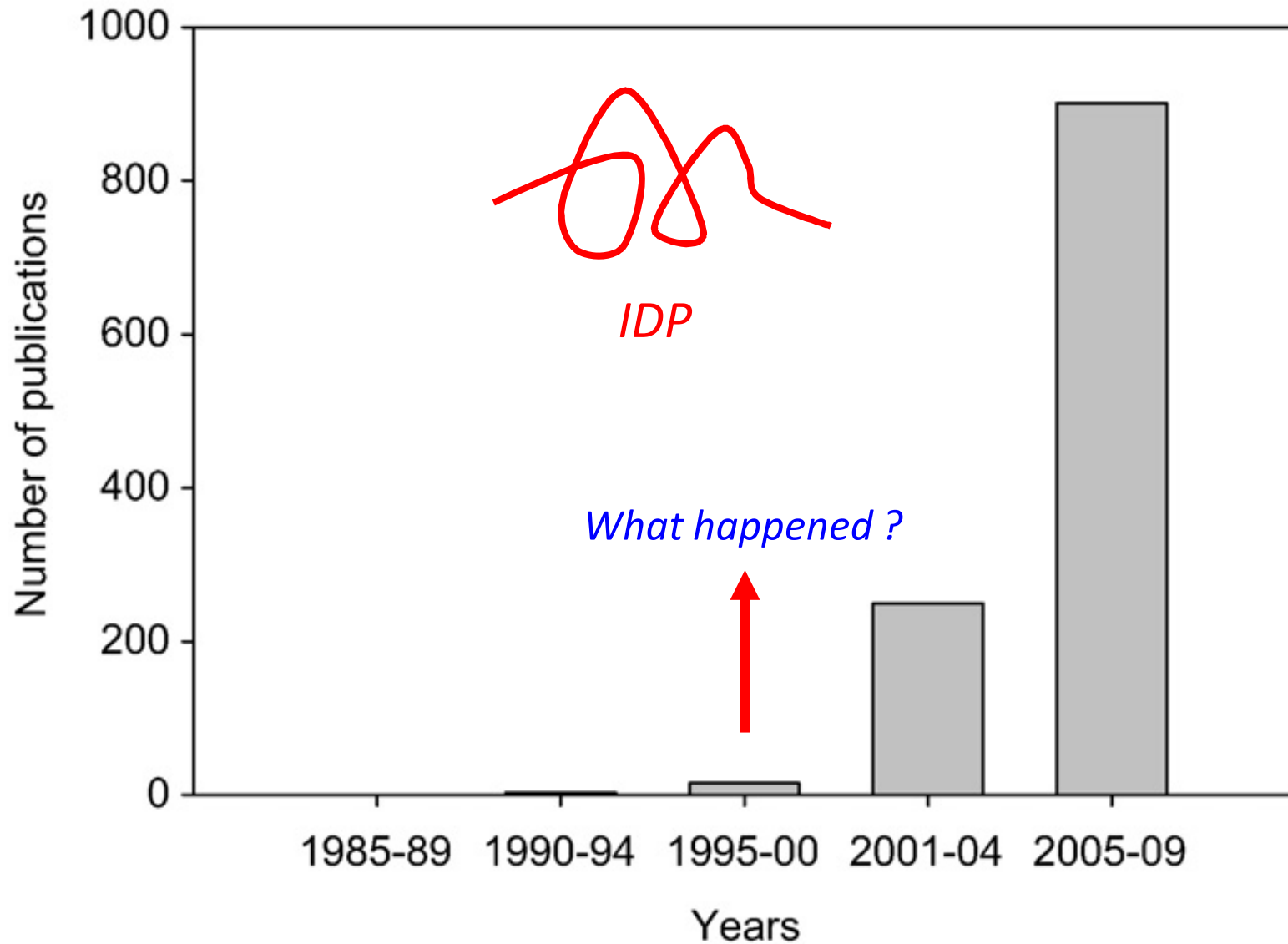
回転平均  
&  
アンサンブル平均

散乱強度:  $I(Q) = \left\langle \left\langle |F(\mathbf{Q})|^2 \right\rangle_{\Omega_{\mathbf{Q}}} \right\rangle_{\text{Ensemble}}$



二次元検出器

# No. of PubMed hits dealing with *IDPs*



*Biochim. Biophys. Acta*, **1804**, 1231-1264 (2010)

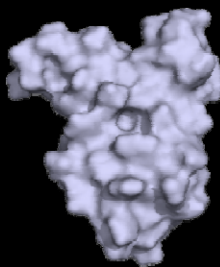


*P. Wright et al. Cell 91, 741-752 (1997)*

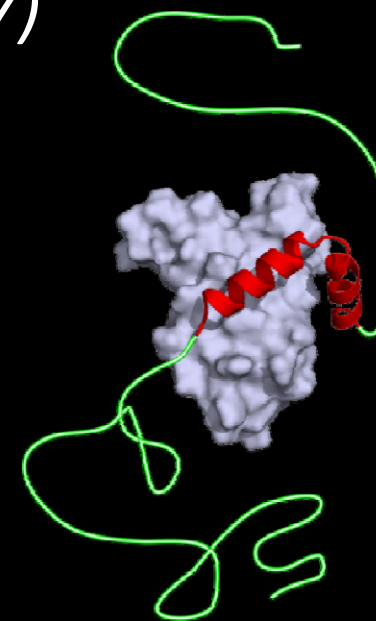


転写因子CREBの  
KIDドメイン  
(天然変性タンパク質)

+



転写のコアクチベータ  
CBPのKIXドメイン



特異的複合体

天然変性タンパク質は他のタンパク質やリガンドと  
相互作用すると立体構造が形成される

*Coupled folding & binding*

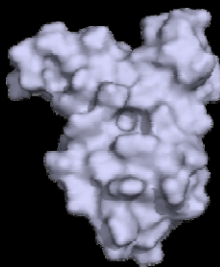


## Coupled Folding & Binding

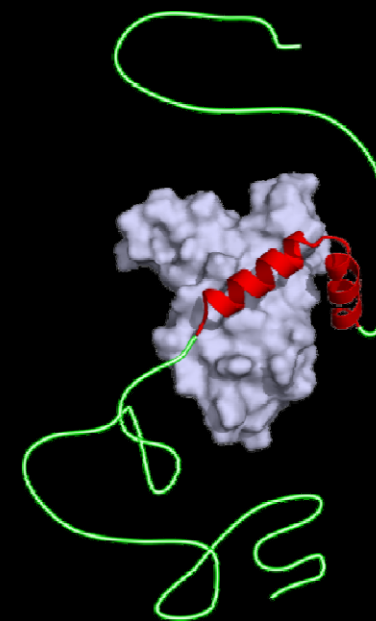


転写因子CREBの  
KIDドメイン  
(天然変性タンパク質)

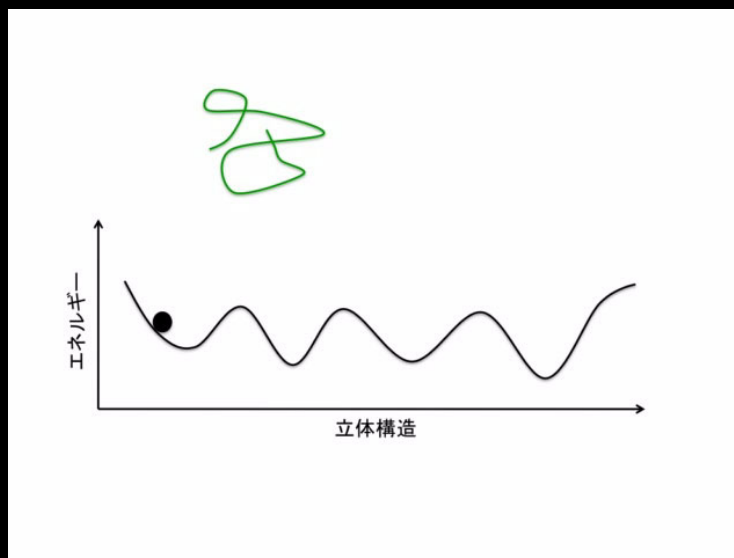
+



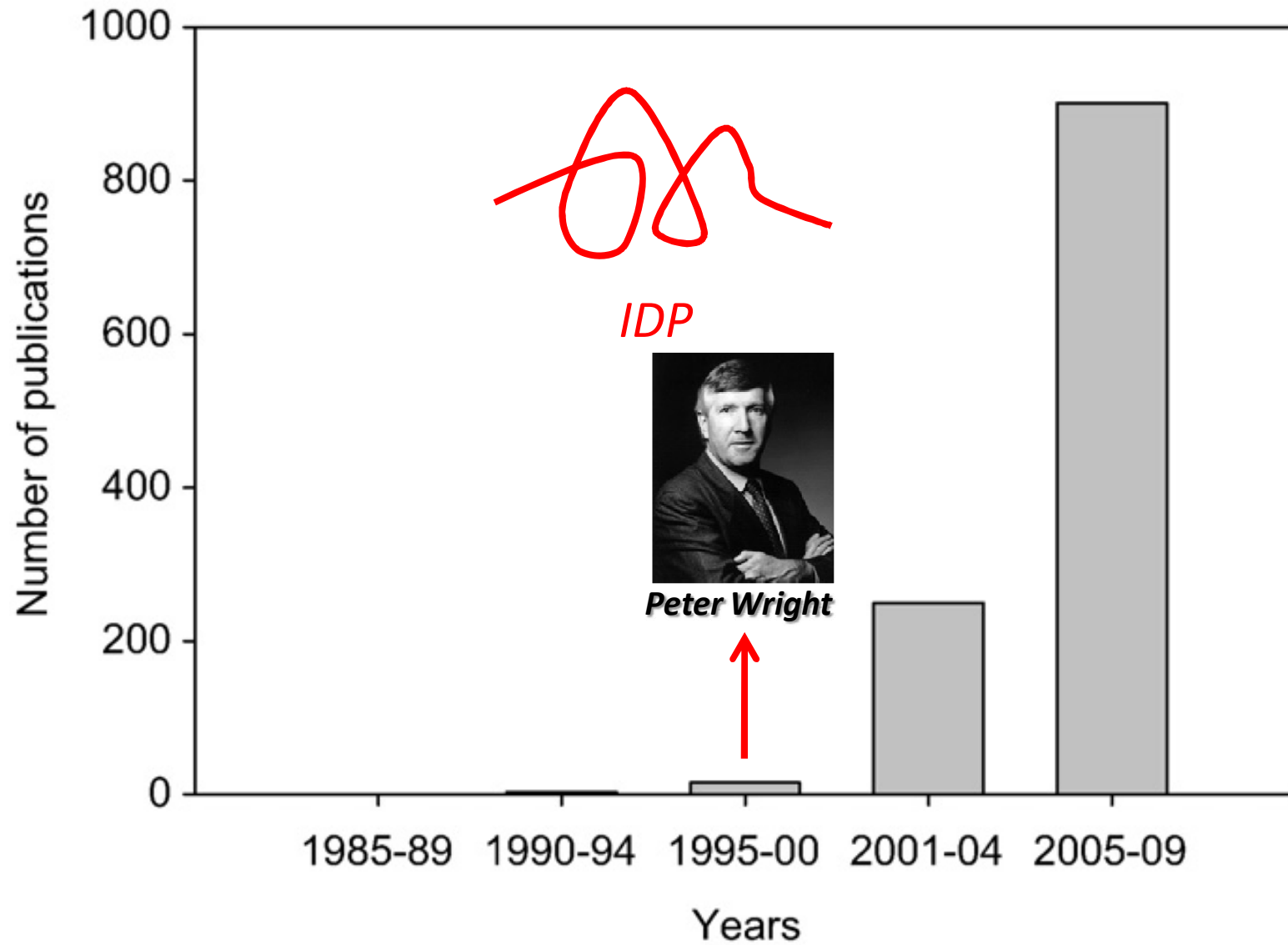
転写のコアクチベータ  
CBPのKIXドメイン



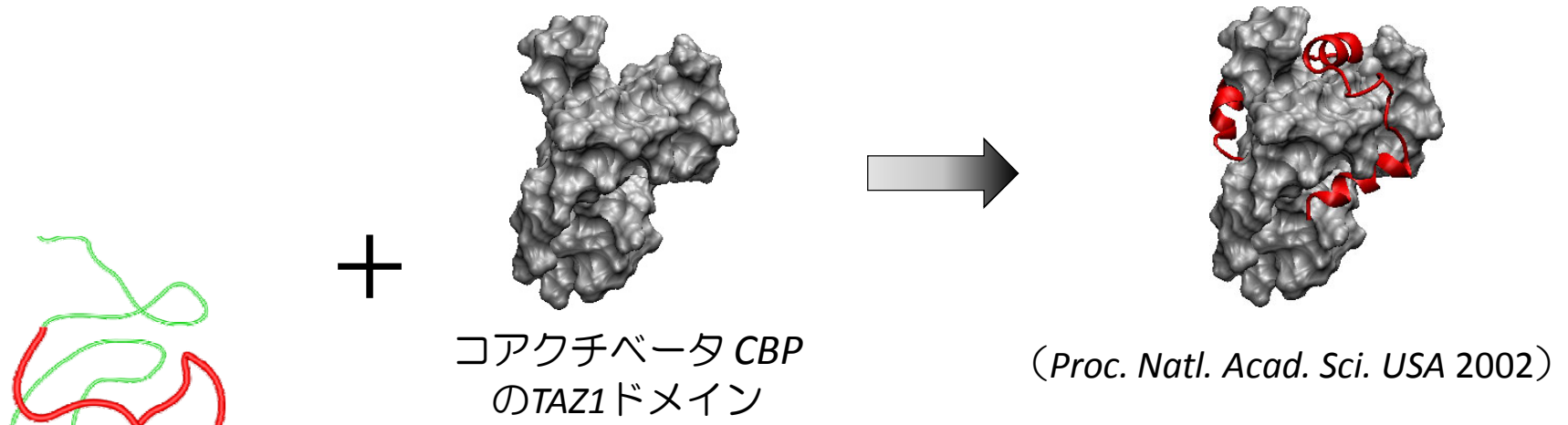
特異的複合体



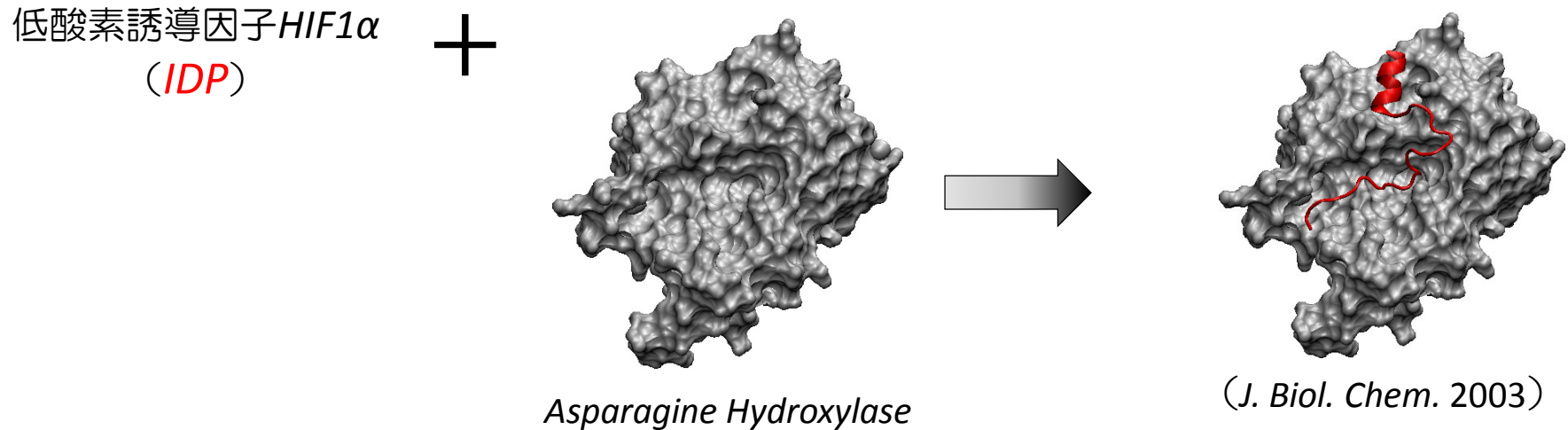
# No. of PubMed hits dealing with *IDPs*



# Intrinsically Disordered Protein: *IDP*



天然変性タンパク質は複数のタンパク質やリガンドと相互作用し  
それぞれ異なる立体構造を形成して機能する



天然変性タンパク質は生体内ネットワークにおける  
ハブ（中核）タンパク質として機能する



# A Recent Review Article on **IDP**

*Biochim. Biophys. Acta*, **1804**, 1231-1264 (2010)



ELSEVIER

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: [www.elsevier.com/locate/bbapap](http://www.elsevier.com/locate/bbapap)



Review

## Understanding protein non-folding

Vladimir N. Uversky<sup>a,b,\*</sup>, A. Keith Dunker<sup>a</sup>

<sup>a</sup> Institute for Intrinsically Disordered Protein Research, Center for Computational Biology and Bioinformatics, Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

<sup>b</sup> Institute for Biological Instrumentation, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia

### ARTICLE INFO

#### Article history:

Received 21 November 2009

Received in revised form 9 January 2010

Accepted 21 January 2010

Available online 1 February 2010

#### Keywords:

Intrinsically disordered protein

Disorder prediction

Partially folded protein

### ABSTRACT

This review describes the family of intrinsically disordered proteins, members of which fail to form rigid 3-D structures under physiological conditions, either along their entire lengths or only in localized regions. Instead, these intriguing proteins/regions exist as dynamic ensembles within which atom positions and backbone Ramachandran angles exhibit extreme temporal fluctuations without specific equilibrium values. Many of these intrinsically disordered proteins are known to carry out important biological functions which, in fact, depend on the absence of a specific 3-D structure. The existence of such proteins does not fit the prevailing structure–function paradigm, which states that a unique 3-D structure is a prerequisite to function. Thus, the protein structure–function paradigm has to be expanded to include intrinsically disordered proteins and alternative relationships among protein sequence, structure, and function. This shift in the paradigm represents a major breakthrough for biochemistry, biophysics and molecular biology, as it opens new levels of understanding with regard to the complex life of proteins. This review will try to answer the following questions: how were intrinsically disordered proteins discovered? Why don't these proteins fold? What is so special about intrinsic disorder? What are the functional advantages of disordered proteins/regions? What is the functional repertoire of these proteins? What are the relationships between intrinsically disordered proteins and human diseases?

© 2010 Elsevier B.V. All rights reserved.



# Gordon Research Conference on **IDP**

The screenshot shows a Windows Internet Explorer browser window displaying the Gordon Research Conferences website. The address bar shows the URL: <http://www.grc.org/programs.aspx?year=2010&program=intrinsic>. The page title is "Gordon Research Conferences - 2010 Program (Intrinsically Disordered Proteins)".

The website header features the GRC logo and the text "Gordon Research Conferences" and "Conference Program".

The main content area is titled "Intrinsically Disordered Proteins" and includes the following information:

- Introducing Unfoldome and Unfoldomics**
- July 11-16, 2010
- Davidson College
- Davidson, NC
- Chairs: **Vladimir Uversky & A. Keith Dunker**
- Vice Chairs: **Rohit V. Pappu & Peter Tompa**

Recent studies revealed that functional proteins without unique 3-D structures are highly abundant in nature. These intrinsically disordered proteins (IDPs) possess a number of crucial biological functions that are complementary to functions of structured (ordered) proteins. In any given organism, IDPs constitute a functionally broad and densely populated unfoldome; i.e., a set of unstructured proteins in a proteome. Being structurally and functionally very different from ordered proteins, IDPs require special experimental and computational tools for their identification and analyses. These specific investigative approaches underlie unfoldomics.

IDPs are common across the three domains of life, being especially abundant in the eukaryotic proteomes. Signaling sequences and sites of posttranslational modifications are located within regions of intrinsic disorder. Disorder-to-order transitions in an IDP are coupled with the adoption of different structures in complexes with different partners. The intrinsic flexibility of IDPs helps different disordered regions to bind to a common binding site on a common partner. This binding diversity plays important roles in both protein-protein interaction networks and likely also in gene regulation networks. Such disorder-based signaling is further modulated in multicellular eukaryotes by alternative splicing, for which splicing events map to regions of disorder much more often than to regions of structure. The combination of disorder and alternative splicing is proposed to provide a mechanism for easily "trying out" different signaling pathways, thereby providing the mechanism for generating signaling diversity and enabling the evolution of cell differentiation and multicellularity. Finally, several small molecules-potential drugs have been shown to act by blocking protein-protein interactions involving intrinsic disorder of one of the partners.

The website also includes a navigation menu on the left with links to HOME, CONFERENCES (Current Meetings (2010), Upcoming Meetings (2011), Past Meetings, Conference Portfolio, Proposing a New Gordon Conference), FOR ATTENDEES, ABOUT GRC, THE ORGANIZATION, FOR CHAIRS, MISCELLANEOUS, and a QUICK SEARCH box with a GO button and a link to [advanced search].

On the right side, there are buttons for MEETING LINKS (Conference History, Chair Contact Info), SITE & TRAVEL LINKS, MEETING FEES, and a link to What is RSS?

# Functional Repertoires of *IDPs*

## *IDPs* as hubs in protein signaling networks

*p53* regulates over 150 gene transcriptions (*p21*, *MDM2*, *BAX* ....)

## *IDPs* and PTMs (post-translational modifications)

PTMs frequently occur in ID regions

Substrates for many PM enzymes such as kinases are ID proteins

## *IDPs* and AS (alternative splicing)

AS region of mRNA code for ID regions

AS is likely to be involved in cell differentiation

## *IDPs* are mainly found in eukaryotes

Eukaryotes contain far more intrinsic disorder than prokaryotes

## *IDPs* and human diseases

Alzheimer disease and amy- $\beta$ , tau,  $\alpha$ -synuclein

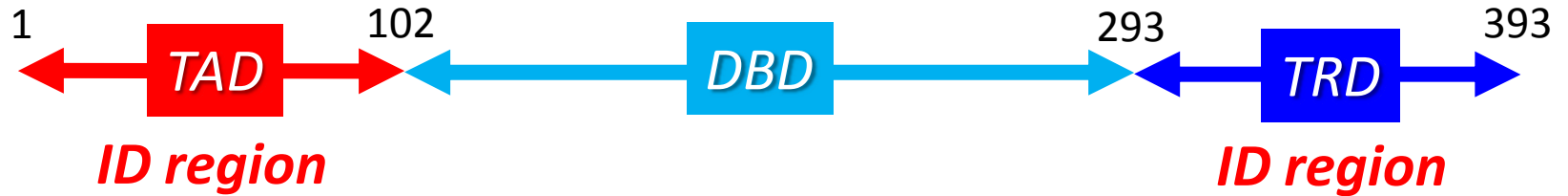
## *IDPs* as novel drug targets

One partner is disordered and the second is structured

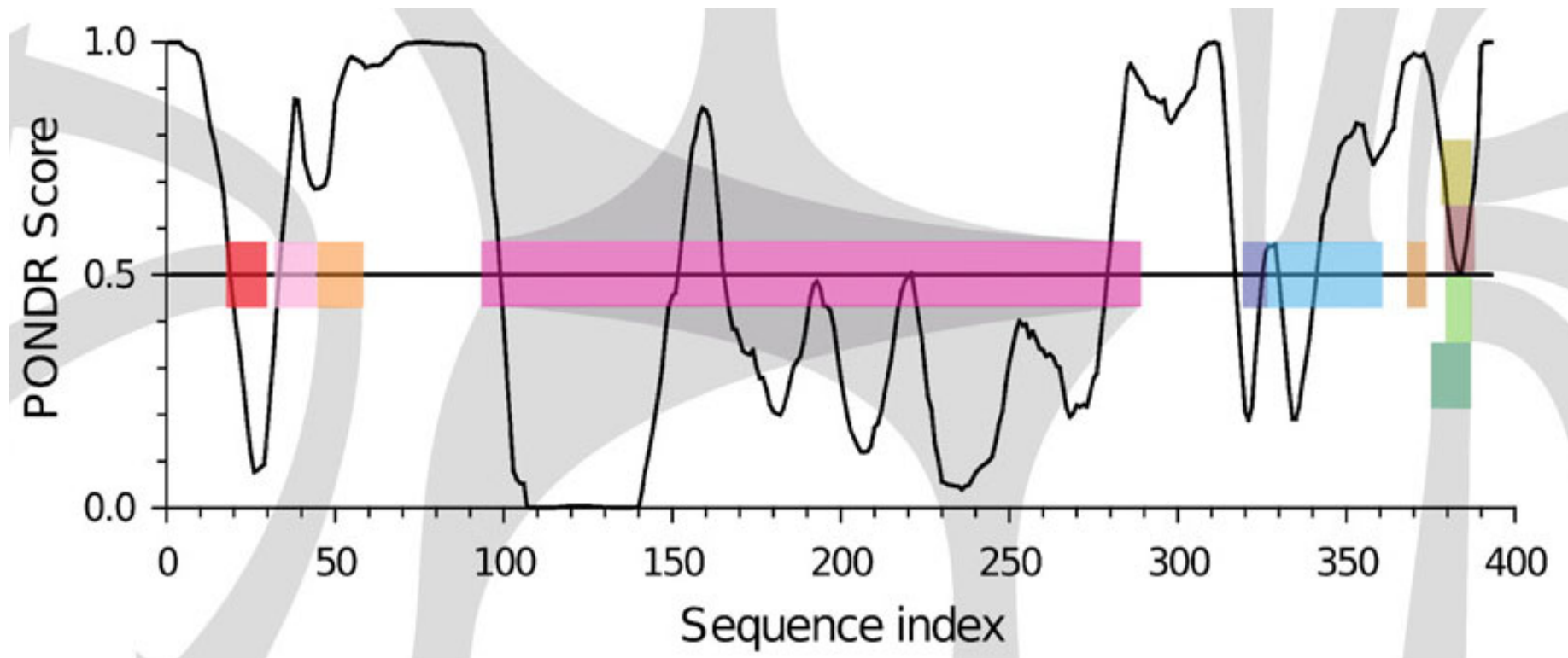
Both partners lacks fixed structures

$D^2$ -concept: disorder in disorders

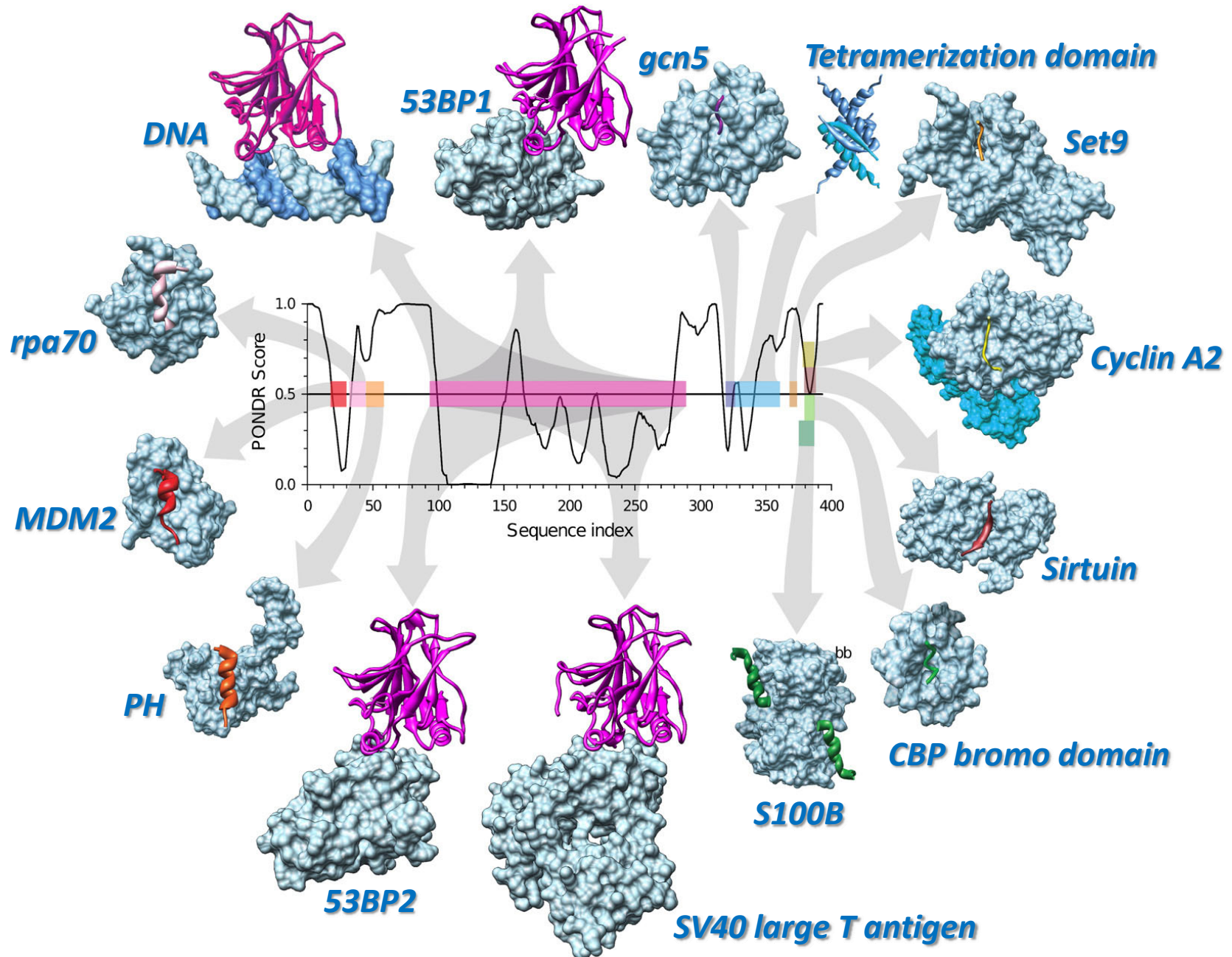
# Domain Structure of p53



## ID Prediction of p53 by PONDR

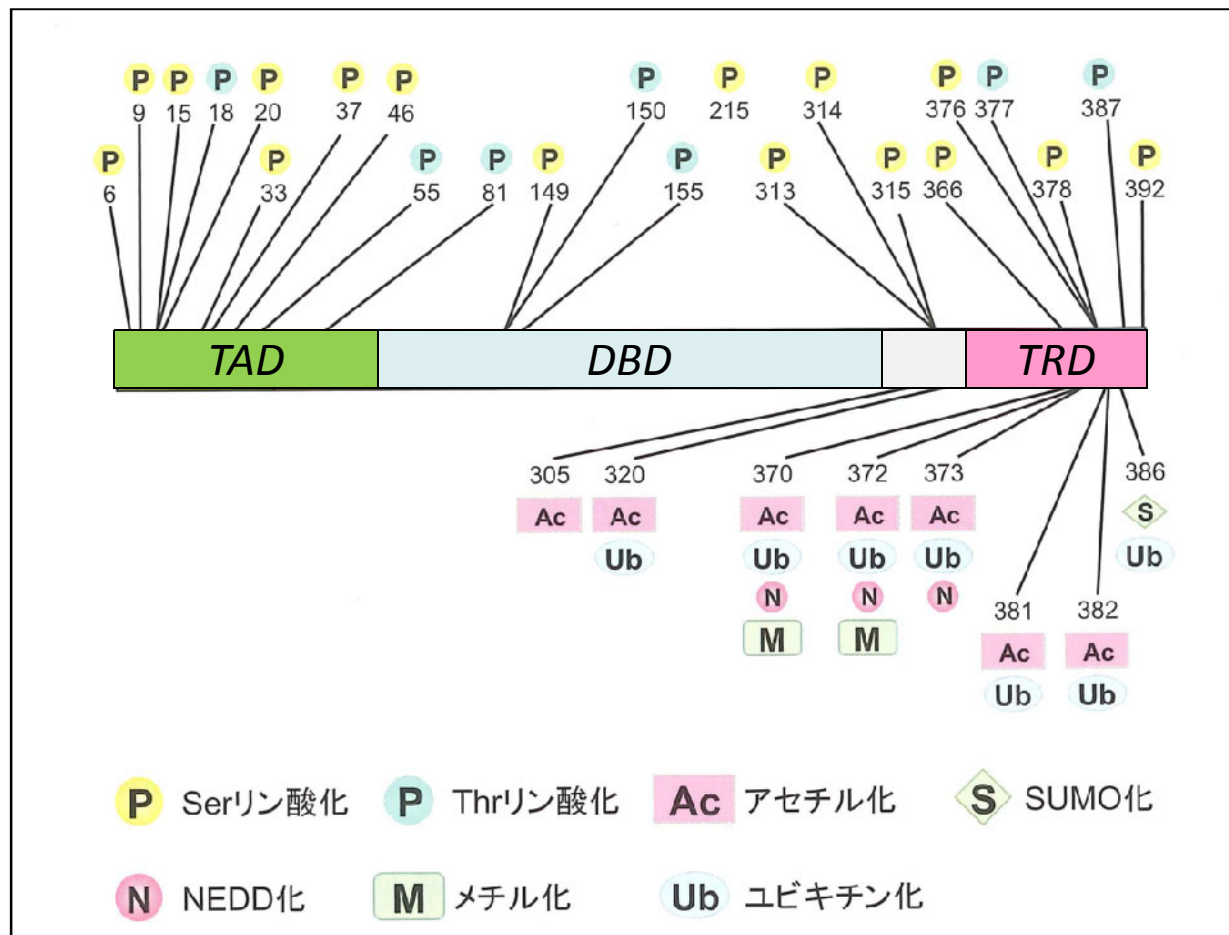


# *p53* interacts with 14 different partners



# PTMs frequently occur in **ID** regions

## PTMs of p53



R. Kamata & K. Sakaguchi, *Seikagaku*, **82**, 484-493 (2010)  
(Partially modified)



# PTMs frequently occur in ID regions

「先端融合領域イノベーション創出拠点の形成」プログラム 平成22年度再審査 コメント(継続課題) : 文部科学省 - Windows Internet Explorer

http://www.mext.go.jp/b\_menu/houdou/23/01/attach/1300988.htm

お気に入り | ホーム 社団法人 日本農... | ホーム 社団法人 日本農... | おすすめサイト | 本日のおすすめアド...

「先端融合領域イノベーション創出拠点の形成...」

文部科学省 MINISTRY OF EDUCATION, CULTURE, SPORTS, SCIENCE AND TECHNOLOGY-JAPAN

平成23年1月21日発表

文字サイズの変更 小 中 大

検索 詳細検索

お知らせ 政策について 白書・統計・出版物 申請・手続き 文部科学省について 教育 科学技術・学術 スポーツ 文化

トップ > お知らせ > 報道発表 > 平成22年度の報道発表 > 科学技術振興調整費「先端融合領域イノベーション創出拠点の形成」プログラム 平成22年度再審査結果の公表について > 「先端融合領域イノベーション創出拠点の形成」プログラム 平成22年度再審査 コメント(継続課題)

## ●「先端融合領域イノベーション創出拠点の形成」プログラム 平成22年度再審査 コメント(継続課題)

### 【継続課題】横浜市立大学「翻訳後修飾プロテオミクス医療研究拠点の形成」

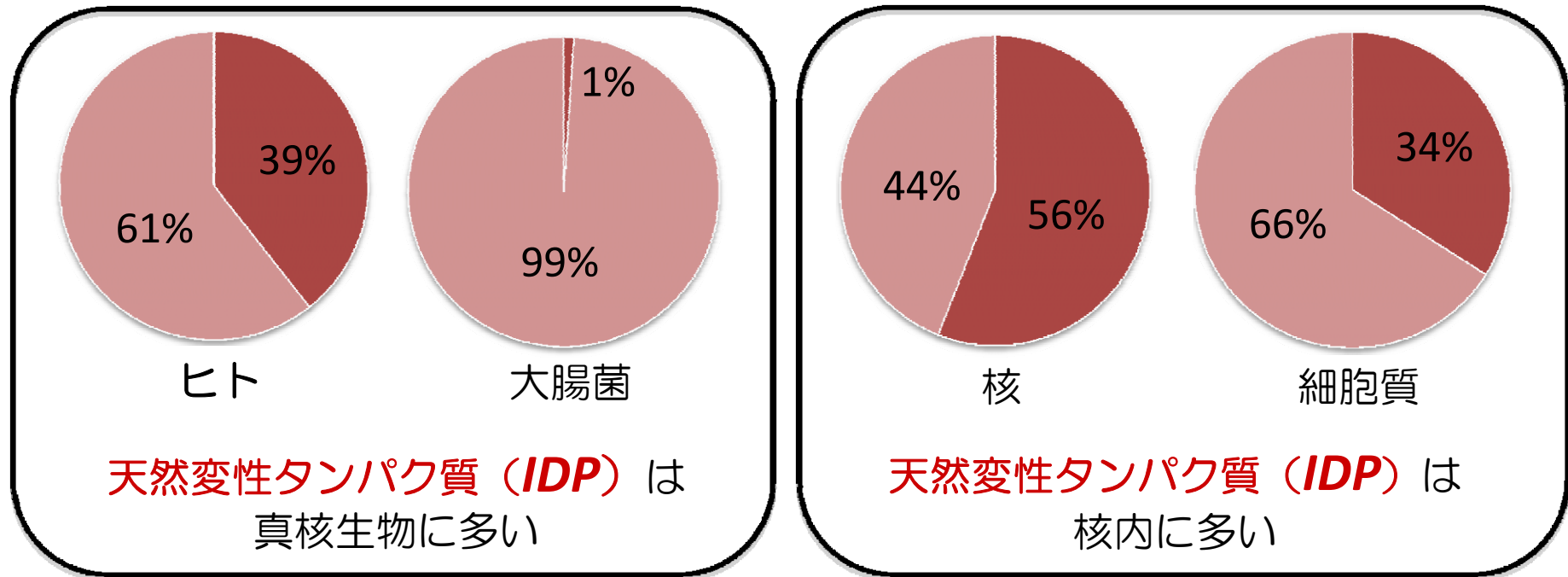
本課題は、タンパク質解析技術を基盤として、翻訳後修飾と疾患の関係を明らかにし、診断マーカーの発見や創薬を産学協働で継続的に実施する拠点形成を目指す取組である。本拠点にはタンパク質の翻訳後修飾の解析成果を医療・創薬・機能性食品・化粧品などへ応用するための協働機関が参加しており、高度なプロテオミクス拠点を形成して産業界の開発研究の推進を図るために必要な解析技術を揃えている。特に大規模分析による診断マーカー探索とそれに基づく診断薬、治療薬の開発は一定の競争力を持つことから、事業化を促進する研究拠点の形成が期待される。プロジェクト全体としての構想や今後の方向性が明確で、協働機関ともよく連携して活動が進められており、協働機関の大学に対する評価や期待も大きい。本課題については、これまでの進捗状況及び今後の見通しから判断して、本プログラムの趣旨に合致した成果を実現することが十分期待されるため、来年度から本格的実施に移行することが適当である。

#### (1) 目標達成度(1. ミッションステートメントの達成度、2. 研究・技術開発の達成度、3. システム改革の進捗状況、4. 人材育成の進捗状況)

タンパク質の翻訳後修飾の解析と医療への応用について、これまで集中的に研究が進められてきたリン酸化ペプチドに関する学術論文の質・量が共に優れており、学術的な研究成果として高く評価される。修飾タンパク質の質量分析等の要素技術基盤の確立に向けた着実な進捗が認められる等の目標は達成されている。

#### (2) 協働機関との関係

## *IDPs are mainly found in eukaryotes*



*Fukuchi et al. unpublished data*

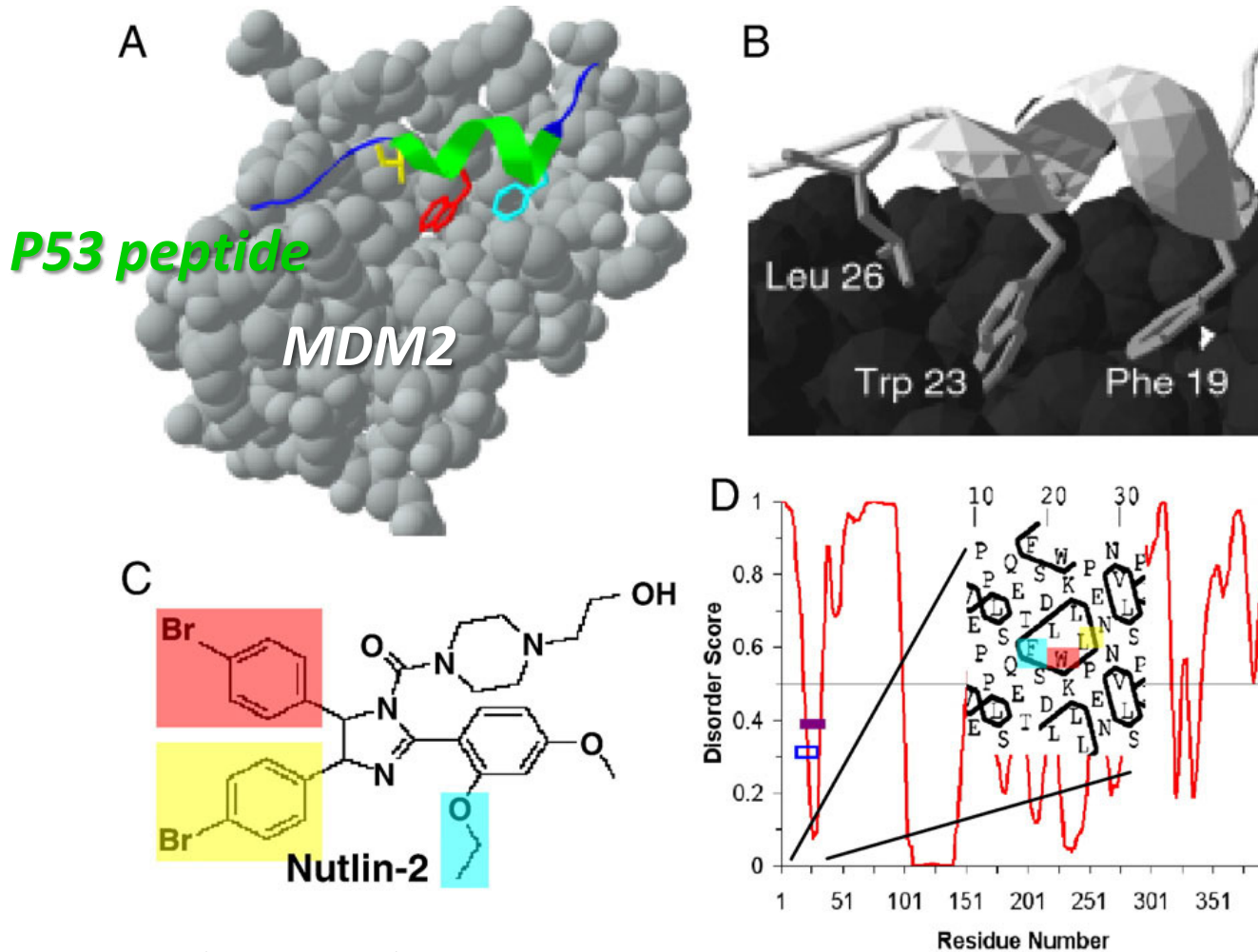
DNAの複製・転写・修復・組換えやヒストン修飾などに関する核内タンパク質の多くは分子内に長大な天然変性領域をもっている

Minezakiらは、膜タンパク質において天然変性領域が膜の細胞質側に顕著に見られることを発見した。

# IDPs as novel drug targets

One partner (p53) is *disordered*

& the second (E3: MDM2) is *structured*



Nutlin-2 (ヌトリン-2) : MDM2 Antagonist

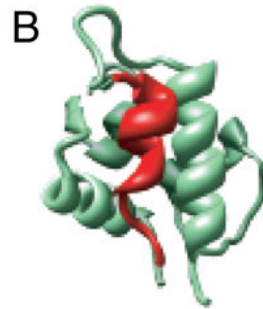
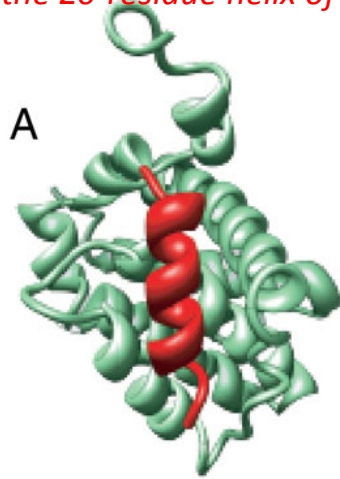


# *IDPs as novel drug targets*

*One partner is disordered & the second is structured*

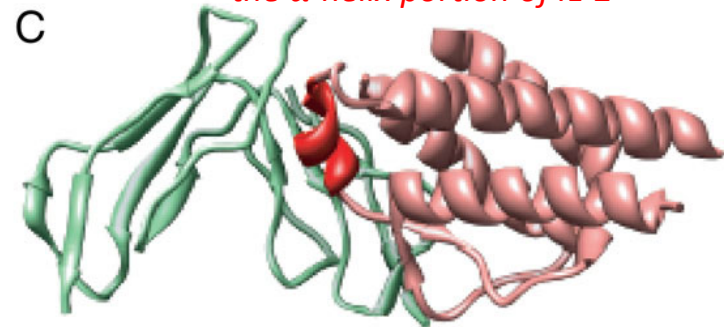
## *Bcl-xL & BAK fragment*

*Small molecules were designed based on the 20-residue helix of BAK*

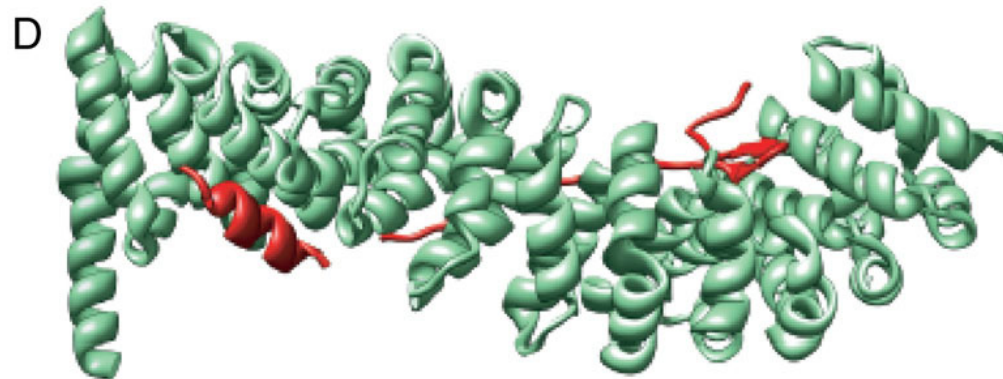


## *IL-2 receptor $\alpha$ & IL-2*

*Small molecules were designed based on the  $\alpha$ -helix portion of IL-2*

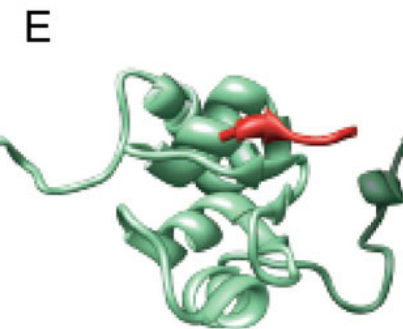


## *MDM2 & P53 fragment*



## *$\beta$ -catenin & T cell factor*

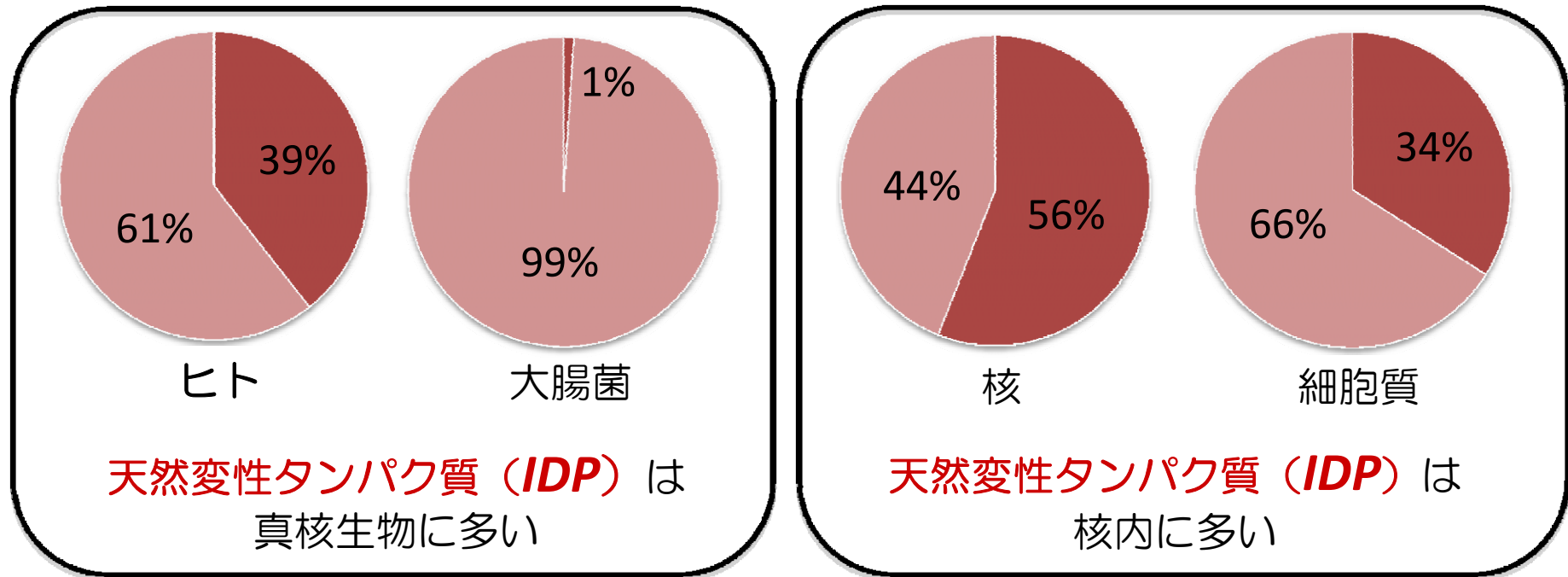
*Binding sites of small molecules were not clear*



## *XIAP & Smac fragment*

*Small molecules were designed based on the  $\beta$ -strand fragment (AVPIAQKSE) of Smac*

## *IDPs are mainly found in eukaryotes*



*Fukuchi et al. unpublished data*

DNAの複製・転写・修復・組換えやヒストン修飾などに関する核内タンパク質の多くは分子内に長大な天然変性領域をもっている

*80% of the proteins in the human cancer-associated proteins database contain ID region (The Scientist 23, 47, 2009)*

# Functional Repertoires of *IDPs*

## *IDPs* as hubs in protein signaling networks

*p53* regulates over 150 gene transcriptions (*p21*, *MDM2*, *BAX* ....)

## *IDPs* and PTMs (post-translational modifications)

PTMs frequently occur in ID regions

Substrates for many PM enzymes such as kinases are ID proteins

## *IDPs* and AS (alternative splicing)

AS region of mRNA code for ID regions

AS is likely to be involved in cell differentiation

## *IDPs* are mainly found in eukaryotes

Eukaryotes contain far more intrinsic disorder than prokaryotes

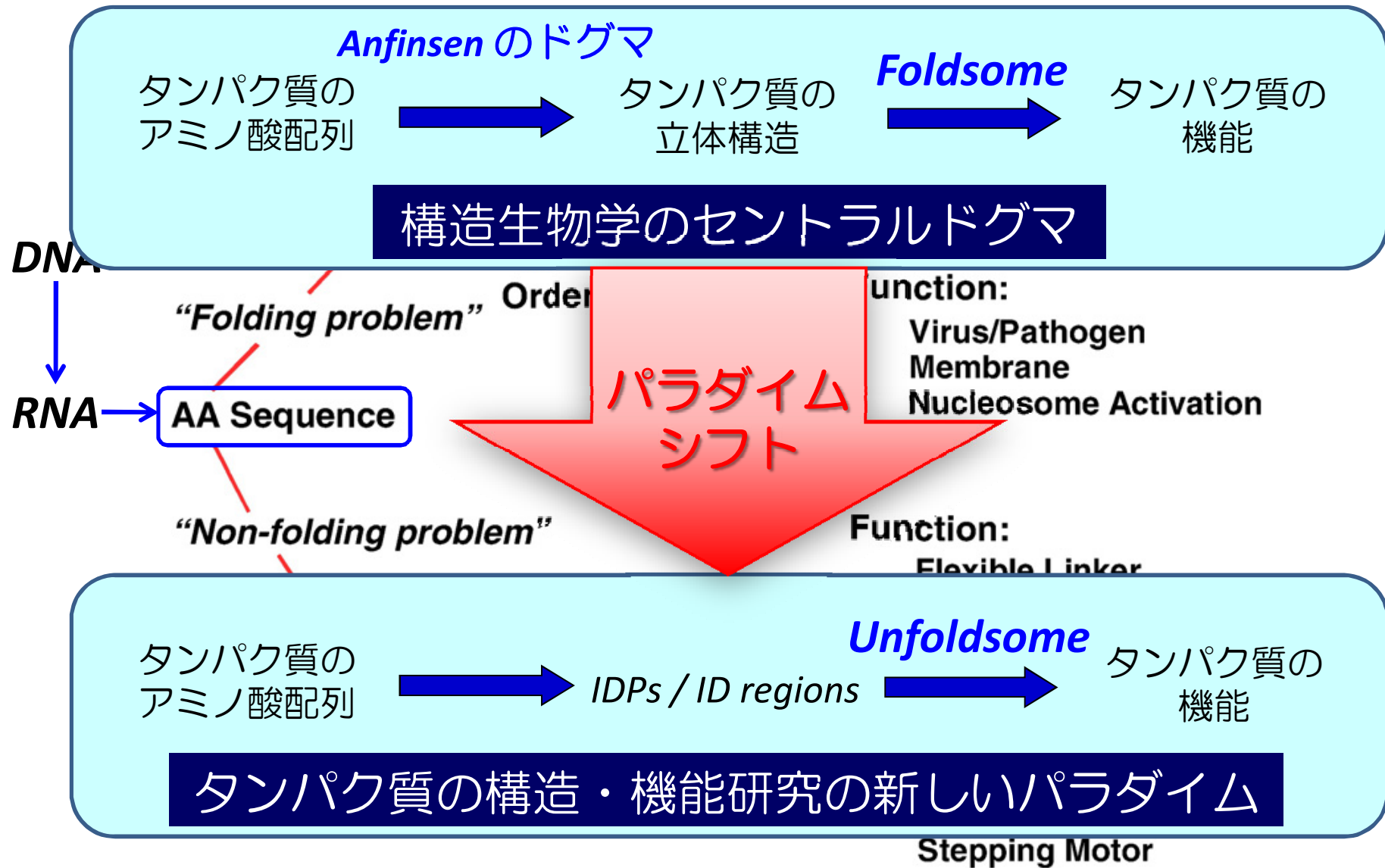
## *IDPs* and human diseases

Alzheimer disease and amy- $\beta$ , tau,  $\alpha$ -synuclein

## *IDPs* as novel drug targets

One partner is disordered and the second is structured

# Protein Structure-Function Analysis



Biochim. Biophys. Acta, **1804**, 1231-1264 (2010)



平成21年度（2009年度） 科研費  
「新学術領域研究」

# 天然変性タンパク質の 分子認識機構と機能発現

（略称：天然変性蛋白質）

研究期間：平成21年度～平成25年度

---

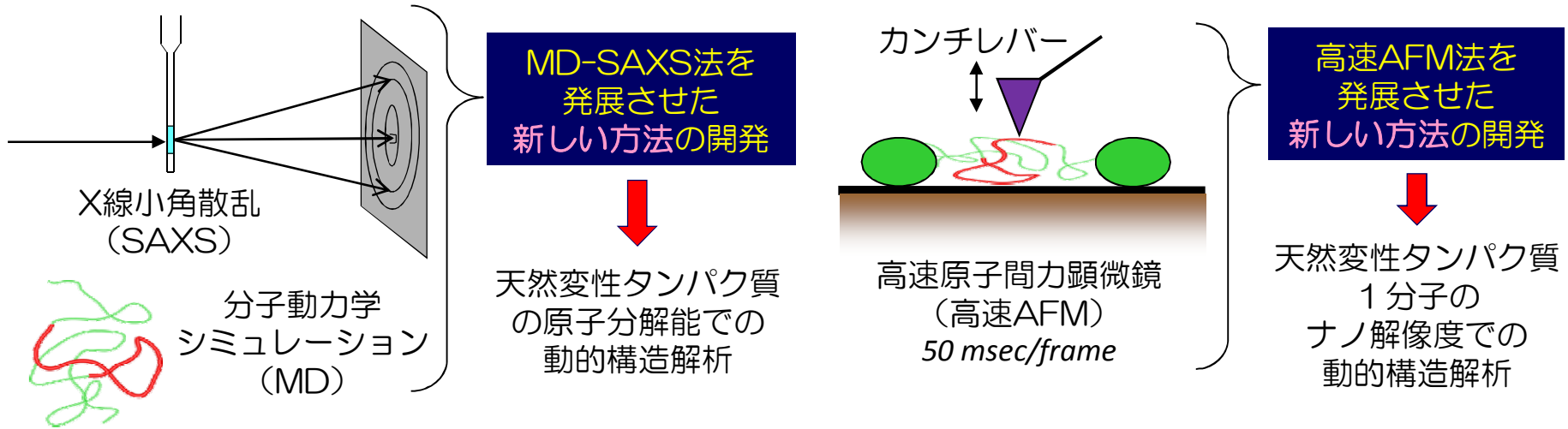
領域代表者：佐藤 衛

（横浜市立大学 大学院 生命ナノシステム科学研究科）

# 研究グループの研究内容

構造生物学グループ **動的構造解析**  
X線小角散乱, 高速AFM, NMR, 質量分析で天然変性タンパク質の動的構造を決定

天然変性タンパク質の動的構造解析はX線結晶構造解析法などの従来手法だけでは困難で、**新しい手法**の開発が不可欠



**分子生物学グループ** **機能解析**  
分子生物学、生化学、分子遺伝学的手法を用いて核内でDNA複製・転写・組換え・修復に関与する天然変性タンパク質の機能を解明

**情報生物学グループ** **構造・機能予測**  
分子動力学シミュレーションなどの計算科学的手法および質・量ともに充実したデータベースを構築して天然変性タンパク質の構造・機能を予測



# Home Page of IDP Research Project in Japan

The image shows a screenshot of a Microsoft Internet Explorer browser window. The address bar displays the URL <http://www.tsurumi.yokohama-cu.ac.jp/IDP/>. The browser window is titled "天然変性タンパク質の分子認識機構と機能発現 - Microsoft Internet Explorer".

In the background, a Google search page is visible. The search term "天然変性蛋白質" is entered in the search box. The search results show the top result as "天然変性タンパク質の分子認識機構と機能発現" from the website [www.tsurumi.yokohama-cu.ac.jp/IDP/](http://www.tsurumi.yokohama-cu.ac.jp/IDP/).

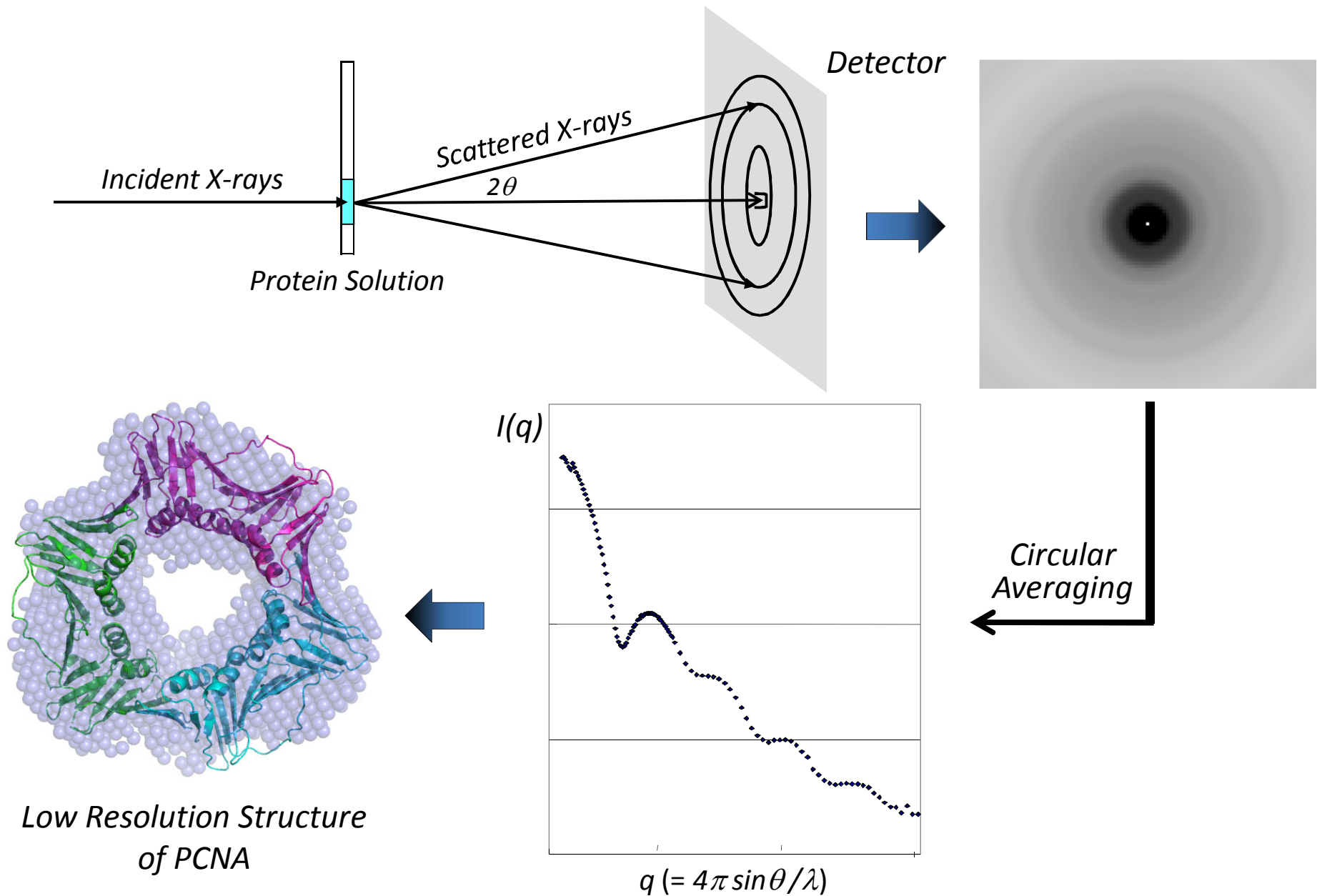
The main content of the IDP website is displayed in the foreground. The header includes the IDP logo and the text "文部科学省科学研究費補助金 (新学術領域研究) 天然変性タンパク質の分子認識機構と機能発現". Below the header, a large banner features the text "当領域の活動状況、シンポジウムのお知らせなどの最新情報をメールマガジン (季刊および速報) にてお知らせします。下記の要領でメールアドレスを登録して下さい。" and the acronym "INTRINSICALLY DISORDERED PROTEINS".

On the left side of the website, there is a navigation menu with the following items: HOME, ニュース, 研究, 概要, and 計画研究. A "Summer School 2010" banner is also present.

On the right side, there is a "メールマガジン" (Email Magazine) button with the text "配信申し込みはこちら".

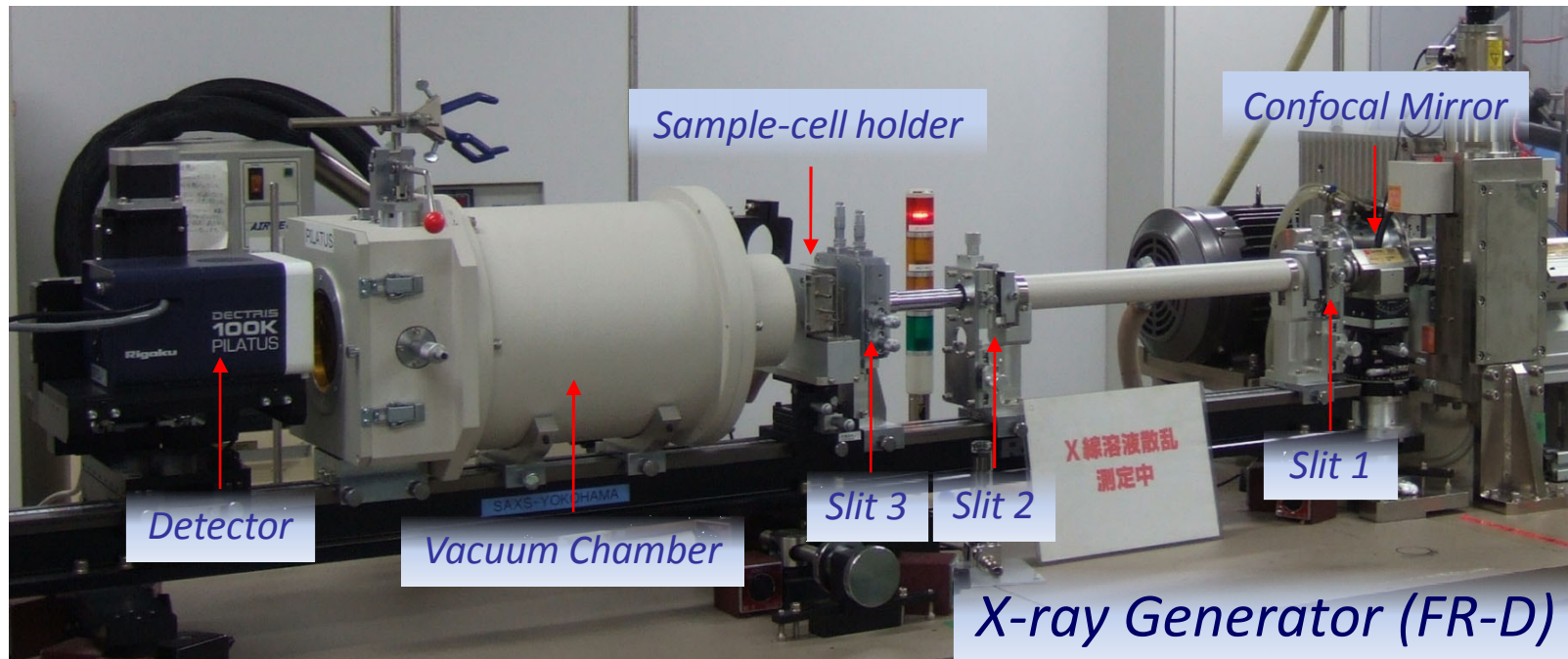
The bottom of the browser window shows the Windows taskbar with the Start button, several open applications, and the system clock displaying 13:24.

# Small-Angle X-ray Scattering (SAXS)



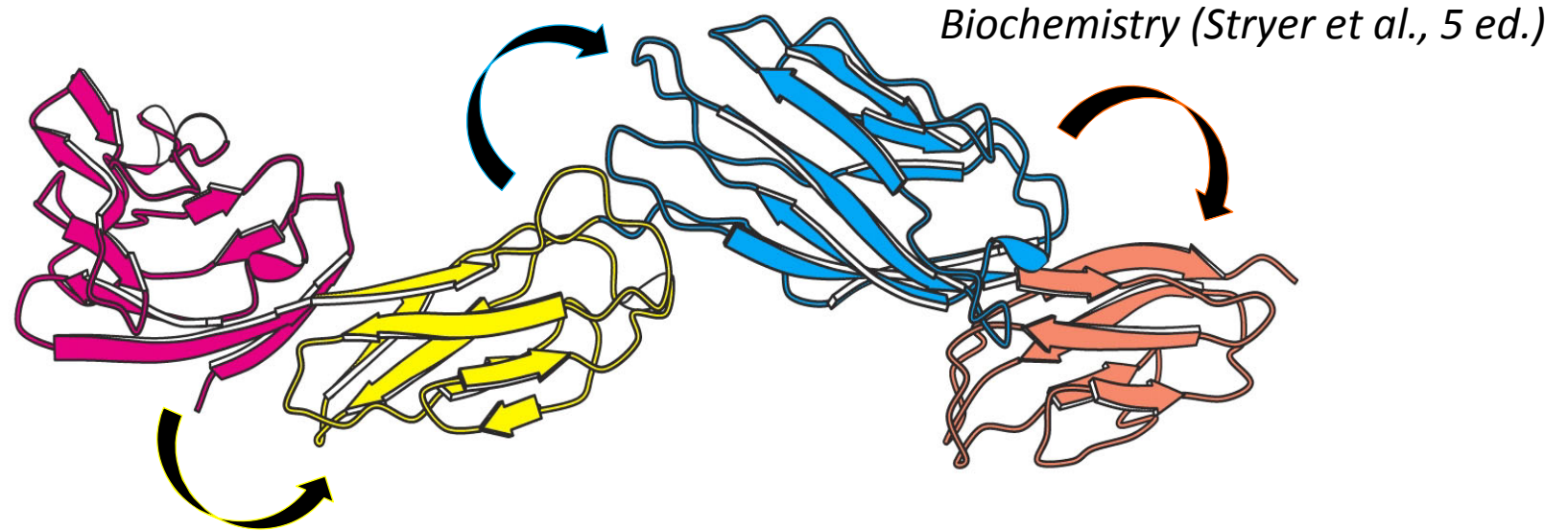


# SAXS in Yokohama City University



<i>Detector</i>	<i>PILATUS 100K</i>
<i>Energy range</i>	<i>3 – 30 keV</i>
<i>Active area</i>	<i>83.8 x 33.5 mm</i>
<i>Pixel number</i>	<i>512 x 256</i>
<i>Efficiency</i>	<i>100% at 8keV 35% at 17keV</i>
<i>Spatial resolution</i>	<i>172 <math>\mu</math>m</i>

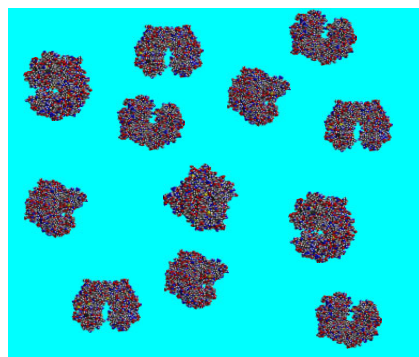
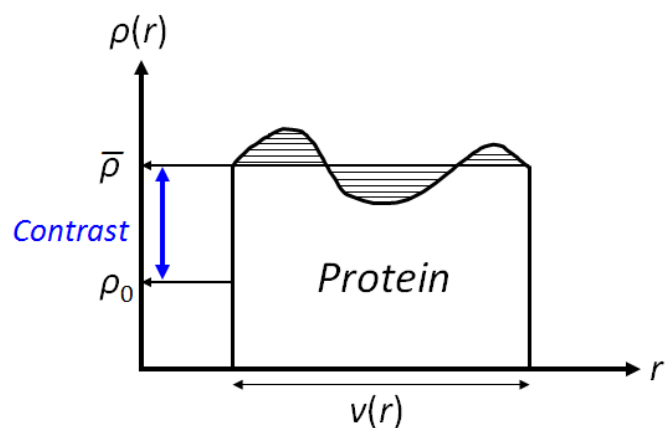
# Small-Angle X-ray Scattering (SAXS)



*Most of multi-domain proteins are  
largely fluctuated in solution*

# Small-Angle X-ray Scattering (SAXS)

## Multi-domain proteins in solution



X-rays

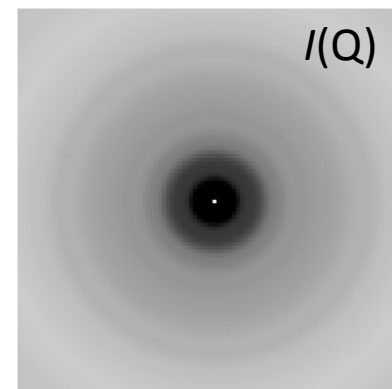
溶媒の電子密度  $\rho_0$

タンパク質  $10^{12}$  分子

散乱因子:  $F(\mathbf{Q}) = \int (\rho(\mathbf{r}) - \rho_0) e^{-i\mathbf{Q}\cdot\mathbf{r}} d^3\mathbf{r}$  (溶媒からの Contrast)

回転平均  
&  
アンサンブル平均

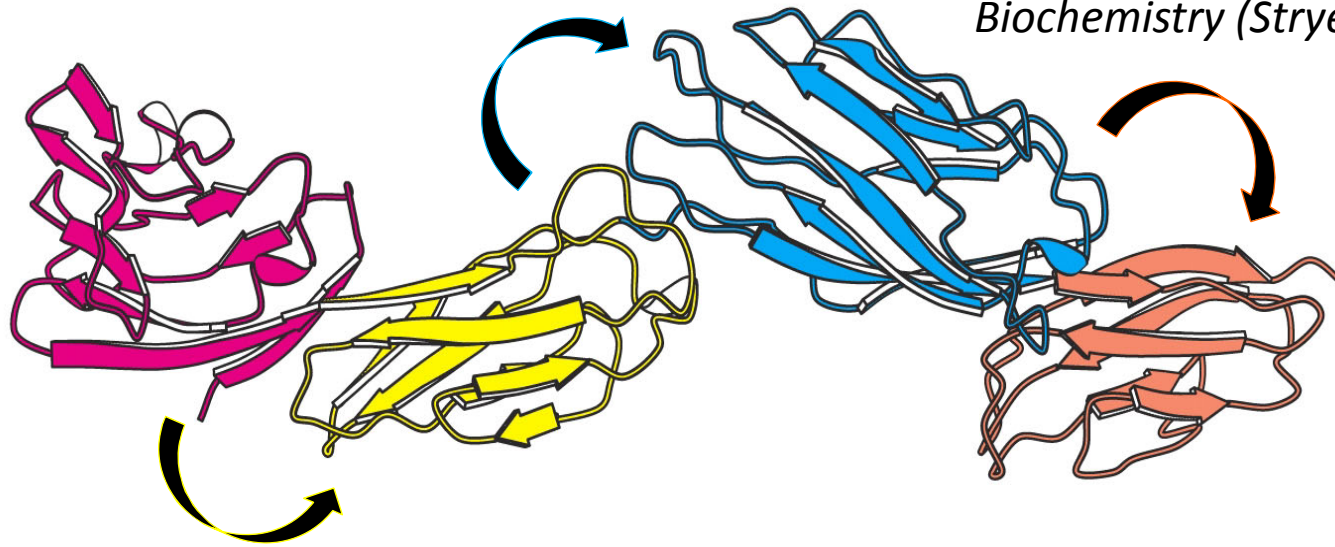
散乱強度:  $I(Q) = \left\langle \left\langle |F(\mathbf{Q})|^2 \right\rangle_{\Omega_{\mathbf{Q}}} \right\rangle_{\text{Ensemble}}$



二次元検出器

# Small-Angle X-ray Scattering (SAXS)

*Biochemistry (Stryer et al., 5 ed.)*

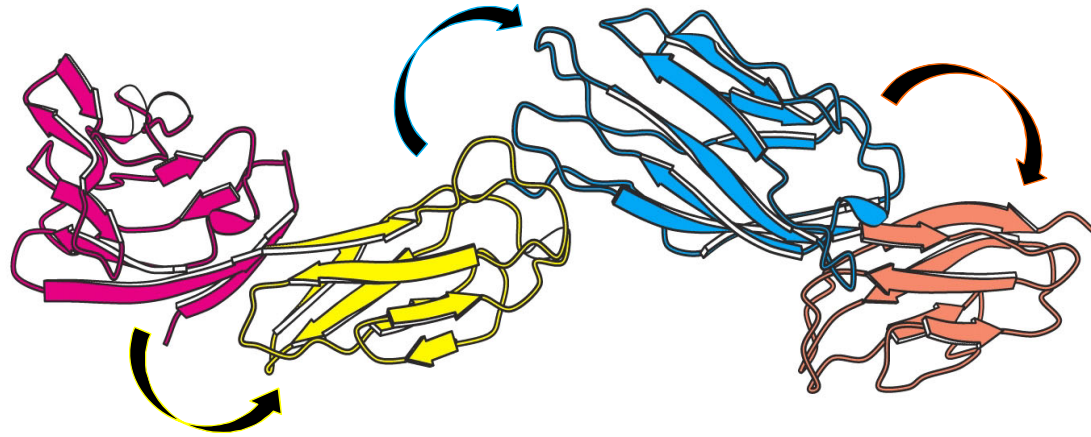


*Most of multi-domain proteins are  
largely fluctuated in solution*

*Time-averaged*



*No structural & functional information !*



*Develop “MD-SAXS” to analyze the dynamical structure of multi-domain protein in solution*



*MD simulation of the multi-domain protein*

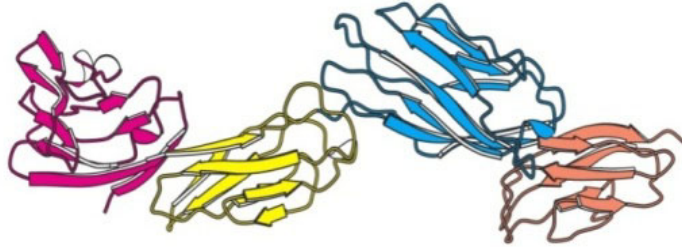
*The initial structure is constructed from the crystal structure of each domain*



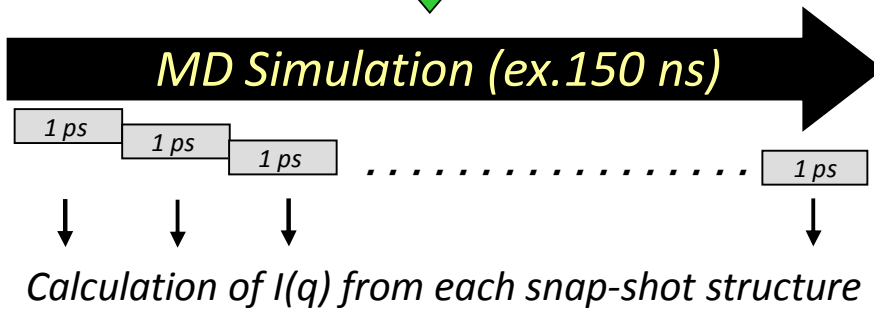
*The result is assessed experimentally by SAXS*

# Principle of MD-SAXS

## Construction of Initial Atomic Structure

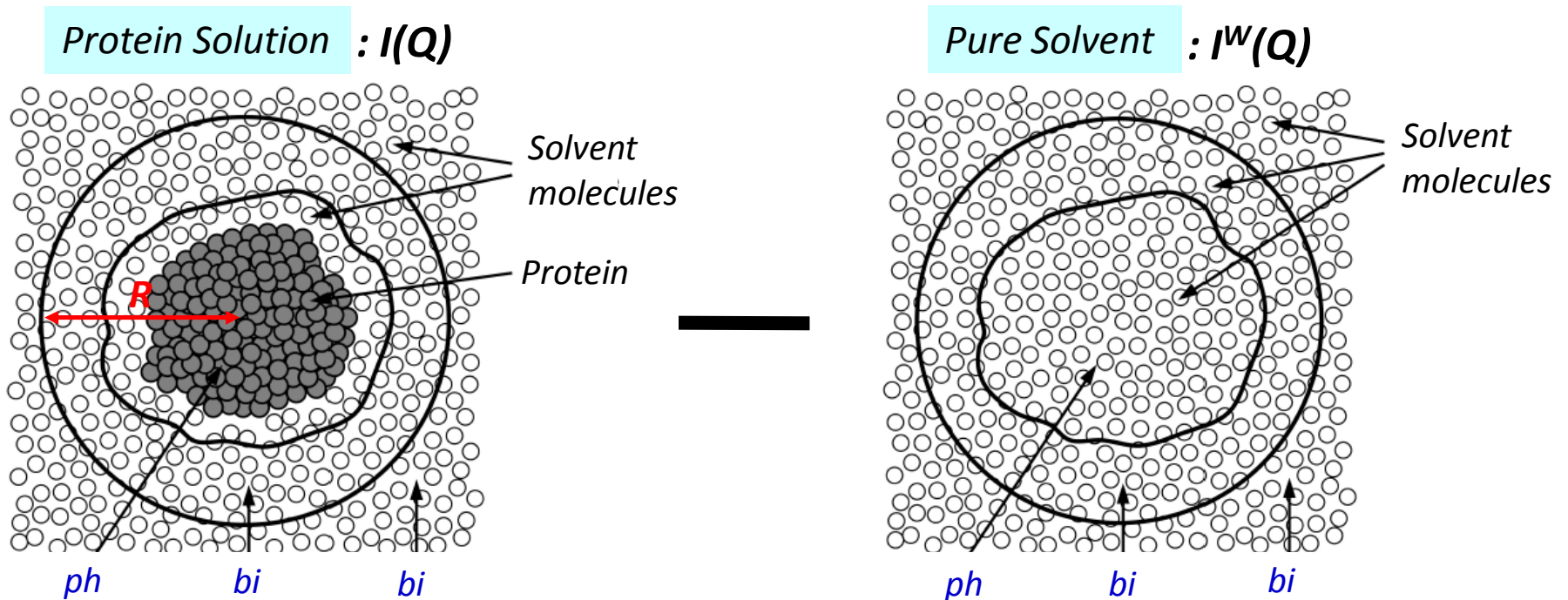


Each domain structure: Crystallographic analysis





# Calculation of $I(q)$ from Each Snap Shot Structure



$ph$  : Region for protein and hydration water molecules

$bi$  : Region for water molecules not perturbed by protein

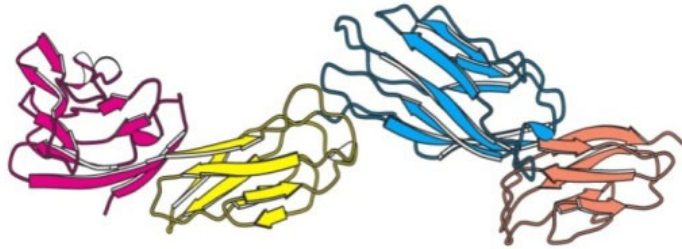
$$I(Q) - I^W(Q) = \Delta I_{ph,ph}(Q) + \Delta I_{ph,bi}(Q) + \Delta I_{ph,bo}(Q) + \Delta I_{bi,bi}(Q) + \Delta I_{bi,bo}(Q) + \Delta I_{bo,bo}(Q)$$

Expansion of Debye formula  
by spherical harmonics

= 0 for  $R \gg$  Region:  $bi$

# Principle of MD-SAXS

Construction of Initial Atomic Structure



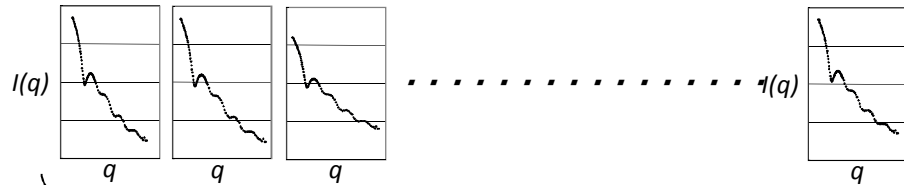
Each domain structure: Crystallographic analysis



**MD Simulation (ex. 150 ns)**



Calculation of  $I(q)$  from each snap-shot structure



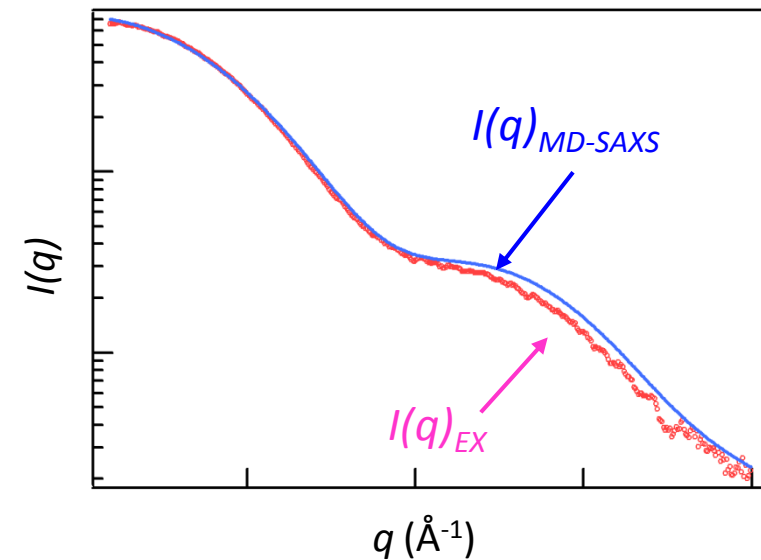
Average the  $I(q)$  data-sets:  $I(q)_{MD-SAXS}$



Inconsistent

Consistent

Each snap-shot structure is the dynamical structure of the protein

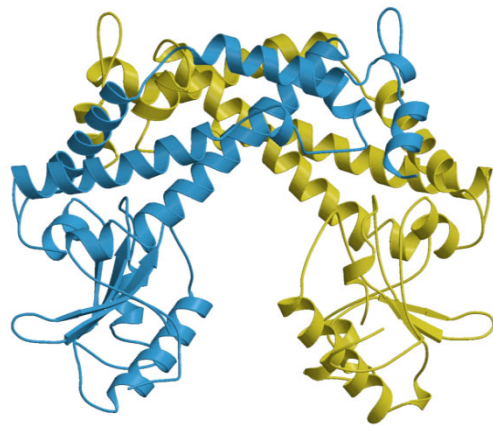


Compare  $I(q)_{MD-SAXS}$  with  $I(q)_{EX}$

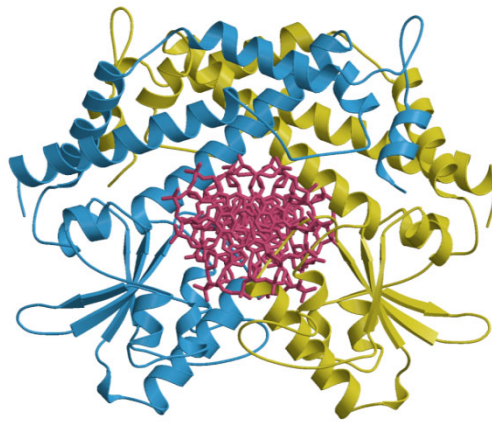


# Example 1: Restriction Endonuclease, *EcoO109I*

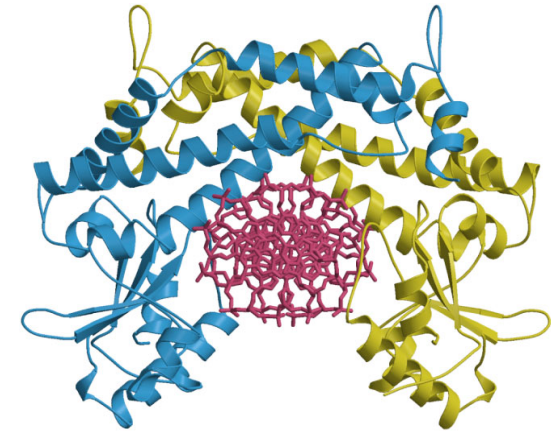
*EcoO109I* consists of 272 residues (Mr 31,000) and exists as a dimer in solution



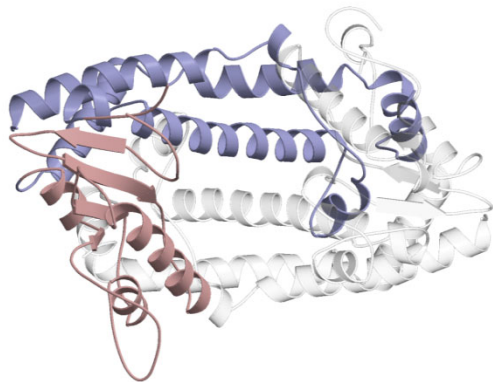
↕<sub>90°</sub>



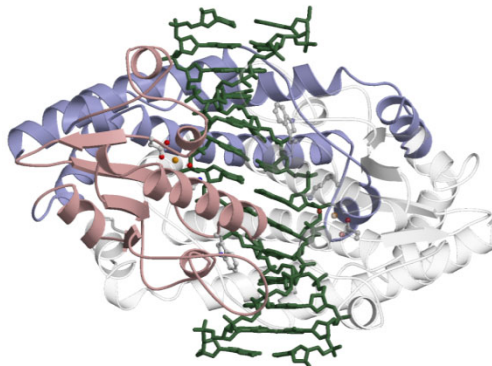
↕<sub>90°</sub>



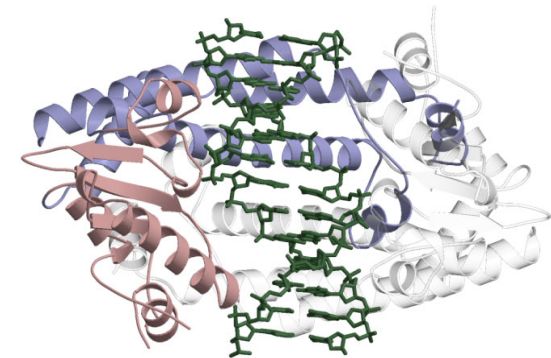
↕<sub>90°</sub>



*EcoO109I*



*EcoO109I* (WT)-DNA

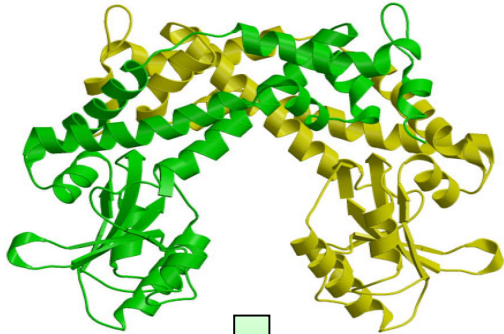


*EcoO109I* (D77A)-DNA

Hashimoto et al. *J. Biol. Chem.* **280**, 5605-5610 (2005)

# Principle of MD-SAXS

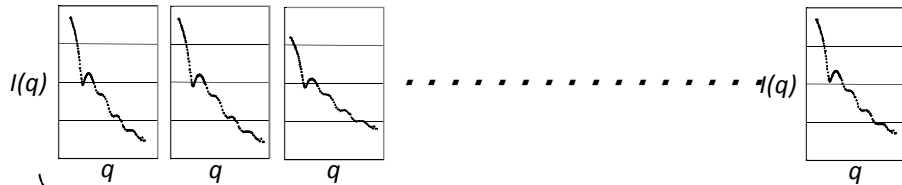
Construction of Initial Atomic Structure



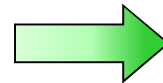
MD Simulation (ex. 150 ns)



Calculation of  $I(q)$  from each snap-shot structure

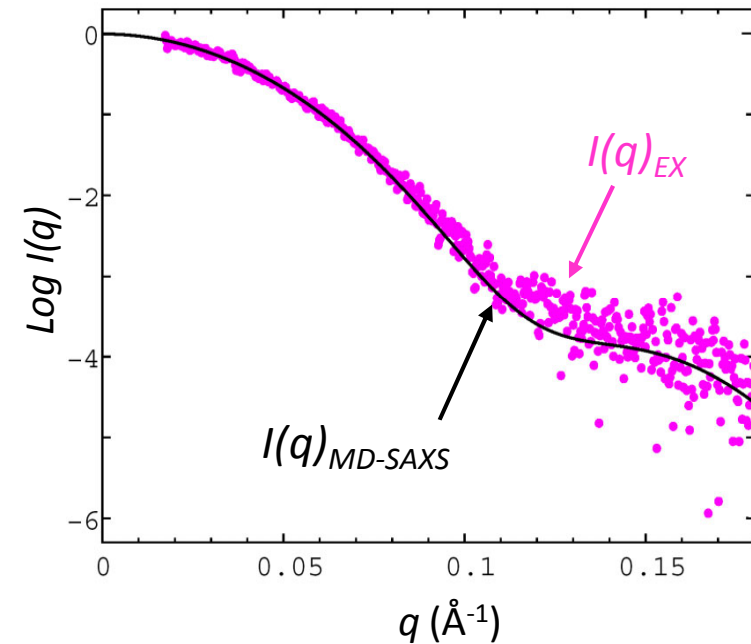


Average the  $I(q)$  data-sets:  $I(q)_{MD-SAXS}$



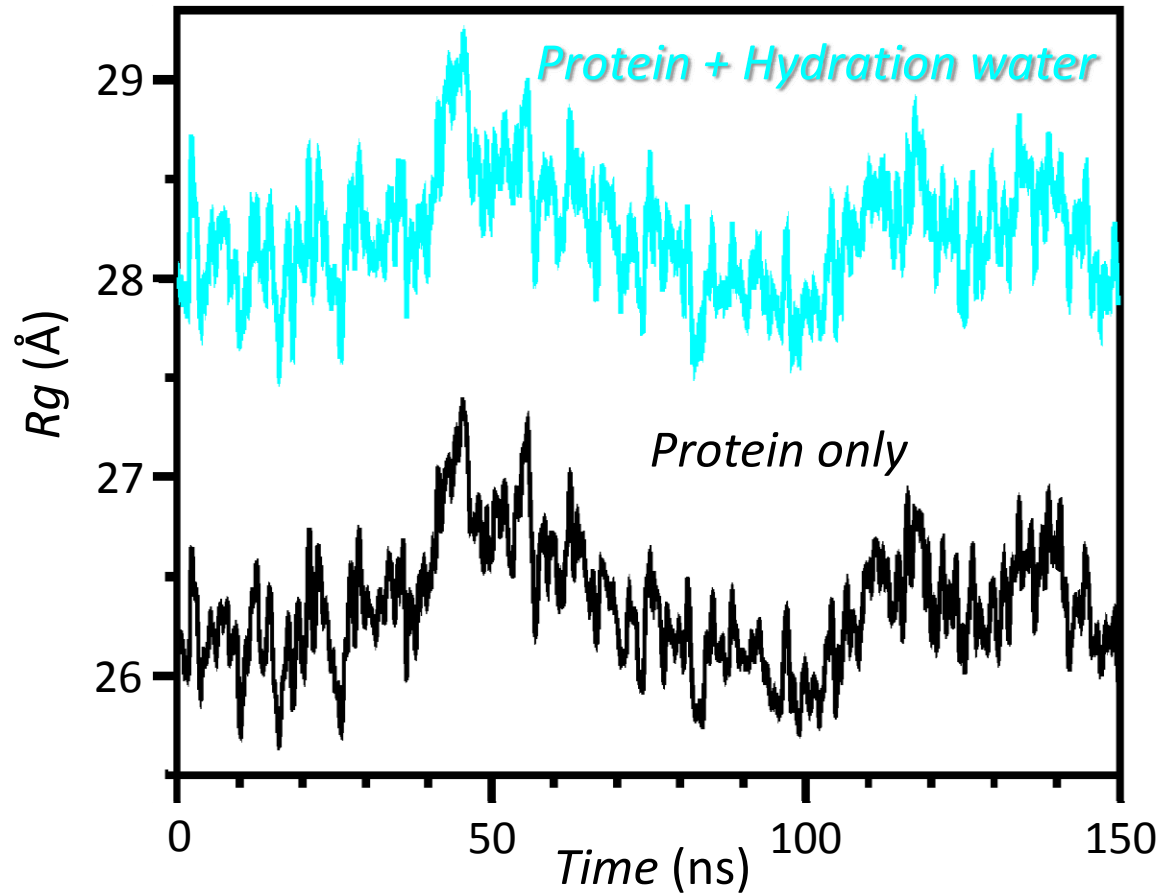
Consistent

Each snap-shot structure is the dynamical structure of the protein



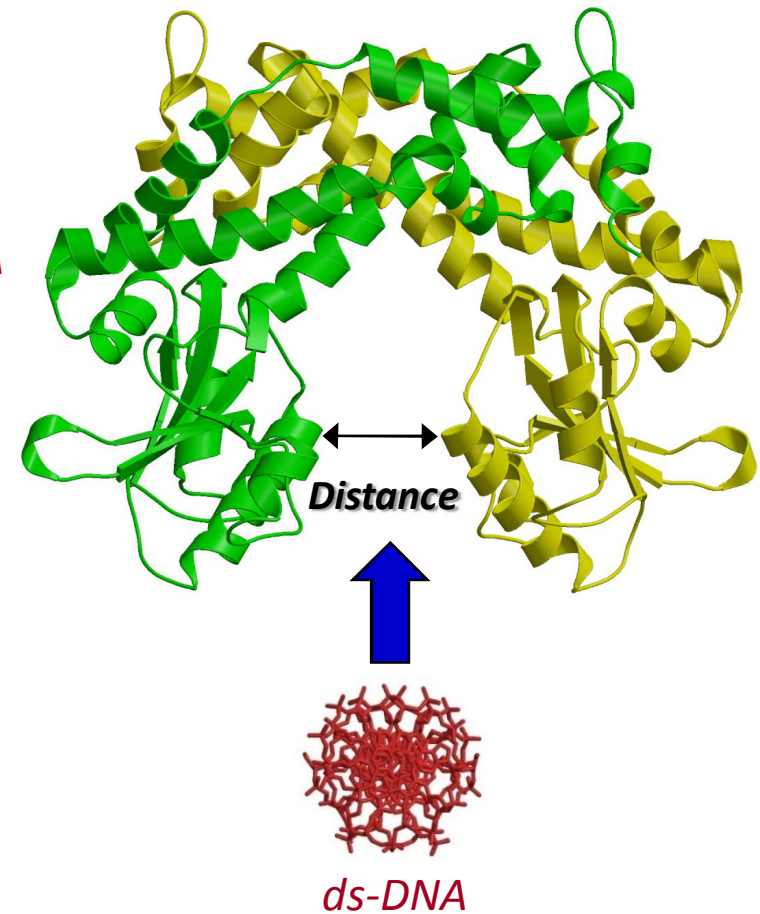
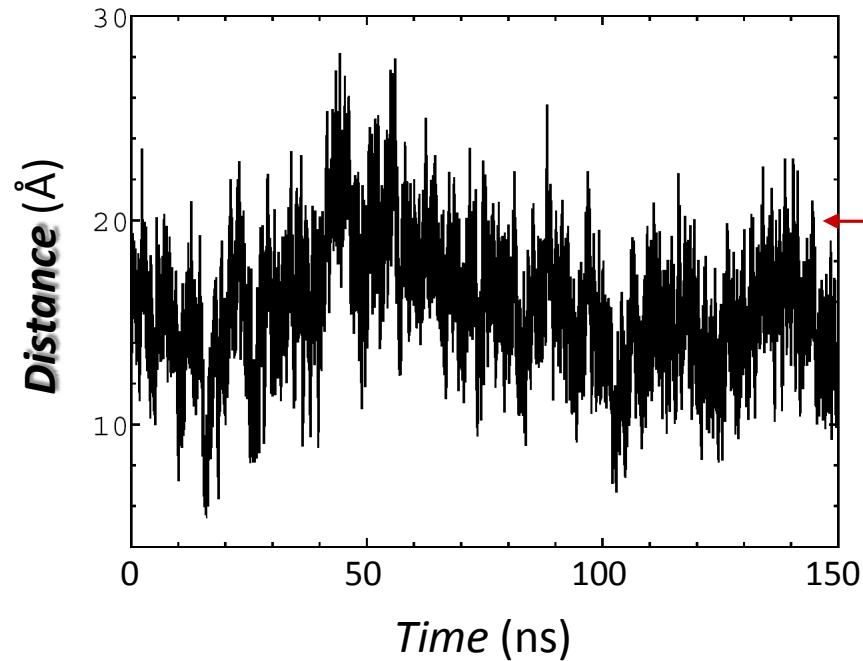
Compare  $I(q)_{MD-SAXS}$  with  $I(q)_{EX}$

## *Rg as a Function of Simulation Time*



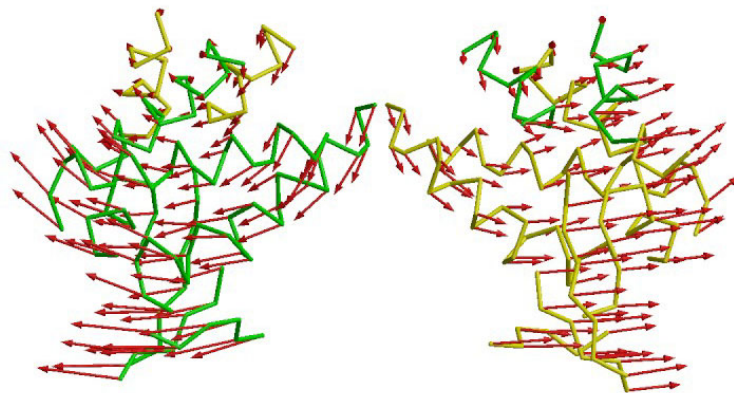
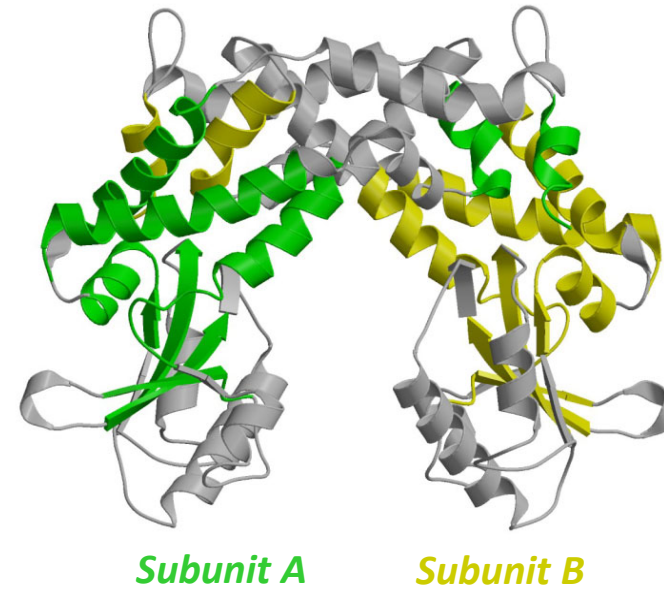
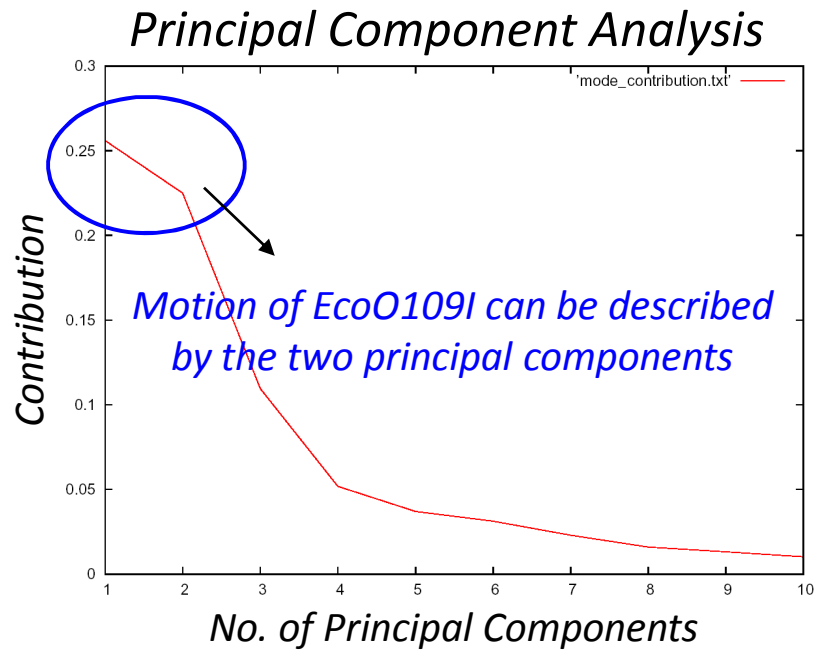
- *Structure ( $R_g$ ) is fluctuated with a cycle of about 100 ns*
- *Hydration shell width is about 2  $\text{\AA}$*

# *Distance as a Function of Simulation Time*

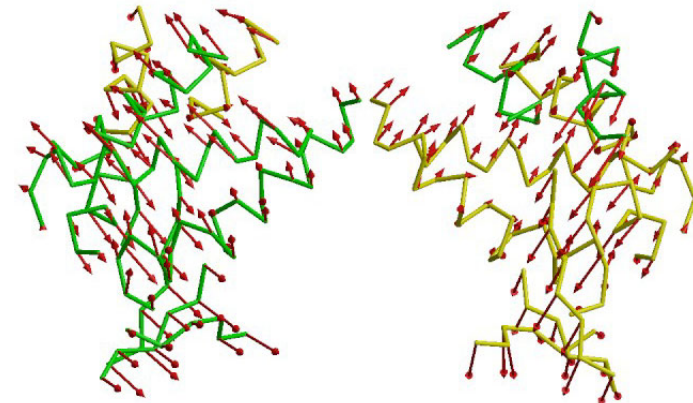


- Structure (**Distance**) is fluctuated with a cycle of about 100 ns
- ds-DNA can enter within EcoO109I during about 100 ns

# Motion between *Subunit A* and *Subunit B*



The first principal component: *open-close motion*

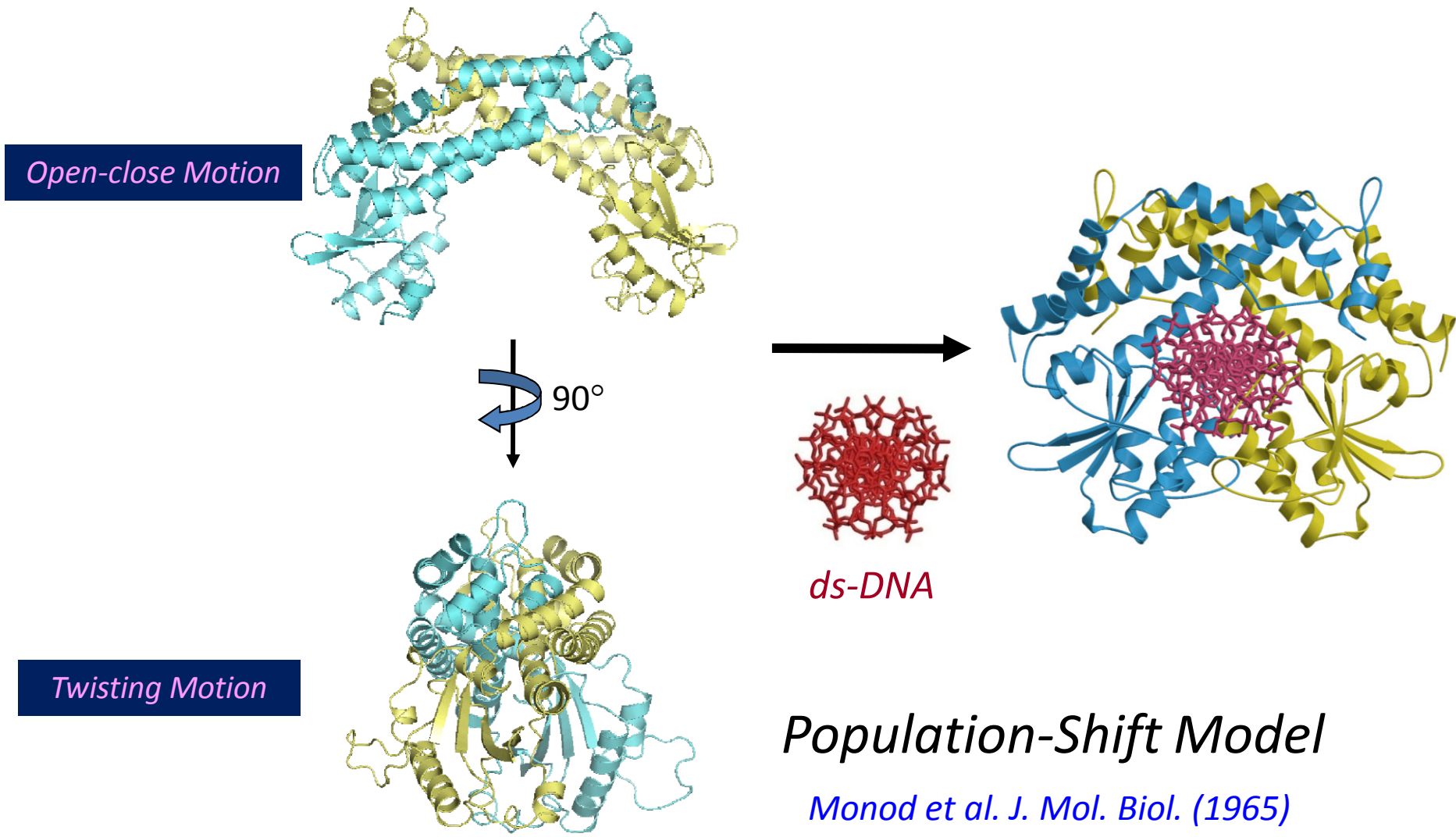


The second principal component: *twisting motion*

Oroguchi et al. *Biophys. J.* **96**, 2808-2822 (2009)



# Motion between *Subunit A* and *Subunit B*

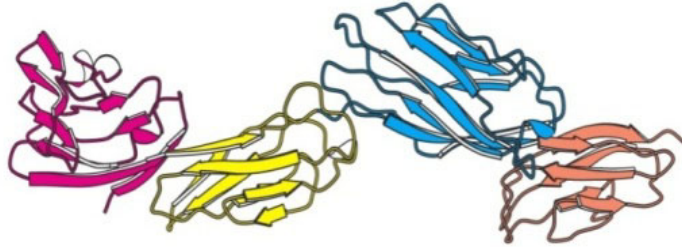


*Population-Shift Model*

*Monod et al. J. Mol. Biol. (1965)*

# Principle of MD-SAXS

Construction of Initial Atomic Structure



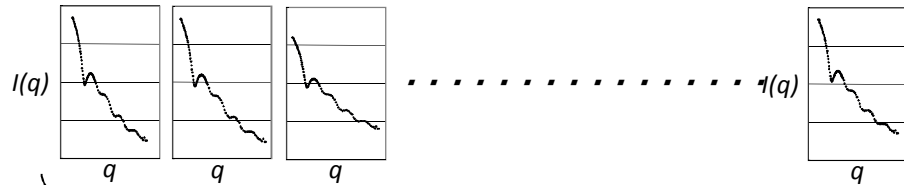
Each domain structure: Crystallographic analysis



**MD Simulation (ex. 150 ns)**



Calculation of  $I(q)$  from each snap-shot structure



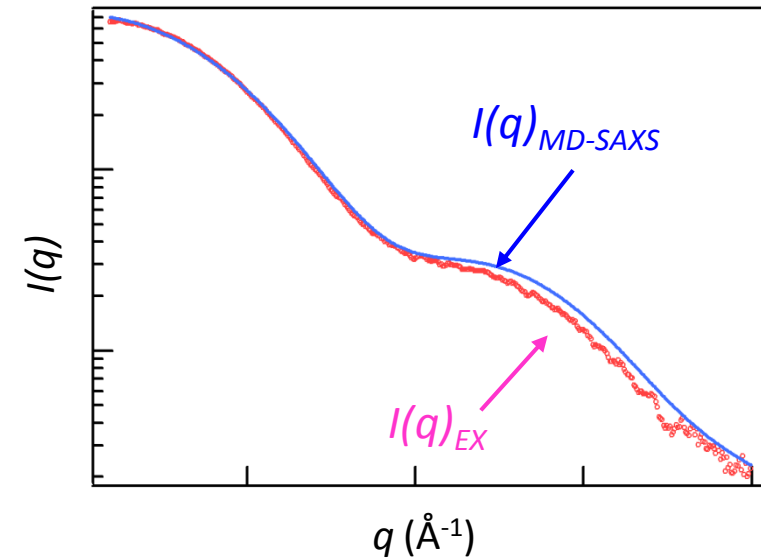
Average the  $I(q)$  data-sets:  $I(q)_{MD-SAXS}$



Inconsistent

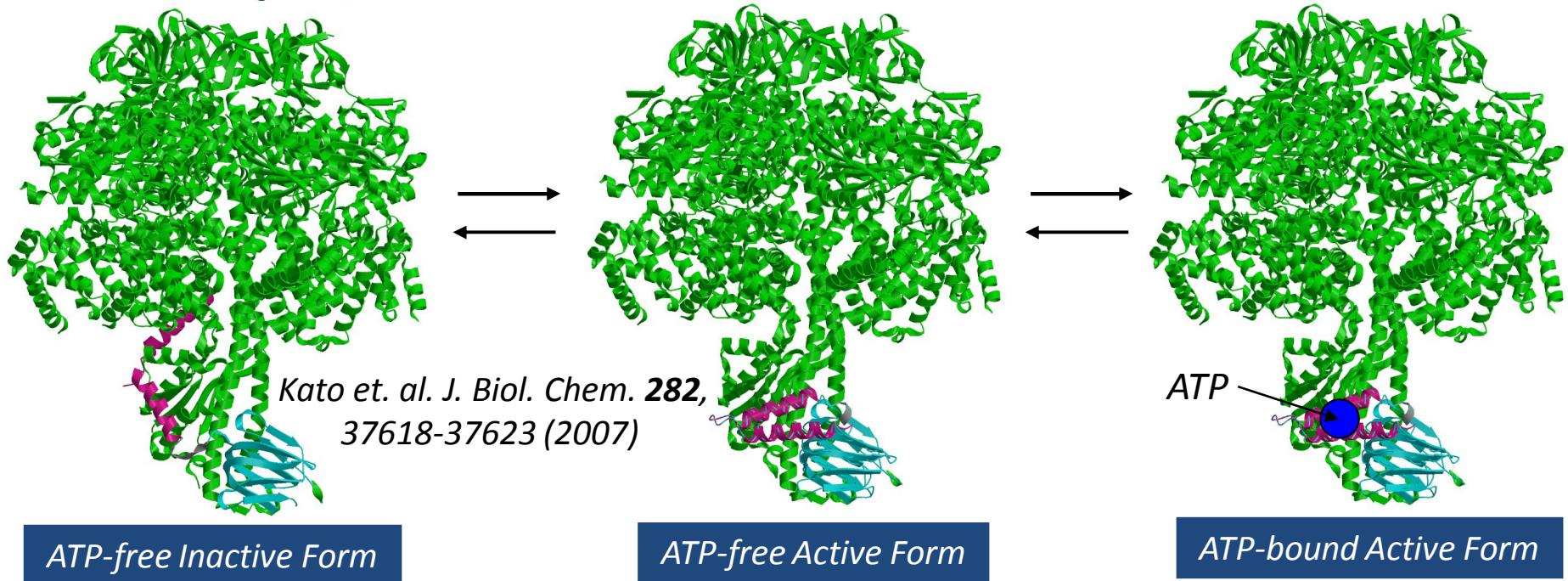
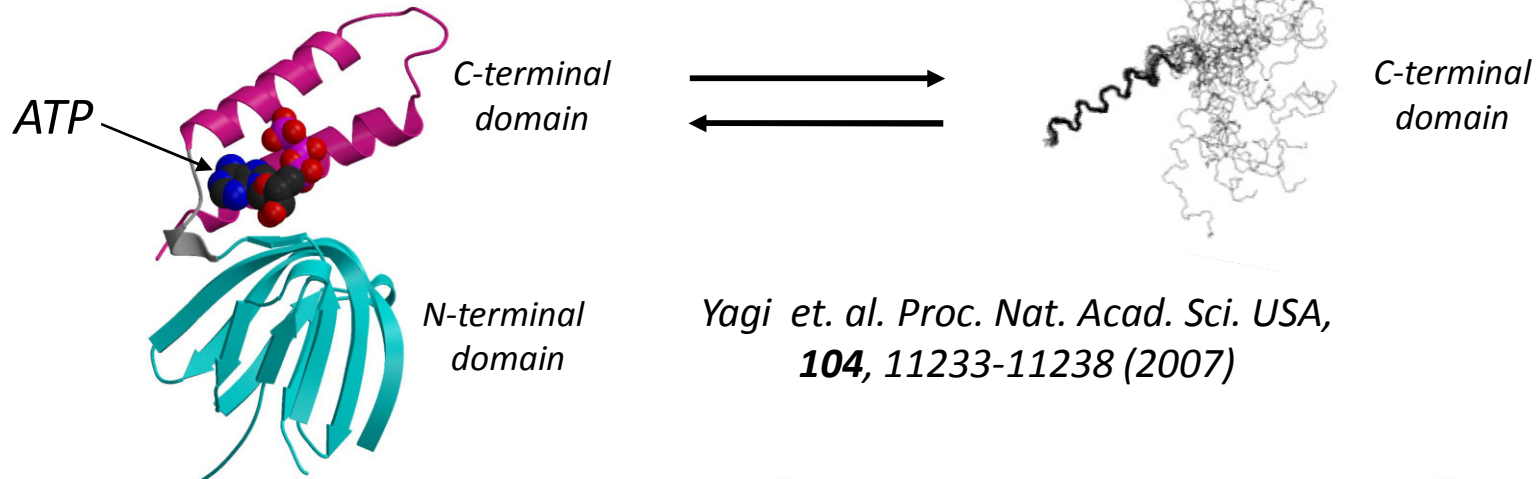
Consistent

Each snap-shot structure is the dynamical structure of the protein



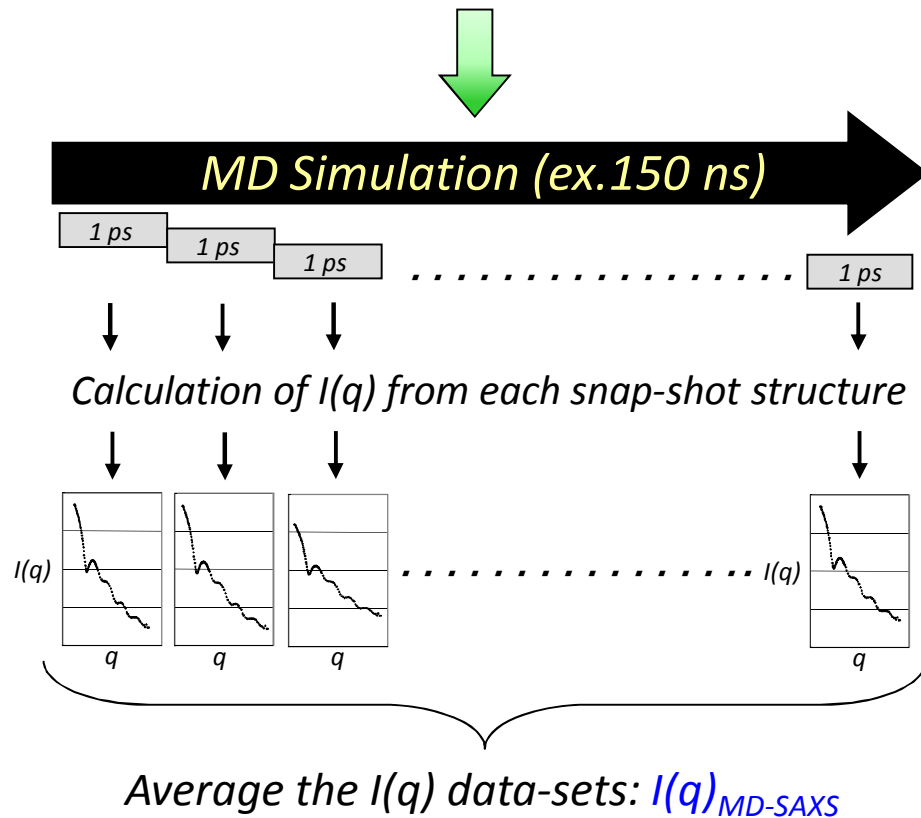
Compare  $I(q)_{MD-SAXS}$  with  $I(q)_{EX}$

## Example 2: $\epsilon$ subunit of thermophilic F1-ATPase

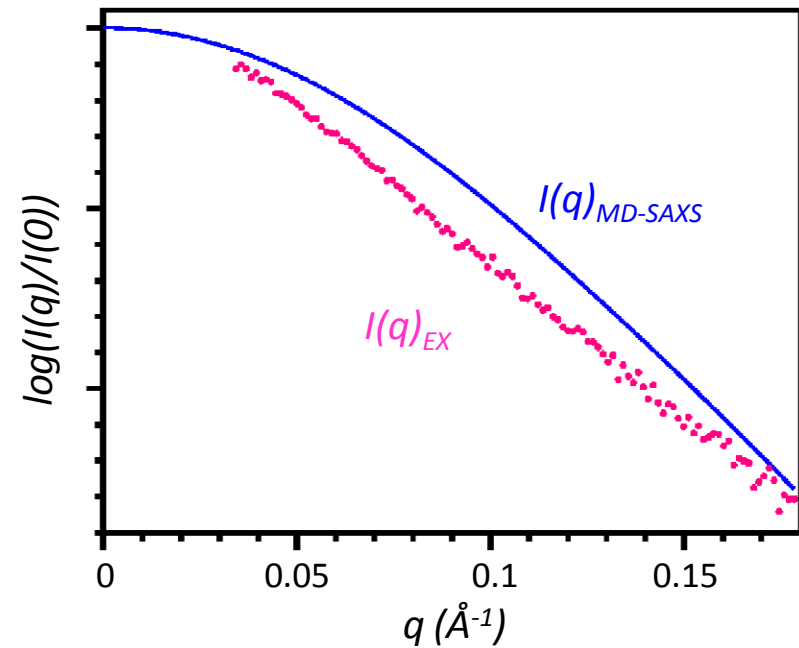


# MD-SAXS of $\epsilon$ subunit of F1-ATPase

## Construction of Initial Atomic Structure



Inconsistent



Compare  $I(q)_{MD-SAXS}$  with  $I(q)_{EX}$

# MD-SAXS of $\epsilon$ subunit of F1-ATPase

Construction of Initial Atomic Structure

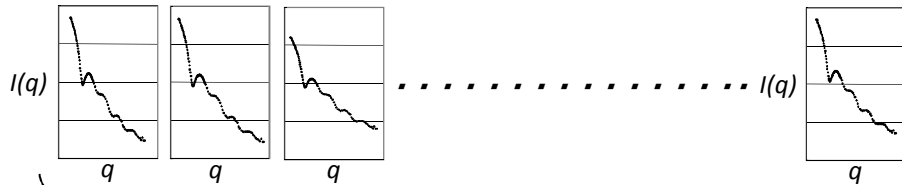


Predicted structure of ATP-free inactive  $\epsilon$  subunit

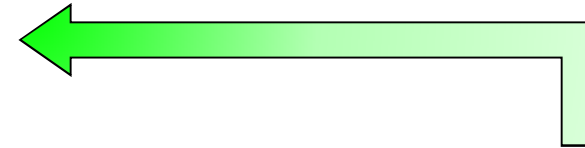
MD Simulation (ex. 150 ns)



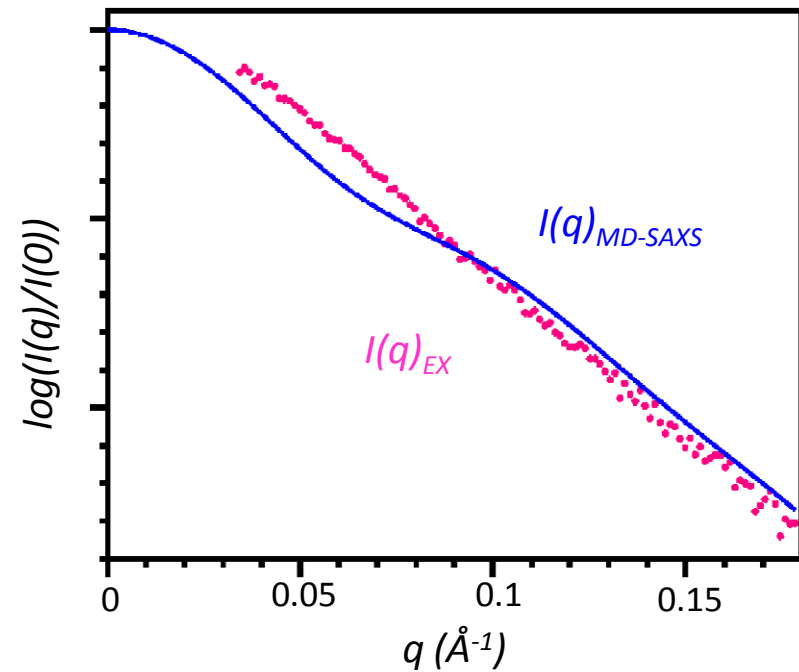
Calculation of  $I(q)$  from each snapshot structure



Average the  $I(q)$  data-sets:  $I(q)_{MD-SAXS}$



Inconsistent

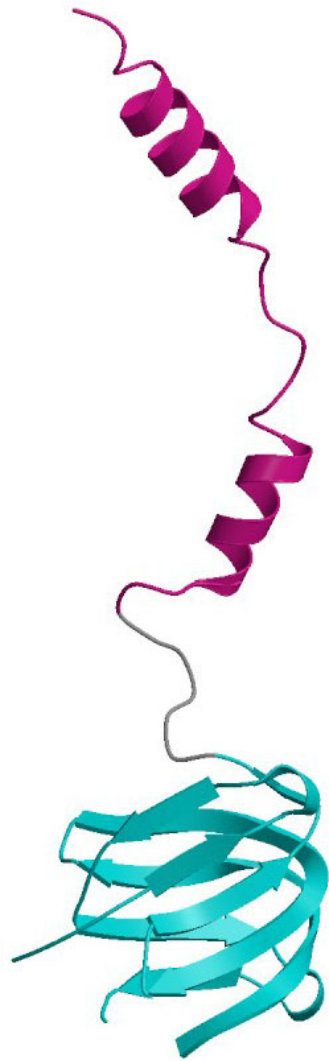


Compare  $I(q)_{MD-SAXS}$  with  $I(q)_{EX}$

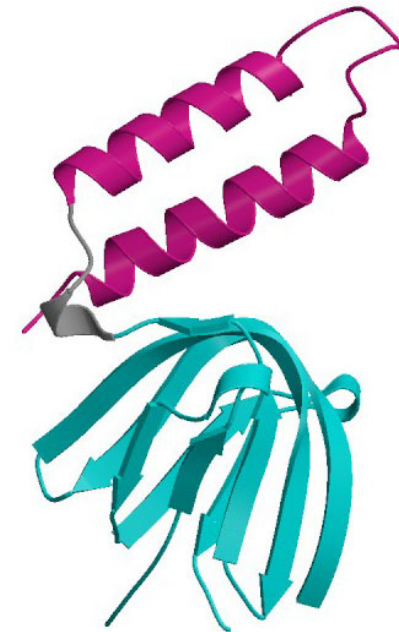
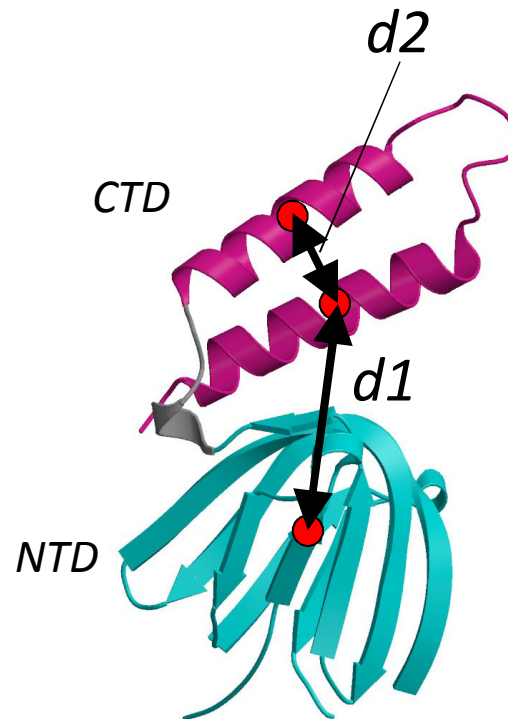




# Reconstruction of Initial Atomic Structure

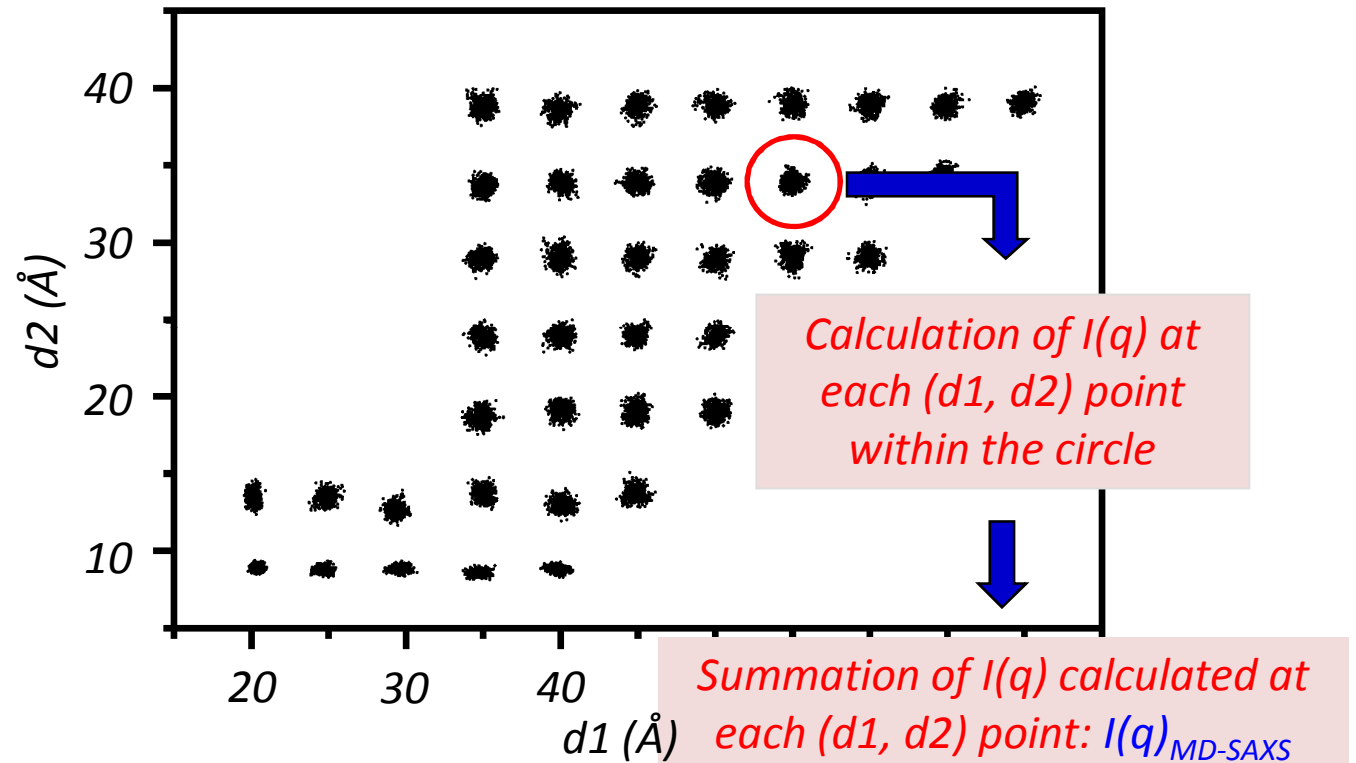
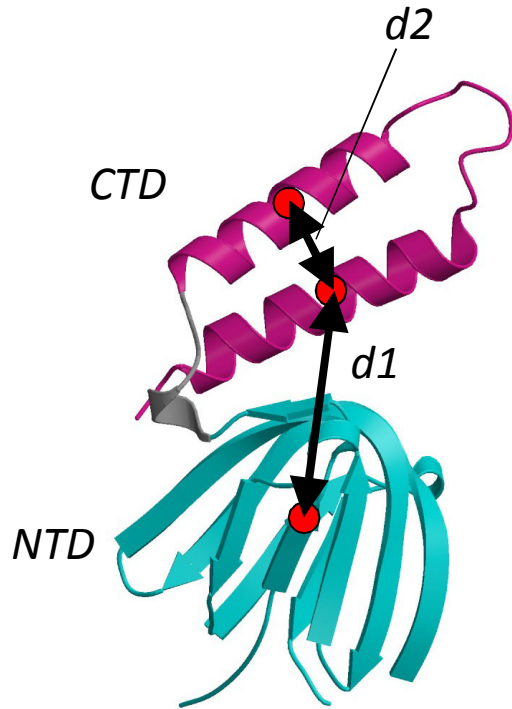


*Predicted structure of  
ATP-free Inactive  $\epsilon$  subunit*



*Crystal structure of  
ATP-free active  $\epsilon$  subunit*

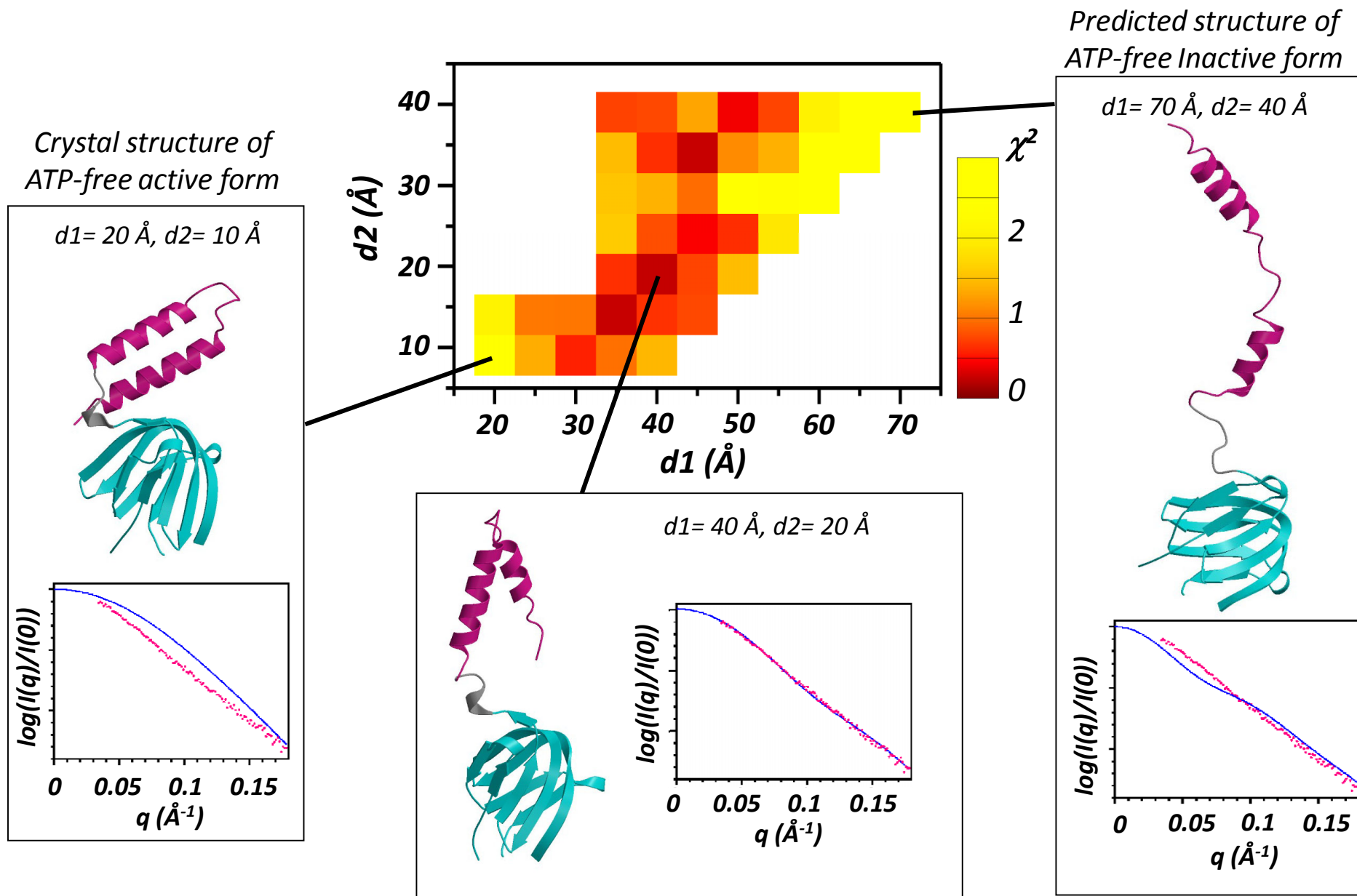
# Reconstruction of Initial Atomic Structures



Each of  $d1$  &  $d2$  was changed at an interval of  $5 \text{ \AA}$  (total 41 initial atomic structures), and MD simulation was carried out using each initial atomic structure.

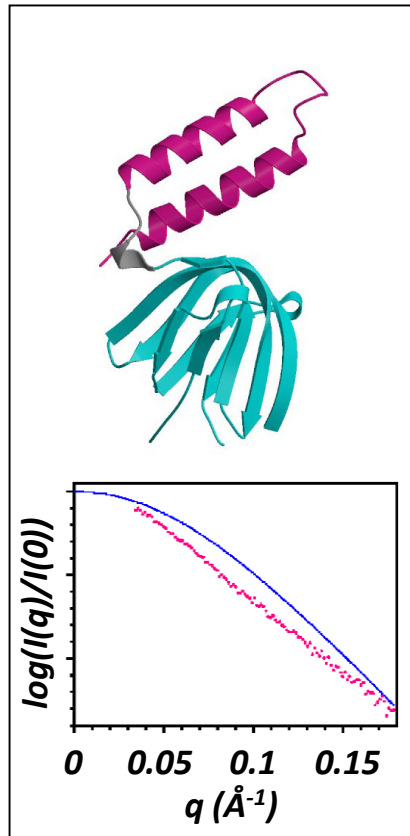
Compare  $I(q)_{MD-SAXS}$  with  $I(q)_{EX}$

# Comparison of $I(q)_{MD-SAXS}$ with $I(q)_{EX}$

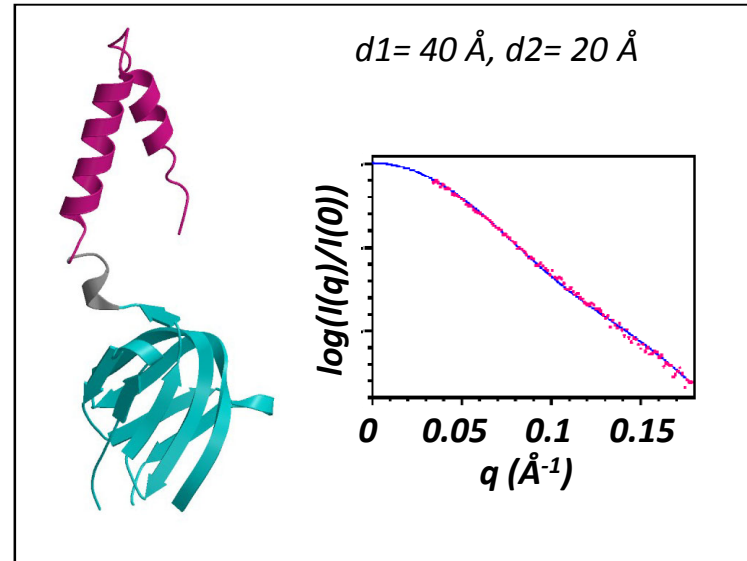


# ATP-free $\epsilon$ subunit structure in solution

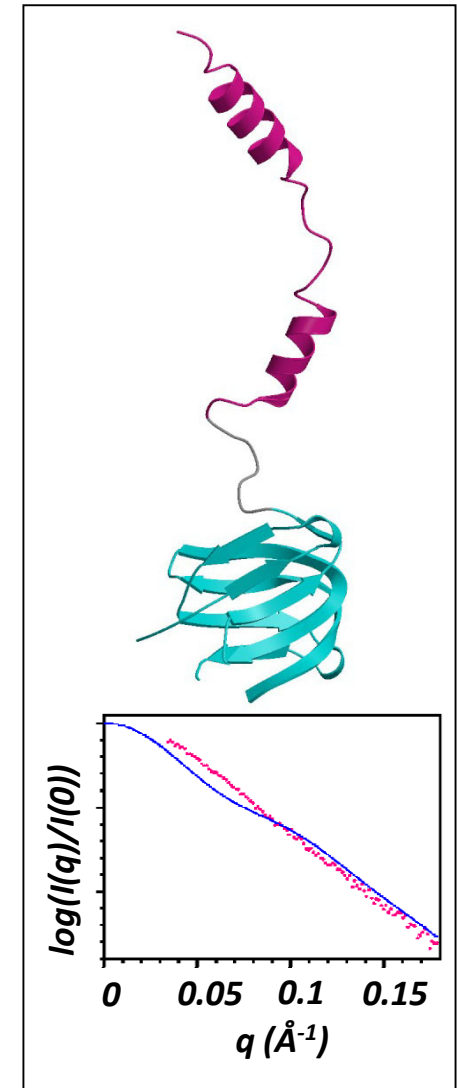
Crystal structure of ATP-free active form



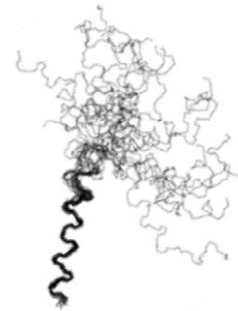
but, fluctuates around



Predicted structure of ATP-free Inactive form

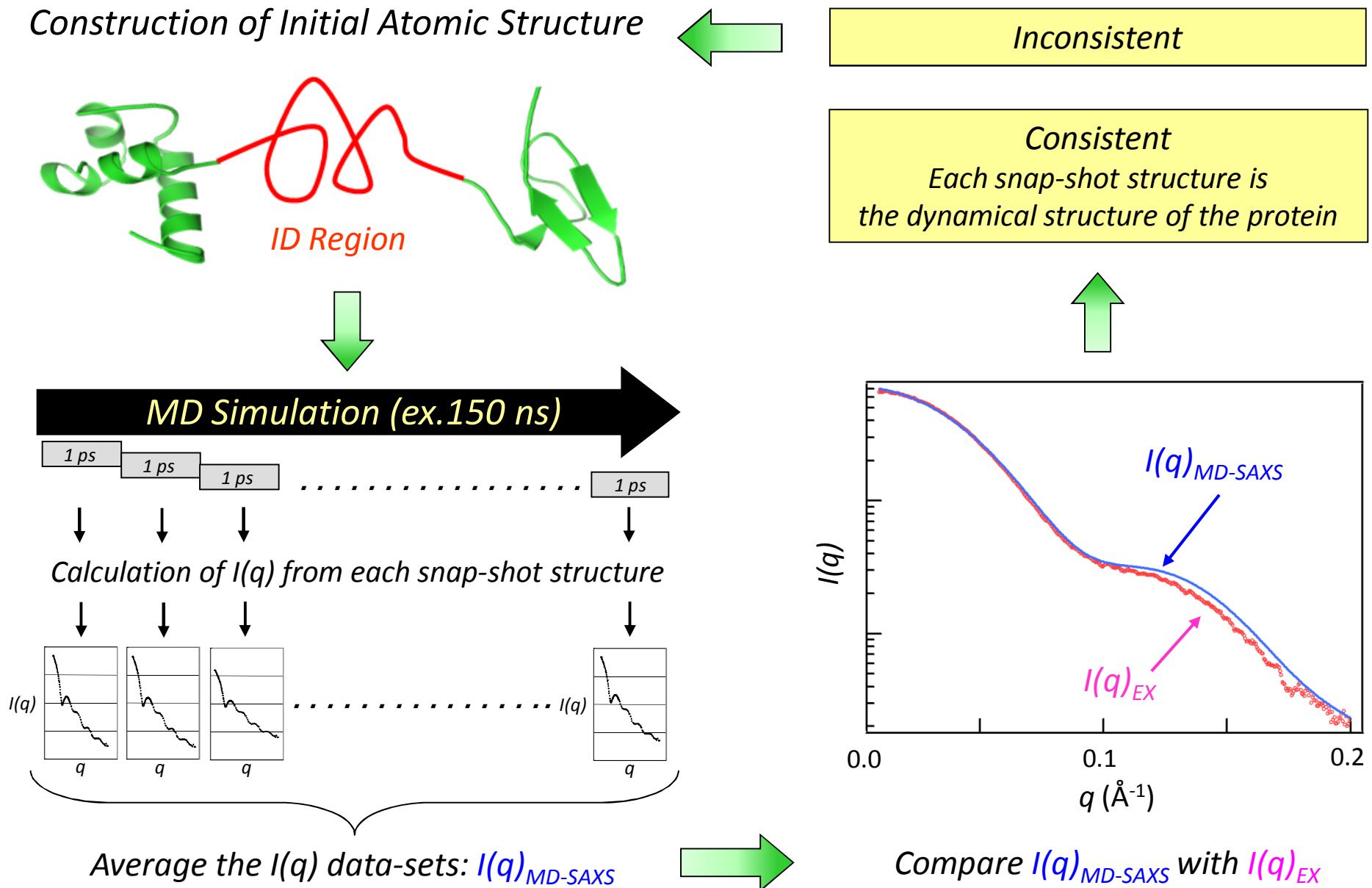


C-terminal domain



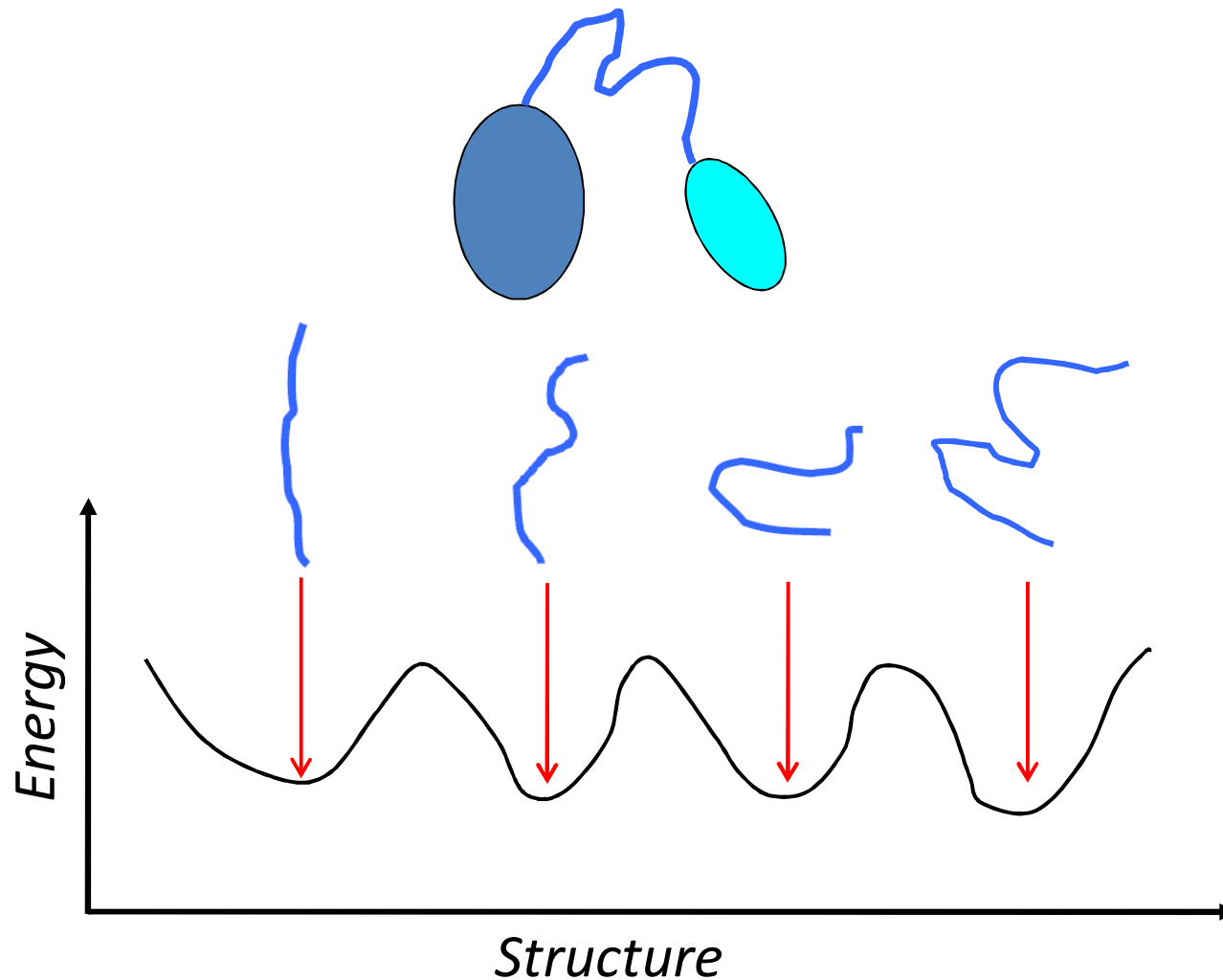
Yagi et. al. Proc. Nat. Acad. Sci. USA, **104**, 11233-11238 (2007)

# MD-SAXS of Multi-Domain Protein with *ID Region*





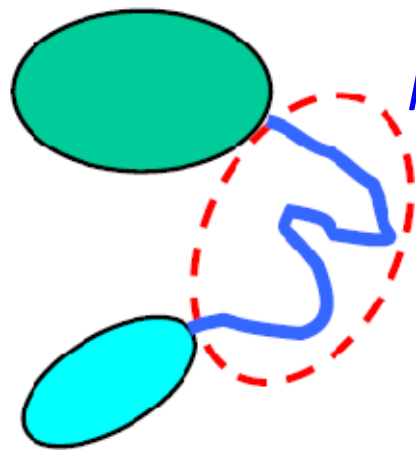
# Construction of Initial Atomic Structure



*ID regions exist as dynamic ensembles and exhibit extremely temporal fluctuations, therefore extending the MD-SAXS method to characterize IDPs (Extended MD-SAXS)*

# Construction of *ID Region* in Extended MD-SAXS

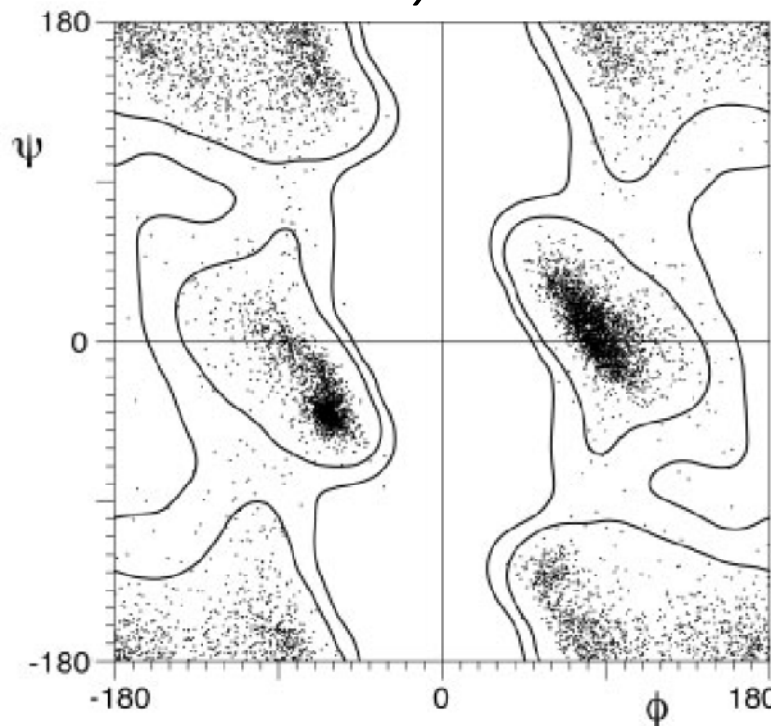
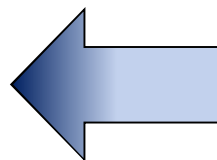
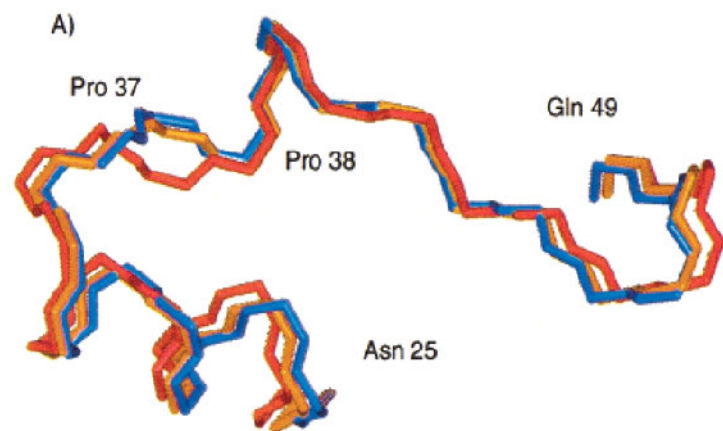
Bernado et. al. *J. Am. Chem. Soc.* **129**, 5656-5664 (2007)



*ID Region:*

is constructed using data-base of dihedral angles for each amino acid residue based on high resolution loop structures in PDB

Ex. Glycine

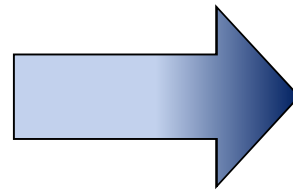


# Construction of Whole Structure in Extended MD-SAXS

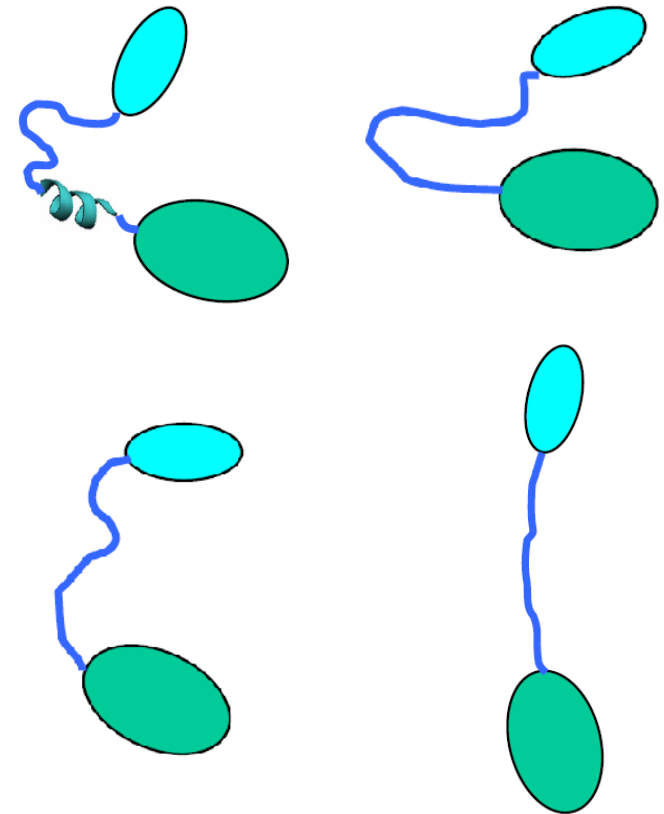
Constructed  
*ID region* structures



+ RDC Data in NMR

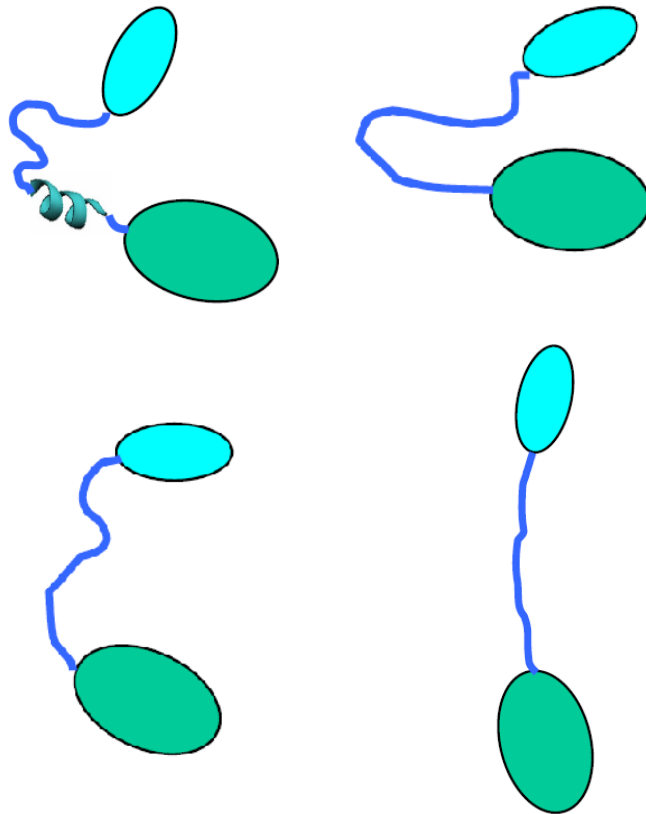


Constructed  
whole structures



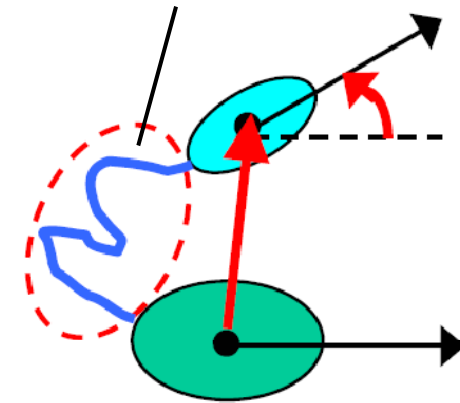
# *Description of the Whole Structure in Extended MD-SAXS*

*Constructed  
whole structures*



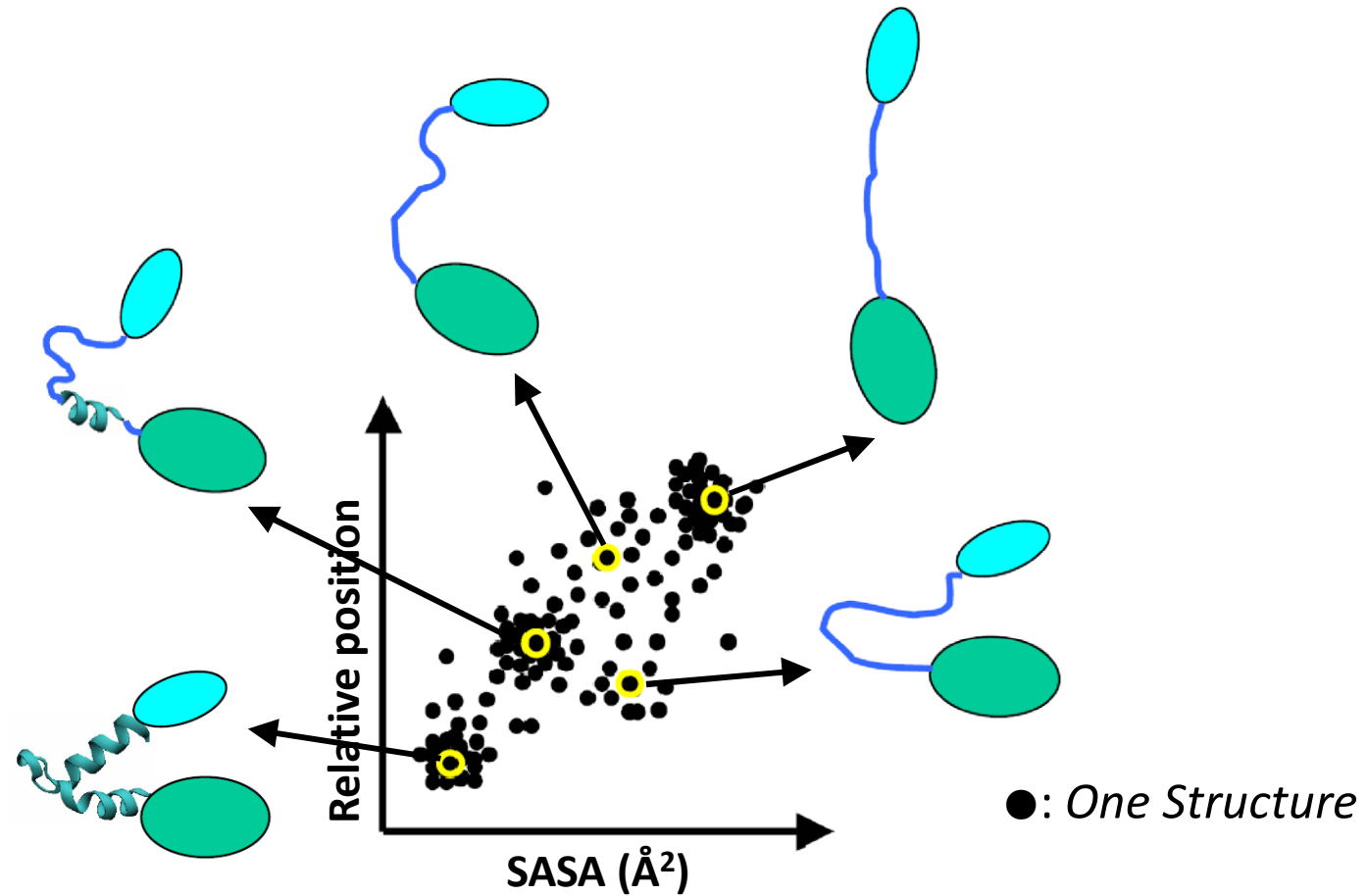
*Description of  
the whole structure by  
the collective coordinate*

*Solvent accessible surface area  
(SASA)*



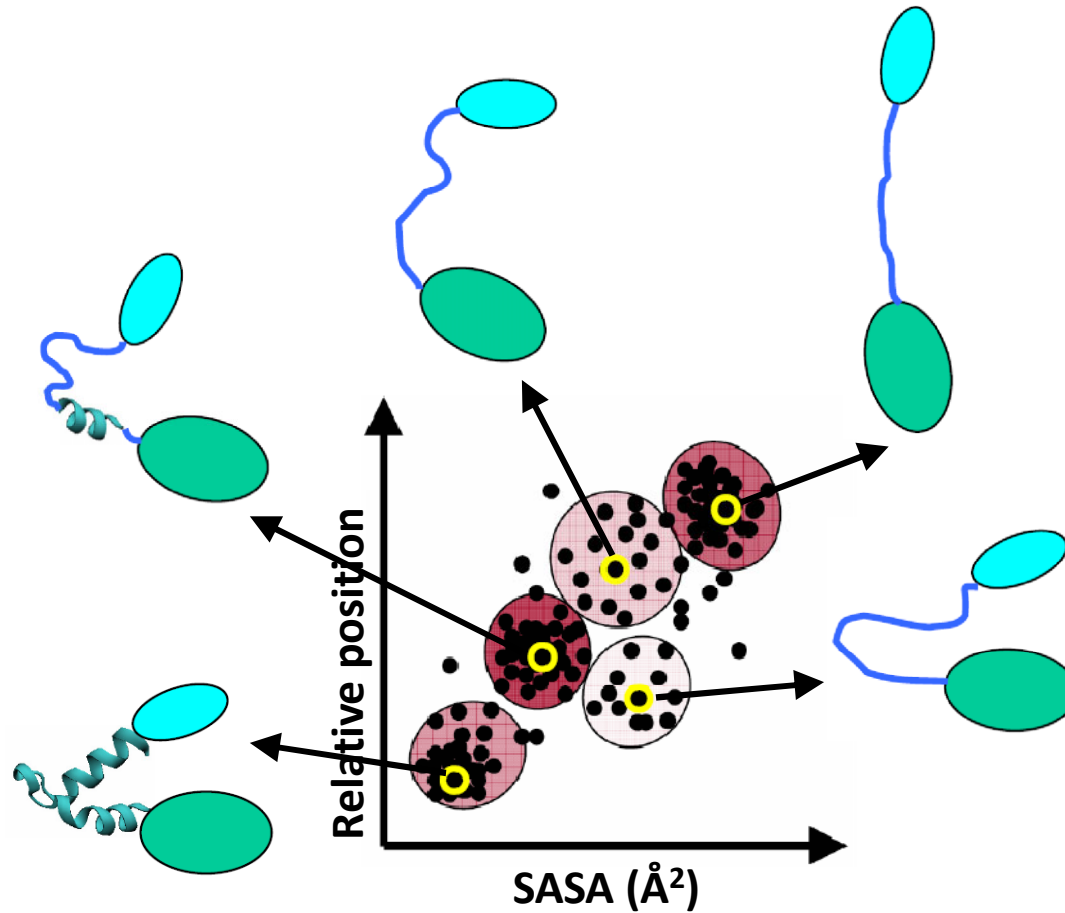
*Relative position  
between the two domains*

# Initial Models of Structural Ensemble in Extended MD-SAXS



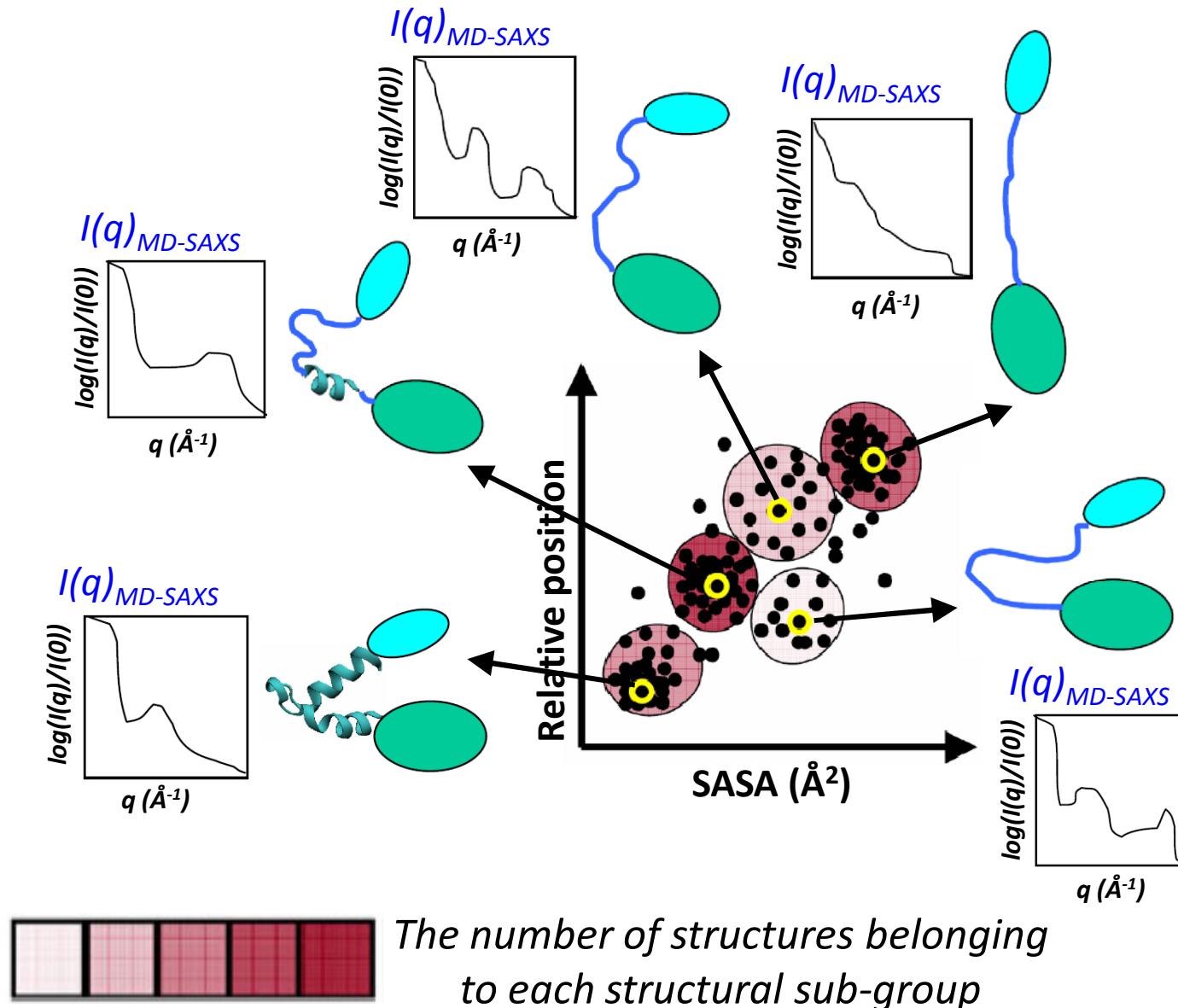


# Classification of the Structural Ensemble into Sub-Groups

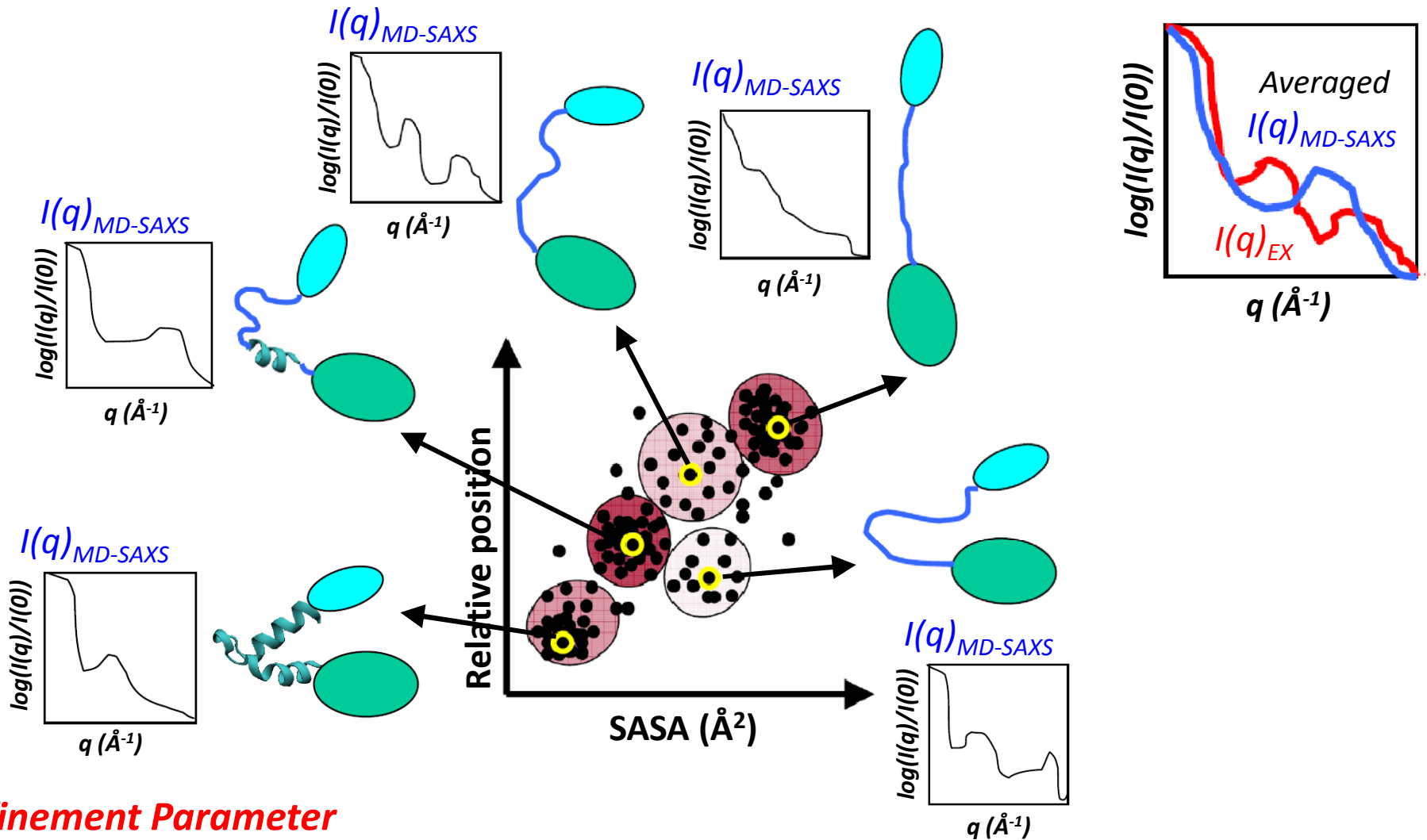


*The number of structures belonging to each structural sub-group*

# Calculation of $I(q)_{MD-SAXS}$ for Each Sub-group



# Average $I(q)_{MD-SAXS}$ & Compare with $I(q)_{EX}$

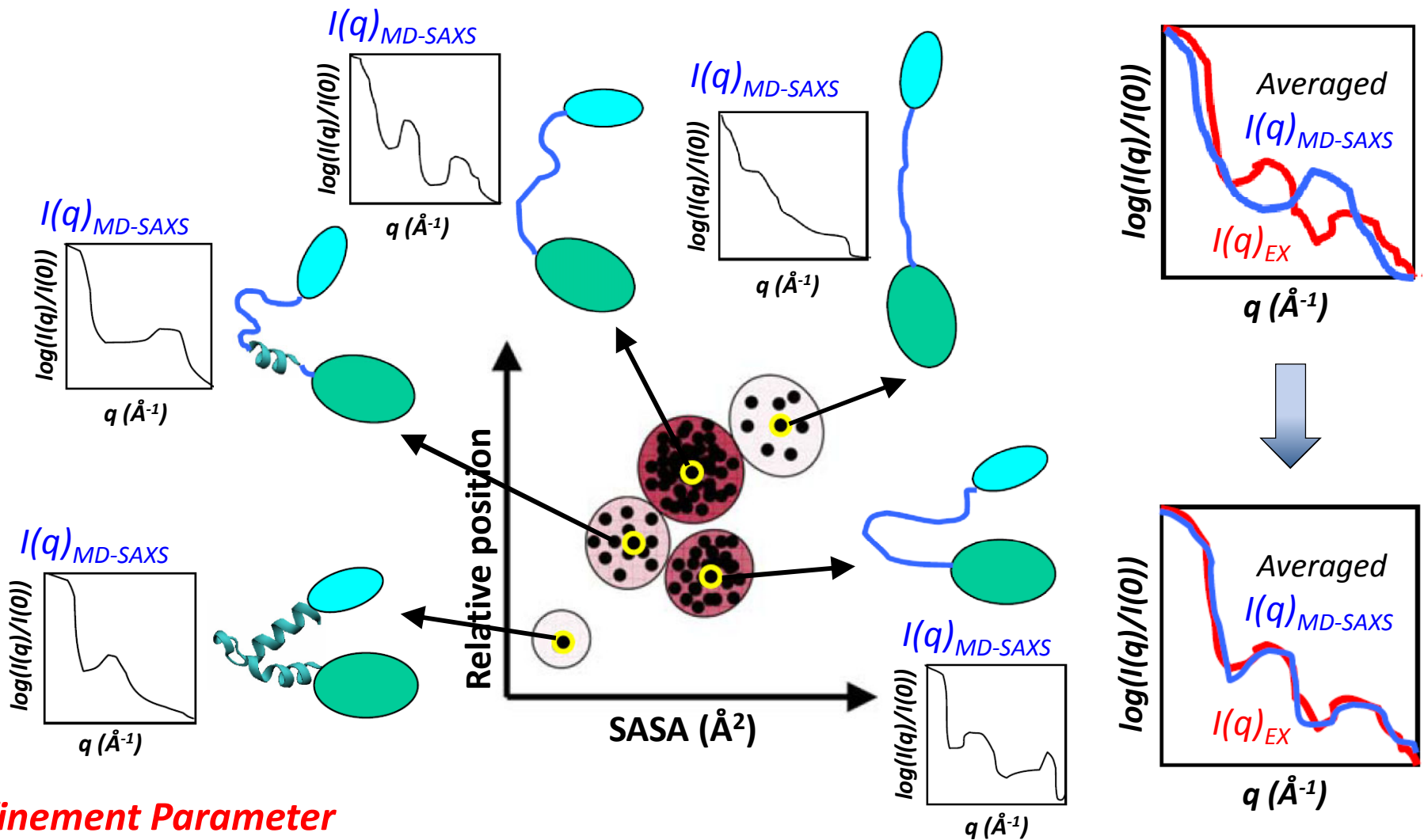


## Refinement Parameter



The number of structures belonging to each structural sub-group

# Refinement of Structural Ensemble

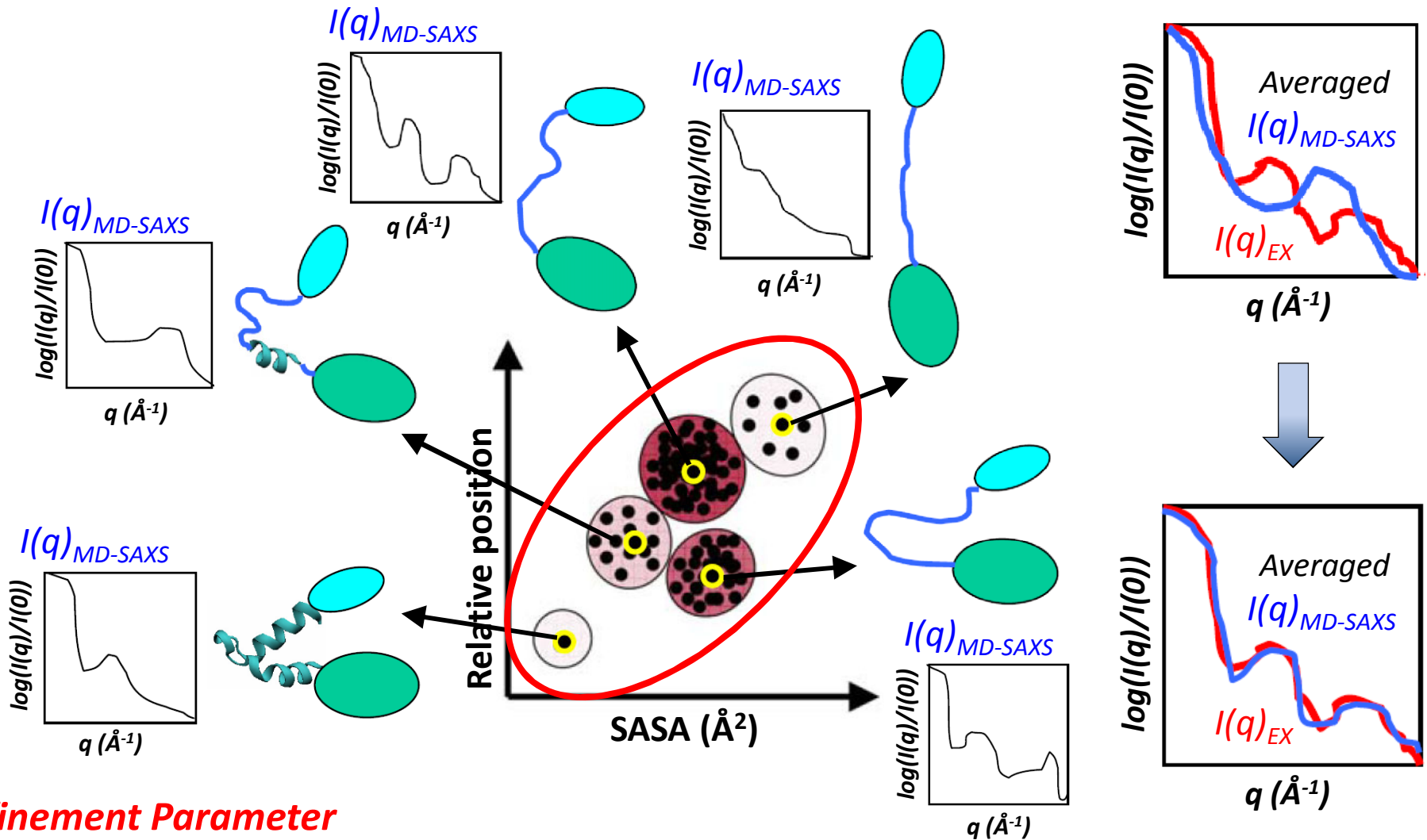


**Refinement Parameter**



The number of structures belonging to each structural sub-group

# Structures (Ensemble) of *IDP* in Solution



## Refinement Parameter



The number of structures belonging to each structural sub-group