Determination of Affinities for *lck* SH2 Binding Peptides Using a Sensitive Fluorescence Assay: Comparison between the pYEEIP and pYQPQP Consensus Sequences Reveals Context-Dependent Binding Specificity

Ruby C. Cousins-Wasti, Richard H. Ingraham, Maurice M. Morelock,* and Christine A. Grygon*

Department of Inflammatory Diseases Research and Development, Section of Biophysics, Boehringer Ingelheim Pharmaceuticals, Inc., 900 Ridgebury Road, P.O. Box 368, Ridgefield, Connecticut 06877

Received August 19, 1996; Revised Manuscript Received October 30, 1996[⊗]

ABSTRACT: The development of a sensitive fluorescence binding assay for evaluating the binding of phosphotyrosyl (pY) peptides to the recombinant SH2 domain of *lck* in solution is described. Several fluorescent peptides containing the consensus sequence of the viral hamster polyoma middle T antigen (pYEEI) were characterized. The peptides contained either the acetamido-anilino-naphthyl sulfonic acid (AANS), acrylodan, or dansyl groups as fluorophores. The spectral features of these probes were characterized in the presence and absence of the *lck* SH2 domain. The binding affinities (K_d) for the fluorescent peptides studied ranged from 40 to 500 nM. The fluorescent peptide containing the sequence FTATEC(AANS)QpYEEIP exhibited the highest binding affinity ($K_d = 3.98 \times 10^{-8}$ M) and largest change in emission intensity (≈8.7-fold) upon binding the SH2 domain. This probe was subsequently used in competitive binding assays to study the interaction of the *lck* SH2 domain with a series of phosphopeptides related to the pYEEIP and pYQPQP (the pY^{505} C-terminal) consensus sequences. The effects of peptide length and substitutions of residues within the pYEEIP sequence are discussed in terms of binding affinities. Comparison between the two peptide series revealed that the contributions of individual substitutions to binding affinity are context-dependent. The data also led to the conclusion that the presence of P at +2 results in a functional "truncation" of the binding sequence; i.e., residues at positions higher than +2 do not participate significantly in binding. This implicit truncation may actually be a desired property for the autoregulatory nature of the pYQPQP sequence, since it retains specificity for the SH2 domain while adjusting the K_d to a value appropriate for maintaining the delicate balance of receptor−ligand interactions that are involved in signal transduction events.

Src-homology 2 (SH2) domains are 100-amino acid modules that are found in many signal transduction proteins (Isakov et al., 1994; Pawson, 1995; Koch et al., 1991). SH2 domains exert their signal-transducing effects through docking interactions involving the specific recognition of amino acid sequences containing phosphotyrosine (pY). Many of these binding interactions have been implicated in cellular proliferation and oncogenic transformation processes and therefore represent potential targets for therapeutic intervention in various immunological diseases and cancers [for a review, see Levitzki (1994)]. The intracellular (non-receptor-type) tyrosine kinase *lck*, which plays a critical role in signal transduction through the T-cell receptor, possesses an SH2 domain for interacting with pY-containing proteins. Part of the regulation of the tyrosine kinase activity of *lck* has been postulated to involve intramolecular binding of the inhibitory pYQPQP sequence from the carboxyl terminus (pY^{505}) to the SH2 domain (Amrein & Sefton, 1988; Marth et al., 1988; Bergman et al., 1992; Payne et al., 1993), possibly by blocking the kinase active site. This SH2 domain has also been found to bind tightly to peptides containing the sequence pYEEI (derived from the viral hamster polyoma middle T antigen protein). Interestingly, the SH2 domain of *lck* displays higher affinities for these pYEEI-containing peptides than for peptides containing the regulatory (native) pYQPQ sequence (Songyang et al., 1993). Other members of the *src* family of tyrosine kinases also display high affinity for the consensus pYEEI sequence (Songyang et al., 1993) and down-regulate their kinase activity via tyrosine phosphorylation of the carboxyl-terminal tail (Bergman et al., 1992).

Several high-resolution crystal structures of *src*-related SH2 domains have been recently reported (Waksman et al., 1992, 1993; Eck et al., 1993, 1994; Mikol et al., 1995; Tong et al., 1996). Among these are structures containing *lck* SH2 domains bound to phosphopeptides containing either the pYEEI or the pYQPQP motif. In both cases, positively charged and hydrophilic amino acid residues of the SH2 domain form a pY binding pocket. With peptides containing the pYEEI sequence, another pronounced structural feature is a well-defined hydrophobic pocket which accommodates insertion of the I side chain. These two residues (pY and I) appear to make the most substantial binding contacts, resembling a "two-pronged plug" (Waksman et al., 1993). In the corresponding structure containing the regulatory sequence, the Q at the +3 position does not associate with the hydrophobic pocket.

The crystallographic details of the interactions between pY-containing peptides and SH2 domains have been used in the rational design of SH2 antagonists. Several methods have previously been employed to quantitate the binding of phosphopeptides to SH2 domains, including ELISA-based techniques (Gilmer et al., 1994), surface plasmon resonance

* Corresponding authors. Telephone: 203-798-9988. Fax: 203-791-6468.

[⊗] Abstract published in *Advance ACS Abstracts,* December 1, 1996.

S0006-2960(96)02086-7 CCC: $12.00

(Payne et al., 1993; Ladbury et al., 1995; Morelock et al., 1995; Felder et al., 1993), and isothermal titration calorimetry (Ladbury et al., 1995; Lemmon & Ladbury, 1994). Binding affinities determined by surface plasmon resonance data can be affected by surface phenomena, i.e., affinity versus avidity (Ladbury et al., 1995), while titration calorimetry requires significant amounts of purified protein. A fluorescence assay has also been reported for determining the binding affinity of the *Grb2* SH2 domain for several pY-containing peptides (Cussac et al., 1994). This assay, which involves monitoring changes in Trp emission of the *Grb2* SH2 domain, is limited due to the following considerations: (1) the total change in emission intensity observed upon peptide binding is less than a factor of 2, and (2) peptide compounds that contain chromophores which absorb strongly in the region of excitation (290 nm) cannot be evaluated easily due to inner filter effects. The fluorescence assay presented herein circumvents the pitfalls of the techniques discussed above and was used to delineate differences in binding affinity for a series of peptides derived from the consensus pYEEIP and the autoregulatory pYQPQP sequences.

THEORY

The following equation is a generalized scheme for receptor−ligand exchange reactions:

$$\mathrm{RP} \leftrightarrows \mathrm{P} + \mathrm{R} + \mathrm{L} \leftrightarrows \mathrm{RL} \tag{1}$$

where R is the receptor, P is the probe ligand, and L is the competing ligand. The individual equilibrium expressions (K_d) for the above generalized reaction scheme are given by the following equations:

$$\mathrm{R} + \mathrm{P} \leftrightarrows \mathrm{RP} \quad \text{where} \quad K_\mathrm{P} = \frac{[\mathrm{R}][\mathrm{P}]}{[\mathrm{RP}]} \tag{2}$$

$$\mathrm{R} + \mathrm{L} \leftrightarrows \mathrm{RL} \quad \text{where} \quad K_\mathrm{L} = \frac{[\mathrm{R}][\mathrm{L}]}{[\mathrm{RL}]} \tag{3}$$

The dissociation constants K_P and K_L can be determined from an analysis of [R], [RP], and [RL] in eq 1 at a time sufficient to allow for equilibrium conditions.

Upon binding the receptor, the probe ligand must (1) display a change in fluorescence signal large enough to effectively detect changes in reactant concentrations and (2) maintain affinity (as this sets the limits of the assay). In addition, the molecular structure and biological activity of the probe should be similar to those of the inhibitor series of interest. Binding of a probe is typically monitored by observing changes in emission signal intensity at a fixed wavelength. A general expression for the total fluorescence observed in such experiments can be written as follows:

$$F_\mathrm{total} = Q_\mathrm{free}[\mathrm{probe}]_\mathrm{free} + Q_\mathrm{bound}[\mathrm{probe}]_\mathrm{bound}$$

or

$$F_\mathrm{T} = Q_\mathrm{f}([\mathrm{P}]_\mathrm{T} - [\mathrm{RP}]) + Q_\mathrm{b}[\mathrm{RP}] + F_\mathrm{bkg} \tag{4}$$

where Q_f and Q_b are the relative quantum yields for the free and bound probe, respectively, and F_bkg is the background fluorescence (due to buffer components, etc.). Note that F_T can be written as a function of only one reaction species, namely RP.

Determination of K_P for the Probe. The characterization of probe binding is typically carried out by adding receptor to a solution containing the probe and measuring the resulting F_T. An expression for bound probe can be obtained by substituting mass balance relationships for [R] and [P] into eq 2. Solving for [RP] leads to the following quadratic equation:

$$a[\mathrm{RP}]^2 + b[\mathrm{RP}] + c = 0 \tag{5}$$

where $a = 1$, $b = -(K_\mathrm{P} + [\mathrm{P}]_\mathrm{T} + [\mathrm{R}]_\mathrm{T})$, and $c = [\mathrm{P}]_\mathrm{T}[\mathrm{R}]_\mathrm{T}$. An equation for fitting the direct binding data (holding $[\mathrm{P}]_\mathrm{T}$ constant) is obtained by substituting the negative root of the quadratic solution for eq 5 into eq 4. Since the SH2 species is added to the solution containing the probe, the trailing emission of protein Trp residues can (depending upon the excitation wavelength) add to the fluorescence intensity. Hence, eq 4 is modified to include the protein's contribution to F_T as follows:

$$F_\mathrm{T} = Q_\mathrm{f}([\mathrm{P}]_\mathrm{T} - [\mathrm{RP}]) + Q_\mathrm{b}[\mathrm{RP}] + Q_\mathrm{e}[\mathrm{R}]_\mathrm{T} + F_\mathrm{bkg} \tag{6}$$

where Q_e is the relative quantum yield for the SH2 species. The parameters K_P, Q_f, Q_b, Q_e, and F_bkg are estimated by nonlinear regression, with $[\mathrm{R}]_\mathrm{T}$ and F_T as the independent and dependent variables, respectively.

Determination of K_L for Peptides (Competitive Binding). Characterization of a series of ligands is commonly performed by adding L to a solution containing R + P and measuring the resulting F_T. Assuming degenerate conditions for P (i.e., $[\mathrm{P}] \approx [\mathrm{P}]_\mathrm{T}$) and nondegenerate for L (i.e., $[\mathrm{L}] \neq [\mathrm{L}]_\mathrm{T}$), a model can be derived for [RP] as follows. Solving eq 3 for [R] using the mass balance relationship for L and substituting into eq 2 yields

$$K_\mathrm{P} = \frac{\dfrac{K_\mathrm{L}[\mathrm{RL}]}{([\mathrm{L}]_\mathrm{T} - [\mathrm{RL}])}[\mathrm{P}]_\mathrm{T}}{[\mathrm{RP}]} \tag{7}$$

Substituting eq 2 and the solution for [RL] from eq 7 into the mass balance expression for R gives

$$[\mathrm{R}]_\mathrm{T} = \frac{K_\mathrm{P}[\mathrm{RP}]}{[\mathrm{P}]_\mathrm{T}} + [\mathrm{RP}] + \frac{K_\mathrm{P}[\mathrm{L}]_\mathrm{T}[\mathrm{RP}]}{K_\mathrm{P}[\mathrm{RP}] + K_\mathrm{L}[\mathrm{P}]_\mathrm{T}} \tag{8}$$

Finally, solving for [RP] results in the following quadratic equation:

$$a[\mathrm{RP}]^2 + b[\mathrm{RP}] + c = 0 \tag{9}$$

where $a = (K_\mathrm{P}/[\mathrm{P}]_\mathrm{T} + 1)K_\mathrm{P}$, $b = (K_\mathrm{P}/[\mathrm{P}]_\mathrm{T} + 1)K_\mathrm{L}[\mathrm{P}]_\mathrm{T} + K_\mathrm{P}[\mathrm{L}]_\mathrm{T} - K_\mathrm{P}[\mathrm{R}]_\mathrm{T}$, and $c = -(K_\mathrm{L}[\mathrm{P}]_\mathrm{T}[\mathrm{R}]_\mathrm{T})$. An expression for fitting the competitive binding data (holding $[\mathrm{R}]_\mathrm{T}$ and $[\mathrm{P}]_\mathrm{T}$ constant) is obtained by substituting the positive root of the quadratic solution for eq 9 into eq 4. The parameters K_L, Q_f, Q_b, and F_bkg are estimated by nonlinear regression (using the K_P previously determined, *vide supra*), with $[\mathrm{L}]_\mathrm{T}$ and F as the independent and dependent variables, respectively.

MATERIALS AND METHODS

lck SH2 Domain. The proteins used in the fluorescence binding assay were (1) a glutathione *S*-transferase (GST) fusion protein, containing a 105-amino acid insert from the *lck* SH2 domain (Voronova & Sefton, 1986) expressed in

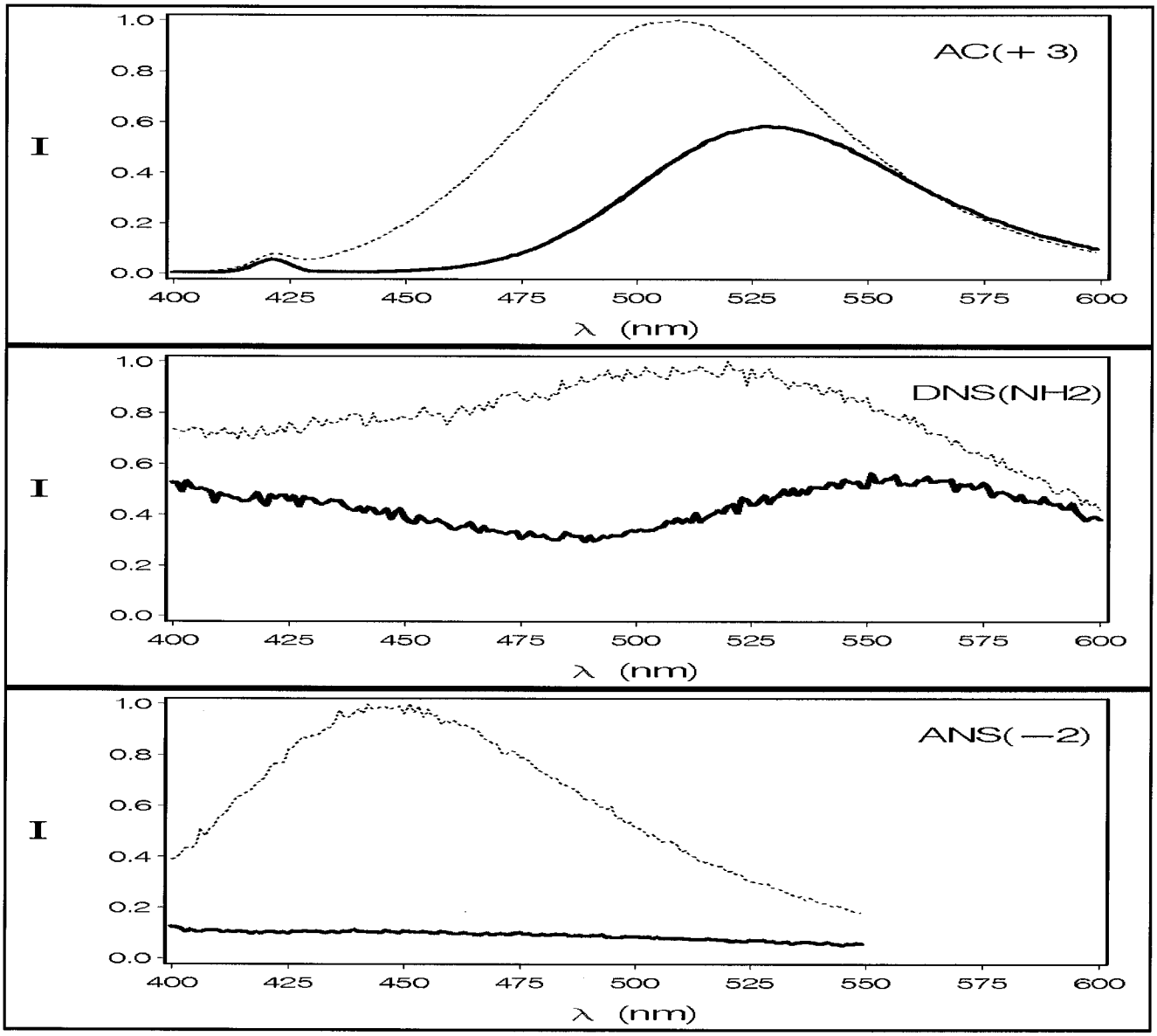

FIGURE 1: Fluorescence emission spectra of fluorophore-labeled peptides in 50 mM Tris buffer at pH 7.8 (solid lines) and in the presence of *lck* GST−SH2 domain (dashed lines): (top panel) [AC(+3)] = 8.6×10^{-8} M, [GST−SH2] = 5.0×10^{-7} M, λ_{ex} = 393 nm; (middle panel) [DNS(NH_2)] = 6.2×10^{-8} M, [GST−SH2] = 8.5×10^{-7} M, λ_{ex} = 335 nm; (bottom panel) [AANS(−2)] = 1.0×10^{-7} M, [GST−SH2] = 5×10^{-7} M, λ_{ex} = 330 nm.

the pGEX-KT vector (Hakes & Dixon, 1992), (2) a R134 to K134 point mutation, and (3) a nonfusion protein, residues 119−226, expressed in the pRSET vector.

Synthesis of Peptides and Fluorescent Peptides. Peptides were synthesized on an Applied Biosystems model 430A peptide synthesizer using Fmoc chemistry and HBTU-HOBT couplings (modified Fast-Moc cycles with extended coupling times). Peptides were cleaved with solutions of 95% TFA, 2.5% water, and 2.5% ethanedithiol followed by ether precipitation, dissolution in water, and lyophilization. Purification was accomplished via RP-HPLC using a TFA/water to TFA/acetonitrile gradient system with a preparative C18 column. Purified peptides were characterized by analytical RP-HPLC and mass spectrometry. Peptides to be used as fluorescent probes were labeled with either iodo-acetamido-anilino-naphthyl sulfonic acid (AANS) or acrylodan (AC, Molecular Probes, Inc.) on a Cys residue at a position either −2 [FTATEC(AANS or AC)QpYEEIP] or +3 [FTATEGQ-pYEEC(AANS or AC)P] relative to the phosphotyrosine or with a dansyl group (DNS) on the amino terminus [(DNS)-pYEE(*n*-methylnorleucine)P] and repurified to ensure homogeneity. Peptides to be labeled on a Cys residue were dissolved in a solution containing 100 mM sodium phosphate (pH 7.8) with 1 mM DTT at a concentration of 2.4 mg/mL peptide and stirred for 30 min at room temperature to ensure reduction of the Cys residue. Aliquots of either AANS or AC (60 mM in dimethylformamide) stock solution were added to peptide to yield a final concentration of 2 mM fluorophore. The labeling reaction proceeded at room temperature in the dark for 3 h. The DNS-labeled peptide was prepared using peptide dissolved in the same solution described above at a concentration of 1.4 mg/mL. Dansyl chloride (5.57 mM in acetonitrile) was added to the peptide solution to yield a final concentration of 1.07 mM. The reaction proceeded at room temperature in the dark for 3 h.

Fluorescence Measurements. Fluorescence assays were performed on an SLM-8100 photon-counting spectrofluorometer at 21.5 °C in 1 cm cuvettes. All fluorescence samples were prepared in a buffer containing 50 mM Tris, 150 mM NaCl, and 1 mM DTT at pH 7.8. The AANS(+3) stock solution was quantitated in methanol using an extinction coefficient of 26 000 M^{-1} cm^{-1} at 329 nm (Molecular

Probes 1992−1994 Catalog); all other labeled peptide concentrations were determined by amino acid analysis. Direct binding data were obtained by titrating *lck* SH2 domain into a solution containing 50−100 nM fluorescently labeled peptide. Fluorescence intensities were measured at the maximum excitation and emission wavelengths observed for the bound peptides. Stock solutions of the unlabeled peptides were prepared in water and quantitated using an extinction coefficient for phosphotyrosine of 652 M^{-1} cm^{-1} at 267 nm. Competition experiments were performed by titrating unlabeled peptides into a solution containing 100 nM GST *lck* SH2 domain and 1 μM AANS(−2). The reaction mixture was stirred for a minimum of 2 min after the addition of each aliquot. The excitation wavelength was 330 nm; emission was monitored at 450 nm. Excitation and emission bandwidths were set at 4 nm.

Nonlinear Regression Analysis. All of the data were analyzed using the SAS statistical software system for Windows (version 6.11, SAS Institute Inc., Cary, NC) on an HP Vectra PC (Hewlett-Packard Co., Palo Alto, CA). Typically, ASCII data files containing fluorescence measurements were converted into SAS data sets which included reactant concentrations. Data analyses were performed by applying ordinary nonlinear least-squares regression techniques to the selected model (see Theory) using the Marquardt−Levenberg minimization method. A grid of the parameters, using the lowest sum of squares criteria, determined the initial estimates.

RESULTS AND DISCUSSION

Spectroscopic methodologies for studying binding events in solution are well established (Lakowicz, 1983; Ward, 1985). Fluorescence spectroscopy is often ideally suited for biological studies due to increased sensitivity over other techniques; i.e., less sample is required for analysis. As a result, tight-binding phenomena which involve low nanomolar concentrations of reactants can be studied. Fluorophores that are spectroscopically sensitive to their local environment have proven particularly useful in ligand binding studies, e.g., AANS, AC, and DNS groups (Prendergast et al., 1983; Haugland, 1992).

Selection of the Probe

AANS or AC labels were incorporated into several peptides containing the pYEEIP consensus sequence at the −2 or +3 positions (relative to pY) through a substituted Cys residue. The DNS derivative was prepared by linking the fluorophore to the amino terminus of a similar peptide. All of the peptides examined displayed substantial spectral changes upon binding to the SH2 domain of *lck* (see Figure 1). The AC-labeled peptides showed a moderate increase in emission intensity and a blue shift in the emission maximum, indicating that the local environment of this fluorophore is more hydrophobic when bound to the SH2 domain (Figure 1). FTATEGQpYEEC(AC)P [subsequently referred to as AC(+3)] exhibits a larger blue shift (≈20 nm) than AC(−2) (≈10 nm). This is consistent with the crystal structure (Eck et al., 1993) in which the side chain of the I(+3) residue was shown to be bound by a hydrophobic pocket on the SH2 domain. The dansyl-labeled peptide displayed similar (but less dramatic) fluorescence changes, with an increase in emission intensity and a significant blue shift accompanying binding (Figure 1).

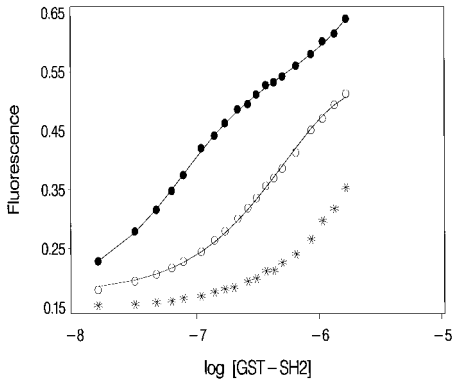

FIGURE 2: Fluorescence intensity of AANS(−2) (recorded at λ_{ex} = 330 nm and λ_{em} = 450 nm) plotted as a function of *lck* GST−SH2 concentration. Data for wild type GST−SH2 (●) were fit by nonlinear regression analysis (applying eqs 5 and 6) which converged to the following solution (solid line): K_d = 4.49(0.21) × 10^{-8} M (where the value in parentheses is the standard error). Data for the R134K mutant (○) resulted in a K_d of 4.33(0.27) × 10^{-7} M. The data given by asterisks (*) are for a nonphosphorylated AANS(−2) species.

Table 1: Dissociation Constants for SH2 Binding Probes[a]

peptide	K_d (M) GST−SH2	K_d (M) SH2
AC(−2)	6.62 (±1.20) × 10^{-8}	3.73 (±0.26) × 10^{-8}
AC(+3)	5.19 (±0.51) × 10^{-8}	1.34 (±0.00) × 10^{-7}
DNS(NH_2)	1.77 (±0.01) × 10^{-7}	nd
AANS(−2)	3.98 (±0.44) × 10^{-8}	5.27 (±3.03) × 10^{-8}
AANS(+3)	3.51 (±0.37) × 10^{-7}	4.97 (±0.45) × 10^{-7}

[a] K_d values listed were determined by nonlinear regression using eqs 4 and 5 for AC and DNS and eqs 5 and 6 for AANS. The values in parentheses are the standard deviations for n = 2; for AANS with GST−SH2, n = 3. nd = not determined.

The AANS derivatives produced the most dramatic spectral changes. AANS-labeled molecules are often employed in probing hydrophobic binding sites since they are relatively non-fluorescent in polar (aqueous) environments but display significant emission intensity increases in nonpolar (hydrophobic pocket) environments. The emission spectra for AANS(−2) in the presence and absence of GST *lck* SH2 domain are shown in Figure 1. Upon binding to the SH2 domain, a large enhancement in emission intensity is observed at ≈450 nm (λ_{ex} = 330 nm), consistent with spectral changes typically reported for AANS probes (Stryer, 1968).

The binding affinity of a probe can be determined by monitoring the change in fluorescence during a titration experiment. Typically, the protein is added to the probe so that the signal due to free probe is minimized; i.e., all probe will be bound at high concentrations of the nonlabeled species, here GST−SH2. The fluorescence changes which occur upon binding can be used to determine equilibrium constants for each probe. Figure 2 shows representative binding curves for the interaction between the AANS(−2) probe and the SH2 domain of *lck*. As GST−SH2 is titrated into the reaction cuvette, a nonlinear increase in fluorescence intensity is observed at 450 nm. The fluorescence does not appear to saturate at high GST−SH2 concentrations but continues to increase due to the trailing emission of protein Trp residues. These data were fit by the model given in eqs 5 and 6, giving a K_d of 3.98 (±0.44) × 10^{-8} M (n = 3, see Table 1). In a control experiment, a nonphosphorylated analog of this peptide exhibited no binding (see Figure 2);

these data were similar to a titration of GST−SH2 into buffer containing no AANS probe (data not shown). Thus, the increase in fluorescence observed at high GST−SH2 concentrations is presumably due to either the trailing emission of protein Trp or a fluorescent impurity present in the protein stock solution. In either case, the increase would be (and is) linear and can be modeled (see eq 6). This result confirms that the AANS moiety on the peptide does not participate in any nonspecific interactions with the SH2 domain.

Both AC-labeled peptides bind GST−SH2 with K_d values of <100 nM. These results indicate that the AC moiety of AC(+3) substitutes for the hydrophobic I side chain at the +3 position. Since the probes with the AC moiety at the −2 and +3 positions displayed good binding characteristics, and since AANS(−2) behaves similarly, AANS at +3 might also be expected to bind tightly. Nevertheless, the K_d for the AANS(+3)-labeled peptide was significantly higher than that of any of the other probes studied. While the hydrophobic pocket in the SH2 binding site can accommodate an AC moiety, it apparently cannot make favorable contacts with an AANS group at this position. The AANS- and AC-labeled peptides were also evaluated against the free *lck* SH2 domain (without the GST fusion protein). The K_d values obtained for the free SH2 were comparable to those for GST−SH2 (see Table 1), indicating that GST does not participate in or alter peptide binding. Consequently, the probe chosen for competitive binding studies was AANS(−2) due to its high affinity for the *lck* SH2 domain (3.98×10^{-8} M) and large fluorescence intensity increase upon binding [$Q_b/Q_f = 8.06$, $Q_f = 7.90\ (\pm 0.92) \times 10^5$, $Q_b = 6.37\ (\pm 0.99) \times 10^6$, $n = 40$].

Determination of Peptide Affinities

Fluorophore-labeled probes which interact with the binding site of the SH2 domain can be used to determine binding affinities of nonlabeled (spectroscopically silent) peptides by carrying out competitive equilibrium reactions. Typically, the GST−SH2 is added to a cuvette containing the AANS(−2) so that [AANS(−2)]/[GST−SH2] ≈ 10. Under the conditions employed, >90% of the GST−SH2 binding sites are bound with AANS(−2). Aliquots of a nonlabeled peptide are then sequentially added to the reaction mixture. Nonlabeled peptide competes with the probe for SH2 binding sites, and the resulting displacement of AANS(−2) is shown by a decrease in fluorescence intensity (see Figure 3). These data are then fit to a competitive binding model given by eqs 4 and 9. Several representative binding curves are shown in Figure 3 for a series of peptides with K_d values ranging from 6×10^{-8} to 5×10^{-5} M.

Peptides Containing the pYEEI Motif. Table 2 summarizes the dissociation constants obtained using the competitive method (described above) for several peptides containing the pYEEI consensus sequence. The starting sequence EGQpYEEIP was determined to bind the SH2 domain of *lck* with a K_d of 6.55×10^{-8} M, which compares well with that of the structurally similar AANS(−2) probe (3.98×10^{-8} M). A nonphosphorylated analog was unable to displace the probe up to a concentration of 50 μM (data not shown), underlining the requirement of the phosphotyrosine for tight binding. The peptide series listed in Table 2A (see also Figure 3) delineates the effect of decreasing the length of this peptide. Deletion of the two amino-terminal residues (E and G) results in a 4-fold loss in binding affinity. If, however, these two

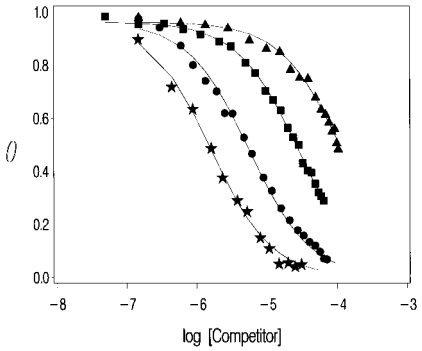

FIGURE 3: Application of the competitive equilibrium method to a series of pYEEI-containing peptides in which residues amino-terminal to pY have been truncated to examine the effect of peptide length on affinity. Representative binding curves with *lck* GST−SH2 were generated by displacement of the probe AANS(−2) ($K_d = 3.98 \times 10^{-8}$ M) with competing nonlabeled peptides. Data for an arbitrary experiment have been normalized to each other by converting fluorescence data to the fraction bound ($\theta = [RP]/[R]_{total}$) and plotted versus competitor concentration. The peptides shown are EGQpYEEIP [stars, $K_d = 5.97(0.57) \times 10^{-8}$ M], QpYEEIP [circles, $K_d = 2.12(0.15) \times 10^{-7}$ M], pYEEIP [squares, $K_d = 1.07(0.06) \times 10^{-6}$ M], and desNHpYEEIP [triangles, $K_d = 4.52(1.00) \times 10^{-5}$ M].

Table 2: Dissociation Constants for pYEEI Peptides[a]

peptide sequence	K_d (M)	relative K_d
(A) Effect of Peptide Length		
EGQpYEEIP	$6.55\ (\pm 0.60) \times 10^{-8}$	1
AcQpYEEIP	$6.46\ (\pm 0.81) \times 10^{-8}$	1
QpYEEIP	$2.41\ (\pm 0.40) \times 10^{-7}$	4
AcpYEEIP	$1.24\ (\pm 0.28) \times 10^{-7}$	2
pYEEIP	$1.08\ (\pm 0.01) \times 10^{-6}$	16
desNHpYEEI	$4.49\ (\pm 0.06) \times 10^{-6}$	69
pYEEI	$2.94\ (\pm 1.00) \times 10^{-6}$	45
EGQpYEEI	$3.11\ (\pm 0.00) \times 10^{-7}$	5
EGQpYEEI-NH_2	$3.52\ (\pm 0.02) \times 10^{-7}$	5
(B) Effect of G Substitution at the +1, +2, and +3 Positions		
EGQpYEEIP	$6.55\ (\pm 0.60) \times 10^{-8}$	1
EGQpY**G**EIP	$8.85\ (\pm 0.30) \times 10^{-6}$	135
EGQpYE**G**IP	$7.66\ (\pm 3.05) \times 10^{-7}$	12
EGQpYEE**G**P	$2.16\ (\pm 0.17) \times 10^{-6}$	33

[a] The equilibrium constants were determined using eqs 4 and 9 with a K_d of 3.98×10^{-8} M for AANS**(**−2**)**. The values in parentheses are the standard deviations for $n = 2$; for EGQpYEEIP, $n = 3$.

residues are replaced with an acetyl group (Ac), virtually no loss in binding affinity is observed. Deleting all three amino-terminal residues (pYEEIP) results in a 16-fold reduction in affinity, pointing out the importance of Q in the −1 position. Cocrystal structures have shown that Q is hydrogen bonded to R134 (Waksman et al., 1993; Eck et al., 1993, 1994; Mikol et al., 1995). It may also be possible that this loss in affinity is due in part to charge repulsion between the protonated amino terminus of pYEEIP and R134. In order to differentiate between these two possibilities, a peptide was synthesized which lacked an amino group on the pY, i.e., desNHpYEEI. As can be seen in Table 2A, removal of the charged amino terminus does not enhance the binding affinity. Since the relative affinities between QpYEEIP and AcpYEEIP are also similar, the carbonyl at the −1 position must make the most significant contact for the amino-terminal residues through a H bond to R134. An alternative approach to answering this question involves the use of the *lck* SH2 mutant R134K. Although the K residue can interact with the phosphate on pY (Waksman et al., 1992), it is incapable of forming a simultaneous H bond to the −1 carbonyl of Q. The K_d for binding of AANS(−2) to

mutant *lck* GST−SH2 was 3.86 (±0.67, $n = 2$) × 10^{-7} M (see Figure 2), about 10-fold higher than that for binding to wild type SH2 (3.98 × 10^{-8} M). This loss in binding affinity is relatively close to that observed upon removal of the Ac group of AcpYEEIP and provides complementary evidence for the importance of the H bond formed between R134 and the carbonyl at the −1 position.

Similar losses in binding affinity are observed upon truncation of the starting peptide EGQpYEEIP at the C terminus. Since the crystal structure also shows a H bond from R95 to the backbone carbonyl at the +4 position (Eck et al., 1993), truncation at the +4 residue (abolishing this H bond) should decrease the binding affinity. As can be seen in Table 2A, deleting P at +4 from the starting peptide results in a 5-fold decrease in binding affinity, consistent with the prediction from crystal structure data. Amidation of this truncated species, i.e., EGQpYEEI-NH_2 (which removes the C-terminal charge), does not significantly improve the binding affinity (see Table 2A). Consequently, the backbone carbonyl at the +4 position makes an important H bond contact with the SH2 domain.

The peptides listed in Table 2B explore the importance of the side chains at the +1, +2, and +3 positions. Each residue was separately replaced with G. Substitution at any of these three positions resulted in at least a 10-fold loss in binding affinity. The same relative binding affinities for this series of peptides have also been reported using a surface plasmon resonance method (Morelock et al., 1995). The most dramatic effects were observed at the +1 (8.85 × 10^{-6} M) and +3 (2.16 × 10^{-6} M) positions. The loss in affinity observed for the +3 G peptide is expected since the I side chain of the parent peptide makes several hydrophobic contacts within a deep binding pocket. Obviously, replacing this side chain with a hydrogen atom precludes these interactions. The decrease in binding affinity for substitution at the +1 position has also been observed for *src* SH3−SH2 (Gilmer et al., 1994). Although the published peptide−SH2 cocrystal structures show that the +1 and +2 glutamic side chains are directed away from the binding cleft, the β carbon in the +1 position extends into the cleft and interacts in a hydrophobic manner with the phenyl ring of Y181, and the COO^- group H bonds with the phenolic group of Y181 through a water molecule (Eck et al., 1993; Tong et al., 1996). While the backbone of a G residue at +1 can interact with the binding cleft, the H side chain cannot make the indicated interactions with Y181.

Comparison of the pYEEIP and pYQPQP Motifs. The SH2 domain of *lck* has been reported (Payne et al., 1993) to bind the hamster middle T antigen sequence (pYEEIP) with higher affinity than its own autoregulatory sequence (pYQPQP). The data reported in this paper show an 11-fold difference in affinity for AcpYEEIP vs AcpYQPQP (see Table 3). Since the side chains of residues at the +2 and +3 positions of these peptides are so dissimilar, this difference was expected to be much larger. To better understand this result, two series of peptides were synthesized to examine the issue of selectivity and how it is manifested within the context of these two series (see Table 3).

In the first series, single residues from the AcpYQPQP sequence were individually substituted into the AcpYEEIP sequence. As expected, due to the conservative nature of the replacement, similar affinities were observed upon substitution of Q for E at the +1 position. This is also consistent with crystal structure data that show that E and Q at +1 make similar contacts with the SH2 binding site for both parent peptides (Eck et al., 1993, 1994). Substitution of P for E at the +2 position results in a 6-fold loss in affinity. It is possible that this loss in binding affinity is due to the elimination of interactions involving the side chain of E. The cocrystal structure of the pYQPQP sequence with SH2−SH3 (Eck et al., 1994), however, shows that P at +2 "kinks" the peptide backbone, preventing the +2 and +3 residues from making close contact with the SH2 binding site. In an analogous manner, P at +2 is also presumed to kink the AcpYEPIP peptide, resulting in poor interactions of I with the *lck* SH2 hydrophobic binding pocket. Since binding between the I at +3 and the hydrophobic pocket involves multiple interactions, it is suprising that the replacement of this residue with the hydrophilic Q decreases the binding affinity only 8-fold. Thus, while the interaction between I at +3 and the hydrophobic binding pocket is significant, the residues at the +2 and particularly the +1 positions (see G at +1 in Table 2B) contribute more to the binding energy than would have been predicted by the two-prong (pYXXI) plug model.

Table 3: Comparison of Binding Constants for pYEEIP- and pYQPQP-Related Peptides[a]

peptide	K_d (M)	relative K_d
AcpYEEIP	1.24 (±0.28) × 10^{-7}	1
AcpYQEIP	2.18 (±0.11) × 10^{-7}	2
AcpYEPIP	7.26 (±0.40) × 10^{-7}	6
AcpYEEQP	9.65 (±0.06) × 10^{-7}	8
AcpYQPQP	1.41 (±0.06) × 10^{-6}	11
AcpYEPQP	8.53 (±0.51) × 10^{-7}	7
AcpYQEQP	1.85 (±0.30) × 10^{-6}	15
AcpYQPIP	8.58 (±0.09) × 10^{-7}	7

[a] The equilibrium constants were determined using eqs 4 and 9 with a K_d of 3.98 × 10^{-8} M for AANS**(**−2**)**. The values in parentheses are the standard deviations for $n = 2$.

In the second series, single residues from the AcpYEEIP sequence were individually substituted into the AcpYQPQP sequence. As expected (see above), substitution of E for Q at +1 was found to have a relatively small effect on the binding affinity (since both side chains are capable of making similar interactions). Substitution of E for P at the +2 position also did not significantly affect the binding affinity, since Q remains in the +3 position (see AcpYEEQP in Table 3). Similarly, replacement of Q with I at +3 did not result in a significant gain in binding affinity, since P remains at the +2 position, kinking the peptide. If the binding energies of the three substitutions QPQ for EEI were cumulative, then the expected increase in K_d on going from AcpYEEIP to AcpYQPQP would be approximately 96-fold (2 × 6 × 8) instead of 11-fold [see also Ingraham et al. (1995)]. It is clear from these results that the contributions of the +2 and +3 side chains to the binding energy are not simply additive but are dependent upon adjoining residues. A recent study reported that peptides containing the consensus (pYEEIP) and autoregulatory (pYQPQP) sequences bind with similar affinities (Payne et al., 1994). This study also noted that the contributions of individual substitutions to binding affinity were context-dependent; e.g., substitution of Q for I at +3 in the parent pYEEI peptide resulted in a larger loss in binding affinity (24-fold) than substitution of the entire consensus sequence (5-fold).

The data presented herein show that the binding affinities for the AcpYEEIP and AcpYQPQP peptides differ by 11-

fold (see Table 3). In addition, peptides having residues truncated from the carboxy terminus, such as AcpYQPQ (from *lck*) and AcpYQPG (from *src*, *yes*, *fyn*, and *fgr*), were found to bind to *lck* SH2 with affinities comparable to that of the parent peptide AcpYQPQP (data not shown), despite the lack of a +4 residue and the significant difference in their +3 side chains. The cocrystallographic data for the *lck* SH2 with two peptides containing the consensus and autoregulatory sequences show the following features: (1) no significant conformational changes occur within the binding cleft and (2) the two peptides bind with similar interactions at the −1 backbone carbonyl, the pY and +1 positions, and (3) dissimilar interactions at ≥+2. The binding motif defined by P at +2 is not limited to the *lck* SH2 domain but occurs in the majority of *src* family members (Cheng et al., 1988). For example, the cocrystal structure for the platelet-derived growth factor peptide (pYVPML) with *src* SH2 (Waksman et al., 1992, 1993) also shows that P at the +2 position kinks the peptide, preventing significant interactions between M and the hydrophobic pocket. This appears to result from a 4.5 Å shift of the peptide backbone away from the pocket, which limits the extent of the interaction between the SH2 domain and the +3 position of the peptide. In view of these crystallographic results, the data presented herein suggest that the AcpYQPQP sequence has been functionally truncated by the presence of P at the +2 position; i.e., residues at positions higher than +2 do not participate significantly in binding.

In conclusion, comparison of peptides related to pYEEIP and pYQPQP sequences revealed that the contributions of individual substitutions to binding affinity are context-dependent. The data are best explained by a model in which the presence of P at +2 results in a functional truncation of the binding sequence. This implicit truncation may actually be a desired property for the autoregulatory nature of this sequence for the following reasons. Firstly, the specificity of pYQPQP (a proposed intramolecular regulator of *lck* kinase activity) for the SH2 domain is retained. Secondly, a tighter-binding sequence (such as pYEEIP) may have led to a constitutively inactive kinase. Thirdly, internally adjusting the K_d of the autoregulatory binding sequence may provide a mechanism for maintaining the delicate balance of receptor−ligand interactions that are involved in signal transduction events.

ACKNOWLEDGMENT

We are especially grateful to Josephine Schembri-King for the cloning, expression, and purification of SH2 domain constructs. We also thank Thomas C. Warren for additional protein purification, Raj Betageri for the synthesis of several peptides, and Daniel Greenwood for measuring the extinction coefficient of phosphotyrosine. We also thank Scott Jakes, Liang Tong, Mario Cardozo, and James Stevenson for helpful discussions and critical reading of the manuscript.

REFERENCES

Amrein, K. E., & Sefton, B. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4247−4251.

Bergman, M., Mustelin, T., Oetken, C., Partanen, J., Flint, N. A., Amrein, K. E., Autero, M., Burn, P., & Alitalo, K. (1992) *EMBO J. 11*, 2919−2924.

Cheng, S. H., Harvey, R., Espino, P. C., Semba, K., Yamamoto, T., Toyoshima, K., & Smith, A. E. (1988) *EMBO J. 7*, 3845−3855.

Cussac, D., Frech, M., & Chardin, P. (1994) *EMBO J. 13*, 4011−4021.

Eck, M. J., Shoelson, S. E., & Harrison, S. C. (1993) *Nature 362*, 87−91.

Eck, M. J., Atwell, S. K., Shoelson, S. E., & Harrison, S. C. (1994) *Nature 368*, 764−769.

Felder, S., Zhou, M., Hu, P., Urena, J., Ullrich, A., Chaudhuri, M., White, M., Shoelson, S. E., & Schlessinger, J. (1993) *Mol. Cell Biol. 13*, 1449−1455.

Gilmer, T., Rodriguez, M., Jordan, S., Crosby, R., Alligood, K., Green, M., Kimery, M., Wagner, C., Kinder, D., Charifson, P., Hassell, A. M., Willard, D., Luther, M., Rusnak, D., Sternbach, D. D., Mehrota, M., Peel, M., Shampine, L., Davis, R., Robbins, J., Patel, I., Kassel, D., Burkhart, W., Moyer, M., Bradshaw, T., & Berman, J. (1994) *J. Biol. Chem. 269*, 31711−31719.

Hakes, D. J., & Dixon, J. E. (1992) *Anal. Biochem. 202*, 293−298.

Haugland, R. P. (1992) *Molecular Probes Handbook of Fluorescent Probes and Research Chemicals*, 5th ed, Molecular Probes Inc., Eugene, OR.

Ingraham, R. H., Wasti, R., Greenwood, D. J., Morelock, M. M., & Grygon, C. A. (1995) *Protein Sci. 4*, 333-S.

Isakov, N., Wange, R. L., & Samelson, L. E. (1994) *J. Leukocyte Biol. 55*, 265−271.

Koch, C. A., Anderson, D., Moran, M. F., Ellis, C., & Pawson, T. (1991) *Science 252*, 668−674.

Ladbury, J. E., Lemmon, M. A., Zhou, M., Green, J., Botfield, M. C., & Schlessinger, J. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 3199−3203.

Lakowicz, J. R. (1983) *Principles of Fluorescence Spectroscopy*, Plenum Press, New York.

Lemmon, M. A., & Ladbury, J. E. (1994) *Biochemistry 33*, 5070−5076.

Levitzki, A. (1994) *Eur. J. Biochem. 226*, 1−13.

Marth, J. D., Cooper, J. A., King, C. S., Ziegler, S. F., Tinker, D. A., Overell, R. W., Krebs, E. G., & Perlmutter, R. M. (1988) *Mol. Cell. Biol. 8*, 540−550.

Mikol, V., Baumann, G., Keller, T. H., Manning, U., & Zurini, M. G. M. (1995) *J. Mol. Biol. 220*, 344−355.

Morelock, M. M., Ingraham, R. H., Betageri, R., & Jakes, S. (1995) *J. Med. Chem. 38*, 1309−1318.

Pawson, T. (1995) *Adv. Cancer Res. 64*, 87−110.

Payne, G., Shoelson, S. E., Gish, G. D., Pawson,T., & Walsh, C. T. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 4902−4906.

Payne, G., Stolz, L. A., Pei, D., Band, H., Shoelson, S. E., & Walsh, C. T. (1994) *Chem. Biol. 1*, 99−105.

Prendergast, F. G., Meyers, M., Carlson, G. L., Iida, S., & Potter, J. D. (1983) *J. Biol. Chem. 258*, 7541−7544.

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., Hanafusa, H., Shaffhausen, B., & Cantley, L. C. (1993) *Cell 72*, 767−778.

Stryer, L. (1968) *Science 162*, 526−533.

Tong, L., Warren, T. C., Schembri-King, J., Betageri, R., Rose, J., & Jakes, S. (1996) *J. Mol. Biol. 256*, 601−610.

Voronova, A. F., & Sefton, B. M. (1986) *Nature 319*, 682−685.

Waksman, G., Kominos, D., Robertson, S. C., Pant, N., Baltimore, D., Birge, R. B., Cowburn, D., Hanafusa, H., Mayer, B. J., Overduin, M., Resh, M. D., Rios, C. B., Silverman, L., & Kuriyan, J. (1992) *Nature 358*, 646−653.

Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D., & Kuriyan, J. (1993) *Cell 72*, 779−790.

Ward, L. D. (1985) *Methods Enzymol. 117*, 400−414.

BI9620868