Volume 15 Number 2 1987 Nucleic Acids Research

The complete coding sequence of the human A-*raf*-1 oncogene and transforming activity of a human A-*raf* carrying retrovirus

T.W.Beck, M.Huleihel, M.Gunnell[1], T.I.Bonner[2] and U.R.Rapp

Laboratory of Viral Carcinogenesis, National Cancer Institute, Frederick, MD 21701, [1]Program Resources Inc., Frederick Cancer Research Facility, Frederick, MD 21701 and [2]Laboratory of Cell Biology, National Institute of Mental Health, Bethesda, MD 20205, USA

Received September 9, 1986; Revised and Accepted December 5, 1986

ABSTRACT

The complete 606 amino acid sequence of the human A-raf oncogene has been deduced from the 2453 nucleotide sequence of a human T cell cDNA. A cysteine-rich region located near the amino terminus, which is highly conserved in A-raf and c-raf, shows significant homology with protein kinase C. A 5' deleted fragment of the cDNA has been incorporated into a murine retrovirus which endows the virus with the ability to transform cells in vivo and in vitro. Functionally, human A-raf is similar to v-raf and v-mos in that transformation is independent of ras gene function.

INTRODUCTION

There are at least four raf-oncogene related genes in man (1,2). Two of these, c-raf-1 and A-raf-1, are active genes (1-8). c-raf-1 encodes a cytoplasmic protein which has an associated serine/threonine specific protein kinase activity similar to the gag-raf fusion protein of 3611-MSV in which the raf oncogene was originally discovered (4,5,7,9-12). Both raf proteins, c-raf and A-raf, appear to function in the signal transduction pathway of growth factors downstream of ras oncogenes (3,4,13). One other gene, c-raf-2 is an inactive pseudogene derived from c-raf-1 which has however gained clinical significance as a RFLP probe due to its chromosomal location on the short arm of chromosome 4, in the vicinity of the gene for Huntingtons Chorea (4). A-raf-1 has probably also given rise to a pseudogene(s), A-raf-2 which is located on human chromosome 7p14-q21 in the wider vicinity of the site implicated in cystic fibrosis (2,4). The complete coding sequence of c-raf-1 (14), the structure of the c-raf-1 gene (1,14) and of the c-raf protein (7,11,12,14) have recently been characterized.

Cytogenetic evidence (15-19), as well as expression studies

(4,8) and transfection of human tumor DNAs (20-23) have implicated c-raf-1 in a variety of common human and rodent carcinomas and, less frequently, other neoplasia. In the case of A-raf-1, a role in human pathology has not yet been established although incorporation of a 5' truncated mouse A-raf cDNA into a retrovirus genome has established its potential transforming ability (4).

A-raf was originally discovered by screening a mouse cDNA library with v-raf as probe and the sequence of an incomplete murine A-raf cDNA has been described (3). Here we report the complete coding sequence of human A-raf-1, its comparison with other raf family oncogenes and its potential transforming function.

MATERIALS AND METHODS

cDNA Libraries, Hybridization Conditions and DNA Sequencing.

Two human cDNA libraries, one from fetal liver (obtained from Dr. E.F. Fritsch, Genetics Institute) and one from T cells (obtained from Dr. Tak Mak, Ontario Cancer Institute) in lamda gt10 vectors were screened using a mouse A-raf (3) and subsequently a human A-raf cDNA as probe. Hybridizations were carried out in 3X SSC and 1X Denhardts at 60° C and the filters were washed in 1X SSC at 60° C as described previously (1,14). cDNA inserts were digested with restriction enzymes and subcloned into pUC 18/19 or M13mp18/19 vectors and sequenced by the dideoxynucleotide chain termination method (1,24).

Construction of Human A-raf-MSV.

The human A-raf-MSV construct virus was prepared as described previously (3). Briefly, a 1580 nucleotide Nco I/ Eco RI fragment of the human A-raf cDNA (nucleotides 227 to 2452) encoding amino acids 227 to 606 was inserted in-frame into the Xho I site of Moloney murine leukemia virus (strain Leuk) DNA by blunt end ligation.

DNA Transfections and Immunoprecipitations.

Transfection of plasmid DNA onto NIH 3T3 cells was carried out by calcium phosphate precipitation as described previously (3). Cells were labelled, lysed and immunoprecipitated as described previously (1,3).

RESULTS

Isolation and Sequencing of human A-raf cDNAs.

A 1.6 kb murine A-raf cDNA was used to screen a human fetal liver cDNA library which yielded a clone with a 1.92 kb cDNA insert. Since the size of the most common A-raf hybridizing transcript was known to be 2.6 kb, a 350 bp fragment from the 5' end of the 1.9 kb fetal liver cDNA was used as a probe to screen a second cDNA library from human T cells. In this second screen a cDNA insert of 2.45 kb was obtained which presumably represents a full length copy since addition of a poly A tail of 100-200 nucleotides would yield the apparent size of the A-raf mRNA (2.6 kb). Both cDNAs were sequenced by the dideoxynucleotide chain termination method and the resulting cDNA sequences and the predicted amino acid sequences are shown in figure 1. The T cell cDNA is 2452 nucleotides in length with a potential initiation codon located at nucleotide 201. This ATG is preceeded by termination codons in all three reading frames. The next termination codon (TTA) is at nucleotide 2019 giving an open reading frame of 1818 nucleotides. Therefore the predicted protein is 606 amino acids in length and has a calculated molecular weight of 67,530. Neither a poly(A) tail nor a polyadenylation signal (AATAAA) is contained in the A-raf cDNA insert. Furthermore, the 3'end of both clones contain an Eco RI site which is not preceded by the CCC of the linkers used in making the library. This suggests that both clones end with an Eco RI site which is located slightly 5' to the poly A addition site. The two cDNA sequences (T cell and fetal liver) differ by 4 nucleotide substitutions (nucleotide 653, G to A; 656, C to G; 1070, G to A; 1297, T to C) which result in 2 changes in the predicted amino acid sequence (amino acid 154, phe to leu; 368, leu to pro). These differences may well represent reverse transcription errors, however, if they are in fact correct, one potentially advantageous consequence would be restriction site polymorphisms for Rsa 1 (at nucleotide 656) and Mbo 2 (at nucleotide 1070).

Amino acid homology with raf family oncogenes and other kinases.

Figure 2 shows a comparison of the amino acid sequences between the raf family oncogenes. The human and mouse A-raf amino acid sequence show 94 percent identity in 438 overlapping

```
                                                                                                                    119
TGACCCAATAAGGGTGGAAGGCTGAGTCCCGCAGAGCCAATAACGAGAGTCCGAGAGGCGACGGAGGCGGACTCTGTGAGGAAACAAGAAGAGAGGCCCAAGATGGAGACGGCGGCGGC

                                                                                                                    227
TGTAGCGGCGTGACAGGAGCCCCATGGCACCTGCCCAGCCCCACCTCAGCCCATCTTGACAAAATCTAAGGCTCC ATG GAG CCA CCA CGG GGC CCC CCT GCC AAT GGG
                                                                            MET Glu Pro Pro Arg Gly Pro Pro Ala Asn Gly
                                                                            1                                       10
                                                                                                                    317
GCC GAG CCA TCC CGG GCA GTG GGC ACC GTC AAA GTA TAC CTG CCC AAC AAG CAA CGC ACG GTG GTG ACT GTC CGG GAT GGC ATG AGT GTC
Ala Glu Pro Ser Arg Ala Val Gly Thr Val Lys Val Tyr Leu Pro Asn Lys Gln Arg Thr Val Val Thr Val Arg Asp Gly MET Ser Val
                                    20                                      30                                      40
                                                                                                                    407
TAC GAC TCT CTA GAC AAG GCC CTG AAG GTG CGG GGT CTA AAT CAG GAC TGC TGT GTG GTC TAC CGA CTC ATC AAG GGA CGA AAG ACG GTC
Tyr Asp Ser Leu Asp Lys Ala Leu Lys Val Arg Gly Leu Asn Gln Asp Cys Cys Val Val Tyr Arg Leu Ile Lys Gly Arg Lys Thr Val
                                    50                                      60                                      70
                                                                                                                    497
ACT GCC TGG GAC ACA GCC ATT GCT CCC CTG GAT GGC GAG GAG CTC ATT GTC GAG GTC CTT GAA GAT GTC CCG CTG ACC ATG CAC AAT TTT
Thr Ala Trp Asp Thr Ala Ile Ala Pro Leu Asp Gly Glu Glu Leu Ile Val Glu Val Leu Glu Asp Val Pro Leu Thr MET His Asn Phe
                                    80                                      90                                      100
                                                                      M+ / NA                                       587
GTA CGG AAG ACC TTC TTC AGC CTG GCG TTC TGT GAC TTC TGC CTT AAG TTT CTG TTC CAT GGC TTC CGT TGC CAA ACC TGT GGC TAC AAG
Val Arg Lys Thr Phe Phe Ser Leu Ala Phe Cys Asp Phe Cys Leu Lys Phe Leu Phe His Gly Phe Arg Cys Gln Thr Cys Gly Tyr Lys
                                    110                                     120                                     130
                                                                                                                    677
TTC CAC CAG CAT TGT TCC TCC AAG GTC CCC ACA GTC TGT GTT GAC ATG AGT ACC AAC CGC CAA CAG TTC TAC CAC AGT GTC CAG GAT TTG
Phe His Gln His Cys Ser Ser Lys Val Pro Thr Val Cys Val Asp MET Ser Thr Asn Arg Gln Gln Phe Tyr His Ser Val Gln Asp Leu
                                    140                                     150                                     160
                                                                                                                    767
TCC GGA GGC TCC AGA CAG CAT GAG GCT CCC TCG AAC CGC CCC CTG AAT GAG TTG CTA ACC CCC CAG GGT CCC AGC CCC CGC ACC CAG CAC
Ser Gly Gly Ser Arg Gln His Glu Ala Pro Ser Asn Arg Pro Leu Asn Glu Leu Leu Thr Pro Gln Gly Pro Ser Pro Arg Thr Gln His
                                    170                                     180                                     190
                                                                                                                    857
TGT GAC CCG GAG CAC TTC CCC TTC CCT GCC CCA GCC AAT GCC CCC CTA CAG CGC ATC CGC TCC ACG TCC ACT CCC AAC GTC CAT ATG GTC
Cys Asp Pro Glu His Phe Pro Phe Pro Ala Pro Ala Asn Ala Pro Leu Gln Arg Ile Arg Ser Thr Ser Thr Pro Asn Val His MET Val
                                    200                                     210                                     220
                                                                                                                    947
AGC ACC ACG GCC CCC ATG GAC TCC AAC CTC ATC CAG CTC ACT GGC CAG AGT TTC AGC ACT GAT GCT GCC GGT AGT AGA GGA GGT AGT GAT
Ser Thr Thr Ala Pro MET Asp Ser Asn Leu Ile Gln Leu Thr Gly Gln Ser Phe Ser Thr Asp Ala Ala Gly Ser Arg Gly Gly Ser Asp
                                    230                                     240                                     250
                                                                                                                   1037
GGA ACC CCC CGG GGG AGC CCC AGC CCA GCC AGC GTG TCC TCG GGG AGG AAG TCC CCA CAT TCC AAG TCA CCA GCA GAG CAG CGC GAG CGG
Gly Thr Pro Arg Gly Ser Pro Ser Pro Ala Ser Val Ser Ser Gly Arg Lys Ser Pro His Ser Lys Ser Pro Ala Glu Gln Arg Glu Arg
                                    260                                     270                                     280
                                                                                                                   1127
AAG TCC TTG GCC GAT GAC AAG AAG AAA GTG AAG AAC CTG GGG TAC CGG GAY TCA GGC TAT TAC TGG GAG GTA CCA CCC AGT GAG GTG CAG
Lys Ser Leu Ala Asp Asp Lys Lys Lys Val Lys Asn Leu Gly Tyr Arg Asp Ser Gly Tyr Tyr Trp Glu Val Pro Pro Ser Glu Val Gln
                                    290                                     300                                     310
                     *       *           *                                                         ↓                   1217
CTG CTG AAG AGG ATC GGG ACG GGC TCG TTT GGC ACC GTG TTT CGA GGG CGG TGG CAT GGC GAT GTG GCC GTG AAG GTG CTC AAG GTG TCC
Leu Leu Lys Arg Ile Gly Thr Gly Ser Phe Gly Thr Val Phe Arg Gly Arg Trp His Gly Asp Val Ala Val Lys Val Leu Lys Val Ser
                                    320                                     330                                     340
                                                                                                                   1307
CAG CCC ACA GCT GAG CAG GCC CAG GCT TTC AAG AAT GAG ATG CAG GTG CTC AGG AAG ACG CGA CAT GTC AAC ATC TTG CTG TTT ATG GGC
Gln Pro Thr Ala Glu Gln Ala Gln Ala Phe Lys Asn Glu MET Gln Val Leu Arg Lys Thr Arg His Val Asn Ile Leu Leu Phe MET Gly
                                    350                                     360                                     370
                                                                                                                   1397
TTC ATG ACC CGG CCG GGA TTT GCC ATC ATC ACA CAG TGG TGT GAG GGC TCC AGC CTC TAC CAT CAC CTG CAT GTG GCC GAC ACA CGC TTC
Phe MET Thr Arg Pro Gly Phe Ala Ile Ile Thr Gln Trp Cys Glu Gly Ser Ser Leu Tyr His His Leu His Val Ala Asp Thr Arg Phe
                                    380                                     390                                     400
                                                                                                                   1487
GAC ATG GTC CAG CTC ATC GAC GTG GCC CGG CAG ACT GCC CAG GGC ATG GAC TAC CTC CAT GCC AAG AAC ATC ATC CAC CGA GAT CTC AAG
Asp MET Val Gln Leu Ile Asp Val Ala Arg Gln Thr Ala Gln Gly MET Asp Tyr Leu His Ala Lys Asn Ile Ile His Arg Asp Leu Lys
                                    410                                     420                                     430
                                                                                                                 ■ 1577
TCT AAC AAC ATC TTC CTA CAT GAG GGG CTC ACG GTG AAG ATC GGT GAC TTT GGC TTG GCC ACA GTG AAG ACT CGA TGG AGC GGG GCC CAG
Ser Asn Asn Ile Phe Leu His Glu Gly Leu Thr Val Lys Ile Gly Asp Phe Gly Leu Ala Thr Val Lys Thr Arg Trp Ser Gly Ala Gln
                                    440                                     450                                     460
                                                  +   +   +                                                        1667
CCC TTG GAG CAG CCC TCA GGA TCT GTG CTG TGG ATG GCA GCT GAG GTG ATC CGT ATG CAG GAC CCG AAC CCC TAC AGC TTC CAG TCA GAC
Pro Leu Glu Gln Pro Ser Gly Ser Val Leu Trp MET Ala Ala Glu Val Ile Arg MET Gln Asp Pro Asn Pro Tyr Ser Phe Gln Ser Asp
                                    470                                     480                                     490
                                                                                                                   1757
GTC TAT GCC TAC GGG GTT GTG CTC TAC GAG CTT ATG ACT GGC TCA CTG CCT TAC AGC CAC ATT GGC TGC CGT GAC CAG ATT ATC TTT ATG
Val Tyr Ala Tyr Gly Val Val Leu Tyr Glu Leu MET Thr Gly Ser Leu Pro Tyr Ser His Ile Gly Cys Arg Asp Gln Ile Ile Phe MET
                                    500                                     510                                     520
                                                                                                                   1847
GTG GGC CGT GGC TAT CTG TCC CCG GAC CTC AGC AAA ATC TCC AGC AAC TGC CCC AAG GCC ATG CGG CGC CTG CTG TCT GAC TGC CTC AAG
Val Gly Arg Gly Tyr Leu Ser Pro Asp Leu Ser Lys Ile Ser Ser Asn Cys Pro Lys Ala MET Arg Arg Leu Leu Ser Asp Cys Leu Lys
                                    530                                     540                                     550
                                                                                                                   1937
TTC CAG CGG GAG GAG CGG CCC CTC TTC CCC CAG ATC CTG GCC ACA ATT GAG CTG CTG CAA CGG TCA CTC CCC AAG ATT GAG CGG AGT GCC
Phe Gln Arg Glu Glu Arg Pro Leu Phe Pro Gln Ile Leu Ala Thr Ile Glu Leu Leu Gln Arg Ser Leu Pro Lys Ile Glu Arg Ser Ala
                                    560                                     570                                     580
                                                                                                                   2030
TCG GAA CCC TCC TTG CAC CGC ACC CAG GCC GAT GAG TTG CCT GCC TGC CTA CTC AGC GCA GCC CGC CTT GTG CCT TAG GCCCCGCCCAAGCCA
Ser Glu Pro Ser Leu His Arg Thr Gln Ala Asp Glu Leu Pro Ala Cys Leu Leu Ser Ala Ala Arg Leu Val Pro
                                    590                                     600                     606
```

```
                                                                    2117
CCAGGGAGCCAATCTCAGCCCTCCACGCCAAGGAGCCTTGCCCACCAGCCAATCAATGTTCGTCTCTGCCCTGATGCTGCCTCAGGATCCCCCATTCCCCACCCTGGGAGATGAGGGGG

                                                                    2268
TCCCCATGTGCTTTTCCAGTTCTTCTGGAATTGGGGGACCCCCGCCAAAGACTGAGCCCCCTGTCTCCTCCATCATTTGGTTTCCTCTTGGCTTTGGGGATACTTCTAAATTTTGGGAG

                                                                    2387
CTCCTCCATCTCCAATGGCTGGGATTTGTGGCAGGGATTCCACTCAGAACCTCTCTGGAATTTGTGCCTGATGTGCCTTCCACTGGATTTTGGGGTTCCCAGCACCCCATGTGGATTTT

GGGGGGTCCCTTTTGTGTCTCCCCCGCCATTCAAGGACTCCTCTCTTTCTTCACCAAGAAGCACAGAATTC
```

Figure 1. Human A-raf cDNA and derived amino acid sequence. Nucleotides are numbered to the right of each line and amino acids below the corresponding residue. Sequences constituting the proposed metal/nucleic acid binding domain, the ATP binding domain and that corresponding to the major phosphorylation site of src are overlined. The G-X-G-X-X-G sequence is indicated by asterisks, the conserved lysine by an arrow, the APE sequence by plus signs, and the position corresponding to Tyr 416 of v-src by a box.

amino acids. Taking into acount conservative amino acid changes the homology is 98 percent. Comparison of human c-raf-1 to human A-raf indicates there is 60 percent identity over 604 overlapping amino acids with the most extensive homology in the carboxy-terminal half (kinase domain) of the molecule (75 percent). Considering conservative amino acid substitutions these two proteins (human A-raf and c-raf) are 89 percent homologous. The amino acid sequence of the viral versions of raf (v-raf and v-mil) are more closely related to c-raf-1 than either murine or human A-raf indicating that they were transduced from c-raf-1 rather than A-raf. Outside the kinase domain, there are two blocks of significant homology located between amino acids 61 to 192 and 253 to 275 of c-raf-1. The former block containing the cysteine-rich sequence C-X2-C-X9-C-X2-C is of the form recently proposed to be a characteristic of metal-binding and/or nucleic acid binding domains (25).

Recently, another protein kinase having specificity for serine and threonine residues has been found to contain a similar cysteine-rich sequence (C-X2-C-X13-C-X2-C) as a tandem repeat (26-28). Figure 3 shows an amino acid sequence comparison of raf proteins (human A-raf and c-raf) and the multiple species of protein kinase C (alpha, beta and gamma). The homology between raf and protein kinase C extends over 46 amino acids (excluding a 4 amino acid insertion to maximize the alignment, roughly one turn of an alpha helix) and shows 30 percent or more identity. Com-

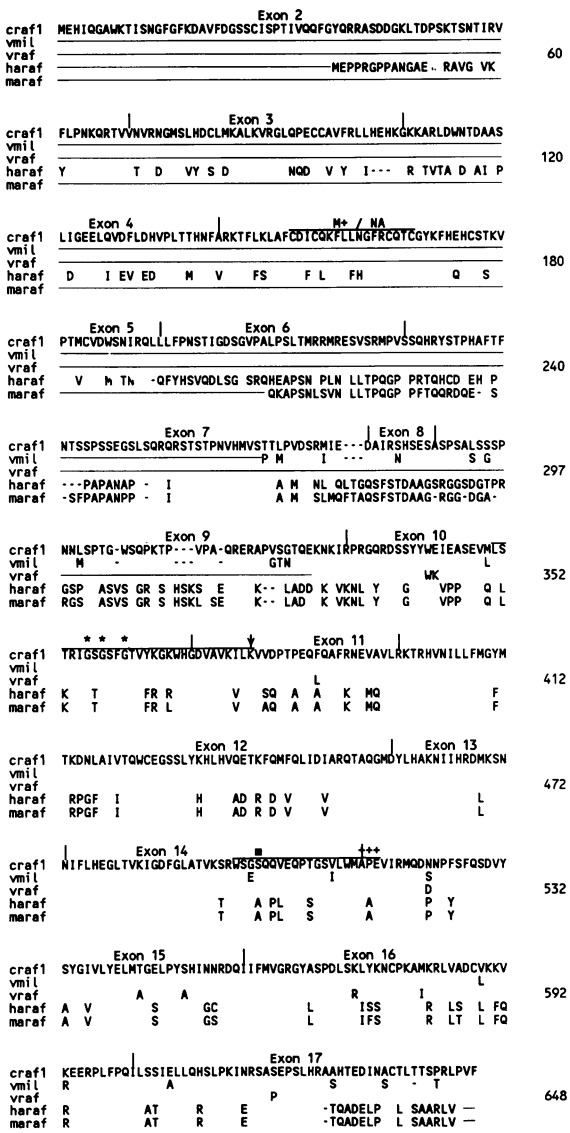

Figure 2. Amino acid sequence homology among raf family oncogenes. Solid lines denote truncated sequences, dashed lines indicate gaps introduced to align sequences, open spaces represent amino acid identities with respect to c-raf-1, and vertical lines demarcate the exon structure of c-raf-1. The single letter amino acid abbreviations used are A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; and Y, tyrosine.

```
HNFARKTFLKLAFCDICQKFL----LNGFRCQTCGYKFHEHCSTKVPTMC   CRAF1 Amino Acids 139 - 184
  v    fs    f L    ----fh          q   s    v      HARAF Amino Acids  99 - 144
 k KIh yGsPt   H Gsl YGLIHq mk d  dMnV kq Vin  sl    BPKCA Amino Acids 102 - 151
 k KIh ySsPt   H Gsl YGLIHq mk d  MMnV kr VMn  sl    BPKCB Amino Acids 102 - 151
 k RLhsySsPt   H Gsl YGLVHq mk SC eMnV Rr VRd  sl    BPKCG Amino Acids 101 - 150
 r IArFFKqPt  sH Td IWGFGKq  q  v CfVV kr HeF tFS    BPKCA Amino Acids  37 -  86
 k tArFFKqPt  sH Td IWGFGKq  q  v CfVV kr HeF tFS    BPKCB Amino Acids  37 -  86
 k tArFFKqPt  sH Td IWGIGKq lq  v sfVV Rr HeF tFE    BPKCG Amino Acids  36 -  85
```

Figure 3. Comparison of the putative metal/nucleic acid binding domains of human A-raf-1, c-raf-1 and bovine protein kinase C alpha, beta and gamma (ref. 27). Amino acid identities with respect to human A-raf-1 are indicated by open spaces, conservative amino acid differences by lower case letters, nonconservative differences by upper case letters and dashed lines as in figure 2.

parison to src family oncogenes indicates that A-raf and c-src are related only in the carboxy terminal half of the molecule with 31 percent identity in 229 overlapping amino acids similar to c-raf (33 percent identity relative to c-src). Similarly low homology scores were obtained in amino acid sequence comparisons with other protein kinases in the National Biomedical Research Foundation protein data bank. Functional regions that are highly conserved in the src family oncogenes are also conserved within the raf family (29). The putative ATP binding domain which includes the G-X-G-X-X-G sequence and the V-A-X-K sequence which is essential for transformation by src (30) is located between amino acids 351 and 379 of c-raf-1. In this region A-raf shows 21/28 identical amino acids relative to c-raf-1. In the domain corresponding to the phosphotyrosine acceptor site of src where the ser/thr specific protein kinases diverge from the tyrosine specific protein kinases (amino acids 505 to 525 of c-raf-1), A-raf has 15/20 amino acid identities relative to c-raf-1 (3/20 relative to c-src). The APE sequence which is essential for transformation by v-src (31) is changed to AAE in human and mouse A-raf, but not in c-raf-1. Another oncogene having protein kinase activity in which the APE site was not conserved is v-erb B which contains the sequence ALE (32).

Potential Transforming Function of human A-raf-1.

To test the transforming potential of human A-raf a recombinant murine sarcoma virus was constructed. A 1.59 kb fragment of human A-raf cDNA encoding amino acids 227 to 606 was incorporated such that the A-raf cDNA would be expressed as a gag-A-raf fusion protein. The specific transforming activity of A-raf-MSV DNA is 6

TABLE 1. Transforming Efficiency of Human A-raf-MSV.

Transfected DNA	Transformation Efficiency FFU/ug of DNA
-	0
Leuk (helper)	0
Moloney MSV (v-mos) + Leuk	10000
3611-MSV (v-raf) + Leuk	4000
hA-raf-MSV + Leuk	6300

FFU - Focus forming units

X 10^3 FFU/ug similar to that of 3611-MSV DNA indicating that point mutations in A-raf DNA are not required for its transforming activity in the context of this virus (Table 1).

To verify that the a gag-A-raf fusion protein was synthesized, A-raf-MSV transformed and uninfected control cells were labelled with ^{35}S-methionine or ^{3}H-myristic acid and cell lysates were immunoprecipitated with either anti-p15 gag or anti-A-raf antisera (figure 4). Both antisera precipitate proteins of 80 kd and 90 kd proteins in S-methionine labelled A-raf-MSV transformed cells. In contrast, precipitation of lysates from H-myristic acid labelled A-raf-MSV transformed cells only identify the 80 kd protein. Neither the 80 kd, the 90 kd nor the expected normal size A-raf protein were precipitated from uninfected cells by these antisera (data not shown). In analogy with the findings with pp79 and gp90 of 3611-MSV, we infer that the 90 kd protein is the glycosylated form of the 80 kd protein which is instead myristilated and phosphorylated (4).

Stocks of infectious human A-raf-MSV were generated by cotransfection of construct DNA and helper virus DNA (leuk strain of MoLV). Supernatants containing 10^4 focus forming units per ml of culture fluid were injected intraparitoneally into newborn NFS/N mice and resulted in the rapid developement of sarcomas and erythroleukemias with a mean latency of 9 weeks (data not shown).

To test whether human A-raf-MSV would overcome the ras antibody block of DNA synthesis in microinjected NIH 3T3 cells as has recently been demonstrated with 3611-MSV (3,13), human A-raf-MSV transformed NIH 3T3 cells were microinjected with ras monoclonal antibody Y13-259. NIH 3T3/A-raf-MSV cells are as insensi-

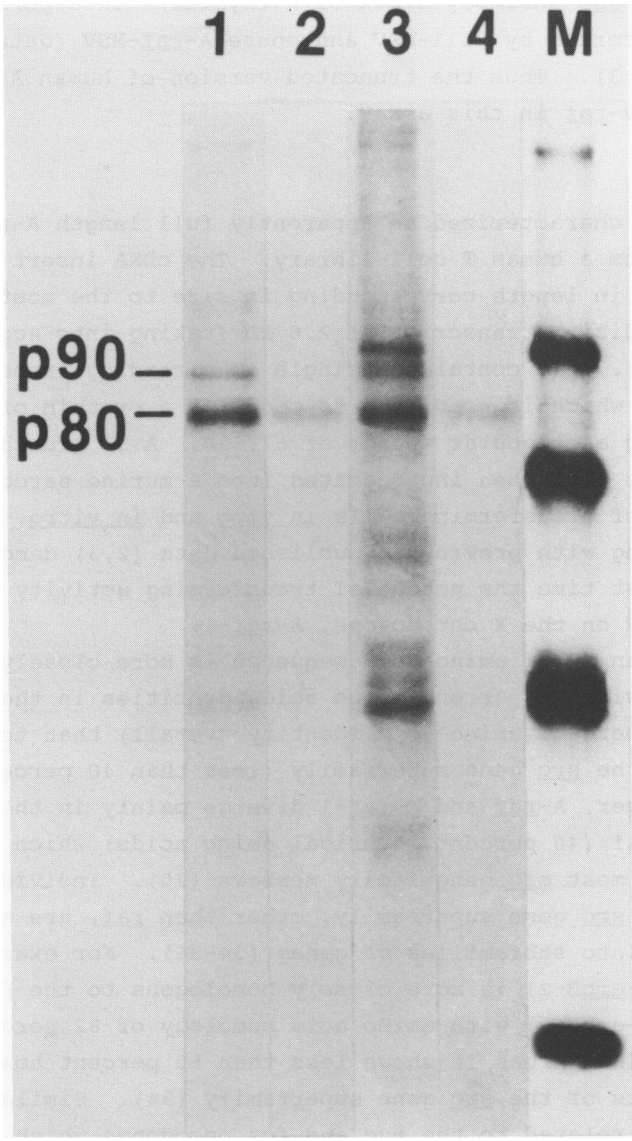

Figure 4. Human A-raf-gag fusion proteins expressed in transformed fibroblasts. Human A-raf-MSV transformed cells were metabolically labelled with ^{35}S-methionine (lanes 1 and 3) or ^{3}H-myristic acid (lanes 2 and 4) and immunoprecipitated with anti-p15gag (lanes 1 and 2) or anti-A-raf (lanes 3 and 4) antisera as described in Materials and Methods. The molecular weight markers are myosin, 200 kd; phosphorylase B, 97.4 kd; albumin, 69 kd; ovalbumin, 46 kd; carbonic anhydrase, 30 kd.

tive to the ras antibody block of H-thymidine incorporation as cells transformed by 3611-MSV and mouse A-raf-MSV (data not shown; ref. 3). Thus the truncated version of human A-raf functions like v-raf in this assay.

DISCUSSION

We have characterized an apparently full length A-raf cDNA isolated from a human T cell library. The cDNA insert is 2453 nucleotides in length corresponding in size to the most common A-raf hybridizing transcript of 2.6 kb (taking into account the poly A tail). It contains a single open reading frame of 1818 nucleotides which is predicted to code for a protein of 606 amino acids having a molecular weight of 67,530. A 5' truncated version of this cDNA when incorporated into a murine sarcoma virus is capable of transforming cells in vivo and in vitro. These results along with previously published data (2,3) demonstrate for the first time the potential transforming activity of a human gene located on the X chromosome, A-raf-1.

The human A-raf amino acid sequence is more closely related to c-raf-1 with 75 percent amino acid identities in the kinase domain (60 percent amino acid identity overall) than to other members of the src gene superfamily (less than 40 percent homology). However, A-raf and c-raf-1 diverge mainly in the amino terminal half (46 percent identical amino acids) which is characteristic of most src gene family members (30). Individual members of the src gene superfamily, other than raf, are also developing into subfamilies of genes (34-36). For example, the neu gene (c-erbB-2) is more closely homologous to the EGF receptor (c-erbB-1) with amino acid homology of 82 percent in the kinase domain whereas it shows less than 50 percent homology with other members of the src gene superfamily (34). Similarly, some new members related to the myc and fos oncogenes which function in the nucleus have recently been identified (37-39).

Although human A-raf and c-raf-1 are structurally similar and share the potential to transform NIH 3T3 cells, A-raf appears to be expressed only in a restricted set of tissues (with highest levels in the epididymis), whereas c-raf-1 expression is more generalized (4,5). This indicates that the expression of A-raf-1

and c-raf-1 mRNA is differentially and independently regulated at the level of transcription or RNA processing. Similarly, differential expression of myc subfamily members in specialized tissues has recently been reported (40).

Additional controls on raf activity may be exerted at the protein level. We have recently proposed that the N-terminal half of the raf protein represents a regulatory domain which if altered or removed would activate the kinase domain and transforming potential of the protein (4). Activation of the full size raf protein (67.5 kd A-raf-1 or 74 kd c-raf-1) might be achieved by the binding of effector molecules to a regulatory domain. An interesting model for the normal regulation of raf enzyme activity may be protein kinase C, another serine/threonine specific enzyme having a molecular weight of approximately 76 kd. The binding of effectors (Ca^2 and diacyl-glycerol) to this enzyme activates the C-terminal kinase domain (41). Although the binding sites for the effectors have not yet been defined, it is interesting that A-raf and c-raf share homology with the tandemly repeated (N-terminal) metal/nucleic acid binding domains of protein kinase C (26-28). Moreover, the homology suggests that this region may play a common functional role in these proteins.

Another kinase which is activated via the binding of effector molecules is cGMP-dependent protein kinase (cGK, ref. 42). This protein exists as a dimer of two identical 76 kd subunits, binding of cGMP to the N-terminal regulatory domain causes a conformation change which activates the C-terminal kinase domain. Comparison of cGK to A-raf reveals a low degree of homology in the N-terminal half of the molecule. An alternative model for raf activation is the double-stranded (ds) RNA dependent protein kinase (43). This protein exists as a dimer of two subunits having molecular weights of 48 and 68 kd. In the presence of dsRNA, the 48 kd subunit phosphorylates the 68 kd subunit which is converted to an active protein kinase capable of phosphorylating exogenous substrates. Whether activation occurs by the binding of a ligand to or phosphorylation of the putative regulatory domain, the apparent amino acid sequence divergence in the N-terminal half suggests that A-raf-1 and c-raf-1 proteins may be regulated independently. This suggests, along with the

expression studies (3), that A-raf plays a unique physiological role in certain tissues.

The location of A-raf on the X chromosome and the demonstration of its potential transforming activity suggests a possible role for A-raf in tumors having a sex-linked mode of inheritance. A recessive tumorigenic allele on the X chromosome would be masked in females and dominant in males such as is the case with certain X-linked lymphoproliferative diseases (44-47) and possibly hairy cell leukemia (48).

In a functional assay where virus transformed NIH 3T3 cells are tested for their ability to overcome a block in DNA synthesis by microinjected anti-ras antibody, human A-raf-MSV, mouse A-raf-MSV and 3611-MSV transformed fibroblasts are able to overcome the antibody block as well as v-mos(3,13). However, fibroblasts transformed with other oncogenes, v-sis, v-fms, v-fes and v-src are unable to enter S phase in the presence of anti-ras antibody (13). Moreover, in a second type of functional assay, morphologically normal revertants of Ki-ras transformed cells (C11our) were found to suppress transformation by v-Ki-ras, v-Ha-ras, v-fes and v-src, whereas mouse A-raf-MSV and 3611-MSV were found to be transforming as well as v-mos (4,49). Together, these results in conjunction with the cytoplasmic location and ser/thr kinase activity of raf suggest that raf family oncogenes act independent of ras either through a signal transduction pathway not involving ras or one in which raf has a position downstream of ras. In this context, knowledge of the complete amino acid sequence of human A-raf will allow the development of A-raf-1 specific antisera which are necessary, together with c-raf-1 specific antisera to determine by antibody microinjection where raf functions in signal transduction.

ACKNOWLEDGEMENTS

We thank Patricia Borchert, Stephen Kerby and Robert Nalewaik for excellent technical assistance, M.R. Smith and D.W. Stacey for the microinjection work, and Walter Kolch for helpful discussions.

This project was funded in part by the Department of Health and Human Services under contract no. NO1-CO-23910 with Program Resources, Inc.

REFERENCES

1. Bonner, T.I., Kerby, S.B., Sutrave, P. Gunnell, M.A., Mark, G.and Rapp, U.R. (1985) Mol. Cell. Biol. 5, 1400-1407.
2. Huebner, K., ar-Rushdi, A., Griffin, C., Isobe, M., Kozak, C., Emmanuel, L. Nagarajan, Cleveland, J., Bonner, T., Goldsborough, M., Croce, C. and Rapp, U.R. (1986) Proc. Natl. Acad. Sci. USA 83, 3934-3938.
3. Huleihel, M., Goldsborough, M., Cleveland, J., Gunnell, M. Bonner, T. and Rapp, U.R. (1986) Mol. Cell. Biol. 6, 2655-2662.
4. Rapp, U.R., Cleveland, J. and Bonner, T.I. In Handbook of Oncogenes, Reddy, P., Curran, T. and Skalka, A., eds., Elsevier, New York. (in press).
5. Rapp, U.R., Bonner, T.I., Mqelling, K., Jansen, H.W., Bister, K. and Ihle, J. (1985) in Recent Results in Cancer Research, Vol. 99, Haveman, K. and Sorenson, G., eds., pp 221-236, New York.
6. Rapp, U.R., Bonner, T.I., and Cleveland, J.L. (1986) in Retroviruses and Human Pathology, Gallo, R.C., Stehelin, D. and Varnier, eds., The Humana Press, New York.
7. Cleveland, J.L., Jansen, H.W., Bister, K., Frederickson, T.N., Morse, H.C., Ihle, J.N., and Rapp, U.R. (1986) J. Cell. Biochem. 30, 195-218.
8. Rapp, U.R., Cleveland, J.L., and Storm, S. (1986) in Modern Trends in Human Leukemia VII, Haematology and Blood Transfusion (in press).
9. Sutrave, P., Bonner, T.I., Rapp, U.R., Jansen, H.W., Patchinsky, T. and Bister, K. (1984) Nature 309, 85-88.
10. Moelling, K., Heimann, B., Beinling, P., Rapp, U.R., and Sander, T. (1984) Nature 312, 558-561.
11. Molders, H, Defesche, I., Muller, D., Bonner, T.I., Rapp, U.R. and Muller, R. (1985) Embo J. 4, 693-698.
12. Schultz, A.M., Oroszlan, S., Mueller, R. Moelling, K. and Rapp, U.R. Mol. Cell. Biol. (submitted).
13. Smith, M. DeGudicibus, S. and Stacey, D. (1986) Nature 320, 540-543.
14. Bonner, T.I., Oppermann, H., Seeburg, P., Kerby, S.B., Gunnell, M.A., Young, A.C. and Rapp, U.R. (1986) Nucl. Acids Res. 14, 1009-1015.
15. Whang-Peng, J., Bunn, P.A., Kao-Shan, C.S., Lee, E.C., Carney, D.N., Gazdar, A. and Minna, J.D. (1982) Cancer Genet. Cytogenet. 6, 119-132.
16. Whang-Peng, J., Kao-Shan, C.S., Lee, E.C., Bunn, P.A., Carney, D.N., Gazdar, A.F., and J.D. Minna. (1982) Science 215, 181-182.
17. Mark, J., Dahlenfors, R., and Ekedahl, C. (1982) Hereditas 96, 141-148.
18. Bonner, T., O'Brien, S.J., Nash, W.G., Rapp, U.R., Morton, C.C. and Leder, P. (1984) Science 223, 71-74.
19. Drabkin, H.A.., Bradly, C., Hart, I., Bleskan, J., Li, F.P. and Patterson, D. (1985) Proc. Natl. Acad. Sci. USA 82, 6980-6984.
20. Fukui, M., Yamamoto, T., Kawai, S., Maruo, K. and Toyoshima, K. (1985) Proc. Natl. Acad. Sci. USA 81, 5954-5958.
21. Ishikawa, F., Takaku, F., Ochiai, M., Hayashi, K., Hirohashi,

S., Terada, M., Takayama, S., Nasao, M. and Susimura, T. (1985) Biochem. Biophys. Res. Commun. 132, 186-192.
22. Ishikawa, F., Takaku, F., Hayashi, K. Nagao, M., and Sugimura, T. (1986) Proc. Natl. Acad. Sci. USA 83, 3209-3212.
23. Shimizu, K., Nakatsu, Y., Sekisuchi, M., Hokamura, K., Tanaka, K., Terada, M. and Sugimura, T. (1985) Proc. Natl. Acad. Sci. USA 82, 5641-5645.
24. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
25. Berg, J.M. (1986) Science 232, 485-487.
26. Parker, P.J., Coussens, L., Totty, N., Rhee, L., Young, S., Chen, E., Stabel, S., Waterfield, M.D. and Ullrich, A. (1986) Science 233, 853-859.
27. Coussens, L., Parker, P.J., Rhee, L., Yang-Feng, T.L., Chen, E., Waterfield, M.D., Franke, U. and Ullrich, A. (1986) Science 233, 859-866.
28. Knopf, J.L., Lee, M., Sultzman, L.A., Kritz, R.W., Loomis, C.R., Hewick, R.M. and Bell, R.M. (1986) Cell 46,491-502.
29. Mark, G.E. and Rapp, U.R. (1984) Science 224, 185-289.
30 Kamps, M.P., Buss, J.E. and Sefton, B.M. (1984) Nature 310, 589-592.
31. Bryant, D.L. and Parsons, J.T. (1984) Mol. Cell. Biol. 4 862-866.
32. Van Beveren, C. and Verma, I.M. (1986) in Current Topics in Microbiology and Immunology Vol 123, Vogt, P.K. and Kaprowski, eds., pp 73-98, Springer-Verlag, New York.
33 Sefton, B.M. (1986) Current Topics in Microbiology and Immunology Vol 123, Vogt, P.K. and Kaprowski, H., eds., pp 40-71 Springer-Verlag, New York.
34. Semba, K., Kamata, N., Toyoshima, K., and Yamamoto, T. (1985) Proc. Natl. AAcad. Sci. USA 82, 6497-6501.
35. King, C.R., Kraus, M.H., and Aaronson, S.A. (1985) Science 229, 974-976.
36. Schechter, A.L., Hung, M., Vaidyanathan, L., Weinberg, R.A., Yang-Feng, T.L., Francke, Y., Ullrich, A., and Coussens, L. (1985) Science 229, 976-978.
37. DePinho, R.A., Legouy, E., Feldman, L.B., Kohl, N.E., Yancopoulos, G.D. and Alt, F.W. (1986) Proc. Natl. Acad. Sci. USA 83, 1827-1831.
38. Stanton, L.W., Schwab, M., and Bishop, J.M. Proc. Natl. Acad. Sci. USA 83, 1772-1776.
39. Cochran, B.H., Zullo, J., Verma, I., and Stiles, C.D.(1986) Science 226, 1080-1083.
40. Zimmerman, K.A., Yancopoulos, G.D., Collum, R.G., Smith, R.K., Kohl, N.E., Denis, K.A., Nau, M.M., Witte, O.N., Toran-Allerand, D., Gee, C.E., Minna, J.D. and Alt, F.W. (1986) Nature 319, 780-783.
41. Bell, R.M. (1986) Cell 45, 631-632.
42. Takio, K., Wade, R.D., Smith, S.B., Krebs, E.G., Walsh, K.A. and Titani, K. (1984) Biochemistry 23, 4207-4218.
43. Galibru, J. and Hovanessian, A.G. (1985) Cell 43, 685-694.
44. Luzzatto, L., Usanga, E.A., Bienzle, U., Esan, G.F.J., and Fasaan, F.A. (1979) Science 205, 1418-1420).
45. Hamilton, J.K., Paquin, L.A., Sullivan, J.L., Maurer, H.S., Cruzi, F.G., Provisor, A.J., Steuber, C.P., Hawkins, E., Yawn, D., Cornet, J., Clausen, K., Finkelstein, G.Z.,

Landing, B., Grunnet, M., and Portilo, D.T. (1980) J. Pediat. 96, 669-673.
46. Li, F.P., Marchetto, D.J., and Vawter, G.F. (1979) Am. J. Hemat. 6, 61-69.
47. Falletta, J.M.,, Fernbach, D.J., Singer, D.B., Smith, M.A., Landing, B.H., Heath, C.W., Jr., Shore, N.A., and Barrett, F.F. (1973) J. Pediat. 83, 549-556.
48. Champlin, R., Gale, R.P., Foon,K.A., and Golde, D.W. (1986) Annals Int. Med. 104, 671-688.
49. Noda, M., Selinger, Z., Scolnick, E.M. and Bassin, R.H. (1983) Proc. Natl. Acad. Sci. USA 80, 5602-5606.