achieved by the very efficient packing of filaments by fimbrin (0.4 mM) and/or villin (0.6 mM) into a bundle more likely to have a supportive function within the cell than a locomotive one.

Received 9 July; accepted 8 November 1982.

1. Matsudaira, P. T. & Burgess, D. R. *J. Cell Biol.* **92,** 657–664 (1982).
2. Mooseker, M. S. & Tilney, L. G. *J. Cell Biol.* **67,** 725–743 (1975).
3. Bretscher, A. & Weber, K. *Expl Cell Res.* **116,** 397–407 (1978).
4. Matsudaira, P. T. & Burgess, D. R. *J. Cell Biol.* **83,** 667–673 (1979).
5. Bretscher, A. & Weber, K. *Cell* **20,** 839–847 (1980).
6. Howe, C. L., Mooseker, M. S. & Graves, T. A. *J. Cell Biol.* **85,** 916–923 (1980).
7. Glenney, J. R. & Weber, K. *J. biol. Chem.* **255,** 10551–10554 (1980).
8. Mooseker, M. S., Graves, T. A., Wharton, K. A., Falco, N. & Howe, C. L. *J. Cell Biol.* **87,** 809–822 (1980).
9. Matsudaira, P. T. & Burgess, D. R. *J. Cell Biol.* **92,** 648–656 (1982).
10. Glenney, J. R., Kaulfus, P., Matsudaira, P. & Weber, K. *J. biol. Chem.* **256,** 9283–9288 (1981).
11. Bretscher, A. *Proc. natn. Acad. Sci. U.S.A.* **78,** 6849–6853 (1982).
12. Hanson, J. *Nature* **213,** 353–356 (1967).
13. Moore, P. B., Huxley, H. E. & DeRosier, D. J. *J. molec. Biol.* **50,** 279–295 (1970).
14. O'Brien, E. J., Bennett, P. M. & Hanson, J. *Phil. Trans. R. Soc.* B**261,** 201–208 (1971).
15. O'Brien, E. J., Gillis, J. M. & Couch, J. *J. molec. Biol.* **99,** 461–475 (1975).
16. Gillis, J. M. & O'Brien, E. J. *J. molec. Biol.* **99,** 445–459 (1975).
17. DeRosier, D., Mandelkow, E., Silliman, A., Tilney, L. G. & Kane, R. E. *J. molec. Biol.* **113,** 679–695 (1977).
18. Spudich, J. A. & Amos, L. A. *J. molec. Biol.* **129,** 319–331 (1979).
19. Tilney, L. G., DeRosier, D. J. & Mulroy, M. J. *J. Cell Biol.* **86,** 244–259 (1980).
20. DeRosier, D. & Censullo, R. *J. molec. Biol.* **146,** 77–99 (1981).
21. Reedy, M. K. *J. molec. Biol.* **31,** 155–176 (1968).

We thank J. Glenney, P. Glenney, W. Koch, V. Gehrke, B. Sintram, E. M. Mandelkow and P. Zingsheim for their help. P.M. is supported by a Muscular Dystrophy Association fellowship and L.K.H. by an EMBO fellowship.

22. Small, J. V. & Squire, J. M. *J. molec. Biol.* **67,** 117–149 (1972).
23. Hanson, J., Lednev, V., O'Brien, E. J. & Bennett, P. M. *Cold Spring Harb. Symp. quant. Biol.* **37,** 311–318 (1972).
24. Hanson, J. *Proc. R. Soc.* B**183,** 39–58 (1973).
25. Selby, C. C. & Bear, R. S. *J. biophys. Biochem. Cytol.* **3,** 71–85 (1956).
26. Cohen, C. & Hanson, J. *Biochim. biophys. Acta* **21,** 177–178 (1956).
27. Hanson, J. *Q. Rev. Biophys.* **1,** 177–216 (1968).
28. Lednev, V. V. *Biofizika* **19,** 116–121 (1974).
29. Vazina, A., Frank, G. & Lemazhikhin, B. *J. molec. Biol.* **14,** 373–380 (1965).
30. Spencer, M. *Nature* **223,** 1361–1362 (1969).
31. Bernal, J. D. & Fankuchen, I. *J. gen. Physiol.* **25,** 111–165 (1941).
32. Huxley, H. E. & Brown, W. *J. molec. Biol.* **30,** 383–434 (1967).
33. Marvin, D. A., Wiseman, R. L. & Wachtel, E. J. *J. molec. Biol.* **82,** 121–138 (1974).
34. Oster, G. & Riley, D. P. *Acta crystallogr.* **5,** 272–276 (1952).
35. Millman, B. M. & Nickel, B. G. *Biophys. J.* **32,** 49–63 (1980).
36. Franklin, R. & Klug, A. *Biochim. biophys. Acta* **19,** 403–416 (1956).
37. Vibert, P. J., Haselgrove, J. C., Lowy, J. & Poulsen, F. R. *J. molec. Biol.* **71,** 757–767 (1972).
38. Parry, D. A. D. & Squire, J. M. *J. molec. Biol.* **75,** 33–55 (1973).
39. Haselgrove, J. C. *Cold Spring Harb. Symp. quant. Biol.* **37,** 341–352 (1972).
40. Miller, A. & Treagear, R. *J. molec. Biol.* **70,** 85–104 (1972).
41. Onsager, L. *Ann. N.Y. Acad. Sci.* **51,** 627–659 (1949).
42. Spudich, J. A. & Watt, S. *J. biol. Chem.* **246,** 4866–4871 (1971).
43. Harmsen, A. thesis, Heidelberg Univ. (1980).
44. Gregory, J. & Holmes, K. C. *J. molec. Biol.* **13,** 796–801 (1965).
45. Suck, D., Kabsch, W. & Mannherz, H. G. *Proc. natn. Acad. Sci. U.S.A.* **78,** 4310–4323 (1981).

Cloning and expression of human tissue-type plasminogen activator cDNA in *E. coli*

Diane Pennica*, William E. Holmes*, William J. Kohr†, Richard N. Harkins†, Gordon A. Vehar†, Carole A. Ward†, William F. Bennett†, Elizabeth Yelverton*, Peter H. Seeburg*, Herbert L. Heyneker* & David V. Goeddel*

Departments of *Molecular Biology and †Protein Biochemistry, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

Desire Collen

Center for Thrombosis and Vascular Research, Department of Medical Research, University of Leuven, B-3000 Leuven, Belgium

Bacterial clones containing human tissue-type plasminogen activator (t-PA) cDNA sequences were identified in a cDNA library prepared using gel-fractionated mRNA from human melanoma cells. A plasmid containing the Escherichia coli trp *promoter and the cDNA sequence coding for the 527-amino acid mature t-PA protein was constructed for expression in* E. coli. *A polypeptide was produced having the fibrinolytic properties characteristic of authentic human t-PA.*

MAMMALIAN plasma contains an enzymatic system capable of dissolving the fibrin in blood clots. One component of this fibrinolytic system, plasminogen activators, generates the active enzyme plasmin by limited proteolysis of the zymogen plasminogen. Plasmin then degrades the fibrin network of a clot to form soluble products[1,2].

There are two drugs commercially available for thrombolytic therapy which function as plasminogen activators[3]: streptokinase, a bacterial protein, and urokinase, a serine protease isolated from human urine. Thrombolysis with these substances, however, is associated with systemic activation of plasminogen which can produce indiscriminate digestion of coagulation proteins, and significantly increase the risk of haemorrhage during treatment[4,5].

Plasminogen activators have been extracted from normal and tumour tissues and are produced by certain cells in culture[6]. The plasminogen activators derived from these sources have been classified into two major groups, urokinase-type plasminogen activators (u-PA) and tissue-type plasminogen activators (t-PA), based on differences in their immunological properties. (The abbreviations t-PA and u-PA are those proposed at the XXVIII Meeting of the International Committee on Thrombosis and Hemostasis, Bergamo, Italy, 27 July 1982.) Recently, a human melanoma cell line has been identified which secretes t-PA. Characterization of this melanoma plasminogen activator has shown it to be indistinguishable both immunologically and in amino acid composition from the plasminogen activator isolated from normal human tissue[7].

Several studies *in vitro*[8–11] and *in vivo*[12–14] which used t-PA purified from the melanoma cell line, have demonstrated its higher affinity for fibrin compared with urokinase-type plasminogen activators. This higher affinity may explain the fact that no systemic activation of plasminogen has been observed to accompany t-PA treatment[13–15]. Preliminary studies in two patients with renal and iliofemoral thrombosis have demonstrated the clinical potential of t-PA as a thrombolytic agent[15].

More extensive investigation of human t-PA as a potential thrombolytic agent has been hampered by its extremely low concentration in blood, tissue extracts, vessel perfusates and cell cultures. The production of t-PA using recombinant DNA techniques should provide sufficient quantities to examine its clinical usefulness in the treatment of pulmonary embolism, deep vein thrombosis, heart attacks and strokes. We report here the isolation and DNA sequence of two recombinant cDNA clones which between them contain the entire coding sequence of human t-PA. We also describe the construction of an expression plasmid which directs the synthesis in *E. coli* of a polypeptide having the fibrinolytic activity of authentic human t-PA.

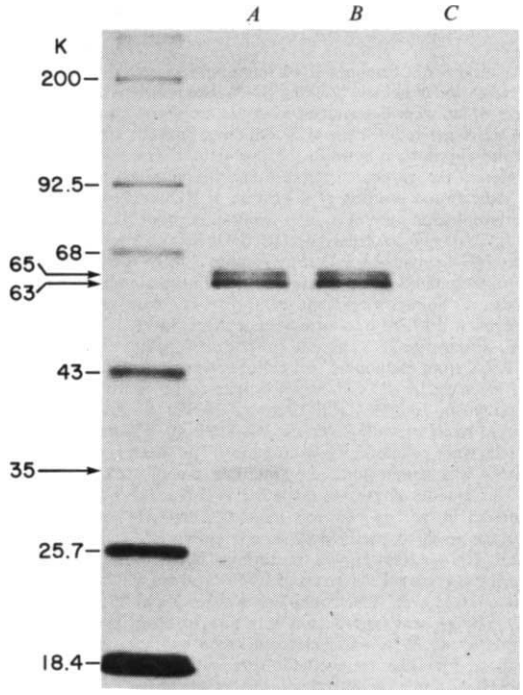


Fig. 1 Immunoprecipitation of melanoma t-PA produced *in vivo*. *A*, t-PA IgG immunoprecipitation; *B*, aprotinin and t-PA IgG immunoprecipitation; *C*, non-immune IgG immunoprecipitation. **Methods:** The Bowes melanoma cell line[6] was cultured to confluency in a 24-well microtitre dish in Earle's minimal essential medium (EMEM) supplemented with 0.12% sodium bicarbonate, 2 mM glutamine and 10% heat-inactivated fetal calf serum. The cells were washed once with phosphate-buffered saline, and 0.3 ml of serum-free methionine-free EMEM was added to each well. ^{35}S-methionine (75 μCi; Amersham) was added and the cells were labelled at 37 °C for 3 h in the presence or absence of 0.33 μM aprotinin (Trasylol; FBA Pharmaceuticals). After the 3-h labelling period, the medium was removed from the cells and incubated with either t-PA-specific rabbit IgG or non-immune rabbit IgG for immunoprecipitation[17]. The immunoprecipitated products were displayed by electrophoresis on a 10% SDS-polyacrylamide gel[55]. The slab gel was fixed, dried and subjected to fluorography[56]. Antisera against tissue-type plasminogen activator were raised in rabbits by multiple-site intradermal injections of 100 μg of purified protein dissolved in 0.5 ml of saline and emulsified with 0.5 ml of complete Freund's adjuvant. Booster injections of 50 μg were given at 2-week intervals and serum was collected after the second or third booster. The IgG fractions of the antisera were isolated by affinity chromatography on protein A–Sepharose and dialysed against 0.15 M NaCl solution.

Table 1 Plasminogen activation by extracts of *E. coli*

E. coli W3110 transformed by	Treatment	ΔA_{405}	Units per A_{550} cells
pt-PAtrp12	—	0.657	2.9
pt-PAtrp12	Non-immune IgG	0.665	2.9
pt-PAtrp12	t-PA IgG	0.059	0.26
pLeIFAtrp103	—	0.055	0.25

E. coli cultures were grown and extracts prepared as described in Fig. 5 legend. A 20-μl aliquot of each extract was incubated at 37 °C for 10 min with 0.15 ml of a 50 mM Tris-HCl *p*H 7.4, 12 mM NaCl solution containing 0.37 mg ml^{-1} plasminogen and 2.6 mg ml^{-1} fibrinogen. 0.35 ml of a 1 mM solution of S2251 (Kabi) in the above buffer was added and the reaction continued for 5–30 min at 37 °C. Acetic acid was added to a final concentration of 0.8 M to stop the reaction. The samples were centrifuged and absorbance at 405 nm measured. Where indicated, ~30 μg of either rabbit non-immune IgG or t-PA-specific IgG was added to the initial incubation. Units were defined by comparison to a urokinase standard (Calbiochem). The published specific activity of pure t-PA is ~60,300 PU mg^{-1} (refs 7, 46).

Immunological characterization of t-PA and mRNA isolation

To characterize the plasminogen activator species secreted from the Bowes human melanoma cell line[7], confluent monolayers of the cells were grown and the proteins labelled for 3 h with ^{35}S-methionine. The cell medium was immunoprecipitated with t-PA-specific rabbit IgG and electrophoresed on a SDS-polyacrylamide gel. Three different forms of melanoma t-PA were observed having molecular weights (MWs) of ~65,000 (65K), 63K and 35K (Fig. 1*A*). These three species correspond to those observed following electrophoresis and Coomassie blue staining of purified melanoma t-PA (data not shown). When the protease inhibitor, aprotinin, was added during the 3-h labelling period, only the 65K and 63K MW species were observed (Fig. 1*B*). Aprotinin has been shown previously to prevent the processing of single-chain t-PA to the two-chain form[7]. No immunoprecipitated polypeptides were observed when non-immune rabbit IgG was substituted for t-PA-specific IgG (Fig. 1*C*).

To determine whether the 65K and 63K MW species are related forms of t-PA, limited peptide mapping was performed using staphylococcal V8 protease[16,17]. Analysis of the peptides on SDS-polyacrylamide gels revealed that the two species are closely related in peptide composition, one probably being a derivative of the other (data not shown). The molecular weight difference might be due to a variation in carbohydrate composition or to the presence of an additional peptide in the 65K species. However, on sequence analysis both forms were found to have the same N-terminal amino acid sequence (W.J.K. and R.N.H., unpublished results).

The Bowes melanoma cell line was used as the source of RNA for cDNA cloning experiments. Polyadenylated RNA was prepared[18] and fractionated by electrophoresis through urea–agarose gels[19–21] to enrich for t-PA mRNA. Aliquots of the individual mRNA fractions were assayed for the presence of t-PA mRNA by *in vitro* translation in a rabbit reticulocyte lysate system[22] supplemented with dog pancreas microsomes[23]. The resulting translation products were immunoprecipitated with t-PA-specific IgG and analysed by SDS-polyacrylamide gel electrophoresis (Fig. 2*A*). One major polypeptide having ~63,000 MW was observed for both RNA fractions 7 and 8. This band was not observed when non-immune rabbit IgG was used for immunoprecipitation (Fig. 2*B*, lanes *a* and *d*) of either total RNA or gel slice 7 RNA translation products, which suggested that the polypeptide was t-PA. The size of the mRNA in fractions 7 and 8 is ~21–24S (2,400–2,700 nucleotides).

Construction and identification of bacterial clones containing t-PA cDNA sequences

RNA from gel slice 7 (Fig. 2*A*) was used to prepare double-stranded cDNA by standard procedures[24,25]. The cDNA was fractionated by size, and material longer than 350 base pairs (bp) was extended with deoxy(C) residues, annealed to deoxy(G)-tailed *Pst*I-cleaved pBR322, and used to transform *E. coli* K-12 strain 294 (ref. 26) as previously described[25]; ~4,600 transformants were obtained (see legend to Fig. 2*B*).

To prepare a specific DNA hybridization probe, we determined the amino acid sequences of several tryptic fragments of t-PA purified from melanoma cells (W.J.K. and R.N.H., unpublished results). This information permitted the design of synthetic deoxynucleotides potentially complementary to a specific region of t-PA mRNA. The amino acid sequence (Trp-Glu-Tyr-Cys-Asp) of a region of a purified tryptic peptide was selected for preparation of a probe because it required the synthesis[27] of only eight tetradecanucleotides ($dTC^{A}_{G}CA^{A}_{G}TA^{C}_{T}TCCCA$) to account for all possible anti-coding sequences. This pool of 14-mers was labelled with ^{32}P and used to screen the 4,600 cDNA clones by *in situ* colony hybridization[28]. Plasmid DNA was isolated[29] from 12 colonies that gave a positive hybridization signal. DNA sequencing was performed on the cDNA inserts from each of these clones by the dideoxy

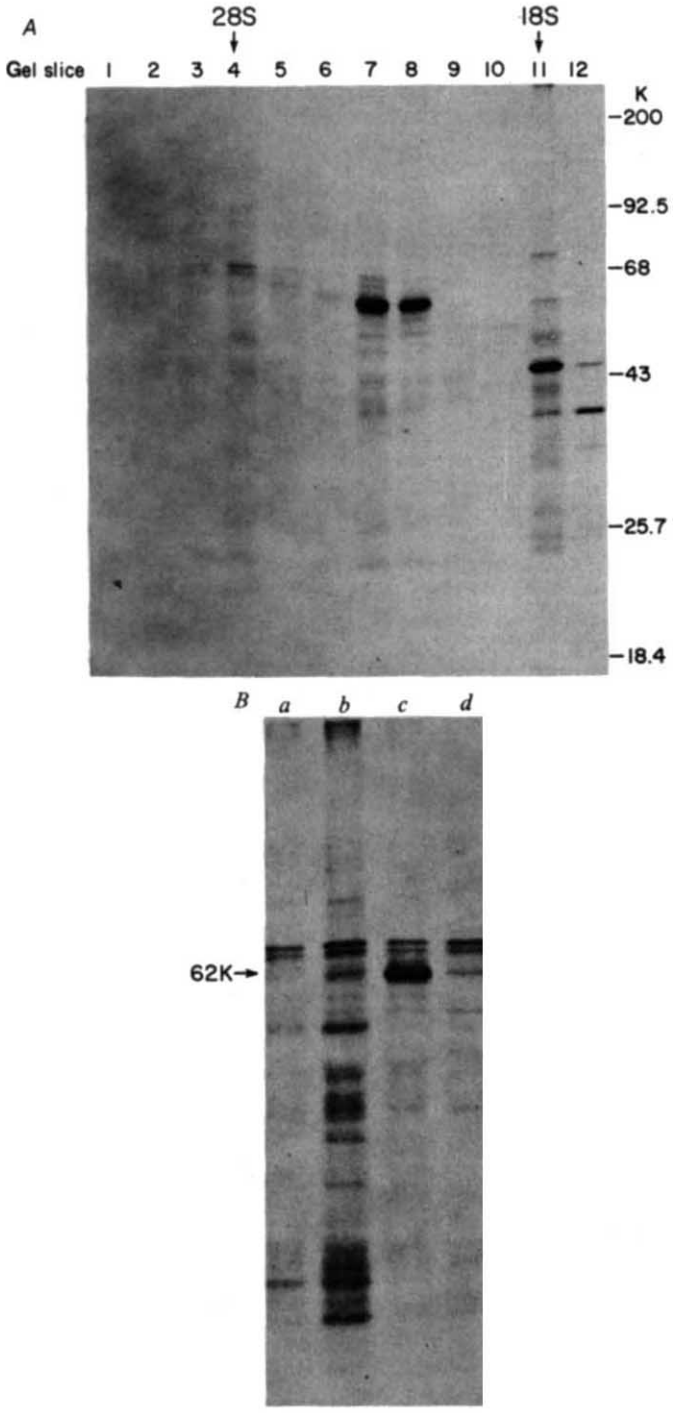


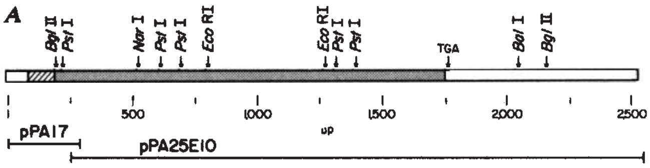
Fig. 2 *A*, gel electrophoresis of immunoprecipitated products from *in vitro* translation of gel-fractionated RNA. The positions of ribosomal RNA markers which were determined after electrophoresis in an adjacent lane on the RNA gel and visualized by ethidium bromide staining are labelled above the appropriate lanes of the protein gel. The numbers above each lane refer to the particular gel slice and the numbers to the right of the figure indicate the position of migration of ^{14}C-labelled protein markers. *B*, immunoprecipitation of *in vitro* translation products using non-immune rabbit IgG. Total RNA translation products immunoprecipitated with non-immune IgG (*a*) or t-PA IgG (*b*). Gel slice 7 mRNA translation products immunoprecipitated with t-PA IgG (*c*) or non-immune IgG (*d*).
Methods: *A*, human melanoma cells (Bowes) were grown to confluent monolayers in EMEM in plastic tissue culture flasks (175 cm^2; style 3028, Falcon Labware) at 37 °C in atmospheric air supplemented with 6% CO_2. Total RNA from melanoma cell cultures was extracted essentially as reported by Ward *et al.*[57]. Cells were pelleted by centrifugation and then resuspended in 10 mM NaCl, 10 mM Tris-HCl *p*H 7.5, 1.5 mM $MgCl_2$. Cells were lysed by adding Nonidet P-40 (NP-40; 1% final concentration), and nuclei were pelleted by centrifugation. The supernatant containing the total RNA was further purified by multiple phenol and chloroform extractions. The aqueous phase was made 0.2 M in NaCl and the total RNA was precipitated by adding two vols ethanol. Oligo(dT)-cellulose chromatography was used to purify poly(A)-containing RNA[17]. Fractionation of poly(A)$^+$ RNA (200 μg) was performed by electrophoresis through a denaturing agarose gel composed of 1.75% agarose, 0.025 M sodium citrate *p*H 3.8 and 6 M urea. Electrophoresis was for 7 h at 25 mA and 4 °C (refs 19–21). The gel was fractionated with a razor blade and individual slices were melted at 70 °C and extracted twice with phenol and once with chloroform. Fractions were ethanol-precipitated and then assayed by *in vitro* translation in a rabbit reticulocyte lysate system[22] supplemented with dog pancreas microsomes[23]. Incubations were at 30 °C for 90 min. The resulting translation products from each gel fraction were immunoprecipitated with rabbit IgG directed against human t-PA and analysed by electrophoresis on 10% SDS-polyacrylamide gels. The gels were fixed, dried and fluorographed[56]. *B*, 11 μg of total melanoma RNA or ~340 ng of gel slice 7 poly(A)$^+$ RNA were translated in a rabbit reticulocyte lysate system[22] supplemented with dog pancreas microsomes[23]; 11 μg of total RNA contains between 330 and 550 ng of mRNA based on the finding that ~3–5% of total RNA is mRNA[58]. The resulting translation products were immunoprecipitated with either rabbit IgG directed against human t-PA or non-immune rabbit IgG and analysed by electrophoresis on 10% SDS-polyacrylamide gels. Double-stranded cDNA was prepared by standard methods[24,25,40] using 5 μg of mRNA from gel fraction 7. The cDNA longer than 350 bp (125 ng) was recovered by electroelution[25] after fractionation on a 6% polyacrylamide gel, 30 ng of cDNA was extended with deoxy(C) residues using terminal deoxynucleotidyl transferase[59] and annealed with 300 ng of the plasmid pBR322[60] which had been similarly tailed with deoxy(G) residues at the *Pst*I site. The annealed mixture was used to transform *E. coli* K-12 strain 294 (ref. 25) and resultant tetracycline-resistant colonies were individually inoculated into wells of microtitre plates containing L broth[61] and 5 μg ml^{-1} tetracycline. The cDNA library of 4,600 transformants was grown up on nitrocellulose filters and the DNA from each colony was fixed to the filter[28]. The eight deoxyoligonucleotides (dTC^A_GCA^A_GTA^C_TTCCCA) were chemically synthesized in two pools of four 14-mers by the solid-phase phosphotriester method[27]. A ^{32}P-labelled probe was prepared from the pool of eight 14-mers as described previously[40]. The set of filters containing the 4,600 transformants was hybridized to 5×10^7 c.p.m. of the labelled probe in 50 mM sodium phosphate *p*H 6.8, 5×SSC, 150 μg ml^{-1} sonicated salmon sperm DNA, 5×Denhardt's solution and 10% formamide. After 16 h at room temperature, the filters were washed extensively at room temperature in 6×SSC and 0.1% SDS, then exposed to X-ray film.

chain termination procedure[30] after subcloning fragments into the M13 vector mp7 (ref. 31), or by the Maxam–Gilbert chemical procedure[32]. Only one cDNA insert, that of colony 25E10, contained sequences which could code for the amino acid sequence of the tryptic peptides of melanoma t-PA (W.J.K. and R.N.H., unpublished results). The cDNA insert of plasmid pPA25E10 was sequenced and found to be 2,304 bp long. The sequence (beginning with nucleotide 243 of Fig. 3*B*) contains an open reading frame encoding a protein of 508 amino acids and a 745-bp 3′ untranslated region.

Preparation of a colony library containing NH_2-terminal t-PA sequences

The cDNA clone pPA25E10 was not a full-length copy of t-PA mRNA as it lacked the t-PA NH_2-terminal coding sequences determined by amino acid sequencing of the purified protein. Therefore, it was necessary to produce cDNA clones containing the 5′ portion of the t-PA mRNA. A 16-base long deoxyoligonucleotide complementary to nucleotides 256–271 of the t-PA mRNA and having the sequence dTTCTGAGCACAGGGCG was synthesized[27]. This hexadecanucleotide was used to prime cDNA synthesis[33] of fraction 7 mRNA; 1,500 clones were obtained. As the cDNA priming reaction was performed using a synthetic fragment that hybridized only 13 bp from the 5′ terminus of the pPA25E10 cDNA, no convenient restriction fragment was available for use as a probe to screen for t-PA cDNA clones containing t-PA sequences 5′ of those in pPA25E10. We therefore isolated a genomic clone for t-PA from a human gene library[34] and used this as a hybridization probe to identify primer-extended cDNA clones containing NH_2-terminal t-PA coding sequences.

The first step in this process was to determine whether only a single homologous t-PA gene was present in human genomic DNA. Southern hybridizations[35] were performed using high-molecular weight human DNA which had been digested with various restriction endonucleases. The ^{32}P-labelled[36] cDNA



B

```
                                                                                    -35                 -30
                                                                                    met asp ala met lys arg gly leu
TTCTGAGCACAGGGCTGGAGAGAAAACCTCTGCGAGGAAAGGGAAGGAGCAAGCCGTGAATTTAAGGGACGCTGTGAAGCAATC        ATG GAT GCA ATG AAG AGA GGG CTC
1                                                50                                                                 100

                            -20                                     -10                                     1
cys cys val leu leu leu cys gly ala val phe val ser pro ser gln glu ile his ala arg phe arg arg gly ala arg SER TYR GLN
TGC TGT GTG CTG CTG CTG TGT GGA GCA GTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA GGA GCC AGA TCT TAC CAA
                                                    150

                        10                                      20                                      30
VAL ILE CYS ARG ASP GLU LYS THR GLN MET ILE TYR GLN GLN HIS GLN SER TRP LEU ARG PRO VAL LEU ARG SER ASN ARG VAL GLU TYR
GTG ATC TGC AGA GAT GAA AAA ACG CAG ATG ATA TAC CAG CAA CAT CAG TCA TGG CTG CGC CCT GTG CTC AGA AGC AAC CGG GTG GAA TAT
200                                                                 250

                        40                                      50                                      60
CYS TRP CYS ASN SER GLY ARG ALA GLN CYS HIS SER VAL PRO VAL LYS SER CYS SER GLU PRO ARG CYS PHE ASN GLY GLY THR CYS GLN
TGC TGG TGC AAC AGT GGC AGG GCA CAG TGC CAC TCA GTG CCT GTC AAA AGT TGC AGC GAG CCA AGG TGT TTC AAC GGG GGC ACC TGC CAG
            300                                                                 350

                        70                                      80                                      90
GLN ALA LEU TYR PHE SER ASP PHE VAL CYS GLN CYS PRO GLU GLY PHE ALA GLY LYS CYS CYS GLU ILE ASP THR ARG ALA THR CYS TYR
CAG GCC CTG TAC TTC TCA GAT TTC GTG TGC CAG TGC CCC GAA GGA TTT GCT GGG AAG TGC TGT GAA ATA GAT ACC AGG GCC ACG TGC TAC
                            400                                                             450

                        100                                     110                                     120
GLU ASP GLN GLY ILE SER TYR ARG GLY THR TRP SER THR ALA GLU SER GLY ALA GLU CYS THR ASN TRP ASN SER SER ALA LEU ALA GLN
GAG GAC CAG GGC ATC AGC TAC AGG GGC ACG TGG AGC ACA GCG GAG AGT GGC GCC GAG TGC ACC AAC TGG AAC AGC AGC GCG TTG GCC CAG
                                        500                                                                 550

                        130                                     140                                     150
LYS PRO TYR SER GLY ARG ARG PRO ASP ALA ILE ARG LEU GLY LEU GLY ASN HIS ASN TYR CYS ARG ASN PRO ASP ARG ASP SER LYS PRO
AAG CCC TAC AGC GGG CGG AGG CCA GAC GCC ATC AGG CTG GGC CTG GGG AAC CAC AAC TAC TGC AGA AAC CCA GAT CGA GAC TCA AAG CCC
                                                    600

                        160                                     170                                     180
TRP CYS TYR VAL PHE LYS ALA GLY LYS TYR SER SER GLU PHE CYS SER THR PRO ALA CYS SER GLU GLY ASN SER ASP CYS TYR PHE GLY
TGG TGC TAC GTC TTT AAG GCG GGG AAG TAC AGC TCA GAG TTC TGC AGC ACC CCT GCC TGC TCT GAG GGA AAC AGT GAC TGC TAC TTT GGG
650                                                                 700

                        190                                     200                                     210
ASN GLY SER ALA TYR ARG GLY THR HIS SER LEU THR GLU SER GLY ALA SER CYS LEU PRO TRP ASN SER MET ILE LEU ILE GLY LYS VAL
AAT GGG TCA GCC TAC CGT GGC ACG CAC AGC CTC ACC GAG TCG GGT GCC TCC TGC CTC CCG TGG AAT TCC ATG ATC CTG ATA GGC AAG GTT
            750                                                                 800

                        220                                     230                                     240
TYR THR ALA GLN ASN PRO SER ALA GLN ALA LEU GLY LEU GLY LYS HIS ASN TYR CYS ARG ASN PRO ASP GLY ASP ALA LYS PRO TRP CYS
TAC ACA GCA CAG AAC CCC AGT GCC CAG GCA CTG GGC CTG GGC AAA CAT AAT TAC TGC CGG AAT CCT GAT GGG GAT GCC AAG CCC TGG TGC
                            850                                                             900

                        250                                     260                                     270
HIS VAL LEU LYS ASN ARG ARG LEU THR TRP GLU TYR CYS ASP VAL PRO SER CYS SER THR CYS GLY LEU ARG GLN TYR SER GLN PRO GLN
CAC GTG CTG AAG AAC CGC AGG CTG ACG TGG GAG TAC TGT GAT GTG CCC TCC TGC TCC ACC TGC GGC CTG AGA CAG TAC AGC CAG CCT CAG
                                        950                                                                 1000

                        280                                     290                                     300
PHE ARG ILE LYS GLY GLY LEU PHE ALA ASP ILE ALA SER HIS PRO TRP GLN ALA ALA ILE PHE ALA LYS HIS ARG ARG SER PRO GLY GLU
TTT CGC ATC AAA GGA GGG CTC TTC GCC GAC ATC GCC TCC CAC CCC TGG CAG GCT GCC ATC TTT GCC AAG CAC AGG AGG TCG CCC GGA GAG
                                                    1050

                        310                                     320                                     330
ARG PHE LEU CYS GLY GLY ILE LEU ILE SER SER CYS TRP ILE LEU SER ALA ALA HIS CYS PHE GLN GLU ARG PHE PRO PRO HIS HIS LEU
CGG TTC CTG TGC GGG GGC ATA CTC ATC AGC TCC TGC TGG ATT CTC TCT GCC GCC CAC TGC TTC CAG GAG AGG TTT CCG CCC CAC CAC CTG
1100                                                                1150

                        340                                     350                                     360
THR VAL ILE LEU GLY ARG THR TYR ARG VAL VAL PRO GLY GLU GLU GLU GLN LYS PHE GLU VAL GLU LYS TYR ILE VAL HIS LYS GLU PHE
ACG GTG ATC TTG GGC AGA ACA TAC CGG GTG GTC CCT GGC GAG GAG GAG CAG AAA TTT GAA GTC GAA AAA TAC ATT GTC CAT AAG GAA TTC
            1200                                                                1250

                        370                                     380                                     390
ASP ASP ASP THR TYR ASP ASN ASP ILE ALA LEU LEU GLN LEU LYS SER ASP SER SER ARG CYS ALA GLN GLU SER SER VAL VAL ARG THR
GAT GAT GAC ACT TAC GAC AAT GAC ATT GCG CTG CTG CAG CTG AAA TCG GAT TCG TCC CGC TGT GCC CAG GAG AGC AGC GTG GTC CGC ACT
                            1300                                                            1350

                        400                                     410                                     420
VAL CYS LEU PRO PRO ALA ASP LEU GLN LEU PRO ASP TRP THR GLU CYS GLU LEU SER GLY TYR GLY LYS HIS GLU ALA LEU SER PRO PHE
GTG TGC CTT CCC CCG GCG GAC CTG CAG CTG CCG GAC TGG ACG GAG TGT GAG CTC TCC GGC TAC GGC AAG CAT GAG GCC TTG TCT CCT TTC
                                        1400                                                                1450

                        430                                     440                                     450
TYR SER GLU ARG LEU LYS GLU ALA HIS VAL ARG LEU TYR PRO SER SER ARG CYS THR SER GLN HIS LEU LEU ASN ARG THR VAL THR ASP
TAT TCG GAG CGG CTG AAG GAG GCT CAT GTC AGA CTG TAC CCA TCC AGC CGC TGC ACA TCA CAA CAT TTA CTT AAC AGA ACA GTC ACC GAC
                                                    1500

                        460                                     470                                     480
ASN MET LEU CYS ALA GLY ASP THR ARG SER GLY GLY PRO GLN ALA ASN LEU HIS ASP ALA CYS GLN GLY ASP SER GLY GLY PRO LEU VAL
AAC ATG CTG TGT GCT GGA GAC ACT CGG AGC GGC GGG CCC CAG GCA AAC TTG CAC GAC GCC TGC CAG GGC GAT TCG GGA GGC CCC CTG GTG
1550                                                                1600

                        490                                     500                                     510
CYS LEU ASN ASP GLY ARG MET THR LEU VAL GLY ILE ILE SER TRP GLY LEU GLY CYS GLY GLN LYS ASP VAL PRO GLY VAL TYR THR LYS
TGT CTG AAC GAT GGC CGC ATG ACT TTG GTG GGC ATC ATC AGC TGG GGC CTG GGC TGT GGA CAG AAG GAT GTC CCG GGT GTG TAC ACA AAG
            1650                                                                1700

                        520                         527
VAL THR ASN TYR LEU ASP TRP ILE ARG ASP ASN MET ARG PRO OP
GTT ACC AAC TAC CTA GAC TGG ATT CGT GAC AAC ATG CGA CCG TGA CCAGGAACACCCGACTCCTCAAAAGCAAATGAGATCCCGCCTCTTCTTCTTCAGAAGACA
                            1750                                                    1800

CTGCAAAGGCGCAGTGCTTCTCTACAGACTTCTCCAGACCCACCACACCGCAGAAGCGGGACGAGACCCTACAGGAGAGGGAAGAGTGCATTTTCCCAGATACTTCCCATTTTGGAAGTT
              1850                                              1900                                              1950

TTCAGGACTTGGTCTGATTTCAGGATACTCTGTCAGATGGGAAGACATGAATGCACACTAGCCTCTCCAGGAATGCCTCCTCCCTGGGCAGAAGTGGCCATGCCACCCTGTTTTCGCTAA
                                              2000                                              2050

AGCCCAACCTCCTGACCTGTCACCGTGAGCAGCTTTGGAAACAGGACCACAAAAATGAAAGCATGTCTCAATAGTAAAAGAAACAAGAGATCTTTCAGGAAAGACGGATTGCATTAGAAA
                        2100                                              2150

TAGACAGTATATTTATAGTCACAAGGGCCCAGCAGGGCTCAAAGTTGGGGCAGGCTGGCTGGCCCGTCATGTTCCTCAAAAGCGCCCTTGACGTCAAGTCTCCTTCCCCTTTCCCCACTC
    2200                                              2250                                              2300

CCTGGCTCTCAGAAGGTATTCCTTTTGAGTACAGTGTGTAAAGTGTAAATCCTTTTTCTTTATAAACTTTAGAGTAGCATGAGAGAATTGTATCATTTGAACAACTAGGCTTCAGCATAT
                                  2350                                              2400

TTATAGCGATCCATCGTTAGTTTTTACTTTCCGTTGCCACAACCCTGTTTTATACCGTACTTAATAAATTCGGATATATTTTTTCACAGTTTTTTCCAAAAAAAAAAAAAA
              2450                                              2500
```

Fig. 3 *A*, restriction endonuclease map of the cDNA insert of pt-PAtrp12. The number and size of fragments produced by restriction endonuclease cleavage were estimated by electrophoresis through 6% polyacrylamide gels. Positions of sites were confirmed by nucleic acid sequencing. The shaded region indicates the coding sequences of the putative mature t-PA protein (527 amino acids), the cross-hatched region represents the putative 35-residue 'prepro' peptide coding sequence, and the open regions show the 3' and 5' noncoding sequences. The 5' end of the mRNA is to the left while the 3' end is to the right. The lines below the restriction map represent the length of the two partial clones pPA17 and pPA25E10. *B*, nucleotide sequence and deduced amino acid sequence of the full-length human tissue-type plasminogen activator cDNA insert. The putative presequence is represented by the residues labelled −35 to −1. The entire sequence was determined by a combination of the dideoxynucleotide chain termination method[30] after subcloning fragments into the M13 vector mp7 (ref. 31) and the chemical method of Maxam and Gilbert[32]. Numbers above each line refer to amino acid position and numbers beneath each line to nucleotide position.

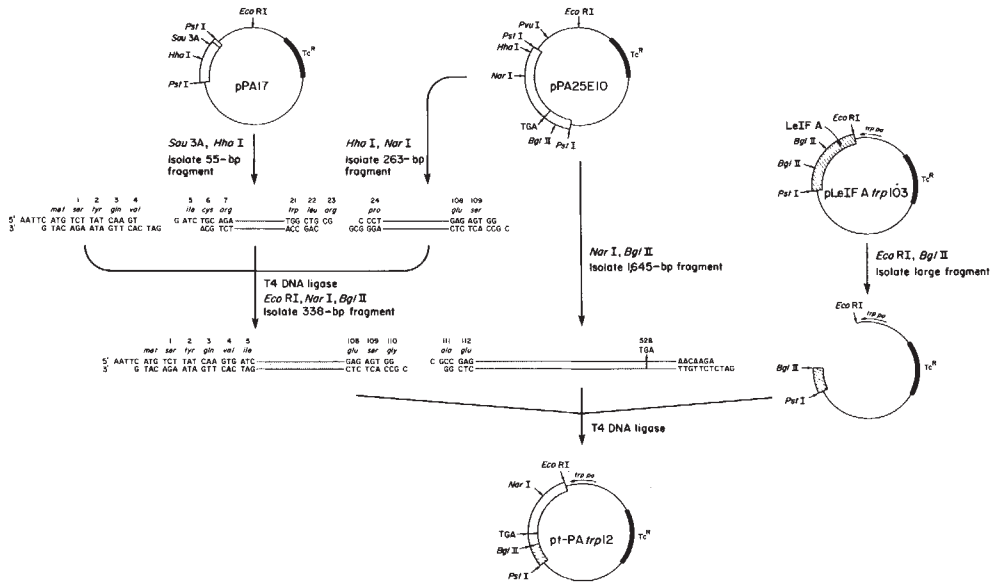


Fig. 4 Construction of a plasmid coding for the direct expression of mature t-PA in *E. coli*. 50 μg of plasmid pPA17 was digested with *Sau*3AI, *Hin*cII and *Hha*I and electrophoresed on a 6% polyacrylamide gel; ~0.5 μg of the 55-bp *Sau*3AI–*Hha*I fragment was recovered. Similarly, ~3 μg of the 263-bp *Hha*I–*Nar*I fragment was purified from 80 μg of clone pPA25E10 by first isolating a 300-bp *Pst*I–*Nar*I fragment and then digesting this fragment with *Hha*I. All digests were performed at 37 °C for 1 h and the reaction products resolved and electroeluted from 6% polyacrylamide gels. The two indicated deoxyoligonucleotides 5′dAATTCATGTCTTATCAAGT (I) and 5′ GATCACTTGATAAGACATG (II) were synthesized by the solid-phase phosphotriester method[27]. 100 pmol of oligonucleotide II was phosphorylated in 30-μl reaction mixture containing 60 mM Tris *p*H 8, 10 mM $MgCl_2$, 15 mM β-mercaptoethanol and 50 μCi [γ-^{32}P]ATP (5,000 Ci mmol^{-1}; Amersham). T4 polynucleotide kinase (12 units) was added and the reaction allowed to proceed at 37 °C for 15 min; 1 μl of 10 mM ATP and 12 units of T4 kinase were then added and the reaction allowed to proceed for an additional 30 min. After phenol and chloroform extraction, the phosphorylated oligomer II and the 5′-hydroxyl oligomer I were combined with 0.5 μg of the eluted 55-bp *Sau*3AI–*Hha*I fragment and 2 μg of the 263-bp *Hha*I–*Nar*I fragment and ethanol-precipitated. These fragments were ligated at room temperature for 4 h in 60 μl of 20 mM Tris-HCl *p*H 7.5, 10 mM $MgCl_2$, 10 mM dithiothreitol, 0.5 mM ATP and 1,000 units of T4 DNA ligase. The mixture was digested for 1 h with 48 units of *Nar*I, 20 units of *Eco*RI and 40 units of *Bgl*II (to eliminate polymerization through ligation of cohesive *Sau*3AI termini) and electrophoresed on a 6% gel. The 338-bp product (~0.1 μg; 1,385 c.p.m.) was recovered by electroelution. The remainder of the t-PA-coding sequences (amino acids 111–528) were isolated on a 1,645-bp fragment by digesting plasmid pPA25E10 with *Nar*I and *Bgl*II. The plasmid pLeIFAtrp103 is a derivative of the plasmid pLeIFA25 (ref. 25) in which the *Eco*RI site distal to the LeIF A gene has been removed[19]. pLeIFAtrp103 (3 μg) was digested with 20 units of *Eco*RI and 20 units of *Bgl*II for 90 min at 37 °C, electrophoresed on a 6% polyacrylamide gel and the large (~4,200 bp) vector fragment was recovered by electroelution. For the final construction, 80 ng of *Eco*RI–*Bgl*II pLeIFAtrp103 was ligated with 100 ng of the 1,645-bp *Nar*I–*Bgl*II fragment and 20 ng of the 338-bp *Eco*RI–*Nar*I fragment at room temperature for 10 h. This ligation mixture was used to transform *E. coli* K-12 strain 294 (ref. 26). Plasmid DNA was prepared from 38 of the transformants and digested with *Eco*RI; 10 of these plasmids contained the desired 600- and 472-bp *Eco*RI fragments. DNA sequence analysis verified that one of these plasmids (pt-PAtrp12) had the desired nucleotide sequence at the junctions between the *trp* promoter, synthetic DNA and cDNA.

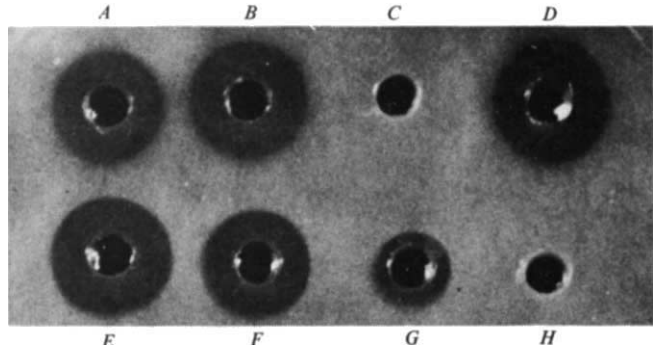


Fig. 5 Characterization of t-PA produced by *E. coli*. Wells *A–D* contain 10 μl of a pt-PAtrp12 cell extract plus the following additions: *A*, none; *B*, non-immune IgG; *C*, t-PA IgG; *D*, urokinase IgG. Wells *E–G* contain, respectively, 0.2, 0.1 and 0.02 unit purified melanoma t-PA. Well *H* contains 10 μl of an extract from cells containing the plasmid pLeIFAtrp103.
Methods: An overnight culture of *E. coli* W3110/pt-PAtrp12 in Luria broth[57] containing 5 μg ml^{-1} tetracycline was diluted 1:100 in M9 medium[57] containing 0.2% glucose, 0.5% casamino acids and 5 μg ml^{-1} tetracycline. The cells were grown at 37 °C to A_{550} of 0.2 and indole acrylic acid was added to a final concentration of 20 μg ml^{-1}. Samples were collected, by centrifugation, at $A_{550} = 0.5–0.6$ (~2×10^8 cells ml^{-1}) and immediately frozen. The cell pellets were suspended in 6 M guanidine hydrochloride at 5×10^8 cells ml^{-1}, sonicated for 10 s, incubated at 24 °C for 30 min and then dialysed (buffer flow rate 250 ml h^{-1}) for 4 h against 50 mM Tris-HCl *p*H 8.0, 250 mM NaCl, 0.25 mM EDTA and 0.01% Tween 80, using a continuous flow microdialysis system (Bethesda Research Labs). After dialysis the samples were centrifuged at 13,000*g* for 2 min and 10-μl samples of the supernatants (1 ml) analysed for t-PA activity. Following the procedure of Granelli-Piperno and Reich[45], the plate was incubated for 3.5 h at 37 °C and lysis zones measured. Activity was approximated by comparison with dilutions of a standard purified melanoma t-PA solution.

probe used was a 232-bp *Rsa*I–*Pst*I fragment (nucleotides 388–620 of Fig. 3*B*) prepared from the 5′ end of the cDNA insert of clone pPA25E10. Two endonuclease digestion patterns provided only a single hybridizing DNA fragment: *Bgl*II (5.7 kilobase pairs, kbp) and *Pvu*II (4.2 kbp). Two hybridizing DNA fragments were observed with *Hin*cII (5.1 and 4.3 kbp) (data not shown). Comparison of these data with the cDNA restriction map (Fig. 3*A*) suggests that there is only one t-PA gene in the human genome and that this gene contains at least one intervening sequence.

Approximately 10^6 plaques of a λ Charon 4A/human genomic library[34] were screened[37] with the ^{32}P-labelled 232-bp *Rsa*I–*Pst*I fragment; 19 individual clones were isolated and phage DNA was prepared[38]. A 4.2-kbp *Pvu*II fragment containing t-PA sequences was isolated from one of these clones, labelled with ^{32}P and used to screen the 1,500 clones of the 5′ primer-extended cDNA library. Plasmid DNAs were prepared from the 18 colonies which gave positive hybridization signals and these DNAs were bound to a nitrocellulose filter[39,40].

To identify clones having cDNA inserts which overlapped with the cDNA insert of pPA25E10, the filter was hybridized with the ^{32}P-labelled synthetic oligonucleotide (16-mer) used for the original priming reaction. Of the 18 selected cDNA clones, 7 hybridized with the ^{32}P-labelled 16-mer; however, on

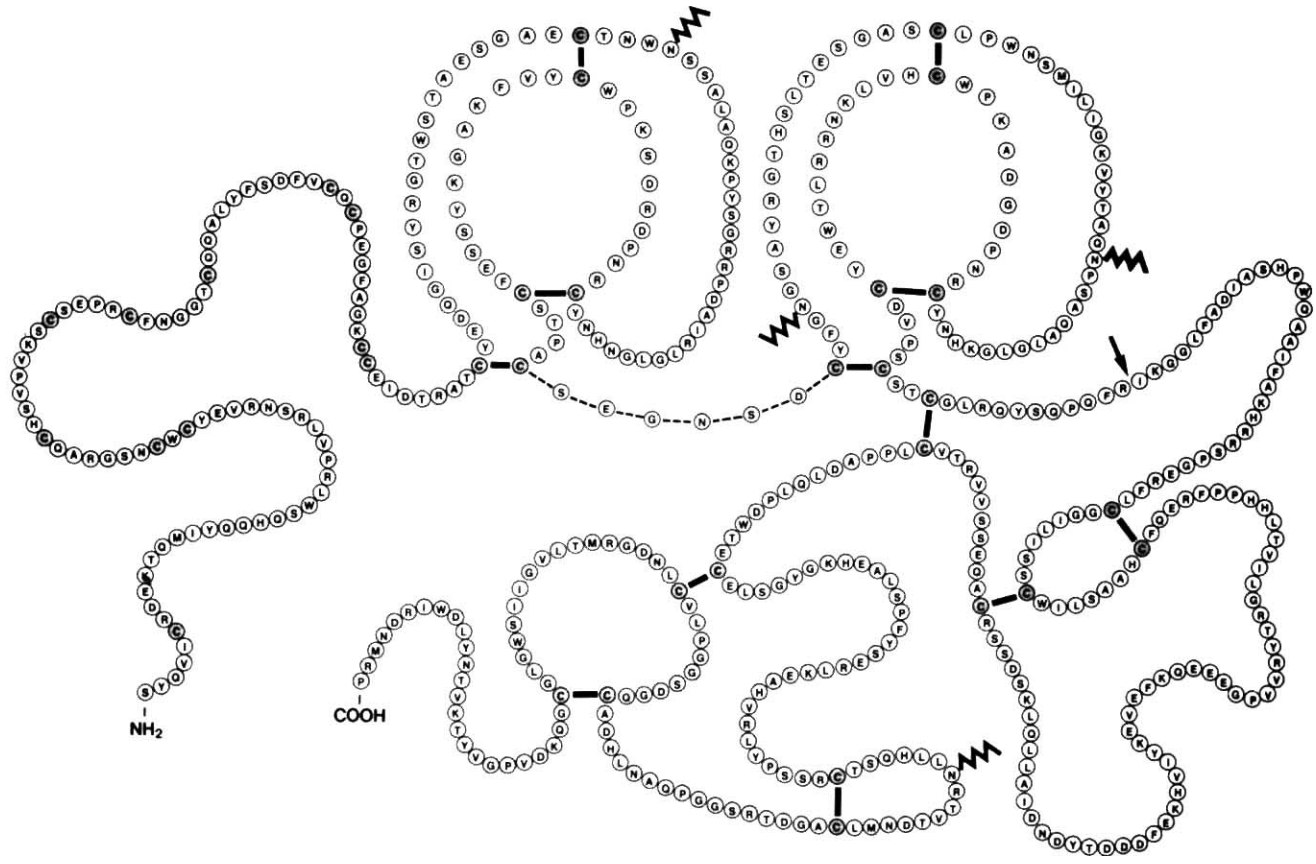


Fig. 6 Schematic diagram of the potential structure of human tissue-type plasminogen activator. The one-letter code for each of the amino acids is given in the open circles. The one-letter abbreviations are those recommended by the IUPAC-IUB Commission on Biochemical Nomenclature[54]. The cysteine residues are shaded. The solid black bars indicate the potential disulphide bridges based on homology with other serine proteases. The arrow indicates the potential cleavage site between arginine and the isoleucine which generates the two-chain molecule from the one-chain form. The zig-zag lines indicate potential glycosylation sites. The broken lines connect the six amino acids between the two kringles.

sequence analysis only one of the cDNA clones, pPA17, was found to contain the complete NH_2-terminal region of t-PA and to overlap with the original clone, 25E10.

The cDNA insert of pPA17 is 271 bp long. It contains the hexadecanucleotide sequence used to prime its synthesis, which permitted the alignment of its DNA sequence with that of pPA25E10. From these two cDNA clones, pPA15E10 and pPA17, we determined the nucleotide sequence and corresponding amino acid sequence of t-PA (Fig. 3*B*).

The complete 2,530-bp cDNA sequence contains a single open reading frame, beginning with the ATG codon at nucleotides 85–87. This ATG is followed, 562 codons later, by a TGA termination triplet at nucleotides 1,771–1,773. This ATG probably serves as the site of translation initiation as it is the first one encountered and it is preceded in phase at nucleotides 4–6 by a termination codon. The serine designated as amino acid number 1 is based on NH_2-terminal sequencing of purified melanoma cell t-PA. This serine is preceded by 35 amino acids, the NH_2-terminal 20–23 of which probably constitute a hydrophobic signal peptide[41] involved in the secretion of t-PA. The remaining 12–15 hydrophilic amino acids immediately preceding the start of mature t-PA may constitute a 'pro' sequence similar to that found for serum albumin[33,42]. The 3′ untranslated region of 759 nucleotides contains the hexanucleotide AATAAA (position 2,496–2,501) which precedes the site of polyadenylation in many eukaryotic mRNAs[43]. The native t-PA molecule has 35 cysteine residues and thus has the potential to be stabilized by 17 disulphide bridges. There are four potential N-glycosylation sites[44], located at Asn 117, Asn 184, Asn 218 and Asn 448 in t-PA. We have confirmed only the carbohydrate on Asn 448 in native t-PA (W.J.K. and R.N.H., unpublished results). Variations in the structure of the oligosaccharide ligands may account for the different molecular forms of t-PA (65K and 63K species) observed.

Synthesis of tissue-type plasminogen activator in *E. coli*

The procedure followed to directly express the full-length cDNA insert of t-PA in *E. coli* is outlined in Fig. 4. The common *Hha*I restriction endonuclease site shared by the cDNA inserts of both partial clones pPA17 and pPA25E10 permitted the reconstruction of the entire mature t-PA coding sequence. A 55-bp *Sau*3AI–*Hha*I restriction fragment corresponding to amino acids 5–23 was isolated from the plasmid pPA17. The *Sau*3AI restriction site is located at codon 4 of the presumed mature coding sequence and was used to remove the 'prepro' peptide-coding region. A 263-bp *Hha*I–*Nar*I fragment (encoding amino acids 14–110) was isolated from plasmid pPA25E10. Two synthetic deoxyoligonucleotides were designed to restore the codons for amino acids 1–4, incorporate an ATG translational initiation codon and create an *Eco*RI cohesive terminus. These fragments were then ligated together to form a 338-bp fragment coding for amino acids 1–110 of t-PA. This fragment and a 1,645-bp *Nar*I–*Bgl*II fragment from pPA15E10 were then ligated between the *Eco*RI and *Bgl*II sites of the plasmid pLeIFAtrp103 (ref. 21) to give the expression plasmid, pt-PAtrp12. The cloned t-PA gene is transcribed under the control of a 300-bp fragment of the *E. coli trp* operon which contains the *trp* promoter, operator and the Shine–Dalgarno sequence of the *trp* leader peptide, but lacks the leader peptide ATG initiation codon[40].

E. coli K-12 strain W3110 (ATCC no. 27325) containing the plasmid pt-PAtrp12, was grown, and extracts prepared for assay of fibrinolytic activity using the fibrin plate technique[45].

This assay involves the use of an agarose plate containing plasminogen and fibrin; when a sample containing a plasminogen activator is added to a well in the agarose, the plasminogen is converted to the active enzyme plasmin. The plasmin so formed then degrades the fibrin to give a clear zone of lysis surrounding the well. The area of this lysis zone can be correlated to the amount of plasminogen activator in the sample. Figure 5 shows that when extracts from pt-PAtrp12 clones were tested for plasminogen activator activity using this assay, a clear zone of lysis was evident (well *A*). This fibrinolytic activity was inhibited by t-PA-specific IgG (well *C*) but not by non-immune IgG (well *B*) or urokinase-specific IgG (well *D*) and no activity was seen in an extract prepared from cells containing as a control the α-interferon expression plasmid pLeIFAtrp103 (well *H*). Using the standard curve of purified melanoma cell t-PA (wells *E–G*) we estimated that ~5 units of t-PA activity could be extracted per A_{550} of culture.

t-PA activity can also be measured using a chromogenic substrate assay. This assay measures the rate of conversion of plasminogen to plasmin by t-PA. The amount of plasmin formed is determined by monitoring the proteolytic cleavage of a tripeptide from the chromogenic substrate, S2251. Table 1 shows the results obtained with this assay for *E. coli* W3110/pt-PAtrp12 extracts. The activity (~3 units per A_{550} of culture) measured by this assay is also neutralized by t-PA-specific IgG but not by non-immune IgG. The published specific activity of melanoma t-PA is 60,300 Plough units (PU) mg^{-1} (refs 7, 46).

Conclusion

Human melanoma cells actively synthesizing and secreting t-PA were used for the preparation of poly(A) mRNA. This mRNA was size-fractionated using agarose gel electrophoresis and the fractions containing t-PA mRNA were used to prepare a cDNA library, which was hybridized with a pool of radiolabelled synthetic deoxynucleotides potentially complementary to mRNA encoding a known amino acid sequence of t-PA. A recombinant plasmid (pPA25E10) containing a partial-length t-PA cDNA insert was identified and sequenced. The remaining 5′ cDNA sequences were generated by preparing a cDNA library primed with a 16-base long synthetic fragment. A t-PA cDNA clone (pPA17) overlapping with the original clone pPA25E10 and containing the NH_2-terminal and presequence regions was identified after specific hybridization with a DNA fragment prepared from a t-PA genomic clone.

The 2,530 bp of t-PA cDNA (excluding the poly(A) sequence) code for a polypeptide of 562 amino acids. The 35 amino acids (−35 to −1) preceding the mature sequence probably constitute a 20–23-amino acid long hydrophobic signal peptide followed by a hydrophilic 'pro' sequence of 12–15 amino acids. This type of prepro structure on secreted proteins has been described previously, for example, with preproalbumin[33,42]. On this assumption, all the secreted one-chain t-PA molecules will have a serine residue at the amino terminus. An alternative possibility is that the hydrophilic 'pro' sequence could be removed in a manner analogous to that observed with plasminogen where a peptide of MW 10,000 can be cleaved from the amino-terminal portion of native plasminogen (Glu-plasminogen, named for the amino terminal residue), resulting in a protein having a new amino terminus, designated Lys-plasminogen[47,48].

It has been demonstrated that two-chain t-PA is formed by proteolytic cleavage of the single-chain molecule into two polypeptides connected by disulphide bonds[7]. NH_2-terminal sequencing of the two-chain molecule suggests that it is generated by cleavage of the Arg 275–Ile 276 peptide bond (W.J.K. and R.N.H., unpublished results) (see Fig. 6). As determined from the primary structure, the heavy chain (MW 30,882) is derived from the NH_2-terminal part and the light chain (MW 28,126) comprises the COOH-terminal region (Fig. 6).

The primary structure of a portion of the heavy-chain region of t-PA has a high degree of sequence homology with the 'kringle' regions of plasminogen[49] and prothrombin[50,51]. 'Kringle' refers to a characteristic triple disulphide structure originally discovered in the 'pro' fragment of prothrombin and described in detail by Magnusson *et al.*[51,52]. From the primary sequence of t-PA, two such regions of 82 amino acids each (residues 92–173 and 180–261; Fig. 6) are apparent that share a high degree of homology with the five kringles of plasminogen. We have also recently determined that human urinary urokinase contains one similar triple disulphide region[53]. The remaining NH_2-terminal 91 amino acids of t-PA share little homology to the conventional kringle region. One can speculate, however, that this region may also assume a structure containing multiple disulphide bonds as 11 additional cysteine residues are found here.

The catalytic site of the light chain of t-PA, as in other serine proteases, is probably formed by the His 322, Asp 371 and Ser 478 residues. Furthermore, the amino acid sequences surrounding these residues are highly homologous to corresponding parts of other serine proteases such as trypsin, prothrombin and plasminogen[54].

The amounts of t-PA activity (3–5 units per A_{550}, depending on the assay used) recovered from the *E. coli* extracts correspond to ~50–80 μg per l of culture at $A_{550} = 1$, or ~1,500–2,400 fully active t-PA molecules per cell. This activity has been shown to activate plasminogen catalytically and is neutralized by t-PA-specific IgG antibody. As only functional activity was measured, it is possible that considerably more t-PA is present in these cells. In addition, because polypeptides synthesized in *E. coli* are not glycosylated, the t-PA synthesized by *E. coli* might have a specific activity different from that of authentic t-PA.

We thank Laurie May for growing melanoma cells, Joel Hayflick for help with the DNA sequencing, John Adelman for assistance in screening the human genomic library and preparation of the G-tailed pBR322 vector, Jane Baldridge for plasmid DNA preparations and the Genentech Organic Chemistry Department for deoxyoligonucleotide synthesis. A portion of this work was presented by D.P. at the sixth International Congress on Fibrinolysis, Lausanne, Switzerland, 23 July 1982. This is Genentech contribution no. 142.

Received 12 October; accepted 23 November 1982.

1. Collen, D. *Thromb. Haemostasis* **43,** 77–89 (1980).
2. Wiman, B. & Collen, D. *Nature* **272,** 549–550 (1978).
3. Verstraete, M. in *Fibrinolysis* (eds Kline, D. L. & Reddy, K. N. N.) 185–200 (CRC, Boca Raton, Florida, 1980).
4. Duckert, F. in *Fibrinolytics and Antifibrinolytics* (ed. Markwardt, F.) 209–237 (Springer, New York, 1978).
5. Vogel, G. in *Fibrinolytics and Antifibrinolytics* (ed. Markwardt, F.) 179–207 (Springer, New York, 1978).
6. Reich, E. in *Proteases and Biological Control* (eds Reich, E., Rifkin, D. B. & Shaw, E.) 333–341 (Cold Spring Harbor Laboratory, New York, 1975).
7. Rijken, D. C. & Collen, D. *J. biol. Chem.* **256,** 7035–7041 (1981).
8. Matsuo, O., Rijken, D. C. & Collen, D. *Thromb. Haemostasis* **45,** 225–229 (1981).
9. Korninger, C. & Collen, D. *Thromb. Haemostasis* **46,** 662–665 (1981).
10. Korninger, C. & Collen, D. *Thromb. Haemostasis* **46,** 561–565 (1981).
11. Hoylaerts, M., Rijkin, D. C., Lijnen, H. R. & Collen, D. *J. biol. Chem.* **257,** 2912–2919 (1982).
12. Korninger, C., Stassen, J. M. & Collen, D. *Thromb. Haemostasis* **46,** 658–661 (1981).
13. Matsuo, O., Rijken, D. C. & Collen, D. *Nature* **291,** 590–591 (1981).
14. Korninger, C., Matsuo, O., Suy, R., Stassen, J. M. & Collen, D. *J. clin. Invest.* **69,** 573–580 (1982).
15. Weimar, W. *et al. Lancet* **ii,** 1018–1020 (1981).
16. Cleveland, D., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. *J. biol. Chem.* **252,** 1102–1106 (1978).
17. Oppermann, H., Levinson, A. D. & Varmus, H. E. *Virology* **108,** 47–70 (1981).
18. Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69,** 1408–1412 (1972).
19. Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. *Biochemistry* **16,** 4743–4751 (1977).
20. Lynch, K. R., Pennica, D., Ennis, H. L. & Cohen, P. S. *Virology* **98,** 251–254 (1979).
21. Gray, P. W. *et al. Nature* **295,** 503–508 (1982).
22. Pelham, H. R. & Jackson, R. J. *Eur. J. Biochem.* **67,** 247–256 (1976).
23. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67,** 852–862 (1975).
24. Wickens, M. P., Buell, G. N. & Schimke, R. T. *J. biol. Chem.* **253,** 2483–2495 (1978).
25. Goeddel, D. V. *et al. Nature* **281,** 544–548 (1979).
26. Backman, K., Ptashne, M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **73,** 4174–4178 (1976).
27. Crea, R. & Horn, T. *Nucleic Acids Res.* **8,** 2331–2348 (1980).
28. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72,** 3961–3965 (1975).

29. Birnboim, H. C. & Doly, J. *Nucleic Acids Res.* **7,** 1513–1523 (1979).
30. Smith, A. J. H. *Meth. Enzym.* **65,** 560–580 (1980).
31. Messing, J., Crea, R. & Seeburg, P. H. *Nucleic Acids Res.* **9,** 309–321 (1981).
32. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65,** 499–560 (1980).
33. Lawn, R. M. *et al. Nucleic Acids Res.* **9,** 6103–6114 (1981).
34. Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake G. & Maniatis, T. *Cell* **15,** 1157–1174 (1978).
35. Southern, E. M. *J. molec. Biol.* **98,** 503–517 (1975).
36. Taylor, J. M., Illmensee, R. & Summers, S. *Biochim. biophys. Acta* **442,** 324–330 (1976).
37. Benton, W. D. & Davis, R. W. *Science* **196,** 180–182 (1977).
38. Maniatis, T. *et al. Cell* **15,** 687–701 (1978).
39. Kafatos, F. C., Jones, C. W. & Efstratiadis, A. *Nucleic Acids Res.* **7,** 1541–1552 (1979).
40. Goeddel, D. V. *et al. Nature* **287,** 411–416 (1980).
41. Blobel, G. *et al. Symp. Soc. exp. Med.* **33,** 9–36 (1979).
42. Patterson, J. E. & Geller, D. M. *Biochem. biophys. Res. Commun.* **74,** 1220–1226 (1977).
43. Proudfoot, N. J. & Brownlee, G. G. *Nature* **263,** 211–214 (1976).
44. Winzler, R. J. in *Hormonal Proteins and Peptides* (ed. Li, C. H.) 1–15 (Academic, New York, 1973).
45. Granelli-Piperno, A. & Reich, E. *J. exp. Med.* **148,** 223–234 (1978).
46. Barlow, G. H. & Robbins, K. C. in *Hemostasis and Thrombosis: Basic Principles and Clinical Practice,* 1028–1036 (Lippincott, Philadelphia, 1982).
47. Wallen, P. & Wiman, B. *Biochim. biophys. Acta* **221,** 20–30 (1970).
48. Wallen, P. & Wiman, B. *Biochim. biophys. Acta* **257,** 122–234 (1972).
49. Sottrup-Jensen, L., Claeys, H., Zajdel, M., Peterson, T. E. & Magnusson, S. in *Progress in Chemical Fibrinolysis and Thrombolysis* Vol. 3 (eds Davidson, J. F., Rowan, R. M., Samana, M. M. & Desnoyers, P.C.) 191–209 (Raven, New York, 1978).
50. Sottrup-Jensen, L., Zajdel, M., Claeys, H., Petersen, T. E. & Magnussón, S. *Proc. natn. Acad. Sci. U.S.A.* **72,** 2577–2581 (1975).
51. Magnusson, S., Sottrup-Jensen, L., Petersen, T. E., Dudek-Wojciechowska, G. & Claeys, H. in *Proteolysis and Physiological Regulation* (eds Ribbons, D. W. & Brew, K.) 203–212 (Academic, New York, 1976).
52. Magnusson, S., Petersen, T. E., Sottrup-Jensen, L. & Claeys, H. in *Proteases and Biological Control* (eds Reich, E., Rifkin, D. B. & Shaw, E.) 123–149 (Cold Spring Harbor Laboratory, New York, 1975).
53. Heyneker, H. L. *et al. Proc. 4th int. Symp. on Genetics of Industrial Microorganisms* June 1982, Kyoto, Japan (in the press).
54. Dayhoff, M. O. *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3 (National Biomedical Research Foundation, Washington DC, 1978).
55. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
56. Bonner, W. M. & Laskey, R. A. *Eur. J. Biochem.* **46,** 83–88 (1974).
57. Ward, R., Banerjee, A. K., LaFiandra, A. & Shatkin, A. J. *J. Virol.* **9,** 61–69 (1972).
58. Singer, R. H. & Penman, S. *J. molec. Biol.* **78,** 321–335 (1973).
59. Chang, A. C. Y. *et al. Nature* **275,** 617–624 (1978).
60. Bolivar, F. *et al.* Gene **2,** 95–113 (1977).
61. Miller, J. *Experiments in Molecular Genetics,* 431–433 (Cold Spring Harbor Laboratory, New York, 1972).

LETTERS TO NATURE

A QSO in a rich, distant cluster of galaxies

Bruce Margon*, Ronald A. Downes† & Hyron Spinrad‡

* Department of Astronomy, University of Washington, FM-20, Seattle, Washington 98185, USA
† Department of Astronomy, University of California, Los Angeles, California 90024, USA
‡ Department of Astronomy, University of California, Berkeley, California 94720, USA

We have previously remarked[1] that there are no confirmed cases of quasistellar objects (QSOs) that are members of rich clusters of galaxies. Several candidates exist (cited in ref. 1), but all suffer from defects. In some of these cases, the linear dimensions may be more appropriate for superclusters than clusters (an interesting result nonetheless), and in other cases only the QSO members of the putative clusters are visible, and the normal galaxies are inferred but not observed. A particularly direct confirmation that QSOs share the same clustering characteristics as other constituents of the Universe would come from the detection of a 'classical' (for example, point-like, high luminosity) QSO in a previously studied, distant rich cluster of visible galaxies. Here we report such a discovery.

The cluster 0016+16, at $z = 0.541$ (ref. 2) is perhaps the best studied distant cluster of galaxies. Detailed multicolour photometry and isopleths[3] show this cluster to be richer than the Coma Cluster, and remarkably normal except for its very large redshift. The brightest members have red magnitudes of ~20.4. The cluster is a luminous, extended X-ray source[4], indicating the presence of an ionized intracluster gas as is common in rich clusters. There is also a detection[5] of a decrement in the microwave background radiation against 0016+16, so the intracluster gas parameters can be studied in some detail, and (at least in principle) the cosmological parameters q_0 and H_0 may be estimated[6].

During X-ray observations of 0016+16, a previously unreported point source near the northern edge of the visible galaxy distribution was detected[4], and the identification of this source with an unstudied $B \sim 18$ blue stellar object has been suggested[3,4]. We have found[7,8] that a large fraction of these serendipitously-observed faint X-ray sources are associated with previously uncatalogued QSOs. With this motivation, we obtained medium resolution (10 Å) spectroscopy of the 18th mag object. On 25 July 1982, the blue spectrum was observed with the SIT Vidicon spectrometer[9] at the 4-m reflector of the Cerro Tololo Interamerican Observatory; these data were converted to absolute fluxes through observation of spectrophotometric standard stars. The red/IR region of the spectrum was observed on 15 October 1982, using the image tube scanner[10] at the Lick Observatory 3-m Shane telescope.

The resulting spectra are shown in Fig. 1. This object is clearly a QSO, which we denote as 0015+162. We identify the prominent broad emission lines at 4,347, 6,749 and 7,557 Å as Mg II λ2,800, Hγ and Hβ, respectively, and derive redshift $z = 0.554 \pm 0.002$. There is a marginally positive detection of [O III] λ5,007 emission, at an atypically weak intensity, less than Hβ. The [O III] λ4,959 line is not separable from the very strong OH λ7,715 night-sky emission at our spectral resolution, and similarly any [O II] λ3,727 emission would coincide very closely with the Hg λ5,790 emission from city lights, which appears very strongly in the Lick data. Our spectrophotometry yields $B \sim 18.2$, $(B-V) \sim 0$, which for $q_0 = 0$, $H_0 = 50$ km s^{-1} Mpc^{-1} implies an absolute V magnitude of −25. Our magnitude combined with the observed X-ray intensity[4] implies a value for α_{ox}, the slope of a line connecting the optical and X-ray fluxes, of 1.32. This is quite close to the mean found for both serendipitously-discovered[7,8] and previously known[11,12] QSOs, and suggests that 0015+162 is quite a normal quasar. Previous radio observations[5] searching for point sources near the cluster indicate the object to be radio-quiet, as is expected for a randomly-selected QSO.

There seems little doubt that the QSO 0015+162 and cluster 0016+16 are related. The surface density of QSOs this bright is very small[13]. We expect less than one QSO this bright to occur by accident at any redshift and at any location on the entire 4-m telescope plate that discovered the cluster, much less at a redshift essentially identical to that of the cluster, co-located on the 10^{-2} of the plate that contains the cluster. The simplest explanation of this association is that the QSO 0015+162 is a member of the rich cluster of galaxies 0016+16. The formal velocity difference between the objects is currently 2,000 km s^{-1} in the observed frame, which must contain components due to the observational uncertainty in our determination of the QSO redshift, the analogous quantity for the exceedingly faint galaxy observed to obtain the cluster redshift, and the velocity dispersion expected for a rich cluster. Each of these three individual quantities, much less their combination, is of magnitude comparable with the observed difference. The core radius of the cluster[3] is measured as 0.75 arc min (~0.4 Mpc), implying that the QSO is 4.5 core radii from the optical and X-ray centre of the cluster. Such an offset is, of course, quite consistent with cluster membership. In the classic study of membership of the Coma Cluster[14], for example, 60% of all galaxies assigned as cluster members are more distant than 4 core radii from the cluster centre. A more recent study[15] finds an even more extreme percentage of members beyond this distance.