Potential of Mean Force Treatment of Salt-Mediated Protein Crystallization

Dikeos Mario Soumpasis* and Yannis Georgalis#
*Max-Planck-Institut für Biophysikalische Chemie, Biocomputation Group, 37018 Göttingen, Germany, and #Institut für Kristallographie, Freie Universität Berlin, 14195 Berlin, Germany

ABSTRACT In the initial stages of crystallization of proteins, monomers aggregate rapidly and form nuclei and large fractal clusters, as previously shown by dynamic light scattering experiments (Georgalis, Y., J. Schüler, J. Frank, D. M. Soumpasis, and W. Saenger. 1995. Protein crystallization screening through scattering techniques. *Adv. Colloid Interface Sci.* 58:57–86). In this communication we initiate an effort to understand the effective interactions controlling charged protein aggregation and crystallization using the potential of mean force (PMF) theory. We compute the PMFs of the system lysozyme-water-NaCl within the framework of the hypernetted chain approximation for a wide range of protein and salt concentrations. We show that the computed effective interactions can rationalize the experimentally observed aggregation behavior of lysozyme under crystallization conditions.

INTRODUCTION

Protein crystallization is the major obstacle on the way to atomic resolution structures that are of paramount importance for establishing structure-function relations in molecular biology, biotechnology, and related areas. The whole discipline is lacking firm physicochemical background, and it is based on empirically formed recipes for conditions vaguely mixed with gross thermodynamic concepts such as supersaturation and solvent activities. From the diagnostics point of view, the extraction of molecular parameters such as size and charge by using power spectroscopy (Kam et al., 1978; Feher and Kam, 1985) dynamic light scattering (DLS) (Eberstein et al., 1994; Muschol and Rosenberger; 1995) is an equally cumbersome issue. It has to be recognized that molecular interactions examined by DLS in concentrated suspensions and high salt molarities are also another partially tractable and not quantitatively understood topic (Pusey and Tough, 1985; Schmitz, 1990).

Both inorganic or small organic molecules and biological macromolecules in a solvent do not crystallize unless the proper environmental conditions are met. Whereas, for simple systems such as salts, the degree of supersaturation may serve as a qualitative measure, for biological molecules stability requirements and interactions with the solvent do not allow for such simplistic considerations.

On general physical grounds we view protein crystallization in an aqueous-electrolytic environment as a process driven by effective, solvent-mediated interactions or potential of mean force (PMF). Even an approximate knowledge of these PMFs as a function of the parameters characterizing crystallization conditions (e.g., concentrations, temperatures, types of ions present) and protein parameters (e.g., size and charge) will provide extremely useful insights and serve to establish semiquantitative criteria for the prediction of optimal crystallization conditions, in the sense that one can enormously reduce the number of crystallization assays usually required with a new protein.

In the case of subtle solvent effects on single biomolecular structures (DNA) in solution, the PMF approach has already produced outstanding results, such as the quantitative description of salt-induced B→Z and B→A transitions (Soumpasis, 1984), ionic effects on the vibrational spectrum of DNA (Garcia and Soumpasis, 1989), and the computations of the detailed 3D distribution of ions and waters around charged biomolecular structures (Klement et al., 1991; Hummer and Soumpasis, 1994; Hummer et al., 1995). We therefore expect the PMF approach to be very useful for probing intermolecular interactions in proteins as well.

Due to the inherent complexity of the systems at hand, it is essential to proceed step by step and start with the simplest cases possible. The various empirical cocktails often chosen to enforce protein crystallization (e.g., complex salts plus alcohols plus water are not physicochemically understood even by themselves, let alone with proteins swimming around). Consequently, we have chosen to work with lysozyme (a well-characterized and stable monomeric protein) in aqueous NaCl solution (the solvent we presently understand best but still not completely!). Despite of this simplicity, computation of the PMFs involved via available "brute-force" simulation techniques (i.e., molecular dynamics or Monte Carlo) and microscopic descriptions are completely out of the question due to the combinatorial explosion (multicomponent system), the presence of long-range Coulombic interactions, and the structural complexity of the lysozyme moiety.

Fortunately, such computations are not needed to obtain a first-order quantitative description. For this purpose it

Received for publication 24 June 1996 and in final form 7 January 1997.
Address reprint requests to Dr. Dikeos Mario Soumpasis, MPI for Biophysical Chemistry, Biocomputational Group and Department of Molecular Biology, Am Fabberg 11, D-37077 Gottingen, Germany. Tel: 49-551-201-1785; Fax: 49-551-201-1467; E-mail: msoumpa@mario01mpibpc.gwdg.de.

0006-3495/97/06/2770/05 $2.00

suffices to treat the protein molecules as hard-charged spheres, the water as a dielectric continuum and the Na^+ and Cl^- ions as small hard-charged spheres. Such a description captures the essential physics of the many body-collective effects (packing and screening) that dominate charged dense liquid phases and has been widely used in the theories of electrolytes and polyelectrolytes.

METHODS

In our previous works, time-resolved DLS studies (Georgalis et al., 1993–1995; Georgalis and Sanger, 1993), we have observed the coexistence of two prominent components in the autocorrelation function. These components can be decoupled with inverse Laplace techniques and identified as lysozyme oligomers and mass fractal clusters (Weitz and Oliveria, 1984; Weitz and Lin, 1986; Meakin, 1988;). The first component does not exhibit any kinetics. Its slowly vanishing amplitude indicates that lysozyme oligomers are consumed to other processes (i.e., nuclei and fractal cluster formation) depending on protein and salt concentration. In contrast, the fractal clusters exhibit kinetics that can be studied with DLS techniques (Klein et al., 1990). The mean hydrodynamic fractal cluster radii exhibit a clear power-law scaling with time. This behavior is characteristic for cluster-cluster diffusion limited (DLCA) aggregation kinetics. From these kinetics and by assuming simple models (Vicsek, 1989) one primarily obtains estimates of the mean cluster radius at zero-lag time, $R_h(0)$ and the fractal dimension of the clusters, d_f. Information on the dynamic exponents z and homogeneity coefficients λ that characterize the cluster reactivity can be best extracted through small-angle light scattering techniques (Asnaghi et al., 1992).

Observables like $R_h(0)$ require some data manipulations before one can predict real trends as a function of protein or salt concentration. We have, therefore, defined an arbitrary observable, the "quasi-stationary" hydro-

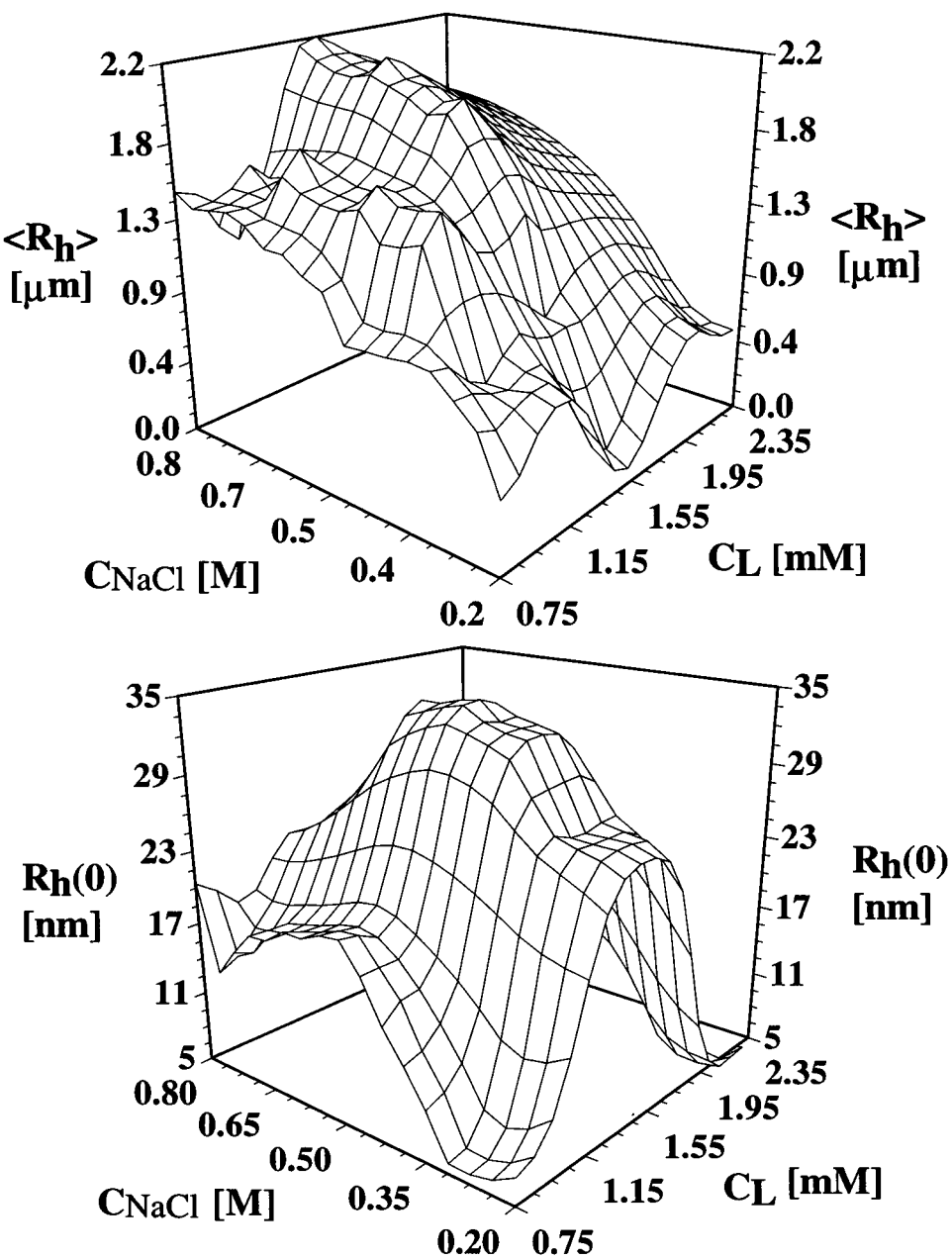

FIGURE 1 Quasi-stationary R_h^{qs} and zero-time extrapolated, $R_h(0)$, hydrodynamic radii of lysozyme fractal clusters plotted as a function of protein and NaCl concentration. The computed surfaces are improved versions of those previously reported (Georgalis et al., 1995).

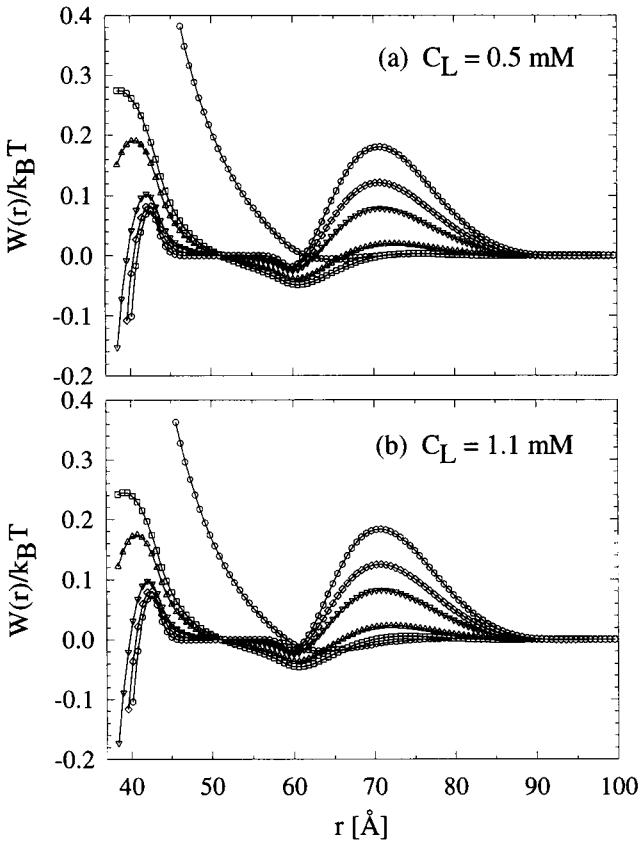

FIGURE 2 PMF computations of (*a*) 0.5 mM and (*b*) 1.0 mM lysozyme in (○) 0.1, (□) 0.32, (△) 0.4, (▽) 0.64, (◇) 0.8, and (◇) 1.0 M NaCl. Note the development of a repulsive barrier as a function of increasing NaCl molarity close to a separation of one lysozyme diameter.

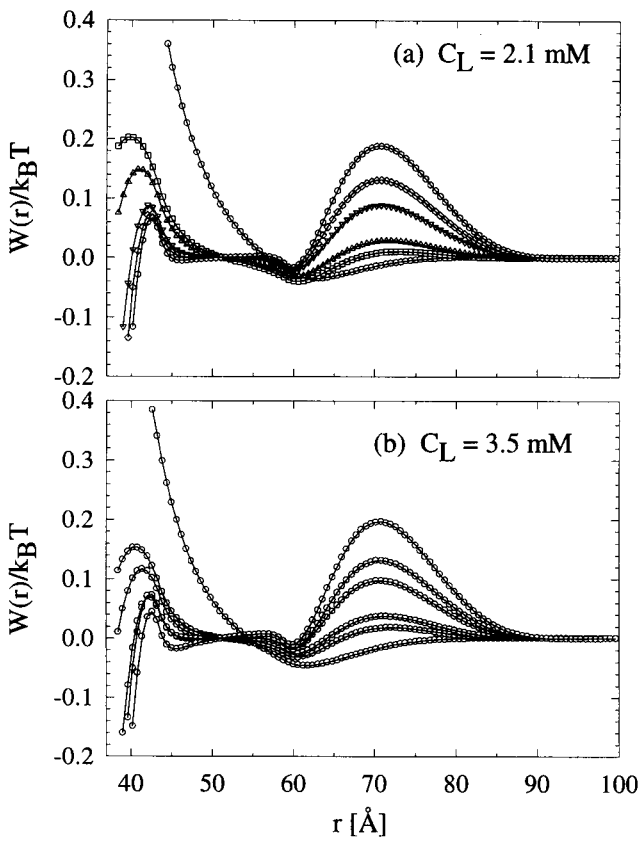

FIGURE 3 PMF computations of (*a*) 2.1 mM and (*b*) 3.5 mM lysozyme in (○) 0.1, (□) 0.32, (△) 0.4, (▽) 0.64, (◇) 0.8, and (◇) 1.0 M NaCl. Note the development of a repulsive barrier as a function of increasing NaCl molarity close to a separation of one lysozyme diameter.

dynamic radius, which is an average of hydrodynamic cluster radii within a discrete time interval; i.e., between 40 and 50 min from the initiation of the experiment. This average can be obtained fairly easily upon completion of the experiments. We usually use it as a first indicator for judging the quality of a series of kinetic experiments conducted at either constant protein or salt concentration.

For the PMF computations the system studied comprises three species; i.e., Na^+ and Cl^- ions and lysozyme- $+7\,|e|$, labeled 1,2,3 involving the five distinct pair PMFs W_{11}, W_{12}, W_{13}, W_{22}, and W_{23} (due to symmetry $W_{11} = W_{22}$, $W_{12} = W_{21}$, $W_{21} = W_{12}$, and $W_{31} = W_{13}$) computed for six thermodynamic states (system compositions) at a temperature of 297.2 K from numerical solution of the system of nonlinear Ornstein-Zernike equations as previously described (Soumpasis, 1993), and for four different molarities of lysozyme covering the range of suboptimal to optimal NaCl-mediated crystallization conditions for the enzyme (Riés-Kautt and Ducruix, 1989).

The statistical mechanical approximation used to calculate the PMFs is the hypernetted chain approximation (HNC) theory (Van Leewen et al., 1959; Morita and Hiroike, 1961) known to yield excellent results for ionic systems (Hafskøld and Stell, 1982; Hummer and Soumpasis, 1993).

Denoting species by Greek indices ($\alpha, \beta, \gamma, \cdots = 1, 2, 3, \cdots$) the five PMFs $W_{\alpha\beta}$ are obtained from numerical solutions of the coupled system of nonlinear integral Ornstein-Zernike equations

$$h_{\alpha\beta}(r) = c_{\alpha\beta}(r) + \sum_{\gamma} \rho_{\gamma} \int c_{\alpha\gamma}(|\mathbf{r} - \mathbf{r}'|) h_{\gamma\beta}(r') d\mathbf{r}' \quad (1)$$

with the HNC-closure relation

$$-\frac{\Psi_{\alpha\beta}(r)}{k_BT} + \frac{W_{\alpha\beta}(r)}{k_BT} = c_{\alpha\beta}(r) - h_{\alpha\beta}(r) \quad (2)$$

where $h_{\alpha\beta}$, $c_{\alpha\beta}$ denote the total and direct correlation functions, respectively. ρ_α is the number density of species α, r is the distance of closest approach, an $\alpha\beta$ ionic pair, and k_B the Boltzmann constant.

$$\Psi_{\alpha\beta} = \begin{cases} \dfrac{q_\alpha q_\beta}{\epsilon r} & : \text{ if } r > \sigma_{\alpha\beta} \\ \infty & : \text{ if } r < \sigma_{\alpha\beta} \end{cases} \quad (3)$$

is the pair interaction potential with $\epsilon = 78.4$, the dielectric constant of water, and $\sigma_{\alpha\beta}$ the distance of closest approach of an α-β pair.

For the Na^+, Cl^- species we use a $\sigma_{11} = \sigma_{12} = \sigma_{22} = 4.90$Å found to yield excellent results in our previous work with NaCl (Soumpasis et al., 1990; Hummer et al., 1994). For the sphere modelling of lysozyme we take a radius of 19.2 Å based on our previous DLS works (Eberstein et al., 1994) yielding $\sigma_{13} = \sigma_{23} = 21.3$ Å and $\sigma_{33} = 38.4$ Å. The charge q_1 of lysozyme was taken to be +7. From the known crystal structure of lysozyme (Protein Data Bank entry 2LYZ) 16 positive and 9 negative charges have been identified. Therefore, a net charge of $+7\,|e|$ is expected. Arginine 5 is completely buried and histidine 15 is partially buried in the polypeptide chain. At low pH another positive charge would be expected at the very most. A value close to that was determined previously (Eberstein et al., 1994; Muschol and Rosenberger, 1995). The HNC equations are solved using numerical techniques described by Belloni (1985), which

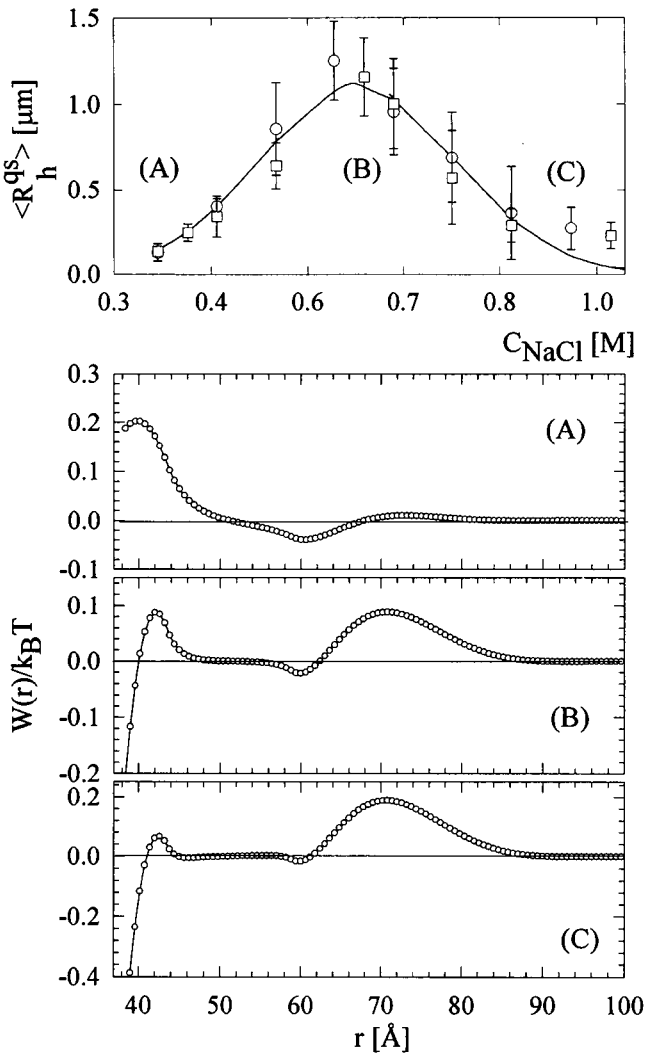

FIGURE 4 Quasi-stationary radii of lysozyme fractal clusters as a function of NaCl molarity. The different symbols denote experiments performed with 2.05 (○) and (□) 3.42 mM lysozyme. The curve through the points is only a guide to the eye. The data are an improved version of those previously reported (Georgalis et al., 1995). (*A*), (*B*), and (*C*) denote those regions where predominantly soluble, microcrystalline, and amorphously precipitated lysozyme with crystals was observed three days after completion of the DLS experiments. In (b), (c) and (d) we display PMFs computed with 2.1 mM lysozyme and 0.32, 0.64, and 1.0 M NaCl; i.e., conditions are those of the 2.05-mM lysozyme experiments appearing in (a). Note the differences in ordinate scales.

were found to yield convergent solutions over a wide range of salt concentrations, i.e., 10^{-3} to 5.0 M NaCl using a discretization of 1024 points.

RESULTS AND DISCUSSION

We have used as observables the mean quasi-stationary hydrodynamic radius of the fractal clusters R_h^{qs} and the zero time-lag size of the clusters $R_h(0)$. Both quantities can be deduced from time-resolved DLS experiments in relatively short times. The three-dimensional plots shown in Fig. 1 were computed from several sets of time-resolved DLS experiments. Lysozyme concentrations varied in the range

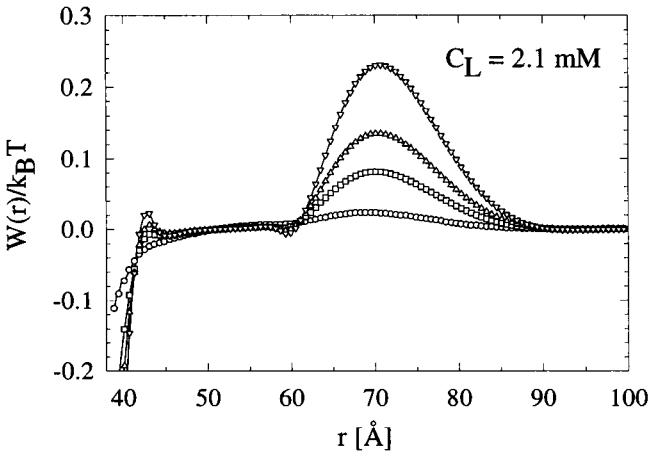

FIGURE 5 PMF computations of 2.1 mM "uncharged" lysozyme in 0.1, 0.4, 0.64, and 1.0 M NaCl. Note the development of the barrier due to packing effects.

between 0.75 and 2.35 mM and NaCl concentrations in the range between 0.2 to 1.0 M. Optimal crystallization conditions are found in the regime of parameters space where the values of the observables R_h^{qs} and $R_h(0)$ are maximal. They indicate a favorable balance of repulsive and attractive forces in the many body systems that we model by means of PMFs.

In Figs. 2 and 3 we display the computed lysozyme-lysozyme PMFs (species 3) for compositions of the system resembling the suboptimal and optimal lysozyme crystallization conditions. Depending on protein and salt concentration, one obtains PMFs exhibiting from purely repulsive (low salt content) to highly non-monotonic (high salt content) behaviors. The structure of the PMF reflects the complex interplay of many body electrostatic and packing forces that cannot be accounted for by traditional theories like DLVO.

Fig. 4 summarizes the conclusions that can be drawn based on our experimental and computational results. At a fixed near-optimal protein concentration and low salt (*A*) one has a purely repulsive PMF and correspondingly little cluster formation, as indicated by R_h^{qs}. Upon increasing the salt concentration up to an optimal value (*B*) one obtains a highly structured PMF with both an attractive short range and a moderately repulsive long range (70-Å peak) barrier. This is accompanied by a maximal R_h^{qs} value and the formation of 1–2 μm fractal clusters. Further increase of the salt content (*C*) leads to the development of an even more pronounced, repulsive barrier and concomitant decrease of the observed cluster size. This latter effect is primarily due to packing and not to Coulombic forces. As can be seen in Fig. 5, the computed PMF for the "uncharged" lysozyme-charged electrolyte system is plotted. Similar qualitative tendencies are obtained after plotting $R_h(0)$ as a function of electrolyte concentration.

The critical salt concentration has been experimentally determined for the system lysozyme-NaCl to vary between 0.5 and 0.65 M NaCl (Georgalis et al., 1993, 1995; Eberstein et al., 1994). At this point fractal and crystal formation

are expected to compete with equal probability. The appearance of the repulsive barrier with increasing amplitude as the concentration of salt increases is in full accordance with experimental findings presented in our previous work and puts, for first time, empirical findings and arbitrarily defined observables onto a sound basis.

CONCLUDING REMARKS

Due to the extreme complexity of the systems involved, we do not think that too fine-grained, "brute-force" simulations are a viable route. The same is also true for traditional mean-field approaches, such as DLVO, that neglects both the non-Coulombic ion-ion repulsion and ion-ion correlations and overestimates the importance of the van der Waals interactions, since it is assumed that they are not modified by the solvent.

Using HNC-PMFs for the lysozyme-NaCl system we find that the concentration dependence of the effective interactions is not monotonic and leads to appreciable aggregation, which is necessary for crystallization, only above a certain concentration and below another, i.e., only in a limited range of salt and protein concentrations. This behavior, which cannot be explained via DLVO-like theories, is in very good agreement with the light scattering data. Therefore, we think that PMF computations combined with time-resolved dynamic light scattering experiments may soon provide a powerful tool for predicting the conditions necessary for crystallization of charged globular proteins induced by simple salts.

Financial support from the Bundesministerium für Bildung und Forschung (to D.M. Soumpasis) and the Deutsche Forschungsgemeinschaft (Sa 196/26-1, to Y.G.) is acknowledged.

REFERENCES

Asnaghi, D., M. Carpineti, M. Giglio, and M. Sozzi. 1992. Coagulation kinetics and aggregate morphology in the intermediate regimes between diffusion-limited and reaction-limited cluster aggregation. *Phys. Rev. A* 45:1018–1023.

Belloni, L. 1985. A hypernetted chain study of highly asymmetrical polyelectrolytes. *Chem. Phys.* 99:43–54.

Eberstein, W., Y. Georgalis, and W. Saenger. 1994. Molecular interactions in crystallizing lysozyme solutions studied by photon correlation spectroscopy. *J. Cryst. Growth.* 143:71–78.

Feher, G., and Z. Kam. 1985. Nucleation and growth of protein crystals: general principles and assays. *Methods Enzymol.* 114:77–111.

Garcia, A. E., and D. M. Soumpasis. 1989. Harmonic vibrations and thermodynamic stability of a DNA oligomer in monovalent salt solution. *Proc. Natl. Acad. Sci. U.S.A.* 86:3160–3164.

Georgalis, Y., and W. Saenger. 1993. Time-resolved light scattering studies on protein precrystallization fractal clusters. *Adv. Colloid Interface Sci.* 46:165–183.

Georgalis Y., J. Schüler, W. Eberstein, and W. Saenger. 1994. Time-resolved light scattering studies on protein precrystallization fractal clusters. *In* Fractals in Nature and Applied Sciences (A-41). M. M. Novak, editor. Elsevier Science, North Holland, Amsterdam. 139–151.

Georgalis, Y., J. Schüler, J. Frank, D. M. Soumpasis, and W. Saenger. 1995. Protein crystallization screening through scattering techniques. *Adv. Colloid Interface Sci.* 58:57–86.

Georgalis, Y., A. Zouni, W. Eberstein, and W. Saenger. 1993. Formation dynamics of protein precrystallization fractal clusters. *J. Cryst. Growth.* 126:245–260.

Hafskøld, B., and G. Stell. 1982. The equilibrium statistical mechanics of simple ionic liquids. *In* The Liquid State of Matter. Fluids, Simple and Complex. E. W. Montroll and J. L. Lebowitz, editors. North Holland, Amsterdam. 175–274.

Hummer, G., A. E. Garcia, and M. D. Soumpasis. 1995. Hydration of nucleic acid fragment: comparison of theory and experiment for high resolution crystal structures of RNA, DNA DNA-drug complexes. *Biophys. J.* 68:1639–1652.

Hummer, G., and M. D. Soumpasis. 1992. An extended RISM study of simple electrolytes: pair correlations in a NaCl-SPC water model. *Molecular Physics.* 75:633–651.

Hummer, G., and M. D. Soumpasis. 1993. Correlations and free energies in restricted primitive model of electrolytes. *J. Chem. Phys.* 98:581–591.

Hummer, G., and M. D. Soumpasis. 1994. Statistical mechanical treatment of the structural hydration of biological macromolecules: results on B-DNA. *Phys. Rev. E.* 50:5085–5095.

Hummer, G., M. D. Soumpasis, and A. E. Garcia. 1995. Potential of mean force description of ion-interactions and structural hydration in biomolecular systems. *In* Non-Linear Excitations in Biomolecules. M. Peyrard, editor. Springer-Verlag, New York. 82–99.

Hummer, G., M. D. Soumpasis, and M. Neumann. 1994. Computer simulation of aqueous NaCl electrolytes. *J. Phys. Condensed Matter.* 6:A141–A144.

Kam, Z., H. B. Shore, and G. Feher. 1978. On the crystallization of proteins. *J. Mol. Biol.* 123:539–555.

Klein, R., D. A. Weitz, M. Y. Lin, H. M. Lindsay, R. C. Ball, and P. Meakin. 1990. Theory of scattering from colloidal aggregates. *Progr. Colloid Polym. Sci.* 81:161–168.

Klement, R., M. D. Soumpasis, and T. Jovin. 1991. Computation of ionic distributions around charged biomolecular structures: results for right-handed and left-handed DNA. *Proc. Natl. Acad. Sci. U.S.A.* 88: 4631–4635.

Meakin, P. 1988. The growth of fractal aggregates and their fractal measures. *In* Phase Transitions and Critical Phenomena, Vol. 12. C. Domb and J. L. Lebowitz, editors. Academic Press, New York. 335–489.

Morita, T., and K. Hiroike. 1961. A new approach to the theory of classical fluids III. General treatment of classical systems. *Progr. Theor. Phys.* 25:537–578.

Muschol, M., and F. Rosenberger. 1995. Interactions in undersaturated and supersaturated lysozyme solutions: static and dynamic light scattering results. *J. Chem. Phys.* 103:10424–10432.

Pusey, P. N., and R. J. A. Tough. 1985. Particle interactions. *In* Dynamic Light Scattering. R. Pecora, editor. Plenum Press, New York. 85–179.

Riés-Kautt, M. M., and A. F. Ducruix. 1989. Relative effectiveness of various ions on the solubility and crystal growth of lysozyme. *J. Biol. Chem.* 264:745–748.

Schmitz, S. K. 1990. An Introduction to Dynamic Light Scattering by Macromolecules. Academic Press, New York.

Soumpasis, D. M. 1984. Statistical mechanics of the B→Z transition of DNA: Contributions of diffuse ionic interactions. *Proc. Natl. Acad. Sci. U.S.A.* 81:5116–5120.

Soumpasis, D. M. 1993. Formal Aspects of the Potential of Mean Force Approach in Computations of Biomolecular Structures, Achievements, Problems and Perspectives. D. M. Soumpasis and T. Jovin, editors. Springer-Verlag, Berlin. 223–239.

Soumpasis, D. M., A. E. Garcia, R. Klement, and T. M. Jovin. 1990. The potential of mean force (PMF) approach for treating ionic effects on biomolecular structures in solution. *In* Theoretical Biochemistry and Molecular Biophysics. D. Beveridge and R. Lavery, editors. Adenine Press Guilderland, New York. 343–360.

Van Leewen, J. M. J., J. Groeneweld, and J. De Boer. 1959. New method for the calculation of the pair correlation function. *Physica.* 25:792–808.

Vicsek, T. 1989. Fractal growth phenomena. World Scientific, Singapore.

Weitz, D. A., and M. Y. Lin. 1986. Dynamic scaling of cluster-mass distributions in kinetic colloid aggregation. *Phys. Rev. Lett.* 57:2037–2040.

Weitz, D. A., and M. Oliveria. 1984. Fractal structures formed by kinetic aggregation of aqueous gold colloids. *Phys. Rev. Lett.* 52:1433–1436.