Selective ion-binding by human lysozyme studied by the statistical mechanical integral equation theory

Young Investigators Symposium
Tennessee, USA
October 13-15 2008

Yutaka Maruyama

Institute for Molecular Science, Japan

Background

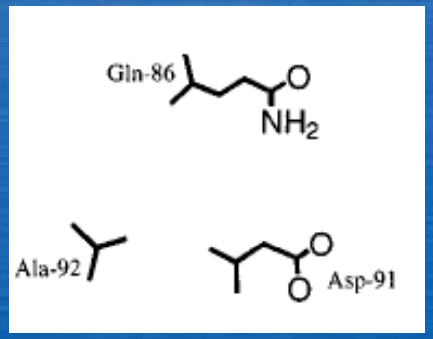
- The molecular recognition by protein is the most fundamental process in living systems to maintain their life.
- One of the most important and essential processes of the molecular recognition is “selective ion-binding”.
- In this work, we try to reproduce and predict such a process by integral equation theory.

Background

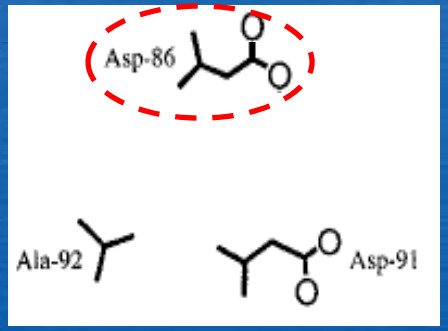
We consider three sequence of human lysozyme

- Wild type
- Q86D
- Q86D/A92D

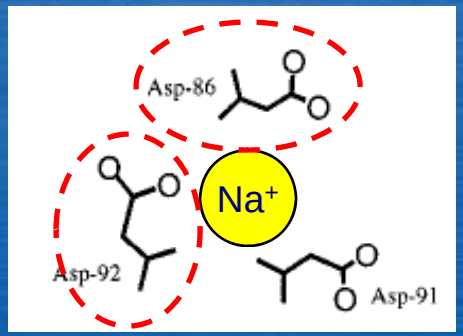
QuickTimeý Ç²
Sorenson Video 3 êLí£ÉvÉçÉOÉâÉÄ
Ç™Ç±ÇÃÉ[illegible]É`ÉÉÇ¾å©ÇÈÇžÇ½Ç…ÇÕïKóvÇ-ÇÅB



wild type



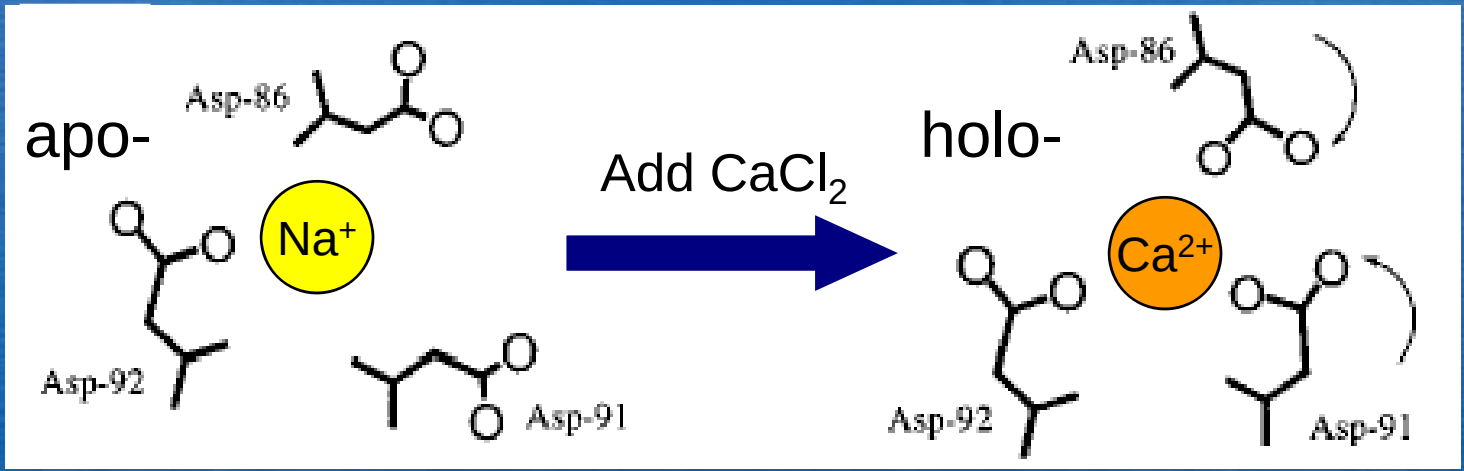
Q86D



Q86D/A92D

Background

- A doubly mutated lysozyme, Q86D/A92D has two isomers, apo and holo.
- apo-Q86D/A92D: in NaCl*aq*
- holo-Q86D/A92D: in NaCl/$CaCl_2$*aq*



Outline

- We consider selective ion-binding by lysozyme and its mutant in two type electrolyte solution.

- Selectivity of ion-binding is evaluated by probability of finding ions in active site.

  $\Rightarrow$ Distribution function of ion

  $\Rightarrow$ 3D-RISM theory

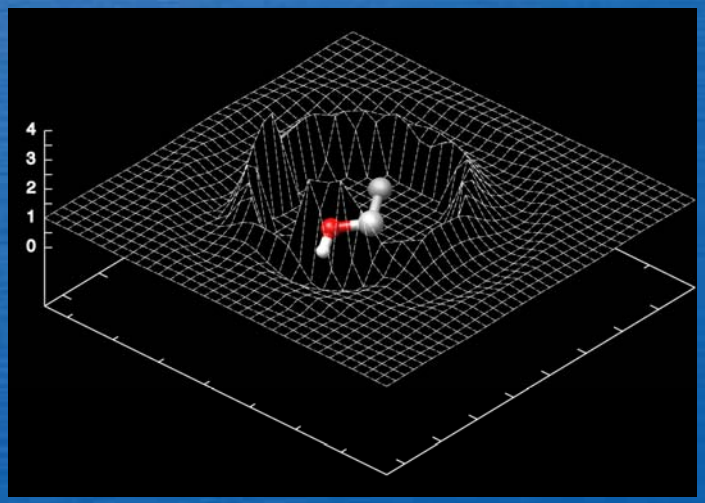
- Which mutant can bind Na^{+} or Ca^{2+} ion?
- Can 3D-RISM reproduce such a selectivity?

3D-RISM Theory

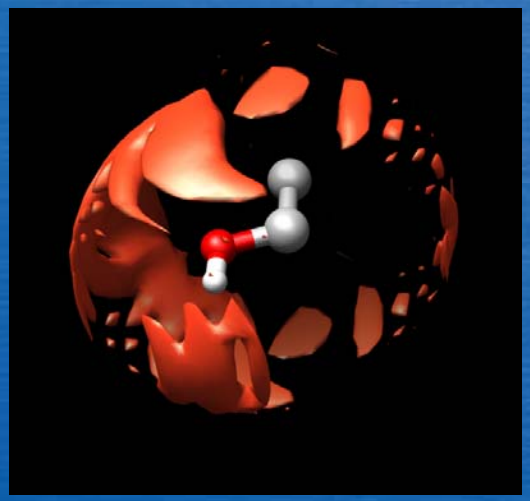
- In order to investigate the solvent distribution around the solute molecule,
  we introduce 3D-distribution functions.

- 3D-Distribution Function (3D-DF): $g_a^s(\mathbf{r}_a)$

  Example: Oxygen of water distribution around ethanol



3D-DF($g_O(r)$) on the X-Y plane



$g_O(r) > 2$ is painted

3D-RISM Theory

- How to evaluate the 3D-DFs?

  $\Rightarrow$ 3D-RISM integral equation theory

- 3D-RISM theory

  3D-RISM theory is one of the statistical mechanics theories.

  This has been successfully applied for large molecular systems such as protein.

3D-RISM Theory

3D-RISM equation

$$h_a^s(\mathbf{r}) = \sum_{s' \in \text{Solvent species}} \sum_{c \in \text{Site on } s'} c_c^{s'}(\mathbf{r}_c) * \left[\varpi_{ca}^s + \rho^{s'} X_{ca}^{s's}(r_{ca}) \right]$$

Convolution integral

Just a multiplication in the Fourier space

Kovalenko-Hirata (KH) closure

$$g_a^s(\mathbf{r}) = \begin{cases} \exp(-\beta u_a^s(\mathbf{r}) + h_a^s(\mathbf{r}) - c_a^s(\mathbf{r})) & \text{for } h_a^s(\mathbf{r}) \le 0 \\ 1 - \beta u_a^s(\mathbf{r}) + h_a^s(\mathbf{r}) - c_a^s(\mathbf{r}) & \text{for } h_a^s(\mathbf{r}) > 0 \end{cases}$$

3D-DF

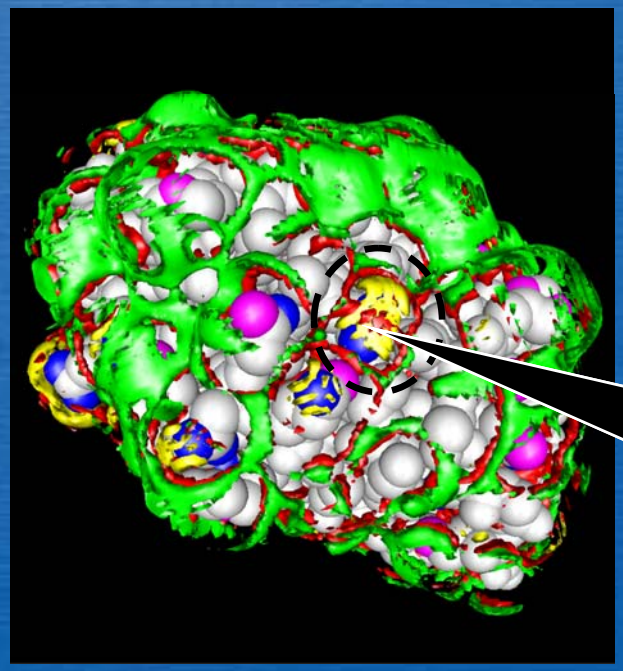
$$g_a^s(\mathbf{r}) = h_a^s(\mathbf{r}) + 1$$

Solute-solvent interaction potential

$$u_a^s(\mathbf{r}) = \sum_{b \in \text{solute}} \frac{q_b^u q_a^s}{\left| \mathbf{r} - \mathbf{r}_b^u \right|} + \sum_{b \in \text{solute}} 4\varepsilon_{ab} \left\{ \left(\frac{\sigma_{ab}}{r_{ab}} \right)^{12} - \left(\frac{\sigma_{ab}}{r_{ab}} \right)^{6} \right\}$$

Results

- 3D-DFs of water oxygen, Na^+ and Cl^- around the mutation of lysozyme are shown.



Cl^- (g(r)>5.0)

O of water (g(r)>3.0)

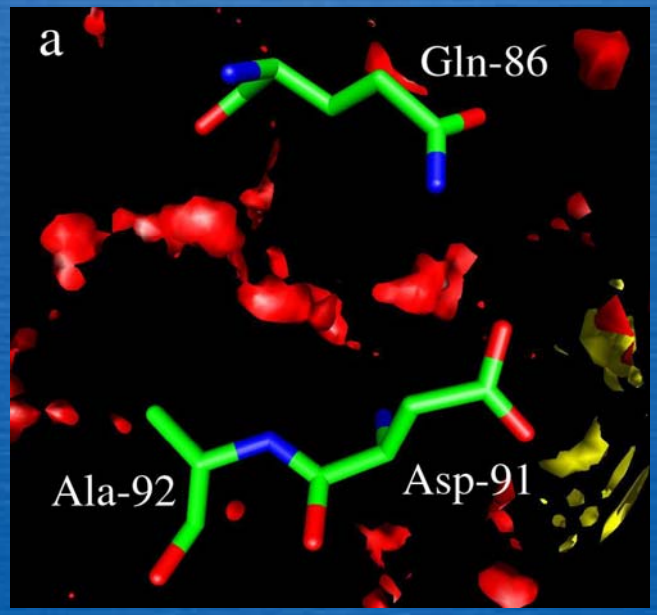
Na^+ (g(r)>3.0)

Active site:
Consists of residues
from Gln-86 to Ala-92

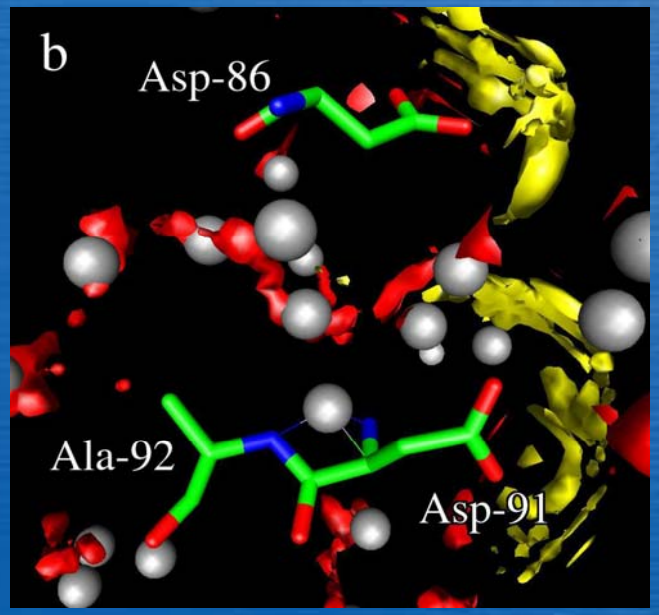
Wild type & Q86D

- In both cases, Na^+ distribution could not be found in active site.
- The reason why is that the negative charged residue are directing outward.
- Water oxygen distribution is consistent with experiment.

Wild type



Q86D



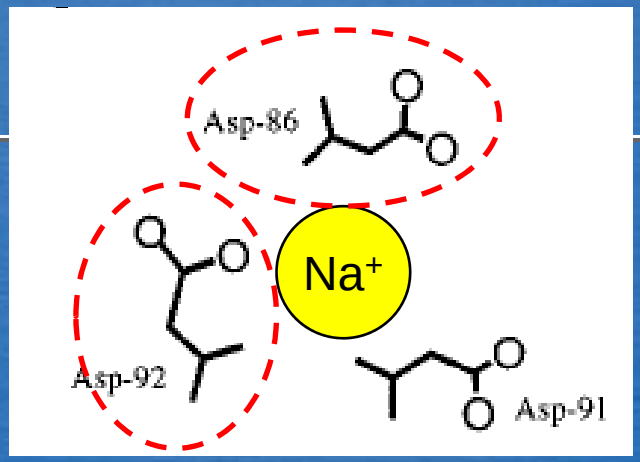
O (g(r)>5.0)

Na^+ (g(r)>5.0)

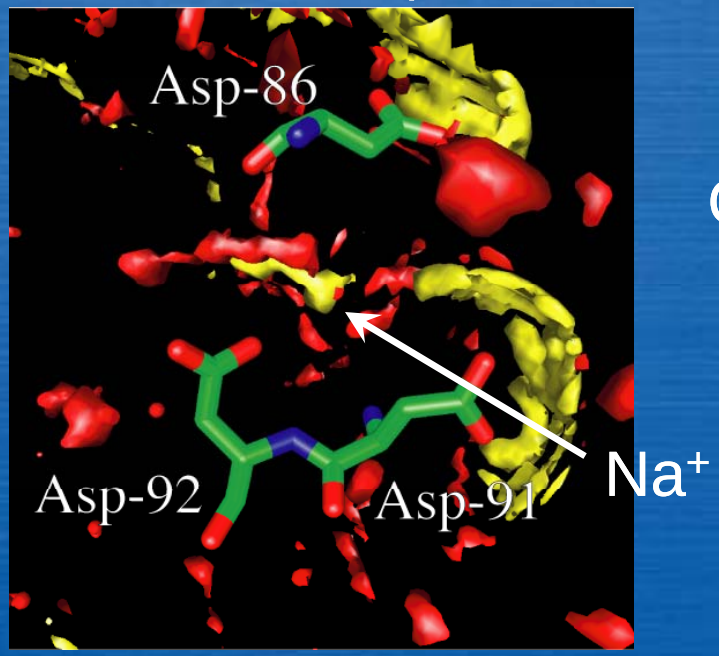
Water (X-ray)

apo-Q86D/A92D

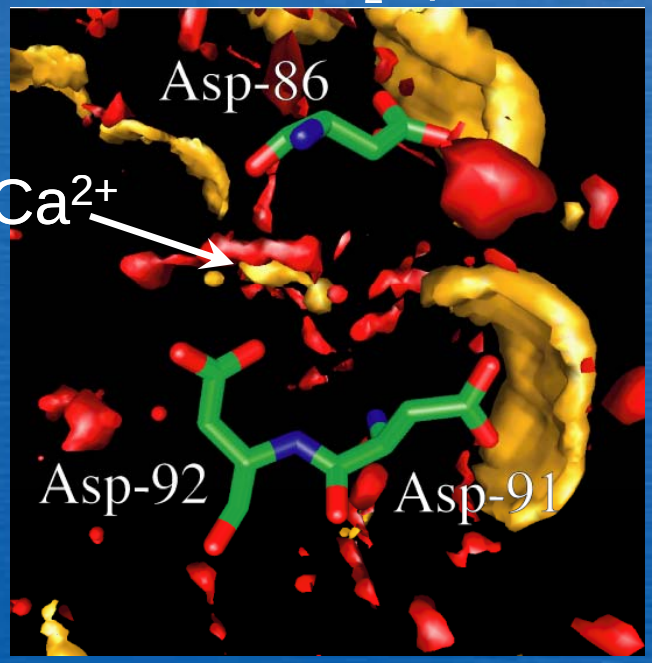
apo mutant have affinity with both cations, Na^+ and Ca^{2+}.



in NaCl*aq*



in $CaCl_2$*aq*



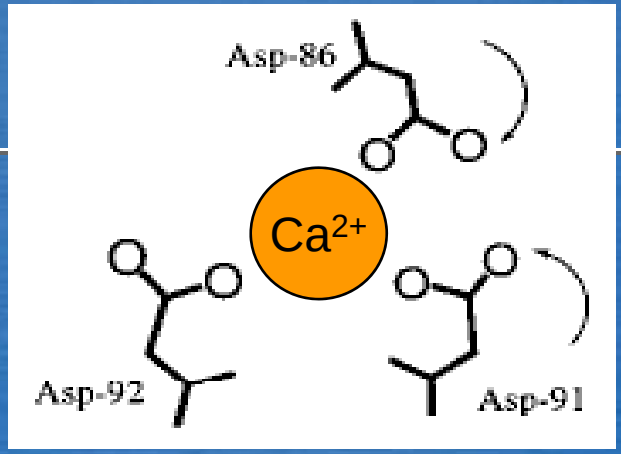
O (g(r)>5.0)

Na^+ (g(r)>5.0)

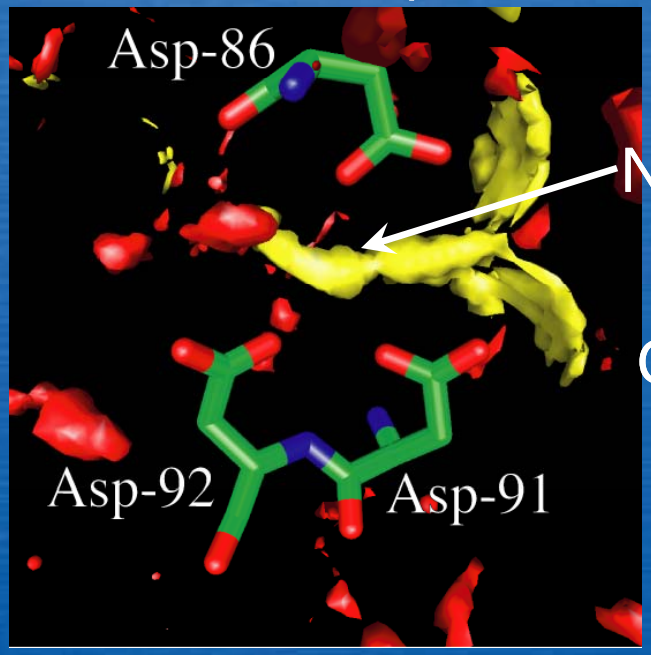
Ca^{2+} (g(r)>5.0)

holo-Q86D/A92D

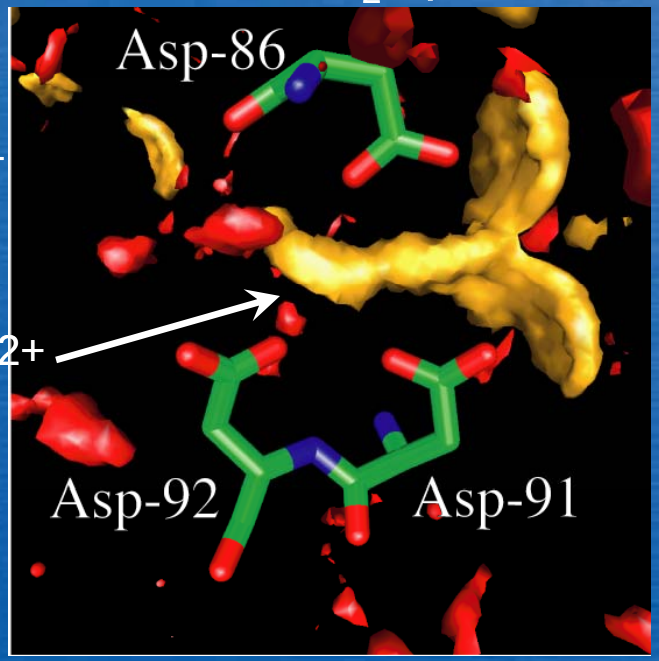
Because three carbonyl oxygen are directing inward, holo mutant can strongly bind the both cation, Na^+ and Ca^{2+}.



in NaCl*aq*



in $CaCl_2$*aq*



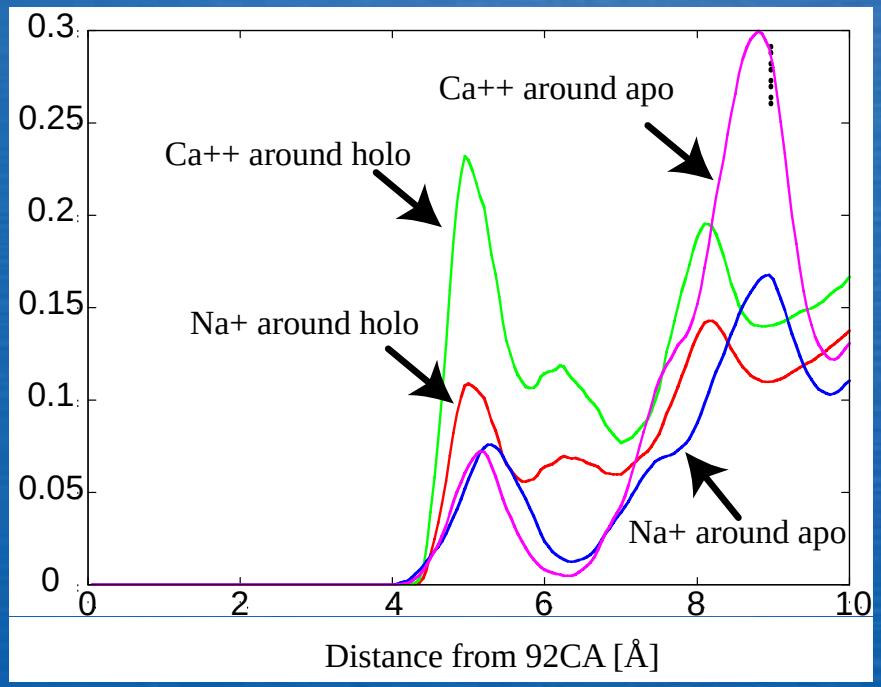
O (g(r)>5.0)

Na^+ (g(r)>5.0)

Ca^{2+} (g(r)>5.0)

apo- & holo-Q86D/A92D

- RDFs of cations for apo and holo mutant in $CaCl_2$*aq* and NaCl*aq*.
- Because of the peak of Ca^{2+} is even the twice as large as that of Na^+, holo mutant selectively binds Ca^{2+} ion.
- In the apo case, the 1st peak for Na^+ is slightly higher and broader than that for Ca^{2+}. ⇒apo has greater affinity to the Na^+.



Summary

- The wild type and Q86D show no cation affinity, and Q86D/A92D have cation affinity in accord with the experiment.
- Driving force of ion-binding is the interaction between ion and active site.
- Although the experimental result indicates that apo-Q86D/A92D binds Na^+ only, the theory predicts that the mutant has ability to recognize both Na^+ and Ca^{2+}.
- The results of this work imply that 3D-RISM theory can predict new mutant which has ability of ion binding.

Future plans

- Quantitative analysis of cation affinity
  - Free energy analysis
- Reproduce the protein structure change due to mutation.
  - Geometry optimization of protein in electrolyte solution.
- Apply the 3D-RISM theory to not only ion-binding problem but also molecular recognition process.

Background

- Target :
  Ion-binding
  by human lysozyme

QuickTimeý Ç²
Sorenson Video 3 êLíÉvÉçÉOÉâÉÄ
Ç™Ç±ÇÄÉsÉNÉ´ÉÉÇ¾å©ÇÈÇžÇ½Ç…ÇÕÍKóvÇ-ÇÅB

holo-Q86D/A92D Human Lysozyme mutant

Reference : Kuroki, R.; Yutani, K. *J. Biol. Chem.* **1998**, *273*, 34310.

3D-RISM Theory

- 3D-RISM equation:

$$h_a^s(\mathbf{r}) = \sum_{\substack{s' \in \text{Solvent} \\ \text{species}}} \sum_{c \in \text{Site on } s'} c_c^{s'}(\mathbf{r}_c) * \left[\varpi_{ca}^s + \rho^{s'} X_{ca}^{s's}(r_{ca}) \right]$$

$$g_a^s(\mathbf{r}) = \begin{cases} \exp(-\beta u_a^s(\mathbf{r}) + h_a^s(\mathbf{r}) - c_a^s(\mathbf{r})) & \text{for } h_a^s(\mathbf{r}) \le 0 \\ 1 - \beta u_a^s(\mathbf{r}) + h_a^s(\mathbf{r}) - c_a^s(\mathbf{r}) & \text{for } h_a^s(\mathbf{r}) > 0 \end{cases}$$

$$g_a^s(\mathbf{r}) = h_a^s(\mathbf{r}) + 1$$

$$u_a^s(\mathbf{r}) = \sum_{b \in \text{solute}} \frac{q_b^u q_a^s}{\left| \mathbf{r} - \mathbf{r}_b^u \right|} + \sum_{b \in \text{solute}} 4\varepsilon_{ab} \left\{ \left(\frac{\sigma_{ab}}{r_{ab}} \right)^{12} - \left(\frac{\sigma_{ab}}{r_{ab}} \right)^{6} \right\}$$

3D-RISM Theory

- Computational Scheme:

1. Solve the solvent-solvent RISM and get $X_{ca}^{s's}(r_{ca})$ and solvent intramolecular correlation function ϖ_{ca}^{s}.

2. Choose appropriate structure and potential parameters of solute molecule.

3. Set the solute-solvent interaction potential $u_a^s(\mathbf{r})$.

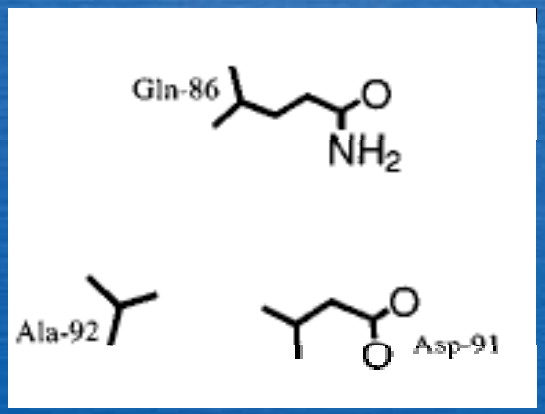
4. Solve the 3D-RISM equations and get 3D-DF, $g_a^s(\mathbf{r}_a)$.

Computational Detail

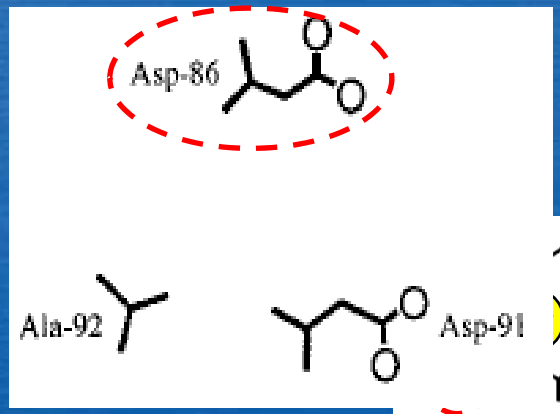
- We considered a human lysozyme in the electrolyte solution at infinite dilution.
- Lysozyme immerges to 0.01 M electrolyte solution. (KCl, NaCl, $CaCl_2$)
- All potential parameters of lysozyme from Amber 99.
- Protein structures are taken from PDB.
- SPC model is employed as solvent water, OPLS parameters are used for ions.

Wild type & Q86D

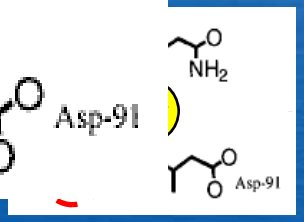
	83	84	85	86	87	88	89	90	91	92
Wild type	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Ala
Q86D	Ala	Leu	Leu	Asp	Asp	Asn	Ile	Ala	Asp	Ala
A92D	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Asp
Q86D/A92D	Ala	Leu	Leu	Asp	Asp	Asn	Ile	Ala	Asp	Asp



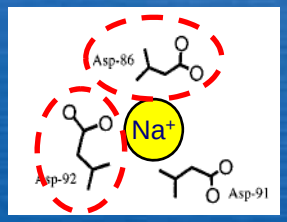
wild type

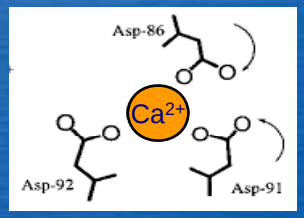


Q86D



A92D

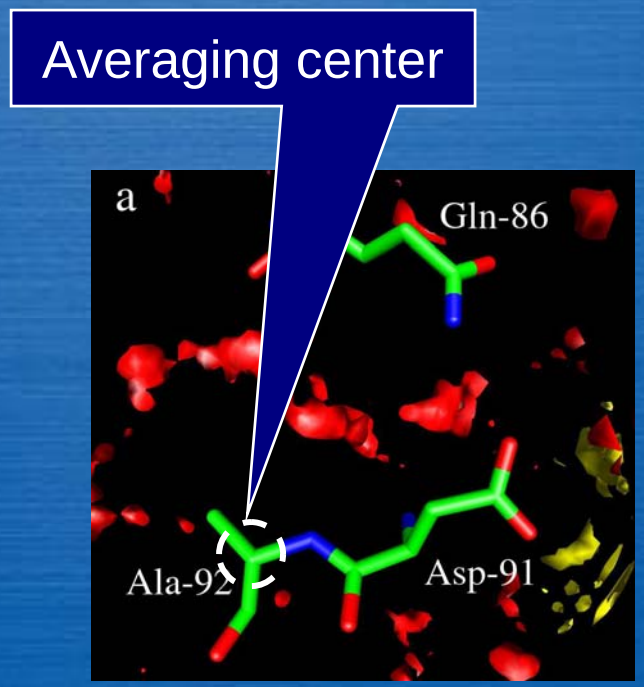


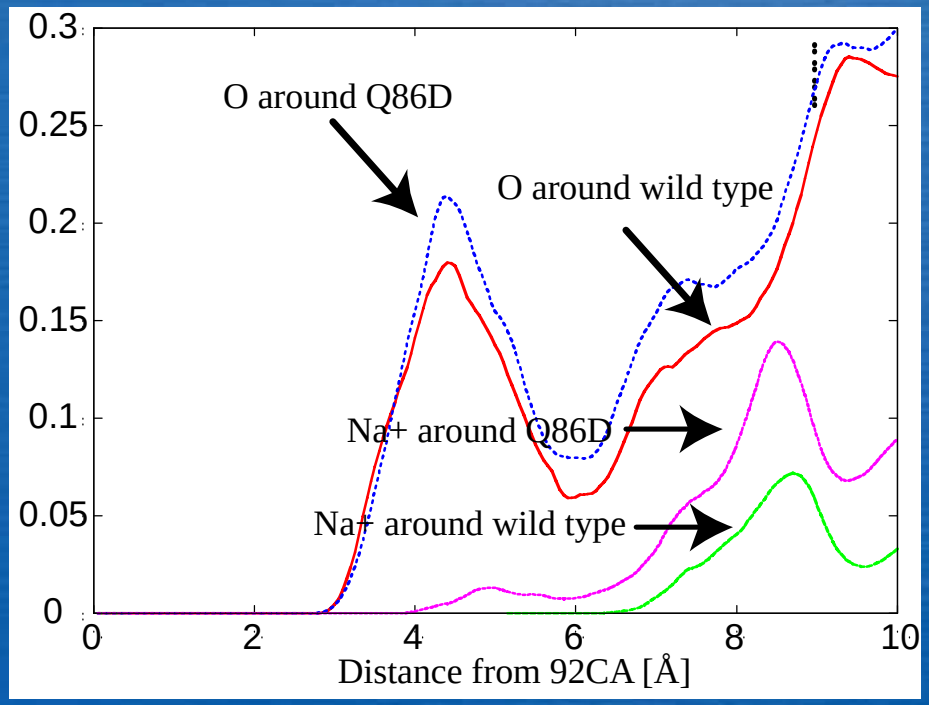


Q86D/A92D

Wild type & Q86D

- This figure shows the 1D radial distribution functions evaluated by the orientational averaging of the 3D-DFs.
  In this figure, there are no peaks of ions.

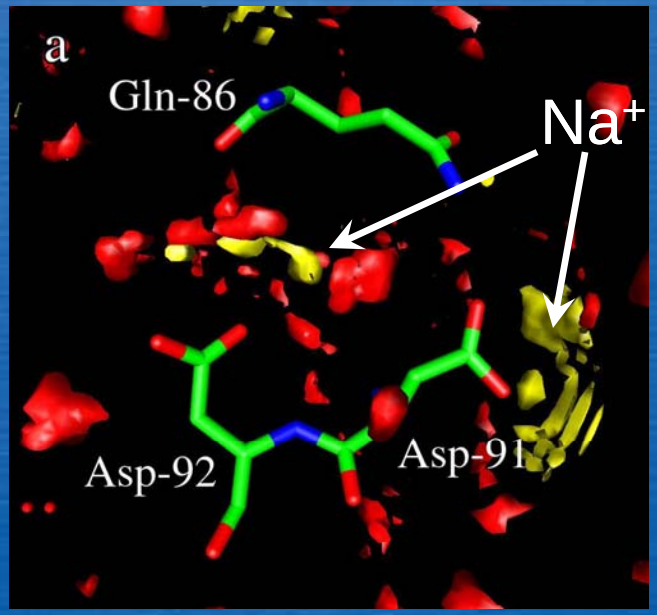




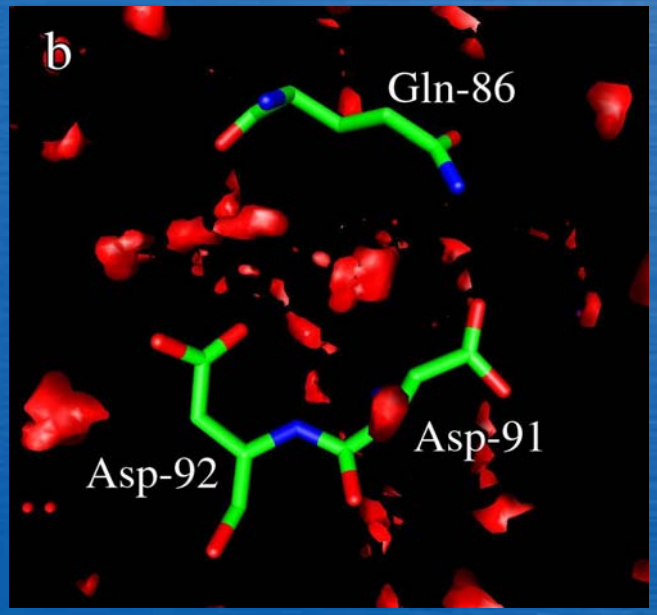
A92D

- By replacing Ala with Asp, charge of the inside of active site becomes negative. ⇒ Na^+ can be binded.
- K^+ cannot be found in active site. ⇒ This selectivity should originate from the difference in size of the ion species.

In NaCl*aq*



In KCl*aq*



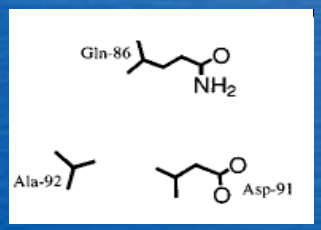
O (g(r)>5.0)

Na^+ (g(r)>5.0)

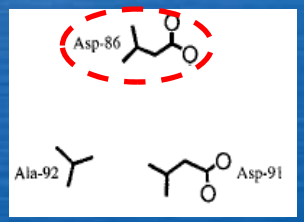
K^+ (g(r)>5.0)

A92D

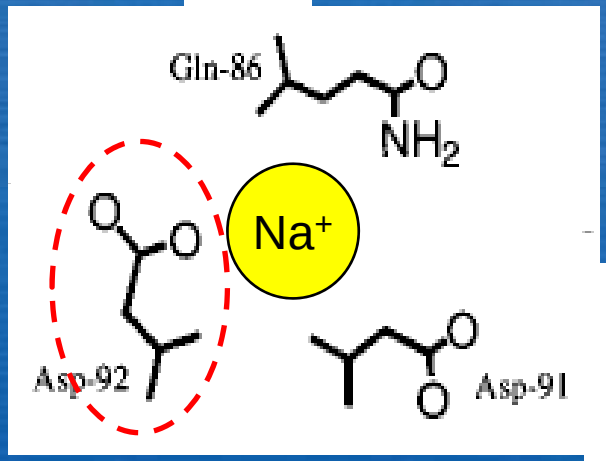
	83	84	85	86	87	88	89	90	91	92
Wild type	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Ala
Q86D	Ala	Leu	Leu	Asp	Asp	Asn	Ile	Ala	Asp	Ala
A92D	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Asp
Q86D/A92D	Ala	Leu	Leu	Asp	Asp	Asn	Ile	Ala	Asp	Asp



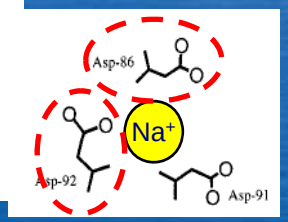
wild type

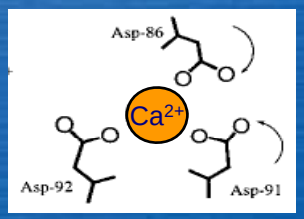


Q86D



A92D

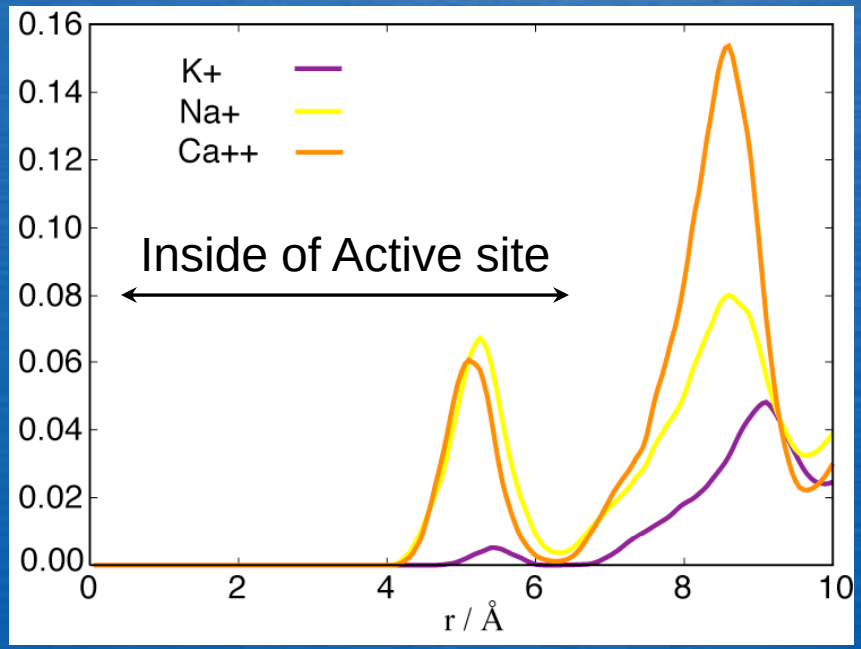




Q86D/A92D

A92D

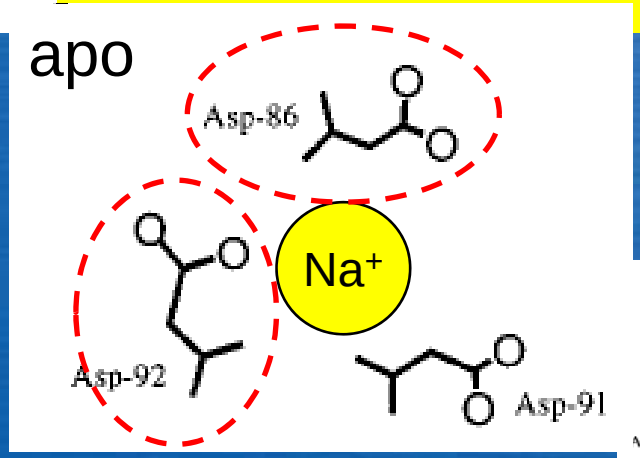
- There are the peaks of Na^+ and Ca^{2+} in the active site, but no peaks of K^+.
- For A92D, the Ca^{2+} ion was found in the cleft contrary to the experimental results.
- The 1st peak for Na^+ is slightly higher and broader than that for Ca^{2+}.
  $\Rightarrow Ca^{2+}$ has less affinity to the active site.

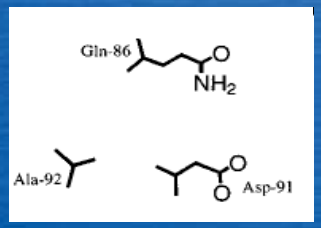


apo-Q86D/A92D

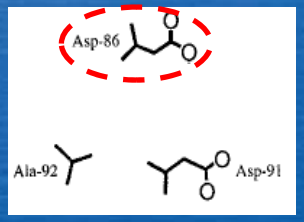
	83	84	85	86	87	88	89	90	91	92
Wild type	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Ala
Q86D	Ala	Leu	Leu	Asp	Asp	Asn	Ile	Ala	Asp	Ala
A92D	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Asp
Q86D/A92D	Ala	Leu	Leu	Asp	Asp	Asn	Ile	Ala	Asp	Asp

apo

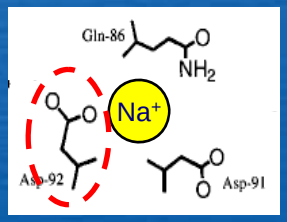




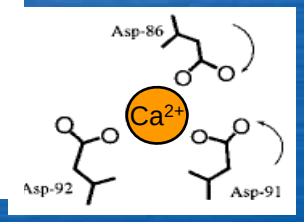
wild type



Q86D



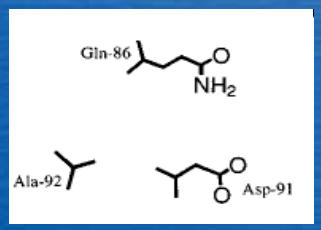
A92D



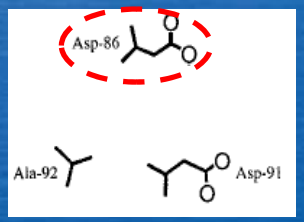
Q86D/A92D

holo-Q86D/A92D

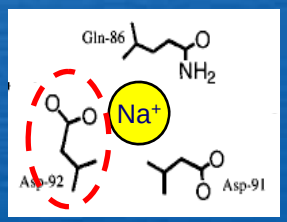
	83	84	85	86	87	88	89	90	91	92
Wild type	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Ala
Q86D	Ala	Leu	Leu	Asp	Asp	Asn	Ile	Ala	Asp	Ala
A92D	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Asp
Q86D/A92D	Ala	Leu	Leu	Asp	Asp	Asn	Ile	Ala	Asp	Asp



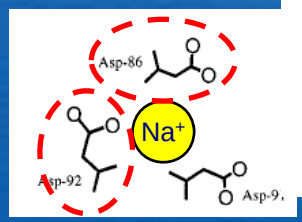
wild type

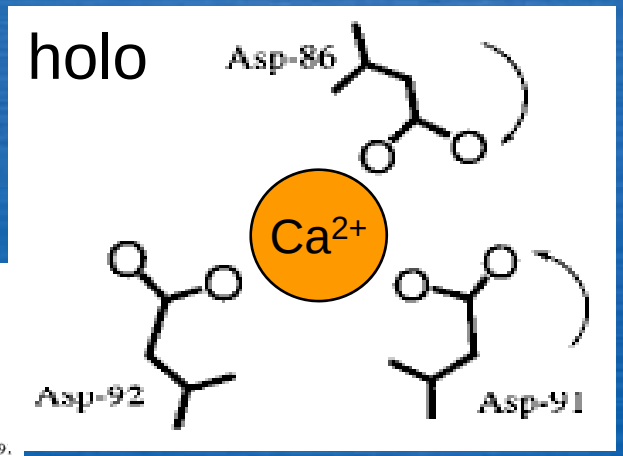


Q86D



A92D





Q86D/A92D

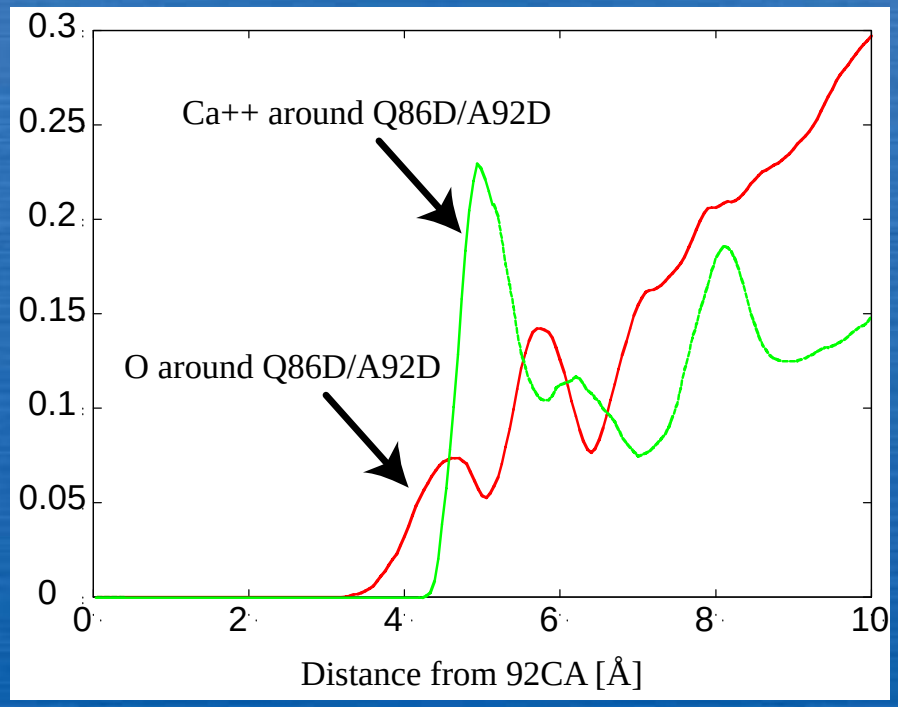
Acknowledgement

- This work is supported by the Grant-in Aid for Scientific Research on Priority Area of 'Water and Biomolecules' from the Ministry of Education in Japan, Culture, Sports, Science and Technology (MONBUKAGAKUSHO).

- We are also grateful to the support by the grant from the Next Generation Supercomputing Project, Nanoscience Program, MEXT, Japan.

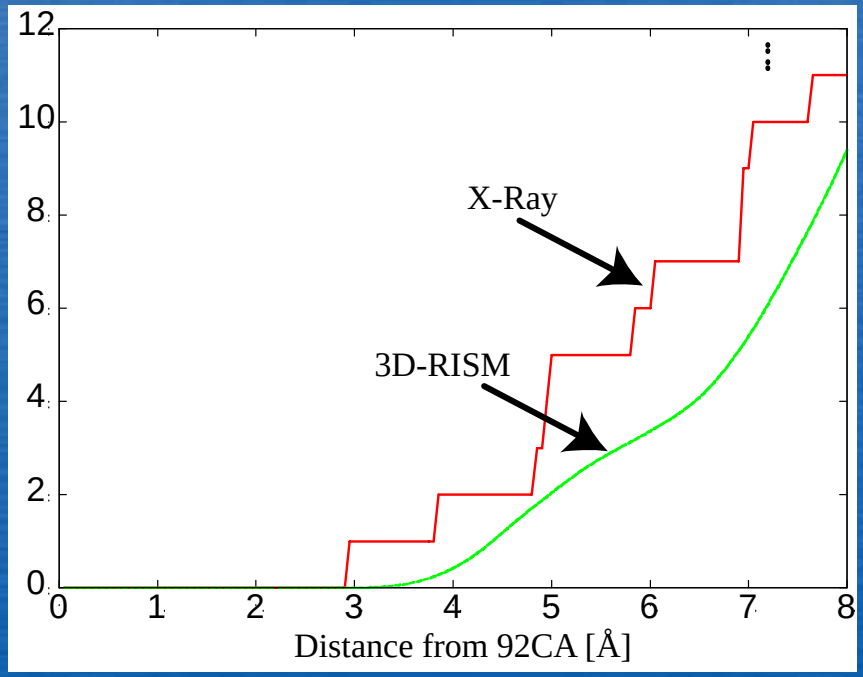
Appendix 1

- 1D-RDFs of oxygen and calcium ions around holo-Q86D/A92D mutant.
- Center of orientation averaging is 92CA.



Appendix 2

- Comparison of coordination number of water oxygen around CA of residu 92 between X-Ray and theory.



Appendix

- Molecular OZ-equation for mixture

$$h^{\alpha\beta}(12) = c^{\alpha\beta}(12) + \sum_{\gamma \in \text{All species}} \rho^{\gamma} \int c^{\alpha\gamma}(13) h^{\gamma\beta}(32) d(3)$$

- Protein (ly□sozyme) is assumed to be□ a solute that infinitely diluted to the electrolyte solution.

$$\rho^{\text{lysozyme}} = 0$$

- This assumption allowed us to split the OZ equation to two type set of equations.
  - □ Solute-solvent system

$$h^{\alpha s}(12) = c^{\alpha s}(12) + \sum_{s' \in \text{Solvent species}} \rho^{s'} \int c^{\alpha s'}(13) h^{s's}(32) d(3)$$

  - □ Solvent-solvent system

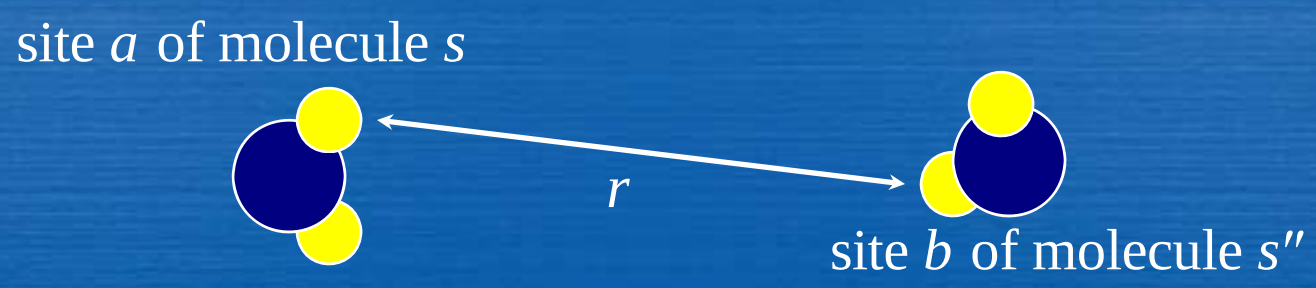
$$h^{ss''}(12) = c^{ss''}(12) + \sum_{s' \in \text{Solvent species}} \rho^{s'} \int c^{ss'}(13) h^{s's''}(32) d(3)$$

Appendix

- Strategy of RISM theory

  The pair correlation functions between two molecules is expressed in terms of those between interaction sites, which can be accomplished by averaging the functions over orientations fixing the distance between the interaction sites:

$$h_{ab}^{ss'}(r) = \frac{1}{\Omega^2}\int d(1)d(2)\delta(\mathbf{R}_1 + \mathbf{l}_1^a)\delta(\mathbf{R}_2 + \mathbf{l}_2^b - \mathbf{r})h^{ss'}(12)$$



Appendix

- RISM equation

By these averaging operation, molecular OZ equation becomes 1D-RISM equation:

$$h_{ab}^{ss''}(r) = \sum_{c,d \in \text{Site on a,b}} \varpi_{ac}^{s} * c_{cd}^{ss''}(r) * \varpi_{db}^{s''} + \sum_{s' \in \text{Solvent species}} \rho^{s'} \sum_{c,d \in \text{Site on s,s}'} \varpi_{ac}^{s} * c_{cd}^{ss'}(r_{cd}) * h_{db}^{s's''}(r_{db})$$

In this study, 1D-RISM theory was employed for solvent-solvent system.

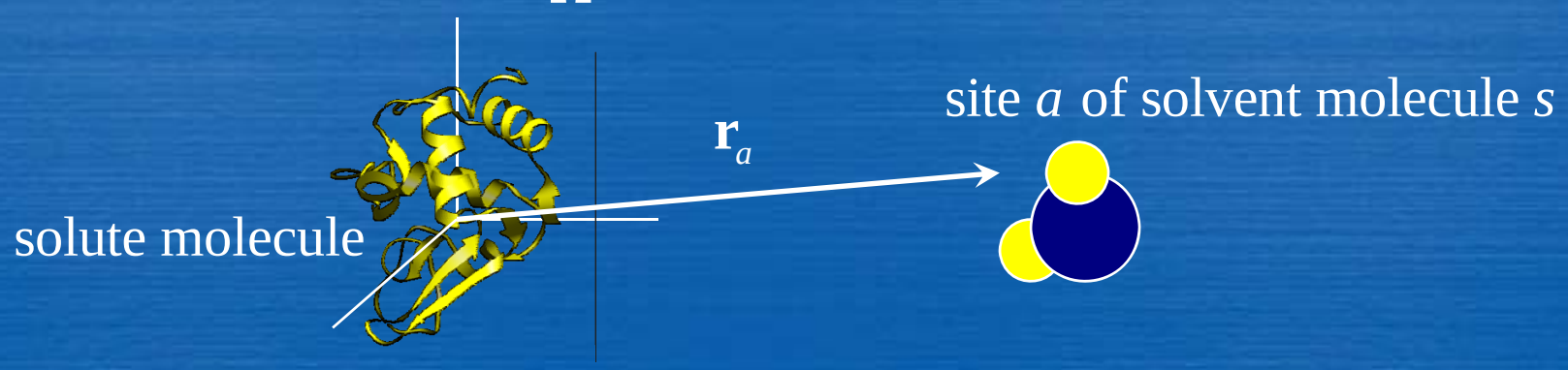
Appendix

- For solute-solvent system

  Because protein (lysozyme) is fully anisotropic molecule, orientation dependency of the solute-solvent correlation functions cannot be ignored.

  For solute-solvent system, □□the orientation average is taken over just for solvent molecules, not for the solute molecule.

$$h_a^s(\mathbf{r}) = \frac{1}{\Omega}\int d(2)\delta(\mathbf{R}_2 + \mathbf{l}_2^a - \mathbf{r})h^s(12)$$



Appendix

- 3D-RISM theory

  By this averaging, the pair correlation functions become molecular-site correlation functions, which depend on the vector connecting solute molecule and solvent site.

  This averaging reduces the solute-solvent MOZ to the 3D-RISM equation:

$$h_a^s(\mathbf{r}) = \sum_{\substack{s' \in \text{Solvent} \\ \text{species}}} \sum_{c \in \text{Site on } s'} c_c^{s'}(\mathbf{r}_c) * \left[\varpi_{ca}^s + \rho^{s'} h_{ca}^{s's}(r_{ca}) \right]$$

Appendix

- Coordination number of cations in binding pocket surrounded residu 86, 91 and 92.
- Coordination number is quite small for all cases. This originates the fact that the density of the ion is low.

	Na^+	Ca^{2+}	Na^+(Mix)	Ca^{2+}(Mix)
A92D	0.00013	0.00010		
apo-Q86D/A92D	0.00018	0.00013	0.00019	0.00014
holo-Q86D/A92D	0.00033	0.00063	0.00034	0.00065

Appendix

- Estimate 1D-RDF from 3D-DF

  The RDFs can be obtained by averaging the 3D-distrubution function over the direction around a specified center:

$$g_a^{1\mathrm{D}}(r, \mathbf{r}_0) = \frac{1}{4\pi} \int g_a(\mathbf{r}_0 + \mathbf{r}) d\hat{\mathbf{r}}$$

$\mathbf{r}_0$: avaraging center (α-carbon of 92)
$\hat{\mathbf{r}}$: direction of $\mathbf{r}$