Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 5802–5806, August 1990
Biochemistry

The complete coding sequence of *arg* defines the Abelson subfamily of cytoplasmic tyrosine kinases

(c-*abl*/c-*src*/c-*fps*/c-*fes*)

GARY D. KRUH, ROBERTO PEREGO, TORU MIKI, AND STUART A. AARONSON

Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, MD 20892

Communicated by David Baltimore, March 7, 1990

ABSTRACT **We have previously described partial genomic sequences of *arg*, a human gene related to c-*abl*, and shown that it is expressed as a 12-kilobase transcript and is located at chromosome position 1q24–25. In this study we elucidate the complete coding sequence of *arg* by characterization of cDNA clones. Analysis of the predicted amino acid sequence of arg revealed that it is indeed closely related to that of c-abl. The two proteins are strikingly similar with regard to overall structural architecture as well as the amino acid sequences of their tyrosine kinase and src homologous 2 and 3 domains. In addition, *arg*, like c-*abl*, is expressed as two transcripts that result from a process of alternative splicing and encode alternative protein forms that differ only in their amino termini. The two genes define the Abelson subfamily of cytoplasmic tyrosine kinases and share a common homolog in *Drosophila*.**

Protein tyrosine kinases can be divided into receptors that are characterized by transmembrane and ligand-binding domains and cytoplasmic molecules that do not cross the plasma membrane. The latter group, based upon characteristic structural features, falls into three subfamilies—c-src (lck, hck, fgr, yes, fyn, and lyn), c-fps/fes, and c-abl (1). Although functions of the cytoplasmic tyrosine kinases have been difficult to investigate because of the inability to trigger them with specific ligands, their role in signal transduction is indicated by the resemblance of their catalytic domains to those of growth factor receptors (2) and the transforming ability of their transduced counterparts in naturally occurring retroviruses (3).

Among the cytoplasmic protein tyrosine kinases, only c-abl has so far been shown to be directly involved in human malignancy (4–6). The translocation involving the c-*abl* locus on chromosome 9 and the *bcr* locus on chromosome 22 results in the cytogenetically recognizable Philadelphia chromosome and is found in 95% of patients with chronic myelogenous leukemia and ≈20% of adult patients with acute lymphoblastic leukemia (for review, see refs. 7 and 8). c-*abl* is also the transduced homolog of the transforming genes of Abelson murine leukemia virus (9, 10) and Hardy–Zuckerman 2 feline sarcoma virus (11), replication-defective retroviruses that cause aggressive lymphomas in mice and fibrosarcomas in cats, respectively. In each of these cases, the activated forms of c-*abl* have in common an elevated level of tyrosine kinase activity (5, 12).

Because of the involvement of c-*abl* in human malignancy, we sought to identify related genes in an effort to define the structure and function of this important subfamily. We have previously reported the identification by reduced stringency hybridization of genomic clones containing sequences related to c-*abl* (13). This gene, designated *arg* (for Abelson-related gene), is located at chromosome position 1q24–25 and is expressed as a 12-kilobase (kb) transcript. In this study we show by elucidation of the *arg* coding sequence* that it is indeed closely related to c-*abl* and that the two genes form a distinct subfamily of cytoplasmic tyrosine kinases represented by a common homolog in *Drosophila melanogaster*.

EXPERIMENTAL PROCEDURES

cDNA Cloning. The K562 cDNA library was kindly provided by O. N. Witte (Department of Microbiology, University of California, Los Angeles). The M426 and A172 cDNA libraries were prepared as described (14). The monocyte cDNA library was prepared as described (14) and kindly provided by S. Gutkind (National Institute of Dental Research, Bethesda, MD). cDNA libraries were screened with radiolabeled *arg* genomic or cDNA probes by using previously described hybridization conditions (13).

Nucleotide and Amino Acid Sequence Analysis. The nucleotide sequences were determined for both strands by the dideoxy chain termination method (15) using supercoiled plasmid DNA as template. Nucleotide sequence comparisons were performed by using the alignment program of Pearson and Lipman (16), and amino acid comparisons were performed by using the University of Wisconsin Genetics Computer Group GAP program.

RNA Blot Analysis. Total (17) and poly(A)$^+$ (18) RNA were prepared as described. RNA was separated by denaturing gel electrophoresis, transferred to nitrocellulose filters, and hybridized with radiolabeled *arg* probes as described (13).

RESULTS

Cloning and Nucleotide Sequence Analysis of *arg* cDNA. Initial *arg* cDNA clones were identified in a phage library prepared from K562 cells by using the previously described *arg* genomic clones as probes (13). The longest *arg* cDNA clone derived from this library, K3, was used as a source of probes to isolate additional cDNA clones. Overlapping cDNA clones extending 5′ or 3′ of clone K3 could not be identified in the K562 library but were found in cDNA libraries prepared from M426 cells, monocytes, and A172 cells. The overlapping *arg* cDNA clones, which together spanned 9.4 kb, are shown in Fig. 1.

Nucleotide sequence analysis of *arg* cDNA revealed a long open reading frame of 3546 base pairs, extending from nucleotide 205 to 3750 in the cDNA. The open reading frame encoded a predicted protein of 1182 amino acids and was preceded by a 5′ untranslated region high in G+C content. The location of the coding sequence at the 5′ end of the cDNA leaves at least 6 kb of 3′ untranslated sequence. Nucleotide

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: SH2, src homologous region 2; SH3, src homologous region 3.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M35296).

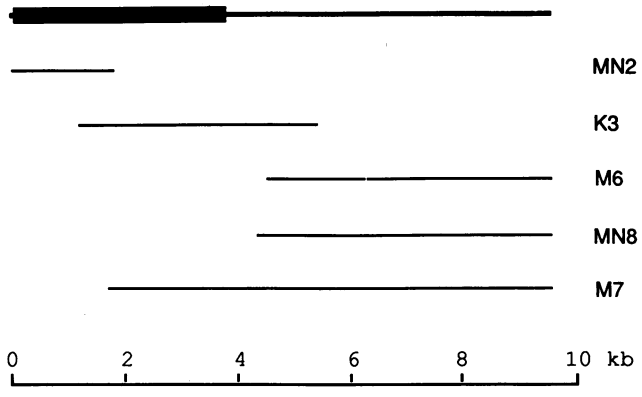

FIG. 1. Physical map of overlapping *arg* cDNA clones. The bar represents the open reading frame. cDNA clones were identified in K562 (K3), M426 (M6 and M7), and monocyte (MN2 and MN8) cDNA libraries. The positions of overlapping cDNA clones were established by restriction mapping.

sequence analysis of cDNA clones M6, M7, and MN8 revealed that their 3′ sequences were identical but did not disclose a typical polyadenylylation consensus motif. Thus it is possible that these three cDNA clones were generated by internal priming and that additional sequence lies further downstream. We have previously demonstrated that *arg* is expressed as a 12-kb transcript (13). Although size estimation of large transcripts by agarose gel electrophoresis can be inaccurate, the presence of additional 3′ sequence could account for the difference between the size estimated by Northern analysis and the 9.4 kb accounted for by these cDNA clones. It is also possible that additional noncoding sequence lies upstream of the *arg* coding region.

The *arg* Gene Product Has a High Degree of Structural Similarity with That of c-*abl*. Analysis of the predicted arg protein revealed features typical of members of the cytoplasmic tyrosine kinase family. A conserved tyrosine kinase domain spanned amino acids 281–532, and hydrophobic regions that could serve as signal peptide or transmembrane-spanning domains were absent (Fig. 2*B*). A schematic comparison of the *arg* gene product with the c-abl, c-src, and c-fps proteins is shown in Fig. 2*A*. The latter three proteins represent the known subfamilies of cytoplasmic tyrosine kinases and are distinguished from one another by well-defined structural features that confer a characteristic topology on each. In addition to conserved catalytic domains common to all tyrosine kinases, two conserved functional domains designated src homologous regions 2 and 3 (SH2 and SH3) have been identified in the c-abl and c-src proteins (for review, see ref. 1). An SH2 domain but not an SH3 domain has been identified in c-fps/fes. These domains are located amino-terminal to the kinase domains and appear to be involved in regulatory interactions. Comparison of the *arg* gene product with these three prototypic proteins showed that its structural architecture was strikingly similar to c-abl. arg and c-abl were readily distinguishable from the other cytoplasmic tyrosine kinases by virtue of their significantly larger sizes (1182 and 1148 amino acids, respectively) as well as the location of their catalytic domains amino-terminal to long carboxyl-terminal domains.

Comparison of the predicted amino acid sequences of arg and c-abl revealed a high degree of identity in their tyrosine kinase, SH2, and SH3 domains (Fig. 2*B*). Their tyrosine kinase domains were 94% identical, as compared with 52% and 44% for c-src and c-fes, respectively (numbers in the boxes, Fig. 2*A*). The major tyrosine residue phosphorylated *in vitro* in c-abl (22) was conserved in arg (arg position 439). The region in arg that extends amino-terminal from its catalytic domain to the position at which the similarity with c-abl breaks (arg position 73) was 85% identical to c-abl. The latter region contains the arg SH2 and SH3 domains, each of which were 90% identical to those of c-abl.

The carboxyl-terminal domains of arg and c-abl, with 652 and 645 amino acids, respectively, were 29% identical. Although the overall degree of similarity was weak, several conserved features were present. The greatest similarity in this domain was located at the extreme carboxyl terminus where the last 60 amino acids were 56% identical. A proposed nuclear localization sequence in murine c-abl consisting of five consecutive lysine residues (23) was completely conserved in human c-abl (underlined sequence, Fig. 2*B*) and was represented by three consecutive basic amino acids in arg (Lys-Lys-Arg). Several nuclear localization sequences consisting of three basic amino acids have been described (24–26), so it is possible that this sequence could be functional in arg. Two protein kinase C phosphorylation sites have been described in murine c-abl (27) and are completely conserved in the human protein (boxed sequences, Fig. 2*B*). One of these sites, along with several residues immediately bordering it, was completely conserved in arg. The second site was only partially conserved and is unlikely to be functional since the serine residue that is apparently phosphorylated in c-abl was not present in the arg sequence. An unusual structural feature of c-abl is a proline-rich (18%) region spanning amino acids 800–1000 (19). The analogous region of arg, spanning amino acids 848–1049, was also proline rich (16%). The location of these several conserved regions in a domain with otherwise poor similarity suggests important functions.

***arg* Is Expressed as Two Transcripts That Encode Proteins with Alternative Amino Termini.** Our analysis of 5′ *arg* cDNA clones demonstrated heterogeneity. Nucleotide sequence analysis of several clones revealed that they could be classified into two distinct types. The two classes of clones diverged from a common 3′ sequence at precisely the same nucleotide and contained one of two types of upstream sequence. The 5′ regions of two *arg* cDNA clones that contained alternative upstream sequences are shown in Fig. 3 *A* and *B*. cDNA clones A8 and MN2 contain open reading frames of 37 and 73 amino acids, respectively, that are joined in frame to the common open reading frame. The most upstream methionine residue was designated as the initiation codon in each of the cDNA clones. The assignment of the initiation codon in clone MN2 was strongly supported by the observation of significant amino acid similarity with c-abl (described below). However, in the absence of preceding in-frame stop codons in these clones, it is possible that further coding sequence lies upstream.

This finding raised the possibility that alternative splicing results in the synthesis of two distinct arg transcripts. When probes from the alternative 5′ sequences were analyzed by Northern blot hybridization, they both recognized ≈12-kb arg transcripts (Fig. 3*C*, lanes 2 and 4). Thus *arg* is expressed as two transcripts whose sizes are too similar to allow them to be distinguished by agarose gel electrophoresis and encode proteins containing alternative amino termini.

Human c-*abl* is expressed as two transcripts of 6 and 7 kb that arise by a process of alternative splicing of 5′ first exons (19). This generates two transcripts that diverge at their 5′ ends but share a common 3′ sequence and results in the synthesis of two proteins that differ only in their amino termini. A similar situation has also been described for murine c-*abl* (28). We have not formally shown by genomic analysis that the *arg* splicing event involves first exons; however, by analogy with c-*abl*, we have designated the alternative 5′ *arg* sequences IA and IB and assumed that they are joined to the *arg* second exon.

The comparison of the IB forms of arg and c-abl is shown in Fig. 2*B*. The similarity between the two proteins begins

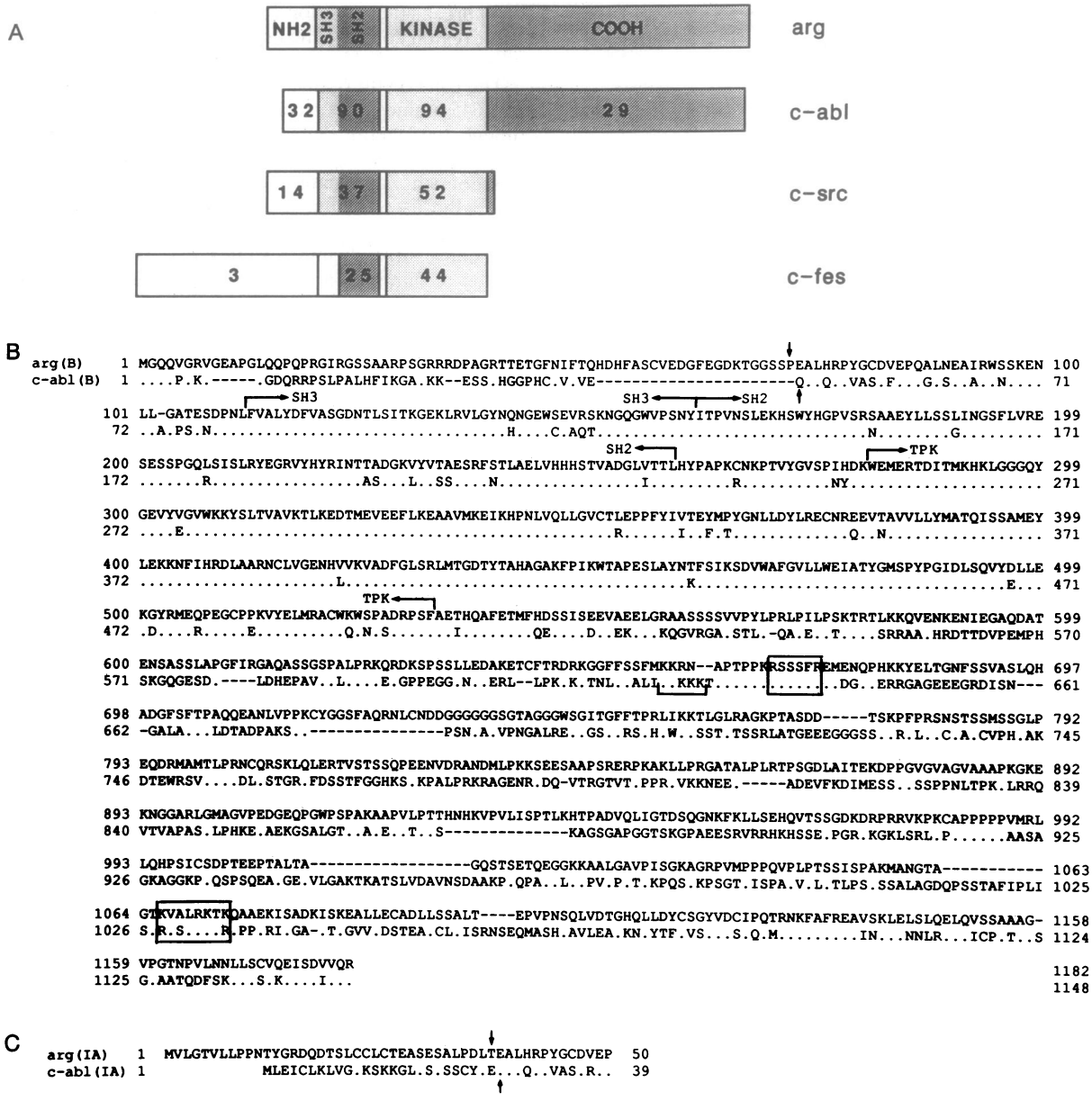

FIG. 2. (*A*) Schematic comparison of the *arg* gene product with the human c-abl (19), human c-src (20), and murine c-fes (21) proteins. The numbers indicate the percent amino acid identity with arg. An SH3 domain has not been defined for c-fes. (*B*) Comparison of the amino acid sequence of arg with human c-abl. The protein forms with IB amino termini are shown. Two suggested protein kinase C phosphorylation sites in c-abl are boxed, and a proposed nuclear localization sequence in c-abl is underlined. (*C*) Comparison of IA amino termini of arg and human c-abl. TPK, tyrosine phosphokinase domain. The vertical arrows indicate the splice junctions of the alternative amino termini. Periods indicate identical amino acids and dashes represent gaps in the alignment.

abruptly at the junctions of the respective alternative amino termini (denoted by arrows). While the IB amino termini have diverged extensively, six of the first eight amino-terminal amino acids were identical and two represented conservative changes. The conservation of this region is significant since the glycine residue at position 2 is myristoylated in c-abl (29), and this modification facilitates the association of proteins with cellular membranes (30–32). Although the comparison of the IA amino termini of arg with c-abl did not reveal significant similarity, they shared the absence of a myristoylation site (Fig. 3*C*).

***Drosophila abl* Is the Common Homolog of Both *arg* and c-*abl*.** The *Drosophila* Abelson homolog, *abl*, has been molecularly characterized and shown to encode a functional protein tyrosine kinase that is essential for development (33). The high degree of similarity between *arg* and c-*abl* raised the possibility that *abl* might be the homolog of *arg* as well. The nucleotide sequence of *abl* was 70% identical to c-*abl* over a 1.4-kb region, and the tyrosine kinase domains within this region were 74% identical. When *abl* was compared to *arg*, the *abl* nucleotide sequence was 66% identical to *arg* over the same 1.4-kb region as c-*abl*, and the tyrosine kinase domains shared 69% identity. The degree of similarity between *abl* and *arg* was thus only slightly less than that observed between *abl* and c-*abl*, strongly suggesting that *abl* was also the homolog of *arg*. To formally confirm this possibility, Southern blot analysis was performed. As expected, when analogous probes from the tyrosine kinase regions of *arg* and c-*abl* were hybridized under reduced stringency conditions to blots containing *Drosophila* DNA, identical *abl* restriction fragments were observed (data not shown). A schematic comparison of abl with arg and c-abl (Fig. 4) illustrates the structural similarities of the proteins. The tyrosine kinase domain and region containing the SH2 and SH3 domains of

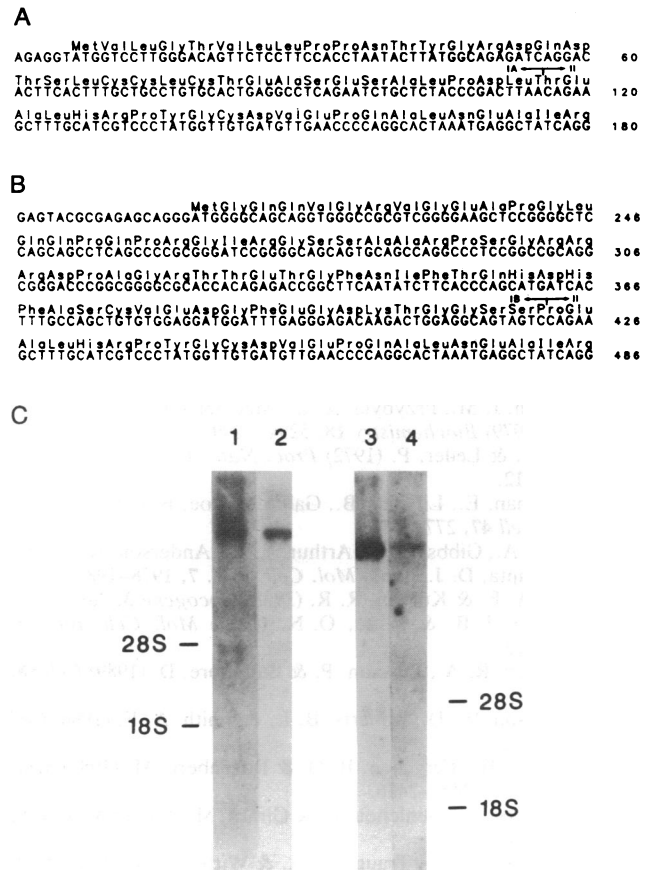

FIG. 3. Nucleotide sequence and predicted amino acid sequence of the 5′ regions of *arg* cDNA clones encoding alternative amino termini. (*A*) cDNA clone A8. (*B*) cDNA clone MN2. The arrows indicate the junctions of the divergent upstream sequences, IA and IB, respectively, with the putative second exon in the downstream open reading frame. (*C*) Hybridization of *arg* probes to total (lanes 1, 3, and 4) or poly(A)$^+$ (lane 2) RNA from A172 cells. A probe from the common *arg* sequence was used in lanes 1 and 3. Probes from the IA sequence of clone A8 or the IB sequence of clone MN2 were used in lane 2 and lane 4, respectively.

abl were about equally well conserved with those of arg (76% and 72%, respectively) and c-abl (78% and 75%, respectively). In contrast, the carboxyl- and amino-terminal domains of abl were extensively, but about equally, divergent from those of the two human proteins.

DISCUSSION

In this report we show that *arg* encodes a protein with features remarkably similar to c-abl. The high degree of nucleotide similarity between the tyrosine kinase domains of these two genes originally facilitated the identification of *arg*. Over the 250-amino acid stretch that comprises this domain, arg has only 16 amino acid differences with c-abl. The striking similarity of the proteins extends amino-terminal to include the SH2 and SH3 domains. Accumulating evidence indicates that these domains are involved in regulatory interactions (1). Mutations of the SH2 domain of v-*fps* and v-*src* render them defective in transformation, without affecting their *in vitro* tyrosine kinase activities. Moreover, these domains are found in proteins involved in signal transduction that are otherwise unrelated to protein tyrosine kinases. Phospholipase C and the transforming protein of v-crk have both SH2 and SH3 motifs, and GTPase activating protein (GAP) and spectrin have an SH2 and SH3 domain, respectively. Direct evidence that this region is involved in the regulation of c-*abl* is provided by the findings that small mutations in the SH3 domain can activate the myristoylated c-abl protein in both *in vitro* and *in vivo* transformation assays (29, 34). Because this regulatory region as well as the catalytic domains of arg and c-abl are so similar, distinct functions of the two proteins may be determined by their dissimilar amino and carboxyl termini.

The *arg* gene product, like that of c-*abl*, is expressed as two forms bearing different amino termini, one of which contains a myristoylation site. This feature distinguishes arg and c-abl from all other tyrosine kinases, which are either myristoylated proteins as in the case of the other cytoplasmic tyrosine kinases or transmembrane proteins in the case of growth factor receptors. The abrupt loss of similarity between arg and c-abl amino-terminal of the junctions of the alternative amino termini is striking. Aside from the conserved myristoylation sites, the analogous amino termini of arg and c-abl are quite distinct. The significance of this dissimilarity is indicated by the observation that the analogous amino termini of mouse and human c-abl are completely conserved in the case of the IA amino termini or conserved in 42 of 44 residues in the case of the IB polypeptide. Such rigid conservation suggests an important function. The distinct amino termini of arg and c-abl could be involved in specifying interactions with molecules that supply a triggering signal or possibly determining substrate interactions.

The transcripts encoding the two human c-abl proteins initiate at two separate promoters and contain alternative 5′ exons spliced to a common set of 3′ exons (19). This splicing process is unusual in that exon IB is located more than 200 kb away from the exon II acceptor site, and the intervening exon IA acceptor site must be skipped. In the chimeric *bcr/abl* transcript, the exon II acceptor site is linked to *bcr* exons located more than 100 kb upstream, and the intervening IA and possibly IB splice donor sites must be skipped. These unique splicing features may be associated with a stretch of pyrimidines located immediately upstream of the exon II splice acceptor site (19). A similar situation has been described for murine c-*abl* in which four transcripts arise by usage of two transcriptional promoters and alternative splicing (28). The genomic structure of *arg* has not yet been analyzed, but it will be interesting to determine whether alternative first exons are spliced over long distances and to characterize the exon II splice acceptor site.

Like their amino termini, the carboxyl termini of arg and c-abl are quite dissimilar. Although this region is not required for catalytic activity (35), several lines of evidence suggest

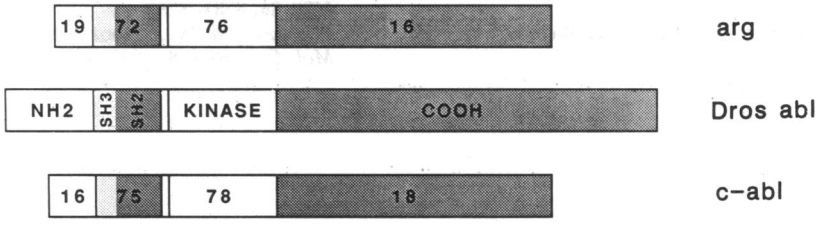

FIG. 4. Schematic comparison of the *Drosophila abl* gene product with the arg and human c-abl proteins. The numbers indicate the percent amino acid identity with abl.

that it can modify the function of abl proteins. Deleterious effects of v-abl on fibroblast growth have been observed, and carboxyl-terminal sequences seem to mediate this property (36). In addition, both of the two naturally occurring Abelson murine leukemia virus strains, p120 and p160, have mutations affecting this region (37, 38). Moreover, murine c-abl is phosphorylated in response to protein kinase C activators *in vivo* and by protein kinase C *in vitro* (27). One of these target sequences is completely conserved in the arg carboxyl terminus. Thus, the dissimilar carboxyl termini of arg and c-abl might function as regulatory regions that allow for distinct modulation of very similar catalytic activities.

The biological functions of the cytoplasmic tyrosine kinases are as yet largely undefined. Unlike growth factor receptors, their function cannot be analyzed by triggering them with specific ligands. For some cytoplasmic tyrosine kinases, however, expression restricted to specific cell types suggests involvement in particular differentiated functions. For example, the lymphocyte-specific protein, lck, may play a role in T-cell activation (39, 40), and c-fgr, a myelomonocytic-specific protein, may be involved in specific granule function (41). Both c-abl (42) and arg (G.D.K., unpublished observations), however, are widely expressed. Hence the signal transduction pathways that they function in are likely present in most cell types.

The subcellular location of a large fraction of the myristoylated murine c-abl protein was reported to be nuclear when overexpressed in NIH 3T3 cells, with the remainder in the cytoplasm and plasma membrane (23). In addition, a possible nuclear localization sequence was mapped to a region located carboxyl-terminal of the c-abl tyrosine kinase domain. These findings have suggested that one role of c-abl in signal transduction might be that of a nuclear protein kinase whose substrates could include molecules involved in growth regulation. The close relationship between arg and c-abl and our observation that the proposed c-abl nuclear localization sequence is partially conserved in arg suggest that arg might also have such a role.

We thank Dr. Matthias Kraus for excellent technical advice. R.P., who is partially supported by an Associazione Italiana per la Ricerca sul Cancro Fellowship, is on leave from the Institute of General Pathology and Consiglio Nazionale delle Ricerce Center for Research on Cellular Pathology, University of Milan.

1. Pawson, T. (1988) *Oncogene* **3**, 491–495.
2. Hanks, S. K., Quinn, A. M. & Hunter, T. (1988) *Science* **241**, 42–52.
3. Bishop, J. M. & Varmus, H. (1985) in *RNA Tumor Viruses*, eds. Weiss, R., Teich, N., Varmus, H. & Coffin, J. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), Vol. 2, pp. 249–356.
4. Shtivelman, E., Lifshitz, B., Gale, R. & Canaani, E. (1985) *Nature (London)* **315**, 550–554.
5. Kanopka, J. B., Watanabe, S. M. & Witte, O. N. (1984) *Cell* **37**, 1035–1042.
6. Mes-Masson, A. M., McLaughlin, J., Daley, G. Q., Paskind, M. & Witte, O. N. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9768–9772.
7. Sandberg, A., Gemmill, R., Hecht, B. & Hecht, F. (1986) *Cancer Genet. Cytogenet.* **21**, 129–146.
8. Kurzrock, R., Gutterman, J. & Talpaz, M. (1988) *N. Engl. J. Med.* **319**, 990–998.
9. Shields, A., Goff, S., Paskind, M., Otto, G. & Baltimore, D. (1980) *Cell* **18**, 955–962.
10. Srinivasan, A., Reddy, E. P. & Aaronson, S. A. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 2077–2081.
11. Besmer, P., Hardy, W. D., Jr., Zuckerman, E. E., Bergold, P., Lederman, L. & Snyder, H. W., Jr. (1983) *Nature (London)* **303**, 825–827.
12. Witte, O. N., Dasgupta, A. & Baltimore, D. (1980) *Nature (London)* **283**, 826–831.
13. Kruh, G. D., King, C. R., Kraus, M. H., Pepescu, N. C., Amsvaugh, S. C., McBride, W. O. & Aaronson, S. A. (1986) *Science* **234**, 1545–1548.
14. Miki, T., Matsui, T., Heidaran, M. A. & Aaronson, S. A. (1989) *Gene* **83**, 137–146.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
16. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
17. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
18. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
19. Shtivelman, E., Lifshitz, B., Gale, R., Roe, B. & Canaani, E. (1986) *Cell* **47**, 277–284.
20. Tanaka, A., Gibbs, C. P., Arthur, R. R., Anderson, K., Kung, H. & Fujita, D. J. (1987) *Mol. Cell. Biol.* **7**, 1978–1983.
21. Wilks, A. F. & Kurban, R. R. (1988) *Oncogene* **3**, 289–294.
22. Konopka, J. B. & Witte, O. N. (1985) *Mol. Cell. Biol.* **5**, 3116–3123.
23. Van Etten, R. A., Jackson, P. & Baltimore, D. (1989) *Cell* **58**, 669–678.
24. Richardson, W. D., Roberts, B. L. & Smith, A. E. (1986) *Cell* **44**, 77–85.
25. Lyons, R. H., Ferguson, B. Q. & Rosenberg, M. (1987) *Mol. Cell. Biol.* **7**, 2451–2456.
26. Wychowski, C., Benichou, D. & Girard, M. (1986) *EMBO J.* **5**, 2569–2576.
27. Pendergast, A. M., Traugh, J. A. & Witte, O. N. (1987) *Mol. Cell. Biol.* **7**, 4280–4289.
28. Ben-Neriah, Y., Bernards, A., Paskind, M., Daley, G. Q. & Baltimore, D. (1986) *Cell* **44**, 577–586.
29. Jackson, P. & Baltimore, D. (1989) *EMBO J.* **8**, 449–456.
30. Cross, F. R., Garber, E. A., Pellman, D. & Hanafusa, H. (1984) *Mol. Cell. Biol.* **4**, 1834–1842.
31. Goddard, C., Arnold, S. T. & Felsted, R. L. (1989) *J. Biol. Chem.* **264**, 15173–15176.
32. Gottlinger, H. G., Sodroski, J. G. & Haseltine, W. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 5781–5785.
33. Henkemeyer, M. J., Bennett, R. L., Gertler, F. B. & Hoffman, F. M. (1988) *Mol. Cell. Biol.* **8**, 843–853.
34. Franz, W. M., Berger, P. & Wang, Y. J. (1989) *EMBO J.* **8**, 137–147.
35. Srinivasan, A., Dunn, C. Y., Yuasa, Y., Devare, S. G., Reddy, E. P. & Aaronson, S. A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5508–5512.
36. Ziegler, S. F., Whitlock, C. A., Gold, S. P., Gifford, A. & Witte, O. N. (1981) *Cell* **27**, 477–486.
37. Reddy, E. P., Smith, M. J. & Srinivasan, A. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3623–3627.
38. Wang, J. Y. J., Ledley, F., Goff, S., Lee, R., Groner, Y. & Baltimore, D. (1984) *Cell* **36**, 349–356.
39. Rudd, C. E., Trevillyan, J. M., Dasgupta, J. D., Wong, L. L. & Schlossman, S. F. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5190–5194.
40. Veillette, A., Bookman, M. A., Horak, E. M. & Bolen, J. B. (1988) *Cell* **55**, 301–308.
41. Gutkind, J. S. & Robbins, K. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8783–8787.
42. Renshaw, M. W., Capozza, M. A. & Wang, J. Y. J. (1988) *Mol. Cell. Biol.* **8**, 4547–4551.