

710
TTC CAG GGT GCG CTC CTC CTC TCA GAT CCC CCA CCC TGT AGG CTG AGC TGG GAC GTG CAG GCC TTT TTG TTT TTT GAG ACA AGT CTC GCT CTG
Phe Gln G
 TCA CCC AGG GTG GAG GCG ACT GGC GCG ATC TGG TCT CA.....intron 7 (~1.0kb).....
C CTC CTC TCC AAC TGG TCC CCG TCG ACG TGA CCC CTG ACC CTC TGC TGG GTT TCA GCT
Ty
 720 730 740 750 760 770 780 790 800 810
TTC TGG AGG AAC AAG TTT GAC CCG AGC CTT ACC CAG AGA GAC TCC TTC CAC CTG GAC GAG CAG TTC ACG GTG CCC GTG GAA ATG ATG CAG GCC CCG
Phe Trp Arg Asn Lys Phe Asp Pro Ser Leu Thr Gln Arg Asp Ser Phe His Leu Asp Gln Cln Phe Thr Val Pro Val Gln Met Met Gln Ala Arg
 210 220 230
 820 830 840 850
ACG TAC CCG CTG GCG TGG TTG TTG CTG GAG CAG CCT GAG ATC CAG GTC ACC CTT GGT TCT CCA GCA GCG TGC C.....intron 8 (~3.0kb).....
Thr Tyr Pro Leu Arg Trp Phe Leu Leu Gln Gln Pro Gln Ile Cln
TG CCT TAG CAG CAC CTG CTG GCC CCA CCC CCA CTT AGC TTC GGG CCT TTC TGT CCT CAT GCT CTT CCC TTC CCT TTT CTG TAG
 860 870 880 890 900 910 920 930 940 950
GTC GCT CAT TTC CCC TTT AAG AAC AAC ATG AGC TTT GTG GTC CTT GTA CCC ACC CAC TTT GAA TGG AAC GTG TCC CAG GTA CTG GCC AAC CTG AGT
Val Ala His Phe Pro Phe Lys Asn Asn Met Ser Phe Val Val Leu Val Pro Thr His Phe Gln Trp Asn Val Ser Gln Val Leu Ala Asn Leu Ser
 250 260 270
 960 970 980 990 1000 1010 1020 1030 1040 1050
TGG GAC ACC CTG CAC CCA CCT CTG GTG TGG GAG AGC CCC ACC AAG CTC GGG CTG CCT AAG CTG TAT CTG AAA CAC CAA ATG GAC CTG GTG GCC ACC
Trp Asp Thr Leu His Pro Pro Leu Val Trp Gln Arg Pro Thr Lys Val Arg Leu Pro Lys Leu Tyr Leu Lys His Gln Met Asp Leu Val Ala Thr
 280 290 300 310
 1060
CTC AGC CAG CTG GGT AAG GAG GAG GGT GCG GCG GAG CCC CGA GGT CAG GCT GGG CAG GCG GGG TAA.....intron 9 (~1.0kb).....TAG GAA
Leu Ser Gln Leu G
 TGA AGC GGT ATC TGT GAG TTC AAG CTG TTC CCT GGC CAG GAT CTC AGA CAC CCT CCA AAG CAC CTC CAG GAG CCT GTG ACC CCA AGG GCA GCT CTG
 1070 1080 1090 1100 1110 1120 1130
ACC ACG CAT CTC TGG CCC TGG GCA GGC CTG CAG GAG TTG TTC CAG GCC CCA GAC CTG CCG GGG ATC TCC GAG CAG ACC CTG GTG GTG TCC GGC GTG
Ty Leu Gln Gln Leu Phe Gln Ala Pro Asp Leu Arg Gln Ile Ser Gln Cln Ser Leu Val Val Ser Gln Val
 320 330 340 350 360 370
 1140 1150 1160 1170 1180 1190 1200 1210 1220 1230
CAG CAT CAG TCC ACC CTG CAG CTC ACC GAG GTC GGC GTG GAG GCG GCG GCC ACC AGC ATT GCC ATG TCC CCG ATG TCC CTG TCC TCC TTC CCG
Gln His Gln Ser Thr Leu Gln Leu Ser Gln Val Gln Val Gln Ala Ala Ala Ala Thr Ser Ile Ala Met Ser Arg Met Ser Leu Ser Ser Phe Ser
 340 350 360 370
 1240 1250 1260 1270 1280 1290 1300 1310 1320 1330
GTG AAC CCG CCC TTC CTC TTC TTC ATC TTC GAG GAC ACC ACA GGC CTT CTT CCC TTC GTG GCG AGC GTG AGG AAC CCC AAC CCC AGT GCA CCG CCG
Val Asn Arg Pro Phe Leu Phe Phe Ile Phe Gln Asp Thr Thr Gln Leu Pro Leu Phe Val Gln Ser Val Arg Asn Pro Asn Pro Ser Ala Pro Arg
 380 390 400 410 420 430 440 450
 1330 1340 1350 1360 1370 1380 1390 1400 1410 1420
GAG CTC AAG GAA CAG CAG GAT TCC CCG GGC AAC AAG GAC TTC CTC CAG AGC CTG AAA GGC TTC CCC CCG GGA GAC AAG CTT TTC GGC CCT GAC TTA
Gln Leu Lys Gln Gln Gln Asp Ser Pro Gln Asn Lys Asp Phe Leu Gln Ser Leu Lys Gln Phe Pro Arg Gln Asp Lys Leu Phe Gln Pro Asp Leu
 410 420 430 440 450
 1430 1440 1450 1460 1470
AAA CTT GTC CCC CCC ATG GAG GAG GAT TAC CCC CAG TTT GGC AGC CCC AAG TGA GGG GCC GTG GCT GTG GCA TCC AGA GTC CCT GCC TGG ACC AGC
Lys Leu Val Pro Pro Met Gln Gln Asp Tyr Pro Gln Phe Gln Ser Pro Lys
 440 450
CTC TCC ACT CAT GTG ACT CTT TCC AAC CCG CTT TGT GGC ACT GGC GCA GGG GCC GGG GGC AGT CTG AGA GAG GCC ATT CTT TCC CAA CAC CTC TTG 1614
GGG AGT TTA GGG TGG GGG GGG GCG CCG CTG GGA GGA GGG CAG GCA TCG GGG AGC CCG GAG CCT GAC CCT CAT CTT TCT TCC AAA CAG GCT CAG AGG 1710
GTG TCC TGC ACC GGG GCC TGG GCA GGA GGG AGG TGC TTC TAG TTC TGC CAG GAG ACA GGT TAG CTG CTC CCC ACG TCA GCT GGG ACA CCC CGA CTT 1806
TTG TTT ACC AGA GAA AAA GGG AGG GGG AGA GGG CTG CCT TTG GAC TTG TCC CCG GAC ACC TAG GCT AGG GTG GGG AGA GAC GGG CCC TGG TGG TGG 1902
CTC GGG AGG CGA AGC GTT GTC CTC AGC CCC GCG TGG AAC TCG TGT CTG GCA CAG CCT GGC TGT GGC CTA ACC TGC CGA GAG TCC ATC AGC CTC CAT 1998
CCT ACC CCC TGT GGC TTG TCA CCG CAG ACT TCC CAC GGC TCC TCG AGA TCC CAA CAC TGC CAG CAT TTG CCT TCC CTC TCC TGT CTT CCT CTT 2094
CTG CCC GGG AGC TCA GGA ACC CAG GCA GGG AAG GAT CCC ATG AGC TCC TTA AGG CTC TTT TGT AAG GTT TTT GTA GTG ATT TTT ATG CCA CCT CTA 2190
TAA TGA ATG AAT GGG CCT GGC TGG TTT GAT GTC ACC GTT CTG GG 2234

FIG. 2. Nucleic acid sequence of the α_2 -plasmin inhibitor gene. Exons are underlined with solid lines. Bases in the exons and the 5'- and 3'-flanking regions are numbered relative to the translation initiation site. Amino acids are numbered from the NH₂-terminal residue in the plasma protein. Regions corresponding to a potential TATA box, the GC boxes, a potential transcriptional start site (-22), and a polyadenylation recognition site (2189-2194) are boxed. The direct repeats of CCAAT box-like sequence are indicated by dots. G + C-rich sequences are indicated by the dashed underlines. The sequence (-123 to -108), similar to the hepatitis B virus enhancer sequence, is indicated by a wavy line. The sequence (-809 to -800), similar to the human immunodeficiency virus enhancer sequence or κ -immunoglobulin light-chain gene enhancer sequence, is bracketed.