

Alignment Annotation

G-Protein Coupled Receptors (GPCRs) all share a common structural core of seven transmembrane helices but they lack significant sequence homology between subfamilies. When modeling GPCRs it is important to get a good alignment between the query and template structure. Each helix has one or more conserved motifs:

Helix 1: GX_3N or GN

Helix 2: $N(S,H)LX_3DX_{7,8,9}P$

Helix 3: $SX_3LX_2IX_2D(E,H)RY$

Helix 4: $WX_{8,9}P$

Helix 5: FX_2PX_7Y

Helix 6: $FX_2CW(Y,F)XP$

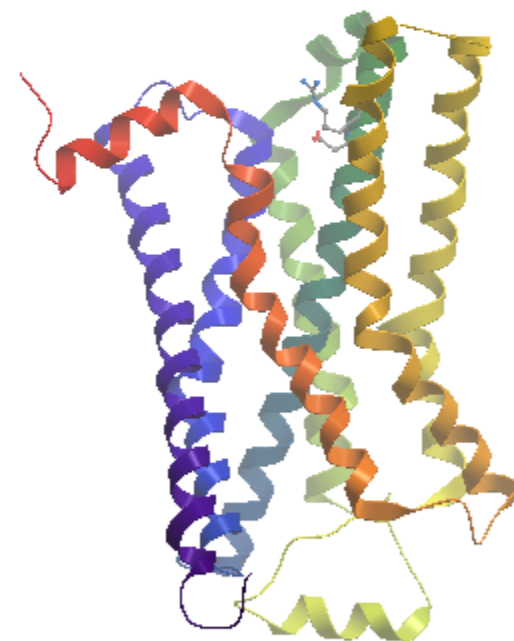
Helix 7/Helix 8: $LX_3NX_3N(D)PX_2YX_{5,6}F$

The ProSite class A alignment <http://prosite.expasy.org/PDOC00210> can be used to guide GPCR alignments.

1. Windows/Alignment Main.

Read in the icb file containing the class A GPCR alignment. To make the alignment panel the main window – choose Window --> Alignment Main.

```
id=18 nSeq=25
5HT1A_HUMAN_53_400 1 -----GNACVVAAIALER
5HT1B_HUMAN_66_369 1 -----SNAFVIATVYRTR
5HT1D_HUMAN_55_356 1 -----SNAFVLTTILLTR
ADA1A_HUMAN_43_326 1 -----GNILVILSVACHR
ADA1B_HUMAN_62_348 1 -----GNILVILSVACNR
DRD2_HUMAN_51_426 1 -----GNVLVCMASREK
DRD3_HUMAN_46_383 1 -----GNGLVCMAVLKER
ADRB1_HUMAN_75_377 1 -----GNVLVIVAIAKTP
ADRB2_HUMAN_50_326 1 -----GNVLVITAIAKFE
ADRB3_HUMAN_54_346 1 -----GNLLVIVAIAWTP
DRD1_HUMAN_40_331 1 -----GNTLVCAAVIRFR
AA1R_HUMAN_26_288 1 -----GNVLVIWAVKVNQ
AA2AR_HUMAN_23_288 1 -----GNVLVCWAVWLNS
GRPR_HUMAN_57_322 1 -----GNITLIKIFCTVK
CCR1_HUMAN_51_301 1 -----GNILVVLVLVQYK
CCR10_HUMAN_58_310 1 -----GNGLVLATHLAAR
CXCR1_HUMAN_56_305 1 -----GNSLVMLVILYSR
CXCR2_HUMAN_65_314 1 -----GNSLVMLVILYSR
CXCR3_HUMAN_70_318 1 -----GNGAVAAVLLSRR
OPRD_HUMAN_66_318 1 -----GNVLVMFGIVRYT
PAR1_HUMAN_119_371 1 -----LNIMAVVFILKM
PAR2_HUMAN_92_344 1 -----SNGMALWVFLFRT
PAR3_HUMAN_111_358 1 -----ANAVTLWMLFFRT
PAR4_HUMAN_94_340 1 -----ANGLALWVLATQA
3p0g_a 23 DVTQQRDEVVVVGMGIVMSLIVLAIVFGNVLVITAIKFE
.. ..###..LA..D###...# .....W.
5HT1A HUMAN 53 400 14 SL-QNVANYLIGSLAVTDLMVSVLVL-PMAALYQVLNKT
```



Display the secondary structure of PDB 3p0G in the alignment.

The screenshot shows the gpcr_a.icb software interface. On the left is a sequence alignment table with columns for sequence ID, name, start, end, and alignment status. The alignment includes sequences like 5HT1A_HUMAN, ADA1A_HUMAN, DRD2_HUMAN, and 3p0g_a. The 3p0g_a sequence is highlighted in blue. On the right is a 3D Graphics window showing a ribbon representation of the protein structure, colored by consensus strength. The interface includes a top menu bar (File, Tools, Edit, View, Bioinfo, Homology, Ch...), a toolbar with various icons, and a central control panel with sections for 'Comment', 'Selection', and 'Preferences'. A status bar at the bottom right indicates '1 Mol 1 Obj'.

id	nSeq	seq1	seq2	seq3	seq4	seq5	seq6	seq7	seq8	seq9	seq10	seq11	seq12	seq13	seq14	seq15	seq16	seq17	seq18	seq19	seq20	seq21	seq22	seq23	seq24	seq25
5HT1A_HUMAN	53	400	1																							
5HT1B_HUMAN	66	369	1																							
5HT1D_HUMAN	55	356	1																							
ADA1A_HUMAN	43	326	1																							
ADA1B_HUMAN	62	348	1																							
DRD2_HUMAN	51	426	1																							
DRD3_HUMAN	46	383	1																							
ADRB1_HUMAN	75	377	1																							
ADRB2_HUMAN	50	326	1																							
ADRB3_HUMAN	54	346	1																							
DRD1_HUMAN	40	331	1																							
AA1R_HUMAN	26	288	1																							
AA2AR_HUMAN	23	288	1																							
GRPR_HUMAN	57	322	1																							
CCR1_HUMAN	51	301	1																							
CCR10_HUMAN	58	310	1																							
CXCR1_HUMAN	56	305	1																							
CXCR2_HUMAN	65	314	1																							
CXCR3_HUMAN	70	314	1																							
OPRD_HUMAN	66	314	1																							
PAR1_HUMAN	11	94	1																							
PAR2_HUMAN	92	94	1																							
PAR3_HUMAN	11	94	1																							
PAR4_HUMAN	94	94	1																							
3p0g_a																										
3p0g_a																										

1. Click here to open tools panel.

2. Click here to display secondary structure.

Secondary structure is displayed colored cylinders and arrows under the alignment. Red cylinder = alpha helix, green arrow = beta sheet, blue cylinder = pi helix, magenta cylinder = 3/10 helix.

The image shows a software interface for protein analysis. On the left, a sequence alignment is displayed with columns for residue numbers and sequence. The alignment includes sequences for 5HT1A_HUMAN, 5HT1B_HUMAN, ADA1A_HUMAN, ADA1B_HUMAN, DRD2_HUMAN, DRD3_HUMAN, ADRB1_HUMAN, ADRB2_HUMAN, ADRB3_HUMAN, DRD1_HUMAN, AA1R_HUMAN, AA2AR_HUMAN, GRPR_HUMAN, CCR1_HUMAN, CCR10_HUMAN, CXCR1_HUMAN, CXCR2_HUMAN, CXCR3_HUMAN, OPRD_HUMAN, PAR1_HUMAN, PAR2_HUMAN, and PAR3_HUMAN. The alignment shows conserved regions highlighted in green and red. A red box highlights the text "Secondary structure" with an arrow pointing to the alignment. To the right, a 3D ribbon diagram of the protein structure is shown, with secondary structure elements colored according to the legend: red for alpha helix, green for beta sheet, blue for pi helix, and magenta for 3/10 helix. The 3D Graphics window shows the protein structure in a ribbon representation. The interface includes a toolbar at the top and bottom, and a status bar at the bottom right indicating "1 Mol 1 Obj".

Residue	5HT1A_HUMAN	5HT1B_HUMAN	ADA1A_HUMAN	ADA1B_HUMAN	DRD2_HUMAN	DRD3_HUMAN	ADRB1_HUMAN	ADRB2_HUMAN	ADRB3_HUMAN	DRD1_HUMAN	AA1R_HUMAN	AA2AR_HUMAN	GRPR_HUMAN	CCR1_HUMAN	CCR10_HUMAN	CXCR1_HUMAN	CXCR2_HUMAN	CXCR3_HUMAN	OPRD_HUMAN	PAR1_HUMAN	PAR2_HUMAN	PAR3_HUMAN	
14	SL-QNVANYLIGSLAVTDL	KL-HTPANYLIASLAVTDL	HL-HSVTHYYIVNLAVAD	HL-RTPTNYFIVNLAMAD	AL-QTTTNYLIVSLAVAD	AL-QTTTNYLVVSLAVAD	RL-QTLTNLFIMSLASAD	RL-QTVTNYFITSLACAD	RL-QTMTNVFVTSLAAAD	HLSKVNTNFFVISLAVSD	AL-RDATFCFIVSLAVAD	NL-QNVTNYFVVSLLAA	SM-RNVPNLFISSALGDL	RL-KNMTSIYLLNLAISD	RAARSPTSALLQLALAD	VG-RSVTDVYLLNLALAD	VG-RSVTDVYLLNLALAD	TA-LSSTDTEFLHLAVAD	EM-KTATNXYIENLALAD
14	LMVSVLVLPMAALYQVLN	LLVSVILVMPISTMYTVT	LLLTSTVLPFSAIFEVLG	LLLSFTVLPFSAALEVLG	LLVATLVPWVVYLEVVG	LLVATLVMPWVVYLEVT	LVMGLLVV-PFGATIVV	LVMGLAVV-PFGAAHIL	LVMGLLVV-PPAATLALT	LLVAVLVPWKAVAEIAG	VAVGALVPLAILINIGP	AVGLAIPFAITISTGFCA	LLITCA-PVDASRYLADR	LLFLFTL--PFWIDYK	LLALTLP---FAAAGALQ	LLFALTLP---IWAASKV	LLFALTLP---IWAASKV	TLVLTLP---LWAVDAAV	ATSTI--PFQSAKYLME
14	KT	GRWT	YWA	YVW	EWK	WN	WE	WT	WP	FWP	QTY	CAA	WL	WV	WS	WI	WI	QVW	TWP
63	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD	RL-QTVTNYFITSLACAD
	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL	LVMGLAVV-PFGAAHIL
	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI	MKMWI

Secondary structure

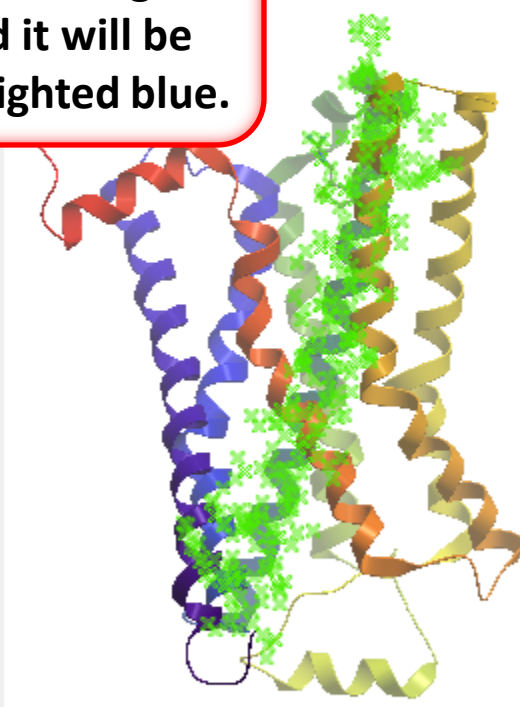
In this example we will box and annotate all 7 of the transmembrane helices.

Multi-line
annotation

1. Click and drag
over the region
and it will be
highlighted blue.

2. Right click on
selection and
choose Box and
Annotation

5HT1A_HUMAN_5			IGSLAVTDLMSVLVL-PMAALYQVINF
5HT1B_HUMAN_66_369	14	KL	HTPANYLIASLAVTDLLVSILVM-PISTMYFVTIGP
5HT1D_HUMAN_55_356	14	KL	HTPANYLIGSLATTDLLVSILVM-PISIAYTITHT
ADA1A_HUMAN_43_326	14	HL	HSVTHYYIVNLAVADLLLTSTVL-PFSAIFEVLGY
ADA1B_HUMAN_62_348	14	HL	RTPTNYFIVNLAMADLLLSFTVL-PFSAALEVLGY
DRD2_HUMAN_51_426	14	AL	QTTTNYLIVSLAVADLLVATLVM-PWVVYLEVVGW
DRD3_HUMAN_46_383	14	AL	QTTTNYLVVSLAVADLLVATLVMPPWVVYLEVTGGVWN
ADRB1_HUMAN_75_377	14	RL	QTLTNLFIMSLASADLVMGLLV-PFGATIVVWGRWE
ADRB2_HUMAN_50_326	14	RL	QVTVNYFITSLACADLVMGLAVV-PFGAAHILMKMWT
ADRB3_HUMAN_54_346	14	RL	QTMVNFVTSLAAADLVMGLLV-PPAATLALTGHWP
DRD1_HUMAN_40_331	14	HL	RSKVTNFFVISLAVSDLLVAVLVM-PWKAVAEIAGFWP
AA1R_HUMAN_26_288	14	AL	RDATFCFIVSLAVADVAVGALVI-PLAILINIGPQTY
AA2AR_HUMAN_23_288	14	NL	QNVVNYFVVSLLAAADIAVGVLAIPFAITISTGFCAA
GRPR_HUMAN_57_322	14	SM	RNVPNLFISSLALGDDLLLLITCA-PVDASRYLADRWL
CCR1_HUMAN_111_358	14	R	KNMYSIYLLNLAISDLLFLFTL--PFWIDYKDKDDWV
CCR10_HUMAN_94_340	14	P	RSPTSALLQLALADLLALTLF---FAAAGALQGW
CXCR1_HUMAN_111_358	14	R	RSVTDVYLLNLALADLLFALTLF---IWAASKVNGWI
CXCR2_HUMAN_111_358	14	R	RSVTDVYLLNLALADLLFALTLF---IWAASKVNGWI
CXCR3_HUMAN_111_358	14	R	LSSTDTFLHLAVADTLLVLTLP---LWAVDAAVQWV
OPRD_HUMAN_111_358	14	R	KTATNIYIFNLALADALATSTL--PFQSAKYLMTWP
PAR1_HUMAN_111_358	14	R	KKPAVVYMLHLATADVLFVSVLP-FKISYYFSGSDWQ
PAR2_HUMAN_111_358	14	R	KHPAVIYMANLALADLLSVIWF-PRIAYHLRGNWV
PAR3_HUMAN_111_358	14	R	SICTTVFYTNLAIAADFLFCVTLF-FKIAHYHLNGN
PAR4_HUMAN_94_340	14	P	RLPSTMLLMNLAAADLLALALP-PRIAYHLRGNWV
3p0g_a	63	RL	QVTVNYFITSLACADLVMGLAVV-PFGAAHILMKMWT
3p0g_a			



Display the sequence profile.

1. Open the tools panel and choose profile.



Sequence	Residue	Position	Conservation	Profile
5HT1A_HUMAN	53			L
5HT1B_HUMAN	66-369			YNYLIGSLAVTDLMVSVLVL-PMAALYQVLNK
5HT1D_HUMAN	55-356	15	L	HTPANYLIGSLAVTDLMVSILVM-PISIAYTITHTWNEF
ADA1A_HUMAN	43-326	15	L	HSVTHYYIVNLAVADLLLTSTVL-PFSAIFEVLGYWAFG
ADA1B_HUMAN	62-348	15	L	RTPTNYFIVNLAMADLLLSFTVL-PFSAALEVLGYWVLG
DRD2_HUMAN	51-426	15	L	QTTTNYLIVSLAVADLLVATLVM-PWVVYLEVVGEWKES
DRD3_HUMAN	46-383	15	L	QTTTNYLVVSLAVADLLVATLVMPWVVYLEVTGGVWNES
ADRB1_HUMAN	75-377	15	L	QTLTNLFIMSLASADLVMGLLVV-PFGATIVVWGRWEYG
ADRB2_HUMAN	50-326	15	L	QTVTNYFITSLACADLVMGLAVV-PFGAAHILMKMWTFG
ADRB3_HUMAN	54-346	15	L	QTMTNVFVTSLAAADLVMGLLVV-PPAATLALTGHWPLG
DRD1_HUMAN	40-331	15	L	RSKVTNFFVISLAVSDLLVAVLVM-PWKAVAEIAGFWPF-
AA1R_HUMAN	26-288	15	L	RDATEFCFIVSLAVADVAVGALVI-PLAILINIGPQTYE-
AA2AR_HUMAN	23-288	15	L	QNVVTNYFVVS LAAADIAVGVLAI-PFAITISTGFCAAC-
GRPR_HUMAN	57-322	15	M	RNVPNLFISSLALG D L L L L I T C A - P V D A S R Y L A D R W L E F G
CCR1_HUMAN	51-301	15	L	KNMTSIYLLNLAISDLLFLFTL--PFWIDYKCLKDDWVFG
CCR10_HUMAN	58-310	15	A	ARSPTSAHLLQLALADLLLALTLP---FAAAGALQGWSLG
CXCR1_HUMAN	56-305	15	G	RSVTDVYLLNLALADLLFALTLP---IWAASKVNGWIFG
CXCR2_HUMAN	65-314	15	G	RSVTDVYLLNLALADLLFALTLP---IWAASKVNGWIFG

To highlight the conserved motifs shade the residues that are 100% conserved.

5HT1A_HUMAN_53_400 12 ERSL-QNVANYLIGSLAVT
5HT1B_HUMAN_66_369 12 TRKL-HTPANYLIASLAVT
5HT1D_HUMAN_55_356 12 TRKL-HTPANYLIGSLATT
ADA1A_HUMAN_43_326 12 HRHL-HSVTHYYIVNLAVA
ADA1B_HUMAN_62_348 12 NRHL-RTPTNYFIVNLAMA
DRD2_HUMAN_51_426 12 IVSLAVA
DRD3_HUMAN_46_383 12 VVSLAVA
ADRB1_HUMAN_75_377 12 IMSLASA
ADRB2_HUMAN_50_326 12 ITSLACA
ADRB3_HUMAN_54_346 12 VTSLAAA
DRD1_HUMAN_40_331 12 VISLAVS
AA1R_HUMAN_26_288 12 NQAL-RDATFCFIVSLAVA
AA2AR_HUMAN_23_288 12 NSNL-QNVTNYFVVS LAAA
GRPR_HUMAN_57_322 12 VKSM-RNVPNLFISSLALG
CCR1_HUMAN_51_301 12 YKRL-KNMTSIYLLNLAIS
CCR10_HUMAN_58_310 12 ARRAARSPTS AHLLQLALA
CXCR1_HUMAN_56_305 12 SRVG-RSVTDVYLLNLALA
CXCR2_HUMAN_65_314 12 SRVG-RSVTDVYLLNLALA
CXCR3_HUMAN_70_318 12 RRTA-LSSTDTFLLHLAVA
OPRD_HUMAN_66_318 12 YTKM-KTATNIYIFNLALA

1. Drag strength to 100%.

3. Right click on the selection and choose shade color.

2. Choose Select.

View options
 title consensus profile
 sequence offset order hide body
Strength (100%)
Color no color
Selection
 Propagate to ALL sequences
 Highlight Selection
By Consensus X
Select Invert Hide
Preferences
Hidden block format clean
Hidden block width 1
Show secondary structure for
 3p0g_a

Calculate sequence similarity and display residues at each position.

ADA1A_HUMAN_43_326 54 RVF⁺NIWAAVDVLCCTASIMGLCIISID⁺RYIGVSYPLRYPT
ADA1B_HUMAN_62_348 54 RIF⁺CDIWAAVDVLCCTASILSLCAISID⁺RYIGVRYSLQYPT
DRD2_HUMAN_51_426 54 RIH⁺CDIFVTL⁺LDVMMCTASILNLCAISID⁺RYTAVAMPMLYNT
DRD3_HUMAN_46_383 55 RIG⁺CDVFEVTL⁺DVMMCTASTILNLCAISID⁺RYTAVVMPVHYQH
ADRB1_HUMAN_75_377 YIALDRYLAITS⁺SPFRYQS
ADRB2_HUMAN_50_326 YIAVD⁺RYFAITS⁺SPFKYQS
ADRB3_HUMAN_54_346 ALAVDRYLA⁺VTNPLRYGA
DRD1_HUMAN_40_331 54 GSE⁺NIWVAEDIMCSTASTILNLCVLSVD⁺RYWALISSPERYER

1. Right click and choose "Calculate sequence similarity".

	p29	p30	p31	p32	p33	p34	p35	p36	p37	p38	p39	
1	N2	A3	C4	V5	V6	A7	A8	I9	A10	L11	E12	
2	N2	A3	F4	V5	I6	A7	T8	V9	Y10	R11	T12	
3	N2	A3	F4	V5	L6	T7	T8	I9	L10	L11	T12	
4	N2	I3	L4	V5	I6	L7	S8	V9	A10	C11	H12	
5	N2	I3	L4	V5	I6	L7	S8	V9	A10	C11	N12	
6	N2	V3	L4	V5	I6	L7	S8	V9	A10	S10	R11	E12
7	N2	G3	L4	V5	I6	L7	S8	V9	A10	L10	K11	E12
8	N2	V3	L4	V5	I6	L7	S8	V9	A10	A10	K11	T12
9	N2	V3	L4	V5	I6	L7	S8	V9	A10	A10	K11	F12
10	N2	L3	L4	V5	I6	L7	S8	V9	A10	A10	W11	T12
11	N2	T3	L4	V5	C6	A7	A8	V9	I10	R11	F12	
12	N2	V3	L4	V5	I6	W7	A8	V9	K10	V11	N12	
13	N2	V3	L4	V5	C6	W7	A8	V9	W10	L11	N12	
14	N2	T2	T4	L5	I6	K7	T8	F9	C10	T11	V12	

2. Percentage similarity and residues at each position.