Nucleotide Sequence of the Rasheed Rat Sarcoma Virus Oncogene: New Mutations

Abstract. *The nucleotide sequence of the oncogene of the Rasheed strain of rat sarcoma virus was determined. The oncogene (Ra-v-*ras*) encodes a 29,000-dalton (p29) transforming protein. This protein is distinct from the immunologically related 21,000-dalton protein (p21) of the Harvey murine sarcoma virus in its amino terminus and in having additional mutations in its carboxyl terminus. Although the functional significance of these changes is unknown, they appear to occur only in rat sarcoma virus.*

Human and animal genomes contain a set of highly conserved genes (called proto-oncogenes or c-*onc*) that are homologous to the transforming genes (v-*onc*) of various oncornaviruses (*1*). Recombination of c-*onc* with the retroviral genes confers a rapid transforming potential on these viruses. Several human tumors contain oncogenes homologous to the viral transforming genes of the Harvey or Kirsten strains of murine sarcoma viruses (Ha-MuSV and Ki-MuSV, respectively) (*2, 3*). The viral oncogenes of both Ha-MuSV and Ki-MuSV (designated Ha-v-*ras* and Ki-v-*ras*) (*4*) were derived from cellular sequences (c-*ras*) during the passage of murine leukemia viruses (MuLV) in rats (*5, 6*). A third, related rat oncogene has been isolated from the Rasheed strain of rat sarcoma virus (RaSV), which was generated in cell cultures of chemically induced rat tumors by the use of an endogenous rat leukemia virus (RaLV) (*7, 8*). This virus is distinct from Ha-MuSV and Ki-MuSV in that it contains only rat-derived viral and cellular sequences and is thus exclusively of rat origin (*7, 8*). Heteroduplex analysis and restriction endonuclease maps of the RaSV oncogene (Ra-v-*ras*) indicated that a 0.7–kilobase pair (kbp) fragment was homologous and colinear with the 1-kbp fragment of Ha-v-*ras* and its two related rat cellular genes designated c-*ras* I and c-*ras* II (*6, 9*). However, the Ra-v-*ras* showed very little homology with Ki-v-*ras* and its normal cellular counterpart (*9*).

Both RaSV-transformed cells and tumors induced in vivo express a transformation-specific protein of 29,000 daltons (p29) that is immunologically related to the Ha-MuSV– and Ki-MuSV–induced proteins of 21,000 daltons (p21) (*8–10*). A related p21 has been detected in several human tumors, which also express Ha-v-*ras*– or Ki-v-*ras*–like oncogenes (*2, 3*). Nucleotide sequence analyses of both human and rat oncogenes have indicated that a critical mutation at the 12th amino acid residue of p21 is probably responsible for the change in the normal gene function (*2*). However, it is not yet clear if this change is a universal property of all human and animal cancers expressing p21. To compare the genetic structure of p29 with p21, we derived the nucleotide sequence of the DNA encoding p29 and deduced the corresponding amino acid sequence. Our results indicate that although there is an extensive homology between proteins p29 and p21, the amino terminus of p29 is totally distinct and contains a sequence resembling p15 of RaLV. In addition, there are several mutations and a DNA region not found in the viral oncogenes of Ha-MuSV or Ki-MuSV genes. Considering the small changes that have been identified between normal genes and viral Ha-v-*ras*–related oncogenes (*2*), these additional mutations in Ra-v-*ras* may be important. Some of these changes could have been induced in c-*ras* by the chemical carcinogen before its transduction to the RaSV genome (*7, 8*) or, alternatively, the RaSV oncogene may represent a variant of the c-*ras* family. In this report we define the transcriptional and translational controls of Ra-v-*ras* and discuss the possible implications of our findings.

A DNA fragment of about 1.7 kbp (Fig. 1) was purified from an RaSV provirus clone (*9*) and subcloned by shotgun and forced-cloning techniques in various single-stranded M13 bacteriophage strains (*11*). Nucleotide sequences were determined in both directions from 30 different overlapping DNA fragments by the dideoxynucleotide chain termination method (*12*).

Translation of most eukaryotic messenger RNA's is initiated at the AUG (A, adenine; U, uracil; G, guanine) triplet nearest the 5′ cap site. We have identified this codon at position 610 in the 5′ end of the Ra-v-*ras* gene by alignment with the p15 initiation codons of the Moloney strain of murine leukemia virus (Mo-MuLV) and feline leukemia virus (FeLV) (*13, 14*). The sequences upstream of this first ATG (T, thymine) triplet conform to the R-U5 region of the retroviral long terminal repeat as identified by the presence of various consensus sequences. The polyadenylated [poly(A)] addition signal AATAAA is located between positions 12 and 19, and a sequence homologous to Mo-MuLV inverted repeat sequence (102 to 117) is followed by the primer binding site (118 to 134) at the 3′ boundary of U5. In Mo-MuLV, the 5′ splice site and a 3′ splice site may be involved in removal of intervening sequences that contain three ATG triplets preceding the initiator codon of *gag* gene (*13*). Although the 5′ splice site is present at positions 176 to 182 in Ra-v-*ras* and several possible 3′ splice sites match with the consensus sequence proposed by Sharp (*15*), splicing in this region of RaSV may not be imperative since only one ATG codon that coincides with the retroviral *gag* gene is located in the 5′ end of the gene (Fig. 2).

A large open reading frame from the first ATG codon is 747 nucleotides long

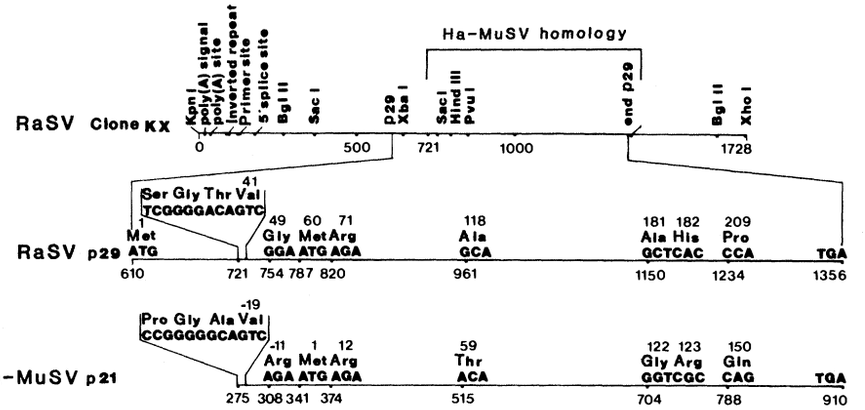

Fig. 1. Schematic diagram summarizing the homologies and differences between p29, encoded by Ra-v-*ras*, and p21 proteins, encoded by Ha-v-*ras*. The numbers below each line refer to the nucleotide positions and those above the line denote the deduced amino acids. The top line represents RaSV proviral DNA between restriction sites Kpn I and Xho I, including the p29-encoding sequence. Homology of Ha-MuSV with RaSV extends from position 721 to 1364. The middle line is a map of RaSV p29 sequences and the bottom line shows the corresponding positions in p21.

and terminates at the TGA triplet at position 1356 (Figs. 1 and 2). No other amber, ochre, or opal codon is present in this protein encoding region (Fig. 2). The first few amino acids encoded by Ra-v-*ras* are highly conserved retroviral sequences and match the sequences encoding the amino terminal sequences of both Mo-MuLV and FeLV p15 proteins (*13, 14*) (Fig. 2). Thus, this first RaLV *gag* gene–related ATG codon is also the first translational signal for the RaSV encoded p29. The Mo-MuLV– and FeLV–derived p15 encoding sequences in the Moloney strain of mouse sarcoma virus (Mo-MSV) and the feline sarcoma virus (FeSV) specify 130 and 127 amino acids, respectively (*13, 14, 16*). However, only 32 of the p15-derived amino acid residues are present in the RaSV genome (Fig. 2). Of these, four residues each are similar to Mo-MSV or FeSV and 18 residues are shared by both FeSV and Mo-MSV p15 sequences, although RaSV-specific amino acids are interspersed at six positions (Fig. 2). Furthermore, the nucleotide triplets encoding the shared amino acids show 12 transitions, mostly by first- or third-base substitutions and seven second- or third-base transversions, probably reflecting generic differences between the mouse, cat, and rat genomes (Fig. 2). The RaLV p15-related sequences are followed by a cluster of six amino acids (residues 33 to 38) that are not shared by Mo-MuSV, FeSV, Ha-MuSV, or Ki-MuSV (*4, 13, 14, 16*). Thus, this sequence is either RaSV-specific or may represent a new rat cellular sequence that has translocated and formed a hybrid gene by recombination with a c-*ras* sequence related to Ha-v-*ras*. The nucleotides downstream from this region form a junction between the RaSV-specific sequence and the sequence that is homologous to the Ha-v-*ras* gene encoding p21. The recombination of these two sequences therefore may have occurred at position 724 or at 730 (Figs. 1 and 2).

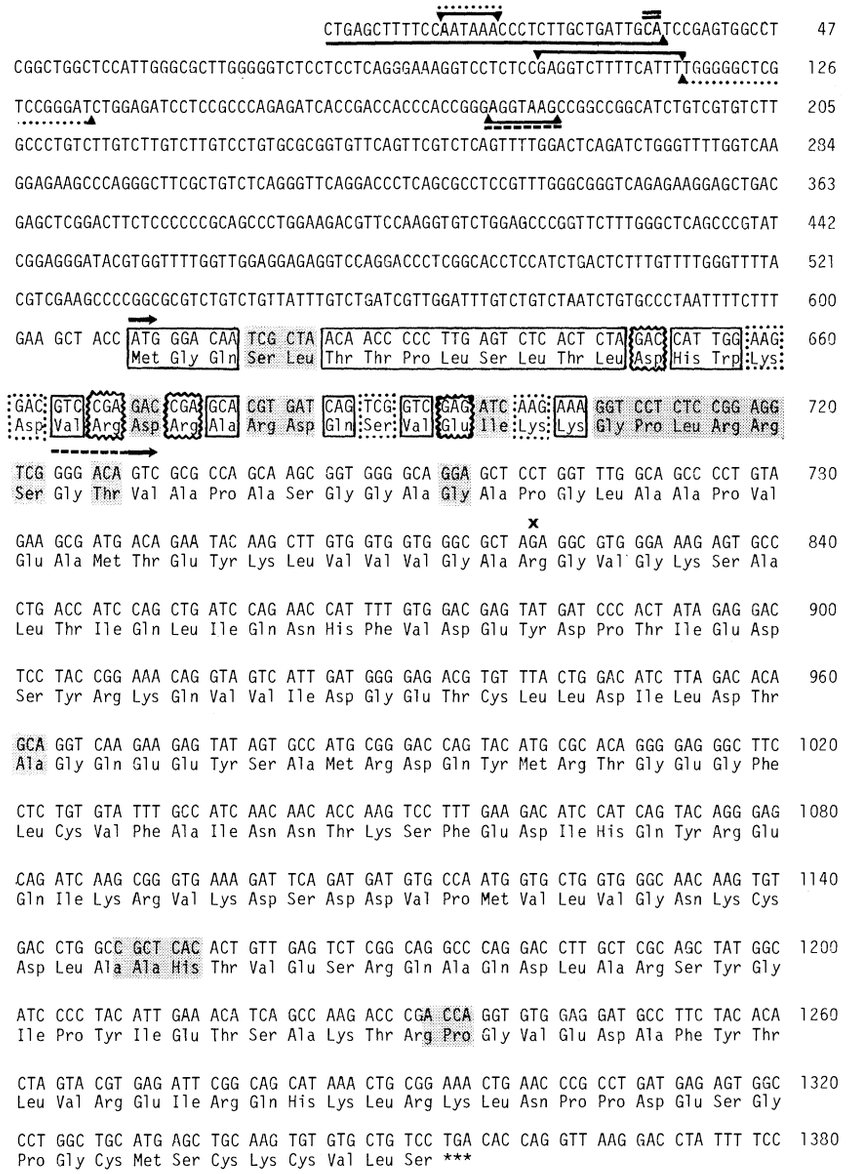

Fig. 2. The sequence of a subgenomic fragment of RaSV containing about 1400 nucleotides representing the region of the genome encoding p29. The numbers of the nucleotides in the sequence are on the right and those of the amino acids are on the left. Features discussed in the text are terminal redundancy (▲——▲), poly(A) signal ( ), poly(A) addition site (=), inverted repeat (▼——▼), primer binding site (▲·····▲), 5′ splice site ( ), and possible 3′ splice sites ending at positions 255, 312, 329, 387, 410, and 416. Homology of RaSV with both Mo-MuLV and FeLV p15 is designated by solid-lined boxes, homology with FeLV alone by wavy-lined boxes, and homology with Mo-MuLV alone by dotted boxes. RaSV-specific sequences are shaded; initiation of p29 (⟶); homology with p21 (⇢); **x** indicates the 12th amino acid of p21-like sequence; and *** underscores the stop codon for the p29 encoding sequence.

Homology with the Ha-v-*ras* sequence begins at position 724 and shows two differences of threonine and glycine (residues 40 and 49) in place of alanine and arginine at positions −19 and −11, respectively, of the p21 protein (Figs. 1 and 2). With the exception of four amino acids and two altered alanine and arginine codons (1150 and 1234), all of the nucleotides and deduced amino acids, beginning with the first methionine of p21-like protein (787), coincide exactly with the Ha-v-*ras* sequence (Fig. 2). This homology includes the critical mutation of arginine at the 12th amino acid residue which in p29 corresponds to residue 71 (Figs. 1 and 2). The four mutations within the p21-like sequence of Ra-v-*ras* are alanine, alanine, histidine, and proline at residues 118, 181, 182, and 209, respectively, compared to the corresponding threonine, glycine, arginine, and glutamine at residues 59, 122, 123, and 150 of Ha-v-*ras*–encoded p21 (Figs. 1 and 2).

Although p21's of Ha-MuSV and Ki-MuSV are encoded by distinct cellular genes, both proteins are autophosphorylated solely at threonine residue 59 and are encoded by the same nucleotides at the same respective map positions (*4, 17*). In p29, this threonine is replaced by alanine (residue 118), but the adjacent nucleotides and their deduced amino acids in the surrounding area are identical to those of p21. This is important because it confirms our results (*18*) and those of others (*19*) that the RaSV-induced p29 is not autophosphorylated. Moreover, the phosphoamino acid of p29 is a serine residue in vivo but these studies have yet to be confirmed in vitro (*19, 20*).

The viral oncogene of Ha-MuSV contains three initiation codons in the same reading frame and can potentially encode proteins p30, p29, and p21. However, the Ha-MuSV–transformed cells produce mainly the p21 protein, rarely show a minor component of p30 protein, and have never expressed the purported p29 (*4, 6, 17, 18*). The first ATG codon of the Ra-v-*ras* is in a large open reading frame of 747 nucleotides, 21 nucleotides larger than the putative p30-encoding sequence of Ha-v-*ras*. Although a second ATG triplet is located at position 787 of Ra-v-*ras*, where homology with p21-like gene begins, we have never isolated a p21 protein from the immunoprecipitates of RaSV-transformed cells (*10*). The translation of p29 therefore must start at the first ATG codon and continue until the end of the protein, indicating that it is not a cleavage product of a larger protein or synthesized as a precursor of a smaller protein.

We have tested the biological activity of several DNA fragments after cleavage of a 6.4-kbp cloned RaSV segment with various enzymes that would cleave within the RaLV long terminal repeat sequence (Kpn I, Sac I, and Xba I). The 6.4-kbp fragment transformed the transfected mouse NIH/3T3 or the rat cells that expressed a p29 protein; the two smaller fragments containing the Ra-v-*ras* gene but not the long terminal repeat sequence did not transform rat or mouse cells. Since Kpn I cleaves within the region of the RaLV long terminal repeat (*18*), these results indicate that the retroviral promoter sequences are required for the efficient transcription and translation of Ra-v-*ras*.

Complete nucleotide sequences of the normal rat c-*ras* genes have not been published. However, the replacement of glycine with arginine in Ha-MuSV, and with serine in Ki-MuSV has been reported to be the major difference between normal and viral oncogenes (*2, 4*). Thus the four new changes within the p21-like sequence of Ra-v-*ras* and the RaSV-specific sequence in the 5′ half of this gene may represent important deviations from the transforming oncogenes of RaSV and Ha-MuSV. This is not surprising because the oncogenes of both Ha-MuSV and Ki-MuSV were transduced by passage of MuLV in normal rats, whereas Ra-v-*ras* was rescued by RaLV from chemically induced tumors (*7*).

After submission of this report, Capon *et al.* (*21*) reported that the T24 human bladder carcinoma and the normal cellular p21 proteins were identical to the viral Ha-v-*ras* proteins at all amino acid residues except three (residues 12, 59, and 122). Two of these three unique amino acids (residues 59 and 122) of the human p21 are identical to those encoded by the p29 protein of Ra-v-*ras* at precisely the same corresponding positions. Although the biological significance of these changes remains to be determined, the Ra-v-*ras*, in this respect, appears to be more closely related to the human gene than to the viral Ha-v-*ras*.

SURAIYA RASHEED
GARY L. NORMAN
University of Southern California School of Medicine, Los Angeles 90033
GISELA HEIDECKER
University of California, Davis, School of Medicine, Davis 95616

References and Notes

1. J. M. Bishop, *Sci. Am.* **247**, 80 (March 1982); D. Stehelin, H. E. Varmus, J. M. Bishop, *Nature (London)* **260**, 170 (1976).
2. P. E. Reddy, R. K. Reynolds, E. Santos, M. Barbacid, *Nature (London)* **300**, 149 (1982); C. J. Tabin *et al.*, *ibid.*, p. 143.
3. L. F. Parada, D. J. Tabin, C. Shih, R. A. Weinberg, *ibid.* **297**, 474 (1982); E. H. Chang, M. E. Furth, E. M. Scolnick, D. R. Lowy, *ibid.*, p. 479; S. Pulciani, E. Santos, A. V. Lauver, L. K. Long, K. C. Robbins, M. Barbacid, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2845 (1982).
4. R. Dhar *et al.*, *Science* **217**, 934 (1982); N. Tsuchida, T. Ryder, E. Ohtsubo, *ibid.*, p. 937.
5. J. J. Harvey, *Nature (London)* **204**, 720 (1964); W. H. Kirsten and L. A. Mayer, *J. Natl. Cancer Inst.* **39**, 311 (1966).
6. R. W. Ellis *et al.*, *Nature (London)* **292**, 506 (1981); D. DeFeo, M. A. Gonda, H. A. Young, E. H. Chang, D. R. Lowy, E. M. Scolnick, R. W. Ellis, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 3328 (1981).
7. S. Rasheed, M. B. Gardner, R. J. Huebner, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 2972 (1978); S. Rasheed, J. Bruszewski, R. W. Rongey, T. Roy-Burman, H. P. Charmin, M. B. Gardner, *J. Virol.* **18**, 799 (1976).
8. S. Rasheed, *Cold Spring Harbor Symp. Quant. Biol.* **54**, 779 (1980).
9. M. A. Gonda *et al.*, *J. Virol.* **44**, 520 (1982).
10. S. Rasheed and H. A. Young, *Virology* **118**, 219 (1982); H. A. Young, T. Y. Shih, E. M. Scolnick, S. Rasheed, M. B. Gardner, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 3523 (1979); H. A. Young *et al.*, *J. Virol.* **38**, 286 (1981).
11. J. Vieira and J. Messing, *Gene* **19**, 259 (1982); J. Messing and J. Vieira, *ibid.*, p. 269.
12. J. Messing, R. Crea, P. H. Seeburg, *Nucleic Acids Res.* **9**, 309 (1981).
13. T. M. Shinnick, R. A. Lerner, J. G. Sutcliffe, *Nature (London)* **293**, 543 (1981).
14. A. Hampe, I. Laprevotte, F. Galibert, L. A. Fedele, C. J. Sherr, *Cell* **20**, 775 (1982).
15. P. A. Sharp, *ibid.* **23**, 643 (1981).
16. C. Van Beveren, F. Van Straaten, J. A. Galleshaw, I. M. Verma, *ibid.* **27**, 97 (1981).
17. T. Y. Shih, P. E. Stokes, G. W. Smythers, R. Dhar, S. Oroszlan, *J. Biol. Chem.* **257**, 11767 (1982).
18. S. Rasheed, unpublished data.
19. T. Y. Shih, personal communication.
20. A. Papageorge, D. R. Lowy, E. M. Scolnick, *J. Virol.* **44**, 509 (1982).
21. D. J. Capon, E. Chen, A. Levinson, P. Seeburg, D. Goeddel, *Nature (London)* **302**, 33 (1983).
22. Supported by grant CA 27246 from the National Cancer Institute. We thank M. B. Gardner for his help and H. A. Young and M. Gonda for the RaSV clone.

31 January 1983; revised 4 April 1983

Chitin in Sea Anemone Shells

Abstract. *Chitin, which is widely distributed among life forms, is well documented in the coelenterate class Hydrozoa and is contained in one member of class Scyphozoa. In class Anthozoa, hard corals synthesize it but soft corals do not. Chitin was identified by infrared spectrophotometry in the trochoid shell of the actinian* Stylobates. *It constitutes 1.7 percent of the shell by weight, the rest probably being protein. The ability of sea anemones to synthesize chitin is thereby confirmed.*

Chitin is a linear polysaccharide similar in many respects to cellulose (*1–3*). Jeuniaux (*1, 4*) concluded that its wide distribution among fungi, plants, protists, and animals is evidence that the capacity to synthesize it evolved early.

Chitin, in the restricted sense, has repeatedly been demonstrated in the perisarcs of both calyptoblastic and gymnoblastic hydroids; it occurs in chondrophore and millepore hydrozoans as well (*1, 2, 4*). Jeuniaux once proposed (*1*) that this feature distinguishes hydrozoans from other coelenterates but abandoned the idea (*4*) when chitin was demonstrated, albeit rarely, in members of other classes. Among scyphozoans, only the podocyst of *Aurelia aurita* is known to contain this substance (*5*).

Chitin is more common among anthozoans, but may be totally absent in some groups of the class; the several species of octocorals that have been analyzed lack it (*1, 6*). The first record of chitin in a nonhydrozoan coelenterate was from the reef-forming coral *Pocillopora damicornis* (*7*). Wainwright (*7*) proposed that this material, which makes up 0.01 to 0.1 percent of the dry coral skeleton by weight, and which was identified by x-ray diffraction and biochemical tests, forms the organic matrix upon which calcification occurs. Initially Jeuniaux (*1*) acknowledged only that the physical and chemical properties of the compound were near to those of chitin, but chitin has since been found in other corals (*6*).

Pelagic anemones of family Minyadidae secrete from their pedal end a "chitinous mass" that keeps them afloat (*8*); a cuticle envelops the column of many species of actinians; sea anemones of the genus *Adamsia* living on gastropod shells inhabited by hermit crabs may extend the shell's lip by secreting what is variously called a cuticle (*9*), "chitine" (*10*), a horny membrane (*10, 11*), or