# Effective Protein-Protein Interaction and Clustering Phenomenon in Solution Studied by Small-Angle Neutron Scattering

Sow-Hsin Chen (sowhsin@mit.edu)

*Department of Nuclear Science and Engineering, MIT*

Collaborators:

Yun Liu (MIT), Emiliano Fratini (U of Florence), Piero Baglioni (U of Florence), Wei-Ren Chen (MIT)

Research supported by Material Sciences Program, BES, US DOE.

Presented at *Symposium on Bio-Materials and Neutron, AVS Meeting in Boston, October 31, 2005*

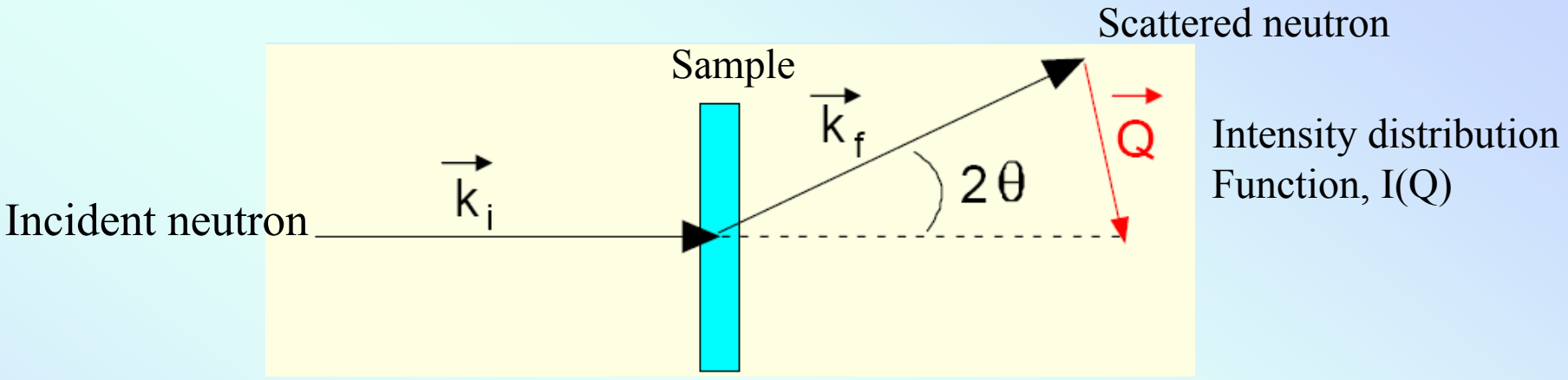
# Outline

1. Introduction to small angle neutron scattering and its applications to study of clustering phenomenon in protein solutions.

2. The effective interaction potential between protein molecules in solution.

- Experimental & theoretical evidence for DLVO effective potential
- Experimental & theoretical evidence of non-DLVO effective potential
- Existence of a new effective long range attraction between protein molecules studied by SANS

# Introduction to small angle neutron scattering (SANS)




SANS is an elastic neutron scattering technique, where

$$\left|\vec{k}_f\right| = \left|\vec{k}_i\right| = k = \frac{2\pi}{\lambda}$$

SANS probes periodic structure in the direction of $\vec{Q}$, where

$$\vec{Q} = \vec{k}_i - \vec{k}_f, \qquad Q = 2k\sin\theta = (2\pi/\lambda)\sin\theta$$

and Q is called the magnitude of scattering wave vector.

For an isotropic system, $I(\vec{Q}) = \frac{1}{V}\frac{d\sigma}{d\Omega} = \frac{N}{V}\sum_{m,n}\overline{b_m b_n}\frac{1}{N}\left\langle e^{-i\vec{Q}\cdot\vec{r}_m} e^{i\vec{Q}\cdot\vec{r}_n}\right\rangle = I(Q)$

# What SANS measures is I(Q)

For a one component isotropic colloidal system, the intensity can be written as

$$I(Q) = \frac{d\Sigma(Q)}{d\Omega} = N_p(\rho_p - \rho_s)^2 V_p^2 \overline{P}(Q) S(Q)$$

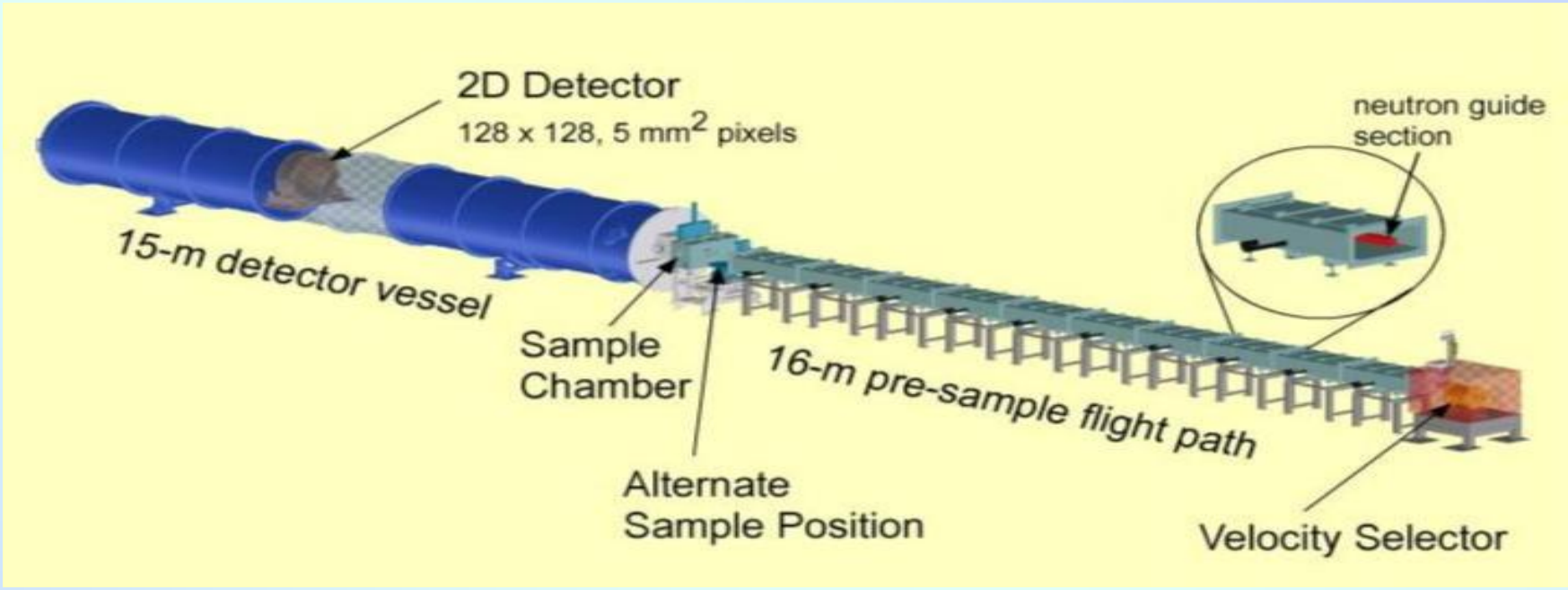
$N_p$: the number density of particles in solution
$\rho_p$, $\rho_s$ : scattering length densities of the colloidal particle and of the solvent
$V_p$: the volume of an individual particle

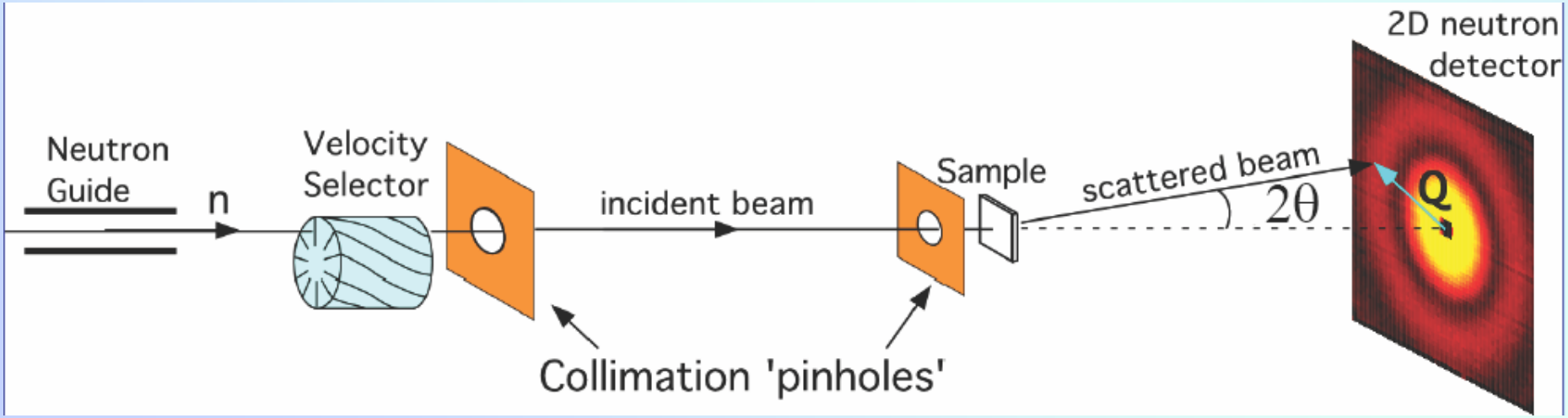
$\overline{P}(Q)$ : the normalized particle structure factor, which is determined by the composition, size and shape of the individual colloidal particle.
$S(Q)$ : the inter-particle structure factor, which is determined by the particle pair-correlation function.

Applications to protein solutions: size, charge number, hydration level
Applications to micelle solutions: volume fraction, aggregation number, etc.

SANS probes nano-structures, typical range: 1 nm ~ 100 nm.

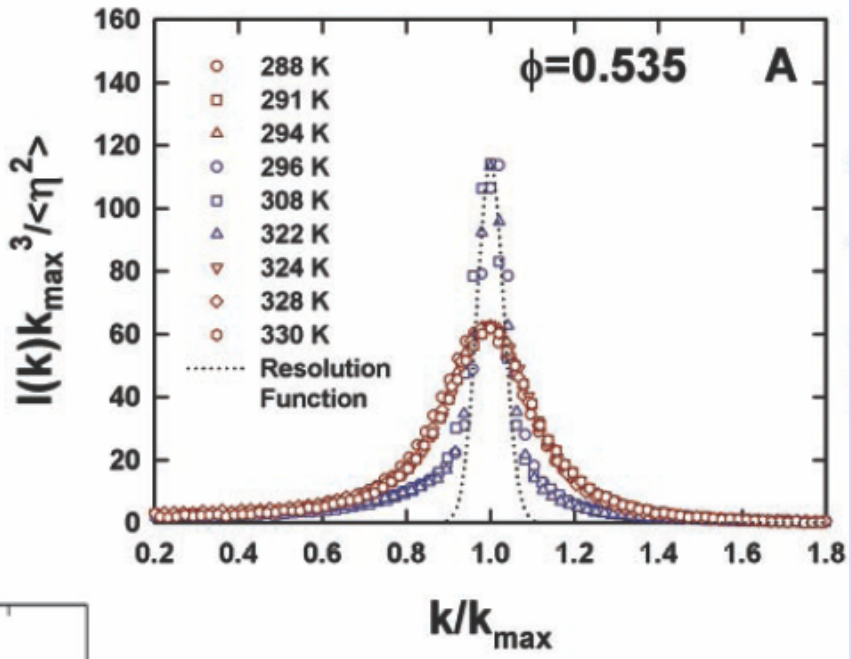
# SANS instrument NG3 at NIST NCNR



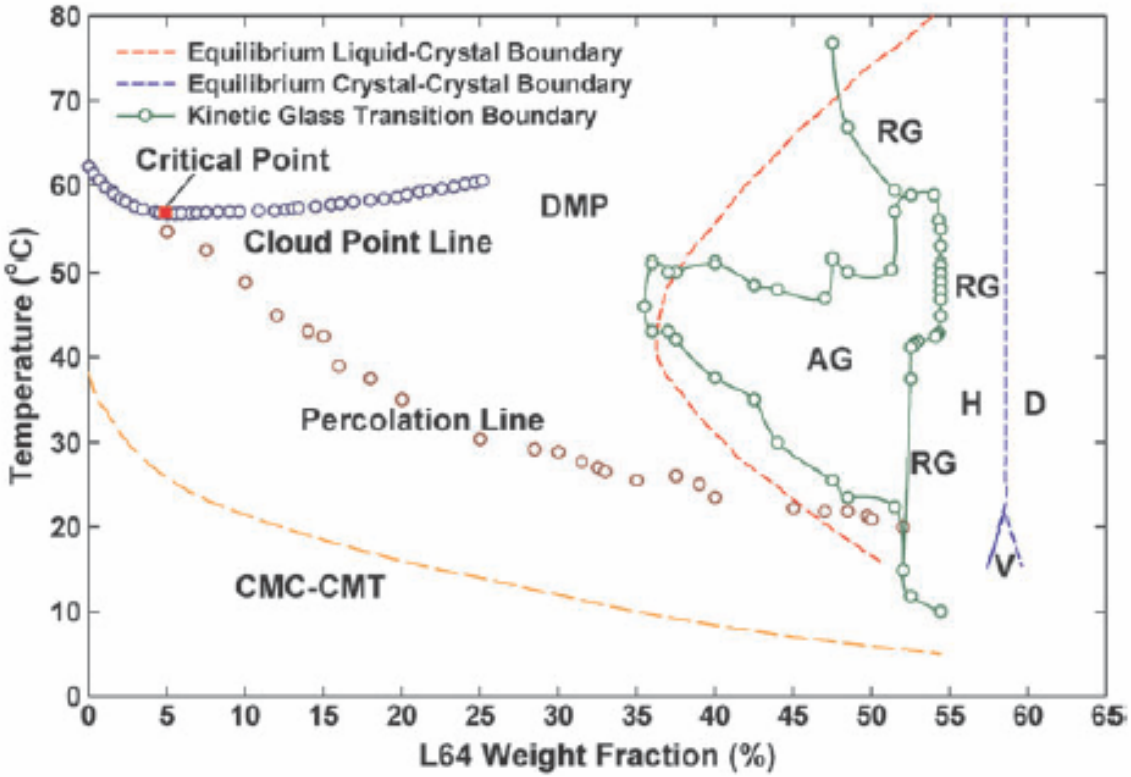



# Applications of SANS (I) – Structural Arrest Transition

S. H. Chen, et al., *Science* **300**, 5619 (2003)



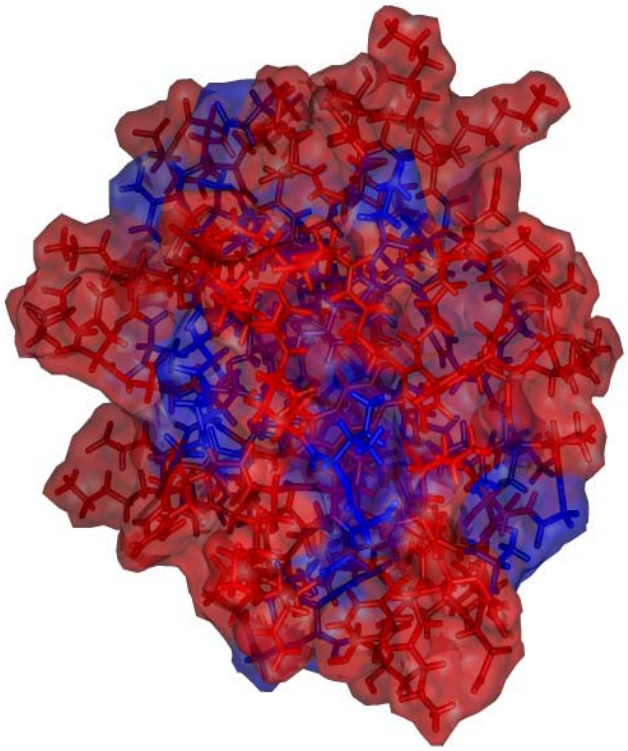

Liquid-to-glass transition




Structural arrest phase diagram

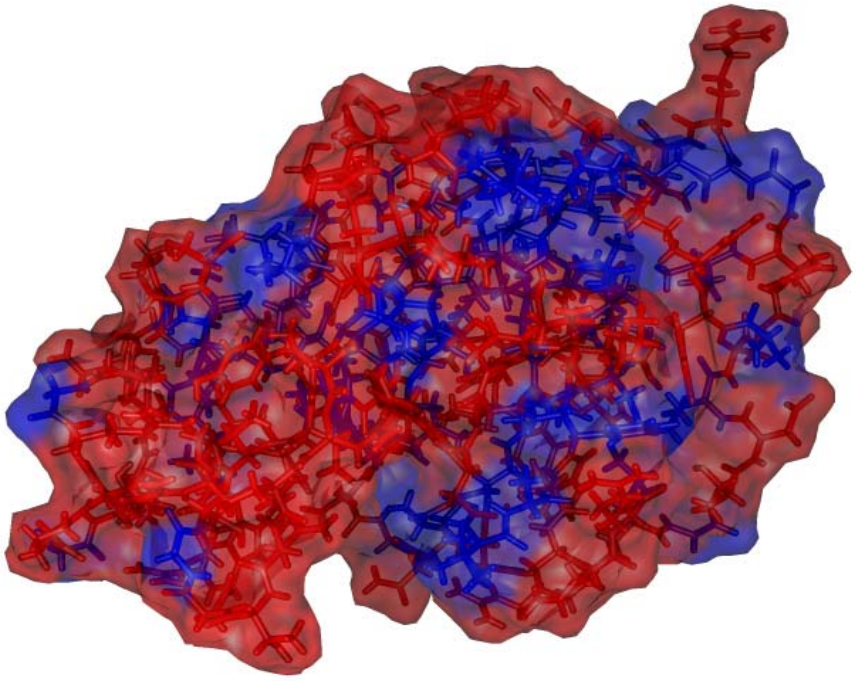
# Spherical proteins: cytochrome C and lysozyme

Dimension: 15×17×17Å$^3$
PI: 10.2  MW: 12,384 Da



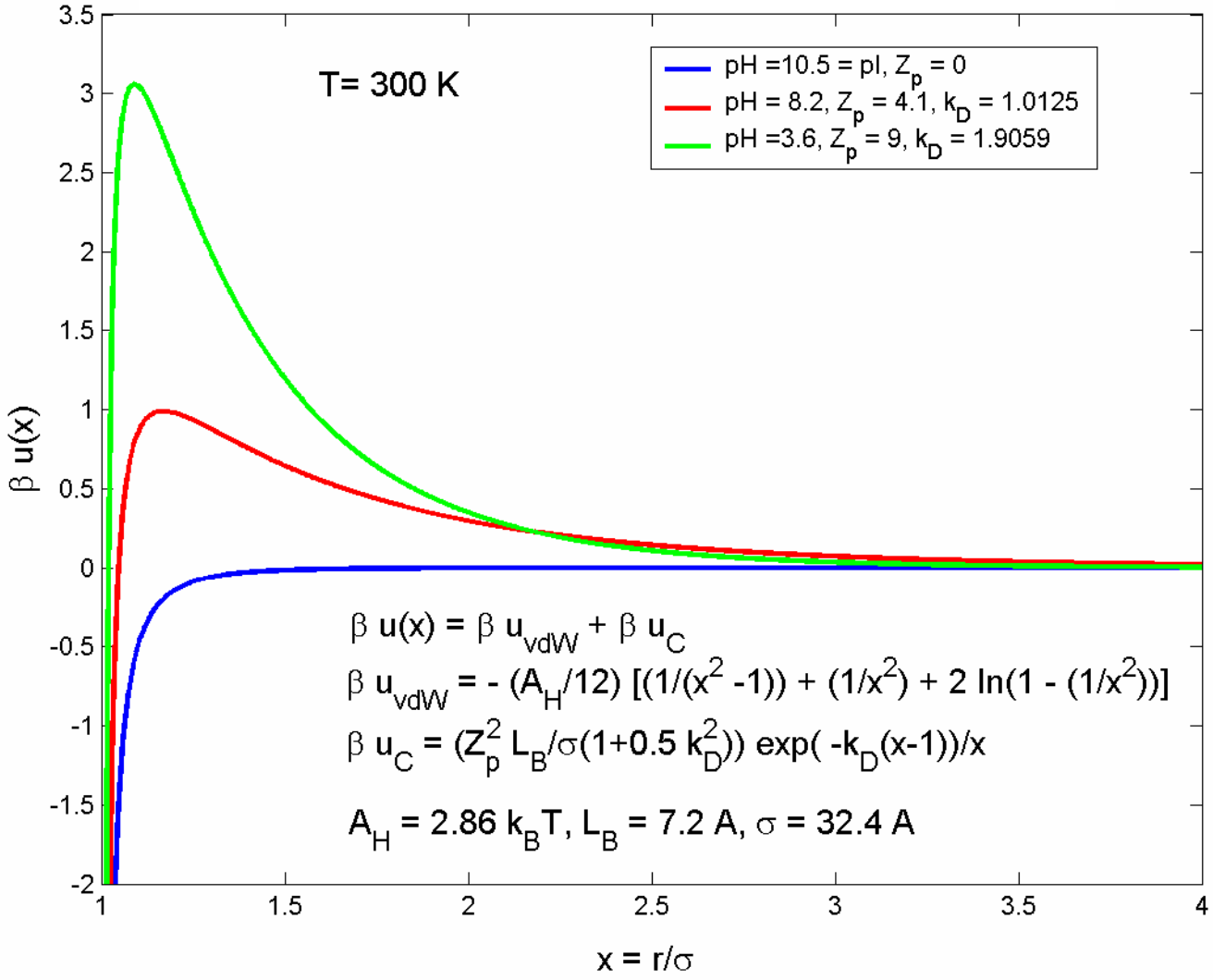
Cytochrome C

Dimension: 22.5×15×15Å$^3$
PI: 10.7  MW: 14,400 Da



Lysozyme

# DLVO (Derjaguin-Landau-Verwey-Overbeek) potential obtained from fitting of SAXS data of lysozyme solution (85 mg/ml)




A. Tardieu et al., J. Crystal Growth 196, 193 (1999)

# Neutron and X-ray scattering studies of protein solutons

## Evidence for the screened Coulomb repulsion

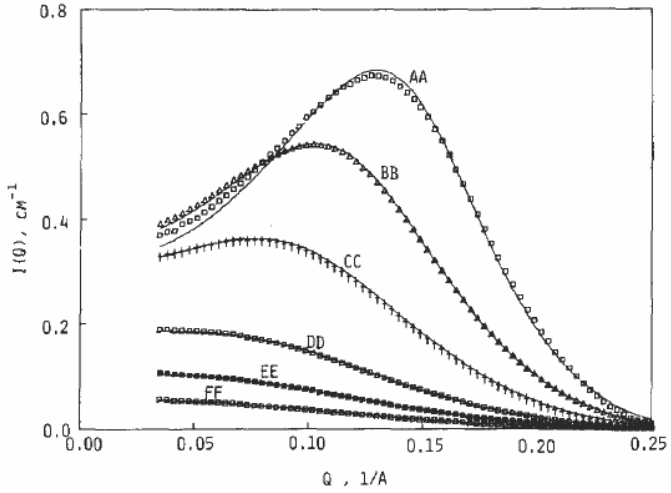
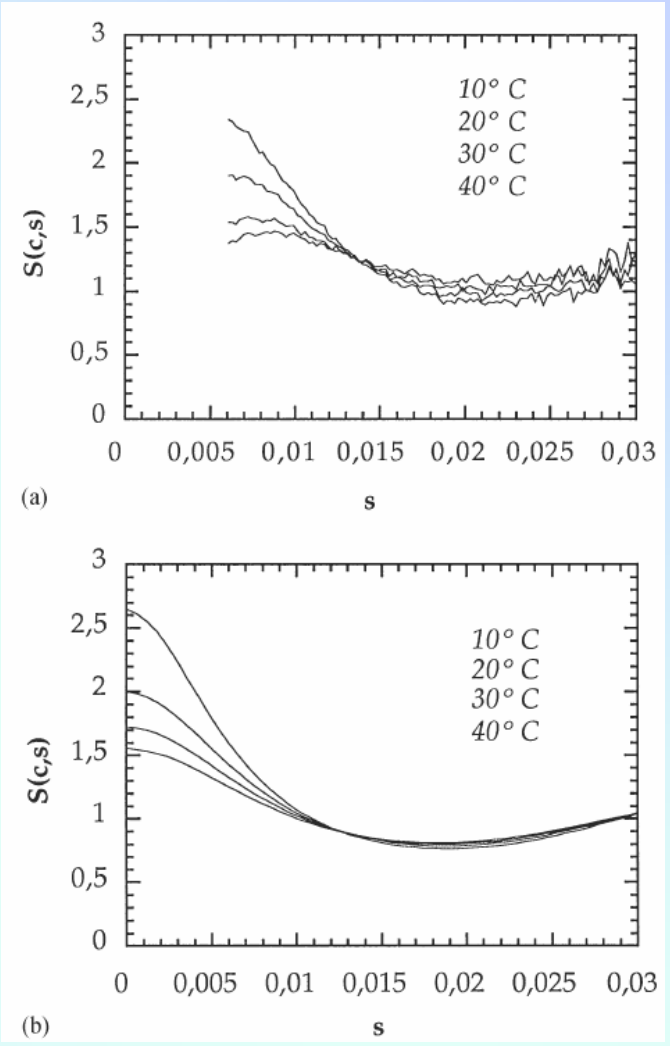



FIG. 5. Experimental scattering intensities (symbols) and GOCM-fitted curves (solid lines) of cytochrome *C* in 100 mM acetate buffer solutions. Volume fractions are AA = 18.13%, BB = 9.06%, CC = 4.53%, DD = 1.81%, EE = 0.91%, and FF = 0.45%.

*A. Tardieu, et al., J. Crystal Growth* ***196****, 193 (1999)*

Evidence for the existence of a short-range attraction

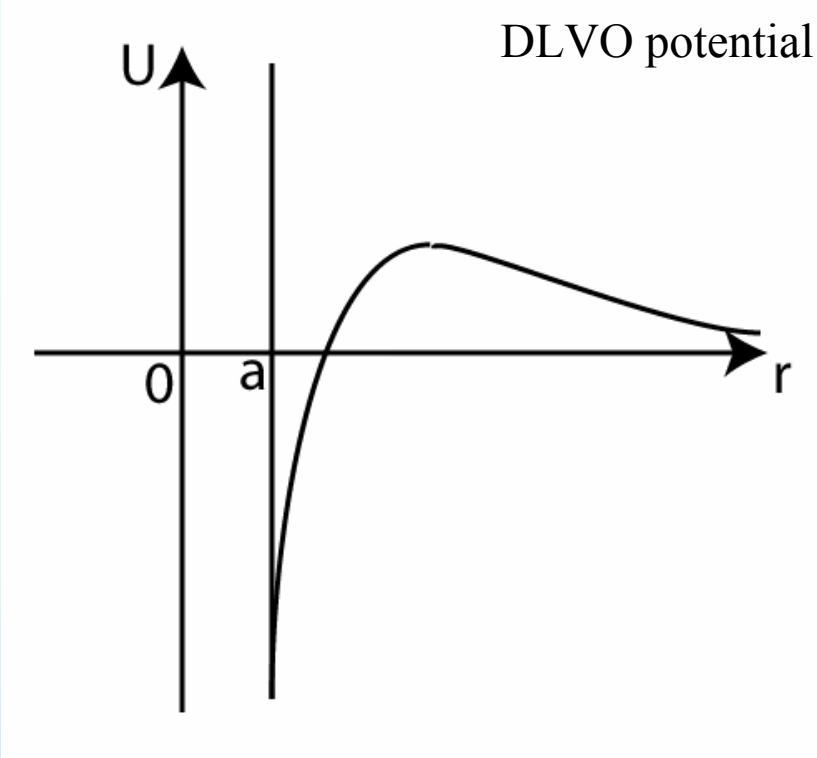



*C. F. Wu, S. H. Chen, J. Chem. Phys.* ***87****, 6199 (1987)*

Chen and coworkers found that by assuming there is only electrostatic repulsion and using the GOCM theory, the small angle neutron scattering intensity distribution, I(Q), of cytochrome C solutions can be fitted with excellent agreement.

By assuming the DLVO potential between protein particles, the small angle X-ray intensity distribution of lysozyme solutions can be reasonably fitted.

# Short-range attraction + long-range repulsion




Is this the complete picture of the effective inter-protein potential?

# Studies of Structures of Two-Yukawa Fluid

- Two Yukawa fluid is a simple fluid system whose effective inter-particle potential, *V(r)*, is in the form of the two Yukawa form plus a hard core,

$$\beta V(r) = \begin{cases} \infty & \text{for } r < 1; \\ -K_1 \dfrac{e^{-Z_1(r-1)}}{r} - K_2 \dfrac{e^{-Z_2(r-1)}}{r} & \text{for } r > 1 \end{cases}$$

- We develop an efficient analytical theory to calculate the inter-particle structure factor, S(Q), of two Yukawa fluid.
- Apply this method to study different colloidal solutions.
- The scattering intensity distribution of SANS and SAXS is proportional to S(Q)
- S(Q) can be used as the input for the mode coupling theory to predict the kinetic phase diagram (structural arrest transition boundary).

J. Wu, Y. Liu, W.R. Chen, J. Cao, S.H. Chen, "Structural Arrest Transitions in Fluids Described by Two Yukawa Potentials", *Physical Review E* Rapid Comm.,**70,** 050401, (2004).

## An efficient method of obtaining an analytical solution of structure factor for two Yukawa Fluids in Mean Spherical Approximation (MSA)

The Orstein-Zernike equation is :

$$h(r) = c(r) + \int c(|\vec{r} - \vec{r}'|)h(\vec{r}')d\vec{r}'$$

The MSA closure is :

$$\begin{cases} h(r) = -1 & \text{for } r < 1; \\ c(r) = K_1 \dfrac{e^{-Z_1(r-1)}}{r} + K_2 \dfrac{e^{-Z_2(r-1)}}{r} & \text{for } r > 1. \end{cases}$$

The direct correlation function *c(r)* inside the core (*r<1*) can be solved from the OZ equation in the form:

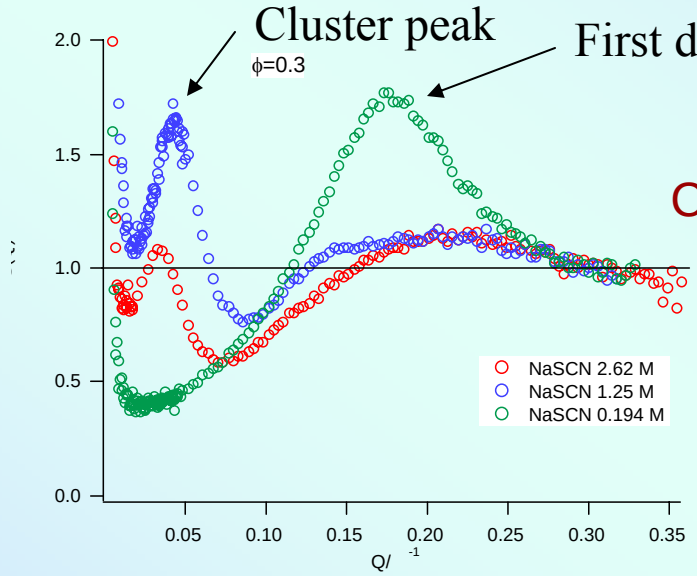
$$-rc(r) = a_0 r + b_0 r^2 + \frac{1}{2}\xi a_0 r^4 + \sum_{i=1,2} \frac{v_i}{z_i}(1 - e^{-z_i r}) + \sum_{i=1,2} \frac{v_i}{2K_i e^{z_i} z_i^2}(\cosh(z_i r) - 1)$$

where $a_0$, $b_0$, $\xi$, and $v_i$ are functions of $K_1$, $K_2$, $Z_1$, $Z_2$ and volume fraction $\varphi$.

Therefore, the structure factor $S(Q) = \dfrac{1}{1 - \rho c(Q)}$

*Y. Liu, W. R. Chen, S. H. Chen, Journal of Chemical Physics, 122, 044507 (2005)*

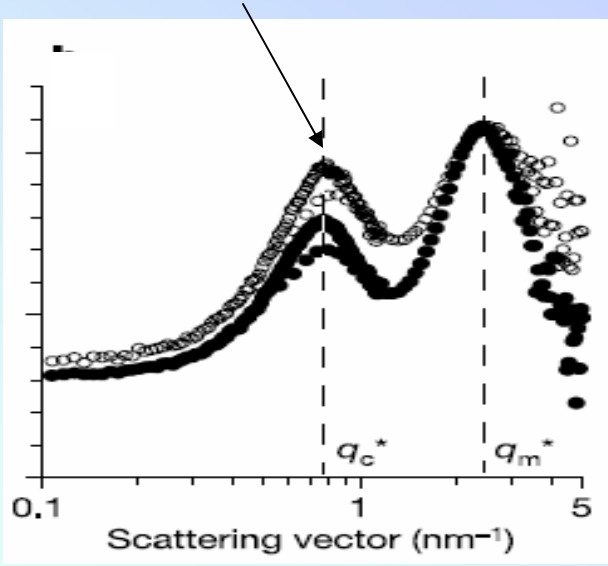
# Equilibrium Cluster Formation Phenomenon




Cytochrome C

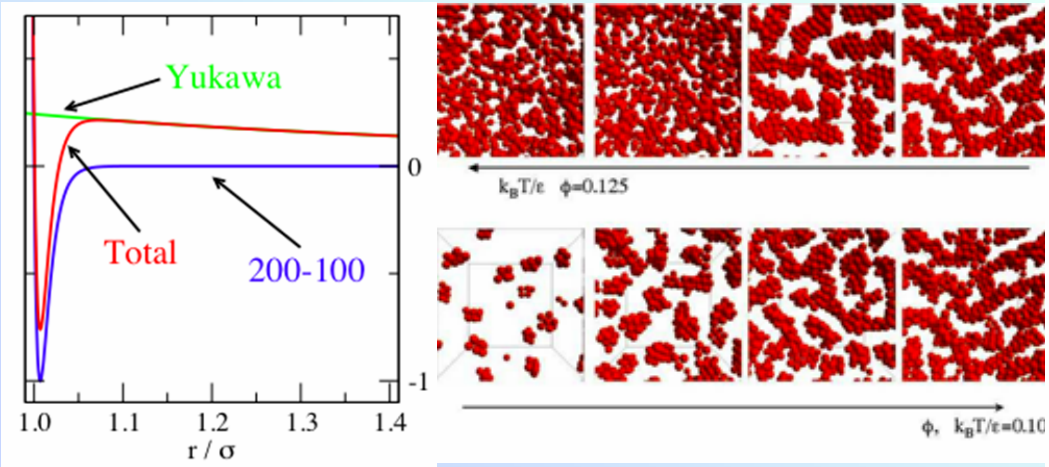
B. Lonetti et al, Phys. Chem. Chem. Phys. **6**, 1388, (2004)

Cluster peak (c independent)

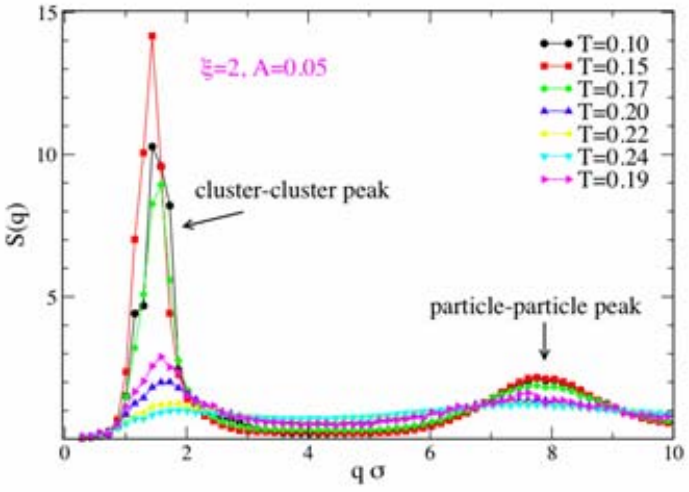



Lysozyme

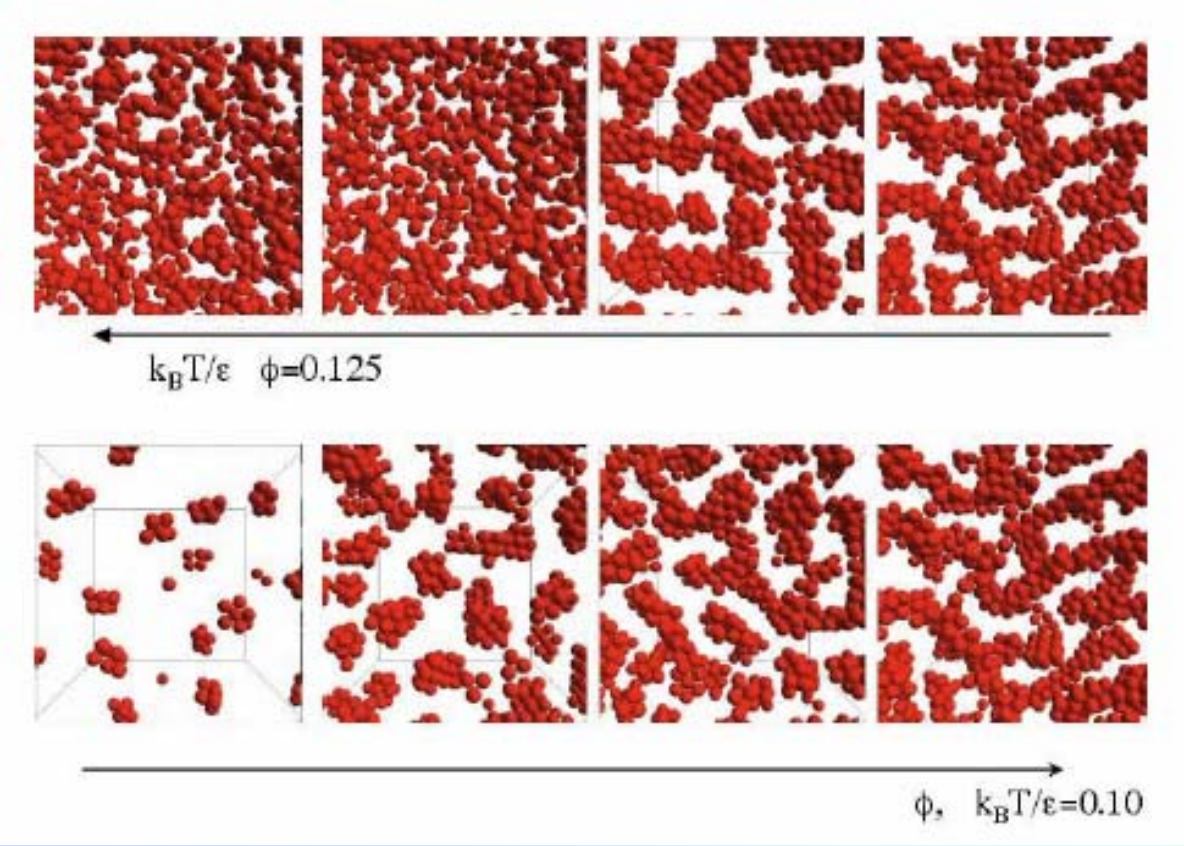
A. Stradner et al, Nature, **432**,492 (2004)



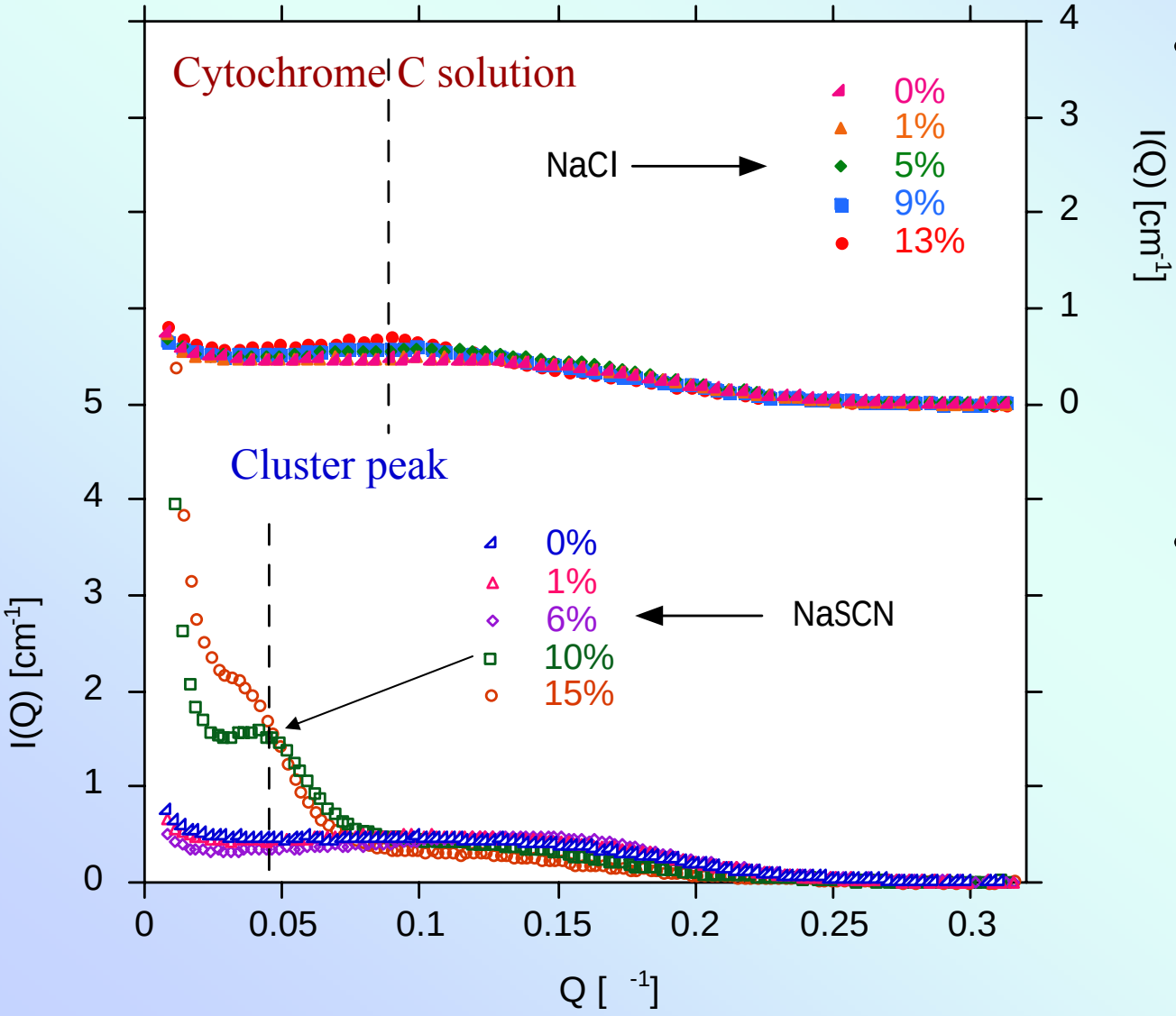

Sciortino et al, Phys. Rev. Lett. **93**, 055701, (2004)



**A clustering phenomenon can be induced by the competition of attraction (which favors cluster growth at low T ) and long range repulsive interaction (which favors low local particle densities and, therefore, small aggregates). With the appropriate balance of the attractive and repulsive potential, an optimal cluster size results from such a competition.**

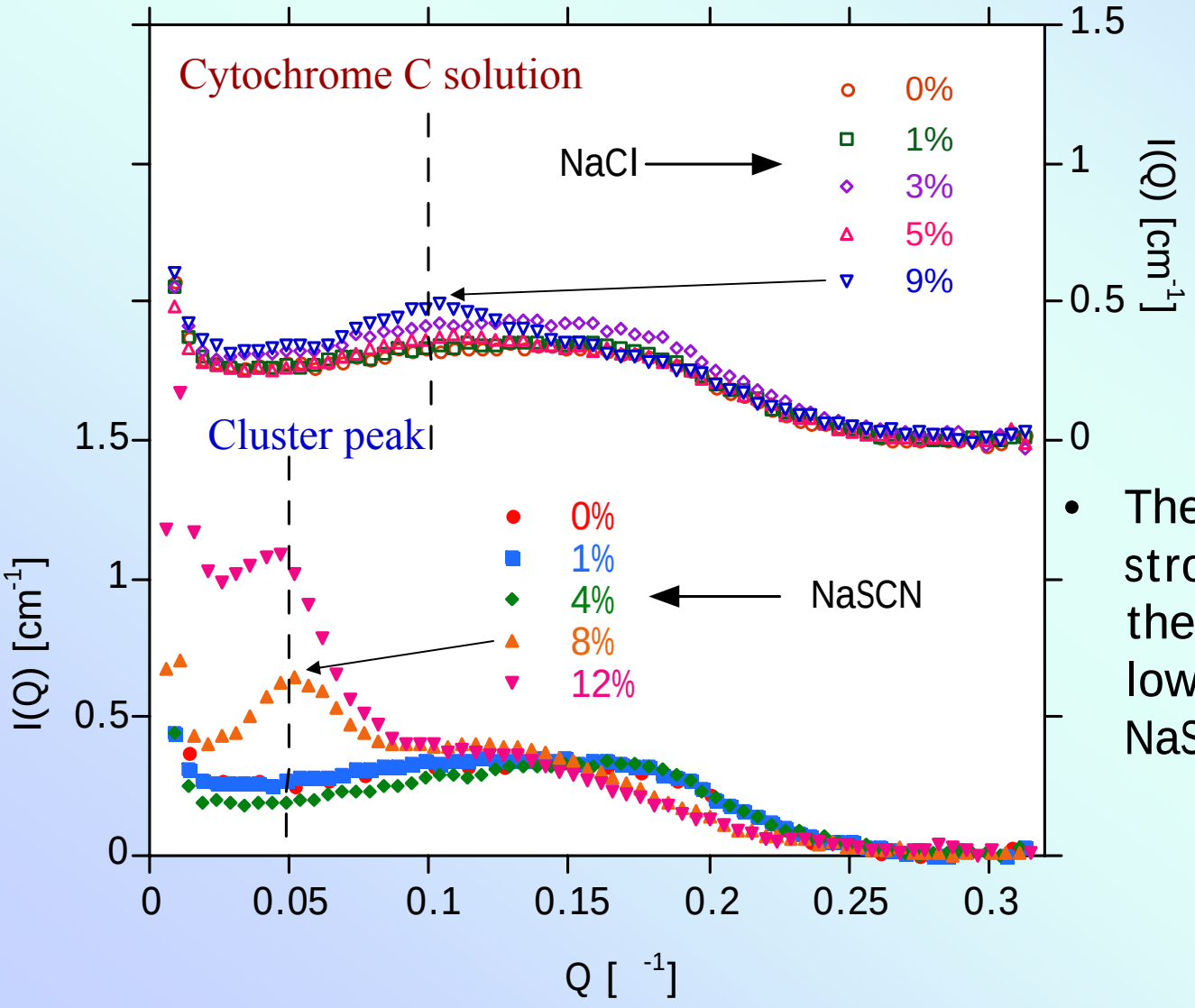


# Salt effect at $\phi = 0.3$



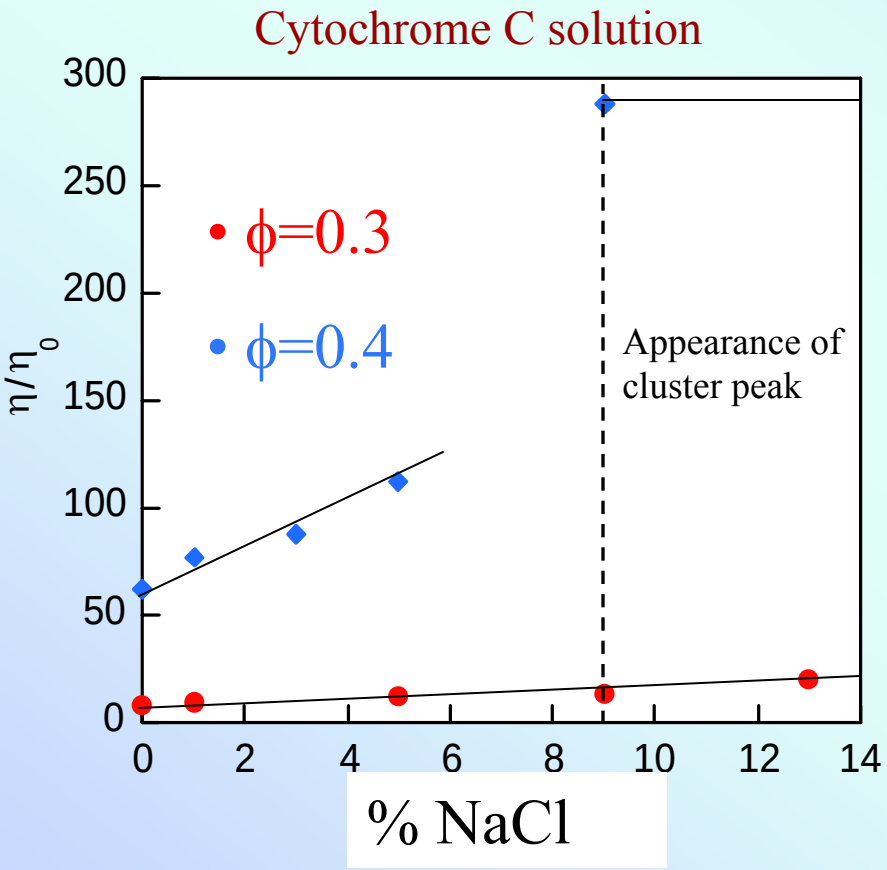

- A cluster peak appears at low Q (for NaCl at around 0.1 $Å^{-1}$, and for NaSCN at around 0.05 $Å^{-1}$).
- NaSCN has a stronger effect on cluster peak and shows a growth of intensity at very low Q for high salt concentrations.

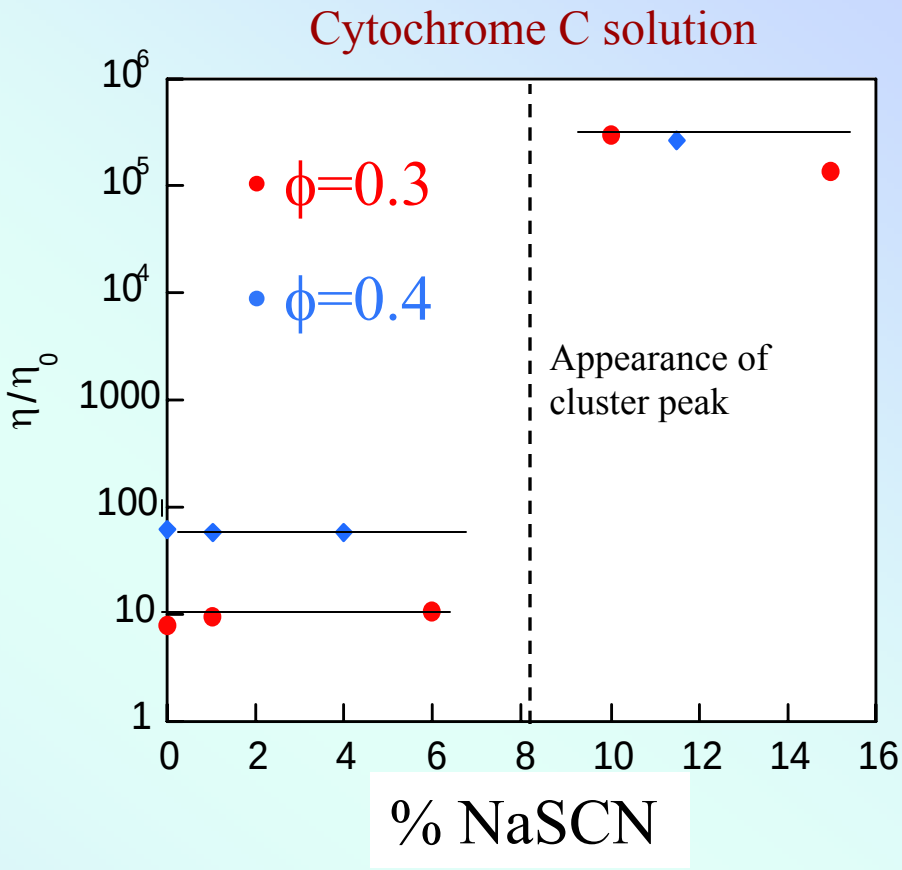
# Salt effect at $\phi = 0.4$




- The salt effect is stronger at $\phi$ = 0.3 and the peaks appear at a lower Q position for NaSCN.

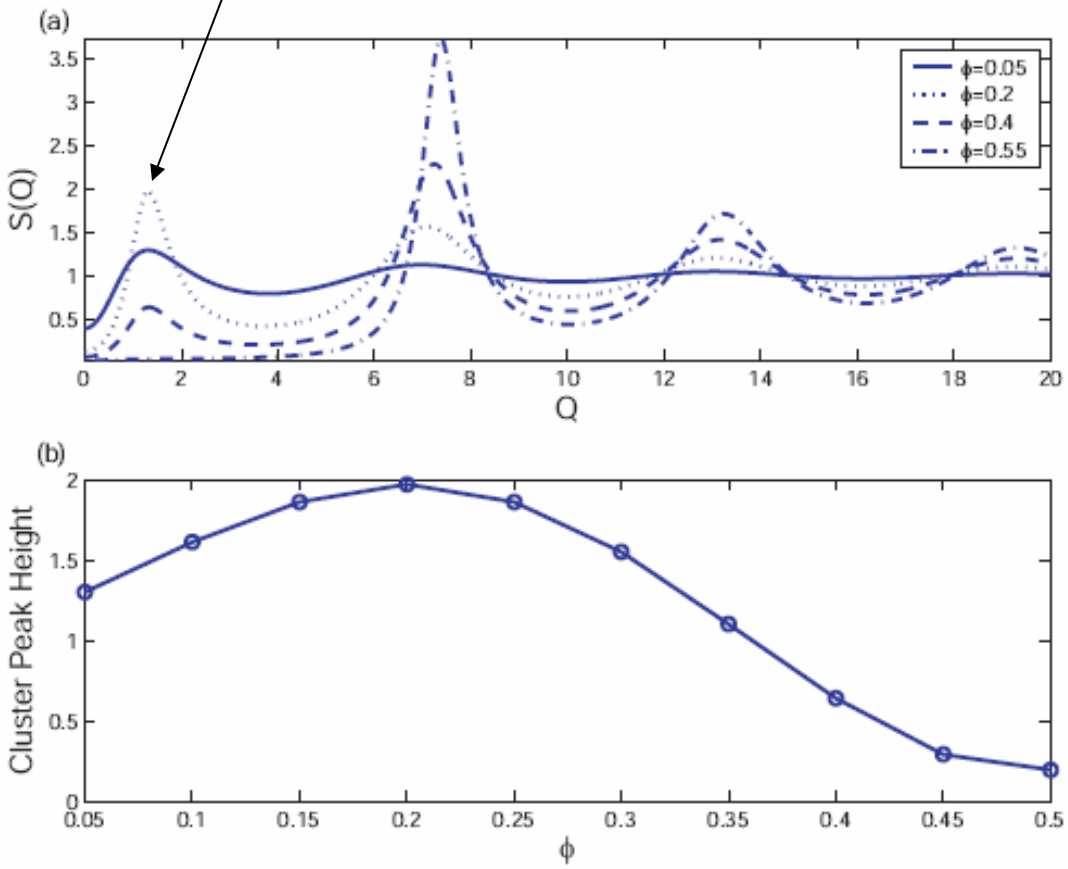
# Appearance of cluster peak on rheological properties



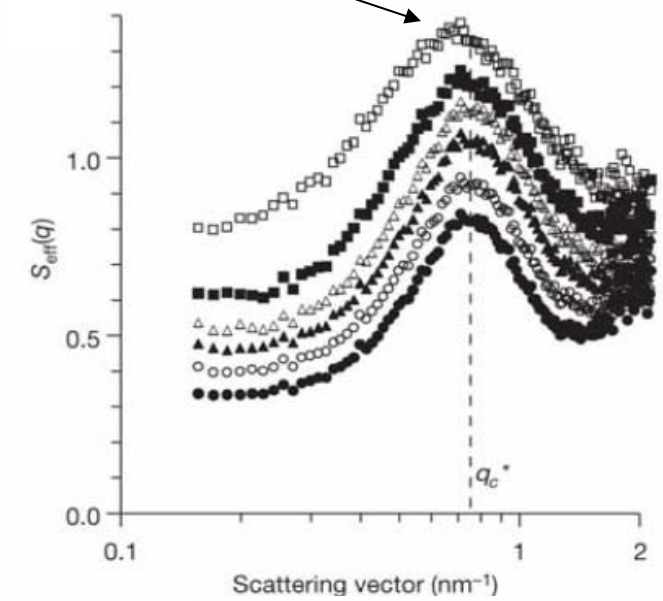



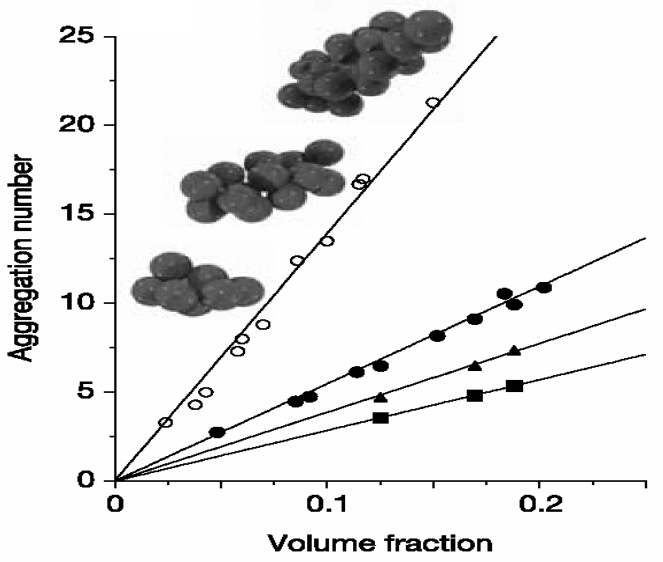
# Cluster aggregation number is proportional to ϕ

Cluster peak as function of ϕ




Cluster peak as function of c



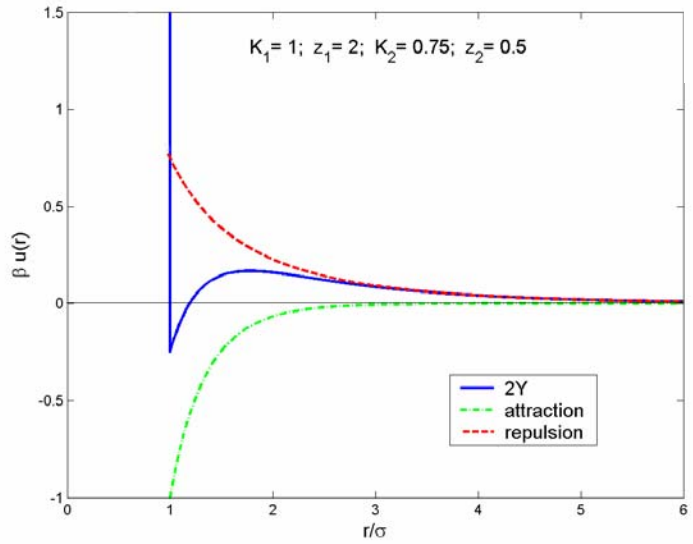

The cluster peak position is independent of volume fraction in a wide range of volume fraction. ($K_1$=10,$Z_1$=10,$K_2$=−1,$Z_2$=0.5 for the left figure)

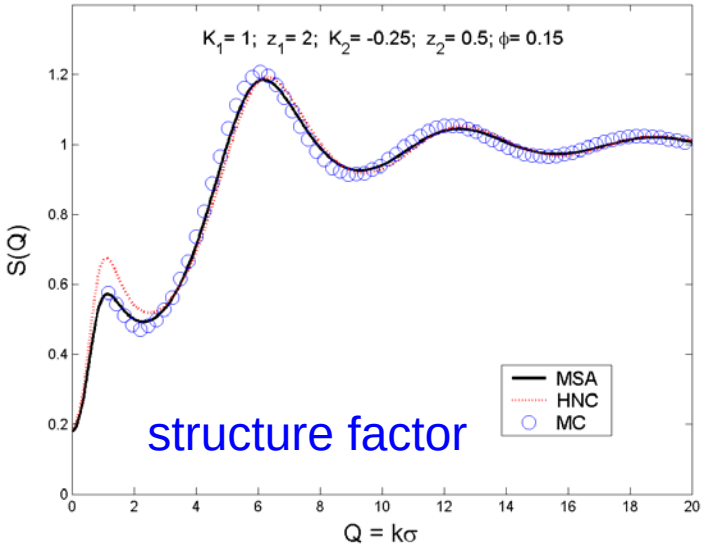
*Y. Liu, W. R. Chen, S. H. Chen, ICP, 122, 044507 (2005)*

# Origin of the cluster peak in two-Yukawa fluid

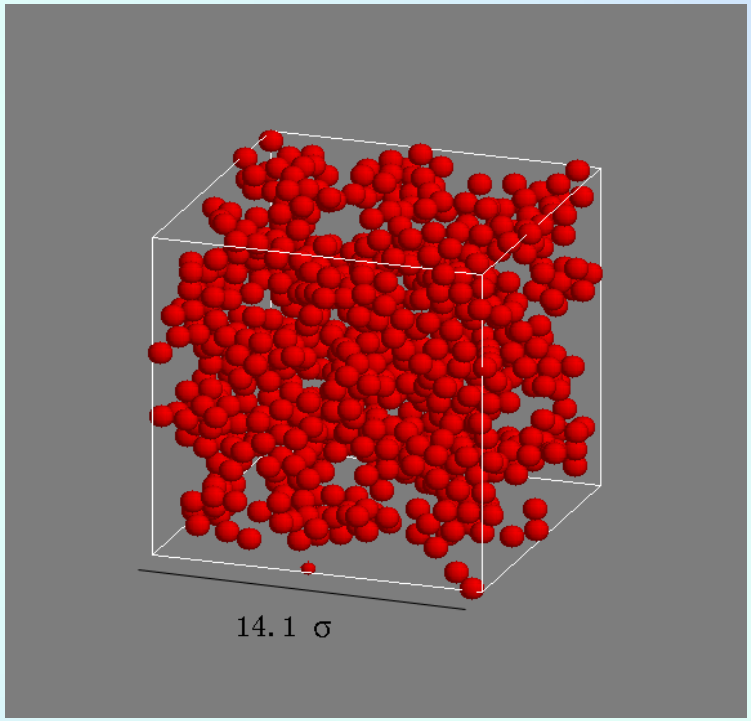
Effective potential

$$\beta u(r) = \begin{cases} \infty, & 0 < r \leq 1 \\ -K_1 \frac{e^{-z_1(r-1)}}{r} + K_2 \frac{e^{-z_2(r-1)}}{r} & r > 1 \end{cases}$$



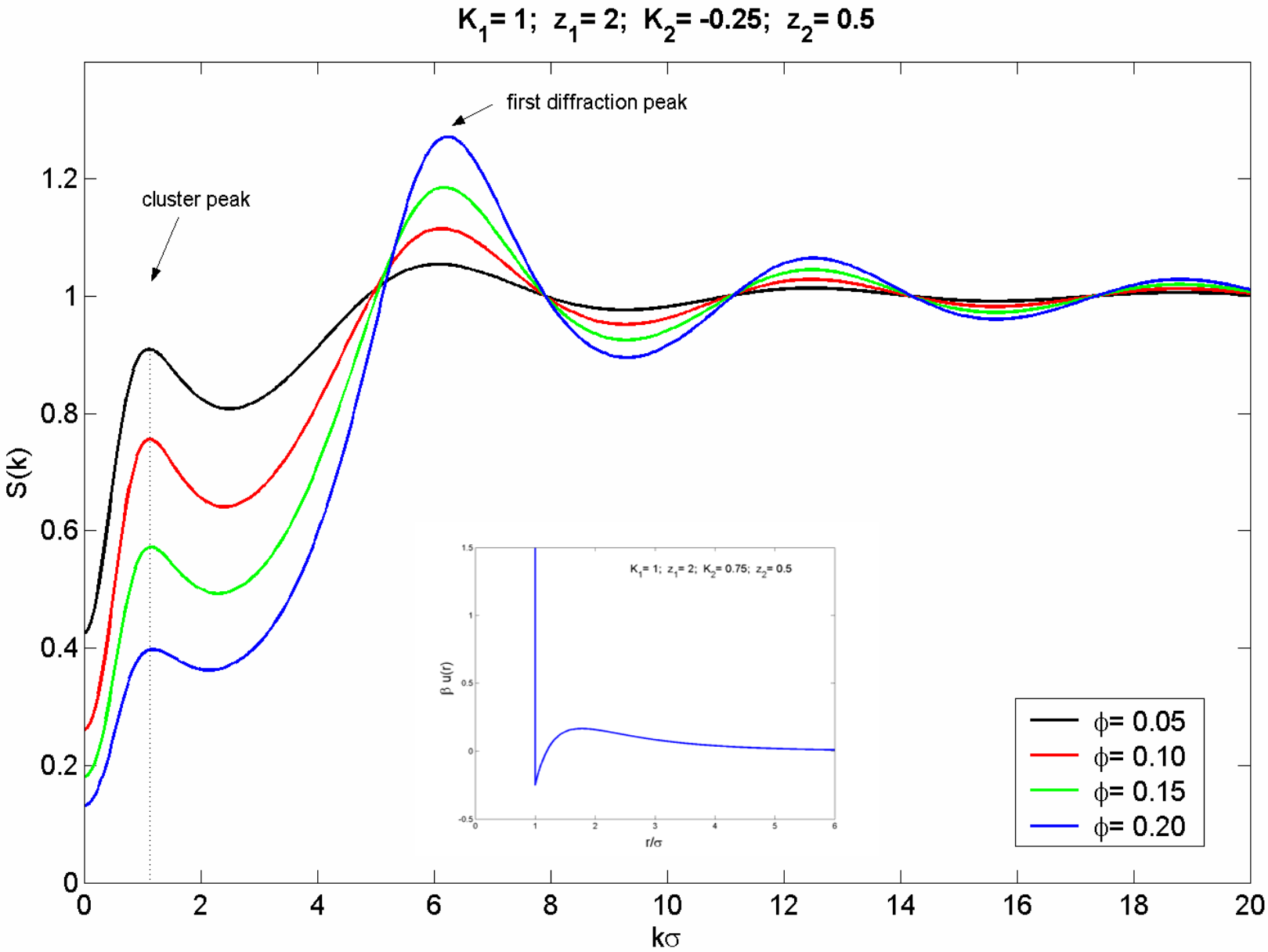




structure factor

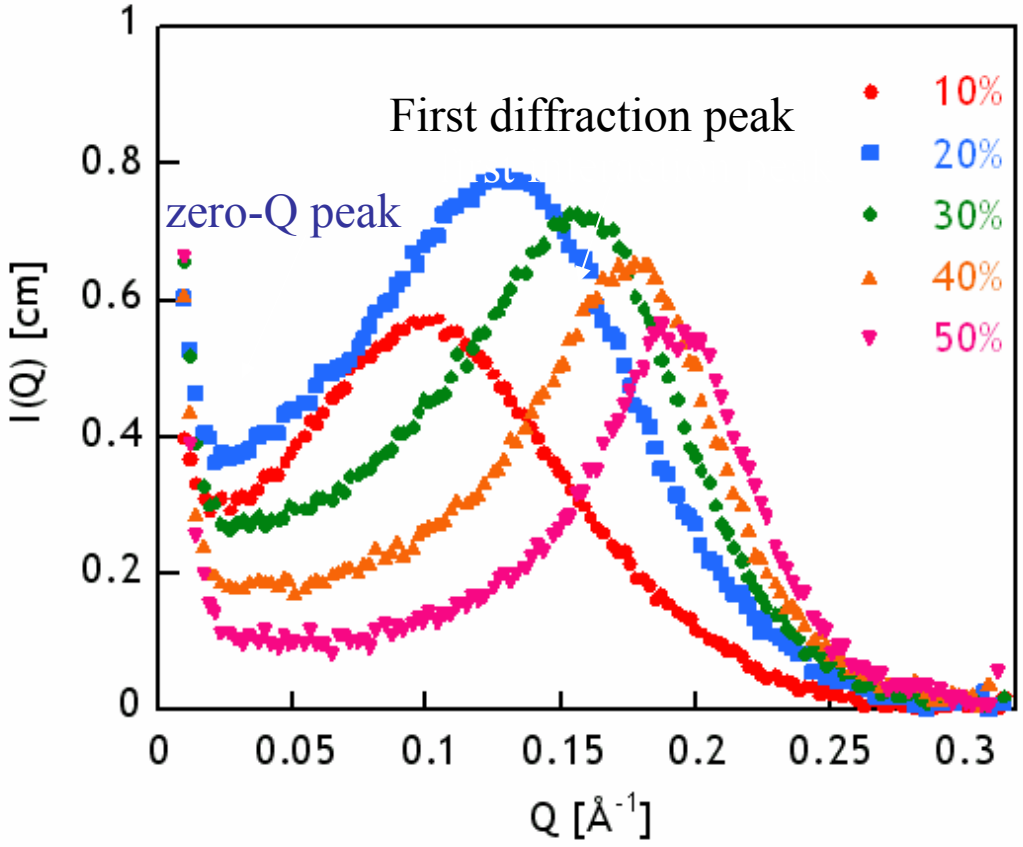



Snapshot from Monte Carlo simulation after $10^6$ steps, using the parameters set: { $K_1$= 1, $z_1$= 2, $K_2$= 0.25, $z_2$= 0.5 }, at 15% volume fraction.

# Effect of volume fraction on the cluster peak



# Cytochrome C at pD=5.4




pI ~ 10.2
Positively charged
As volume fraction increases the interaction peak moves to higher Q (i.e. molecules are forced to stay closer)
Spectra show a sharp rise at low Q (zero-Q peak).

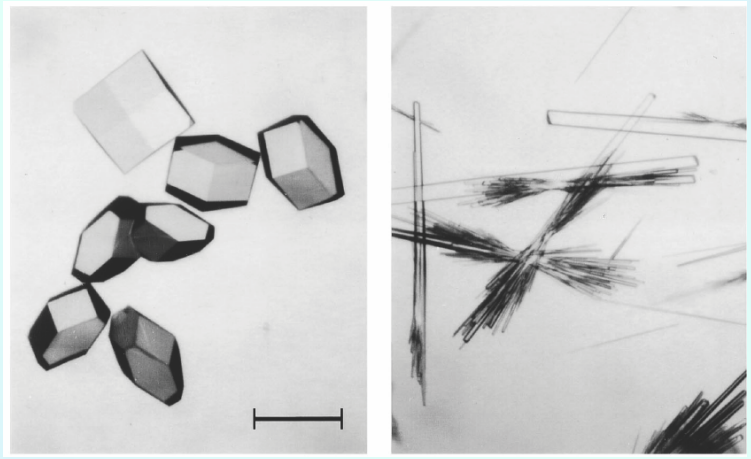
$$I(Q) = N_p(\rho_p - \rho_s)^2 V_p^2 \overline{P}(Q) S(Q)$$

Evidence for the existence of a long-range attraction

*B. Lonetti, P. Baglioni, S. H. Chen, et al., Phys. Chem. Chem. Phys.,* ***6****, 1388(2004)*

# Motivations

A general comprehension of the effective inter-protein potential and phase behavior is essential for developing a systematic method to grow protein crystals.



An example of two lysozyme crystals

*Picture taken from M. L. Broide, et al., Phys. Rev. E 53, 6325 (1996)*

"Protein interactions and association: an open challenge for colloid science"

*R. Piazza, Curr. Opin. Colloid Interface Sci.* ***8****, 515 (2004)*

# $B_{22}$ of protein solutions at crystallization conditions

*George A. et al., Acta Crystallogr. D* ***50****, 361 (1994)*

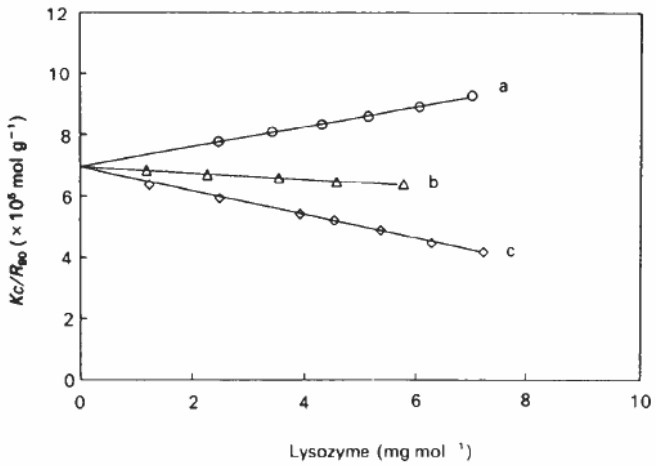



Fig. 3. SLS data for lysozyme in a (*a*) non-crystallizing solvent, (*b*) crystallizing solvent, (*c*) precipitating solvent. See text for solvent conditions.

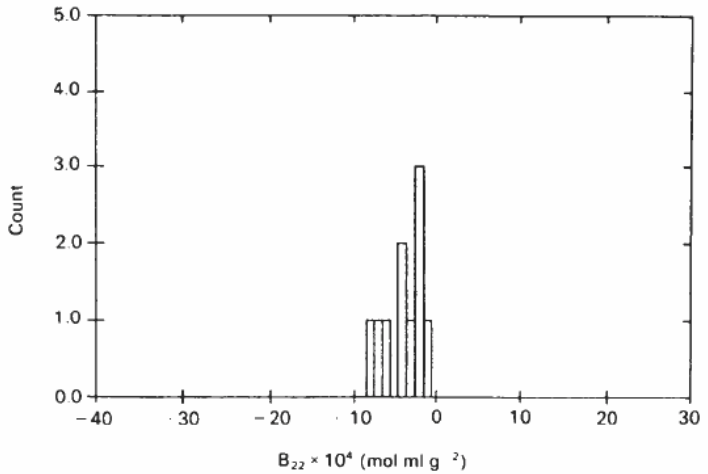
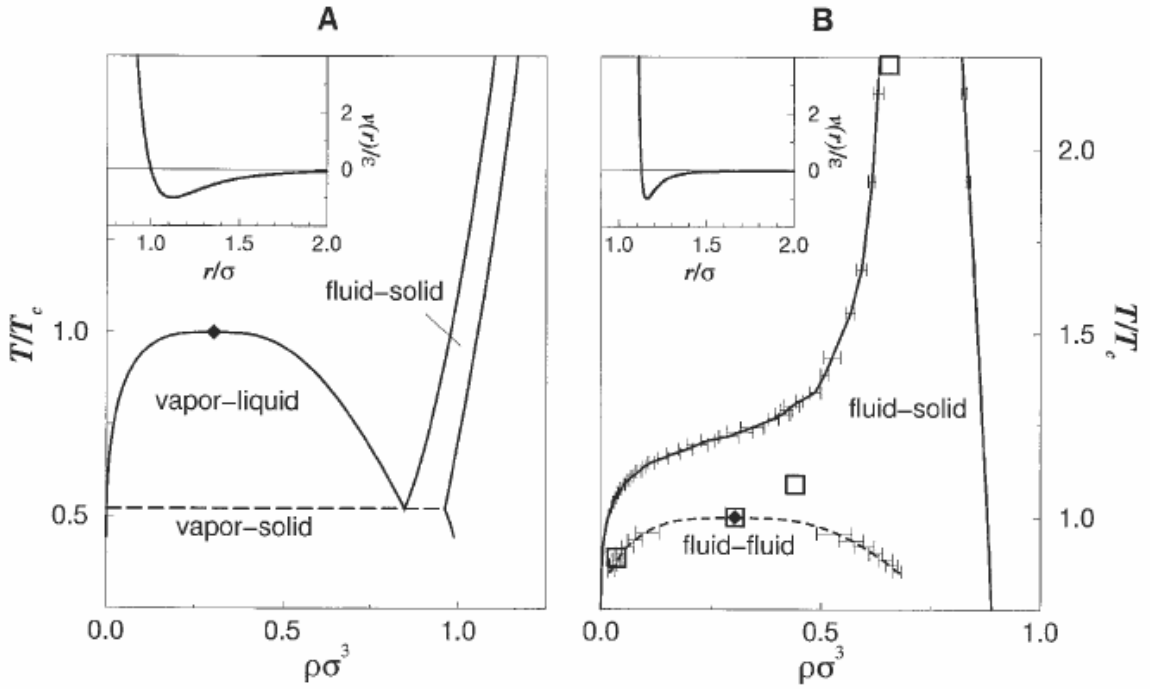



Fig. 2. Histogram representing crystallization slot obtained from second virial coefficient measurements on various proteins in crystallizing solvents.

Table 1. *Crystallization details*

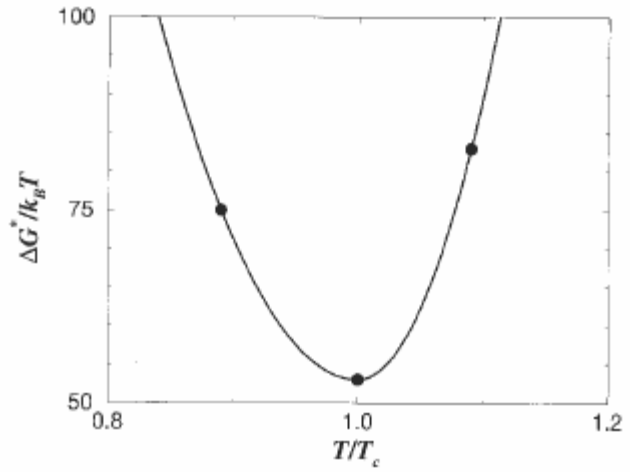
| Protein | $E^{1\%}_{280}$ | Crystallization conditions | $B_{22} \times 10^4$ (mol ml g$^{-2}$) | References |
|---|---|---|---|---|
| Lysozyme | 26.3 | 40 m*M* NaAc, pH = 4.6, 2% NaCl, 298 K | −3.0 | Mikol *et al.* (1990) |
| Canavalin | 6.8 | 50 m*M* Phosphate, pH = 7.0, 0.7% NaCl, 298 K | −0.8 | McPherson (1982) |
| Concanavalin A | 13.0 | 50 m*M* Tris–Ac, pH = 7.0, 1.0 *M* $(NH_4)_2SO_4$, 298 K | −2.5 | Mikol *et al.* (1990) |
| Concanavalin A | 13.0 | 10 m*M* Sodium cacodylate, pH = 6.0, 0.1 *M* NaCl, 298 K | −1.9 | Mikol *et al.* (1990) |
| BSA | 6.6 | 50 m*M* Potassium phosphate, pH = 6.2, 52% saturated $(NH_4)_2SO_4$, 298 K | −2.0 | Carter (1992) |
| Ovostatin | 7.98 | 0.1 *M* Imidazole, pH = 7.5, 7.5% PEG 8000, 293 K | −7.1 | Pusey (1992) |
| Ribonuclease A | 22.0 | 50% *n*-Propanol, pH = 5.0, 297 K | −4.1 | King *et al.* (1956) |
| α-Chymotrypsin | 20.0 | 0.1 *M* NaAC, pH = 4.6, 10% PEG 3350, 298 K | −8.4 | Gaier *et al.* (1981) |
| STMV | | 12.5% SAS, pH = 6.5, 298 K | −1.8 | Malkin & McPherson (1993*a*) |
| Ovalbumin | 26.9 | 50 m*M* Sodium cacodylate, pH = 5.4, 43% SAS, 2% methanol | -6.1 | Miller *et al.* (1983) |

# Enhancement of protein crystal nucleation by density fluctuation




(A) Typical phase diagram of a molecular substance with a long-range attraction.

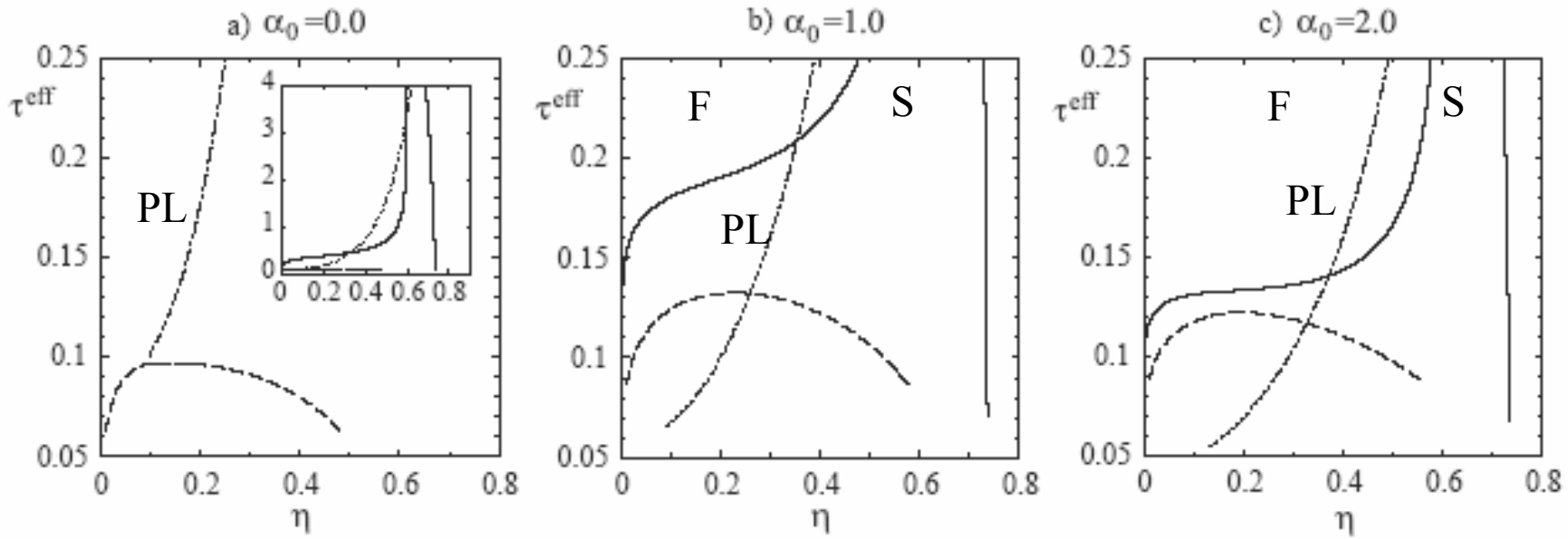
(B) Typical phase diagram of colloids with short-range attraction.




From the simulation of a globular proteins with short-range attraction, the crystal nucleation rate increases by many order of magnitude due to the presence of the metastable fluid-fluid critical point.

*Ten Wolde P. R., et al, Science* ***277****, 1975 (1997)*

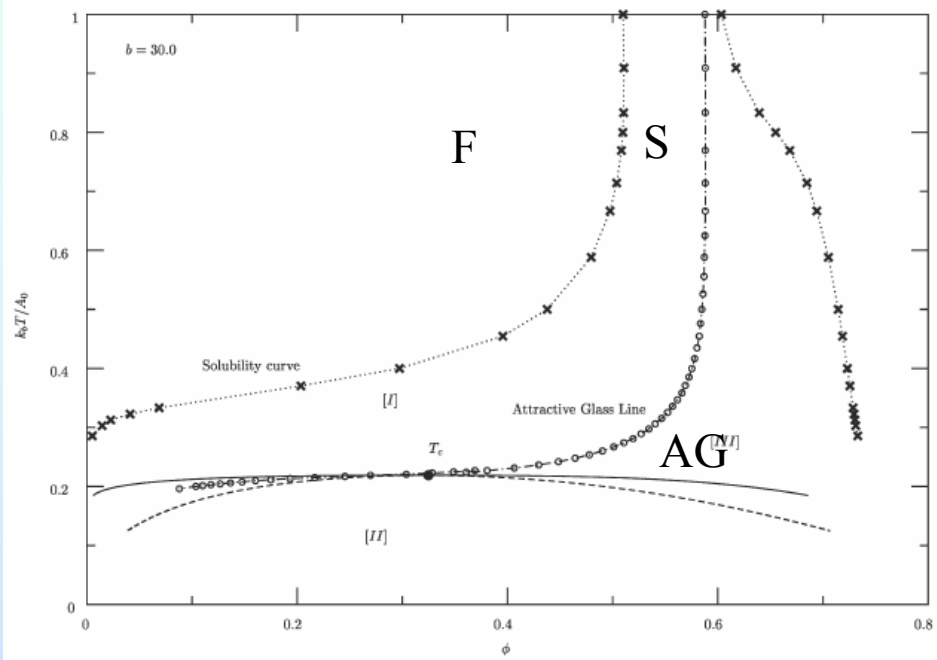
# Role of long-range interaction on the phase behavior of protein solution

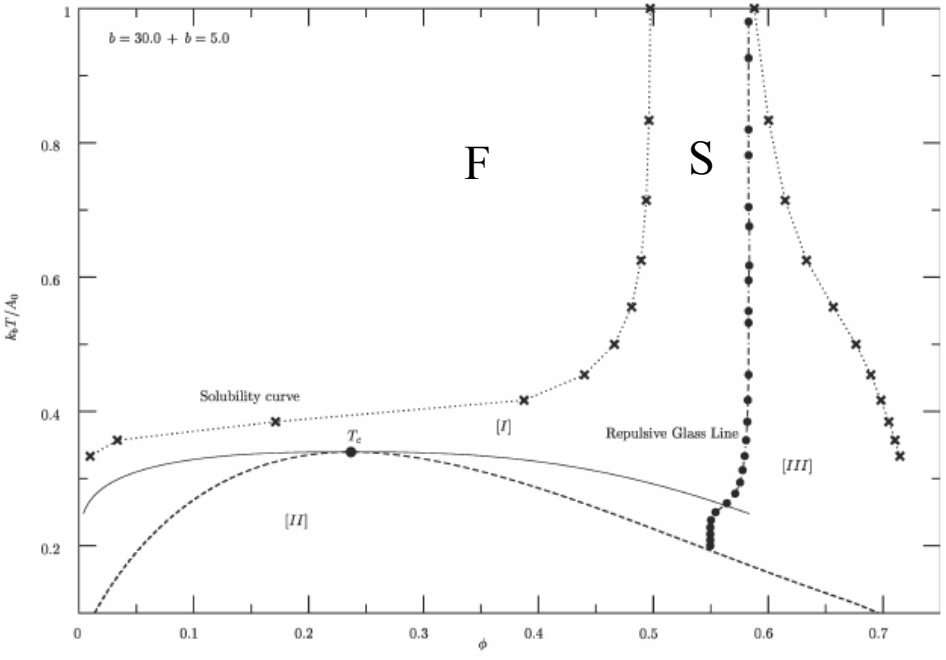



*M. G. Noro, et al., Europhys. Lett.* ***48****, 332 (1999)*

The addition of a long-range attraction can shift the percolation line below the meta-stable fluid-fluid transition line so that the protein crystallization can be favored instead of gel phase. Notice that with the increasing strength of long-range attraction, the feature of the phase of a short-range attraction system is still preserved.

# Role of long-range interaction on the phase behavior of protein solution




$$v(r)=\begin{cases}\infty & r<\sigma \\ -A_0\dfrac{e^{-b(r-\sigma)}}{r/\sigma} & r\geq\sigma\end{cases}$$

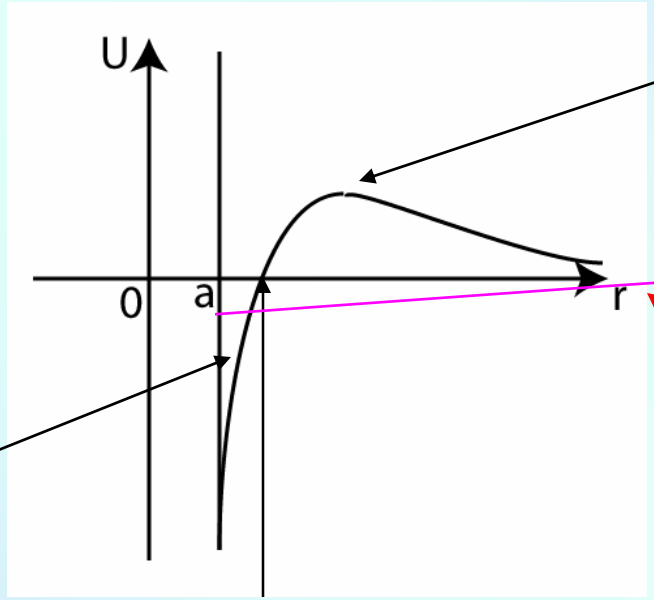
$$v(r)=\begin{cases}\infty & r<\sigma \\ -\dfrac{1}{A_1+A_2}\left[A_1\dfrac{e^{-b_1(r-\sigma)}}{r/\sigma}+A_2\dfrac{e^{-b_2(r-\sigma)}}{r/\sigma}\right] & r\geq\sigma\end{cases}$$

With only a short-range attraction, the fluid-fluid transition is preempted by a disordered solid state, attractive glass, in a wide range of volume fraction.

The addition of a long-range attraction suppress the attractive glass transition while keeping the equilibrium features of a short-range attraction system.

*A. Lawlor, et al., Progr. Colloid Polym. Sci.* ***123****, 104(2004)*

# Effective inter-protein potential in solution



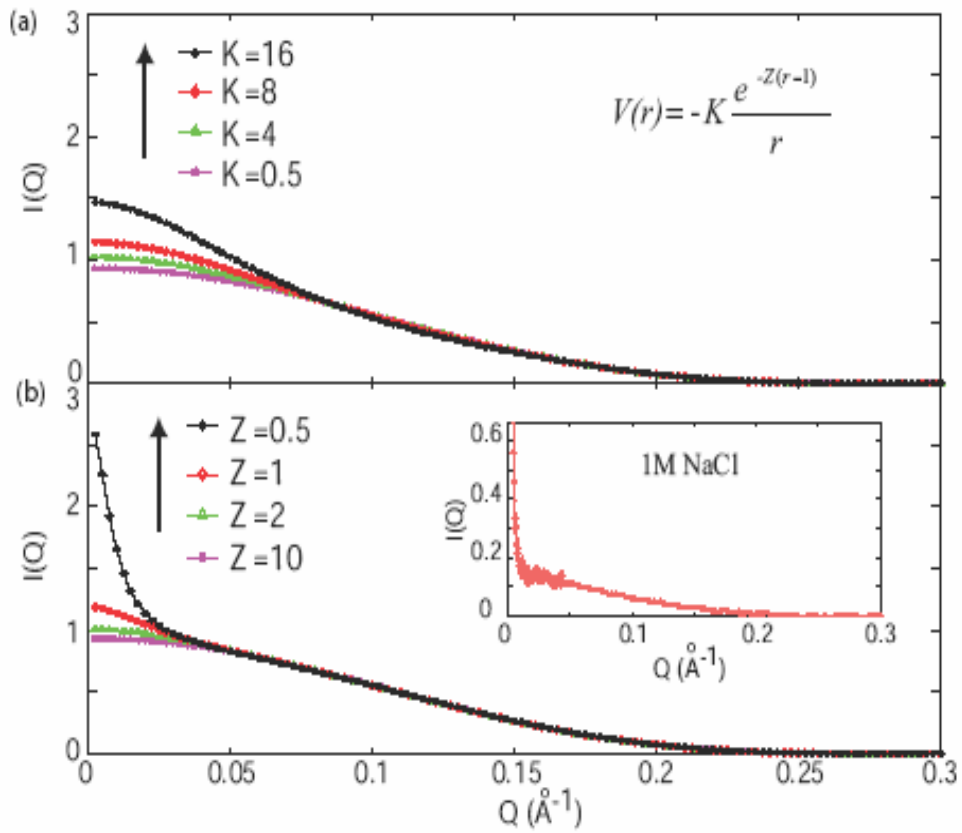

Electrostatic repulsion

Weak long-range attraction

Short-range attraction

Range: ~10%

# Cytochrome C at pD=11 with 1M NaCl




- Electrostatic repulsion is screened out.
- Only attraction needs to be considered.

- A sharp rising up at very low Q appears in the experimental data (zero-Q peak).
- A short-range attraction itself can not explain the zero-Q peak.

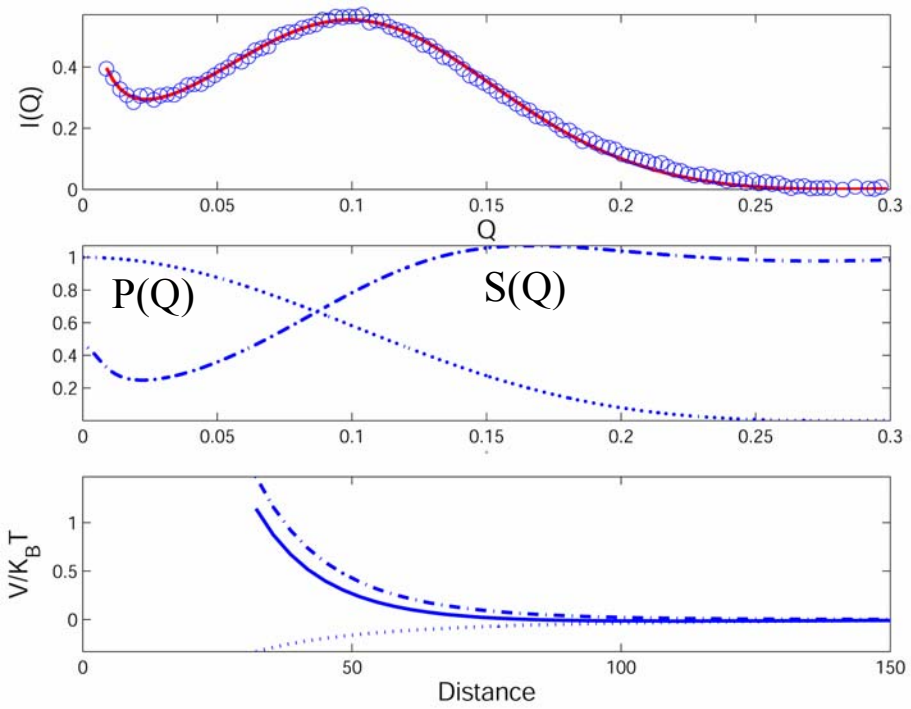
- In order to explain the zero-Q peak, we have to introduce a weak long-range attraction.

- The static light scattering technique is measuring the height of the zero-Q peak.

*Y. Liu, SH Chen, et. al. PRL 95, 118102 (2005)*

# Cytochrome C at pD=5.4

The zero-Q peak is induced by a long-range attraction




pD=5.4, $\phi$=10%

Fitted results

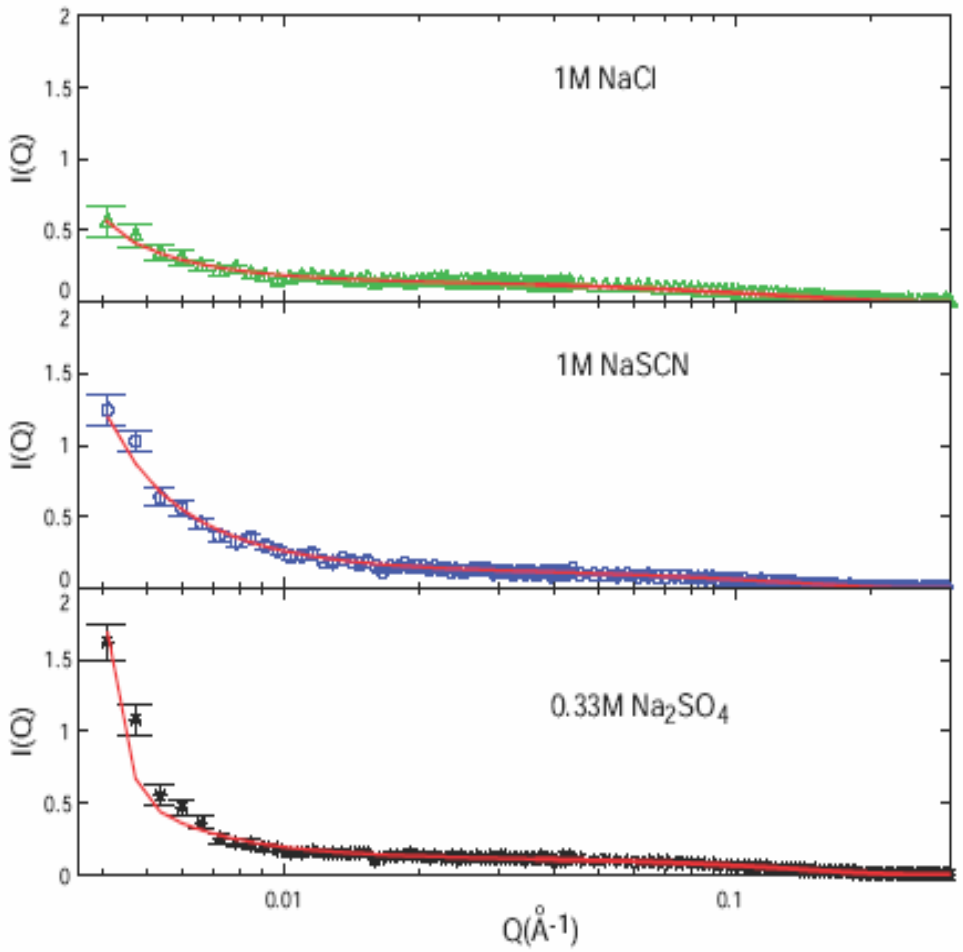
| Protein Concentration | $K_1$ | $Z_1$ | $z_p$ | a(Å) |
|---|---|---|---|---|
| 10.18 wt% | 0.33 | 0.51 | 4.2 | 14.8 |
| 20.40 wt% | 0.18 | 0.39 | 3.6 | 14.9 |

$$\beta V(r) = \begin{cases} \infty & \text{for } r < 1 \\ -K_1 \dfrac{e^{-Z_1(r-1)}}{r} - K_2 \dfrac{e^{-Z_2(r-1)}}{r} & \text{for } r > 1 \end{cases}$$

$z_p$: the charge number of a protein
a: the semi axis of a protein molecule. According to X-ray crystallography, a=15 Å.

*Y. Liu, SH Chen, et. al. PRL 95, 118102 (2005)*

# Cytochrome C at pD=11 with different salts




Fitted results

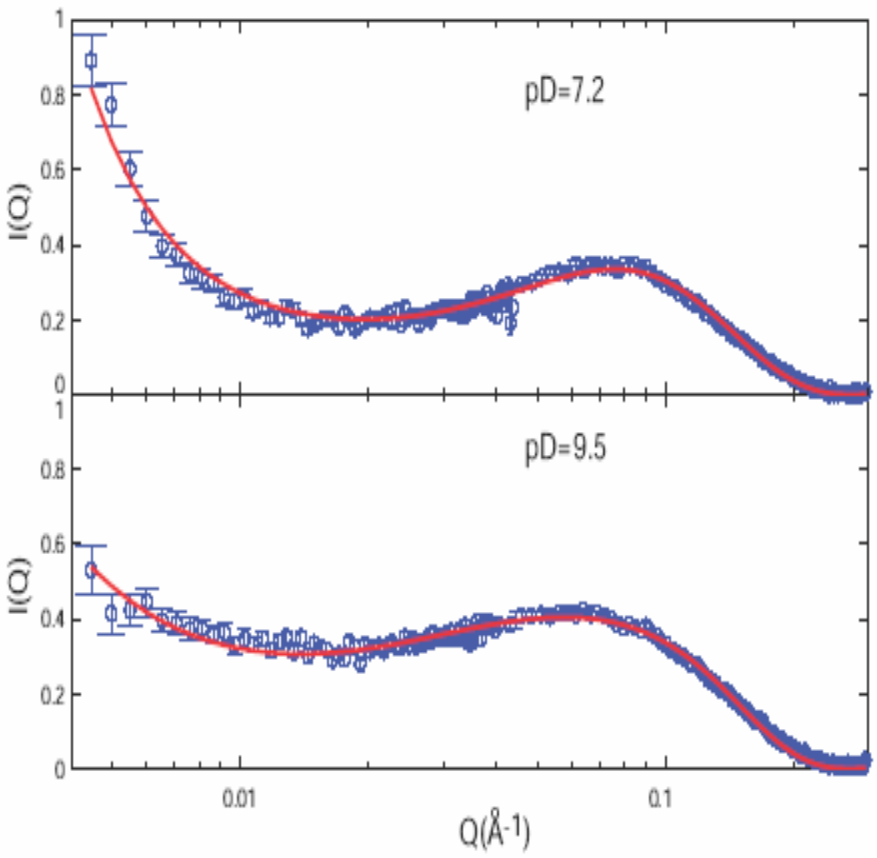
| | $a(\AA)$ | $K_1$ | $Z_1$ | $K_2$ | $Z_2$ |
|---|---|---|---|---|---|
| $NaCl$ | 14.7 | $7 \pm 3$ | $7 \pm 3$ | $0.11 \pm 0.02$ | $0.20 \pm 0.01$ |
| $NaSCN$ | 14.6 | $8 \pm 5$ | $10 \pm 6$ | $0.33 \pm 0.02$ | $0.35 \pm 0.01$ |
| $Na_2SO_4$ | 14.7 | $4 \pm 5$ | $10 \pm 10$ | $0.24 \pm 0.02$ | $0.27 \pm 0.02$ |

$$\beta V(r) = \begin{cases} \infty & \text{for } r < 1 \\ -K_1 \dfrac{e^{-Z_1(r-1)}}{r} - K_2 \dfrac{e^{-Z_2(r-1)}}{r} & \text{for } r > 1 \end{cases}$$

- zero-Q peak depends on different type of added salts.
- The range of the long range attraction has even longer range than the electrostatic repulsion.
- The range and the strength of the long range attraction is about the same value as the long range attraction between like-charge particles reported in literature.

*Y. Liu, SH Chen, et. al. PRL 95, 118102 (2005)*

# Cytochrome C at different pD values without salts




Fitted results

| | $a(\AA)$ | charge number | $K_2$ | $Z_2$ |
|---|---|---|---|---|
| $pD = 7.2$ | 15.3 | $3.8 \pm 0.1$ | $0.37 \pm 0.02$ | $0.34 \pm 0.01$ |
| $pD = 9.5$ | 15.4 | $1.7 \pm 0.1$ | $0.08 \pm 0.01$ | $0.21 \pm 0.01$ |

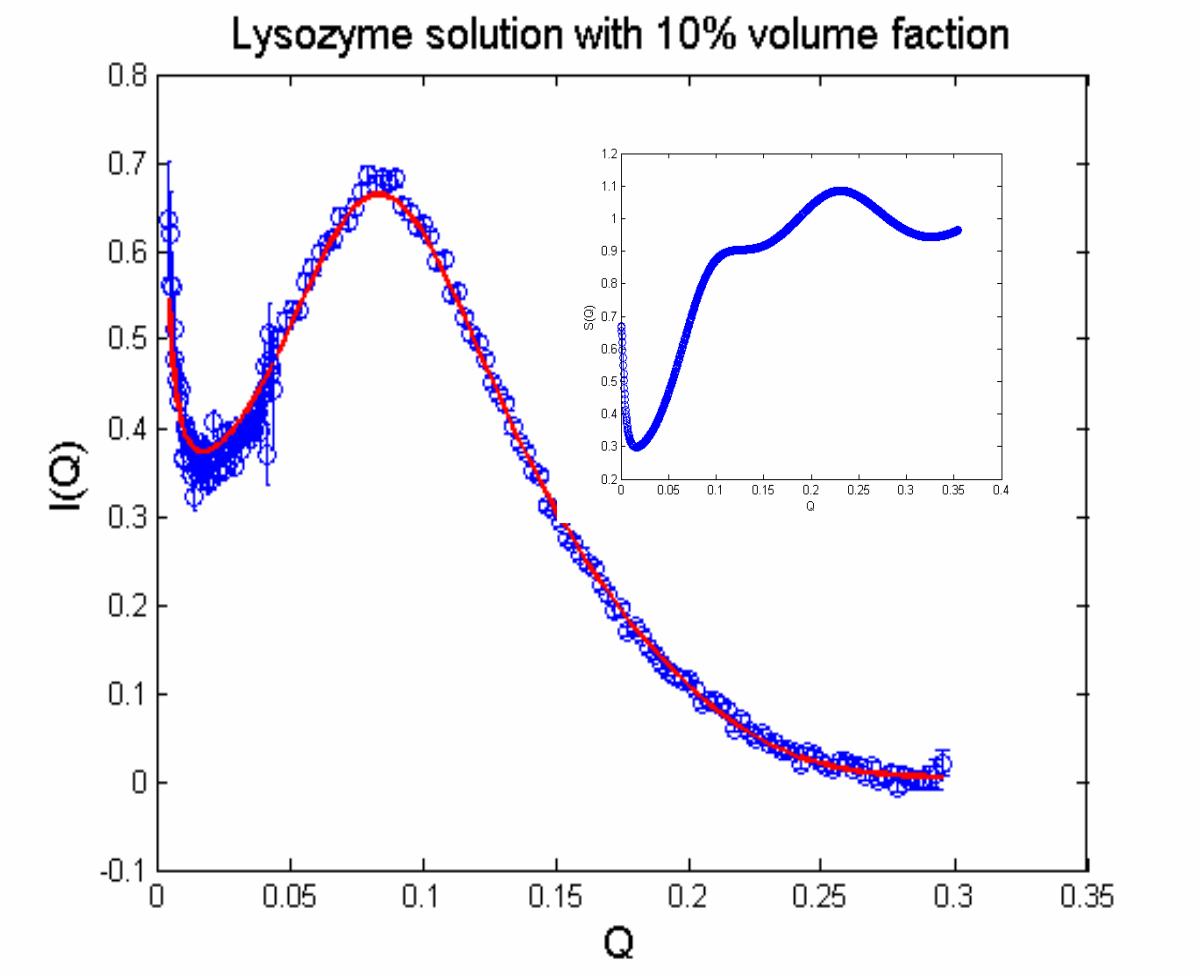
$$\beta V(r) = \begin{cases} \infty & \text{for } r < 1 \\ -K_1 \dfrac{e^{-Z_1(r-1)}}{r} - K_2 \dfrac{e^{-Z_2(r-1)}}{r} & \text{for } r > 1 \end{cases}$$

- Ignore the short-range attraction and only consider the long-range attraction and the electrostatic repulsion.
- The the isoelectric point, pI, of cytochrome C is about 10.
- At pD=9.5, the zero-Q peak is smaller showing a result of weaker attraction.

*Y. Liu, SH Chen, et. al. PRL 95, 118102 (2005)*

# Lysozyme protein solutions

## HNC Analysis of Three Yukawa Potential




Dimension of lysozyme:
a×b×b=22.5×15×15Å

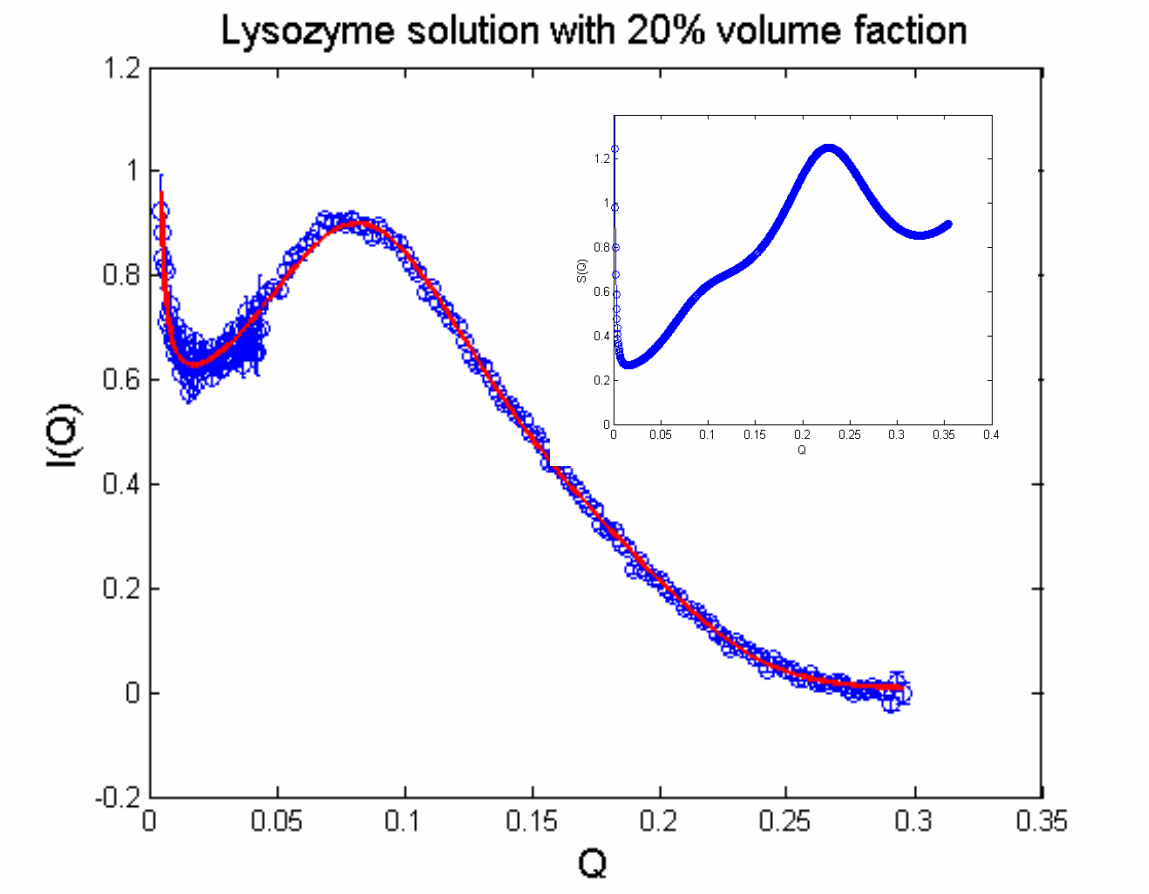
Charge number=7.1±1.3
a=21.1±0.1

Short-range attraction:
K1=4.3±0.6
Z1=14±3

Long-range attraction:
K2=0.03±0.006
Z2=0.2±0.04

# Lysozyme protein solutions

## HNC Analysis of Three Yukawa Potential




Dimension of lysozyme:
a×b×b=22.5×15×15Å

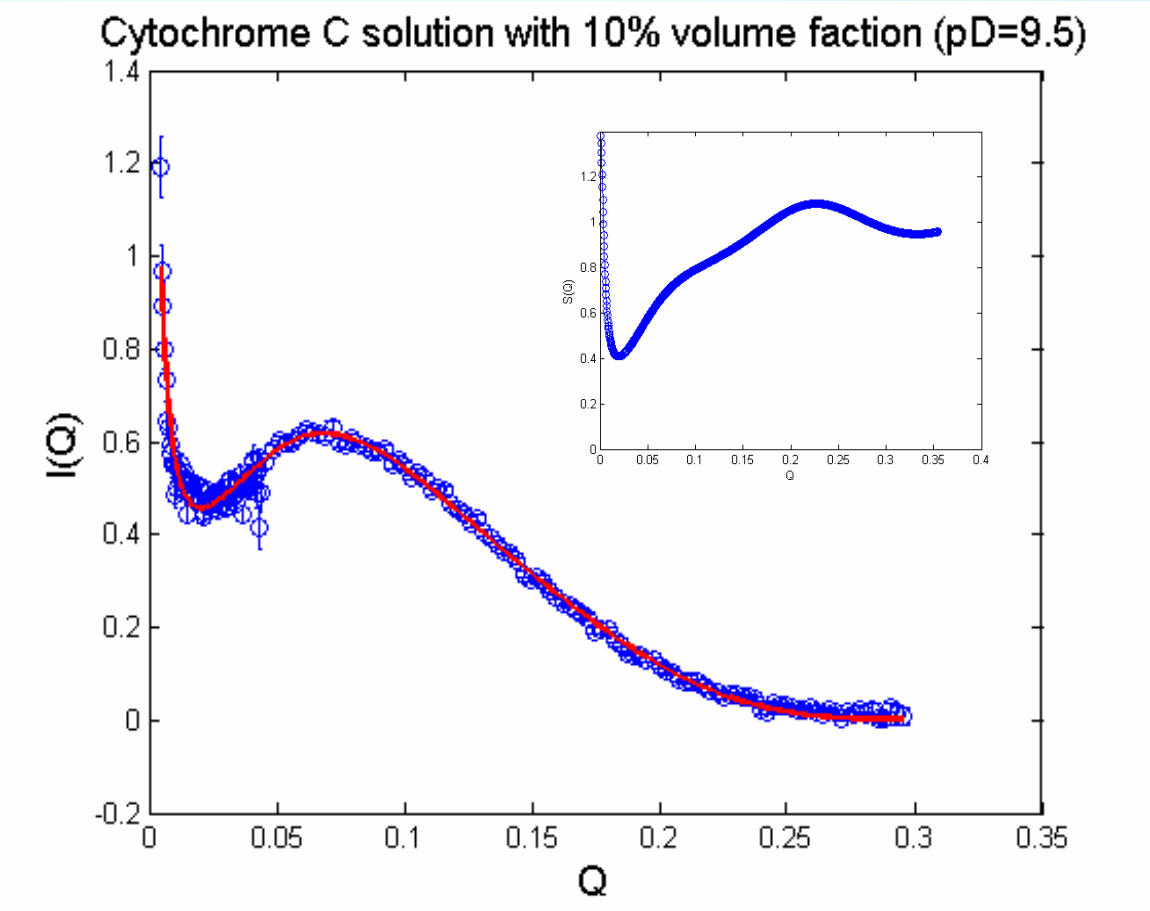
Charge number=6.7±0.5
a=21.1±0.1

Short-range attraction:
K1=3.9±0.06
Z1=13±2

Long-range attraction:
K2=0.01±0.0005
Z2=0.12±0.007

# Cytochrome C protein solutions

## HNC Analysis of Three Yukawa Potential




Dimension of cytochrome C:
a×b×b=15×17×17Å

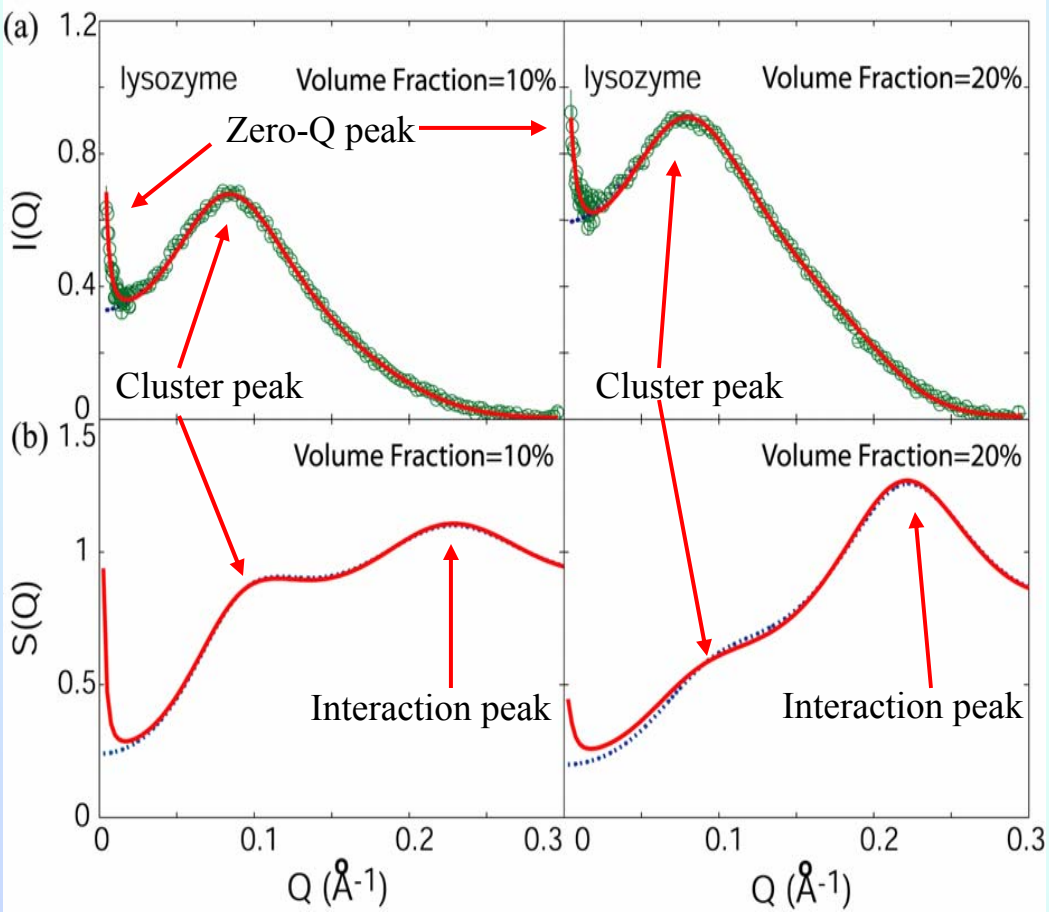
Charge number=2.3±0.1
a=14.2±0.06

Short-range attraction:
K1=2.77±0.7
Z1=27.4±19

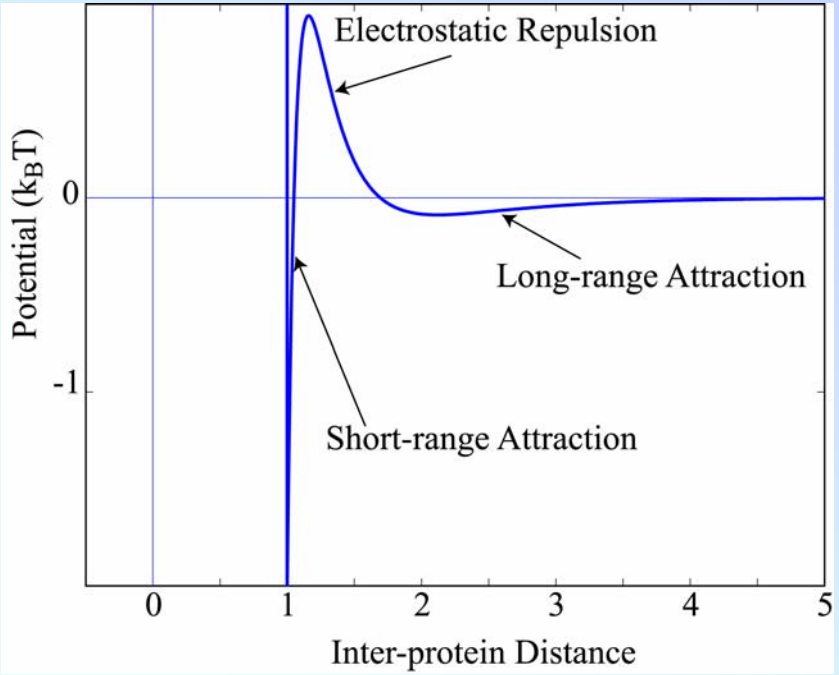
Long-range attraction:
K2=0.08±0.007
Z2=0.31±0.02

Note: For short-range attraction, the individual fitting gives K1~2.7.
If simultaneously fit both experimental results at Vol=10% and Vol=15%,
the short-range attraction becomes about 1.7.

SANS Intensity from Lysozyme Protein Solutions




Schematic Inter-protein Potential




*Y. Liu, SH Chen, et. al. PRL 95, 118102 (2005)*

Careful SANS measurements on lysozyme and cytochome C solutions, by Chen's Group at MIT, find that the effective inter-protein potential consists of three features, a short-range attraction, an intermediate-range repulsion, and a weak long-range attraction. The latter interaction is important for protein crystallization.

# Conclusions

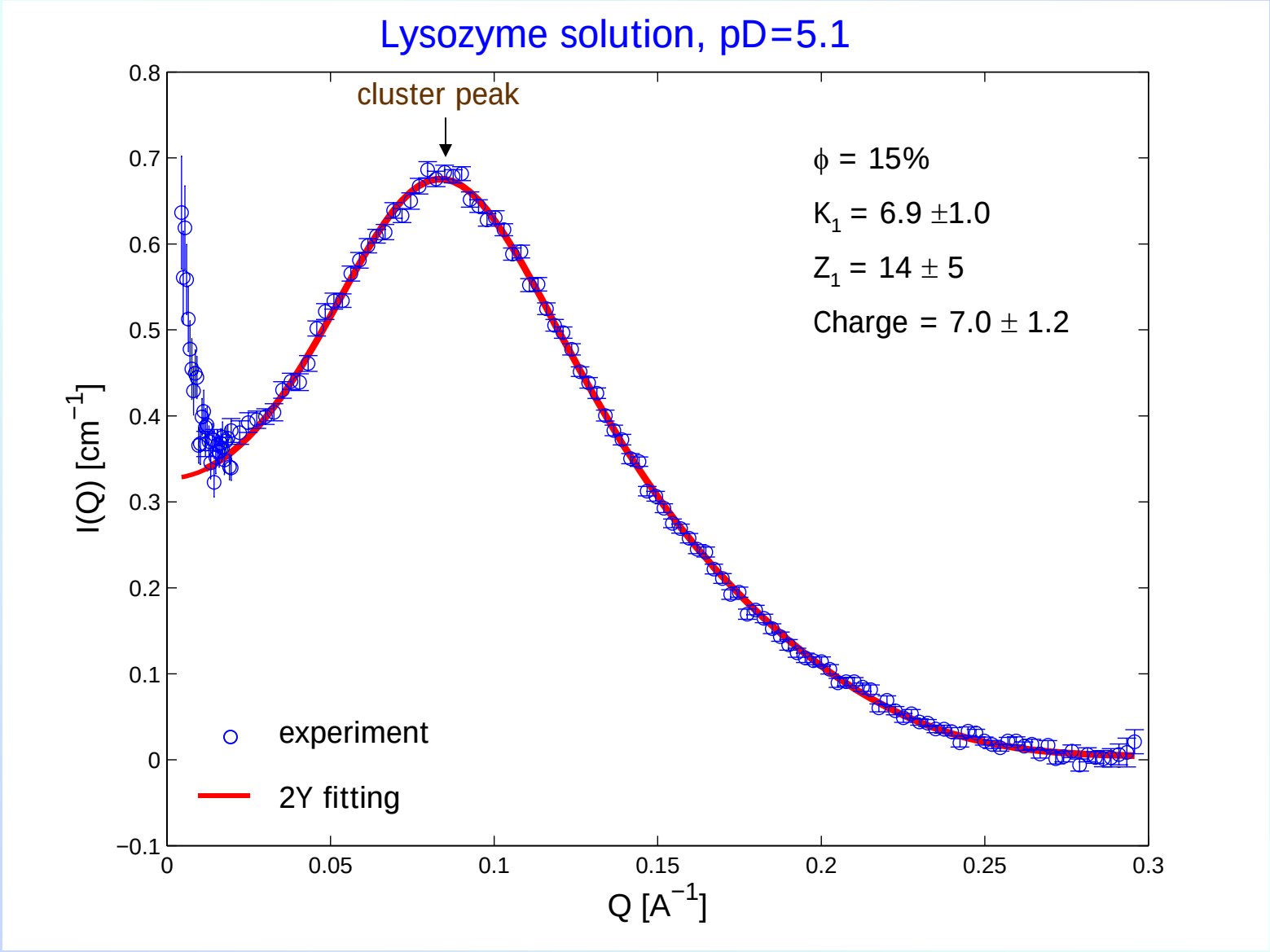
- The appearance of the zero-Q peak in our protein solutions have to be interpreted as due to an existence of a weak long-range attraction.

- The existence of zero-Q peak in both Cytochrome C and Lysozyme solutions implies that the zero-Q peak, i.e., the long-range attraction, is universal for all proteins.

- This long-range attraction is closely related to the type of ion cloud around a protein molecule.

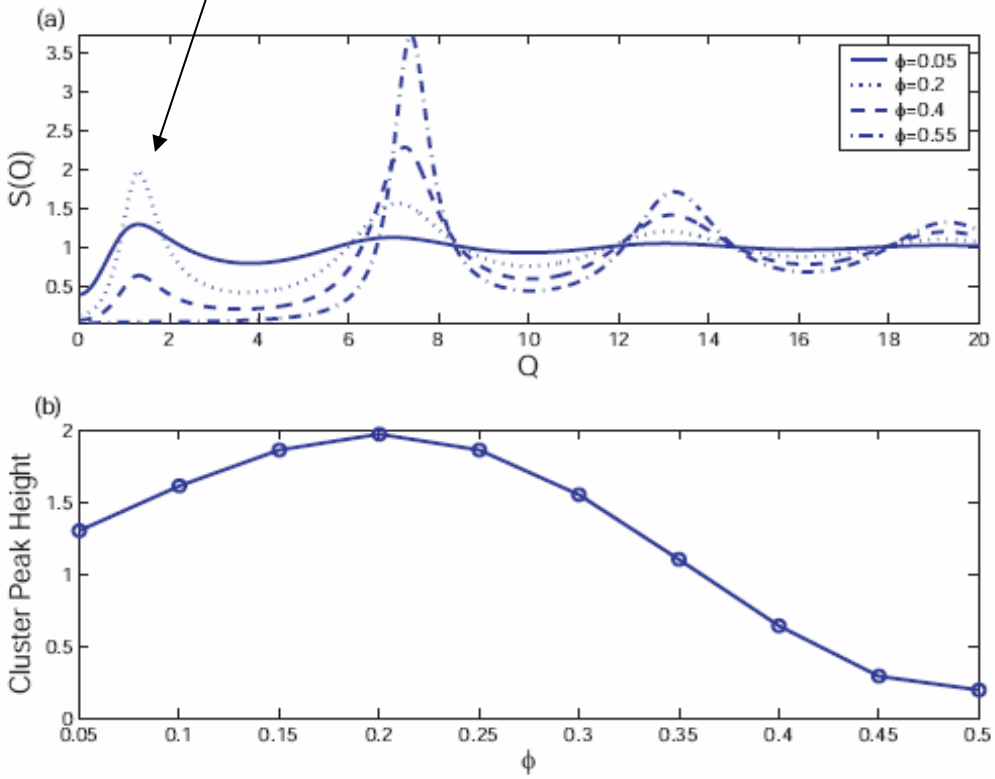
- Since the static light scattering is measuring the height of the zero-Q peak due to the very small Q value in the light scattering, the interpretation of the measured results need to consider effects of the long-range attraction.

# Application of 2Y model to SANS data analysis

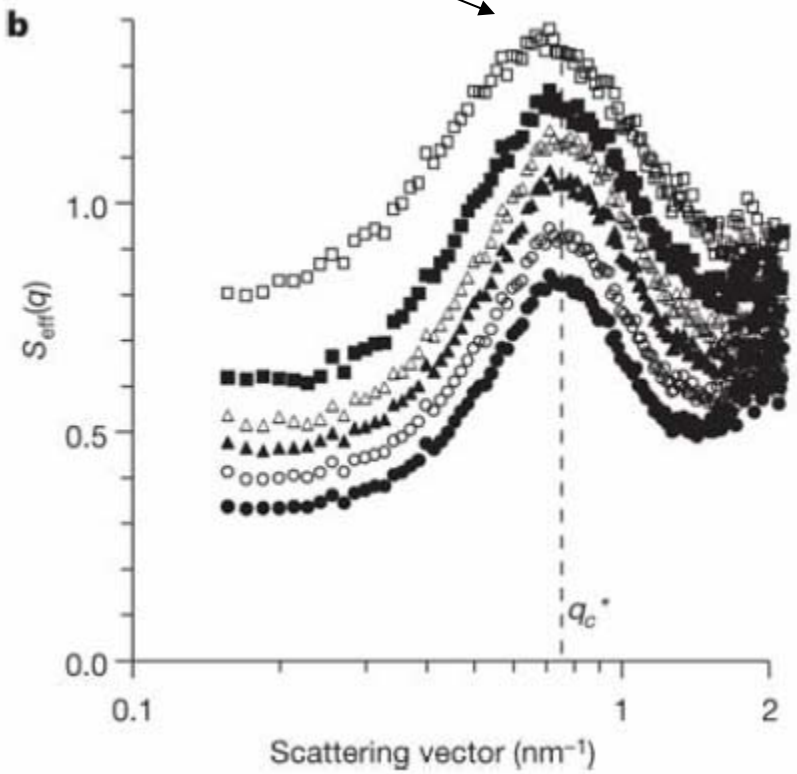


# Cluster aggregation number is proportional to ϕ

Cluster peak as function of ϕ




Cluster peak as function of c

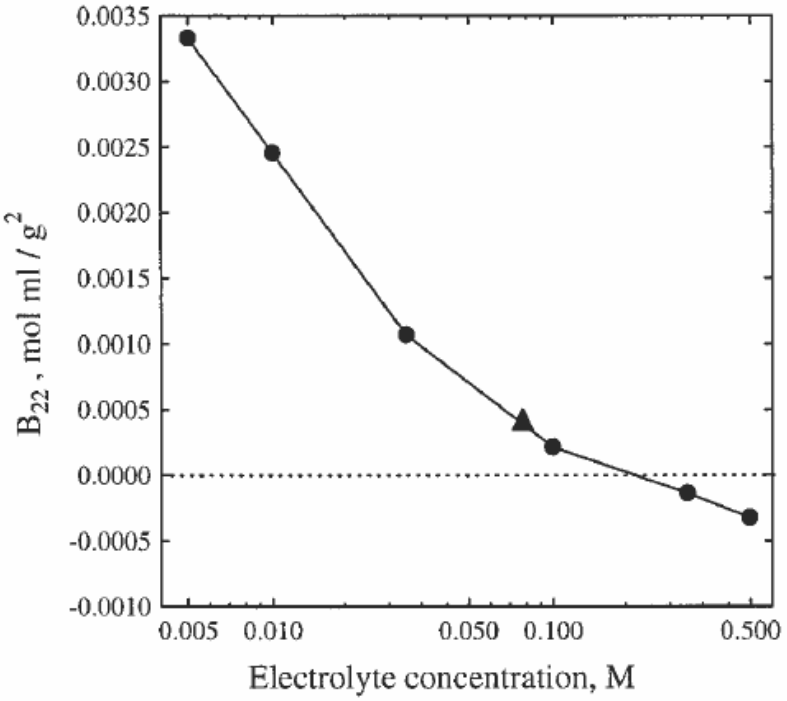



The cluster peak position is independent of volume fraction in a wide range of volume fraction. ($K_1$=10,$Z_1$=10,$K_2$=−1,$Z_2$=0.5 for the left figure)

*Y. Liu, W. R. Chen, S. H. Chen, Journal of Chemical Physics, 122, 044507 (2005)*

# Evidence of Non-DLVO potential in protein solutions

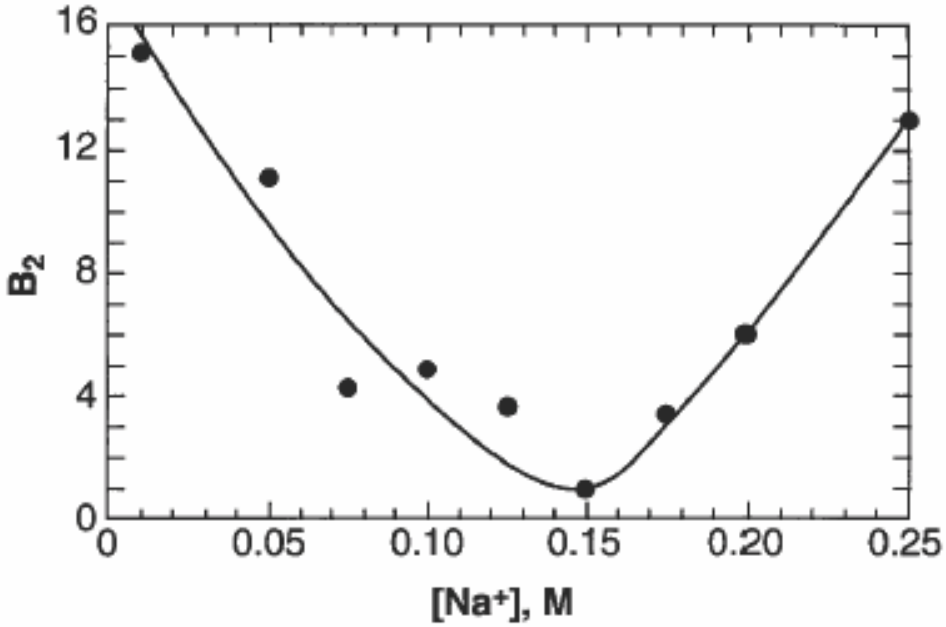
DLVO-like potential




The second virial coefficient of Lysozyme measured at pH=4.5 as a function of electrolyte concentrations.

*O. D. Velev, et al., Biophysical J.* ***75****, 2682 (1998)*

Non-DLVO-like potential




Dependence of the second virial coefficients of apoferritin protein solutions on ion concentration .

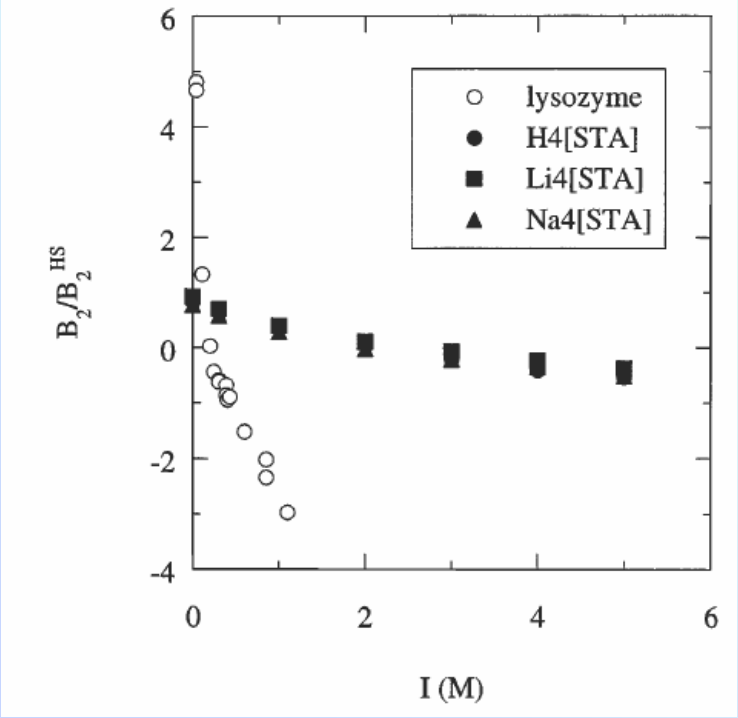
*D. N. Petsev, et al., Phys. Rev. Lett.* ***84****, 1339 (2000)*

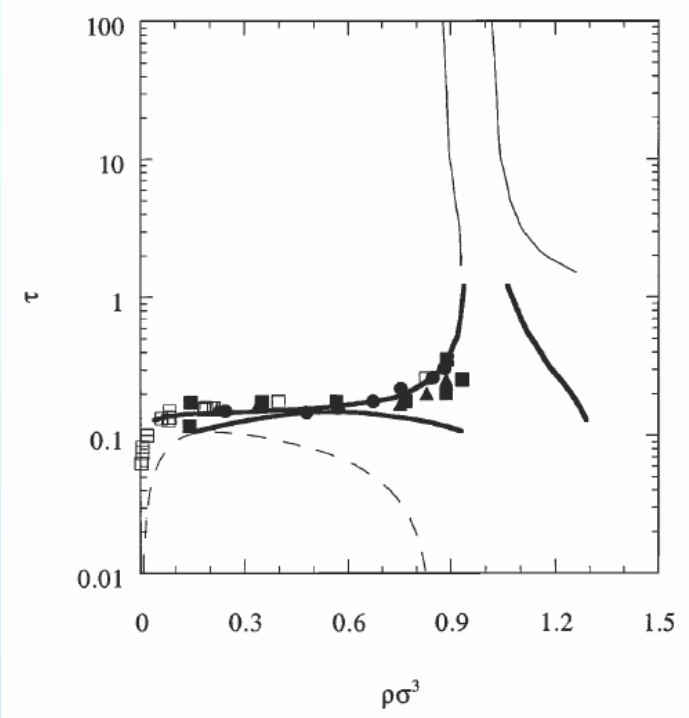
# Phase behaviors of protein solutions

Baxter's adhesive hard sphere model:

$$\frac{V(r)}{kT} = \begin{cases} \infty, & r < \sigma, \\ \ln[12\tau\delta/(\sigma + \delta)], & \sigma \le r \le \sigma + \delta, \\ 0, & r > \sigma, \end{cases}$$

The crystallization boundaries of small, attractive particles are found to be identical when Interactions are modeled as adhesive hard spheres.




Solid symbols:
Silicon tungstate anion solutions

Open symbols:
Lysozyme solutions

$$\frac{B_2}{B_2^{HS}} = 1 - \frac{1}{4\tau}$$

*D. Rosenbaum, et al., PRL* ***76****, 150 (1996)*