Appendix I

Annotated protein sequence alignments of human SLC4s (see Figures 3 and 15)

This appendix includes two sequence alignments. **Alignment 1** compares the protein sequences of example variants of all ten human SLC4s. Note that AE4 and BTR1 include unusual sequence that can complicate alignment of the remaining 8 SLC4s in some regions, thus we align the AEs and NCBTs without AE4 and BTR1 in **Alignment 2**. Both **Alignments 1** and **2** display the putative positions of transmembrane spans (gray shaded areas), substantial extracellular loops (ELs), putative glycosylation sites (dark green shaded motifs), conserved cysteines in EL3 (C), lysine residues in TM5 that form a DIDS-interaction site in some SLC4s (K), NBCe1 residues substituted in proximal renal tubular acidosis (X), splice cassettes (yellow=cassette I, cyan=cassette II, green=cassette A, blue=cassette III). **Alignment 2** displays, in addition, the boundaries of NCBT subdomains defined in the review and depicted in Figure 15. Sequence accession numbers are provided in Appendix IV.

Alignment 1. Clustal W alignment of human SLC4 proteins sequences.

```
AE2a            MSSAPRRPAKGADSFCTPEPESLGPGTPGFPEQEEDELHRTLGVERFEEILQEAGSRGGE 60
bAE3            MANGVIPPPGGASPLPQVRVPLEEPPLSPDVEEEDDDLGKTLAVSRFGDLISKPPAWDPE 60
eAE1            ------------------------------------------------------------
NBCe1-A         ------------------------------------------------------------
NBCe2-c         ------------------------------------------------------------
AE4             ------------------------------------------------------------
NDCBE-A         ------------------------------------------------------------
NBCn2-B         ------------------------------------------------------------
NBCn1-D         ------------------------------------------------------------
BTR1            ------------------------------------------------------------
```

```
AE2a            EPGRSYGEEDFEYHRQSSHHIHHPLSTHLPPDARRRKTPQGPGRKPRRR-----PGASPT 115
bAE3            KPSRSYSERDFEFHRHTSHHTHHPLSARLPPPHKLRRLPPTSARHTRRKRKKEKTSAPPS 120
eAE1            ------------------------------------------------------------
NBCe1-A         ------------------------------------------------------------
NBCe2-c         ------------------------------------------------------------
AE4             ------------------------------------------------------------
NDCBE-A         ------------------------------------------------------------
NBCn2-B         ------------------------------------------------------------
NBCn1-D         ------------------------------------------------------------
BTR1            --------------------------------------------------------MSQV 4
```

```
AE2a            GETPTIEEGEEDEDEASEAEGARALTQPSPVSTPSSVQFFLQEDDSADRKAERTSPSSPA 175
bAE3            EGTPPIQEEGGAGVDEEEEEEEEEEEGESEAEPVEPPHSGTPQKAKFSIGSDEDDSPGLPG 180
eAE1            -------MEELQDDYEDMMEENLEQEEYEDPDIPESQMEEPAAHDTEATATDYHTTSHPG 53
NBCe1-A         -----MSTENVEGKPSNLGERGRARSSTFLRVVQPMFN---------------------- 33
NBCe2-c         -----MKVKEEKAGVGKLDHTNHRRRFPDQKECPPIHIGLPVPTYPQRKTDQKGHLSGLQ 55
AE4             ---MEMKLPGQEGFEASSAPRNIPSGELDSNPDPGTGPSPDGPSDTESKELG-------- 49
NDCBE-A         -MPAAGSNEPDGVLSYQRPDEEAVVDQGGTSTILNIHYEKEELEGHRTLYVGVRMPLG-R 58
NBCn2-B         -MEIKDQGAQMEPLLPTRNDEEAVVDRGGTRSILKTHFEKEDLEGHRTLFIGVHVPLGGR 59
NBCn1-D         -MEADGAGEQMRPLLTRGPDEEAVVDLGKTSSTVNTKFEKEELESHRAVYIGVHVPFS-K 58
BTR1            GGRGDRCTQEVQGLVHGAGDLSASLAENSPTMSQNGYFEDSSYYKCDTDDTFEAREEILG 64


AE2a            PLPHQEATPRASKGAQAGTQVEEAEAEAVAVASGTAGGDDGGASGRPLPKAQPGHRSYNL 235
bAE3            RAAVTKPLPSVGPHTDKSPQHSSSSPSPRARASRLAGEKS---------RPWSPSASYDL 231
eAE1            ------------------------------------------------------------
NBCe1-A         ----------------------------HSIFTSAVSPAA-------------------- 45
NBCe2-c         KVHWGLRPDQPQQELTGPGSGASSQDSSMDLISRTRSPAA-------------------- 95
AE4             -----------------------------VPKDPLLFIQL-------------------- 60
NDCBE-A         QSHRHHRTHGQKHRRRGRGKGASQGEEGLEALAHDTPSQR-------------------- 98
NBCn2-B         KSHRRHRHRGHKHRKRDRER-DSGLEDGRESPSFDTPSQR-------------------- 98
NBCn1-D         ESRRRHRHRGHKHHHRRRKDKESDKEDGRESPSYDTPSQR-------------------- 98
BTR1            ------------------------------------------------------------


AE2a            QERRRIGSMTGAEQALLPRVPTDEIEAQTLATADLDLMKSHRFEDVPGVRRHLVRKNAKG 295
bAE3            RERLCPGSALGNPGGPEQQVPTDEAEAQMLGSADLDDMKSHRLEDNPGVRRHLVKKPSRT 291
eAE1            ------------------------------------------------------------
NBCe1-A         ------------------------------------------------------------
NBCe2-c         ------------------------------------------------------------
AE4             ------------------------------------------------------------
NDCBE-A         ------------------------------------------------------------
NBCn2-B         ------------------------------------------------------------
NBCn1-D         ------------------------------------------------------------
BTR1            ------------------------------------------------------------


AE2a            STQSGREGREPGPTPRARPR--APHKPHEVFVELNELLLD-KNQEPQWRETARWIKFEED 352
bAE3            QGGRGSPSGLAPILRRKKKKKLDRRPHEVFVELNELMLD-RSQEPHWRETARWIKFEED 350
eAE1            --------------------------THKVYVELQELVMDEKNQELRWMEAARWVQLEEN 87
NBCe1-A         -----------ERIRFILGEEDDSPAPPQLFTELDELLAV-DGQEMEWKETARWIKFEEK 93
NBCe2-c         -----------EQLQDILGEEDEAPNP--TLFTEMDTLQH-DGDQMEWKESARWIKFEEK 141
AE4             -----------NELLGWPQALEWRETGSSSASLLLDMGEM-PSITLSTHLHHRWVLFEEK 108
NDCBE-A         -----------VQFILGTEED-EEHVPHELFTELDEICMK-EGEDAEWKETARWLKFEED 145
NBCn2-B         -----------VQFILGTEDDDEEHIPHDLFTELDEICWR-EGEDAEWRETARWLKFEED 146
NBCn1-D         -----------VQFILGTEDDDEEHIPHDLFTEMDELCYR-DGEEYEWKETARWLKFEED 146
BTR1            ------------------DEAFDTANSSIVSGESIRFFVNVNLEMQATNTENEATSGGCV 106
                                                                    .

AE2a            VEEETERWGKPHVASLSFRSLLELRRTLAHGAVLLDLDQQTLPGVAHQVVEQMVISDQIK 412
bAE3            VEEETERWGKPHVASLSFRSLLELRRTIAHGAALLDLEQTTLPGIAHLVVETMIVSDQIR 410
eAE1            LGEN-GAWGRPHLSHLTFWSLLELRRVFTKGTVLLDLQETSLAGVANQLLDRFIFEDQIR 146
NBCe1-A         VEQGGERWSKPHVATLSLHSLFELRTCMEKGSIMLDREASSLPQLVEMIVDHQIETGLLK 153
NBCe2-c         VEEGGERWSKPHVSTLSLHSLFELRTCLQTGTVLLDLDSGSLPQIIDDVIEKQIEDGLLR 201
AE4             LEVAAGRWSAPHVPTLALPSLQKLRSLLAEGLVLLDCPAQSLLELVEQVTRVESLS---- 164
NDCBE-A         VEDGGERWSKPYVATLSLHSLFELRSCLINGTVLLDMHANSIEEISDLILDQQELSSDLN 205
NBCn2-B         VEDGGERWSKPYVATLSLHSLFELRSCILNGTVLLDMHANTLEEIADMVLDQQVSSGQLN 206
NBCn1-D         VEDGGDRWSKPYVATLSLHSLFELRSCILNGTVMLDMRASTLDEIADMVLDNMIASGQLD 206
BTR1            LLHTSRKYLKLKNFKEEIRAHRDLDGFLAQASIVLNETATSLDNVLRTMLRRFARDPDNN 166
                :      :         : :  .*   :  .  :*:    ::  :   :
```

```
                                                      Cassette I
AE2a            AEDRANVLRALLLKHSHPSDEKD-FSFPRNISAGSLGSLLGHHHGQGAESDPHVTEPLMG 471
bAE3            PEDRASVLRTLLLKHSHPNDDKDSGFFPRNPSSSSMNSVLGNHHPTPSHGPDGAVPTMAD 470
eAE1            PQDREELLRALLLKHSHAG-------------------------ELEALGGVKPAVLTRS 181
NBCe1-A         PELKDKVTYTLLRKHRHQTKK------SNLRSLADIGKTVSSASRMFTNPDNGSPAMTHR 207
NBCe2-c         PELRERVSYVLLRRHRHQTKK------PIHRSLADIGKSVSTTNRSPARSPGAGPSLHHS 255
AE4             PELRGQLQALLLQRPQHYNQT------TGTRPCWGS------------------------ 194
NDCBE-A         DSMRVKVREALLKKHHHQNEKKRNNLIPIVRSFAEVGKKQSDPHLMDKHGQTVSPQSVPT 265
NBCn2-B         EDVRHRVHEALMKQHHHQNQKKLTNRIPIVRSFADIGKKQSEPNSMDKNAGQVVSPQSAP 266
NBCn1-D         ESIRENVREALLKRHHHQNEKRFTSRIPLVRSFADIGKKHSDPHLLERNGEGLSASRHSL 266
BTR1            EPNCNLDLLMAMLFTDAGAPMRG------------------------------------- 189
                           :


                                Cassette II
AE2a            GVPETRLEVERERE---------------------------------------------- 485
bAE3            DLGEPAPLWPHDPD---------------------------------------------- 484
eAE1            GDPSQPLLPQHSSL---------------------------------------------- 195
NBCe1-A         ------------------------------------------------------------
NBCe2-c         TEDLRMRQSANYGR---------------------------------------------- 269
AE4             ------------------------------------------------------------
NDCBE-A         TNLEVKNGVNCEHS---------------------------------------------- 279
NBCn2-B         ACVENKNDVSRENS---------------------------------------------- 280
NBCn1-D         RTGLSASNLSLRGESPLSLLLGHLLPSSRAGTPAGSRCTTPVPTPQNSPPSSPSISRLTS 326
BTR1            ------------------------------------------------------------


AE2a            ------------------------------------------------------------
bAE3            ------------------------------------------------------------
eAE1            ------------------------------------------------------------
NBCe1-A         ------------------------------------------------------------
NBCe2-c         ------------------------------------------------------------
AE4             ------------------------------------------------------------
NDCBE-A         --------------------------------------------------------PVDL 283
NBCn2-B         --------------------------------------------------------TVDF 284
NBCn1-D         RSSQKSQRQAPELLVSPASDDIPTVVIHPPEEDLEAALKGEEQKNEENVDLTPGILASPQ 386
BTR1            ------------------------------------------------------------

                       Cassette A
AE2a            ----------------LPPPAPPAGITRSKSKHELKLLEKIPENAEATVVLVGCVEFLSR 529
bAE3            ----------------AKEKPLHMPGGDGHRGKSLKLLEKIPEDAEATVVLVGCVPFLEQ 528
eAE1            ----------------ETQLFCEQGDGGTEGHSPSGILEKIPPDSEATLVLVGRADFLEQ 239
NBCe1-A         --------------NLTSSSLNDISDKPEKDQLKNKFMKKLPRDAEASNVLVGEVDFLDT 253
NBCe2-c         ------------LCHAQSRSMNDISLTPNTDQRKNKFMKKIPKDSEASNVLVGEVDFLDQ 317
AE4             ---------------THPRKASDNEEAPLREQCQNPLRQKLPPGAEAGTVLAGELGFLAQ 239
NDCBE-A         SK------------------------------VDLHFMKKIPTGAEASNVLVGEVDILDR 313
NBCn2-B         SKGLGGQQKGHTSPCGMKQRHEKGPPHQQEREVDLHFMKKIPPGAEASNILVGELEFLDR 344
NBCn1-D         SAPGNLDNSKSGEIKGNGSGGSRENSTVDFSKVDMNFMRKIPTGAEASNVLVGEVDFLER 446
BTR1            --------------------------KVHLLSDTIQGVTATVTGVRYQQSWLCIICTMKA 223
                                                           . .           :

AE2a            PTMAFVRLREAVELDAVLEVPVPVRFLFLLLGP--SSANMDYHEIGRSISTLMSDKQFHE 587
bAE3            PAAAFVRLNEAVLLESVLEVPVPVRFLFVMLGP--SHTSTDYHELGRSIATLMSDKLFHE 586
eAE1            PVLGFVRLQEAAELEAVELP-VPIRFLFVLLGP--EAPHIDYTQLGRAAATLMSERVFRI 296
NBCe1-A         PFIAFVRLQQAVMLGALTEVPVPTRFLFILLGP--KGKAKSYHEIGRAIATLMSDEVFHD 311
NBCe2-c         PFIAFVRLIQSAMLGGVTEVPVPTRFLFILLGP--SGRAKSYNEIGRAIATLMVDDLFSD 375
AE4             PLGAFVRLRNPVVLGSLTEVSLPSRFFCLLLGP--CMLGKGYHEMGRAAAVLLSDPQFQW 297
NDCBE-A         PIVAFVRLSPAVLLSGLTEVPIPTRFLFILLGP--VGKGQQYHEIGRSMATIMTDEIFHD 371
NBCn2-B         TVVAFVRLSPAVLLQGLAEVPIPTRFLFILLGP--LGKGQQYHEIGRSIATLMTDEVFHD 402
NBCn1-D         PIIAFVRLAPAVLLTGLTEVPVPTRFLFLLLGP--AGKAPQYHEIGRSIATLMTDEIFHD 504
```

```
BTR1            LQKRHVCISRLVRPQNWGENSCEVRFVILVLAPPKMKSTKTAMEVARTFATMFSDIAFRQ 283
                    .* :   .            **. ::*.*          ::.*: :.:: :  *

AE2a            AAYLADEREDLLTAINAFLDCSVVLPPSEVQGEELLRSVAHFQRQMLKKREEQGRLLPTG 647
bAE3            AAYQADDRQDLLSAISEFLDGSIVIPPSEVEGRDLLRSVAAFQRELLRKRREREQTKVEM 646
eAE1            DAYMAQSRGELLHSLEGFLDCSLVLPPTDAPSEQALLSLVPVQRELLRRRYQSSPAKPDS 356
NBCe1-A         IAYKAKDRHDLIAGIDEFLDEVIVLPPGEWDPAIRIEPPKSLPSSDKRKNMYSGGENVQM 371
NBCe2-c         VAYKARNREDLIAGIDEFLDEVIVLPPGEWDPNIRIEPPKKVPSADKRKSVFSLAELGQM 435
AE4             SVRRASNLHDLLAALDAFLEEVTVLPPGRWDPTARIPPPKCLPSQHKRLPSQQREIRGPA 357
NDCBE-A         VAYKAKERDDLLAGIDEFLDQVTVLPPGEWDPSIRIEPPKNVPSQEKRKMPGVPNGNVCH 431
NBCn2-B         VAYKAKDRNDLVSGIDEFLDQVTVLPPGEWDPSIRIEPPKNVPSQEKRKIPAVPNGTAAH 462
NBCn1-D         VAYKAKDRNDLLSGIDEFLDQVTVLPPGEWDPSIRIEPPKSVPSQEKRKIPVFHNGSTPT 564
BTR1            KLLETRTEEEFKEALVHQRQLLTMVSHGPVAPRTKERSTVSLPAHRHPEPPKCKDFVPFG 343
                    :    ::  .:    :   ::.           .   .


AE2a            AGLEPKSAQDKALLQMVEAAGAA------------------------E-DDPLRRTGRPF 682
bAE3            TTRGGYTAPGKELSLELGGSEAT------------------------PEDDPLLRTGSVF 682
eAE1            SFYKGLDLNGG-------------------------------------PDDPLQQTGQLF 379
NBCe1-A         NGDTPHDGGHGGGGHG--------------------------------DCEELQRTGRFC 399
NBCe2-c         NGSVGGGGGAPGGGNGGGGGGGSGGGAGSGGAGGTSSGDDGEMPAMHEIGEELIWTGRFF 495
AE4             VPRLTSAEDRHRHGPHAHS-------------------------------PELQRTGRLF 386
NDCBE-A         IEQEP--HGGHS-------------------------------------GPELQRTGRLF 452
NBCn2-B         GEAEP--HGGHS-------------------------------------GPELQRTGRIF 483
NBCn1-D         LGETPKEAAHHA-------------------------------------GPELQRTGRLF 587
BTR1            ------------------------------------------------------------

                        Nt                                TM1
AE2a            GGLIRDVRRRYPHYLSDFRDALD------PQCLAAVIFIYFAALSPAITFGGLLGEKTQD 736
bAE3            GGLVRDVRRRYPHYPSDLRDALH------SQCVAAVLFIYFAALSPAITFGGLLGEKTEG 736
eAE1            GGLVRDIRRRYPYYLSDITDAFS------PQVLAAVIFIYFAALSPAITFGGLLGEKTRN 433
NBCe1-A         GGLIKDIKRKAPFFASDFYDALN------IQALSAILFIYLATVTNAITFGGLLGDATDN 453
NBCe2-c         GGLCLDIKRKLPWFPSDFYDGFH------IQSISAILFIYLGCITNAITFGGLLGDATDN 549
AE4             GGLIQDVRRKVPWYPSDFLDALH------LQCFSAVLYIYLATVTNAITFGGLLGDATDG 440
NDCBE-A         GGLVLDIKRKAPWYWSDYRDALS------LQCLASFLFLYCACMSPVITFGGLLGEATEG 506
NBCn2-B         GGLILDIKRKAPYFWSDFRDAFS------LQCLASFLFLYCACMSPVITFGGLLGEATEG 537
NBCn1-D         GGLILDIKRKAPFFLSDFKDALS------LQCLASILFLYCACMSPVITFGGLLGEATEG 641
BTR1            KGIREDIARRFPLYPLDFTDGIIGKNKAVGKYITTTLFLYFACLLPTIAFGSLNDENTDG 403
                 *:  *: *: * :  *  *.:        : .:: :::* . :  .*:**.* .: * .

                         TM2                          TM3                TM4
AE2a            LIGVSELIMSTALQGVVFCLLGAQPLLVIGFSGPLLVFEEAFFSFCSSNHLEYLVGRVWI 796
bAE3            LMGVSELIVSTAVLGVLFSLLGAQPLLVVGFSGPLLVFEEAFFKFCRAQDLEYLTGRVWV 796
eAE1            QMGVSELLISTAVQGILFALLGAQPLLVVGFSGPLLVFEEAFFSFCETNGLEYIVGRVWI 493
NBCe1-A         MQGVLESFLGTAVSGAIFCLFAGQPLTILSSTGPVLVFERLLFNFSKDNNFDYLEFRLWI 513
NBCe2-c         YQGVMESFLGTAMAGSLFCLFSGQPLIILSSTGPILIFEKLLFDFSKGNGLDYMEFRLWI 609
AE4             AQGVLESFLGTAVAGAAFCLMAGQPLTILSSTGPVLVFERLLFSFSRDYSLDYLPFRLWV 500
NDCBE-A         RISAIESLFGASMTGIAYSLFAGQALTILGSTGPVLVFEKILFKFCKDYALSYLSLRACI 566
NBCn2-B         RISAIESLFGASMTGIAYSLFGGQPLTILGSTGPVLVFEKILFKFCKEYGLSYLSLRASI 597
NBCn1-D         RISAIESLFGASLTGIAYSLFAGQPLTILGSTGPVLVFEKILYKFCRDYQLSYLSLRTSI 701
BTR1            AIDVQKTIAGQSIGGLLYALFSGQPLVILLTTAPLALYIQVIRVICDDYDLDFNSFYAWT 463
                  .. : : . :: *  :.*:..*.* ::  :.*: :: . :  :.    :.:
```

```
                                                        TM5
AE2a            GFWLVFLALLMVALEGSFLVRFVSRFTQEIFAFLISLIFIYETFYKLVKIFQEHPLHGCS 856
bAE3            GLWLVVFVLALVAAEGSFLVRYISPFTQEIFAFLISLIFIYETFYKLYKVFTEHPLLPFY 856
eAE1            GFWLILLVVLVVAFEGSFLVRFISRYTQEIFSFLISLIFIYETFSKLIKIFQDHPLQKTY 553
NBCe1-A         GLWSAFLCLILVATDASFLVQYFTRFTEEGFSSLISFIFIYDAFKKMIKLADYYPINSNF 573
NBCe2-c         GLHSAVQCLILVATDASFIIKYITRFTEEGFSTLISFIFIYDAIKKMIGAFKYYPINMDF 669
AE4             GIWVATFCLVLVATEASVLVRYFTRFTEEGFCALISLIFIYDAVGKMLNLTHTYPIQKPG 560
NDCBE-A         GLWTAFLCIVLVATDASSLVCYITRFTEEAFASLICIIFIYEAIEKLIHLAETYPIHMHS 626
NBCn2-B         GLWTATLCIILVATDASSLVCYITRFTEEAFASLICIIFIYEALEKLFELSEAYPINMHN 657
NBCn1-D         GLWTSFLCIVLVATDASSLVCYITRFTEEAFAALICIIFIYEALEKLFDLGETYAFNMHN 761
BTR1            GLWNSFFLALYAFFNLSLVMSLFKRSTEEIIALFISITFVLDAVKGTVKIFWKYYYGHYL 523
                *:         .  : * ::  ..  *:* :. :*.: *: ::.         :

                                          EL3
AE2a            ASNSSEVDGGENMTWAGA-----------------------------------RPTLGPG 881
bAE3            PP--------EGALEGSL-----------------------------------DAGLEPN 873
eAE1            NYN--------------------------------------------------------- 556
NBCe1-A         KVGYNTLFSCTCVPPDPANISISNDTTLAPEYLPTMSSTDMYHNTTFDWAFLSKKECSKY 633
NBCe2-c         KPNFITTYKCECVAPDTVNTTVFNASAPLAPDT-NASLYNLLNLTALDWSLLSKKECLSY 728
AE4             SS----AYGCLCQYPGPG----GNESQWIRTRPKDRDDIVSMDLGLINASLLPPPECTRQ 612
NDCBE-A         QLDHLSLYYCRCTLPENP-----------NNHTLQYWKDHNIVTAEVHWANLTVSECQEM 675
NBCn2-B         DLELLTQYSCNCVEPHNP-----------SNGTLKEWRESNISASDIIWENLTVSECKSL 706
NBCn1-D         NLDKLTSYSCVCTEPPNP-----------SNETLAQWKKDNITAHNISWRNLTVSECKKL 810
BTR1            DDYHTKRTSSLVSLSGLG-----------------------------ASLNASLHTALNA 554

                                         TM6                             TM7
AE2a            NRSLAGQSGQGKPRGQPNTALLSLVLMAGTFFIAFFLRKFKNSRFFPGRIRRVIGDFGVP 941
bAE3            GSALPPTEGPPSPRNQPNTALLSLILMLGTFFIAFFLRKFRNSRFLGGKARRIIGDFGIP 933
eAE1            -----VLMVPKPQGPLPNTALLSLVLMAGTFFFAMMLRKFKNSSYFPGKLRRVIGDFGVP 611
NBCe1-A         GGNLVGNNC----NFVPDITLMSFILFLGTYTSSMALKKFKTSPYFPTTARKLISDFAII 689
NBCe2-c         GGRLLGNSC----KFIPDLALMSFILFFGTYSMTLTLKKFKFSRYFPTKVRALVADFSIV 784
AE4             GGHPRGPGC----HTVPDIAFFSLLLFLTSFFFAMALKCVKTSRFFPSVVRKGLSDFSSV 668
NDCBE-A         HGEFMGSACGHHGPYTPDVLFWSCILFFTTFILSSTLKTFKTSRYFPTRVRSMVSDFAVF 735
NBCn2-B         HGEYVGRACGHDHPYVPDVLFWSVILFFSTVTLSATLKQFKTSRYFPTKVRSIVSDFAVF 766
NBCn1-D         RGVFLGSACGHHGPYIPDVLFWCVILFFTTFFLSSFLKQFKTKRYFPTKVRSTISDFAVF 870
BTR1            SFLASPTELPSATHSGQATAVLSLLIMLGTLWLGYTLYQFKKSPYLHPCVREILSDCALP 614
                                    . . :::  :      *  .: . ::    *  :.* .

                                               EL4                    TM8
AE2a            IAILIMVLVDYSIE---DTYTQKLSVPSGFSVTAPEKRGWVINPLGEKSPFPVWMMVASL 998
bAE3            ISILVMVLVDYSIT---DTYTQKLTVPTGLSVTSPDKRSWFIPPLGSARPFPPWMMVAAA 990
eAE1            ISILIMVLVDFFIQ---DTYTQKLSVPDGFKVSNSSARGWVIHPLGLRSEFPIWMMFASA 668
NBCe1-A         LSILIFCVIDALVG----VDTPKLIVPSEFKPTSPN-RGWFVPPFG---ENPWWVCLAAA 741
NBCe2-c         FSILMFCGIDACFG----LETPKLHVPSVIKPTRPD-RGWFVAPFG---KNPWWVYPASI 836
AE4             LAILLGCGLDAFLG----LATPKLMVPREFKPTLPG-RGWLVSPFG---ANPWWSVAAA 720
NDCBE-A         LTIFTMVIIDFLIG----VPSPKLQVPSVFKPTRDD-RGWIINPIG---PNPWWTVIAAI 787
NBCn2-B         LTILCMVLIDYAIG----IPSPKLQVPSVFKPTRDD-RGWFVTPLG---PNPWWTVIAAI 818
NBCn1-D         LTIVIMVTIDYLVG----VPSPKLHVPEKFEPTHPE-RGWIISPLG---DNPWWTLLIAA 922
BTR1            IAVLAFSLISSHGFREIEMS------KFRYNPSESPFAMAQIQSLS------LRAVSGAM 662
                :::.    :.                    . :        : .:.            :
```

```
                                                               TM9
AE2a            LPAILVFILIFMETQITTLIISKKERMLQKGSGFHLDLLLIVAMGGICALFGLPWLAAAT 1058
bAE3            VPALLVLILIFMETQITALIVSQKARRLLKGSGFHLDLLLIGSLGGLCGLFGLPWLTAAT 1050
eAE1            LPALLVFILIFLESQITTLIVSKPERKMVKGSGFHLDLLLVVGMGGVAALFGMPWLSATT 728
NBCe1-A         IPALLVTILIFMDQQITAVIVNRKEHKLKKGAGYHLDLFWVAILMVICSLMALPWYVAAT 801
NBCe2-c         LPALLVTILIFMDQQITAVIVNRKENKLKKAAGYHLDLFWVGILMALCSFMGLPWYVAAT 896
AE4             LPALLLSILIFMDQQITAVILNRMEYRLQKGAGFHLDLFCVAVLMLLTSALGLPWYVSAT 780
NDCBE-A         IPALLCTILIFMDQQITAVIINRKEHKLKKGCGYHLDLLMVAIMLGVCSIMGLPWFVAAT 847
NBCn2-B         IPALLCTILIFMDQQITAVIINRKEHKLKKGCGYHLDLLMVAVMLGVCSIMGLPWFVAAT 878
NBCn1-D         IPALLCTILIFMDQQITAVIINRKEHKLKKGAGYHLDLLMVGVMLGVCSVMGLPWFVAAT 982
BTR1            GLGFLLSMLFFIEQNLVAALVNAPENRLVKGTAYHWDLLLLAIINTGLSLFGLPWIHAAY 722
                  .:*  :*:*:: ::.: ::.     : *. .:* **: :  :    . :.:**  ::

                                                TM10                 TM11
AE2a            VRSVTHANALTVMSKAVAPGDKPK-IQEVKEQRVTGLLVALLVGLSIVIGDL-LRQIPLA 1116
bAE3            VRSVTHVNALTVMRTAIAPGDKPQ-IQEVREQRVTGVLIASLVGLSIVMGAV-LRRIPLA 1108
eAE1            VRSVTHANALTVMGKASTPGAAAQ-IQEVKEQRISGLLVAVLVGLSILMEPI-LSRIPLA 786
NBCe1-A         VISIAHIDSLKMETETSAPGEQPK-FLGVREQRVTGTLVFILTGLSVFMAPI-LKFIPMP 859
NBCe2-c         VISIAHIDSLKMETETSAPGEQPQ-FLGVREQRVTGIIVFILTGISVFLAPI-LKCIPLP 954
AE4             VISLAHMDSLRRESRACAPGERPN-FLGIREQRLTGLVVFILTGASIFLAPV-LKFIPMP 838
NDCBE-A         VLSITHVNSLKLESECSAPGEQPK-FLGIREQRVTGLMIFVLMGCSVFMTAI-LKFIPMP 905
NBCn2-B         VLSITHVNSLKLESECSAPGEQPK-FLGIREQRVTGLMIFILMGSSVFMTSI-LKFIPMP 936
NBCn1-D         VLSISHVNSLKVESECSAPGEQPK-FLGIREQRVTGLMIFILMGLSVFMTSV-LKFIPMP 1040
BTR1            PHSPLHVRALALVEERVENGHIYDTIVNVKETRLTSLGASVLVGLSLLLLPVPLQWIPKP 782
                  *  *  :*         *   . :  ::* *::.     * * *:.:  : *  ** .

                                         Span 12                      TM13
AE2a            VLFGIFLYMGVTSLNGIQFYERLHLLLMPPKHHPDVTYVKKVRTLRMHLFTALQLLCLAL 1176
bAE3            VLFGIFLYMGVTSLSGIQLSQRLLLILMPAKHHPEQPYVTKVKTWRMHLFTCIQLGCIAL 1168
eAE1            VLFGIFLYMGVTSLSGIQLFDRILLLFKPPKYHPDVPYVKRVKTWRMHLFTGIQIICLAV 846
NBCe1-A         VLYGVFLYMGVASLNGVQFMDRLKLLLMPLKHQPDFIYLRHVPLRRVHLFTFLQVLCLAL 919
NBCe2-c         VLYGVFLYMGVASLNGIQFWERCKLFLMPAKHQPDHAFLRHVPLRRIHLFTLVQILCLAV 1014
AE4             VLYGIFLYMGVAALSSIQFTNRVKLLLMPAKHQPDLLLLRHVPLTRVHLFTAIQLACLGL 898
NDCBE-A         VLYGVFLYMGVSSLQGIQFFDRLKLFGMPAKHQPDFIYLRHVPLRKVHLFTLIQLTCLVL 965
NBCn2-B         VLYGVFLYMGASSLKGIQFFDRIKLFWMPAKHQPDFIYLRHVPLRKVHLFTIIQMSCLGL 996
NBCn1-D         VLYGVFLYMGVSSLKGIQLFDRIKLFGMPAKHQPDLIYLRYVPLWKVHIFTVIQLTCLVL 1100
BTR1            VLYGLFLYIALTSLDGNQLVQRVALLLKEQTAYPPTHYIRRVPQRKIHYFTGLQVLQLLL 842
                **:*:***:. ::*.. *: :*  *:    .  *    :  *   ::* ** :*:  : :

                                  TM14              Ct
AE2a            LWAVMST---AASLAFPFILILTVPLRMVVLTRIFTDREMKCLDANEAEPVFDEREGVDE 1233
bAE3            LWVVKST---AASLAFPFLLLLTVPLRHCLLPRLFQDRELQALDSEDAEPNFDE-DGQDE 1224
eAE1            LWVVKST---PASLALPFVLILTVPLRRVLLPLIFRNVELQCLDADDAKATFDEEEGRDE 903
NBCe1-A         LWILKST---VAAIIFPVMILALVAVRK-GMDYLFSQHDLSFLDDVIPEKDKKKKEDEKK 975
NBCe2-c         LWILKST---VAAIIFPVMILGLIIVRR-LLDFIFSQHDLAWIDNILPEKEKKETDKKRK 1070
AE4             LWIIKST---PAAIIFPLMLLGLVGVRK-ALERVFSPQELLWLDELMPEEERSIPEKGLE 954
NDCBE-A         LWVIKAS---PAAIVFPMMVLALVFVRK-VMDLCFSKRELSWLDDLMPESKKKKLDDAKK 1021
NBCn2-B         LWIIKVS---RAAIVFPMMVLALVFVRK-LMDLLFTKRELSWLDDLMPESKKKKLEDAEK 1052
NBCn1-D         LWVIKVS---AAAVVFPMMVLALVFVRK-LMDLCFTKRELSWLDDLMPESKKKKEDDKKK 1156
BTR1            LCAFGMSSLPYMKMIFPLIMIAMIPIRYILLPRIIEAKYLDVMDAEHRP----------- 891
                *  .  :      : :*.:::  : :*   :   :    :  :*
```

```
                                                           Cassette III
AE2a            YNEMPMPV---------------------------------------------------- 1241
bAE3            YNELHMPV---------------------------------------------------- 1232
eAE1            YDEVAMPV---------------------------------------------------- 911
NBCe1-A         KKKKKGSLDSDNDD-------------------SDCPYSEKVPSIKIPMDIMEQQPFLSD 1016
NBCe2-c         RKKGAHEDCDEEPQ-------------------FPPPSVIKIPMESVQSDPQNGIHCIAR 1111
AE4             PEHSFSGSDSEDSE-------------------LMYQPKAPEINISVN------------ 983
NDCBE-A         KAKEEEEAEKMLEIGGDKFPLESRKLLSSPGKNISCRCDPSEINISDEMPKTTVWKALSM 1081
NBCn2-B         EE--EQSMLAMEDEGTVQLPLEG-----------HYRDDPSVINISDEMSKTALWRNLLI 1099
NBCn1-D         KEKEEAERMLQDDDDTVHLPFEGGSLLQIPVKALKYSVDPSIVNISDEMAKTAQWKALSM 1216
BTR1            ------------------------------------------------------------

AE2a            ---------------------------------------
bAE3            ---------------------------------------
eAE1            ---------------------------------------
NBCe1-A         SKPSDRERSPTFLERHTSC-------------------- 1035
NBCe2-c         KRSSSWSYSL----------------------------- 1121
AE4             ---------------------------------------
NDCBE-A         NSGNAKEKSLFN--------------------------- 1093
NBCn2-B         TADNSKDKESSFPSKSSPS-------------------- 1118
NBCn1-D         NTENAKVTRSNMSPDKPVSVKISFEDEPRKKYVDAETSL 1255
BTR1            ---------------------------------------
```

Alignment 2. Clustal W alignment of human AE and NCBT proteins sequences, showing NCBT subdomains.

```
                ------------------Nt_appendage------------------------------
NDCBE-A         ------------------------------------------------------------
NBCn2-B         ------------------------------------------------------------
NBCn1-D         ------------------------------------------------------------
NBCe1-A         ------------------------------------------------------------
NBCe2-c         ------------------------------------------------------------
AE2a            MSSAPRRPAKGADSFCTPEPESLGPGTPGFPEQEEDELHRTLGVERFEEILQEAGSRGGE 60
bAE3            MANGVIPPPGGASPLPQVRVPLEEPPLSPDVEEEDDDLGKTLAVSRFGDLISKPPAWDPE 60
eAE1            ------------------------------------------------------------

                ------------------Nt_appendage------------------------------

NDCBE-A         ---------------------------------------------------MPAAGSNEP 9
NBCn2-B         ---------------------------------------------------MEIKDQGAQ 9
NBCn1-D         ---------------------------------------------------MEADGAGEQ 9
NBCe1-A         -----------------------------------------------------MSTENVE 7
NBCe2-c         -----------------------------------------------------MKVKEEK 7
AE2a            EPGRSYGEEDFEYHRQSSHHIHHPLSTHLPPDARRRKTPQGPGRKPRRR-----PGASPT 115
bAE3            KPSRSYSERDFEFHRHTSHHTHHPLSARLPPPHKLRRLPPTSARHTRRKRKKEKTSAPPS 120
eAE1            ------------------------------------------------------------

                ------------------Nt_appendage------------------------------
NDCBE-A         DGVLSYQRPDEEAVVDQGGTSTILNIHYEKEELEGHRTLYVGVRMPLG-RQSHRHHRTHG 68
NBCn2-B         MEPLLPTRNDEEAVVDRGGTRSILKTHFEKEDLEGHRTLFIGVHVPLGGRKSHRRHRHRG 69
NBCn1-D         MRPLLTRGPDEEAVVDLGKTSSTVNTKFEKEELESHRAVYIGVHVPFS-KESRRRHRHRG 68
NBCe1-A         GKPSNLGERGRARSSTFLRVVQPMFN---------------------------------- 33
NBCe2-c         AGVGKLDHTNHRRRFPDQKECPPIHIGLPVPTYPQRKTDQKGHLSGLQKVHWGLRPDQPQ 67
AE2a            GETPTIEEGEEDEDEASEAEGARALTQPSPVSTPSSVQFFLQEDDSADRKAERTSPSSPA 175
bAE3            EGTPPIQEEGGAGVDEEEEEEEEEEGESEAEPVEPPHSGTPQKAKFSIGSDEDDSPGLPG 180
eAE1            -------MEELQDDYEDMMEENLEQEEYEDPDIPESQMEEPAAHDTEATATDYHTTSHPG 53

                ------------------Nt_appendage------------------------------
NDCBE-A         QKHRRRGRGKGASQGEEGLEALAH------------------------------------ 92
NBCn2-B         HKHRKRDRERDSG-LEDGRESPSF------------------------------------ 92
NBCn1-D         HKHHHRRRKDKESDKEDGRESPSY------------------------------------ 92
NBCe1-A         ----------------HSIFTSAV------------------------------------ 41
NBCe2-c         QELTGPGSGASSQDSSMDLISRTR------------------------------------ 91
AE2a            PLPHQEATPRASKGAQAGTQVEEAEAEAVAVASGTAGGDDGGASGRPLPKAQPGHRSYNL 235
bAE3            RAAVTKPLPSVGPHTDKSPQHSSSSPSPRARASRLAGEKS---------RPWSPSASYDL 231
eAE1            ------------------------------------------------------------

                ------------------Nt_appendage------------------------------
NDCBE-A         ------------------------------------------------------------
NBCn2-B         ------------------------------------------------------------
NBCn1-D         ------------------------------------------------------------
NBCe1-A         ------------------------------------------------------------
NBCe2-c         ------------------------------------------------------------
AE2a            QERRRIGSMTGAEQALLPRVPTDEIEAQTLATADLDLMKSHRFEDVPGVRRHLVRKNAKG 295
bAE3            RERLCPGSALGNPGGPEQQVPTDEAEAQMLGSADLDDMKSHRLEDNPGVRRHLVKKPSRT 291
eAE1            ------------------------------------------------------------
```

```
                -----------------------Nt_core_1------------------------------
NDCBE-A         -----DTPSQRVQFILGTEED-EEHVPHELFTELDEICMK-EGEDAEWKETARWLKFEED 145
NBCn2-B         -----DTPSQRVQFILGTEDDDEEHIPHDLFTELDEICWR-EGEDAEWRETARWLKFEED 146
NBCn1-D         -----DTPSQRVQFILGTEDDDEEHIPHDLFTEMDELCYR-DGEEYEWKETARWLKFEED 146
NBCe1-A         -----SPAAERIRFILGEEDD--SPAPPQLFTELDELLAV-DGQEMEWKETARWIKFEEK 93
NBCe2-c         -----SPAAEQLQDILGEEDE--APNP--TLFTEMDTLQH-DGDQMEWKESARWIKFEEK 141
AE2a            STQSGREGREPGPTPRARPR--APHKPHEVFVELNELLLD-KNQEPQWRETARWIKFEED 352
bAE3            QGGRGSPSGLAPILRRKKKKKKLDRRPHEVFVELNELMLD-RSQEPHWRETARWIKFEED 350
eAE1            --------------------------THKVYVELQELVMDEKNQELRWMEAARWVQLEEN 87
                                          .        :      .:: .* *:***:::**.

                ---------------------Nt_core_1------------------------------
NDCBE-A         VEDGGERWSKPYVATLSLHSLFELRSCLINGTVLLDMHANSIEEISDLILDQQELSSDLN 205
NBCn2-B         VEDGGERWSKPYVATLSLHSLFELRSCILNGTVLLDMHANTLEEIADMVLDQQVSSGQLN 206
NBCn1-D         VEDGGDRWSKPYVATLSLHSLFELRSCILNGTVMLDMRASTLDEIADMVLDNMIASGQLD 206
NBCe1-A         VEQGGERWSKPHVATLSLHSLFELRTCMEKGSIMLDREASSLPQLVEMIVDHQIETGLLK 153
NBCe2-c         VEEGGERWSKPHVSTLSLHSLFELRTCLQTGTVLLDLDSGSLPQIIDDVIEKQIEDGLLR 201
AE2a            VEEETERWGKPHVASLSFRSLLELRRTLAHGAVLLDLDQQTLPGVAHQVVEQMVISDQIK 412
bAE3            VEEETERWGKPHVASLSFRSLLELRRTIAHGAALLDLEQTTLPGIAHLVVETMIVSDQIR 410
eAE1            LGEN-GAWGRPHLSHLTFWSLLELRRVFTKGTVLLDLQETSLAGVANQLLDRFIFEDQIR 146
                : :    *.:*::: *:: **:***  :  *: :**    ::  : . :::     . :

                                                     Cassette I
                ------Nt_core_1--------------------------Nt_loop------------
NDCBE-A         DSMRVKVREALLKKHHHQNEKKRNNLIPIVRSFAEVGKKQSDPHLMDKHGQTVSPQSVPT 265
NBCn2-B         EDVRHRVHEALMKQHHHQNQKKLTNRIPIVRSFADIGKKQSEPNSMDKNAGQVVSPQSAP 266
NBCn1-D         ESIRENVREALLKRHHHQNEKRFTSRIPLVRSFADIGKKHSDPHLLERNGEGLSASRHSL 266
NBCe1-A         PELKDKVTYTLLRKHRHQTKK------SNLRSLADIGKTVSSASRMFTNPDNGSPAMTHR 207
NBCe2-c         PELRERVSYVLLRRHRHQTKK------PIHRSLADIGKSVSTTNRSPARSPGAGPSLHHS 255
AE2a            AEDRANVLRALLLKHSHPSDEKD-FSFPRNISAGSLGSLLGHHHGQGAESDPHVTEPLMG 471
bAE3            PEDRASVLRTLLLKHSHPNDDKDSGFFPRNPSSSSMNSVLGNHHPTPSHGPDGAVPTMAD 470
eAE1            PQDREELLRALLLKHSHAG-------------------------ELEALGGVKPAVLTRS 181
                 . :  :  .*: :* *


                        Cassette II
                ------------------------Nt_loop---------------------------
NDCBE-A         TNLEVKNGVNCEHS---------------------------------------------- 279
NBCn2-B         ACVENKNDVSRENS---------------------------------------------- 280
NBCn1-D         RTGLSASNLSLRGESPLSLLLGHLLPSSRAGTPAGSRCTTPVPTPQNSPPSSPSISRLTS 326
NBCe1-A         ------------------------------------------------------------
NBCe2-c         TEDLRMRQSANYGR---------------------------------------------- 269
AE2a            GVPETRLEVERERE---------------------------------------------- 485
bAE3            DLGEPAPLWPHDPD---------------------------------------------- 484
eAE1            GDPSQPLLPQHSSL---------------------------------------------- 195


                -------------------------Nt_loop----------------------------
NDCBE-A         --------------------------------------------------------PVDL 283
NBCn2-B         --------------------------------------------------------TVDF 284
NBCn1-D         RSSQKSQRQAPELLVSPASDDIPTVVIHPPEEDLEAALKGEEQKNEENVDLTPGILASPQ 386
NBCe1-A         ------------------------------------------------------------
NBCe2-c         ------------------------------------------------------------
AE2a            ------------------------------------------------------------
bAE3            ------------------------------------------------------------
eAE1            ------------------------------------------------------------
```

```
                           Cassette A
                ---Nt_loop----------------------------------Nt_core_2-------
NDCBE-A         SK------------------------------VDLHFMKKIPTGAEASNVLVGEVDILDR 313
NBCn2-B         SKGLGGQQKGHTSPCGMKQRHEKGPPHQQEREVDLHFMKKIPPGAEASNILVGELEFLDR 344
NBCn1-D         SAPGNLDNSKSGEIKGNGSGGSRENSTVDFSKVDMNFMRKIPTGAEASNVLVGEVDFLER 446
NBCe1-A         --------------NLTSSSLNDISDKPEKDQLKNKFMKKLPRDAEASNVLVGEVDFLDT 253
NBCe2-c         ------------LCHAQSRSMNDISLTPNTDQRKNKFMKKIPKDSEASNVLVGEVDFLDQ 317
AE2a            ----------------LPPPAPPAGITRSKSKHELKLLEKIPENAEATVVLVGCVEFLSR 529
bAE3            ----------------AKEKPLHMPGGDGHRGKSLKLLEKIPEDAEATVVLVGCVPFLEQ 528
eAE1            ----------------ETQLFCEQGDGGTEGHSPSGILEKIPPDSEATLVLVGRADFLEQ 239
                                                    ::.*:* .:**: :***   :*.

                ---------------------Nt_core_2------------------------------
NDCBE-A         PIVAFVRLSPAVLLSGLTEVPIPTRFLFILLGPVGKGQQYHEIGRSMATIMTDEIFHDVA 373
NBCn2-B         TVVAFVRLSPAVLLQGLAEVPIPTRFLFILLGPLGKGQQYHEIGRSIATLMTDEVFHDVA 404
NBCn1-D         PIIAFVRLAPAVLLTGLTEVPVPTRFLFLLLGPAGKAPQYHEIGRSIATLMTDEIFHDVA 506
NBCe1-A         PFIAFVRLQQAVMLGALTEVPVPTRFLFILLGPKGKAKSYHEIGRAIATLMSDEVFHDIA 313
NBCe2-c         PFIAFVRLIQSAMLGGVTEVPVPTRFLFILLGPSGRAKSYNEIGRAIATLMVDDLFSDVA 377
AE2a            PTMAFVRLREAVELDAVLEVPVPVRFLFLLLGPSSANMDYHEIGRSISTLMSDKQFHEAA 589
bAE3            PAAAFVRLNEAVLLESVLEVPVPVRFLFVMLGPSHTSTDYHELGRSIATLMSDKLFHEAA 588
eAE1            PVLGFVRLQEAAELEAVELP-VPIRFLFVLLGPEAPHIDYTQLGRAAATLMSERVFRIDA 298
                .  .****  :. * .:    :* ****::***     .* ::**: :*:* :  *   *

                ---------------------Nt_core_2------------------------------
NDCBE-A         YKAKERDDLLAGIDEFLDQVTVLPPGEWDPSIRIEPPKNVPSQEKRKMPGVPNGNVCHIE 433
NBCn2-B         YKAKDRNDLVSGIDEFLDQVTVLPPGEWDPSIRIEPPKNVPSQEKRKIPAVPNGTAAHGE 464
NBCn1-D         YKAKDRNDLLSGIDEFLDQVTVLPPGEWDPSIRIEPPKSVPSQEKRKIPVFHNGSTPTLG 566
NBCe1-A         YKAKDRHDLIAGIDEFLDEVIVLPPGEWDPAIRIEPPKSLPSSDKRKNMYSGGENVQMNG 373
NBCe2-c         YKARNREDLIAGIDEFLDEVIVLPPGEWDPNIRIEPPKKVPSADKRKSVFSLAELGQMNG 437
AE2a            YLADEREDLLTAINAFLDCSVVLPPSEVQGEELLRSVAHFQRQMLKKREEQGRLLPTGAG 649
bAE3            YQADDRQDLLSAISEFLDGSIVIPPSEVEGRDLLRSVAAFQRELLRKRREREQTKVEMTT 648
eAE1            YMAQSRGELLHSLEGFLDCSLVLPPTDAPSEQALLSLVPVQRELLRRRYQSSPAKPDSSF 358
                * * .* :*: .:. ***   *:** :      : .   .     ::

                --------------------Nt_TMD_linker---------------------------
NDCBE-A         QEP--HGGHS-------------------------------------GPELQRTGRLFGG 454
NBCn2-B         AEP--HGGHS-------------------------------------GPELQRTGRIFGG 485
NBCn1-D         ETPKEAAHHA-------------------------------------GPELQRTGRLFGG 589
NBCe1-A         DTPHDGGHGGGGHG--------------------------------DCEELQRTGRFCGG 401
NBCe2-c         SVGGGGGAPGGGNGGGGGGGSGGGAGSGGAGGTSSGDDGEMPAMHEIGEELIWTGRFFGG 497
AE2a            LEPKSAQDKALLQMVEAAGAA------------------------E-DDPLRRTGRPFGG 684
bAE3            RGGYTAPGKELSLELGGSEAT------------------------PEDDPLLRTGSVFGG 684
eAE1            YKGLDLNGG-------------------------------------PDDPLQQTGQLFGG 381
                                                                  *  **   **

                ---------------------------------TM1------------EL1---------
NDCBE-A         LVLDIKRKAPWYWSDYRDALSLQCLASFLFLYCACMSPVITFGGLLGEATEGRISAIESL 514
NBCn2-B         LILDIKRKAPYFWSDFRDAFSLQCLASFLFLYCACMSPVITFGGLLGEATEGRISAIESL 545
NBCn1-D         LILDIKRKAPFFLSDFKDALSLQCLASILFLYCACMSPVITFGGLLGEATEGRISAIESL 649
NBCe1-A         LIKDIKRKAPFFASDFYDALNIQALSAILFIYLATVTNAITFGGLLGDATDNMQGVLESF 461
NBCe2-c         LCLDIKRKLPWFPSDFYDGFHIQSISAILFIYLGCITNAITFGGLLGDATDNYQGVMESF 557
AE2a            LIRDVRRRYPHYLSDFRDALDPQCLAAVIFIYFAALSPAITFGGLLGEKTQDLIGVSELI 744
bAE3            LVRDVRRRYPHYPSDLRDALHSQCVAAVLFIYFAALSPAITFGGLLGEKTEGLMGVSELI 744
eAE1            LVRDIRRRYPYYLSDITDAFSPQVLAAVIFIYFAALSPAITFGGLLGEKTRNQMGVSELL 441
                *  *::*: * : **  *.:  * :::.:*:* . :: .********: * .  .. * :
```

```
                -TM2---------     ---------TM3-----------EL2------------TM4-
NDCBE-A         FGASMTGIAYSLFAGQALTILGSTGPVLVFEKILFKFCKDYALSYLSLRACIGLWTAFLC 574
NBCn2-B         FGASMTGIAYSLFGGQPLTILGSTGPVLVFEKILFKFCKEYGLSYLSLRASIGLWTATLC 605
NBCn1-D         FGASLTGIAYSLFAGQPLTILGSTGPVLVFEKILYKFCRDYQLSYLSLRTSIGLWTSFLC 709
NBCe1-A         LGTAVSGAIFCLFAGQPLTILSSTGPVLVFERLLFNFSKDNNFDYLEFRLWIGLWSAFLC 521
NBCe2-c         LGTAMAGSLFCLFSGQPLIILSSTGPILIFEKLLFDFSKGNGLDYMEFRLWIGLHSAVQC 617
AE2a            MSTALQGVVFCLLGAQPLLVIGFSGPLLVFEEAFFSFCSSNHLEYLVGRVWIGFWLVFLA 804
bAE3            VSTAVLGVLFSLLGAQPLLVVGFSGPLLVFEEAFFKFCRAQDLEYLTGRVWVGLWLVVFV 804
eAE1            ISTAVQGILFALLGAQPLLVVGFSGPLLVFEEAFFSFCETNGLEYIVGRVWIGFWLILLV 501
                ..::: *  :.*:..*.* ::. :**:*:**. ::.*.    :.*:  *  :*:

                ----------           ----------TM5--------------------------
NDCBE-A         IVLVATDASSLVCYITRFTEEAFASLICIIFIYEAIEKLIHLAETYPIHMHSQLDHLSLY 634
NBCn2-B         IILVATDASSLVCYITRFTEEAFASLICIIFIYEALEKLFELSEAYPINMHNDLELLTQY 665
NBCn1-D         IVLVATDASSLVCYITRFTEEAFAALICIIFIYEALEKLFDLGETYAFNMHNNLDKLTSY 769
NBCe1-A         LILVATDASFLVQYFTRFTEEGFSSLISFIFIYDAFKKMIKLADYYPINSNFKVGYNTLF 581
NBCe2-c         LILVATDASFIIKYITRFTEEGFSTLISFIFIYDAIKKMIGAFKYYPINMDFKPNFITTY 677
AE2a            LLMVALEGSFLVRFVSRFTQEIFAFLISLIFIYETFYKLVKIFQEHPLHGCSASNSSEVD 864
bAE3            LALVAAEGSFLVRYISPFTQEIFAFLISLIFIYETFYKLYKVFTEHPLLPFYPP------ 858
eAE1            VLVVAFEGSFLVRFISRYTQEIFSFLISLIFIYETFSKLIKIFQDHPLQKTYNYN----- 556
                : :** :.* :: :.: :*:* *: **.:****::: *:      :.:

                --------------------------------EL3-------------------------
NDCBE-A         YCRCTLPENP----NNHTLQYWKDH-NIV------TAEVHWANLTVSECQEMHGEFMGSA 683
NBCn2-B         SCNCVEPHNP----SNGTLKEWRES-NIS------ASDIIWENLTVSECKSLHGEYVGRA 714
NBCn1-D         SCVCTEPPNP----SNETLAQWKKD-NIT----AHNIS--WRNLTVSECKKLRGVFLGSA 818
NBCe1-A         SCTCVPPDPANISISNDTTLAPEYLPTMSSTDMYHNTTFDWAFLSKKECSKYGGNLVGNN 641
NBCe2-c         KCECVAPDTVNTTVFNASAPLAPDT-NASLYNLLNLTALDWSLLSKKECLSYGGRLLGNS 736
AE2a            GGENMTWAGA----------------------------------RPTLGPGNRSLAGQS 889
bAE3            --EGALEGSL----------------------------------DAGLEPNGSALPPTE 881
eAE1            ---------------------------------------------------------VLM 559

                ------------------TM6----------             ----------TM7---
NDCBE-A         CGHHGPYTPDVLFWSCILFFTTFILSSTLKTFKTSRYFPTRVRSMVSDFAVFLTIFTMVI 743
NBCn2-B         CGHDHPYVPDVLFWSVILFFSTVTLSATLKQFKTSRYFPTKVRSIVSDFAVFLTILCMVL 774
NBCn1-D         CGHHGPYIPDVLFWCVILFFTTFFLSSFLKQFKTKRYFPTKVRSTISDFAVFLTIVIMVT 878
NBCe1-A         C----NFVPDITLMSFILFLGTYTSSMALKKFKTSPYFPTTARKLISDFAIILSILIFCV 697
NBCe2-c         C----KFIPDLALMSFILFFGTYSMTLTLKKFKFSRYFPTKVRALVADFSIVFSILMFCG 792
AE2a            GQGKPRGQPNTALLSLVLMAGTFFIAFFLRKFKNSRFFPGRIRRVIGDFGVPIAILIMVL 949
bAE3            GPPSPRNQPNTALLSLILMLGTFFIAFFLRKFRNSRFLGGKARRIIGDFGIPISILVMVL 941
eAE1            VPKPQGPLPNTALLSLVLMAGTFFFAMMLRKFKNSSYFPGKLRRVIGDFGVPISILIMVL 619
                        *:  : . :*:  *   :  *: *: . ::    *  :.**.: ::*. :

                ----------------------EL4---------------------------TM8-----
NDCBE-A         IDFLIG-VPSPKLQVPSVFKPTRDD-RGWIINPIG---PNPWWTVIAAIIPALLCTILIF 798
NBCn2-B         IDYAIG-IPSPKLQVPSVFKPTRDD-RGWFVTPLG---PNPWWTVIAAIIPALLCTILIF 829
NBCn1-D         IDYLVG-VPSPKLHVPEKFEPTHPE-RGWIISPLG---DNPWWTLLIAAIPALLCTILIF 933
NBCe1-A         IDALVG-VDTPKLIVPSEFKPTSPN-RGWFVPPFG---ENPWWVCLAAAIPALLVTILIF 752
NBCe2-c         IDACFG-LETPKLHVPSVIKPTRPD-RGWFVAPFG---KNPWWVYPASILPALLVTILIF 847
AE2a            VDYSIEDTYTQKLSVPSGFSVTAPEKRGWVINPLGEKSPFPVWMMVASLLPAILVFILIF 1009
bAE3            VDYSITDTYTQKLTVPTGLSVTSPDKRSWFIPPLGSARPFPPWMMVAAAVPALLVLILIF 1001
eAE1            VDFFIQDTYTQKLSVPDGFKVSNSSARGWVIHPLGLRSEFPIWMMFASALPALLVFILIF 679
                :*  .    : ** **  :. :  . *.*.: *:*     * *    : :**:*  ****
```

```
                -----                      --------TM9---------
NDCBE-A         MDQQITAVIINRKEHKLKKGCGYHLDLLMVAIMLGVCSIMGLPWFVAATVLSITHVNSLK 858
NBCn2-B         MDQQITAVIINRKEHKLKKGCGYHLDLLMVAVMLGVCSIMGLPWFVAATVLSITHVNSLK 889
NBCn1-D         MDQQITAVIINRKEHKLKKGAGYHLDLLMVGVMLGVCSVMGLPWFVAATVLSISHVNSLK 993
NBCe1-A         MDQQITAVIVNRKEHKLKKGAGYHLDLFWVAILMVICSLMALPWYVAATVISIAHIDSLK 812
NBCe2-c         MDQQITAVIVNRKENKLKKAAGYHLDLFWVGILMALCSFMGLPWYVAATVISIAHIDSLK 907
AE2a            METQITTLIISKKERMLQKGSGFHLDLLLIVAMGGICALFGLPWLAAATVRSVTHANALT 1069
bAE3            METQITALIVSQKARRLLKGSGFHLDLLLIGSLGGLCGLFGLPWLTAATVRSVTHVNALT 1061
eAE1            LESQITTLIVSKPERKMVKGSGFHLDLLLVVGMGGVAALFGMPWLSATTVRSVTHANALT 739
                :: ***::*:.:  . : *..*:****: :  :  :...:.:**  *:** *::* ::*.

                               ---------TM10---------   --------TM11--------
NDCBE-A         LESECSAPGEQPKFLGIREQRVTGLMIFVLMGCSVFMTAILKFIPMPVLYGVFLYMGVSS 918
NBCn2-B         LESECSAPGEQPKFLGIREQRVTGLMIFILMGSSVFMTSILKFIPMPVLYGVFLYMGASS 949
NBCn1-D         VESECSAPGEQPKFLGIREQRVTGLMIFILMGLSVFMTSVLKFIPMPVLYGVFLYMGVSS 1053
NBCe1-A         METETSAPGEQPKFLGVREQRVTGTLVFILTGLSVFMAPILKFIPMPVLYGVFLYMGVAS 872
NBCe2-c         METETSAPGEQPQFLGVREQRVTGIIVFILTGISVFLAPILKCIPLPVLYGVFLYMGVAS 967
AE2a            VMSKAVAPGDKPKIQEVKEQRVTGLLVALLVGLSIVIGDLLRQIPLAVLFGIFLYMGVTS 1129
bAE3            VMRTAIAPGDKPQIQEVREQRVTGVLIASLVGLSIVMGAVLRRIPLAVLFGIFLYMGVTS 1121
eAE1            VMGKASTPGAAAQIQEVKEQRISGLLVAVLVGLSILMEPILSRIPLAVLFGIFLYMGVTS 799
                :     :**  .::  ::***::* ::  * * *:.:  :*  **:.**:*:*****.:*

                -----------Span 12----------------------TM13------    ----
NDCBE-A         LQGIQFFDRLKLFGMPAKHQPDFIYLRHVPLRKVHLFTLIQLTCLVLLWVIKASPAAIVF 978
NBCn2-B         LKGIQFFDRIKLFWMPAKHQPDFIYLRHVPLRKVHLFTIIQMSCLGLLWIIKVSRAAIVF 1009
NBCn1-D         LKGIQLFDRIKLFGMPAKHQPDLIYLRYVPLWKVHIFTVIQLTCLVLLWVIKVSAAAVVF 1113
NBCe1-A         LNGVQFMDRLKLLLMPLKHQPDFIYLRHVPLRRVHLFTFLQVLCLALLWILKSTVAAIIF 932
NBCe2-c         LNGIQFWERCKLFLMPAKHQPDHAFLRHVPLRRIHLFTLVQILCLAVLWILKSTVAAIIF 1027
AE2a            LNGIQFYERLHLLLMPPKHHPDVTYVKKVRTLRMHLFTALQLLCLALLWAVMSTAASLAF 1189
bAE3            LSGIQLSQRLLLILMPAKHHPEQPYVTKVKTWRMHLFTCIQLGCIALLWVVKSTAASLAF 1181
eAE1            LSGIQLFDRILLLFKPPKYHPDVPYVKRVKTWRMHLFTGIQIICLAVLWVVKSTPASLAL 859
                *.*:*: :*  *: * *::*:  ::  *   ::*:** :*: *: :** :  : *:: :

                -TM14-----------------------Ct_core-------------------------
NDCBE-A         PMMVLALVFVRK-VMDLCFSKRELSWLDDLMPESKKKKLDDAKKKAKEEEEAEKMLEIGG 1037
NBCn2-B         PMMVLALVFVRK-LMDLLFTKRELSWLDDLMPESKKKKLEDAEKEE--EQSMLAMEDEGT 1066
NBCn1-D         PMMVLALVFVRK-LMDLCFTKRELSWLDDLMPESKKKKEDDKKKKEKEEAERMLQDDDDT 1172
NBCe1-A         PVMILALVAVRK-GMDYLFSQHDLSFLDDVIPEKDKKKKEDEKKKKKKKGSLDSDNDD-- 989
NBCe2-c         PVMILGLIIVRR-LLDFIFSQHDLAWIDNILPEKEKKETDKKRKRKKGAHEDCDEEPQ-- 1084
AE2a            PFILILTVPLRMVVLTRIFTDREMKCLDANEAEPVFDEREGVDEYNEMPMPV-------- 1241
bAE3            PFLLLLTVPLRHCLLPRLFQDRELQALDSEDAEPNFDE-DGQDEYNELHMPV-------- 1232
eAE1            PFVLILTVPLRRVLLPLIFRNVELQCLDADDAKATFDEEEGRDEYDEVAMPV-------- 911
                *.:::  : :*   :   * . ::  :*   .:   .: :   :

                                               Cassette III
                ---------------------Ct_appendage---------------------------
NDCBE-A         DKFPLESRKLLSSPGKNISCRCDPSEINISDEMPKTTVWKALSMNSGNAKEKSLFN---- 1093
NBCn2-B         VQLPLEG-----------HYRDDPSVINISDEMSKTALWRNLLITADNSKDKESSFPSKS 1115
NBCn1-D         VHLPFEGGSLLQIPVKALKYSVDPSIVNISDEMAKTAQWKALSMNTENAKVTRSNMSPDK 1232
NBCe1-A         --------------------SDCPYSEKVPSIKIPMDIMEQQPFLSDSKPSDRERSPTFL 1029
NBCe2-c         --------------------FPPPSVIKIPMESVQSDPQNGIHCIARKRSSSWSYSL--- 1121
AE2a            ------------------------------------------------------------
bAE3            ------------------------------------------------------------
eAE1            ------------------------------------------------------------
```

```
                ---Ct_appendage--------
NDCBE-A         -----------------------
NBCn2-B         SPS-------------------- 1118
NBCn1-D         PVSVKISFEDEPRKKYVDAETSL 1255
NBCe1-A         ERHTSC----------------- 1035
NBCe2-c         -----------------------
AE2a            -----------------------
bAE3            -----------------------
eAE1            -----------------------
```