

iBio-GATS - a semi-automated workflow for structural modelling of insect odorant receptors

by Vaanathi Chidambara Thanu, Amara Jabeen, and Shoba Ranganathan

SUPPLEMENTARY FIGURES AND TABLES

Figure S1. Hydrophobicity plots for the *AbOrco* sequence with the template *MhOR5*

Figure S2. Helical wheel plots for the *AbOrco* sequence with the template *MhOR5*

Figure S3. iBio-GATS helix-wise alignment for the target sequence, *AbOrco* with the template, *MhOR5*

Figure S4. Full alignment for the target sequence, *AbOrco* with the template, *MhOR5*

Figure S5. Predicted models for *AbOrco* and *MhOR5*

Table S1. Helix definitions for *DmOR59b* (UniProt ID: Q9W1P8)

Figure S6. Helical wheel plots for *DmOR59b* compared with templates, *MhOR5* and *AbOrco*

Figure S7. Final pairwise alignment of *DmOR59b* with *MhOR5*

Figure S8. Predicted models for wild *DmOR59b*

Figure S9. Ramachandran plots

Figure S10. Binding pockets for predicted models of *DmOR59b* with DEET

Figure S11. Final pairwise alignment of *DmelOrco* with *AbOrco* (PDB ID:6C70)

Table S2. Helix definitions for *DmelOrco* (UniProt ID: Q9VNB5)

Figure S12. Hydrophobicity plots for the *DmelOrco* sequence with templates *AbOrco* and *MhOR5*

Figure S13. Helical wheel plots for the *DmelOrco* sequence compared templates, *AbOrco* and *MhOR5*

Figure S14. Structural overlay of predicted *DmelOrco* model with its template *AbOrco*.

Figure S15. RMSD plots of all four *DmelOrco* complexes over 100ns of MD simulations

Figure S16. The structure-derived sequences of the template structures

Table S3. Helix positions for *AbOrco* (PDB ID:6C70)

Table S4. Helix positions for *MhOR5* (PDB ID:7LIG).

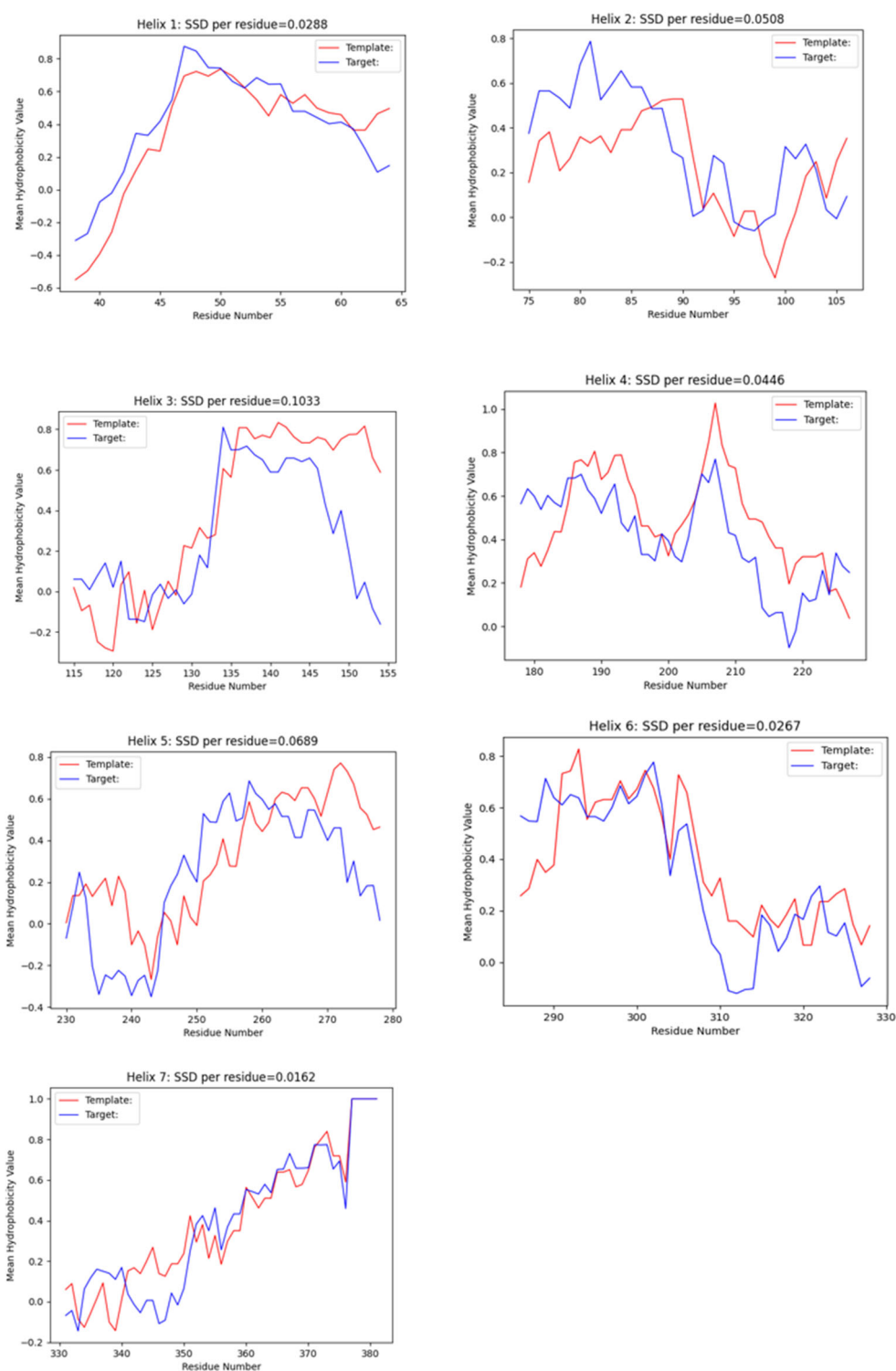


Figure S1. Hydrophobicity plots for the *AbOrco* sequence with the template *MhOR5*. Template PDB ID:7LIG .

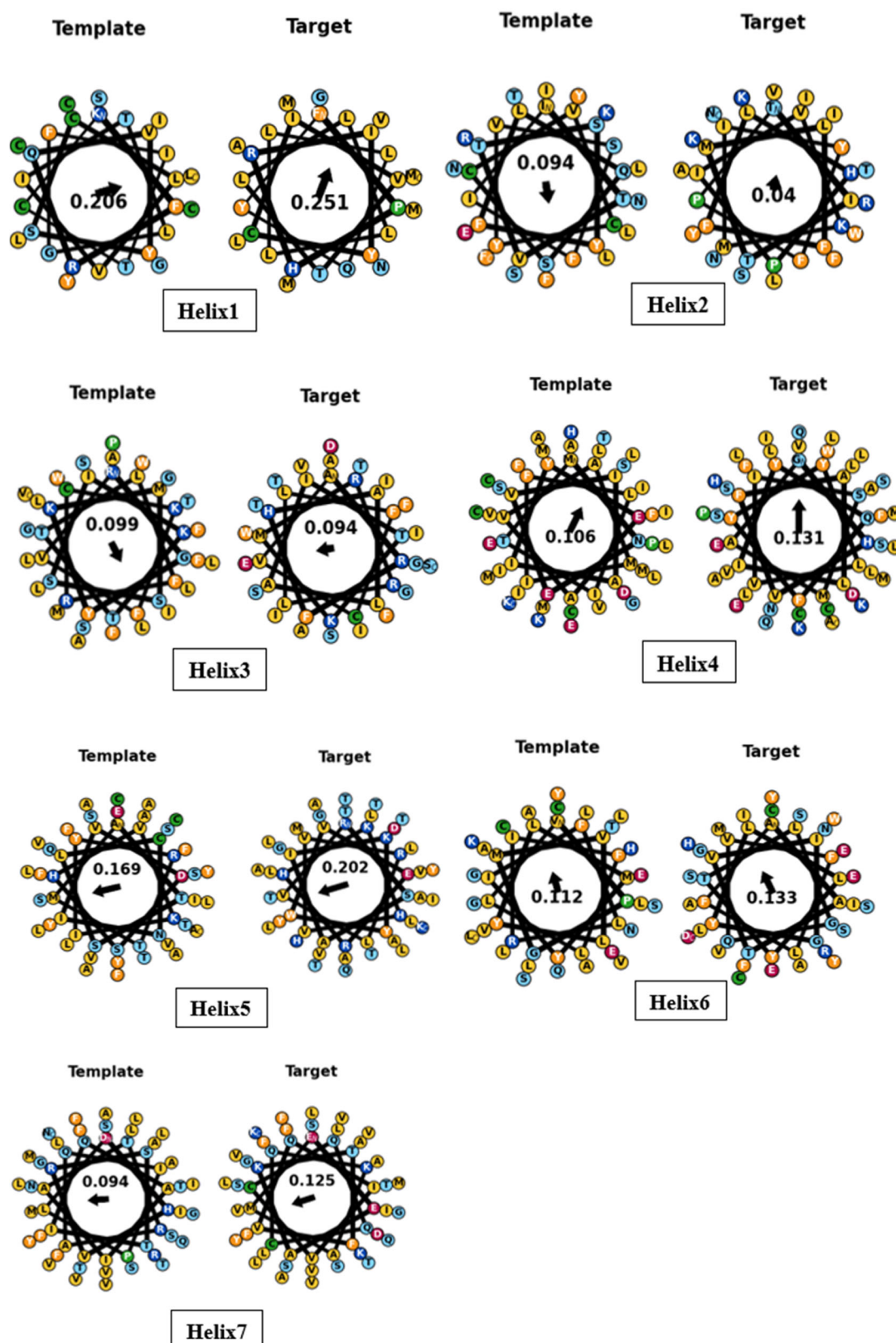


Figure S2. Helical wheel plots for the *AbOrco* sequence with the template *MhOR5*. Template PDB ID:7LIG.

```

Helix1
Template: KFRQVYSCLVITLGFITCSCYICGLCL
Target:   FPHRIYCIIVTLLLLLQYGMMAVNLM

Helix2
Template: ITVTSYFLQSCVCYVSFIINSRKLETLFNYLF-
Target:   TITMLFFLHPIVKMIYFPVRSKIFYKT LAIWN

Helix3
Template: RGYKMSSIKTTLFRCKFVAFSLGILSFFGWLMTLLPLAV
Target:   -ARFHALAITKMRRLLFCVAGATIFSVISWTGITFIEDS-

Helix4
Template: MNEVIAIYEAVAMIFLITAPMSSDIMFCVLMIFIVEHLKCLGMAIECTLK---
Target:   GHVFALIYQFYLVISMAVSNLDFCSWLLFACEQLQHLKAIMKPLMELSA

Helix5
Template: ATSLCNIVDSHVKIYRTMEIVQSVYSSYFATLFFTSCLAVCALAYFLAA
Target:   RSAIKYWVERHKHVRLVTAVGDAYGVALLHMLTTTITLTLAYQATK

Helix6
Template: VPGMVLYLMIYIFLRIFLLCLLATEVAEQGLNLCHAGYSSKLVL
Target:   AATVIGYLLYTLGQVFLFCIFGNRLIEESSVMEAAYSCHWYD

Helix7
Template: DHVRSTIQAIATRAQIPLSITGARFFT VNLSFLASMAGVMLTYFIVLLQVN
Target:   EEAKTFVQIVCQCQKAMSISGAKFFT VSLDLFASVLGAVVTYFMVLVQLK

```

Figure S3. iBio-GATS helix-wise alignment for the *AbOrco* target, with the, *MhOR5* template.

```

>Pl;Template
structureX:7LIG:: :: :: 0.00: 0.00
DDYIHLR-KWI-K-RIGIILRISGHW--PF--R-LPHEK--R---N---Q-HKSKFRQVYSCLVITLGFITCSCYICGLC
LSES--IAQALN---NITVTSYFLQSCVCYVSFIINSRKLETLFNYLF-ENEVVGCP-----RGYKMSSIKTTLFRCK
FVAFSLGILSFFGWLMTLLPLAVLV--V-D-QT-S--L---RF--VEAW-YPFDT-TTSPMNEVIAIYEAVAMIFLIT
APMSSDIMFCVLMIFIVEHLKCLGMAIECTLK-----G-----
----D-----ATSLCNIVDSHVKIYRTMEIVQSVYSSYFATLFFTSCLAVCALAY
FLAATSTS---FTRVPGMVLYLMIYIFLRIFLLCLLATEVAEQGLNLCHAGYSSKLVLASDHVRSTIQAIATRAQIPLSIT
GARFFT VNLSFLASMAGVMLTYFIVLLQVN*
>Pl;Target
sequence:6C70:: :: :: 0.00: 0.00
---GPGRAK--FKHQ-G--L-V-AD-LLP-NIRVM-QG-VGHFMFNYYSEGKK-FPHRIYCIIVTLLLLLQYGMMAVNLM
M-ESDDV-D--DLTANTITMLFFLHPIVKMIYFPVRSKIFYKT LAIWNPN---S-HPLFAESN-ARFHALAITKMRRLL
FCVAGATIFSVISWTGITFIEDS--VKRITDPETNETTIIPR-LMIRT-FYPFNAMSGA-GHVFALIYQFYLVISMA
VSNLDFCSWLLFACEQLQHLKAIMKPLMELSATLDTVPNSGELFKAGSADHLRESQGVQPSGNGDNVLDVLRGIY
SNRQDFTATFRPTAGTTFNNGVGPNGLTKKQEMLVRSIAIKYWVERHKHVRLVTAVGDAYGVALLHMLTTTITLTLAY
QATK-VNGVNVY--AATVIGYLLYTLGQVFLFCIFGNRLIEESSVMEAAYSCHWYDGSEAKTFVQIVCQCQKAMSIS
GAKFFT VSLDLFASVLGAVVTYFMVLVQLK*

```

Figure S4. Full alignment for the *AbOrco* sequence, with the *MhOR5* template.

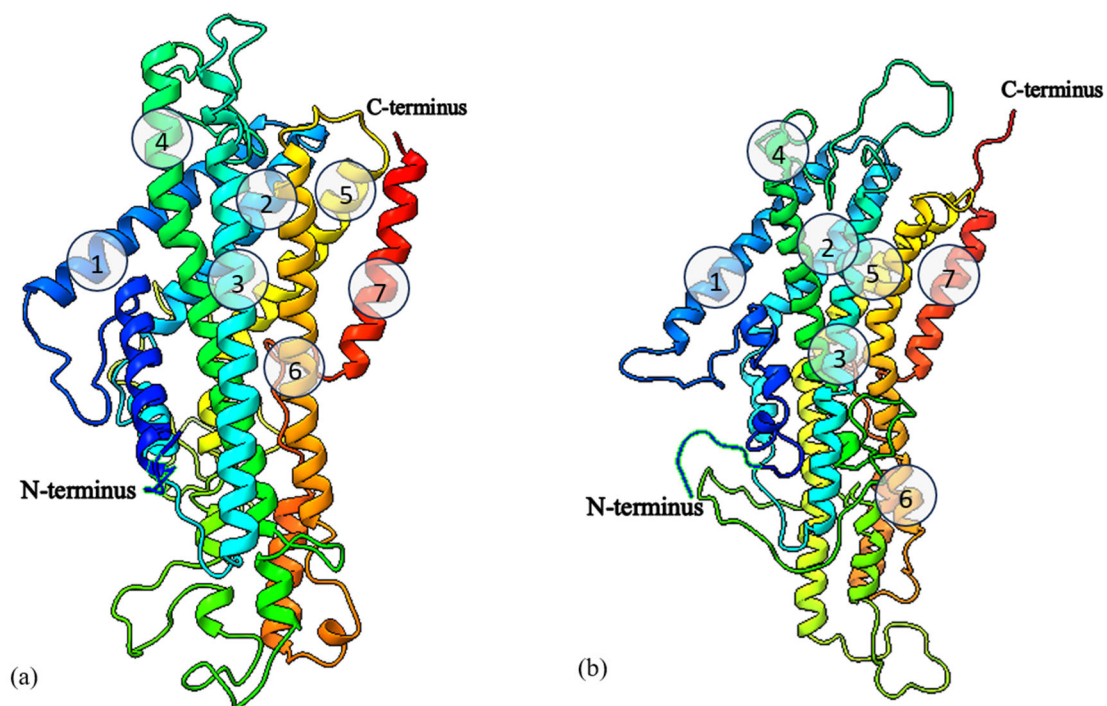


Figure S5. Predicted models for *AbOrco* and *MhOR5*. (a) The model predicted for *AbOrco* sequence using the template *MhOR5* and (b) the model predicted for *MhOR5* using the template *AbOrco*. The seven helices, C-terminus and N-terminus are labelled.

Table S1. Helix definitions for *DmOR59b* (UniProt ID: Q9W1P8)

Helix	Helix start	Helix end	Centre residue
1	47	73	Y60
2	88	119	F104
3	121	160	S146
4	183	232	D206
5	242	290	S267
6	297	339	L321
7	342	392	F367

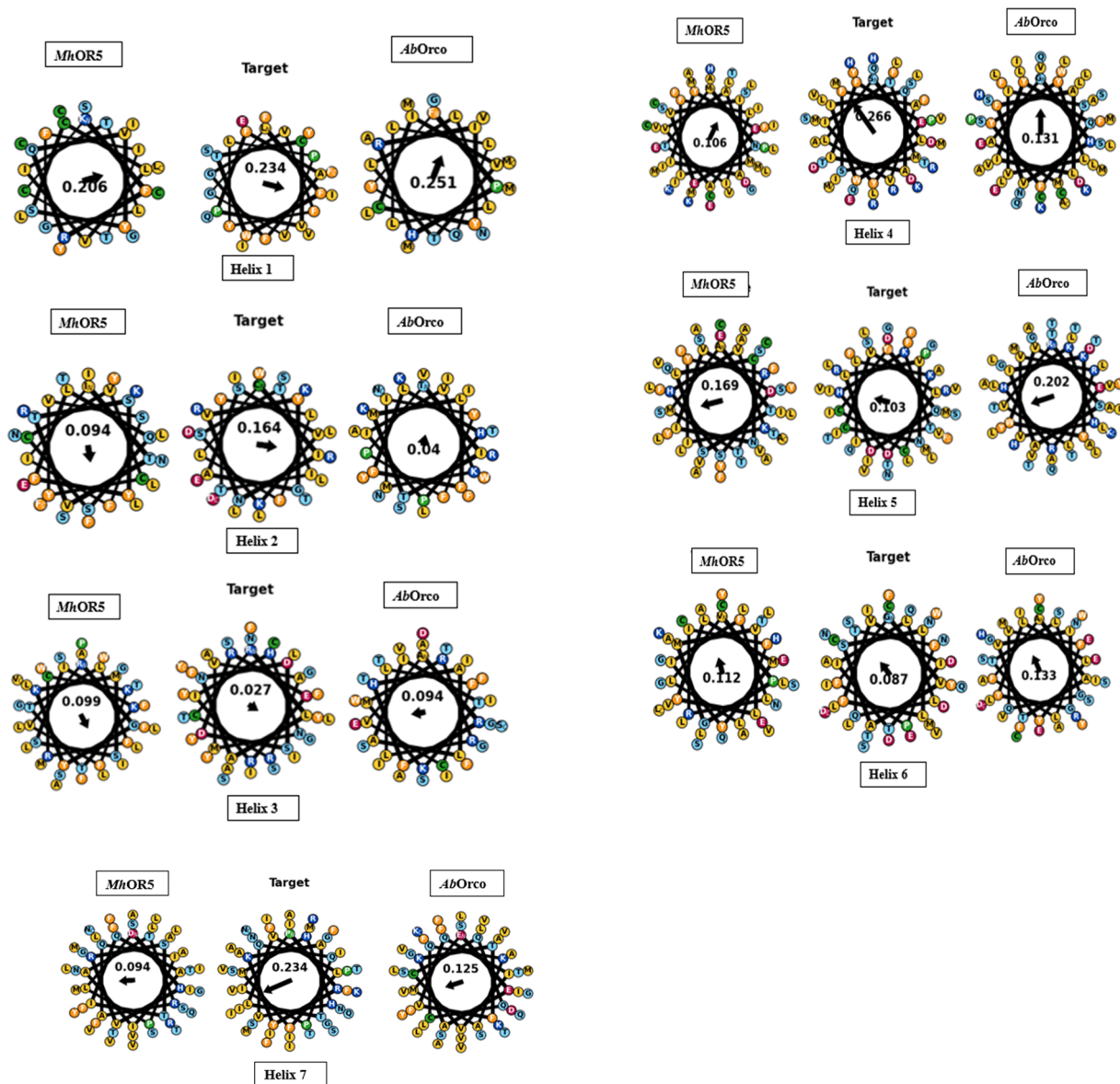


Figure S6. Helical wheel plots for *DmOR59b* compared with templates, *MhOR5* and *AbOrco*. The helical moments of *DmOR59b* are better aligned to those of *MhOR5* than those of *AbOrco*.

	10	20	30	40	50	He1	60
7LIG_1Chain_A	-----DDYIHLRKWIKRIGIILRISGHWPFRLPHEKRNQHKSKFRQVYSCLVI						
NP_523822.1od	MAVFKLIKPAPLTEKVQSRQGNIIYLRAMWLIGWIP---PKEGVLRYVYLFWTCVPFAFG						
	70	80	90	100	He2	110	120
7LIG_1Chain_A	TLGFITCSCYICGLCLS---ESIAQALNNITVTSYFLQSCVCVVSFIINSRKLETLFNY						
NP_523822.1od	VFYLPVGFIIISYVQEFKNFTPGFELTSLQVCINVGASVKSTITYLFLWRLRKTEILLDS						
	130	140	150	He3	160	170	180
7LIG_1Chain_A	LFENEVVGCPRGYKMSSIKTTLFRCCKFVAFSLGILSFFGWLMTLLPLAVLVVDQTSRFR						
NP_523822.1od	LDK-----RLANDSDRERIHNMAVRCNYAFLIYSFIYCGYAGSTFLSYALSGRPPWSV						
	190	200	He4	210	220	230	240
7LIG_1Chain_A	VEAWYPFDTTSPMNEVIAIYEAVAMIFLITAPMSSDIMFCVLMIFIVEHLKCLGMAIEC						
NP_523822.1od	YNPFIDWR-DGMGSLWIQAIFEYITMSFAVLQDQLSDTYPLMFTIMFRAHMEVLKDHVRS						
	250	260	270	He5	280	290	300
7LIG_1Chain_A	TLKGD-----ATSLCNIVDSHVKIYRTMEIVQSVYSSYFATLFFTSCLAVCALAYFLA						
NP_523822.1od	LRMDPERSEADNYQDLVNCVLDHKTILKCCDMIRPMISRTIFVQFALIGSVLGLTLVNVF						
	310	320	330	He6	340	350	360
7LIG_1Chain_A	ATSTSFTRVPGMVLVLMYIFLRIFLLCLLATEVAEQGLNLCHAGYSSKLVLASDHVRSTI						
NP_523822.1od	FFS-NFWKGVASLLFVITILLQTFPFCYTCNMLIDDAQDLSNEIFQSNWVDAEPRYKATL						
	370	380	390	He7	400	410	
7LIG_1Chain_A	QAIATRAQIPLSITGARFFTVNLSFLASMAGVMLTYFIVLLQVN-----						
NP_523822.1od	VLFMHVQQPIIFIAGGIFPISMNSNITVAKFAFSIITIVRQMNLAEQFQ						

Figure S7. Final pairwise alignment of *Dm*OR59b with *Mh*OR5. The AlignMe 1.2 output was edited manually to remove gaps in the helical regions using the MEGA 11 sequence viewer.

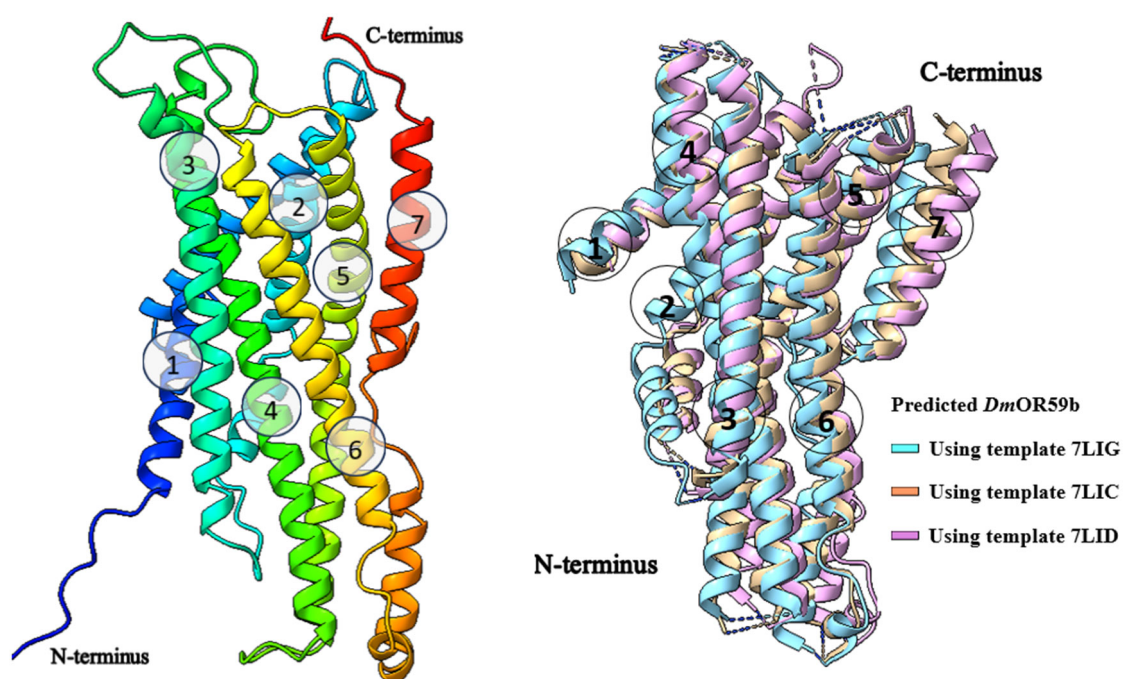


Figure S8. Predicted models for *DmOR59b* sequence. (a) The top predicted model for *DmOR59b* (WfWtOR59b) using the template *MhOR5* (PDB:7LIG). (b) Structural overlay of predicted models (showing only helices) of *DmOR59b* (WfWtOR59b) using all three templates of *MhOR5*. The seven helices are labelled.

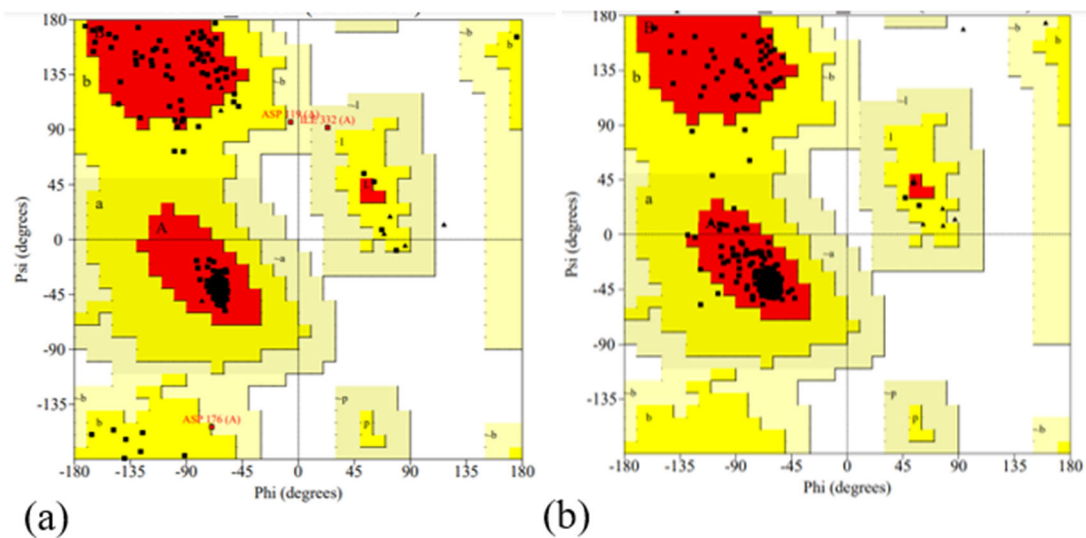


Figure S9. Ramachandran plots for *DmOR59b*. (a) iBio-GATS workflow model, WfWtOR59b and (b) AlphaFold2 model, AfWtOR59b.

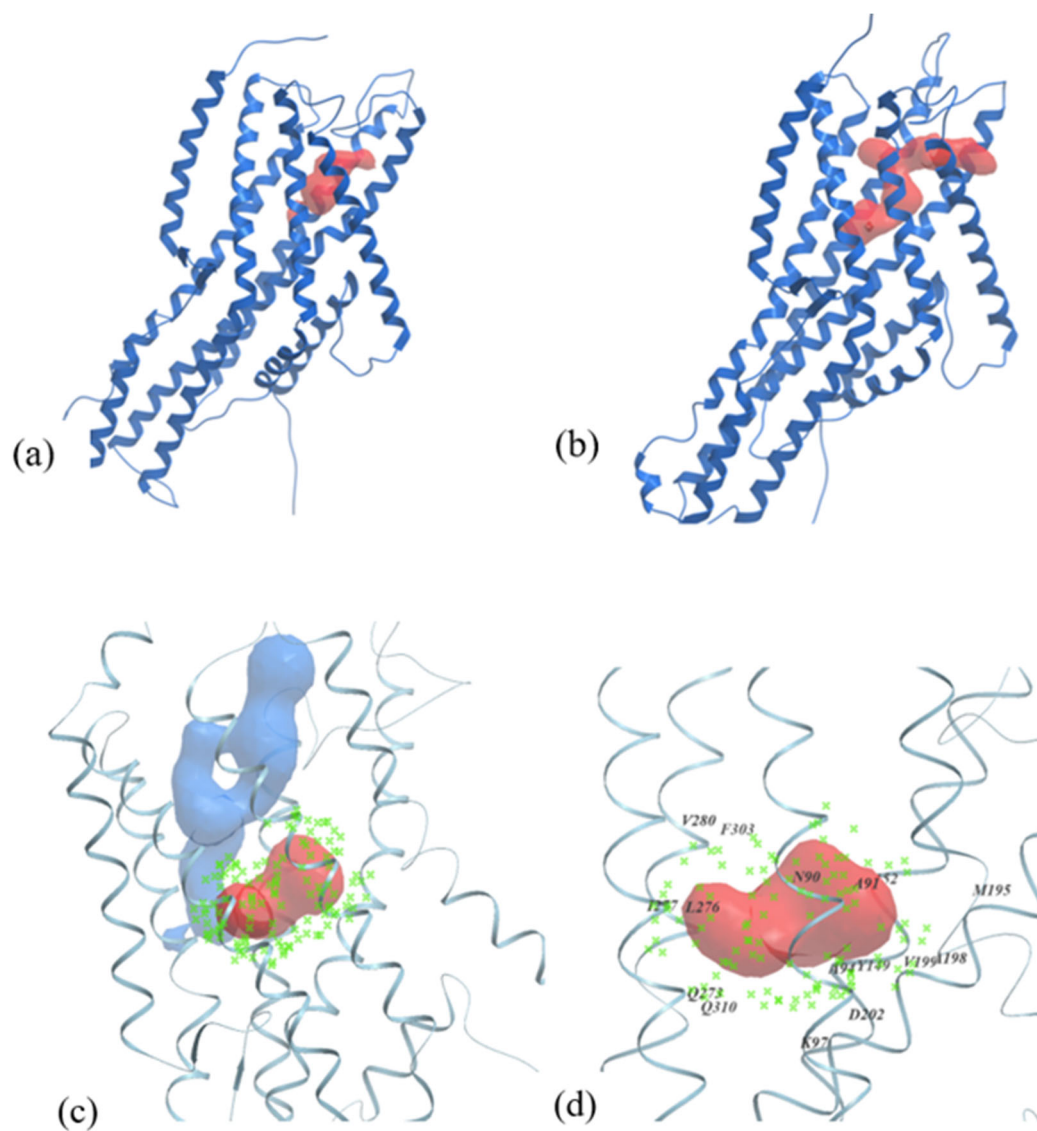


Figure S10. Binding pockets for predicted models of *DmOR59b* with DEET

The ICM binding pockets of the workflow (Wf) and Alpha Fold (Af) wild type (Wt) and mutant (Mt) models of *DmOR59b*. (a) WfWtOR59b, (b) WfMtOR59b, (c) AfWtOR59b and (d). AfMtOR59b.

	10	20	30	40	50	He1	60
6C70_1chainsA	-----FKHQGLVADLLPNIRVMQGVGHFMFNYYSEGKK	FPHRIYCI	VTLLLLLL	QYGM			
spQ9VNB5ORCO_	MTTSMQPSKYTGLVADLMPNIRAMKYSGLFMHNF-TGGS	AFMKKVYSSV	HLVFLLMQ	FTF			
	70	80	He2	90	100	110	120
6C70_1chainsA	MAVNLMMESDDVDDL	TANTITMLF	FLHP	IVKMIYFP	VRSKIFYK	TLAIWNN	PNSHPLFAE
spQ9VNB5ORCO_	ILVNMALNAEEVNELSGN	TITTLFF	THCITK	FIYLA	VNQKNFYR	TLNIWNQ	VNTHPLFAE
	130	140	He3	150	160	170	180
6C70_1chainsA	SNARFHALAITKMRRLL	FCVAGATIF	SVISWTG	ITFIEDS	-----	PIPR	L
spQ9VNB5ORCO_	SDARYHSIALAKMRKLF	FLVMLTT	VASATAW	TTITFFG	DSVKM	VVDHETN	SSIPVEIPRL
	190	200	He4	210	220	230	240
6C70_1chainsA	MIRTFYFPFNAMSGA	GHVFAL	IYQFY	YLVISMA	VNSLDV	LCFCSW	LLFACEQLQHLKAIMK
spQ9VNB5ORCO_	PIKSFYPWNASHGM	FYMISFA	FIQIYY	VLFSMI	HSNLC	DMFCSW	LIFACEQLQHLKGIMK
	250	260	270	280	290	300	
6C70_1chainsA	PLMELSAT	-----					
spQ9VNB5ORCO_	PLMELSA	SLDTYR	PNNSAAL	FRSL	SANSK	SELIHNEE	KDPGTDMDMSGIYSSKADWGAQFR
	310	320	330	He5	340	350	360
6C70_1chainsA	-----	GLTKKQ	EMLVR	SAIKY	WVERH	KHVVRL	VTAVGDAY
spQ9VNB5ORCO_	APSTLQSF	GGNGGG	GNGLV	NGANP	NGLTKK	QEMMV	RSIAIKYWVERHKKHVRLVAAIGDTY
	370	380	390	He6	400	410	420
6C70_1chainsA	GVALLLHMLTTTITL	TLLAYQ	ATKVN	GVNVY	AATV	IGYLLY	TLGQVFLFCIFGNRLIEES
spQ9VNB5ORCO_	GAALLHMLTSTIKL	TLLAYQ	ATKING	VNVYA	FTTV	GYLGY	ALAQVFHFCIFGNRLIEES
	430	440	450	He7	460	470	480
6C70_1chainsA	SSVMEAAYSCHWYD	GSEEA	KTFVQ	IVCQQ	CQKAMS	ISGAKFF	TVSLDLFASVLGAVVTYF
spQ9VNB5ORCO_	SSVMEAAYSCHWYD	GSEEA	KTFVQ	IVCQQ	CQKAMS	ISGAKFF	TVSLDLFASVLGAVVTYF
	487						
6C70_1chainsA	MVLVQLK						
spQ9VNB5ORCO_	MVLVQLK						

Figure S11. Final pairwise alignment of *Dmel*Orco with *Ab*Orco (PDB ID:6C70) obtained from AlignMe 1.2. The sequence identity between the two sequence is 55.2%.

Table S2. Helix definitions for *Dmel*Orco (UniProt ID: Q9VNB5)

Helix	Helix start	Helix end	Centre residue
1	40	66	L54
2	78	110	L94
3	122	159	S146
4	194	246	D217
5	335	383	G260
6	391	433	L415
7	436	486	F461

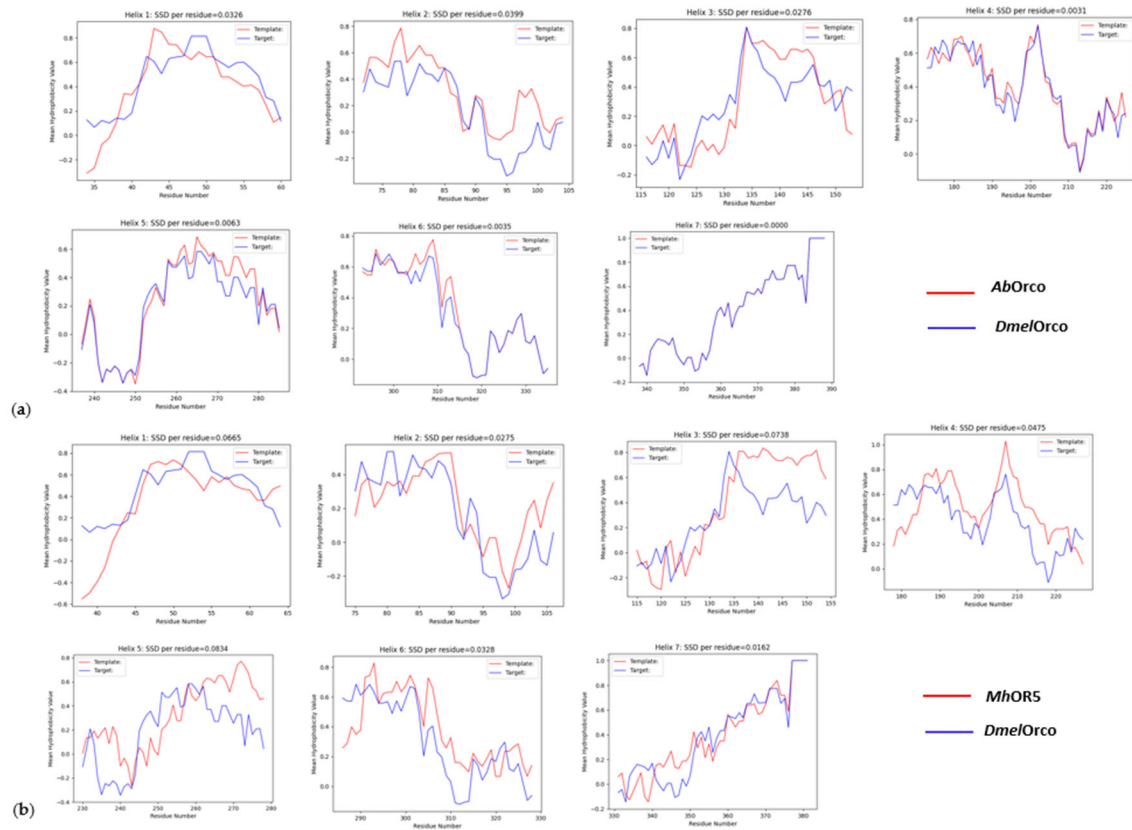


Figure S12. The result summary output from workflow shows the hydrophobicity plots for all the seven helices for *DmelOrco* compared with both the templates, *AbOrco* and *MhOR5*

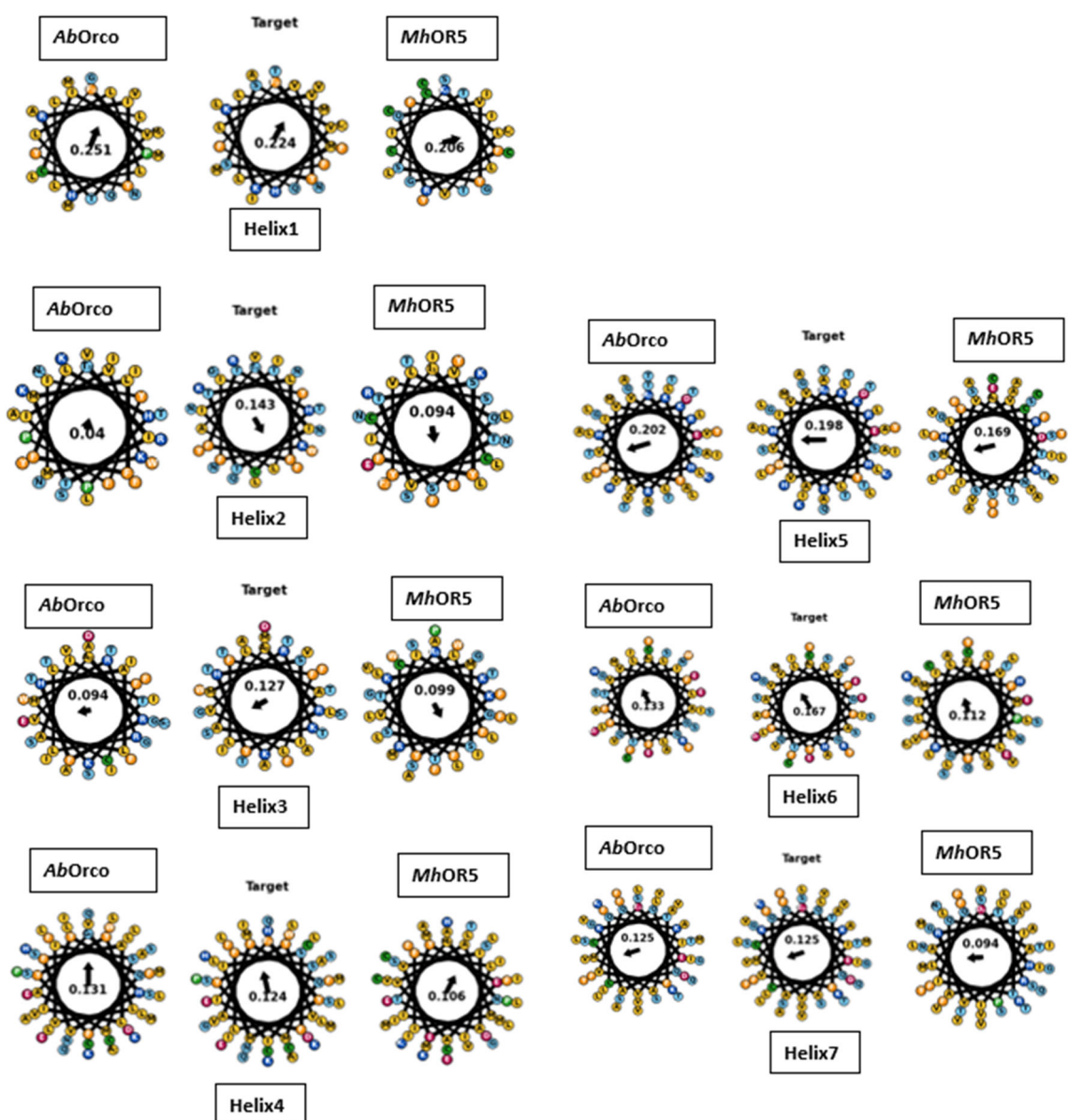


Figure S13. The result summary output from workflow shows the helical plots for all the seven helices for *DmelOrco* compared with both the templates, *AbOrco* and *MhOR5*. The helical moments of *DmelOrco* are better aligned to those of *AbOrco* than those of *MhOR5*.

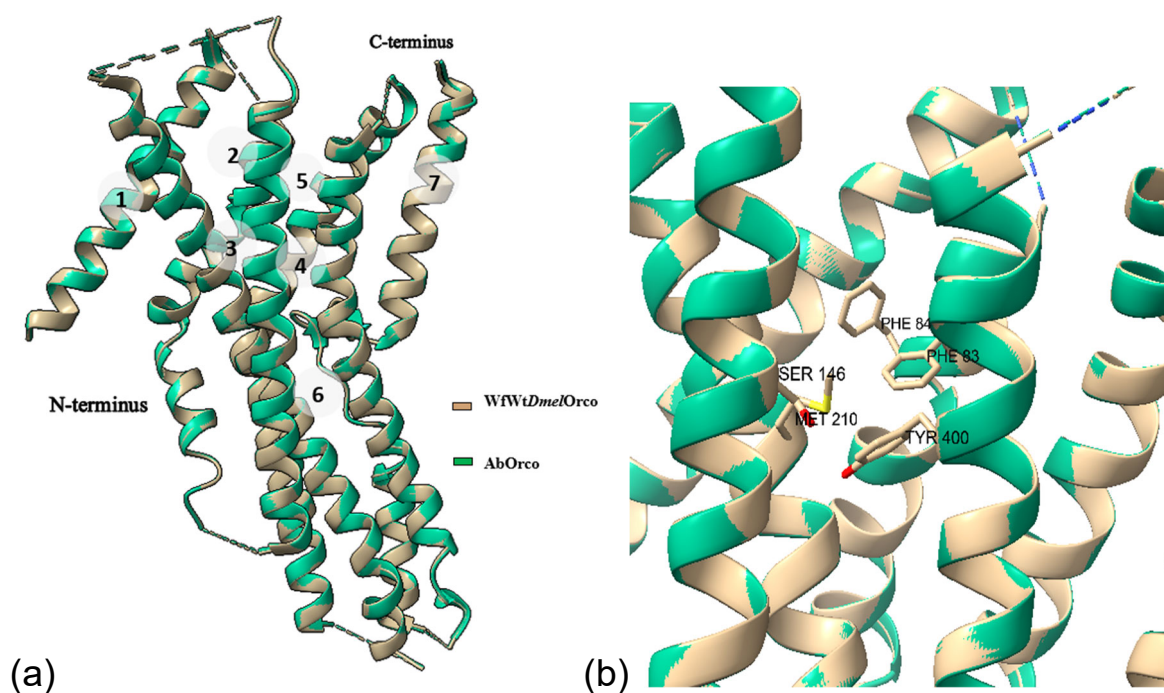


Figure S14. Structural overlay of predicted WfWtDmelOrco model with its template AbOrco. (PDB ID:6C70). (a) Superposed helical regions, with an RMSD of 0.1 Å, and the seven helices numbered. (b) The five residues (F83, F84, M210, S146, Y400) that were frequently identified to interact with VUAA1 in the binding site are captured.

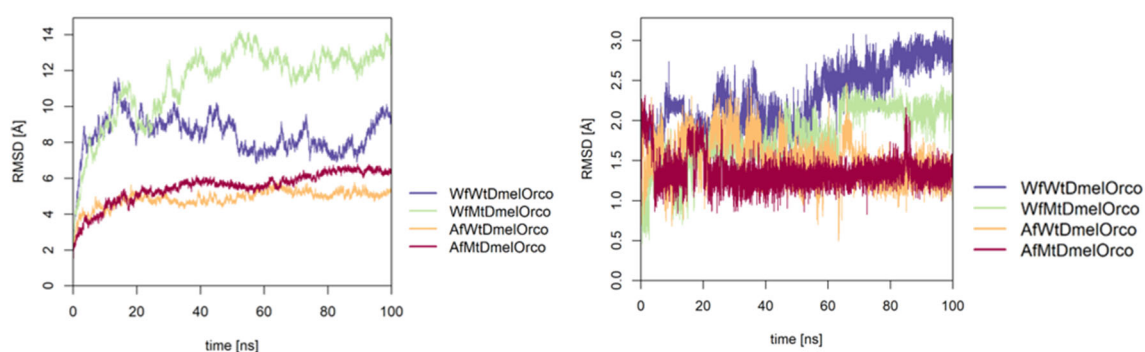


Figure S15. RMSD plots of all four *DmelOrco* complexes over 100 ns of MD simulations. (a) *DmelOrco* complex with bound VUAA1 and (b) the ligand VUAA1 alone.

```

>7LIG|Chains A||MhOR5|Machilis hrabei (438506)
DDYIHLRKWIKRIGIILRISGHWPFRLPHEKRNQHKSKFRQVYVSLVITLGFITCSCYIGLCLSEISIAQALNNITVTS
YFLQSCVCYVSFIINSRKLETLFNYLFENEVVGCPRGYKMSSIKTTLFRCKFVAFSLGILSFFGWLMWTLLPLAVLVVD
QTSLRFVEAWYPFDTTTSPMNEVIAIYEAVAMIFLITAPMSSDIMFCVLMIFIVEHLKCLGMAIECTLKGDATSLCNIV
DSHVKIYRTMEIVQSVYSSYFATLFFTCLAVCALAYFLAATSTSFTRVPGMVLYLMYIFLRIFLLCLLATEVAEQGLN
LCHAGYSSKLVLASDHVRSTIQAIATRAQIPLSITGARFFTVNLSFLASMAGVMLTYFIVLLQVN
>6C70|Chains A|Odorant receptor|Apocrypta bakeri (490712)
FKHQGLVADLLPNIRVMQGVGHFMFNYYSEGKKFPHRIYCIVTLLLLLLQYGMMAVNLMMESDDVDDLTTANTITMLFFL
HPIVKMIYFPVRSKIFYKTLAIWNPNNSHPLFAESNARFHALAITKMRRLLFCVAGATIFSVISWTGITFIEDSPIPRL
MIRTFYPFNAMSGAGHVFAIYQFYLVISMAVSNLSDLVFCSWLLFACEQLQHLKAIMKPLMELSATGLTKKQEMLVR
SAIKYWVERHKHVRLVTAVGDAYGVALLHMLTTTITLTLLAYQATKVNNGVNVYAATVIGYLLYTLGQVFLFCIFGNR
LIEESSVMEAAYSCHWYDGSEEAKTFVQIVCQQCQKAMSISGAKFFTVSLLDFASVLGAVVTYFMVLVQLK

```

Figure S16. The structure-derived sequences of the template structures. *MhOR5*: PDB ID: 7LIG and *AbOrco*: PDB ID: 6C70.

Table S3. Helix positions for *AbOrco* (PDB ID:6C70)

Helix	Helix start	Helix end	Centre residue
1	34	60	L47
2	72	104	F88
3	116	153	S140
4	173	225	D196
5	237	285	G262
6	293	335	L317
7	338	388	F363

Table S4. Helix positions for *MhOR5* (PDB ID:7LIG)

Helix	Helix start	Helix end	Centre residue
1	38	64	G51
2	75	106	F91
3	115	154	S140
4	178	227	D201
5	230	278	S255
6	286	328	V310
7	331	381	F356