Exploration of the Sequence Specificity of pp60<sup>−src</sup> Tyrosine Kinase

MINIMAL PEPTIDE SEQUENCE REQUIRED FOR MAXIMAL ACTIVITY*

(Received for publication, March 6, 1995, and in revised form, June 1, 1995)

Ann M. Edison, Sean C. Barker, Daniel B. Kassel, Michael A. Luther, and Wilson B. Knight‡

From the Division of Molecular Sciences, Glaxo Research Institute, Research Triangle Park, North Carolina 27709

The minimum length required for phosphorylation of a peptide by pp60<sup>−src</sup> tyrosine kinase (srcTK) was delineated in this work. Budde (M. D. Anderson University of Texas, personal communication) suggested that the peptide (FGE)Y(GEF)2GD (peptide I) was a “good” srcTK substrate. Peptide I yielded a 251-fold higher \( k_{cat}/K_m \) than RRLIEDAYAARG, a peptide substrate based upon the autophosphorylation site of srcTK. This was due to a 38-fold lower \( K_m \) and a 6.6-fold increase in \( k_{cat} \). N-terminal truncation of up to 8 residues in a series of peptides yielded only a 3-fold decrease in activity. Removal of the final N-terminal residue resulted in a 10-fold decrease in substrate activity, primarily as a result of an increase in the \( K_m \). C-terminal truncations ending in the amide yielded no significant loss in activity until the Y+3 residue was removed, which resulted in a 73-fold decrease in \( k_{cat}/K_m \) relative to peptide I. The latter was due primarily to an increase in \( K_m \). The results from peptides truncated on both termini suggest that substrate recognition N- and C-terminal relative to the site of phosphorylation can be examined independently. In addition, the observation that only 5 residues are required for significant substrate activity suggests that small molecule inhibitors based upon interactions with the phosphoacceptor site may be developed.

Protein tyrosine kinases (TKs)<sup>4</sup> were initially discovered as either oncogenes or proto-oncogene products, pointing to their therapeutic potential as targets in cancer (for a review of oncogenes, see Pimentel (1989a and 1989b)). The phosphorylation equilibrium of protein tyrosine residues is regulated by cytokines and growth factors, pointing to roles for these equilibria in signal transduction pathways. Therefore, control of these equilibria has the potential for therapeutic intervention in diseases including cancer, inflammatory diseases, and diabetes, to name a few.

While initially TKs were thought to be nonspecific, more recent work has demonstrated that they do indeed phosphorylate specific substrates in vivo (for example, see Ogawa et al. (1994)). The study of these enzymes has been hampered by the lack of specific peptide substrates. The specificity requirements for amino acid sequences remain to be elucidated. While several groups proposed that the specificity of TKs was not governed by the amino acid sequence surrounding the tyrosine (Tinker et al., 1988; Radziejewski et al., 1989), more recent studies have suggested that there is specific recognition of proximal residues. For example, Garcia et al. (1993) reported differences in the peptide sequences recognized by pp60<sup>−src</sup> and v-abl TK. These workers found that substitution of N<sup>3</sup> for D in the Y–1 position of KKSRGDYMMTMQIG, a peptide based upon a phosphorylation site of insulin receptor substrate-1, resulted in complete loss of substrate activity for both TKs. Substitution of I for M in the Y+1 position resulted in a 4-fold loss in catalytic efficiency for pp60<sup>−src</sup> but a 10-fold increase in the catalytic efficiency versus v-abl-TK. Till et al. (1994) have suggested that I is preferred by 6-fold over E and L at the Y–1 site of v-abl-TK substrates. Tinker et al. (1992) suggested that acidic residues N-terminal to the tyrosine were important for pp6<sup>−-src</sup> activity. The Y–3 and Y–4 positions displayed the greatest sensitivity to E versus A at these positions. A clear-cut correlation for pp6<sup>−src</sup> activity was not obtained in those studies. Songyang et al. (1995) derived some specificity requirements governing substrates for 9 TKs using peptide libraries. These workers report that the optimal substrate for srcTK would contain the sequence EELYGFF, although the scoring of additional residues in various positions suggests considerable flexibility. Barker et al. suggested that srcTK prefers smaller hydrophobic residues at the Y+1 position when angiotensin analogs were used as substrates. This proposal is consistent with the conclusion drawn from the peptide library studies. Wong and Goldberg (1983) reported that pp60<sup>−src</sup> displayed similar activity toward the two angiotensin analogs. Knowledge of the sequence specificity of TKs could allow the development of specific inhibitors that interact with the protein binding subsites.

Typically, members of the srcTK family have been assayed with either peptides based upon the autophosphorylation site of srcTK (Hunter, 1982; Casnellie et al., 1982; Wong and Goldberg, 1983) or peptides based upon angiotensin (Wong and Goldberg, 1983). The utility of these and other reported peptide TK substrates is limited by inefficient kinetic constants and the size of the peptides (from 8 to 13 residues). The only relatively systematic examination of the length of peptide required for TK substrate activity was reported by Cola et al. (1989). These

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: the Dept. of Enzymology, Glaxo Wellcome Inc., 5 Moore Dr., Research Triangle Park, NC 27709.

The abbreviations used are: TK, tyrosine kinase; HPLC, high pressure liquid chromatography; srcTK, pp60<sup>−src</sup> tyrosine kinase.

2 srcTK is used to refer to protein constructs containing at least the kinase domain of human pp60<sup>−src</sup>. pp60<sup>−src</sup> refers to the gene product of the Rous sarcoma virus. The mutant protein used in this work was N-85-srcTK, which refers to the 85-residue N-terminal deletion mutant that contains the SH3, SH2, and tyrosine kinase domains of pp60<sup>−src</sup>. This mutant protein displays similar kinetic parameters toward substrates essentially identical to those of the full-length protein (Footnote 5).

3 Peptides are named according to the 1-letter designations for amino acid residues. Phosphotyrosine residues are abbreviated pY.

4 The peptides are numbered according to the following system. The residues C-terminal to the tyrosine are given positive values beginning with Y equal to 0. The residues N-terminal to the tyrosine are given negative values beginning with Y equal to 0.

5 Barker, S. C., Kassid, D., Weigl, D., Huang, X., Luther, M. A., and Knight, W. B. (1995) Biochemistry, in press.
The dependence of srcTK kinetic parameters on the length of the peptide substrate

| Peptide                          | $K_m$  | $k_{cat}$ | $k_{cat}/K_m$ |
|---------------------------------|--------|----------|---------------|
| RRLIEDAeyaarrg                  | 2.80 x 10^-3 | 3.49 | 1.25 x 10^4 |
| AEEEEEYGFEEAKK                  | 3.3 x 10^-5 | 0.95 | 2.9 x 10^4 |
| FGEFGFGEYGFEGFEGD              | 7.30 x 10^-5 | 23.0 | 3.14 x 10^4 |
| AcGFGEFGFGEYGFEGFEGD           | 8.37 x 10^-5 | 21.0 | 2.52 x 10^4 |
| AcGFGEFGFGEYGFEGFEGD           | 9.05 x 10^-5 | 22.0 | 2.43 x 10^4 |
| AcGFGEFGFGEYGFEGFEGD           | 6.49 x 10^-5 | 16.5 | 2.54 x 10^4 |
| AcGFGEFGFGEYGFEGFEGD           | 8.25 x 10^-5 | 26.5 | 3.22 x 10^4 |
| AcGFGEFGFGEYGFEGFEGD           | 1.32 x 10^-4 | 24.0 | 1.82 x 10^4 |
| AcGFGEFGFGEYGFEGFEGD           | 2.20 x 10^-4 | 35.6 | 1.62 x 10^4 |
| AcGFGEFGFGEYGFEGFEGD           | 3.05 x 10^-4 | 34.8 | 1.14 x 10^4 |
| AcGFGEFGFGEYGFEGFEGD           | 4.35 x 10^-4 | 44.5 | 1.02 x 10^4 |
| AcGFGEFGFGEYGFEGFEGD           | 2.30 x 10^-3 | 27.4 | 1.2 x 10^4 |
| FGEFGFGEYGFEGFEGFEGD-amide     | 5.10 x 10^-5 | 16.7 | 3.28 x 10^3 |
| FGEFGFGEYGFEGFEGFEGD-amide     | 4.40 x 10^-5 | 19.1 | 4.35 x 10^3 |
| FGEFGFGEYGFEGFEGFEGD-amide     | 1.25 x 10^-3 | 24.1 | 1.93 x 10^3 |
| FGEFGFGEYGFEGFEGFEGD-amide     | 4.10 x 10^-5 | 25.9 | 6.32 x 10^2 |
| FGEFGFGEYGFEGFEGFEGD-amide     | 9.27 x 10^-5 | 24.1 | 2.6 x 10^2 |
| FGEFGFGEYGFEGFEGFEGD-amide     | 2.49 x 10^-3 | 10.7 | 4.29 x 10^1 |
| FGEFGFGEYGFEGFEGFEGD-amide     | 4.07 x 10^-4 | 7.3 | 1.8 x 10^1 |
| FGEFGFGEYGFEGFEGFEGD-amide     | 1.96 x 10^-4 | 3.80 | 3.78 x 10^0 |
| FGEFGFGEYGFEGFEGFEGD-amide     | 2.79 x 10^-4 | 35.3 | 1.27 x 10^1 |
| AcGFGEYGF-amide                 | 4.79 x 10^-4 | 33.1 | 6.91 x 10^0 |
| AcGFGEYGF-amide                 | 3.55 x 10^-4 | 16.9 | 4.77 x 10^0 |
| AcGFGEYGF-amide                 | 4.68 x 10^-4 | 31.8 | 6.79 x 10^0 |
| AcGFGEYGF-amide                 | 3.96 x 10^-4 | 71.8 | 1.81 x 10^1 |

a These values were reported in Barker et al. Essentially identical results were obtained when these experiments were repeated during the course of this study.

b These values are calculated from the data reported by Songyang et al. (1995). These workers used a GST-SH2-srcTK mutant protein. A molecular weight of 71,000 Da was used to calculate $k_{cat}$ and $k_{cat}/K_m$.

c This peptide displayed substrate inhibition at concentrations greater than 3 x $K_m$.

---

EXPERIMENTAL PROCEDURES

Materials—The peptides listed in Table I were purchased from Zencia (Wilmington, DE), with the exception of RRLIEDAeyaarrg, which was purchased from Bachem Biosciences. The peptides were dissolved in 50 mM HEPES, pH 7.5. Peptide stock concentrations were based upon the weight and peptide content. All other reagents were purchased from Sigma and used without further purification. Buffers were titrated to the appropriate pH prior to use with NaOH. Stock solutions of ATP, NADH, and phosphoenolpyruvate were titrated to pH 7.2 with NaOH or HCl. Pyruvate kinase and lactate dehydrogenase were dissolved at concentrations of 1-4 mg/ml in 50 mM HEPES, pH 7.5. The purification of srcTK from a baculovirus expression system will be published elsewhere, but was analogous to similar isolations (Ellis et al., 1994; Zhang et al., 1994; Saya et al., 1993; Buddle et al., 1993). Prior to use, the enzyme (10-30 µM, pH 7.5) was incubated with 1 mM ATP, 20 mM MgCl2 at 4°C for 20 min. This treatment results in autoactivation of srcTK and produces linear progress curves over the first 10% of substrate converted to product. The enzyme was then diluted 10-fold into 2 mg/ml bovine serum albumin, 1 mM ATP, and 20 mM MgCl2 to yield a final stock concentration of 2 µM. The final concentration of srcTK in the assay was 0.02 µM.

srcTK Assay—The phosphorylation of peptides and concomitant production of ADP was coupled to the oxidation of NADH using phosphoenolpyruvate, pyruvate kinase, and lactate dehydrogenase according to Barker et al. The reactions were monitored at 340 nm at 25°C for 10 min using a Cary-4 spectrophotometer (Varian Instruments). The rates were calculated from the linear progress curves according to Equation 1 using a linear least squares routine. Each reaction contained 100 mM HEPES, pH 7.5, 20 mM MgCl2, 100 µM ATP, 1 mM phosphoenolpyruvate, 100 µM dithiothreitol, 0.24 mM NADH, 44 µg/ml pyruvate kinase, and 64 µg/ml lactate dehydrogenase in a final volume of 1 ml. The initial rates as a function of substrate concentration. Data were fit to Equation 2 using GraFit (Leatherbarrow, 1992). The peptide substrate concentration was varied from 0.3 to 3 x $K_m$. The $K_m$ values for MgATP were determined at two concentrations of AcGEYGF(EF)2GD and (FGE)Y-amide by varying the nucleotide similarly.

$$ v = v_m[S] / (K_m + [S]) $$

Phosphorylation of peptides (FGE)Y(YEF)GD, Ac-FGE)YGE-amide, and Ac-FGEYGF-amide was confirmed by HPLC-electrospray ionization mass spectrometry. In a typical experiment, activated N-85-srcTK (0.125 µM) was incubated at room temperature with 1 mM Ac-FGEYGF-amide, 1 mM ATP, and 20 mM MgCl2 in 100 mM HEPES, pH 7.5. After 5 min, a 4-µl aliquot of the reaction mixture was added to 46 µl of 0.2% trifluoroacetic acid, and a 3-µl aliquot of this solution was injected onto a Poros R2H 800-µm x 10-cm perfusion column (LC Packings, San Francisco). The unreacted peptide and phosphopeptide product were eluted at 80 µl/min (Hewlett Packard-1090 microbore pump system) using a gradient consisting of 1% to 51% eluant B over 6 min (0.035% trifluoroacetic acid in 90:10 acetoni/trifluoroacetic acid; eluant B was 0.05% trifluoroacetic acid). The column eluant was monitored at 215 nm.

---

The dependence of srcTK kinetic parameters on the length of the peptide substrate

The standard errors from the fit of the data were ±20% of the calculated values. A molecular weight of 52,000, based upon the mass of N-85-srcTK obtained by electrospray ionization mass spectrometry, was used to calculate $k_{cat}$ and $k_{cat}/K_m$. The data were normalized for slight variations in enzyme specific activity based upon the activity versus (FGE)Y(YEF)GD.
with an Applied Biosystems Instruments model 789A UV-visible detector equipped with a capillary Z-flow cell (LC Packings) and a API-III triple quadrupole mass spectrometer equipped with an electrospray ion source (PE-Sciex, Thornhill, Ontario). The mass spectrometer was scanned from 300 to 1000 Da in 2.3 s using a 0.3-Aa steSD step and a 1-ms dwell time. Mass spectra were acquired using an orifice potential of 80 V and an ion multiplier voltage of -4000 V.

RESULTS AND DISCUSSION

Comparison of RRLIEDEAAYARG and (FGE)3Y(GEF)2GD As Substrates for srcTK—(FGE)3Y(GEF)2GD (peptide I) was reported to be a “good” substrate for srcTK.7 This sequence contains the Y + 1 to Y + 3 residues predicted by Songyang et al. (1995) to be optimal for srcTK. The srcTK phosphorylation of representative members of the series of peptides in Table I was confirmed by HPLC-electrospray ionization mass spectrometry. For example, the mass and HPLC retention times of AcFGEYGEF-amide were 889.4 Da and 5.52 min, respectively. Upon incubation with N-85-srcTK, a new species was evident that displayed a HPLC retention time and mass of 5.04 min and 969.5 Da, respectively. The 80.1-Da increase in mass and increase in polarity indicated by the decreased HPLC retention time are indicative of phosphorylation to produce AcFGE-pY-GEF-amide. The srcTK substrate activities of peptide I and RRLIEDEAAYARG are compared in Table I.8 Peptide I yielded a 251-fold greater catalytic efficiency, 38-fold lower Km, and a 6.6-fold increase in kcat than did RRLIEDEAAYARG.9 The activity of peptide I compares favorably to that reported for AEEEYGFEEAKKKK (peptide II), which contains the optimal sequence derived from peptide library work (Songyang et al., 1995). kcat and kcat/Km were 24-fold and 11-fold greater for peptide I than the values reported for peptide II. This reflects either absolute differences in the catalytic efficiency of the two substrates or differences in the specific activity of the mutant enzyme preparations. The reported Km for peptide II was 2-fold lower than the value obtained in this work for peptide I.

Phosphorylation of Truncated Versions of Peptide I by srcTK—To determine the minimum number of amino acids required for srcTK for activity, the kinetic parameters were determined for a series of peptide substrates with 1-residue truncations at each terminus. The results shown in Table I indicate that with respect to the N terminus, 8 N-terminal residues can be removed with only a 3-fold loss in catalytic efficiency relative to the activity of peptide I. Removal of the Y – 1 residue resulted in a significant additional loss of activity (9-fold), indicating that only a single N-terminal residue is sufficient for substrate activity. There was a slight increase in kcat as the N terminus was shortened. This trend could be the result of nonproductive binding of the longer peptide sequences, since the Km also increased similarly through AcFGEY(GEF)2GD. Nonproductive binding will lower both kcat and Km, by the factor 1 + Ks/Km, where Ks is the dissociation constant for the nonproductive binding mode but will not affect kcat/Km (Fersht, 1977). The repetitive nature of the sequence of peptide I is likely to lead to nonproductive binding modes that would preclude phosphorylation of the tyrosine. For example, recognition of the GEF or GGE sequence by the Y + 1 to Y + 3 binding subsites in the enzyme active site (Songyang et al., 1995) could lead to the placement of F in the Y binding site until the Y – 4 residue was removed. The requirement for a Y – 1 residue would preclude FGEYGEFGEF from binding in that mode. The potential for nonproductive binding modes with larger peptide sequences complicates the interpretation of the specificity requirements unless one consistently compares kcat/Km.

An absolute requirement for residues C-terminal to the tyrosine was not evident until the amino acid in the Y + 3 position was removed. Removal of this residue resulted in a 73-fold decrease in kcat/Km, relative to the value obtained with peptide I. This was primarily due to a 34-fold increase in Km.9 Further removal of the Y + 1 and Y + 2 residues resulted in only slight additional decreases in activity. There is not an obvious trend indicative of the relief of nonproductive binding modes upon truncating the C terminus. This may reflect the statistics of the effect of a single non-productive binding mode (GEF in Y + 3 to Y + 6) due to the C terminus of the peptide relative to two to three possible nonproductive modes due to the N-terminal sequence of these peptides.

Combining the results from the N- and C-terminal truncations into two peptides produced essentially additive results. AcFGEYGEF-amide had a similar catalytic efficiency to truncations on each terminus singularly. However, removal of the Y – 4 residue of AcFGEYGEF-amide yielded approximately a 2-fold increase in Km, and a 2-fold lower kcat/Km relative to the values obtained with the octamer. A similar trend in the Km was seen when AcFGEY(GEF)2GD is compared to AcFGEY(GEF)3GD. Further N-terminal truncation of the 7-mer to produce AcGEYGEF-amide had no effect on the kinetic parameters, although a similar truncation in the context of AcFGEY(GEF)3GD to AcEY(GEF)2GD resulted in a 2-fold decrease in the Km. These data demonstrate to a first approximation that substrate specificity can be explored by assuming that substrate recognition, N- and C-terminal relative to the site of phosphorylation, can be examined independently.

Songyang et al. (1995) suggested a preference for a hydrophobic residue, particularly an isoleucine in the Y – 1 position. AcFGEYGEF-amide and AcGFGEY-amide were chosen with the hope that the E to L substitution in the Y – 1 position might restore optimal activity. In the case of the 5-mer, this substitution resulted in a 2-fold increase in kcat and kcat/Km, but in the case of the 7-mer, the same substitution resulted in a 2-fold decrease in kcat. In the latter case, kcat/Km was only marginally affected. Comparison of these results and those obtained from the peptide library work suggests that the specificity for particular subsites are affected by the presence and/or identity of residues occupying other subsites N-terminal to the tyrosine. These results suggest that the equivalent of local minima may be obtained when determining subsite specificity using peptide libraries. In other words, the optimal sequence determined may depend upon the context and sequence length that the peptide library is based upon. Future work will examine the subsite specificity of srcTK using peptide libraries based upon AcFGEYGEF-amide and AdYGEF-amide to explore this possibility.

In conclusion, several short, highly active substrates for
srcTK were developed. The observation that only 5–7 residues are required for significant substrate activity suggests that small molecule inhibitors based upon interaction with the phosphoacceptor site may be developed.

Acknowledgments—We acknowledge M. Moyer of Glaxo Research Institute (GRI) for peptide amino acid analysis, R. Boerner for critically reading the manuscript, and the GRI Protein Engineering Group for preparation of the srcTK used in these studies.

REFERENCES

Budde, J. A., Ramdas, L. & Ke, S. (1993) Prep. Biochem. 23, 493–515
Casnellie, J. E., Harrison, M. L., Hellstrom, K. E. & Krebs, E. G. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 282–286
Cola, C., Brunati, A. M., Borin, G., Ruzza, P., Calderan, A., Castiglione, R. D. & Pinna, L. A. (1989) Biochim. Biophys. Acta 1012, 191–195
Ellis, B. E., De-Lacy, P., Weigl, D., Patel, I., Wisely, G. B., Lewis, K., Overton, L., Cadwell, S., Kost, T., Hoffman, C., Barrett, G., Knight, B., Edison, A., Huang, X., Kassel, D., Berman, J., Rodriquez, M. & Luther, M. (1994) J. Cell. Biochem. 18, Suppl. B, 276
Fersht, A. R. (1977) Enzyme Structure and Mechanism, pp. 109–111, Freeman, San Francisco

Garcia, P., Shoelson, S. E., George, S. T., Hinds, D. A., Goldberg, A. R., and Miller, W. T. (1993) J. Biol. Chem. 268, 25146–25151
Hunter, T. (1982) J. Biol. Chem. 257, 4843–4848

Leatherbarrow, R. J. (1992) GraFit Version 3.0, Erithacus Software Ltd., Staines, U. K.

Ogawa, W., Hosomi, Y., Shi, K. & Roth, R. (1994) J. Biol. Chem. 269, 29602–29608

Pimentel, E. (1989a) Oncogenes, Vol. 1, CRC Press, Boca Raton, FL

Pimentel, E. (1989b) Oncogenes, Vol. 2, CRC Press, Boca Raton, FL

Radziejewski, C., Miller, W. T., Mobashery, S., Goldberg, A. R. & Kaiser, E. T. (1989) Biochemistry 28, 9047–9052

Saya, H., Lee, P. S. Y., Nishi, T., Izawa, I., Nakajima, M., Gallick, G. E. & Levin, V. A. (1993) FEBS Lett. 327, 224–230

Songyang, Z., Carraway, K. L., Eck, M., Harrison, S. C., Feldman, R. A., Mohammad, M., Schisenger, J., Hubbard, S. R., Smith, D. P., Eng, C., Lorenzo, M. J., Ponder, B. A., Mayer, B. J. & Cantley, L. C. (1995) Nature 373, 536–539

Till, J. H., Annan, R. S., Carr, S. A. & Miller, W. T. (1994) J. Biol. Chem. 269, 7423–7428

Tinker, D. A., Krebs, E. G., Feltham, I. C., Attah-Poku, S. K. & Ananthanarayanan, V. S. (1988) J. Biol. Chem. 263, 5024–5026

Tinker, D. A., Carton, J.-L., McMurtry, J. S. & Levin, V. A. (1992) Anticancer Res. 12, 123–128

Wong, T. W. & Goldberg, A. R. (1983) J. Biol. Chem. 258, 1022–1025

Zhang, S., El-Genady, K., Abdel-Ghany, M., Clark, R., McCormick, F. & Racker, E. (1991) Cell. Physiol. Biochem. 1, 24–30