Determination of Receptor–Ligand Kinetic and Equilibrium Binding Constants using Surface Plasmon Resonance: Application to the *lck* SH2 Domain and Phosphotyrosyl Peptides

M. M. Morelock,*,† R. H. Ingraham,† R. Betageri,‡ and S. Jakes§

Sections of Biophysics and Medicinal Chemistry, Department of Inflammatory Diseases, and Section of Cellular Inflammation, Department of Immunological Diseases, Boehringer Ingelheim Pharmaceuticals, Inc., Research and Development Center, 900 Ridgebury Road/P.O. Box 368, Ridgefield, Connecticut 06877-0368

Received November 15, 1994®

Experimental and computational methods were developed for surface plasmon resonance (SPR) measurements involving interactions between a solution-binding component and a surface-immobilized ligand. These protocols were used to distinguish differences in affinity between the SH2 domain of *lck* and phosphotyrosyl peptides. The surface-immobilized ligand was the phosphotyrosyl peptide EPQpYEEIPIA, which contains a consensus sequence (pYEEI) for binding *lck* SH2. In the kinetic experiment, SPR phenomena were measured during association and dissociation reactions for a series of glutathione-*S*-transferase (GST)-SH2 concentrations, generating a set of SPR curves. A global computational analysis using an $A + B \leftrightarrows AB$ model resulted in single set of parameter estimates and statistics. In an abbreviated format, an equilibrium experiment was designed so that equilibrium constants (K_{eq}) could be determined rapidly and accurately. A competitive equilibrium assay was developed for GST-SH2 in which K_{eq} values for a series of phosphotyrosyl peptides (derived from the pYEEI sequence) varied over 3 orders of magnitude. Interestingly, these results highlighted the significance of the +1 glutamate in providing high-affinity binding to the SH2 domain. For most drug discovery programs, these K_{eq} determinations are a sufficient measure of potency for the primary screen, with k_{off} and k_{on} determined in a secondary assay. Thus, the application of these techniques to SPR binding phenomena should prove valuable in the discovery and design of receptor–ligand antagonists.

SH2 domains have been identified as 100-amino acid motifs that are included in the noncatalytic regions of signaling proteins such as *src*, *fps*, PLC-γ, and *v-crk*.[1] Interest in SH2 domains has been heightened in recent years with the discoveries that SH2 domains bind specifically to phosphotyrosine-containing peptides and that many signaling pathways, from antigen stimulation to interferon transcription, operate through SH2 domain binding.[2] Many SH2 domains recognize specific amino acid sequences surrounding phosphotyrosine (pY), as is the case for $p56^{lck}$ which binds to the sequence pYEEI, derived from a degenerate peptide library.[3] Because of the clear connection between SH2 domains and mitogenic signaling pathways, there are obvious opportunities for pharmacological intervention in disease states characterized by excessive cellular proliferation. By perturbing SH2 domain interactions, many signal transduction pathways can be specifically interrupted without resorting to inhibitors of enzymatic activity. In order to pursue SH2 domains as a drug target, however, precise and reproducible binding assays need to be developed.

Experimental studies of receptor–ligand interactions are often contingent upon the presence of a probe whose physical characteristics can be used to monitor the binding event. Conventional spectroscopic methods have been used to delineate structural changes in proteins which occur upon binding, but these spectroscopic changes are typically small and difficult to interpret. Consequently, most SAR (structure–activity relationship) studies carried out to date have involved the use of extrinsic probes, e.g., fluorophor or radioactive labeled ligand, which involve a significant change upon binding and/or are capable of being detected at very low concentrations. These systems, however, typically involve equilibrium measurements and are further compromised by the use of structurally modified native components.

Due to the advent of surface plasmon resonance (SPR) technology, interactions between solubilized and surface-immobilized biological counterparts, e.g., cytokines, growth factors, transcription factors, and their cell-surface receptors, can now be measured in real-time without the introduction of an extrinsic probe.[4] The SPR experiment measures small changes in refractive index at a surface which are the direct result of mass changes in the approximate medium. Consequently, interactions between a soluble-binding component (analyte) and an immobilized ligand on the surface can be measured in real-time. In a typical experiment, detectors such as the BIAcore (Pharmacia Biosensor) measure the SPR phenomena during association and dissociation reactions for an analyte–ligand pair (at several different concentrations of analyte). Appropriate kinetic models are then applied to the data to give the respective rate constants. In an abbreviated format, the experiment can also be designed to give equilibrium information. SPR methodology is particularly suitable for use in drug design and SAR studies, since the experimental design affords fast data acquisition times and yields the equilibrium constant (K_{eq}). The theoreti-

* To whom correspondence should be addressed.
† Section of Biophysics.
‡ Section of Medicinal Chemistry.
§ Section of Cellular Inflammation.
® Abstract published in *Advance ACS Abstracts*, March 15, 1995.

0022-2623/95/1838-1309$09.00/0

cal models and experimental protocols presented herein were used to distinguish the differences in affinity between the SH2 domain of *lck* and a series of phosphotyrosyl peptides derived from the pYEEI sequence.

Theory

Since the SPR experiment involves solution–surface reactions, kinetic models cannot be strictly developed using equilibrium relationships for solutions. For example, on rates for the analyte will be slower as a result of diffusion-limited mass transport to the surface; in addition, off rates will also be slower due to the relatively high concentrations of immobilized ligand at the surface.[5] Nevertheless, it is precisely this solution–surface interaction that is of interest in considering the *in vivo* environment for many receptor–ligand interactions. Consequently, we have developed models based on Langmuir equilibrium expressions and have noted that parameters estimated by such models are *apparent* binding constants. These models postulate that the immobilized binding component is uniformly distributed on the surface and that binding of one analyte does not affect the binding of another. Since higher molecular weight species induce larger changes in the SPR signal upon binding the surface matrix, it is advantageous to design the experiment such that the immobilized ligand is the lower molecular weight component. A schematic of the working surface used in this study is given in Figure 1.

The following equation is a generalized scheme for simple binding between an analyte (A) and a BIAcore sensor chip-immobilized ligand (C):

$$\mathrm{A} + \mathrm{C} \underset{k_d}{\overset{k_a}{\leftrightarrow}} \mathrm{AC} \quad \text{and} \quad K_C = \frac{k_d}{k_a} = \frac{[\mathrm{A}][\mathrm{C}]}{[\mathrm{AC}]} \tag{1}$$

where k_a and k_d are the association and dissociation rate constants, respectively, and K_C is the equilibrium dissociation constant. The BIAcore instrument measures the surface plasmon resonance of this interaction in terms of resonance units (RU) at various analyte concentrations generating BIAcore sensorgrams (see Figure 2, top panel). The relative response at a given time (R_t) can be expressed as

$$R_t = R_{\mathrm{base}} + R_{\mathrm{bulk}} + R_{\mathrm{AC}} \tag{2}$$

R_{base} is the base line average value prior to the association reaction and a function of the chip-immobilized ligand and running buffer; R_{bulk} arises from the difference in refractive index between the analyte solution and the running buffer; and R_{AC} is the nonlinear response resulting from association (increasing RU) or dissociation (decreasing RU) reactions of analyte with the chip-immobilized ligand. It should be noted that in the SPR experiment the data obtained are a macroscopic or bulk measurement; consequently, the designation of the reaction phases as 'association' or 'dissociation' is descriptive only of the dominant phenomenon, since both association and dissociation are clearly occurring simultaneously on the microscopic level in both phases.

Since only changes in R_{AC} are of interest, eq 2 can be rearranged and applied to the raw BIAcore data (see

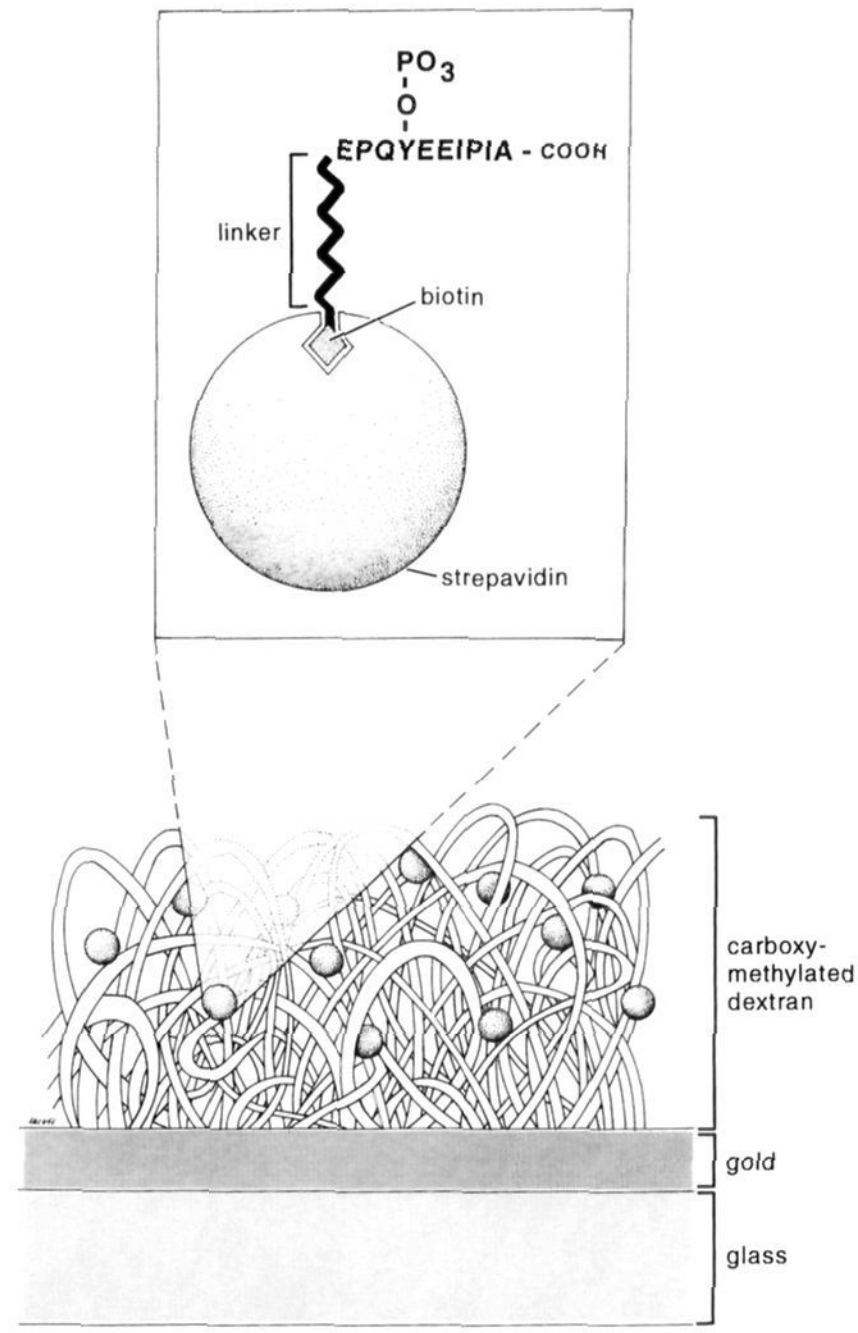

Figure 1. Schematic representation of the working surface used for SPR measurements of SH2 domain binding. The immobilized ligand EPQpYEEIPIA is a tyrosine-phosphorylated peptide derived from the degenerate library given in ref 3 for binding *lck* SH2.

Figure 2, bottom panel). The association phase will be governed by mass transport and the surface-binding event. Since both of these phenomena are a function of the analyte concentration, the observed rates will be $k_{\mathrm{MT}}[\mathrm{A}]$ (for mass transport) and $k_a[\mathrm{A}][\mathrm{C}]$ (for the kinetic event of interest), where [A] and [C] are the analyte and chip concentrations, respectively. Since both reactions are a function of [A], the relative rates are given by k_{MT} and $k_a[\mathrm{C}]$. Due to the relatively high concentration of chip ligand at early reaction times, the initial phase of the association reaction is usually dominated by mass transport since $k_a[\mathrm{C}] > k_{\mathrm{MT}}$, i.e., mass transport is rate limiting. As the reaction proceeds, [C] decreases such that the rates approach each other and finally to the point where $k_{\mathrm{MT}} > k_a[\mathrm{C}]$, i.e., the surface binding event is rate limiting. Thus, mass transport can be minimized by decreasing the surface concentration of the immobilized ligand and increasing the flow rate of analyte across the chip surface. The main problem encountered in the dissociation phase is the phenomenon of rebinding, i.e., once an analyte molecule dissociates from the chip surface it can rebind if there is an accessible, available site. Rebinding of analyte (which always occurs to some degree) is less problematic during the early stages of the reaction, since the number of open sites for rebinding is decreased at the end of the

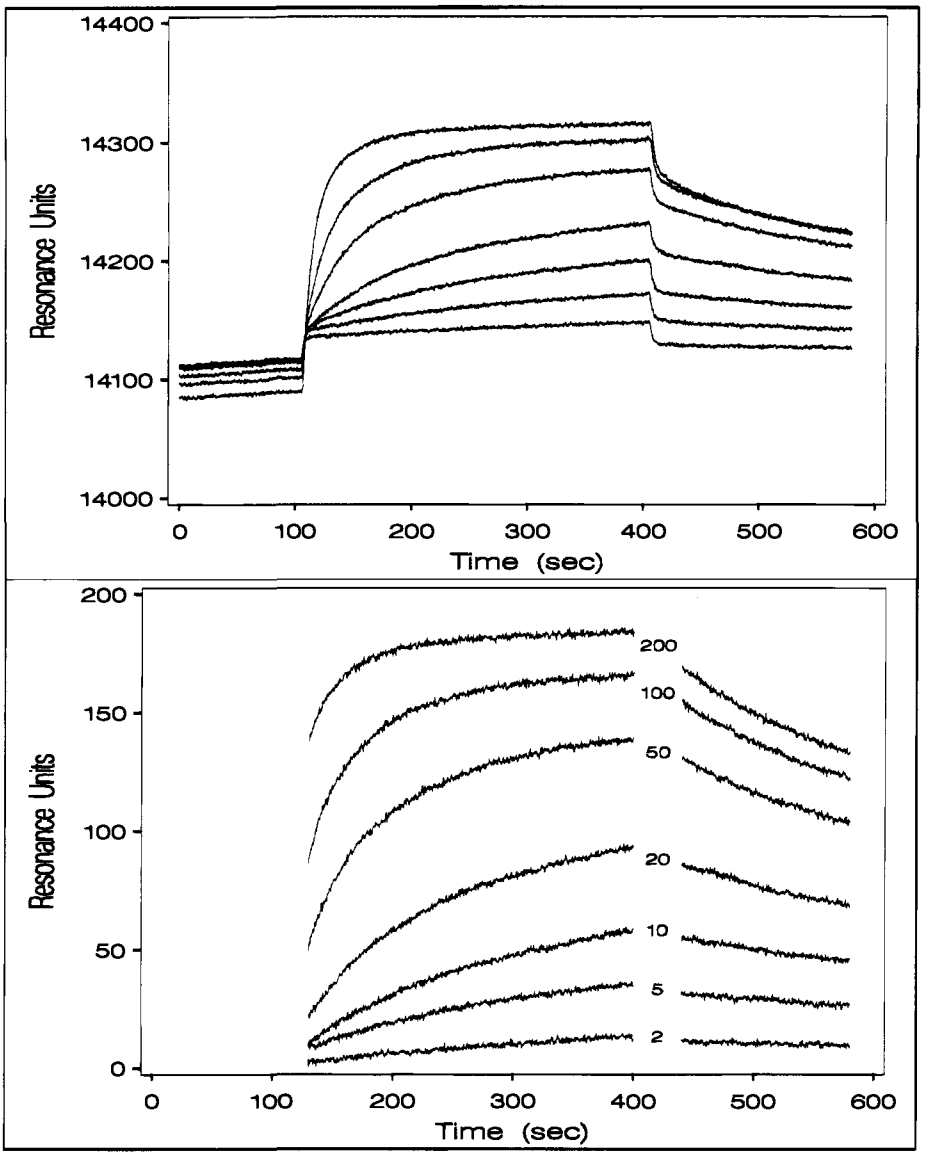

Figure 2. SPR data taken by the BIAcore instrument for the interaction of GST-SH2 and the chip-immobilized ligand EPQpYEEIPIA. Top panel, the ASCII raw data files are plotted for reactions involving a series of GST-SH2 concentrations. Bottom panel, these data have been converted into a SAS data file and corrected by eq 2.

association phase. As dissociation progresses, however, more and more sites become available for rebinding. Thus, the analysis of data involving short reaction times will be less affected by the rebinding phenomena. It should also be noted that rebinding can be minimized by decreasing the surface concentration of immobilized ligand, coinjecting solubilized ligand, and increasing the flow rate of the analyte across the chip surface. Even after optimization of the experimental technique, it may still be necessary to edit portions of the data. In the present example, the first 30 s of the association data given in the top panel of Figure 2 was found to drop off steeply below a fitted surface, *vide infra* (presumably due to the mass transport), and were edited out in the bottom panel.

R_{AC} and the concentration of surface-bound analyte are related to one another by the proportionality constant (k_p), which has been shown to be valid for proteins over a wide response range.[6]

$$[AC] = k_p R_{AC} \tag{3}$$

In addition, the maximum value for R_{AC} (defined as R_{max}) occurs during the association phase when the chip is saturated with analyte, i.e., $[AC] = [C]_T$; the minimum value, i.e., R_{min}, will be observed when $[AC] = 0$. Thus, the relationship can be defined as follows:

$$[C]_T = k_p(R_{max} - R_{min}) \tag{4}$$

Theoretically, R_{min} should equal zero but is included since each sensorgram is independently corrected by eq 2.

Determination of Rate Constants

The following differential rate equation can be written for the reaction described in eq 1:

$$\frac{d[AC]}{dt} = k_a[A][C] - k_d[AC] \tag{5}$$

For the association reaction, the mobile phase is switched from running buffer to sample. As analyte binds the surface-immobilized ligand, it is continuously replenished by fresh sample solution, so that [A], the concen-

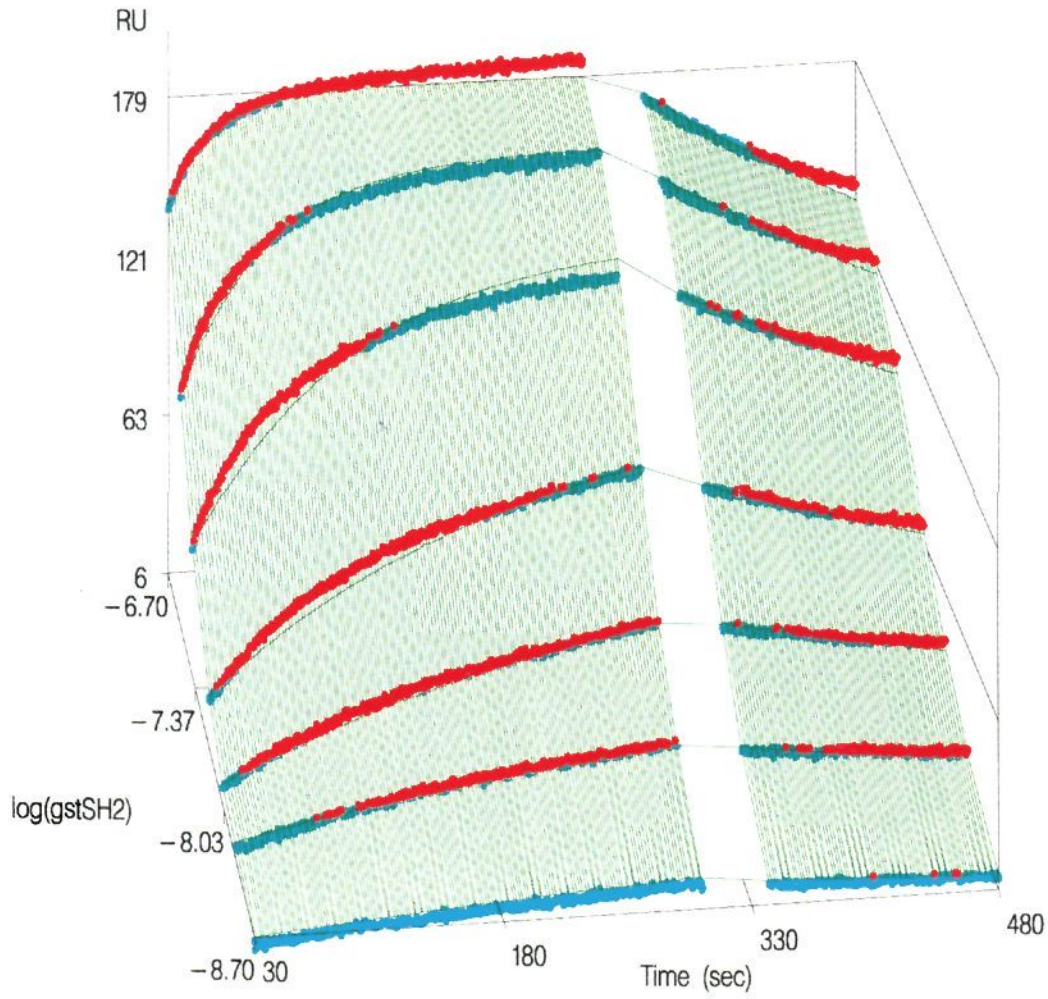

Figure 3. Determination of the apparent kinetic rate constants for the interaction between GST−SH2 and the chip-immobilized ligand EPQpYEEIPIA. The data are taken from the bottom panel of Figure 2. Nonlinear regression global analysis (applying eq 7 to the kinetic data, see bottom panel of Figure 2) converged the following solution (green grid): $k_a = 1.66\ (0.004) \times 10^5\ M^{-1}\ s^{-1}$, $k_d = 2.09\ (0.01) \times 10^{-3}\ s^{-1}$, $R_{max} = 1.90\ (0.002) \times 10^2$ RU, and $R_{min} = 1.84\ (0.06)$ RU, where the values in parentheses are the standard errors. The data in red and cyan lie above and below the fitted surface, respectively.

tration of analyte available for binding to the surface, is always equal to the total concentration of analyte in the sample, i.e., $[A]_T$. Substituting the mass balance relationship for [C] and integrating for [AC] at time t leads to

$$[AC]_t = [AC]_\infty + ([AC]_0 - [AC]_\infty)e^{-\alpha t} \qquad (6)$$

where $\alpha = (k_a[A]_T) + k_d$, $\beta = k_a[A]_T[C]_T$, and $[AC]_\infty = \beta/\alpha$.

For the dissociation reaction, the mobile phase is switched back to running buffer and dissociated analyte is removed on a continuous basis so that $[A]_T = 0$. Thus, eq 6 is a real-time expression for the amount of chip-bound analyte in the association and dissociation reactions. Substituting eq 3 for $[AC]_t$ and eq 4 for $[C]_T$, eq 6 can now be rewritten in terms of R as follows:

$$R_{AC} = \left\{\frac{k_a[A]_T(R_{max} - R_{min})}{\alpha}\right\} + \left\{R'_0 - \frac{k_a[A]_T(R_{max} - R_{min})}{\alpha}\right\}e^{-\alpha t} \qquad (7)$$

R_0' refers to the corrected relative response at the point on the respective reaction curve where t is defined to be zero; the proportionality constant (k_p), which is a part of each term in eq 7, does not appear here because it has been canceled out. The direct binding experiment is typically carried out by obtaining a series of sensorgrams (containing association and dissociation curves) at several concentrations of analyte, resulting in a response surface (see Figures 2 and 3). Equation 7 is then applied to all of the curves simultaneously, estimating the parameters k_a, k_d, R_{max}, and R_{min} with $[A]_T$ and t as the independent and R_{AC} as the dependent variables. It should be pointed out that although the model given by eqs 5−7 represents the simplest (and most common) model, other models (involving an increased number of components, binding valency, etc.) can be substituted.

Traditional analysis[6] would involve carrying out linear regressions on log-transformed dissociation data for each analyte concentration. The resulting estimates for k_d would then be averaged and carried forward into the analysis of the association data for k_a, which also involves linear regressions and estimate averaging. Obviously, linear transformation of the data and averaging of parameter estimates make the statistical significance of each estimate difficult to determine.[7] O'Shannessy[8] improved this procedure by introducing a more accurate nonlinear treatment of the kinetic data, but the method was still a multistep procedure, solving for k_d and then for k_a. More recently Fisher[9] developed an elegant global analysis in which nonlinear models were applied to the all the data simultaneously, resulting in a single set of parameter estimates and statistics for the rate constants of interest. Although Fisher's method is the most comprehensive to date, the complexity of the models requires considerable computational power. In the method presented here, a global analysis is also applied but the model is computationally less intensive and can be run on a PC.

Figure 3 shows the global fit for the direct binding of glutathione-*S*-transferase (GST)-SH2 with the chip-immobilized peptide EPQpYEEIPlA where pY is phosphotyrosine. For the association phase, the rate of

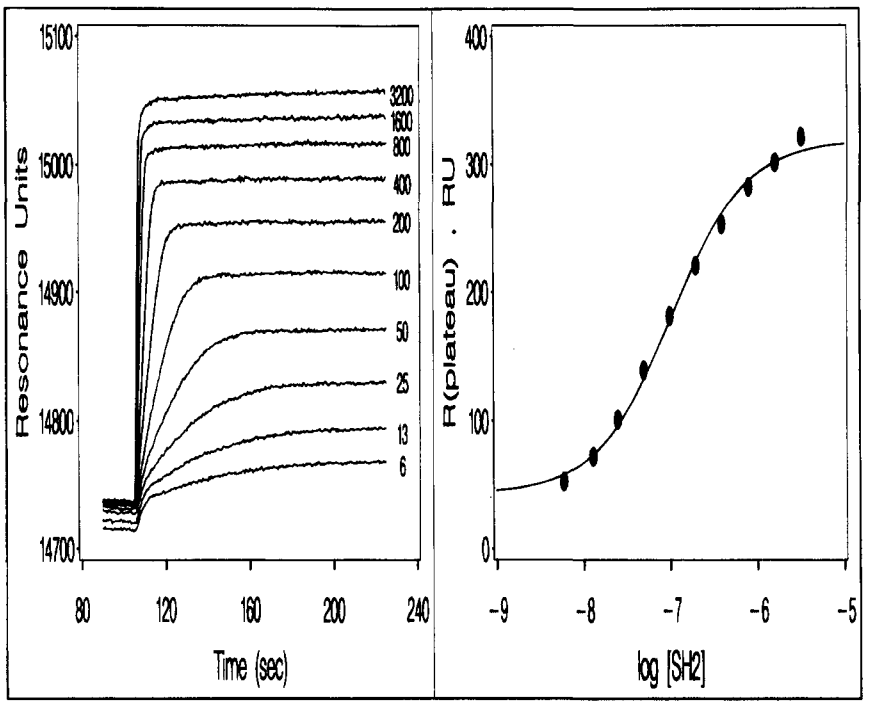

Figure 4. Determination of K_C for the interaction between the free (nonfusion) SH2 domain and the chip-immobilized peptide EPQpYEEIPIA. Left panel, raw data are shown for the association reaction between SH2 (concentration numbers given in nM) and the immobilized ligand EPQpYEEIPIA. Right panel, the $R_{plateau}^{[A]_T}$ values [obtained by taking the highest RU value for each concentration of analyte (free SH2) in the left panel] are plotted versus $[A]_T$. Nonlinear regression analysis (applying eq 11) converged the following solution (solid line): $K_C = 1.05\ (0.12) \times 10^{-7}$ M, $R_{max} = 3.19\ (0.05) \times 10^2$ RU, and $R_{min} = 4.29\ (0.66) \times 10^1$ RU, where the values in the parentheses are the standard errors.

reaction and the total RU's reached increase with increasing concentrations of GST-SH2. Looking down the [GST-SH2] axis at about 300 s, the surface has a sigmoid shape characteristic of equilibrium binding (with the value of K_C being given by the concentration of GST-SH2 at the midpoint). Since the system is not affected by mass transport or rebinding phenomena at equilibrium, the K_C of the system is computationally locked in by this portion of the data and serves to constrain the estimation for the rate constants. Thus, as the data are fitted to eq 7, areas in the data affected by mass transport or rebinding (as evidenced by departure above or below the fitted surface) can be easily visualized using the global analysis. In this example, the initial phase of the association data given in the top panel of Figure 2 was found to drop off steeply below a fitted surface (presumably due to mass transport). After editing the first 30 s of the association data (see Figure 2, bottom panel), the data conformed more uniformly to the applied rate equation (see Figure 3). The model fits the data quite well with an $R^2 = 0.9977$, P values on all parameters < 0.0001, and standard errors on the rate constants less than 1%. Although the residuals (i.e., the difference between the data and the fitted surface) are small, they are not randomly distributed on a per reaction basis but rather are evenly distributed on a global basis. This global distribution of residuals results from fitting a series of reactions for a set of common rate constants where experimental error is evident. For example, if experimental error occurs in reactant preparation, some reactions may appear slightly above or below the fitted surface, since reactant concentrations are hard-coded assuming perfect dilution. Finally, it should be noted that k_a and k_d are *apparent* rate constants due to the solution–surface reaction format.

Determination of Equilibrium Constants

Direct Binding. As analyte at a fixed concentration is injected over the chip matrix, R_{AC} increases nonlinearly until a plateau value ($R_{plateau}^{[A]_T}$) is reached (see Figure 4, left panel), indicating that analyte is in equilibrium with immobilized ligand. The response amplitude ($R_{plateau}^{[A]_T} - R_{min}$) is always some fractional value of the quantity $R_{max} - R_{min}$. From equilibrium considerations, the fraction of free and bound chip-immobilized ligand can be expressed by first dividing the top and bottom of eq 1 by $[C]_T$ to give

$$K_c = \frac{[A]\frac{[C]}{[C]_T}}{\frac{[AC]}{[C]_T}} = \frac{[A]\theta_{free}}{\theta_{bound}} \tag{8}$$

where θ_{free} and θ_{bound} are the fractions of free and bound chip-immobilized ligand, respectively. Since $\theta_{free} + \theta_{bound} = 1$ and $[A] = [A]_T$ (analyte is continuously replenished in the BIAcore experiment), substitution and rearrangement leads to

$$\theta_{bound} = \frac{[A]_T}{K_c + [A]_T} \tag{9}$$

Note that when $[A]_T = K_C$, $\theta_{bound} = 0.5$. Since the fraction of bound immobilized ligand in the BIAcore

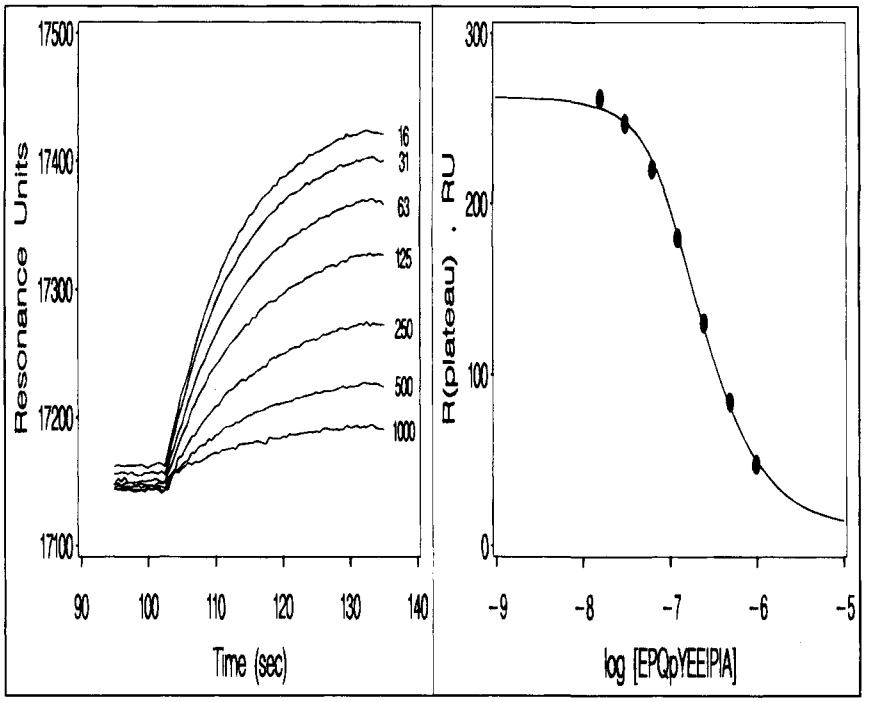

Figure 5. Determination of K_S for the interaction between the GST-SH2 domain (50 nM) and the solubilized peptide EPQpYEEIPIA. Left panel, raw data are shown for the association reaction between GST-SH2 and the immobilized ligand EPQpYEEIPIA in the presence of the solubilized ligand (concentration numbers given in nM). Right panel, the $R^{[S]_T}_{plateau}$ values (obtained by taking the highest RU value for each concentration of solubilized ligand in the left panel) are plotted versus the concentration of solubilized ligand, $[S]_T$. Nonlinear regression analysis (applying eq 15 with K_C = 7.76 nM) converged the following solution (solid line): K_S = 2.38 (0.23) × 10^{-8} M, R_{max} = 3.05 (0.03) × 10^2 RU, and R_{min} = 9.72(6.96) RU, where the values in parentheses are the standard errors.

experiment can be expressed as $(R^{[A]_T}_{plateau} - R_{min})/(R_{max} - R_{min})$, these data can be fitted to the following general equation:

$$R^{[A]_T}_{plateau} = \theta_{bound}(R_{max} - R_{min}) + R_{min} \quad (10)$$

Substitution of eq 9 for θ_{bound} gives

$$R^{[A]_T}_{plateau} = \frac{[A]_T}{K_C + [A]_T}(R_{max} - R_{min}) + R_{min} \quad (11)$$

The parameters K_C, R_{max}, and R_{min} are estimated by nonlinear regression using eq 11 with $[A]_T$ as the independent and $R^{[A]_T}_{plateau}$ as the dependent variables. The $R^{[A]_T}_{plateau}$ values are plotted versus $[A]_T$ in the right panel of Figure 4. The solid line shows the nonlinear regression fit to the data for the direct binding of free SH2 with the chip-immobilized peptide EPQpYEEIPIA. It should be noted that K_C is an *apparent* dissociation constant due to the solution–surface reaction format.

Competitive Binding. In the competitive BIAcore experiment, a solubilized ligand (S) is preincubated with the analyte (A) and then injected over the chip matrix. A generalized scheme for competitive binding can be written as follows:

$$AC \overset{K_C}{\leftrightarrows} C + A + S \overset{K_S}{\leftrightarrows} AS \quad (12)$$

where K_S is the equilibrium binding constant between analyte and solubilized ligand. The equilibrium involving K_S will not be shifted during the experiment since the sample solution (A + S) is the mobile phase during the association reaction. As the concentration of S is increased (with $[A]_T$ held constant), the concentration of *free* A will decrease, resulting in a lower absorption of analyte to the chip matrix. Equation 9 can be modified for the competitive binding experiment as follows:

$$\theta_{bound} = \frac{[A]_T - [AS]}{K_C + [A]_T - [AS]} \quad (13)$$

Substituting mass balance relationships into the equilibrium expression for K_S gives

$$K_S = \frac{[A][S]}{[AS]} = \frac{([A]_T - [AS])([S]_T - [AS])}{[AS]} \quad (14)$$

The quadratic solution Q for [AS] is a function of $[A]_T$, $[S]_T$, and K_S. Substitution of Q and eq 13 into eq 10 gives

$$R^{[S]_T}_{plateau} = \frac{[A]_T - Q}{K_C + [A]_T - Q}(R_{max} - R_{min}) + R_{min} \quad (15)$$

In a manner similar to the equilibrium method described in the preceding section, $R^{[S]_T}_{plateau}$ values (holding $[A]_T$ constant) are obtained for competitive reactions of GST-SH2 with the solubilized and chip-immobilized peptide EPQpYEEIPIA (see Figure 5). Note that the values of $R^{[S]_T}_{plateau}$ for the competitive experiment decrease with increasing concentrations of solubilized peptide (left panel). These $R^{[S]_T}_{plateau}$ values are plotted versus $[S]_T$ in the right panel of Figure 5.

The parameters K_S, R_{max}, and R_{min} are estimated by nonlinear regression using eq 15 with K_C (determined from direct binding experiments, *vide supra*) and $[A]_T$ constant and $[S]_T$ as the independent and $R^{[S]_T}_{plateau}$ as the dependent variables. The solid line in the right panel of Figure 5 shows the nonlinear regression fit to the data. In contrast to the apparent equilibrium constant K_C, K_S is a true affinity constant (albeit apparent due to nonphysiological conditions) since its measurement concerns a solution phase reaction which is independent of the considerations affecting solution–surface interactions.

Since many investigators have already analyzed competitive equilibrium data by fitting for the midpoint of the curve, i.e., where $\theta = 0.5$, the following relationships were developed so that the K_S can be calculated from these data. Substituting eq 14 for [AS] into eq 13 gives

$$\theta_{bound} = \frac{[A]_T - \frac{[A][S]}{K_S}}{K_C + [A]_T - \frac{[A][S]}{K_S}} \quad (16)$$

Since $[A] = K_C$ at $\theta = 0.5$ (see eq 9), the following mass balance relationships can be written:

$$[AS] = [A]_T - K_C \quad (17)$$

and

$$[S] = [S]_T - ([A]_T - K_C) \quad (18)$$

Substitution of eq 18 for [S] into eq 16 and K_C for [A] and solving for K_S at $\theta = 0.5$ gives

$$K_S = \frac{[S]_T - [A]_T + K_C}{\frac{[A]_T}{K_C} - 1} \quad (19)$$

where K_C is the previously determined midpoint of the direct binding experiment and $[S]_T$ is the previously established midpoint in the competitive binding experiment (with $[A]_T$ constant).

Results and Discussion

The SPR experiment measures small changes in refractive index at a surface which are the direct result of mass changes in the approximate medium. Consequently, receptor–ligand interactions can be measured in real-time without the introduction of an extrinsic probe. In a typical experiment, the BIAcore instrument measures the SPR phenomena during association and dissociation reactions for an analyte–ligand pair. The working SPR surface used in these studies involved the immobilization of the phosphotyrosyl peptide EPQpY-EEIPIA (which contains a consensus sequence for binding *lck* SH2) via a biotin–strepavidin linkage (see Figure 1). Since the surface can be titrated using this biotin–strepavidin format, the surface density of immobilized ligand can be adjusted for the maximum signal with the least amount of mass transport and rebinding; in addition, the affinity of the biotin–strepavidin complex is such that the ligand remains bound during the acid regeneration cycle, providing a well-characterized and reproducible binding surface.

In the direct binding experiment, association and dissociation curves were recorded for a series of GST-SH2 concentrations, generating a response surface (see Figures 2 and 3). The data were fitted with a simple binding model ($A + B \rightleftharpoons AB$) using a global computational method which resulted in a single set of parameter estimates and statistics [$k_a = 1.66\ (0.004) \times 10^5\ M^{-1}\ s^{-1}$, $k_d = 2.09\ (0.01) \times 10^{-3}\ s^{-1}$ (where values in parenthesis are standard errors); calculated $K_c = 12.6$ nM]. It should be emphasized that these are *apparent* constants since the SPR measurement involves a solution–surface reaction. Nevertheless, the global analysis (based on Langmuir theory) fits the response surface quite well with standard errors on the kinetic estimates being less than 1%. Additional determinations resulted in a mean value for K_C of 7.76 (±3.77) nM ($n = 5$, value in parentheses is the standard deviation); this value was used in all further calculations involving GST-SH2 equilibrium reactions.

As mentioned, the density of the immobilized ligand on the chip surface must be adjusted to maximize the total signal while minimizing the effects of mass transport and rebinding. In an equilibrium experiment, however, mass transport and rebinding phenomena do not affect the steady state SPR signal of the analyte–ligand pair. Although it may be difficult to kinetically measure K_C for analytes that have very fast on and off rates, this information can be reliably determined from equilibrium measurements. To illustrate this point, an experiment was performed using a high density of immobilized ligand and free SH2 (which exhibits fast on and off rates), rather than the GST fusion protein (see the Experimental Section). The left panel of Figure 4 shows the reaction between the free SH2 domain and a high-density surface. Most of the reactions appear to have reached equilibrium within a few minutes, with mass transport clearly observable for the higher SH2 concentrations (compare with Figure 2, top panel). The equilibrium RU values are plotted versus [SH2] in the right panel. The model developed for direct binding fits the data quite well, giving a K_C of 105 (12) nM (value in parentheses is the standard error). This value is comparable with the calculated K_C of 67.5 nM determined from the kinetic analysis (data not shown). Since GST-SH2 is expressed as a dimer, it is possible that the higher affinity observed for the GST construct (7.76 nM, *vide supra*) is due to either an avidity effect of the bivalent dimer or the influence of the fusion partner on the intrinsic binding ability of the SH2 domain.

In the competitive binding experiment, a solubilized ligand is preincubated with the SH2 domain and then injected over the chip matrix. The equilibrium RU values and the apparent K_C for the chip-immobilized ligand (determined from direct binding experiments) can be used to calculate the concentration of SH2 available for binding to the surface and thus the concentration of SH2 bound by solubilized ligand (by mass balance). In contrast to the direct binding experiments, however, characterization of these reactions gives a true affinity constant, K_S (albeit apparent due to nonphysiological conditions), since they are a solution-binding phenom-

ena. Figure 5 shows an example involving GST-SH2 in which both the solubilized and immobilized ligands are EPQpYEEIPIA. Note that the equilibrium RU values decrease as a function of increasing solubilized peptide (left panel). The equilibrium model developed for this experiment fits the data quite well (right panel) giving a K_S of 23.8 (2.30) nM (value in parentheses is the standard error). This value is somewhat higher than the value of the calculated K_C [7.76 (±3.77) nM] from the direct binding kinetic experiments, *vide supra*. It is possible that the higher affinity observed for the surface-immobilized (versus the solubilized) peptide is due in part to an avidity effect of the bivalent GST-SH2 dimer.

Previous reports applying SPR technology to the analysis of SH2 competitive binding experiments have noted that the relative affinities for the phosphotyrosyl peptides studied appear to be weaker than the value determined by conventional direct binding measurements.[10] It should be pointed out that if the midpoint of the curve was taken to be K_S in the present example (≈300 nM, see Figure 5), the apparent affinity of the SH2–phosphotyrosyl peptide interaction (24 nM) would also have been underestimated (in this case by 1 order of magnitude). Nevertheless, if K_C is known for a given system, K_S can be calculated from midpoint data using eq 19.

Karlsson[11] recently reported a method to evaluate competitive BIAcore reactions in which the initial rates of association reactions were used to estimate the concentration of free analyte (not bound by solubilized ligand). When k_a is fast, k_d slow, and the surface density employed high, the initial rate of the reaction will typically be limited by mass transport and, thus, linearly proportional to the concentration of free analyte. When k_a is slow and/or k_d fast, however, it may not be possible to achieve mass transport-limited kinetics at analyte concentrations relevant to the competitive binding phenomena (even with high-density surfaces); in such cases, the initial rates will be nonlinear and difficult to estimate. For the method presented here, however, the only requirement is to achieve a steady state SPR signal (flow rate, surface density, reaction time, etc., can be optimized in the absence of solubilized ligand).

The competitive equilibrium method was used to distinguish the differences in affinity between the SH2 domain of *lck* and a series of phosphotyrosyl peptides based on the sequence EGQpYEEIP. Using a degenerate peptide library,[3] a consensus sequence for *lck* SH2 domain binding has previously been identified as containing the sequence pYEEI. The X-ray structure of the *lck* SH2 domain[12] revealed that the EPQpYEEIPIYL peptide, which lies in an extended conformation across the face of the SH2 domain, makes the most extensive contacts by inserting phosphotyrosine and isoleucine at +3 into specific pockets of the SH2 domain. From these studies it could be assumed that of the residues at +1, +2, and +3, the greatest binding energy would be provided by the isoleucine at +3. To test this, a glycine scan was done on positions +1, +2, and +3 relative to the pY. A set of competitive binding curves for the peptide series is shown in Figure 6 and reflects the sensitivity of the assay. The K_S values for five experiments are given in Table 1. While variability can be seen between runs (less than 2-fold in most cases), the relative order of affinity remains the same. While substitution of each of the three residues causes a major increase in K_S, the greatest increases occur with the substitution for the glutamate at +1 and the isoleucine at +3. Interestingly, these results highlight the significance of the +1 glutamate in providing high-affinity binding to the SH2 domain.

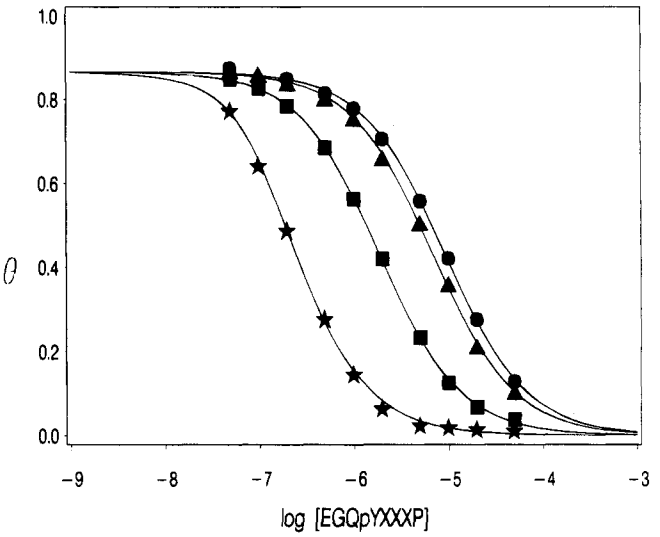

Figure 6. Application of the competitive equilibrium method to a series of peptides in which a glycine has been substituted to assess the contribution of each position to the binding affinity for GST-SH2. The data (for an arbitrary experiment) have been normalized to each other by converting RU data to θ (applying eq 15 with K_C = 7.76 nM). The peptides are EGQpYEEIP (stars; K_S = 26.9 nM), EGQpYGEIP (dots; K_S = 1220 nM), EGQpYEGIP (squares; K_S = 242 nM), and EGQpYEEGP (triangles; K_S = 908 nM). See Table 1, run 3 for statistics.

Although kinetic experiments give additional information concerning k_{off} and k_{on}, the chip optimization process is time consuming and the selection of data not affected by mass transport or rebinding phenomena can be subjective. In contrast, the equilibrium experiment does not require chip optimization, and equilibrium constants for a series of molecules can be determined rapidly and accurately. Summarizing the steps of the protocol: (1) optimize the working surface for the maximum signal with the least amount of mass transport and rebinding, (2) determine the apparent rate (k_a, k_d) and equilibrium (K_C) constants for the interaction of analyte with the chip-immobilized ligand (direct binding), (3) determine the equilibrium constant (K_S) for a solubilized analog of the immobilized ligand (competitive binding), and (4) determine K_S for the series of interest using the competitive binding format.

In conclusion, experimental and computational methods were developed for the analysis of SPR binding data involving receptor–ligand interactions. Determinations of the equilibrium constant for the interaction of the free SH2 domain and the chip-immobilized ligand (K_C) gave comparable results using the kinetic and equilibrium methods. A competitive assay was developed in which equilibrium constants for solubilized ligands (K_S) could be determined. K_S for the binding interaction between a solubilized analog of the chip-immobilized ligand and the GST-SH2 construct was found to be comparable with the K_C determined by direct binding techniques.

Table 1. Peptide Recognition of *lck* SH2 Domain: K_S Values for Several SPR Determinations[a]

peptide sequence	run 1	2	3	4	5
EGQpYEEIP	$1.44\ (0.15) \times 10^{-8}$	$2.28\ (0.11) \times 10^{-8}$	$2.69\ (0.15) \times 10^{-8}$	$2.78\ (0.26) \times 10^{-8}$	na
EQGpYGEIP	$1.32\ (0.04) \times 10^{-6}$	$9.82\ (2.04) \times 10^{-7}$	$1.22\ (0.06) \times 10^{-6}$	$1.25\ (0.05) \times 10^{-6}$	$2.85\ (0.22) \times 10^{-6}$
EGQpYEGIP	$1.71\ (0.12) \times 10^{-7}$	$1.67\ (0.15) \times 10^{-7}$	$2.42\ (0.07) \times 10^{-7}$	$2.61\ (0.10) \times 10^{-7}$	$5.36\ (0.41) \times 10^{-7}$
EGQpYEEGP	$8.33\ (0.79) \times 10^{-7}$	$9.16\ (1.18) \times 10^{-7}$	$9.08\ (0.40) \times 10^{-7}$	$9.38\ (0.26) \times 10^{-7}$	$2.16\ (0.09) \times 10^{-6}$

[a] K_S values (M) determined using eq 15 with $K_C = 7.76$ nM. The values in parentheses are the standard errors, na = not analyzable.

The competitive assay was then applied to a series of SH2-binding peptides containing the pYEEI consensus sequence; K_S was found to vary over 3 orders of magnitude (relative to the immobilized peptide), with the significance of the +1 glutamate for high-affinity binding revealed. For most drug discovery programs, these K_S determinations are a sufficient measure of potency for the primary screen, with k_{off} and k_{on} determined in a secondary assay. Thus, the application of these experimental and computational techniques should prove valuable in the discovery and design of receptor–ligand antagonists.

Experimental Section

SH2 Domain. The protein used in the binding assays was a glutathione-*S*-transferase (GST) fusion protein containing a 105-amino acid insert from the *lck* SH2 domain,[13] expressed in the pGEX-KT vector.[14] The fusion protein was cleaved with thrombin to generate the free SH2 domain, which was then repurified over glutathione–Sepharose.

Synthesis of Phosphotyrosine-Containing Peptides. All the peptides were synthesized by solid phase methodology using Fmoc chemistry. The side-chain-protecting groups were *tert*-butyl ester for Glu and trityl for Gln. The solid phases used were Fmoc-L-Ala-HMPA resin (0.62 mmol/g) and NH-(L)-Pro-(2-chloro)-trityl resin (0.7 mmol/g). The coupling protocol was carried out with amino acid:HBTU:HOBT:DIEA in a mole ratio of 1:1:1:2. Coupling times were typically 30 min. The same protocol was used to couple Fmoc-L-Tyr(PO_3H_2)OH (Advanced ChemTech, Louisville, KY) to introduce phosphotyrosine. Cleavage of the peptide from the resin was achieved by treating the dried resin with either trifluoroacetic acid: ethanedithiol in a v/v ratio of 95:5 or trifluoroacetic acid: anisole:ethanedithiol:ethyl methyl sulfide in a v/v ratio of 95: 3:1:1 for 2 h at room temperature. The resin was filtered and washed with the cleavage cocktail. After evaporation of trifluoroacetic acid under vacuum, the residue was treated with the cold diethyl ether. The precipitated solid was centrifuged; the solid pellet obtained was washed with ether and dissolved in aqueous acetonitrile and lyophilized. The crude peptides were purified by reverse phase HPLC on a Vydac C-18 preparative column (300 Å pore size, 10 μm particle size, 2.2 cm × 25 cm) with UV monitoring at 215 nm and a flow rate of 15 mL/min. The gradient conditions were 5–100% B over 25 min for EPQpYEEIPIA and 30–100% B over 25 min for biotin-ε-aminohexanoic acid-EPQpYEEIPIA (A, 0.1% TFA/H_2O; B, 60% acetonitrile/water containing 0.1% TFA). Peptides EGQpYEEIP, EGQpYGEIP, EGQpYEGIP, and EGQpYEEGP were purified using the gradient conditions 1–30% B over 30 min (A, 0.1% TFA/H_2O; B, 80% acetonitrile/water containing 0.087% TFA). The major peak in each case was collected and characterized by FABMS, giving the desired MH^+ for the product.

BIAcore System and Chip Surface. The BIAcore instrument used in these studies was manufactured by Pharmacia Biosensor AB (Uppsala, Sweden). For the chip surface, strepavidin at 0.3 mg/mL in 10 mM sodium acetate, pH 4.5, was injected over a NHS/EDC-prepared CM5 sensor chip and then blocked with 1.0 M ethanolamine, pH 8.5 (Pharmacia Biosensor coupling kit). Biotin-ε-aminohexanoic acid-EPQpY-EEIPIA at 50 (for kinetic experiments) or 500 (for equilibrium experiments) nM in running buffer (10 mM HEPES, 150 mM NaCl, 0.005% P-20, pH 7.4) was injected at a flow rate of 5 μL/min for 48 s.

Direct Binding Assays. In the kinetic experiment, the GST-SH2 domain was diluted from 200 to 2 nM in running buffer containing 1 mM DTT and 0.2 mg/mL BSA; 25 μL of each concentration was passed over the working BIAcore chip surface (that provided a R_{max} of approximately 200 RU) with looped injections at a flow rate of 5 μL/min, with 4 μL of 10 mM HCl regenerations. In the equilibrium experiment, the free SH2 domain was diluted from 3200 to 6 nM in running buffer containing 1 mM DTT and 0.2 mg/mL BSA. Sample injection and regeneration followed the kinetic procedure (with the exception of $R_{max} \sim 350$ RU).

Competitive Binding Assay. The GST-SH2 construct was diluted to 50 nM in running buffer containing 1 mM DTT and 0.2 mg/mL BSA. Peptide was titrated into GST-SH2 from 50 μM to 5 nM and allowed to come to equilibrium; 25 μL of each reaction mixture was passed over the surface with looped injections at a flow rate of 5 μL/min, with 4 μL of 10 mM HCl regenerations.

Computational Analyses. The data were analyzed using the SAS statistical software system (version 6.08, SAS Institute Incorporated, Cary, NC). For the kinetic experiments, ASCII data files containing BIAcore SPR measurements were converted into a single SAS data set and corrected as per eq 2. Data analyses were then performed by applying the appropriate models (see eq 7) to each subset of the data simultaneously, using the procedure MODEL with the Marquardt–Levenberg minimization method. A single set of parameter estimates and statistics was obtained, with SAS-GRAPH being used to display the data overlaid with the theoretical curve or surface. Data analyses for the equilibrium experiments were performed by applying either eq 11 (direct binding) or eq 15 (competitive binding) to the data using the SAS procedure NLIN with the Marquardt–Levenberg minimization method.

Acknowledgment. The authors thank Edward T. Graham for his helpful advice and valuable discussions concerning the SAS programming of the computational models and graphical analyses, Josephine King for the cloning and expression of the *lck* SH2 domain, Daniel Greenwood for assistance in peptide purification, Leigh Rondano for the illustration of the working surface, and Dr. James Stevenson for his critical reading of the manuscript.

References

(1) Koch, C. A.; Anderson, D.; Moran, M. F.; Ellis, C.; Pawson, T. SH2 and SH3 domains: Elements that control interactions of cytoplasmic signaling proteins. *Science* **1991**, *252*, 668–674.
(2) Shuai, K.; Horvath, C. M.; Tsai Huang, L. H.; Qureshi, S. A.; Cowburn, D.; Darnell, J. E., Jr. Interferon Activation of the Transcription Factor Stat91 Involves Dimerization through SH2-Phosphotyrosyl Peptide Interactions. *Cell* **1994**, *76*, 821–828. Wange, R. L.; Malek, S. N.; Desiderio, S.; Samelson, L. E. Tandem SH2 Domains of ZAP-70 Bind to T Cell Antigen Receptor ζ and CD3ε from Activated Jurkat T Cells. **1993**, *268*, 19797–19801.
(3) Songyang, Z.; Shoelson, S. E.; Chaudhuri, M.; Gish, G.; Pawson, T.; Haser, W. G.; King, F.; Roberts, T.; Ratnofsky, S.; Lechleider, R. J.; Neel, B. G.; Birge, R. B.; Fajardo, J. E.; Chou, M. M.; Hidesaburo, H.; Schaffhausen, B.; Cantley, L. C. SH2 Domains Recognize Specific Phosphopeptide Sequences. *Cell* **1993**, *72*, 767–778.

(4) O'Shannessy, D. J. Determination of Kinetic Rate and Equilibrium Binding Constants for Macromolecular Interactions: A Critique on the Surface Plasmon Resonance Literature. *Curr. Opin. Biotechnol.* **1994**, *5*, 65–71. Malmqvist, M. Surface Plasmon Resonance for Detection and Measurement of Antibody-Antigen Affinity and Kinetics. *Curr. Opin. Immunol.* **1993**, *5*, 282–286.

(5) Glaser, R. W. Antigen-Antibody Binding and Mass Transport by Convection and Diffusion to a Surface: A Two-Dimensional Computer Model of Binding and Dissociation Kinetics. *Anal. Biochem.* **1993**, *213*, 152–161. Sandana, A.; Sii, D. Binding Kinetics of Antigen by Immobilized Antibody: Influence of Reaction Order and External Diffusion Limitations. *Biosens. Bioelectron.* **1992**, *7*, 559–568.

(6) Karlsson, R.; Michaelsson, A.; Mattsson, L. Kinetic Analysis of Monoclonal Antibody-Antigen Interactions with a New Biosensor Based Analytical System. *J. Immunol. Methods* **1991**, *145*, 229–240.

(7) Johnson, M. L. Why, When, and How Biochemists Should Use Least Squares. *Anal. Biochem.* **1992**, *206*, 215–225.

(8) O'Shannessy, D. J.; Brigham-Burke, M.; Soneson, K. K.; Hensley, P.; Brooks, I. Determination of Rate and Equilibrium Binding Constants for Macromolecular Interactions Using Surface Plasmon Resonance: Use of Nonlinear Least Squares Analysis Methods. *Anal. Biochem.* **1993**, *212*, 457–468.

(9) Fisher, R. J.; Fivash, M.; Casas-Finet, J.; Erickson, J. W.; Kondoh, A.; Bladen, S. V.; Fisher, C.; Watson, D. K.; Papas, T. Real-time DNA Binding Measurements of the ETS1 Recombinant Oncoproteins Reveal Significant Kinetic Differences Between the p42 and p51 Isoforms. *Protein Sci.* **1994**, *3*, 257–266. Fisher, R. J.; Fivash, M.; Casas-Finet, J.; Bladen, S.; McNitt, K. L. Real-time BIAcore Measurements of *Escherichia coli* Single-Stranded DNA Binding (SSB) Protein to Polydeoxythymidylic Acid Reveal Single-State Kinetics with Steric Cooperativity. *METHODS: Companion Methods Enzymol.* **1994**, *6*, 1–13.

(10) Payne, G.; Shoelson, S. E.; Gish, G. D.; Pawson, T.; Walsh, C. T. Kinetics of $p56^{lck}$ and $p60^{src}$ Src Homology 2 Domain Binding to Tyrosine-phosphorylated Peptides Determined by a Competition Assay or Surface Plasmon Resonance. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 4902–4906.

(11) Karlsson, R. Real-Time Competitive Kinetic Analysis of Interactions between Low-Molecular-Weight Ligands in Solution and Surface-Immobilized Receptors. *Anal. Biochem.* **1994**, *221*, 142–151.

(12) Eck, M. J.; Shoelson, S. E.; Harrison, S. C. Recognition of a High Affinity Phosphotyrosyl Peptide by the Src Homology-2 Domain of $p56^{lck}$. *Nature* **1993**, *362*, 87–91.

(13) Voronova, A. F.; Sefton, B. M. Expression of a New Tyrosine Protein Kinase is Stimulated by Retrovirus Promoter Insertion. *Nature* **1986**, *319*, 682–685.

(14) Guan, K.; Dixon, J. E. Eukaryotic Proteins Expressed in *Escherichia coli*: An Improved Thrombin Cleavage and Purification Procedure of Fusion Proteins with Glutathione S-Transferase. *Anal. Biochem.* **1991**, *192*, 262–267.

JM9407724