Tight-binding inhibitory sequences against pp60^{c-src} identified using a random 15-amino-acid peptide library

Toru Nishi[a,b], Raymond J.A. Budde[a], John S. McMurray[a], Nihal U. Obeyesekere[a], Naueen Safdar[a], Victor A. Levin[a], Hideyuki Saya[a,c,*]

[a]*Department of Neuro-Oncology, The University of Texas, M.D. Anderson Cancer Center, Houston, TX 77030, USA*
[b]*Department Neurosurgery, Kumamoto University School of Medicine, Kumamoto 860, Japan*
[c]*Department of Tumor Genetics and Biology, Kumamoto University School of Medicine, Kumamoto 860, Japan*

Received 25 September 1996; revised version received 28 October 1996

Abstract A bacteriophage peptide library containing a random 15-amino-acid insert was screened for identification of peptide sequence(s) that bind pp60^{c-src}. Sequencing the random insert from more than 100 virions indicated that more than 60% of the phage virions that bound to this enzyme contained a GXXG sequence motif in which X was frequently a hydrophobic residue. The GXXG sequence was often repeated as GXXGXXG. Two nonameric peptides were synthesized to determine whether or not the peptide inhibits pp60^{c-src} tyrosine kinase activity and the importance of the glycine residues within this sequence. The peptide containing glycine had a K_i of 24 μM, whereas replacing the glycines with proline increased the K_i value to 3.1 mM.

Key words: Random phage library; Peptide inhibitor; pp60^{c-src}

1. Introduction

Protein phosphorylation plays a central role in the signal-transduction pathways regulating cellular activity. Elevated protein tyrosine phosphorylation level correlates with cellular proliferation and growth. pp60^{c-src} is a protein tyrosine kinase which couples diverse signals mediated through membrane receptors. In many types of cancers including neuroectodermal, breast and colon tumors, the activity of pp60^{c-src} is significantly activated [1–5]. The use of anti-sense oligonucleotides to reduce the expression and activity of pp60^{c-src} inhibits the growth of malignant cells. Hence, developing specific inhibitors to PTKs such as pp60^{c-src} will have a significant impact on medical research associated with proliferative disorders.

The search for 'lead compounds' from which to develop specific enzyme inhibitors is generally performed by randomly screening natural products or synthetic compounds. However, this method is time consuming, expensive and labor intensive. Since the natural substrate of pp60^{c-src} is supposed to be a cellular protein, a tightly binding peptide provides a rational starting point from which to design an inhibitor against pp60^{c-src}. However, the number of possible sequences within a peptide sequence of length n, containing the 20 naturally occurring amino acids, is equivalent to 20 to the nth power, a number which cannot be approached by standard peptide synthesis techniques. Phage peptide libraries have been developed by several groups [6–8]. A phage peptide library is a collection of phage virions in which each virion expresses a different peptide sequence on its surface as part of its coat protein. This library can be produced by ligating a mixture of synthetic random degenerate oligonucleotides that encode random sequences of amino acids, into the phage pIII gene. Expression of peptides as random sequences within the pIII coat protein allows their physical interaction with other molecules. Therefore, the peptide sequences that bind to a target protein can be screened by affinity purifying of protein-bound phage from the library, propagating the selected phage in *E. coli*, and determining the amino acid sequence of the displayed peptide by DNA sequencing the random insert of each clone.

In this study, we constructed and characterized a library with a random 15-amino-acid sequence with which to identify a tightly binding peptide sequence(s) to pp60^{c-src}. Screening showed that more than 60% of the clones have GXXG sequences within the 15-amino-acid insert. X was frequently a hydrophobic residue (F, V or L). This motif was often repeated as GXXGXXG. Based on these results we synthesized a model peptide to determine whether or not this motif would generate a tight binding peptide. The peptides were tested as inhibitors of the kinase activity of pp60^{c-src}. The model peptide containing glycine (FVGFLGFLG) had a K_i of 24 μM, whereas the control peptide (FVPFLPFLP), in which the glycines were replaced with proline, had a K_i of 3.1 mM. This finding demonstrates that the glycine within this sequence is important for the binding of this peptide to pp60^{c-src}.

2. Materials and methods

2.1. Reagents

T4 DNA ligase was purchased from Boehringer Mannheim and Sequenase 2.0 was from US Biochemical. Oligonucleotides were synthesized using an Applied Biosystems 392 DNA/RNA synthesizer and purified on PD-10 columns. Bacteriophage fUSE5, *E. coli* K91kan and MC1061 were provided by G. Smith. Baculovirus containing the full length open reading frame of c-*src* cDNA was obtained from H. Varmus. A hybridoma expressing monoclonal antibody clone 327 against pp60^{c-src} was obtained from J. Brugge.

2.2. Construction of the 15-amino-acid random peptide library

The 15-amino-acid random peptide library was constructed as described [8] with the modifications indicated in Fig. 1. A mixture of oligonucleotides encoding a 15-amino-acid peptide was synthesized using the sequence 5′-ACTCGGCCGACGGGGCT (NNK)$_{15}$ GGGGCCGCTGGGGCCGAA-3′ (TN-1). Within the NNK codon, 'N' is an equimolar mixture of G, A, T and C whereas 'K' is an equimolar mixture of G and T. Two 5′-biotinylated oligonucleotide primers, TN-2 (5′-ACTCGGCCGACGGGGC-3′) and TN-3 (5′-TTCGGCCCCA-G CGGCCC-3′), converted the TN-1 single strand degenerative oligonucleotide to double stranded DNA by polymerase chain reaction

*Corresponding author. Department of Tumor Genetics and Biology, Kumamoto University School of Medicine, 2-2-1 Honjo, Kumamoto 860, Japan. Fax: (81) (96) 373-5120.
E-mail: hsaya@gpo.kumamoto-u.ac.jp

PII S0014-5793(96)01329-4

(PCR) amplification. The PCR reaction (1 ml) was performed for five temperature cycles (2.5 min, 95°C; 4 min, 42°C; 4.4 min, 72°C; and 7 min; 72°C) and contained 1 μg of template, 5 μg of each biotinylated primer and 25 units of *Taq* DNA polymerase (Boehringer Mannheim). The product was purified by phenol:chloroform extraction and ethanol precipitation and dissolved in 100 μl of TE (10 mM Tris-Cl, 1 mM Na_2EDTA, pH adjusted to 8.0 with NaOH).

A 30 μl aliquot of the PCR-amplified 63-bp oligonucleotide mixture was digested (3 h, 37°C) with 40 μl of *BglI* (80 U/ml, Promega) in a total volume of 500 μl. Streptavidin-agarose (S-A) beads (200 μl of a 50% v/v suspension) were washed with 0.1 M NaCl in TE, and suspended in the *BglI* digest to remove the biotinylated primers. After rocking for 1 h, the S-A beads were removed by centrifugation, washed twice with 200 μl of water and the supernatants were pooled. The final product was extracted from the supernatant with phenol:chloroform and evaporated to 100 μl under reduced pressure using a Speed-Vac concentrator.

The degenerate 63-bp insert (36 μl) was ligated (15°C, overnight) into an *SfiI* (10 μg, Promega) digest of the fUSE5 fusion phage. Electroporation of 2.5 μl of ligation product into MC1061 cells (50 μl) was performed by discharging a 40 μF capacitor charged to 2.5 kV across a 0.2 cm gap in parallel with a 200 Ω resistor. After 1 h in SOC medium at 37°C, transformants were pooled, and an aliquot (5 ml of a 10^{-3} dilution) was removed and plated on tetracycline (Tc) agar plates (20 μg/ml of Tc in LB) to quantify Tc-resistant cells. The remainder was added to 1 l of LB containing Tc (20 μg/ml) and grown through approx. 10 doublings at 37°C to amplify the library. Phage from liquid culture was obtained by clearing the supernatant twice by centrifugation (10000×*g*, 10 min, 4°C), precipitating phage virions with PEG (3.3% polyethylene glycol 8000 in 0.4 M NaCl), and centrifugation as above. The phage was resuspended in TBS (50 mM Tris-Cl, pH 7.5 with 0.15 M NaCl) and stored at 4°C. The final phage pellet was dissolved in 1.5 ml TBS buffer containing 0.02% NaN_3. The transducing unit (TU) of this library was determined by infecting starved K91kan cells with the diluted library and seeding on plates containing 20 μg/ml Tc. The calculated TU was 2.5×10^{14}/ml.

2.3. Sequencing of phage library clones

98 Tc resistant clones were picked and cultured (37°C, 24 h) in 2 ml medium with 20 μg/ml Tc to analyze the quality of the library. Each phage ssDNA was purified and used as a sequencing template. The occurrence of each amino acid was calculated and compared to that predicted on the basis of codon frequency.

2.4. Library screening

The library was screened against antibody-bound $pp60^{c-src}$. Purified recombinant $pp60^{c-src}$ (24 μg) was mixed with 9.2 μg of biotinylated anti-$pp60^{c-src}$ monoclonal antibody (clone 327), and rocked at 4°C for 1 h. To block the non-specific binding of phage to $pp60^{c-src}$ and/or the monoclonal antibody, 2.4×10^{12} TU of wild-type phage, which expresses the same surface proteins as our phage library except for the random 15-amino-acid peptide insert within the pIII protein, was added to the mixture and rocked at 4°C for 3 h. Subsequently, 50 μl of our library (2.5×10^{14} TU/ml) was added to the mixture and rocked at 4°C.

This mixture was diluted with 3.2 ml of TBS-Tween (TBS with 0.5% Tween 20) and incubated (10 min, room temperature) with a streptavidin coated polystyrene Petri dish (10 μg/35 mm) to facilitate binding between streptavidin and the $pp60^{c-src}$-antibody complex. The streptavidin-coated plate was blocked with 5 mg/ml BSA prior to incubation with the $pp60^{c-src}$-antibody complex. The dish was washed 10 times with TBS-Tween to remove phage that was not bound to $pp60^{c-src}$. Bound phage was removed with 800 μl of elution buffer (0.1 N HCl, pH adjusted to 2.2 with glycine, 1 mg/ml BSA). The eluate was immediately neutralized with 150 μl of 1 M Tris-NaOH, pH 9.1 and amplified by infecting starved K91kan cells. This amplified eluate underwent a second round of bio-panning using the same amount of purified $pp60^{c-src}$ and monoclonal antibody. However, only 0.12 mg of purified $pp60^{c-src}$ was used in the third round of bio-panning.

Thereafter, starved K91kan cells were infected with an aliquot of the eluate and plated on Petri dishes containing 20 μg/ml Tc. More than 100 colonies were picked and the ssDNA from each colony was purified for DNA sequencing.

2.5. DNA sequencing

The DNA sequence of the random insert of each clone determined by dideoxynucleotide chain termination was converted to an amino acid sequence and analyzed collectively.

2.6. Peptide synthesis

The peptides were synthesized by the sequential addition of amino acids using standard Fmoc continuous flow solid-phase chemistry. Peptides were purified to homogeneity by reverse-phase HPLC and the amino acid composition and molecular weight were confirmed by FAB-MS.

2.7. Purification and assay of $pp60^{c-src}$

Recombinant avian $pp60^{c-src}$ was expressed using the baculovirus-insect cell system and purified (final specific activity: 3.9 μmol min^{-1} mg^{-1}) as described [9]. The tyrosine kinase activity of $pp60^{c-src}$ was assayed radiometrically using poly E_4Y as the substrate [10]. The kinetics of inhibition were determined with variable concentrations of peptide inhibitor against 0–200 μg/ml poly E_4Y. K_i values were determined from secondary plots of the slope versus the inhibitor concentration.

3. Results

3.1. Characterization of the phage library

The number of different 15-amino-acid sequences (2.5×10^8) within the library was estimated by determining the number of recombinants recovered as Tc-resistant colonies. To determine the quality of the library, we picked 98 individual colonies and sequenced the DNA of their random insert. Table 1 compares the actual occurrence of each amino acid in the library to that predicted on the basis of codon frequency. There was an overabundance of Phe and Cys, and a low frequency of Lys, Met and Gln. The frequency of the other amino acids was similar to that predicted, indicating that the quality of this library was sufficient for screening. We did not find any identical 15-amino-acid insert sequences within the 98 clones examined.

3.2. Sequences of $pp60^{c-src}$ binding peptides

After the third round of bio-panning, 107 colonies were picked from Tc plates and sequenced. As shown in Table 2,

Table 1
Amino acid composition analysis of the phage peptide library

Amino acid	Codon	Frequency (%)	
		Predicted	Observed
Phe	TTT	3.13	8.4
Gly	GGG GGT	6.25	9.5
Val	GTG GTT	6.25	10.2
Ala	GCG GCT	6.25	6.7
Cys	TGT	3.13	6.8
Asp	GAT	3.13	5.5
His	CAT	3.13	1.7
Ile	ATT	3.13	3.9
Leu	TTG CTG CTT	9.38	10.5
Trp	TGG	3.13	2.8
Ser	AGT TCG TCT	9.38	9.9
Thr	ACG ACT	6.25	3.8
Glu	GAG	3.13	2.0
Pro	CCG CCT	6.25	6.3
Lys	AAG	3.13	0.4
Met	ATG	3.13	1.5
Asn	AAT	3.13	1.6
Gln	CAG	3.13	0.9
Arg	AGG CGG CGT	9.38	5.5
Tyr	TAT	3.13	1.9

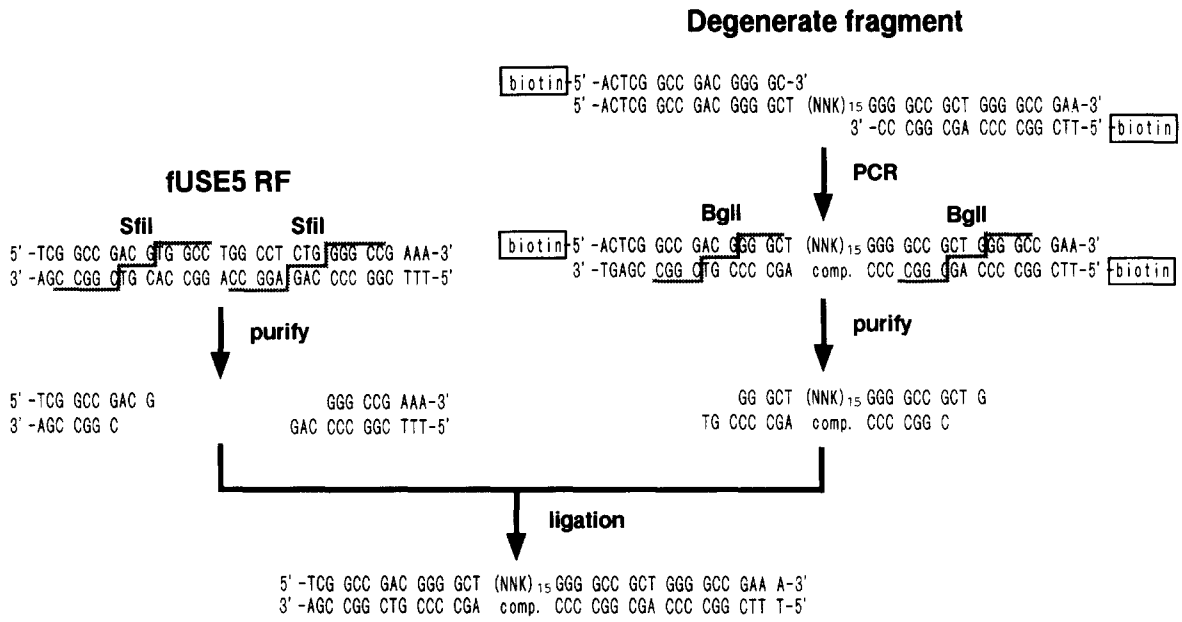

Fig. 1. Construction of a 15-amino-acid phage peptide library. (Left) Nucleotide sequence of fUSE5 around *SfiI* cloning sites. The vector was provided by Scott and Smith [8]. (Right) Sequence of the degenerate 63-bp *BglI* insert prepared by PCR amplification and *BglI* digestion.

67 of 107 clones (63%) have a GXXG sequence in their 15-amino-acid random region. Furthermore, 63 of those 67 clones (94%) have repeating sequences including GXXGXXG. In the original library, 21 of 98 clones (21%) contained a GXXG sequence and only 5 had repeating GXXG sequences. This result indicates that $pp60^{c-src}$ preferably interacts with the repeated GXXG motifs.

3. Testing the GXXGXXG sequence motif

To determine whether or not the glycine-rich GXXGXXG sequence motif can tightly bind to $pp60^{c-src}$, we synthesized FVGFLGFLG and examined its ability to inhibit the tyrosine kinase activity of $pp60^{c-src}$. This peptide was a competitive inhibitor (K_i=24 μM) that became a mixed-type inhibition profile at higher (double the K_i value) concentrations of inhibitor (Table 3). Considering that the peptide was a competitive-type inhibitor, we made tyrosine analogs of the FVGFLGFLG peptide, by replacing each Phe with a Tyr, to determine if it could serve as a tyrosine kinase substrate for $pp60^{c-src}$. Peptide phosphorylation was analyzed using a PEI column. However, there was no apparent phosphorylation of this peptide.

The multiple glycines, along with their spacing within this motif, suggest that peptide conformation plays an important role in its binding. To address this issue, we synthesized FVPFLPFLP. In comparison to the analog containing glycine, the proline peptide would be more conformationally constrained but would retain the hydrophobic nature of FVGFLGFLG. The FVPFLPFLP peptide was a mixed-type inhibitor with over 100-fold lower affinity (K_i = 3100 μM) for $pp60^{c-src}$.

4. Discussion

A bacteriophage library of 2.5×10^8 clones, with each clone expressing a potentially different 15-amino-acid peptide as part of its pIII coat protein, was constructed to identify a tightly binding peptide sequence that may inhibit the tyrosine kinase activity of $pp60^{c-src}$. The library was characterized with respect to its randomness and was found to be slightly biased with a greater than expected occurrence of two amino acids (Phe and Cys) and a lower than expected occurrence of three (Lys, Met, and Gln). This bias may have been introduced during the chemical synthesis of the degenerate oligonucleotide due to variability in the oligonucleotide synthesizer in simultaneously dispensing equivalent amounts of each nu-

Table 2
The $pp60^{c-src}$ binding peptide motifs identified from 107 clones after 3rd round panning

	$pp60^{c-src}$ binding clones (after 3rd round panning)	Original library (randomly selected)
GXXG (−)	40 (37.4%)	77 (78.6%)
GXXG (+)	67 (62.6%)	21 (21.4%)
GXXG (no repeat)	4 (3.7%)	16 (16.3%)
GXGXXG	6 (5.6%)	1 (1.0%)
GXXGXXG	20 (18.7%)	1 (1.0%)
GXXXGXXG	9 (8.4%)	0 (0.0%)
GXXXXGXXG	7 (6.5%)	0 (0.0%)
GXXXXXGXXG	6 (5.6%)	0 (0.0%)
GXXGXG	2 (1.9%)	1 (1.0%)
GXXGXXXG	7 (6.5%)	2 (2.0%)
GXXGXXXXG	5 (4.7%)	0 (0.0%)
GXXGXXXXXG	1 (0.9%)	0 (0.0%)

Table 3
Inhibition of $pp60^{c-src}$ kinase activity by a glycine-rich peptide

Peptide sequence	K_i (μM)
FVGFLGFLG	24
FVPFLPFLP	3100

cleotide and/or an effect of biological selection. This library was used to identify the peptide structures that bind to HLA-DR4-DQ4 [11] or HLA-DR9 [12]. Consensus peptide structures that bind to those HLA molecules were isolated from the library. Screening for an HLA-DR4-DQ4 binding structure showed that over 85% of the clones had consensus amino-acid sequences after the 3rd round of bio-screening [11]. Thus, the quality of this library is sufficient for screening.

Although several small organic compounds inhibit $pp60^{c-src}$ kinase activity, there are few active site-directed peptide-based inhibitors of $pp60^{c-src}$. The best peptidic inhibitor ($K_i = 8\mu M$) of $pp60^{c-src}$ consists of 21-amino acids, which is too long to serve as a lead sequence for developing an inhibitor [13]. The peptidic substrates are too large (10–15 residues) and their affinity is moderate ($K_m = 0.1$–6 mM). The peptide sequence described within is a 9-mer with moderate affinity. Although tyrosine analogs of this peptide did not serve as substrates for $pp60^{c-src}$ kinase activity, we believe that the GXXGXXG motif binds at the active site, since the mode of inhibition was competitive and we identified this motif as an excellent substrate for $pp60^{c-src}$ [14]. The peptide $(YGE)_5YGD$ which contains the GXXGXXG motif throughout is a tight-binding substrate with a K_m of 2 μM. ATP-binding sites of all protein kinases contain a sequence characterized by the Rossman motif, GXGXXG [15]. Cys-cdc-2(8–20), a synthetic peptide corresponding to the ATP binding site of $p34^{cdc2}$, is a specific and efficient substrate for a of $pp60^{c-src}$-related tyrosine kinase isolated from bovine spleen [16]. This result suggested that peptides containing repeated Gly-X or Gly-X-X can be a good substrate for src family kinases. The lack of observable phosphorylation of our peptides may be due to either a very low V_{max} and/or the orientation in which the peptides bind is not conducive to phosphate transfer.

Two independent groups have constructed and applied phage libraries to identify a tight-binding peptide ligand to the SH_3 domain of $pp60^{c-src}$ [17,18]. Both groups identified a PLPXXP sequence motif with binding constants ranging from 180 nM to 5 mM. Although we used full-length $pp60^{c-src}$ in bio-panning we did not identify any PLPXXP sequences. This could be due to use of the antibody-bound $pp60^{c-src}$. Binding of the antibody to the enzyme could inhibit interaction between the SH_3 domain and PLPXXP peptides. Moreover, since our washes were more stringent (300 mM NaCl) at the third-round screening than theirs (150 mM NaCl), GXXG peptides may preferentially bind to the full-length $pp60^{c-src}$.

In conclusion, our 15-amino-acid random peptide phage library is a powerful tool with which to identify peptide sequences that bind and/or inhibit the catalytic activity of various protein kinases.

Acknowledgements: We wish to thank G. Smith and J. Scott (University of Missouri-Columbia) for the phage strains and protocols, P. Lee (M.D. Anderson Cancer Center) for synthesizing the oligonucleotides and Shi Ke (M.D. Anderson Cancer Center) for technical assistance, K. Uriuda for secretarial assistance, and D. Mirozek for reading the manuscript.

References

[1] Jacobs, C. and Rubsamen, H. (1983) Cancer Res. 43, 1696–1702.
[2] Bolen, J., Rosen, N. and Israel, M. (1985) Proc. Natl. Acad. Sci. USA 82, 7275–7279.
[3] Barnekow, A., Paul, E. and Scharath, M. (1987) Cancer Res 47, 235–240.
[4] Bolen, J., Veilette, A. and Schwartz, A. (1987) Proc. Natl. Acad. Sci. USA 84, 2251–2255.
[5] Cartwright, C.A., Eckart, W., Simon, S. and Kaplan, P. (1987) Cell 49, 83–91.
[6] Cwirla, S.E., Peters, E.A., Barrett, R.W. and Dower, W.J. (1990) Proc. Natl. Acad. Sci. USA 87, 6378–6382.
[7] Devlin, J.J., Panganiban, L.C. and Devlin, P.E. (1990) Science 249, 404–406.
[8] Scott, J.K. and Smith, G.P. (1990) Science 249, 386–390.
[9] Budde, R.J., Ramdas, L. and Ke, S. (1993) Prep. Biochem. 23, 493–515.
[10] Ramdas, L., Obeyesekere, N.U., McMurray, J.S. and Budde, R.J. (1996) Arch. Biochem. Biopys. 326, 73–78.
[11] Matsushita, S., Nishi, T., Oiso, M., Yamaoka, K., Yone, K., Kanai, T. and Nishimura, Y. (1996) Int. Immunol. 8, 757–764.
[12] Fujisao, S., Matsushita, S., Nishi, T. and Nishimura, Y. (1996) Hum. Immunol. 45, 131–136.
[13] Sato, K., Miki, S., Tachibana, H., Hayashi, F., Akiyama, T. and Fukami, Y. (1990) Biochem. Biophys. Res. Commun. 171, 1152–1159.
[14] Budde, R., Obeyesekere, N.U., Ke, S. and McMurray, J.S. (1995) Biochim. Biophys. Acta 1248, 50–56.
[15] Rossmann, M.G., Moras, D. and Olsen, K.W. (1974) Nature 250, 194–199.
[16] Litwin, C.M.E., Cheng, H. and Wang, J.H. (1991) J. Biol. Chem. 266, 2557–2566.
[17] Cheadle, C., Ivashchenko, Y., South, V., Searfoss, G.H., French, S., Howk, R., Ricca, G.A. and Jaye, M. (1994) J. Biol. Chem. 269, 24034–24039.
[18] Sparks, A.B., Quilliam, L.A., Thorn, J.M., Der, C.J. and Kay, B.K. (1994) J. Biol. Chem. 269, 23853–23856.