

6 APPENDIX

Collection of Proline-Kinked Transmembrane Helices

Because of influences on transmembrane helices by different proline insertions, we collected structural data based on transmembrane helices with proline-kinks. To decide, which structure (first four alphanumeric letters of entries represent entries of Protein Data Base, PDB) finally should be used for modeling our transmembrane bundle based on rhodopsin template, the following sequences were aligned against the respective sequence from our receptors.

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/prokink_helix.ps Date: Tue Oct 17 16:00:27 2006
Page 1 of 2

```

  la87          -WGPLLLEVESWII
  laij_1_1      -HIPPAFAFAILAYLTFLVLFPRVM
  laij_1_2      -HYNPAHMIAISFFFTHALALALHGALVLSAA
  laij_1_3      -MRTPDHEDTFFRDL
  laij_1_4      -GWNPAVPLR
  laij_1_5      -IRPIL
  laij_1_6      -FYNPFHGLSIAFLYGSALLFAMHGATILA
  laij_1_7      -CFFPHAPTGDPMK
  larl_a_1      -VIPALFGGFGNYFMPMLH
  larl_a_2      -APRRLNNLSSYNNMYVCQVALGVASLL
  larl_a_3      -LSPVLVLAGAITMLLMD
  larl_a_4      -GDPVLYQHLNFFGHPEVYIILPGFGIISHVIST
  larl_a_5      -YLPMVLLAMAAGIL
  larl_a_6      -AVPTGIEKVSWIAT
  larl_a_7      -KTPMLWAFGLFLFTVGVTGVUL
  lbcc_c_1      -LTRFFALHFLPPAIGITIHLTLFLHE
  lbcc_c_2      -D1IPPHYYSFKDILGLTLMLTPFLTLA
  lbcc_c_3      -LILPFLFLH
  lbcc_c_4      -FPFLSQSTLFWLLVANLILTWIGSQPVENPFIIIGQMASL
  lbcc_c_5      -LPFTICLTENRM
  lbcc_e        -ESDPSRKIGFSYLVTAITLGVAYAAKNNVTQFVSSM
  lbcc_g_1      -GVPVNW
  lbcc_g_2      -VAPPFLAFYLLYTWTQEF
  lbgy_c_1      -FILPPIIMAIAMVHLLFL
  lbgy_c_2      -DKIFFHPYYTIKNDILGALL
  lbgy_c_3      -LIPLL
  lbgy_c_4      -GGQPVEHPYITIGQLASVLYFLLILVLMPTAGTIENKL
  lbgy_d        -PEHDHRKRKMGLFMLLMMGLLLPLVYAMKRHKWSVLKS
  lbgy_f        -YLEPYLKVEIRERKERKEEWM
  lbgy_g_1      -YFSKQIPNPVLR
  lbgy_g_2      -ILRVAPPFVAFYLVYTWTQEF
  lc3w_a_1      -SDPDAKKFYAITTLV
  lc3w_a_2      -TTPLLLLDDALL
  lc3w_a_3      -MAPEVASTFKVLRNVTVVLMWSAYPVVWL
  Icii          -WPLFLWATE
  lcol_a        -WGPLMLLEVESWVL
  idxx_a        -ERMTLSISVLLSLTVFLLVIVELPST
  lel2_a_1      -GRPLRIWATLMIPLVSISSYLGLLS
  lel2_a_2      -STPMILLALGLL
  lel2_a_3      -AEIFDTLRLVLTVVWLGYPIVNAV
  lezy_d        -EPEHDERKRLGLKTIIILSSLYLLSINVKKPFWAGIK
  lezv_f_1      -VLSKLCVPVANQFINL
  lezv_f_2      -ENPIMQTALR
  lezv_f_3      -LLPYILEAEAAAKEKDEL
  lezv_g        -FLYVLIPIAGIYIWYWKNGNEYNEFL
  lfx8_a_1      -LNPAVTIALML
  lfx8_a_2      -VIFPIIVSQQVAGAFCAAALVYGL
  lfx8_a_3      -GNGVPRGFLAPLLIGLLIAIVIGASMGLPT
  lfx8_a_4      -MNPARDFGPKVFAWL
  lfx8_a_5      -FGPIVGAIVGAFAYRK
  lih5_a        -VQPFIGGAL
  ljb0_a_1      -SLPINILL
  ljb0_a_2      -DIPLPHEFILNPSLMAE
  ljb0_a_3      -TDIPTQSLSTHMMWIGGFLVVGGAHGAIFMV
  ljb0_b_1      -AIPES
  ljb0_b_2      -STMPHPAGLAFF
  ljb0_b_3      -GGFHFQTES
  ljb0_b_4      -MLPGMLDBA
  ljb0_b_5      -IOPGDFLVHHAIALGLHTTLLILVNGAL
  liggj_a_1     -FVPRYIDWILTTPLIVYFLQLLA
  liggj_a_2     -VGPMTES
  liggj_a_3     -SGIKSLYVRLRNLTVVWAIYPPFINL
  liggj_a_4     -LPTPTVDVALIVYLDLV
  lkzu_a        -NQGKINTVVVHFAIGIPIALLGSVTVIAILVHLAILSHTTWFPAIWQGGV
  llgh_a        -SKPDDYKMLVINPNSTWLPIVIWIVATVVVIAVHAVALAAGPFWNIALGAAKSAAK
  lprc_1_1      -SQGPTMDPFAI
  lprc_1_2      -HVPLAFCVPPIPMFCVQLQVFPRPL
  lprc_1_3      -HYNPGHMSVSFLFVHAMALGLHGGLILSV
  lprc_1_4      -GMPEW
  lprc_1_5      -CINPTLV
  lprc_1_6      -GVPFGIMPHIDWLTAFPSIR
  lprc_1_7      -IYCPWHGCFSIGFAYGCGLLFAAHGATILA
  lprc_1_8      -EGAPLQPTGNGFLV

  ruler 1.....10.....20.....30.....40.....50.....60.....70

```

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/prokink_helix.ps Date: Tue Oct 17 16:00:27 2006
 Page 2 of 2

	*	
lprc_1_9	LYATPERA	8
lgle_c_1	IWPFFGAI	8
lgle_c_2	EHTPVVIRIGLQYGFILFIMSEVMFFFVAWFW	30
lvsg_a_1	QAFWQPLCQVSEELDDQPKGALFTLQAASKIQNMNRDAALRASIYAEI	48
lvsg_a_2	HQPNIDANRAKK	12
2occ_a_1	VMPIM	5
2occ_a_2	AFPRMNMSFWLLPPSFLLLAS	23
2occ_a_3	SLPVLAAG	8
2occ_a_4	GGDPILYQHLF	12
2occ_a_5	CHPEVYLILPGFGMISHIVTY	22
2occ_a_6	AIPTGVKVFSWLATL	15
2occ_a_7	WSPAMMMALGFIFLFTVGGLTGIV	24
2occ_a_8	WFPLF	5
2occ_c_1	VNHPSPWNFLTGALSALLMTSGLTMMNF	25
2occ_c_2	HHTPAVQKGLRYGMILFIISSEVLFFTGFYWAFYHSS	36
2occ_c_3	IHPLNPLEVFLLNTSVLLA	19
2vsg_a	WQPECELTAEI	11
lc17_m_1a	LIPLLRTQFFIVMGLVDRAIPMIAVGLGLYVVMFAVA	35
lc17_m_1b	RSKLIAPLAUTIFVNVFLMNLMDLLPIDLLP	31
lc17_m_2	GLPALRVVPSADVNVTLSMALGVFILILFYSI	31
lc17_n_1	SLLSKPVSLGLRLFGNMYAGELIFILIAGL	30
lc17_n_2	QWILNVPWAIFHILIIITLQAFIFMVLTIVYLS	32
ruler	1.....10.....20.....30.....40.....50.....60.....70.....	

Multiple Sequence Alignment of Endothelin Receptors by Subtypes

Endothelin receptor subtypes ETA and ETB from different species were sorted by subtype, and were aligned using typical transmembrane arrangement of G protein-coupled receptors.

- et1r – endothelin receptor subtype ETA
- etbr – endothelin receptor subtype ETB
- OPSD – rhodopsin

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/etr_builder.ps Date: Tue Oct 17 15:53:32 2006
Page 1 of 5

etlr_bovin	ME---TFWLRLSFWVALVGGVI	SDNPESY	STNL	SIHVDSVATFHGT	ELSF	47		
etlr_chicken	ME---ALYLRVSS	LLLLLG	FVLCDSSD	KYATNWSDVG	IYLPTSSGIESSL	47		
etlr_mouse_fragment	-----	-----	-----	-----	-----			
etlr_pig	ME---TFCFRVSFWALL	GCVI	SDNPESH	STNL	STHVDDFTTFRGTEFSL	47		
etlr_rat	MG---VLCFLASF	FWLALV	GCAIA	ADNAERY	SANLSSHVEDFTPFGTEFN	47		
etlr_rabbit_fragment	---VISDNPERY	STNL	SNHMD	DEF	TTFHPGP	ELNLV	41	
etlr_sheep	ME---TFWLRVSF	FWVALV	GGVI	SDNPESY	STNL	IHVDSVTTFRGTEFSL	47	
etlr_human	ME---TLCLRAS	FWLALV	GCVI	SDNPESY	STNL	IHVDSVTTFRGTEFSL	47	
etar_selectiv	TKT	SATAQY	FGRFD	HNDFF	SNQLIS	FVFPFEQHTNATSE	9DGCIINNSRSFK	50
etbr_selectiv	I	EKETLSKHYE	D	EVASKKWL	VDTMSY	LIHPVQTAQKTS	IFRSQNLQNL	50
etbr_human	MQPPPSLCG	RALVALVL	-ACGLS	SRIWGEER	GFP	PDRAT	-PLLQTAEIMTP	48
etbr_bovin	MQPLPSLCG	RALVALVL	-ACGVAGI	QAAERE	REFFP	AGATQPLPGT	GEMMET	49
etbr_canfa	MQPPPSLCG	RALVALVL	-ACGLS	SRIWGEER	GFP	PDRAT	-PLLQTAEIMTP	48
etbr_horse	MQPLPTLCGRV	LVALVIL	-ACGVAGV	QGEER	RFPP	PARATP	PLLGFEEIMTP	49
etbr_mouse	MQSPASRCG	RALVALLL	-ACGFL	GVNGE	KRGFP	PAQATL	SLLGTKEVMTP	49
etbr_pig	MQPLRSCLCGR	RALVALIF	-ACGVAGV	QSEER	GFP	PAGATPP	PALRTGEIVAP	49
etbr_rat	MQSSASRCG	RALVALLL	-ACGLL	GVNGE	KRGFP	PAQATPS	SLLGTKEVMTP	49
etbr_cotja_fragment	---	EIKETFKY	INTVVS	CLVFVLGI	I	GNSTLLRI	IYKNKC-MRN	43
etbr_macfa_fragment	---	P	FGAEMCKL	VLPFI	QKASVG	ITVLSLCAL	SIDRYRAVAS	38
etbr_rabbit	MQPPPSLCG	IALLALVL	-ACGMAEV	VNGEER	REMP	SAPATP	PLLGASE	49
et_r_oryla	-MRASVLLL	CVFMT	TEVGTS	TRLVRD	SQSAP	-EALEV	SETNSSVQLPL	45
	-----	SSTM-TMT	TM	-----	SSTM	TM		19
OPSD	-----	-----	-----	-----	-----	-----		
ruler	1.....	10.....	20.....	30.....	40.....	50		



etlr_bovin	VV-----	TTHQPTN	LALPSNGSMHN-----	-----	67	
etlr_chicken	LP-----	TTRRPP-----	VSNQTVK-----	-----	62	
etlr_mouse_fragment	-----	-----	-----	ATLLRIIYQNK	11	
etlr_pig	VV-----	TTHRPTN	LALPSNGSMHN-----	-----	67	
etlr_rat	LG-----	TTLQPPN	LALPSNGSMHG-----	-----	67	
etlr_rabbit_fragment	LPSNGS	RHN-----	-----	YCPQQTKITSAFKYINT	67	
etlr_sheep	VV-----	TTHQPTN	LALPSNGSMHN-----	-----	67	
etlr_human	LV-----	TTHQPTN	LVLPSNGSMHN-----	-----	67	
etar_selectiv	EMKFQ	-----	-----	-----	55	
etbr_selectiv	K	-----	-----	-----	51	
etbr_human	PT-----	KTLWP-----	KGSNASLARSLAPAEV	PKGDRTAGSPPRTISPP	88	
etbr_bovin	PT-----	ETSWP-----	GRSNASDPRSSATPQ	I	PRGGRMAGIPPR--TPP	87
etbr_canfa	PT-----	KTLWP-----	KGSNASLARSLAPAEV	PKGDRTAGSPPRTISPP	88	
etbr_horse	PT-----	KTSWP-----	TGSNASVPRLSAPPQ	OMPQMKAGR	TAGAQRRTLPPP	89
etbr_mouse	PT-----	KTSWT-----	RGSNSL	MRSSA	PAEVTKGGRGAGVPPRSF-PP	88
etbr_pig	PT-----	KTFWP-----	RGSNASL	PRSSFPQ	MPKGGRMAGPPARTLTTPP	89
etbr_rat	PT-----	KTSWT-----	RGSNSL	MRSSA	PAEVTKGGRVAGVPPRSF-PP	88
etbr_cotja_fragment	ILIASLAL	GDLLIII	IDIPISVY	KLLAEDWP-----	FGVEMCKLVPFIQK	88
etbr_macfa_fragment	WSRIK	GIGIPK	WTAVEIVLIVWV	SVVLAVPEAIGFDMIT	YKGSYLRIC	88
etbr_rabbit	ST-----	KTSWP-----	RDSNASL	PRSSA	PAEIPKEGRTAGAPRRT--PP	87
et_r_oryla	PQPGR-----	PRSPPPP	MCVKPTEIKDAFKY	VNTIISCLIFVV	GIIGNST	90
OPSD	-----	-----	MNGTEGP	NFYVPF	FSNKTGVVRSPF	19
ruler60.....	70.....	80.....	90.....	100	24



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/etr_builder.ps Date: Tue Oct 17 15:53:32 2006
 Page 2 of 5

etlr_bovin	YCPQQTKITSASFYINTVISCTIFIVGMVGNATLLRIIYQNK-C-MRNGP-	115
etlr_chicken	-CSQQTKIAETFKYINTVVS CA IFIVGMVGNATLLRIIYQNK-C-MRNGP-	109
etlr_mouse_fragment	C-MRNGP-NALIASLALGDLIYVVIDLPINVFKLLAGRWPFDHNDFGVFL	59
etlr_pig	YCPQQTKITSASFYINTVISCTIFIVGMVGNATLLRIIYQNK-C-MRNGP-	115
etlr_rat	YCPQQTKITTAFKYINTVISCTIFIVGMVGNATLLRIIYQNK-C-MRNGP-	115
etlr_rabbit_fragment	VISCTIFIVGMVGNATLLRIIYQNK-C-MRNGP-NALIASLALGDLIYVVI	115
etlr_sheep	YCPQQTKITSASFYINTVISCTIFIVGMVGNATLLRIIYQNK-C-MRNGP-	115
etlr_human	YCPQQTKITSASFYINTVISCTIFIVGMVGNATLLRIIYQNK-C-MRNGP-	115
etar_selectiv		55
etbr_selectiv		51
etbr_human	PCQGPIEIKETFKYINTVVSCLVFVLGIIGNSTLLRIIYQNK-C-MRNGP-	136
etbr_bovin	PDCGPIEIKETFKYINTVVSCLVFVLGIIGNSTLLRIIYQNK-C-MRNGP-	135
etbr_canfa	PCEGSIEIKETFKYINTVVSCLVFVLGIIGNSTLLRIIYQNK-C-MRNGP-	136
etbr_horse	PCERTIEIKETFKYINTVVSCLVFVLGIIGNSTLLRIIYQNK-C-MRNGP-	137
etbr_mouse	PCQRNIEISKTFKYINTIVSCLVFVLGIIGNSTLLRIIYQNK-C-MRNGP-	136
etbr_pig	PCEGPIEIKDTFKYINTVVSCLVFVLGIIGNSTLLRIIYQNK-C-MRNGP-	137
etbr_rat	PCQRKIEINKTFKYINTIVSCLVFVLGIIGNSTLLRIIYQNK-C-MRNGP-	136
etbr_cotja_fragment	ASVGITVLSLCALSIDRYRAVASWSRIKGIGVPKWIAVEIVLIWVISVVL	138
etbr_macfa_fragment	LLHPVQKTAFM	99
etbr_rabbit	PCQRPTEIKDTFKYINTVVSCLVFVLGIIGNSTLLRIIYQNK-C-MRNGP-	135
et_r_oryla	LLRIIYRNKC-MRNGP-NVLIGSLALGDLIYIIIAFVINVYKLIAEDWP-	137
OPSD	EAPQYYLAEPWQFSMLAAYMFLLIMLGPFINFLTLVYTQHKK-LRT-PL	72
ruler110.....120.....130.....140.....150	

etlr_bovin	NALIASLALGDLIYVVIDLPINVFKLLAGRWPFEQNDFGVFLCKLFPFLQ	165
etlr_chicken	NALIASLALGDLIYIVIDIPIIIVYKLLAQWKPFGDSEFGQFLCKLFPFIQ	159
etlr_mouse_fragment	CKLFFLQKSSVGITVNL CAL SVDRYRAVASWSRVQGIGIPITAIEIV	109
etlr_pig	NALIASLALGDLIYVVIDLPINVFKLLAGRWPFEQNDFGVFLCKLFPFLQ	165
etlr_rat	NALIASLALGDLIYVVIDLPINVFKLLAGRWPFDHNDFGVFLCKLFPFLQ	165
etlr_rabbit_fragment	DLPINVFKLLAGRWPFDHNDFGVFLCKLFPFLQKSSVGITVNL CAL SVD	165
etlr_sheep	NALIASLALGDLIYVVIDLPINVFKLLAGRWPFEQNDFGVFLCKLFPFLQ	165
etlr_human	NALIASLALGDLIYVVIDLPINVFKLLAGRWPFDHNDFGVFLCKLFPFLQ	165
etar_selectiv		55
etbr_selectiv		51
etbr_human	NILIASLALGDLHIVIDIPINVYKLLAEDWP-----FGAEMCKLVPPFIQ	181
etbr_bovin	NILIASLALGDLHIIIIDIPINTYKLLAEDWP-----FGVEMCKLVPPFIQ	180
etbr_canfa	NILIASLALGDLHIIIIDIPINVYKLLAEDWP-----FGVEMCKLVPPFIQ	181
etbr_horse	NILIASLALGDLHIIIIDIPINVYKLLAEDWP-----FGVEMCKLVPPFIQ	182
etbr_mouse	NILIASLALGDLHIIIIDIPINVYKLLAEDWP-----FGAEMCKLVPPFIQ	181
etbr_pig	NILIASLALGDLHIIIIDIPINVYKLLAEDWP-----FGVEMCKLVPPFIQ	182
etbr_rat	NILIASLALGDLHIIIIDIPINVYKLLAEDWP-----FGAEMCKLVPPFIQ	181
etbr_cotja_fragment	AVPEAIAFDMITMEYRGKDLRICLLHPTQKTSFMMFYKQAKDWLFSFYF	188
etbr_macfa_fragment		99
etbr_rabbit	NILIASLALGDLHIIIIDIPINVYKLLAEDWP-----FGAEMCKLVPPFIQ	180
et_r_oryla	----FGVYICKLMPFIQKASVGITVLSLCALSIDRYHAVTSWSRVKGMI	183
OPSD	NYILLNLAVADLFMVFGGFTTLYTSLHGYFV----FGPTGCNLEGFFA	117
ruler160.....170.....180.....190.....200	

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/etr_builder.ps Date: Tue Oct 17 15:53:32 2006
 Page 3 of 5

etlr_bovin	KSSVGITVNL CALSVD RAVASWSRV QGIGIPLVTAIEIVSIWILSFI	215
etlr_chicken	KASVGITVNL CALSVD RAVASWSRV QGIGIPMIAIEIVSIWILSFI	209
etlr_mouse_fragment	SIWILSFI LAIPEAIGFVMVPFEYKGEHLTCMLNATSK--FMEFYQDVK	157
etlr_pig	KSSVGITVNL CALSVD RAVASWSRV QGIGIPLVTAIEIVSIWILSFI	215
etlr_rat	KSSVGITVNL CALSVD RAVASWSRV QGIGIPLITAEIVSIWILSFI	215
etlr_rabbit_fragment	RYRAVASWSRV QGIGIPLITAEIVSIWILSFI LAIPEAIGFVMVPFEYR	215
etlr_sheep	KSSVGITVNL CALSVD RAVASWSRV QGIGIPLVTAIEIVSIWILSFI	215
etlr_human	KSSVGITVNL CALSVD RAVASWSRV QGIGIPLVTAIEIVSIWILSFI	215
etar_selectiv		55
etbr_selectiv		51
etbr_human	KASVGITVLSL CALSIDRY RAVASWSRIKGIGVPKWTAVEIVLIWVVSVV	231
etbr_bovin	KASVGITVLSL CALSIDRY RAVASWSRIKGIGVPKWTAVEIVLIWVVSVV	230
etbr_canfa	KASVGITVLSL CALSIDRY RAVASWSRIKGIGVPKWTAVEIVLIWVVSVV	231
etbr_horse	KASVGITVLSL CALSIDRY RAVASWSRIKGIGVPKWTAVEIVLIWVVSVV	232
etbr_mouse	KASVGITVLSL CALSIDRY RAVASWSRIKGIGVPKWTAVEIVLIWVVSVV	231
etbr_pig	KASVGITVLSL CALSIDRY RAVASWSRIKGIGVPKWTAVEIVLIWVVSVV	232
etbr_rat	KASVGITVLSL CALSIDRY RAVASWSRIKGIGVPKWTAVEIVLIWVVSVV	231
etbr_cotja_fragment	CLPLAITALFYTLMTCEMLRKSG-MQIALNDHLKQRREVAKTVFCLVLV	237
etbr_macfa_fragment		99
etbr_rabbit	KASVGITVLSL CALSIDRY RAVASWSRIKGIGVPKWTAVEIVLIWVVSVI	230
et_r_oryla	PLWKAVEVTLIMLVAVVLAPEALAFDMLEMPYRGNKLRICLLHPEQPTV	233
		19
OPSD	TLGGEIALWSLVVLAIERYVVVCKPMSNF-RFGENHAIMGVAFTWVMALA	166
ruler210.....220.....230.....240.....250	

etlr_bovin	LAIPEAIGFVMVPFEYKGAQHRTCMLNATSK--FMEFYQDVKDWWLFGFY	263
etlr_chicken	LAIPEAIGFAVVPFRYKDES YVTCMLNPTNK--FMLFYKDAKDWWLFGFY	257
etlr_mouse_fragment	DNWLFGFYFCMLPVCTAIFYTLMTCEMLNRRNGSLRIALSEHLKQRREVA	207
etlr_pig	LAIPEAIGFVMVPFEYKGEENHKTCLMNATSK--FMEFYQDVKDWWLFGFY	263
etlr_rat	LAIPEAIGFVMVPFEYKGEQHRTCLMNATSK--FMEFYQDVKDWWLFGFY	263
etlr_rabbit_fragment	GEQHKTCLMNATSK--FMEFYQDVKDWWLFGFYFCMLPVCTAIFYTLMTC	263
etlr_sheep	LAIPEAIGFVMVPFEYKGAQHRTCLMNATSK--FMEFYQDVKDWWLFGFY	263
etlr_human	LAIPEAIGFVMVPFEYRGEQHKTCLMNATSK--FMEFYQDVKDWWLFGFY	263
etar_selectiv		55
etbr_selectiv		51
etbr_human	LAVPEAIGFDIITMDYKGSYLRICLLHPVQKTA FMQFYKTAKDWWLF SFY	281
etbr_bovin	LAVPEAVGFDIITS DHIGNKLRICLLHTQKTA FMQFYKTAKDWWLF SFY	280
etbr_canfa	LAVPEAVGFDMITIDYKGRLRICLLHTQKTA FMQFYKTAKDWWLF SFY	281
etbr_horse	LAVPEAVGFDMITADYKGSYLRICLLHTQKTA FMQFYKNAKDWWLF SFY	282
etbr_mouse	LAVPEAIGFDIMITSDYKGKPLRVCMNPFQKTA FMQFYKTAKDWWLF SFY	281
etbr_pig	LAVPEALGFDMITTSDYKGRLRICLLHTQKTA FMQFYKTAKDWWLF SFY	282
etbr_rat	LAVPEAIGFDVITS DYKGKPLRVCMNPFQKTA FMQFYKTAKDWWLF SFY	281
etbr_cotja_fragment	FALCWPLHLSRILKLTIYDQKDPNRCELLSFFLVMDYIGINMASLNSCI	287
etbr_macfa_fragment		99
etbr_rabbit	LAVPEAIGFNLVTIDYKGSYLRICLLNPTQKTA FMQFYKTAKDWWLF SFY	280
et_r_oryla	YMKFYQEAKDNWLFGFYFCLPLACTGIFYTLMCMLSRKKG-MRIALND	282
		19
OPSD	CAAPPLVGW----SRYIPEGMQCSCGIDYYTPH--EETNNESFVIYMFV	209
ruler260.....270.....280.....290.....300	

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/etr_builder.ps Date: Tue Oct 17 15:53:32 2006
 Page 4 of 5

etlr_bovin	FCMPLVCTAIFYTLMTCEMLNRRNGSLRIALSEHLKQRREVAKTVFCLVV	313
etlr_chicken	FCMPLACTAIFYTLMTCEMLNRRNSNLRIALSEHLKQRREVAKTVFCLVV	307
etlr_mouse_fragment	KTVFCLVVIFALCWFLHLSRILKKTVYDEMDKNRCELLSFLLMDYIGI	257
etlr_pig	FCMPLVCTAIFYTLMTCEMLNRRNGSLRIALSEHLKQRREVAKTVFCLVV	313
etlr_rat	FCMPLVCTAIFYTLMTCEMLNRRNGSLRIALSEHLKQRREVAKTVFCLVV	313
etlr_rabbit_fragment	EMLNRRNGSLRIALSEHLKQRREVAKTVFCLVVIFALCWFFLHLSRILKK	313
etlr_sheep	FCMPLVCTAIFYTLMTCEMLNRRNGSLRIALSEHLKQRREVAKTVFCLVV	313
etlr_human	FCMPLVCTAIFYTLMTCEMLNRRNGSLRIALSEHLKQRREVAKTVFCLVV	313
etar_selectiv		55
etbr_selectiv		51
etbr_human	FCLPLAITAFFYTLMTCEMLRKSG-MQIALNDHLKQRREVAKTVFCLVL	330
etbr_bovin	FCLPLAITALFYTLMTCEMLRKSG-MQIALNDHLKQRREVAKTVFCLVL	329
etbr_canfa	FCLPLAITAFFYTLMTCEMLRKSG-MQIALNDHLKQRREVAKTVFCLVL	330
etbr_horse	FCLPLAITAFFYLETCEMLRKSG-MQIALNDHLKQRREVAKTVFCLVL	331
etbr_mouse	FCLPLAITAVFYTLMTCEMLRKSG-MQIALNDHLKQRREVAKTVFCLVL	330
etbr_pig	FCLPLAITAFFYTLMTCEMLRKSG-MQIALNDHLKQRREVAKTVFCLVL	331
etbr_rat	FCLPLAITAIFYTLMTCEMLRKSG-MQIALNDHLKQRREVAKTVFCLVL	330
etbr_cotja_fragment	NPIALYLVSKRFQNCFKSCLCCWC-QS-KDLSLEERQSCLKFKANDHG	335
etbr_macfa_fragment		99
etbr_rabbit	FCLPLAITAFFYTLMTCEMLRKSG-MQIALNDHLKQRREVAKTVFCLVL	329
et_r_oryla	HMKQRREVAKTVFCLVLIFACWNLPLHLSRILKKTVYNENDPNRCELLSF	332
		19
OPSD	VHFIIPLIVIFFCYGQLVFTVKE--AAAQQQESATTQKAKEVTRMVIIM	257
ruler310.....320.....330.....340.....350	



etlr_bovin	IFALCWFPPLHLSRILKKTVYDEMDTNRCELLSFLLMDYIGINLATMNSC	363
etlr_chicken	IFALCWFPPLHLSRILKKMVYNERDPGRCELLSFLLPLDYISINLATMNSC	357
etlr_mouse_fragment	NLATMNSCINPIA	270
etlr_pig	IFALCWFPPLHLSRILKKTVYDEMDKNRCELLSFLLMDYIGINLATMNSC	363
etlr_rat	IFALCWFPPLHLSRILKKTVYDEMDKNRCELLSFLLMDYIGINLATMNSC	363
etlr_rabbit_fragment	TVYDEMDKNRCELLSFLLMDYIGINLATMNSCINPIALYFVSKPKNCF	363
etlr_sheep	IFALCWFPPLHLSRILKKTVYDEMDTNRCELLSFLLMDYIGINLATMNSC	363
etlr_human	IFALCWFPPLHLSRILKKTVYNEMDKNRCELLSFLLMDYIGINLATMNSC	363
etar_selectiv		55
etbr_selectiv		51
etbr_human	VFALCWPLPLHLSRILKLTYLNQNQDPNRCELLSFLLVLDYIGINMASLNSC	380
etbr_bovin	VFALCWPLPLHLSRILKLTYDQHDPRRCEFLSFLLVLDYIGINMASLNSC	379
etbr_canfa	VFALCWPLPLHLSRILKLTIYDQNDPNRCELLSFLLVLDYIGINMASLNSC	380
etbr_horse	VFALCWPLPLHLSRILKHTLYDQNDPHRCELLSFLLVLDYIGINMASLNSC	381
etbr_mouse	VFALCWPLPLHLSRILKLTYDQSNPCHRCELLSFLLVLDYIGINMASLNSC	380
etbr_pig	VFALCWPLPLHLSRILKLTYDQNDSNRCELLSFLLVLDYIGINMASLNSC	381
etbr_rat	VFALCWPLPLHLSRILKLTYDQSNPQRCELLSFLLVLDYIGINMASLNSC	380
etbr_cotja_fragment	DNFRSSNKYSSS	347
etbr_macfa_fragment		99
etbr_rabbit	VFGLCWLALHLSRILKLTYDQNDPNRCELLSFLLVLDYIGINMASLNSC	379
et_r_oryla	LLVMFYIGINMASLNSCINPIALYFVSQFKNCFQSCCLCCWC-Y---RTS	378
		19
OPSD	VIAFLICWL PYAGVAFYIFTHQGSD-----PGPIFM TI PAFFAKTS	298
ruler360.....370.....380.....390.....400	



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/etr_builder.ps Date: Tue Oct 17 15:53:32 2006
 Page 5 of 5

etlr_bovin	INPIALYFVSKFKNCFQSCLCCCCYQS	QSKSLMTSVP	MNGTSI	QWKNHEQN	413	
etlr_chicken	INPIALYFVSKFKNCFQSCLCCCC	SQS	QSKSLATSV	MNGTSI	QWKNQELN	407
etlr_mouse_fragment					270	
etlr_pig	INPIALYFVSKFKNCFQSCLCCCCY	QSKSLMTSVP	MNGTSI	QWKNHEQN	413	
etlr_rat	INPIALYFVSKFKNCFQSCLCCCC	QSKSLMTSVP	MNGTSI	QWKNQEQQ-	412	
etlr_rabbit_fragment	QSCLCCCCHQSKSLMTSVP	MNGTSI	QWKNH		393	
etlr_sheep	INPIALYFVSKFKNCFQSCLCCCCY	QSKSLMTSVP	MNGTSI	QWKNP	413	
etlr_human	INPIALYFVSKFKNCFQSCLCCCCY	QSKSLMTSVP	MNGTSI	QWKNH	413	
etar_selectiv					55	
etbr_selectiv					51	
etbr_human	INPIALYLVSKRFKNCFKSCLCCWC	-QSFEEKQS	LEEKQ	SCLKF	KANDHG	429
etbr_bovin	INPIALYLVSKRFKNCFKSCLCCWC	-QSFEEKQS	LEEKQ	SCLKF	KANDHG	428
etbr_canfa	INPIALYLVSKRFKNCFKSCLCCWC	-QSFEEKQS	LEEKQ	SCLKF	KAN	426
etbr_horse	INPIALYLVSKRFKNCFKWCLCCWC	-QSFEEKQS	LEEKQ	SCLKF	KANDHG	430
etbr_mouse	INPIALYLVSKRFKNCFKSCLCCWC	-QTFEEKQS	LEEKQ	SCLKF	KANDHG	429
etbr_pig	INPIALYLVSKRFKNCFKSCLCCWC	-QSFEEKQS	LEEKQ	SCLKF	KANDHG	430
etbr_rat	INPIALYLVSKRFKNCFKSCLCCWC	-QTFEEKQS	LEEKQ	SCLKF	KANDHG	429
etbr_cotja_fragment					347	
etbr_macfa_fragment					99	
etbr_rabbit	INPIALYLVSKRFKNCFKSCLCCWC	-QSFEEKQS	LEEKQ	SCLKF	KANDHG	428
et_r_oryla	PLDERGSGGRWKGSCQVNGLDRTSSR-SS				406	
					19	
<hr/>					342	
OPSD	AVYNPVIYIMMNKQFRNCMVTLCC	-----	GKNP-LGDDEASTTVSKTET			
ruler410.....420.....430.....440.....450					

etlr_bovin	NHNTERSSHKD\$IN	427
etlr_chicken	NHNTDRSSHKD\$IN	421
etlr_mouse_fragment		270
etlr_pig	NHNTERSSHKD\$IN	427
etlr_rat	NHNTERSSHKD\$MN	426
etlr_rabbit_fragment		393
etlr_sheep	NHNTERSSHKD\$IN	427
etlr_human	NHNTDRSSHKD\$MN	427
etar_selectiv		55
etbr_selectiv		51
etbr_human	YDNFRSSNKYSSS	442
etbr_bovin	YDNFRSSNKYSSS	441
etbr_canfa		426
etbr_horse	YDNFRSSNKYSSS	443
etbr_mouse	YDNFRSSNKYSSS	442
etbr_pig	YDNFRSSNKYSSS	443
etbr_rat	YDNFRSSNKYSSS	442
etbr_cotja_fragment		347
etbr_macfa_fragment		99
etbr_rabbit	YDNFRSSNKYSSS	441
et_r_oryla		406
		19
<hr/>		348
OPSD	SQVAPA	
ruler460....	

Multiple Sequence Alignment of Nicotinic Acid Receptors GPR109A/B

Because of the very recent de-orphanization of GPR109A only few experimental data were available. As a result, GPR109A and GPR109B were aligned against GPCRs that are intensively investigated.

HM74A – GPR109A	OPSD – rhodopsin	EDNRA – ETA receptor
HM74 – GPR109B	OxytocR – Oxytocin receptor	ETBR – ETB receptor
TSHR – Thyroid-stimulating hormone receptor	LSHR – Luteinizing hormone receptor	
FSHR – Follicle-stimulating hormone receptor	etc.	

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/hm74_gpcr.1.ps Date: Tue Oct 17 16:20:43 2006
Page 1 of 4

```

OPSD -----
HM74A_1 -----
HM74 -----
PUMA_G_mouse -----
FKSG80_hum -----
HUMEDNRA -----
ETBR_hum -----
V2R_human -----
OxytocR_hu -----
VasRV1A_hu -----
vibr_human_1 -----
grhr_human_2 -----
grhr_catf1_1
    tshru MRPADLLQLVLLLDLFRDLGNGCSSPPCECHQEEEDFRVICKDIQRIPSLLPPSTQTLKLIETHLRTIPSHAFSNLPNISR 80
    lshru MKQRFSAIQQ--LLKLLLLLQPPLPRA--LREALCPEPCN-CVPDGALRCPGP-TAGLTFLSLAYLPVKVIPSQAFRGLNE 74
    fshrum MALLLVSLLAFLSLGSGCCHHRICHCNSNVFLCQESKVTEIPSDLPRNAIELRFVLTKLIVIQKGAFSCFGDLEKIEISQN 80
    ruler 1.....10.....20.....30.....40.....50.....60.....70.....80

-----
OPSD -----
HM74A_1 -----
HM74 -----
PUMA_G_mouse -----
FKSG80_hum -----
HUMEDNRA -----
ETBR_hum -----
V2R_human -----
OxytocR_hu -----
VasRV1A_hu -----
vibr_human_1 -----
grhr_human_2 -----
grhr_catf1_1
    tshru IYVSIDVTLQQLESHSFYNLSKVTHIEIRNTRNLTYIDPDALNELLPLKKFLGIFNTGLNMFPDLTKVVSTDIFFILEITD 160
    lshru VIKIEISQIDSLERIEANAFDNILLNLSEILIQNTKLNRYIEPGQAFINLPGCLWYLISICNTGIREFPDVTKVFSSSENFILE 154
    fshrum DVLEVIEADVFSNLPKLMHEIRIEKANNLLYINPEAFQNLFNLQYLLISNTGIKHLPDVHNHISLQ-KVLLDIQDININT 159
    ruler .....90.....100.....110.....120.....130.....140.....150.....160

-----
OPSD -----
HM74A_1 -----
HM74 -----
PUMA_G_mouse -----
FKSG80_hum -----
HUMEDNRA -----
ETBR_hum -----
V2R_human -----
OxytocR_hu -----
VasRV1A_hu -----
vibr_human_1 -----
grhr_human_2 -----
grhr_catf1_1
    tshru NPYMTSIPVNAPQGLCNETLTLALYNNGFTSVQGYTAFNOTKLDAVILNKKNKLITVIDKDAFGCGVYSGPSLLDVSQTSVIA 240
    lshru ICDNLHITTIPGNAFQMNNESVTLKLYGNOFEEVQSHAFNGTTLTSLELKENVHLEKMHNGAFRGA-TGPKTLDISSTK 233
    fshrum IERNFSVULSFESVILWLNKNGIQEIHNCAFNGTQLDDELNLSDNNNLSEELPHDVFHGA-SGPVILDISRTPIHSLPSYGL 238
    ruler .....170.....180.....190.....200.....210.....220.....230.....240

```

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/hm74_gper.1.ps Date: Tue Oct 17 16:20:43 2006
Page 2 of 4

```

OPSD -----
HM74A_1 -----
HM74 -----
PUMA_G_mouse -----
FKSG60_hum -----
HUMEDNRA -----
ETBR_hum -----
V2R_human -----
OxytocR_hu -----
VasRV1A_hu -----
vibr_human_1 -----
grhr_human_2 -----
grhr_catf1_i -----
    tshrhru LPSKQGLEBLKELIARNTWILKKLPLSLSFLSHLTRADLSYFISHCCAFKNQKNIKGILESILMCNESSMQSLRQQRKSVNAALNS 320
    lshrhru LQALPSYQLESIQRLIATSSYSLKKLPSRETIVNLLERATLTYPSHCCAFRNLPTQEKFHSH-ISENFS--KQCEST 307
    fshrhru ENKLKKLRLARASTYLNKKLPLTKEVALMEASLYTPSHECCAFANMRQISELPPI-CNKSLIL--EQEVDYMT- 305
ruler .....250.....260.....270.....280.....290.....300.....310.....320

```

OPSD	TGVVRSPEA P	-Q-YYLAEPWQFS-MLAA YMFLLIMLQFFPINFVFLTYLVTVQH-----KKLRTPLNYILLNL	79
HM74_A_1	-----MNRHHLQDHFLIEIDKKNCCVFRDDFIWVKVLPPVLCGLEFIFGQLLGNGNGLALWIFCFH-----	LKSWKSSRIFLFNL	69
HM74	-----MNRHHLQDHFLIEIDKKNCCVFRDDFIWVKVLPPVLCGLEFIFGQLLGNGNGLALWIFCFH-----	LKSWKSSRIFLFNL	69
PUMA_G_mouse	-MSKSDHFLLVINGRNCCVFRDENIAKVLPPVLCLEFVF GQLLGNGNGLALWIFCFH-----	LKSWKSSRIFLFNL	66
FKSGS60_hum	MYNSQCCRIEOTDQSIVPMSPLLIVAFVFLQALGNQVALCGFCFHR-----MKYKWPSTVYLFLNL	MKYKWPSTVYLFLNL	57
HUMEDNA_R	LPSNGSMHNYC-----PQQNTIATSKY-INTVVISCTIFVGCWGNATLIIYIQN-----KCMRNPMLALISL	KCMRNPMLALISL	122
ETBR_hum	SPPRTTISPPPC-----QGPIEIKEKFCKY-INTVVSCLVPVLGIIIGNSTLLRIIYIN-----KCMRNPNIILIASL	KCMRNPNIILIASL	143
V2R_human	SPLSNSQSQERP-----LDTDRFLALAAEALLSIVFVAFLNNQSLVLAALRRGR-----RGHWAPBFVFIQHL	RGHWAPBFVFIQHL	81
OxytocR_bu	GAEGNRATACPP-----BRNEALAR-VEAVAVCLCLILLALSGNACVALLRHT-----RQKHSRLFFMMKL	RQKHSRLFFMMKL	81
VasRVIa_hu	AEGALGEGGNGFP-----RDVRNEELAK-LEIAVLAUTFAVAVLGNNSVLLALHHT-----PRATSRMHLFIRHL	PRATSRMHLFIRHL	93
vibr_human_1	TLTSAPNNAITP-----MLGRDEELAK-VEIGVLAUTVVLVLTAGGNLAVLLTLQGL-----GRKRSRLMFVFLHL	GRKRSRLMFVFLHL	76
grhr_human_2	-IPLM-----GNLPLTLLTSKCRIVTFVFLFLSATTFNNSFLKLQRKWTOKKEKGKLLSRMLLKKRL	GNLPLTLLTSKCRIVTFVFLFLSATTFNNSFLKLQRKWTOKKEKGKLLSRMLLKKRL	83
grhr_catfi_1	-VLNVSARSPV-----LKWEPTPTTAARPFWAAATLVLVFPFAAASLSLSSLVSVTRGR-----GRLASHRPLIASL	GRLASHRPLIASL	86
tshrbu	KSDEFNP C-----EDIMGYKFLRIVWVFWSLLALLGQNVFLVLLTSH-----KLNVPRFLMCNLAFA	KLNVPRFLMCNLAFA	459
lshrbu	CAPEDPAFNPC-----EDIMGYDPLFWLVLIMLNLIAIMGMNMTFLVFLVLLTSRY-----KLTVPFRFLMCNL	KLTVPFRFLMCNL	401
fshrbu	PC-----EDIMGYNILRVLWLFNISLAIYTONIIVLVLVLLTTSQY-----KLTVPFRFLMCNLAFAADLIGI	KLTVPFRFLMCNLAFAADLIGI	413
ruler	.410.....420.....430.....440.....450.....460.....470.....480		

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/hm74_gpcr.1.ps Date: Tue Oct 17 16:20:43 2006
 Page 3 of 4

OPSD	AVADLFMVFGGFTTLYTSLHGYFVFG-----	PTG-----	CNLEGFFATLGGIEIALMWSLVLAIERYVVVC	KPMNSNFRFG	149
HM74A_1	AVADFLLIICLPLFLMDNYVRRMDKFC-----	DIP-----	CRLMLFMLAMNNRQGSIIIFLT	VAVDRYFRVVHPHHALNKI	139
HM74	AVADFLLIICLPLFLMDNYVRRSDWFG-----	DIP-----	CRLVLFMAMNNRQGSIIIFLT	VAVDRYFRVVHPHHALNKI	139
PUMA_G_mouse	AVADFLLIICLPLTDNYVHNMDWFG-----	DIP-----	CRVMLFMLAMNNRQGSIIIFLT	VAVDRYFRVVHPHHFLNKI	136
FKSG80_hum	AVADFLLIICLPLTDYVLRRRHWAFG-----	DIP-----	CRVGLFLPFLQKSSVGIIVLNL	LCRALDVRRAVASWSRVQCIG	127
HUMEDNRA	ALGDLIYVVIDPINVFKNLLAGRNPFH-----	DIP-----	CRVGLFLPFLQKSSVGIIVLNL	LCRALDVRRAVASWSRVQCIG	197
ETBR_hum	ALGDLIYVVIDPINVFKNLLAGRNPFH-----	DIP-----	GAEMCKLWPFTQKASVGITVLSLCALS	IDRYRAVASWSRICKGIG	213
V2R_human	CLADFLVAVQVLPQLLWDITFRFKGP-----	DAL-----	CRAVKYLQMVCOMYASSYMLIL	DRHRAICRPMLAYRHG	151
OxytocR_hu	SIADLTVVAVQVLPQLLWDITFRFKGP-----	DLL-----	CRVLKYLQVVCFCMFASTYLLLML	LDRCLAICQPLRSLRRL	150
VasRVI1A_hu	SLADLAVAFQVLPQNCWDITYRFKGP-----	DNL-----	CRVVKHLQVFCMFASAYMLVMM	TADRYIAVCHPLKYLQQ-	162
vibr_human_1	ALTDLAVAFQVLPQNCWDITYRFKGP-----	DNL-----	CRVVKYLQVLCMFASAYMLVMM	TADRYIAVCHPLKYLQQ-	145
grhr_human_2	TLANLLETLIVMPLDGMWNTIVQWYAG-----	ELL-----	CKVLSYLFLFSMIAFAMMVWIS	LDRLSLAITRPL-ALKSN	152
grhr_catf1_1	ASADLFLMVFFVVMPLDAWNVNTVQWYAG-----	DAM-----	CKLMCFLNKFLAMHESAFWVSLD	RHHAILHPL-DTLDAA	155
tshru	DFCNGMYLLIASVUDLYTHSEYYNHAI	DAM-----	DKLQVWHDAMFLLEFFPLPQGII	NDWQTGPG-CNTAGFTTVFASELSVY	204
lshru	SFADFCMOLYLLIASVDSQTKHQYNNHAI	DAM-----	DKLQVWHDAMFLLEFFPLPQGII	DWQTGSG-CSTAGFTTVFASELSVY	156
fshrhmu	YLLLIASVDIHTRSQYHNYAIDWQTGAC-----	DAM-----	DKLQVWHDAMFLLEFFPLPQGII	CDANGFTTVFASELSVY	157
ruler490.....500.....510.....520.....530.....540.....550.....560				



OPSD	-ENHAIMGVAPINVMALACAAPPLVGMSRYIP-----	EGMQCSCCGIDYY-TPHEE-----	TNNESFVIYMFVVHFIIPLIVI		219
HM74A_1	SNRTAAIISCLLNGITIGLT	TVHLLKKRNPFIQN-----	GGANLC-SSFS-----	I CHTFWHEAMFLLEFFPLPQGII	204
HM74	SNNTAAIISCLLNGITV	GQVLT	GGANLC-SSFS-----	I CHTFWHEAMFLLEFFPLPQGII	204
PUMA_G_mouse	SNRTAAIISCLLNGITV	HLILYTNBTKN-----	GGANLC-SSFS-----	I CHTFWHEAMFLLEFFPLPQGII	201
FKSG80_hum	STRVAAGIVCTLWALVILQ	TVYVLLLNHL	CGANLC-SSFS-----	MEANSQWHDIMFQLEFFPMLPQGII	192
HUMEDNRA	IPLVTAIISIASIWL	FLILAPEAIGFVWVFE-----	CGANVC-ESFI-----	MESANQWHDIMFQLEFFPMLPQGII	271
ETBR_hum	VPKNTAVI	IVLINVVS	CGANVC-ESFI-----	YRGQHHTCMLNATSR-----	289
V2R_human	SGAHWNRPV	VLVAMAFS	CGANVC-ESFI-----	YKGSKYDQVTCMLNATSR-----	221
OxytocR_hu	SGAHWNRPV	VLVAMAFS	CGANVC-ESFI-----	YKGSKYDQVTCMLNATSR-----	216
VasRVI1A_hu	FARKSRLMIAAA	MVLSFVLS	CGANVC-ESFI-----	YKGSKYDQVTCMLNATSR-----	232
vibr_human_1	PGQSTYLLIAAP	MLLAISFLPQ	CGANVC-ESFI-----	YKGSKYDQVTCMLNATSR-----	215
grhr_human_2	SKVQPSMVGLAMIL	SSVFAQGPQLI	CGANVC-ESFI-----	YKGSKYDQVTCMLNATSR-----	227
grhr_catf1_1	GRBNRRMLLTAMIL	SLLLASPQLL	CGANVC-ESFI-----	YKGSKYDQVTCMLNATSR-----	225
tshru	HACAIMVGGWVCC	FLALLPLVQ	CGANVC-ESFI-----	YKGSKYDQVTCMLNATSR-----	598
lshru	RLRHALILMLGGWLF	SSLLIAMLP	CGANVC-ESFI-----	YKGSKYDQVTCMLNATSR-----	540
fshrhmu	VMGWIFAFAA	LLFPPIFGISSY	CGANVC-ESFI-----	YKGSKYDQVTCMLNATSR-----	552
ruler570.....580.....590.....600.....610.....620.....630.....640				



OPSD	FFCYGQLVFTVKE-----	AAAQQQE-----	SATTOQRAKEVTRMVIIMVIAFLICW		265
HM74A_1	LFCSARIISLWLRQ-----	-----	MDRHAKIKRAISFIMVVAIVFVICP		244
HM74	LFCSARIISLWLRQ-----	-----	MDRHAKIKRAISFIMVVAIVFVICP		244
PUMA_G_mouse	LFCSARIISLWLRQ-----	-----	MDRHAKIKRAISFIMVVAIVFVICP		241
FKSG80_hum	LFCSFKIVWSSLRQ-----	-----	MDRHAKIKRAISFIMVVAIVFVICP		233
HUMEDNRA	AIFYTILMTCEMLRKKS-----	-----	LARQARMKKATFIMVVAIVFVICY		319
ETBR_hum	AFFYTLMTCEMLRKKS-----	-----	GSLQIALSEHLNQRRREVAKTVFC	LVLVFALCW	336
V2R_human	AACQVLFIREIHASLV-----	-----	GMP=IALNDHLNQRRREVAKTVFC	LVLVFALCW	284
OxytocR_hu	ATCYGLISFKIMQNLILKT	AAAAAAE-----	GDPGRKVALARVSSVNLISRAKIR	TVFVMTFIIIVLAFIVCW	288
VasRVI1A_hu	GTCYGFICYNIMC	NVNRC	GDPGRKVALARVSSVNLISRAKIR	TVFVMTFIIIVLAFIVCW	304
vibr_human_1	TACYSLICH	EICKKNLKVKTQAMRVGGCW-----	RTWDRPSPSTLAATRGLPSV	VSSINTISRAKIRTVFVMTFIVL	294
grhr_human_2	LICHAKIIFT	TRVLHQDREL-----	RTWDRPSPSTLAATRGLPSV	VSSINTISRAKIRTVFVMTFIVL	280
grhr_catf1_1	SLC1	TYRILVEINRQMHRSKDRA-----	RTWDRPSPSTLAATRGLPSV	VSSINTISRAKIRTVFVMTFIVL	282
tshru	CCHVRIYIIT	TVRNPOY-----	RTWDRPSPSTLAATRGLPSV	VSSINTISRAKIRTVFVMTFIVL	640
lshru	IICACYI	YIYFAVRNPELMAT-----	RTWDRPSPSTLAATRGLPSV	VSSINTISRAKIRTVFVMTFIVL	582
fshrhmu	YLT	TVRNPHIVSS-----	RTWDRPSPSTLAATRGLPSV	VSSINTISRAKIRTVFVMTFIVL	594
ruler650.....660.....670.....680.....690.....700.....710.....720				



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/hm74_gpcr.1.ps Date: Tue Oct 17 16:20:43 2006
 Page 4 of 4

OPSD	L PY A G V A F Y I F ----- T H Q G S D F C P ----- I F M T I P A F F A K T S A V Y N P V I Y Y I M M N K Q F R N C M V - T T L C C - G K H P ----- L G - D	330
HM74A_1	L P S V V V R I R I F W L L H T S G T Q N C E V Y R S V D L A F F I T L S F T Y M N S M L D P V V Y Y F S S P S F P M F F S T L I N R C L Q R M T G E P O N N	324
HM74	L P S V V V R I R I F W L L H T S G T Q N C E V Y R S V D L A F F I T L S F T Y M N S M L D P V V Y Y F S S P S F P M F F S T L I N R C L Q R M T G E P O N N	324
PUMA_G_mouse	L P S V A V R I R I F W L L Y K Y N V R N C D I Y S S V D L A F F I T L S F T Y M N S M L D P V V Y Y F S S P S F P M F F S T C I N R C L R K K T L G E P O N N	321
FKSG80_hum	L P S V S A R L Y F L M N V P P S ----- A C D P S V H G A L H I T L S F T Y M N S M L D P L V Y Y F S S P S F P M F F S T C I N R C L R K K T L G E P O N N	308
HUMEDNRA	F P L H L S R L I N K I T V Y N E M D K N R C E L L S F I L L M D Y I G I N L A T M N S C I N P I A L Y F V S K K F K N C F Q S C L C C C W C Q S F E E K Q S - L E	398
ETBR_hum	L P L H L S R L I N K I T V Y N E M D K N R C E L L S F I L L M D Y I G I N M A S L N S C I N P I A L Y L V S K R F K N C F Q S C L C C C W C Q S F E E K Q S - L E	415
V2R_human	A P F F L V Q L W A A M ----- D P E A P ----- L E G A P F V V G A L H I T L S F T Y M N S M L D P L V Y Y F S S P S F P M F F S T C I N R C L R K K T L G E P O N N	354
OxytocR_hu	T P F F F V Q M W S V M ----- D A R A P ----- K E A S A F I I V V M L L A S L N S C C N P M I Y M L F T C H L F H E L V O R F L C C S A S Y L K G R R L G -	358
VasRVIa_hu	A P F F I I Q M W S V M ----- D P M S V W T E S E N P T I T I T I T A L L G S L N S C C N P M I Y M F F S C H L L Q D C V Q S F P C C Q N M K E K F N K E D T	378
vibr_human_1	A P F F I I Q M W S V M ----- D K N A P D E D S T V A F T I S M L L G M L N S C C N P M I Y M F F S C H L L Q D C V Q S F P C C Q N M K E K F N K E D T	368
grhr_human_2	T P Y V L G I W Y W F ----- D P M S V W T E S E N P T I T I T I T A L L G S L N S C C N P M I Y M F F S C H L L Q D C V Q S F P C C Q N M K E K F N K E D T	328
grhr_catf1_1	T P Y I L L G I W Y W F ----- Q P Q M L - H V I P D Y V H H V F F V G M L N T C C D F V I Y G F T P S F R A D L S R C F C W N N Q N A S A K S L P H F	355
tshru	S F Y A L S A I L N K ----- P L I T V S N S K I L L V L F Y P L N S C A N F P L Y A I F T K T F Q R D F L L L S K F G C C R R A E L Y R R	711
lshru	A P I S F F A I S A A F K V ----- P L I T V T I N S K V L L V L F Y P I N S C A N F P L Y A I F T K T F Q R D F L L L S K F G C C R R A E L Y R R	653
fshrhmu	A S L K V ----- P L I T V S R A K I L L V L F E P I N S C A N F P L Y A I F T K T F Q R D F L L L S K F G C C R R A E L Y R R	665
ruler730.....740.....750.....760.....770.....780.....790.....800	



OPSD	D E A S T T V S --- K T E T S Q V A P A	348
HM74A_1	R S T S V E L T G D P N K T R G A P E A L M A N S G E P W S P S Y L G P T S P	363
HM74	R S T S V E L T G D P N K T R G A P E A L M A N S G E P W S P S Y L G P T S P	387
PUMA_G_mouse	R S T S V E L T G D P S T T R S I P G A L M A D P S E P G S P P Y L G P T S P	360
FKSG80_hum	R S T S V E L T G D P S T T R S I P G A L M A D P S E P G S P P Y L G P T S P	346
HUMEDNRA	F P M N G T S I Q W K N H D Q Q N N H N T D R S S H K D S M N	427
ETBR_hum	E K Q S C L K F X A N D H G Y D N F - R S S N H Y S S S	442
V2R_human	D E S C T T A S S L A N D T S S	371
OxytocR_hu	E T S A S K K S N S S S F V L S H R S S S Q R S C S Q ----- P S T A	389
VasRVIa_hu	D S - M S R R Q - T F Y S N N R S P T N S T G M - W R D S P K S S K S I K F I P V S T	418
vibr_human_1	D S L S S R H T T L	378
grhr_human_2	D S L S S R H T T L	328
grhr_catf1_1	S G H R R E V S G E A E S D L G S G D Q P S G Q	379
tshru	F P K N S T D I Q V Q R V T H O M R Q C L H N M E D V Y E L I E N S H L T P K K Q Q I S E E Y M Q T V L	764
lshru	K D F S - A Y T S N C R N G F T G S S N K P S Q S T L L S T L H C Q G T A L L D R T Y I T E C	699
fshrhmu	H P R N G H C S S A P R V T N G S T V I L V P L S H L A Q N	695
ruler810.....820.....830.....840.....850.....860...	



Multiple Sequence Alignment of G α Proteins

Investigations of selectivity patterns for G protein interaction have been done by using sequence and structural data of several G protein families. The codes on the left resemble the entries of SwissProt database.

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/Ga_ali2.ps

Date: Tue Oct 17 16:03:52 2006

Page 1 of 6

```

gpii_human -----MGCTLSAEDKAAVERSKIMIDRLNLRDEGEKAAREVNLLLLAGESGKSTIVKQMKII 56
gbq_human -----MAC CLESEEAKEARINDEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 56
gbq_mouse -----MAC CLESEEAKEARINDEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 56
gbq_xenla -----MAC CLESEEAKEARINDEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 56
    lgg2 -----LSAEDKAAVERSKIMIDRLNLRDEGEKAAREVNLLLLAGESGKSTIVKQMKII 52
1yp2_template -----LASEDKAAVERSKIMIDRLNLRDEGEKAAREVNLLLLAGESGKSTIVKQMKII 52
gbil_rat -----MGCTLSAEDKAAVERSKIMIDRLNLRDEGEKAAREVNLLLLAGESGKSTIVKQMKII 56
gbil_xenla -----MGCTLSAEDKAAVERSKIMIDRLNLRDEGEKAAREVNLLLLAGESGKSTIVKQMKII 56
gbil3_human -----MGCTLSAEDKAAVERSKIMIDRLNLRDEGEKAAREVNLLLLAGESGKSTIVKQMKII 56
    gbi3_rat -----MGCTLSAEDKAAVERSKIMIDRLNLRDEGEKAAREVNLLLLAGESGKSTIVKQMKII 56
gbil3_xenla -----REAAEERSKIMIDRLNLRDEGEASKEVNLLLLAGESGKSTIVKQMKII 47
gbil2_mouse -----MGCTVSaedKAAMERSKIMIDKLNLREDGEKAAREVNLLLLAGESGKSTIVKQMKII 56
    gbi2_rat -----MGCTVSaedKAAMERSKIMIDKLNLREDGEKAAREVNLLLLAGESGKSTIVKQMKII 56
gbil2_human -----MGCTVSaedKAAMERSKIMIDKLNLREDGEKAAREVNLLLLAGESGKSTIVKQMKII 56
gbaz_human -----MGCRQSSSEEKAJARRSRRIDDRHLRSESSQQRREIKLLLLGTNSGKSTIVKQMKII 56
    gbaz_rat -----MGCRQSSSEEKAJARRSRRIDDRHLRSESSQQRREIKLLLLGTNSGKSTIVKQMKII 56
gb01_mouse -----MGCTLSAEEAALERSKAIENKLKEDGISIAAKDVNLLLLAGESGKSTIVKQMKII 56
    gb01_rat -----MGCTLSAEEAALERSKAIENKLKEDGISIAAKDVNLLLLAGESGKSTIVKQMKII 56
gb01_human -----MGCTLSAEEAALERSKAIENKLKEDGISIAAKDVNLLLLAGESGKSTIVKQMKII 56
gb02_mouse -----MGCTLSAEEAALERSKAIENKLKEDGISIAAKDVNLLLLAGESGKSTIVKQMKII 56
    gb02_rat -----MGCTLSAEEAALERSKAIENKLKEDGISIAAKDVNLLLLAGESGKSTIVKQMKII 56
gb02_criilo -----MGCTLSAEEAALERSKAIENKLKEDGISIAAKDVNLLLLAGESGKSTIVKQMKII 56
gb02_human -----MGCTLSAEEAALERSKAIENKLKEDGISIAAKDVNLLLLAGESGKSTIVKQMKII 56
gb02_drome -----MGCTTSAAEPAAIQSNEIQLKEAIDKQIAANDIKLLLLAGESGKSTIVKQMKII 56
gbt1_bovin -----MGAGASAEKK----HSELEKKLKDADKEADARVNLLLLAGESGKSTIVKQMKII 52
gbt1_mouse -----MGAGASAEKK----HSELEKKLKDADKEADARVNLLLLAGESGKSTIVKQMKII 52
gbt1_human -----MGAGASAEKK----HSELEKKLKDADKEADARVNLLLLAGESGKSTIVKQMKII 52
gbt2_bovin -----LENQTHGSASAEDKEELAKRSKELEKKLQEDADKEAVNLLLLAGESGKSTIVKQMKII 61
gbt2_mouse -----MGSCASAEKK----HSELEKKLKDADKEADARVNLLLLAGESGKSTIVKQMKII 56
gbt2_human -----MGSCASAEKK----HSELEKKLKDADKEADARVNLLLLAGESGKSTIVKQMKII 56
gbt2_bovin -----MGSGASAEKK----HSELEKKLKDADKEADARVNLLLLAGESGKSTIVKQMKII 56
gbt3_mouse -----MGSGISSESKEASRSKKELEKKLQEDADKEAVNLLLLAGESGKSTIVKQMKII 56
    gbt3_rat -----MTLESMMAC-----CLSDDEVKESKRINAIEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 62
gb11_bovin -----MTLESMMAC-----CLSDDEVKESKRINAIEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 62
gb11_mouse -----MTLESMMAC-----CLSDDEVKESKRINAIEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 62
gb11_human -----MAC CLESEEAKEARINDEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 56
gbq_human_1 -----MAC CLESEEAKEARINDEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 56
gbq_mouse_1 -----MAC CLESEEAKEARINDEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 56
gbq_xenla_1 -----MAC CLESEEAKEARINDEIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 56
gb14_mouse -----MA GCC CLSAEEEKESQRISAIIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 58
gb14_bovin -----MA GCC CLSAEEEKESQRISAIIEQLRRDKDARRELKLNLLLGTGESGKSTFIQKMRRII 58
gb15_human -----MARSLTWCCP-----WCLTEDEJAAAVIDQEINNILLQEKKQDQGELKLNLLLGPGECSGKSTFIQKMRRII 65
gb15_mouse -----MARSLTWCCP-----WCLTEDEJAAAVIDQEINNILLQEKKQDQGELKLNLLLGPGECSGKSTFIQKMRRII 65
gb12_human -----MSGVVURTLSCRLPFAEAGGARERRAGSGARDAREAMRSRIDALLARERAVRALVKILLNGAGESGKSTFLKQMRRII 80
gb12_mouse -----MSGVVURTLSCRLPFAEAGGARERRAGSGARDAREAMRSRIDALLARERAVRALVKILLNGAGESGKSTFLKQMRRII 78
gb13_mouse -----MADFLP--SRSVLSCVFP-----GC_VLTNGEAEQQRKSKEDICLRSERKTYVKLVKILLNGAGESGKSTFLKQMRRII 71
gbas_human -----MGCLGNSK-----TEDQNEEKAQREANNNIEKQLQKDQKVYRATHRLLLLAGESGKSTIVKQMRRII 63
gbas_bovin -----MGCLGNSK-----TEDQNEEKAQREANNNIEKQLQKDQKVYRATHRLLLLAGESGKSTIVKQMRRII 63
gbas_mouse -----MGCLGNSK-----TEDQNEEKAQREANNNIEKQLQKDQKVYRATHRLLLLAGESGKSTIVKQMRRII 63
    gbas_pig -----MGCLGNSK-----TEDQNEEKAQREANNNIEKQLQKDQKVYRATHRLLLLAGESGKSTIVKQMRRII 63
    gbas_xenla -----MGCLGNSK-----TEDQNEEKAQREANNNIEKQLQKDQKVYRATHRLLLLAGESGKSTIVKQMRRII 63
gbolf_human -----MGCLGNSK-----TTEDQVDEKERREANNNIEKQLQKDQKVYRATHRLLLLAGESGKSTIVKQMRRII 65
gbolf_rat -----MGCLGNSK-----TAEDQVDEKERREANNNIEKQLQKDQKVYRATHRLLLLAGESGKSTIVKQMRRII 65
    ruler 1.....10.....20.....30.....40.....50.....60.....70.....80

```



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/Ga_ali2.ps
Page 2 of 6

Date: Tue Oct 17 16:03:52 2006

```

*   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *
gb1l_human HEAGYSEE-----ECKQYKAVVYSNTIQSIIAIIRAMGRLK--IDFGDSARADDARQLFVLAGAAEE--- 116
gbq_human HGCGYSDE-----DKRGFTKLVYQNIIFTANQAMIRAMDTLK--IPYKYEHNKAHAQQLVREVDVEKVSafe 119
gbq_mouse HGCGYSDE-----DKRGFTKLVYQNIIFTANQAMIRAMDTLK--IPYKYEHNKAHAQQLVREVDVEKVSafe 119
gbq_xenla HGCGYSDE-----DKRGFTKLVYQNIIFTANQAMIRAMDTLK--IPYKYEHNKAHAQQLVREVDVEKVSafe 119
gbq_xenla_l HGCGYSDE-----DKRGFTKLVYQNIIFTANQAMIRAMDTLK--IPYKYEHNKAHAQQLVREVDVEKVSafe 119
lgg2 HEAGYSEE-----ECKQYKAVVYSNTIQSIIAIIRAMGRLK--IDFGDSARADDARQLFVLAGAAEE--- 112
1gp2_template HEAGYSEE-----ECKQYKAVVYSNTIQSIIAIIRAMGRLK--IDFGDSARADDARQLFVLAGAAEE--- 112
gbil_rat HEAGYSEE-----ECKQYKAVVYSNTIQSIIAIIRAMGRLK--IDFGDSARADDARQLFVLAGAAEE--- 116
gbil_xenla HEAGYSEE-----ECKQYKAVVYSNTIQSIIAIIRAMGRLK--IDFGDSARADDARQLFVLAGAAEE--- 116
gbii_human HEDGYSEE-----ECKQYKAVVYSNTIQSIIAIIRAMGRLK--IDFGDSARADDARQLFVLAGAAEE--- 116
gbii_xenla HEDGYSEE-----ECKQYKAVVYSNTIQSIIAIIRAMGRLK--IDFGDSARADDARQLFVLAGAAEE--- 116
gbii_xenla_l HEDGYSEE-----ECKQYKAVVYSNTIQSIIAIIRAMGRLK--IDFGDSARADDARQLFVLAGAAEE--- 116
gbii3_rat HEDGYSEE-----ECKQYKAVVYSNTIQSIIAIIRAMGRLK--IDFGDSARADDARQLFVLAGAAEE--- 116
gbii3_xenla HEDGYSEE-----ECRQYKAVVYSNTIQSIIAIIRAMGRLR--IDFGDVARADDARQLFVLAGAAEE--- 107
gbii2_mouse HEDGYSEE-----ECRQYKAVVYSNTIQSIIAIIRAMGRLQ--IDFADPQRADDARQLFALSCAAEEQ--- 117
gbii2_rat HEDGYSEE-----ECRQYKAVVYSNTIQSIIAIIRAMGRLQ--IDFADPQRADDARQLFALSCAAEEQ--- 117
gbii2_human HEDGYSEE-----ECRQYKAVVYSNTIQSIIAIIRAMGRNLQ--IDFADPQRADDARQLFALSCAAEEQ--- 117
gbaz_human HSGGFNL-----ACKEYNPLIIYNAIDSILTRIIRALNALR--IDFHNPDRADDAVOLFALTGPAESK--- 117
gbaz_rat HSGGFNL-----ACKEYNPLIIYNAIDSILTRIIRALNALR--IDFHNPDRADDAVOLFALTGPAESK--- 117
gb01_mouse HEDGFSGE-----DVKQYKPVVYSNTIQSIIAIIRAMDTLG--VEYGDKERKDSDHVKCDVVSRMEDT--- 117
gb01_rat HEDGFSGE-----DVKQYKPVVYSNTIQSIIAIIRAMDTLG--VEYGDKERKDSDHVKCDVVSRMEDT--- 117
gb01_human HEDGFSGE-----DVKQYKPVVYSNTIQSIIAIIRAMDTLG--VEYGDKERKDSDHVKCDVVSRMEDT--- 117
gb02_mouse HEDGFSGE-----DVKQYKPVVYSNTIQSIIAIIRAMDTLG--VEYGDKERKDSDHVKCDVVSRMEDT--- 117
gb02_rat HEDGFSGE-----DVKQYKPVVYSNTIQSIIAIIRAMDTLG--VEYGDKERKDSDHVKCDVVSRMEDT--- 117
gb02_cribo HEDGFSGE-----DVKQYKPVVYSNTIQSIIAIIRAMDTLG--VEYGDKERKDSDHVKCDVVSRMEDT--- 117
gb02_human HEDGFSGE-----DVKQYKPVVYSNTIQSIIAIIRAMDTLG--VEYGDKERKDSDHVKCDVVSRMEDT--- 117
gb02_drome HESOFTAE-----DFKQYKPVVYSNTIQSIIAIIRAMDTLS--IQYSNNERESDANMVFDVCQRMHD--- 117
gbt1_bovin HQDGYSLE-----ECLEFIAIYYGNTLQSIILAIVRAMTTLN--IQYGDSSARQDDARFLMHMADTIEE--- 112
gbt1_mouse HQDGYSLE-----ECLEFIAIYYGNTLQSIILAIVRAMTTLN--IQYGDSSARQDDARFLMHMADTIEE--- 112
gbt1_human HQDGYSLE-----ECLEFIAIYYGNTLQSIILAIVRAMTTLN--IQYGDSSARQDDARFLMHMADTIEE--- 112
gbt2_bovin HQDGYSPE-----ECLEFKAIYYGVVLQSIILAIIRAMPTLG--IDYAEPSCADDGRQLNSNLADSIEE--- 121
gbt2_human HQDGYSPE-----ECLEFKAIYYGVVLQSIILAIIRAMPTLG--IDYAEPSCADDGRQLNSNLADSIEE--- 116
gbt2_mouse HQDGYSPE-----ECLEFKASIVYGVNLQSIILAIVRAMSTLG--IDYAEPSCADAGRQLNSNLADSIEE--- 116
gbt3_rat HKNGYSKQ-----ECMFKAVVYSNTIQSIIAIIVRAMTTLG--IDYVNPRSRSEDQQLLSMLNTLED--- 116
gbt1_bovin HGAGYSEE-----DKRGFTKLVYQNIIFTANQAMIRAMETLK--ILYKYEQNKNANALLIREVDVEKVTTFE 125
gbt1_mouse HGAGYSEE-----DKRGFTKLVYQNIIFTANQAMIRAMETLK--ILYKYEQNKNANALLIREVDVEKVTTFE 125
gbt1_human HGAGYSEE-----DKRGFTKLVYQNIIFTANQAMIRAMETLK--ILYKYEQNKNANALLIREVDVEKVTTFE 119
gbq_human_1 HGCGYSDE-----DKRGFTKLVYQNIIFTANQAMIRAMDTLK--IPYKYEHNKAHAQQLVREVDVEKVSafe 119
gbq_mouse_1 HGCGYSDE-----DKRGFTKLVYQNIIFTANQAMIRAMDTLK--IPYKYEHNKAHAQQLVREVDVEKVSafe 119
gbq_xenla_1 HGCGYSDE-----DKRGFTKLVYQNIIFTANQAMIRAMDTLK--IPYKYEHNKAHAQQLVREVDVEKVSafe 119
gb14_mouse HGCGYSDE-----DKRGFTKLVYQNIIFTANQAMIRAMDTLK--IQYMCEQNKEENAQIIIREVEVDKVTALS 121
gb14_bovin HGCGYSDE-----DKRGFTKLVYQNIIFTANQAMIRAMDTLK--IQYCEQNKEENAQIIREVEVDKVTALS 121
gb15_human HGAGYSEE-----ERKAFPLVYQNIIFTANQAMIRAMDTLK--IQYCEQNKEENAQIIREVEVDKVTALS 128
gb15_mouse HGAGYSEE-----DRRAFRLLYQNIIFTANQAMIRAMDTLK--IPFSRDPDKQHASLVMTOQDPYKVSTFE 128
gb12_human HGREFDQK-----ALLEFRDTIFDNLKGSRVLVBDARDKLG--IPWQYSENENHGMPFLMAFENKAGLP-- 141
gb12_mouse HGREFDQK-----ALLEFRDTIFDNLKGSRVLVBDARDKLG--IPWQHSENENHGMPFLMAFENKAGLP-- 139
gb13_mouse HQQDFDQR-----AREEFPRPIIYSNVIKGNRSLVLDAREKLH--IPWQHSENENHGMPFLMAFENKAGLP-- 134
gbas_human HVNGFNGE-----PQAARSNSDGEKATVQD1KNNLKEA1ET1VAAMSNLVPPVELANPENQFRVDYLILSVMNVPDFD-- 141
gbas_bovin HVNGFNGE-----PQAARSNSDGEKATVQD1KNNLKEA1ET1VAAMSNLVPPVELANPENQFRVDYLILSVMNVPDFD-- 141
gbas_mouse HVNGFNGE-----PQAARSNSDGEKATVQD1KNNLKEA1ET1VAAMSNLVPPVELANPENQFRVDYLILSVMNVPDFD-- 141
gbas_pig HVNGFNGD-----EKATVQD1KNNLKEA1ET1VAAMSNLVPPVELANPENQFRVDYLILSVMNVPDFD-- 126
gbas_xenla HVNGFNAE-----EKKTKVQD1KNNLKEA1ET1VAMGNLSPVVELANPENQFRVDYLILSVMNVPDFD-- 126
gbolf_human HVNGFNPE-----EKKQNLDIRKNUKDAIVTIVSAMSTIIPPVPLANPENQFRSDYIKSIAPITDFE-- 128
gbolf_rat HVNGFNPE-----EKKQNLDIRKNUKDALVTIISAMSTIIPPVPLANPENQFRSDYIKSIAPITDFE-- 128
ruler .....90.....100.....110.....120.....130.....140.....150.....160

```



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/Ga_ali2.ps
Page 3 of 6

Date: Tue Oct 17 16:03:52 2006

gbil_human	----GFM T AELAGVIKRLNW D SGVQACFNR S REYQLNDSAA Y YLNDLDR I AQPN H YIPTQ Q DVL R TR V T T GIVE E HT F PK	192
gbq __ human	----NPYVDAIKS L W N DPG I Q E CYD R REYQLSD S T K Y L NDLDR V ADP A Y L P T Q Q DVL R VR V T T G I I E Y P FDL Q	191
gbq __ mouse	----NPYVDAIKS L W N DPG I Q E CYD R REYQLSD S T K Y L NDLDR V ADP S Y L P T Q Q DVL R VR V T T G I I E Y P FDL Q	191
gbq __ xenla	----NPYVDAIKS L W N DPG I Q E CYD R REYQLSD S T K Y L NDLDR V ADP S Y L P T Q Q DVL R VR V T T G I I E Y P FDL Q	191
lgg2	----GFM T AEL N AGVIKRLNW D SGVQACFNR S REYQLNDSAA Y YLNDLDR I AQPN H YIPTQ Q DVL R TR V T T GIVE E HT F PK	188
1gp2 __ template	----GFM T AEL N AGVIKRLNW D SGVQACFNR S REYQLNDSAA Y YLNDLDR I AQPN H YIPTQ Q DVL R TR V T T GIVE E HT F PK	188
gbil __ rat	----GFM T AEL N AGVIKRLNW D SGVQACFNR S REYQLNDSAA Y YLNDLDR I AQPN H YIPTQ Q DVL R TR V T T GIVE E HT F PK	192
gbil __ xenla	----GFM T AEL N AGVIKRLNW D GGVQACFNR S REYQLNDSAA Y YLNDLDR I AQNSY I PTQ Q DVL R TR V T T GIVE E HT F PK	192
gbil __ human	----GFM T AEL N AGVIKRLNW D GGVQACFNR S REYQLNDSAA Y YLNDLDR I AQNSY I PTQ Q DVL R TR V T T GIVE E HT F PK	192
gbil __ rat	----GFM T AEL N AGVIKRLNW D GGVQACFNR S REYQLNDSAA Y YLNDLDR I AQNSY I PTQ Q DVL R TR V T T GIVE E HT F PK	192
gbil __ xenla	----GFM T AEL N AGVIKRLNW D GGVQACFNR S REYQLNDSAA Y YLNDLDR I AQNSY I PTQ Q DVL R TR V T T GIVE E HT F PK	192
gbil __ mouse	----GFM T AEL N AGVIKRLNW D GGVQACFNR S REYQLNDSAA Y YLNDLDR I AQNSY I PTQ Q DVL R TR V T T GIVE E HT F PK	192
gbil __ rat	----GFM T AEL N AGVIKRLNW D GGVQACFNR S REYQLNDSAA Y YLNDLDR I AQNSY I PTQ Q DVL R TR V T T GIVE E HT F PK	192
gbil __ xenla	----GFM T AEL N AGVIKRLNW D GGVQACFNR S REYQLNDSAA Y YLNDLDR I AQNSY I PTQ Q DVL R TR V T T GIVE E HT F PK	192
gbil __ human	----GFM T AEL N AGVIKRLNW D GGVQACFNR S REYQLNDSAA Y YLNDLDR I AQNSY I PTQ Q DVL R TR V T T GIVE E HT F PK	192
gbaz __ rat	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gbaz __ mouse	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gb01 __ rat	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gb01 __ rat	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gb01 __ human	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gb02 __ mouse	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gb02 __ rat	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gb02 __ crilo	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gb02 __ human	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gb02 __ drome	----GFM T AEL N AGVIKRLNW D GGVQACFNR S SEYHLEDNA A YY L NDLDR I AQD P YIPTVEDIL R SRDM T GIVENKF PK	193
gbt1 __ bovin	----GTM P KEM S IIQRLNW D DSG I Q E CFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	188
gbt1 __ mouse	----GTM P KEM S IIQRLNW D DSG I Q E CFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	188
gbt1 __ human	----GTM P KEM S IIQRLNW D DSG I Q E CFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	188
gbt2 __ bovin	----GTM P PELV E VI R RLNW D GGVQACFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	197
gbt2 __ human	----GTM P PELV E VI R RLNW D GGVQACFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	197
gbt2 __ mouse	----GTM P PELV E VI R RLNW D GGVQACFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	197
gbt3 __ rat	----GDT P ELLV D VIR R LNWDGGVQACFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	192
gb11 __ bovin	----HRYVSAIK T RLNW D DSG I Q E CFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	197
gb11 __ mouse	----HRYVSAIK T RLNW D DSG I Q E CFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	197
gb11 __ human	----HRYVSAIK T RLNW D DSG I Q E CFDR A EYQLNDSAKY L LDLDR I AQD P YIPTQ Q DIL R TR V T T GIVE E HF PK	197
gbq __ human __ 1	----NPYVDAIKS L W N DPG I Q E CYD R REYQLSD S T K Y L NDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	191
gbq __ mouse __ 1	----NPYVDAIKS L W N DPG I Q E CYD R REYQLSD S T K Y L NDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	191
gbq __ xenla __ 1	----NPYVDAIKYLNW D DPG I Q E CYD R REYQLSD S T K Y L NDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	191
gb14 __ mouse	----RDQVAAIKQLNLDPG I Q E CYD R REYQLSD S AKY L LDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	193
gb14 __ bovin	----RDQVEAIKQLNLDPG I Q E CYD R REYQLSD S AKY L LDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	193
gb15 __ human	----KRYAAAMQYLNW D DSG I Q E CFDR A EYQLSD S AKY L LDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	200
gb15 __ mouse	----KPYAVAMQYLNW D DSG I Q E CFDR A EYQLSD S AKY L LDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	200
gb12 __ human	----VFPATFQLY V PAL A LNWL D DSG I Q E CFDR A EYQLSD S AKY L LDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	219
gb12 __ mouse	----VFPATFQLY V PAL A LNWL D DSG I Q E CFDR A EYQLSD S AKY L LDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	217
gb13 __ mouse	----GMVETR V FLQ Y PAIK A LNWL D DSG I Q E CFDR A EYQLSD S AKY L LDLDR I AQD P YIPTQ Q DVL R TR V T T GIVE E HF PK	214
gbas __ human	----FPPEFYEH A LNWL D DEGV V RA C YERSNEYQLIDC A Q Y FLD K ID V I Q DD V Y P SP Q DL L RC V L T SC I FE T KFQ V D	215
gbas __ bovin	----FPPEFYEH A LNWL D DEGV V RA C YERSNEYQLIDC A Q Y FLD K ID V I Q DD V Y P SP Q DL L RC V L T SC I FE T KFQ V D	215
gbas __ mouse	----FPPEFYEH A LNWL D DEGV V RA C YERSNEYQLIDC A Q Y FLD K ID V I Q DD V Y P SP Q DL L RC V L T SC I FE T KFQ V D	215
gbas __ pig	----FPPEFYEH A LNWL D DEGV V RA C YERSNEYQLIDC A Q Y FLD K ID V I Q DD V Y P SP Q DL L RC V L T SC I FE T KFQ V D	200
gbas __ xenla	----FPPEFYEH A LNWL D DEGV V RA C YERSNEYQLIDC A Q Y FLD K ID V I Q DD V Y P SP Q DL L RC V L T SC I FE T KFQ V D	200
gbolf __ human	----YSQEFFDHVKKLW D DEGV V RA C YERSNEYQLIDC A Q Y FLD K ID V I Q DD V Y P SP Q DL L RC V L T SC I FE T KFQ V D	202
gbolf __ rat	----YSQEFFDHVKKLW D DEGV V RA C YERSNEYQLIDC A Q Y FLD K ID V I Q DD V Y P SP Q DL L RC V L T SC I FE T KFQ V D	202
ruler170.....180.....190.....200.....210.....220.....230.....240	



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/Ga_ali2.ps
Page 4 of 6

Date: Tue Oct 17 16:03:52 2006



CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/Ga_ali2.ps
Page 5 of 6

Date: Tue Oct 17 16:03:52 2006

gbil_human	LFEEKIKK-S	PLTICYPEYA-	-GS-NTYEEAAAYIQCQFEDLN	311
gbq_human	LLEEKIMY-S	HLVDYFPEYD-	-GPQRDAQAAREFILKMFVDLN	311
gbq_mouse	LLEEKIMY-S	HLVDYFPEYD-	-GPQRDAQAAREFILKMFVDLN	311
gbq_xenla	LLEEKIMY-S	HLVDYFPEYD-	-GPQRDAQAAREFILKMFVDLN	311
1gg2	LFEEKIKK-S	PLTICYPEYA-	-GS-NTYEEAAAYIQCQFEDLN	307
1gp2_template	LFEEKIKK-S	PLTICYPEYA-	-GS-NTYEEAAAYIQCQFEDLN	307
gbil_rat	LFEEKIKK-S	PLTICYPEYA-	-GS-NTYEEAAAYIQCQFEDLN	311
gbil_xenla	LFEEKIKR-S	PLTICYPEYP-	-GS-NTYEEAAAYIQCQFEDLN	311
gbil_humaa	LLEEKIKR-S	PLTICYPEYT-	-GS-NTYEEAAAYIQCQFEDLN	311
gbil_rat	LFEEKIKR-S	PLTICYPEYT-	-GS-NTYEEAAAYIQCQFEDLN	311
gbil_xenla	LFEEKIKR-S	PLTICYPEYS-	-GS-NTYEEAAAYIQCQFEDLN	302
gbil2_mouse	LFEERIKTQ-S	SLTICFPEYT-	-GA-NKYDEEASYIQSIFEDLN	312
gbil2_rat	LFEERIKTQ-S	PLTICFPEYT-	-GA-NKYDEEASYIQSIFEDLN	312
gbil2_human	LFEERIKTH-S	PLTICFPEYT-	-GA-NKYDEEASYIQSIFEDLN	312
gbaz_human	LLAEKIRR-I	PLTICFPEYK-	-GQ-NTYEEAVAYIQRQFEDLN	312
gbaz_rat	LLSERKIRR-I	PLSVCFFPEYK-	-GQ-NTYEEAVAYIQRQFEDLN	312
gb01_mouse	LFGEKIKK-S	PLTICFPEYP-	-GS-NTYEDAAAYIQTQFESKN	312
gb01_rat	LFGEKIKK-S	PLTICFPEYP-	-GS-NTYEDAAAYIQTQFESKN	312
gb01_humaa	LFGEKIKK-S	PLTICFPEYT-	-GP-NTYEDAAAYIQAQFESKN	312
gb02_mouse	IFEEKIKK-S	PLTICFPEYT-	-GP-SATFEAVAMIQQYESKN	312
gb02_rat	IFEEKIKK-S	PLTICFPEYT-	-GP-SATFEAVAMIQQYESKN	312
gb02_cirilo	IFEEKIKT-S	PLTICFPEYT-	-GP-SATFEAVAMIQQYESKN	312
gb02_human	IFEEKIKK-S	PLTICFPEYT-	-GP-SATFEAVAYIQAQYESKN	312
gb02_drome	LFEEKIKK-S	PLTICFPEYT-	-GG-QEYGEAAYIQAQFEAKN	312
gbt1_bovin	VFSEKIKK-A	HLSICFPDYN-	-GP-NTYEDAGNYIKVQFLELN	307
gbt1_mouse	VFSEKIKK-A	HLSICFPDYD-	-GP-NTYEDAGNYIKVQFLELN	307
gbt1_humaa	VFFEKIKK-A	HLSICFPDYD-	-GP-NTYEDAGNYIKVQFLELN	307
gbt2_human	LFEEKIKK-V	HLSICFPFYD-	-GN-NSYEDAGNYIKSFLDLN	316
gbt2_bovin	LFEEKIKK-V	HLSICFPFYD-	-GN-NSYEDAGNYIKSFLDLN	311
gbt2_mouse	LFEEKIKK-V	HLSICFPFYD-	-GN-NSYEDAGNYIKSFLDLN	311
gbt3_rat	LFQEENVKI-V	HLSICFPFYT-	-GP-NTYEDAGNYIKNQFLDLN	311
gbil_bovin	LLEDKILH-S	HLVDYFPEFD-	-GPQRDAQAAREFILKMFVDLN	317
gbil_mouse	LLEDKILH-S	HLVDYFPEFD-	-GPQRDAQAAREFILKMFVDLN	317
gbil_humaa	LLEDKILY-S	HLVDYFPEFD-	-GPQRDAQAAREFILKMFVDLN	311
gbq_human_1	LLEEKIMY-S	HLVDYFPEYD-	-GPQRDAQAAREFILKMFVDLN	311
gbq_mouse_1	LLEEKIMY-S	HLVDYFPEYD-	-GPQRDAQAAREFILKMFVDLN	311
gbq_xenla_1	LLEEKIMY-S	HLVDYFPEYD-	-GPQRDAQAAREFILKMFVDLN	311
gbil4_mouse	LLEEKIMY-S	HLISYFPEYT-	-GPKQDVKAARDFILKLYQDQN	313
gbil4_bovin	LLEEKIMY-S	HLISYFPEYT-	-GPKQDVKAARDFILKLYQDQN	313
gbil5_human	ILEEKIPT-S	HLATYFPSFQ-	-GPQKQDAEAAKRFILDMDYTRMTGCVDPGE	328
gbil5_mouse	ILLEDKINT-S	HLATYFPSFQ-	-GPQKQDAEAAKRFILDMDYTRMTGCVDPGE	328
gbil2_humaa	LLVEKVKT-V	SINKHFDDFK-	-GDPHQLEDDVQRYLVQCFDRKK	339
gbil2_mouse	LLVEKVKS-V	SINKHFDDFK-	-GDPHQLEDDVQRYLVQCFDRKK	337
gbil3_mouse	LLEEKVQV-V	SINKYFLEFE-	-GDPHQLRDVQMFVLECFRGRK	334
gbas_humaa	LLAEKVLAGKS-	KIEDYFPEFA-YTTPEDATPPEGEDP-RVTAKYFIRDEFLRIS	349	
gbas_bovin	LLAEKVLAGKS-	KIEDYFPEFA-YTTPEDATPPEGEDP-RVTAKYFIRDEFLRIS	349	
gbas_mouse	LLAEKVLAGKS-	KIEDYFPEFA-YTTPEDATPPEGEDP-RVTAKYFIRDEFLRIS	349	
gbas_pig	LLAEKVLAGKS-KIELFWLDDRLFQERFFSF	KIEDYFPEFA-YTTPEDATPPEGEDP-RVTAKYFIRDEFLRIS	352	
gbas_xenla	LLAEKVLAGKS-	KIEDYFPEFA-YTTPEDATPPEGEDP-RVTAKYFIRDEFLRIS	334	
gbolf_human	MLAEKVLAGKS-	KIEDYFPEFA-YTTPEDATPPEGEDP-RVTAKYFIRDEFLRIS	336	
gbolf_rat	MLAEKVLAGKS-	KIEDYFPEFA-YTTPEDATPPEGEDP-RVTAKYFIRDEFLRIS	336	
ruler330.....340.....350.....360.....370.....380.....390.....400			

CLUSTAL X (1.83) MULTIPLE SEQUENCE ALIGNMENT

File: /Users/FMP/Desktop/Ga_ali2.ps
Page 6 of 6

Date: Tue Oct 17 16:03:52 2006

6

gb11_human	---KRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHLKDQGLF	354
gbq11_human	---PDSDKIIYSHFTCATDTENIKFVFVFAAVDTILQLNLKEYNAV	353
gbq11_mouse	---PDSDKIIYSHFTCATDTENIKFVFVFAAVDTILQLNLKEYNVL	353
gbq11_xenla	---PDSDKIIYSHFTCATDTENIKFVFVFAAVDTILQLNLKEYNVL	353
lgg2	---KRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHL	341
lgp2_template	---KRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHLKDQGLF	344
gbil1_rat	---KRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHLKDQGLF	354
gbil1_xenla	---KRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHLKDQGLF	354
gb113_human	---RRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHLKECLY	354
gb113_rat	---RRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHLKECLY	354
gb113_xenla	---RRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHLMECLY	345
gb112_mouse	---RRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKSNLMECLY	355
gb112_rat	---RRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHLKDQGLF	355
gb112_human	---RRKDTEIYTHFTCATDTEKVNQFVFDATDVIIKKNHLKDQGLF	355
gbaz11_human	---RNKETEIYSHFTCATDTNSIQFVFPAVDTIVIIQNLNLKYIGL	355
gbaz11_rat	---RNKETEIYSHFTCATDTNSIQFVFPAVDTIVIIQNLNLKYIGL	355
gb011_mouse	---RSP-NKEIYCHMTCATDTNNIQUVFPAVDTIDIIIANHLRGCGLY	354
gb011_rat	---RSP-NKEIYCHMTCATDTNNIQUVFPAVDTIVIIAKNLRGCGLY	354
gb011_human	---RSP-NKEIYCHMTCATDTNNIQUVFPAVDTIVIIAKNLRGCGLY	354
gb022_mouse	---KSA-HMEVSYHVTCATDTNNIQFVFPAVDTIVIIAKNLRGCGLY	354
gb022_rat	---KSA-HMEVSYHVTCATDTNNIQFVFPAVDTIVIIAKNLRGCGLY	354
gb022_cribo	---KSA-HMEIYTHFTCATDTNNIQFVFPAVDTIVIIAKNLRGCGLY	354
gb022_human	---KSA-HMEIYSHVTCATDTNNIQFVFPAVDTIVIIAKNLRGCGLY	354
gb022_dromo	---KST-SKEIYCHMTCATDTNNIQFVFPAVDTIVIIAKNLRGCGLY	354
gbt11_bovin	---MRRDVKIEIYSHMTCATDTQNVKVFVFDATDVIIKEKLKDQGLF	350
gbt11_mouse	---MRRDVKIEIYSHMTCATDTQNVKVFVFDATDVIIKEKLKDQGLF	350
gbt11_human	---MRRDVKIEIYSHMTCATDTQNVKVFVFDATDVIIKEKLKDQGLF	350
gbt212_human	---MRRDVKIEIYSHMTCATDTQNVKVFVFDATDVIIKEKLKDQGLF	359
gbt212_bovin	---MRRDVKIEIYSHMTCATDTQNVKVFVFDATDVIIKEKLKDQGLF	354
gbt212_mouse	---MRRDVKIEIYSHMTCATDTQNVKVFVFDATDVIIKEKLKDQGLF	354
gbt31_rat	---LKEDNEIYSHMTCATDTQNVKVFVFDATDVIIKEKLKDQGLF	354
gb111_bovin	---PDSDKIIYSHFTCATDTENIFVFVAAVDTILQLNLKEYNVL	359
gb111_mouse	---PDSDKIIYSHFTCATDTENIFVFVAAVDTILQLNLKEYNVL	359
gb111_human	---PDSDKIIYSHFTCATDTENIFVFVAAVDTILQLNLKEYNVL	353
gbq111_human_1	---PDSDKIIYSHFTCATDTENIRFVFVAAVDTILQLNLKEYNAV	353
gbq111_mouse_1	---PDSDKIIYSHFTCATDTENIRFVFVAAVDTILQLNLKEYNVL	353
gbq111_xenla_1	---PDSDKIIYSHFTCATDTENIRFVFVAAVDTILQLNLKEYNVL	353
gb114_mouse	---PDKEWVIIYSHFTCATDTENIRFVFVAAVDTILQLNLREFRLV	355
gb114_bovin	---PDKEWVIIYSHFTCATDTENIRFVFVAAVDTILQLNLREFRLV	355
gb115_human	GSKKGRASRRLPHSYTCATDTQNRKVFVFDATDVSVLARYLDIINLL	374
gb115_mouse	GGRKGRASRRLPHSYTCATDTQNRKVFVFDATDVSVLARYLDIINLL	374
gb112_human	--ENRS--=KPLFHHTTAIOTENIRFVFVAAVDTILQLNLKEYMLQ	381
gb112_mouse	--ENRS--=KPLFHHTTAIOTENIRFVFVAAVDTILQLNLKEYMLQ	379
gb113_mouse	--RDQQQ--=RFLYHHHTTAIOTENIRFVFVAAVDTILQLNLKEYMLQ	377
gbas11_human	--TASGDCGRHYCYPHTCAVDTENIRRNFNDCRDIIQRMHLRQYELL	394
gbas11_bovin	--TASGDCGRHYCYPHTCAVDTENIRRNFNDCRDIIQRMHLRQYELL	394
gbas11_mouse	--TASGDCGRHYCYPHTCAVDTENIRRNFNDCRDIIQRMHLRQYELL	394
gbas11_pig	--TASGDCGRHYCYPHTCAVDTENIRRNFNDCRDIIQRMHLRQYELL	397
gbas11_xenla	--TASGDCGRHYCYPHTCAVDTENIRRNFNDCRDIIQRMHLRQYELL	379
gbolf11_human	--TATGDCGRHYCYPHTCAVDTENIRRNFNDCRDIIQRMHLRQYELL	381
gbolf11_rat	--TATGDCGRHYCYPHTCAVDTENIRRNFNDCRDIIQRMHLRQYELL	381



Summary

Using homology of GPCR subtypes as well as their differences in biological function, we characterized ligand binding for endothelin receptors ETA and ETB.

For the first time, endothelin receptor-selective peptide ligands were subdivided into 4 regions (addressor, hook, core, modulator), explaining all existing data in literature as well as data from our co-operation partners. This is based on interactions with 4 complementary regions (gateway, edge, neck, binding cleft), found at the endothelin receptors. The pairwise interactions of addressor and gateway, hook and edge, core and neck, as well as modulator and binding cleft accurately explain the peptide ligands' selectivity for endothelin receptors and endothelin receptor subtypes. We impressively demonstrated this by the design and experimental validation of a new ETB-selective peptide ligand. Additionally, for the first time we described the necessities and differences in activation and in inhibition of both endothelin receptors. According to our models the molecular differences of peptide ligand-induced receptor activation and inhibition can be explained by the existence and absence of negatively charged residues at the peptide ligand's hook region, as counterpart of the receptor's edge region. In the case of receptor activation we suggest a scenario where the ligand's hook region catches the receptor's edge region. In succession, the ligand's modulator region is restrained in a position where its C-terminus orients in between transmembrane helices TMH3, TMH6 and TMH7 and induces reorientation of several side chains, finally leading to receptor activation. In the case of inhibition, the ligands lack the charged hook region as well as bulky moieties at the N-terminal portion. Such antagonistic peptides slide beyond the edge down into the transmembrane binding cleft and additionally constrain the existing interactions between the transmembrane helices of the inactive state, leading to inhibition. One of the most impressive results is the orientation of the last two residues of the antagonist's C-terminal modulator region at the same site of the inverse agonist 11-cis-retinal in inactive rhodopsin, demonstrating conserved mechanisms in many GPCRs.

Additionally, our models explain the so far unclear formation of the super-stable complex between ET-1 and ETB occurring in many mammals. A lid-like mechanism on this receptor's N-terminus, which is different to ETA in sequence and structure, covers ET-1 and restrains it tight into the binding site.

In a second project, we utilized the high degree of sequence homology between the human nicotinic acid receptor GPR109A and its homologue GPR109B, which has only low affinity for nicotinic acid. For binding of nicotinic acid to GPR109A we could identify crucial residues. The binding site is positioned between transmembrane helices TMH2, TMH3 and TMH7, and is therefore an interaction site different than most other rhodopsin-like GPCRs, where the binding site is located between TMH4, TMH5 and TMH6. Validated experimental data on predicted aromatic residues within TMH5 and TMH7, as well as at ECL2, clearly confirm this result. Combining mutagenesis data and comparative structural modeling allowed us to identify five residues located in close spatial proximity to the main interaction points for nicotinic acid. The characterization of the structural determinants and complementary pharmacophoric patterns for nicotinic acid binding in GPR109A is of general importance for understanding the binding mechanism of small molecule ligands to GPCRs. It is also important in the design and the development of new drugs, acting via GPR109A to treat dyslipidemic disorders. Comparisons of both binding sites explained the differences in ligand selectivity. Furthermore, they led to the proposal of 2-oxo-octanoic acid as a ligand selectively for GPR109B but not GPR109A, which was experimentally proven by our collaborators. Necessary interaction features, which were identified for 2-oxo-octanoic acid at GPR109B, could be unexceptionally applied to the recently published family of GPR109B-compounds based on benzotriazole-5-carboxylic

acid.

Taken together, nicotinic acid possesses the optimal hydrogen-bond patterns to bind to the rather hydrophilic patterns of the small binding site in GPR109A, whereas 2-oxo-octanoic acid possesses optimal hydrophobic patterns to bind to the additional specific hydrophobic cleft of GPR109B. Both receptor sites consist of hydrophilic and hydrophobic patterns, and are in spatial proximity. The binding of Acifran, which is an unselective ligand for both receptors, possesses the hydrophobic patterns (benzene, methyl) to bind to the hydrophobic binding cleft of GPR109B. Acifran also contains hydrophilic patterns (furan ring oxygen and carbonyl oxygen), allowing the interaction with the hydrophilic site in the GPR109A.

In a third approach, investigations on interactions of G proteins with their small-, medium- and large-sized ligands identified a common pattern for recognition and/or interaction, namely the interaction of negative and positive charges on G proteins and their ligands. This pattern was clearly identified by studies of small-sized alkyl-substituted amino acid derivatives (so called lipoamines) as well as by medium-sized secretagogues mastoparan-X and mastoparan-S. It contains specific distances of positive charges within G protein-ligands, recognized by negatively charged residues of complementary distances in G proteins. Investigations of four different G protein-subtypes ($\text{G}\alpha_i$, $\text{G}\alpha_o$, $\text{G}\alpha_q$, $\text{G}\alpha_s$) clearly showed different localizations and, therefore, distances of negatively charged residues in close proximity to the C-terminus (below 12 Å for $\text{G}\alpha_i$, around 15 Å for $\text{G}\alpha_q$, and greater than 18 Å for $\text{G}\alpha_s$). As a result, it was possible to design a new compound (FU244) that exclusively interacts with $\text{G}\alpha_i$ but not with $\text{G}\alpha_q$ or $\text{G}\alpha_s$. The application of these patterns to the interaction interface of GPCRs and G proteins resulted in identification of two residues within ICL2 that are required for $\text{G}\alpha_i$ selectivity of ETA and ETB. This way, we demonstrated that also in GPCRs complementary charge patterns determine G protein-selectivity.

Zusammenfassung

In der hier vorliegenden Studie wurden die Interaktionsmechanismen und Hintergründe für selektive Wechselwirkungen zwischen G-Protein gekoppelten Rezeptoren, deren Liganden und G-Proteinen untersucht. Ausgangspunkt dafür waren homologe Rezeptorsubtypen, die sich durch Unterschiede in ihren Funktionen auszeichnen.

Die Ursachen für Selektivität und hochaffine Ligandenbindung der Endothelin-Rezeptorsubtypen ETA und ETB war vor der Arbeit nicht aufgeklärt. Basierend auf unseren Untersuchungen, präsentieren wir hier erstmals Liganden-Rezeptor-Wechselwirkungen, die auf vier unterschiedlichen aber zueinander komplementären Erkennungsregionen in Liganden und Rezeptoren beruhen. Vergleiche von Sequenz- und Funktionsinformationen von endothelin-rezeptorselektiven Peptiden führten schlussendlich zur Unterteilung der Peptidliganden in vier Regionen (*Addressor*, *Hook*, *Core*, *Modulator*), welche die existierenden Daten aus der Literatur und von unseren Kooperationspartnern erklären. Gleichsam zu diesen vier Regionen an Peptidliganden wurden komplementäre Erkennungsregionen an beiden Endothelin-Rezeptorsubtypen identifiziert: *Gateway*, *Edge*, *Neck*, *Binding Cleft*. Paarweise Wechselwirkungen von *Addressor* und *Gateway*, *Hook* und *Edge*, *Core* und *Neck* sowie *Modulator* und *Binding Cleft*, erklären dabei vollständig die Ursachen und Hintergründe der Selektivität für Endothelin-Rezeptoren sowie deren Subtypen, wie wir eindrucksvoll durch die Entwicklung eines neuen ETB-selektiven Peptidliganden zeigen konnten.

Weiterhin war es uns dadurch möglich, die Notwendigkeiten und Unterschiede für Aktivierung und Inhibierung beider Endothelin-Rezeptorsubtypen aufzuzeigen. Dabei liegt der Hauptunterschied zwischen peptidligandenabhängiger Aktivierung und Inhibierung am Vorhandensein negativer Ladungen innerhalb der *Hook*-Region der Peptidliganden. Sind solcherart Ladungen innerhalb des *Hook* vorhanden, können sie mit entsprechenden Ladungen der *Edge*-Region des Rezeptors wechselwirken, wodurch der aktivierende Ligand nicht weiter in die *Binding Cleft* zwischen den Transmembran-Helices TMH3, TMH6 und TMH7 hinabtaucht. Dadurch wird eine Neuorientierung wichtiger Rezeptorseitenketten innerhalb des Transmembran-Helixbündels hervorgerufen, die den Rezeptor in den aktiven Zustand überführt.

Fehlen negative Ladungen innerhalb der *Hook*-Region, wie dies bei Inhibitoren der Fall ist, gleitet der Ligand tief in die *Binding Cleft* und sorgt dort durch zusätzliche Interaktionen mit Rezeptorseitenketten zu einer Stabilisierung des inaktiven Zustands des Rezeptors – der daraufhin inhibiert wurde.

Eins der am beeindruckendsten Ergebnisse dieser Untersuchungen ist die Orientierung der beiden letzten Aminosäuren des C-Terminus (innerhalb des *Modulator*) von Inhibitoren, die in identischer Position zu der bekannten Struktur von 11-cis-Retinal in Rinderrhodopsin zu liegen kommen. Dadurch lässt sich eine Konserviertheit dieses Mechanismus in vielen G-Protein gekoppelten Rezeptoren vermuten.

Ferner war es uns durch unsere Modelle möglich, die bisher unverstandene Formierung des superstabilen Komplexes zwischen Endothelin-1 und Endothelin-Rezeptorsubtyp ETB zu erklären. Während der Ligandenbindung an ETB werden Sequenz- und Strukturbereiche des N-Terminus dieses Rezeptors genutzt, um ET-1 wie mit einem Deckel zu bedecken. Da diese N-terminalen Bereiche in ETA anders formiert und gestaltet sind, ist dort dieser Mechanismus und damit auch die Bildung eines superstabilen Komplexes nicht möglich.

Des Weiteren wurde die Ligandenbindung an Nikotinsäure-Rezeptoren GPR109A und GPR109B untersucht. Obwohl beide Rezeptoren einen hohen Grad an Sequenz- und Strukturhomologie aufweisen, bindet GPR109B Nikotinsäure mit sehr geringer Affinität, während GPR109A für die Bindung dieses Liganden optimiert ist.

Durch die Nutzung der hohen Homologie und der funktionellen Unterschiede war es uns möglich, Aminosäuren mit kritischer Bedeutung für die Ligandenbindung an beiden Rezeptoren zu identifizieren. Die Bindungsstelle befindet sich zwischen den Transmembran-Helices TMH2, TMH3 und TMH7 und ist somit unterschiedlich zu den meisten anderen rhodopsin-ähnlichen Rezeptoren, deren Bindungstaschen zwischen TMH4, TMH5 und TMH6 liegt. Dass diese Bindungsstelle dennoch richtig identifiziert wurde, bestätigen experimentelle Ergebnisse von mutierten Aromatenresten in TMH5 und TMH7, die aufgrund unserer Vorschläge gemacht wurden.

Durch die korrekte Identifizierung der Bindungstasche von Nikotinsäure konnten wir ebenfalls komplementäre pharmakophore Muster zwischen Rezeptor und Ligand ableiten, die nun für die Identifizierung und Entwicklung weiterer Medikamente genutzt werden können. Dass dies so möglich ist, wurde von uns und den experimentellen Daten unserer Kooperationspartner mit der Identifizierung von 2-Oxo-oktansäure als selektiven Liganden für GPR109B und nicht für GPR109A bestätigt.

Ferner sind die für 2-Oxo-oktansäure identifizierten pharmakophoren Muster ebenso auffindbar in den ebenso GPR109B-selektiven Verbindungen basierend auf Benzotriazol-5-carbonsäure, welche vor kurzem veröffentlicht wurden.

In einem dritten Projekt wurden die Wechselwirkungen von G-Protein sowie deren klein-, mittel- und großmolekularen Interaktionspartnern untersucht. Dabei identifizierten wir ein generelles Muster für G-Protein-Selektivität, welches auf komplementären Ladungsabständen beruht. Diese Muster wurden klar für die Wechselwirkungen der alkyl-substituierten Aminosäurederivate (auch Lipoamine genannt) und der sekretierten Peptide (z.B. Mastoparan-X und Mastoparan-S) herausgearbeitet.

Im Allgemeinen interagieren positive Ladungen von G-Protein-Liganden an negativen Ladungen vergleichbarer Abstände an benachbarten Strukturen des C-terminus von G-Proteinen. Unsere Untersuchungen von vier verschiedenen G-Protein-Subtypen ($\text{G}\alpha_i$, $\text{G}\alpha_o$, $\text{G}\alpha_q$, $\text{G}\alpha_s$) identifizierten verschiedene Verteilungen und Ladungsabstände negativer Aminosäuren in diesen Regionen (unter 12 Å für $\text{G}\alpha_i$, etwa 15 Å für $\text{G}\alpha_q$, und über 18 Å für $\text{G}\alpha_s$). Basierende auf diesen Erkenntnissen entwickelten wir einen neuen Lipoamin mit optimierten Ladungsabständen für $\text{G}\alpha_i$. Tatsächlich wurde für die entsprechende Struktur FU244 experimentell gezeigt, dass dieser Ligand mit $\text{G}\alpha_i$ interagiert, nicht aber mit $\text{G}\alpha_q$ oder $\text{G}\alpha_s$.

Anwendung dieser Ladungsmuster auf die Interaktionen zwischen G-Proteinen und G-Protein gekoppelten Rezeptoren resultierte schlussendlich in der Identifizierung von zwei Aminosäuren innerhalb des intrazellulären Loops ICL2 an ETA und ETB, die absolut notwendig sind für die Wechselwirkung mit $\text{G}\alpha_i$.

Daraus schliessen wir, dass distanzabhängige Ladungsmuster sowohl für niedermolekulare als auch für hochmolekulare Interaktionspartner wie G-Protein gekoppelte Rezeptoren zur Bestimmung für G-Protein-Selektivität notwendig sind.

Zusammenfassend lässt sich festhalten, dass die Kombination aus Computer-Untersuchungen homologer Rezeptor Subtypen unterschiedlicher, biologischer Funktion und experimenteller Absicherung zu stimmigen Hypothesen und Theorien über die Mechanismen von G-Protein gekoppelten Rezeptoren führt. Dadurch stellt die hier vorgestellte Methode ein wertvolles Werkzeug in der Charakterisierung von GPCRs, deren Ligandenbindung und G-Proteinkopplung dar.

Publications

Publications in Journals

Kleinau G, Jäschke H, Neumann S, **Lättig J**, Paschke R and Krause G (2004), Identification of a Novel Epitope in the TSH Receptor Ectodomain Acting as Intramolecular Signalling Interface. *J Biol Chem.* 279 (49): 51590-51600

Tunaru S, **Lättig J**, Kero J, Krause G and Offermanns S (2005), Characterization of Determinants of Ligand Binding in the Nicotinic Acid Receptor GPR109A (HM74A/PUMA-G). *Mol. Pharmacol.* 68 (5): 1271-1280

Lättig J, Böhl M, Fischer P, Tischer S, Tietböhl C, Menschikowski M, Gutzeit HO, Metz P and Pisabarro MT (2007), Mechanism of Inhibition of Human Secretory Phospholipase A2 by Flavonoids – Rationale for Lead Design. *J Comput Aided Mol Des.* 21 (8): 473-483

Lättig J, Ferkert J, Oksche A and Krause G, Structural Determinants for Ligand Selectivity at EtA- and EtB-Receptors. *Biophys J.* (submitted)

Lättig J, Böhl M, Böhl J, Pisabarro MT and Gutzeit HO, An experimental and theoretical approach to study inhibitor binding to phospholipase A2. *J Comput Aided Mol Des.* (in preparation)

Lättig J, Tunaru S, Kero J, Krause G and Offermanns S, Selectivity Patterns to Differentiate Between Nicotinic Acid Receptor GPR109A and Related Receptor GPR109B: First Ligand for GPR109B. (in preparation)

Lättig J, Ferkert J, Oksche A and Krause G, Residues of the 2nd Intracellular Loop of Endothelin Receptor Subtypes A and B Selectively for Interactions with G-protein Subunit Gαi. (in preparation)

Contributions to Books

Lättig J, Schunack W, Nürnberg B and Krause G (2003), Towards 3D-patterns for G-protein Specificity of GPCRs by G-protein Interacting Small Molecules. Designing Drugs and Crop Protectants: Processes, Problems and Solutions. Editors: M Ford, D Livingston, J Dearden, H van de Waterbeemd Blackwell publishing Inc., Oxford, UK

Lendeckel U, Bukowska A, **Lättig J** and Brandt W, Alanyl-Aminopeptidase in Human T Cells: Structure and Function. Aminopeptidases in Biology and Disease Editors: Hopper NM, Lendeckel U Kluwer Academic/Plenum Publishers, New York, USA

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Referenzen und Hilfsmittel verfasst habe.

Ich habe diese Arbeit bisher noch an keiner Universität vorgelegt.

Ich erkläre weiter, dass ich mich bisher nicht an einer anderen Einrichtung um einen Doktorgrad beworben habe und ebenso keine derartigen Titel besitze.

Berlin, 20. Juli 2007

Jens Lättig