Computer simulation of protein self-association during small-zone gel filtration

Estimation of equilibrium constants

Fred J. STEVENS and Marianne SCHIFFER
Division of Biological and Medical Research, Argonne National Laboratory, Argonne, IL 60439, U.S.A.

(Received 15 October 1980/Accepted 5 January 1981)

A simulation is developed that qualitatively describes the small-zone-gel-filtration behaviour of a reversibly associating protein. The results reflect the dependence of the apparent molecular weight of a reversibly associating protein on the equilibrium constant ($K_D$) and initial concentration of the protein as well as the column length. The behaviour of a protein on an individual column is characterized and thus a means is provided for estimation of $K_D$. The procedure is extended to describe the behaviour of a mixture of two proteins capable of heterologous as well as homologous association. This computer simulation has been applied in association studies of immunoglobulin light chains [Stevens, Westholm, Solomon & Schiffer (1980) Proc. Natl. Acad. Sci. 77, 1144–1148]. The $K_D$ value determined for the Bence–Jones protein Au ($10^4\text{m}^{-1}$) is close to the value ($6.6 \times 10^4\text{m}^{-1}$) determined by other methods [Maeda, Steffen & Engel (1978) Biophys. Chem. 9, 57–64].

The association constant ($K_D$) governing protein self-association is typically determined by equilibrium sedimentation (Steiner, 1952; Klotz et al., 1975) or chromatographically by large-zone (plateau) gel filtration (Winzor & Scheraga, 1963, 1964; Ackers, 1975). Certain cases in which the environment of an aromatic amino acid residue is altered by the aggregation of the polypeptides permit the use of spectrophotometric methods to monitor changes in circular dichroism, fluorescence polarization or absorbance (Azuma et al., 1978; Maeda et al., 1976, 1978) to characterize the association process. The value of computer simulation to assist the interpretation of sedimentation equilibrium was illustrated by Cox (1965a,b), and simulation techniques have been extensively applied to large-zone gel filtration in studies employing numerical solution to transport equations characterizing the chromatography process (for reviews, see Ackers, 1970, 1975).

Small-zone filtration substantially differs from large-zone filtration in that no stable equilibrium composition is attained within the sample. As a consequence, the apparent molecular weight of the eluted protein does not represent a weighted average of dimer and monomer molecular weights and cannot be directly related to $K_D$, as observed by Zimmerman & Ackers (1971) in evaluating their small-zone-gel-filtration simulation. Accordingly, the procedure described in the present paper calibrates the chromatography column with respect to $K_D$, characterizing the gel-filtration behaviour of the self-associating protein on the basis of monomer and dimer migration rates and taking into consideration band spreading and column length. We have found this simplified simulation to be useful in studies comparing and characterizing the association properties of immunoglobulin light chains, Bence–Jones proteins (Stevens et al., 1980), which exist as stable (non-associating) monomers, non-covalent dimers and covalent dimers (Solomon, 1981).

Materials and methods

Experimental

Protein purification and gel-filtration chromatography were performed as detailed previously (Stevens et al., 1980).

Computer simulation

Single protein. The simulation developed to model the gel-filtration behaviour of the reversibly associating protein was based on the assumption that the only molecular parameters necessary to describe the elution behaviour of the sample were the relative rates of migration of the monomeric and dimeric components of the protein mixture and the equili-
brium constant characterizing association of monomers.

The monomer–dimer equilibrium is described by the relationship:

\[ c_D = K_D c_M^2 \]

where \( c_M \) and \( c_D \) are the molar monomer and dimer concentrations respectively. Substituting \( c_0 = c_M + 2c_D \), it can be shown that:

\[ c_M = \frac{(-1 + \sqrt{8K_D c_0 + 1})}{4K_D} \]  
(1)

where \( c_0 \) is the total molar protomer concentration.

For the purpose of monitoring the passage of the protein through the column, the column was divided into imaginary cells. Loading of the sample is simulated by assigning the original protein concentration to the initial cells required to represent the sample size. By use of eqn. (1) and implicitly assuming complete re-equilibration, the monomer–dimer composition of the protein contained within each cell is calculated. Protein migration is effected by separately advancing the monomer and dimer content of each cell. The ratio of dimer advance to monomer advance is set by the relative times of elution of monomer and covalent dimer observed on the column being simulated.

After the migration step, a dispersion cycle is performed to approximate zone broadening. Again, an empirical method is used. Axial dispersion is simulated by defining a dispersion factor \( f_D \) reflecting the degree of dispersion equilibrium attained per dispersion cycle. Hence, if \( c_i \) represents the protein concentration of cell \( i \) after a cycle of dispersion, then:

\[ c_i^* = c_i - f_D(c_i - c_{i-1}) + f_D(c_{i+1} - c_i) \]

or

\[ c_i^* = f_D(c_{i-1} + c_{i+1}) + c_i(1 - 2f_D) \]  
(2)

For \( f_D = 0.5 \), dispersion equilibrium is reached and \( c_i^* \) is the average of \( c_{i-1} \) and \( c_{i+1} \). The dispersion cycle is repeated several times for each migration step. After the dispersion process is completed, the protein concentration at the cell corresponding to the terminus of the column is recorded, analogous to continuously monitoring the column eluant experimentally. At this point, the monomer–dimer composition of each cell is again determined and the complete cycle is repeated.

**Two proteins.** The simulation was extended to include the case of simultaneous gel filtration of two proteins capable of association with each other (heterologous association) as well as self association (homologous association).

This process is represented below:

\[ A_2 \xrightarrow{k_{1+}} A + A + B + B \xrightarrow{k_{1-}} B_2 \]

\[ \text{AB} \]

At equilibrium, concentrations have reached a steady state and hence the usual equilibrium equations apply,

\[ a_D = K_1 a^2 \]  
(3)

\[ b_D = K_2 b^2 \]  
(4)

\[ (ab)_D = K_{12} a \cdot b \]  
(5)

where \([A] = a \) and \( K_1 = k_{1+}/k_{1-} \), etc.

The free monomer concentrations are given by:

\[ a = a_0 - 2a_D - (ab)_D \]

\[ b = b_0 - 2b_D - (ab)_D \]

where \( a_0 \) and \( b_0 \) are the total protomer molar concentrations of each species.

Substituting eqns. (3), (4) and (5),

\[ a = a_0 - 2K_1 a^2 - K_{12} a \cdot b \]  
(6)

\[ b = b_0 - 2K_2 b^2 - K_{12} a \cdot b \]  
(7)

Direct solution of eqns. (6) and (7) to determine \( a \) and \( b \) requires solution of a fourth-order equation. Accordingly, the simulation determined these concentrations by use of a simple numerical procedure. A preliminary value of \( a \) is first estimated by use of eqn. (1), thus effectively setting \( K_{12} = 0 \). This overestimated value for \( a \) is then used in eqn. (7), with \( K_{12} \neq 0 \), and an estimate for \( b \) is obtained. Subsequent substitution of this \( b \) value into eqn. (6) results in a new estimate of \( a \). By repetition of this process, values of \( a \) and \( b \) are determined that simultaneously satisfy eqns. (6) and (7).

The procedure employed to determine the monomer–dimer composition of each cell is the principal difference between the one- and two-protein simulations. In all other aspects, the two programs used the same strategies.

**Estimation of simulation parameters.** The parameters required by the simulation to represent the gel-filtration column are rates of monomer and dimer migration, column length, sample volume, dispersion factor and the number of repetitions of the dispersion cycle. The monomer and dimer migration rates in the examples described here were set to correspond to the (inverse) elution times for Bence–Jones covalent dimer (mol.wt. 45,000) and monomer (22,500) observed during chromatography on a column (about 90 cm long) of Sephadex G-75 (Superfine grade) (Stevens et al., 1980). These elution times were approx. 13 and 17h respectively. The relative rates were satisfied by setting the monomer migration rate at 10 and that of the dimer at 13. The determinations of column length and sample volume were also based on characteristics of the experimental column. At a flow rate of 15 ml/h, the 1 ml sample required 1/15 h to load and hence the ratio of elution time to load time was 225 at the average of the elution times noted.
above. The values of the two parameters depend on each other in the simulation; thus, for a sample volume set at eight cells, the corresponding column length is 2000 cells.

Dilution and band spreading were accomplished by setting the dispersion factor and the number of repetitions of the dispersion cycle. The band spreading obtained by eight repetitions of a cycle using a dispersion factor of 0.5 was similar to that observed during chromatography of covalent dimer and stable monomer.

The simulated column was calibrated by the elution positions of the covalent dimer and stable monomer. It is assumed that the covalent and non-covalent dimers have the same Stoke's radii, as we have observed for Bence–Jones proteins with high $K_D$ values. On the basis of the standard logarithmic relationship between molecular weight and elution position, the apparent molecular weight of a self-associating monomer was then estimated as the molecular weight corresponding to the position of its elution peak.

Results and discussion

In small-zone filtration, the observed elution position for a reversibly associating protein is determined by the length of the column and the gel-partitioning properties of the matrix as well as the initial protein concentration ($c_0$) and $K_D$. The $K_D$ and $c_0$ values determine the monomer–dimer equilibrium composition of the sample as applied to the column. However, this initial equilibrium does not persist as the protein migrates. Since dimers migrate more rapidly than monomers at a relative rate dependent on the partitioning properties of the gel, the result is that the non-covalently associated monomers from the leading edge of the zone migrate into an area of low protein concentration. The equilibrium is thus shifted to dissociation of the dimers to form the more slowly migrating monomer. Accordingly, a sharp boundary is formed at the leading edge of the zone; on the other hand, at the trailing edge the boundary is extended as dilution by diffusion and forward dimer migration continually shifts the sample composition to a higher monomer content (Winzor & Scheraga, 1963).

The reproduction of the general properties of chromatographic columns was tested in the simulation by varying the sample-volume and column-length parameters. If the total protein applied is held constant while the sample volume is varied, the resulting elution profiles are relatively invariant until the sample volume approaches approx. 10% of the column size. At this point, plateau behaviour characteristic of large-zone chromatography begins to become apparent (Fig. 1). The initial (loading) concentrations of the samples depicted in Fig. 1 range from 20 mg/ml down to 0.2 mg/ml. Accordingly, at a $K_D$ of $1 \times 10^4 \text{M}^{-1}$, the weight-average molecular weight varies from 40300 to 25500 at loading, whereas the position of the eluted protein is nearly constant, with the apparent molecular weight only shifting from 28500 to 26000. Thus the simulation exhibited the expected behaviour that the weight-average molecular weight agrees with apparent molecular weight only for large-zone chromatography. The apparent molecular weight (26000) of the most concentrated fraction of the sample exhibiting a plateau is close to the calculated weight-average molecular weight (25500). These results agree with Zimmerman & Ackers (1971) finding that neither the weight-average molecular weight of the load sample nor that estimated by the position of the eluted protein provides a valid analysis of small-zone filtration.

In Fig. 2 the effect of column length on the observed elution pattern is demonstrated in the case of a protein having $K_D = 1 \times 10^4 \text{M}^{-1}$. The original sample has a concentration of 10 mg/ml and is composed of 50% covalent dimer by weight. As expected, with increasing column length, dilution
The protein (Bence-Jones protein) is assumed to have a mol.wt. of 22500. The samples consist of covalent dimer and monomer at a concentration of 5 mg/ml each. The association constant \( (K_D) \) is \( 1 \times 10^4 \text{M}^{-1} \). Column lengths: ----, 500 cells; ---, 1000; ------, 1500; ----, 2000; ----, 2500. Inset: the dependence of apparent molecular weight (mol.wt. app.) on column length. \( K \) values; ----, \( 1 \times 10^4 \text{M}^{-1} \); ---, \( 5 \times 10^4 \text{M}^{-1} \); ----, \( 1 \times 10^5 \text{M}^{-1} \).

Moreover, it follows from eqn. (1), and considering that apparent molecular weight is determined by the fractional monomer content, \( c_m/c_0 \), that the relevant
variable is neither \( K_D \) nor \( c_p \) but the product \( c_pK_D \). In Fig. 4, apparent molecular weight is plotted as a function of \( \log(c_pK_D) \) and exhibits a sigmoidal relationship with plateaux at the monomer and dimer molecular weights. Fig. 4 also shows that a nearly linear relationship of apparent molecular weight to \( \log(c_pK_D) \) is obtained in the domain centred at approx. \( \log(c_pK_D) = 1 \), and comprising almost two orders of magnitude of \( c_pK_D \).

Fig. 5 depicts the elution profile obtained at 20°C for the Bence–Jones protein Au. The sample is composed of a mixture of immunoglobulin light chains of mol.wt. \( \sim 22,500 \), disulphide-linked light chains forming covalent dimers of mol.wt. \( \sim 45,000 \) and a light-chain fragment comprising the variable domain and having mol.wt. \( \sim 12,000 \). From Fig. 3, and assuming an initial monomer concentration of 5 mg/ml, a \( K_D \) of \( 10^5\text{M}^{-1} \) may be estimated from the observed light-chain elution at a position corresponding to an apparent mol.wt. of 35,000. This compares favourably with a \( K_D \) of \( 6.6 \times 10^4\text{M}^{-1} \) determined spectrophotometrically for protein Au at 20°C by Maeda et al. (1978).

The simulated gel filtration of a mixture of two proteins is of interest in this laboratory for experiments in which the heterologous association of Bence–Jones proteins is examined. Fig. 6 shows the simulated elution patterns of two proteins each having homologous association constants of \( 1 \times 10^4\text{M}^{-1} \). Heterologous association has little noticeable effect on the observed elution pattern until the heterologous association constant is approximately one-half of an order of magnitude greater than the \( K_D \) values governing self-association. The origin of this limitation is apparent in Fig. 7, which shows the percentage of total dimers in a heterologous protein mixture as a function of \( K_{12} \) for fixed \( K_1 \) and \( K_2 \). Formation of the heterologous dimer is compensated for by a decrease in the content of homologous dimer and thus, unless \( K_{12} \) dominates, the net increase in dimer content is negligible and results in no significant perturbation in the chromatographic profile.

The results of gel filtration of a mixture of the \( \kappa_1 \) Bence–Jones proteins Kin and Und are shown in Fig. 8(b). Fig. 8(a) shows simulated elution profiles of proteins having the estimated \( K_D \) values and concentrations of proteins Kin and Und. As shown in Fig. 8(b), the monomeric component of the Kin and Und mixture eluted at a higher apparent
molecular weight than either monomer chromatographed individually. By comparison with the simulation shown in Fig. 8(a), the elution behaviour of the Kin and Und monomer mixture suggests a $K_{12}$ of at least $1 \times 10^4 \text{M}^{-1}$.

The procedure developed in the present study to model the small-zone-gel-filtration behaviour of a reversibly associating protein and to estimate the $K_D$ governing this process appears to have several areas of application. The reasonably close value of the $K_D$ determined for the Bence-Jones protein Au compared with the $K_D$ reported by Maeda et al. (1978) suggests the simulation provides a means for relatively rapid estimation of association constants valid within an order of magnitude. Furthermore, the dependence of elution profile on $K_D$ and $c_0$ is as expected, and the formation of a plateau as the sample is increased is observed. Thus, although this simulation is not intended to replace the more quantitative procedures available in laboratories specializing in the analysis of protein–protein interactions, it may be useful in other protein studies more directly aimed at aspects such as enzymic or structural properties in which information on the association properties is of value. In addition to its convenience as a routine laboratory procedure, an advantage of small-zone gel filtration is that expenditures of sample may be minimized relative to the requirements of large-zone filtration. The upper limit of $K_D$ that may be evaluated by this procedure is limited by the lower limits of protein concentration detectable. However, dimerization constants of at least $10^8 \text{M}^{-1}$ appear measurable with sensitive spectroscopic monitors or by effectively increasing column length by recycling techniques. The lower limit of $K_D$ that can be estimated is determined by the maximum quantity of protein that may be applied to the column. Since excessively high protein concentration results in osmotic shrinkage of the Sephadex gel (Nichol et al., 1973), a $K_D$ value of $5 \times 10^2 \text{M}^{-1}$ appears to be approximately the lower limit for this technique when a Sephadex matrix is used.

We thank Dr. S. Peter Spragg (Department of Chemistry, University of Birmingham, Birmingham, U.K.) for helpful discussions. Dr. Norbert Hilschmann
Simulation of small-zone gel filtration

(Max-Plank-Institut, Göttingen, Germany) for protein Au, Dr. Alan Solomon (University of Tennessee, Knoxville, TN, U.S.A.) for immunochemically characterized Bence–Jones proteins Kin and Und, and Mrs. Florence Westholm for expert technical assistance. This work was supported by the U.S. Department of Energy (contract no. W-31-109-ENG-38).

References

Cox, D. J. (1965a) Arch. Biochem. Biophys. 112, 249–258
Cox, D. J. (1965b) Arch. Biochem. Biophys. 112, 259–266