The EMBO Journal vol.6 no.11 pp.3341-3351, 1987

Human proto-oncogene c-*kit*: a new cell surface receptor tyrosine kinase for an unidentified ligand

Yosef Yarden[1,4], Wun-Jing Kuang[1], Teresa Yang-Feng[2], Lisa Coussens[1], Susan Munemitsu[1], Thomas J.Dull[1], Ellson Chen[1], Joseph Schlessinger[3], Uta Francke[2] and Axel Ullrich[1]

[1]Department of Developmental Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080, [2]Department of Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510 and [3]Biotechnology Research Center, Meloy Laboratories, 4 Research Court, Rockville, MD 20850, USA

[4]Present address: Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

Communicated by J.Schlessinger

Structural features of v-*kit*, the oncogene of HZ4 feline sarcoma virus, suggested that this gene arose by transduction and truncation of cellular sequences. Complementary DNA cloning of the human proto-oncogene coding for a receptor tyrosine kinase confirmed this possibility: c-*kit* encodes a transmembrane glycoprotein that is structurally related to the receptor for macrophage growth factor (CSF-1) and the receptor for platelet-derived growth factor. The c-*kit* gene is widely expressed as a single, 5-kb transcript, and it is localized to human chromosome 4 and to mouse chromosome 5. A c-*kit* peptide antibody permitted the identification of a 145 000 dalton c-*kit* gene product that is inserted in the cellular plasma membrane and is capable of self-phosphorylation on tyrosine residues in both human glioblastoma cells and transfected mouse fibroblasts. Our results suggest that p145$^{c\text{-}kit}$ functions as a cell surface receptor for an as yet unidentified ligand. Furthermore, carboxy- and amino-terminal truncations that occurred during the viral transduction process are likely to have generated the transformation potential of v-*kit*.

Key words: receptor/proto-oncogene/tyrosine kinase

Introduction

The receptors for epidermal growth factor (EGF) (Ullrich *et al.*, 1984), insulin (Ebina *et al.*, 1985; Ullrich *et al.*, 1985), platelet-derived growth factor (PDGF) (Yarden *et al.*, 1986), insulin-like growth factor (IGF-I) (Ullrich *et al.*, 1986), and macrophage colony-stimulating factor-1 (CSF-1) (Sherr *et al.*, 1985; Coussens *et al.*, 1986) represent a family of closely related proteins that transduce growth regulatory signals across the plasma membrane. The structural organization of these glycoproteins includes an extracellular ligand binding domain, connected by a single membrane-spanning segment to an intracellular domain that possesses tyrosine-specific protein kinase activity. The oncogenes of certain transforming retroviruses appear to be derived from the proto-oncogenes that encode some of these receptors. Thus v-*erbB*, the transforming gene of avian erythroblastosis virus (AEV), is thought to encode a truncated form of the EGF receptor (EGFR) (Downward *et al.*, 1984a; Ullrich *et al.*, 1985), whereas the v-*fms* oncogene of the McDonough strain of feline sarcoma virus (SM-FeSV) specifies an altered receptor for CSF-1 (Sherr *et al.*, 1985; Coussens *et al.*, 1986; Sacca *et al.*, 1986). Similarly, the non-viral oncogene HER2/*neu* codes for a receptor tyrosine kinase (RTK) for an as yet unknown ligand (Coussens *et al.*, 1985; Bargmann *et al.*, 1986a; Yamamoto *et al.*, 1986). In each case, structural alterations in these oncogenes appear to deregulate the receptor kinase activity, which results in uncontrolled cell growth and transformation.

Characterization of the primary structures of six members of the RTK family (Ullrich *et al.*, 1984, 1985, 1986; Bargmann *et al.*, 1986b; Coussens *et al.*, 1986; Yamamoto *et al.*, 1986; Yarden *et al.*, 1986) revealed the existence of three distinct structural subtypes which may have independent evolutionary origins (Y.Yarden and A.Ullrich, submitted). One group (type III) is represented by the receptors for PDGF and CSF-1, which share overall sequence homology including a unique distribution of cysteine residues within their ligand binding domains (Coussens *et al.*, 1985; Yarden *et al.*, 1986). The most unusual feature of this group of RTKs is that their cytoplasmic domains are interrupted by long hydrophilic insertion sequences. This organization of the tyrosine-specific kinase domain was recently found to be shared by v-*kit*, a new member of the tyrosine kinase family of oncogenes, present in the genome of the Harvey-Zuckerman-4 strain of feline sarcoma virus (HZ4-FeSV) (Besmer *et al.*, 1986). The predicted v-*kit* amino acid sequence displays the highest homology with the kinase domain of the v-*fms* oncogene product and to a lesser extent with the PDGF receptor (Yarden *et al.*, 1986).

Although v-*kit* encodes a protein that lacks extracellular and membrane-spanning sequences, its unique structural features suggest that it was generated by transduction and truncation of the cellular *kit* sequence, which itself might encode a receptor-like molecule. To test this hypothesis, we isolated and characterized cDNA clones that encode the human c-*kit* proto-oncogene. The deduced amino acid sequence of c-*kit* reveals that in addition to the tyrosine kinase region which is shared by the viral and the cellular forms of *kit*, the proto-oncogene encodes a 976 amino acid polypeptide sharing major structural features with the CSF-1 and PDGF receptor subfamily. We demonstrate that the c-*kit*-encoded polypeptide is expressed in human glioblastoma and transfected NIH-3T3 cells as a 145 000 M_r cell surface molecule with tyrosine-specific autophosphorylation activity. Our findings further emphasize the importance of amino- and carboxy-terminal truncations in the conversion of normal receptors to oncogenic proteins.

Results

Human c-kit proto-oncogene cDNA and deduced protein sequence

Oligonucleotide probes based on v-*kit* sequences (Besmer *et al.*, 1986) were synthesized for the isolation of human c-*kit* cDNA clones. Probe sequences were chosen from regions that are highly variable within the RTK family, but conserved between species homologues (Coussens *et al.*, 1985; Yarden *et al.*, 1986) to minimize the probability of detecting related RTK clones distinct from c-*kit*. To compensate for likely species-specific sequence differences between feline v-*kit*-derived probes and human c-*kit*

cDNA target sequences, we employed long probes of 90–96 nucleotides (ON-1, ON-2, ON-3; Figure 1).

Northern blot analysis was carried out using a mixture of all three oligonucleotide probes to identify tissues expressing c-*kit* mRNA. At low stringency, one strong hybridization signal of ~5 kb was obtained in poly(A)$^+$ RNA preparations from a variety of human, rat, bovine and murine tissues (not shown). The largest quantities of the 5-kb mRNA were found in human glioblastoma cells and term placenta as well as in rat, simian and bovine brain. A human term placenta cDNA library in λgt10 (Huynh *et al.*, 1985) was therefore screened at low stringency using the pooled oligonucleotide probes. One strongly hybridizing clone containing an 0.9 kb cDNA insert was isolated (λHP-ckit9; Figure 1). When nick-translated and used as a probe on Northern blots with human placental poly(A)$^+$ RNA, it also hybridized to a single transcript of 5 kb (Figure 2). Hybridization analysis with individually labeled oligonucleotides demonstrated that only sequences homologous to probe ON-1 (see Figure 1) were present in λHP-ckit9. Nucleotide sequence analysis of this clone revealed homology with a portion of v-*kit* that lacked sequences homologous to probes ON-2 and ON-3.

The initial clone (λHP-ckit9) was used to obtain overlapping cDNA fragments from human term placenta and fetal brain cDNA libraries (Figure 1) which span 5085 nucleotides, and include the entire c-*kit* coding region in addition to 5′ and 3′ untranslated sequences. Fetal brain cDNA clones λHFB-ckit/171 and λHFB-ckit/1 contain almost the entire c-*kit* cDNA sequence. A gap of 262 bp is filled by placental clone λHP-ckit63, which overlaps λHFB-ckit/1 and 171 and confirms their sequence. The fetal brain sequence was further partially confirmed by analysis of placental clones λHP-ckit9 and 33 (Figure 1).

The nucleotide sequence of human c-*kit* cDNA contains an open reading frame of 976 codons beginning with a possible initial methionine ATG codon (Figure 3). Although the open reading frame extends further upstream, our assignment of the initiation methionine codon is supported by (i) the downstream amino acid sequence having all the features characteristic of a signal sequence (Watson, 1984), (ii) the colinearity with CSF-1 and PDGF receptor sequences, and (iii) the size of our cDNA slightly exceeding the estimated c-*kit* mRNA size. Furthermore, the ATG at nucleotide position 23 is flanked by sequences that fulfil Kozak's criteria for an authentic initiation codon (Kozak, 1984). The 2928 nucleotide coding region is flanked by 22 nucleotides of 5′ and 2135 nucleotides of 3′ untranslated sequences. No poly(A) tail is present at the 3′ end of our sequence, suggesting that our cDNA clones are incomplete.

The amino-terminal stretch of most hydrophobic residues is likely to represent cleavable signal peptide necessary for polypeptide translocation into the lumen of the endoplasmic reticulum (Watson, 1984). By analogy with other signal peptidase cleavage sites, we predict the amino terminus of the mature c-*kit* polypeptide product to be Ser 24. This putative amino-terminal serine residue is followed by the 953 amino acid human c-*kit* sequence with a calculated mol. wt of 109 740. A hydrophobic stretch of 23 amino acids (residues 521–543) characteristic of receptor membrane-spanning regions divides the sequence into two major domains (Figures 3 and 4). The amino-terminal, 491-residue portion has previously described structural features of extracellular receptor ligand binding domains (Coussens *et al.*, 1985; Yarden *et al.*, 1986), including six regularly spaced cysteine residues (Cys 58, 97, 136, 186, 233, 290) and nine potential N-linked glycosylation sites (Figure 3). The carboxy-terminal 439 residues distal to the putative transmembrane domain include a group of polar and positively charged amino acids that directly flank this hydrophobic region, an ATP binding site consensus sequence (Hannink and Donoghue, 1985) 53 residues downstream, sequences homologous to receptors and oncogene products of the tyrosine kinase family, and a hydrophilic carboxy-terminal tail.

p145$^{c\text{-}kit}$ autophosphorylation and membrane association

To identify and characterize the c-*kit* gene encoded polypeptide, a polyclonal rabbit antibody was prepared against the C-terminal c-*kit* hexadecapeptide GSTASSSQPLLVHDDV (Figure 3) conjugated to soybean trypsin inhibitor. A172 glioblastoma cells which express c-*kit* mRNA (Figure 2) were metabolically labeled with [^{35}S]methionine and [^{35}S]cysteine and analysed by immunoprecipitation. Figure 5A shows that a polypeptide of mol. wt 145 000 reacted with the antibody (Figure 5A, lanes 2 and 3). The specificity of interaction was demonstrated by competition experiments employing excess carrier-conjugated peptide, which abolished the signal (Figure 5A, lane 4). The increased size of the c-*kit* protein is probably due to N-linked and possibly O-linked glycosylation, by analogy with its close relatives, the PDGF receptor and c-*fms*/CSF-1 receptor. As predicted from this homology, incubation of immunoprecipitated c-*kit* protein with [γ-^{32}P]ATP led to phosphorylation of the 145 000 dalton protein on tyrosine residues (Figure 5B and C).

To demonstrate unequivocally that this protein was encoded by our c-*kit* cDNA clones, we expressed the cDNA in mammalian cells. Transfection of NIH-3T3 mouse fibroblasts with the mam-

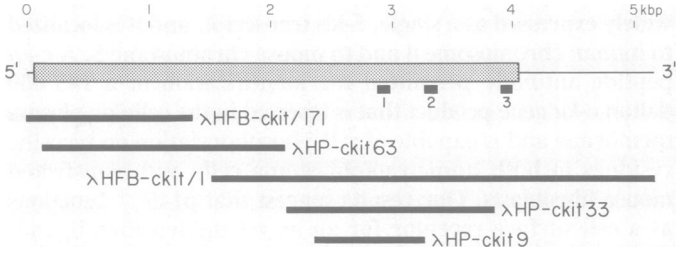

Fig. 1. Human c-*kit* cDNA clones. Human c-*kit* cDNA clones (solid bars) from human fetal brain (HFB) and placenta (HP) λgt10 cDNA libraries are shown with a schematic diagram of the predicted c-*kit* mRNA. Coding sequences (shaded box) and untranslated regions (solid line) are indicated. Solix boxes demarcate the locations of v-*kit* (Besmer *et al.*, 1986) derived oligonucleotide probes ON-1 (v-*kit* 1020–1110), ON-2 (v-*kit* 1405–1500) and ON-3 (v-*kit* 2041–2130) within the human c-*kit* sequence.

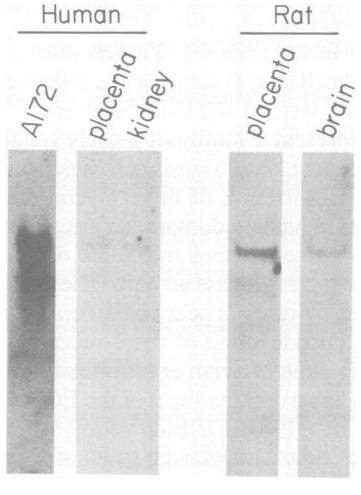

Fig. 2. Identification of c-*kit* mRNA by Northern blot analysis.

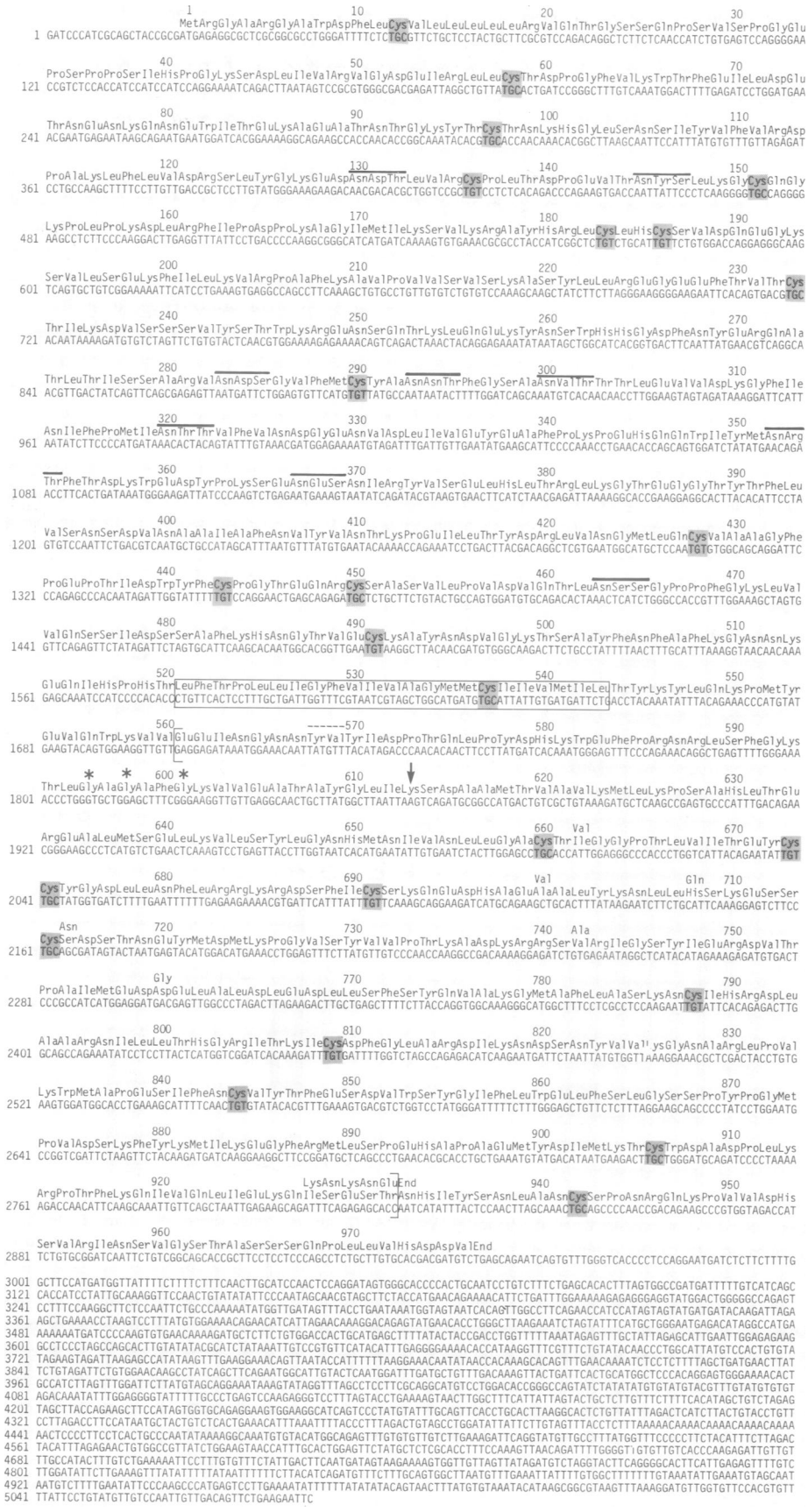

Fig. 3. Nucleotide sequence and deduced amino acid sequence of the human c-*kit* proto-oncogene in comparison with v-*kit* sequences. Nucleotides are numbered at the left and amino acids are numbered above the line, starting at the putative initiation codon. Potential sites of N-linked glycosylation are overlined and cysteine residues are shaded. The potential transmembrane region (see Figure 4) is boxed. Residues involved in ATP binding are indicated by asterisks and an arrow (Lys). Brackets indicate the ends of the aligned v-*kit* sequence and amino acid sequence differences are indicated above the line.

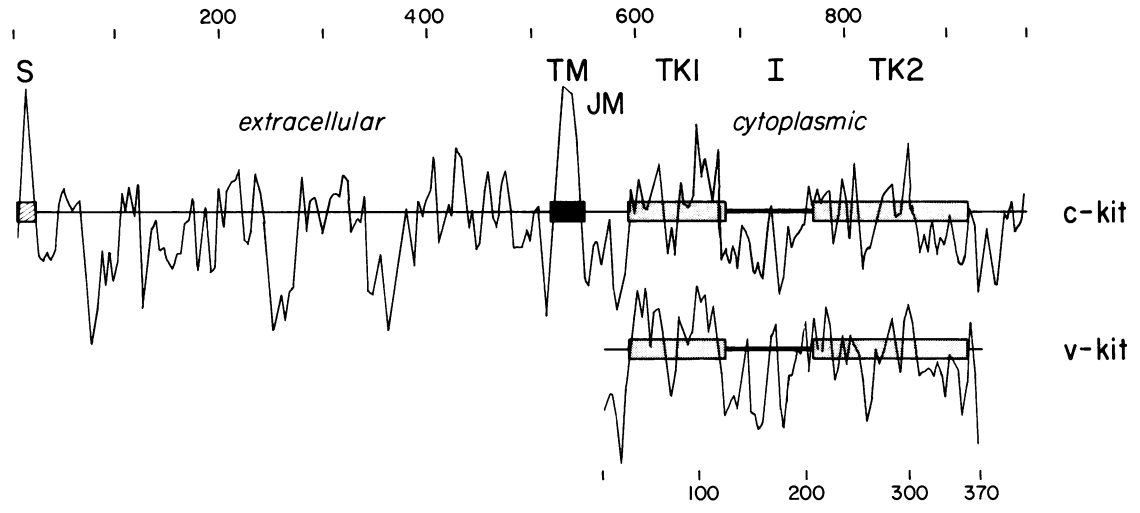

Fig. 4. Hydropathy profile of cellular and viral *kit* amino acid sequences and schematic diagrams of predicted protein domains. The method of Kyte and Doolittle (1982) was used. Hydrophobic signals are above the line. Shaded areas indicate tyrosine kinase sequences (TK1, TK2), the black box represents the transmembrane (TM) domain and the signal sequence is indicated as a hatched box (S). Tyrosine kinase insertion regions are shown as heavy lines and the juxtamembrane domain is indicated (JM).

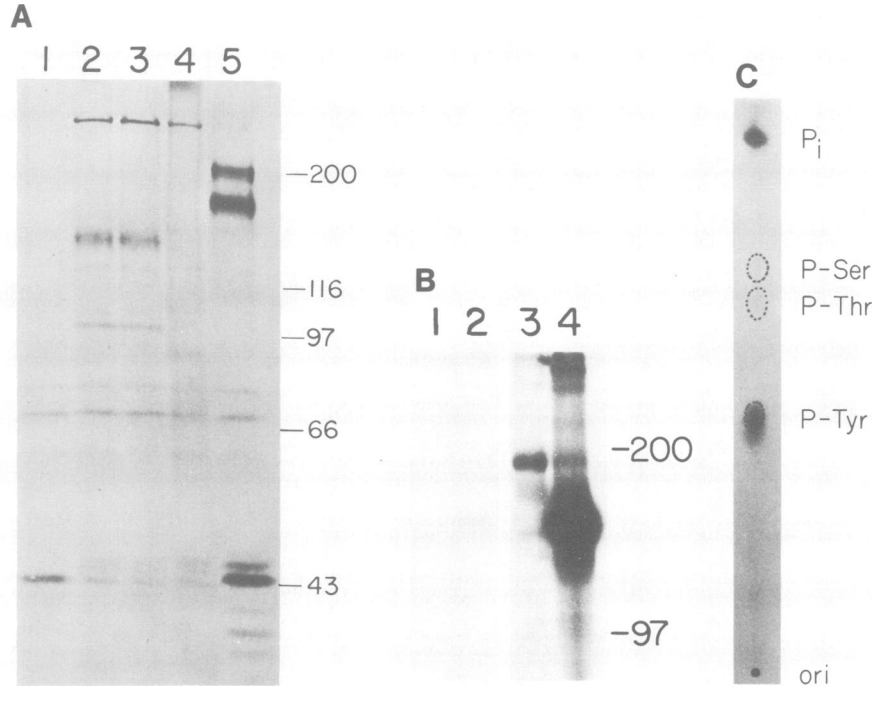

Fig. 5. (**A**) Immunoprecipitation of c-*kit* gene product from metabolically labeled A172 human glioblastoma cells. **Lane 1**, pre-immune serum; **lanes 2** and **3**, c-*kit* peptide antiserum (LJ11) (two different bleeds); **lane 4**, as in lane 3 except that the immunizing peptide coupled to soybean trypsin inhibitor was present during immunoprecipitation at 0.5 mg/ml peptide concentration; **lane 5**, EGFR antipeptide antibody RK2 (Kris *et al.*, 1985). Due to EGFR gene rearrangements in the A172 cell genome, a normal and an altered larger receptor polypeptide are detected with RK2. Mol. wt markers are indicated in kd. (**B**) Autophosphorylation of $p145^{c\text{-}kit}$. Plasma membrane preparations from A172 human glioblastoma cells were solubilized and immunoprecipitated with pre-immune rabbit antiserum (**lanes 1** and **3**) or anti-c-*kit* peptide antiserum (LJ11, **lanes 2** and **4**). Extensively washed immunoprecipitates were labeled by incubation with [γ-^{32}P]ATP (15 μM, **lanes 1** and **2** and 50 nM, **lanes 3** and **4**) and $MnCl_2$ (15 mM). Reaction was terminated after 30 min at 22°C, and the immune complexes washed and subjected to gel electrophoresis. Sizes of protein standards are indicated in kd. (**C**) Phosphoamino acid analysis of *in vitro* phosphorylated $p145^{c\text{-}kit}$. The radioactive 145 K band shown in Figure 2B (**lane 4**) was subjected to phosphoamino acid analysis (Kobayashi *et al.*, 1981). Non-radioactive phosphoamino acids were mixed with the radioactive sample prior to electrophoresis to allow location of the three phosphorylated amino acids (indicated by arrows). An identical pattern was obtained with the radioactive 145 K band derived from lane 2 of Figure 2B.

malian expression plasmid, CVNhckit, and amplification by methotrexate selection led to the generation of stable cell lines expressing human c-*kit*. Immunoprecipitation experiments with metabolically labeled CVNhckit-N250, CVNhckit-N500 and NIH-3T3 control cells, transfected with the CVN expression vector, showed that a polypeptide of mol. wt 145 000, identical in size to the A172 c-*kit* product (Figure 6A), was produced in the transfected cell lines. This polypeptide incorporated ^{32}P when incubated with [γ-^{32}P]ATP and Mn^{2+} ions, presumably by autophosphorylation (Figure 6B). ^{125}I-Surface labeling of transfected mouse NIH-3T3 cells demonstrated that the c-*kit* protein is indeed a cell surface protein that contains an extra-

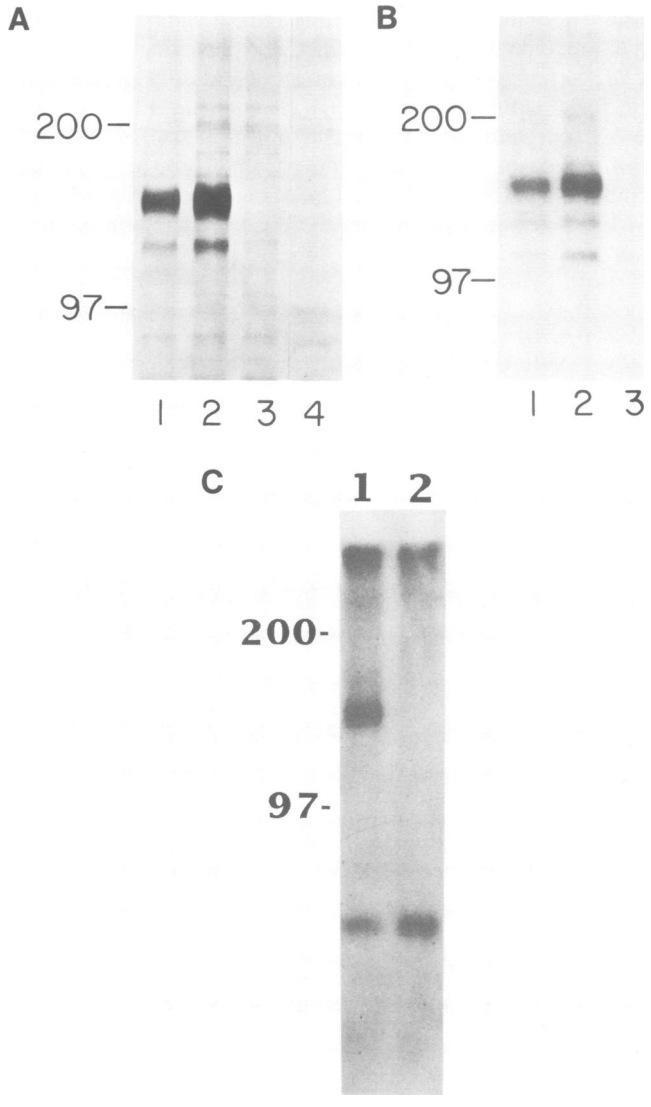

Fig. 6. **(A)** Immunoprecipitation of the c-*kit* gene product from metabolically labeled NIH-3T3 cells transfected with CVNhckit. [^{35}S]Methionine-labeled cells were solubilized and the cell lysates subjected to immunoprecipitation with pre-immune sera followed by immunoprecipitation with the anti-c-*kit* peptide (LJ11), except in lane 4 where only the first immunoprecipitation with pre-immune sera was performed. **Lanes 1** and **4**, CVNhckit-N250; **lane 2**, CVNhckit-N500; **lane 3**, CVN-N500 (NIH-3T3 cells transfected with the expression plasmid CVN and maintained in 500 nM methotrexate). Mol. wt markers are indicated in kd. **(B)** Autophosphorylation of p145$^{c\text{-}kit}$. Immunoprecipitated autophosphorylation products were analyzed by SDS–PAGE. **Lane 1**, CVNhckit-N250; **lane 2**, CVNhckit-N500; **lane 3**, CVN-N500. Mol. wt markers are indicated in kd. **(C)** Immunoprecipitation of ^{125}I surface-labeled NIH-3T3 cells transfected with CVNhckit. Surface iodinated cells were subjected to immunoprecipitation with pre-immune sera followed by immunoprecipitation with anti-c-*kit* peptide (LJ11). **Lane 1**, CVNhckit-N500; **lane 2**, CVN-N500. Mol. wt markers are indicated in kd.

cellular domain as predicted from the primary sequence (Figure 6C).

Chromosomal location of c-kit proto-oncogene

The chromosomal location of human c-*kit* was determined by *in situ* hybridization of the ^{3}H-labeled 0.9-kb λHP-ckit9 cDNA insert to normal human prometaphase chromosomes from two normal individuals. In 110 cells, a total of 233 grains were asociated with chromosomes. Of these, 22 (9.5%) were located at 4q11 → q21 (Figure 7). In addition, however, 17 grains (7.9%)

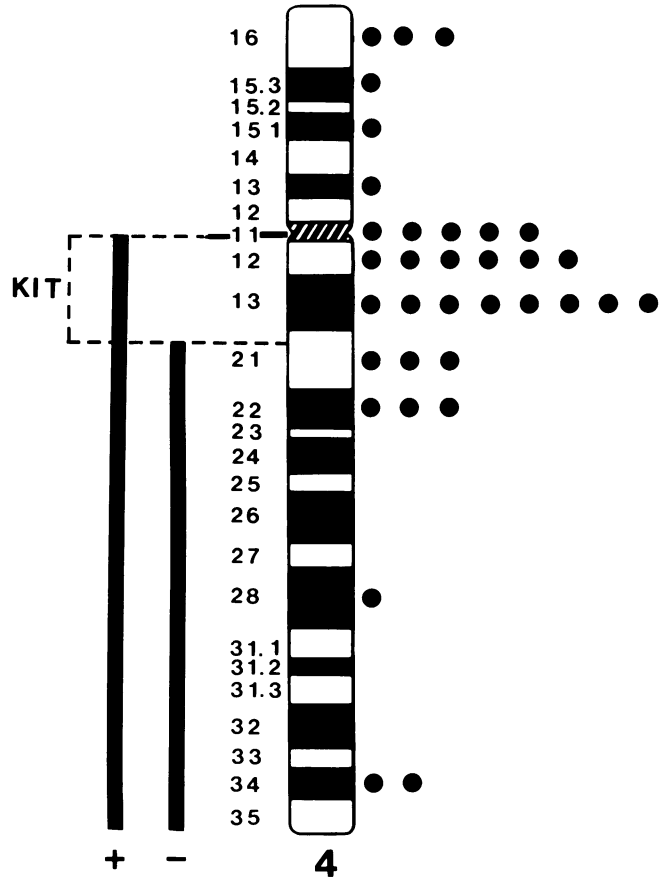

Fig. 7. Autoradiographic silver grain distribution on chromosome 4 after *in situ* hybridization with λHP-ckit9 cDNA insert (ideogram from ISCN, 1985). Left, vertical bars represent regions of low chromosome 4 present in hamster somatic cell hybrids that do (+) or do not (–) contain human c-*kit* sequences as determined by Southern blot analysis.

were found at a second site on chromosome 5, region q23 → q34. This second site of hybridization coincides with the location of related receptor genes, including the PDGF receptor (Yarden *et al.*, 1986) and c-*fms*/CSF-1 receptor (Groffen *et al.*, 1983; Le Beau *et al.*, 1986).

Southern blot analysis of DNA from 25 somatic cell hybrid clones with the 0.9-kb probe confirmed the localization of the human c-*kit* proto-oncogene to chromosome 4. The hybrids were derived from nine different series of Chinese hamster × human hybrids, one mouse × human and two rat × human hybrids (Yang-Feng *et al.*, 1985, 1986). *Eco*RI and *Pvu*II digested hybrid DNA was analyzed (not shown). Two human bands were detected in both *Pvu*II (9.0 kbp, 6.5 kbp) and *Eco*RI (6.3 kbp, 2.1 kbp) digests and were clearly distinguishable from Chinese hamster and mouse signals. Of the 25 hybrids, 11 were positive for both human *Pvu*II and *Eco*RI fragments. Correlation of the human c-*kit* restriction fragments with the presence or absence of human chromosomes in 25 rodent × human somatic cell hybrids revealed perfect synteny with chromosome 4, with all other chromosomes excluded by five or more discordant hybrids. Two hybrids contained parts of chromosome 4 translocated to a Chinese hamster chromosome (Brissenden *et al.*, 1984). The hybrid containing the long arm of chromosome 4 was positive for the human c-*kit* sequence and the hybrid containing only region 4q21 → 4qter was negative. These results place c-*kit* in the region cen → q21 of chromosome 4 (Figure 7).

For mapping of the mouse c-*kit* locus, a panel of 11 Chinese

hamster × mouse hybrids and one rat × mouse hybrid segregating mouse chromosomes was used. *Eco*RI and *Hin*dIII digested DNA immobilized on filters was hybridized with the human c-*kit* cDNA probe. Concordancy was found between c-*kit* and mouse chromosome 5, with all other mouse chromosomes excluded by two or more discordant hybrids (data not shown).

The assignments of c-*kit* genes to human and mouse chromosomes indicate their location within known syntenic groups. The conserved regions on human chromosome 4 and mouse chromosome 5 include the following gene loci: phosphoglucomutase-2 (*PGM2* in human, *Pgm-1* in mouse), peptidase S (*PEPS* in human, *Pep-7* in mouse), albumin (*ALB* in human, *Alb-1* in mouse) and α-fetoprotein (*AFP* in human, *Afp* in mouse). *PEPS, AFP* and *ALB* have been assigned to the promixal long arm of human chromosome 4 (cen → q13) and *PGM2* has been assigned to p14 → q13. Thus, it is likely that the conserved region is on the

A

```
c-kit     1  MRGARGAWDFLCVLL--LLLRVQT----GSSQPSVSPGEPSPPSIHPGKS
c-fms     1  M-G-PG------VLL--LLL-VATAWH-G-QGI--PVIEPSVPE------
PDGF-R  -31  M-GLPGVIPAL-VLRGQLLLSVL--WLLGPQTSRGLVITPPGPE------

c-kit    45  DLIVRVGDEIRLLCTDPGFVKWTFEILDETNENKQNEWI-------TEKA
c-fms    30  -LVVKPGATVTLRCVGNGSVEWDGPPS-PHWTLYS--DGS-SSILSTNNA
PDGF-R   10  -FVLNISSTFVLTCSGSAPVMWEQMSQVPWQEAAMNQDGTFSSVLTLTNV

c-kit    88  EATNTGKYTCTNKHG----LSNS--IYVFVRDPAKLFLVDRSLYGK-EDN
c-fms    75  TFQNTGTYRCTEPGD--P-LGGSAAIHLYVKDPARPWNVLAQEVVVFEDQ
PDGF-R   59  TGGDTGEYFCVYNNSLGPELSERKRIYIFVPDPTMGFLPMDSEDLFIFVT

c-kit   131  DTLVR---CPLTDPEVTNY-SLKGCQGKPLPKDLRFIPDPKAGIMIKSVK
c-fms   122  DAL---LPCLLTDPVLEAGVSLVRVRGRPLMRHTNYSFSPWHGFTIHRAK
PDGF-R  109  DVTETTIPCRVTDPQLEVTLHEKKVDI-PL--HVPY--DHQRGFTG---T

c-kit   177  RAYHRLCLHCSVDQEGKSVLSEKFILKVRPAFKAVPVVSVSKASYLLREG
c-fms   169  FIQSQDYQ-CSALMGGRKVMSISIRLKVQKVIPGPPALTLVPAELVRIRG
PDGF-R  151  FEDKT-YI-CKTTIGDREVDSDTYYVYSLQVSSINVSVNAVQTV-VRQ-G

c-kit   227  EEFTVTCTIKDVSSSVYSTWKRENSQTKLQEKYNSWHHGDFNYERQATLT
c-fms   218  EAAQIVCSASSVD-VNFDVFLQHNN-TKLAIPQQSDFHNNRYQKVL-TLN
PDGF-R  197  ESITIRCIVMGNDVVNFQWTYPRMKSGRLVEPVT-DYLFGVPSRIGSILH

c-kit   277  ISSARVNDSGVFMCYANNTFGSANVTTTLE-VVDK-GFINIFPMINTTVF
c-fms   265  LDQVDFQHAGNYSCVASN-VQGKHSTSMFFRVV-ESAYLNLSSEQNLIQE
PDGF-R  246  IPTAELSDSGTYTCNVSVSVNDHGDEKAINISVIENGYVRLLETLGDVEI

c-kit   325  VNDGENVDLIVE-YEAFPKPEHQQWIYMNRTFT--DKWEDYPKSENESNI
c-fms   313  VTVGEGLNLKVMV-EAYPGLQGFNWTYLGPFSDHQPEPKLANATTKDTY-
PDGF-R  296  AELHRSRTLRV-VFEAYPMPSVL-WLKDNRTLGDSGAGELVLSTRNMSET

c-kit   372  RYVSELHLTRLKGTEGGTYTFLVSNSDVNAAIAFN--V-YVNTKPEILTY
c-fms   361  RHTFTLSLPRLKPSEAGRYSFLARNPGGWRALTFELTLRYPPEVSVIWTF
PDGF-R  344  RYVSELILVRVKVSEAGYYTMRAFHEDDEVQLSFKLQVNVPVRVLELSES

c-kit   419  DRLVNGM-LQCVAAGFPEPTIDWYFCPGTEQRCSASVLPVDVQTLNSSGP
c-fms   411  INGSG--TLLCAASGYPQPNVTWLQCSGHTDRCDEAQVLQVWDDPYPEVL
PDGF-R  394  HPANGEQTIRCRGRGMPQPNVTWSTCRDLK-RCPRKLSPTPLGNSSKEE-

c-kit   468  -----PFGKL-VVQSSIDSSAFK--HN---GTVECKAYNDVGKTSAYFNF
c-fms   459  SQ-E-PFHKVT-VQSLLTVETLE--HN-Q--TYECRAHNSVGSG-SWAFI
PDGF-R  442  SQLETNVTFWEEDQEYEVVSTLRLRHVDQPLSVRCMLQNSMG-GDSQEV-

c-kit   507  AFKGNNKEQIHPHT
c-fms   500  PISAGAHTHP-PDE
PDGF-R  490  --TVVPHSLP-FK
```

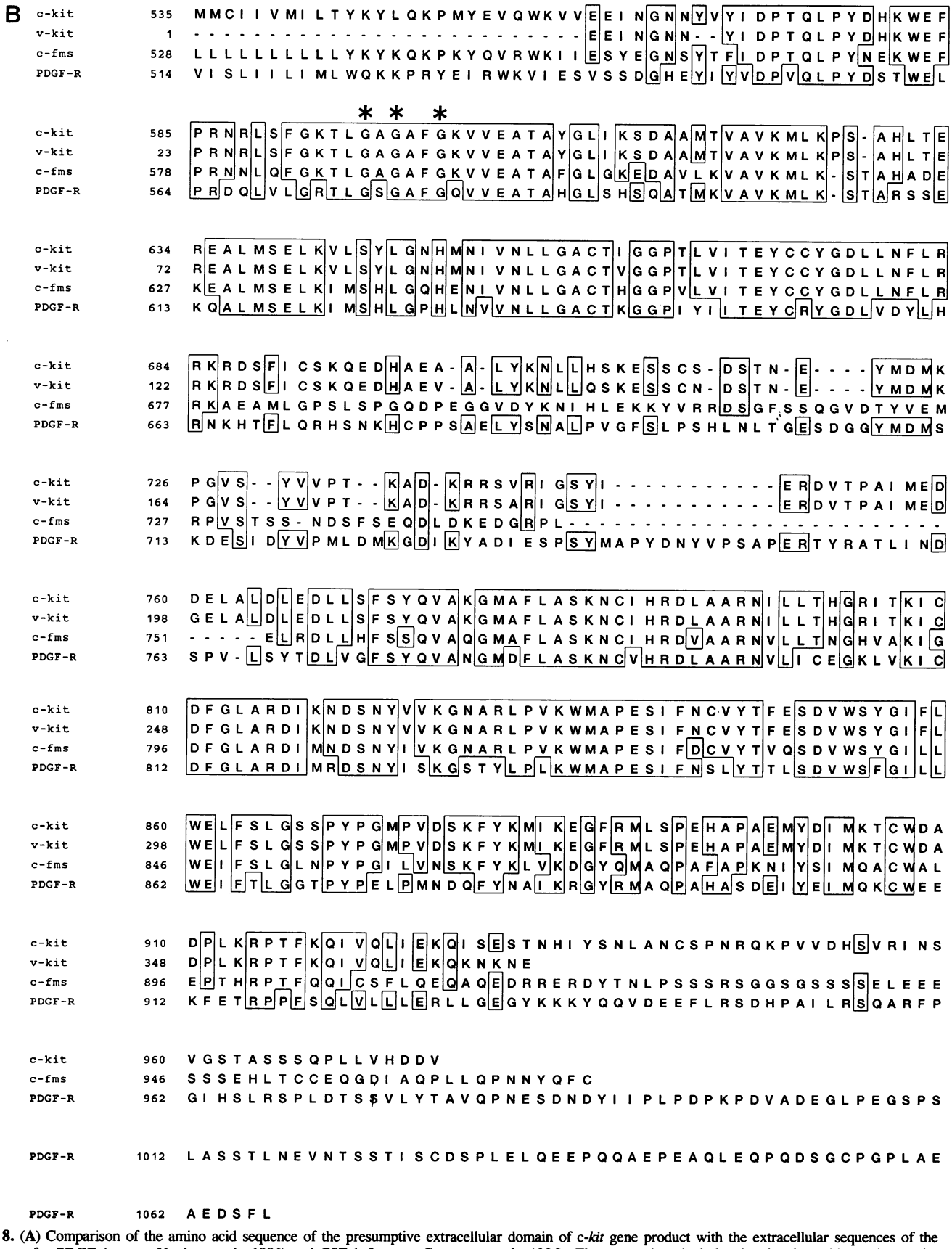

Fig. 8. **(A)** Comparison of the amino acid sequence of the presumptive extracellular domain of c-*kit* gene product with the extracellular sequences of the receptors for PDGF (mouse; Yarden *et al.*, 1986) and CSF-1 (human; Coussens *et al.*, 1986). The comparison includes the signal peptides and extends to the transmembrane domains. For better alignment, gaps (hyphens) were introduced. Conserved amino acids are boxed and cysteine residues are shaded. **(B)** Comparison of cytoplasmic domain of c-*kit* amino acid sequence with its viral homologue v-*kit* and the cytoplasmic domains of the receptors for PDGF and CSF-1. Only the cytoplasmic domains of the receptors starting with their transmembrane sequences are shown. Residues which are identical in three or more of the proteins are boxed, and gaps (shown by hyphens) were introduced for optimal alignment. Asterisks indicate the location of the nucleotide binding site consensus sequence Gly-X-Gly-X-X-Gly.

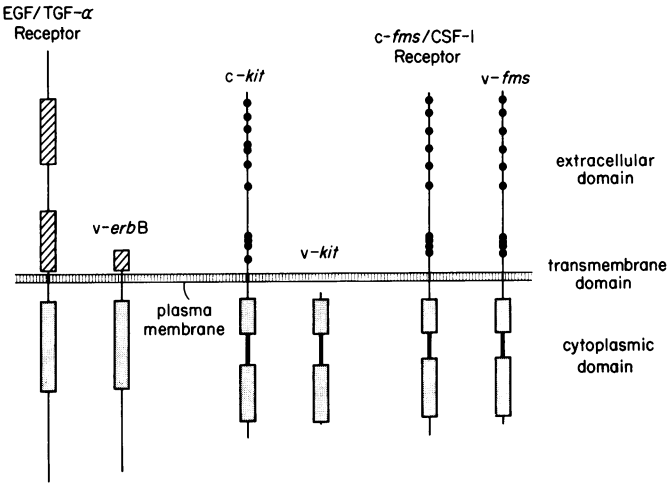

Fig. 9. Schematic comparison of receptor–oncogene product pairs EGFR/*erbB*, c-*kit*/v-*kit* and c-*fms*/v-*fms*. Tyrosine kinase domains (shaded), cysteine-rich regions (hatched) and single Cys residues (black circles) are indicated.

proximal long arm of human chromosome 4. Genes that map more distally on 4q, such as loci for alcohol dehydrogenase and EGF, belong to a syntenic group on mouse chromosome 3 (Lalley and McKusick, 1985). On mouse chromosome 5, the loci *Pgm-1*, *Pep-7*, *Afp* and *Alb-1*, are located in the central part of the chromosome spanning a region of 11 cM. Therefore, even though we have not physically mapped the c-*kit* locus on mouse chromosome 5, we would expect it to reside in the vicinity of these other conserved genes.

Discussion

We have reported here the identification and characterization of c-*kit*, the human cellular homologue of the HZ4 feline sarcoma virus oncogene v-*kit*. Isolation and characterization of cloned human c-*kit* cDNA showed that the c-*kit* polypeptide is synthesized by translation of a single 5-kb mRNA, which contains an open reading frame coding for a 976 amino acid polypeptide. Comparison of the c-*kit* structural features with previously characterized growth factor receptors strongly suggests that c-*kit* encodes a transmembrane receptor for an as yet unidentified ligand of unknown function. The c-*kit* gene product is highly homologous to the receptors for PDGF and CSF-1, which comprise a structural subtype within the family of RTKs. Analogous to these receptors, it is likely that c-*kit* utilizes its ligand-induced tyrosine kinase activity to generate an intracellular signal that eventually triggers mitogenesis. Expression of the endogenous c-*kit* gene in human glioblastoma cells generates a 145 000 M_r surface glycoprotein $p145^{c\text{-}kit}$ that possesses autophosphorylation activity. The putative cell surface receptor polypeptide $p145^{c\text{-}kit}$ is also expressed by mouse NIH-3T3 cells containing amplified mammalian expression plasmids with c-*kit* cDNA sequences under SV40 early promoter control.

Comparative structural analysis of human c-*kit* suggests that a 23 amino acid signal sequence precedes an amino-terminal, putative extracellular ligand binding domain of 497 amino acid residues. The presence of nine of 11 potential N-linked glycosylation sites in this region of c-*kit* further supports this assignment. A low level of overall sequence homology is found between this c-*kit* domain and corresponding sequences of the CSF-1 (27%) and PDGF (19%) receptors (Figure 8A), as expected for structures defining specific binding of different ligands. Furthermore, the extracellular domains of these three proteins share a common pattern of cysteine distribution (Figure 8A) that is distinct from the pattern of cysteine-rich domains found in other known RTKs such as the EGFR (Ullrich *et al.*, 1984).

A hydrophobic stretch of 23 residues that presumably serves to anchor the c-*kit* polypeptide in the plasma membrane (Figures 3 and 4) connects the extracellular ligand binding domain with sequences containing the tyrosine kinase enzymatic domain and other structures involved in generation, control, and possibly modulation of an intracellular signal. The cytoplasmic portion of c-*kit* shows a subdomain organization that is highly homologous to its closest relatives, the c-*fms*/CSF-1 receptor and the PDGF receptor. The 52 amino acid residues closest to the membrane are ~70% homologous in sequence to the closest relative of c-*kit*, the c-*fms*/CSF-1 receptor (Figure 8B). In the EGFR, this region contains a threonine residue (Thr 654) which plays a role in protein kinase C-mediated modulation of receptor activity (Hunter *et al.*, 1984; Downward *et al.*, 1985). No obviously homologous threonine residue is found in the c-*kit* polypeptide or in CSF-1 or PDGF receptors. Nevertheless, the CSF-1 receptor down-modulates in response to phorbol ester (Wheeler *et al.*, 1986), suggesting that this subfamily of RTKs contains alternative target sites for protein kinase C phosphorylation.

The portion of the cytoplasmic domain that includes the ATP binding site and tyrosine-specific kinase activity is highly conserved in the tyrosine kinase family of growth factor receptors and oncogenes. However, in c-*kit* as well as in the related receptors for CSF-1 and PDGF, this domain is bisected by unique insertion sequences that measure 77 amino acids in c-*kit* and ~70 and 100 residues in the CSF-1 and PDGF receptors respectively. All three inserted sequences are highly hydrophilic and could therefore form a surface structure important for receptor-specific protein–protein interactions, such as binding of distinct substrates. That these inserted sequences play an important role in signal generation is supported by their remarkable conservation: this region of the human and feline *kit* products show 94% homology, whereas human and feline *fms* sequences are 89% homologous in this region (Coussens *et al.*, 1986). Similarly, the distal carboxy-terminal sequences, found to be generally hydrophilic and highly divergent in previously characterized RTKs, show the same characteristics in human c-*kit*. We speculate that this region may play a role in determining either the specificity of receptor signals or in regulating the enzymatic activity of the tyrosine kinase domain.

Characterization of the cellular homologue of the oncogene v-*kit* as coding for a potential growth factor receptor provides another example of conversion of a gene involved in normal growth control into an oncogene. When compared with its human receptor-like homologue, the feline v-*kit* has lost the entire extracellular ligand binding domain, the transmembrane domain and 17 amino acids from the juxtamembrane region (Figures 3, 4 and 9). The transduction event that generated the HZ4-FeSV resulted in another sequence deletion at the carboxy terminus where 49 residues were replaced by five amino acids as a consequence of sequence fusion with the FeLV polymerase gene. These characteristics of the v-*kit*/c-*kit* system resemble those found in the v-*erbB*/EGFR oncogene-receptor pair (Figure 9). In both cases, elimination of most or all of the ligand binding domain sequences may lead to deregulation and possibly constitutive activation of the tyrosine kinase activity. In v-*fms*, where the extracellular ligand binding domain is still intact (Figure 9), autocrine stimulation of the altered cytoplasmic domain may be necessary to generate an aberrant growth signal resulting in cell

transformation (Roussel *et al.*, 1987).

Surface association of viral tyrosine kinases seems to be essential for their transforming ability (Roussel *et al.*, 1984; Pellman *et al.*, 1985). In contrast to v-*erbB* and v-*fms*, which are inserted in the plasma membrane via their own membrane-spanning sequences, v-*kit* lost its transmembrane domain and much of its amino terminus. Therefore, it appears likely that the fusion protein gag-kit associates with cellular membranes via gag-linked myristilation (Sefton *et al.*, 1982; Schultz *et al.*, 1985; Besmer *et al.*, 1986).

Deletion of carboxy-terminal sequences is the only distinctive feature thus far identified that is shared by all characterized proto-oncogene−viral oncogene pairs of the tyrosine kinase type, including EGFR/v-*erbB* (Ullrich *et al.*, 1984), c-*fms*/v-*fms* (Coussens *et al.*, 1986) and c-*src*/v-*src* (Takeya and Hanafusa, 1983). In the case of *erbB* and *src*, the deletions remove tyrosine residues whose phosphorylation appears to regulate receptor kinase activity (Downward *et al.*, 1984b; Iba *et al.*, 1985; Courtneidge, 1985; Cooper *et al.*, 1986). For the CSF-1R/v-*fms* system, a 40 amino acid deletion in v-*fms* that includes removal of tyrosine 969 may be a crucial event in the generation of a transforming gene from the normal receptor gene (Roussel *et al.*, 1987). Analogously, the 49-residue deletion found in the c-*kit*/v-*kit* pair includes a tyrosine in c-*kit* at position 936, which may represent a regulatory autophosphorylation site.

In addition to deletions, point mutations are also known to play a role in the generation and enhancement of the *src* gene transforming potential (Levy *et al.*, 1986) and an analogous situation has been suggested for the *fms* system (Roussel *et al.*, 1987). In another receptor-like tyrosine kinase, the rat *neu* gene product, a single conversion of a valine to a glutamic acid residue in the transmembrane domain has been shown to unmask the transforming potential of this molecule (Bargmann *et al.*, 1986b). Human c-*kit* and feline v-*kit* sequences display only six, mostly conservative, sequence differences and a two amino acid comparative deletion (Figure 3). Some of these minor sequence differences may represent mutations involved in activation of the v-*kit* transforming activity rather than species-specific neutral differences.

In addition to providing new insight into strategies of development of oncogenic potential, the isolation, characterization and expression of c-*kit* sequences offers the opportunity to unravel a new biological communication system by isolating an as yet unknown polypeptide ligand and investigating its function in normal growth and development.

Materials and methods

Isolation and characterization of cDNA clones

Three regions of the v-*kit* (Besmer *et al.*, 1986) nucleotide sequence were selected for the synthesis of 90−96 nucleotide-long probes ON-1 (v-*kit* 1020−1110), ON-2 (v-*kit* 1405−1500) and ON-3 (v-*kit* 2041−2130). Oligonucleotide probes were labeled by phosphorylation with T4 polynucleotide kinase (Pharmacia) and [γ-^{32}P]ATP (Amersham; >5000 Ci/mmol) and used to screen human placental and fetal brain cDNA libraries (~2 × 10^6 clones each) in λgt10 as described previously (Mostov *et al.*, 1984). Hybridization to nitrocellulose filter replicas of recombinant λgt10 plaques was carried out at low stringency conditions (12 h at 42°C in 5 × SSC, 20% formamide, 5 × Denhardt's solution). Filters were subsequently washed at 42°C with 0.2 × SSC, 0.1% SDS. Phage plaque purification and cDNA preparation was carried out as described elsewhere (Maniatis *et al.*, 1982). Nucleotide sequence analysis was carried out by subcloning into either M13 vectors or pUC119 followed by primed DNA synthesis on single-stranded DNA templates in the presence of dideoxynucleotide triphosphates (Sanger *et al.*, 1977; Messing *et al.*, 1981).

Northern blot hybridization analysis

RNA was isolated by the guanidine−monothiocyanate−LiCl method (Cathala *et al.*, 1983) and chromatographed on oligo(dT) cellulose. Five micrograms of poly(A)$^+$ RNA were denatured at 65°C in the presence of formamide and formaldehyde and then subjected to electrophoresis on a vertical 1.2% agarose gel in the presence of formaldehyde (Maniatis *et al.*, 1982) and thereafter blotted to nitrocellulose filter with 10 × SSC. Blots were prehybridized for 4 h at 42°C in 5 × SSC, 50% formamide and salmon sperm DNA (50 μg/ml) and hybridized with ^{32}P-labeled cDNA (λHP-ckit9 insert) under the same conditions for 14 h. The filters were then washed at 45°C with 0.2 × SSC and 0.1% SDS.

Rabbit c-kit peptide antiserum

Antiserum LJ11 was obtained after immunization with the peptide GSTASSSQPLLVHDDV, a sequence located at the C terminus of human c-*kit* gene product. The peptide was synthesized using the solid phase method of Barany and Merrifield (1980). An amino-terminal cysteine residue was added to the original sequence to enable thioester linkage to soybean trypsin inhibitor at a ratio of 5−10 mol peptide/mol carrier protein. Rabbits were injected intradermally with 1 mg conjugate emulsified with complete Freund's adjuvant. Additional injections of 200 μg of the immunogen in incomplete Freund's adjuvant were given subcutaneously after 4, 6 and 8 weeks and serum obtained prior to each injection. Serum was evaluated by solid phase radioimmunoassay for antibody titer. Positive antibodies to the carrier protein were found in all bleeds. Antipeptide antibodies appeared afer 6 weeks, but reaction with the c-*kit* gene product was not positive before the 8th week's bleeding.

c-kit expression vector and transfection of NIH-3T3 cells

A pUC119/hckit subclone containing sequences from nucleotide 1−4769 of the cDNA sequence (Figure 3) was constructed by combining cDNA fragments of clones λHP-ckit/171, λHP-ckit63 and λHP-ckit/1 in pUC119. A 4804 bp *Bam*HI fragment containing the entire human c-*kit* insert from this plasmid was inserted into the *Bam*HI site of a mammalian expression vector (Rosenthal *et al.*, 1986), which results in hc-*kit* expression under the control of the SV40 early (SVE) promoter. The vector also includes genes for gentamycin resistance and dihydrofolate reductase (Simonsen and Levinson, 1983) under SVE control, and sequences from *Escherichia coli* plasmid pML (Lusky and Botchan, 1981) allowing for replication and selection in *E. coli*.

NIH-3T3 cells were grown in Dulbecco's modified Eagle's medium (DMEM) with 10% fetal calf serum (FCS). Five micrograms of cloned plasmid were used to transfect a 10-cm plate of cells (5 × 10^5 cells) by calcium phosphate precipitation. Next day the cells were seeded in one 10-cm dish and maintained in DMEM supplemented with 10% FCS and 50 μg/ml gentamycin (G418) for 2−3 weeks. The cells were then maintained in DMEM:F-12 (50:50) containing 10% dialyzed FCS and 250 nM methotrexate for 3 weeks before transferring to DMEM:F-12 containing 10% dialyzed FCS and either 250 nM or 500 nM methotrexate which resulted in cell lines CVNhckit-N250 and CVNhckit-N500 respectively. NIH-3T3 control cells were transfected with the CVN expression vector only and grown in 500 nM methotrexate containing serum.

Immunoprecipitations

A172 human glioblastoma cells were grown to confluence in 90-mm dishes. The cells were washed with methionine- and cysteine-free DME medium and then labeled for 18 h in the same medium containing 100 μCi of [^{35}S]methionine and 50 μCi of [^{35}S]cysteine (Amersham) per ml, and supplemented with 10% FCS, which was dialyzed against phosphate-buffered saline (PBS). The labeled cells were washed twice with ice-cold PBS which contained 1 mM EGTA and scraped by means of a rubber policeman in the same buffer. Cells were pelleted by short centrifugation and then solubilized in 1 ml of a modified solubilization buffer containing 50 mM Hepes, pH 7.5, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1% Na-deoxycholate, 0.1% SDS, 1.5 mM $MgCl_2$, 5 mM EGTA, 2 mM benzamidine, 2 mM PMSF, 1% aprotinin (0.15 TIU/ml) and 10 μg/ml leupeptin. Solubilized cells were centrifuged (12 000 *g*) for 10 min at 4°C. The supernatant was diluted with 50 mM Hepes, pH 7.5, 100 mM NaCl, 10% glycerol, 0.1% bovine serum albumin. Equal aliquots were then mixed with an adsorbing antibody bound to Sepharose−Protein A (Pharmacia). The adsorbing antibody used was generated in rabbit by exactly the same protocol used to generate the specific anti-c-*kit* antibody except that a completely unrelated sequence was used for the immunizing peptide. The reaction with the adsorbing antibody was terminated after 1 h at 4°C, and the unadsorbed material was mixed with immune or preimmune serum (7 μl and 12 μl respectively) bound to Sepharose−Protein A. Immune complexes were allowed to form for 2 h at 4°C with constant agitation. Immunoprecipitates were washed twice with 1 ml of each of the following solutions: modified solubilization buffer, high-salt buffer (50 mM Tris−HCl, pH 8.0, 500 mM NaCl, 0.1% SDS, 0.2% Triton X-100, 5 mM EGTA), medium-salt buffer (50 mM Tris−HCl, pH 8.0, 150 mM NaCl, 7.5 mM EDTA, 0.1% SDS, 0.2% Triton X-100) and low-salt buffer (10 mM Tris−HCl, pH 8.0, 0.1% Triton X-100). To allow dissociation of loosely attached proteins, a 30-min incubation at 4°C in 1 ml low-salt buffer was performed and followed by replacement of the tube and final wash. The washed beads were heated at 95°C for 5 min after the addition of sample buffer. The samples were analzyed by gel electrophoresis on 7.5% polyacrylamide gels containing SDS. The fixed and stained

gel was treated for 1 h with ENHANCE (New England Nuclear), dried and autoradiographed on Agfa X-ray film for 60 h at −70°C.

NIH-3T3 cells transfected with CVNhckit and maintained in DMEM:F-12 supplemented with 10% dialzyed FCS and either 250 nM methotrexate or 500 nM methotrexate were grown in 100-mm dishes. Cells were washed with PBS and then labeled for 16 h at 37°C in methionine-free MEM containing 100 μCi/ml [^{35}S]methionine and 5% FCS. The labeled cells were washed with ice-cold PBS and lysed in solubilization buffer [50 mM Hepes, pH 7.5, 0.15 M NaCl, 1 mM EDTA, 1 mM EGTA, 1.5 mM $MgCl_2$, 10% glycerol, 1% Triton X-100, 2 mM PMSF, 1% aprotinin (0.15 TIU/ml)]. The cell lysates were centrifuged for 2 min and the supernatants subjected to immunoprecipitation. The cell lysates were mixed with Protein A−Sepharose (3 mg dry wt), which were precoupled to 10 μl preimmune serum and incubated for 1 h at 4°C with constant agitation. The beads were pelleted and the supernatant was incubated with 5 μl of the anti-c-*kit* antibody for 2 h at 4°C with constant agitation. Sepharose− Protein A (4 mg dry wt) was added and immune complexes were allowed to form for 1 h at 4°C with constant agitation. Immunoprecipitates were washed seven times with 20 mM Hepes, pH 7.5, 0.15 M NaCl, 0.1% Triton X-100, 10% glycerol (HNTG). The washed beads were heated at 95°C for 5 min after the addition of sample buffer. The samples were then analyzed by polyacrylamide gel electrophoresis (PAGE).

A172 cell membrane preparation

A172 cells were grown to confluence in 15-cm diameter dishes. Cell monolayers were washed twice with PBS containing 1 mM EGTA. The cells were scraped by means of a rubber policeman into the same buffer and pelleted by 10 min centrifugation at 1000 *g*. The pellet was resuspended in 10 volumes of hypotonic buffer (20 mM Hepes, pH 7.5, 1.5 mM $MgCl_2$, 1 mM PMSF, 2 mM benzamidine, 1 mM EGTA, 10 μg/ml leupeptin and aprotinin at 0.2 TIU/ml). Cells were left to lyse for 10 min at 4°C and then disrupted with 12 strokes in a Dounce homogenizer. The lysate was centrifuged at 1000 *g* for 10 min, and the resulting supernatant centrifuged for 30 min at 40 000 *g*. The microsomal pellet was resuspended in isotonic solution (20 mM Hepes, pH 7.5, 150 mM NaCl, 1 mM EGTA), briefly homogenized in a Dounce homogenizer and loaded on top of a 35% (w/v) sucrose cushion (in PBS). The interface fraction was collected after 30 min centrifugation at 24 000 *g*, diluted in 10 volumes of 10 mM Hepes solution (pH 7.3) and re-centrifuged for 30 min at 40 000 *g*. The final membrane pellet was resuspended in 10 mm Hepes buffer (pH 7.5) and stored at −70°C.

Immune complex kinase assays

Isolated plasma membranes (25 μg) from A172 cells were dissolved in 0.25 ml of the solubilization buffer (see above) and cleared by 10 min centrifugation at 12 000 *g* at 4°C. The supernatant was brought to 1 ml with dilution buffer (20 mM Hepes, pH 7.5, 150 mM NaCl, 1 mM EGTA) and mixed with Protein A− Sepharose beads (5 mg dry wt), which were pre-coupled to 7 μl of the antiserum. For NIH-3T3 cells transfected with CVNhckit, confluent cells were washed twice with ice-cold PBS and lysed in solubilization buffer. The cell lysate was centrifuged for 3 min and the supernatant was mixed with Protein A− Sepharose beads (3 mg dry wt), which were pre-coupled to 7 μl of the antiserum. Immune complexes were allowed to form for 1 h at 4°C under agitation, and then washed by centrifugation and resuspension (four times) in HNTG buffer. The final pellet was resuspended in 50 μl of the HNTG buffer which contained 15 mM $MnCl_2$, 50 μCi [γ-^{32}P]ATP (New England Nuclear) and non-labeled ATP to yield the indicated concentrations. Samples were incubated at 22°C for 30 min, and the immune complexes washed once with high-, medium- and low-salt buffers containing 1 mM ATP. The final pellet was resuspended in 2× electrophoresis sample buffer, heated at 95°C for 5 min and analyzed by PAGE.

Phosphoamino acid analysis

The protocol of Kobayashi *et al.* (1981) was used with slight modifications. All glassware was siliconized prior to use. ^{32}P-Labeled protein bands were cut out of the dry gel, incubated for 4 h at 37°C with 10% methanol, and then freeze-dried. The dry acryalmide piece was crushed briefly and hydrated in 75 μl of 1 mg/ml tryspin solution in 50 mM NH_4HCO_3 buffer (pH 8.0). One milliliter of the same buffer was added and trypsin allowed to cleave the protein at 37°C for 10 h. The mixture was then centrifuged in an Eppendorf centrifuge and the supernatant was freeze-dried. After lyophilization, samples were subjected to acid hydrolysis in 0.2 ml 6 N HCl at 110°C for 2 h under N_2 gas. An equal volume of water was added and the samples were lyophilized again. Samples of the acid hydrolysates were then mixed with 2 μl solution of phosphoserine, phosphothreonine and phosphotyrosine (1 mg/ml each) and spotted onto cellulose-coated plastic plates (20 × 20 cm, Merck). Thin layer electrophoresis of the samples was carried out at 4°C at pH 3.5 (acetic acid/pyridine/H_2O, 10:2:88, by volume) for 60 min at 1000 V. Ninhydrin (0.3% in acetone) was used to visualize the phosphoamino acids on the thin layer plates.

^{125}I surface iodination

Transfected NIH-3T3 cells were grown to confluency in 10-cm plates and surface-labeled by Na^{125}I (Amersham, 1.4 mCi/plate) by the lactoperoxidase− glucose oxidase system (Enzymobead, BioRad) as recommended by the manufacturer. Following iodination, cells were washed, lysed and subjected to immunoprecipitation essentially as described above.

Acknowledgements

We thank John Burnier for synthesizing the c-*kit* peptide, and Mark Vasser, Parkash Jhurani and Peter Ng for oligonucleotide synthesis. We acknowledge Sanjoy Sundaresan for his valuable contributions to the chromosomal localization of the mouse c-*kit* gene. We are grateful to Suzanne Pfeffer and Charles Sherr for editorial suggestions, Jeanne Arch for preparation of this manuscript, and Alane Gray and James Lee for preparing the figures. The chromosome mapping work was supported by NIH grant GM26105 (to U.F.).

References

Barany,G. and Merrifield,R.B. (1980) In *Gross,E. and Meienhofer,J. (eds), The Peptides II.* Academic Press, New York, pp. 1−284.

Bargmann,C.I., Hung,M.-C. and Weinberg,R.A. (1986a) *Nature,* **319**, 226−230.

Bargmann,C.I., Hung,M.-C. and Weinberg,R.A. (1986b) *Cell,* **45**, 649−657.

Besmer,P., Murphy,J.E., George,P.C., Qiu,F., Bergold,P.J., Lederman,L., Snyder,H.W.,Jr, Brodeur,D., Zuckerman,E.E. and Hardy,W.D. (1986) *Nature,* **320**, 415−421.

Brissenden,J.E., Ullrich,A. and Francke,U. (1984) *Nature,* **310**, 781−784.

Cathala,G., Savouret,J.-F., Mendez,B., West,B.L., Karin,M., Martial,J.A. and Baxter,J.D. (1983) *DNA,* **2**, 329−335.

Cooper,J.A., Gould,K.L., Cartwright,C.A. and Hunter,T. (1986) *Science,* **231**, 1431−1434.

Courtneidge,S. (1985) *EMBO J.,* **4**, 1471−1477.

Coussens,L., Yang-Feng,T.L., Liao,Y.-C., Chen,E., Gray,A., McGrath,J., Seeburg,P.H., Libermann,T.A., Schlessinger,J., Francke,U., Levinson,A. and Ullrich,A. (1985) *Science,* **230**, 1132−1139.

Coussens,L., Van Beveren,C., Smith,D., Chen,E., Mitchell,R.L., Isacke,C.M., Verma,I.M. and Ullrich.A. (1986) *Nature,* **320**, 277−280.

Downward,J., Yarden,Y., Mayes,E., Scrace,G., Totty,N., Stockwell,P., Ullrich, A., Schlessinger,J. and Waterfield,M.D. (1984a) *Nature,* **307**, 521−527.

Downward,J., Parker,P. and Waterfield,M.D. (1984b) *Nature,* **311**, 483−485.

Downward,J., Parker,P. and Waterfield,M.D. (1985) *J. Biol. Chem.,* **260**, 14538−14546.

Ebina,Y., Ellis,L., Jarnagin,K., Edery,M., Graf,L., Clauser,E., Ou,J.-H., Masiarz,F., Kan,Y.W., Goldfine,I.D., Roth,R.A. and Rutter,W.J. (1985) *Cell,* **40**, 747−758.

Groffen,J., Heisterkamp,N., Spurr,N., Dana,S., Wasmuth,J.J. and Stephenson, J.R. (1983) *Nucleic Acids Res.,* **11**, 6331−6339.

Hannink,M. and Donoghue,D.J. (1985) *Proc. Natl. Acad. Sci. USA,* **82**, 7894−7898.

Human Gene Mapping 8 (1985) *Cytogenet. Cell Genet.,* **40**, 1−823.

Hunter,T., Ling,N. and Cooper,J.A. (1984) *Nature,* **311**, 480−483.

Huynh,T., Young,R. and Davis,R. (1985) In Glover,D. (ed.), *DNA Cloning: a Practical Approach.* IRL Press Ltd, Oxford, Vol. 1, pp. 49−78.

ISCN (1985) *An International System for Human Cytogenetic Nomenclature.* S.Krager, Basel.

Iba,H., Cross,F., Garber,E.A. and Hanafusa,H. (1985) *Mol. Cell. Biol.,* **5**, 1058−1066.

Kobayashi,N., Tanaka,A. and Kaji,A. (1981) *J. Biol. Chem.,* **256**, 3053−3058.

Kozak,M. (1984) *Nucleic Acids Res.,* **12**, 857−872.

Kris,R.M., Lax,I., Gullick,W., Waterfield,M.D., Ullrich,A., Fridkin.,M. and Schlessinger,J. (1985) *Cell,* **40**, 619−625.

Kyte,J. and Doolittle,R.F. (1982) *J. Mol. Biol.,* **157**, 105−132.

Lalley,P.A. and McKusick,V.A. (1985) *Cytogenet. Cell Genet.,* **40**, 536−566.

Le Beau,M.M., Westbrook,C.A., Diaz,M.O., Larson,R.A., Rowley,J.D., Gasson,J.C., Golde,D.W. and Sherr,C.J. (1986) *Science,* **231**, 984−987.

Levy,J.B., Iba,H. and Hanafusa,H. (1986) *Proc. Natl. Acad. Sci. USA,* **83**, 4228−4232.

Lusky,M. and Botchan,M. (1981) *Nature,* **293**, 79−81.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.

Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.,* **9**, 309−321.

Mostov,K.E., Friedlander,M. and Blobel,G. (1984) *Nature,* **318**, 37−43.

Pellman,D., Garber,E.A., Cross,F.R. and Hanafusa,H. (1985) *Nature,* **314**, 374−377.

Rosenthal,A., Lindquist,P.B., Bringman,T.S., Goeddel,D.V. and Derynck,R. (1986) *Cell,* **46**, 301−309.

Roussel,M.F., Rettenmier,C.W., Look,A.T. and Sherr,C.J. (1984) *Mol. Cell. Biol.,* **4**, 1999−2009.

Roussel,M.F., Dull,T.J., Rettenmier,C.W., Ralph,P., Ullrich,A. and Sherr,C.J. (1987) *Nature,* **325**, 549−552.

Sacca,R., Stanley,E.R., Sherr,C.J. and Rettenmier,C.W. (1986) *Proc. Natl. Acad. Sci. USA,* **83**, 3331−3335.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schultz,A.M., Henderson,L.E., Oroszlan,S., Garber,E.A. and Hanafusa,H. (1985) *Science*, **227**, 427–429.
Sefton,B.M., Trowbridge,I.S. and Cooper,J.A. (1982) *Cell*, **31**, 465–474.
Sherr,C.J., Rettenmier,C.W., Sacca,R., Roussel,M.F., Look,A.T. and Stanley, E.R. (1985) *Cell*, **41**, 665–676.
Simonsen,C.S. and Levinson,A.D. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 2495–2499.
Takeya,T. and Hanafusa,H. (1983) *Cell*, **32**, 881–890.
Ullrich,A., Coussens,L., Hayflick,J.S., Dull,T.J., Gray,A., Tam,A.W., Lee,J., Yarden,Y., Libermann,T.A., Schlessinger,J., Downward,J., Bye,J., Whittle, N., Waterfield,M.D. and Seeburg,P.H. (1984) *Nature*, **309**, 418–425.
Ullrich,A., Bell,J.R., Chen,E.Y., Herrera,R., Petruzzelli,L.M., Dull,T.J., Gray, A., Coussens,L., Liao,Y.-C., Tsubokawa,M., Mason,A., Seeburg,P.H., Grunfeld,C., Rosen,O.M. and Ramachandran,J. (1985) *Nature*, **313**, 756–761.
Ullrich,A., Gray,A., Tam,A.W., Yang-Feng,T., Tsubokawa,M., Collins,C., Henzel,W., Le Bon,T., Kathuria,S., Chen,E., Jacobs,S., Francke,U., Ramachandran,J. and Fujita-Yamaguchi,Y. (1986) *EMBO J.*, **5**, 2503–2512.
Watson,M.E.E. (1984) *Nucleic Acids Res.*, **12**, 5145–5164.
Wheeler,E.F., Rettenmier,C.W., Look,A.T. and Sherr,C.J. (1996) *Nature*, **324**, 377–380.
Yamamoto,T., Ikawa,S., Akiyama,T., Semba,K., Nomura,N., Miyajima,N., Saito,T. and Toyoshima,K. (1986) *Nature*, **319**, 230–234.
Yang-Feng,T.L., Floyd-Smith,G., Nemer,M., Drouin,J. and Francke,U. (1985) *Am. J. Hum. Genet.*, **37**, 1117–1128.
Yang-Feng,T.L., Landlu,N.R., Baltimore,D. and Francke,U. (1986) *Cytogenet. Cell Genet.*, **43**, 121–126.
Yarden,Y., Escobedo,J.A., Kuang,W.-J., Yang-Feng,T.L., Daniel,T.O., Tremble,P.M., Chen,E.Y., Ando,M.E., Harkins,R.A., Francke,U., Fried,V.A., Ullrich,A. and Williams,L.T. (1986) *Nature*, **323**, 226–232.

Received on July 28, 1987