Determination of Active Concentrations and Association and Dissociation Rate Constants of Interacting Biomolecules: An Analytical Solution to the Theory for Kinetic and Mass Transport Limitations in Biosensor Technology and Its Experimental Verification†

Kristmundur Sigmundsson,*,‡ Gísli Másson,§ Richard Rice,|| Nicole Beauchemin,⊥ and Björn Öbrink‡

Department of Cell and Molecular Biology, Medical Nobel Institute, Karolinska Institutet, SE-171 77 Stockholm, Sweden, Department of Mathematics, University of Stockholm, SE-106 91 Stockholm, Sweden, Louisiana State University, 15625 Antietam Avenue, Baton Rouge, Louisiana 70817, and McGill Cancer Center, McGill University, Montreal, Quebec H3G 1Y6, Canada

Received February 4, 2002

ABSTRACT: Accurate determination of kinetic rate constants for interacting biomolecules requires knowledge of the active concentrations of the participating molecules. Also, in other biomedical and clinical applications, sensitive, precise and accurate methods are needed to determine the concentration of biologically active molecules, which frequently constitute only a fraction of the total molecular pool. Here we report a novel development of the approach to determining active concentrations based on surface plasmon resonance (SPR) technology. The method relies on changes in binding rates with varying flow rates under conditions of partial mass transport, and does not require standards of known concentrations, given that the molecular mass of the molecule of interest is known. We introduce an analytical solution to the differential equations describing the formation of a 1:1 bimolecular complex, taking into account both the association and dissociation reactions, under partial mass transport limitations. This solution can be used in global fitting to binding curves obtained at different flow rates. The accuracy, precision, and sensitivity of this approach were determined in experiments involving binding of tyrosine-phosphorylated recombinant proteins to anti-phosphotyrosine antibodies, where the active concentration could be determined independently by in vitro phosphorylation with ^{33}P. There was an excellent agreement between the active concentrations determined by the analytical SPR-based method and by determination of the level of radioactivity of the phosphorylated protein. The SPR-based method allows determination of protein concentrations at picomolar levels. A procedure for accurate determinations of association and dissociation rate constants, based on the analytical solution of the mass transport and binding theory, is outlined.

Biological functions result from myriads of specific interactions between molecules, particularly proteins, where selective binding is the basis for specificity. Thus, in both practical applications and biomedical research, it is crucial to have access to methods for determination of specific binding interactions. A prerequisite for such determinations is knowledge of the concentrations of the participating molecular species. However, since only a fraction of the proteins in a population, which might appear to be identical when analyzed by chemical or physicochemical techniques, may have the conformation or structure needed for the specific binding, it is not enough to determine the total concentration of the chemically identical components. What is needed is the concentration of the components that actively participate in each type of reaction, i.e., the active concentration of that molecular species. With the increased use of recombinant proteins and monoclonal antibodies, both in basic research and in clinical medicine, this is a problem of significant magnitude, as both categories are characterized by the presence of large fractions of molecules, which do not have the right folding and accordingly are inactive.

It is not a trivial task to determine the active concentration of a specific biomolecule. The potency of a biologically active substance or a therapeutic product can be demonstrated by testing its biological activity, using cell cultures, tissue preparations, or whole animals. Most biological assays are, however, not as precise and reproducible as physicochemical methods. Until now, there has been only one physicochemical method described that can be used to determine active concentrations of biomolecules. This method is based on the optical phenomenon of surface plasmon resonance (SPR)[1] that is recorded under conditions of limited mass transport in a flow cell (*1*) where molecular binding is assessed in real time (*2*). A unique feature of this method is that it does not require comparison of unknown samples with standards

† This work was supported by grants to B.Ö. (The Swedish Medical Research Council, Projects 05200 and 13113; The Swedish Natural Science Council, Project B5107; Swedish Cancer Foundation, Project 3957; Polysackaridforskning AB) and N.B. (Canadian Institute for Health Research).

* To whom correspondence should be addressed.
‡ Karolinska Institutet.
§ University of Stockholm.
|| Louisiana State University.
⊥ McGill University.

10.1021/bi020099h CCC: $22.00
Published on Web 06/06/2002

of known concentrations, i.e., a standard calibration curve is not needed. Furthermore, this method can be used to determine active concentrations of biomolecules in a crude mixture of many different molecular species.

In a typical experiment utilizing SPR detection, a selective surface is created by covalent immobilization of a specific ligand to the surface of a sensor chip. A sample (analyte) solution is then injected at a constant flow rate. The detected signal is based on changes in the refractive index, which are proportional to the level of accumulation of mass at the binding surface (*3*). Because the data are recorded versus time, it is possible to determine both the analyte concentration and the rate constants for the analyte−ligand interaction since the signal recorded from the flow cell detector results from a combination of both chemical processes (i.e., binding kinetics) and transport processes (diffusion and flow). The sensorgram, depending on the experimental setup, may reflect the following three conditions: (i) rate-limiting binding, (ii) rate-limiting mass transport, and (iii) a combination of (i) and (ii). Condition (i) applies when the mass transport is much faster than the binding reaction. Condition (ii) describes the opposite situation, when the binding is much faster than the transport of molecules to the surface. A method for determination of antibody concentration, utilizing this condition, has been described by Karlsson et al. (*4*). Conditions (i) and (ii) can be regarded as special cases, because experimentally they are not easily obtained in a flow cell. For a continuous-flow system, e.g., a BIAcore instrument, condition (iii), a partial mass transport limitation, describes a more general situation. At flow rates that are operative in a BIAcore flow cell, the flow is laminar and follows a parabolic velocity profile within the detection chamber, giving rise to a concentration gradient from the center of the channel toward the interacting surface. Thus, during sample injection, the analyte concentration at the surface will be lower than in the bulk flow. If one does not properly account for this, systematic errors can occur in the determination of kinetic binding constants. However, this situation can also be advantageous for active concentration determinations.

Significant progress has been made in the development of methods for obtaining kinetic information from biosensor data. The major approaches, in historical order of appearance, have been linearization (*5*), nonlinear least-squares curve fitting with analytical integration (*6*), and numerical integration in combination with global curve fitting (*7*). A meritorious comparison of these three methods has been performed by Morton et al. (*8*). The first two methods apply for a simple bimolecular interaction ($A + B \leftrightarrow AB$) without mass transport limitations, which considerably limits their versatility for practical analysis, where mass transport and more complex interaction mechanisms may be involved. Until now, an analytical solution has only been worked out for a simple bimolecular interaction (*6*). It has even been stated that "for reaction schemes more complex than a simple bimolecular interaction, analytical integration becomes difficult or impossible" (*8*). The alternative has been to approximate an analytical solution by numerical integration and to combine this with global curve fitting (*7*). With such an approach, interactions can be analyzed in which both chemical binding and mass transport are operating. A simple model based on this approach has been proposed by Myzska et al. (*9*), where the observed binding is treated as a two-step process: transport of analyte to the surface, followed by a reaction with the immobilized ligand. This has been termed the "two-compartment model", which at present serves as the basic module for the majority of the models used to analyze kinetic data obtained by flow cell SPR detection.

Less attention has been devoted to the subject of active concentrations of the molecules subjected to kinetic interaction studies. This problem is usually underestimated in the literature, but for accurate determination of true kinetic constants, it is mandatory to know the real, active concentrations of the interacting components. A theory for molecular binding at conditions of partial mass transport in a two-compartment model was elegantly worked out by Christensen (*1*), who formulated an analytical solution. His approach has been used to determine active concentrations of both purified and complex mixtures of proteins (*10*, *11*). However, in the analytical solution of the differential equations resulting from this theory, Christensen restricted his analysis to the initial binding phase where molecular association dominates, neglecting the effects of molecular dissociation on the overall reaction. Unfortunately, this limits the usefulness of the method, particularly with respect to judging the validity of the concentration determination. We have now been able to solve this problem, and present here a more general and complete analytical solution to the process of macromolecular binding at conditions of partial mass transport, accounting for the whole binding process, including the dissociation. We compared our analytical solution, which is based on the same quasi-steady state approximation that Christensen used, with a numerical solution of the equations describing the two-compartment model, and found an excellent agreement between the two approaches. We also present an experimental verification of the concentration determination based on the analytical approach. Because we now can perform a global analysis of the entire binding phase, we demonstrate that this approach can be used to determine the concentration ranges in which the analysis gives accurate, valid results. Conditions under which the molecular system does not behave ideally, e.g., due to aggregation, can accordingly be identified and sorted out.

THEORY

Basic Theory. The basic theory has been described in detail by Christensen (*1*), and is briefly summarized here. The physical system is composed of a sensor surface bearing a specific ligand, and a ligand-binding analyte in a soluble phase, which is flushed over the surface. In a two-compartment model (*9*), the interaction between the analyte and the ligand can be described by the following two-step process

$$A_{bulk} \underset{k_c}{\overset{k_c}{\rightleftharpoons}} A_{surface} + B \underset{k_d}{\overset{k_a}{\rightleftharpoons}} AB \qquad (1)$$

[1] Abbreviations: SPR, surface plasmon resonance; CEACAM1, carcinoembryonic antigen-related cell adhesion molecule 1; A-TC-QSSA, analytical solution of the two-compartment model with a quasi steady state assumption; NI-TC, numerical integration of the two-compartment model; GST, glutathione *S*-transferase; IRK, insulin receptor kinase; cRU, corrected response units; kRU, kilo response units.

where the first step describes the mass transport of the analyte, A, between the bulk flow and the surface. The second step describes the chemical interaction leading to the formation of a complex between A and the ligand, B, at the surface of the sensor. In eq 1, k_c is used to denote the mass transport coefficient, k_a is the association rate constant, and k_d is the dissociation rate constant. Experimentally, conditions where either mass transport or binding kinetics are exclusively rate-limiting are, however, difficult to reach. Thus, under most circumstances, the system will operate under conditions of partially limiting mass transport, meaning that both steps in eq 1 have to be considered, which can be described by the following equations

$$\frac{d[A_{surface}]}{dt} = \frac{k_c([A_{bulk}] - [A_{surface}]) - k_a[A_{surface}][B] + k_d[AB]}{h_{diff}} \quad (2)$$

and

$$\frac{d[AB]}{dt} = k_a[A_{surface}][B] - k_d[AB] \quad (3)$$

where

$$h_{diff} = \sqrt[3]{\frac{Dh^2wl}{F}} \quad (4)$$

is a characteristic height of the diffusion layer that links the change in concentration expressed per surface area ([AB]) and per volume ([$A_{surface}$]) (*1*). The diffusion coefficient of the analyte is denoted *D*, while *F* is the bulk flow rate; *h*, *w*, and *l* are the height, width, and length of the flow chamber, respectively. If the diffusion coefficient is not known, it can be estimated from the molecular mass of the analyte, utilizing Stoke's law together with the Einstein−Sutherland equation (*12*).

If quasi-steady state conditions are assumed (*13*), i.e., $d[A_{surface}]/dt = 0$, and with rearrangement of eq 2, the following expression for the analyte concentration at the surface is obtained:

$$[A_{surface}] = \frac{k_c[A_{bulk}] + k_d[AB]}{k_c + k_a[B]} \quad (5)$$

where

$$k_c = C_{k_c}\sqrt[3]{\frac{D^2F}{h^2wl_2}} \quad (6)$$

and

$$C_{k_c} = 1.47\frac{1 - (l_1/l_2)^{2/3}}{1 - l_1/l_2} \quad (7)$$

where l_1 and l_2 are the lengths to the start and end, respectively, of the detection area from the inlet of the flow cell (*1*). Equation 5 describes the average of a concentration gradient that forms along the surface of the flow chamber. Insertion of eq 5 into eq 3 gives

$$\frac{d[AB]}{dt} = k_a\left(\frac{k_c[A_{bulk}] + k_d[AB]}{k_c + k_a[B]}\right)[B] - k_d[AB] \quad (8)$$

which describes the formation of a 1:1 complex between A and B under partially limiting mass transport conditions.

Analytical Solution to the Association Phase. Under the experimental conditions in a flow cell biosensor (referring to a BIAcore instrument), the analyte bulk concentration is kept constant by a continuous injection of fresh analyte solution at a constant flow rate. The amount of unoccupied ligand, i.e., [B], decreases with time until equilibrium is reached. Thus, the free ligand concentration, [B], at a particular moment is given by

$$[B] = [B_0] - [AB] \quad (9)$$

where $[B_0]$ is the total amount of ligand on the surface (i.e., available binding sites) at time zero and [AB] is the amount of complex formed at time *t* of analyte injection. The response signal, *R*, measured by an SPR-based sensor is proportional to the amount of complex formed at the detector surface multiplied by the factor M_wG, i.e., $R = M_wG[AB]$, where M_w is the molecular mass of the analyte and *G* is a factor converting the concentration to *R* values ($G = 10000R$ cm²/g of protein) (*3*). When all the immobilized ligand has been fully bound into a complex, $[AB] = [B_0]$ and $R = R_{max}$ (i.e., a theoretical maximum value). Equation 8 may now be written in *R* units by inserting eq 9, multiplying through by the factor M_wG, and setting $[A_{bulk}]$ equal to C_A

$$\frac{dR}{dt} = k_a(R_{max} - R)\left[\frac{M_wGk_cC_A + k_dR}{M_wGk_c + k_a(R_{max} - R)}\right] - k_dR \quad (10)$$

Dividing through by R_{max} and setting θ equal to R/R_{max} transforms this equation into

$$\begin{aligned}\frac{d\theta}{dt} &= k_a(1-\theta)\left[\frac{(M_wGk_cC_A)/R_{max} + k_d\theta}{(M_wGk_c)/R_{max} + k_a(1-\theta)}\right] - k_d\theta \\ &= k_d(1-\theta)\left[\frac{(M_wGk_cC_A)/(k_dR_{max}) + \theta}{(M_wGk_c)/(k_aR_{max}) + 1 - \theta}\right] - k_d\theta \\ &= k_d(1-\theta)\left(\frac{\alpha + \theta}{\beta + 1 - \theta}\right) - k_d\theta \end{aligned} \quad (11)$$

where α and β are the dimensionless parameters

$$\alpha = \frac{M_wGk_cC_A}{k_dR_{max}}, \beta = \frac{M_wGk_c}{k_aR_{max}} \quad (12)$$

Next, when $\tau = k_dt$ and one divides through by k_d, the pure dimensionless form of the equation becomes

$$\frac{d\theta}{d\tau} = (1-\theta)\left(\frac{\alpha + \theta}{\beta + 1 - \theta}\right) - \theta \quad (13)$$

This is a separable differential equation that may be solved by integration (*14*), giving the solution

$$\tau = \frac{\theta}{\alpha + \beta} - \frac{\beta(1 + \alpha + \beta)}{(\alpha + \beta)^2}\ln\left[\frac{\alpha - (\alpha + \beta)\theta}{\alpha}\right] \quad (14)$$

explicit in time, taking into account the equalities $\theta = 0$ at $\tau = 0$. This form is not the most desirable for parameter estimation purposes, so we set out to transform this result to yield a form explicit in θ. This can be done by first making the variable substitution

$$\phi = \frac{\alpha - (\alpha + \beta)\theta}{\beta(1 + \alpha + \beta)} \tag{15}$$

which implies that $\alpha - (\alpha + \beta)\theta = \beta(1 + \alpha + \beta)\phi$ and $\theta = [\alpha - \beta(1 + \alpha + \beta)\phi]/(\alpha + \beta)$ and using this substitution in eq 14 to give

$$\tau = \frac{\alpha - \beta(1 + \alpha + \beta)\phi}{(\alpha + \beta)^2} - \frac{\beta(1 + \alpha + \beta)}{(\alpha + \beta)^2} \ln\left[\frac{\beta(1 + \alpha + \beta)\phi}{\alpha}\right]$$

$$= \frac{1}{(\alpha + \beta)^2}\left(\alpha - \beta(1 + \alpha + \beta)\left\{\phi + \ln\left[\frac{\beta(1 + \alpha + \beta)\phi}{\alpha}\right]\right\}\right)$$

which is equivalent to

$$\frac{\alpha - (\alpha + \beta)^2\tau}{\beta(1 + \alpha + \beta)} = \phi + \ln\left[\frac{\beta(1 + \alpha + \beta)\phi}{\alpha}\right] \tag{16}$$

Applying $\exp(x) = e^x$ to both sides of eq 16 then gives

$$\exp\left[\frac{\alpha - (\alpha + \beta)^2\tau}{\beta(1 + \alpha + \beta)}\right] = \exp\left\{\phi + \ln\left[\frac{\beta(1 + \alpha + \beta)\phi}{\alpha}\right]\right\} = \frac{\beta(1 + \alpha + \beta)\phi}{\alpha} e^{\phi}$$

or

$$\phi e^{\phi} = \frac{\alpha}{\beta(1 + \alpha + \beta)} \exp\left[\frac{\alpha - (\alpha + \beta)^2\tau}{\beta(1 + \alpha + \beta)}\right] \tag{17}$$

At this stage, it should be noted that the function $\psi = \phi e^{\phi}$ is invertible for $\phi \geq -1$ and the inverse is usually termed Lambert's *W* function (*15*). In other words, the function $W(\psi)$ is defined for $\psi \geq -1/e$, and it has the property that $W(\phi e^{\phi}) = \phi$ when $\phi \geq -1$. The algorithm for computing the values of $W(x)$ when $x \geq -1/e$, with an implementation in the computer language C, has been worked out by Briggs (*16*). Applying $W(\psi)$ to both sides of eq 17 gives

$$\phi = W(\phi e^{\phi}) = W\left\{\frac{\alpha}{\beta(1 + \alpha + \beta)} \exp\left[\frac{\alpha - (\alpha + \beta)^2\tau}{\beta(1 + \alpha + \beta)}\right]\right\}$$

which after elimination of ϕ, using the substitution (eq 15), becomes

$$\frac{\alpha - (\alpha + \beta)\theta}{\beta(1 + \alpha + \beta)} = W\left\{\frac{\alpha}{\beta(1 + \alpha + \beta)} \exp\left[\frac{\alpha - (\alpha + \beta)^2\tau}{\beta(1 + \alpha + \beta)}\right]\right\}$$

or equivalently

$$\theta = \frac{\alpha}{\alpha + \beta}\left(1 - \frac{\beta(1 + \alpha + \beta)}{\alpha} \times W\left\{\frac{\alpha}{\beta(1 + \alpha + \beta)} \exp\left[\frac{\alpha - (\alpha + \beta)^2\tau}{\beta(1 + \alpha + \beta)}\right]\right\}\right) \tag{18}$$

which is another fully equivalent solution to eq 13, explicit in θ as required.

With a definition of the dimensionless parameters

$$K_1 = \frac{\alpha}{\alpha + \beta}, K_2 = \frac{\alpha}{\beta(1 + \alpha + \beta)}, K_3 = \frac{(\alpha + \beta)^2}{\beta(1 + \alpha + \beta)} \tag{19}$$

and given that $\theta = R/R_{max}$ and $\tau = k_d t$, the solution (eq 18) to the original eq 10 becomes

$$R = R_{max}K_1\left[1 - \frac{1}{K_2}W(K_2 e^{K_2 - K_3 k_d t})\right] \tag{20}$$

where

$$K_1 = \frac{C_A k_a}{C_A k_a + k_d} \tag{21}$$

$$K_2 = \frac{C_A k_a{}^2 R_{max}}{(C_A k_a + k_d)k_c M_w G + k_a k_d R_{max}} \tag{22}$$

$$K_3 = \frac{(C_A k_a + k_d)^2 k_c M_w G}{k_d[(C_A k_a + k_d)k_c M_w G + k_a k_d R_{max}]} \tag{23}$$

The solution is somewhat complicated, partly due to the fact that it contains the function $W(x)$ and it has the four parameters C_A, k_a, k_d, and R_{max} in rather complex combinations, but it is an exact solution to the original eq 10 that should be valid whenever mass transport is partially limiting.

Analytical Solution to the Dissociation Phase. In the dissociation phase, differential eq 13 still holds, but since $C_A = 0$, it follows that $\alpha = 0$ which implies that the equation in this case is

$$\frac{d\theta}{d\tau} = (1 - \theta)\left(\frac{\theta}{\beta + 1 - \theta}\right) - \theta \tag{24}$$

which after simplification becomes

$$\frac{d\theta}{d\tau} = -\left(\frac{\beta + 1}{\beta}\frac{1}{\theta} - \frac{1}{\beta}\right)^{-1}$$

This is again a separable differential equation that can be solved by integration, giving

$$\frac{\theta}{\beta} - \frac{\beta + 1}{\beta}\ln\theta - \frac{\theta_0}{\beta} + \frac{\beta + 1}{\beta}\ln\theta_0 = \tau - \tau_0$$

where τ_0 is the time of the start of the dissociation phase and θ_0 is the value of θ at that time, giving

$$\frac{\theta}{\beta} - \frac{\beta + 1}{\beta}\ln\theta = \tau + L \tag{25}$$

where

$$L = \frac{\theta_0}{\beta} - \frac{\beta + 1}{\beta}\ln \theta_0 - \tau_0$$

By using the substitution $\theta = -(1 + \beta)\phi$, eq 25 becomes

$$\frac{\beta + 1}{\beta}\phi + \frac{\beta + 1}{\beta}\ln[-(\beta + 1)\phi] = -(\tau + L)$$

or

$$\phi + \ln(-\phi) = -(\tau + L)\frac{\beta}{\beta + 1} - \ln(\beta + 1)$$

By applying exp to both sides of this equation, we then obtain

$$-\phi e^{\phi} = e^{\phi + \ln(-\phi)} = \frac{1}{\beta + 1}e^{-(\tau + L)\beta/(\beta + 1)}$$

or

$$\phi e^{\phi} = -\frac{1}{\beta + 1}e^{-(\tau + L)\beta/(\beta + 1)} \quad (26)$$

Now we can apply Lambert's *W* function to both sides of this equation to obtain

$$\phi = W(\phi e^{\phi}) = W\left[-\frac{1}{\beta + 1}e^{-(\tau + L)\beta/(\beta + 1)}\right] \quad (27)$$

and substituting back ϕ gives the solution

$$\theta = -(\beta + 1)W\left[-\frac{1}{\beta + 1}e^{-(\tau + L)\beta/(\beta + 1)}\right] \quad (28)$$

The negative signs in eq 28 are not an error. It can be proved mathematically that

$$\frac{1}{\beta + 1}e^{-(\tau + L)\beta/(\beta + 1)} \leq e^{-1} \quad (29)$$

whenever $\tau \geq \tau_0$. Therefore, the argument to $W(\psi)$ in the equation always lies between $-e^{-1}$ and 0, implying that the value is defined and negative (see Appendix A). The negative sign at the start of the expression then makes the total result positive.

Via definition of the parameters

$$K_4 = \beta + 1,\ K_5 = \frac{\beta}{\beta + 1} \quad (30)$$

and given that $\theta = R/R_{max}$, $\tau = k_d t$, and $R = R_0$ at t_0, which is the start of the dissociation phase, the solution (eq 28) to the original eq 24 becomes

$$R = -R_{max}K_4 W\left\{-\frac{R_0}{R_{max}K_4}e^{-[R_0/R_{max}K_4 + K_5 k_d(t - t_0)]}\right\} \quad (31)$$

or

$$R = -K_6 W\left\{-\frac{R_0}{K_6}e^{-[R_0/K_6 + K_5 k_d(t - t_0)]}\right\} \quad (32)$$

where

$$K_4 = \frac{M_w G k_c + k_a R_{max}}{k_a R_{max}} \quad (33)$$

$$K_5 = \frac{M_w G k_c}{M_w G k_c + k_a R_{max}} \quad (34)$$

$$K_6 = \frac{M_w G k_c + k_a R_{max}}{k_a} \quad (35)$$

EXPERIMENTAL PROCEDURES

Reagents. All reagents for interaction analysis, including sensor chips CM5 and SA, surfactant P20, *N*-hydroxysuccinimide (NHS), *N*-ethyl-*N*′-[(3-dimethylamino)propyl]carbodiimide hydrochloride (EDC), etanolamine-HCl, polyclonal rabbit anti-mouse Fcγ antibody (RAM Fcγ:BR-1000-57), and rabbit anti-mouse IgG1 (α-mouse IgG1:BR-1000-55), were obtained from Biacore AB (Uppsala, Sweden). Glutathione-Sepharose, Hi-Trap protein G columns, and adenosine 5′-[γ-^{33}P]triphosphate ($\geq$2500 Ci/mmol) were from Amersham Pharmacia Biotech (Uppsala, Sweden). Yeast extract was from Merck KGaA (Darmstadt, Germany). Isopropyl thio-β-D-galactoside (IPTG) was from Saven Biotech AB (Malmö, Sweden). Ampicillin was from DUCHEFA (Haarlem, The Netherlands). β-Insulin receptor kinase (206104) was from Stratagene.

Preparation of Recombinant Proteins. The construction of recombinant proteins of wild-type and mutated cytoplasmic domains of mouse CEACAM1-4L (*17*), herein termed CEACAM1-L [Lcyt-wt, Lcyt-Y515F (herein denoted YF), Lcyt-F488Y (herein denoted FY), and Lcyt-Y488F/Y515F (herein denoted FF)], fused with GST using the pGEX-2T vector system, has been described previously (*18*). Proteins were produced in *Escherichia coli* BL21. Protein synthesis was induced with IPTG (0.2 mM). Purification of the GST fusion proteins was performed by affinity adsorption on glutathione-Sepharose according to a standard protocol from the manufacturer (Pharmacia). Buffer exchange and further purification of recombinant proteins was carried out on a Superose 12 prepacked column attached to a FPLC 500 system (Pharmacia AB). Concentrations of purified GST−Lcyt proteins were determined with absorption spectroscopy, using the following molar absorptivity values (280 nm): 43 480 M^{-1} cm^{-1} for GST−Lcyt-wt, 42 200 M^{-1} cm^{-1} for the singly mutated forms GST−Lcyt-Y488F and GST−Lcyt-Y515F, and 40 920 M^{-1} cm^{-1} for the doubly mutated form GST−Lcyt-Y488F/Y515F and for GST. Further characterization was performed by SDS−PAGE on 10% gels (*19*) stained with Coomassie brilliant blue.

Protein Phosphorylation and Concentration Determination. Phosphorylation of GST−Lcyt proteins was catalyzed by a recombinant kinase domain of the human insulin receptor [β-insulin receptor kinase (IRK)]. A freshly prepared aqueous stock solution of 5 mM ATP (determined from a molar absorptivity of 15.4×10^3 at 259 nm and pH 7.0) and a 10-fold dilution of the [γ-^{33}P]ATP stock solution were analyzed on a reverse phase R2/M 4.6/100 column (PerSeptive Biosystems) attached to a Waters 2690 chromatography system. Both stocks were verified to be >98% pure ATP with regard to hydrolyzed contaminants. Substrate concentra-

tions in the phosphorylation reaction were 28.7 μM (1.0 mg/mL) GST−Lcyt and 0.50 mM ATP; the radiochemical concentration of [γ-^{33}P]ATP was 0.25 μCi/μL. The phosphorylation was carried out in a solution consisting of 2 mM $MnCl_2$, 10 mM $MgCl_2$, 50 mM Tris-HCl (pH 7.4), 1 mg/mL BSA, and 0.5 unit/μL IRK for 2 h at 37 °C. The reaction solutions were then aliquoted on ice, frozen in liquid nitrogen, and stored at −72 °C. Quantification of the GST−Lcyt phosphorylation levels was performed by SDS−PAGE on 10% polyacrylamide gels (*19*), followed by determination of the amount of radioactivity in the separated protein bands. Samples (1 μg of GST−Lcyt/well) were loaded onto a 16 cm × 16 cm flat bed gel in a Protean II xi 2-D cell (Bio-Rad) together with prestained molecular mass markers (GIBCO BRL, Life Technologies, catalog no. 10748-010). Following electrophoresis, the gels were fixed at room temperature for 10 min in 5% acetic acid and 20% methanol, washed in 1% glycerol, and dried onto Whatman 3 MM paper. Radioactive standards were prepared directly from the [γ-^{33}P]ATP stock solution, by pipetting quantified amounts (in triplicate) onto Whatman 3 MM paper. After drying at room temperature, the standards were exposed together with the dried gels to a FujiFilm imaging plate for 10 h and analyzed in a Fujifilm Bas-2500 Bio-Imaging Analyzer utilizing the Imaging Gauge (version 3.41) analyzing software (Fuji Photo Film Co., Ltd., Tokyo, Japan). Concentrations (means and standard deviations) of the radiolabeled proteins were calculated from the standard curves that were analyzed by a least-squares curve fitting procedure with IGOR Pro (version 4.01.A, WaveMetrics, Inc.).

Monoclonal Antibodies. Two mouse monoclonal antibodies were used: Mab 4G10 which recognizes protein phosphotyrosine motifs and Mab 3G1 which was raised against a ssDNA binding protein (*20*; a kind gift from B. Daneholt, Department of Cell and Molecular Biology, Karolinska Institutet). Both monoclonal antibodies were of subclass IgG1 and were purified from hybridoma supernatants by chromatography on protein G columns.

Peptides. N-Terminally biotinylated dodecameric peptides spanning the Y488 and Y515 regions of mouse CEACAM1-L, VDDVAY(488)TVLNFN and ATETVY(515)SEVKKK, respectively, were purchased from K. J. Ross-Petersen AS (Horsholm, Denmark). Both unphosphorylated and tyrosine-phosphorylated peptides of both kinds were produced. All four peptides were highly homogeneous and >95% pure as demonstrated by amino acid analysis, HPLC, and MALDI-TOF mass spectrometry.

Surface Plasmon Resonance Measurements. Measurements based on surface plasmon resonanse (SPR) technology were carried out with a BIAcore 2000 instrument (Biacore AB). Protein interaction analyses were performed at 25 °C in 10 mM Hepes, 150 mM NaCl, 3.4 mM EDTA, and 0.005% surfactant P20 (pH 7.4) (HBS). Samples were kept at 2 °C, prior to injection. Data modification (including scale transformation and background subtraction) was performed with the program Bia-Evaluation 3.0.

Data Analyses. All calculations based on numerical approaches, including simulations and global curve fitting for estimation of active concentrations, were performed with IGOR Pro (version 4.01.A, WaveMetrics, Inc.). Special curve fit functions based on eq 20 were created and used together with the Global Fit Procedure in IGOR Pro, based on a nonlinear least-squares method, utilizing the Levenberg−Marquard algorithm (*21*). For numerical integration, a fifth-order Runge−Kutta−Fehlberg method was applied (*21*). Simulations of binding interactions were performed according to the two-compartment model, by numerical integration of the two coupled ordinary differential equations (eqs 2 and 3). The flow cell diffusion layer height (h_{diff}) and the mass transport coefficient (k_c) at a given flow rate were obtained from eqs 4, 6, and 7, using the following values for the flow cell dimensions, as specified by the manufacturer: $l = 2.4$ mm, with a detection interval defined by $l_1 = 0.3$ mm and $l_2 = 2.0$ mm; $h = 0.05$ mm; and $w = 0.5$ mm. Since h_{diff} changes continually along the flow cell, we used an average value for h_{diff} in the computational analyses. The average $(h_{diff})_{av}$ was defined as $(h_{diff})_{av} = [(h_{diff})_{start} + (h_{diff})_{stop}]/2$. The l values used to calculate $(h_{diff})_{start}$ and $(h_{diff})_{stop}$ according to eq 4 were l_1 and l_2, respectively.

For calculations of C_A, the M_wG and k_c parameters were given fixed values (calculated as described above), while R_{max}, k_a, k_d, and C_A were defined as free-running parameters to be optimized by the global fit routine. Correlations of C_A, R_{max}, k_a, and k_d were calculated from a covariance matrix, obtained with the Levenberg−Marquard algorithm (*21*).

RESULTS

Theory Evaluation and Simulations. Our solution is based on a two-compartment (TC) model, which Myszka et al. (*22*) have demonstrated accurately and reliably describes the mass transport and binding processes occurring inside a BIAcore flow cell. A crucial assumption in our analytical solution was that the surface concentration, $[A_{surface}]$, can be regarded as being in a quasi steady state (*13*). We refer to this as the A-TC-QSSA model (analytical solution of the two-compartment model with a quasi steady state assumption). We then asked if the quasi steady state assumption would hold for practical applications, i.e., if the A-TC-QSSA model would correctly describe the binding events in the TC model. To answer this question, we simulated the events occurring in the TC model by numerical integration of the coupled differential equations (eqs 2 and 3). We refer to this procedure as NI-TC (numerical integration of the two-compartment model). The NI-TC procedure was used to simulate both the change in analyte surface concentration during the period of sample injection and formation of the analyte−ligand complex. We then simulated the formation of the analyte−ligand complexes with the A-TC-QSSA model, and compared the binding curves with those generated by the NI-TC procedure. Parameters were selected to favor dominance of mass transport limitation for the simulated reactions (i.e., high surface binding capacity together with relatively high k_a values). Figure 1 illustrates three sets of binding interactions, simulated for identical analyte molecular mass (35 kDa), concentration (10 nM), flow rate (50 μL/min), and surface binding capacity ($R_{max} = 1200$ RU), but differing with regard to the rate constants, k_a and k_d. Figure 1A illustrates the changes in analyte surface concentration during sample injection, from NI-TC simulations. The simulations describe a range of conditions, from a partially limited mass transport ($k_a = 6 \times 10^6$ M^{-1} s^{-1}, $k_d = 0.02$ s^{-1}) to a dominant mass transport limitation (with a 10-fold higher association rate constant, $k_a = 6 \times 10^7$, $k_d = 0.02$ s^{-1}). Figure 1B represents the corresponding changes in the

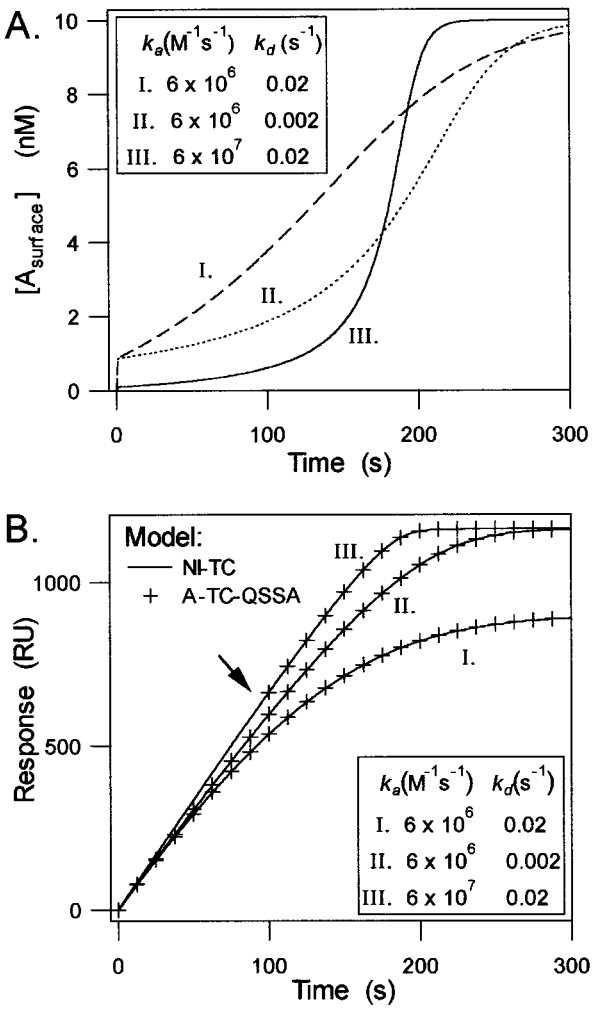

FIGURE 1: Simulation of binding responses using two different mathematical approaches. Three parameter combinations (I, II, and III) with increasing mass transport limitations were analyzed. (A) The average free analyte concentration, [$A_{surface}$], at the sensor surface was calculated by numerical integration using the NI-TC model. (B) The formation of complexes between analyte and ligand on the sensor surface was assessed using the NI-TC model (shown as continuous lines) or by analytical solution of the A-TC-QSSA model (shown as crosses). The parameter combinations were as follows. Shared parameters in combinations I, II, and III were as follows: $M_w = 35$ kDa, $D = 9.09 \times 10^{-7}$ cm^2 s^{-1}, $F = 50$ μL/min, $C_A = 10$ nM, and $R_{max} = 1200$ RU. The unique parameters in combinations I, II, and III were k_a and k_d (see the figure). Note that no analytical solution is defined for condition III, at the time interval of <100 s (indicated with an arrow in panel B).

formation of the analyte−ligand complex for these conditions. Overlaid on the results simulated by the NI-TC procedure are results for the same parameter sets calculated with the A-TC-QSSA model. As demonstrated in Figure 1B, the two models return identical results. A limitation for the A-TC-QSSA model is observed for the more extreme case ($k_a = 6 \times 10^7$ M^{-1} s^{-1}), where the numerical values of ψ in $W(\psi)$ become less than -1/e in the time interval of 0−100 s. For this particular parameter combination, an A-TC-QSSA solution is not defined (curve III in Figure 1B). This situation can be avoided by altering the experimental conditions. In this case, it is enough to decrease the sample concentration 10-fold, i.e., to 1 nM, to bypass the limitation (data not shown). In conclusion, the results of these simulations demonstrate that the quasi steady state assumption for [$A_{surface}$] is a valid simplification with regard to the rate of complex formation, d[AB]/dt.

Next, we investigated how well our analytical model recovered the values of C_A, R_{max}, k_a, and k_d by global curve fitting to noisy data generated by the NI-TC simulation process (Figure 2). Two conditions with different R_{max} values were introduced, keeping the other variables constant. The condition of a weak partial mass transport was simulated by the introduction of a low-capacity binding surface, setting R_{max} to 50 RU (Figure 2A). For the same set of parameters, a condition of dominant mass transport was simulated by setting R_{max} equal to 500 RU (Figure 2B). Time courses for the average free analyte concentrations at the sensor surface, [$A_{surface}$], are shown in Figure 2C. Curve fitting to simulated data, free of noise, resulted in recovery of true values for all variables (data not shown). To see how noise affected the recovery of the original parameters, a pseudonoise of 0.4 RU with random distribution was added to both data sets (a noise level of 0.4 RU is a realistic value for the Biacore 2000 instrument as operated in this investigation). For comparison, we also applied the NI-TC model as a fit routine for the global fit procedure. Global fitting by both models gave almost identical results (Figure 2A,B,D). It should be noted that both graphs (panels A and B of Figure 2) have two sets of curve fits, resulting from fitting the data to the A-TC-QSSA and NI-TC models, respectively, which overlap completely and therefore appear as a single fit for each parameter set. A plot of the residuals from one-third of the data set from the weak mass transport limitation (sample flow of 100 μL/min, from Figure 2A) demonstrated an almost identical fit result for both models (Figure 2D).

To perform the global curve fitting, one has to provide an initial set of parameter values from which to start. When simulations were made without added noise, the global curve fitting analyses returned correct values of C_A, R_{max}, k_a, and k_d for any set of initial values. However, with noisy data, different parameter combinations may give equally good curve fittings, especially when the parameters have a high degree of coupling. To analyze how the models behaved in this respect, and how the parameters were correlated with each other, we analyzed the same set of simulated data, starting from extremely different initial values for C_A, R_{max}, k_a, and k_d (Tables 1 and 2). Again, both models exhibited similar characteristics; however, the four variables behaved differently. Both models returned realistic values for C_A regardless of the initial values given, although the A-TC-QSSA model gave somewhat more accurate values. The calculated values for the other three variables, k_a, k_d, and R_{max}, varied extensively, especially the values for k_a and R_{max}. Meaningful values for these parameters were only obtained in product combinations. The correlations were complex; however, a clear feature was the consistent high level of correlation between k_a and R_{max}. These properties and behavior of the four parameters were predicted by mathematical analysis of the A-TC-QSSA model, as demonstrated in Appendix B.

Determination of Active Concentrations by the Global Curve Fitting Routines. The validity of the A-TC-QSSA model for determining active concentrations was tested experimentally. For this purpose, a GST fusion protein of the mouse CEACAM1-L cytoplasmic domain (GST−Lcyt)

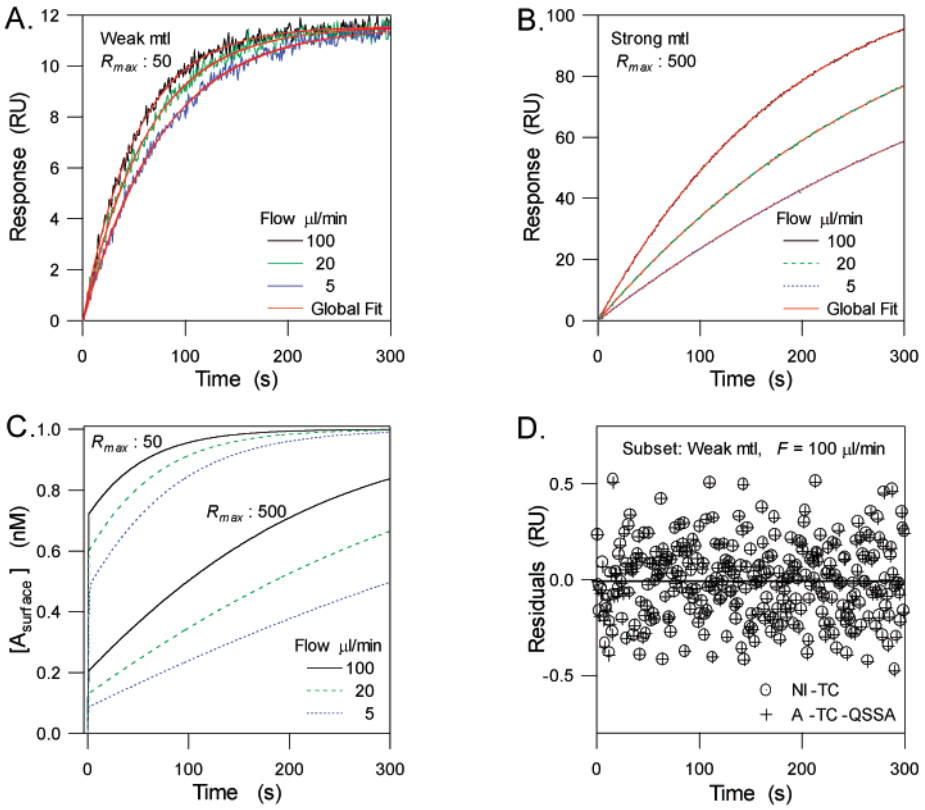

FIGURE 2: Effect of signal noise on the recovery of original parameters by the different mathematical models. Binding data (sensorgrams and free analyte surface concentrations) were generated by simulations using the NI-TC model for two different conditions distinguished by weak [(A) $R_{max} = 50$ RU] or strong [(B) $R_{max} = 500$ RU] mass transport limitation. Shared parameters were as follows: $M_w = 35$ kDa, $D = 9.09 \times 10^{-7}$ cm^2 s^{-1}, $C_A = 1$ nM, $k_a = 6.5 \times 10^6$ M^{-1} s^{-1}, and $k_d = 0.02$ s^{-1}. A random noise of 0.4 RU was added to the simulations, which were carried out for three different flow rates (F) at 5, 20, and 100 μL/min, respectively. The simulated data were analyzed by global curve fitting using both the NI-TC model and the A-TC-QSSA model. The following initial values were used in the global curve fitting analysis: $C_A = 20$ nM, $R_{max} = 1000$ RU, $k_a = 6.0 \times 10^5$ M^{-1} s^{-1}, and $k_d = 1.0 \times 10^{-3}$ s^{-1}. Fit results are listed in Tables 1 and 2. (C) Average free analyte concentration, [$A_{surface}$], at the sensor surface calculated by the NI-TC model at three different flow rates under strong and weak mass transport limitation conditions. (D) Residuals, calculated with the NI-TC (○) and A-TC-QSSA (+) models for the subset of data belonging to a flow rate of 100 μL/min, at conditions of a weak mass transport limitation (from data shown in panel A).

Table 1: Parameter Recovery Analysis for a Simulated Interaction Describing a Weak Partial Mass Transport Limitation[a]

				correlation coefficient		
parameter	initial	estimate	SD	C_A	R_{max}	k_d
A-TC-QSSA						
C_A (nM)	20	1.053	2.91×10^{-2}			
R_{max} (RU)	1000	423	823	0.80		
k_d (s^{-1})	1.0×10^{-3}	2.53×10^{-2}	1.35×10^{-3}	0.67	0.97	
k_a (M^{-1} s^{-1})	6.0×10^5	6.80×10^5	1.33×10^6	−0.81	1.0	−0.97
C_A (nM)	5.0×10^{-2}	0.954	2.3×10^{-2}			
R_{max} (RU)	12	42.46	6.35	0.74		
k_d (s^{-1})	1.0×10^{-3}	1.93×10^{-2}	1.1×10^{-3}	0.59	0.97	
k_a (M^{-1} s^{-1})	6.0×10^7	7.6×10^6	1.3×10^6	−0.84	−0.99	−0.93
NI-TC						
C_A (nM)	20	1.106	1.9×10^{-2}			
R_{max} (RU)	1000	395	73.2	0.09		
k_d (s^{-1})	1.0×10^{-3}	2.50×10^{-2}	3.3×10^{-4}	−0.72	0.46	
k_a (M^{-1} s^{-1})	6.0×10^5	6.83×10^5	1.29×10^5	−0.23	−0.99	−0.34
C_A (nM)	5.0×10^{-2}	1.001	2.5×10^{-2}			
R_{max} (RU)	12	47.18	7.57	0.72		
k_d (s^{-1})	1.0×10^{-3}	1.98×10^{-2}	1.01×10^{-3}	0.56	0.97	
k_a (M^{-1} s^{-1})	6.0×10^7	6.39×10^6	1.16×10^4	−0.82	−0.99	−0.93

[a] The table shows results obtained from global fit analysis based on the A-TC-QSSA and NI-TC models when two different sets of initial values were used in the global fit procedures. The true parameter values used in the simulations were as follows: $C_A = 1.0$ nM, $R_{max} = 50$ RU, $k_d = 0.020$ s^{-1}, and $k_a = 6.0 \times 10^6$ M^{-1} s^{-1}, with an added random noise of 0.4 RU. Standard deviations and correlation coefficients were calculated with IGOR Pro.

was chosen as the analyte, because the active concentration of this protein could be determined by an alternative method based on radioactivity.

The Lcyt domain consisting of 73 amino acids has two tyrosine residues at positions 488 and 515 (*18*), which could be phosphorylated by the human insulin receptor kinase.

Table 2: Parameter Recovery Analysis for a Simulated Interaction Describing a Strong Partial Mass Transport Limitation[a]

				correlation coefficient		
parameter	initial	estimate	SD	C_A	R_{max}	k_d
A-TC-QSSA						
C_A (nM)	20	0.9973	1.5×10^{-3}			
R_{max} (RU)	1000	515.1	15.2	0.84		
k_d (s^{-1})	1.0×10^{-3}	2.02×10^{-2}	1.5×10^{-4}	0.38	0.74	
k_a (M^{-1} s^{-1})	6.0×10^{5}	5.87×10^{6}	1.85×10^{5}	−0.90	−0.99	−0.63
C_A (nM)	0.1	0.997	1.17×10^{-2}			
R_{max} (RU)	60	506.58	14.7	0.84		
k_d (s^{-1})	1.0×10^{-3}	2.02×10^{-2}	1.5×10^{-4}	0.38	0.74	
k_a (M^{-1} s^{-1})	6.0×10^{7}	5.98×10^{6}	1.85×10^{-5}	−0.90	−0.99	−0.63
NI-TC						
C_A (nM)	20	1.0047	6.6×10^{-4}			
R_{max} (RU)	1000	572	1.8	0.17		
k_d (s^{-1})	1.0×10^{-3}	2.04×10^{-2}	1.0×10^{-4}	−0.66	0.08	
k_a (M^{-1} s^{-1})	6.0×10^{5}	5.87×10^{6}	1.9×10^{5}	−0.75	−0.57	0.75
C_A (nM)	5.0×10^{-2}	0.99477	6.6×10^{-4}			
R_{max} (RU)	12	432.5	1.5	0.22		
k_d (s^{-1})	1.0×10^{-3}	1.93×10^{-2}	1.0×10^{-4}	−0.63	0.15	
k_a (M^{-1} s^{-1})	6.0×10^{7}	6.97×10^{6}	4.3×10^{4}	−0.76	−0.62	0.67

[a] The table shows results obtained from global fit analysis based on the A-TC-QSSA and NI-TC models when two different sets of initial values were used in the global fit procedures. The true parameter values used in the simulations were as follows: $C_A = 1.0$ nM, $R_{max} = 500$ RU, $k_d = 0.020$ s^{-1}, and $k_a = 6.0 \times 10^6$ M^{-1} s^{-1}, with an added random noise of 0.4 RU. Standard deviations and correlation coefficients were calculated with IGOR Pro.

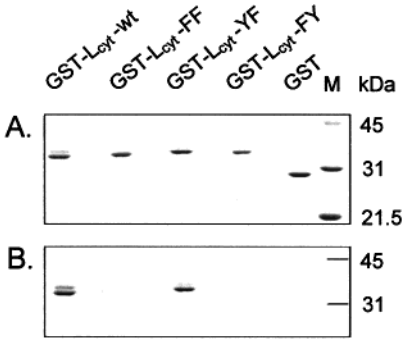

FIGURE 3: Characterization of in vitro-phosphorylated GST−Lcyt proteins. Wild-type and mutated GST−Lcyt fusion proteins were in vitro-phosphorylated with insulin receptor kinase and [^{33}P]ATP. The proteins were analyzed by SDS−PAGE (1 μg/well), and the amount of radioactivity was determined in a Fuji Bio-Imaging analyzer. (A) Proteins stained with Coomassie brilliant blue. (B) Bio-Imaging analyzer-produced autoradiogram of the gel shown in the top panel. Note that only the GST fusion proteins of Lcyt-wt, Lcyt-YF, and Lcyt-FY were phosphorylated, while the double mutant Lcyt-FF or GST did not incorporate any phosphate. M denotes the lane that contained molecular mass marker proteins.

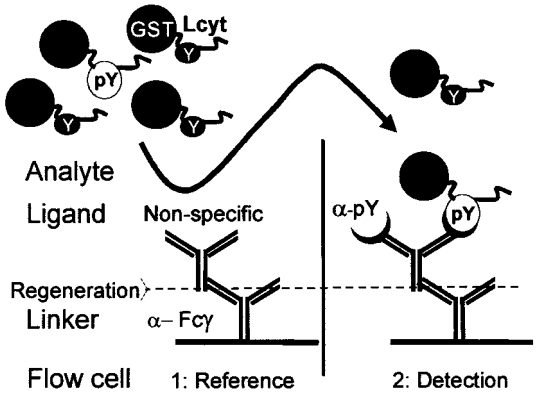

FIGURE 4: Experimental setup for determination of active concentrations. The active concentration in this setup was defined as the fraction of the analyte, GST−Lcyt-YF, that was monophosphorylated with ^{33}P at Y488. Two sensor surfaces in serial alignment were created, flow cell 1 (Fc1) with a nonspecific antibody (Mab 3G1) and flow cell 2 (Fc2) with an anti-phosphotyrosine antibody (Mab 4G10). The analyte solution, containing a mixture of phosphorylated and unphosphorylated GST−Lcyt-YF, was injected over the sensor surface. The active concentration was determined by surface plasmon resonance (SPR) detection of the specific binding of GST−Lcyt-YF to the anti-phosphotyrosine antibody, and compared with the active concentration measured by radioactivity determination.

Phosphorylation utilizing [^{33}P]ATP showed that the wild type (GST−Lcyt-wt) and the Y515F mutant (GST−Lcyt-YF) were effectively phosphorylated; on the other hand, the Y488F mutant (GST−Lcyt-FY) incorporated very little phosphate, and the Y488F/Y515F double mutant (GST−Lcyt-FF) and the GST protein were not phosphorylated at all (Figure 3). Accordingly, the radioactive GST−Lcyt-YF represents a monophosphorylated protein, the concentration of which could be accurately determined from the specific activity of the [^{33}P]ATP and quantification of the incorporated radioactivity. The tyrosine-phosphorylated GST−Lcyt binds specifically to anti-phosphotyrosine antibodies, while unphosphorylated fusion proteins show no binding. Thus, we could use the anti-phosphotyrosine antibody Mab 4G10 as a ligand and the monophosphorylated GST−Lcyt-YF as an analyte in BIAcore experiments. In this system, the concentration of the monophosphorylated GST−Lcyt-YF represents the active concentration of the whole population of GST−Lcyt-YF, which also contains unphosphorylated protein since the in vitro phosphorylation does not result in phosphorylation of all protein molecules.

The experimental setup is described in Figure 4. A covalently bound antibody (α-Fcγ Ab) was used to trap the 4G10 antibody that served as the ligand. Serial dilutions of quantified in vitro-phosphorylated GST−Lcyt-YF were run over a sensor, loaded with a nonspecific antibody (Mab 3G1) in flow cell 1 (Fc1) and an identical amount of the specific ligand antibody (Mab 4G10) in the downstream flow cell (Fc2). Each analyte dilution was injected at three different flow rates (5, 20, and 50 μL/min), to generate a set of mass transport-dependent response curves (sensorgrams), which

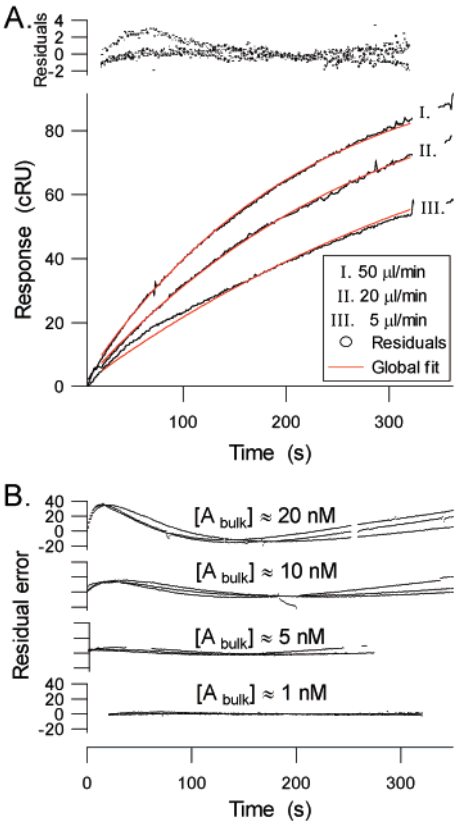

FIGURE 5: Determination of the active concentration of GST−Lcyt-^{33}P-YF. Different concentrations of GST−Lcyt-^{33}P-YF (the analyte) were analyzed with the experimental setup depicted in Figure 4. (A) The analyte, at an active concentration of 1 ± 0.15 nM (determined radiometrically), was run at three different flow rates (5, 20, and 50 μL/min). Global fit analysis based on the A-TC-QSSA model gave a concentration of 0.891 ± 0.006 nM. Plots of the residual errors are shown above the sensorgrams. (B) Plots of residual errors for runs of the analyte at four different concentrations, (1, 5, 10, and 20 nM, determined radiometrically). Each concentration of analyte was analyzed at three different flow rates (5, 20, and 50 μL/min). The residual plots were obtained from the binding curves calculated by global fit analysis according to the A-TC-QSSA model. Note the increasing deviation of binding data from the global fits with increasing analyte concentration.

could be used as input data for the active concentration determination. The curve fitting data for the active concentration determination of in vitro-phosphorylated GST−Lcyt-YF are shown in Figure 5. In these experiments, the active concentration of GST−Lcyt-YF was determined to be 1.0 ± 0.15 nM by the radiometric procedure, which was ~5% of the total concentration (phosphorylated and unphosphorylated protein). The global curve fitting routine utilizing the A-TC-QSSA model yielded curves that coincided perfectly with sensorgrams recorded at low protein concentrations, except at early time points at the lowest flow rate (see curve III in Figure 5A). This deviation is due to a technical limitation of the instrument and can be avoided if flow rates higher than 5 μL/min are used. The concentration determined from the curve fit at these low protein concentrations (taking into account the appropriate dilution factor) resulted in a value of 0.891 ± 0.006 nM, which is in excellent agreement with the active concentration determined by the radiometric analysis. More concentrated solutions yielded sensorgrams that showed systematic deviations from the curves generated by the global analyses (Figure 5B). Values of the active concentrations, as well as of R_{max}, k_d, and k_a, determined by the A-TC-QSSA procedure are listed in Table 3, which also contains pairwise correlation coefficients for the different variables. As in the simulations (Tables 1 and 2), the highest correlation occurred between k_a and R_{max}.

Table 3: Determination of the Active Concentration of GST−Lcyt-^{33}P-YF by the A-TC-QSSA Model[a]

			correlation coefficient		
parameter	estimate	SD	C_A	R_{max}	k_d
C_A (nM)	0.891	0.006			
R_{max} (RU)	1200	508	0.79		
k_d (s^{-1})	0.050	3.8×10^{-3}	−0.23	0.32	
k_a (M^{-1} s^{-1})	6.18×10^6	5.6×10^{-2}	−0.87	−0.99	−0.16

[a] The analyte GST−Lcyt-^{33}P-YF was analyzed at an active concentration of 1 ± 0.15 nM (determined radiometrically) at three different flow rates (5, 20, and 50 μL/min). The parameters were estimated by a global fit analysis of the sensorgrams according to the A-TC-QSSA model. The results are shown graphically in Figure 5A. Pairwise correlations between the different parameters were calculated with the Global Fit Procedure in IGOR Pro.

The failure to fit the sensorgrams obtained at higher concentrations to theoretical binding curves suggested nonideal behavior of the analyte (GST−Lcyt-YF). This is presumably due to dimerization or oligomerization of GST−Lcyt-YF, since it is known that both the Lcyt domain and GST itself can form dimers.

The A-TC-QSSA method was tested additionally, with a different analyte−ligand combination. Now the ligands were synthetic tyrosine-phosphorylated dodecapeptides with amino acid sequences corresponding to the regions around tyrosine residues Y488 and Y515 of the CEACAM1-L cytoplasmic domain. The specific analyte was the phosphotyrosine-binding monoclonal antibody, Mab 4G10. The sensorgrams in Figure 6A demonstrate that while 4G10 bound efficiently to phosphorylated but not to unphosphorylated peptides, neither the unrelated Mab 3G1 nor Mab 5.4 that recognizes the extracellular domain of CEACAM1 bound to the phosphorylated peptides. Determination of the active concentration of protein G-purified Mab 4G10 was analyzed at various concentrations. Dilutions were prepared from a stock solution with a total concentration of 1 mg/mL Mab 4G10 (OD_{280} = 1.45). An assay with a solution diluted 1:10000 (0.1 μg/mL) is shown in Figure 6B. Three runs per flow rate were performed by random order injection. All runs were used for the curve fit calculation, which gave a value for the active concentration of 72 ± 3 pM. Since an IgG concentration of 0.1 μg/mL corresponds to a total molar concentration of approximately 667 pM, this shows that only ~11% of the IgG fraction was active with regard to phosphotyrosine binding. Active concentration determination of a 1:1000 dilution gave a value of 730 ± 3 pM (data not shown), and even at a total concentration of 10 μg/mL there was an excellent agreement with the determinations carried out at higher dilutions.

However, analyses at higher analyte concentrations demonstrated a fundamental problem, not only for determination of active concentrations but also for estimation of R_{max} by a binding saturation approach (Figure 6C). At a total protein concentration of 500 μg/mL, a saturation value of 1420 RU was obtained after injection for 200 s. Injection of a 2-fold higher concentration (1 mg/mL) confirmed the saturation level of 1420 RU. All the lower analyte concentrations

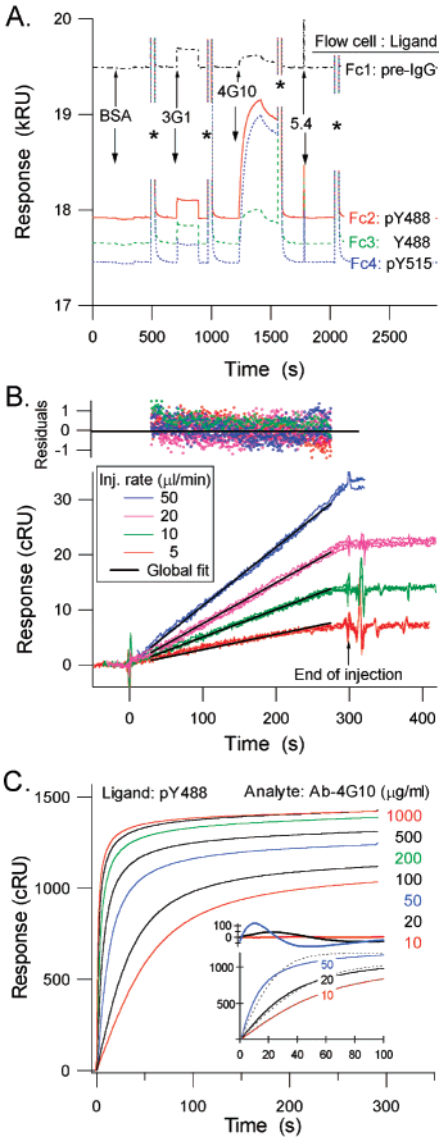

FIGURE 6: Determination of the active concentration of the monoclonal antibody 4G10. The flow cells were derivatized with the following four ligands: mouse IgG for Fc1, a tyrosine-phosphorylated dodecapeptide, pY488-peptide, for Fc2, a nonphosphorylated dodecapeptide, Y488-peptide, for Fc3, and a tyrosine-phosphorylated dodecapeptide, pY515-peptide, for Fc4. (A) Sensorgrams from the four flow cells obtained by consecutive injections of bovine serum albumin (BSA), monoclonal antibody 3G1, monoclonal antibody 4G10, and monoclonal antibody 5.4, at a flow rate of 20 μL/min. Asterisks denote regeneration cycles. (B) Global fit analysis according to the A-TC-QSSA model of Mab 4G10 binding to a surface derivatized with pY488-peptide as ligand. The total protein concentration of 4G10 was 100 ng/mL (total IgG concentration of $\approx$0.67 nM). The Y488-peptide was used as a nonspecific ligand in a different flow cell to monitor background binding. The analyte (Mab 4G10) was run in random order at four different flow rates (5, 10, 20, and 50 μL/min), for three times at each flow rate. The displayed binding curves show data after background subtraction. Residual plots are shown in the same diagram, above the sensorgrams. Global fit analysis gave an analyte concentration of 73 pM. (C) Sensorgrams of the analyte Mab 4G10 run at seven different protein concentrations, ranging from 10 to 1000 μg/mL, at 20 μL/min over a ligand surface consisting of the pY488-peptide. The inset shows sensorgrams (solid lines), A-TC-QSSA global fit curves (dashed lines), and the corresponding residual plots obtained for 10, 20, and 50 μg/mL.

reached values close to equilibrium at an injection time of 300 s. Thus, these measurements indicate a value for R_{max} of 1420 RU. However, when the sensorgrams obtained at different analyte concentrations were analyzed individually according to the A-TC-QSSA model, it was found that only the lowest concentration (10 μg/mL) could be fitted to a 1:1 model of molar binding between analyte and ligand (Figure 6C inset). Neither of the higher analyte concentrations conformed to a 1:1 molar binding model, despite mass transport being included in the model. In fact, the sensorgrams obtained at all concentrations above 10 μg/mL Mab 4G10 did not fit to any model that was tested. The most likely explanation to this behavior is that the analyte, Mab 4G10, forms aggregates of various sizes at higher concentrations. Accordingly, the saturation level of 1420 RU does not represent a true R_{max} value corresponding to binding of the analyte to the ligand in a simple 1:1 molar ratio, but in addition reflects multi-equilibrium binding complexes within the analyte.

DISCUSSION

Flow cell biosensor analysis has become the method of choice for determining kinetic rate constants for interacting biomolecules. A potential problem is that, in this method, the observed binding kinetics reflect not only the chemical binding but also the transport of molecules from the moving fluid phase to the binding surface. However, this can be turned into an advantage, since it can be used to determine the concentration of the molecular species that participates actively in the binding interaction. Here we present a complete analytical solution to the entire process of mass transport and binding kinetics, which allow both determination of the active concentration and accurate and precise measurements of the association and dissociation rate constants of the binding interaction. This solution has two advantages over existing solutions. First, since we used the entire processes of mass transport, association, and dissociation in our solution, there are no approximations or limiting assumptions made that restrict the applications of our procedure. Second, potential inaccuracies and algorithmic problems connected with solutions based on numerical integration can be avoided.

To evaluate the performance of our solution for determination of active concentrations of biomolecules, we compared its outcome with the results from another, independent method for determining the active concentration. To that end, we devised a radiochemical procedure in which we could determine the active concentration of a tyrosine-monophosphorylated protein binding to an antibody specifically recognizing phosphotyrosine. The two methods gave values in excellent agreement with each other, demonstrating the validity of our analytical solution.

Our method of determining active concentrations is precise and extremely sensitive. Picomolar concentrations of proteins can be determined, and there is no requirement for a standard curve made from a reference protein. Another advantage is that the curve fitting procedure explicitly shows if the results are reliable. If the residuals are not randomly distributed but show a systematic deviation along the time scale, as in Figure 5B at concentrations of 5−20 nM, there is an indication of nonideal behavior of the analyte, with regard to a predicted 1:1 binding, for which the model accounts. When this occurs, the measured concentration values cannot be trusted, and the

determination should be repeated at a lower analyte concentration.

Application of our analytical solution to global analysis of sensorgrams obtained at increasing analyte concentrations also demonstrated a potential problem in experimental determinations of R_{max}. Thus, when increasing concentrations of an anti-phosphotyrosine antibody were injected, an increasing deviation from the theoretical binding curve was observed, which can be explained by oligomerization of the antibody, occurring at higher concentrations. When this happens, the binding plateau that is reached does not represent a true value of R_{max} that can be translated to the number of binding sites in the solid phase for a monomeric analyte in the fluid phase. This has far-reaching practical consequences for determinations of equilibrium binding constants by the commonly used Scatchard analysis. In such analyses, the binding equilibrium is determined at various concentrations of the soluble binding partner. From these binding data, the equilibrium dissociation constant, K_D, as well as the maximal number of binding sites in the solid phase can be determined. However, this is valid only as long as the soluble binding molecules behave as an ideal monomeric species. If oligomerization takes place at higher concentrations, an excessively large binding will be recorded. In many instances, curvilinear plots have been observed with a variety of different interacting molecules. These plots have generally been interpreted as the occurrence of different classes of binding sites or receptors with different binding affinities, or as negative cooperativity. However, one has to consider oligomerization of the soluble binding partner at increasing concentrations as an alternative explanation in certain cases. Our analytical procedure can be used to determine the concentration ranges in which the binding partners behave ideally.

For determination of active concentrations, only the injection phase of the binding experiments performed at different flow rates is used. As shown in Appendix B, analysis of the injection phase alone at different flow rates gives a meaningful value only for C_A, i.e., the active concentration, whereas meaningful values for the other unknown parameters, k_a, k_d, and R_{max}, can only be obtained as the combined expression k_aR_{max} or $(k_aR_{max})/k_d$. Therefore, the active concentration and the association and dissociation binding rate constants must be determined separately in different experimental setups. We propose the following procedure for this purpose. First, the active concentration of the soluble analyte is determined by global analysis of sensorgrams obtained at different flow rates, low analyte concentrations, and high density of the immobilized ligand. In this step, only the injection phase is analyzed. Several concentrations of the analyte should be analyzed to determine the concentration range in which the analyte behaves ideally. Then, several different analyte concentrations within the ideal concentration range are analyzed at one single high flow rate, and a low density of the immobilized ligand. The association and dissociation rate constants, k_a and k_d, respectively, are determined by global analysis of both the injection (association) and elution (dissociation) phases, utilizing the value of the active concentration determined in the previous step. In our opinion, this procedure seems to be the best and most accurate approach presently available for determining binding rate constants as well as binding equilibrium constants for interacting biomolecular systems.

APPENDIX

(A) The Argument to W(ψ) in the Dissociation Phase

In the solution to the dissociation phase (eq 28), the argument to W is negative. This makes sense as long as the argument lies between $-e^{-1}$ and 0, giving a negative answer that in turn explains the negative sign in front of the expression in the solution. That this is the case is shown as follows.

The solution (eq 28)

$$\theta = -(\beta + 1)W\left[-\frac{1}{\beta + 1}e^{-(\tau+L)\beta/(\beta+1)}\right]$$

makes sense if and only if

$$-\frac{1}{\beta + 1}e^{-(\tau+L)\beta/(\beta+1)} \geq -e^{-1}$$

if and only if

$$\frac{1}{\beta + 1}e^{-(\tau+L)\beta/(\beta+1)} \leq e^{-1}$$

if and only if

$$e^{-(\tau+L)\beta/(\beta+1)} \leq (\beta + 1)e^{-1}$$

if and only if

$$-(\tau + L)\beta/(\beta + 1) \leq \ln(\beta + 1) - 1$$

if and only if

$$\tau + L \geq \frac{\beta + 1}{\beta}[1 - \ln(\beta + 1)]$$

if and only if [recalling that $L = \theta_0/\beta - \ln(\theta_0)\ (\beta + 1)/\beta - \tau_0$].

$$\tau - \tau_0 + \frac{\theta_0}{\beta} - \frac{\beta + 1}{\beta}\ln\theta_0 \geq \frac{\beta + 1}{\beta}[1 - \ln(\beta + 1)]$$

Since τ is always greater than or equal to τ_0, this implies that it is sufficient that

$$\frac{\theta_0}{\beta} - \frac{\beta + 1}{\beta}\ln\theta_0 \geq \frac{\beta + 1}{\beta}[1 - \ln(\beta + 1)]$$

which holds if and only if

$$\frac{\theta_0}{\beta + 1} - \ln\theta_0 \geq 1 - \ln(\beta + 1)$$

if and only if

$$\frac{\theta_0}{\beta + 1} - \ln\theta_0 - 1 + \ln(\beta + 1) \geq 0 \qquad \text{(A1)}$$

For a sufficiently large β, this is obviously true, and by differentiating the left-hand side with respect to β, we obtain

$$-\frac{\theta_0}{(\beta+1)^2}+\frac{1}{\beta+1}=\frac{1+\beta-\theta_0}{(\beta+1)^2}$$

implying that for a fixed θ_0 the expression has its minimum at $\beta = \theta_0 - 1$ and the value in that case is

$$\frac{\theta_0}{\theta_0-1+1}-\ln\theta_0-1+\ln(\theta_0-1+1)=0$$

This implies that eq A1 holds for all positive β and all $0 \leq \theta_0 \leq 1$, and thus, the argument in the solution is always between $-e^{-1}$ and 0.

(B) Simplifications of the Model

Low Concentration of the Bulk Analyte ($C_A k_a \ll k_d$). If the concentration of the bulk analyte is sufficiently low, then $C_A k_a \ll k_d$, implying that $C_A k_a + k_d \approx k_d$ and the constants K_1 and K_3 become

$$K_1 \approx \frac{C_A k_a}{k_d}$$

$$K_3 \approx \frac{k_c M_w G}{k_c M_w G + k_a R_{max}} = \frac{(k_c M_w G)/(k_a R_{max})}{(k_c M_w G)/(k_a R_{max}) + 1}$$

Furthermore, in this case

$$K_2 \approx \frac{C_A k_a{}^2 R_{max}}{k_d(k_c M_w G + k_a R_{max})} \leq \frac{C_A k_a{}^2 R_{max}}{k_d k_a R_{max}} = \frac{C_A k_a}{k_d} \ll 1$$

and thus $K_2 e^{K_2 - K_3 k_d t} \ll 1$. From this, it follows that

$$W(K_2 e^{K_2 - K_3 k_d t}) \approx K_2 e^{K_2 - K_3 k_d t} = K_2 e^{K_2} e^{-K_3 k_d t} \approx K_2 e^{-K_3 k_d t}$$

since $W(x) \approx x$ and $e^x \approx 1$ for small values of x. Using these approximations, the parameter K_2 disappears from the original model and it becomes

$$R = R_{max} K_1 (1 - e^{-K_3 k_d t})$$

which in terms of the original parameters becomes

$$R = \frac{C_A(k_a R_{max})}{k_d}\left\{1 - \exp\left[-\frac{(k_c M_w G)/(k_a R_{max})}{(k_c M_w G)/(k_a R_{max}) + 1} k_d t\right]\right\}$$

Note that after this simplification the model has only three parameters, namely, C_A, k_d, and the product $k_a R_{max}$. This implies that if the original model is used to fit measured data, and the condition $C_A k_a \ll k_d$ happens to hold, the calculated values of the parameters R_{max} and k_a should not be expected to be meaningful while their product will be.

Small Flow Rates ($k_c \ll k_a[B_0]$). If the mass transport coefficient k_c is negligible compared to $k_a[B_0]$, a condition that holds if the flow rate is sufficiently small, then $k_c M_w G \ll k_a[B_0] M_w G = k_a R_{max}$ and thus $(k_c M_w G)/(k_a R_{max}) \ll 1$. If in addition the bulk concentration is sufficiently low that the previous approximation applies, the model may be simplified even further by using $(k_c M_w G)/(k_a R_{max}) + 1 \approx 1$ and the original model may be approximated by the simplification

$$R = C_A \frac{k_a R_{max}}{k_d}\left[1 - \exp\left(-\frac{k_d}{k_a R_{max}} M_w G k_c t\right)\right]$$

Now the simplified model has only two parameters, namely, C_A and $(k_a R_{max})/k_d$. This implies as before that if one uses the original model to fit the measured data and the conditions for making the simplification shown here are present, one should not expect to obtain meaningful values for the parameters R_{max}, k_a, and k_d. Instead, only the resulting value of $(k_a R_{max})/k_d$ will be meaningful. Most importantly, however, curve fitting to the original model will always yield meaningful values of C_A, i.e., the active concentration of the analyte in the bulk solution.

REFERENCES

1. Christensen, L. L. H. (1997) Theoretical analysis of protein concentration determination using biosensor technology under conditions of partial mass transport limitation, *Anal. Biochem. 249*, 153−164.
2. Jönsson, U., Fägerstam, L., Ivarsson, B., Johnsson, B., Karlsson, R., Lundh, K., Löfås, S., Persson, B., Roos, H., and Rönnberg, I. (1991) Real-time biospecific interaction analysis using surface plasmon resonance and sensor chip technology, *Biotechniques 11*, 620−627.
3. Stenberg, E., Persson, B., Roos, H., and Urbaniczky, C. (1991) Quantitative determination of surface concentration of protein with surface plasmon resonance using radiolabelled proteins, *J. Colloid Interface Sci. 143*, 513−526.
4. Karlsson, R., Fägerstam, L. G., Nilshans, H., and Persson, B. (1993) Analysis of active antibody concentration. Separation of affinity and concentration parameters, *J. Immunol. Methods 166*, 75−84.
5. Karlsson, R., Michaelsson, A., and Mattson, L. (1991) Kinetic analysis of monoclonal antibody−antigen interactions with a new biosensor based analytical system, *J. Immunol. Methods 145*, 229−240.
6. O'Shannessy, D. J., Brighamburke, M., Soneson, K. K., Hensley, P., and Brooks, I. (1993) Determination of rate and equilibrium binding constants for macromolecular interactions using surface plasmon resonance: use of nonlinear least squares analysis methods, *Anal. Biochem. 212*, 457−468.
7. Fisher, R. J., Fivash, M., Casas-Finet, J., Erickson, J. W., Kondoh, A., Bladen, S. V., Fisher, C., Watson, D. K., and Papas, T. (1994) Real-time DNA binding measurements of the ETS1 recombinant oncoproteins reveal significant kinetic differences between the p42 and p51 isoforms, *Protein Sci. 3*, 257−266.
8. Morton, T. A., Myszka, D. G., and Chaiken, I. M. (1995) Interpreting complex binding kinetics from optical biosensors: a comparison of analysis by linearization, the integrated rate equation, and numerical integration, *Anal. Biochem. 227*, 176−185.
9. Myszka, D. G., Morton, M., Doyle, L., and Chaiken, I. M. (1997) Kinetic analysis of a protein antigen−antibody interaction limited by mass transport on an optical biosensor, *Biophys. Chem. 64*, 127−137.
10. Richalet-Sécordel, P. M., Rauffer-Bruyère, N., Christensen, L. L. H., Ofenloch-Haehnle, B., Seidel, C., and van Regenmortel, M. H. V. (1997) Concentration measurement of unpurified proteins using biosensor technology under conditions of partial mass transport limitation, *Anal. Biochem. 249*, 165−173.
11. Zeder-Lutz, G., Benito, A., and van Regenmortel, M. H. V. (1999) Active concentration measurements of recombinant biomolecules using biosensor technology, *J. Mol. Recognit. 12*, 300−309.
12. Cantor, C. R., and Schimmel, P. R. (1980) Techniques for the study of biological structure and function, in *Biophysical Chemistry* (McCombs, L. W., Ed.) 2nd ed., pp 539−590, W. H. Freeman and Co., San Francisco.
13. Segel, L. A., and Slemrod, M. (1989) The quasi-steady-state assumption: a case study in perturbation, *SIAM Rev. 31*, 446−477.

14. Rice, R. G., and Duong, D. D. (1995) *Applied Mathematics and Modeling for Chemical Engineers*, pp 37−97, John Wiley & Sons, New York.
15. Euler, L. (1783) De serie Lambertina plurimisque eivs insignibvs proprietatibvs, *Acta Acad. Sci. Petropol. 2*, 29−51.
16. Briggs, K. (2000) W-ology, or some exactly solvable growth models, http://www.btexact.com/people/briggsk2/W-ology.html.
17. Beauchemin, N., Draber, P., Dveksler, G., Gold, P., Gray-Owen, S., Grunert, F., Hammarström, S., Holmes, K. V., Karlsson, A., Kuroki, M., Lin, S.-H., Lucka, L., Najjar, S. M., Neumaier, M., Öbrink, B., Shively, J. E., Skubitz, K. M., Stanners, C. P., Thomas, P., Thompson, J. A., Virji, M., von Kleist, S., Wagener, C., Watts, S., and Zimmermann, W. (1999) Redefined nomenclature for members of the carcinoembryonic antigen family, *Exp. Cell Res. 252*, 243−249.
18. Beauchemin, N., Kunath, T., Robitaille, J., Chow, B., Turbide, C., Daniels, E., and Veillette, A. (1997) Association of biliary glycoprotein with protein tyrosine phosphatase SHP-1 in malignant colon epithelial cells, *Oncogene 14*, 783−790.
19. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4, *Nature 227*, 680−685.
20. Wurtz, T., Kiseleva, E., Nacheva, G., Alzhanova-Ericson, A., Rosén, A., and Daneholt, B. (1996) Identification of two RNA-binding proteins in Balbiani ring premessenger ribonucleoprotein granules and presence of these proteins in specific subsets of heterogeneous nuclear ribonucleoprotein particles, *Mol. Cell. Biol. 16*, 1425−1435.
21. Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1999) *Numerical Recipies in C*, 2nd ed., pp 683−688, Cambridge University Press, Cambridge, U.K.
22. Myszka, D. G., He, X., Dembo, M., Morton, T. A., and Goldstein, B. (1998) Extending the range of rate constants available from BIACORE: interpreting mass transport-influenced binding data, *Biophys. J. 75*, 583−594.

BI020099H