

0 GACGCTCTGTGCCTTCGGAGGTCTTTCTGCCTGCCTGTCCTCATGCCTCTCCTCCTCTTG
M P L L L L

60 CTGCTCCTGCTGCCAAGCCCCTTACACCCCCACCCCATCTGTGAGGTCTCCAAAGTGGCC
L L L L P S P L H P H P I C E V S K V A

120 AGCCACCTAGAAGTGAAGTGTGACAAGAGGAATCTGACAGCGCTGCCTCCAGACCTGCCG
S H L E V N C D K R N L T A L P P D L P

180 AAAGACACAACCATCCTCCACCTGAGTGAGAACCTCCTGTACACCTTCTCCCTGGCAACC
K D T T I L H L S E N L L Y T F S L A T

240 CTGATGCCTTACACTCGCCTCACTCAGCTGAACCTAGATAGGTGCGAGCTCACCAAGCTC
L M P Y T R L T Q L N L D R C E L T K L

300 CAGGTCGATGGGACGCTGCCAGTGCTGGGGACCCTGGATCTATCCCACAATCAGCTGCAA
Q V D G T L P V L G T L D L S H N Q L Q

360 AGCCTGCCCTTGCTAGGGCAGACACTGCCTGCTCTCACCGTCCTGGACGTCTCCTTCAAC
S L P L L G Q T L P A L T V L D V S F N

420 CGGCTGACCTCGCTGCCTCTTGGTGCCCTGCGTGGTCTTGGCGAACTCCAAGAGCTCTAC
R L T S L P L G A L R G L G E L Q E L Y

480 CTGAAAGGCAATGAGCTGAAGACCCTGCCCCAGGGCTCCTGACGCCACACCCAAGCTG
L K G N E L K T L P P G L L T P T P K L

540 GAGAAGCTCAGTCTGGCTAACAACA ACTTGACTGAGCTCCCCGCTGGGCTCCTGAATGGG
E K L S L A N N N L T E L P A G L L N G

600 CTGGAGAATCTCGACACCCTTCTCCTCCAAGAGAACTCGCTGTATAACAATAACCAAAGGGC
L E N L D T L L L Q E N S L Y T I P K G

660 TTTTTGGGTCCCACCTCCTGCCTTTTGCTTTTCTCCACGGGAACCCCTGGTTATGCAAC
F F G S H L L P F A F L H G N P W L C N

720 TGTGAGATCCTCTATTTTCGTCGCTGGCTGCAGGACAATGCTGAAAATGTCTACGTATGG
C E I L Y F R R W L Q D N A E N V Y V W

Gly249Val; Gly249Ser Met255Val
780 AAGCAAGGTGTGGACGTCAAGGCCATGACCTCTAACGTGGCCAGTGTGCAGTGTGACAAT
K Q G V D V K A M T S N V A S V Q C D N
Asp 251Tyr
840 TCAGACAAGTTTCCCGTCTACAAATACCCAGGAAAGGGGTGCCCCACCCTTGGTGATGAA
S D K F P V Y K Y P G K G C P T L G D E
900 GGTGACACAGACCTATATGATTACTACCCAGAAGAGGACACTGAGGGCGATAAGGTGCGT
G D T D L Y D Y Y P E E D T E G D K V R
960 GCCACAAGGACTGTGGTCAAGTTCCCCACCAAAGCCCATAACAACCCCTGGGGTCTATTC
A T R T V V K F P T K A H T T P W G L F
1020 TACTCATGGTCCACTGCTTCTCTAGACAGCCAAATGCCCTCCTCCTTGCATCCAACACAA
Y S W S T A S L D S Q M P S S L H P T Q
1080 GAATCCACTAAGGAGCAGACCACATTCCCACCTAGATGGACCCCAAATTCACACTTCAC
E S T K E Q T T F P P R W T P N F T L H
1140 ATGGAATCCATCACATTCTCCAAAATCCAAAATCCACTACTGAACCAACCCCAAGCCCG
M E S I T F S K T P K S T T E P T P S P
1200 ACCACCTCAGAGCCCGTCCCGGAGCCCGCCCCAAACATGACCACCCTGGAGCCCACTCCA
T T S E P V P E P A P N M T T L E P T P
1260 AGCCCGACCACCCAGAGCCCACCTCAGAGCCCGCCCCAGCCCGACCACCCCGGAGCCC
S P T T P E P T S E P A P S P T T P E P
1320 ACCTCAGAGCCCGCCCCAGCCCGACCACCCAGAGCCCACCTCAGAGCCCGCCCCAGC
T S E P A P S P T T P E P T S E P A P S
1380 CCGACCACCCCGGAGCCCACCCAATCCCGACCATCGCCACAAGCCCGACCATCCTGGTG
P T T P E P T P I P T I A T S P T I L V
1440 TCTGCCACAAGCCTGATCACTCCAAAAGCACATTTTAACTACCACAAAACCCGTATCA
S A T S L I T P K S T F L T T T K P V S

1500 CTCTTAGAATCCACCAAAAAAACCATCCCTGAACTTGATCAGCCACCAAAGCTCCGTGGG
L L E S T K K T I P E L D Q P P K L R G
1560 GTGCTCCAAGGGCATTGAGAGCTCCAGAAATGACCCTTTTCTCCACCCCGACTTTTGC
V L Q G H L E S S R N D P F L H P D F C
1620 TGCCTCCTCCCCCTGGGCTTCTATGTCCTGGGTCTCTTCTGGCTGCTCTTTGCCTCTGTG
C L L P L G F Y V L G L F W L L F A S V
1680 GTCCTCATCCTGCTGCTGAGCTGGGTTGGGCATGTGAAACCACAGGCCCTGGACTCTGGC
V L I L L L S W V G H V K P Q A L D S G
1740 CAAGGTGCTGCTCTGACCACAGCCACACAAACCACACACCTGGAGCTGCAGAGGGGACGG
Q G A A L T T A T Q T T H L E L Q R G R
1800 CAAGTGACAGTGCCCCGGGCCTGGCTGCTCTTCCTTCGAGGTTTCGCTTCCCACTTTCCGC
Q V T V P R A W L L F L R G S L P T F R
1860 TCCAGCCTCTTCCTGTGGGTACGGCCTAATGGCCGTGTGGGGCCTCTAGTGGCAGGAAGG
S S L F L W V R P N G R V G P L V A G R
1920 AGGCCCTCAGCTCTGAGTCAGGGTCGTGGTCAGGACCTGCTGAGCACAGTGAGCATTAGG
R P S A L S Q G R G Q D L L S T V S I R
1980 TACTCTGGCCACAGCCTCTGAGGGTGGGAGGTTTGGGGACCTTGAGAGAAGAGCCTGTGG
Y S G H S L STOP
2040 GCTCTCCTATTGGAATCTAGTTGGGGGTTGGAGGGGTAAGGAACACAGGGTGATAGGGAG
2100 GGGTCTTAGTTCCTTTTTTCTGTATCAGAAGCCCTGTCTTCACAACACAGGCACACAATTT
2160 CAGTCCCAGCCAAAGCAGAAGGGGTAATGACATGGACTTGGCGGGGGGACAAGACAAAGC
2220 TCCCGATGCTGCATGGGGCGCTGCCAGATCTCACGGTGAACCATTTTGGCAGAATACAGC
2280 ATGGTTCCACATGCATCTATGCACAGAAGAAAATCTGGAAAGTGATTTATCAGGATGTG

2340 AGCACTCGTTGTGTCTGGATGTTACAAATATGGGTGGTTTTATTTTCTTTTCCCTGTTT

2400 AGCATTTTCTAGTTTTCCACTATTATTGTATATTATCTGTATAATAAAAAATAATTTTAG

2460 GGTTGGGA

Nucleotide sequence and predicted amino acid sequence for GPIBA showing the known mutation sites.

Solid Diamonds – Potential N-linked glycosylation sites

Single Underline – Nucleotide sequences obtained by Edman degradation

Double Underline – Potential transmembrane region