Biophysical Chemistry 151 (2010) 22–28

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Thermodynamic study of protein phases formation and clustering in model water–protein–salt solutions

Sergey P. Rozhkov *, Andrey S. Goryunov

Institute of Biology, Karelian Research Center RAS, Pushkinskaya 11, 185910, Petrozavodsk, Russia

ARTICLE INFO

Article history:
Received 3 February 2010
Received in revised form 26 March 2010
Accepted 25 April 2010
Available online 4 May 2010

Keywords:
Spinodal
Phase diagram
Dynamic cluster
Metastable dense liquid phase
Solubility
Lysozyme

ABSTRACT

Thermodynamic analysis of the water–protein–salt system, based on the description of the spinodal curve, has been carried out in various coordinate systems: (water chemical potential, protein concentration m_2); (protein "solubility" log S, salt concentration m_3); (effective temperature, critical composition of the system m_2/m_3). Such presentations explain the existence of diagrams with normal and retrograde protein solubility as a result of straightforward effect of ions present in solution as well as some features of the widely used phase diagram in coordinates (temperature, protein concentration). Analytic expressions for coefficients K and b of the salting out equation $\log S = -K \cdot m_3 + b$ as functions of protein charge and protein adsorbed ions have been obtained and identified with the spinodal characteristic points reflecting quasi-equilibrium between protein-lean phase and dense protein-rich phase. Liquid–liquid, liquid–solid phase transitions, dynamic protein clusters and second virial coefficient that characterize interaction between solution components have been thus interrelated. The results of our thermodynamic analysis have been compared with the data reported for lysozyme .

1. Introduction

The growing interest in the phenomenon of formation of dense liquid phases, gels and clusters in protein solutions [1–6] is caused by its relevance for protein solubility and crystallization, drug delivery, certain diseases, industrial separations, cell cytoplasm function. Theoretical consideration of a (L–L) phase transition and accompanying kinetic processes of equilibration in globular protein solutions often lead to the prediction of metastable with respect to crystal protein-rich liquid as well as of various supramolecular protein structures in the high salt concentration range [7], the nonequilibrium kinetically arrested states (gels) among them [2]. Usually the mechanism of dense liquid phase and cluster formation is treated on the basis of the short-range attractive and long-range repulsive protein-protein interaction. Physical approach employing the concept of force field, faces difficulties in describing the effects of salt lyotropic series and anisotropy in protein–protein interactions [8,9] as well as unusual stability of protein solutions in the high salt region [10]. In this connection substantial efforts have been applied to examine the specific effects of ion distribution on the phase interfaces of various polarities [11,12], the protein–electrolyte interaction with regard for anion polarizability [13], their hydration [14] and water structuring effect [15]. Molecular dynamic simulations of protein–salt interaction are also in progress [16,17].

On the other hand, the interaction of proteins between each other, with water and salt ions is related directly to the problems dissolution capability of water and evolution of biological solutions, which are of fundamental importance for physical/chemical biology. Water structure and its specific surface energy are the very factors that determine the solubility of proteins and stability of protein solutions [15,18]. Such an approach is not frequently employed in considering protein dense liquid phase formation and clustering [19,20] although protein solubility is a thermodynamic quantity. Various contributions to the Gibbs (Helmholtz) free energy of the protein–protein interaction cannot be taken into account easily and the role of the solvent free energy in the protein–protein interaction is rarely studied [21]. Thermodynamic approach in studying the solution properties is based on the idea of activity that compensate the lack of information on the structural changes of solution components. The component activity a_i is its effective concentration. It is determined by real concentration m_i due to the activity coefficient γ_i, and in its turn determines the chemical potential of the component $\mu_i - \mu_i^0 = RT\ln a_i = RT\ln\gamma_i m_i$. Osmotic pressure P is an integral parameter of solution: $P = (\mu_1 - \mu_1^0)/V_1$. P is described empirically through the virial series. The second virial coefficient B_2 is characteristic of protein–protein and protein–salt interaction at molecular level. Its positive or negative values indicate the repulsive or attractive potential of interaction between protein molecules, respectively.

Previously [22,23] we used the concentration (m_i) derivatives of the activity coefficient of the components, water (1), protein (2) and salt (3), to evaluate in the Prigogine criterion for phase instability of the ternary solution to fluctuation: $(\partial\mu_1/\partial m_2) = 0$ [24], resulted in the

* Corresponding author. Fax: +7 8142 769810.
E-mail address: rozhkov@krc.karelia.ru (S.P. Rozhkov).

0301-4622/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2010.04.007

spinodal curve for the ternary system. The needed derivatives can be computed by combining the concentration dependence of the chemical potential ($\partial\mu_1/\partial m_3$) of aqueous salt with the required composition derivatives of the chemical potential of protein in aqueous salt ($\partial\mu_2/\partial m_3$), as determined from the classic light scattering and osmotic pressure measurements performed on water + albumin + NaCl by Scatchard and Edsall and their collaborators in the 1950s [25].

Such an approach allows to present the phase diagram of protein solution in various types of coordinates, the coordinates of water chemical potential versus protein concentration $\mu_1(m_2)_{T,m3}$, as an example [26], in order to elucidate some features of the protein–salt solution phase diagram in commonly used coordinates temperature–protein concentration $T(m_2)_{m3}$, which are only partially thermodynamically understood. Among them are the metastability of the L–L phase equilibrium with respect to solution and to crystal [3,27]; the effect of salts on the temperature of L–L equilibrium [15]; the existence of proteins with normal and retrograde solubility [5,21]; the frustration of aggregation (clustering) processes for some proteins by the addition of salt in the molar ionic strength range [10]; the correlation between the solubility, crystallization and the second osmotic virial coefficient [28–30]; the mechanism of gelation in the spinodal region [2]; the existence of two types of protein clusters of sizes ~10 nm and of sizes ~100 nm in the gap between the liquidus and the L–L binodal in low salt lysozyme solutions [31,32]. Stradner et al. assume the lysozyme clusters of sizes ~10 nm are of an equilibrium character [33], whereas Sukla et al. find they are not evident at all [34]. Most probably the clusters are of dynamic nature with a limited life time [3,27,35]. Protein clusters are supposed [27] not to comply with Gibbs's definition of phase, their sizes being determined by the kinetics, and not by thermodynamics of growth and decay.

In this work various types of phase diagrams and methods of their interconversion reported previously in part [23,36,37], are discussed in order to explain the above mentioned thermodynamic features of the phase diagram. We are aiming mainly at an analysis of the theoretical phase diagrams of model water–protein–salt solutions in the presence of salts with a parabolic dependence of the salt activity coefficient on concentration. Strong hydration of these salts causes a decrease of their activity coefficient on dissolution at low salt and an increase of activity at high salt, i.e. the parabolic dependence of salt activity coefficient on concentration. The rise of activity at high salt indicates a decrease of affinity to solvent and a tendency to form ion pairs. This could favor ion adsorption in heterogeneous systems. Among these salts are physiologically important calcium chloride, magnesium chloride, and even sodium chloride at high concentration. The type of the dependence comes into importance within the higher (high salt) Debye-Huckel approximation also used in our approach.

2. Theory

2.1. Phase diagram in the (μ_1, m_2)-plane

The equation we obtained previously for the characteristic coefficient of stability μ_{12} of the water–protein–salt system was of the form [22,23]:

$$\mu_{12} = (\partial\mu_1/\partial m_2)_{T,P,m3} = (\partial\mu_2/\partial m_3)(\partial\mu_1/\partial m_3)(\partial\mu_3/\partial m_3)^{-1}, \quad (1)$$

where m_1, m_2 and m_3 are the molar concentrations of solution components water, protein, salt, accordingly; μ_1, μ_2, μ_3 are the components chemical potential. This equation can be derived by taking the stability determinant and its principal minors equal to zero: $|\partial^2 G/\partial m_i \partial m_j| = 0$ and $\mu_{11}\mu_{22} - \mu^2{}_{12} = 0$; $\mu_{11}\mu_{33} - \mu^2{}_{13} = 0$; $\mu_{22}\mu_{33} - \mu^2{}_{23} = 0$ [24]. Expressions for derivative chemical potentials for corresponding components can be obtained from Scatchard's thermodynamic theory of multicomponent systems [25]: $\partial\mu_2/\partial m_3 = RT(-z^2 m_2/2m_3^2\varepsilon + \beta_{23})$, $\partial\mu_3/\partial m_3 = RT(2/m_3\varepsilon + \beta_{33})$, $\partial\mu_1/\partial m_3 = -(2RT/m_1)(1 + m_3\beta_{33})$ [23].

where $\varepsilon = 1-(zm_2/2m_3)^2$, and $\beta_{23} = (-\nu/2m_3)(2 + \beta_{33}m_3)$.

Here z is the charge number of biopolymer, obtained from titration curve, $(-\nu)$ is the number of bound anions, $\beta_{33} = \partial \ln\gamma_3/\partial m_3$, and γ_3 is the salt activity coefficient. Correspondingly $\beta_{33}m_3$ may be presented as [23]:

$$m_3\beta_{33} \approx -m_3 \frac{\partial}{\partial m_3}\left(\frac{A\zeta^{1/2}}{1 + r\kappa\zeta^{1/2}} - \Sigma\alpha_i\zeta^i\right) \equiv \Delta. \quad (2)$$

In Eq. (2) Debye-Huckel equation allowing for high salt concentration appears under the derivative sign, where A, r, κ are the Debye-Huckel coefficients, ζ is the ionic strength, α is the experimental constant. The range of higher salt concentrations is described using various, mostly empirical approximations of the theory. One of the approaches is to use the experimental correction factors (coefficients α_i in Eq. (2)). So, Δ (m_3, temperature) is the function that characterizes the dependence of electrolyte activity on its concentration in terms of Debye-Huckel theory. The difference of salts in the dependence of their activity coefficient on concentration results in the fact that parameter Δ can take both negative ($\gamma_\pm$ decreases at low salt concentration) and positive ($\gamma_\pm$ grows up at high concentration) values for the salts NaCl, $CaCl_2$, and $MgCl_2$ with parabolic type dependence of activity coefficient. The condition $\Delta = 0$, corresponds to the minimum activity coefficient and determines an important point at the phase diagram.

By rearranging Eq. (1) can be given in an explicit form [23]:

$$\left(\frac{\partial\mu_1}{\partial m_2}\right) = \frac{\left(m_2^2 - \frac{4m_3}{\nu(2+\Delta)}m_2 - \frac{4m_3^2}{z^2}\right)}{4m_3^2(2+\Delta) - \Delta\cdot z^2 m_2^2}(1+\Delta) = 0. \quad (3)$$

Two factors appear in this expression and the analysis of the condition $(1+\Delta) = 0$ is beyond the scope of this study.

The first factor when taken equal to zero gives the equation of spinodal in coordinates (m_2,m_3) [23]:

$$\left(\frac{m_2}{m_3}\right) = 2\frac{1 + \sqrt{1 + \frac{(2+\Delta)^2\nu^2}{Z^2}}}{\nu(2+\Delta)}. \quad (4)$$

The condition $(\partial^2\mu_1/\partial m_2^2)_T = 0$ indicates that Maxwell's rule for phase transition is satisfied [24]. It also gives the equation for nodes in the same coordinates. The nodes are two points on two binodal branches. These points correspond to the same temperature and two different compositions of two coexisting phases. In the present context the node is the point on the spinodal, given as follows [23]:

$$\left(\frac{m_2}{m_3}\right) = 2\frac{\nu(2+\Delta)}{\Delta z^2}\left(1 \pm \sqrt{1 - \frac{\Delta z^2}{(2+\Delta)\nu^2}}\right). \quad (5)$$

Eq. (1) or (3) gives the equation of the stability boundary of solid phase F when $(\partial\mu_1/\partial m_2) \to \infty$ [23,38]:

$$\left(\frac{m_2}{m_3}\right) = \frac{2}{z}\sqrt{\frac{2+\Delta}{\Delta}}. \quad (6)$$

The integral form of Eq. (1) or (3): $\mu_1 - \mu_1^0 = \int(\partial\mu_1/\partial m_2)dm_2$ gives an expression for the curves on the phase diagram in coordinates (water chemical potential, protein concentration) that discriminate the thermodynamically stable (above the curve) and unstable (beneath the curve) zones with respect to protein concentration fluctuations [23] (Fig. 1).

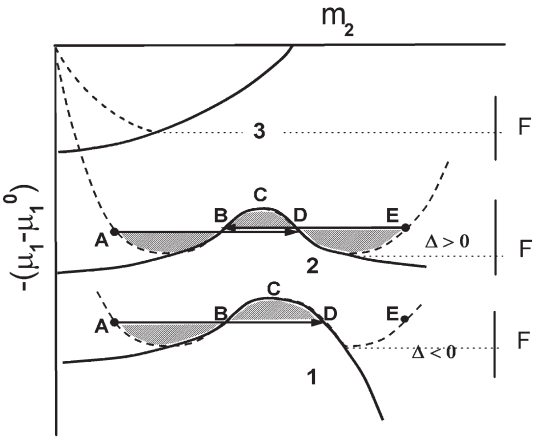

Fig. 1. Schematic presentation of three types of stability boundaries in (water chemical potential μ_1, protein concentration m_2)-plane for different salt concentrations: 1. $\Delta<0$. Boundary is characterized by one inflection point (B) and one critical point (C). Single metastable state A↔D (quasi-equilibrium protein solution ↔ dense liquid phase) is the case. 2. $\Delta>0$. Boundary is characterized by two inflection points (B) and (D) and one critical point (C). A↔D and E↔B states metastable with respect to crystal and to solution and representing continuous fluctuation phase transition are the case. 3. $\Delta>2\nu^2/(z^2-\nu^2)$. No characteristic points present.

Using two terms of the virial series for osmotic pressure Timasheff and coworkers have shown [26] that the water chemical potential first decreases (left dashed lines on Fig. 1) with rising protein concentration at constant temperature and salt concentration, but later increases. Parabola crosses the stability boundary (solid lines on Fig. 1) where the chemical potential starts rising, and the solution stability lowers. Fig. 1 shows schematically that the isotherm $\mu_1(m_2)_{m3}$ can be presented as a Van der Waals loop with a critical point. The practice of analysis of the van der Waals isotherms shows that the ascending segments of isotherm do not perform. Thus dense metastable protein phase of concentration corresponding to the point D start forming at point A of the isotherm. The condition for that is also the equality of shaded areas (Fig. 1, curve 1,2) in accordance with the Maxwell equal-area rule that takes into account the energy expenses of phase formation. Such phase formation or L–L phase separation is a 1st order phase transition. The water chemical potential remains unaltered provided that the Maxwell's rule is satisfied.

The points A and D lie on the descending part of the isotherm. Their energy is somewhat higher with respect to the point F that corresponds to the solid state, and they are separated with the barrier as well. This means that the dense phase appear to be metastable with respect to solid state and stable with respect to solution and allows for the formation of crystal nucleus in the phase.

The total set of points A,D and B,E makes binodal line representing quasi equilibrium of the corresponding phases (Fig.1). Binodal line cannot be derived analytically [24], but it can be determined experimentally. On the other hand, the points B and D taken as middle points of the corresponding node lines can be used to characterize binodal position on the phase diagram. The curves formed by the inflection points of the stability boundaries (points B and D at $\Delta>0$ and point B at $\Delta<0$ on Fig. 1) can be derived from Eq. (1) under the condition $(\partial^2\mu_1/\partial m_2^2)_T=0$ and corresponds to Eq. (5).

As follows from Eq. 5 there can be one or two inflection points depending on the sign of parameter Δ (Fig. 1). The condition $\Delta>0$ (two symmetrical inflection points) gives rise for the continuous critical type phase transition. The protein solution in this case corresponds to the continuous conversion of the system with (A↔D) and (E↔B) type fluctuations in the interval of intermediate protein concentrations. Such dynamic clusters are metastable both with respect to crystal and with respect to solution and their life time is limited.

2.2. Relationship between temperature and critical composition

The critical point C on Fig. 1 can be reached both by temperature and composition changes m_2/m_3. The parameter Δ is related to temperature due to the coefficients κ and A of Debye-Huckel equation, where the dominating contribution is given by $A\sim(\varepsilon T)^{-3/2}$. Here ε is the dielectric permittivity. The relationship between temperature and the system critical composition $X=(m_2/m_3)_{cr}$ at point C and points B and D on Fig. 1 can be obtained using the expression for the parameter Δ from Eqs. (4) and (5):

$$\pm 1/\Delta = (4-X^2z^2)/(2X^2z^2-4Xz^2/\nu-8) \tag{7}$$

$$\pm 1/\Delta = \left(X^2z^2-4X\nu+4\right)/(8X\nu-8). \tag{8}$$

Fig. 2 shows qualitatively the relationship between the critical parameter $(m_2/m_3)_{cr}$ and effective temperature: $|1/\Delta|\sim(\varepsilon T)^{3/2}$. Given the biopolymer concentration m_2 (sufficient to bring the system to the spinodal region), the curves 1 and 2 described by Eq. (7) in Fig. 2 correspond to low (to the right of point $\Delta=0$) and high (to the left of point $\Delta=0$) salt concentration. The curve 3 (Eq. (8)) gives the relationship of temperature and critical composition for nodes. From Fig. 2, it follows that the critical point temperature rises with increasing electrolyte concentration at $\Delta<0$ (curve 1), while the greater electrolyte concentration at $\Delta>0$ results in the decrease of the temperature of the critical point (curve 2). Right and left curves (1,2) approach asymptotically the extreme temperature and reach it at $\Delta=0$. This means that the rise of salt concentration can result in alteration of the trend of cloud-point temperature dependence when the type of the dependence of salt activity coefficient on salt concentration is parabolic.

The curve 3 intersects with the curve 1 at critical composition $(m_2/m_3)_{cr}\equiv Q$ and $\Delta<0$, i.e. at low salt concentration. The temperature of node appears to exceed the critical point temperature at $(m_2/m_3)_{cr}>Q$, whereas it falls short of critical temperature at $(m_2/m_3)_{cr}<Q$. This suggests the temperature inversion of binodal and spinodal; that is the system with the lower critical temperature of solubility at low salt concentrations (or high protein concentration) changes into the

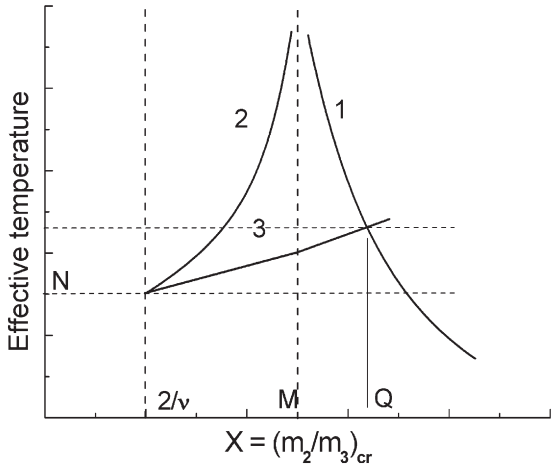

Fig. 2. Relationship of the critical composition $X=(m_2/m_3)_{cr}$ of the system water–protein–salt with the effective temperature. Salt concentration is an independent variable. Curve 1 is for low salt concentrations, $\Delta<0$; curve 2 is for high salt concentrations, $\Delta>0$. Curves 1,2 characterize critical point and curve 3 characterizes line of nodes on spinodal (binodal). $(m_2/m_3)_{cr}\equiv M$ corresponds to maximum temperature of the critical point. Condition $(z^2-\nu^2)/2\nu^2\equiv N$ corresponds to $(m_2/m_3)_{cr}=2/\nu$ and temperature of the triple point. The system with lower critical temperature of solubility changes into the system with the upper critical temperature at $(m_2/m_3)_{cr}=Q$.

system with the upper critical temperature of solubility at high salt concentrations.

The greater the ratio z/ν, the higher the salt concentration of the curves intersection.

Curve 3 (nodes) and critical point curve 2 of Fig. 2 descend monotonously with rising salt concentration and meet at $m_2/m_3 = 2/\nu$ and effective temperature that satisfies the condition $1/\Delta = (z^2-\nu^2)/2\nu^2$. This is a triple point, i.e. the point of thermodynamic equilibrium of all the solution phases: protein-lean, dense liquid and crystal. Consider the range of the second virial coefficient values corresponding to such point. There is thermodynamic expression for interaction coefficient B_2 as follows [25]:

$$B_2 = \frac{1000}{2M_2^2}\left(\frac{z^2}{2m_3} + \beta_{22} - \frac{\beta_{23}^2}{2+\beta_{33}m_3}m_3\right). \quad (9)$$

The first term in the brackets is the Donnan term, β_{22} is the protein activity coefficient, which is positive for asymmetric molecules. β_{22} can be neglected for strongly charged molecules and strong protein–salt interaction. The third term is characteristic of protein interaction with electrolyte and its value is negative. Hence, the negative sign of B_2 reflects primarily the protein–salt attractive interaction. Here $\beta_{23} = (-\nu/2\ m_3)(2+\beta_{33}m_3)$ [25], and $\beta_{33}m_3 = \Delta$. At the triple point, where spinodal transforms into straight line, merges node curve and meets the solid phase boundary, the condition $1/\Delta = (z^2-\nu^2)/2\nu^2$ is satisfied. Then, neglect of β_{22} and substitution of corresponding expressions into Eq. (9) give:

$$B_2 = \frac{1000}{4M_2^2 m_3}\cdot\left[z^2\cdot\frac{(z^2-2\nu^2)}{z^2-\nu^2}\right]. \quad (10)$$

As the liquid–solid phase transition is approached, the interaction coefficient is thus seen to be negative, and changes within narrow limits determined by the parameters variations: $\nu^2 < z^2 < 2\nu^2$.

2.3. Phase diagram in the (log S, m_3)-plane

Eqs. (4)–(6) can be rearranged into a more convenient form that permits presenting data in coordinates (protein solubility, salt concentration). Protein solubility S in salt solutions can be qualitatively described in the form of Debye-Huckel equation [39]:

$$\log(S) = \frac{A\zeta^{1/2}}{1+kr\zeta^{1/2}} - \sum\alpha_i\zeta^i.$$

This allows formal converting of Eqs. (4)–(6) into solubility coordinates. Eq. (2) can be transformed as follows:

$$\log S = -\int\frac{\Delta}{m_3}dm_3 + const \quad (11)$$

The parameter Δ can easily be expressed from Eqs. (4)–(6) as a function of m_3 (see Eqs. (7) and (8)).We consider the behavior of the nodes (Eq. (5)) which takes the following analytic form on the assumption that $\nu^2 < z^2$ [36]:

$$\log S = ln\left(m_3^2 - m_2\nu m_3 + \frac{m_2^2 z^2}{4}\right) - \frac{2}{\sqrt{\frac{z^2}{\nu^2}-1}}arctg\frac{2m_3 - m_2\nu}{m_2\nu\sqrt{\frac{z^2}{\nu^2}-1}} + const. \quad (12)$$

Eq. (12) assumes almost linear descending dependence of log S on electrolyte concentration and presents in a sense the equation of protein salting-out by electrolyte. Eq. (12) can be presented in the form of the dependence of the dimensionless protein solubility on the dimensionless salt concentration in the vicinity of the critical point, where $m_3 = m_2\nu/n$, $n>2$ are the rational positive numbers, and $\nu^2 < z^2$. Then, ignoring small terms $O(1/n^2)$ one obtains:

$$\log S - ln(m_2^2) + const \approx -\frac{4\nu^2(2z^2-\nu^2)}{z^2(z^2-\nu^2)}\cdot\frac{1}{n} + \frac{2\nu^2}{z^2-\nu^2} + ln\left(\frac{z^2}{4}\right) \approx K_s\left(\frac{1}{n}\right) + b. \quad (13)$$

Thus, Eq. (13) has been set to the salting-out curve in the form given by Setchenov equation, where 1/n is a dimensionless value, analogous to the ionic strength. (In what follows below we discuss the relationship between 1/n and salt concentration).K_s determines the slope of extrapolation straight line in the range of salting-out and characterizes the salting-out efficiency of the salt:

$$K_s = -\frac{4\nu^2(2z^2-\nu^2)}{z^2(z^2-\nu^2)} \quad (14)$$

b is the segment intercepted on the solubility axis by the extrapolation of the linear section of the curve to $m_3 = 0$ (apparent solubility in the absence of electrolyte):

$$b = \frac{2\nu^2}{(z^2-\nu^2)} + ln\left(\frac{z^2}{4}\right) \quad (15)$$

The term $\ln(m_2^2)$ accounts for the protein concentration in the metastable dense liquid phase at quasi-equilibrium with diluted phase where protein concentration is S.

2.4. Verification of Eq. (13) using solubility diagram for lysozyme

Fig. 3 presents the comparison of log S calculated theoretically with the use of Eq. (13) with experimental dependence of log S on NaCl concentration at various pH (protein charge z) taken from the experimental results of Retailleau, Ries-Kautt and Ducruix [40]. The relationship between m_3 and 1/n can be obtained on the basis of Eq. (13) by solving the transcendent equation

$$2ln\left(\frac{m_3 n}{\nu}\right) - \frac{4\nu^2\left(2z^2-\nu^2\right)}{z^2\left(z^2-\nu^2\right)}\cdot\frac{1}{n} = -\frac{2\nu^2}{z^2-\nu^2} - ln\left(\frac{z^2}{4}\right).$$

Lysozyme is known to bind no fewer than 4–5 Cl^- ions, and even in salt-less solutions lysozyme is the salt Prot nCl^- [40]. According to Eq. (13) this means no salting-in effect known for proteins, which

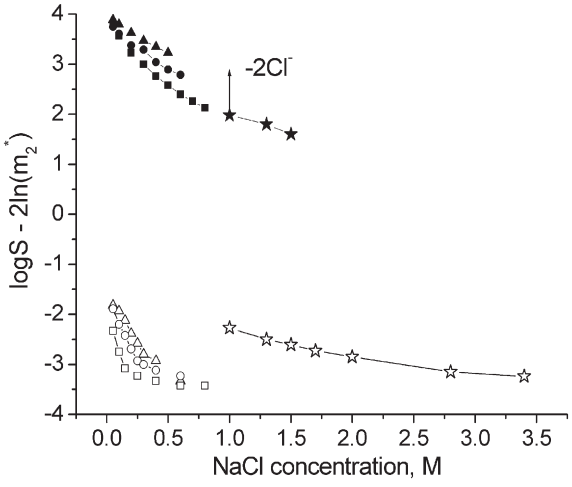

Fig. 3. Theoretical (closed symbols) and experimental (open symbols) values of lysozyme solubility as a function of NaCl concentration at pH 7 (z = 8, closed squares), pH 5 (z = 10, closed circles), pH 4 (z = 11, closed triangles), pH 4.3 (open triangles), pH 6.5 (open circles), pH 8.4 (open squares). Theoretical values for NaCl concentration up to 1 M are calculated at $\nu = 5$; for NaCl concentration over 1 M ν was taken equal 6 and 7. Experimental values are converted from [40] and [2].

increases anion binding with rising salt concentration (increasing ν). Theoretical values of solubility in Fig. 3 exceed the experimental data by $\ln(m_2^2)$ + const. Here, the constant allows for the fact that the node is on the spinodal and not on the binodal. For many of the protein systems, the classical spinodal curve is quite close to the binodal, so the constant may be taken equal to zero at a first approximation. From Fig. 3 it follows that m_2 equals 0.05 M (or 700 g/l) in that case and corresponds to the molar protein concentration in the dense liquid. The general run of the experimental curves is consistent with theoretical dependencies.

The data of Retailleau et.al. on solubility [40] have been obtained in equilibrium with crystal and not in salting-out conditions. In Fig. 3 are also shown the data on "solubility" obtained using the "cloud point" or salting-out method (from Fig. 2 of Dumetz et al. [2]). It is seen that the "salting-out" values are higher. The theoretical solubility jump analogous to that observed on experimental dependences can be obtained on assumption that the formation of dense liquid phase indicated by solution turbidity is accompanied by partial exclusion of chlorine ions from proteins in this dense liquid phase.

2.5. Second virial coefficient for lysozyme

The sign and value of the second virial coefficient B of osmotic pressure (interaction coefficient) has been experimentally approved as a convenient criterion related to protein crystallization [28–30]. The coefficient has been shown negative for proteins under crystallization conditions and belongs to narrow limits from zero to $-8\cdot10^{-4}$ M ml g^{-2} [30]. Table 1 presents the values of B_2 calculated with the use of Eq. (10) for lysozyme-type protein ($M_2 = 14$ kD) at NaCl concentrations 2 and 7% and a number of z values, taken at corresponding pH values from Fig. 1 of Ref. [41]. The values of ν (usually anions of 1:1 electrolyte) have been taken arbitrarily to fit the condition $\nu^2 < z^2 < 2\nu^2$ on the assumption that ν increases with rising NaCl concentration. From Table 1 it is seen that the values calculated enter the relatively narrow interval from -0.6 to $-5\cdot10^{-4}$ M ml g^{-2}. They agree closely with the experimental values for lysozyme in corresponding buffer solution (appear bracketed in Table 1).

3. Discussion

3.1. Phase diagram in the (T, m_2)-plane

The solubility phase diagram can be discussed in the commonly used (temperature, protein concentration)-plane with salt concentration as a parameter using temperature dependencies of critical composition of Fig. 2. In Fig. 4 the liquidus line (1) and the solidus line (2) are shown. Hence, A and B reflect protein crystal in equilibrium with supernatant liquid. Solid line 3 corresponds to the binodal, where S and m_2^* are the protein lean phase and the metastable dense liquid phase. The metastability with respect to crystal phase is caused by the energy barrier (Fig. 1). Nucleation and crystallization in dense liquid phase are caused by overcoming the energy barrier.

Experimentally measured parameter of the system phase equilibrium is binodal. Spinodal can only be determined theoretically, as was done in this study. Spinodal and binodal are similar in shape and the spinodal is embedded into the binodal. Their maxima coincide and form the only common point. Line 4 is the spinodal and CP is the critical point.

Table 1
Interaction coefficient B calculated for lysozyme.

Protein charge, Z	2% NaCl		7% NaCl	
	Number of Cl$^-$, ν	B·10^4, M·ml·g^{-2}	Number of Cl$^-$, ν	B·10^4, M·ml·g^{-2}
9 (pH 5.4)	7	−1.9 (−2.8)	8	−2.4
10 (pH 5.0)	8	−3.0 (−2.9)	9	−3.5
11 (pH 4.6)	8	−0.6 (−1.7)	10	−4.9
13 (pH 4.0)	10	−2.9 (0.1)	12	−5.0 (−7.4)

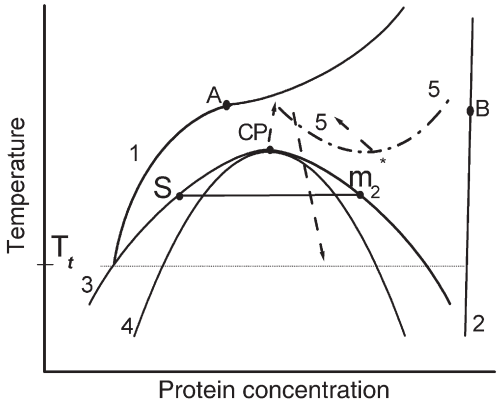

Fig. 4. Schematic phase diagram in (temperature, protein concentration)-plane. Curve 1 is the liquidus line, curve 2 is the solidus line, curves 3 and 4 are binodal and spinodal accordingly for the system with the upper critical temperature of solubility, curve 5 is spinodal for the system with the lower critical temperature of solubility. (A–B) is crystal — supernatant equilibrium. S and m_2^* are metastable protein-lean phase and dense liquid phase. CP is the critical point. Arrows show the shift of spinodal and its extremum CP with rising salt concentration; T_t is the temperature of the triple point.

The zone above the critical point CP is homogeneous. At the same time, it is theorized that critical and supercritical phase transitions cannot be completed in one point [38]. Because of this, fluctuation protein associates (dynamic clusters) may persist in supercritical zone. Thermodynamic prove for the their existence is presented in Fig. 1, curve 2. It follows that near-critical phase transitions can be metastable both with respect to crystal, and with respect to solution.

The dashed arrows show the rise of the critical point temperature (spinodal maximum) with rising salt concentration at $\Delta<0$, but the critical temperature lowers at $\Delta>0$ and approaches the triple point. In the latter case the second virial coefficient B_2 that characterizes protein–salt interaction (Eq. (10)), takes negative values. Experimental values for cloud-point temperatures for lysozyme in $CaCl_2$ and $MgCl_2$ solutions are in good agreement with such trend [15] but not in NaCl solutions. The possible reason is that the activity coefficient of the first two salts is minimum at concentration of 0.4 to 0.6 M, whereas the minimum of activity coefficient for NaCl can hardly be distinguished within the range of 1.5–2 M and the effect is difficult to reveal.

Dot-dashed line 5 in Fig. 4 is spinodal at very low salt concentration (or high protein concentration) in the range of m_2/m_3 values to the right of the intersection point (Q) of the lines 1 and 3 in Fig. 2. By this is meant that in theory the equilibrium protein dense liquid phase could likely have formed in that low-salt (or high protein) region of phase diagram on rising temperature, and this is an example of retrograde solubility [5]. Homogeneous solution forms here on lowering temperature. However, one can expect the existence of dynamic clusters in that supercritical zone as well.

Dynamic clusters in the systems with lower (LCPS) and upper (UCPS) critical points of solubility may differ in size and density, as the reasons and the mechanisms of phase separation in these systems are different. In the system with LCPS the thermodynamic functions of mixing are negative $\Delta H_{mix}<0$ and $\Delta S_{mix}<0$ [42] and functions of homogenizing at lowering temperature are known to be caused by cluster nature of water and hydrophobic interaction that condition the semiclathrate hydration preventing the association of protein molecules. In this connection, the dynamic protein cluster should be compact and not large.

Various aggregates, oligomers, associates widely reported at low salt concentrations [27,31–33] can probably be explained by such clusters.

$\Delta H_{mix} > 0$ and $\Delta S_{mix} > 0$ in the system with UCPS [42] that is caused by the mechanism of preferential hydration and leads to the expulsion of salt ions from the protein hydration shells and association of those shells. Protein molecules coalescence in one phase leads in this case to rise of entropy in diluted phase and favors the phase separation at lowering temperature. Thus, protein clusters can be larger and less dense.

From the preceding consideration and temperature dependences of critical points at Fig. 4, it may be speculated that gradual increase of the salt concentration can lead to "melting" of low-salt clusters and formation of less dense and larger high-salt clusters. This can explain the existence of two types of clusters of ~10 and ~100 nm size at low salt concentrations [31,32] and dissociation of protein aggregates induced by addition of salt [10]. Indirect evidence for that was presented in our paper on serum albumin [43].

3.2. Salting out and Setchenov equation

Salting-out is one of the well known methods of protein isolation and crystallization. At the same time, this method is empirical in many respects because of the difficulties of theoretical consideration of about 20 variable parameters affecting the thermodynamic and aggregation stability of protein solution.

Effect of protein salting out by electrolyte is described experimentally by Setchenov equation $\log S = b\text{-}Km$, where m is the salt concentration (ionic strength), b and K are the empiric (fitting) coefficients, which, as is believed, characterize protein hydration and solvent surface tension and have not yet been derived analytically [44]. The value of S is often thought of as solubility since protein crystals can sometimes be obtained during protein salting out.

The results of the above analysis of the phase diagrams of globular protein solutions and Eq. (13) call into question the relationship between the term "solubility" and the value of S. The latter characterizes the protein concentration in diluted phase in equilibrium with dense liquid phase of protein concentration (m_2^*), which is metastable with respect to crystal. That is why the magnitude of S is always higher than the true solubility (curve 1 in Fig. 4). Protein concentration can be an order of magnitude higher in the dense liquid phase (m_2^*) than in diluted phase (S).

Protein concentration in dense liquid phase can be calculated at certain protein charge, number of associated salt anions and protein concentration of its salting out, using Eq. (13) that characterizes the relationship between two protein concentrations S and m_2^*. Sometimes protein concentration in dense liquid phase can be evaluated using centrifugation which makes the droplets of dense liquid phase to form macrophase [1,2]. When protein concentration in dense liquid phase is known, the number of protein associated electrolyte ions can be calculated. For example, the data on the dependence of lysozyme concentration in condensed phase on NaCl concentration (Fig. 11 C in [2]) can be used to assess the value of ν for that protein. In this case, direct Eq. (12) gives 4.5 Cl^- ions at 1 M NaCl, 6.5 at 2 M, and 6.9 at 3 M in good agreement with the existing data [40].

3.3. Solubility and protein–salt interaction

Kinetics of the protein phases formation may differ depending on the way the system reaches the region of spinodal and binodal. The rise of thermodynamically unstable and kinetically stabilized state of gel is plausible on that way [2]. Still, the mechanism of protein gel formation is not clear.

By virtue of Eq. (13), the greater the protein charge z and the ν value, the higher the S at low electrolyte concentrations m_3. This can be accounted for by a rise of osmotic pressure in water–protein matrix due to hydration of adsorbed ions. Surface tension at protein–solvent interface decreases and the conditions discussed in [6,23] for protein cluster formation disappear. This can understand the physical nature of the salting-in effect. Protein solubility increases as the sites accessible for ion adsorption are filled. The salting-in effects do no arise, when the sites are always filled, i.e. the protein is in the form of salt Prot nCl^- even in salt-less solutions. In this case protein aggregates (oligomers) do not necessarily form at low salt concentrations as well.

In the range of high electrolyte concentrations m_3, the solubility is decreased due to the increasing contribution of the term $K_s(1/n)$, which compensate by far the coefficient b in Eq. (13). The true meaning of the salting-out coefficient has thus far remained uncertain for the general case. There are assumptions of its relation to protein hydration and/or protein–solvent surface tension [44]. The complex combination of protein charges and adsorbed ions in the expression for the coefficient may imply that the adsorbed ions compete for water with protein charged groups and with ions in bulk solvent reducing the thermodynamic stability of the system. That results in the unfavorable effects of preferential hydration [26], general increase of the specific surface energy and the system phase destabilization. On the other hand, protein association into dense liquid phase or large cluster displaces ions from the surface of protein–protein contact where the decrease of dielectric permittivity causes the essential rise of the electrostatic energy of adsorbed ions. The tightly bound ions (heavy metal ions, for example) remaining in the protein structure during phase formation, can cause significant denaturation and abnormal aggregation or polymerization of proteins in the dense liquid phase. Probably this can cause protein gel formation. Alternatively, nucleation of protein polymers as a model of amyloid fibril formation has been shown to follow two-step mechanism, where metastable dense liquid clusters appear as nuclei precursors [45]. Interaction with Cl^- and other anions is characteristic of globular proteins [46]. The magnitude of ν depends on m_3, ion activity, charge z and the number of sites of ion association. Interaction of ions with nonpolar surface segments of protein depends on their polarizability in many aspects [13]. Polarizability along with size, hydration energy, etc. allows arranging anions in Hofmeister series according to their effect on protein solubility.

4. Conclusions

This thermodynamic approach based on the Prigogine criterion of ternary system water–protein–salt stability to the diffusion, Scatchard thermodynamic theory for multicomponent systems, theory of continuous critical and subcritical phase transitions allows to reveal several types of protein-lean and protein-rich phases and dynamic clusters on a solubility phase diagrams shown in different coordinates. The study of the problem presented provides a thermodynamic basis for understanding that metastability of the L–L phase equilibrium with respect not only to the crystal, but also to the solution, causes the existence of protein clusters in the gap between the liquidus and the L–L binodal. Behavior of the salt activity coefficient have been shown an important factor in the formation of dynamic protein clusters.

It has been found that water–salt protein solution can have upper and lower critical temperatures of solubility depending on the concentration ratio of its components and temperature. This is suggestive of two types of protein dense liquid phases corresponding to normal and retrograde solubility and dynamic clusters in subcritical region between liquidus and solidus lines: low-salt aggregates (oligomers) and high-salt dense liquid clusters of submicron size.

The formation of high-salt dense liquid phase as a beginning of protein salting out can be described using the linear salting-out equation representing relationship between protein concentration in diluted phase and in dense phase. Parameters of the analytically obtained equation depend on the protein net charge and the number of electrolyte ions adsorbed. Given the protein concentration in dense liquid phase and in diluted phase one can assess the number of ions adsorbed by protein using the equation obtained. The relation between protein charge and number of ions adsorbed also determines the

position of triple point and the numerical value of the second virial coefficient characteristic of protein–salt interaction.

Acknowledgment

We gratefully acknowledge RFBI for financial support N 08-04-98825.

References

[1] N. Asherie, Protein crystallization and phase diagrams, Methods 34 (2004) 266–272.
[2] A.C. Dumetz, A.M. Chockla, E.W. Kaler, A.M. Lenhoff, Protein phase behavior in aqueous solutions: crystallization, liquid–liquid phase separation, gels, and aggregates, Biophys.J. 94 (2008) 570–583 M.
[3] O. Glico, W. Pan, P. Katsonis, N. Neumaier, O. Galkin, S. Weinkauf, P.G. Vekilov, Metastable liquid clusters in super and undersaturated protein solutions, J.Phys. Chem.B 111 (2007) 3106–3114.
[4] J.D. Gunton, A. Shiryayev, D.L. Pagan, Protein condensation, Cambridge University Press, 2007.
[5] M. Muschol, F. Rozenberger, Liquid–liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization, J. Chem. Phys. 107 (1997) 1953–1962.
[6] S.P. Rozhkov, Three-component system water-biopolymers-ions as a model of molecular mechanisms of osmotic homeostasis, Biophysics 46 (2001) 53–59.
[7] Y. Georgalis, P. Umbach, D.M. Soumpasis, W. Saenger, Dynamic and microstructure formation during nucleation of lysozyme solutions, J.Am.Chem.Soc. 120 (1998) 5539–5548.
[8] A. Tardieu, A. Le Verge, M. Malfois, F. Bonnete, S. Finet, M. Ries-Kautt, L. Belloni, Proteins in solution: from X-ray scattering intensities to interaction potentials, J. Cryst. Growth 196 (1999) 193–203.
[9] A. Lomakin, N. Asherie, G.B. Benedek, Aeolotropic interactions of globular proteins, Proc.Natl.Acad.Sci. USA 96 (1999) 9465–9468.
[10] R. Piazza, Protein interactions and association: an open challenge for colloid science, Curr.Opin.Colloid Interface Sci. 8 (2004) 515–522.
[11] H. Wennerstrom, Ion binding to interfaces, Curr.Opin, Colloid Interface Sci. 9 (1–2) (2004) 163–164.
[12] G. Luo, S. Malkova, J. Yoon, D.G. Schultz, B. Lin, M. Meron, I. Benjamin, P. Vanysek, M.L. Schlossman, Ion distribution near a liquid–liquid interface, Science 311 (2006) 216–218.
[13] M. Bostrom, F.W. Tavares, B.W. Ninham, J.M. Prausnitz, Effect of salt identity on the phase diagram for a globular protein in aqueous electrolyte solution, J. Phys. Chem. B 110 (2006) 24757–24760.
[14] K.D. Collins, Ion hydration: Implications for cellular function, poltelectrolytes, and protein crystallization, Biophys.Chem. 119 (2006) 271–281.
[15] J.J. Grigsby, H.W. Blanch, J.M. Prausnitz, Cloud-point temperatures for lysozyme in electrolyte solutions: effect of salt type, salt concentration and pH, Biophys.Chem. 91 (2001) 231–243.
[16] D. Horinek, R.R. Netz, Specific ion adsorption at hydrophobic solid surfaces, Phys. Rev.Lett. 99 (2007) 226104.
[17] M. Lund, P. Jungwirth, C.E. Woodward, Ion specific protein assembly and hydrophobic surface forces, Phys.Rev.Lett. 100 (2008) 258105.
[18] W. Melander, C. Horvath, Salt effects on hydrophobic interactions in precipitation and chromatography of proteins: An interpretation of the lyotropic series, Arch. Biochem.Biophys. 183 (1977) 200–215.
[19] R. Tuinier, G.J. Fleer, Critical endpoint and analytical phase diagram of attractive Hard-core Yukawa spheres, J. Phys. Chem. B 110 (2006) 20540–20545.
[20] C. Haas, J. Drenth, The protein–water phase diagram and the growth of protein crystals from aqueous solution, J. Phys. Chem. B. 102 (1998) 4226–4232.
[21] A. Shiryayev, D.L. Pagan, J.D. Gunton, D.S. Rhen, A. Saxena, T. Lookman, Role of solvent for globular proteins in solution, J. Chem. Phys. 122 (2005) 234911.
[22] S.P. Rozhkov, Spinodal in concentrated water–salt solutions of serum albumin molecules, Zh.fiz.khim. (in russian) 62 (1988) 1925–1928.
[23] S.P. Rozhkov, Phase transitions and precrystallization processes in a water–protein–electrolyte system, J.Cryst.Growth 273 (2004) 266–279.
[24] I. Prigogine, R. Defay, Thermodynamique Chimique, Desoer, Liege, 1950 English translation by H.D. Everett, Longmans-Green, London, 1954.
[25] J.T. Edsall, H. Edelhoch, R. Lontie, P.R. Morrison, Light scattering in solutions of Serum albumin: effects of charge and ionic strength, J.Am.Chem.Soc. 72 (1950) 4641–4656.
[26] T. Arakawa, S.N. Timasheff, Mechanism of protein salting in and salting out by divalent cation salts: balance between hydration and salt binding, Biochemistry 23 (1984) 5912–5923.
[27] P.G. Vekilov, Metastable mesoscopic phases in concentrated protein solutions, Ann.N.Y.Acad.Sci. 1161 (2009) 377–386.
[28] A.C. Dumetz, A.M. Snellinger-O'brien, E.W. Kaler, A.M. Lenhoff, Patterns of protein protein interactions in salt solutions and implications for protein crystallization, Protein Sci. 16 (2007) 1867–1877.
[29] F. Bonnete, S. Finet, A. Tardieu, Second virial coefficient: variations with lysozyme crystallization conditions, J. Cryst. Growth 196 (1999) 403–414.
[30] B. Guo, S. Kao, H. McDonald, A. Asanov, L.L. Combs, W.W. Wilson, Correlation of second virial coefficients and solubilities useful in protein crystal growth, J.Cryst. Growth 196 (1999) 424–433.
[31] W. Pan, O. Galkin, L. Filobelo, P.G. Vekilov, Metastable mesoscopic clusters in low-ionic strength protein solutin, The 2005 Annual Meeting (Cincinatti, OH), #42.
[32] K. Igarashi, M. Azuma, J. Kato, H. Ooshima, The initial stage of crystallization of lysozyme: a differential scanning calorimetric (DSC) study, J. Crystal Growth 204 (1999) 191–200.
[33] A. Stradner, F. Cardinaux, P. Schurtenberger, A small-angle scattering study on equilibrium clusters in lysozyme solutions, J.Phys.Chem.B. 110 (2006) 21222–21231.
[34] A. Shukla, E. Mylonas, E. Di Cola, S. Finet, P. Timmins, T. Narayanan, D.I. Svergun, Absence of equilibrium cluster phase in concentrated lysozyme solutions, Proc. Natl.Acad.Sci.U.S.A 105 (2008) 5075–5080.
[35] L. Porcar, P. Falus W-R, A. Faraone Chen, E. Fratini, K. Hong, P. Baglioni, Y. Liu, Formation of the dynamic clusters in concentrated lysozyme protein solutions, Phys. Chem. Lett. 1 (2010) 126–129.
[36] S.P. Rozhkov, Solubility and phase transitions in a water–protein–salt system, Biophysics 51 (2006) 822.
[37] S.P. Rozhkov, Critical phenomena in water–salt solutions of biopolymers, Russ.J. Phys.Chem. 70 (1996) 1982–1986.
[38] I.P. Bazarov, Termodinamika (Thermodynamics), High School, Moscow, 1983.
[39] T.L. Blundell, L.N. Johnson, Protein crystallography, Academic Press, New York, 1976.
[40] P. Retailleau, M. Ries-Kautt, A. Ducruix, No salting in of lysozyme chloride observed at low ionic strength over a large range of pH, Biophys.J. 73 (1997) 2156–2163.
[41] J.K. Baird, S.C. Scott, Y.W. Kim, Theory of the effect of pH and ionic strength on the nucleation of protein crystals, J.Cryst.Growth 232 (2001) 50–62.
[42] H. Morawetz,, Macromolecules in solution, Interscience Publishers, New York, 1965.
[43] S.P. Rozhkov, A.S. Goryunov, Effects of inorganic salts on the structural heterogeneity of serum albumin solutions, Eur. Biophys. J. 28 (2000) 639–647.
[44] W.T. Jenkins, Three solutions of the protein solubility problem, Protein Sci. 7 (1998) 376–382.
[45] O. Galkin, W. Pan, L. Filobelo, R.E. Hirsch, R.L. Nagel, P.G. Vekilov, Two-stop mechanism of homogeneous nucleation of sickle cell hemoglobin polymers, Biophys.J. 93 (2007) 902–913.
[46] T.M. Record, C.F. Anderson, T.M. Lohman, Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening and ion effects on water activity, Quart. Revs.Biophys. 11 (1978) 103–178.