Molecular dynamics and small-angle neutron scattering of lysozyme aqueous solutions

Maria Concetta Abramo, Carlo Caccamo, Massimo Calvo, Valeria Conti Nibali, Dino Costa, Rita Giordano, Giuseppe Pellicane, Romina Ruberto, Ulderico Wanderlingh

Molecular dynamics and small-angle neutron scattering of lysozyme aqueous solutions

Journal:	*Philosophical Magazine & Philosophical Magazine Letters*
Manuscript ID:	TPHM-10-Jun-0282.R2
Journal Selection:	Philosophical Magazine
Date Submitted by the Author:	27-Jan-2011
Complete List of Authors:	Abramo, Maria Concetta; Universita' degli Studi di Messina, Fisica Caccamo, Carlo; Universita' degli Studi di Messina, Fisica Calvo, Massimo; Universita' degli Studi di Messina, Dip. di Fisica Conti Nibali, Valeria; University of Messina, physics Costa, Dino; Universita' degli studi di Messina, Fisica Giordano, Rita; University of Messina, Physics Pellicane, Giuseppe; University of Kwazulu-Natal, School of Physics Ruberto, Romina; Universita' degli Studi di Messina, Fisica Wanderlingh, Ulderico; University of Messina, Physics
Keywords:	proteins, molecular dynamics, SANS
Keywords (user supplied):	coarse-grained models

Molecular dynamics and small-angle neutron scattering of lysozyme aqueous solutions

M. C. Abramo[a], C. Caccamo[a], M. Calvo[a], V. Conti Nibali[a], D. Costa[a]*, R. Giordano[a], G. Pellicane[b], R. Ruberto[a], U. Wanderlingh[a]

[a] *Dipartimento di Fisica, Università degli Studi di Messina and CNISM (Consorzio Nazionale Interuniversitario di Struttura della Materia) Viale F. Stagno d'Alcontres 31, 98166 Messina, Italy*

[b] *School of Physics, University of Kwazulu-Natal Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa*

Abstract

Molecular dynamics simulations of a coarse-grained, embedded-charge model of lysozyme aqueous solutions are compared with small-angle neutron scattering experiments. Measures concern different solutions with a 10% by weight protein concentration and an increasing pH in the range 2-6. The model is based on a soft-core modification of the original Carlsson, Malmsten and Linse model [J. Phys. Chem. B 105 (2001) p.9040], where in particular all residues carrying an appreciable amount of residual charge, as a function of the pH, are explicitly taken into account into the overall macromolecular interaction. Simulations reproduce qualitatively the experimental trend of the structure factor as, in particular, the observed change from a low-pH regime, dominated by repulsive interactions, to a behaviour mainly determined by attractive forces, at higher pH. Possible improvements of the model, toward a better reproduction of the structural properties of the real solution are proposed.

*Corresponding author. Email: dino.costa@unime.it

1 Introduction

The theoretical description of protein solutions in terms of interactions among the microscopic species is a major scientific task, lively addressed in the current literature (see e.g. our recent paper [1] for an up-to-date bibliography). Modelling diluted protein solutions typically regards to in-vitro experiments, where solvent-mediated and screened long-range interactions play a major role. The study of concentrated solutions is relevant instead for understanding the protein aggregation processes, which have a critical importance in many biological phenomena, and are of crucial interest in the preparation and growth of single protein crystals, employed for the investigation of the internal macromolecular structure [2,3]. In a larger perspective, such studies may shed light on the physico-chemical properties of protein matter in-vivo, in the molecularly crowded cellular environment, characterized by sparse solvent and strong contact interactions.

In this scenario, coarse-grained models, in which the protein-protein interactions are described in terms of distributed interaction sites, play a particularly relevant role. In fact, such models place at an intermediate level between a fully atomistic description of the solution, and a bare central potential representation, where the macroparticle (solvent-mediated) interactions are usually approximated by simple short-range forces. Fully atomistic models can determine quite accurately specific properties of the protein system, like the kinetic of protein folding or free energy surfaces (see e.g. [4] and the recent reviews [5–7]), but their use is typically restricted to two or few more proteins in the solvent environment; central models provide instead only a generic description of the structure and phase diagram of protein solutions [8–16]. For these reasons, "mesoscopic" effective models could in principle provide a viable path to study on a quantitative basis the bulk properties of protein solutions. A second key feature of coarse-grained models is that they embody the intrinsic anisotropy of protein-protein interactions; in fact, distributed-site interaction effects, often described in terms of patch-patch interactions [17, 18], are known to be relevant both in the crystal [19] and in the disordered phase [20] of protein systems. Models of this type have been studied by several authors (see e.g. [18, 20–29] and references) and should be more appropriate in order to reproduce the structure and detailed shape of the phase diagram, as well as more complex phenomena as, for instance, self-assembly. Among coarse-grained models, the de-

scription of lysozyme aqueous solutions introduced by Carlsson, Malmsten and Linse [21, 22] (henceforth CML) is based on a tailoring of experimental solution conditions. In the CML picture, lysozymes are represented by hard spheres surrounded by a spherically symmetric short-range hydrophobic attraction; the amino acid residues are explicitly taken into account as a collection of charged sites whose spatial distribution reproduces that of the real protein; the aqueous solvent and added salt effects enter such description only implicitly, contributing to the overall effective protein-protein interactions. CML-like models are appealing also because they are flexible enough to allow in principle for fractional charges, for a refined description of hydrophobic effects, as well as for non-spherical molecular shapes.

In this work we carry out a molecular dynamics (MD) simulation study of the CML model, in which a continuous, soft-core interaction, suitable for MD, replaces the hard-core repulsion adopted in the original prescription [21]. We compare MD data, at various pH and fixed temperature and protein concentration, with newly generated small-angle neutron scattering (SANS) results. Our purpose is to ascertain whether the CML class is able to capture the structural correlations observed in real lysozyme solutions. This topic was not explicitly addressed in previous studies: in particular, in [22] the authors studied through Monte Carlo simulations the general structure and clustering properties of the model, and only alluded to an agreement with the experimental trends observed in the structure factor [30, 31], though no detailed comparison was then presented. Later, Rosch and Errington studied the liquid-vapor coexistence, for a different choice of potential parameters [27] and found a binodal line narrower than the experimental one; the phase behaviour of the model in a confined geometry was also investigated by the same authors [28]. Finally, very recently, we have addressed the possibility to describe from a dynamical viewpoint the approach to the liquid-vapor phase separation of the model [1]. The comparison between simulated and experimental structural correlations presented in this work allows in principle for a systematic refinement of the model, a task to be seen as a preliminary step towards the more challenging calculations of the thermodynamic properties, possibly including the overall appearance of the phase portrait.

The paper is organized as follows: the experimental procedure and data handling are described in Section 2; the model and MD simulations are introduced in Section 3; results are reported and discussed in Section 4, while conclusions follow in Section 5.

2 Experimental procedure and data handling

Lysozyme is a protein formed by 129 amino acids, folding into a compact globular structure with a molecular weight of 14.3 kD and isoelectric point at pH = 11. The protein was purchased from Sigma, further dialysed against D_2O, and repeatedly freeze dried, so to substitute the protein exchangeable hydrogen with deuterium, in order to avoid pollution of the D_2O buffers used in the preparation. Samples were prepared at a concentration of 10% by weight in D_2O citrate-phosphate buffer at pH = 2.8, 4.2, 5.07 and 6.04. A further sample, concentrated at 2% by weight and higher ionic strength, was also prepared by using a 0.07 M NaCl solution, in order determine the form factor of the protein, i.e. the Fourier transform of the density distribution of the isolated protein; in fact, in this case, the low concentration and the high ionic strength (whose purpose is to screen the charged polar groups of the protein) make practically negligible the contributions of the lysozyme-lysozyme structure factor to the scattered intensity.

SANS experiments were carried out on the PAXE spectrometer at the ORPHEE reactor of the Laboratoire Lèon Brillouin in Saclay, France. Samples were contained, at room temperature, in quartz cells of path length of 2 mm. Two different spectrometer configurations were used, namely one with a sample-to-detector distance of 5.0 m and $\lambda = 7$ Å, and another one with a detector distance of 1.4 m and $\lambda = 5$ Å. In this way we covered two partially overlapping Q wave-vector ranges: one going from 0.007 to 0.098 $Å^{-1}$ and the other one from 0.045 to 0.325 $Å^{-1}$.

All data, collected by a bidimensional detector, were converted in intensity *vs* exchanged wave-vectors and corrected for the background and the empty cell. In order to remove the solvent contribution, each buffer was measured in the same conditions of the corresponding solution and, after correction, subtracted with the appropriate transmission coefficient. Finally the data were normalized to absolute scale by dividing for the intensity of a secondary standard of known cross sections [32]. After such procedures the spectra collected in the two different Q-ranges match very well in the region with overlapping Q (see Figure 5 in section 4).

In a SANS experiment the scattered intensity is generally expressed as

$$I(Q) = n\Delta\rho^2 V^2 F(Q) S(Q)\,, \tag{1}$$

where n is the number density of the scatterers, V is the volume of a single

scatterer, $\Delta\rho$ is the contrast factor accounting for the difference between the average (over the volume V) scattering cross section of the protein and the solvent, $F(Q)$ and $S(Q)$ are, respectively, the form factor and the structure factor. Equation (1) holds strictly for monodisperse spherical particles [33], but corrections turn out to be in our case negligibly small [34].

3 Model and molecular dynamics simulations

As anticipated in the Introduction, a lysozyme solution at various protein concentrations, pH, and ionic strengths is represented in this work by an effective one-component fluid. The interaction between two macromolecules is given by a suitable modification of the original CML prescription [21, 22], within the parameterization proposed by Rosch and Errington [27]. We refer the reader to our recent paper [1], for a detailed description of the procedure leading to the present model from previous prescriptions [21, 22, 27]. We shortly recall that the total interaction potential is given by the sum of two different contributions. The first one is a central soft-core interaction in the form

$$U^{\rm soft} = \varepsilon \left[\left(\frac{\sigma}{r}\right)^{48} - \left(\frac{\sigma}{r}\right)^{6} \right] , \qquad (2)$$

that is characterized by a minimum $U^{\rm soft}_{\rm min} = -7\varepsilon/(8\ \times 2^{3/7})$ at $r_{\rm min} = \sigma\ \sqrt[14]{2}$. The r^{-6} decay gives the same (attractive) long-range behaviour of previous models, whereas the 48-power repulsion sets as a reasonable compromise between *(i)* the most common choice of a 12-6 Lennard-Jones potential, which yields too soft a repulsion with respect to the original hard-core CML model, and *(ii)* an even steeper repulsive barrier, which would yield too intense impulsive forces at close contact, thus requiring exceedingly small time steps to integrate the equation of motions. In Equation (2) $\sigma = 36.72\,$Å; the attractive contribution, with $\varepsilon = 4.84\,$kJ/mol (yielding $U^{\rm soft}_{\rm min} = -3.15\,$kJ/mol), takes into account the dispersion forces between proteins in the aqueous solution; the resulting potential is shown in Figure 1. According to the Noro-Frenkel law of corresponding states [35], holding for a large class of central potentials, the bare central model (2) is characterized by a critical temperature $T_{\rm cr} = 300\,$K.

As in previous works [21, 22, 27], the second contribution to the total protein-protein interaction is represented by several interacting sites (up

to 32), representing the amino acid residues which carry an appreciable amount of electrostatic charges, as determined by the experimental titration curve [36]; the number of charges (all bearing one positive or negative electronic charge) depends on the pH of the solution (see Table 1); the position of various sites is adjusted so to lie on a sphere 2 Å beneath the hard-sphere surface. Charges are then treated in a linearized Poisson-Boltzmann framework, giving rise to a (Debye-Hückel) screened interaction:

$$U_{\mathrm{DH}}(r) = \frac{Z_i Z_j e^2}{4\pi\epsilon_0\epsilon_r}\frac{1}{r}\exp(-\kappa r)\,, \tag{3}$$

here r is the distance between two charged sites, Z_i is the charge of site i (having the values ± 1, and to be switched on/off depending on the pH of the solution), e is the electronic charge, ϵ_0 the permittivity of vacuum, and $\epsilon_{\mathrm{r}} = 86.765 - 0.3232 \times T\,(^{\circ}\mathrm{C})$ the relative permittivity of the solvent. The Debye screening length κ^{-1} is given by:

$$\kappa^{-1} = \left[\frac{2I}{\epsilon_0\epsilon_r k_{\mathrm{B}} T}\right]\,, \tag{4}$$

where I is the ionic strength. In the model, the aqueous solvent is represented only implicitly, as a continuous medium of relative permittivity ϵ_r, entering Equation (3); equally, the presence of added salts is taken into account only through the ionic strength appearing in Equation (4). A schematic illustration of the procedure that leads to the coarse-grained picture, starting from the complete protein description, is given in Figure 2.

MD simulations have been carried out by using the program `MOLDY` [37]. We have employed a sample composed by one thousand particles (with a total number of $\sim 30\,000$ interaction sites), enclosed in a cubic box with standard periodic boundary conditions. The interaction cutoff has been fixed to 130 Å, roughly corresponding to 3.5 times the protein diameter, with standard long-range corrections. A time step of $\Delta t = 20\,\mathrm{fs}$ has been adopted, with cumulations times of about 10 ns, divided in sub-blocks of 2 ns, in order to calculate the variance over the calculated average values.

4 Results and discussion

Simulations have been carried out at constant temperature $T = 293\,\mathrm{K}$ and constant protein concentration, $0.1\,\mathrm{gr/cm^3}$, corresponding to a number den-

sity $\rho = 4.2 \cdot 10^{-6}$ Å^{-3}. Four different pH have been investigated; the number of corresponding active positive and negative charges is reported in Table 1. The ionic strength of the solution in the experimental setup, entering the screening length in Equation (4), has been imposed by the buffer, and is also reported in Table 1.

Results for the MD radial distribution function $g(r)$ and for the structure factor $S(Q)$ are reported in Figure 3; statistical uncertainties are practically negligible on the scale of the figure. As visible, at low pH the net repulsive charge gives rise to a correlation depletion around the distance of the first neighbours in the $g(r)$; this function appears almost structureless but for a broad and low peak around 60/70 Å. The repulsive interactions are then progressively reduced as the net charge decreases (i.e. as the pH increases); this fact is reflected by the progressive structuring of a well defined peak in the $g(r)$ at $r \sim \sigma$; besides this feature, the correlation function remains almost flat, as can be expected from the low density of the samples. The structure factor, also reported in Figure 3, faithfully reproduces in the Q space the features of $g(r)$; one can observe in particular the shift of the main peak to larger Q as the pH increases, signalling that different correlation distances appear in the sample. A crossover between two different behaviours of $S(Q \to 0)$ occurs at pH $\sim 4/5$, signalling that the fluid passes from a regime mostly dominated by repulsive forces to a different situation where attractive interactions become more and more relevant, with the ensuing development of progressively enhanced density fluctuations.

As for the experimental results, investigations using several techniques have shown that lysozyme has roughly the shape of a prolate ellipsoid [38,39]; the form factor of an ellipsoidal particle with axes a, b, b and volume $V = 4\pi(abc)/3$, averaged over all possible orientation, is given by [40]:

$$F(Q) = \int_0^1 \left(\frac{3J_1(u)}{u} \right)^2 \mathrm{d}u \qquad (5)$$

where J_1 denotes the 1st-order spherical Bessel function and u is a function of the angle α between Q and a defined as $u = Q\sqrt{a^2\alpha^2 + b^2(1-\alpha^2)}$. Figure 4 shows the evaluation of $F(Q)$ based on a bestfit of the experimental data obtained from the 2% concentrated sample (see section 2). As a standard procedure, error bars reported in the Figure are estimated as $\sqrt{c}/c$, where c is the number of detector counts collected at a given Q. Since residual

interaction effects influence the low-Q region, the fit has been performed only for $Q > 0.1$, where, as visible in the inset of Figure 4, the expected Guinier behaviour [i.e. $I(Q) \propto \exp(-Q^2 R_g^2/3)$ at low Q, where R_g is the gyration radius] is well reproduced. The bestfit values for a and b are, respectively, 13.5 Å and 21.9 Å, in good agreement with previous findings [41, 42].

The experimental and MD scattering intensities are compared in Figure 5. The choice to compare real and simulated intensities, rather than using the structure factors, depends on issues generally related to the calculation of the experimental structure factor. In fact, such correlation is obtained by dividing the experimental intensity by the fitted analytical form factor, see Equation (1). In the high-Q region (the so called "Porod regime") the intensity drops as $1/Q^4$, and it is related to the interface between sample and solvent. Now *(i)*, whereas the analytical form factor shows such exact asymptotic behaviour, the experimental intensity behaves as $1/Q^4$ only "on average", due the the roughness of the real protein/solvent interface and to statistical uncertainties. Moreover *(ii)*, at high Q, the experimental intensity and the analytical form factor are both close to zero. As a result of *(i)* and *(ii)*, the ratio of such quantities can become quite noisy in the high Q regime. A workaround to this problem usually consists in e.g. smoothing the data or mixing experimental and analytical form factor, procedures which rely on some degree of arbitrariness and may possibly affect an accurate determination of the structure factor over a substantially extended Q range. The comparison in Figure 5 is then carried out between the experimental intensity and the simulated one, immediately obtained by multiplying the structure factor in Figure 3 with the analytical form factor in Figure 4 [the prefactor in Equation (1) has been estimated as $\approx 1.02\,\mathrm{cm}^{-1}$]. It turns out that at low pH the position and height of the experimental main peak is reasonably reproduced by MD. Upon increasing pH, also the experimental data show a change in the behaviour of $S(Q)$, though this crossover manifests at lower pH with respect to the simulation results; consequently, the increase of the scattering intensity at low-Q is less marked in the MD than in the experimental observation. This discrepancy can be ascribed to different sources: the attractive part in the overall theoretical interaction is probably underestimated; at the same time, our implicit solvent description could poorly account for the effective counterion condensation around the charged sites. On the other hand, the experimental results may be affected by the formation of spurious nanoaggregates, which are more likely to appear with

increasing pH; using a low protein concentration and/or a lower temperature may help to minimize such effects.

5 Conclusions

We have performed a molecular dynamics investigation of a coarse-grained multi-site model of aqueous lysozyme solutions, and compared the simulated static correlations with corresponding small-angle neutron scattering data. Our comparison shows that the MD scattering intensities follow the experimental trend only qualitatively. In particular, concentration fluctuations are generally underestimated by the present model.

Coarse-grained models, like the CML one investigated in this work, play a key role for the possibility they offer to reproduce the distributed-site interactions observed in real protein solutions and to simulate large-scale samples, suitable for the investigation of bulk properties. Besides such advantages, the systematic comparison with experiments and the specific way the CML model is built, allow one to improve the observed basic performances by means of systematic refinements of the model itself. In particular we plan to introduce a better description of attractive effects and possibly a more realistic representation of the ellipsoidal shape of the molecule, as well as of its charge distribution. Hence, the analysis presented in this work represents a preliminary survey of the accuracy of the model, necessarily preluding a (much more demanding) future attempt to calculate the thermodynamic properties of the model, including the overall aspect of the phase diagram.

Acknowledgements

We gratefully acknowledge the computer time and technical support provided by the Centro di Calcolo Elettronico dell'Università di Messina "A. Villari". and by the PI2S2 Project, managed by the Consorzio COMETA (more information available at `http://www.pi2s2.it` and `http://www.consorzio-cometa.it`).

References

[1] M.C. Abramo, C. Caccamo, D. Costa, G. Pellicane and R. Ruberto, J. Phys. Chem. B 114 (2010) p.9109.

[2] A. McPherson, *Crystallization of Biological Macromolecules*, Cold Spring Harbor Laboratory Press, New York, 1999.

[3] K.F. Kelton, *Crystal Nucleation in Liquids and Glasses* in *Solid State Physics*, Vol. 45, H. Ehrenreich and D. Turnbull eds, Academic Press, London, 1991.

[4] G. Pellicane, G. Smith and L. Sarkisov, Phys. Rev. Lett. 101, (2008) p.248102.

[5] J.M. Shea and C.L. Brooks III, Ann. Rev. Phys. Chem. 52 (2001) p.499.

[6] C.D. Snow, E.J. Sorin, Y.M. Rhee and V.S. Pande, Ann. Rev. Biophys. and Biomolecular Stuct. 34 (2005) p.43.

[7] P.F.N. Faisca, J. Phys.: Condens. Matter 21 (2009) p.373102.

[8] D. Rosenbaum, P.C. Zamora and C.F. Zukoski, Phys. Rev. Lett. 76 (1996) p.150.

[9] A. Lomakin, N. Asherie and G.B. Benedek, J. Chem. Phys. 104 (1996) p.1646.

[10] D.F. Rosenbaum, A. Kulkarni, S. Ramakrishnan and C.F. Zukoski J. Chem. Phys. 111 (1999) p.9882.

[11] D. Costa, P. Ballone and C. Caccamo, J. Chem. Phys. 116 (2002) p.3327.

[12] A. Stradner, H. Sedgwick, F. Cardinaux, W.C.K. Poon, S.U. Egelhaaf and P. Schurtenberger Nature (London) 432 (2004) p.492.

[13] M. Broccio, D. Costa, Y. Liu and S.-H. Chen, J. Chem. Phys. 124 (2006) p.084501.

[14] F. Cardinaux, A. Stradner, P. Schurtenberger, F. Sciortino and E. Zaccarelli, Europhys. Lett. 77 (2007) p.48004.

[15] G. Pellicane, D. Costa and C. Caccamo, J. Phys.: Condens. Matter 16 (2004) p.S4923.

[16] G. Pellicane, D. Costa and C. Caccamo, J. Phys. Chem. B 108 (2004) p.7538.

[17] M. Hloucha, J.F.M. Lodge, A.M. Lenhoff and S.I. Sandler, J. Crystal Growth 232 (2001) p.195.

[18] H. Liu, S.K. Kumar and F. Sciortino, J. Chem. Phys. 127 (2007) p.084902.

[19] J. Janin and F. Rodier, Proteins: Struct., Funct. Genet. 23 (1995) p.580.

[20] A. Lomakin, N. Asherie and G.B. Benedek, Proc. Natl. Acad. Sci. USA 96 (1999) p.9465.

[21] F. Carlsson, M. Malmsten and P. Linse J. Phys. Chem. B 105 (2001) p.9040.

[22] F. Carlsson, M. Malmsten and P. Linse J. Phys. Chem. B 105 (2001) p.12189.

[23] N.M. Dixit and C.F. Zukoski, J. Chem. Phys. 117 (2002) p.8540.

[24] X. Song, Phys. Rev. E 66 (2002) p.011909.

[25] E. Allahyarov, H. Löwen, J.-P. Hansen and A.A. Louis, Europhys. Lett. 57 (2002) p.731.

[26] J.M. Cheung V.K. Shen, J.R. Errington and T.M. Truskett, Biophys. J. 92 (2007) p.4316.

[27] T.W. Rosch and J.R. Errington, J. Phys. Chem. B 111 (2007) p.12591.

[28] T.W. Rosch and J.R. Errington, J. Phys. Chem. B 112 (2008) p.14911.

[29] X. Li, J.D. Gunton and A. Chakrabarti, J. Chem. Phys. 131 (2009) p.115101.

[30] U. Wanderlingh, R. Giordano and G. Giunta, Nuovo Cimento 16 (1994) p.1493.

[31] Y. Georgalis, P. Umbach, W. Saenger, B. Ihmel and D.M. Soumpasis, J. Am. Chem. Soc. 121 (1999) p.1627.

[32] J. Teixeira in *Structure and dynamics of supramolecular aggregates and strongly interacting colloids*, S.-H. Chen, J.S. Huang and P. Tartaglia eds, Kluwer, Dordecht, 1992.

[33] M. Kotlarchyk and S.-H. Chen, J. Chem. Phys. 79 (1983) p.2461.

[34] R. Giordano, A. Grasso, J. Teixeira, F. Wanderlingh and U. Wanderlingh, Phys. Rev. A 43 (1991) p.6894.

[35] M.G. Noro and D. Frenkel, J. Chem. Phys. 113 (2000) p.2941.

[36] M. Ramanadham, L.C. Sieker and L.H. Jepsen, Acta Crystallogr. Sect. B 46 (1990) p.63.

[37] K. Refson, Comp. Phys. Comm. 126 (2000) p.310; software available at: `http://ccpforge.cse.rl.ac.uk/frs/?group_id=34`.

[38] N.C.J. Strynadka and M.N.G. James, J. Mol. Biology 220 (1991) p.401.

[39] H. Schwalbe *et al* Protein Sci. 10 (2001) p.677.

[40] A. Guinier and G. Fournet, *Small Angle Scattering of X-rays,* Wiley, New York, 1955.

[41] B. Jacrot, Rep. Progr. Phys. 39 (1976) p.942.

[42] G.A. Bentley *et al* J. Chem. Phys. 76 (1979) p.817.

TABLES AND TABLE CAPTIONS

Table 1: The four different pH analysed in this work, with the corresponding positive and negative charges on the coarse-grained model (in units of e). The ionic strength of the solution (in mM) is also reported.

pH	2.8	4.2	5.07	6.04
Z^+	19	19	19	18
Z^-	5	7	9	9
I	35	81	102	124

FIGURE CAPTIONS

Figure 1. The central 48/6 interaction potential adopted in this work, see Equation (2).

Figure 2. A schematic representation of the full lysozyme protein in terms of the coarse-grained model. The picture on the right is taken from [22]; positive charges are in blue, negative charges in red.

Figure 3. MD behaviour of $g(r)$ (a) and $S(Q)$ (b) for all pH investigated in this work. Error bars are systematically smaller than the size of the symbols.

Figure 4. Experimental determination of the form factor with error bars (symbols) and corresponding bestfit (line). Data are obtained from the 2% concentrated sample. Inset: Guiner plot in the low-Q region (see text).

Figure 5. MD (a) and experimental (b) scattering intensities (in cm^{-1}); note the different vertical scales. The legend in (a) holds for both panels. In panel (b) open and full symbols refer to different spectrometer configurations. Error bars on experimental data are reported at pH = 4.2, at two representative low and intermediate Q-vectors.

FIGURES

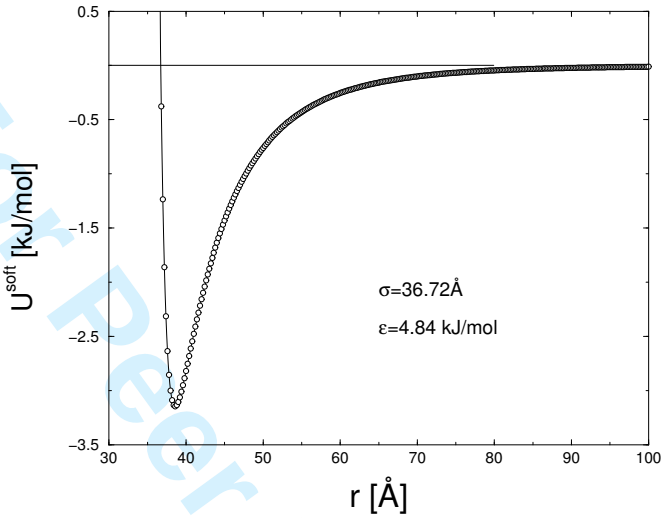


Figure 1:

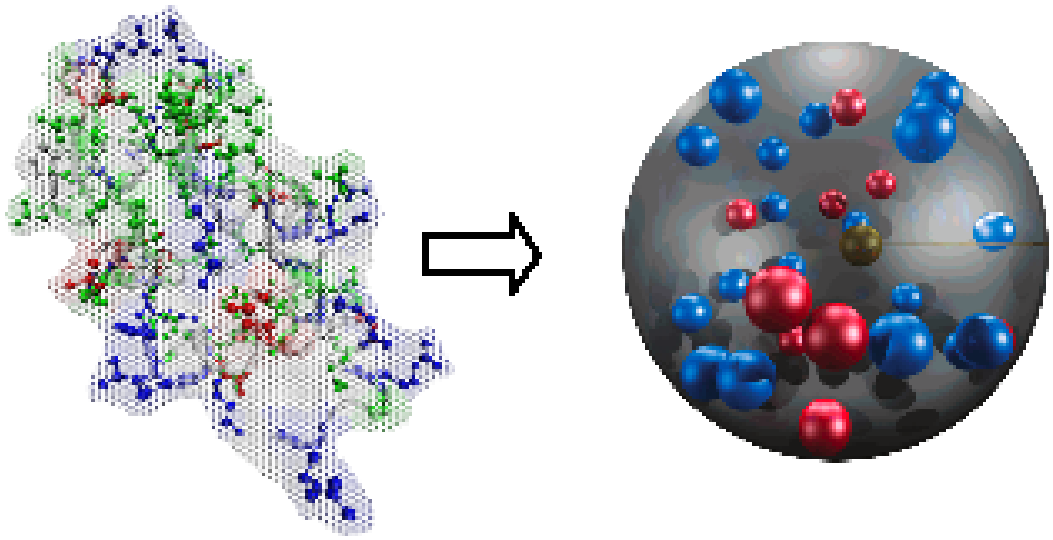
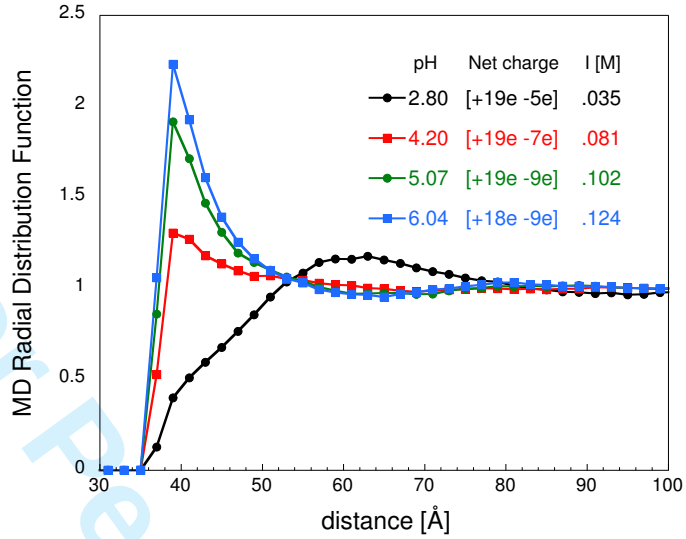
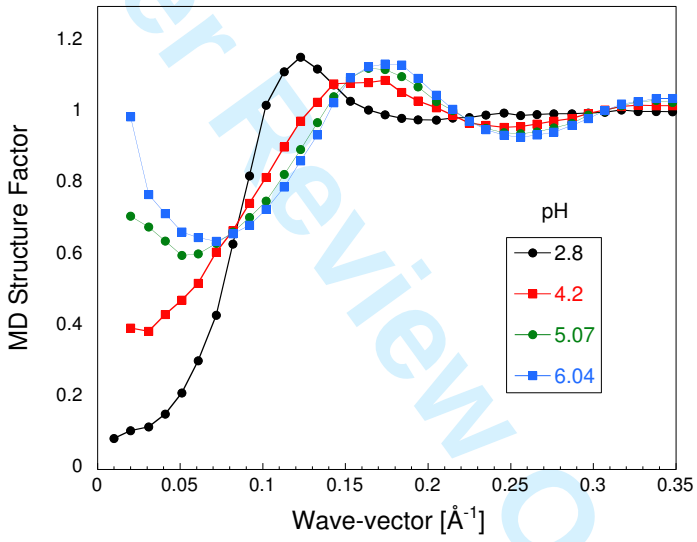


Figure 2:



(a)



(b)

Figure 3:

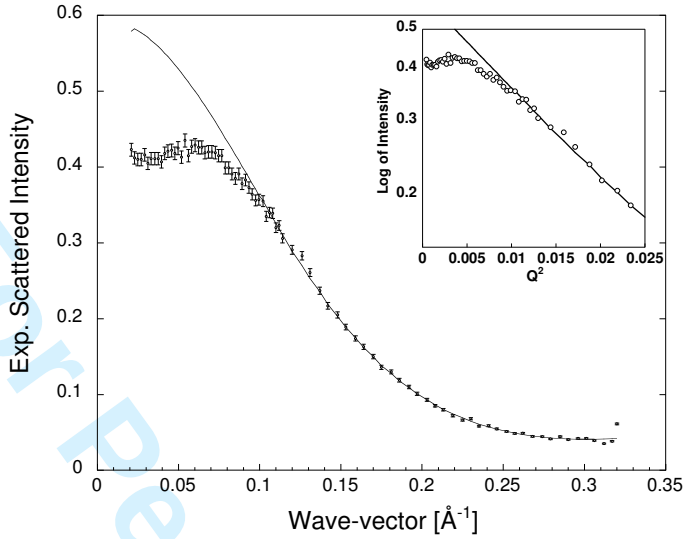
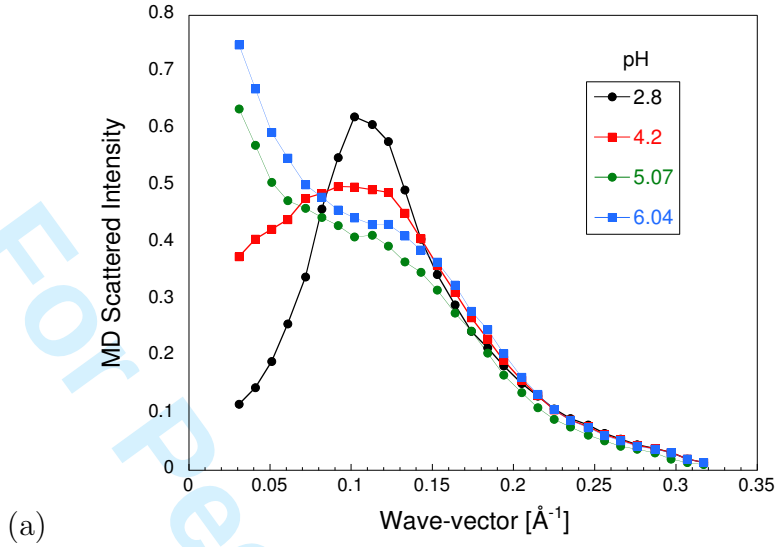
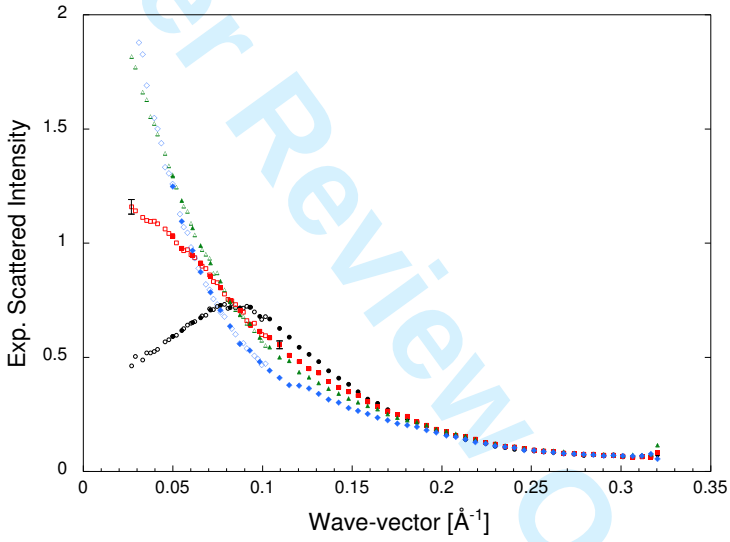


Figure 4:



(a)



(b)

Figure 5: