



(22) **Date de dépôt/Filing Date:** 2004/02/10

(41) **Mise à la disp. pub./Open to Public Insp.:** 2004/08/26

(62) **Demande originale/Original Application:** 2 515 699

(30) **Priorité/Priority:** 2003/02/10 (US60/446,633)

(51) **Cl.Int./Int.Cl. C12N 15/12** (2006.01),
A61K 38/17 (2006.01), **A61K 39/00** (2006.01),
A61K 39/395 (2006.01), **A61K 47/48** (2006.01),
A61K 51/10 (2006.01), **A61P 35/00** (2006.01),
A61P 35/02 (2006.01), **A61P 37/04** (2006.01),
C07K 14/47 (2006.01), **C07K 14/82** (2006.01),
C07K 16/18 (2006.01), ...

(71) **Demandeur/Applicant:**
AGENSYS, INC., US

(72) **Inventeurs/Inventors:**
JAKOBOVITS, AYA, US;
RAITANO, ARTHUR B., US;
CHALLITA-EID, PIA M., US; ...

(54) **Titre : ACIDE NUCLEIQUE ET PROTEINE CORRESPONDANTE DITE 158P1D7, UTILES POUR LE TRAITEMENT ET LA
DETECTION DE CANCERS DE LA VESSIE ET AUTRES**

(54) **Title: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 158P1D7 USEFUL IN THE TREATMENT AND DETECTION
OF BLADDER AND OTHER CANCERS**

158P1D7 SSH_sequence {SEQ ID NO: 1}.

```
1  GATCTGATAA  GCTTTCAATG  TTGCGCTCCT  GACAATGTAT  TAGAAGTCCT  GATGGGGATA
61  GGACTTTGCA  GTTACAAGGA  ATAGGGCAGA  AAGGTCCTGG  AAGTTGAGTG  GATGGCTTTG
121 TAATATAAGG  TATCAAACCT  GGTGCTTTGG  TGGGTAGTTT  TAGAATGGAC  GTGGTCTTAG
181 TTGACATGCG  ACTATCATTT  ATTGAAGATG  TTGCTGCCAG  ATGTAATGAT  C
```

(57) **Abrégé/Abstract:**

A novel gene (designated 158P1D7) and its encoded protein are described. While 158P1D7 exhibits tissue specific expression in normal adult tissue, it is aberrantly expressed in multiple cancers including set forth in Table 1. Consequently, 158P1D7 provides a diagnostic and/or therapeutic target for cancers. The 158P1D7 gene or fragment thereof, or its encoded protein or a fragment thereof, can be used to elicit an immune response.

(51) **Cl.Int./Int.Cl. (suite/continued)** *C12N 15/11* (2006.01), *C12N 5/16* (2006.01), *C12Q 1/68* (2006.01),
G01N 33/574 (2006.01)

(72) **Inventeurs(suite)/Inventors(continued):** PEREZ-VILLAR, JUAN J., US; MORRISON, KAREN JANE MEYRICK, US;
FARIS, MARY, US; GE, WANGMAO, US; GUDAS, JEAN, US; KANNER, STEVEN B., US;
MORRISON, ROBERT KENDALL, US

(74) **Agent:** FETHERSTONHAUGH & CO.

Abstract

A novel gene (designated 158P1D7) and its encoded protein are described. While 158P1D7 exhibits tissue specific expression in normal adult tissue, it is aberrantly expressed in multiple cancers including set forth in Table 1. Consequently, 158P1D7 provides a diagnostic and/or therapeutic target for cancers. The 158P1D7 gene or fragment thereof, or its encoded protein or a fragment thereof, can be used to elicit an immune response.

CLAIMS

1. An isolated polynucleotide that encodes a 158P1D7 protein, wherein the polynucleotide is:
 - (a) a polynucleotide comprising the sequence of SEQ ID NO: 76;
 - (b) a polynucleotide comprising the sequence of SEQ ID NO: 84; or
 - (c) a polynucleotide comprising the sequence of SEQ ID NO: 90.
2. An isolated polynucleotide that encodes a 158P1D7 protein, wherein the polynucleotide encodes the polypeptide sequence shown in SEQ ID NO: 81, SEQ ID NO: 85 or SEQ ID NO: 91.
3. A polynucleotide that is fully complementary to the polynucleotide defined in claim 1 or 2.
4. A recombinant expression vector comprising a polynucleotide as defined in claim 1, 2 or 3.
5. The recombinant expression vector of claim 4, wherein the vector is a viral vector.
6. The recombinant expression vector of claim 5, wherein the viral vector encodes a virus that is: vaccinia, fowlpox, canarypox, adenovirus, influenza, poliovirus, adeno-associated virus, lentivirus, or sindbis virus.
7. A host cell that contains an expression vector as defined in claim 4, 5 or 6.
8. A process for producing a protein encoded by a polynucleotide as defined in claim 1 or 2, comprising culturing a host cell containing an expression vector comprising said polynucleotide, under conditions sufficient for the production of the protein.
9. The process of claim 8, further comprising recovering the protein so produced.
10. The process of claim 9, wherein the protein is recovered using chromatography.
11. An isolated protein, wherein the protein has the amino acid sequence of SEQ ID NO: 81, SEQ ID NO: 85, or SEQ ID NO: 91.
12. A composition comprising a pharmaceutically acceptable carrier and a protein as defined in claim 11.
13. An antibody or antigen binding fragment thereof that immunospecifically binds to a protein as defined in claim 11.

14. The antibody or fragment thereof of claim 13, which is monoclonal.
15. The antibody or fragment thereof of claim 13 or claim 14, wherein the fragment is an Fab, F(ab')₂, or Fv.
16. The antibody or fragment thereof of claim 13, 14 or 15, which is a human antibody or fragment thereof.
17. The antibody or fragment thereof of any one of claims 13 to 16, which is labeled with an agent.
18. The antibody or fragment thereof of claim 17, wherein the agent is a radioactive isotope, a chemotherapeutic agent or a toxin.
19. The antibody or fragment thereof of claim 17, wherein the agent is a radioactive isotope that is: ²¹¹At, ¹³¹I, ¹²⁵I, ⁹⁰Y, ¹⁸⁶Re, ¹⁸⁸Re, ¹⁵³Sm, ²¹²Bi, ³²P or a radioactive isotope of Lu.
20. The antibody or fragment thereof of claim 17, wherein the agent is a chemotherapeutic agent and the chemotherapeutic agent is: taxol, actinomycin, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicine, gelonin, or calicheamicin.
21. The antibody or fragment thereof of claim 17, wherein the agent is a toxin and the toxin is: diphtheria toxin, enomycin, phenomycin, Pseudomonas exotoxin (PE) A, PE40, abrin, abrin A chain, mitogellin, modeccin A chain, or alpha-sarcin.
22. A composition comprising an antibody or fragment thereof as defined in any one of claims 13 to 21, and a pharmaceutically acceptable carrier.
23. A hybridoma that produces an antibody as defined in claim 14.
24. An *in vitro* method for detecting presence of a protein as defined in claim 11 or a polynucleotide encoding said protein in a test sample, said method comprising:
 - contacting the test sample with an antibody that specifically binds to the protein or a polynucleotide that is complimentary to the polypeptide that encodes the protein; and
 - detecting binding in the test sample of the antibody or the complementary polynucleotide, thereby detecting the presence of the protein or the polynucleotide encoding the protein in the test sample.
25. The method of claim 24, wherein the detecting is of said polynucleotide encoding the protein, which is an mRNA.

26. The method of claim 24, wherein the detecting is of said polynucleotide encoding the protein, which is a cDNA produced from the sample by reverse transcription.

27. The method of claim 24, 25 or 26, wherein the detecting step comprises comparing an amount of binding of the antibody or polynucleotide to the protein or the polynucleotide encoding the protein, respectively, in the test sample, and an amount of binding of the antibody or polynucleotide to the protein or the polynucleotide encoding the protein, respectively, in a corresponding normal sample.

28. The method of claim 27, wherein an elevated amount of binding of the antibody or polynucleotide to the protein or the polynucleotide encoding the protein, respectively, in the test sample relative to the normal sample provides an indication of the presence of cancer in the test sample.

29. The method of claim 28, wherein the cancer is leukemia or cancer of the prostate, testis, kidney, brain, bone, skin, ovary, breast, pancreas, colon, or lung, and the test and normal tissue samples are serum, blood, urine, or tissues of the prostate, testis, kidney, brain, bone, skin, ovary, breast, pancreas, colon, or lung.

30. Use of a antibody or fragment thereof as defined in any one of claims 13 to 16 for delivering a cytotoxic agent to a cell expressing a protein as defined in claim 11.

31. Use of an antibody or fragment thereof as defined in any one of claims 13 to 16 in preparing a medicament for delivering a cytotoxic agent to a cell expressing a protein as defined in claim 11.

32. The antibody or fragment thereof of any one of claims 13 to 16 for use in delivering a cytotoxic agent to a cell expressing a protein as defined in claim 11.

33. The antibody or fragment thereof of any one of claims 13 to 16 for use in preparing a medicament for delivering a cytotoxic agent to a cell expressing a protein as defined in claim 11.

34. Use of a protein as defined in claim 11 for inducing an immune response to the protein in a subject.

35. Use of a protein as defined in claim 11 in preparing a medicament for inducing an immune response to the protein in a subject.

36. The use of claim 34 or claim 35, wherein the immune response comprises activation of B cells, wherein the activated B cells generate antibodies that specifically bind to the protein.

37. The use of claim 34 or claim 35, wherein the immune response comprises activation of a T cell, wherein the activated T cell is a cytotoxic T cell (CTL), which, when activated kills an autologous cell that expresses the protein.

38. The use of claim 34 or claim 35, wherein the immune response comprises activation of a T cell, wherein the activated T cell is a helper T cell (HTL), which, when activated secretes cytokines that facilitate cytotoxic activity of a CTL or antibody producing activity of a B cell.

39. The protein of claim 11 for use in inducing an immune response to the protein in a subject.

40. The protein of claim 11 for use in preparing a medicament for inducing an immune response to the protein in a subject.

41. The protein of claim 39 or claim 40, wherein the immune response comprises activation of B cells, wherein the activated B cells generate antibodies that specifically bind to the protein.

42. The protein of claim 39 or claim 40, wherein the immune response comprises activation of a T cell, wherein the activated T cell is a cytotoxic T cell (CTL), which, when activated kills an autologous cell that expresses the protein.

43. The protein of claim 39 or claim 40, wherein the immune response comprises activation of a T cell, wherein the activated T cell is a helper T cell (HTL), which, when activated secretes cytokines that facilitate cytotoxic activity of a CTL or antibody producing activity of a B cell.

44. Use of an antibody as defined in any one of claims 13 to 16 in preparation of a medicament for delivering an agent to a cell expressing a protein as defined in claim 11.

45. Use of an antibody as defined in any one of claims 13 to 16 for delivering an agent to a cell expressing a protein as defined in claim 11.

46. The antibody of any one of claims 13 to 16 for use in delivering an agent to a cell expressing a protein as defined in claim 11.

47. The antibody of any one of claims 13 to 16 for use in preparing a medicament for delivering an agent to a cell expressing a protein as defined in claim 11.

Figure 1: 158P1D7 SSH_sequence (SEQ ID NO: 1).

```
1 GATCTGATAA GCTTTCAATG TTGCGCTCCT GACAATGTAT TAGAAGTCCT GATGGGGATA
61 GGACTTTGCA GTTACAAGGA ATAGGGCAGA AAGGTCCTGG AAGTTGAGTG GATGGCTTTG
121 TAATATAAGG TATCAAACCT GGTGCTTTGG TGGGTAGTTT TAGAATGGAC GTGGTCTTAG
181 TTGACATGCG ACTATCATTT ATTGAAGATG TTGCTGCCAG ATGTAATGAT C
```

Figure 2:

Figure 2A. The cDNA (SEQ ID NO: 2) and amino acid sequence (SEQ ID NO: 3) of 158P1D7 v.1.
The start methionine is underlined. The open reading frame extends from nucleic acid 23-2548 including the stop codon.

```
1           M K L W I H L F Y S S L L
1 tcggatttcacatcacatgacaacATGAAGCTGTGGATTCATCTCTTTTATTCATCTCTCCT
14 A C I S L H S Q T P V L S S R G S C D S
61 TGCCTGTATATCTTTACACTCCCAAACCTCCAGTGCTCTCATCCAGAGGCTCTTGTGATTC
34 L C N C E E K D G T M L I N C E A K G I
121 TCTTTGCAATTGTGAGGAAAAAGATGGCACAATGCTAATAAATTGTGAAGCAAAAGGTAT
54 K M V S E I S V P P S R P F Q L S L L N
181 CAAGATGGTATCTGAAATAAGTGTGCCACCATCACGACCTTTCCAACCTAAGCTTATTAA
74 N G L T M L H T N D F S G L T N A I S I
241 TAACGGCTTGACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTATTTCAAT
94 H L G F N N I A D I E I G A F N G L G L
301 ACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCATTTAATGGCCTTGGCCT
114 L K Q L H I N H N S L E I L K E D T F H
361 CCTGAAACAACCTTCATATCAATCACAATTCTTTAGAAATTCTTAAAGAGGATACTTTCCA
134 G L E N L E F L Q A D N N F I T V I E P
421 TGGACTGGAAAACCTGGAATTCCTGCAAGCAGATAACAATTTTATCACAGTGATTGAACC
154 S A F S K L N R L K V L I L N D N A I E
481 AAGTGCCTTTAGCAAGCTCAACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGA
174 S L P P N I F R F V P L T H L D L R G N
541 GAGTCTTCCTCCAAACATCTTCCGATTTGTTCTTTAACCCATCTAGATCTTCGTGGAAA
194 Q L Q T L P Y V G F L E H I G R I L D L
601 TCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATTGGCCGAATATTGGATCT
214 Q L E D N K W A C N C D L L Q L K T W L
661 TCAGTTGGAGGACAACAAATGGGCCTGCAATTGTGACTTATTGCAGTTAAAACTTGGTT
234 E N M P P Q S I I G D V V C N S P P F F
721 GGAGAACATGCCTCCACAGTCTATAATTGGTGATGTTGTCTGCAACAGCCCTCCATTTT
254 K G S I L S R L K K E S I C P T P P V Y
781 TAAAGGAAGTATACTCAGTAGACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTA
274 E E H E D P S G S L H L A A T S S I N D
841 TGAAGAACATGAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAATAAATGA
294 S R M S T K T T S I L K L P T K A P G L
901 TAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACCTACCCACCAAAGCACCAGGTTT
314 I P Y I T K P S T Q L P G P Y C P I P C
961 GATACCTTATATTACAAAGCCATCCACTCAACTTCCAGGACCTTACTGCCCTATTCCTTG
334 N C K V L S P S G L L I H C Q E R N I E
1021 TAACTGCAAAGTCCTATCCCCATCAGGACTTCTAATACATTGTCAGGAGCGCAACATTGA
354 S L S D L R P P P Q N P R K L I L A G N
```

1081 AAGCTTATCAGATCTGAGACCTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAA
374 I I H S L M K S D L V E Y F T L E M L H
1141 TATTATTCACAGTTTAATGAAGTCTGATCTAGTGAATATTTCACTTTGGAAATGCTTCA
394 L G N N R I E V L E E G S F M N L T R L
1201 CTTGGGAAACAATCGTATTGAAGTTCTTGAAGAAGGATCGTTTATGAACCTAACGAGATT
414 Q K L Y L N G N H L T K L S K G M F L G
1261 ACAAAAACCTCTATCTAAATGGTAACCACCTGACCAAATTAAGTAAAGGCATGTTCCCTGG
434 L H N L E Y L Y L E Y N A I K E I L P G
1321 TCTCCATAATCTTGAATACTTATATCTTGAATACAATGCCATTAAGGAAATACTGCCAGG
454 T F N P M P K L K V L Y L N N N L L Q V
1381 AACCTTTAATCCAATGCCTAAACTTAAAGTCCTGTATTTAAATAACAACCTCCTCCAAGT
474 L P P H I F S G V P L T K V N L K T N Q
1441 TTTACCACCACATATTTTTTCAGGGGTTCTCTAACTAAGGTAAATCTTAAAACAAACCA
494 F T H L P V S N I L D D L D L L T Q I D
1501 GTTTACCCATCTACCTGTAAGTAATATTTTGGATGATCTTGATTACTAACCCAGATTGA
514 L E D N P W D C S C D L V G L Q Q W I Q
1561 CCTTGAGGATAACCCCTGGGACTGCTCCTGTGACCTGGTGGACTGCAGCAATGGATACA
534 K L S K N T V T D D I L C T S P G H L D
1621 AAAGTTAAGCAAGAACACAGTGACAGATGACATCCTCTGCACTTCCCCCGGCATCTCGA
554 K K E L K A L N S E I L C P G L V N N P
1681 CAAAAGGAATTGAAAGCCCTAAATAGTGAAATTCTCTGTCCAGGTTTAGTAAATAACCC
574 S M P T Q T S Y L M V T T P A T T T N T
1741 ATCCATGCCAACACAGACTAGTTACCTTATGGTCACCACTCCTGCAACAACAACAAATAC
594 A D T I L R S L T D A V P L S V L I L G
1801 GGCTGATACTATTTTACGATCTCTTACGGACGCTGTGCCACTGTCTGTTCTAATATTGGG
614 L L I M F I T I V F C A A G I V V L V L
1861 ACTTCTGATTATGTTTCATCACTATTGTTTTCTGTGCTGCAGGGATAGTGGTTCTTGTCT
634 H R R R R Y K K K Q V D E Q M R D N S P
1921 TCACCGCAGGAGAAGATACAAAAGAAACAAGTAGATGAGCAAATGAGAGACAACAGTCC
654 V H L Q Y S M Y G H K T T H H T T E R P
1981 TGTGCATCTTCAGTACAGCATGTATGGCCATAAAACCACTCATCACACTACTGAAAGACC
674 S A S L Y E Q H M V S P M V H V Y R S P
2041 CTCTGCCTCACTCTATGAACAGCACATGGTGAGCCCCATGGTTCATGTCTATAGAAGTCC
694 S F G P K H L E E E E E R N E K E G S D
2101 ATCCTTTGGTCCAAAGCATCTGGAAGAGGAAGAAGAGAGGAATGAGAAAGAAGGAAGTGA
714 A K H L Q R S L L E Q E N H S P L T G S
2161 TGCAAAACATCTCCAAGAAGTCTTTTGGAACAGGAAAATCATTACCACTCACAGGGTC
734 N M K Y K T T N Q S T E F L S F Q D A S
2221 AAATATGAAATACAAAACCACGAACCAATCAACAGAATTTTTATCCTTCCAAGATGCCAG
754 S L Y R N I L E K E R E L Q Q L G I T E
2281 CTCATTGTACAGAAACATTTTAGAAAAAGAAAGGGAACCTCAGCAACTGGGAATCACAGA
774 Y L R K N I A Q L Q P D M E A H Y P G A

2341 ATACCTAAGGAAAAACATTGCTCAGCTCCAGCCTGATATGGAGGCACATTATCCTGGAGC
 794 H E E L K L M E T L M Y S R P R K V L V
 2401 CCACGAAGAGCTGAAGTTAATGGAAACATTAATGTACTCACGTCCAAGGAAGGTATTAGT
 814 E Q T K N E Y F E L K A N L H A E P D Y
 2461 GGAACAGACAAAAAATGAGTATTTTGAACCTTAAAGCTAATTTACATGCTGAACCTGACTA
 834 L E V L E Q Q T *
 2521 TTTAGAAGTCCTGGAGCAGCAAACATAGatggaga

Figure 2B. The cDNA (SEQ ID NO: 4) and amino acid sequence (SEQ ID NO: 5) of 158P1D7 v.2.
 The start methionine is underlined. The open reading frame extends from nucleic acid 23-2548 including the stop codon.

1 M K L W I H L F Y S S L L
 1 tcggatttcatcacatgacaacATGAAGCTGTGGATTCATCTCTTTTATTCATCTCTCCT
 14 A C I S L H S Q T P V L S S R G S C D S
 61 TGCCTGTATATCTTTACACTCCCAACTCCAGTGCTCTCATCCAGAGGCTCTTGATTC
 34 L C N C E E K D G T M L I N C E A K G I
 121 TCTTTGCAATTGTGAGGAAAAAGATGGCACAATGCTAATAAATTGTGAAGCAAAGGTAT
 54 K M V S E I S V P P S R P F Q L S L L N
 181 CAAGATGGTATCTGAAATAAGTGTGCCACCATCACGACCTTTCCAATAAGCTTATTAAA
 74 N G L T M L H T N D F S G L T N A I S I
 241 TAACGGCTTGACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTATTTCAAT
 94 H L G F N N I A D I E I G A F N G L G L
 301 ACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCATTTAATGGCCTTGGCCT
 114 L K Q L H I N H N S L E I L K E D T F H
 361 CCTGAAACAACTTCATATCAATCACAATTCTTTAGAAATTCTTAAAGAGGATACTTTCCA
 134 G L E N L E F L Q A D N N F I T V I E P
 421 TGGACTGGAAAACCTGGAATTCCTGCAAGCAGATAACAATTTTATCACAGTGATTGAACC
 154 S A F S K L N R L K V L I L N D N A I E
 481 AAGTGCCTTTAGCAAGCTCAACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGA
 174 S L P P N I F R F V P L T H L D L R G N
 541 GAGTCTTCCTCCAAACATCTTCCGATTTGTTCTTTAACCCTATCTAGATCTTCGTGGAAA
 194 Q L Q T L P Y V G F L E H I G R I L D L
 601 TCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATTGGCCGAATATTGGATCT
 214 Q L E D N K W A C N C D L L Q L K T W L
 661 TCAGTTGGAGGACAACAAATGGGCCTGCAATTGTGACTTATTGCAGTTAAAACTTGGTT
 234 E N M P P Q S I I G D V V C N S P P F F
 721 GGAGAACATGCCTCCACAGTCTATAATTGGTGATGTTGTCTGCAACAGCCCTCCATTTTT
 254 K G S I L S R L K K E S I C P T P P V Y
 781 TAAAGGAAGTATACTCAGTAGACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTA
 274 E E H E D P S G S L H L A A T S S I N D
 841 TGAAGAACATGAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAATAAATGA

294 S R M S T K T T S I L K L P T K A P G L
901 TAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACTACCCACCAAGCACCAGGTTT
314 I P Y I T K P S T Q L P G P Y C P I P C
961 GATACCTTATATTACAAAGCCATCCACTCAACTTCCAGGACCTTACTGCCCTATTCCTTG
334 N C K V L S P S G L L I H C Q E R N I E
1021 TAACTGCAAAGTCCTATCCCCATCAGGACTTCTAATACATTGTCAGGAGCGCAACATTGA
354 S L S D L R P P P Q N P R K L I L A G N
1081 AAGCTTATCAGATCTGAGACCTCCTCCGCAAATCCTAGAAAGCTCATTCTAGCGGGAAA
374 I I H S L M K S D L V E Y F T L E M L H
1141 TATTATTCACAGTTTAATGAAGTCTGATCTAGTGAATATTTCACTTTGGAAATGCTTCA
394 L G N N R I E V L E E G S F M N L T R L
1201 CTTGGGAAACAATCGTATTGAAGTTCTTGAAGAAGGATCGTTTATGAACCTAACGAGATT
414 Q K L Y L N G N H L T K L S K G M F L G
1261 ACAAAACTCTATCTAAATGGTAACCACCTGACCAAATTAAGTAAAGGCATGTTCTTGG
434 L H N L E Y L Y L E Y N A I K E I L P G
1321 TCTCCATAATCTTGAATACTTATATCTTGAATACAATGCCATTAAGGAAATACTGCCAGG
454 T F N P M P K L K V L Y L N N N L L Q V
1381 AACCTTTAATCCAATGCCTAAACTTAAAGTCCTGTATTTAAATAACAACCTCCTCCAAGT
474 L P P H I F S G V P L T K V N L K T N Q
1441 TTTACCACCACATATTTTTTTCAGGGGTCCTCTAACTAAGGTAAATCTTAAACAAACCA
494 F T H L P V S N I L D D L D L L T Q I D
1501 GTTTACCCATCTACCTGTAAGTAATATTTTGGATGATCTTGATTTGCTAACCAGATTGA
514 L E D N P W D C S C D L V G L Q Q W I Q
1561 CCTTGAGGATAACCCCTGGGACTGCTCCTGTGACCTGGTTGGACTGCAGCAATGGATACA
534 K L S K N T V T D D I L C T S P G H L D
1621 AAAGTTAAGCAAGAACACAGTGACAGATGACATCCTCTGCACTTCCCCCGGCATCTCGA
554 K K E L K A L N S E I L C P G L V N N P
1681 CAAAAGGAATTGAAAGCCCTAAATAGTGAAATCTCTGTCCAGGTTTAGTAAATAACCC
574 S M P T Q T S Y L M V T T P A T T T N T
1741 ATCCATGCCAACACAGACTAGTTACCTTATGGTCACCACTCCTGCAACAACAACAAATAC
594 A D T I L R S L T D A V P L S V L I L G
1801 GGCTGATACTATTTTACGATCTCTTACGGACGCTGTGCCACTGTCTGTTCTAATATTGGG
614 L L I M F I T I V F C A A G I V V L V L
1861 ACTTCTGATTATGTTTCATCACTATTGTTTTCTGTGCTGCAGGGATAGTGGTTCTTGTTCT
634 H R R R R Y K K K Q V D E Q M R D N S P
1921 TCACCGCAGGAGAAGATACAAAAGAAACAAGTAGATGAGCAAATGAGAGACAACAGTCC
654 V H L Q Y S M Y G H K T T H H T T E R P
1981 TGTGCATCTTCAGTACAGCATGTATGGCCATAAAACCACTCATCACACTACTGAAAGACC
674 S A S L Y E Q H M V S P M V H V Y R S P
2041 CTCTGCCTCACTCTATGAACAGCACATGGTGAGCCCCATGGTTCATGTCTATAGAAGTCC
694 S F G P K H L E E E E E R N E K E G S D
2101 ATCCTTTGGTCCAAAGCATCTGGAAGAGGAAGAAGAGAGGAATGAGAAAGAAGGAAGTGA

```

714 A K H L Q R S L L E Q E N H S P L T G S
2161 TGCAAAACATCTCCAAAGAAGTCTTTTGGAAACAGGAAAATCATTCACTACTCACAGGGTC
734 N M K Y K T T N Q S T E F L S F Q D A S
2221 AAATATGAAATACAAAACCACGAACCAATCAACAGAATTTTATCCTTCCAAGATGCCAG
754 S L Y R N I L E K E R E L Q Q L G I T E
2281 CTCATTGTACAGAAACATTTTAGAAAAAGAAAGGGAACCTTCAGCAACTGGGAATCACAGA
774 Y L R K N I A Q L Q P D M E A H Y P G A
2341 ATACCTAAGGAAAAACATTGCTCAGCTCCAGCCTGATATGGAGGCACATTATCCTGGAGC
794 H E E L K L M E T L M Y S R P R K V L V
2401 CCACGAAGAGCTGAAGTTAATGGAAACATTAATGTACTCACGTCCAAGGAAGGTATTAGT
814 E Q T K N E Y F E L K A N L H A E P D Y
2461 GGAACAGACAAAAAATGAGTATTTTGAACCTTAAAGCTAATTTACATGCTGAACCTGACTA
834 L E V L E Q Q T *
2521 TTTAGAAGTCCTGGAGCAGCAAACATAGatggaga

```

Figure 2C. The cDNA (SEQ ID NO: 6) and amino acid sequence (SEQ ID NO: 7) of 158P1D7 v.3.
The start methionine is underlined. The open reading frame extends from nucleic acid 23-2221 including the stop codon.

```

1 M K L W I H L F Y S S L L
1 tcggatttcatcacatgacaacATGAAGCTGTGGATTCATCTCTTTTATTCATCTCTCCT
14 A C I S L H S Q T P V L S S R G S C D S
61 TGCCTGTATATCTTTACACTCCCAAACTCCAGTGCTCTCATCCAGAGGCTCTTGTGATTC
34 L C N C E E K D G T M L I N C E A K G I
121 TCTTTGCAATTGTGAGGAAAAAGATGGCACAATGCTAATAAATTGTGAAGCAAAAGGTAT
54 K M V S E I S V P P S R P F Q L S L L N
181 CAAGATGGTATCTGAAATAAGTGTGCCACCATCACGACCTTTCCAACCTAAGCTTATTAAA
74 N G L T M L H T N D F S G L T N A I S I
241 TAACGGCTTGACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTATTTCAAT
94 H L G F N N I A D I E I G A F N G L G L
301 ACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCATTTAATGGCCTTGGCCT
114 L K Q L H I N H N S L E I L K E D T F H
361 CCTGAAACAACTTCATATCAATCACAATTCTTTAGAAATTCTTAAAGAGGATACTTTCCA
134 G L E N L E F L Q A D N N F I T V I E P
421 TGGACTGGAAAACCTGGAATTCCTGCAAGCAGATAACAATTTTATCACAGTGATTGAACC
154 S A F S K L N R L K V L I L N D N A I E
481 AAGTGCCTTTAGCAAGCTCAACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGA
174 S L P P N I F R F V P L T H L D L R G N
541 GAGTCTTCCTCCAAACATCTTCGATTTGTTCCCTTAACCCATCTAGATCTTCGTGGAAA
194 Q L Q T L P Y V G F L E H I G R I L D L
601 TCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATTGGCCGAATATTGGATCT
214 Q L E D N K W A C N C D L L Q L K T W L

```

661 TCAGTTGGAGGACAACAAATGGGCCTGCAATTGTGACTTATTGCAGTTAAAACTTGGTT
234 E N M P P Q S I I G D V V C N S P P F F
721 GGAGAACATGCCTCCACAGTCTATAATTGGTGATGTTGTCTGCAACAGCCCTCCATTTTT
254 K G S I L S R L K K E S I C P T P P V Y
781 TAAAGGAAGTATACTCAGTAGACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTA
274 E E H E D P S G S L H L A A T S S I N D
841 TGAAGAACATGAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAATAAATGA
294 S R M S T K T T S I L K L P T K A P G L
901 TAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACTACCCACCAAAGCACCAGGTTT
314 I P Y I T K P S T Q L P G P Y C P I P C
961 GATACCTTATATTACAAAGCCATCCACTCAACTTCCAGGACCTTACTGCCCTATTCCTTG
334 N C K V L S P S G L L I H C Q E R N I E
1021 TAACTGCAAAGTCCTATCCCCATCAGGACTTCTAATACATTGTCAGGAGCGCAACATTGA
354 S L S D L R P P P Q N P R K L I L A G N
1081 AAGCTTATCAGATCTGAGACCTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAA
374 I I H S L M K S D L V E Y F T L E M L H
1141 TATTATTACAGTTTAAATGAAGTCTGATCTAGTGGAAATATTTCACTTTGGAAATGCTTCA
394 L G N N R I E V L E E G S F M N L T R L
1201 CTTGGGAAACAATCGTATTGAAGTTCTTGAAGAAGGATCGTTTATGAACCTAACGAGATT
414 Q K L Y L N G N H L T K L S K G M F L G
1261 ACAAAAACCTCTATCTAAATGGTAACCACCTGACCAAATTAAGTAAAGGCATGTTCCCTTGG
434 L H N L E Y L Y L E Y N A I K E I L P G
1321 TCTCCATAATCTTGAATACTTATATCTTGAATACAATGCCATTAAGGAAATACTGCCAGG
454 T F N P M P K L K V L Y L N N N L L Q V
1381 AACCTTTAATCCAATGCCTAAACTTAAAGTCCTGTATTTAAATAACAACCTCCTCCAAGT
474 L P P H I F S G V P L T K V N L K T N Q
1441 TTTACCACCACATATTTTTTTCAGGGGTTCTCTAACTAAGGTAAATCTTAAACAAACCA
494 F T H L P V S N I L D D L D L L T Q I D
1501 GTTTACCCATCTACCTGTAAGTAATTTTTGGATGATCTTGATTTACTAACCCAGATTGA
514 L E D N P W D C S C D L V G L Q Q W I Q
1561 CCTTGAGGATAACCCCTGGGACTGCTCCTGTGACCTGGTTGGACTGCAGCAATGGATACA
534 K L S K N T V T D D I L C T S P G H L D
1621 AAAGTTAAGCAAGAACACAGTGACAGATGACATCCTCTGCACTTCCCCCGGGCATCTCGA
554 K K E L K A L N S E I L C P G L V N N P
1681 CAAAAGGAATTGAAAGCCCTAAATAGTGAAATTCTCTGTCCAGGTTTAGTAAATAACCC
574 S M P T Q T S Y L M V T T P A T T T N T
1741 ATCCATGCCAACACAGACTAGTTACCTTATGGTCACCACTCCTGCAACAACAACAAATAC
594 A D T I L R S L T D A V P L S V L I L G
1801 GGCTGATACTATTTTACGATCTCTTACGGACGCTGTGCCACTGTCTGTTCTAATATTGGG
614 L L I M F I T I V F C A A G I V V L V L
1861 ACTTCTGATTATGTTTCATCACTATTGTTTTCTGTGCTGCAGGGATAGTGGTTCTTGTCT
634 H R R R R Y K K K Q V D E Q M R D N S P

1921 TCACCGCAGGAGAAGATACAAAAAGAAACAAGTAGATGAGCAAATGAGAGACAACAGTCC
 654 V H L Q Y S M Y G H K T T H H T T E R P
 1981 TGTGCATCTTCAGTACAGCATGTATGGCCATAAAACCACTCATCACACTACTGAAAGACC
 674 S A S L Y E Q H M G A H E E L K L M E T
 2041 CTCTGCCTCACTCTATGAACAGCACATGGGAGCCCACGAAGAGCTGAAGTTAATGGAAAC
 694 L M Y S R P R K V L V E Q T K N E Y F E
 2101 ATTAATGTACTCAGTCCAAGGAAGGTATTAGTGGAACAGACAAAAAATGAGTATTTTGA
 714 L K A N L H A E P D Y L E V L E Q Q T *
 2161 ACTTAAAGCTAATTTACATGCTGAACCTGACTATTTAGAAGTCCTGGAGCAGCAAACATA
 2221 Gatggaga

Figure 2D. The cDNA (SEQ ID NO: 8) and amino acid sequence (SEQ ID NO: 9) of 158P1D7 v.4.
 The start methionine is underlined. The open reading frame extends from nucleic acid 23-1210 including the stop codon.

1 M K L W I H L F Y S S L L
 1 tcggatttcatcacatgacaacATGAAGCTGTGGATTCATCTCTTTTATTCATCTCTCCT
 14 A C I S L H S Q T P V L S S R G S C D S
 61 TGCCTGTATATCTTTACACTCCCAAACCTCCAGTGCTCTCATCCAGAGGCTCTTGTGATTC
 34 L C N C E E K D G T M L I N C E A K G I
 121 TCTTTGCAATTGTGAGGAAAAAGATGGCACAATGCTAATAAATTGTGAAGCAAAAGGTAT
 54 K M V S E I S V P P S R P F Q L S L L N
 181 CAAGATGGTATCTGAAATAAGTGTGCCACCATCAGACCTTTCCAACCTAAGCTTATTAAA
 74 N G L T M L H T N D F S G L T N A I S I
 241 TAACGGCTTGACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTATTTCAAT
 94 H L G F N N I A D I E I G A F N G L G L
 301 ACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCATTTAATGGCCTTGGCCT
 114 L K Q L H I N H N S L E I L K E D T F H
 361 CCTGAAACAACCTTCATATCAATCACAATTCTTTAGAAATTCTTAAAGAGGATACTTTCCA
 134 G L E N L E F L Q A D N N F I T V I E P
 421 TGGACTGGAAAACCTGGAATTCCTGCAAGCAGATAACAATTTTATCACAGTGATTGAACC
 154 S A F S K L N R L K V L I L N D N A I E
 481 AAGTGCCTTTAGCAAGCTCAACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGA
 174 S L P P N I F R F V P L T H L D L R G N
 541 GAGTCTTCCTCCAAACATCTTCCGATTGTTCCTTTAACCCATCTAGATCTTCGTGGAAA
 194 Q L Q T L P Y V G F L E H I G R I L D L
 601 TCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATTGGCCGAATATTGGATCT
 214 Q L E D N K W A C N C D L L Q L K T W L
 661 TCAGTTGGAGGACAACAAATGGGCCTGCAATTGTGACTTATTGCAGTTAAAACTTGTT
 234 E N M P P Q S I I G D V V C N S P P F F
 721 GGAGAACATGCCTCCACAGTCTATAATTGGTGATGTTGTCTGCAACAGCCCTCCATTTTT
 254 K G S I L S R L K K E S I C P T P P V Y

781 TAAAGGAAGTATACTCAGTAGACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTA
 274 E E H E D P S G S L H L A A T S S I N D
 841 TGAAGAACATGAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAATAAATGA
 294 S R M S T K T T S I L K L P T K A P G L
 901 TAGTCGCATGTCAACTAAGACCACGTCCATTCTAAACTACCCACCAAAGCACCAGGTTT
 314 I P Y I T K P S T Q L P G P Y C P I P C
 961 GATACCTTATATTACAAAGCCATCCACTCAACTTCCAGGACCTTACTGCCCTATTCCTTG
 334 N C K V L S P S G L L I H C Q E R N I E
 1021 TAACTGCAAAGTCCTATCCCCATCAGGACTTCTAATACATTGTCAGGAGCGCAACATTGA
 354 S L S D L R P P P Q N P R K L I L A G N
 1081 AAGCTTATCAGATCTGAGACCTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAA
 374 I I H S L M K S I L W S K A S G R G R R
 1141 TATTATTACAGTTTAATGAAGTCCATCCTTTGGTCCAAAGCATCTGGAAGAGGAAGAAG
 394 E E *
 1201 AGAGGAATGAgaaagaaggaagtgatgcaaaacatctccaaagaagtcttttggaacagg
 1261 aaaatcattcaccactcacaggggtcaaatatgaaatacaaaaccacgaaccaatcaacag
 1321 aatTTTTatccttccaagatgccagctcattgtacagaaacattttagaaaaagaaaggg
 1381 aacttcagcaactgggaatcacagaatacctaaggaaaaacattgctcagctccagcctg
 1441 atatggaggcacattatcctggagcccacgaagagctgaagttaatggaaacattaatgt
 1501 actcacgtccaaggaaggtattagtggaacagacaaaaaatgagtattttgaacttaaag
 1561 ctaatttacatgctgaacctgactatttagaagtcctggagcagcaaacatagatggaga

Figure 2E. The cDNA (SEQ ID NO: 10) and amino acid sequence (SEQ ID NO: 11) of 158P1D7 v.5.
 The start methionine is underlined. The open reading frame extends from nucleic acid 480-3005 including the stop codon.

1 gcgtcgacaacaagaaataactagaaaaggaggaaggagaaacattgctgcagcttggtatct
 61 acaacctaaagaaagcaagagtgatcaatctcagctctgttaaacatcttgtttacttact
 121 gcattcagcagcttgcaaatggttaactatatgcaaaaaagtcagcatagctgtgaagta
 181 tgccgtgaatttttaattgagggaaaaaggacaattgcttcaggatgctctagtatgcact
 241 ctgcttgaaatattttcaatgaaatgctcagttatctatctttgaccagaggttttaact
 301 ttatgaagctatgggacttgacaaaaagtgatatttgagaagaaagtacgcagtggttgg
 361 tgttttcttttttttaataaaggaattgaattactttgaacacctcttcagctgtgcat
 1 M
 421 tacagataacgtcaggaagagtctctgctttacagaatcggatttcatcacatgacaacA
 2 K L W I H L F Y S S L L A C I S L H S Q
 481 TGAAGCTGTGGATTCATCTCTTTTATTCATCTCTCCTTGCCTGTATATCTTTACACTCCC
 22 T P V L S S R G S C D S L C N C E E K D
 541 AAATCCAGTGCTCTCATCCAGAGGCTCTTGATTCCTTTGCAATTGTGAGGAAAAAG
 42 G T M L I N C E A K G I K M V S E I S V
 601 ATGGCACAATGCTAATAAATTGTGAAGCAAAAGGTATCAAGATGGTATCTGAAATAAGTG
 62 P P S R P F Q L S L L N N G L T M L H T

661 TGCCACCATCACGACCTTTCCAACCTAAGCTTATTAAATAACGGCTTGACGATGCTTCACA
82 N D F S G L T N A I S I H L G F N N I A
721 CAAATGACTTTTCTGGGCTTACCAATGCTATTTCAATACACCTTGGATTTAACAATATTG
102 D I E I G A F N G L G L L K Q L H I N H
781 CAGATATTGAGATAGGTGCATTTAATGGCCTTGGCCTCCTGAAACAACCTTCATATCAATC
122 N S L E I L K E D T F H G L E N L E F L
841 ACAATTCTTTAGAAATTCTTAAAGAGGATACTTTCCATGGACTGGAAAACCTGGAATTCC
142 Q A D N N F I T V I E P S A F S K L N R
901 TGCAAGCAGATAACAATTTTATCACAGTGATTGAACCAAGTGCCTTTAGCAAGCTCAACA
162 L K V L I L N D N A I E S L P P N I F R
961 GACTCAAAGTGTTAATTTTAAATGACAATGCTATTGAGAGTCTTCCTCCAAACATCTTCC
182 F V P L T H L D L R G N Q L Q T L P Y V
1021 GATTTGTTCTTTAACCCTAGATCTTCGTGGAAATCAATTACAAACATTGCCTTATG
202 G F L E H I G R I L D L Q L E D N K W A
1081 TTGGTTTTCTCGAACACATTGGCCGAATATTGGATCTTCAGTTGGAGGACAACAATGGG
222 C N C D L L Q L K T W L E N M P P Q S I
1141 CCTGCAATTGTGACTTATTGCAGTTAAAACTTGGTTGGAGAACATGCCTCCACAGTCTA
242 I G D V V C N S P P F F K G S I L S R L
1201 TAATTGGTGATGTTGTCTGCAACAGCCCTCCATTTTTTAAAGGAAGTATACTCAGTAGAC
262 K K E S I C P T P P V Y E E H E D P S G
1261 TAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTATGAAGAACATGAGGATCCTTCAG
282 S L H L A A T S S I N D S R M S T K T T
1321 GATCATTACATCTGGCAGCAACATCTTCAATAAATGATAGTCGCATGTCAACTAAGACCA
302 S I L K L P T K A P G L I P Y I T K P S
1381 CGTCCATTCTAAAACTACCCACCAAAGCACCAGGTTTGATACCTTATATTACAAAGCCAT
322 T Q L P G P Y C P I P C N C K V L S P S
1441 CCACTCAACTTCCAGGACCTTACTGCCCTATTCCTTGTAAGTCAAAGTCCTATCCCCAT
342 G L L I H C Q E R N I E S L S D L R P P
1501 CAGGACTTCTAATACATTGTGAGGAGCGCAACATTGAAAGCTTATCAGATCTGAGACCTC
362 P Q N P R K L I L A G N I I H S L M K S
1561 CTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAATATTATTACAGTTTAATGAAGT
382 D L V E Y F T L E M L H L G N N R I E V
1621 CTGATCTAGTGGAATATTTCACTTTGGAAATGCTTCACTTGGGAAACAATCGTATTGAAG
402 L E E G S F M N L T R L Q K L Y L N G N
1681 TTCTTGAAGAAGGATCGTTTATGAACCTAACGAGATTACAAAACCTCTATCTAAATGGTA
422 H L T K L S K G M F L G L H N L E Y L Y
1741 ACCACCTGACCAAATTAAGTAAAGGCATGTTCTTGGTCTCCATAATCTTGAATACTTAT
442 L E Y N A I K E I L P G T F N P M P K L
1801 ATCTTGAATACAATGCCATTAAGGAAATACTGCCAGGAACCTTTAATCCAATGCCTAAAC
462 K V L Y L N N N L L Q V L P P H I F S G
1861 TTAAAGTCCTGTATTTAAATAACAACCTCCTCCAAGTTTTACCACCACATATTTTTTCAG
482 V P L T K V N L K T N Q F T H L P V S N

1921 GGGTTCCTCTAACTAAGGTAAATCTTAAACAAACCAGTTTACCCATCTACCTGTAAGTA
502 I L D D L D L L T Q I D L E D N P W D C
1981 ATATTTTGGATGATCTTGATTTACTAACCCAGATTGACCTTGAGGATAACCCCTGGGACT
522 S C D L V G L Q Q W I Q K L S K N T V T
2041 GCTCCTGTGACCTGGTTGGACTGCAGCAATGGATACAAAAGTTAAGCAAGAACACAGTGA
542 D D I L C T S P G H L D K K E L K A L N
2101 CAGATGACATCCTCTGCACTTCCCCGGGCATCTCGACAAAAAGGAATTGAAAGCCCTAA
562 S E I L C P G L V N N P S M P T Q T S Y
2161 ATAGTGAAATTCTCTGTCCAGGTTTAGTAAATAACCCATCCATGCCAACACAGACTAGTT
582 L M V T T P A T T T N T A D T I L R S L
2221 ACCTTATGGTCACCACTCCTGCAACAACAACAATACGGCTGATACTATTTTACGATCTC
602 T D A V P L S V L I L G L L I M F I T I
2281 TTACGGACGCTGTGCCACTGTCTGTTCTAATATTGGGACTTCTGATTATGTTTCATCACTA
622 V F C A A G I V V L V L H R R R R Y K K
2341 TTGTTTTCTGTGCTGCAGGGATAGTGGTTCTTGTTCCTCACCGCAGGAGAAGATACAAA
642 K Q V D E Q M R D N S P V H L Q Y S M Y
2401 AGAAACAAGTAGATGAGCAAATGAGAGACAACAGTCCTGTGCATCTTCAGTACAGCATGT
662 G H K T T H H T T E R P S A S L Y E Q H
2461 ATGGCCATAAAACCACTCATCACTACTGAAAGACCCTCTGCCTCACTCTATGAACAGC
682 M V S P M V H V Y R S P S F G P K H L E
2521 ACATGGTGAGCCCCATGGTTCATGTCTATAGAAGTCCATCCTTTGGTCCAAAGCATCTGG
702 E E E E R N E K E G S D A K H L Q R S L
2581 AAGAGGAAGAAGAGAGGAATGAGAAAGAAGGAAGTGATGCAAAACATCTCCAAAGAAGTC
722 L E Q E N H S P L T G S N M K Y K T T N
2641 TTTTGGAACAGGAAAATCATTCACTCACTACAGGGTCAAATATGAAATACAAAACCACGA
742 Q S T E F L S F Q D A S S L Y R N I L E
2701 ACCAATCAACAGAAATTTTATCCTTCCAAGATGCCAGCTCATTTGTACAGAAACATTTTAG
762 K E R E L Q Q L G I T E Y L R K N I A Q
2761 AAAAAGAAAGGGAACCTTCAGCAACTGGGAATCACAGAATACCTAAGGAAAAACATTGCTC
782 L Q P D M E A H Y P G A H E E L K L M E
2821 AGCTCCAGCCTGATATGGAGGCACATTATCCTGGAGCCCACGAAGAGCTGAAGTTAATGG
802 T L M Y S R P R K V L V E Q T K N E Y F
2881 AAACATTAATGTACTCACGTCCAAGGAAGGTATTAGTGGAACAGACAAAAAATGAGTATT
822 E L K A N L H A E P D Y L E V L E Q Q T
2941 TTGAACCTAAAGCTAATTTACATGCTGAACCTGACTATTTAGAAGTCCTGGAGCAGCAAA
842 *
3001 CATAGatggagagttgagggctttcgccagaaatgctgtgattctgttattaagtccata
3061 ccttgtaaataagtgccttacgtgagtggtgtcatcaatcagaacctaacgacagagtaaa
3121 ctatggggaaaaaaaaaagaagacgaaacagaaactcagggatcactgggagaagccatgg
3181 cataatcttcaggcaatttagtctgtcccaaataaacatacatccttggcatgtaaatca
3241 tcaagggtaatagtaatatcatatacctgaaacgtgtctcataggagtcctctctgcac

Figure 2F. The cDNA (SEQ ID NO: 12) and amino acid sequence (SEQ ID NO: 13) of 158P1D7 v.6.
The start methionine is underlined. The open reading frame extends from nucleic acid 23-1612 including the stop codon.

```
1           M K L W I H L F Y S S L L
1  tcggatttcatcacatgacaacATGAAGCTGTGGATTCATCTCTTTTATTCATCTCTCCT
14  A C I S L H S Q T P V L S S R G S C D S
61  TGCCTGTATATCTTTACACTCCCAAACCTCCAGTGCTCTCATCCAGAGGCTCTTGTGATTC
34  L C N C E E K D G T M L I N C E A K G I
121 TCTTTGCAATTGTGAGGAAAAAGATGGCACAATGCTAATAAATTGTGAAGCAAAAGGTAT
54  K M V S E I S V P P S R P F Q L S L L N
181 CAAGATGGTATCTGAAATAAGTGTGCCACCATCACGACCTTTCCAACCTAAGCTTATTAAA
74  N G L T M L H T N D F S G L T N A I S I
241 TAACGGCTTGACGATGCTTCACACAAATGACTTTTCTGGGCTTACCAATGCTATTTC AAT
94  H L G F N N I A D I E I G A F N G L G L
301 ACACCTTGGATTTAACAATATTGCAGATATTGAGATAGGTGCATTTAATGGCCTTGGCCT
114 L K Q L H I N H N S L E I L K E D T F H
361 CCTGAAACAACCTTCATATCAATCACAATTCTTTAGAAATTCTTAAAGAGGATACTTTCCA
134 G L E N L E F L Q A D N N F I T V I E P
421 TGGACTGGAAAACCTGGAATTCCTGCAAGCAGATAACAATTTTATCACAGTGATTGAACC
154 S A F S K L N R L K V L I L N D N A I E
481 AAGTGCCTTTAGCAAGCTCAACAGACTCAAAGTGTTAATTTTAAATGACAATGCTATTGA
174 S L P P N I F R F V P L T H L D L R G N
541 GAGTCTTCCTCCAAACATCTTCCGATTTGTTCCCTTAAACCCATCTAGATCTTCGTGGAAA
194 Q L Q T L P Y V G F L E H I G R I L D L
601 TCAATTACAAACATTGCCTTATGTTGGTTTTCTCGAACACATTGGCCGAATATTGGATCT
214 Q L E D N K W A C N C D L L Q L K T W L
661 TCAGTTGGAGGACAACAAATGGGCCTGCAATTGTGACTTATTGCAGTTAAAAACTTGGTT
234 E N M P P Q S I I G D V V C N S P P F F
721 GGAGAACATGCCTCCACAGTCTATAATTGGTGATGTTGTCTGCAACAGCCCTCCATTTT
254 K G S I L S R L K K E S I C P T P P V Y
781 TAAAGGAAGTATACTCAGTAGACTAAAGAAGGAATCTATTTGCCCTACTCCACCAGTGTA
274 E E H E D P S G S L H L A A T S S I N D
841 TGAAGAACATGAGGATCCTTCAGGATCATTACATCTGGCAGCAACATCTTCAATAAATGA
294 S R M S T K T T S I L K L P T K A P G L
901 TAGTCGCATGTCAACTAAGACCACGTCCATTCTAAAACTACCCACCAAAGCACCAGGTTT
314 I P Y I T K P S T Q L P G P Y C P I P C
961 GATACCTTATATTACAAAGCCATCCACTCAACTTCCAGGACCTTACTGCCCTATTCCTTG
334 N C K V L S P S G L L I H C Q E R N I E
1021 TAACTGCAAAGTCCTATCCCCATCAGGACTTCTAATACATTGTTCAGGAGCGCAACATTGA
354 S L S D L R P P P Q N P R K L I L A G N
1081 AAGCTTATCAGATCTGAGACCTCCTCCGCAAAATCCTAGAAAGCTCATTCTAGCGGGAAA
```

374 I I H S L M N P S F G P K H L E E E E E
1141 TATTATTCACAGTTTAATGAATCCATCCTTTGGTCCAAAGCATCTGGAAGAGGAAGAAGA
394 R N E K E G S D A K H L Q R S L L E Q E
1201 GAGGAATGAGAAAGAAGGAAGTGATGCAAAACATCTCCAAAGAAGTCTTTTGGAACAGGA
414 N H S P L T G S N M K Y K T T N Q S T E
1261 AAATCATTCACCACTCACAGGGTCAAATATGAAATACAAAACCACGAACCAATCAACAGA
434 F L S F Q D A S S L Y R N I L E K E R E
1321 ATTTTATCCTTCCAAGATGCCAGCTCATTGTACAGAAACATTTTAGAAAAAGAAAGGGA
454 L Q Q L G I T E Y L R K N I A Q L Q P D
1381 ACTTCAGCAACTGGGAATCACAGAATACCTAAGGAAAAACATTGCTCAGCTCCAGCCTGA
474 M E A H Y P G A H E E L K L M E T L M Y
1441 TATGGAGGCACATTATCCTGGAGCCCACGAAGAGCTGAAGTTAATGGAAACATTAATGTA
494 S R P R K V L V E Q T K N E Y F E L K A
1501 CTCACGTCCAAGGAAGGTATTAGTGGAACAGACAAAAAATGAGTATTTTGAACCTTAAAGC
514 N L H A E P D Y L E V L E Q Q T *
1561 TAATTTACATGCTGAACCTGACTATTTAGAAGTCCTGGAGCAGCAAACATAGatggaga

Figure 3:

Figure 3A. Amino acid sequence 158P1D7 v.1 (SEQ ID NO: 14). The 158P1D7 v.1 protein has 841 amino acids.

```

1  MKLWIHLFYS SLLACISLHS QTPVLSSRGS CDSLNCNEEK DGTMLINCEA KGIKMOVSEIS
61  VPPSRPFQLS LLNNGLTMLH TNDFSGLTNA ISIH LGFN NI ADIEIGAFNG LGLLKQLHIN
121 HNSLEILKED TFHGLENLEF LQADNNFITV IEPSAFSKLN RLKVLILNDN AIESLPPNIF
181 RFVPLTHLDL RGNQLQTLPY VGFLEHIGRI LDLQLEDNKW ACNCDLLQLK TWLENMPPQS
241 IIGDVVCNSP PFFKGSILSR LKKESICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT
301 TSILKLPTKA PGLIPYITKP STQLPGPYCP IPCNCKVLSP SGLLIHCQER NIESLSDLRP
361 PPQNPRKLIL AGNIIHSLMK SDLVEYFTLE MLHLGNNRIE VLEEGSFMNL TRLOKLYLNG
421 NHLTKLSKGM FLGLHNLEYL YLEYNAIKEI LPGTFNMPK LKVLYLNNNL LQVLPPHIFS
481 GVPLTKVNLK TNQFTHLPVS NILDDLDLLT QIDLEDNPWD CSCDLVGLQQ WIOKLSKNTV
541 TDDILCTSPG HLDKKELKAL NSEILCPGLV NNPSMPTQTS YLMVTTPATT TNTADTILRS
601 LTDAVPLSVL ILGLLIMFIT IVFCAAGIVV LVLHRRRRYK KKQVDEQMRD NSPVHLQYSM
661 YGHKTTHHTT ERPSASLYEQ H MVSPMVHVY RSPSFGPKHL EEEEEERNEKE GSDAKHLQRS
721 LLEQENHSPL TGSNMKYKTT NQSTEFLSFQ DASSLYRNIL EKERELQQLG ITEYLRKNIA
781 QLQPDMEAHY PGAHEELKLM ETLMYSRPRK VLVEQTKNEY FELKANLHAE PDYLEVLEQQ
841 T

```

Figure 3B. Amino acid sequence 158P1D7 v.3 (SEQ ID NO: 15). The 158P1D7 v.3 protein has 732 amino acids.

```

1  MKLWIHLFYS SLLACISLHS QTPVLSSRGS CDSLNCNEEK DGTMLINCEA KGIKMOVSEIS
61  VPPSRPFQLS LLNNGLTMLH TNDFSGLTNA ISIH LGFN NI ADIEIGAFNG LGLLKQLHIN
121 HNSLEILKED TFHGLENLEF LQADNNFITV IEPSAFSKLN RLKVLILNDN AIESLPPNIF
181 RFVPLTHLDL RGNQLQTLPY VGFLEHIGRI LDLQLEDNKW ACNCDLLQLK TWLENMPPQS
241 IIGDVVCNSP PFFKGSILSR LKKESICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT
301 TSILKLPTKA PGLIPYITKP STQLPGPYCP IPCNCKVLSP SGLLIHCQER NIESLSDLRP
361 PPQNPRKLIL AGNIIHSLMK SDLVEYFTLE MLHLGNNRIE VLEEGSFMNL TRLOKLYLNG
421 NHLTKLSKGM FLGLHNLEYL YLEYNAIKEI LPGTFNMPK LKVLYLNNNL LQVLPPHIFS
481 GVPLTKVNLK TNQFTHLPVS NILDDLDLLT QIDLEDNPWD CSCDLVGLQQ WIOKLSKNTV
541 TDDILCTSPG HLDKKELKAL NSEILCPGLV NNPSMPTQTS YLMVTTPATT TNTADTILRS
601 LTDAVPLSVL ILGLLIMFIT IVFCAAGIVV LVLHRRRRYK KKQVDEQMRD NSPVHLQYSM
661 YGHKTTHHTT ERPSASLYEQ H MGAHEELKL METLMYSRPR KVLVEQTKNE YFELKANLHA
721 EPDYLEVLEQ QT

```

Figure 3C. Amino acid sequence 158P1D7 v.4 (SEQ ID NO: 16). The 158P1D7 v.4 protein has 395 amino acids.

```

1  MKLWIHLFYS SLLACISLHS QTPVLSSRGS CDSLNCNEEK DGTMLINCEA KGIKMOVSEIS
61  VPPSRPFQLS LLNNGLTMLH TNDFSGLTNA ISIH LGFN NI ADIEIGAFNG LGLLKQLHIN
121 HNSLEILKED TFHGLENLEF LQADNNFITV IEPSAFSKLN RLKVLILNDN AIESLPPNIF
181 RFVPLTHLDL RGNQLQTLPY VGFLEHIGRI LDLQLEDNKW ACNCDLLQLK TWLENMPPQS
241 IIGDVVCNSP PFFKGSILSR LKKESICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT

```

301 TSILKLPTKA PGLIPYITKP STQLPGPYCP IPCNCKVLSP SGLLIHCQER NIESLSDLRP
361 PPQNPRKLIL AGNIIHSLMK SILWSKASGR

Figure 3D. Amino acid sequence 158P1D7 v.6 (SEQ ID NO: 17). The 158P1D7 v.6 protein has 529 amino acids.

1 MKLWIHLFYS SLLACISLHS QTPVLSSRGS CDSLNCNEEK DGTMLINCEA KGIKMOVSEIS
61 VPPSRPFQLS LLNNGLTMLH TNDFSGLTNA ISIH LGFN NI ADIEIGAFNG LGLLKQLHIN
121 HNSLEILKED TFHGLENLEF LQADNNFITV IEPSAFSKLN RLKVLILNDN AIESLPPNIF
181 RFVPLTHLDL RGNQLQTL PY VGFLEHIGRI LDLQLEDNKW ACNCDLLQLK TWLENMPPQS
241 IIGDVVCNSP PFFKGSILSR LKESICPTP PVYEEHEDPS GSLHLAATSS INDSRMSTKT
301 TSILKLPTKA PGLIPYITKP STQLPGPYCP IPCNCKVLSP SGLLIHCQER NIESLSDLRP
361 PPQNPRKLIL AGNIIHSLMN PSFGPKHLEE EEERNEKEGS DAKHLQRSLL EQENHSPLTG
421 SNMKYKTTNQ STEFLSFQDA SSLYRNILEK ERELQQLGIT EYLRKNIAQL QPDMEAHYPG
481 AHEELKLMET LMYSRPRKVL VEQTKNEYFE LKANLHAEPD YLEVLEQQT

Figure 4: 158P1D7 v.1 amino acid (SEQ ID NO: 18) BLAST homology to hypothetical protein FLJ22774 (SEQ ID NO: 19).

Identities = 798/798 (100%)

Query: 44 MLINCEAKGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIH LGFN NIADI 103
MLINCEAKGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIH LGFN NIADI
Sbjct: 1 MLINCEAKGIKMVSEISVPPSRPFQLSLLNNGLTMLHTNDFSGLTNAISIH LGFN NIADI 60

Query: 104 EIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITVIEPSAFSKLNRLK 163
EIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITVIEPSAFSKLNRLK
Sbjct: 61 EIGAFNGLGLLKQLHINHNSLEILKEDTFHGLENLEFLQADNNFITVIEPSAFSKLNRLK 120

Query: 164 VLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPYVGFLEHIGRIIDLQLEDNKWACN 223
VLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPYVGFLEHIGRIIDLQLEDNKWACN
Sbjct: 121 VLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPYVGFLEHIGRIIDLQLEDNKWACN 180

Query: 224 CDLLQLKTWLENMPPQSIIGDVVCNSPPFFKGSILSRLKKESICPTPPVYEEHEDPSGSL 283
CDLLQLKTWLENMPPQSIIGDVVCNSPPFFKGSILSRLKKESICPTPPVYEEHEDPSGSL
Sbjct: 181 CDLLQLKTWLENMPPQSIIGDVVCNSPPFFKGSILSRLKKESICPTPPVYEEHEDPSGSL 240

Query: 284 HLAATSSINDSRMSTKTTSSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGL 343
HLAATSSINDSRMSTKTTSSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGL
Sbjct: 241 HLAATSSINDSRMSTKTTSSILKLPTKAPGLIPYITKPSTQLPGPYCPIPCNCKVLSPSGL 300

Query: 344 LIHCQERNIESLSDLRPPPPQNPRKLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIEVLE 403
LIHCQERNIESLSDLRPPPPQNPRKLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIEVLE
Sbjct: 301 LIHCQERNIESLSDLRPPPPQNPRKLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIEVLE 360

Query: 404 EGSFMNLTRLQKLYLNGNHLTKLSKGMFLGLHNLEYLYLEYNAIKEILPGTFNPMPLKV 463
EGSFMNLTRLQKLYLNGNHLTKLSKGMFLGLHNLEYLYLEYNAIKEILPGTFNPMPLKV
Sbjct: 361 EGSFMNLTRLQKLYLNGNHLTKLSKGMFLGLHNLEYLYLEYNAIKEILPGTFNPMPLKV 420

Query: 464 LYLNNNLLQVLPPHIFSGVPLTKVNLKTNQFTHLPVSNILDDLDLLTQIDLEDNPWDCSC 523
LYLNNNLLQVLPPHIFSGVPLTKVNLKTNQFTHLPVSNILDDLDLLTQIDLEDNPWDCSC
Sbjct: 421 LYLNNNLLQVLPPHIFSGVPLTKVNLKTNQFTHLPVSNILDDLDLLTQIDLEDNPWDCSC 480

Query: 524 DLVGLQQWIQKLSKNTVTDDILCTSPGHLDDKKELKALNSEILCPGLVNNPSMPTQTSYLM 583
DLVGLQQWIQKLSKNTVTDDILCTSPGHLDDKKELKALNSEILCPGLVNNPSMPTQTSYLM
Sbjct: 481 DLVGLQQWIQKLSKNTVTDDILCTSPGHLDDKKELKALNSEILCPGLVNNPSMPTQTSYLM 540

Query: 584 VTTPATTTNTADTILRSLTDAVPLSVLILGLLIMFITIVFCAAGIVVLVHRRRRYKKKQ 643
VTTPATTTNTADTILRSLTDAVPLSVLILGLLIMFITIVFCAAGIVVLVHRRRRYKKKQ
Sbjct: 541 VTTPATTTNTADTILRSLTDAVPLSVLILGLLIMFITIVFCAAGIVVLVHRRRRYKKKQ 600

Query: 644 VDEQMRDNSPVHLQYSMYGHKTTHHTTERPSASLYEQHMVSPMVHVYRSPSFGPKHLEEE 703
VDEQMRDNSPVHLQYSMYGHKTTHHTTERPSASLYEQHMVSPMVHVYRSPSFGPKHLEEE
Sbjct: 601 VDEQMRDNSPVHLQYSMYGHKTTHHTTERPSASLYEQHMVSPMVHVYRSPSFGPKHLEEE 660

Query: 704 EERNEKEGSDAKHLQRSLLQEENHSPLTGSNMKYKTTNQSTEFLSFQDASSLYRNILEKE 763
EERNEKEGSDAKHLQRSLLQEENHSPLTGSNMKYKTTNQSTEFLSFQDASSLYRNILEKE
Sbjct: 661 EERNEKEGSDAKHLQRSLLQEENHSPLTGSNMKYKTTNQSTEFLSFQDASSLYRNILEKE 720

Query: 764 RELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLMETLMYSRPRKVLVEQTKNEYFEL 823
RELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLMETLMYSRPRKVLVEQTKNEYFEL
Sbjct: 721 RELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLMETLMYSRPRKVLVEQTKNEYFEL 780

Query: 824 KANLHAEPDYLEVLEQQT 841
KANLHAEPDYLEVLEQQT
Sbjct: 781 KANLHAEPDYLEVLEQQT 798

Figure 5:

Figure 5A: Alignment of 158P1D7 v.1 (SEQ ID NO: 20) with human FLJ22774, CLONE KAIA1575.[Homo sapiens] (SEQ ID NO: 21)

Identities = 405/415 (97%), Positives = 405/415 (97%)

158P1D7: 44 MLINCEAKGIKMVSEISVPPSRPFQSLNNGLTMLHTNDFSGLTNAISIHGFGNNIADI 103
 MLINCEAKGIKMVSEISVPPSRPFQSLNNGLTMLHTNDFSGLTNAISIHGFGNNIADI
 Sbjct: 1 MLINCEAKGIKMVSEISVPPSRPFQSLNNGLTMLHTNDFSGLTNAISIHGFGNNIADI 60

158P1D7:104 EIGAENGLGLLKQLHINHNLSLEILKEDTFHGLNLEFLQADNNFITVIEPSAFSKLNRLK 163
 EIGAENGLGLLKQLHINHNLSLEILKEDTFHGLNLEFLQADNNFITVIEPSAFSKLNRLK
 Sbjct: 61 EIGAENGLGLLKQLHINHNLSLEILKEDTFHGLNLEFLQADNNFITVIEPSAFSKLNRLK 120

158P1D7:164 VLIINDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPIYVGFEHIGRILDQLQEDNKWACN 223
 VLIINDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPIYVGFEHIGRILDQLQEDNKWACN
 Sbjct: 121 VLIINDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPIYVGFEHIGRILDQLQEDNKWACN 180

158P1D7:224 CDLLQLKTWLENMPPQSIIGDVVCNSPPFFKGSILSRKKESICPTPPVYEEHEDPSGSL 283
 CDLLQLKTWLENMPPQSIIGDVVCNSPPFFKGSILSRKKESICPTPPVYEEHEDPSGSL
 Sbjct: 181 CDLLQLKTWLENMPPQSIIGDVVCNSPPFFKGSILSRKKESICPTPPVYEEHEDPSGSL 240

158P1D7:284 HLAATSSINDSRMSTKTTSSILKLPKAPGLIPYITKPSQPLPGPYCPIPCNCKVLSPSGL 343
 HLAATSSINDSRMSTKTTSSILKLPKAPGLIPYITKPSQPLPGPYCPIPCNCKVLSPSGL
 Sbjct: 241 HLAATSSINDSRMSTKTTSSILKLPKAPGLIPYITKPSQPLPGPYCPIPCNCKVLSPSGL 300

158P1D7:344 LIHCQERNIESLSDLRPPQNPRLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIEVL 403
 LIHCQERNIESLSDLRPPQNPRLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIEVL
 Sbjct: 301 LIHCQERNIESLSDLRPPQNPRLILAGNIIHSLMKSDLVEYFTLEMLHLGNNRIEVL 360

158P1D7:404 EGSEFMNLTRELQKLYLNGNHLTKLSKGMFLGLHXXXXXXXXXXXXAIKEILPGTFNPM 458
 EGSEFMNLTRELQKLYLNGNHLTKLSKGMFLGLH AIKEILPGTFNPM
 Sbjct: 361 EGSEFMNLTRELQKLYLNGNHLTKLSKGMFLGLHNLLEYLYLEYNAIKEILPGTFNPM 415

Figure 5b: Alignment of 158P1D7 v.1 protein (SEQ ID NO: 22) with a human protein similar to IGFALS (SEQ ID NO: 23)

Identities = 316/864 (36%), Positives = 459/864 (52%)

158P1D7:1 MKLWIHLFYSSLLACISLHSQTPVLSSRGSCDSLNCCEKDGTMLINCEAKGIKMVSEIS 60
 M LW+ L S+L++ + S V ++C+C + + +NCE + +++
 Sbjct: 17 MFLWFLILSALISSTNADSDISV-----EICNVSCSVSVENVLYVNCEKVSRYRPNQLK 71

158P1D7:61 VPPSRPFQSLNNGLTMLHTNDFSGLTNAISIHGFGNNIADI EIGAENGLGLLKQLHIN 120
 P S + L+ NN L +L+ N F ++A+S+HLG N + +IE GAF GL LKQLH+N
 Sbjct: 72 PPWSNFYHLNFQNNFLNLYPNTFLNFESHAVSLHLGNNKLQNI EGGAFLGLSALKQLHLN 131

158P1D7:121 HNSLEILKEDTFHGLNLEFLQADNNFITVIEPSAFSKLNRLKVLIINDNAIESLPPNIF 180
 +N L+IL+ DTF G+ENLE+LQAD N I IE AF+KL++LKVLIINDN I LP NIF
 Sbjct: 132 NNELKILRADTFGLIENLEYLQADYNLIK YIERGAFNKLHLKVLIINDNLISFLPDNIF 191

158P1D7:181 RFVPLTHLDLRGNQLQTLPIYVGFEHIGRILDQLQEDNKWACNCDLLQLKTWLENMPPQS 240
 RF LTHLD+RGN++Q LPY+G LEHIGR+++LQLEDN W C+CDLL LK WLENMP
 Sbjct: 192 RFASLTHLDIRGNRIQKLPYIGVLEHIGRVVELQLEDNPNWCSDLLPLKAWLENMPYNI 251

158P1D7:241 IIGDVVCNSPPFFKGSILSRKKESICP-----TPPVYEEHEDPSGSLHLAATS 289
 IG+ +C +P G +L K+ +CP PP E+ + + H TS
 Sbjct: 252 YIGEAICETPSDLYGRLLKETNKQELCPMGTGSDFDVRILPPSQLENGYTTPNGHTTQTS 311

158P1D7:290 SINDSRMSTKTTSSILKLPKAPGLI-----PYITKPSQPLPG-PYCPIPCNCNV- 337
 KTT+ P+K G++ I T++P CP PC CK
 Sbjct: 312 LHRLVTKPPKTTN----PSKISGIVAGKALSNNRNLISQIVSYQTRVPPLTPCPAPCFCKTH 367

158P1D7:338 LSPSGLLIHCQERNIESLSDLRPPQNPRLILAGNIIHSLMKSDLVEYFTLEMLHLGNN 397
 S GL ++CQE+NI+S+S+L P P N +KL + GN I + SD ++ L++LHLG+N
 Sbjct: 368 PSDLGLSVNCQEKNIQSMSELIPKPLNAKKLHVNGNSIKDQDVSDFTDFEGLDLLHLGNS 427

158P1D7:398 RIEVLEEGSFMNLTRELQKLYLNGNHLTKLSKGMFLGLHXXXXXXXXXXXXAIKEILPGTFNP 457
 +I V++ F NLT L++LYLNGN + +L +F GLH IKEI GTF+
 Sbjct: 428 QITVIKGDVFNLTNLRRLYLNGNQIERLYPEIFSGLHNLQYLYLEYNLIKEISAGTFDS 487

WO 2004/072263

PCT/US2004/003984

18/73

158P1D7:458 MXXXXXXXXXXXXXXXXXHIIFSGVPLTKVNLKTNQFTHLPVSNIXXXXXXXXXXXXXXN 517
M +IFSG PL ++NL+ N+F +LPVS + N
Sbjct: 488 MPNLQLLYLNNNLLKSLPVYIFSGAPLARLNLNNKFMYPVSGVLDQLQSLTQIDLEGN 547

158P1D7:518 PWDCSCDLVGLQQWIKLSKNTVTDDILCTSPGHLDKKEKALNSEILCPGLVNNPSMPT 577
PWDC+CDLV L+ W++KLS V ++ C +P ELK+L +EILCP L+N PS P
Sbjct: 548 PWDCTCDLVALKLWVEKLSDGIVVKELKCETPVQFANIELKSLKNEILCPKLLNKPSAP- 606

158P1D7:578 QTSYLMVXXXXXXXXXXXXXILRSLTDAVPLSVLILGLLIMFITIVFCAAGIVVLVLRHRR 637
+ I VPLS+LIL +L++ I VF A ++V VL R +
Sbjct: 607 ---FTSPAPAITFTTPLGPIRSPPGPVPLSILILSVLILTVFVAFCLLVFVLRNRK 663

158P1D7:638 RYKKKQVDEQMRDNSPVHLQYSMYGHKTTHHTTERPSASLYEQHVMVSPMVHVYRSPSFGP 697
+ K D + LQ + HK T + E + + +S + G
Sbjct: 664 KPTVKHEGLGNPDCGSMQLQLRKHDHK-----TNKKDGLSTEAFIPQTIEQMSKSHTCGL 718

158P1D7:698 KHLXXXXXXXXXXGSDAKHLQRSLLLEQENHSPLTGSNMKYKTTNQSTEFLSFQDASSLYR 757
K G K + R++ ++E + + T ++ E +D++ +
Sbjct: 719 KESETGFMFSDPPGQ--KVVMRNVADKEKDLLHVDTRKRLSTIDELDELFPSSRDSNVFIQ 776

158P1D7:758 NILEKERELQQLGITEYLRKNIAQLQPDMEAHYPGAHEELKLMETLMYSRPRKVLVEQTK 817
N LE ++E +G++ + E YP + K ++L+ K++VEQ K
Sbjct: 777 NFLESKKEYNSIGVSGF-----EIRYPEKQPDKKSCKSLIGGNHISKIVVEQRK 824

158P1D7:818 NEYFELKANLHAEPDYLEVLEQQT 841
+EYFELKA L + PDYL+VLE+QT
Sbjct: 825 SEYFELKAKLQSSPDYLQVLEEQT 848

Figure 6. Expression of 158P1D7 by RT-PCR

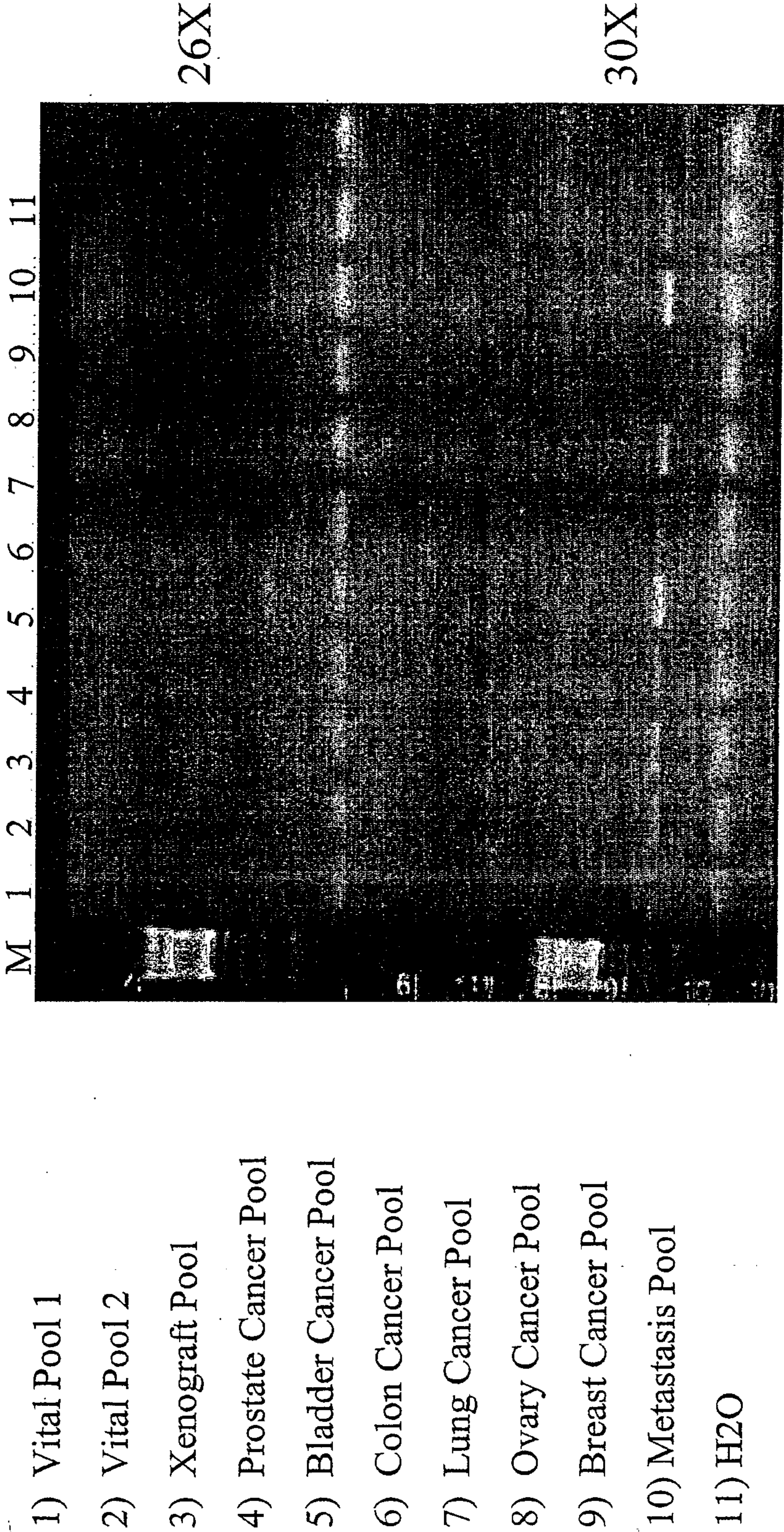


Figure 7. Expression of 158P1D7 in Normal Tissues

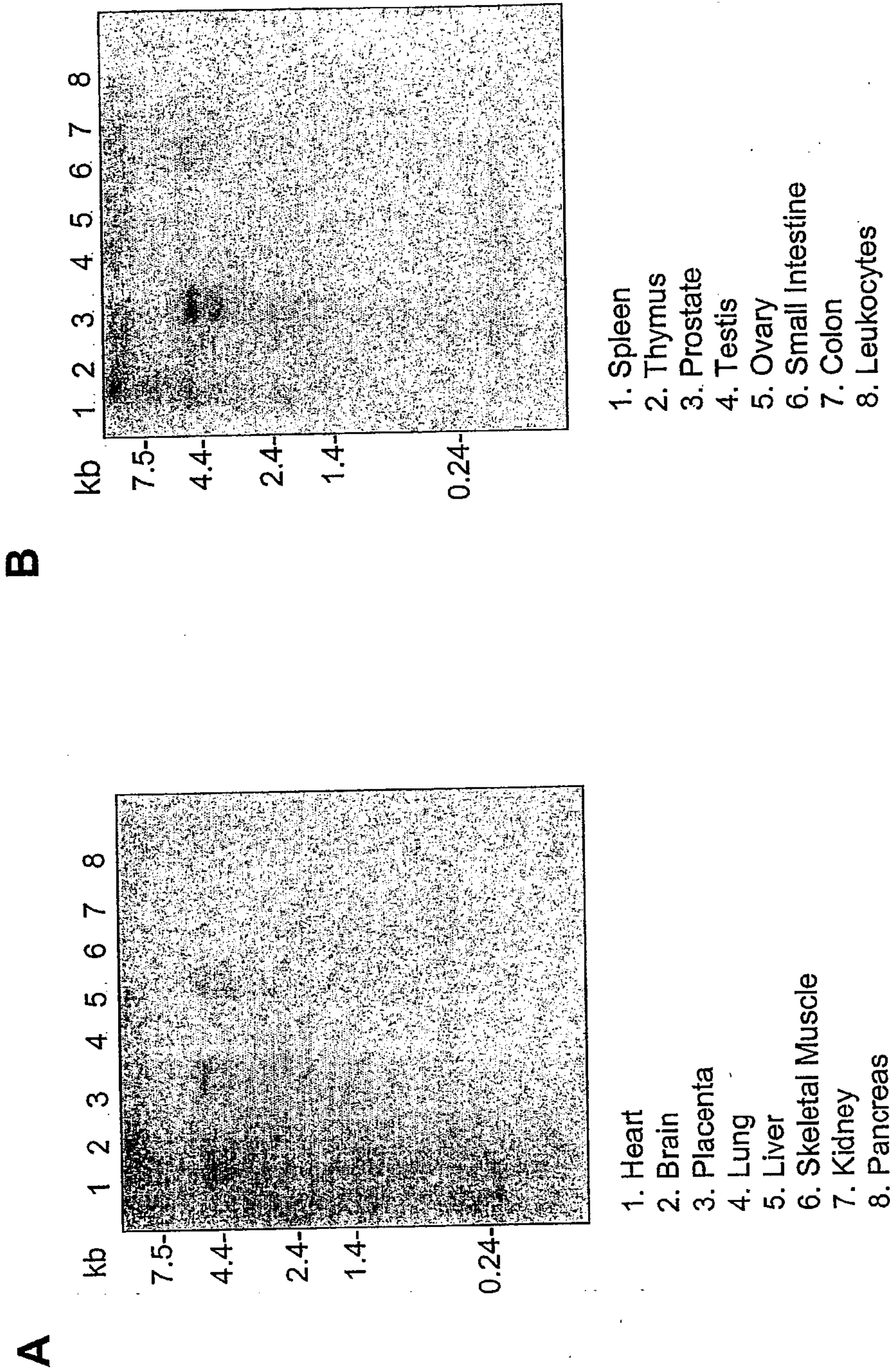
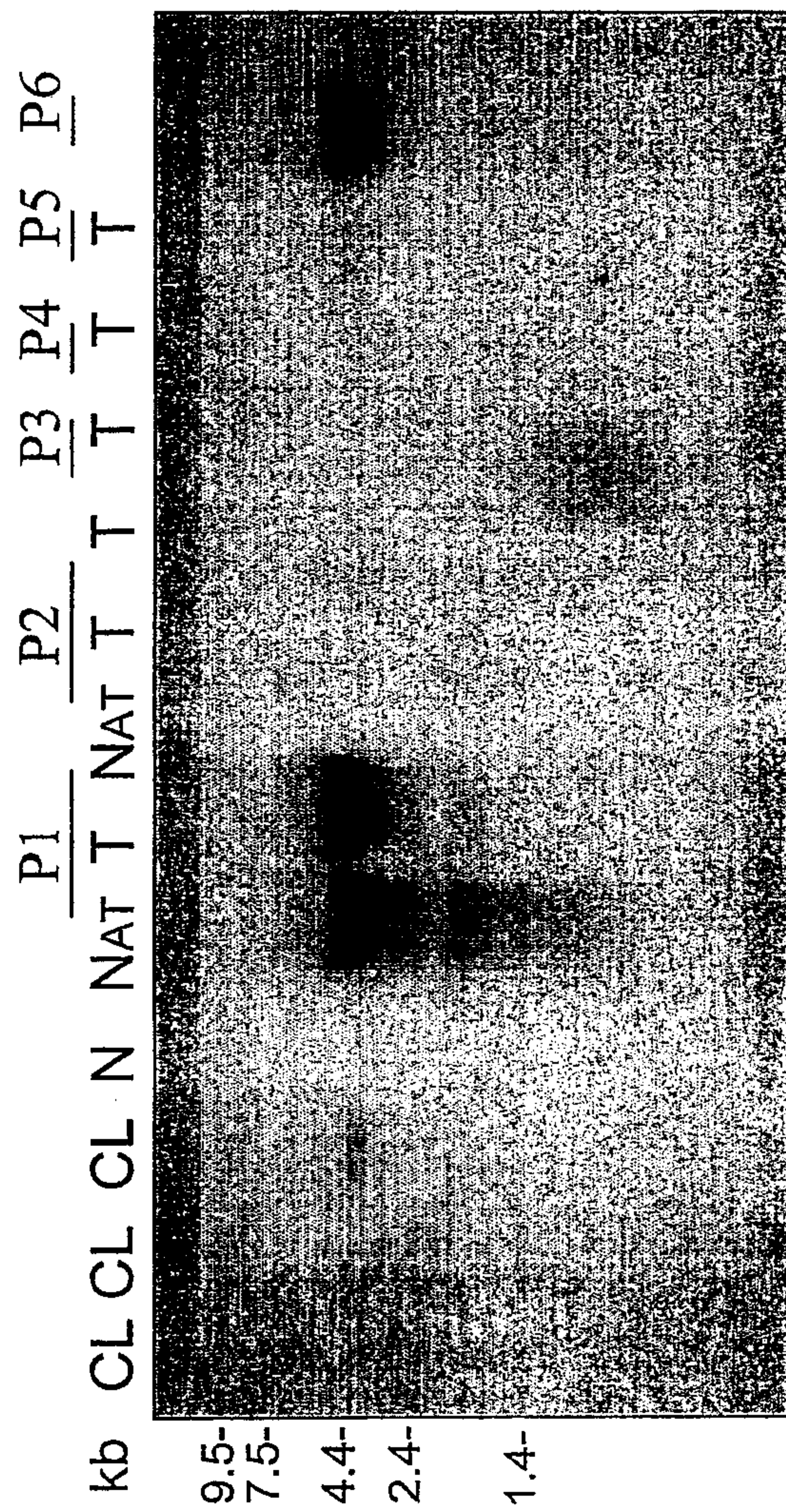


Figure 8A. Expression of 158P1D7 in Bladder Cancer Patient Specimens



- P1 - Transitional, grade 4
- P2 - Squamous inv.
- P3 - Transitional, grade 3
- P4 - Papillary non-inv, grade 1/3
- P5 - Papillary, grade 3/3
- P6 - Transitional, grade 3/2

CL = Cell lines (listed in order): UM-UC-3, J82, SCaBER
P = Patient
N = Normal Bladder
NAT = Normal adjacent tissue
T = Tumor

Figure 8B. Expression of 158P1D7 in Bladder Cancer Patient Specimens

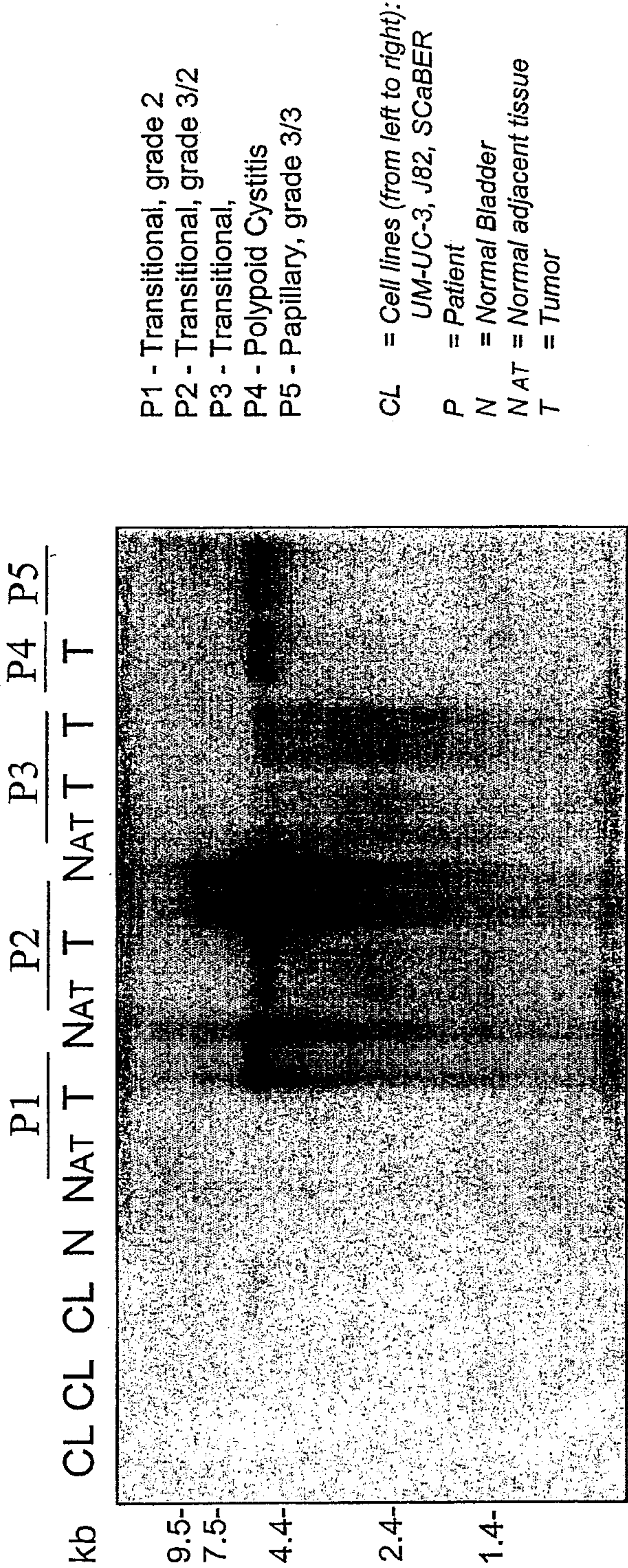


Figure 9. Expression of 158P1D7 in Lung Cancer Patient Specimens

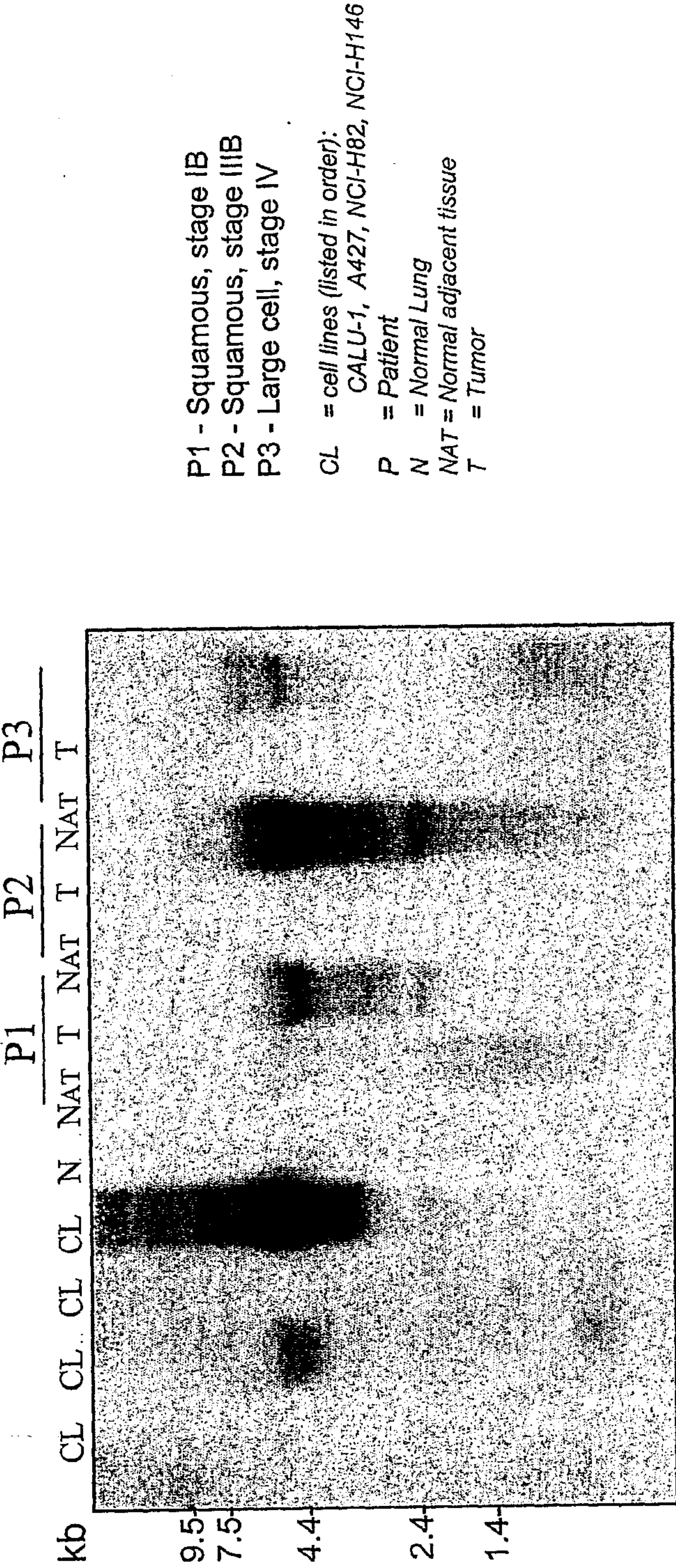


Figure 10. Expression of 158P1D7 in Breast Cancer Patient Specimens

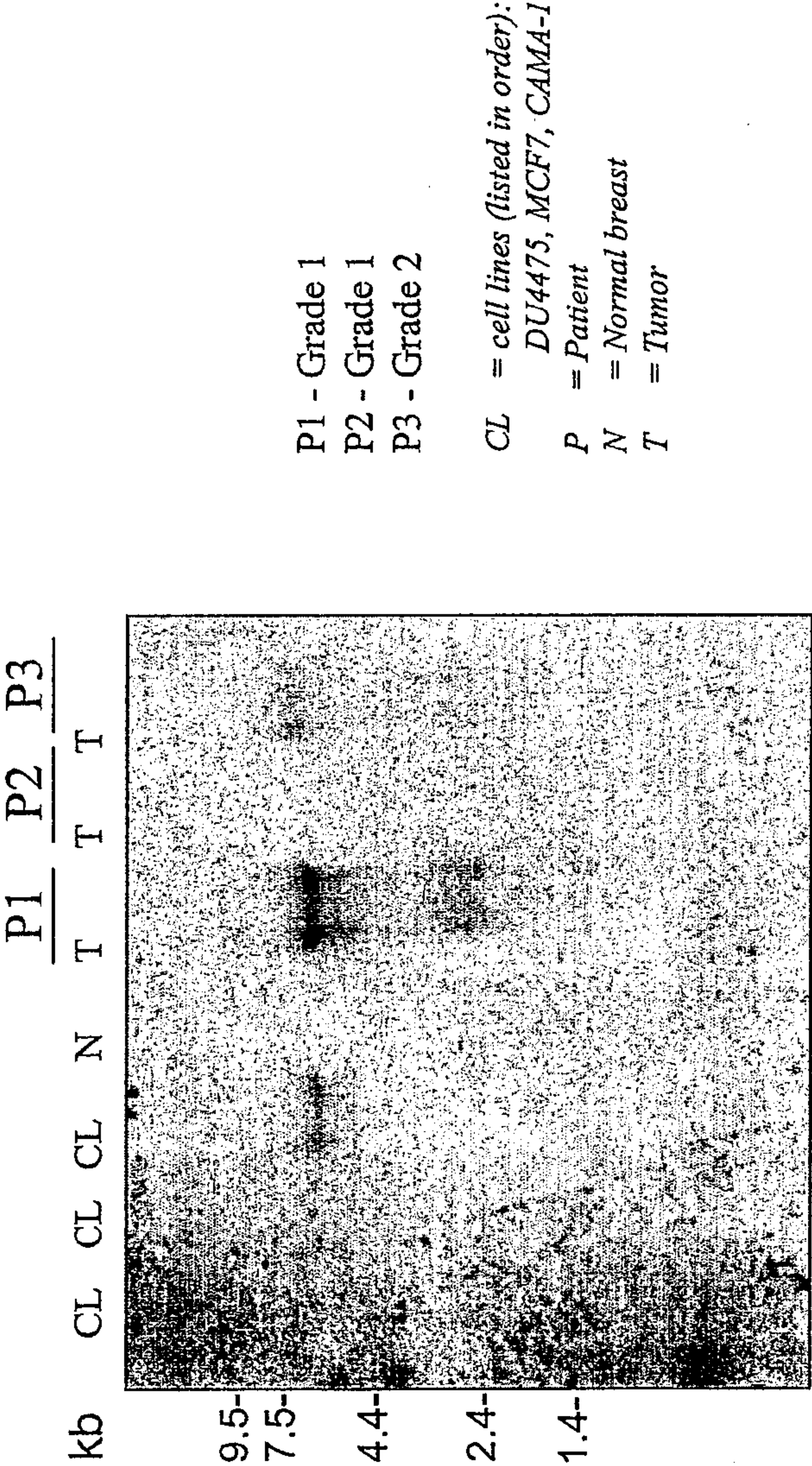


Figure 11a - 158P1D7 variant 1
Hydrophilicity profile
(Hopp T.P., Woods K.R., 1981.
Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828)

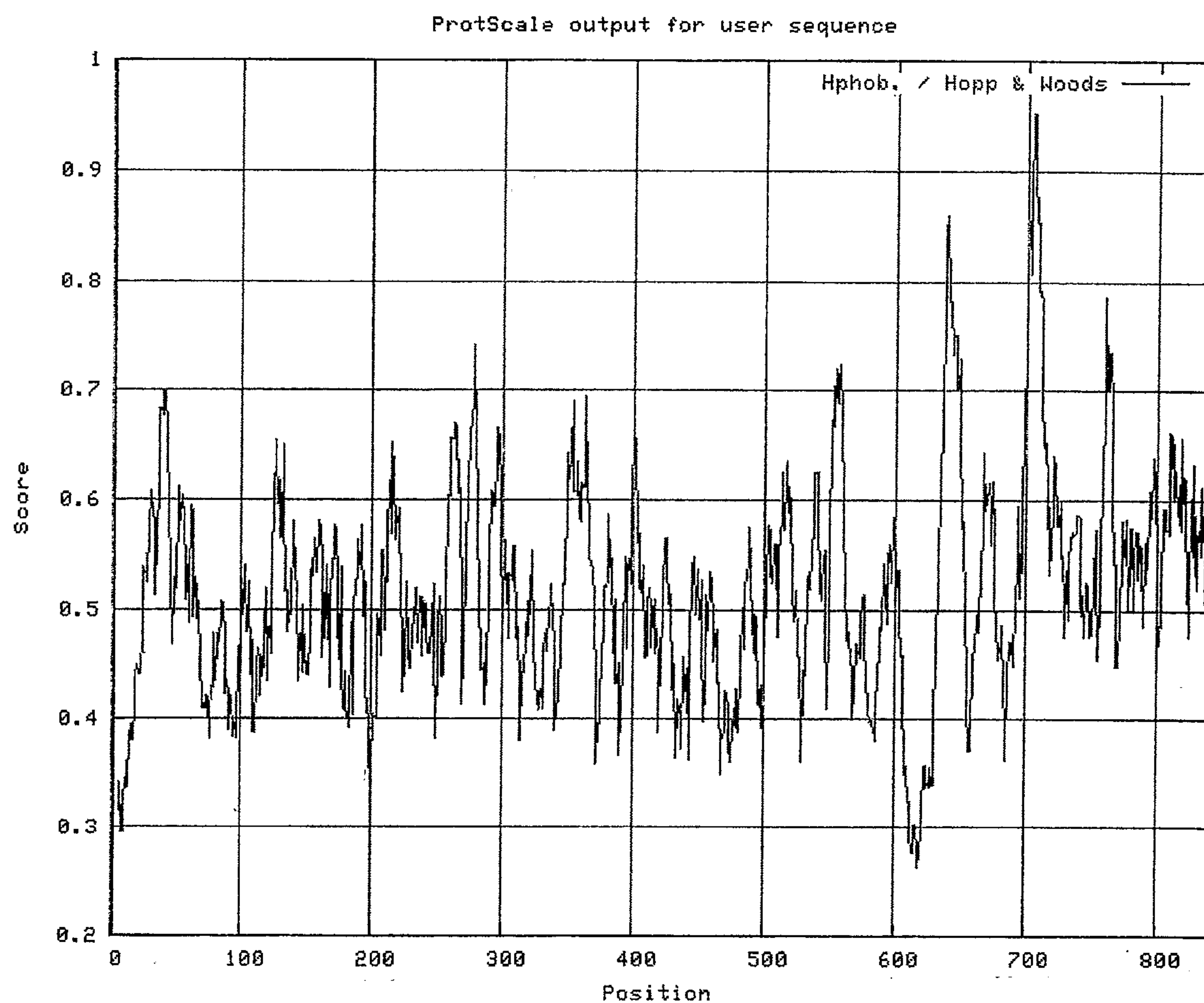


Figure 11b - 158P1D7 variant 3
Hydrophilicity profile
(Hopp T.P., Woods K.R., 1981.
Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828)

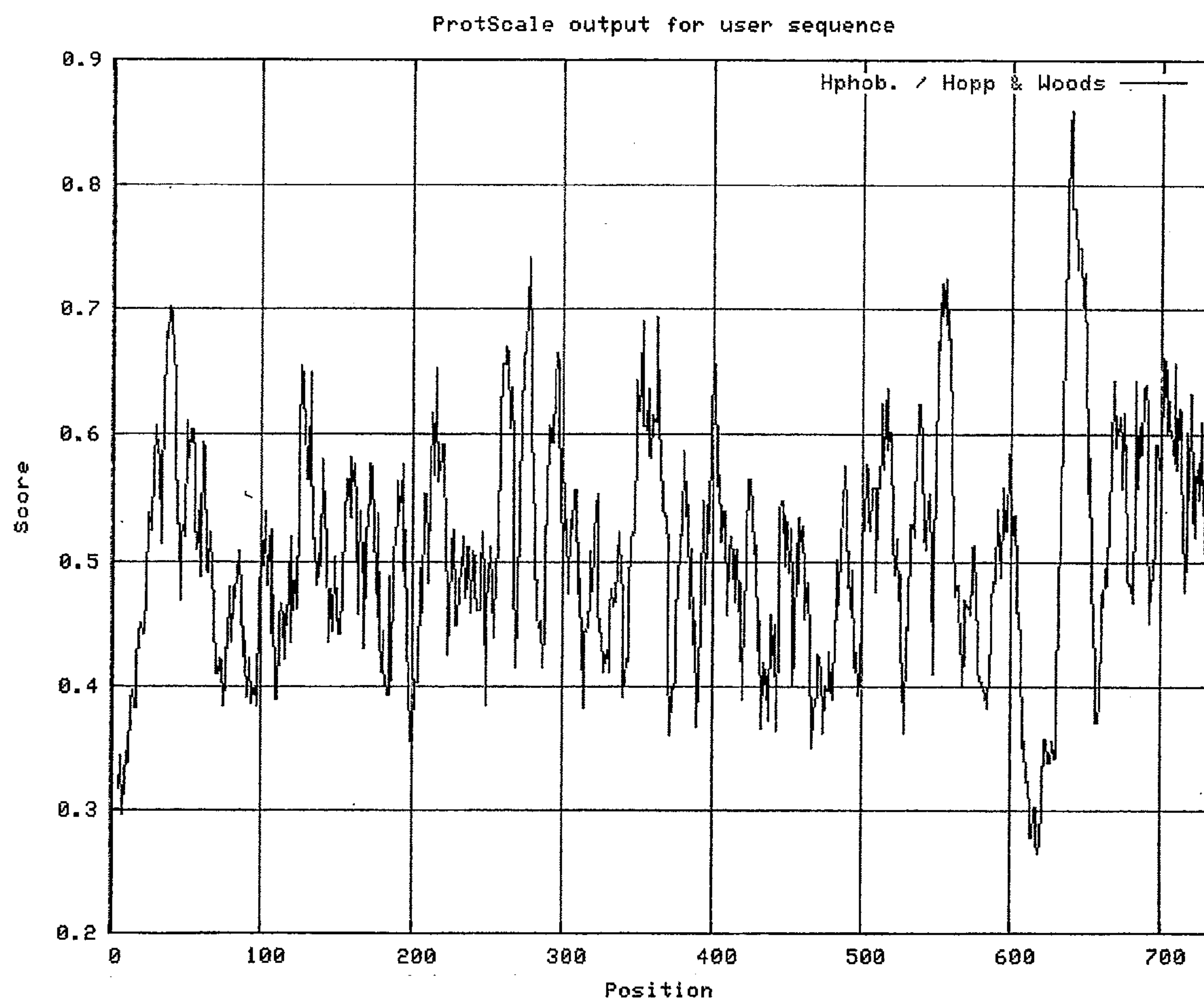


Figure 11c - 158P1D7 variant 4
Hydrophilicity profile
(Hopp T.P., Woods K.R., 1981.
Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828)

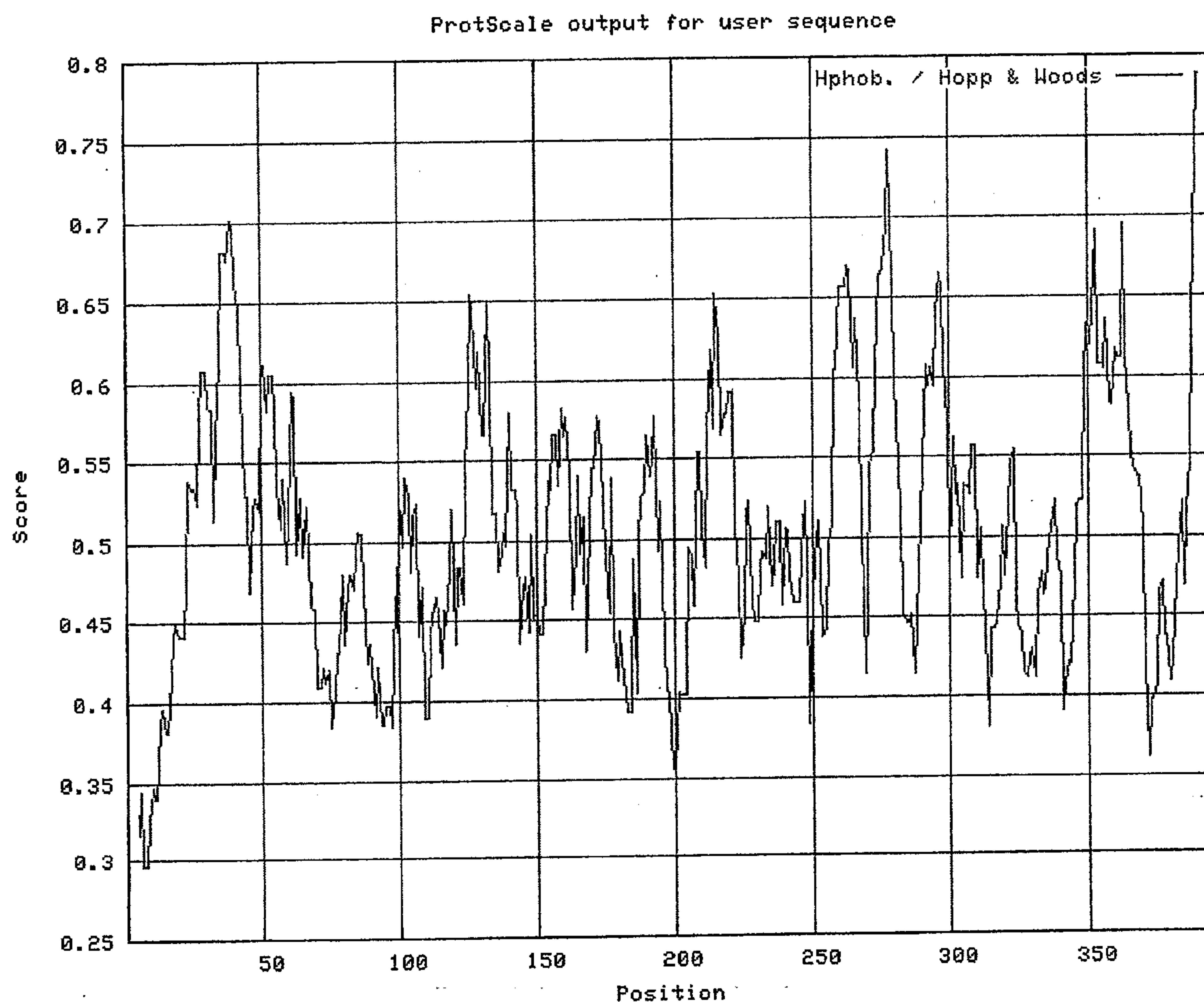


Figure 11d - 158P1D7 variant 6
Hydrophilicity profile
(Hopp T.P., Woods K.R., 1981.
Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828)

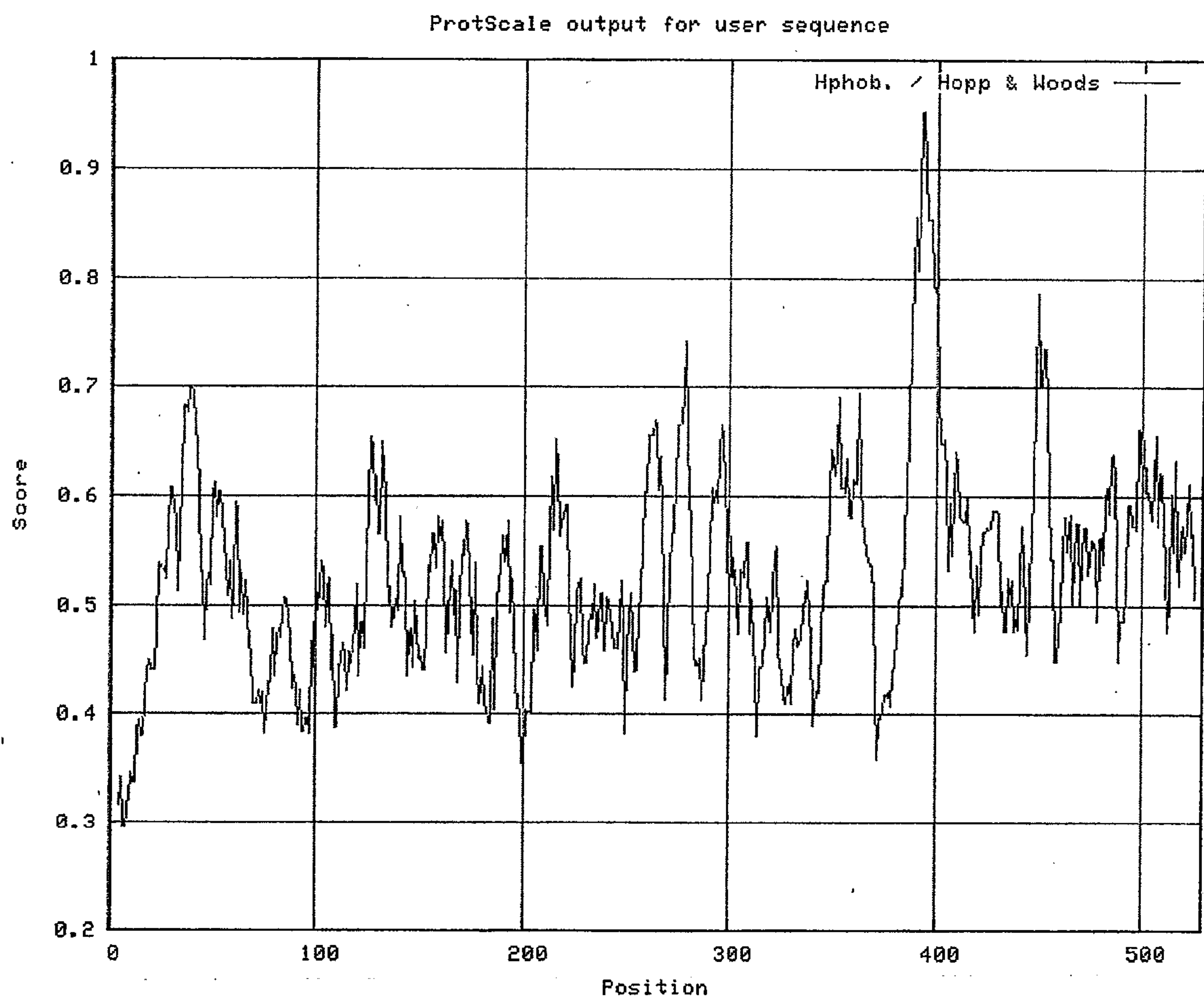


Figure 12a - 158P1D7 variant 1
Hydropathicity Profile
(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132)

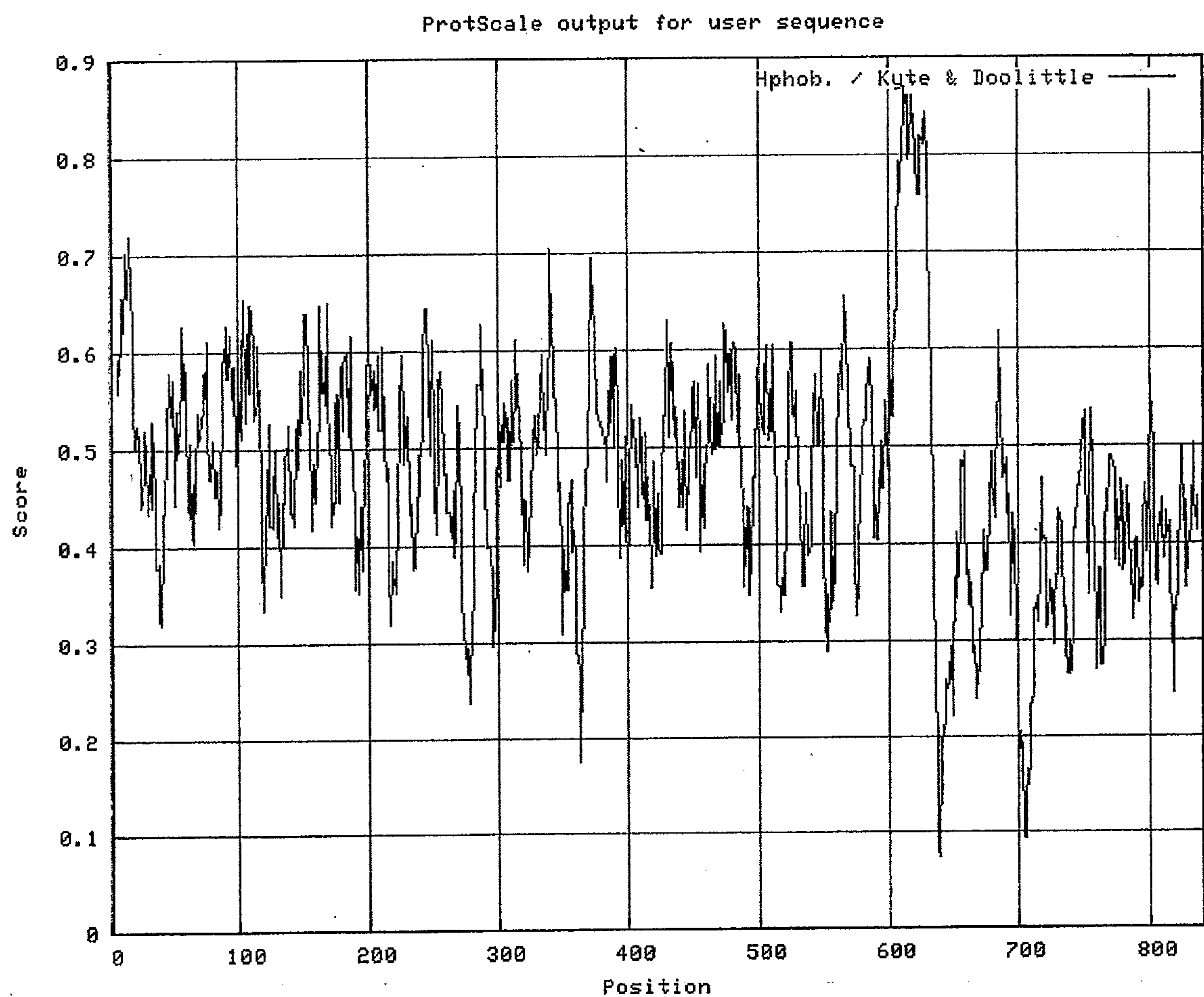


Figure 12b - 158P1D7 variant 3
Hydropathicity Profile
(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132)

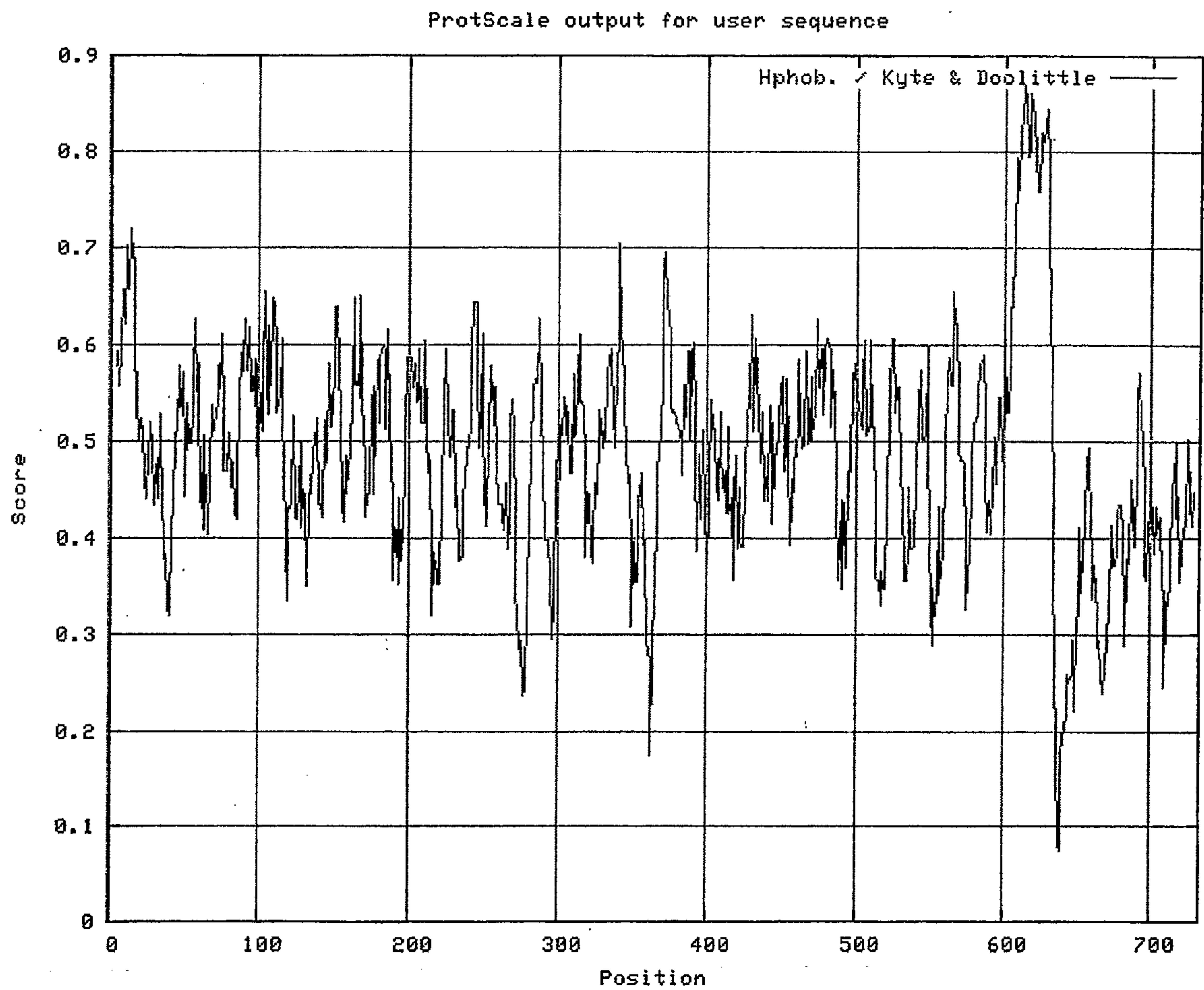


Figure 12c - 158P1D7 variant 4
Hydropathicity Profile
(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132)

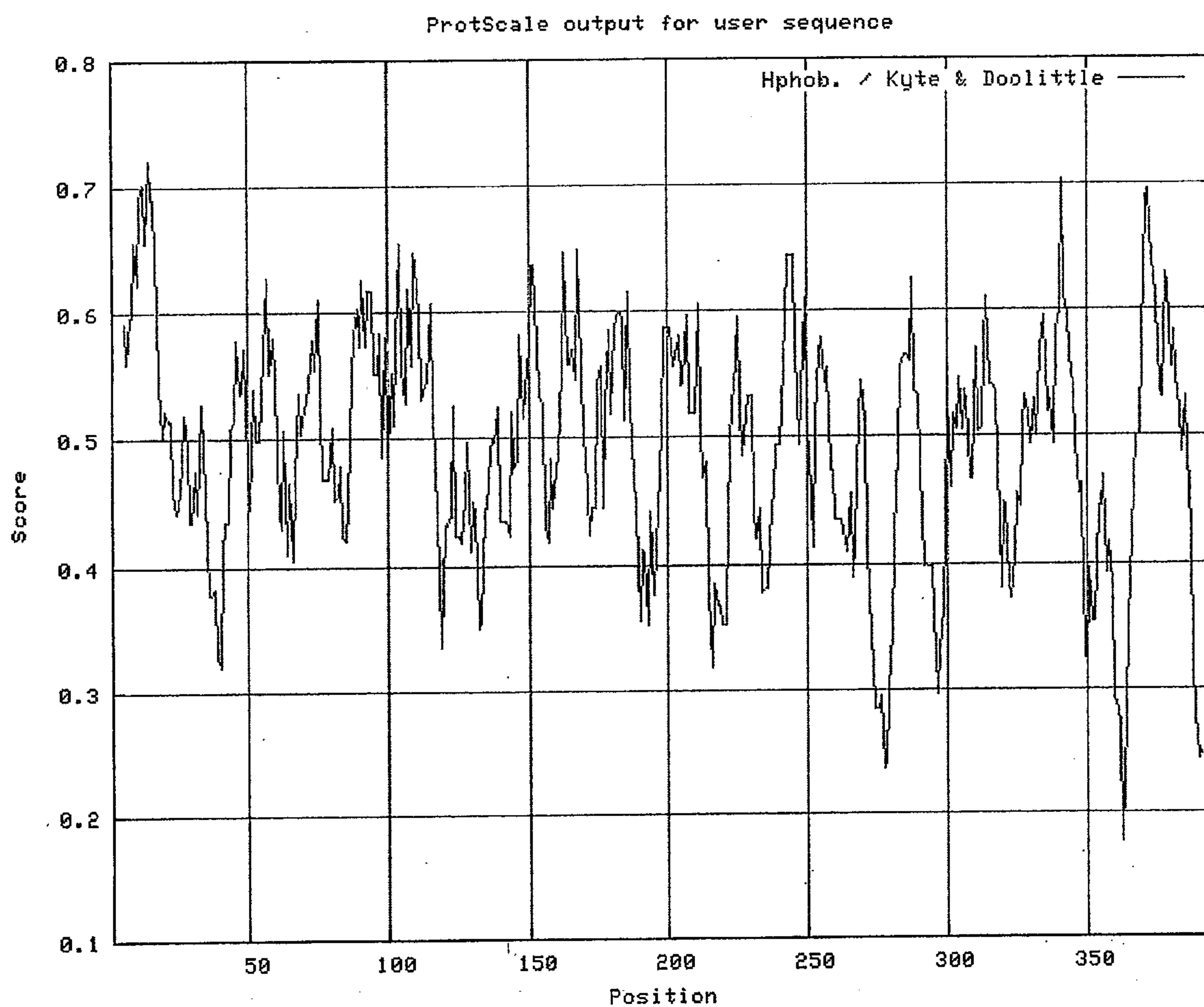


Figure 12d - 158P1D7 variant 6
Hydropathicity Profile
(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132)

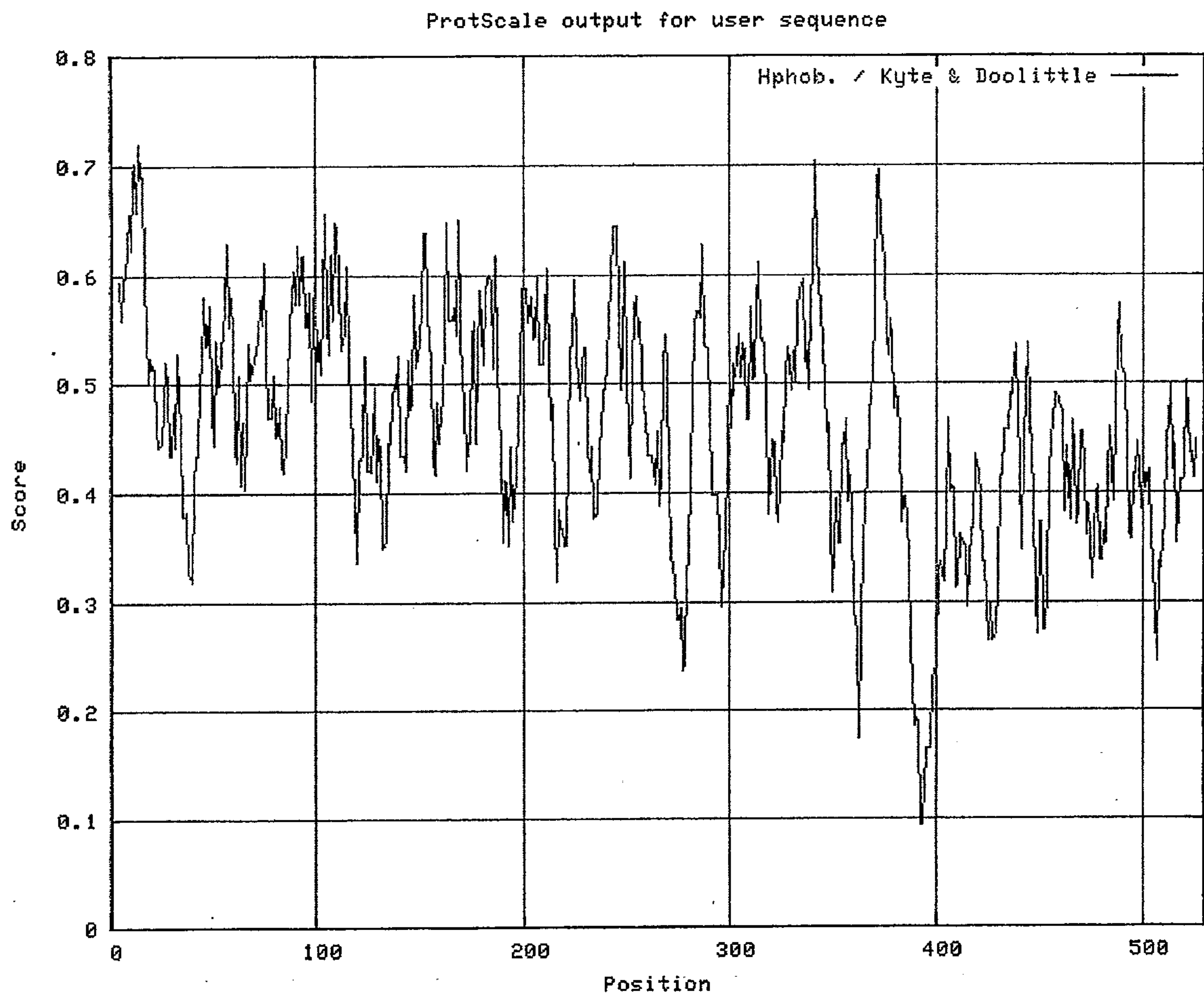


Figure 13a - 158P1D7 variant 1 %
Accessible Residues Profile
(Janin J., 1979. Nature 277:491-492)

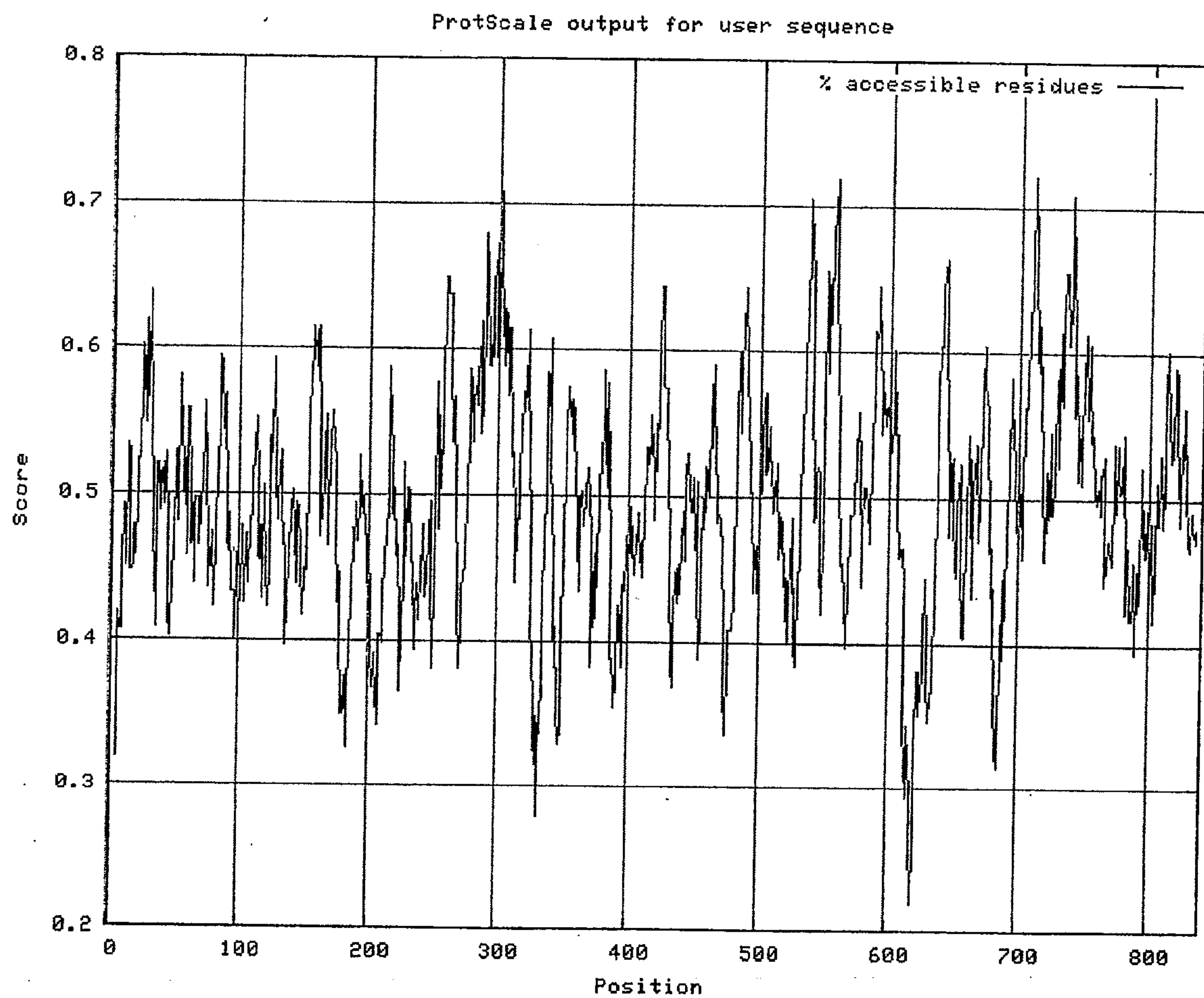


Figure 13b - 158P1D7 variant 3 %
Accessible Residues Profile
(Janin J., 1979. Nature 277:491-492)

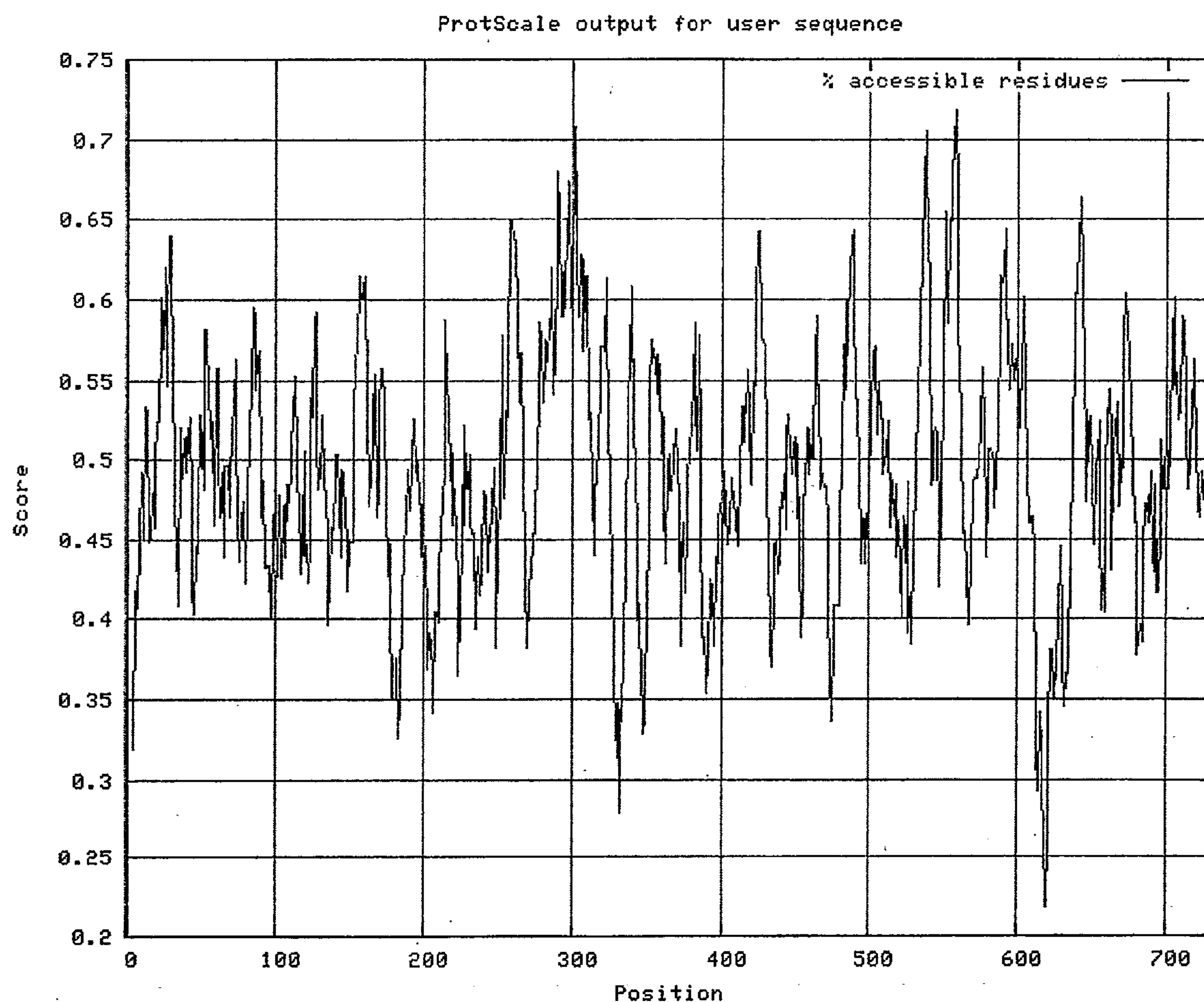


Figure 13c - 158P1D7 variant 4 %
Accessible Residues Profile
(Janin J., 1979. Nature 277:491-492)

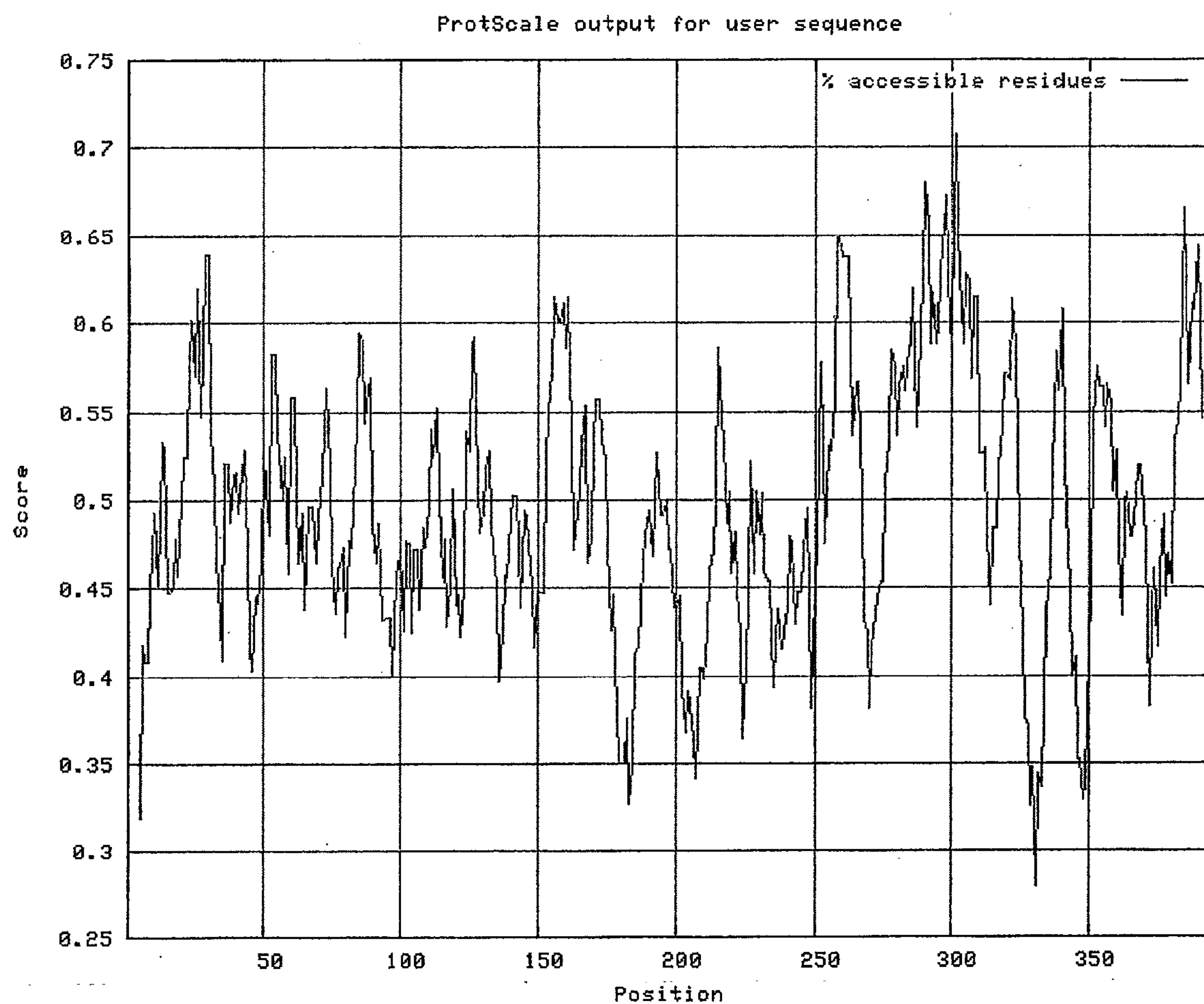


Figure 13d - 158P1D7 variant 6 %
Accessible Residues Profile
(Janin J., 1979. Nature 277:491-492)

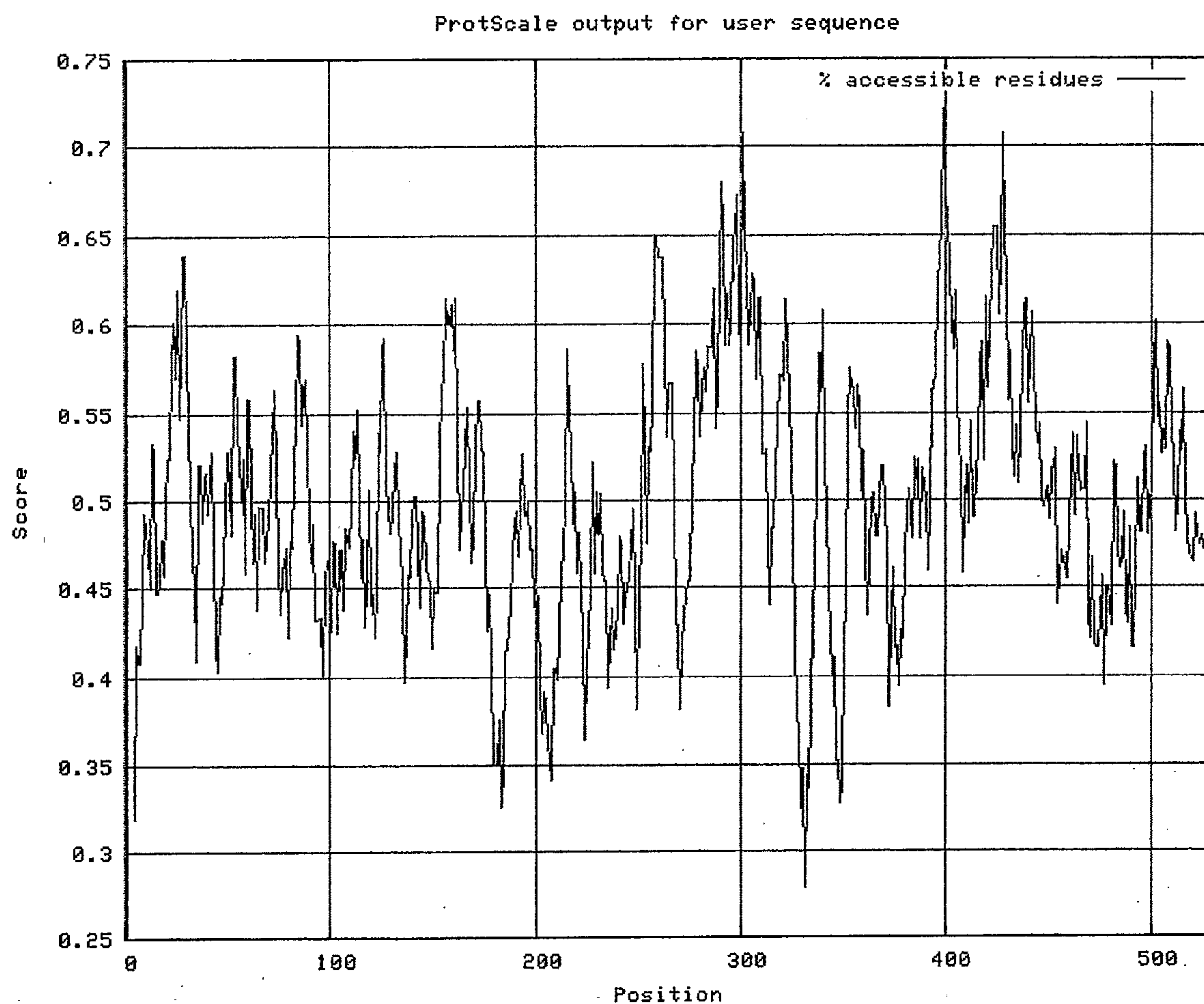


Figure 14a - 158P1D7 variant 1
Average Flexibility Profile
(Bhaskaran R., Ponnuswamy P.K., 1988.
Int. J. Pept. Protein Res. 32:242-255)

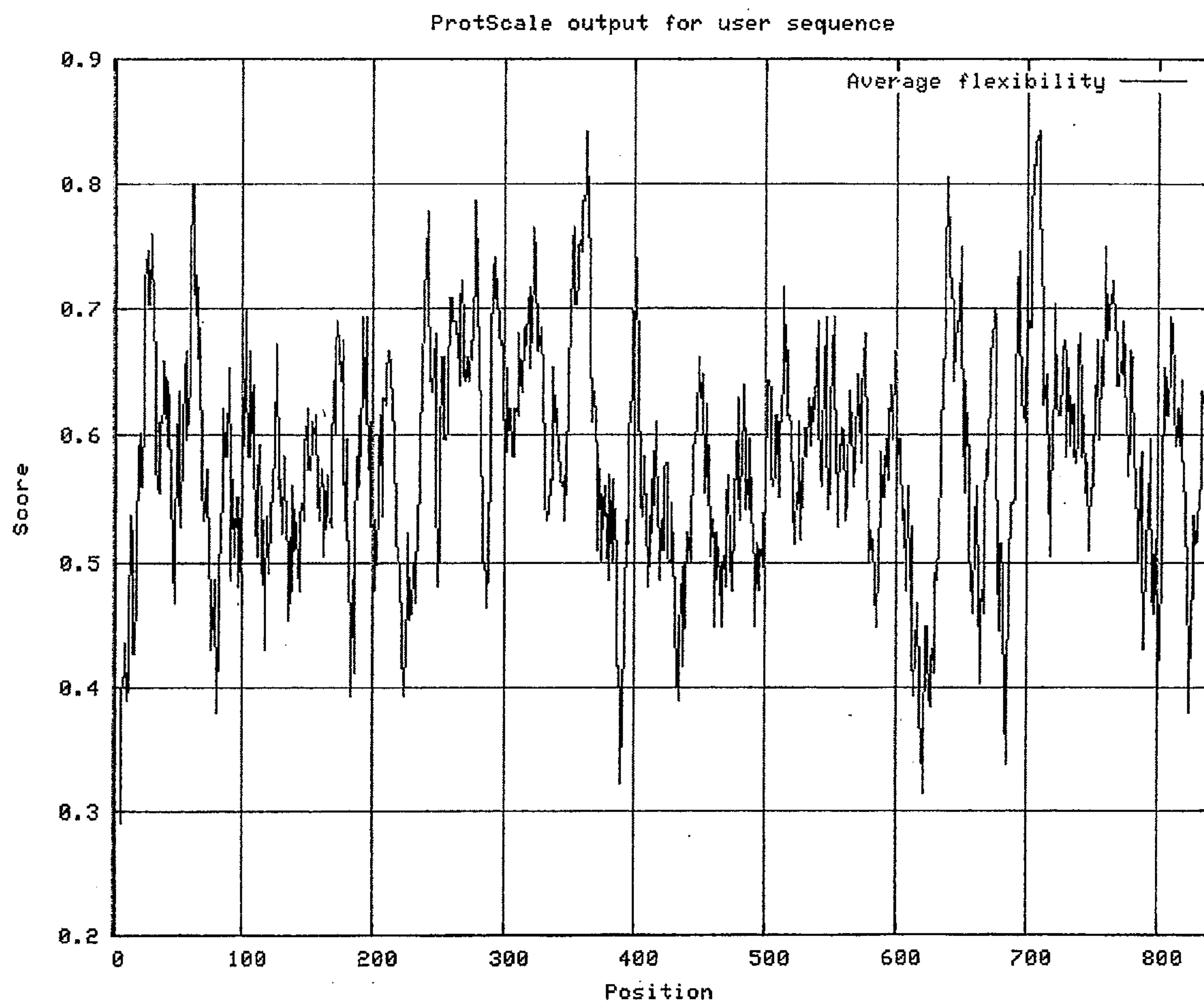


Figure 14B - 158P1D7 variant 3
Average Flexibility Profile
(Bhaskaran R., Ponnuswamy P.K., 1988.
Int. J. Pept. Protein Res. 32:242-255)

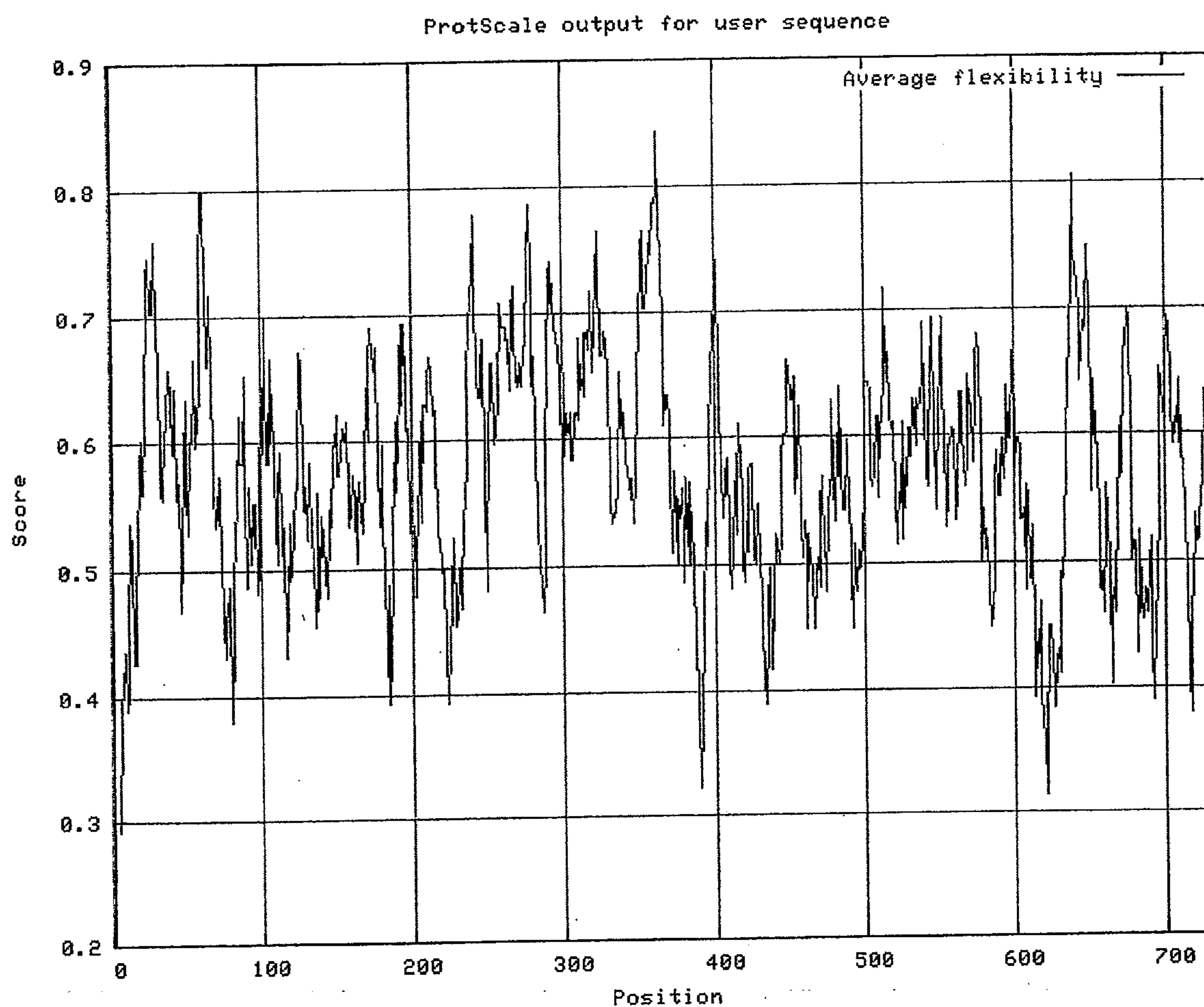


Figure 14c - 158P1D7 variant 4
Average Flexibility Profile
(Bhaskaran R., Ponnuswamy P.K., 1988.
Int. J. Pept. Protein Res. 32:242-255)

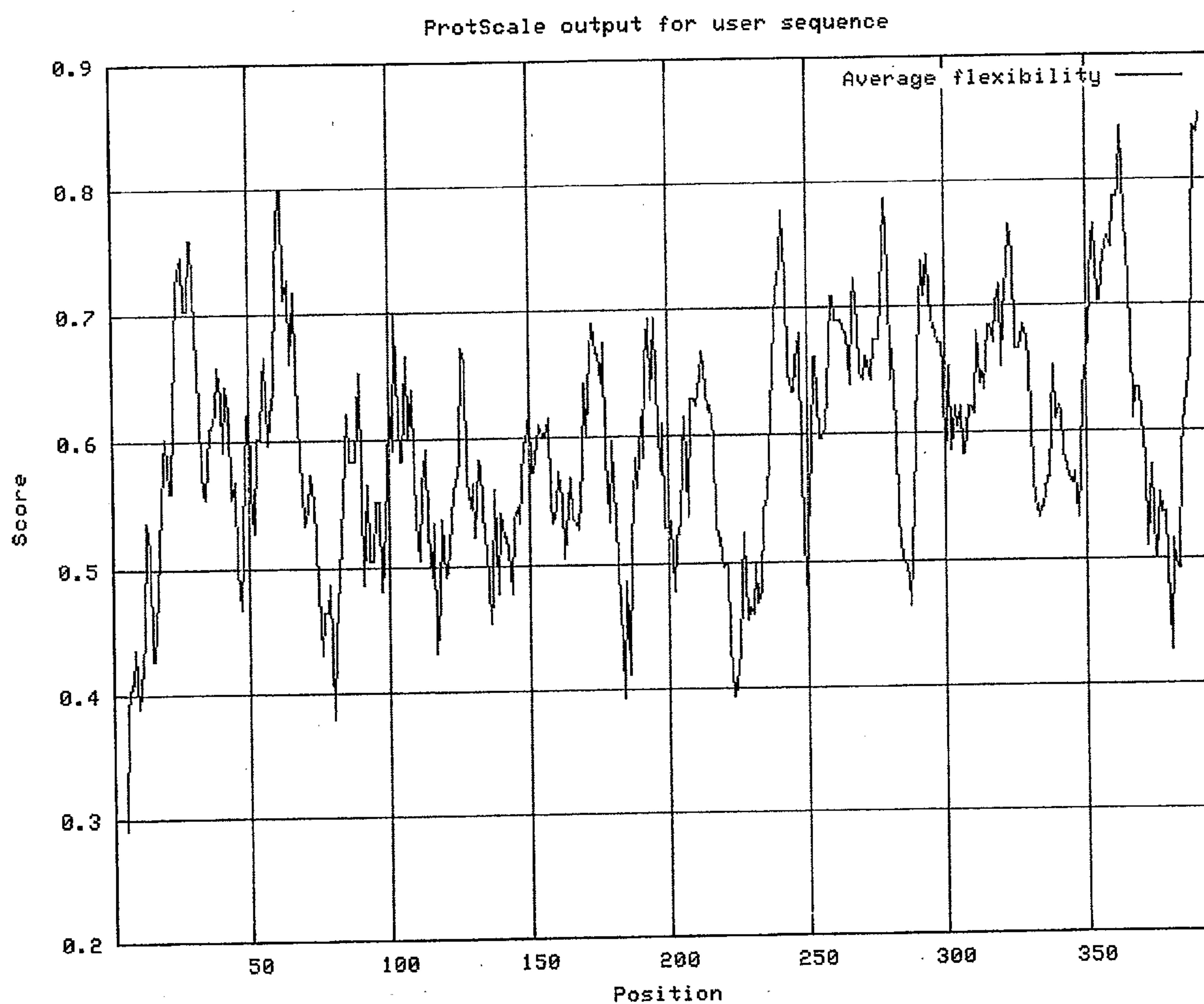


Figure 14d - 158P1D7 variant 6
Average Flexibility Profile
(Bhaskaran R., Ponnuswamy P.K., 1988.
Int. J. Pept. Protein Res. 32:242-255)

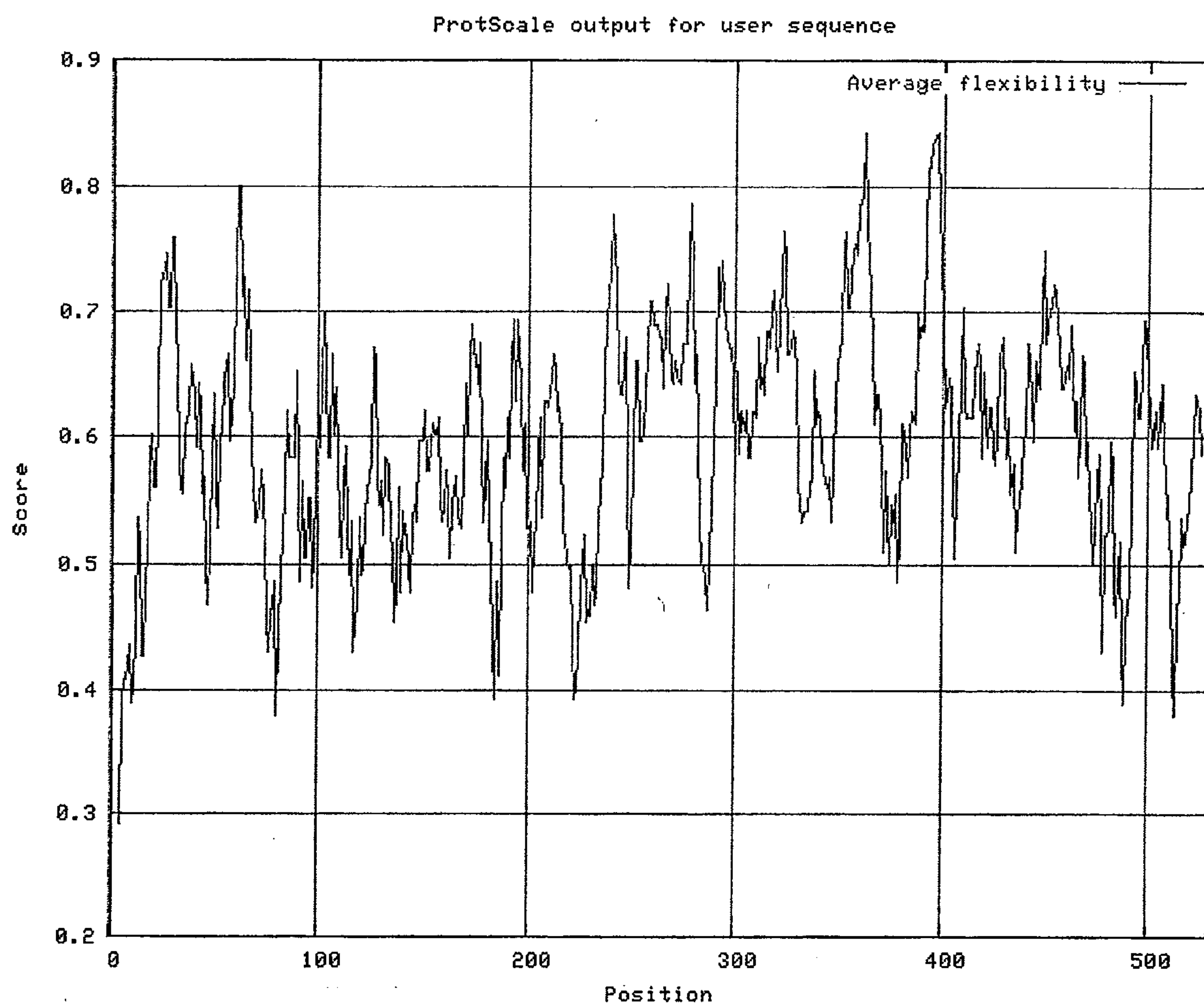


Figure 15a - 158P1D7
variant 1 Beta-turn Profile
(Deleage, G., Roux B. 1987.
Protein Engineering 1:289-294)

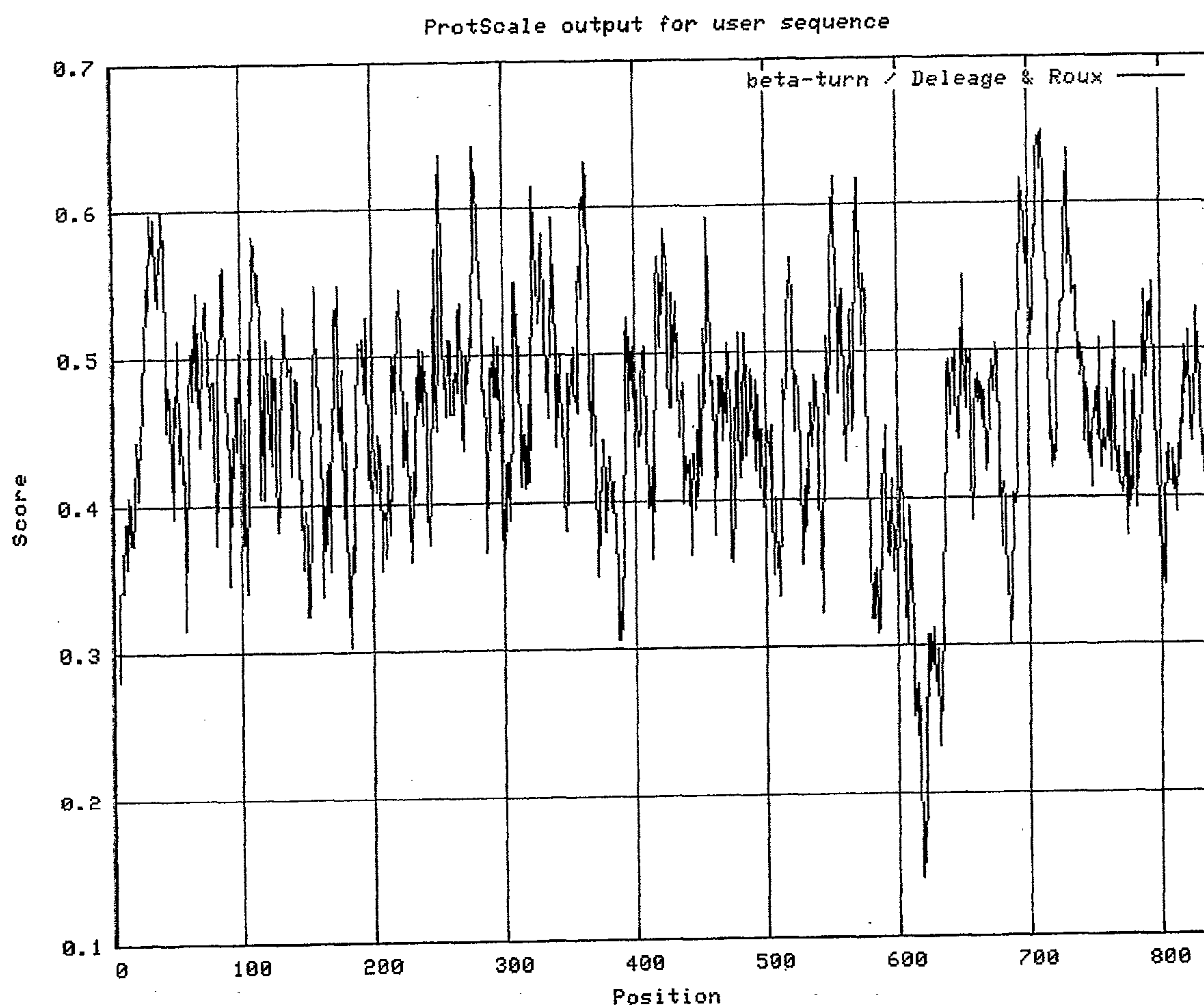


Figure 15b - 158P1D7 variant 3
Beta-turn Profile
(Deleage, G., Roux B. 1987.
Protein Engineering 1:289-294)

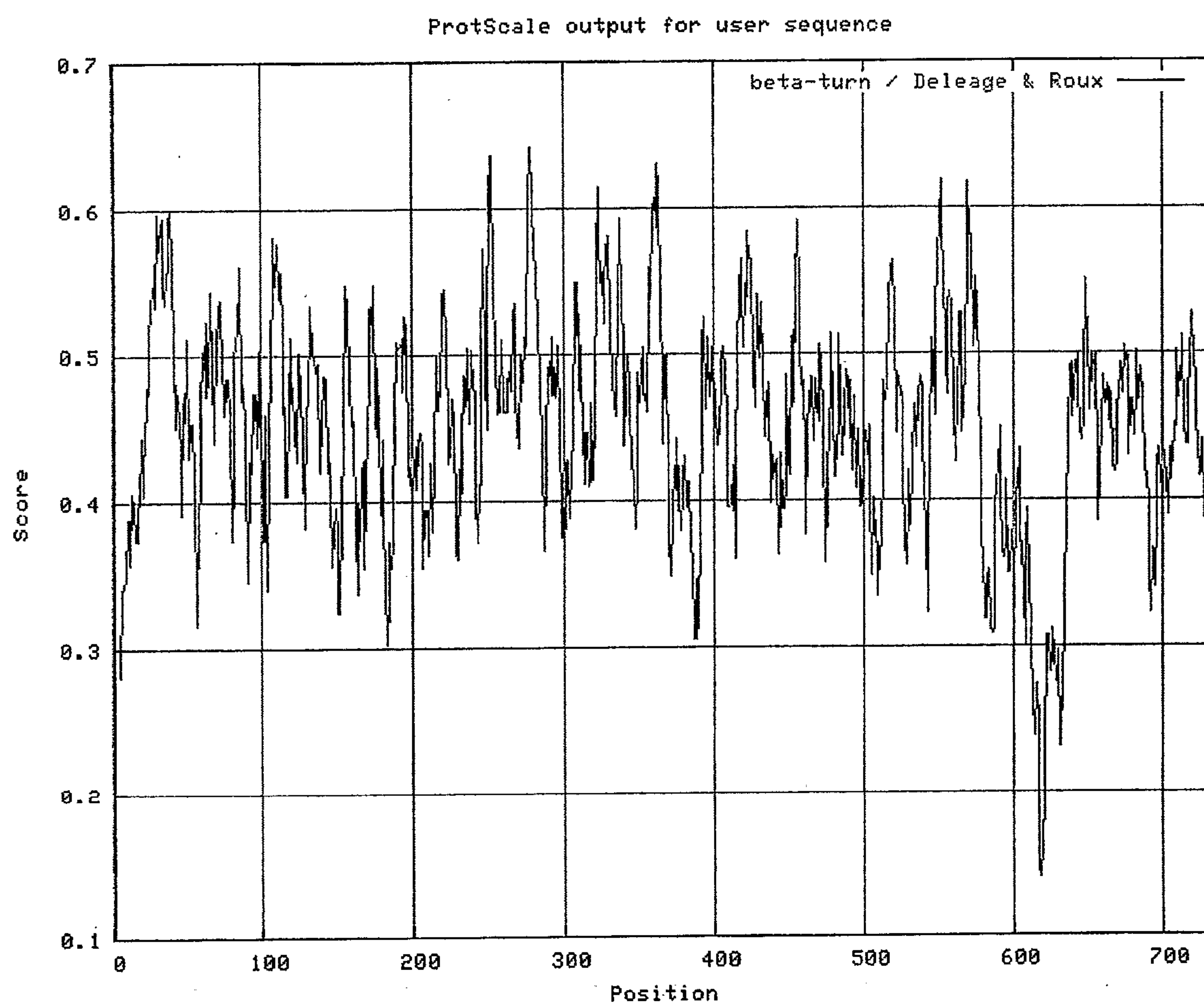


Figure 15c - 158P1D7 variant 4
Beta-turn Profile
(Deleage, G., Roux B. 1987.
Protein Engineering 1:289-294)

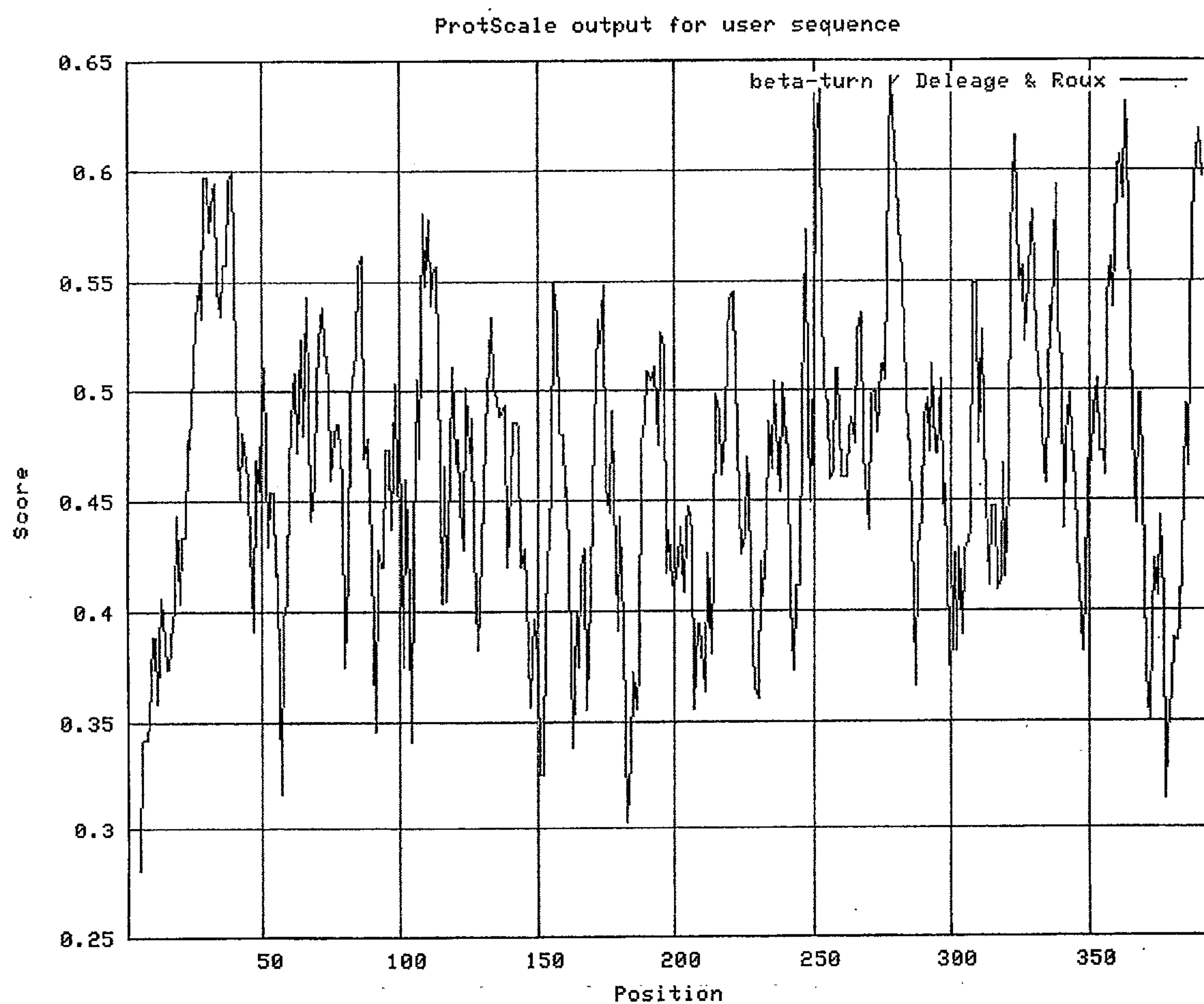


Figure 15d - 158P1D7 variant 6
Beta-turn Profile
(Deleage, G., Roux B. 1987.
Protein Engineering 1:289-294)

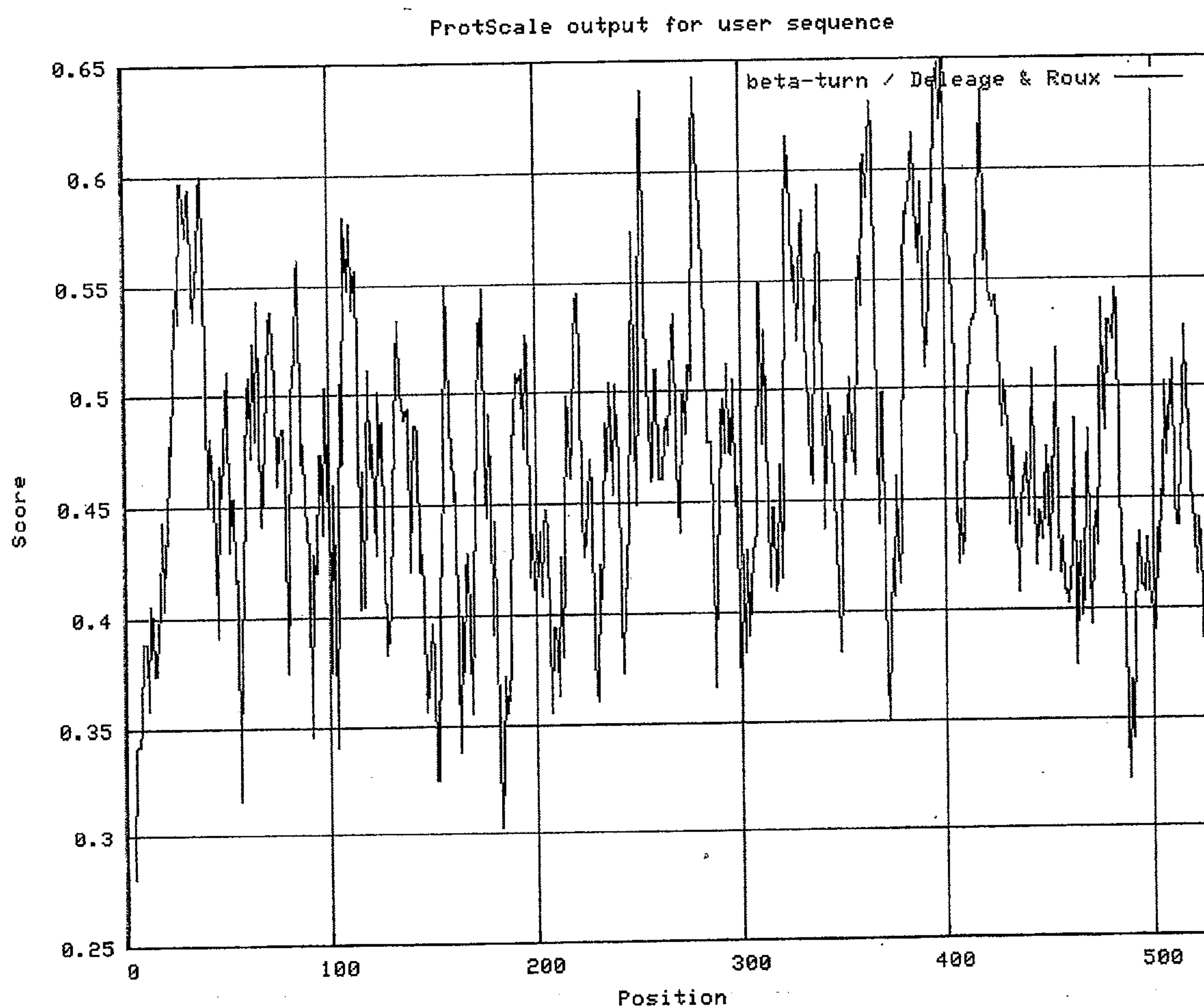
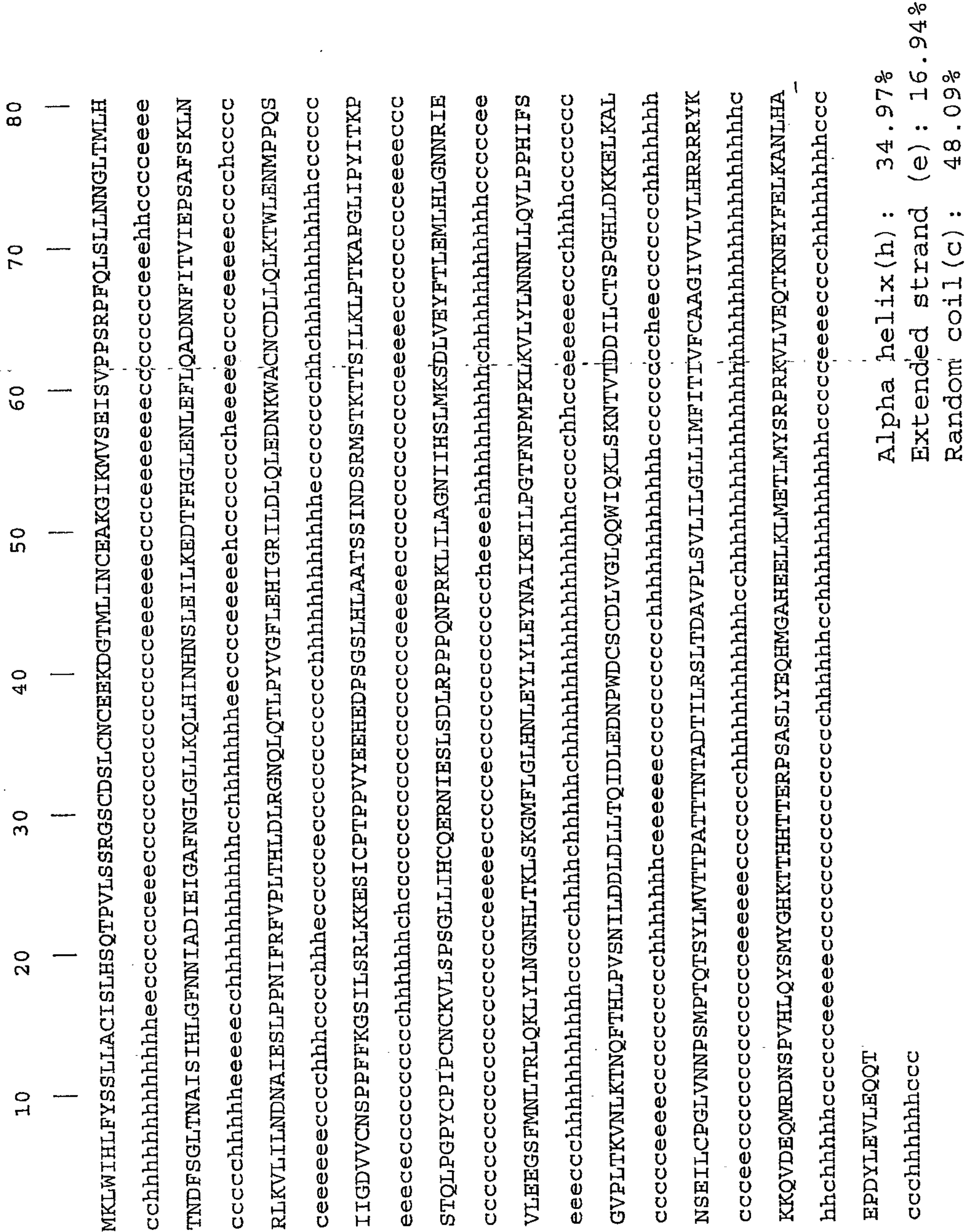


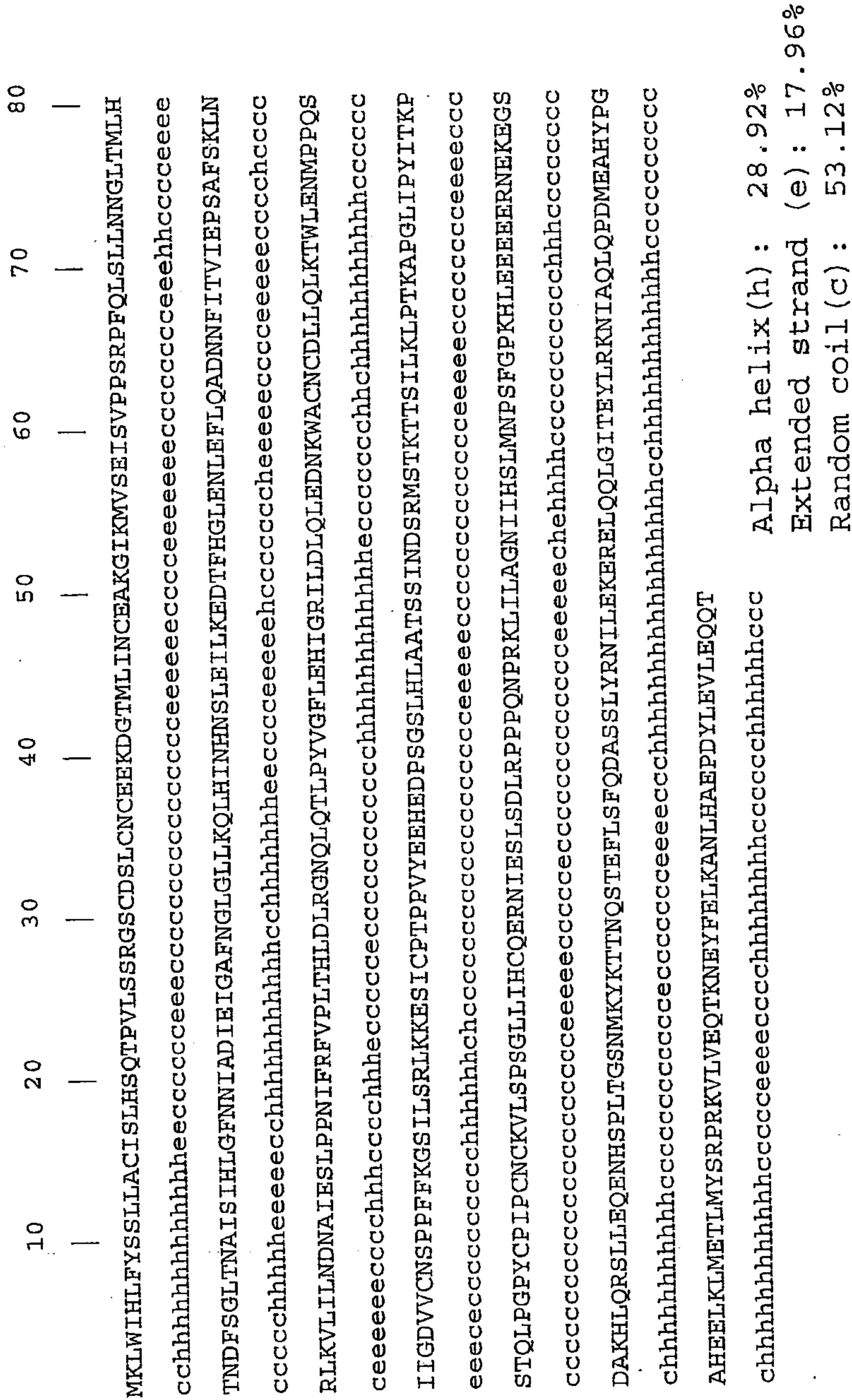
Fig 16B Secondary structure prediction of 158P1D7 variant 3



	10	20	30	40	50	60	70	80
MKLWIHLFYSSLLACISLHSQTPVLSRSGSCDSLNCNCEKDGTMLINCEAKGIMVSEISVPPSRPFQLSLLNNGLTMLH								
ccchhhhhhhhhhecccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc								
TNDFSGLTNAISIHGFENNIADIEIGAFNGLGLLKQHLHNHNSLEILKEDTFHGLENLEFLQADNNFITVIEPSAFSKLN								
ccccchhhheeeeeccchhhhhhhhhccchhhhhhhhecccccccccccccccccccccccccccccccccccccccc								
RLKVLILNDNAIESLPPNIFRFVPLTHLDLRGNQLQTLPYVGFEHIGRILDQLEDNKWACNCDLLQLKTWLENMPPQS								
ceeeeeccchhhcccccchhhhecccccccccccccccccccccccccccccccccccccccccccccccccccc								
IIGDVVCNSPFFKGSILRLKKEICPTPPVYEEHEDPSGSLHLAATSSINDSRMSTKTTISILKLPTKAPGLIPYITKP								
eeeeccccccccchhhhhchcccccccccccccccccccccccccccccccccccccccccccccccccccccccc								
STQLPGPYCIPCNCKVLSPSGLLIHCQERNIESLSDLRPPQNPRKLILAGNIHSLMKSILWSKASGRGREE								
cccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccccc								

Alpha helix(h): 24.56%
Extended strand (e): 20.76%
Random coil(c): 54.68%

Fig 16D Secondary structure prediction of 158P1D7 variant 6



Transmembrane prediction for 158P1D7 variant 1

Fig 16E

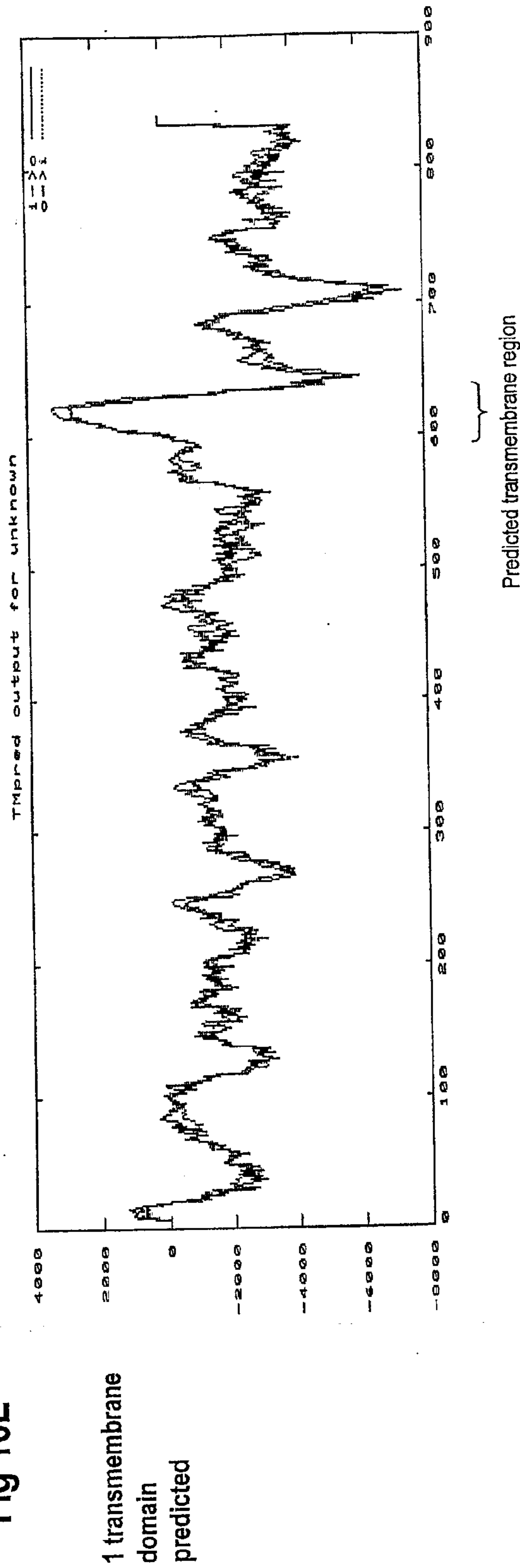
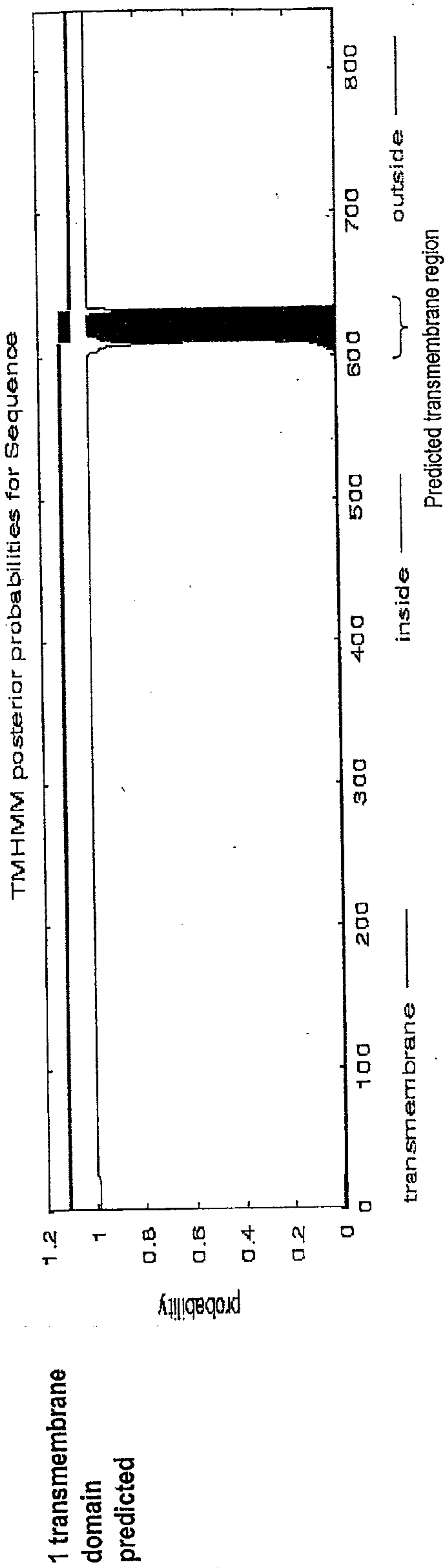


Fig 16F



Transmembrane prediction for 158P1D7 variant 3

Fig 16G

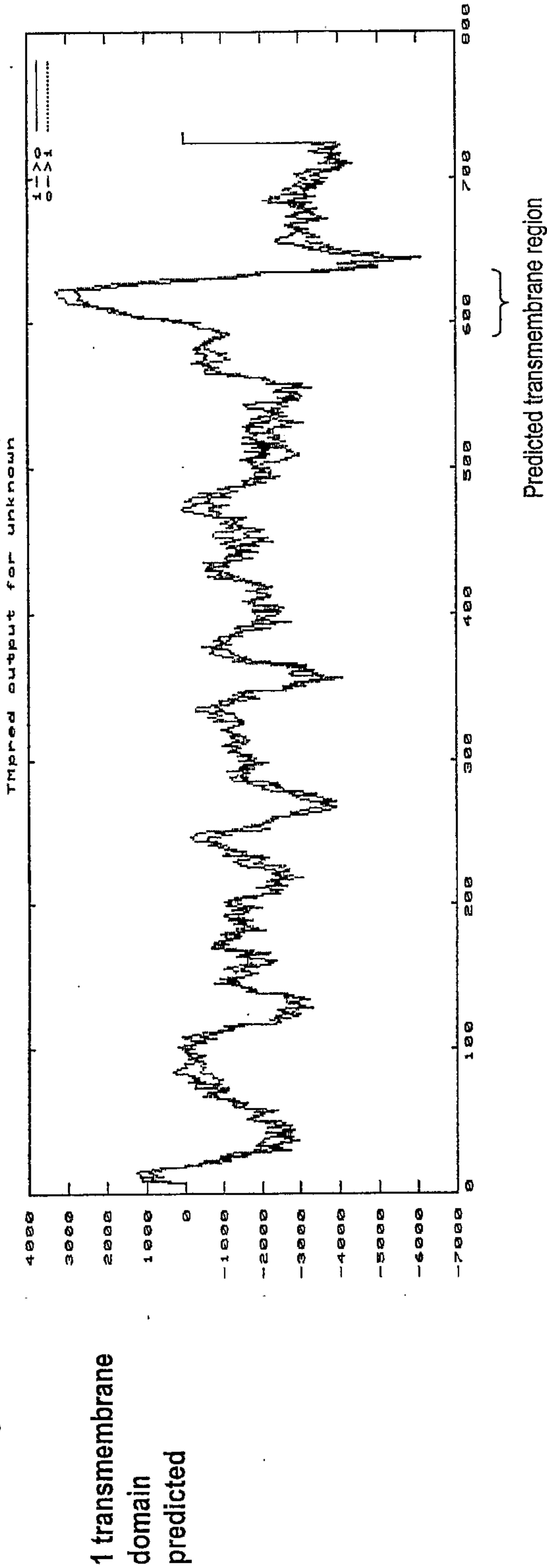


Fig 16H

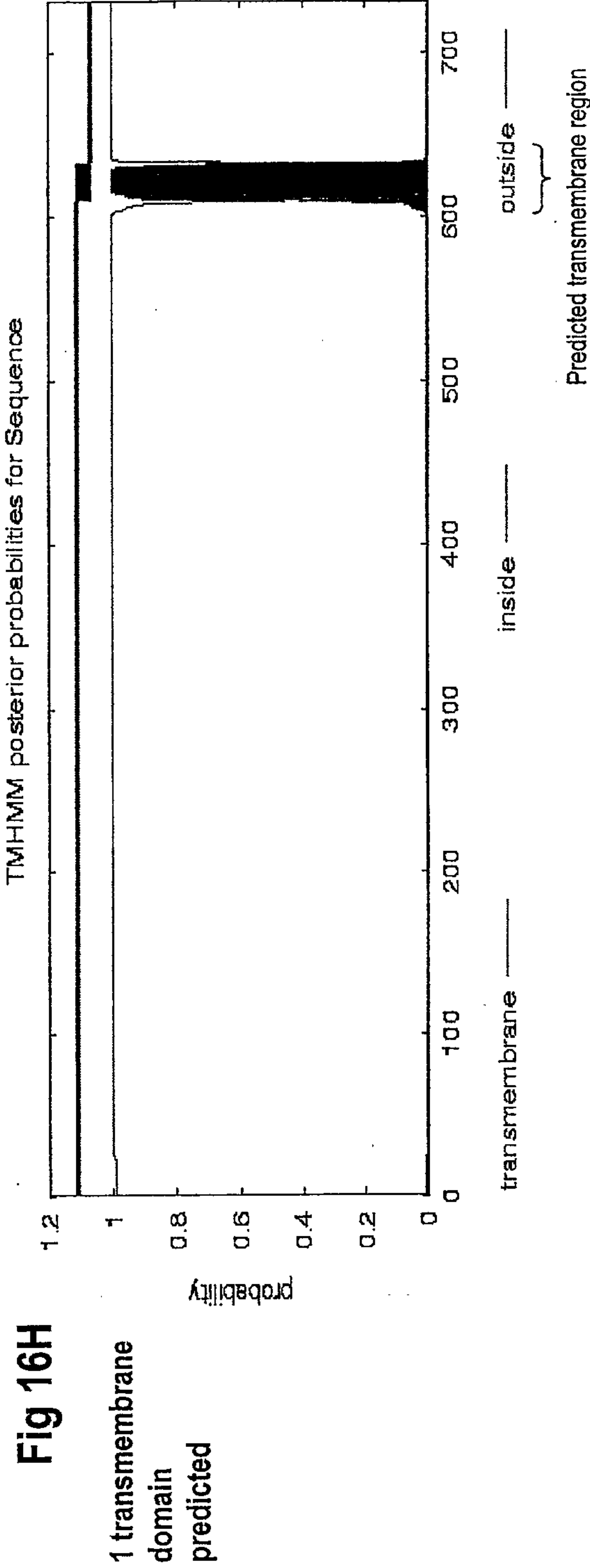


Fig 16I

Transmembrane prediction for 158P1D7 variant 4

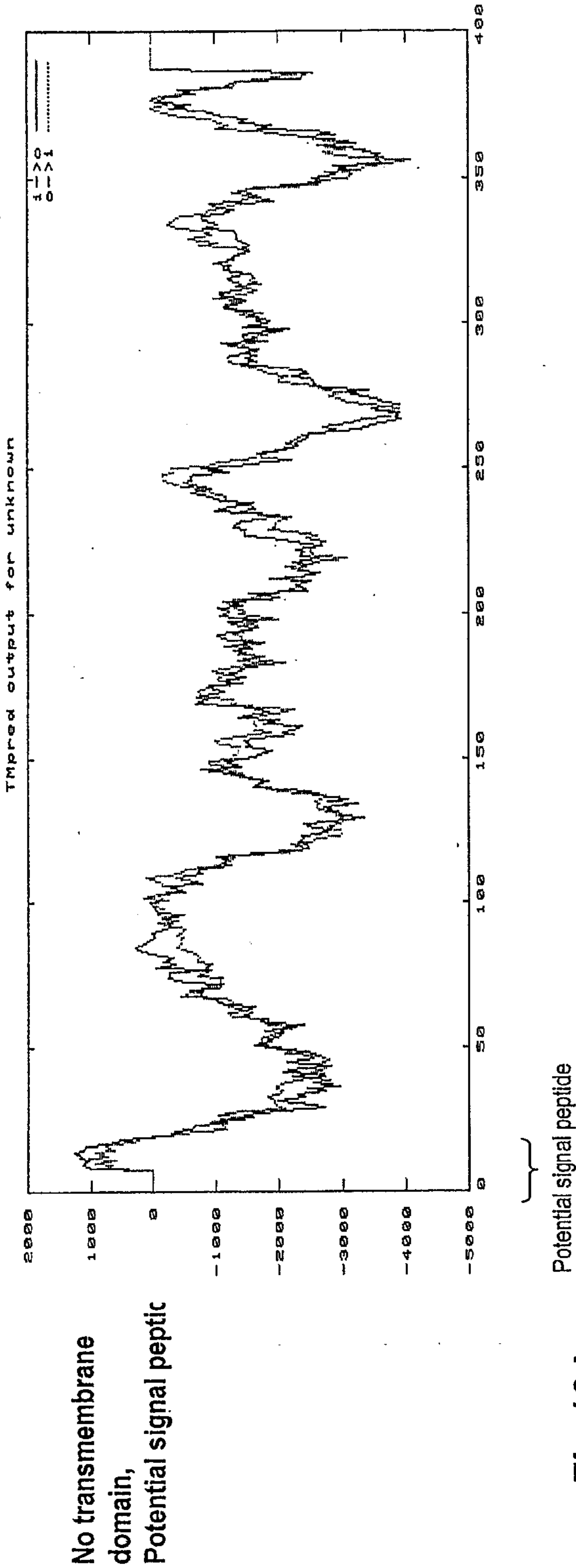
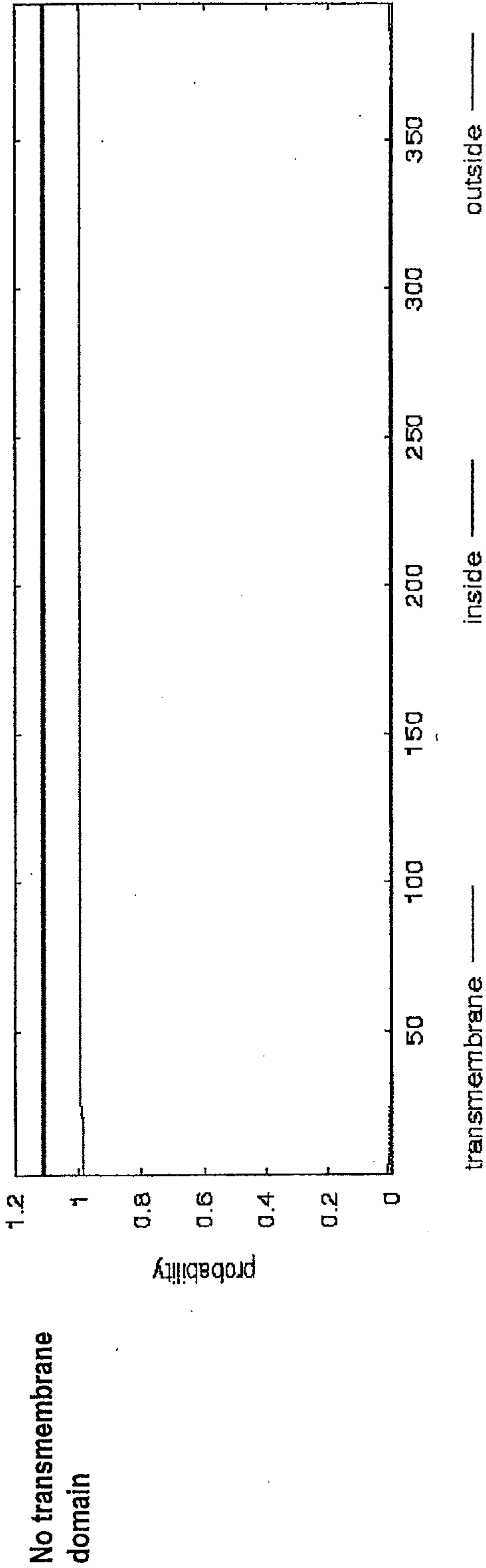


Fig 16J

TMHMM posterior probabilities for Sequence



Transmembrane prediction for 158P1D7 variant 6

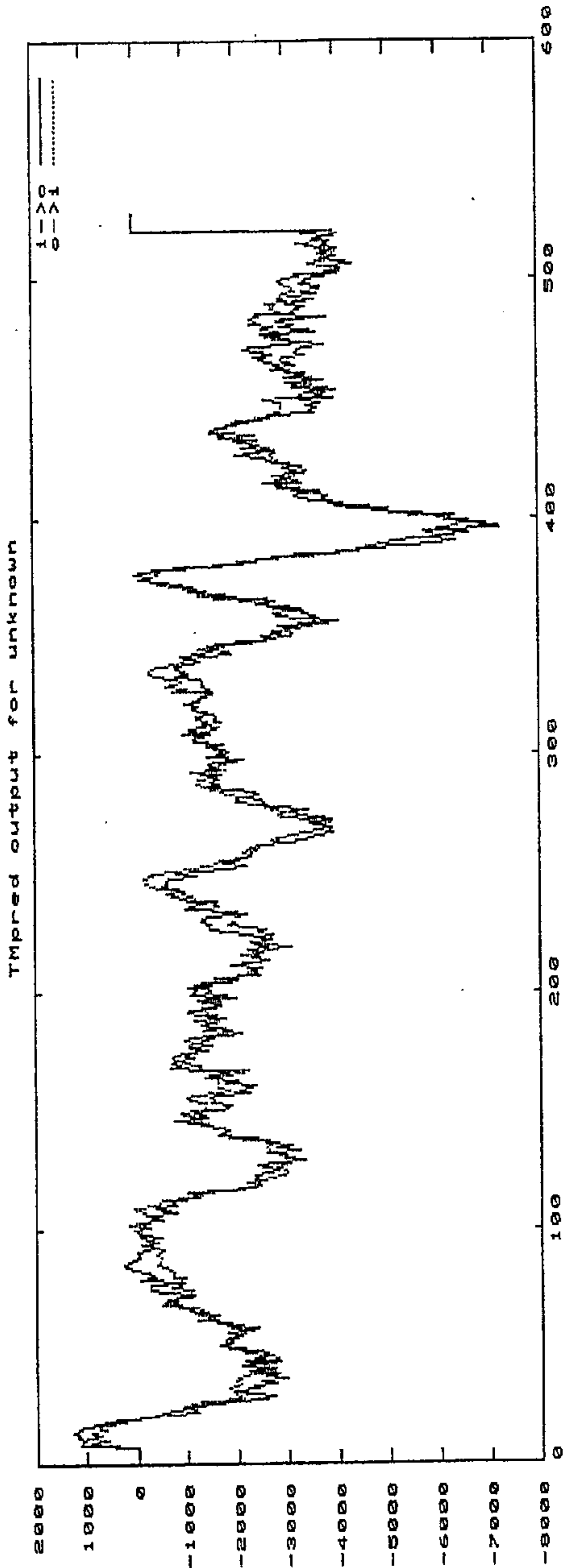


Fig 16K

No transmembrane
domain,
Potential signal peptide

TMHMM posterior probabilities for Sequence

Potential signal peptide

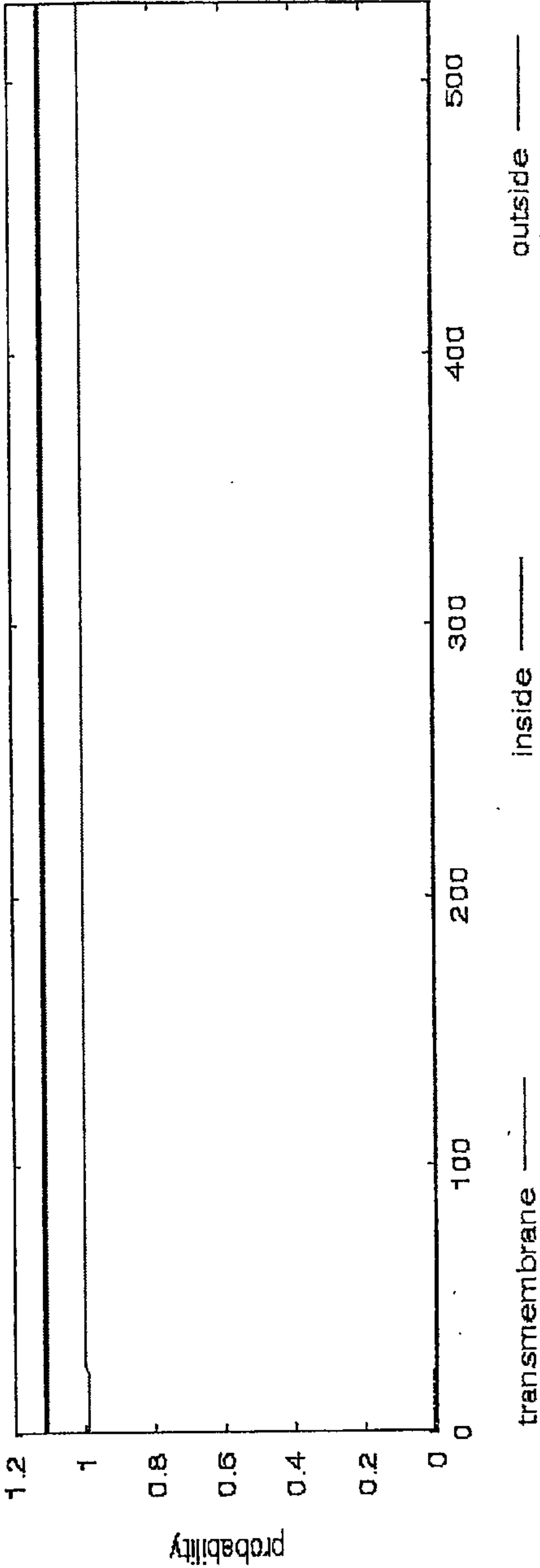


Fig 16L

No transmembrane
domain

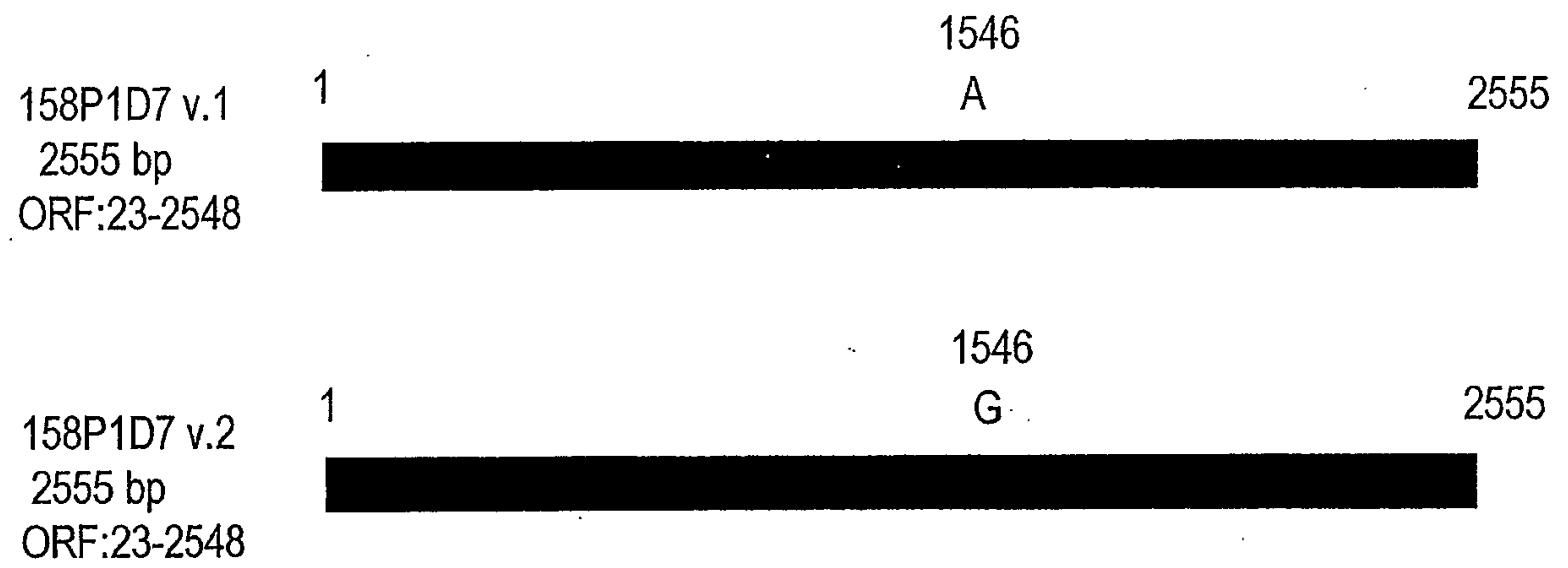
Figure 17

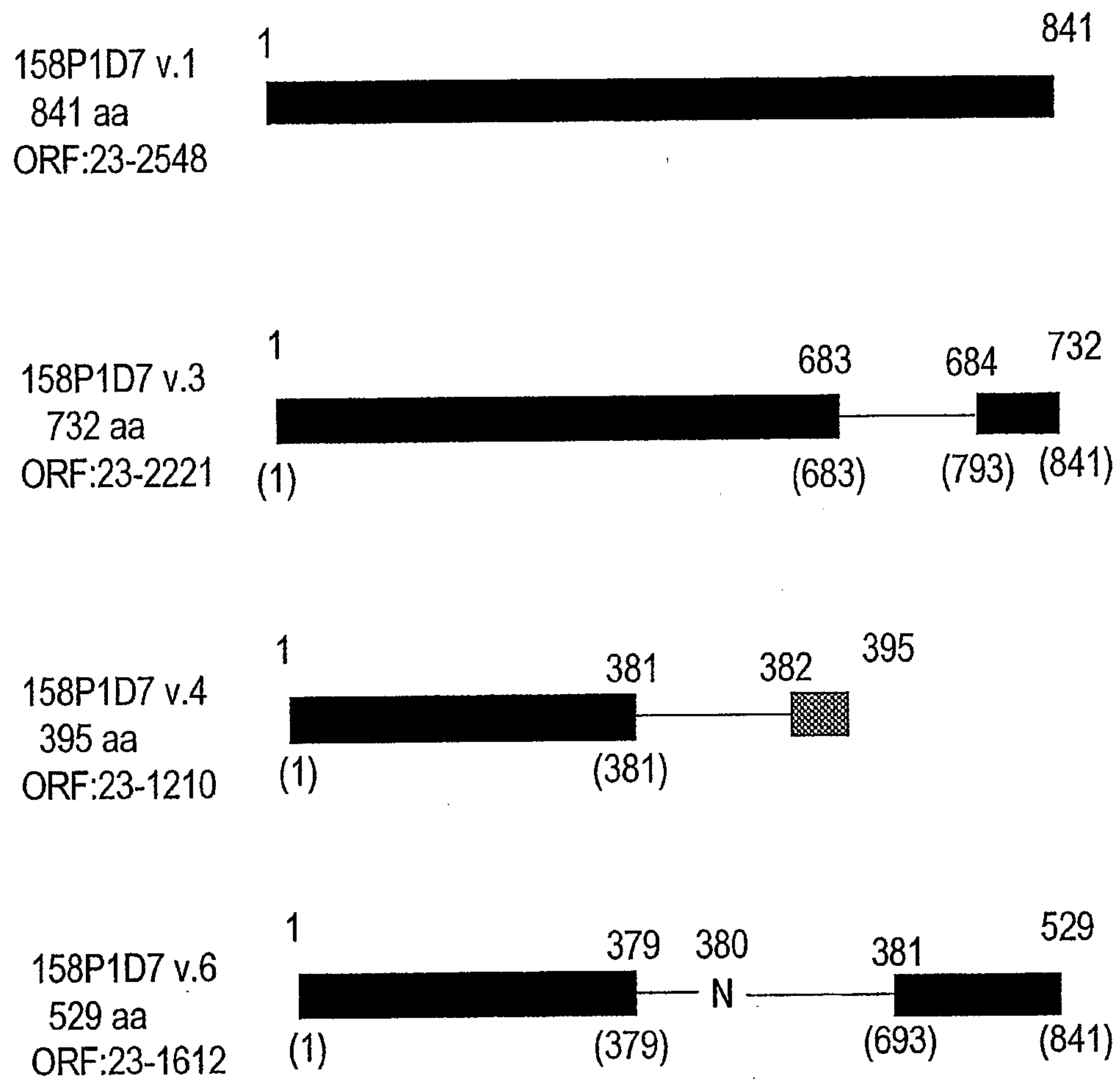
Figure 18

Figure 19

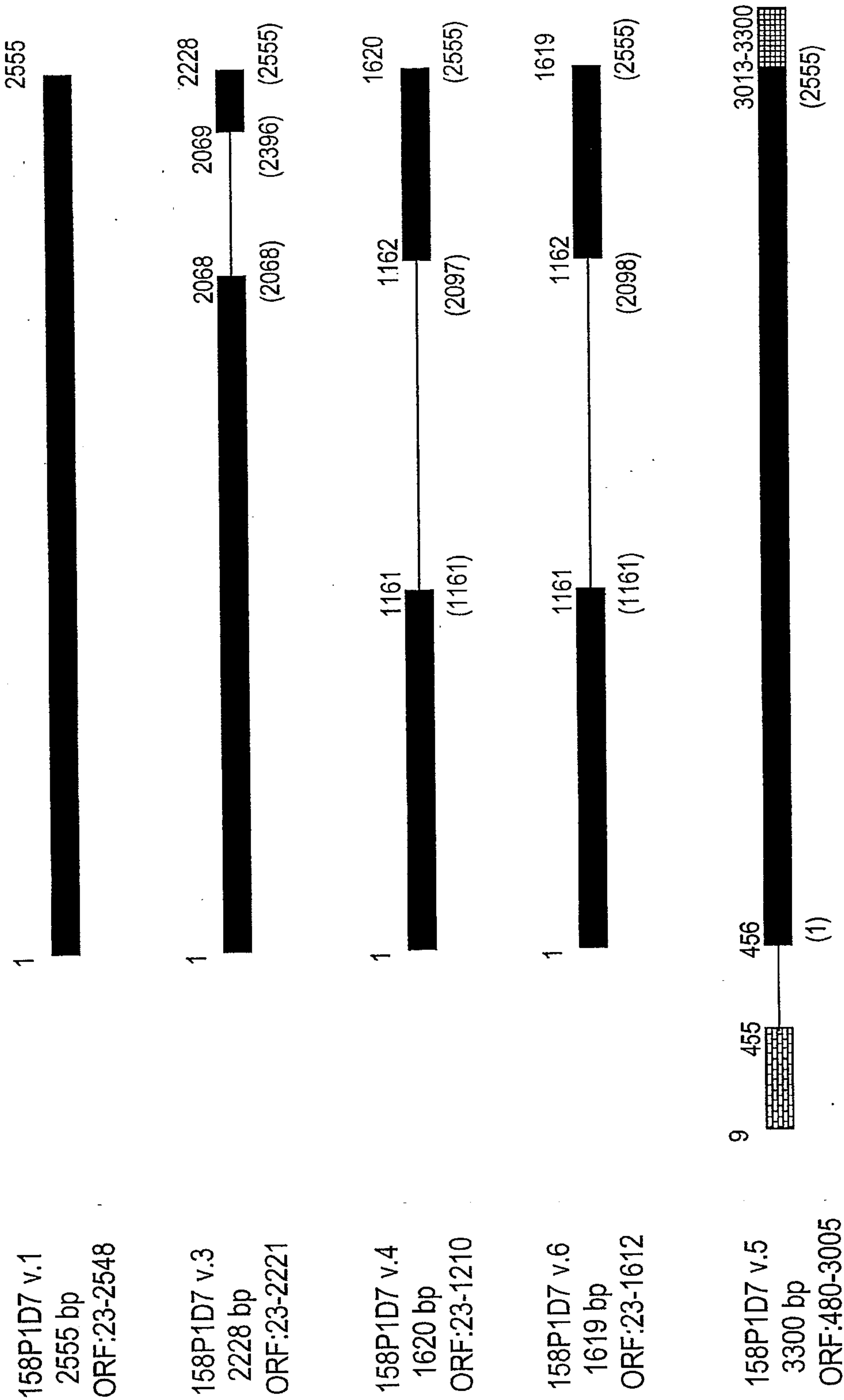


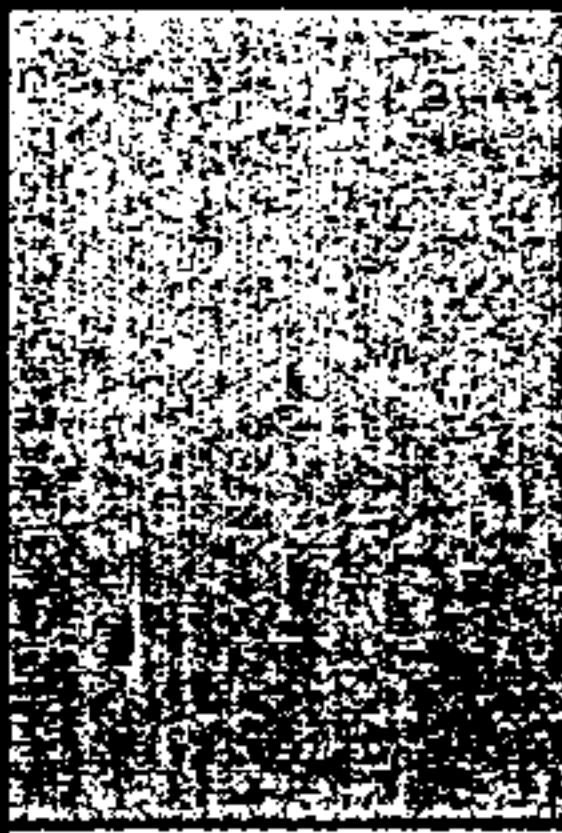
Figure 20 158P1D7 Expression in Melanoma

1 2



1. Normal skin cell line - Detroit-551
2. Melanoma cancer cell line - A375

Figure 21 158P1D7 Expression in Cervical Cancer Patient Specimens

Panel#	Patient ID#	Diagnosis	Grade	Stage	Cervix CA
1	Normal Cervix	(Ambion)			
2	HeLa	Cell Line			
3	USA-00281-D01	Intraepithelial neoplasia	2-3	T3AN0MX	
4	VNM-00266	AdenoCA	1	IIA	
5	VNM-00376	AdenoCA	1	IIA	
6	IND-00396	Mucinous AdenoCA	2	IIB	
7	A0098	Adenosquamous	2B	T2bNXM0	

No expression

Positive expression

Figure 22: Detection of AGS15 protein in recombinant cells
with monoclonal antibodies

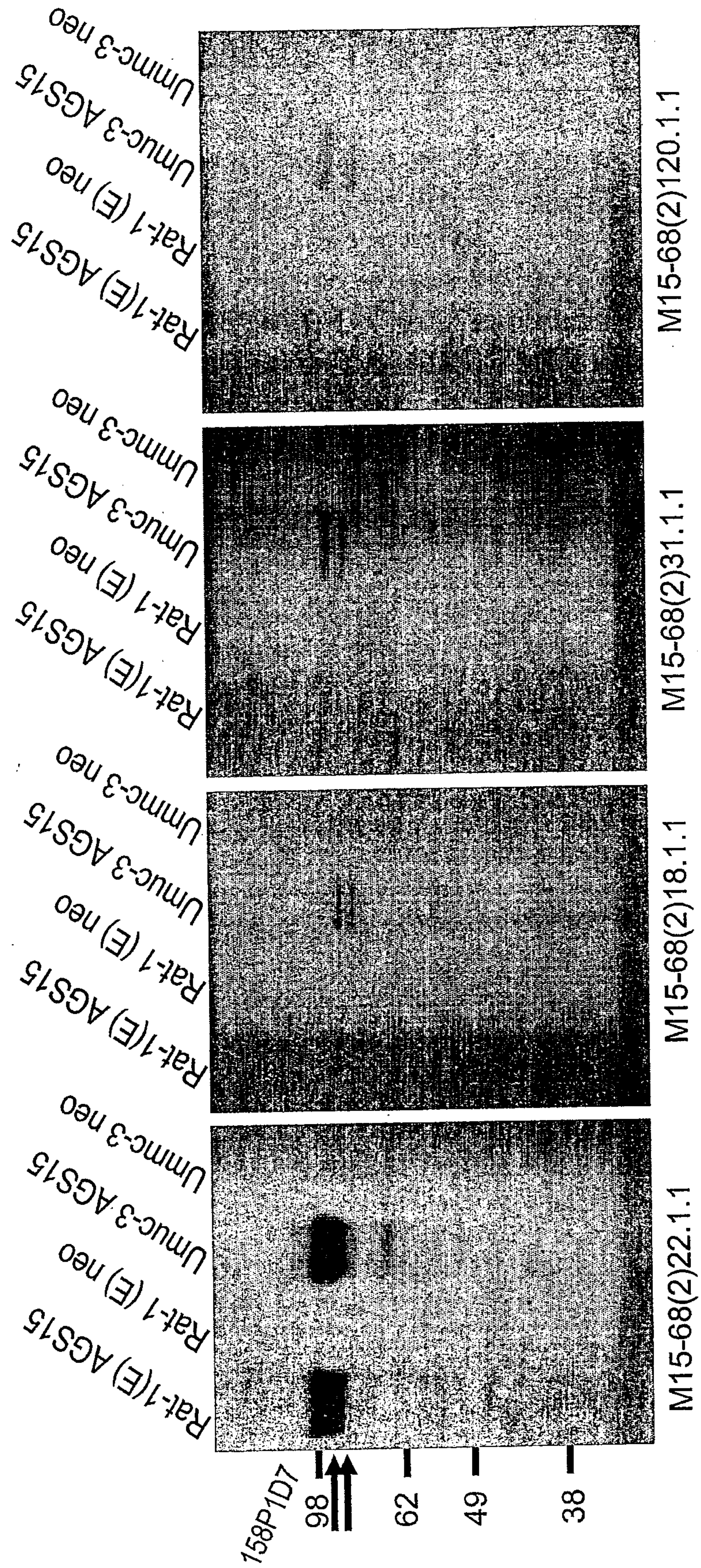


Figure 23: Surface staining of AGS15-expressing 293T and UMUC cells with anti-AGS15 monoclonal antibodies

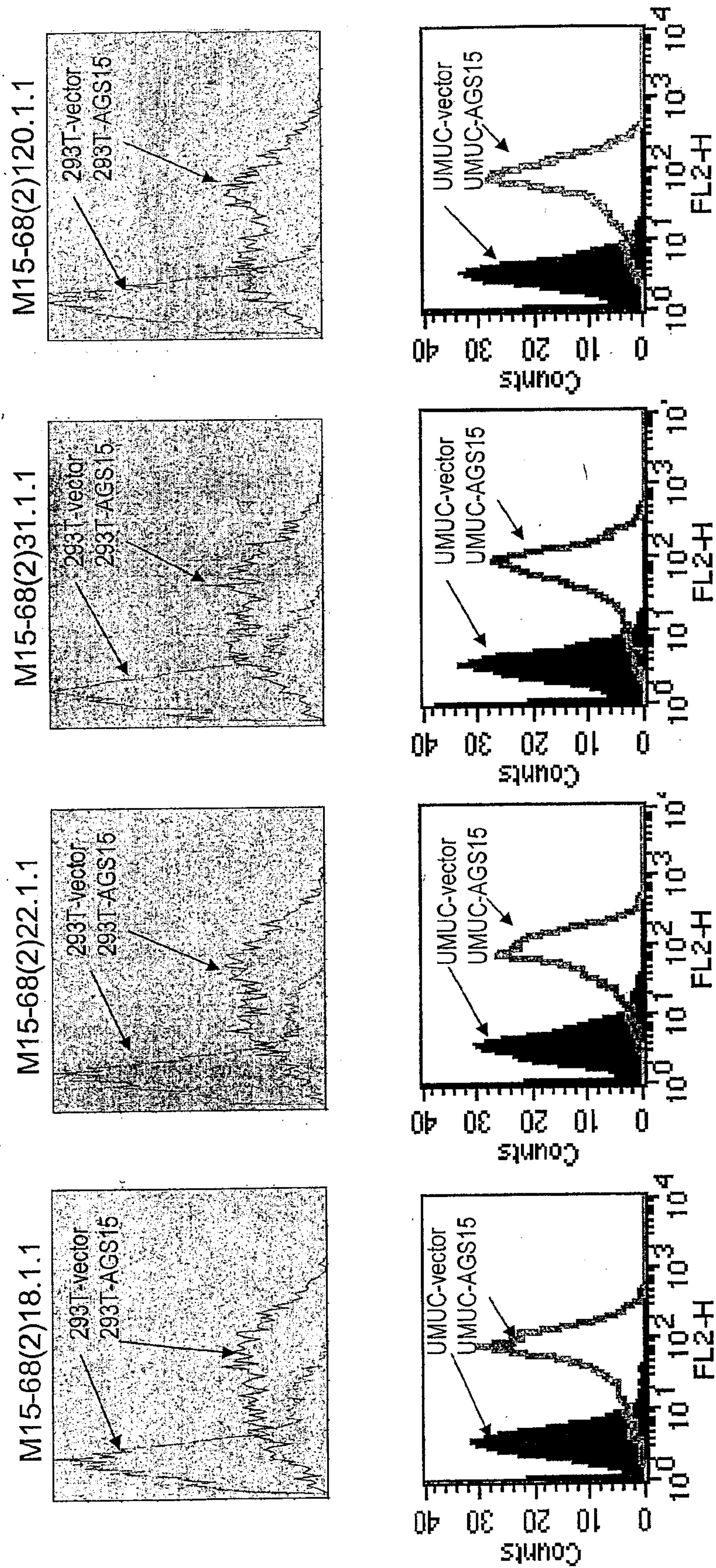


Figure 24: Surface staining of endogenous AGS15-expressing LAPC9 prostate cancer and UGB1 bladder cancer xenograft cells with MAb M15-68(2)22.1.1

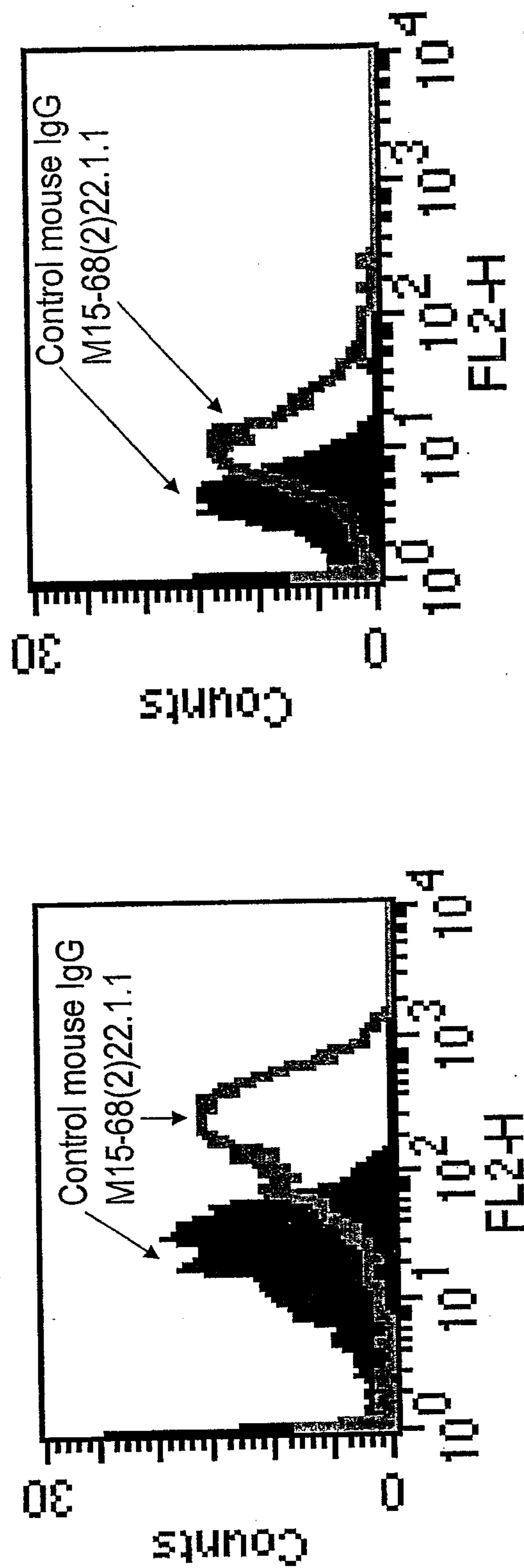


Figure 25: Monoclonal antibody-mediated internalization of endogenous surface 158P1D7 in NCI-H146 small cell lung cancer cells

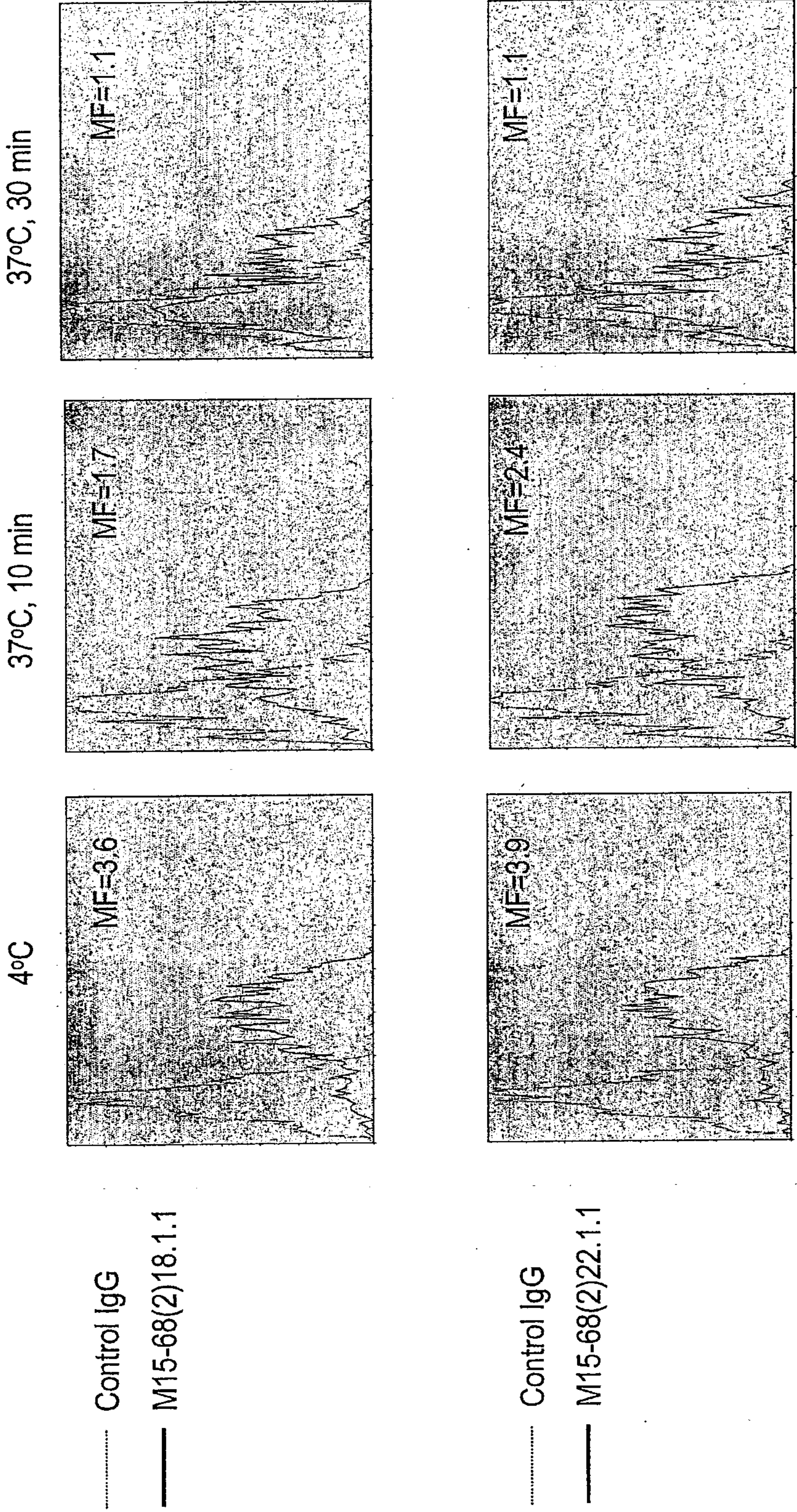


Figure 25 Cont.

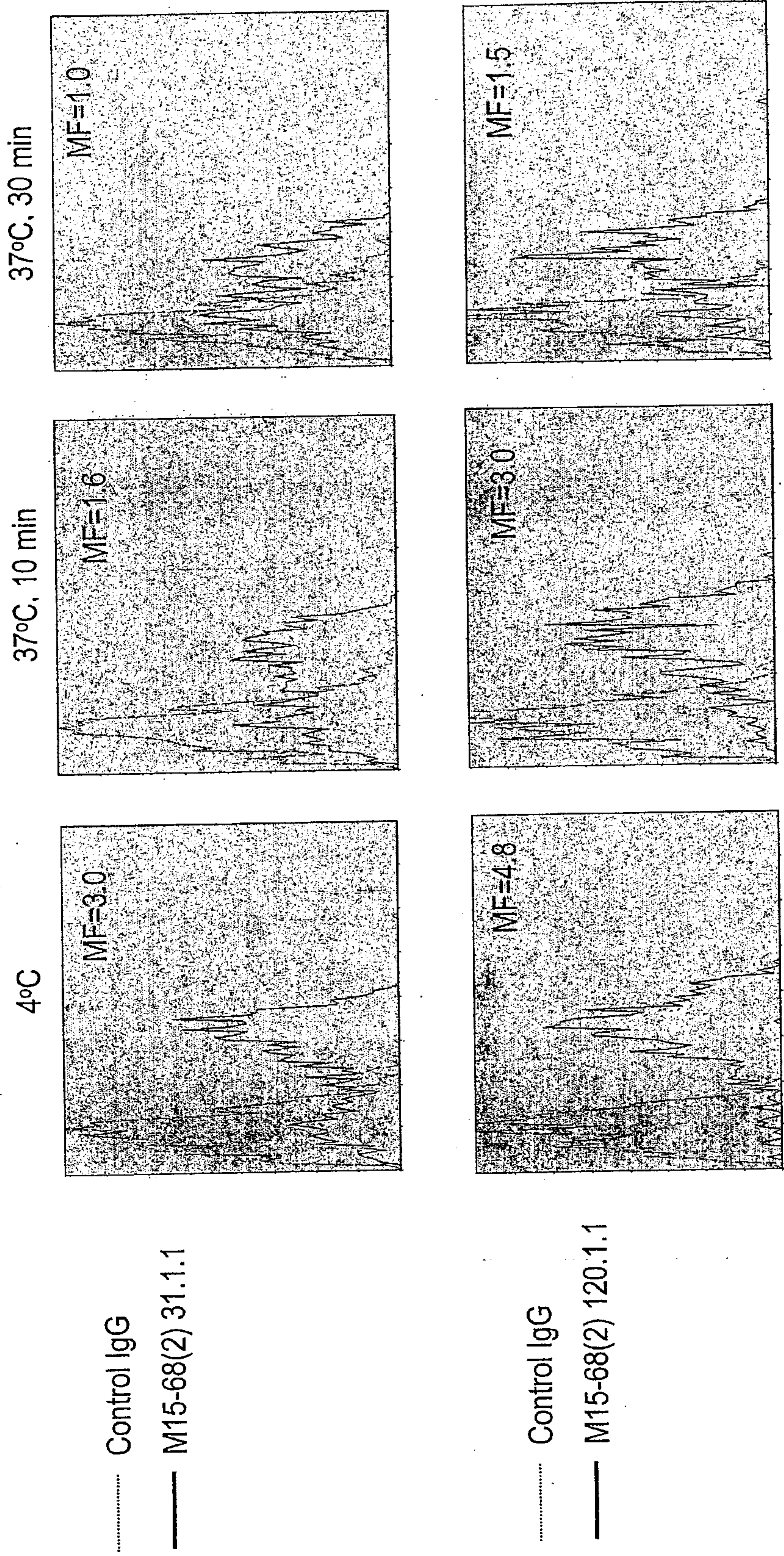


Figure 26: Binding of the 158P1D7 extracellular domain to human umbilical vein endothelial cells (HUVEC)

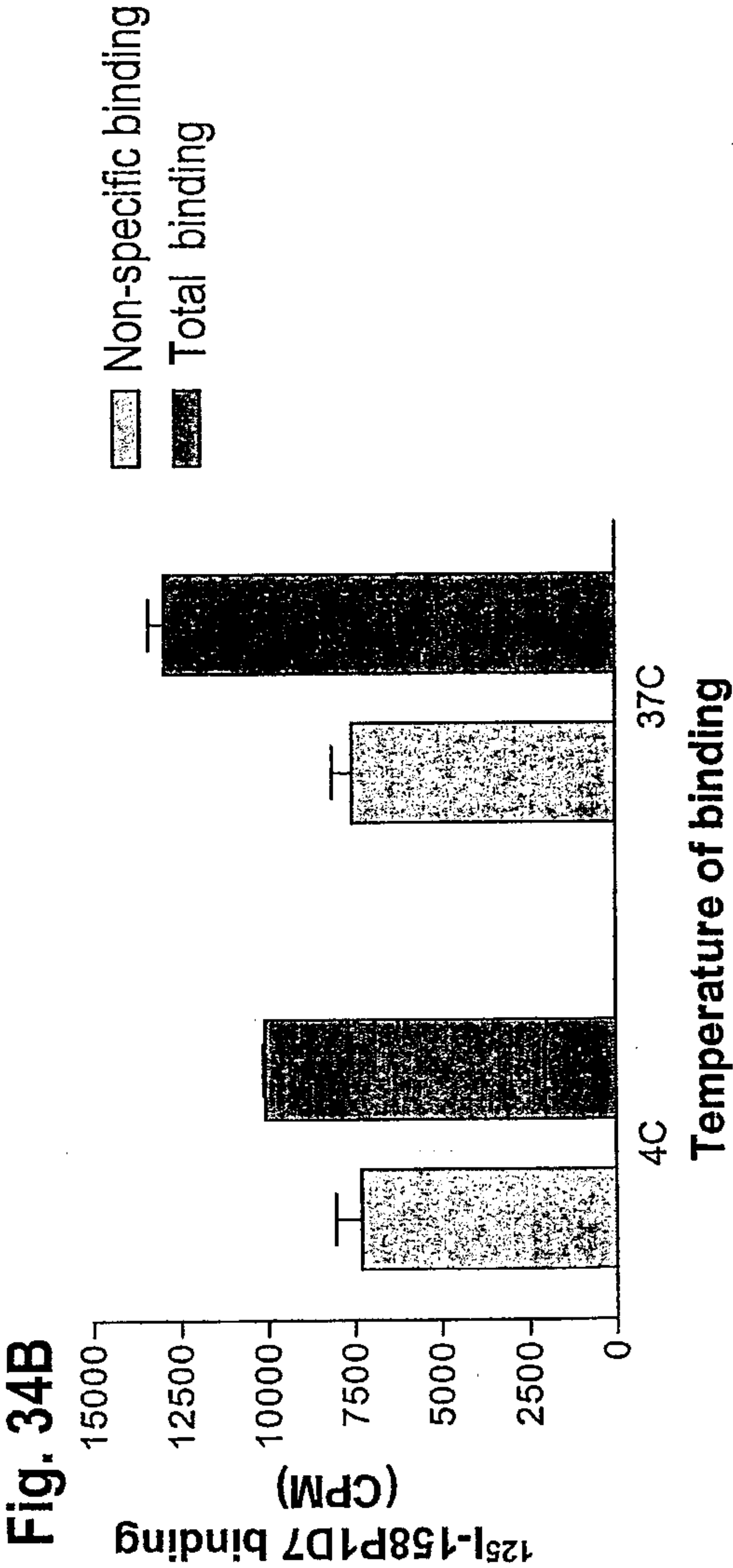
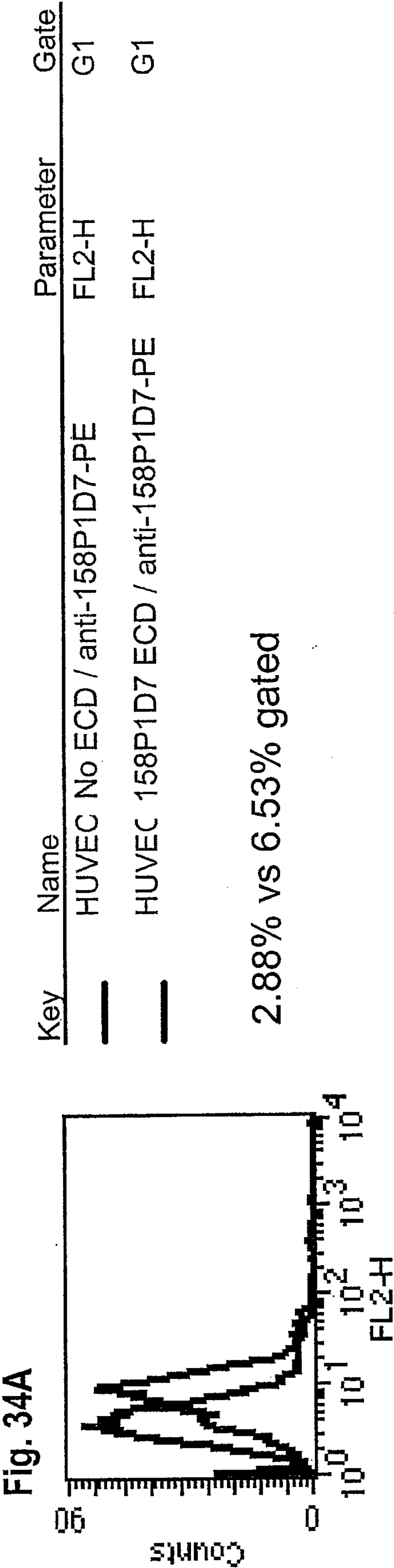


Figure 27 - 158P1D7 Enhances the Growth of Bladder Cancer UM-UC-3 Cells in Mice

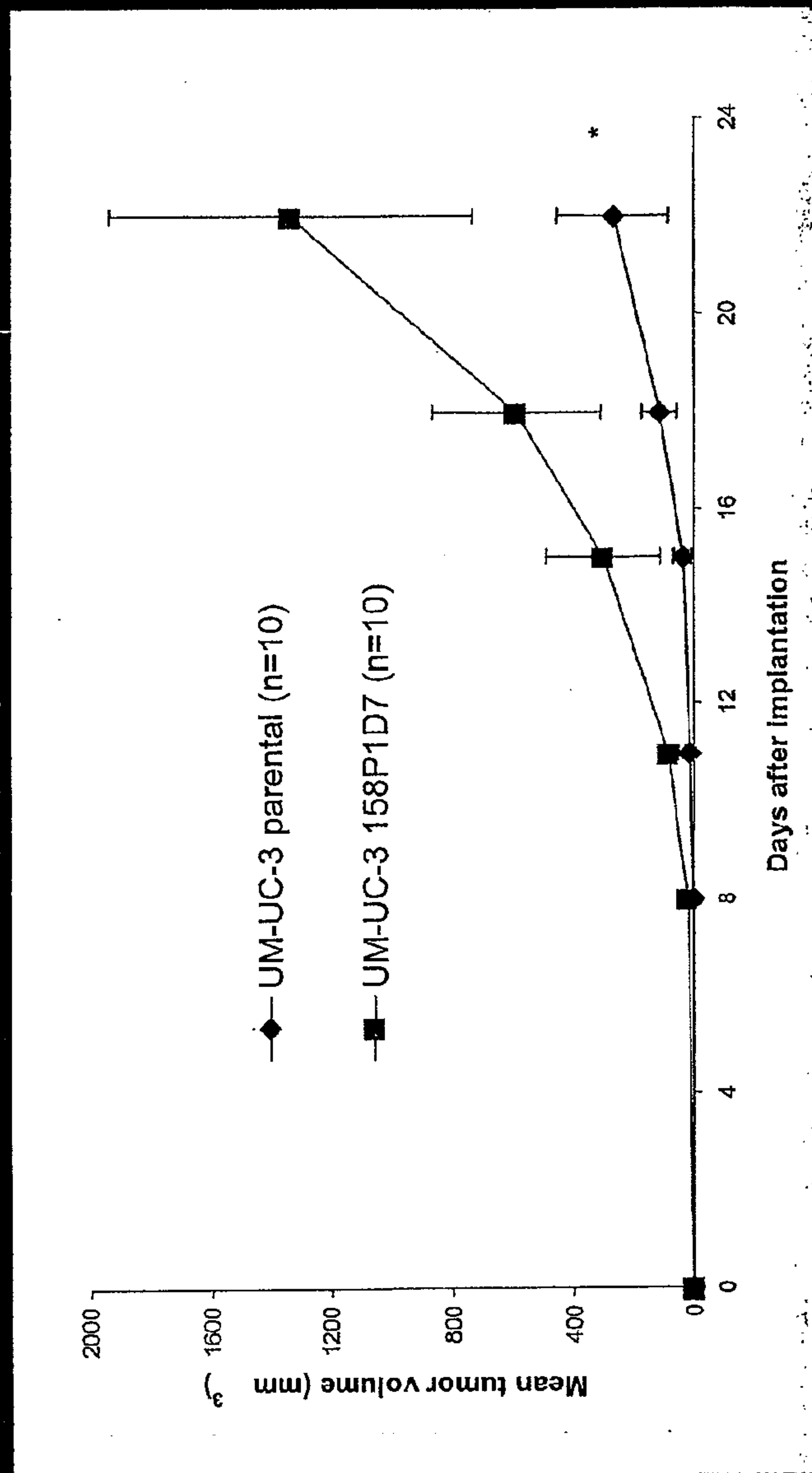
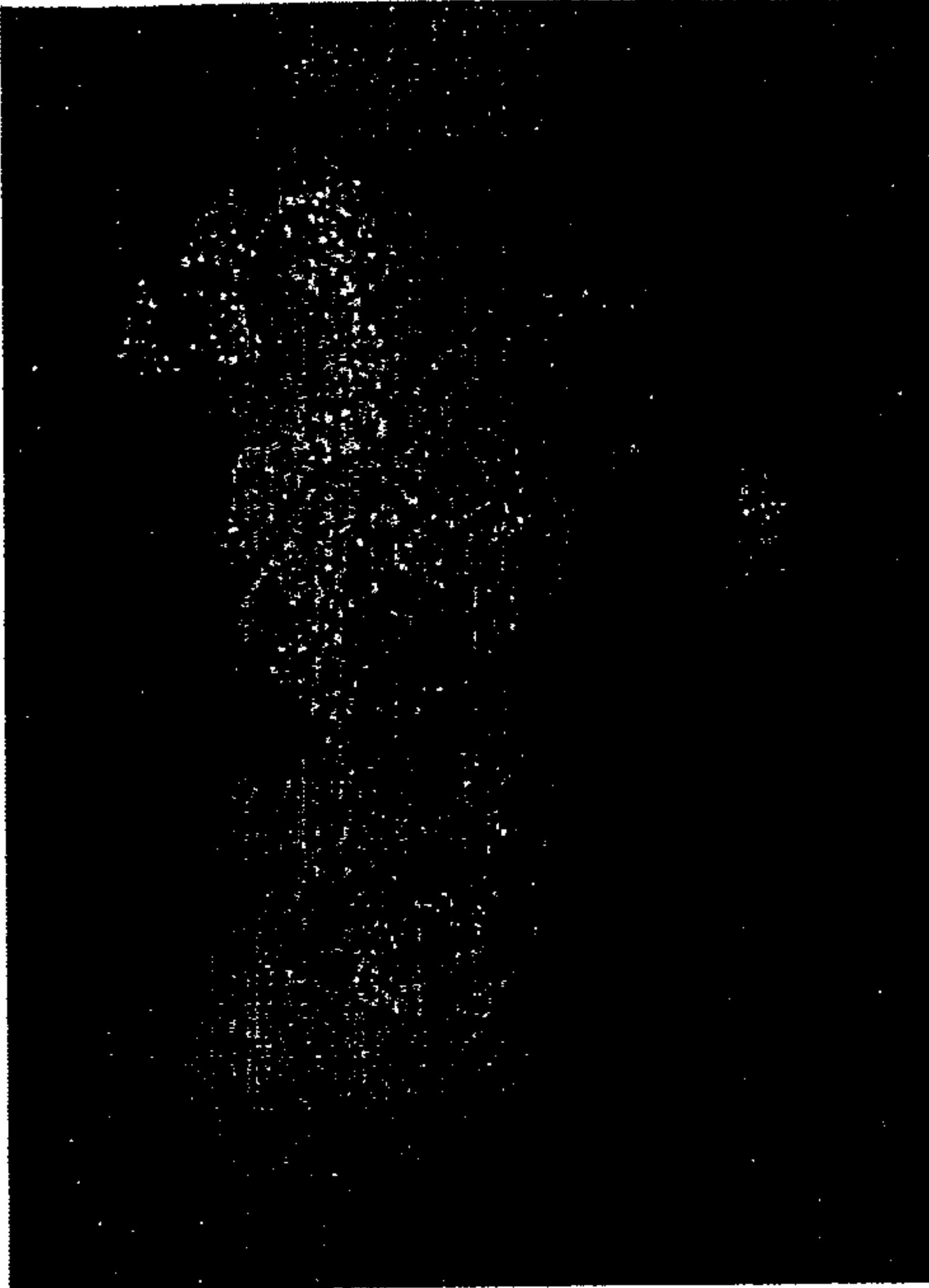
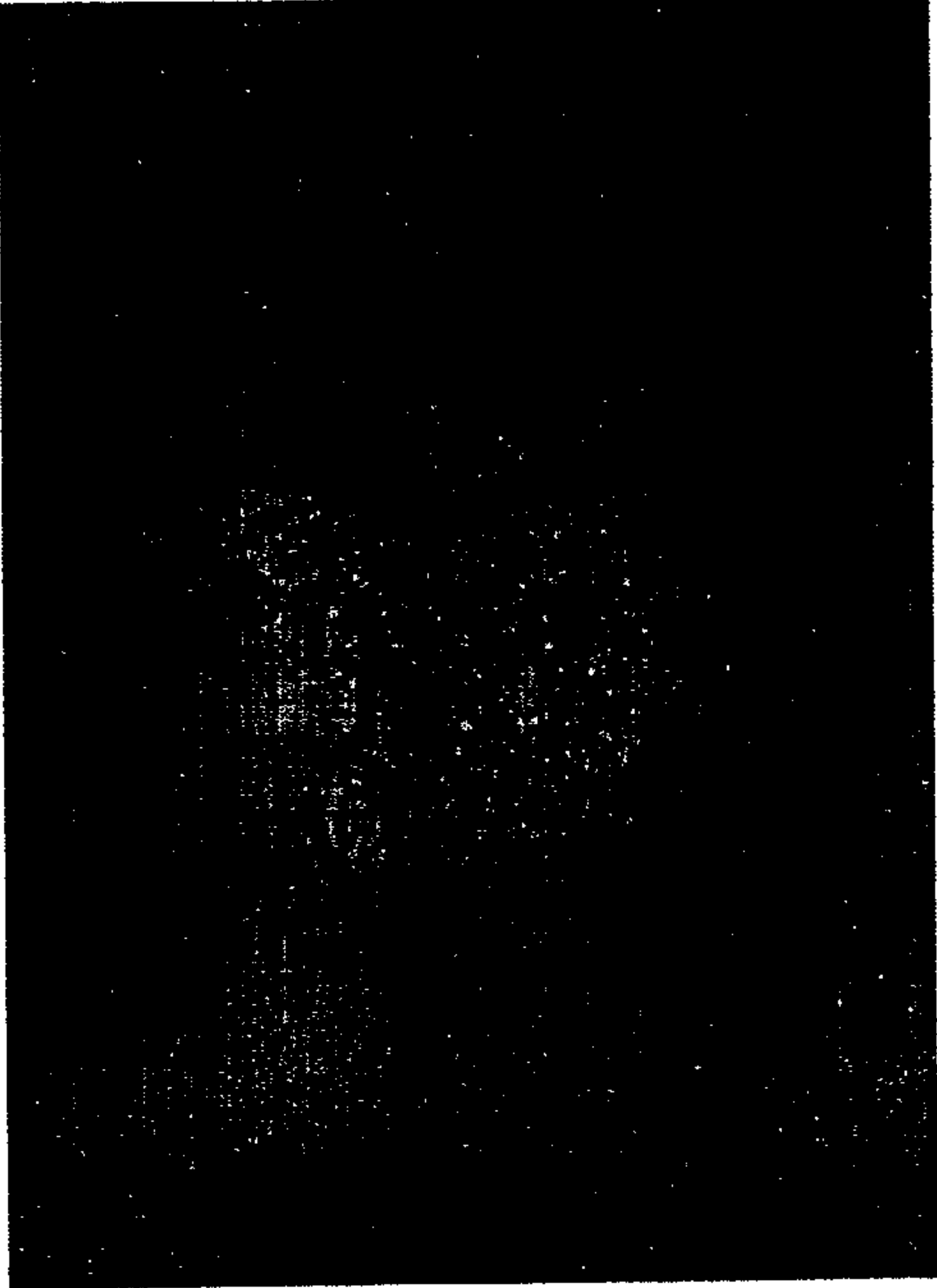
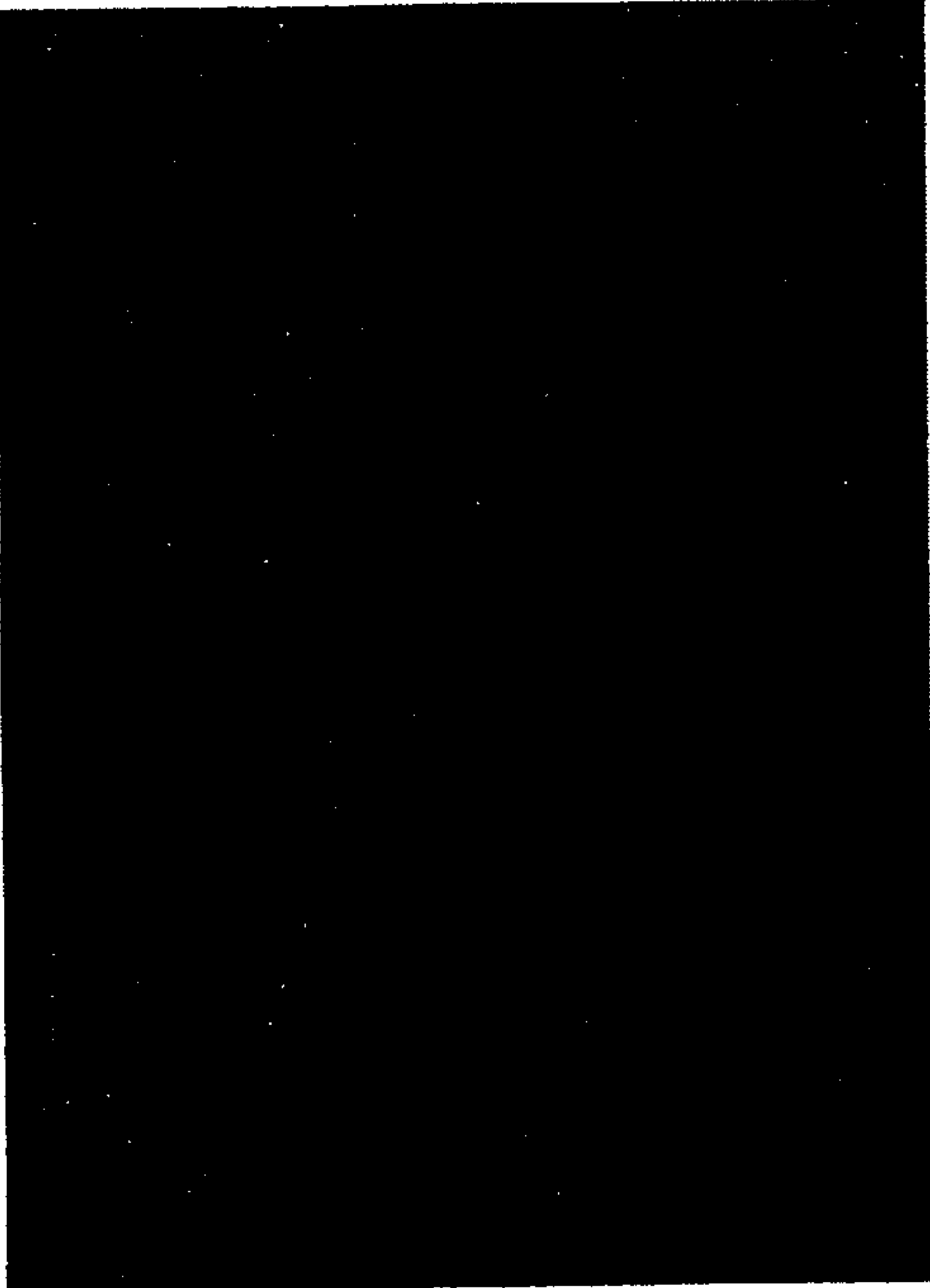
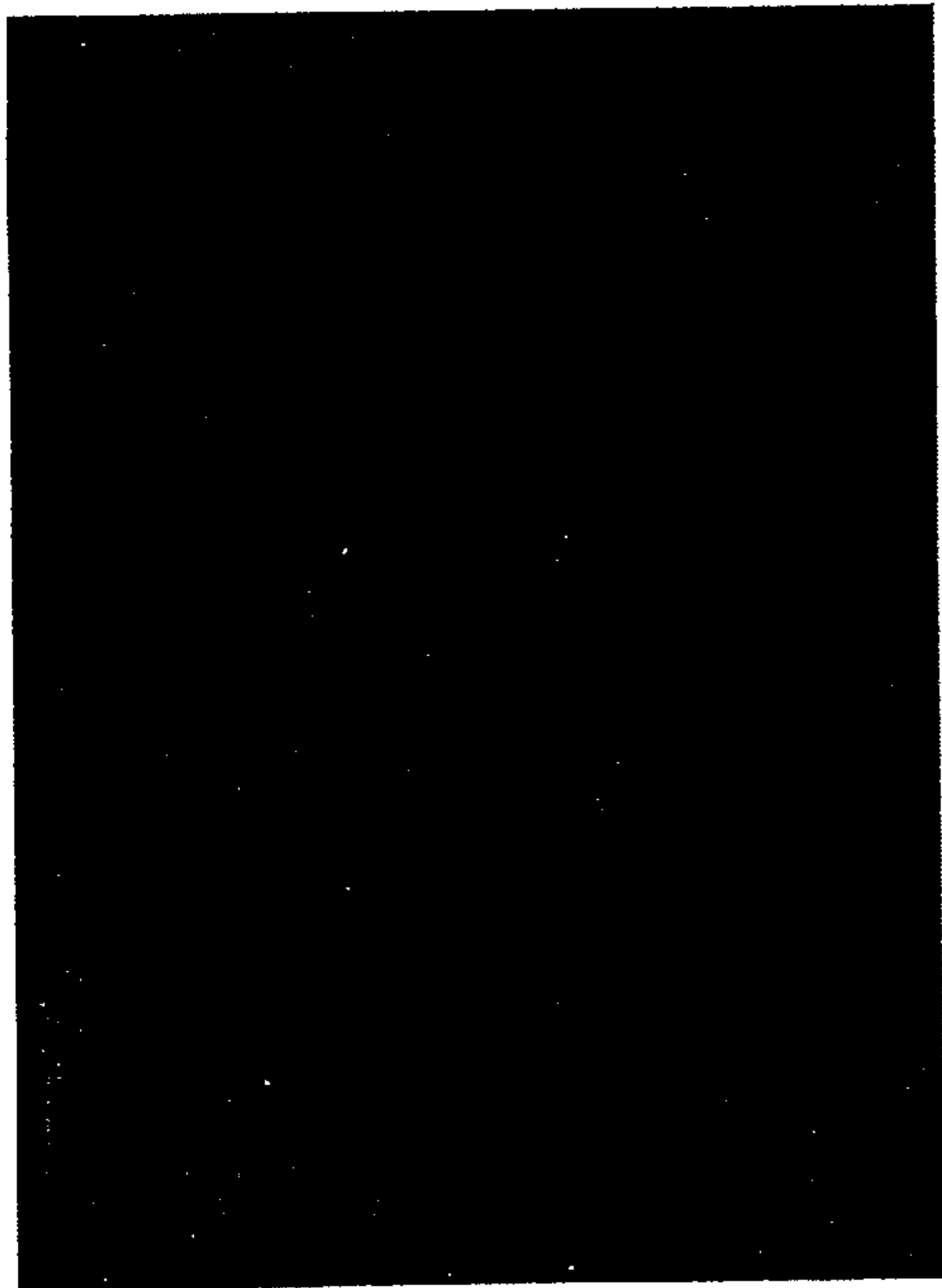


Figure 28: Internalization of MAb M15-68(2).31.1.1 in NCI-H146 cells

4C

4C → 37C for 30 min



Fluorescence

Bright Field

Figure 29: Effect of 158P1D7 RNAi on cell survival

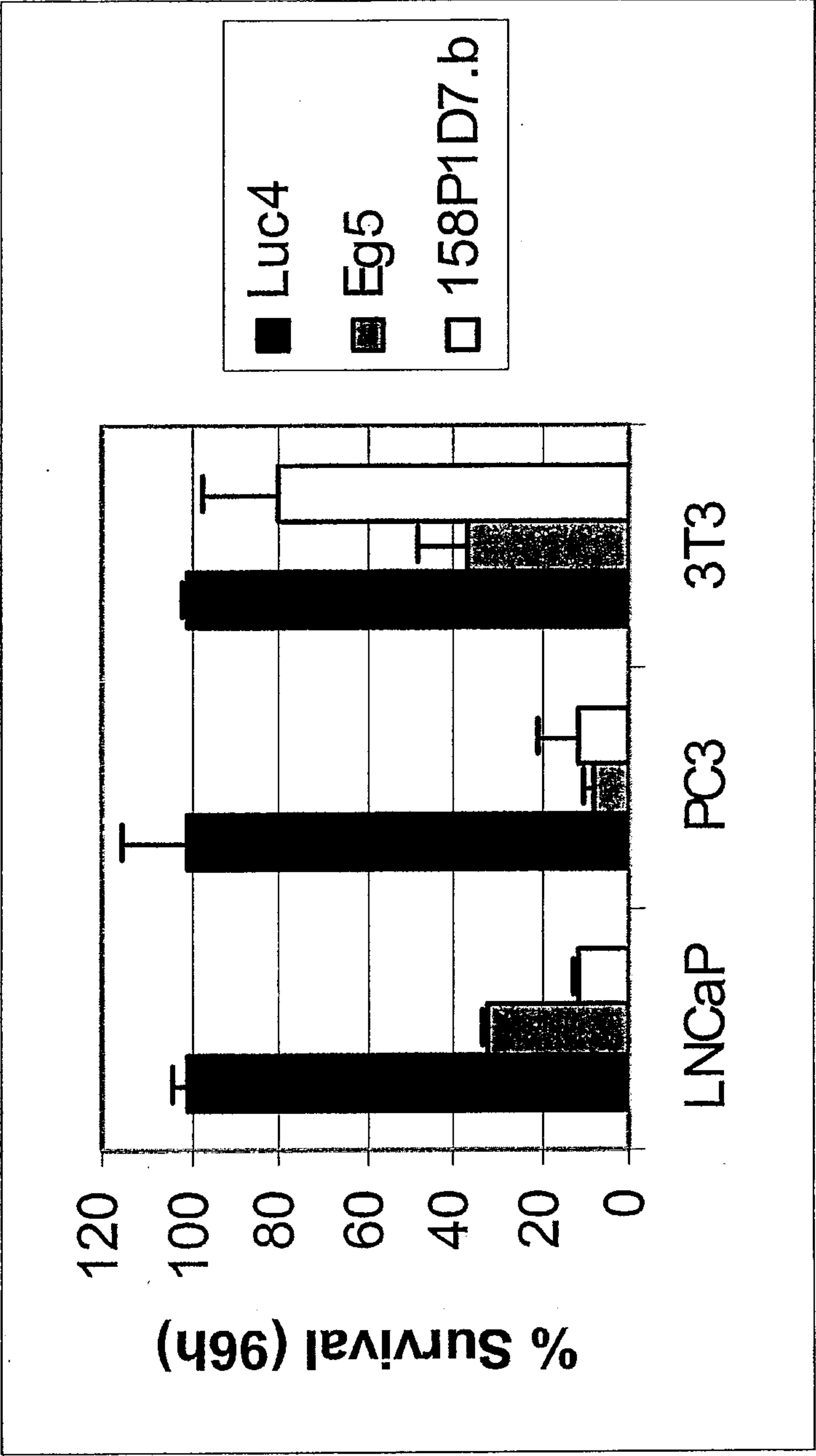


Figure 30 - 158P1D7 MAbs Retard the Growth of Human Bladder Cancer Xenografts in Mice

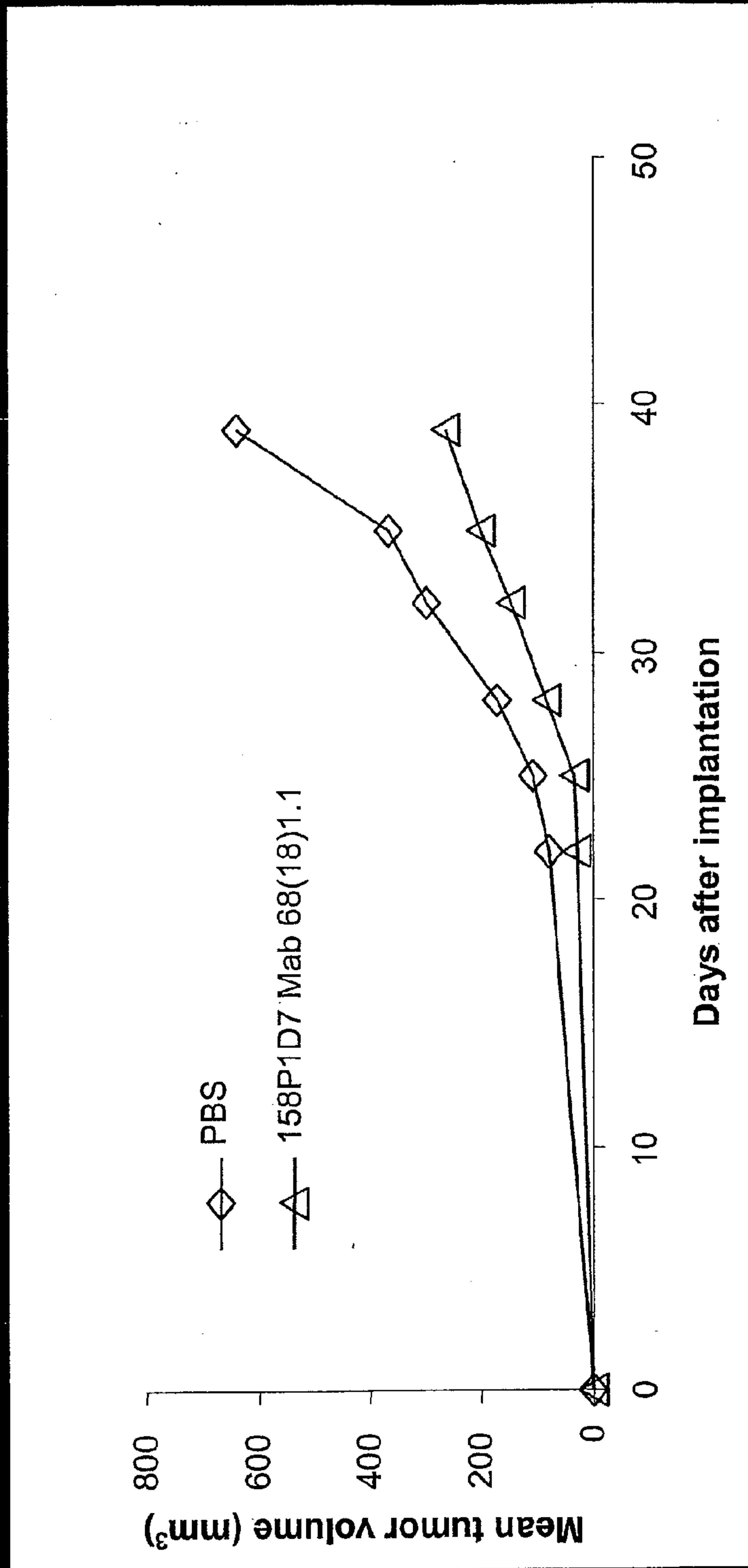


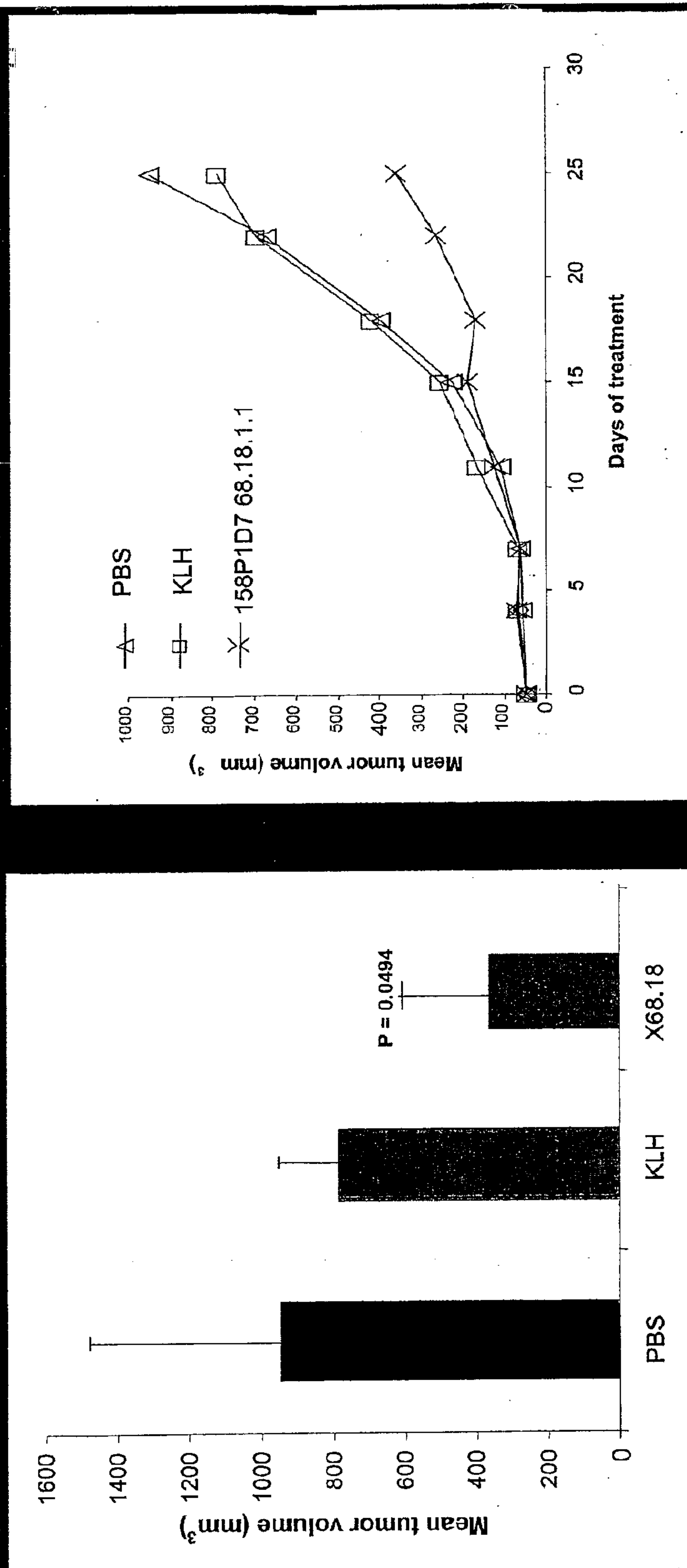
Figure 31 - 158P1D7 MAbs Retard Growth of Human Prostate Cancer Xenografts in Mice**Patient-derived LAPC9 xenograft**

Figure 32: Effect of 158P1D7 on Proliferation of Rat1 cells

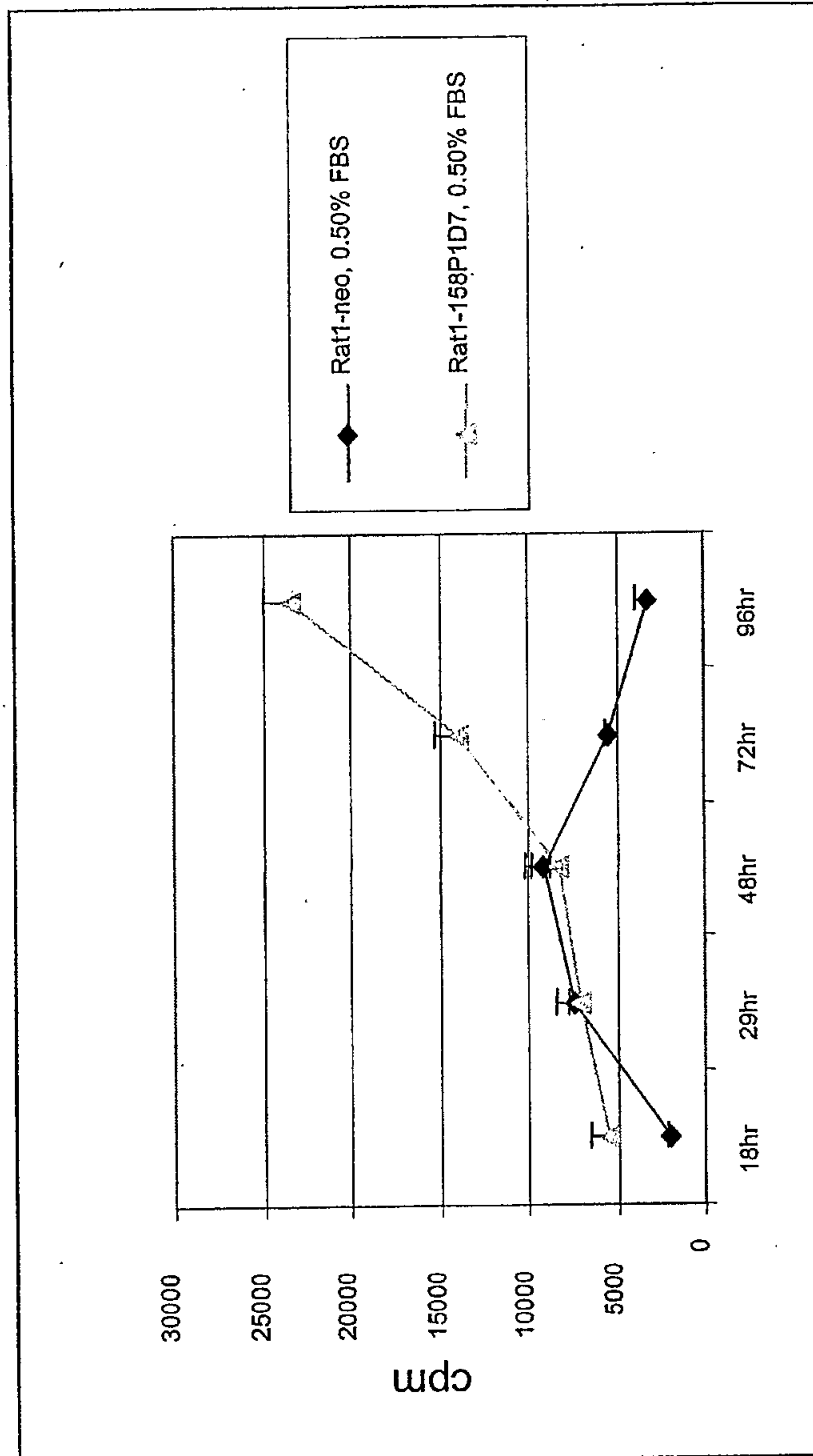


Figure 33: 158P1D7 Enhances Entry Into the S Phase

Cells	Treatment	Percent Cells		
		G1	S	G2
3T3	0.5% FBS	92.7	2.6	2.2
	10% FBS	72.8	11.4	14.7
3T3-neo	0.5% FBS	95.1	1.4	2.3
	10% FBS	59.6	14.1	18.3
3T3-158P1D7	0.5% FBS	90.1	3.3	4.4
	10% FBS	68.4	21.2	1.7

Figure 34A. The cDNA and amino acid sequence of M15/X68(2)18 VH clone #1.

```

1 Q T A G V R S W P G G A L T E P V H H M
1 caaactgcaggagtcaggagttggcctggtggcgccctcacagagcctgtccatcacatg
21 H R L R I L I D R L W C K L G S P A S R
61 caccgtctcaggattctcattgaccggctatggtgttaaactgggttcgccagcctccagg
41 K G S G V A G N D L G R W K H R L Y F S
121 aaagggctctgggggtggctgggaatgatttggggcgatggaagcacagattatacttcagc
61 S P I Q T E H Q E G Q F K S Q T F L K N
181 tctccaatccagactgagcatcaggaaggacaattcaagagccaaactttcttaaaaaat
81 N S L Q T D D T A R Y Y C A R D E G R G
241 aacagtctgcaaactgatgacacagccaggtattactgtgccagagatgaagggagggga
101 L C L I A G A K G P R S P S P
301 ctctgtttgattgctggggccaagggaccacgggtcaccgtctcctca

```

Figure 34B. The cDNA and amino acid sequence of M15/X68(2)18 VL clone #2.

```

1 D I Q L T Q S P A S L A V S L G Q R A T
1 gacattcagctgacccagtctcctgcttccttagctgtatctctggggcagagggccacc
21 I S Y R A S K S V S T S G Y S Y M H W N
61 atctcatacagggccagcaaaagtgtcagtacatctggctatagttatatgcactggaac
41 Q Q K P G Q P P R L L I Y L V S N L E S
121 caacagaaaccaggacagccacccagactcctcatctatcttgtatccaacctagaatct
61 G V P A R F S G S G S G T D F T L N I H
181 ggggtccctgccaggttcagtggcagtggttctgggacagacttcaccctcaacatccat
81 P V E E E D A A T Y Y C Q H I R E L T R
241 cctgtggaggaggaggatgctgcaacctattactgtcagcacattagggagcttacacgt
101 S E G G P S W R S N
301 tcggaggggggaccaagctggagatctaac

```

Figure 35A: The amino acid sequence of M15/X68(2)18 VH clone #1.

1 QTAGVRSWPG GALTEPVHHM HRLRILIDRL WCKLGSPASR KGSGVAGNDL
51 GRWKHRLYFS SPIQTEHQEG QFKSQTFLLKN NSLQTDDTAR YYCARDEGRG
101 LCLIAGAKGP RSPSP

Figure 35B: The amino acid sequence of M15/X68(2)18 VL clone #2.

1 DIQLTQSPAS LAVSLGQRAT ISYRASKSVS TSGYSYMHWN QOKPGQPPRL
51 LIYLVSNLES GVPARFSGSG SGTDFTLNIH PVEEEDAATY YCQHIRELTR
101 SEGGPSWRSN

Figure 36: Detection of 158P1D7 protein by immunohistochemistry in various cancer patient specimens.

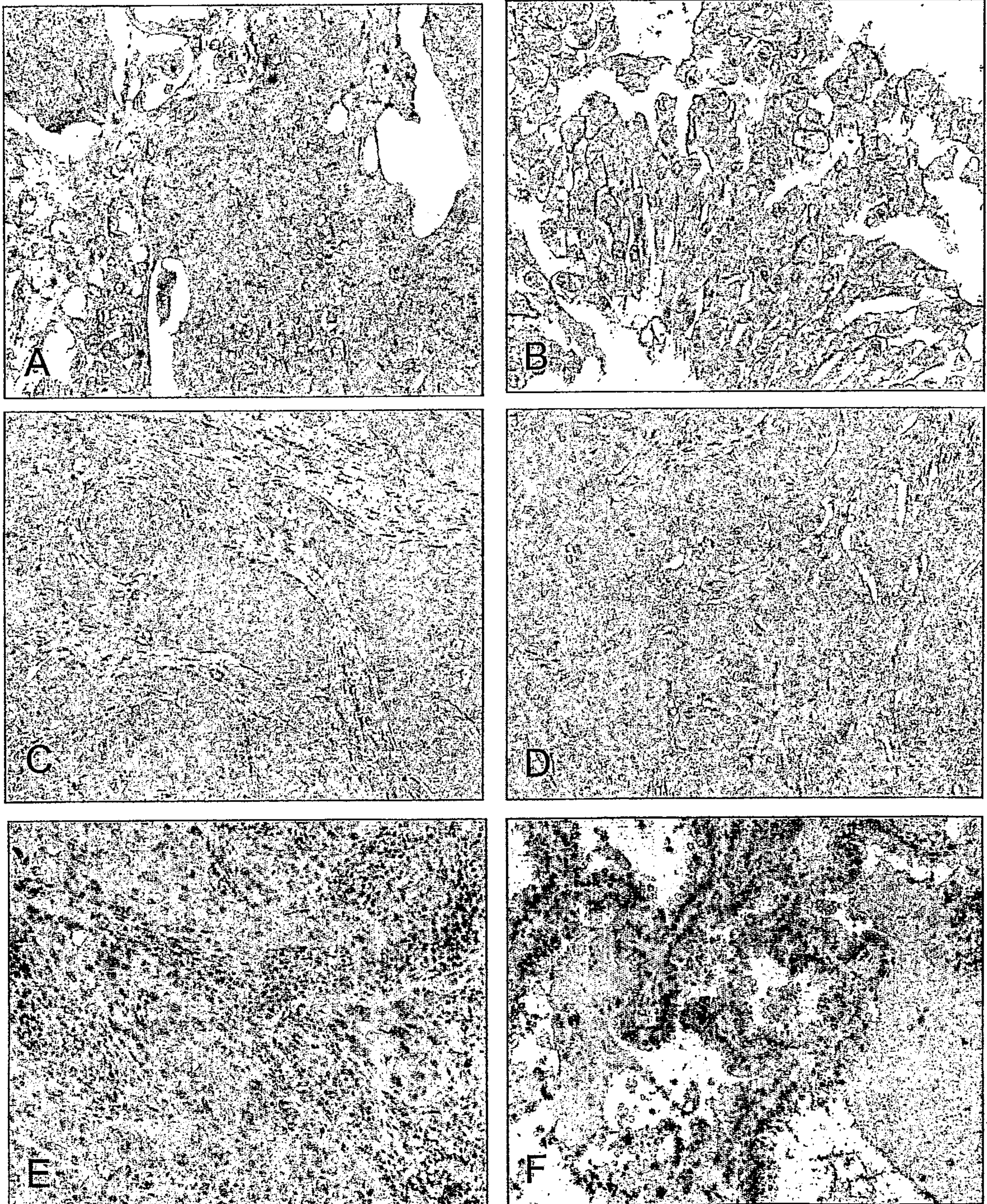


Figure 1: 158P1D7 SSH_sequence (SEQ ID NO: 1) .

```

      1  GATCTGATAA  GCTTTCAATG  TTGCGCTCCT  GACAATGTAT  TAGAAGTCCT  GATGGGGATA
     61  GGACTTTGCA  GTTACAAGGA  ATAGGGGCAGA  AAGGTCCTGG  AAGTTGAGTG  GATGGCTTTG
    121  TAATATAAGG  TATCAAACCT  GGTGCTTTGG  TGGGTAGTTT  TAGAATGGAC  GTGGTCTTAG
    181  TTGACATGCG  ACTATCATTT  ATTGAAGATG  TTGCTGCCAG  ATGTAATGAT  C

```