SEQUENCE LISTING

SEQ ID NO: 1 (native CRM197)

GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNE
NPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPF
AEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVIR
DKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAA
WAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGEL
VDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKI
TAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVA
FHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS

SEQ ID NO: 2 (CRM197 with Arg-Asn substitution)

GADDVVDSSKSFVMENFSSYHGTKPGYVDSIQKGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDNE
NPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLPF
AEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRNSVGSSLSCINLDWDVIR
DKTKTKIESLKEHGPIKNKMSESPNKTVSEEKAKQYLEEFHQTALEHPELSELKTVTGTNPVFAGANYAA
WAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGEL
VDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHKTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIKI
TAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHVA
FHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSKLSLFFEIKS

SEQ ID NO: 3 (CRM197 with 6 preferred nnAA sites and N-terminus Met)

MGADDVVDSSKSFVMENFSSYHGTKPGYVDSIQXGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDN
ENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLP
FAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRRSVGSSLSCINLDWDVI
RDXTKTKIESLKEHGPIKNKMSESPNKTVSEEKAXQYLEEFHQTALEHPELSELXTVTGTNPVFAGANYA
AWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGE
LVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHXTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIK
ITAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHV
AFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSXLSLFFEIKS

SEQ ID NO: 4 (CRM197 with Arg-Asn subst^{n}, 6 preferred nnAA sites, and N-terminus Met)

MGADDVVDSSKSFVMENFSSYHGTKPGYVDSIQXGIQKPKSGTQGNYDDDWKEFYSTDNKYDAAGYSVDN
ENPLSGKAGGVVKVTYPGLTKVLALKVDNAETIKKELGLSLTEPLMEQVGTEEFIKRFGDGASRVVLSLP
FAEGSSSVEYINNWEQAKALSVELEINFETRGKRGQDAMYEYMAQACAGNRVRNSVGSSLSCINLDWDVI
RDXTKTKIESLKEHGPIKNKMSESPNKTVSEEKAXQYLEEFHQTALEHPELSELXTVTGTNPVFAGANYA
AWAVNVAQVIDSETADNLEKTTAALSILPGIGSVMGIADGAVHHNTEEIVAQSIALSSLMVAQAIPLVGE
LVDIGFAAYNFVESIINLFQVVHNSYNRPAYSPGHXTQPFLHDGYAVSWNTVEDSIIRTGFQGESGHDIK
ITAENTPLPIAGVLLPTIPGKLDVNKSKTHISVNGRKIRMRCRAIDGDVTFCRPKSPVYVGNGVHANLHV
AFHRSSSEKIHSNEISSDSIGVLGYQKTVDHTKVNSXLSLFFEIKS

SEQ ID NO: 5 (H.influenzae protein D)

CSSHSSNMANTQMKSDKIIIAHRGASGYLPEHTLESKALAFAQQADYLEQDLAMTKDGRLVVIHDHFLDG
LTDVAKKFPHRHRKDGRYYVIDFTLKEIQSLEMTENFETKDGKQAQVYPNRFPLWKSHFRIHTFEDEIEF
IQGLEKSTGKKVGIYPEIKAPWFHHQNGKDIAAETLKVLKKYGYDKKTDMVYLQTFDFNELKRIKTELLP

QMGMDLKLVQLIAYTDWKETQEKDPKGYWVNYNYDWMFKPGAMAEVVKYADGVGPGWYMLVNKEESKPDN
IVYTPLVKELAQYNVEVHPYTVRKDALPEFFTDVNQMYDALLNKSGATGVFTDFPDTGVEFLKGIK