Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 973–977, February 1982
Biochemistry

Analysis of the sequence of amino acids surrounding sites of tyrosine phosphorylation

(Rous sarcoma virus/p60^{src}/avian sarcoma viruses Y73 and PRCII/Abelson virus)

TILO PATSCHINSKY*, TONY HUNTER*, FRED S. ESCH†, JONATHAN A. COOPER*, AND BARTHOLOMEW M. SEFTON*

*Tumor Virology Laboratory and †Laboratories of Endocrinology, The Salk Institute, P.O. Box 85800, San Diego, California 92138

Communicated by Robert W. Holley, October 13, 1981

ABSTRACT **We have identified the single phosphorylated tyrosine in p60^{src}, the transforming protein of Rous sarcoma virus, as part of the sequence**

NH_2-Arg-Leu-Ile-Glu-Asp-Asn-Glu-Tyr(*P*)-Thr-Ala-Arg-COOH.

Therefore, this is a sequence that is recognized efficiently by a tyrosine protein kinase *in vivo*. Phosphorylation of tyrosine in cellular proteins appears to play a role in malignant transformation by four classes of genetically distinct RNA tumor viruses. Phosphorylated tyrosines in several other proteins resemble the tyrosine in p60^{src} in that they are located 7 residues to the COOH-terminal side of a basic amino acid and either 4 residues to the COOH-terminal side of, or in close proximity to, a glutamic acid residue. Therefore, it is possible that these features play a role in the selection of sites of phosphorylation by some tyrosine protein kinases. However, several clear exceptions to this rule exist.

Rous sarcoma virus (RSV) transforms cells to a malignant state through the expression of a single gene, *src* (reviewed in ref. 1). This transforming gene encodes a single 60,000-dalton phosphoprotein, p60^{src} (2). p60^{src} is a protein kinase with the unusual specificity of phosphorylating tyrosine residues in its target proteins (3–6). This unscheduled phosphorylation of tyrosine almost certainly plays a role in cellular transformation by this virus (7) and probably also by three other groups of RNA tumor viruses genetically distinct from RSV (8–10). Tyrosine protein kinases may also play an important role in normal metabolism because they are implicated in the cellular response to epidermal growth factor (11).

An important question is how tyrosine protein kinases select and recognize their substrates. In the case of the well-characterized protein kinases that phosphorylate serine and threonine, the amino acid sequence of the substrate plays a crucial role in selection of the phosphorylation site. The cAMP-dependent protein kinase has a strong preference for serines that are located 2–5 residues to the COOH-terminal side of one or two basic amino acids, most commonly arginine (12, 13). In contrast, the casein protein kinases as a class show a strong preference for sites in the vicinity of acidic residues (14, 15). In casein itself, all but one of the sites of phosphorylation are located 2 residues to the NH_2-terminal side of an acidic amino acid (14).

Because of these precedents, we have characterized the sequence of amino acids at sites of tyrosine phosphorylation. We deduced the complete amino acid sequence surrounding the single phosphorylated tyrosine in p60^{src} and analyzed indirectly the sequence at sites of tyrosine phosphorylation in a number of other proteins.

MATERIALS AND METHODS

Cells and Viruses. Preparation and infection of chicken fibroblasts has been described (16). The Schmidt–Ruppin strain of RSV of subgroup A (SR-RSV-A) originated in the laboratory of H. Hanafusa, The Rockefeller University. Avian sarcoma virus Y73 was obtained from K. Toyoshima, the University of Tokyo, and the avian sarcoma virus PRCII was a gift from J. Neil, the University of Southern California. The ANN-1 line of NIH/3T3 cells nonproductively infected with Abelson virus (17) was the source of p120. Mink lung cells nonproductively transformed by the Snyder–Theilin strain of feline sarcoma virus (ST-FeSV) (18) were provided by M. Barbacid, the National Cancer Institute.

Radioactive Labeling. Labeling of infected chicken cells and of the ANN-1 cells with [^{32}P]orthophosphate (ICN, carrier-free) was performed as described (19). RSV-infected chicken cells, growing on 100-mm plastic Petri dishes, were labeled with individual radioactive amino acids overnight at 41°C in Dulbecco's modified Eagle's medium containing only 5% of the normal concentration of the labeled amino acid and supplemented with 4% complete calf serum. Each sample was from a single culture. Labeling was in 5 ml of medium containing one of the following: 2 mCi of [^{35}S]cysteine (470 Ci/mmol; 1 Ci = 3.7×10^{10} becquerels); 10 mCi of [^{3}H]isoleucine (84 Ci/mmol); 3 mCi of [^{35}S]methionine (1000 Ci/mmol); 5 mCi of [^{3}H]tryptophan (26 Ci/mmol); 5 mCi of [^{3}H]tyrosine (53 Ci/mmol). [^{35}S]Cysteine and [^{3}H]tyrosine were from New England Nuclear; all others were from Amersham/Searle.

Cell Lysis and Immunoprecipitation. The transforming proteins of RSV, Y73 virus, PRCII virus, ST-FeSV, and Abelson virus were isolated by immunoprecipitation from transformed cells as described (4, 16, 19–21).

Gel Electrophoresis, Elution of Proteins, and Tryptic Digestion. Immunoprecipitated proteins were purified on preparative $NaDodSO_4$/polyacrylamide gels (22) and localized by autoradiography. The position of ^{3}H-labeled proteins was identified by ^{32}P- or ^{35}S- labeled markers that were run in parallel. All other procedures were as described (22).

Peptide Mapping. Two-dimensional mapping was done by using electrophoresis at pH 8.9 in the first dimension and ascending chromatography in the second, as described (4). ^{32}P-Labeled peptides were detected by using preflashed film and an intensifying screen. ^{3}H- and ^{35}S-labeled peptides were visualized by fluorography. Analysis of mixtures of ^{32}P-labeled and

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

Abbreviations: RSV, Rous sarcoma virus; SR-RSV-A, Schmidt–Ruppin RSV of subgroup A; Tyr(*P*), phosphotyrosine; ST-FeSV, Snyder–Theilin feline sarcoma virus.

^{3}H- or ^{35}S-labeled peptides involved the use of just enough ^{32}P-labeled material to give an image after 3–6 days of exposure with an intensifying screen. The thin-layer plates were then dipped in molten 2-methylnaphthalene containing 0.4% diphenyloxazole and re-exposed without an intensifying screen.

***Staphylococcus aureus* V8 Protease Digestion.** Whole tryptic digests or isolated tryptic peptides were digested with 5–10 μg of *S. aureus* protease (Miles) at pH 4.0 or pH 7.8 as described (19).

Amino Acid Sequence Analysis. Automated Edman degradation was performed as described (23). The residues from each step were dried and assayed directly by using Cerenkov radiation. In some cases, the anilinothiazolinone derivatives were converted into phenylthiohydantoin derivatives and separated by electrophoresis at pH 3.5 in the presence of phenylthiohydantoin-phosphotyrosine [Tyr(*P*)] marker to verify that the released radioactivity was indeed associated with tyrosine.

***In Vitro* Kinase Assay.** This assay was performed as described (20) except that the buffer used contained 0.01 M sodium 1,4-piperazinebis(ethanesulfonic acid) at pH 7.0 and 0.01 M $MnCl_2$.

Isolation of p36 and p50. p50 was isolated from preparative gels containing immunoprecipitated p60^{src} as described (4). p36 was purified by column chromatography and gel electrophoresis as described (24).

Analysis of the Synthetic Peptide. A peptide with the sequence Leu-Ile-Glu-Asp-Asn-Glu-Tyr-Thr-Ala-Arg was synthesized and purified (unpublished data). The peptide was phosphorylated *in vitro* by incubation with immunoprecipitates containing p90 of Y73 virus as described above. The phosphorylated peptide was purified from supernatant of the kinase reaction by electrophoresis at pH 3.5.

RESULTS AND DISCUSSION

Identification of the Phosphorylated Tyrosine in p60^{src}. The Tyr(*P*) in p60^{src} is contained in a single tryptic peptide that is located in the COOH-terminal 26,000 daltons of the polypeptide (4, 25). Czernilofsky *et al.* (26) deduced a complete amino acid sequence for the p60^{src} of SR-RSV-A. In this sequence of 530 amino acids, each of the potential tyrosine-containing tryptic peptides in the COOH-terminal half of the molecule has a unique amino acid composition. Determination of the amino acid composition of the phosphorylated tryptic peptide should therefore allow unambiguous identification of the tyrosine that is phosphorylated.

Because only trace amounts of this peptide can be obtained, chemical determination of its amino acid composition was not attempted. Instead, p60^{src} was labeled biosynthetically with one of the following amino acids: [^{35}S]methionine, [^{35}S]cysteine, [^{3}H]isoleucine, [^{3}H]histidine, [^{3}H]glycine, [^{3}H]tryptophan, or [^{3}H]tyrosine. p60^{src} was then isolated by immunoprecipitation, purified by $NaDodSO_4$/polyacrylamide gel electrophoresis, subjected to digestion with trypsin, and analyzed by two-dimensional peptide mapping (Fig. 1). We detected two phosphorylated tryptic peptides in p60^{src} of SR-RSV-A (4). Peptide α contains only phosphoserine and peptide β, only Tyr(*P*) (4). Peptide β was labeled with [^{3}H]tyrosine, as expected, and with [^{3}H]isoleucine (Fig. 1). It was not labeled detectably with [^{35}S]methionine, [^{35}S]cysteine, [^{3}H]glycine, [^{3}H]histidine, or [^{3}H]tryptophan.

Additionally, we knew from earlier work (19) that peptide β was susceptible to digestion with the protease from *S. aureus* V8 under conditions in which cleavage of glutamyl bonds was favored (27). This strongly suggested that peptide β also contained glutamic acid; however, we have not tested this by biosynthetic labeling. Only one tryptic peptide in the deduced sequence of the COOH-terminal half of p60^{src} contained isoleucine and tyrosine and was devoid of methionine, cysteine, glycine, tryptophan, and histidine—the tryptic peptide containing residues 413–422. This tryptic peptide contained only a single tyrosine, tyrosine-419 (Fig. 2).

This conclusion was also supported by automated Edman degradation of peptide β. Phenylthiohydantoin-Tyr(*P*) was released at the seventh cycle of degradation of the tryptic peptide and at the first cycle of degradation of the product of secondary digestion of the peptide with *S. aureus* protease (Table 1). These results are consistent with the positions of both arginine-412 and glutamic acid-418 (Fig. 2). It should be noted that the predicted sequence of peptide β contains two glutamic acids, at positions 415 and 418. However, complete digestion of peptide β with *S. aureus* protease yielded only one phosphorylated product—that resulting from cleavage at glutamic acid-418. Smart *et al.* (30) also concluded that tyrosine-419 is phosphorylated in p60^{src}.

The approach we used to identify tyrosine-419 as the site of tyrosine phosphorylation in p60^{src} is indirect and requires that the deduced sequence of the protein is correct. However, additional evidence substantiates this conclusion. A synthetic peptide, identical in sequence to the deduced sequence of peptide β, can be phosphorylated on tyrosine *in vitro* by a protein kinase present in immunoprecipitates containing either p60^{src} or the putative transforming protein of Y73 virus (unpublished data). This phosphorylated synthetic peptide comigrated with biosynthetically labeled peptide β from p60^{src} when analyzed by peptide mapping (Fig. 1). Additionally, the phosphorylated fragment of this synthetic peptide, generated by digestion with *S. aureus* protease, comigrated precisely with the *S. aureus* protease cleavage fragment of biosynthetically phosphorylated peptide β when analyzed by peptide mapping (data not shown). Therefore, there seems to be little doubt that the sequence of the synthetic peptide corresponds to the sequence of amino acids surrounding the tyrosine in p60^{src} that is phosphorylated *in vivo*.

Analysis of Other Sites of Tyrosine Phosphorylation. Peptide β of p60^{src} is apparently identical in sequence to the Tyr(*P*)-containing tryptic peptide II of p90 (19, 31), the putative transforming protein of Y73 virus (10). Because the p60^{src} and p90 polypeptides appear otherwise largely unrelated (19), the identity of these two decapeptides suggested that the primary sequence of the substrate might be important in the interaction of tyrosine protein kinases with their substrates. Even though only 1 of the more than 10 Tyr(*P*)-containing tryptic peptides we analyzed, peptide II of p90, is identical to peptide β of p60^{src}, the possibility existed that these other sites might resemble peptide β with respect to the placement of basic residues or the presence of glutamic acid. Again, because only miniscule amounts of these proteins can be obtained, the sequence surrounding the Tyr(*P*) residues in these proteins could not be determined directly. Instead, we deduced certain features of such sequences by performing automated Edman degradation on ^{32}P-labeled peptides generated by proteolysis. For example, analysis of a tryptic peptide tells the number of residues between the Tyr(*P*) and the nearest lysine or arginine on the NH_2-terminal side. Likewise, the location of the closest glutamic acid residue was determined by sequence analysis of peptides resulting from secondary digestion of ^{32}P-labeled tryptic peptides with *S. aureus* protease under conditions that strongly favor cleavage of glutamyl bonds (27).

We first examined other viral transforming proteins. PRCII virus, ST-FeSV, and Abelson virus each encodes a single Tyr(*P*)-containing polypeptide, p105 (32), p85 (33), and p120 (34, 35), respectively. Although PRCII virus and ST-FeSV clearly orig-

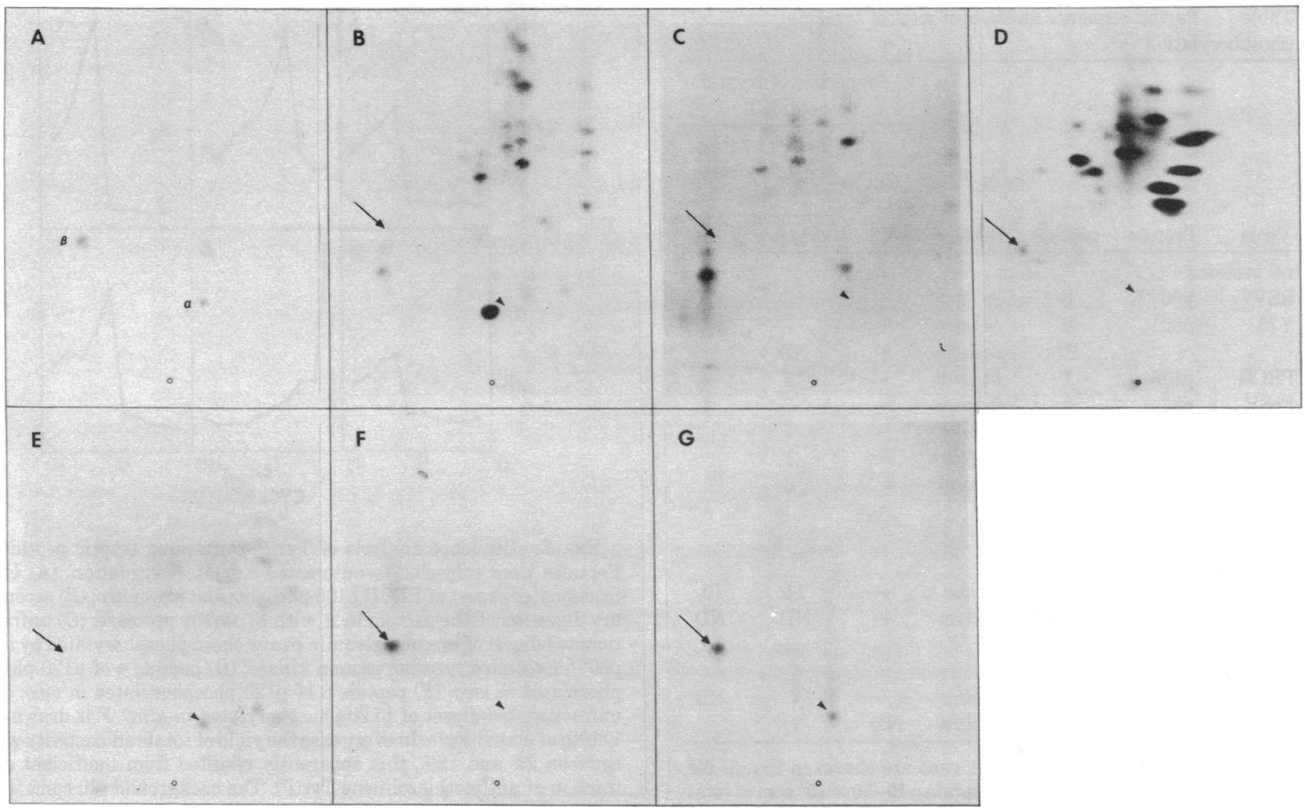

FIG. 1. Amino acid composition of the Tyr(*P*)-containing tryptic peptide of $p60^{src}$ of SR-RSV-A. Tryptic digests of $p60^{src}$ were labeled with [^{32}P]orthophosphate or individual amino acids and analyzed. *Src* IV, a synthetic peptide corresponding in sequence to that deduced for peptide β, was phosphorylated *in vitro* by the tyrosine protein kinase associated with p90 of Y73 virus. The source of the digests, the amount of radioactivity loaded, and the exposure times were as follows: (*A*) $p60^{src}$, ^{32}P, 1200 cpm, 16 hr; (*B*) $p60^{src}$, [^{35}S]methionine, 10,400 cpm, 5 days; (*C*) $p60^{src}$, [^{35}S]cysteine, 14,000 cpm, 3 days; (*D*) $p60^{src}$, [^{3}H]isoleucine, 48,200 cpm, 51 days; (*E*) $p60^{src}$, [^{3}H]glycine, 5200 cpm, 46 days; (*F*) *src* IV peptide, ^{32}P, 160 cpm, 3 days; (*G*) mixture of *src* IV, ^{32}P, 80 cpm, and $p60^{src}$, ^{32}P, 260 cpm, 5 days. Arrows and arrowheads indicate the positions of peptides β and α, respectively, determined from maps of mixtures run in parallel. The origin is marked o.

inated independently, their transforming proteins are structurally related (21). However, p120 is unrelated to p105 or p85, and none of these proteins has homology with $p60^{src}$ or p90 of Y73 virus (19, 21). All three proteins undergo "autophosphorylation" *in vitro* (8, 33, 36). In the case of PRCII virus and ST-FeSV, phosphorylation occurs on a tyrosine that is normally phosphorylated *in vivo* (29, 36). Such is not true for p120 of Abelson virus; as many as 10 new sites become phosphorylated on tyrosine *in vitro* with little, if any, labeling of the 2 sites phosphorylated *in vivo* (28).

Sequence analysis revealed that the Tyr(*P*) in p105 of PRCII virus phosphorylated *in vitro* was 7 residues to the COOH-terminal side of a basic amino acid (Table 1; Fig. 3). The Tyr(*P*)-containing peptide of p105 was cleaved by *S. aureus* protease to yield a peptide in which the Tyr(*P*) was now 4 residues from the NH_2 terminus (Table 1; Fig. 3). The Tyr(*P*) in p85 of ST-FeSV was 7 residues from a lysine or arginine and 5 residues from a glutamic acid (Table 1). Thus the Tyr(*P*) residues in p105

FIG. 2. Sequence of amino acids surrounding the phosphorylated tyrosine in $p60^{src}$. The amino acid sequence and the numbering of the residues are from Czernilofsky *et al.* (26). Solid arrows, sites of cleavage with trypsin; open arrow, identified site of cleavage of tryptic peptide β with *S. aureus* protease.

and p85 both resemble those in $p60^{src}$ and in p90 of Y73 virus in being 7 residues to the COOH-terminal side of a tryptic cleavage site. However, the placement of glutamic acid residues in the neighborhood of the Tyr(*P*) was only partly homologous to that in $p60^{src}$ and p90. In p105, as in $p60^{src}$, there is a glutamic acid 4 residues upstream of the Tyr(*P*), but digestion with *S. aureus* did not reveal a glutamic acid adjacent to the Tyr(*P*). In p85 the nearest detectable glutamic acid was 5 residues away from the Tyr(*P*).

The Ig heavy chain of rabbit anti-RSV tumor serum is phosphorylated exclusively on tyrosine by the $p60^{src}$-associated protein kinase activity *in vitro* (4). Sequence analysis of an unfractionated tryptic digest of heavy chains phosphorylated *in vitro* revealed that essentially all of the phosphorylated tyrosines were located 7 residues at the COOH-terminal side of a basic amino acid (Table 1; Fig. 3). This is striking in view of the fact that the tryptic digest subjected to sequence analysis contained a heterogeneous mixture of tryptic peptides (data not shown). The heavy chain of rabbit anti-RSV tumor serum phosphorylated *in vitro* with immunoprecipitates containing p90 of Y73 virus (10) gave a similar result (Table 1). Despite this similarity between sites of phosphorylation in the viral polypeptides and those in the heavy chains, the large majority of the tryptic phosphopeptides from the heavy chain could not be digested with *S. aureus* protease (Table 1).

The Tyr(*P*) in peptide 4 (28) of p120 of Abelson virus, phosphorylated *in vivo*, was also located 7 residues to the COOH-terminal side of a basic amino acid (Table 1; Fig. 3). However, the homology with peptide β was again only partial. This tryptic

Table 1. Partial sequence analysis of sites of tyrosine phosphorylation

Virus	Protein	Tryptic peptide	Labeled	Properties of tryptic peptide: Cut by V8	Tyr(P) location, no. of residues from Lys/Arg	Tyr(P) location, no. of residues from Glu
Viral proteins						
RSV	$p60^{src}$	β	*In vivo*	+	7	1,4*
Y73	p90	II	*In vitro*	+	7	1,4*
		III	*In vivo*	−	ND	
PRCII	p105	I	*In vitro*	+	7	4
FeSV	p85	c	*In vitro*	+	7	5
A-MLV	p120	4	*In vivo*	−	7	
		8	*In vivo*	−	1	
		Total digest	*In vitro*	−	6	
Cellular proteins						
RSV	p36	A	*In vivo*	+	14	10
RSV	p50	I	*In vivo*	+	ND	ND
RSV	Ig heavy chain	Total digest	*In vitro*	+/−	7	ND
Y73	Ig heavy chain	Total digest	*In vitro*	ND	7	

Typical examples of the sequenator runs are shown in Fig. 3. We analyzed 3–10 × 10^3 cpm of purified peptides; 10–50 × 10^3 cpm of total tryptic digests were used. Peptide II of p90, peptide I of p105, and peptide c of p85 are phosphorylated both *in vivo* and during "autophosphorylation" *in vitro*. The samples subjected to sequence analysis were labeled *in vitro*. Unless indicated, all other samples were labeled biosynthetically. Direct sequence analysis of the limit digests of peptide β of $p60^{src}$ and peptide II of p90 with *S. aureus* protease placed the Tyr(P) next to a glutamic acid. The deduced sequence of these peptides (Fig. 2) reveals an extra glutamic acid 4 residues to the NH_2-terminal side of the Tyr(P). Because this was not detected by sequence analysis it is designated 4*. FeSV and A-MLV, ST-FeSV and Abelson virus, respectively; ND, not determined; +, sensitivity to secondary digestion; −, no cleavage with *S. aureus* protease (V8). Peptides are designated as described (4, 19, 28, 29).

peptide was not sensitive to digestion with *S. aureus* protease (Table 1). The four other sites of tyrosine phosphorylation in viral polypeptides which we obtained in sufficient quantity to study had no obvious homology with peptide β. Peptide 8 of p120 phosphorylated *in vivo* (28), three of the major Tyr(P)-containing tryptic peptides from p120 phosphorylated *in vitro* in the immune complex (28), and peptide III, the second Tyr(P)-containing tryptic peptide from biosynthetically labeled p90 of Y73 virus (19), were all insensitive to secondary digestion with *S. aureus* protease (Table 1). Additionally, the phosphorylated tyrosine in peptide 8 of p120 phosphorylated *in vivo* adjoined directly, on the COOH-terminal side, a basic amino acid. Most of the label from an unfractionated tryptic digest of p120 phosphorylated *in vitro* (a mixture of as many as 10 peptides) was released at cycle 6 during sequence analysis (Table 1; Fig. 3). Peptide III of p90 was not subjected to sequence analysis.

At least seven cellular proteins have been found to contain elevated amounts of Tyr(P) in RSV-transformed cells (37). We have studied two of these, p36, the 36,000-dalton protein first described by Radke and Martin (38), and p50, the 50,000-dalton protein found complexed with $p60^{src}$ (4, 16). The Tyr(P) in the major tryptic phosphopeptide of p36 was located 14 residues to the COOH-terminal side of a tryptic cleavage site (Table 1). This peptide was sensitive to secondary digestion with *S. aureus* protease and thus appears to contain glutamic acid. Sequence

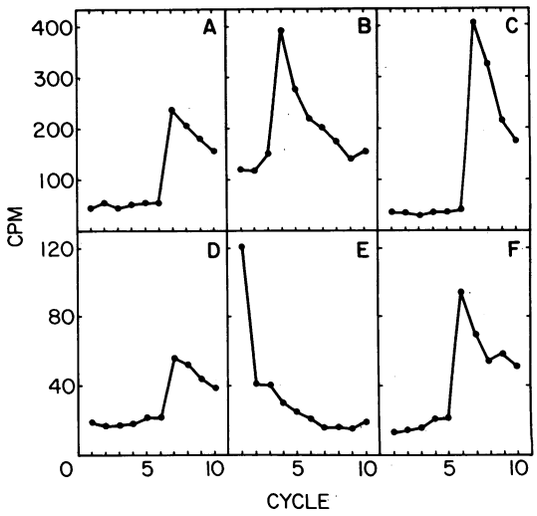

FIG. 3. Sequence analysis of Tyr(P)-containing tryptic peptides. Peptides were subjected to automated Edman degradation. (*A*) Unfractionated digest of PRCII p105 phosphorylated *in vitro*; (*B*) secondary digestion of the sample in *A* with *S. aureus* protease; (*C*) unfractionated digest of immunoglobulin heavy chain phosphorylated by the $p60^{src}$-associated tyrosine protein kinase; (*D*) peptide 4 of p120 phosphorylated *in vivo*; (*E*) peptide 8 of p120 phosphorylated *in vivo*; (*F*) unfractionated digest of p120 phosphorylated *in vitro*. *F* is drawn to 1/10th of actual scale. In every case the yield of total radioactivity was between 2% and 13%; this apparently resulted from inefficient extraction of anilinothiazolinone Tyr(P). The background (20 cpm) was not subtracted.

data suggest that the nearest glutamic acid is located 10 residues away from the Tyr(P). Thus at this level of analysis the tyrosine phosphorylation site in p36 does not bear striking homology with that in $p60^{src}$. The Tyr(P)-containing peptide of p50 was also susceptible to digestion with *S. aureus* protease, but insufficient quantities of the protein were obtained to locate this site by sequence analysis.

CONCLUSIONS

Does the primary amino acid sequence at the site of tyrosine phosphorylation play an important role in the recognition of this site by a tyrosine protein kinase? Four viral transforming proteins—$p60^{src}$, p90 of Y73 virus, p105 of PRCII virus, and p85 of ST-FeSV—have similar sequences on the NH_2-terminal side of the phosphorylated tyrosine, with those in $p60^{src}$ and p90 being identical (Fig. 2). All four sites share the placement of a lysine or arginine 7 residues to the NH_2-terminal side of the phosphorylated tyrosine. In addition, all four sites have at least one acidic amino acid among the residues between the Tyr(P) and the lysine or arginine on the NH_2-terminal side. $p60^{src}$, p90, and p105 have a glutamic acid 4 residues away, and $p60^{src}$ and p90 have an additional glutamic acid adjacent to the Tyr(P). p85 does not appear to have a glutamic acid at either of these sites but has at least one glutamic acid 5 residues to the NH_2-terminal side of the Tyr(P). Aspartic acids are found 3 residues to the NH_2-terminal side of the Tyr(P) in $p60^{src}$ and p90 and also 3 residues to the NH_2-terminal side of the presumptive site of tyrosine phosphorylation in p85 (Arg-Glu-Glu-Ala-Asp-Gly-Val-Tyr; A. Hampe, F. Galibert, and C. J. Sherr, personal communication). These homologies suggest that a basic amino acid 7 residues to the NH_2-terminal side of the target tyrosine together with one or more acidic residues between 1 and 6 residues away may be important in the recognition of the tyrosine in these substrates by the relevant tyrosine protein kinases.

However, some sites of tyrosine phosphorylation appear not to have both these characteristic features. It is striking that essentially all of the sites in the Ig heavy chain phosphorylated *in vitro* by either $p60^{src}$ or p90 of Y73 virus and the phosphorylated tyrosine in peptide 4 of p120 of Abelson virus phosphorylated *in vivo* are also 7 residues to the COOH-terminal side of a basic amino acid. Nevertheless, neither peptide 4 of p120 nor the large majority of the phosphorylated tryptic peptides from the heavy chain are sensitive to digestion with *S. aureus* protease. Additionally, the Tyr(*P*) in p36, which appears to be a major *in vivo* substrate of $p60^{src}$ (38), is at least 10 residues to the COOH-terminal side of sites of cleavage by either trypsin or *S. aureus* protease. Finally, at least three sites in p120 of Abelson virus that are phosphorylated on tyrosine either *in vivo* or *in vitro* do not obviously resemble the others studied here with regard to placement of sites of cleavage by either *S. aureus* protease or trypsin.

The possibility exists that these exceptional sites are more homologous with the sites in the viral transforming proteins than they appear. They could contain arginines, lysines, or glutamic acids that are positioned appropriately but are present in linkages that are resistant to digestion by trypsin or *S. aureus* protease. Additionally, a requirement for basic or acidic amino acids could be fulfilled by histidine or aspartic acid, respectively, neither of which would have been detected by the techniques used here. The presence of aspartic acid in the phosphorylation sites of viral transforming proteins clearly makes examination of other sites for this residue desirable.

It must be emphasized that local primary sequence is unlikely to be the sole basis on which substrates of tyrosine protein kinases are selected. Regions of the protein not in the immediate vicinity of the target tyrosine may play a role in substrate recognition. In addition, there is reason to believe that the intracellular localization of the kinases and their potential substrates plays an additional and important role in substrate selection.

The identities of the protein kinases phosphorylating the sites characterized here are uncertain. If tyrosine protein kinases with different specificities exist, this might account for some of the sequence variability at these sites. In this context the pronounced sequence homology of the sites of tyrosine phosphorylation in $p60^{src}$, p90, p105, and p85 warrants further discussion. All four viral proteins undergo an apparent autophosphorylation *in vitro*. If this reaction is indeed autophosphorylation, the homology detected at these sites would suggest that the sequence specificity and perhaps the protein substrate specificity of these four viral protein kinases *in vivo* are very similar or identical. Alternatively, the homology may suggest that each viral protein is a substrate of, and therefore possibly regulated by, the same cellular protein kinase.

Note Added in Proof. The epidermal growth factor receptor contains a single major site of tyrosine phosphorylation *in vivo*. We have found that the Tyr(*P*) in this site has a lysine or arginine 9 residues and glutamic acids 1 and 4 residues to the NH_2-terminal side. This site therefore has homology with the sites in the viral transforming proteins.

We thank John Smart and Jim Neil for discussion of their work, Jean Rivier for synthesis of the peptide, and Gernot Walter for reading the manuscript. This work was supported by Grants CA 14195, CA 17096, and CA 17289 from the National Institutes of Health. T.P. was supported by a fellowship from the Deutsche Forschungsgemeinschaft.

1. Hanafusa, H. (1977) in *Comprehensive Virology*, eds. Fraenkel-Conrat, H. & Wagner, R. R. (Plenum, New York), Vol. 10, pp. 401–483.
2. Purchio, A. F., Erikson, E., Brugge, J. S. & Erikson, R. L. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1567–1571.
3. Erikson, R. L., Collett, M. S., Erikson, E. L. & Purchio, A. F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6260–6264.
4. Hunter, T. & Sefton, B. M (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1311–1315.
5. Collett, M. S., Purchio, A. F. & Erikson, R. L. (1980) *Nature (London)* **285**, 167–169.
6. Levinson, A. D., Oppermann, H., Varmus, H. E. & Bishop, J. M. (1980) *J. Biol. Chem.* **255**, 11973–11980.
7. Sefton, B. M., Hunter, T., Beemon, K. & Eckhart, W. (1980) *Cell* **20**, 807–816.
8. Witte, O. N., Dasgupta, A. & Baltimore, D. (1980) *Nature (London)* **283**, 826–831.
9. Feldman, R. A., Hanafusa, T. & Hanafusa, H. (1980) *Cell* **21**, 757–766.
10. Kawai, S., Yoshida, M., Segawa, K., Sugiyama, H., Ishizaki, R. & Toyoshima, K. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 6199–6203.
11. Ushiro, H. & Cohen, S. (1980) *J. Biol. Chem.* **255**, 8363–8365.
12. Kemp, B. E., Bylund, D. B., Huang, T.-S. & Krebs, E. G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3448–3452.
13. Kemp, B. E., Graves, D. J., Benjamini, E. & Krebs, E. G. (1977) *J. Biol. Chem.* **252**, 4888–4894.
14. Mercier, J.-C., Grosclavde, G. & Ribadeau-Dumas, B. (1971) *Eur. J. Biochem.* **23**, 41–51.
15. Tuazon, P. T., Bingham, E. W. & Traugh, J. A. (1979) *Eur. J. Biochem.* **94**, 497–504.
16. Sefton, B. M., Beemon, K. & Hunter, T. (1978) *J. Virol.* **28**, 957–971.
17. Scher, C. D. & Siegler, R. (1975) *Nature (London)* **253**, 729–731.
18. Porzig, K. J., Barbacid, M. & Aaronson, S. A. (1979) *Virology* **92**, 91–107.
19. Patschinsky, T. & Sefton, B. M. (1981) *J. Virol.* **39**, 104–114.
20. Sefton, B. M., Hunter, T. & Beemon, K. (1980) *J. Virol.* **33**, 220–229.
21. Beemon, K. (1981) *Cell* **24**, 145–153.
22. Beemon, K. & Hunter, T. (1978) *J. Virol.* **28**, 551–566.
23. Welch, W. J., Sefton, B. M. & Esch, F. S. (1981) *J. Virol.* **38**, 968–972.
24. Cooper, J. A. & Hunter, T. (1981) *Mol. Cell. Biol.* **1**, 394–407.
25. Collett, M. S., Erikson, E. & Erikson, R. L. (1979) *J. Virol.* **29**, 770–781.
26. Czernilofsky, A. P., Levinson, A. D., Varmus, H. E., Bishop, J. M., Tischer, E. & Goodman, H. M. (1980) *Nature (London)* **287**, 198–203.
27. Houmard, J. & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 3506–3509.
28. Sefton, B. M., Hunter, T. & Raschke, W. C. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1552–1556.
29. Blomberg, J., van de Ven, W. J. M., Reynolds, F. H., Nalewaik, R. P. & Stephenson, J. R. (1981) *J. Virol.* **38**, 886–894.
30. Smart, J. E., Oppermann, H., Czernilofsky, A. P., Purchio, A. F., Erikson, R. L. & Bishop, J. M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6013–6017.
31. Neil, J. C., Ghysdael, J., Vogt, P. K. & Smart, J. (1981) *Nature (London)* **291**, 675–677.
32. Neil, J. C., Breitman, M. L. & Vogt, P. K. (1980) *Virology* **108**, 98–110.
33. Barbacid, M., Beemon, K. & Devare, S. G. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5158–5162.
34. Witte, O. N., Rosenberg, N., Paskind, M., Shields, A. & Baltimore, D. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2488–2492.
35. Reynolds, F. H., Sacks, T. L., Deobagkar, D. N. & Stephenson, J. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3974–3978.
36. Neil, J. C., Ghysdael, J. & Vogt, P. K. (1981) *Virology* **109**, 223–228.
37. Cooper, J. A. & Hunter, T. (1981) *Mol. Cell. Biol.* **1**, 165–178.
38. Radke, K. & Martin, G. S. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5212–5216.