Turk J Chem
26 (2002) , 877 – 887.

Modeling of TaqI Endonuclease Purification by HPLC

Mehmet DÜZENLİ and Kutlu Ö. ÜLGEN*
Department of Chemical Engineering, Boğaziçi University, Bebek, 80815 İstanbul-TURKEY
e-mail: ulgenk@boun.edu.tr

Received 12.11.2001

The present work is based on the previously reported purification study of TaqI endonuclease[5] and aimed to simulate the experimentally obtained chromatograms in order to reduce the time and cost in the design, optimization and scaling up of purification processes. Three different models depending on equilibrium, rate and plate theories are used in the mathematical description of experimental elution profiles and the solutions of the model equations are compared. The equilibrium-dispersive theory provides excellent results in predicting the experimental chromatograms. It is thus recognized that band dispersion takes place in the column through axial dispersion and nonequilibrium effects. The user-friendly software written in Visual Basic can be used in the design of purification processes as well as in the determination of optimal separation conditions for many other enzymes.

Key Words: TaqI endonuclease, modeling, design, equilibrium theory, equilibrium-dispersive theory, plate theory

Introduction

TaqI endonuclease is a thermostable Type II restriction enzyme that recognizes and cleaves the duplex DNA sequence

$$\begin{array}{c} \downarrow \\ 5' \bullet\bullet\bullet T-C-G-A \bullet\bullet\bullet 3' \\ 3' \bullet\bullet\bullet A-G-C-T \bullet\bullet\bullet 5' \\ \uparrow \end{array}$$

leaving a $3'$ hydroxyl on the T and a $5'$ phosphate on the C[1]. It has extensive applications in molecular biology research and diagnostics.

Several purification procedures for TaqI endonuclease have been reported[1−4]. A short two-step protocol developed by the sequential use of the cation-anion HPLC system resulted a 24-fold purification of TaqI endonuclease[4]. A further investigation of the chromatographic media revealed that sodium acetate

*to whom all correspondence should be addressed

(pH 5.0) and phosphate (pH 7.0) buffers with cation exchange columns and L-histidine (pH 6.0) and Tris (pH 8.0) buffers with anion exchange columns are appropriate for use as the mobile phases in the HPLC purification of TaqI endonuclease[5]. The above mentioned buffer systems result in purification factors of 3.56, 10.02, 6.16 and 3.52, respectively. It is generally accepted that optimization work based on simulation with mathematical models is cheaper, more comprehensive and more versatile than can be achieved with repetitive laboratory experiments.

In the present work, the objective was to develop a user-friendly approach and associated software that could be used for simulating the purification of TaqI endonuclease, as well as for observation of the effects of the experimental parameters on the performance of the developed models. A computer program written in Visual Basic is used to simulate the experimental chromatograms. The program is based on the solution of the mass balance equations in a chromatography column. Three different models depending on equilibrium, equilibrium-dispersive and plate theories are used in predictions. Different numerical methods in solving the model equations are also compared. The parameters used in the calculations are optimized in order to obtain the best fitted chromatogram.

Theory of Modeling

All theories of chromatography are based on the numerical solutions of the mass balance equations obtained from the column under certain fundamental assumptions[6]:

1. The compressibility of the mobile phase is negligible and the velocity profile is flat (plug flow). The influence of the compressibility of liquids on the retention times of the components in liquid chromatography is small. The mobile phase velocity (u) is constant along the column.

2. The axial dispersion coefficient, D_L, is constant. It includes the contributions to the axial dispersion due to molecular diffusion and to the nonhomogeneity of the flow, i.e. eddy diffusion.

3. The partial molar volumes of sample components are the same in both phases. Thus, the mass transfers taking place during the adsorption process are made at constant volume.

4. The mobile phase is not adsorbed.

Mass balance for the differential volume can be written as

$$\frac{\partial(uC)}{\partial z} + \frac{\partial C}{\partial t} + F\frac{\partial C_S}{\partial t} = D_L\frac{\partial^2 C}{\partial z^2} \tag{1}$$

where D_L is the axial dispersion coefficient of the compound in the mobile phase and F is the ratio of volumes of mobile to stationary phases.

In the present work, three general categories of theories proposed for chromatography will be considered: equilibrium, equilibrium-dispersive and plate.

Equilibrium theory

The equilibrium theory is based on ideal cases with the following assumptions:

1. The column efficiency is infinite.
2. The axial dispersion is negligibly small. Thus, zone spreading effects are ignored.
3. The rate of mass transfer kinetics is infinite.
4. The equilibrium of the protein/enzyme concentration between the mobile and stationary phases is always established during elution. This assumption allows the elimination of mass transfer kinetics.

Under these assumptions, Eq. (1) becomes

$$\frac{\partial C}{\partial t} + F\frac{\partial C_S}{\partial t} + u\frac{\partial C}{\partial z} = 0 \tag{2}$$

which is the basic equation for equilibrium theory.

Although the equilibrium theory is too simplified to describe the actual chromatographic processes, it does give important information on the movement of the zone of the solute in the column.

Equilibrium-dispersive theory

The equilibrium-dispersive model is valid if mass transfer in a chromatographic column is controlled only by molecular diffusion across the mobile phase flowing around the packing particles and if the exchange of solutes between the stationary and mobile phases is very fast. In this model the zone spreading effects are taken into consideration. The assumptions are as follows:

1. The mobile and stationary phases are always in equilibrium.
2. The contributions of all the non-equilibrium effects can be lumped into an apparent axial dispersion coefficient (D_L).

The system of mass balance equations can be rewritten as

$$\frac{\partial C}{\partial t} + F\frac{\partial C_S}{\partial t} + u\frac{\partial C}{\partial z} = D_L\frac{\partial^2 C}{\partial z^2} \tag{3}$$

Guiochon et al.[6] gave a simple equation for the estimation of D_L as

$$D_L = \frac{(HETP)u}{2} \tag{4}$$

where u is the mobile phase velocity and HETP is the column height equivalent to a theoretical plate. The mobile phase velocity can be calculated as

$$u = \frac{F_V Z}{V_0} \tag{5}$$

where F_V is the volumetric flowrate, Z is the column length and V_o is the void volume which is calculated as

$$V_0 = V_t.\varepsilon \text{ and } V_t = \frac{\pi d_c^2}{4}Z \tag{6}$$

where d_c is the column diameter and ε is the void fraction. HETP can be determined experimentally from analytical chromatograms or it can be calculated by using the general equation given in books on mass transfer as

$$HETP = \frac{Z}{Np} \tag{7}$$

where Np is the plate number of the chromatography column.

As there is no information in the literature on the physical properties of TaqI endonuclease, Np is approximated as 360 using the expression provided by Yamamoto et al.[7,8] and the properties of BSA, which has a similar molecular weight of 66 kDa to TaqI endonuclease. Using Eqs. (4)-(7), D_L is calculated as 0.0445 cm^2/min. This value was kept constant for all the simulations based on the equilibrium-dispersive model.

Plate theory

The plate model described by Yamamoto et al.[7] is based on the continuous flow plate theory where the column is divided into a series of artificial plates, each with complete mixing. The following assumptions are made:

1. The column consists of a certain number of equivalent theoretical plates, in each of which the ratio of the volume of the stationary phase to that of the mobile phase is the same.
2. The flow is continuous without mixing between the plates.
3. The equilibrium of solutes between the two phases is instantaneously attained.

The mass balance for n^{th} plate is

$$\left\{\begin{array}{c}\text{Solute Flow}\\ \text{into n}^{th}\text{ plate}\end{array}\right\} - \left\{\begin{array}{c}\text{Solute Flow}\\ \text{out of n}^{th}\text{ plate}\end{array}\right\} = \left\{\begin{array}{c}\text{Accumulation in mobile}\\ \text{phase of n}^{th}\text{ plate}\end{array}\right\} + \left\{\begin{array}{c}\text{Accumulation in stationary}\\ \text{phase of } n^{th}\text{ plate}\end{array}\right\}$$

$$F_V C_{n-1} - F_V C_n = \frac{V_0}{N_p}\frac{dC_n}{dt} + \frac{V_t - V_0}{N_p}\frac{dC_{Sn}}{dt} \tag{8}$$

By this scheme, zone-spreading effects are represented by the number of theoretical plates, Np. Np is calculated using the expression suggested by Yamamoto et al.[7–9] Zone sharpening effects are also incorporated in the model by considering a change in ionic strength in each plate. The plate theory is an easy and quick way of simulating the chromatograms in liquid chromatography if the efficiency of the column is large enough so that its effect on the bandwidth and the band profile can be neglected. This model is not directly applicable when the protein concentration is so high that the distribution coefficient is dependent on it as well as on the ionic strength.[7]

Isotherm equations

The mass balance equations of all theories above include two unknown concentrations: concentration of protein (enzyme) in the mobile phase, C, and in the stationary phase, Cs. The isotherm equation gives the relation between these phases at equilibrium.

In linear chromatography, the retained amount of a compound is strictly proportional to its concentration in mobile phase and the fraction of molecules adsorbed is constant. Consequently, the mass transfer coefficient is constant. Therefore, the slope of the isotherm is assumed to be equal to mass distribution coefficient, K

$$C_S = KC \tag{9}$$

The distribution coefficient, K, is influenced by various factors such as ionic strength of a solution, protein concentration, pH and types of ion exchangers. One of the popular estimations of K is suggested by

Yamamoto et al.[7,8] that K is a function of ionic strength (I)

$$K = A * I^B + K_{CRT} \tag{10}$$

where A and B are empirical constants and K_{CRT} is the critical distribution coefficient of the protein at high ionic strength where electrostatic interaction can be ignored. This estimation gives good results especially in ion exchange chromatography when the linear gradient elution is applied.

In nonlinear chromatography, the phase equilibrium isotherms are not linear, i.e. the concentration of a component in the stationary phase at equilibrium is not proportional to its concentration in the mobile phase. The nonlinear isotherms in ion exchange chromatography are usually favorable, i.e. K is a decreasing function of concentration, and expressed by the Langmuir equation.

$$C_S = q = \frac{aC}{1 + bC} \tag{11}$$

Computational

The main partial differential equations to be solved are Eq. (2) for the equilibrium theory, Eq. (3) for the equilibrium-dispersive theory and Eq. (8) for the plate theory; all are combined with isotherm Eqs. (9) to (11). The computer program is written in Visual Basic. Different numerical calculation procedures, i.e. orthogonal collocation method and some special cases of finite difference approximation method (Crank-Nicholson, Lax-Wendroff and Propagation through Grid), are used in solving the governing differential equations. In these calculations, the chromatography column is divided into slices of width, Δz, time is given in multiples of Δt and they are defined as grid spacing in space and time, respectively[10]. Thus, space and time are discretized. When choosing the correct grid spacing in the explicit methods, the general stability criterion of the modulus being less than 0.5 is used for the relative choice of these parameters. The ordinary differential equations in the plate models are solved by the Runge-Kutta-Gill method. All the models and the numerical solution methods are summarized in Table 1.

Table 1. Models and related mathematical equations.

Theory	Isotherm	Method
Equilibrium	Linear, constant K	Crank Nicholson
Equilibrium	Linear, constant K	Lax-Wendroff
Equilibrium	Linear, constant K	Propagation through Grid
Equilibrium	Linear, variable K(I)	Propagation through Grid
Equilibrium	Nonlinear, Langmuir	Propagation through Grid
Eq.-Dispersive	Linear, constant K	Lax-Wendroff
Eq.-Dispersive	Linear, constant K	Orthogonal Collocation
Eq.-Dispersive	Linear, variable K(I)	Lax-Wendroff
Eq.-Dispersive	Nonlinear, Langmuir	Orthogonal Collocation
Plate	Linear, constant K	Runge-Kutta-Gill
Plate	Linear, variable K(I)	Runge-Kutta-Gill
Plate	Nonlinear, Langmuir	Runge-Kutta-Gill

The agreement between the experimental data and the simulations are tested by statistical analysis. The residual is defined as the sum of the squares (SSR) of the differences between experimental and calculated

data and is given by

$$SSR \underset{m=1}{\overset{Nd}{=}} \sum (C_m^{\text{obs}} - C_m^{\text{cal}})^2 \tag{12}$$

where m is observation number and N_d is total number of observations. The estimated variance of the error (= population variance) is calculated by the SSR at its minimum divided by its degrees of freedom

$$\sigma^2 \approx s^2 = (SSR)_{\min}/(m-p) \tag{13}$$

where p is the number of parameters and $s^2$is the variance. The standard error (= the estimated standard deviation) is calculated by taking the square root of the estimated variance of the error.

Results and Discussion

An experimental study[5] on the optimization of the mobile phase conditions for the purification of TaqI endonuclease revealed that sodium acetate (pH 5.0) and phosphate (pH 7.0) buffers with cation exchange columns and L-histidine (pH 6.0) and Tris (pH 8.0) buffers with anion exchange columns resulted in satisfactory separations. Among these chromatographic media, the anion exchange system gave better results in which over 80% of the injected TaqI endonuclease was recovered in elution. The TaqI endonuclease activity was mostly obtained in fractions collected between 32-36 min and at 0.21-0.28 M NaCl gradient when sodium acetate, L-histidine and Tris buffers were used. The present work is based on this previously reported study[5] and aimed to simulate the experimentally obtained chromatograms in order to reduce the time and cost in the design, optimization and scaling up of purification processes.

Modeling studies

The simulations of the experimental chromatograms are performed by different methods depending on the theory, isotherm and calculation procedure chosen. The aim of using various methods for the simulation of the chromatograms is not only to construct a model for TaqI endonuclease elution, but to construct a general model that can be applicable for many restriction enzymes.

In order to simulate each chromatogram containing active TaqI endonuclease, a user-friendly program was written in Visual Basic. The program is based on the algorithm given in Figure 1. The computer program asks the user to enter the values for column length (z), column diameter (dc), flow rate of mobile phase (F), concentration of the injected sample (C_0), grid spacing in length (Δz), grid spacing in time (Δt) and mass distribution coefficient (K) or the Langmuir parameters. The first four are the experimental parameters and, therefore, their experimental values are used. The next two mathematical parameters, Δz and Δt, are space and time increments, and are changed in order to obtain the best-fitted chromatogram. All the parameters are listed in Table 2. In addition, the experimentally obtained active fractions that met the peaks on the HPLC outputs are assumed to contain TaqI endonuclease only, whereas the fractions belonging to other peaks are assumed to contain other intracellular proteins.

Table 2. Experimental (Tris buffer) and theoretical parameters

Experimental Parameters	Co (Tris buffer)	10.9917 mg/ml
	F	1 cm^3/min
	Z	10 cm
	d_C	1 cm
	K (Tris buffer)	1.3 mg/ml
	a (Langmuir par.)	1.3
	b (Langmuir par.)	0.5
	t_S	20 min
	t_G	80 min
	I_O	0.1 M
	I_f	1 M
	ε	0.4
Theoretical Parameters	Np	360
	D_L	0.0445cm^2/min
	A	4.9 x 10^{-6}
	B	-7.51
	K_{CRT}	0.44-0.60
	K_{SALT}	0.9

All of the theories were found to underestimate the retention time of the peaks because of the assumptions used in the models as well as because of the time and volume lags during the experiments with gradient elution. The same situation has been reported by Carta et al.[11]. Therefore, an additional parameter called "time-lag" is included in the computer program, which can easily be adjusted by the program user. An increase in time-lag parameter causes the peak to shift to the right. It should be noted that the value of time-lag is not used in any model for the calculation of the chromatogram, therefore the shape of the peaks (height and width) is independent of it. This parameter is only helpful in precisely adjusting the retention time.

Comparison of equilibrium, equilibrium-dispersive and plate theories

Three different approaches have been used to simulate the elution profiles in the ion exchange chromatography column. Figures 2 a-c show the experimental data and the simulated chromatograms belonging to the HPLC system with anion exchange column and Tris buffer (pH 8.0) as the mobile phase.

The predictions obtained using the equilibrium theory, which completely neglects the influence of the mass transfer kinetics and of axial dispersion on the band profiles, are presented in Figure 2a. The agreement with experimental results is found to be poor with all the numerical methods used. It is reported that when the column efficiency is low and the sample size is low to moderate, this ideal model should be avoided since in these cases the contributions of the mass transfer kinetics and the axial dispersion to the band profiles becomes significant compared to the influence of a linear isotherm[6].

The combination of kinetics and thermodynamics is then illustrated by several numerical solutions of another simple model, i.e. the equilibrium-dispersive model. The equilibrium dispersive model assumes that all contributions due to nonequilibrium can be lumped into an apparent axial dispersion term. Axial dispersion in a chromatography column is mainly caused by two factors: (1) molecular diffusion along a column length and (2) nonuniformity of linear velocity in the mobile phase (eddy diffusion). As the column diameter used in the present study is more than 100 times the particle diameter, a practically uniform velocity

profile should be obtained across the cross section of the column. However, the experimental asymmetric elution curve with a tail may be ascribed to a diffusion coefficient in stationary phase that is much lower than the molecular diffusion coefficient.

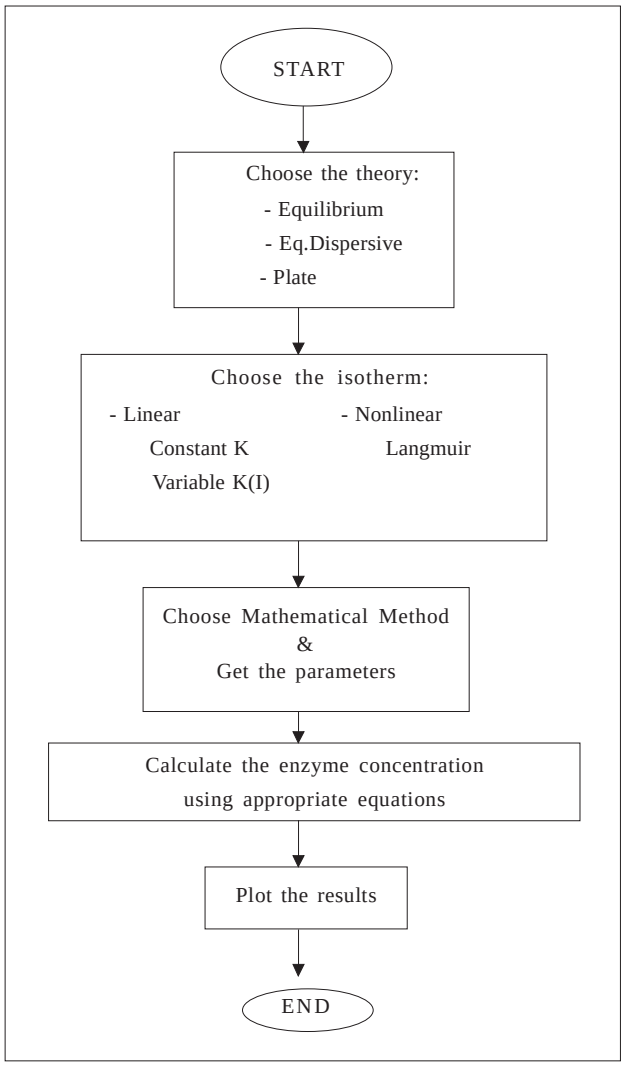

Figure 1. Calculation strategy

The results of the equilibrium-dispersive models with different isotherms and numerical solution methods are presented in Figure 2b. If the equilibrium-dispersive theory with variable distribution coefficient, K(I), is used in predicting the experimental chromatogram, a highly dispersed profile is obtained. The Lax-Wendroff approach with constant K (eq-dispersive theory), on the other hand, resulted in a sharp front and a dispersed tailing, which indicates the effect of apparent dispersion term on the band profile and that the resistance to mass transfer in the column is significant. In addition, the shape of this simulated band profile indicates that the injected sample is overloaded. However, the protein capacity of the anion exchange column was given as 320 mg protein by the manufacturer. A 1 ml injection of the sample contained 9-11 mg protein, i.e 3-4% loading, only. Therefore, the overloaded profile with concentration discontinuity is not correct. The experimental elution data with an asymmetric chromatogram indicates a nonlinear isotherm. Among all the theories, the equilibrium dispersive model combined with a nonlinear isotherm results in a

perfect fit to the experimental data as in Figure 2b. The estimated variance of the error is 0.0044. This model further assumes that the apparent dispersion coefficients of solutes remain constant independent of the concentration of sample components. As concentrations in nonlinear chromatography remain low, this is a reasonable assumption in most cases[6].

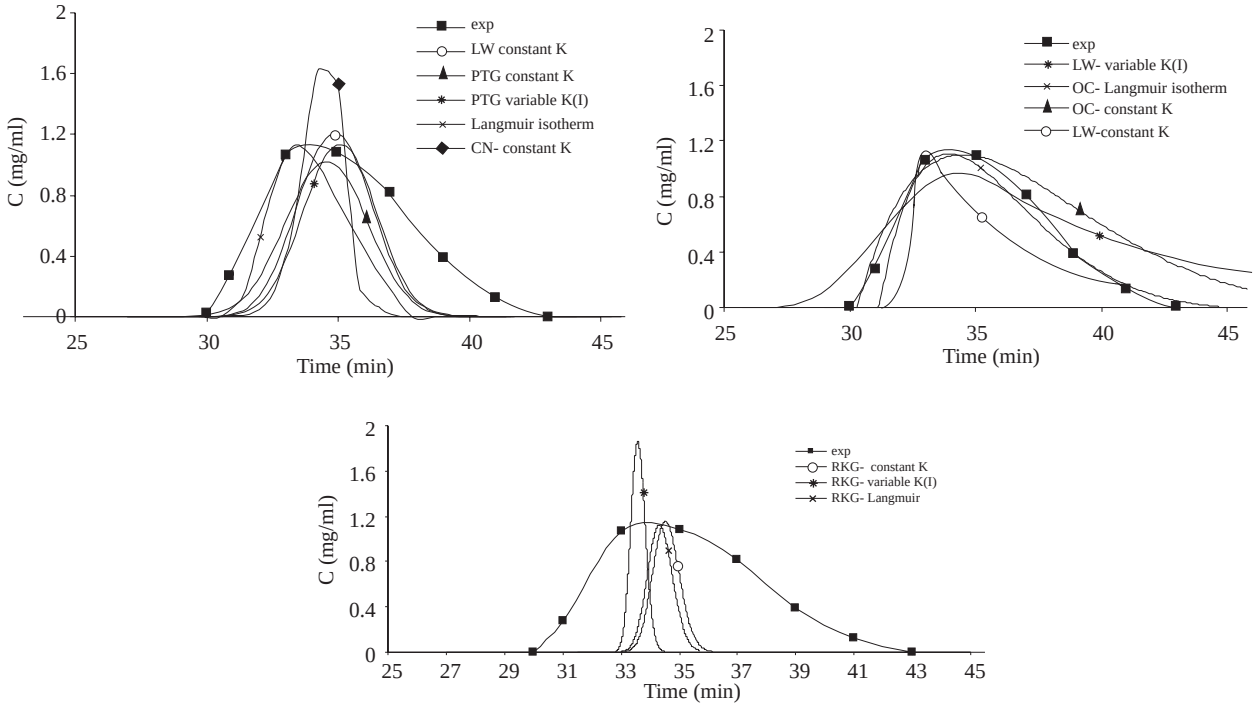

Figure 2. Comparison of experimental and simulated chromatograms in Tris buffer using (a) equilibrium theory, (b) equilibrium-dispersive theory and (c) plate theory.

Figure 2c shows the chromatograms obtained with the plate models. The plate theory results in symmetrical peaks over time and all the calculated chromatograms are narrower than the experimentally obtained elution profiles. This indicates that zone-spreading effects play an important role and the assumption of infinite column efficiency (i.e. infinite number of theoretical plates) is not valid. Np calculated as 360 plates is also too far from accepting this assumption. The equation for Np estimation includes many experimental parameters that have to be known according to the column and the enzyme of interest. Since there are no such physical and chemical property data available for TaqI endonuclease in the literature, the values provided by Yamamoto et al.[7] for BSA and DEAE column were used, which might have caused the deviation. This model is not applicable to cases in which the effect of mass transport is significant, i.e. highly asymmetric elution curves due to low stationary phase diffusion coefficients[7].

Validation of the model

In order to validate the equilibrium dispersive model, it is applied to another anion exchange system with 20 mM L-histidine as the mobile phase. The pH of the buffer is 6.0. The Langmuir parameters, a and b, are changed to 0.7 and 1.5, respectively. The model equations are solved by the orthogonal collocation (OC) numerical method. The agreement with the experimental data is excellent (Fig. 3). The estimated variance

of error is 0.017. The model is further tested on the cation exchange system with the 50 mM Na acetate buffer (pH 5.0) as the mobile phase. Figure 4 shows a perfect fit of experimental and simulated elution profiles. The estimated variance of error is 0.00349.

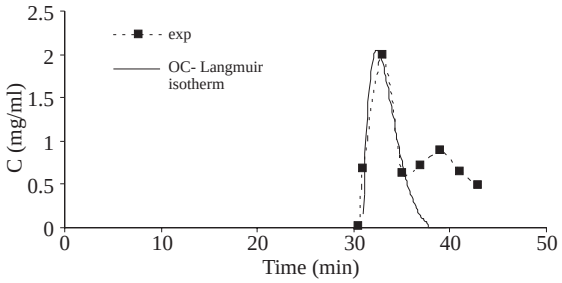

Figure 3. Comparison of experimental and simulated chromatogram in L-histidine buffer using equilibrium-dispersive theory.

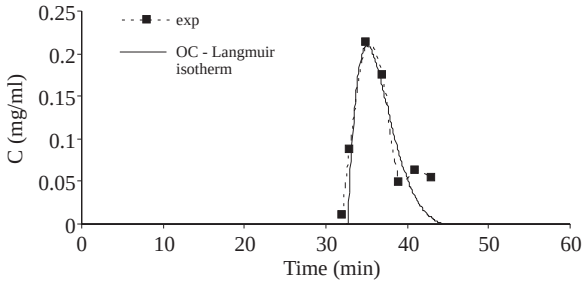

Figure 4. Comparison of experimental and simulated chromatogram in Na-acetate buffer using equilibrium-dispersive theory.

Conclusion

The main diffuculty in modeling the chromatographic purification process of TaqI endonuclease is the lack of data in the literature about the physical and chemical properties of this restriction enzyme. This limits the use of various methods and calculation procedures and also restricts full investigation of the effects of the parameters. Nevertheless, the developed computer program is very satisfactory in predicting the experimental chromatograms of TaqI endonuclease obtained with HPLC purification. This user-friendly software can be used in the design of purification processes as well as in the determination of optimal separation conditions for many other enzymes. The integration of experimental results with computer modeling reduces the number of highly expensive experiments and consequently enables the researcher to accelerate the development of chromatographic separation processes.

Acknowledgment

This research was supported by the Boğaziçi University Research Fund through projects 99HA501D2 and 01A502.

References

1. S. Sato, C.A. Hutchison, and J.A. Harris, **Proc. Natl. Acad. Sci. USA., 74(2),** 542-546 (1977).
2. P.J. Greene, H.L. Heyneker, F. Bolivar, R.L. Rodriguez, M.C. Bedlach, A.A. Covarribuas, K. Backman, D.J. Russel, R. Tait, and H.W. Boyer, **Nucleic Acids Research. 5(7),** 2373-2380 (1978).
3. M.M. Altıntaş, and K.Ö. Ülgen, **Acta Biotechnologica 59,** 46-57 (1999).
4. M.M. Altıntaş, N. Özer, and K. Ö. Ülgen, **Journal of Chromatography A, 828,** 373-381 (1998).
5. M. Düzenli, and K.Ö. Ülgen, **Chromatographia, 52,** 146-151 (2000).
6. G. Guiochon, S.G. Shirazi, and A.M. Katti, New York, Academic Press, 1996.
7. S. Yamamoto, K. Nakanishi, and R. Matsuno, Ion-Exchange Chromatography of Proteins, New York, Marcel Dekker Inc., 1988.
8. S. Yamamoto, M. Nomura, and Y. Sano, **AIChE Journal, 33,** 1426-1434 (1987).
9. S. Yamamoto, **Biotechnology and Bioengineering, 48,** 444-451 (1995).
10. M. Czok, and G. Guiochon, **Analytical Chemistry 62,** 189-200 (1990).
11. G. Carta, and W.B. Stringfield, **Journal of Chromatography 605,** 151-159 (1992).