THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1985 by The American Society of Biological Chemists, Inc.

Vol. 260, No. 1, Issue of January 10, pp. 64–71, 1985
Printed in U.S.A.

Localization of Tyrosine Kinase-coding Region in v-*abl* Oncogene by the Expression of v-*abl*-encoded Proteins in Bacteria*

(Received for publication, July 23, 1984)

Jean Y. J. Wang‡ and David Baltimore§

From the ‡Department of Biology, C-016, University of California San Diego, La Jolla, California 92093 and the §Whitehead Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

A series of plasmids containing different segments of the v-*abl* oncogene have been constructed to express different portions of the v-*abl* protein in bacteria. The tyrosine kinase activity of these proteins was determined by an *in vitro* assay employing histones or angiotensin II as substrates for the v-*abl*-encoded tyrosine kinase. These experiments show that the 5′−1.2 kilobases of v-*abl* is necessary and sufficient to produce an active tyrosine kinase which is functional as a monomeric soluble protein. The kinase-coding region corresponds to the minimal region of v-*abl* required for the transformation of fibroblasts. The kinase-coding region also coincides with the conserved protein sequences which are found in other tyrosine kinases. A compact domain of the v-*abl* protein including this kinase-coding region can accumulate to high levels in bacteria. The C-terminal region of the v-*abl* protein is not needed for the kinase activity and is rapidly degraded in bacteria.

The protein encoded by the oncogene v-*abl* of Abelson murine leukemia virus has been shown to contain tyrosine kinase activity (1, 2). The 3880-base pair sequence of v-*abl* has been determined.[1] The nucleotide sequence shows that v-*abl* encodes 1008 amino acids. The apparent molecular weight of the v-*abl* protein is 130,000. In the A-MuLV[2] genome, this 130-kDa v-*abl* coding sequence is preceded by 30 kDa of *gag* coding sequence which is derived from the *gag* gene of Moloney-MuLV. Thus, the fusion protein derived from A-MuLV has a M_r = 160,000. This strain of A-MuLV is referred to as A-MuLV(P160). Several spontaneous mutations in the v-*abl* of A-MuLV have been isolated (3). An A-MuLV producing a M_r = 120,000 fusion protein (A-MuLV(P120)) has a v-*abl* that lacks 780 bp of some internal coding sequence (3, 4). This deleted v-*abl* codes for a protein of only M_r = 90,000. The nucleotide sequence of this deleted v-*abl* in A-MuLV(P120) has been published (5). The deletion in v-*abl* is believed to have occurred during the passage of the original A-MuLV(P160) (6). The deleted v-*abl* has gone through a further point mutation to give rise to A-MuLV(P90) (3).[1] The v-*abl* of this P90 virus can only produce a M_r = 60,000 protein. All three types of A-MuLV can transform NIH3T3 cells, and the transformed cells all contain elevated levels of tyrosine phosphates (3). These A-MuLV mutants indicate that only a portion of the 3.9-kb sequence of v-*abl* is required for the expression of tyrosine kinase activity.

We have previously expressed v-*abl* from A-MuLV(P90) in bacteria and demonstrated that this coding sequence gives rise to a M_r = 62,000 protein which is phosphorylated on tyrosine residues in bacterial cells (1). To verify that the v-*abl* protein acts as a tyrosine-specific protein kinase, we have developed an *in vitro* assay for the v-*abl*-encoded kinase using exogenously added proteins as substrates. Combining this assay with the expression of different regions of v-*abl* in bacteria, we have localized in the v-*abl* sequence a minimal region required to code for an active tyrosine kinase. These studies also show that the v-*abl*-encoded protein consists of a tyrosine kinase domain and a large segment of polypeptide which is not needed for this enzymatic activity.

EXPERIMENTAL PROCEDURES

Preparation of Bacterial Extracts

Bacteria containing the expression plasmids were grown on L broth with 50 μg/ml ampicillin. Usually an overnight culture was grown at 30 °C and diluted 1/100 into fresh medium the next morning. The culture was incubated at 30 °C with adequate aeration for 2 h and then transferred to 40 °C to induce the expression of the v-*abl* protein (1). Bacteria were grown at the induced temperature for 4–5 h. They were then chilled immediately on ice, collected by centrifugation (5,000 × *g*, 10 min), and washed once with sonication buffer (20 mM Tris-HCl, pH 8.0, 2 mM EDTA, 100 mM NaCl, 10 mM β-mercaptoethanol, 4 °C). The washed cell pellet was resuspended in sonication buffer at 1/100 the volume of the original culture and sonicated for 3 min (six times at 30-s intervals). The sonicate was then centrifuged at 45,000 × *g* for 30 min to remove cell debris as well as membranes. At least 80% of the tyrosine kinase activity in the crude sonicate was shown to present in the supernatant of this high-speed spin. This supernatant was stored in aliquots at −20 °C and used in the *in vitro* assay of tyrosine kinase.

Tyrosine Kinase Assay

Assay of Histone Phosphorylation—Histone phosphorylation reaction was carried out in 25 mM Tris-HCl, pH 8, 10 mM $MgCl_2$, 1 mM dithiothreitol, 0.5 mM [γ-^{32}P]ATP (from New England Nuclear), 5,000 cpm/pmol, 2 mg/ml calf thymus histones (Boehringer Mannheim), and a given amount of bacterial extracts. Reaction was initiated by the addition of ATP and incubated at 30 °C. The rate of phosphorylation was linear with time under these conditions for 15 min. Reaction was terminated by the addition of EDTA to 50 mM final concentration. Bovine serum albumin was then added to a final concentration of 4 mg/ml. Proteins in the reaction mixture were precipitated by the addition of 10 volumes of 10% trichloroacetic acid, 0.1 M Na pyrophosphate; this precipitation step was allowed to proceed at 4 °C for 6 h to overnight. The trichloroacetic acid pellet was collected by centrifugation (10,000 × *g*, 2 min), dissolved in 100

* This work was supported by grants from the National Cancer Institute and a grant from the Camille and Henry Dreyfus Foundation to J. Y. J. W. and a National Science Foundation Grant PCM-8314300. Part of the work was completed while J. Y. J. W. was a postdoctoral fellow of the Jane Coffin Childs Memorial Fund for Medical Research. 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.

[1] R. Lee, M. Paskind, J. Y. J. Wang, and D. Baltimore, unpublished observations.

[2] The abbreviations used are: A-MuLV, Abelson murine leukemia virus; SDS, sodium dodecyl sulfate; kb, kilobase; bp, base pairs.

μl of 1 N NaOH, incubated at 37 °C for 1.5 h, and then neutralized with 8 μl of concentrated HCl, followed by the addition of 30 μl 50% trichloroacetic acid to reprecipitate the proteins. This final trichloroacetic acid pellet was washed three times each with 1 ml of 10% trichloroacetic acid, 0.1 M Na pyrophosphate. The amount of NaOH-resistant, trichloroacetic acid-precipitable ^{32}P incorporation was measured by scintillation counting.

Assay of Peptide Phosphorylation—Phosphorylation reactions were carried out at 30 °C in the same buffer conditions as histone phosphorylation, except 2.5 mM peptide substrates were used in the place of histones. Three different peptides could be used in this assay for the v-*abl*-encoded kinase: RRILEDAEYAARG or RRILDEADYAARG and angiotensin II (from Sigma). Reactions (usually 30 μl) were terminated by the addition of 70 μl of 3.2% trichloroacetic acid and 10 μl of 10 mg/ml bovine serum albumin at room temperature. Proteins were then removed by centrifugation in Eppendorf centrifuge for 2 min, and the supernatant was spotted in 40-μl aliquots onto a square (2 × 2 cm) of phosphocellulose paper (Whatman P81). The filter paper squares were washed five times with 400 ml each of 75 mM phosphoric acid, according to the method of Roskoski (7). These washed filters were air-dried and the amount of phosphopeptide was quantitated by scintillation counting.

Immunoprecipitation

Pulse-labeling with [^{35}S]Methionine—Bacteria were grown on M9 medium containing 0.4% glucose, 50 μg/ml of all amino acids except methionine, and 50 μg/ml ampicillin at 30 °C. After a 30-min induction period at 40 °C, [^{35}S]methionine was added at 100 μCi/ml for 5 min. Bacteria were quickly chilled and collected by centrifugation in the cold. For immunoprecipitation with the hybridoma antibody, bacteria were solubilized in hot SDS (2%), and the lysates were diluted with immunoprecipitation buffer (1) and then incubated with the culture medium from hybridoma pAb108 (from Dr. E. Gurney). When the anti-AbT sera were used to do immunoprecipitation, bacteria were sonicated and the soluble sonicates were used as described before (1).

Pulse-labeling with $^{32}P_i$—Bacteria were grown on minimal medium containing 0.4% glucose, 50 μg/ml of all amino acids, 50 μg/ml ampicillin, and 1 mM phosphate at 30 °C overnight. The next day, these cultures were diluted 1/20 into minimal medium without phosphate. These cultures were grown at 30 °C for 2 h followed by 30 min at 40 °C for induction. Carrier-free $^{32}P_i$ was then added at 500 μCi/ml for 5 min. Bacteria were either solubilized immediately in hot SDS (2%) or sonicated depending on the antibody to be used for immunoprecipitation.

Immunoprecipitation—Lysates were precleared with *Staphylococcus aureus* and incubated with antibody overnight. Immunocomplexes were collected onto formaldehyde fixed *S. aureus* and then solubilized for SDS-polyacrylamide gel electrophoresis as previously described (1).

Plasmid Construction

The v-*abl* sequences used in these constructions were obtained from DNA clones of the A-MuLV(P160) and the A-MuLV(P90) genomes (4, 8). A restriction map of the v-*abl* region used is shown in Fig. 4. Expression vectors pCQV2 and pCS4 were obtained from Queen (12). The v-*abl* regions were linked to an initiation codon of pCQV2 using the *Bam*HI site of the vector and the *Hinc*II site of v-*abl* as described in Ref. 1. The t/*abl* fusion plasmids were constructed as follows: a *Bgl*II site has been placed at the end of the 80-amino acid t-coding region in vector pCS4[3]; digestion of pCS4 with *Bgl*II followed by Klenow polymerase treatment generates a blunt-end that stops at the end of a codon. As shown in Ref. 1, *Hinc*II cuts v-*abl* in between the fourth and fifth codon. Therefore, ligation of the *Hinc*II site of v-*abl* to the blunted *Bgl*II site of pCS4 produced an in frame sequence that encodes a t/*abl*-fusion protein. Plasmid p*tabl*130 was constructed using pCS4 and pAB160 (4) and plasmid p*tabl*60 was from pCS4 and pAB90 (8). The deletions in v-*abl* were generated in p*tabl*130 using the appropriate restriction enzymes as shown in Fig. 4. Plasmid p*tsac* was prepared by linking a S1 nuclease-treated *Bgl*II site of pCS4 to the first v-*abl* *Sac*I site which was blunted by the 3′-exonuclease activity of Klenow polymerase. This arrangement put the 3′–v-*abl* sequence in frame with the t-coding sequence.

[3] C. Queen, personal communication.

RESULTS

In Vitro Assay of v-abl-encoded Tyrosine Kinase—Previously, we have shown that a v-*abl*-encoded protein can phosphorylate some bacterial proteins as well as itself on tyrosine residues when the v-*abl* coding sequence is expressed in *Escherichia coli* (1). To demonstrate that this *in vivo* activity can be assayed *in vitro*, we sought to find protein substrates for the v-*abl*-encoded kinase. Cell-free extracts prepared from bacteria expressing the v-*abl* protein were used as a source of the enzyme. Caseins and tubulins which are substrates for pp60^{src} or the epidermal growth factor receptor tyrosine kinases (9, 10) cannot be phosphorylated by these bacterial extracts. Calf thymus histones, however, can be phosphorylated by bacterial extracts containing the v-*abl* protein (Fig. 1*A*). Extracts prepared from bacteria containing the vector plasmid without v-*abl* could not carry out this phosphorylation. The phosphorylated histones corresponded to histones H2B and H4 based on their mobility on SDS-polyacrylamide gels. The identities of the histones were further demonstrated using purified subfractions of histones H2B and H4 as substrates. In these experiments, histone H1 was not phosphorylated by the v-*abl* kinase. A time course of histone phosphorylation at 30 °C is shown in Fig. 1*B*. The amount of phosphorylation is proportional to the amount of bacterial extracts used (Fig. 1*C*). The phosphate incorporation measured under the assay conditions is in the form of tyrosine phosphate as determined by phosphoamino acid analysis (Fig. 1*D*). These results demonstrate that the v-*abl*-encoded protein can phosphorylate histones H2B and H4 on tyrosine residues *in vitro*. In the absence of added protein substrates, the extracts from p*ABL*60 (1) did not produce significant amounts of phosphotyrosine when incubated at 30 °C with [α-^{32}P]ATP (Fig. 1*D*). This lack of endogenous tyrosine phosphorylation *in vitro* is most likely due to the lack of turnover of phosphotyrosine in bacterial cells. When the tyrosine-phosphorylated v-*abl* proteins and bacterial proteins were pulsed-labeled *in vivo* and then chased with unlabeled phosphates, they did not lose their label after a prolonged chase period (data not shown). It seems that phosphotyrosine residues cannot be dephosphorylated in bacteria. Therefore, the extracts prepared from these bacteria contained no endogenous substrates for the detection of the v-*abl*-encoded tyrosine kinase activity in an *in vitro* reaction.

The *in vitro* histone phosphorylation reaction differs from the *in vitro* autophosphorylation reaction of the Abelson viral protein as described by Witte *et al.* (2). The histone phosphorylation reaction prefers Mg^{2+} to Mn^{2+} (rate in 10 mM Mg^{2+} is about 10 times that in 10 mM or 2 mM Mn^{2+}), and its rate is negligible at 4 °C. On the contrary, autophosphorylation of the A-MuLV protein in an immunoprecipitated form proceeds rapidly at 4 °C, and Mn^{2+} is a much better cofactor than Mg^{2+} for that reaction (2).

The bacterial extracts containing the expression plasmid p*ABL*60 (1) were fractionated on a Sephadex G-100 gel filtration column where the histone kinase activity was found to migrate as a protein with a M_r ~ 60,000 (data not shown). Because the p*ABL*60-derived protein is 62,000, based on immunoprecipitation results (1), the v-*abl*-encoded tyrosine kinase is active as a monomer.

We have also found that two peptides can be used as substrates to measure the v-*abl* kinase activity. One of them is a peptide related to the Tyr^{416} tryptic peptide of pp60^{src}, RRILEDAEYAARG, developed by Casnelli *et al.* (11), the other is angiotensin II (DRVYIHPF). Peptide phosphorylation can be measured by phosphocellulose paper-binding assay (7) which is much more convenient than the trichloroacetic

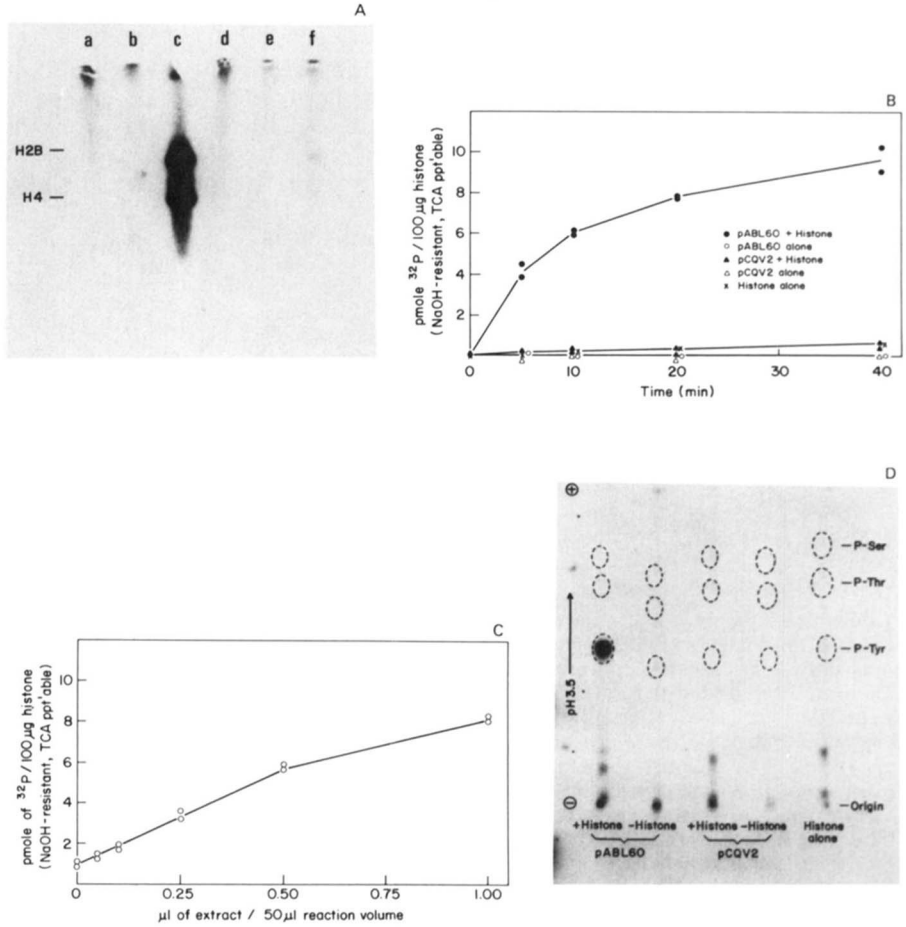

FIG. 1. **Tyrosine phosphorylation of histones by v-*abl*-encoded kinase in bacterial extracts.** *A*, histones were incubated with extracts from bacteria expressing p*ABL*60 or pCQV2 under the conditions described under "Experimental Procedures." A 50-μl reaction mixture was mixed with 25 μl of a 3-fold concentrated SDS buffer (1) and heated at 100 °C for 5 min. The SDS-solubilized material was loaded onto a 15% SDS-polyacrylamide gel. After electrophoresis and fixation, the gel was soaked in 1 N NaOH at 55 °C for 1 h, refixed with 10% acetic acid, dried, and autoradiographed. *Lane a*, p*ABL*60 extract alone; *lane b*, p*ABL*60 extract plus casein; *lane c*, p*ABL*60 extract plus histones; *lane d*, pCQV2 extract alone; *lane e*, pCQV2 extract plus casein; *lane f*, pCQV2 extract plus histones. *B*, time course of histone phosphorylation. Reactions were carried out with 200 μg/ml of the different bacterial extracts indicated in the figure and 2 mg/ml calf thymus histones as described under "Experimental Procedures." Portions were removed from the reaction at indicated times, and the amount of NaOH-resistant, trichloroacetic acid (TCA)-precipitable ^{32}P incorporation was measured as described under "Experimental Procedures." *C*, histone phosphorylation as a function of the amount of bacterial extract used. Reactions were carried out with an extract prepared from a 40 °C culture of bacteria containing p*ABL*60. The extract had a protein concentration of 15 mg/ml. Incubations and measurement of ^{32}P incorporation were as described above. *D*, phosphoamino acid analysis of phosphorylated histones. Phosphorylation reactions were carried out under the standard conditions containing histones and/or bacterial extracts as indicated. The final trichloroacetic acid pellet recovered after NaOH treatment was resuspended in 100 μl of 6 N HCl and hydrolyzed at 100 °C for 2.5 h. The acid hydrolysates were dried, washed with H_2O, and subjected to thin-layer electrophoresis at pH 3.5 as described (1). Authentic phosphoserine, phosphothreonine, and phosphotyrosine were added to the sample, and their positions of migration were revealed by ninhydrin staining.

acid precipitation assay of histone phosphorylation described here. The angiotensin II kinase encoded by a *Hinc*II to *Pst*I fragment of the v-*abl* sequence (see below) has been purified to homogeneity from bacterial extracts.[4]

Expression of Different Segments of v-abl Using pCQV2 Vector—As described above, a M_r = 60,000 protein encoded by the v-*abl* of A-MuLV(P90) was previously expressed in bacteria (p*ABL*60) (1). We have since constructed three other plasmids using the pCQV2 vector (1). The first contains 3.7 kb of the undeleted v-*abl* (from *Hinc*II to *Hin*dIII) which has the capacity to encode a 130,000-dalton protein; this plasmid is called p*ABL*130 (see Fig. 4 for restriction map of v-*abl*). The second p*ABL*90 was derived from p*ABL*130 by deleting the v-*abl* sequence 3′ to the unique *Sal*I site. The v-*abl* in plasmids p*ABL*90 and p*ABL*60 terminate at the same *Sal*I site, but p*ABL*90 contains the 780-bp internal sequence which is not present in p*ABL*60. The third plasmid p*ABL*40 was derived from p*ABL*60 by deleting sequences downstream from the first *Sac*I site of v-*abl* (Fig. 4), leaving only the first 1.1 kb of v-*abl*.

The products of these four expression plasmids p*ABL*40, p*ABL*60, p*ABL*90, and p*ABL*130 were identified by immunoprecipitation of [^{35}S]methionine-labeled bacterial lysates with anti-AbT sera (1). As shown in Fig. 2, a series of three bands of 62, 55, and 48 kDa were precipitated from p*ABL*60 lysates as observed before (1). The two lower bands are most likely

[4] J. Fulkes, A. Gurka, and D. Baltimore, manuscript in preparation.

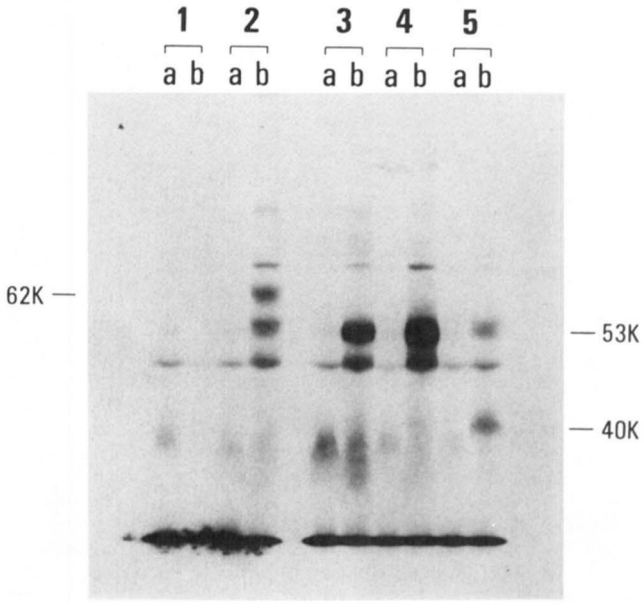

FIG. 2. **Products of four different v-*abl* expression plasmids in vector pCQV2.** Bacteria were grown at 30 or 40 °C in methionine-free medium. Pulse-labeling with [^{35}S]methionine, preparation of lysates by sonication, and immunoprecipitations with anti-AbT sera were as described (1). *1*, pCQV2; *2*, p*ABL*60; *3*, p*ABL*90; *4*, p*ABL*130; *5*, p*ABL*40; *a*, 30 °C; *b*, 40 °C.

degraded forms of the 62-kDa product (*lane 2b*). The 1.1-kb v-*abl* of p*ABL*40 can encode 39,000 daltons of protein. Immunoprecipitation of p*ABL*40 lysates did produce a 40-kDa band and another protein of 53 kDa (*lane 5b*). Because there is no termination codon in the 1.1-kb v-*abl*, the 53-kDa protein could be the primary product of p*ABL*40, where the v-*abl*-encoded protein terminated in the vector sequence. In that case, the 40-kDa band could be a degraded form of the 53-kDa product. It is also possible that the 53-kDa band is a bacterial protein which was coprecipitated with the 40-kDa v-*abl* protein. Both p*ABL*90 and p*ABL*130 plasmids produced a 53-kDa protein which is much smaller than the predicted size (*lanes 3b* and *4b*). Because these two plasmids generated the same protein, the 53-kDa product must not be encoded by the sequences 3′ to the *Sal*I site of v-*abl*. These immunoprecipitated proteins could only be found in bacteria grown at 40 °C, showing that they were derived from the heat-inducible promoter of the pCQV2 expression vector. *In vivo* pulse-labeling with $^{32}P_i$ showed that the p*ABL*60, 90, and 130 products were tyrosine phosphorylated. The proteins from p*ABL*40 were not phosphorylated, and in bacteria expressing p*ABL*40, there is no tyrosine phosphorylation of bacterial proteins (data not shown).

Over-expression of v-abl Proteins Using PCS4 Vector—Although the v-*abl* expression through PCQV2 gave rise to enough protein for the detection of tyrosine kinase activity, the amount of v-*abl* protein produced is less than 0.005% of the total bacterial proteins. In an attempt to overproduce the v-*abl* protein in *E. coli*, we constructed plasmids to express a small t-antigen/v-*abl* fusion protein because the small t-antigen of SV40 can be overproduced in *E. coli*, using the pCQV2 vector (12). The t/*abl* fusion plasmids were prepared using a derivative of pCQV2 named pCS4 which was constructed by Queen.[3] Vector pCS4 contains the P_R promoter of phage λ, the ribosome-binding sequence, and the initiation codon of the *cro* protein followed by the first exon sequence of SV40 t-antigen. A *Bgl*II site is placed at the end of this large T/small t-antigen common coding sequence so that other coding sequences can be appended to produce a t-fusion protein (see "Experimental Procedures"). The fusion protein contains only the first 80 amino acids of t-antigen. This vector also has the temperature-sensitive c*I* gene of λ phage to regulate expression through the P_R promoter (12).

As shown in Table I, the histone kinase activity in extracts containing the t/*abl*60 fusion protein is six times higher than that of extracts containing the p*ABL*60-encoded protein. This shows that the addition of 80 amino acids of the SV40 small t-antigen to the v-*abl* protein does not abolish its tyrosine kinase activity. The increase in activity can be accounted for by the increased expression of the t/*abl* fusion protein, shown in Fig. 3. Bacteria containing pCS4, p*ABL*40, p*tabl*40, p*ABL*60, and p*tabl*60 were pulse-labeled with [^{35}S]methionine for 5 min at 40 °C, and the total lysates were loaded on a SDS-polyacrylamide gel. The p*ABL*40- and p*ABL*60-encoded proteins are not visible among the bacterial protein bands (*lanes b* and *d*), although they could be detected after immunoprecipitation (see Fig. 2). However, a 51-kDa band is found

TABLE I

Histone kinase activity of pABL60 and ptabl60 extracts

Soluble extracts from bacteria expressing the 62-kDa v-*abl* protein or the 72-kDa t/*abl* fusion protein of these two plasmids were prepared as described under "Experimental Procedures." Histone kinase activity was measured as described.

Plasmid	Histone kinase activity
	pmol/min/mg protein
p*ABL*60	9.5
p*tabl*60	61.2

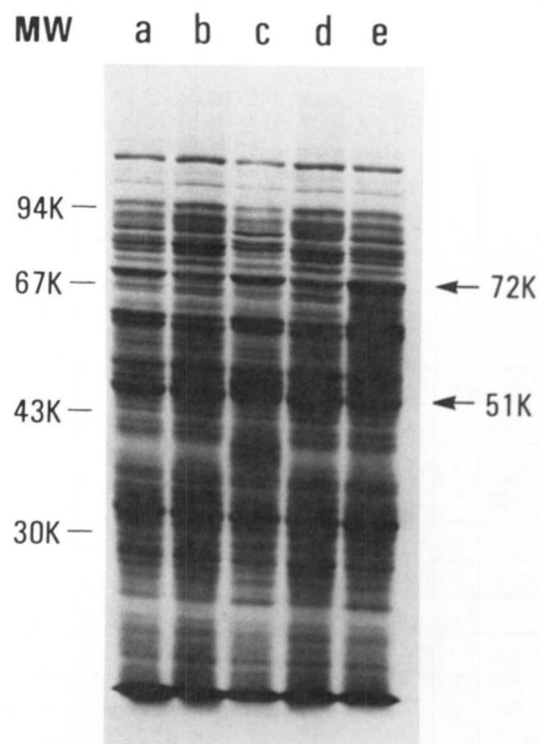

FIG. 3. **Addition of small t-coding sequence enhances production of v-*abl* proteins in bacteria.** Bacteria were grown on minimal medium without methionine at 30 °C and then induced for half an hour at 40 °C. Fifty μCi/ml [^{35}S]methionine was given to each culture; bacteria were collected after a 5-min labeling period, solubilized in SDS-sample buffer, and loaded onto a 10% SDS-polyacrylamide gel. *a*, pCS4 (vector); *b*, p*ABL*40; *c*, p*tabl*40; *d*, p*ABL*60; *e*, p*tabl*60. *Arrows* indicate the products of p*tabl*40 (51 kDa) and p*tabl*60 (72 kDa).

in bacteria expressing p*tabl* 40 (*lane c*) and a 72-kDa band is synthesized in bacteria containing p*tabl*60. These two proteins are the t/*abl* fusion proteins because they could be immunoprecipitated by antibodies against small t-antigen (see below). The t/*abl* fusion proteins are the predominant proteins synthesized at 40 °C. The increased protein production could be due to an increased stability of the t/*abl* protein or a more efficient translation of the hybrid RNA.

Because the pCS4 vector gives a higher level of expression, different segments of the v-*abl* coding sequence were placed in this vector to localize the tyrosine kinase-coding region.

Mapping the Kinase-coding Region in v-abl—Seven different fusion plasmids were constructed; their v-*abl* inserts are shown in Fig. 4. Six of them begin at the 5′-end (*Hin*cII site) of v-*abl*. Plasmid *tabl*40 ends on the first *Sac*I site; p*tabl*60 ends on the *Sal*I site and it lacks the 780-bp internal sequence; p*tabl*90 ends on the same *Sal*I site but contains the 780-bp internal sequence; p*tabl*130 ends on the *Hin*dIII site. Plasmid *tabl*106 was constructed by deleting an internal *Sac*I-*Sac*I fragment of p*tabl*130. The *Sac*I sticky ends were blunted before ligation so that this deletion does not interrupt the reading frame. The seventh p*tsac* plasmid contains sequence from the first *Sac*I site to the *Hin*dIII site; thus, it lacks the 5′–1.1 kb of v-*abl*. The expected molecular weight of the products of these plasmids, based on the coding capacity of their small t and v-*abl* segments, are given in Table II. These fusion proteins were identified by immunoprecipitation, using a hybridoma antibody pAb108 which recognizes a determinant in the 80-amino acid region of the small t-protein. (This hybridoma was prepared and kindly given to us by Dr. E. Gurney at the University of Utah.) Plasmids t-60, t-106, t-130 and *tsac* contain termination codons in their v-*abl* segments (Fig. 4). Plasmids t-40, t-50, and t-90 do not contain

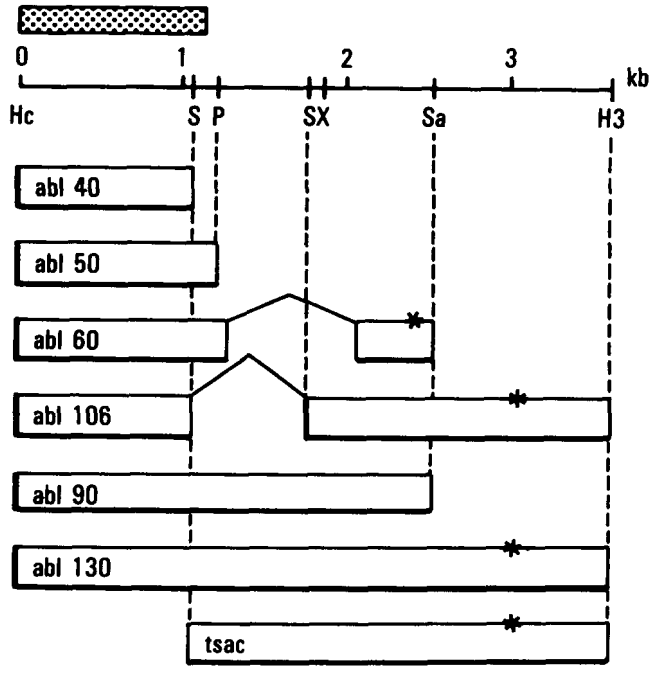

FIG. 4. **Seven t/*abl* fusion plasmids expressing different regions of the v-*abl* coding sequence in *E. coli*.** The different v-*abl* sequences were linked to 240 bp of a small t-coding region in frame. Six of the seven plasmids begin at the *Hin*cII site of v-*abl*. This *Hin*cII digestion removes the first four codons of v-*abl* (1). The *hatched box* above the v-*abl* sequence indicates the location of the homology between v-*abl* and other tyrosine kinase genes. The *asterisk* marks the position of termination codons. *Hc*, *Hin*cII; *S*, *Sac*I; *P*, *Pst*I; *X*, *Xho*I; *Sa*, *Sal*I; *H3*, *Hin*dIII.

TABLE II

Summary of protein products of the seven expression plasmids

Plasmid	t/*abl* protein		Tyrosine kinase[a]
	Expected[b]	Observed[c]	
	kDa		
p*tabl*40	49	51	–
p*tabl*50	54	70,60	+
p*tabl*60	70	72	+
p*abl*106	95	60–69	–
p*tabl*90	102	62	+
p*tabl*130	120	62	+
p*tsac*	82	22–24	–

[a] Tyrosine kinase activity was determined both *in vivo* by pulse-labeling with $^{32}P_i$ (1) and *in vitro* as shown in Table III.

[b] Expected molecular weight was calculated using 110 as an average weight of an amino acid. The number of amino acids encoded by a given segment of v-*abl* was deduced from the number of base pairs in that sequence. The apparent molecular mass of the small t-fragment is 10,000 daltons. The expected size is a sum of 10,000 and the calculated v-*abl*-encoded molecular mass.

[c] Observed molecular weight was determined on SDS-polyacrylamide gel as in Fig. 5.

termination codons in their v-*abl* segments; so, their products can be larger than the predicted size, depending on the location of the first available termination signal from the vector sequence.

A typical immunoprecipitation result is shown in Fig. 5*A*. The product of pCS4 (the 10-kDa fragment of t-antigen) migrates at the front of the 10% gel (*lane a*). Plasmid t-40 produced a 51-kDa protein which is slightly larger than the predicted 49-kDa, due to the need to use a termination codon outside of its v-*abl* sequence (*lane b*). Plasmid t-50 gave rise to a band of M_r = 70,000, also larger than the predicted size, and several degradative products of 52, 54, 56, and 60 kDa (*lane c*). The product of t-60 plasmid is an expected 72-kDa protein, and some of its degraded forms were also precipitated (*lane d*). The products of those plasmids which contain the 3′ sequence of v-*abl* are much smaller than their expected sizes. The fusion proteins from plasmids t-90 and t-130 are 62 kDa (*lanes f* and *g*). The products of these two plasmids were exactly the same in size and in the pattern of their degraded forms. This observation is consistent with those obtained with plasmids p*ABL*90 and p*ABL*130, both of which gave rise to a 53-kDa band (Fig. 2). The 53-kDa v-*abl* protein and the corresponding 62-kDa t/*abl* fusion protein must be encoded by the 5′-region of v-*abl* because (i) deletion of up to 2 kb from the 3′-end does not affect the size of this product; (ii) the addition of t-coding sequence on the 5′-end of v-*abl* causes an expected increase of 10 kDa in the molecular mass of this product. The 3.7-kb fragment of v-*abl* present in t-130 (*Hin*cII-*Hin*dIII) does give rise to a 130,000-dalton protein when it is expressed in mammalian cells. That 130-kDa protein is unstable in bacteria, but an N-terminal 53-kDa region is stable. The instability of the C-terminal region of the v-*abl* protein is demonstrated by the expression results obtained with plasmids t-106 and *tsac*. As shown in Fig. 4, pt-106 is equivalent to t-40 plus a truncated C-terminal portion of v-*abl*. Although the product of t-40 is a discrete band (Fig. 5*A*, *lane b*), t-106 does not produce an expected M_r = 95,000 protein but gives rise to a series of degraded proteins ranging from 58 to 69 kDa (Fig. 5*A*, *lane e*). The truncation at the *Sac*I site prevented the appearance of a stable 62-kDa fusion protein found in t-90 and t-130, and the appendage of the 3′-v-*abl* sequence to t-40 seemed to increase the rate of degradation so that no discrete bands were found. Plasmid *tsac* totally lacks the 5′-region (1.1 kb) of v-*abl*. The immunopre-

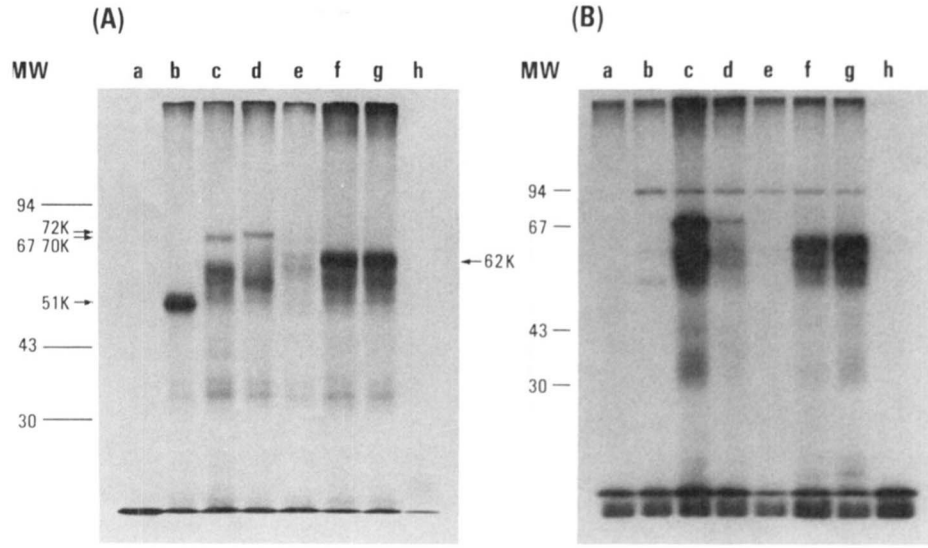

FIG. 5. **Protein products of seven t/*abl* fusion plasmids in noprecipitated with hybridom tibody against t-antigen.** *A*, ba containing the plasmids grow pulse-labeled with [^{35}S]methionin bacteria were grown and pulse-la with $^{32}P_i$. Pulse labeling, preparati lysates, and immunoprecipitation as described under "Experimental cedures." *a*, pCS4; *b*, p*tabl*40; *c*, p*t d*, p*tabl*60; *e*, p*tabl*106; *f*, p*tabl*9 p*tabl*130; *h*, p*tsac*. Protein mole mass (MW) standards were phosph ase *a* (94 kDa), bovine serum alb (67 kDa), ovalbumin (43 kDa), and bonic anhydrase (30 kDa).

cipitated proteins from *tsac*-producing bacteria were a faint smear at molecular mass 22–24 kDa, although the predicted size is 82 kDa (Fig. 5*A*, *lane h*). The half-life of the C-terminal portion must be less than 30 s because the full-length v-*abl* protein could not be found even in a 30-s pulse labeling with [^{35}S]methionine. The results of pulse-chase experiments indicated that the newly synthesized v-*abl* protein was rapidly cleaved to generate the N-terminal domain and some fragments from the C-terminal which were quickly degraded. The rapid degradation of the C-terminal region is observed even in protease mutant (*lon*$^-$) bacteria (data not shown).

Immunoprecipitation of these t/*abl* fusion proteins with anti-AbT sera gave results similar to those shown in Fig. 5*A*, except for the products of p*tabl*40 where, in addition to the 51-kDa band, anti-AbT sera also reacted with a 40-kDa protein (data not shown). This 40-kDa protein can be seen in Fig. 3, *lane c*, as a major product in bacteria expressing p*tabl*40. Because it reacted with anti-AbT sera but not with anti-t antibody, this 40-kDa protein could be a degraded product of the 51-kDa fusion protein where the t-antigen sequence was separated from the v-*abl* protein. Alternatively, the 40-kDa protein could be synthesized from an internal methionine codon (there are six in-frame methionine codons in the t-coding sequence) thus lacking the antigenic determinant recognized by the anti-t monoclonal antibody. The products of the other plasmids may also have a smaller derivative recognized only by anti-AbT sera. These smaller derivatives were not readily discernible because the products of these other plasmids all contained many different degradated forms which were precipitated by anti-t antibody (Fig. 5*A*). The multiple bands could have obscured the differences between immunoprecipitations with anti-AbT or anti-t antibodies.

The tyrosine kinase activity of these different v-*abl* products in bacterial extracts was measured using both histones and angiotensin II as substrates. As shown in Table III, four of the plasmids produced an active tyrosine kinase: pt-50, pt-60, pt-90, and pt-130. The extracts prepared from bacteria containing pt-40, pt-106, and p*tsac* had no detectable kinase activity. The difference in the kinase activities of the different extracts is consistently observed. The lower activity of the pt-90 and pt-130 extracts may be explained by the fact that the 62-kDa fusion protein is overproduced, and most of the proteins are removed from the soluble extracts by centrifugation (see Fig. 6).

The proteins encoded by t-50, t-60, t-90, and t-130 are

TABLE III

In vitro kinase activity of different v-abl proteins

Kinase activity was measured using extracts of bacteria containing the different plasmids. Activity is expressed in picomoles of ^{32}P incorporated per mg of protein/min at 30 °C. UD, undetected.

Plasmid	Histone kinase	Angiotensin II kinase
pCS4	UD	UD
p*tabl*40	UD	UD
p*tabl*50	52.3	169.0
p*tabl*60	51.4	166.0
p*tabl*106	UD	UD
p*tabl*90	41.7	97.0
p*tabl*130	8.1	47.0

themselves phosphorylated on tyrosine. As shown in Fig. 5*B*, the pAb108 hybridoma antibody recognized phosphorylated proteins from bacteria expressing these plasmids (lanes *c*, *d*, *f*, and *g*). The sizes of these phosphorylated proteins corresponded to those labeled by [^{35}S]methionine except for one 92-kDa band which was present in all the immunoprecipitations containing the v-*abl* proteins (*lanes b* through *g*). The 92-kDa and two other minor phosphoprotein bands were not labeled with [^{35}S]methionine, indicating that their synthesis was not induced by heat as were the t/*abl* fusion proteins. These three protein bands are not consistently found in immunoprecipitation, and they do not contain phosphotyrosine. The products of plasmids t-40, t-106, and t*sac* are not phosphorylated (lanes *b*, *e*, and *h*).

The 72-kDa protein of p*tabl*60 was phosphorylated to a lesser extent than the 70-kDa protein of p*tabl*50, although these two proteins were synthesized at a comparable level during the 5-min pulse-labeling period (Fig. 5, *A* and *B*, *lanes c* and *d*). The 72-kDa protein contains some truncated C-terminal sequence which is not present in p*tabl*50 (Fig. 4). These C-terminal sequences might have caused a reduction in the rate of autophosphorylation of the 72-kDa protein.

Overproduction of Two t/abl Fusion Proteins—Although the addition of t-antigen coding sequence to the 5′-end of v-*abl* enhances the expression of the v-*abl* proteins, this enhancement does not always result in overproduction. The products of the fusion plasmids failed to accumulate in bacteria, except for the 62-kDa N-terminal domain derived from plasmids p*tabl*90 and p*tabl*130. Although the 62-kDa fusion protein accumulated to high levels and formed aggregates which could be collected by a low-speed centrifugation (Fig. 6, *lane 10*),

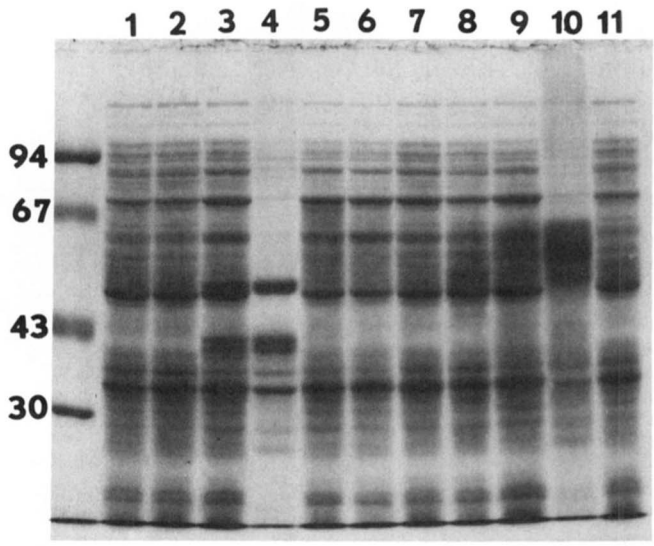

FIG. 6. **Overproduction of two t/*abl* fusion proteins.** Bacteria containing the plasmids were grown on L broth plus ampicillin and induced at 40 °C for 4 h as described (1). Cells were collected (1.5 ml) by centrifugation, solubilized in SDS-sample buffer (50 to 80 μl), and loaded onto a SDS-10% polyacrylamide gel. To prepare insoluble pellets, 1 liter of culture of p*t40a* or p*tabl* 130 was induced for 4 h at 40 °C. Bacteria were collected by centrifugation, washed with buffer A (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 100 mM NaCl), and sonicated. The crude sonicate was centrifuged at 5,000 × *g* for 10 min, and the insoluble protein pellets were resuspended in 10 ml of buffer A containing 3% Triton. After several hours at 4 °C, the Triton X-100 pellets were collected by low-speed centrifugation and resuspended in 1 ml of buffer A. The protein pellets were solubilized in hot SDS, and 0.4 μl was loaded in a SDS-polyacrylamide gel. The proteins were stained with Coomassie Blue. *1*), pCS4; *2*, p*tabl*40; *3*, p*t40a*; *4*, insoluble pellet from p*t40a*; *5*, p*tabl* 50; *6*, p*tabl*60; *7*, p*tabl*106; *8*, p*tabl*90; *9*, p*tabl*130; *10*, insoluble pellet from p*tabl*130; *11*, p*tsac*.

the overproduced proteins were not clearly visible among the bacterial protein bands (Fig. 6, *lanes 8* and *9*). As shown in *lane 10*, the precipitated proteins were a set of four bands ranging in size of 56 to 62 kDa. This is perhaps due to a limited degradation of the N-terminal domain. Approximately 8 mg of the aggregated protein could be obtained from 1 liter of bacteria; thus, the t/*abl* fusion protein made up about 5% of the total bacterial protein. The precipitated proteins are phosphorylated on tyrosine.

Overproduction of an unphosphorylated t/*abl* protein was achieved by the construction of a plasmid that has the same *Sac*I-*Sac*I deletion of pt-106 (see Fig. 4). Unlike pt-106, however, this plasmid p*t40a* was ligated with the *Sac*I sticky ends intact so that the downstream coding sequence was placed out of frame. This expression plasmid produced proteins of 51 and 40 kDa similar in sizes as those produced by pt-40. Although the pt-40 product does not accumulate to high levels (Fig. 6, *lane 2*), the p*t40a* product is overproduced (Fig. 6, *lane 3*). In these bacteria, the overproduced 51- and 40-kDa proteins were clearly visible among the bacterial bands (*lane 3*). These proteins were also aggregated and could be collected by low-speed centrifugation (*lane 4*). The level of overproduction of the p*t40a* product is comparable to that of the p*t130* product. Because of the truncation at the *Sac*I site, the p*t40a* proteins are not tyrosine phosphorylated.

DISCUSSION

Coding Region for the Tyrosine Kinase in v-abl—We have defined a minimal region in v-*abl* which is sufficient to encode an active tyrosine kinase. The coding sequence in a *Hinc*II to *Pst*I fragment (p*tabl*50), the first 1.2 kb of v-*abl*, is able to produce an active tyrosine kinase that phosphorylates bacterial proteins and *in vitro* substrates. This v-*abl* protein fragment is itself phosphorylated on tyrosine residues. Removal of 159 bp from the 3′-end of this minimal region by digestion with *Sac*I resulted in a loss of enzyme activity. The *Hinc*II to *Sac*I fragment (p*tabl*40) gives rise to a stable protein fragment which is not phosphorylated in bacteria, and extracts containing this protein fail to phosphorylate the *in vitro* substrates. There are no tyrosine residues in the 159-bp (53 amino acids)-deleted sequence; thus, the lack of phosphorylation of the *Hinc*II-*Sac*I-encoded fragment is not due to the removal of a phosphorylation site. Rather, the removal of those 53 amino acids eliminated the kinase activity as well as the autophosphorylation reaction.

Interestingly, the minimal kinase-coding region (*Hinc*II-*Pst*I fragment) corresponds to the minimal transforming region of the v-*abl* oncogene. A retrovirus constructed *in vitro* containing 100 bp of MLV *gag* linked to the same *Hinc*II-*Pst*I fragment can transform NIH3T3 cells (13). The *Hinc*II-*Sac*I fragment failed to induce transformation in a similar construction. However, this 40-kDa protein fragment was very unstable in mammalian cells, and we could not rule out the possibility that the lack of transformation may be due to the low level production of this protein.

It has been shown that the v-*abl*-encoded protein shares homology with proteins derived from several other oncogenes: v-*src*, v-*yes*, v-*fps*, v-*fes*, v-*ros*, v-*raf*, v-*erbB*, and v-*fms* (5, 14–16). This homologous region coincides with the minimal tyrosine kinase-coding region defined here. The homology between v-*src* and v-*abl* begins at the *Hinc*II site, extends beyond the *Sac*I site, and ends before the *Pst*I site (5). In fact, the *Sac*I digestion removes a peptide sequence which is highly conserved in all these oncogenes. This peptide sequence (corresponding to v-*src* amino acids 494–512) is apparently crucial to tyrosine kinase activity. Antibody against this peptide (residue 498–512) in v-*src* has been shown to inhibit the kinase activity of $pp60^{src}$ (17). Thus, the genetic and immunochemical results both indicate the importance of this small peptide sequence to the activity of tyrosine kinase.

Our results demonstrate that in v-*abl* this conserved sequence by itself can produce an active tyrosine kinase in bacteria. However, the v-*src* kinase is not active in bacteria (9), and the products of v-*erbB*, v-*raf* and v-*fms* do not contain detectable tyrosine kinase activity (14–16). These observations indicate that differences in the conserved region of these genes can influence the tyrosine kinase activity.

Domains of v-abl Protein—Expression of the v-*abl* sequence in bacteria also provided some indication of a domain structure in the v-*abl*-encoded protein. Of the M_r = 130,000 protein produced by v-*abl*, only a M_r = 53,000 *N*-terminal fragment is stable in bacterial cells. The rest of the protein of approximately 80 kDa is rapidly degraded. The resistance of the 53-kDa fragment to proteolytic digestion suggests that this portion of the v-*abl* protein forms a compact domain. This N-terminal domain is an active tyrosine kinase capable of autophosphorylation and phosphorylation of other protein substrates. It would require 1.4–1.6 kb of coding sequence to produce a 53,000-dalton protein. Although only 1.2 kb on the 5′-end of v-*abl* can give rise to an active tyrosine kinase, the stable domain is slightly larger than the minimal region. This is supported by the fact that the tyrosine kinase produced by the minimal region does not accumulate to high levels in bacterial cells, whereas the N-terminal domain is overproduced at 5% level. A *Hinc*II to *Xho*I fragment (containing the first 1.86 kb of v-*abl*) does produce the same 53,000-dalton-stable protein (data not shown), showing that this N-terminal

domain is derived from the first 1.5 kb of v-*abl*. The v-*abl* protein thus consists of a compact tyrosine kinase domain appended to a large C-terminal portion. Although the tyrosine kinase is completely independent of the C-terminal portion, this region may have a role regulating the kinase activity. Because this portion of the v-*abl* protein is not stable in bacteria, we have not been able to examine its function using the bacterial expression system.

In Vitro Activity of the v-abl-encoded Tyrosine Kinase—The product of the v-*abl* oncogene can phosphorylate proteins and peptides on tyrosine residues *in vitro*. With the limited proteins and peptides tested, we could not determine the specificity of this kinase. The two peptides (RRILEDAEYAARG and DRVYIHPF) have very different compositions, yet they are phosphorylated equally well by the v-*abl* kinase produced in bacteria. However, the v-*abl* kinase does not phosphorylate caseins or histone H1 (which contains one tyrosine residue), showing that it is not completely nonspecific. In the *in vitro* phosphorylation reactions, the v-*abl* kinase prefers Mg^{2+} as a cofactor. The active tyrosine kinase in bacterial extracts is a soluble monomeric protein. Thus, membrane association is not necessary for the manifestation of the enzyme activity.

Overproduction of v-abl Proteins in E. coli—At least two factors are important in achieving overproduction of the v-*abl* proteins. First, the addition of 240 bp of t-antigen-coding sequence enhanced the expression. However, that alone is not sufficient to cause overproduction of the fusion protein. A second factor which we refer to as the "intrinsic stability" of a protein is also crucial to its overproduction. Our results suggest that the stability of a protein fragment in *E. coli* can be augmented by its terminal sequences. The enhanced expression provided by small t-protein may partly be due to a stabilization of the fusion proteins. Sequences on the C-terminal of a protein can also change is stability. This is best demonstrated by the expression of the *Hinc*II-*Sac*I fragment from v-*abl*. In plasmid pt-40, this fragment is isolated from other v-*abl* sequence, and termination codon from the vector sequence is used to produce the 51-kDa fusion protein. The product of this plasmid does not accumulate to a high level. In plasmid p*t40a*, this fragment is linked to some v-*abl* sequences which are out of frame. The 51-kDa product of p*t40a* is different from that of p*t40* only at the extreme C-terminal end, and it is overproduced. Thus, a minute difference on the C-terminal results in the overproduction of one but not the other protein. The rates of synthesis of these two 51-kDa proteins are virtually the same as determined by pulse-labeling experiments. Similarly, the compact domain can accumulate to 5% level, but proteins smaller on the C-terminal end cannot be overproduced at the same level. Because an active tyrosine kinase can be overproduced, there does not seem to be a selection against the expression of this enzymatic activity in bacteria. The intrinsic stability of these t/*abl* fusion proteins was not altered by mutation in the *lon* locus of *E. coli*. The same overproduction results were obtained in *lon*$^-$ bacteria, *i.e.* only the products of p*tabl* 130 and p*t40a* were accumulated to high levels.

REFERENCES

1. Wang, J. Y. J., Queen, C., and Baltimore, D. (1982) *J. Biol. Chem.* **257,** 13181–13184
2. Witte, O. N., Dasgupta, A., and Baltimore, D. (1980) *Nature* (*Lond.*) **283,** 826–831
3. Goff, S. P., Witte, O. N., Gilboa, E., Rosenberg, N., and Baltimore, D. (1981) *J. Virol.* **38,** 460–468
4. Latt, S. A., Goff, S. P., Tabin, C., Paskind, M., Wang, J. Y. J., and Baltimore, D. (1983) *J. Virol.* **45,** 1195–1199
5. Reddy, E. P., Smith, M. J., and Srinivasan, A. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 3623–3627
6. Wang, J. Y. J., and Baltimore, D. (1983) *Mol. Cell. Biol.* **3,** 773–779
7. Roskoski, R., Jr. (1983) *Methods Enzymol.* **99,** 3–7
8. Goff, S. P., Gilboa, E., Witte, O. N., and Baltimore, D. (1980) *Cell* **22,** 777–785
9. Gilmer, T. M., and Erikson, R. L. (1981) *Nature* (*Lond.*) **294,** 771–773
10. Buhrow, S. A., Cohen, S., and Staros, J. V. (1982) *J. Biol. Chem.* **257,** 4019–4022
11. Casnelli, J. E., Hanson, M. L., Pike, L. J., Hellstrom, K. E., and Krebes, E. G. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79,** 282–286
12. Queen, C. (1983) *J. Mol. Appl. Genet.* **2,** 1–10
13. Prywes, R., Foulkes, J. G., Rosenberg, N., and Baltimore, D. (1983) *Cell* **34,** 569–579
14. Hampe, A., Gobet, M., Scherr, C. J., and Galibert, F. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 85–89
15. Privalsky, M. L., Ralston, R., and Bishop, J. M. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81,** 704–707
16. Mark, G. E., and Rapp, U. R. (1984) *Science* (*Wash. D. C.*) **224,** 285–289
17. Gentry, L. E., Rohrschneider, L. R., Casnelli, J. E., and Krebs, E. G. (1983) *J. Biol. Chem.* **258,** 11219–11228