Supplementary Information

Structural bioinformatic study of six human olfactory receptors and their AlphaFold3 predicted water-soluble QTY variants and OR1A2 with an odorant

Finn Johnsson¹, Taner Karagöl^{2,¶}, Alper Karagöl^{2,¶}, Shuguang Zhang^{3,*}

Email:

Finn Johnsson, finn.johnsson@icloud.com	ORCID: <u>0009-0009-9132-9488</u>
Taner Karagöl, taner.karagol@gmail.com	ORCID: <u>0009-0005-1011-7661</u>
Alper Karagöl, alper.karagol@gmail.com	ORCID: <u>0009-0001-7864-0732</u>
Shuguang Zhang, Shuguang@MIT.EDU	ORCID: <u>0000-0002-3856-3752</u>

Table S1. SwissParam scores of octanoic acid - receptor complexes.

Receptor	SwissParam Score ¹
OR52cs	-6.1091
OR1A1	-6.0243
OR1A2	-6.1848
OR51E1	-6.0030
OR51E2	-5.7291

¹SwissParam scores were computed by SwissDock web server. Best scores for each complex were listed.

¹St Paul's School, London, Lonsdale Road, London, SW13 9JT, UK

²Istanbul University Istanbul Medical Faculty, Istanbul, Turkey

³Laboratory of Molecular Architecture, Media Lab, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA, 02139, USA

These authors contribute equally.

^{*}To whom the correspondence should be addressed.

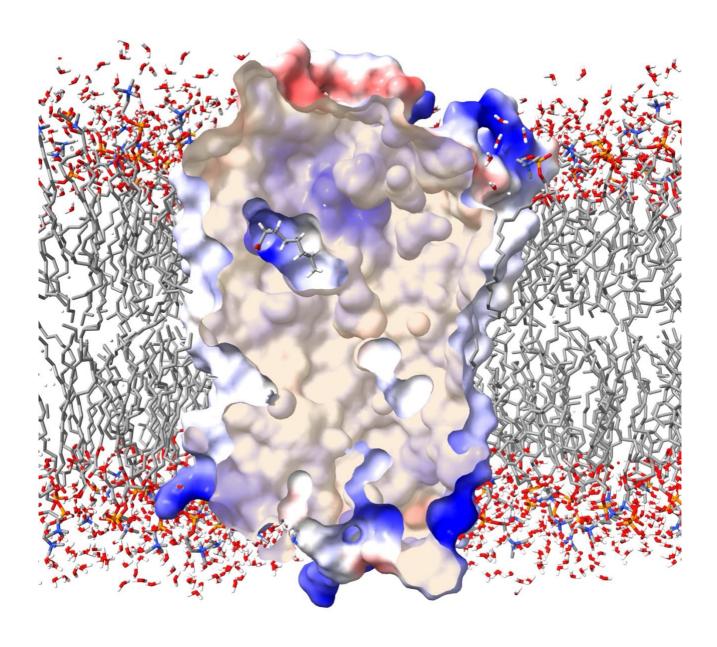
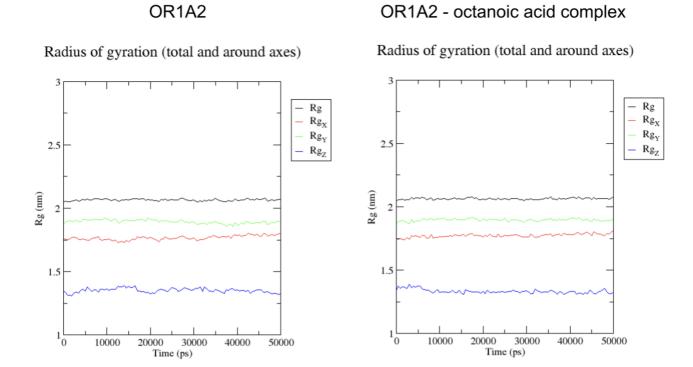


Figure S1. OR1A2-octanoic acid complex in the membrane environment. Octanoic acid (OCA) is positioned within a hydrophobic cleft. The electrostatic potential surface highlights the compatibility of the binding pocket for the hydrophobic tail and polar head group of OCA.



OR1A2 - octanoate complex

Radius of gyration (total and around axes)

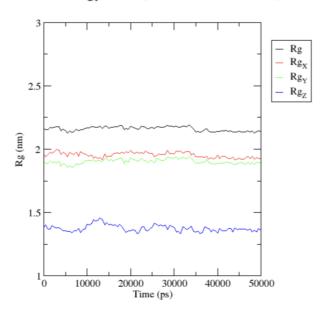
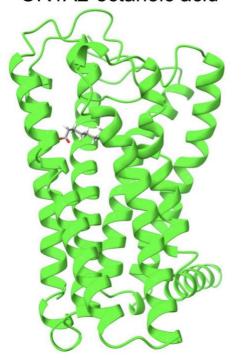


Figure S2. Radius of gyration during the 50ns MD simulation.

OR1A2-octanoic acid



OR1A2-octanoate

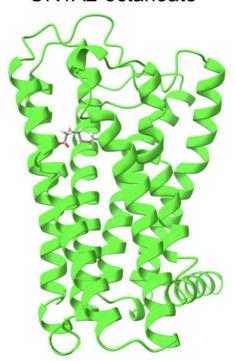


Figure S3. OR1A2 in complex with octanoic acid (left) and octanoate (right). Structural representation of the OR1A2 complex bound with octanoic acid (eft) and octanoate (right). The AlphaFold predicted protein structure in green cartoon representation, with the ligand octanoic acid shown as a stick model.

Energetic Components [Per-residue] Prot-Memb | NORMAL | PB | Delta | TDC

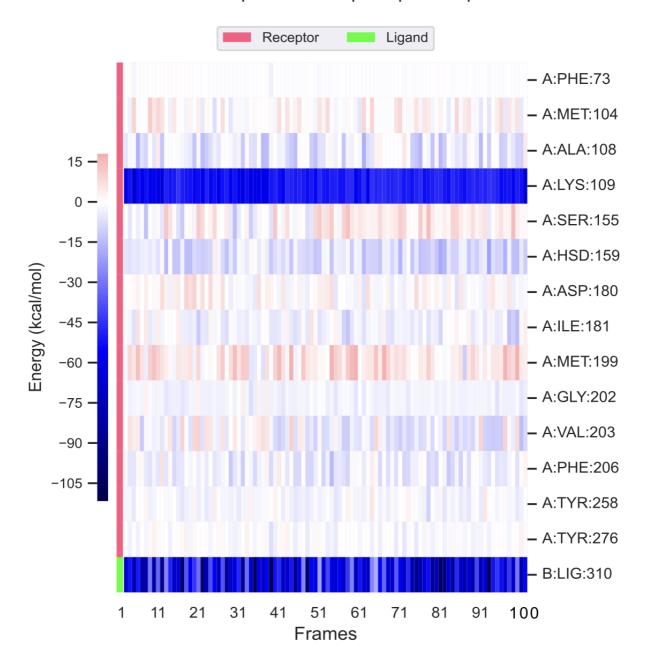


Figure S4. Residue decomposition analysis of OR1A2-octanoate complex during 50ns MD simulation. Heatmap of per-residue energetic contributions to octanoate binding (B:LIG) across 100 simulation frames (50ns). The heatmap shows the contributions of individual residues (on the y-axis) to the binding energy (indicated by the color scale, red to blue).

Energetic Components Prot-Memb | Normal | PB | Delta | TOTAL

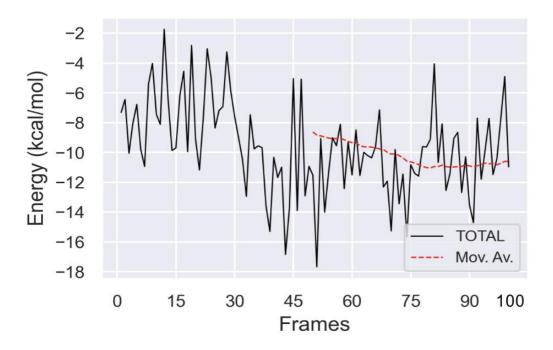
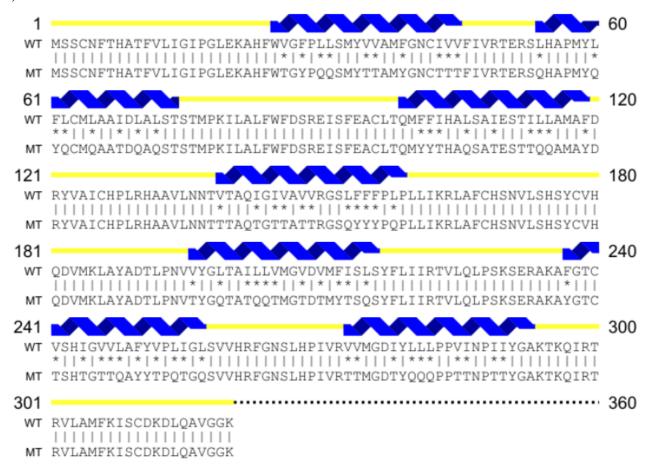


Figure S5. MMPSA calculated binding free energy of OR1A2-octanoate complex during 50ns MD simulation. MMPBSA calculated total binding energy of octanoate to the OR1A2 over 100 simulation frames (50ns). The black line represents the binding energy, while the red dashed line indicates the moving average trend across the frames.

Figure S6a-f. For clarity and easy visualization, the enlarged protein sequence alignments of six native olfactory receptor proteins with their water-soluble QTY variants. The symbols | and * indicate whether amino acids are identical or different, respectively. Note the Q, T, and Y amino acids replacing L, V and I, and F, respectively. The alpha helices (blue) are shown above the protein sequences. The characteristics of natural and QTY variants listed are isoelectric focusing (pI), molecular weight (MW), total variation %, and transmembrane variation %. The alignments are: a) OR51E2 *vs* OR51E2^{QTY}, b) OR52cs *vs* OR52cs ^{QTY}, c) TAAR9 *vs* TAAR9^{QTY}, d) OR51E1 *vs* OR51E1^{QTY}, e) OR1A1 *vs* OR1A1^{QTY}, and f) OR1A2 *vs* OR1A2^{QTY}.

a) OR51E2 vs OR51E2^{QTY}



b) OR52cs vs OR52cs QTY,

1		60
WT	MPTSNHTSFHPSSFLLVGIPGLESVHIWISIPFCAMYLIALLGNSTLLFVIKTERSLHEP	
MT	MPTSNHTSFHPSSFLLVGIPGLESVHIWTSTPYCAMYQTAQQGNSTQQYTTKTERSLHEP	
61		120
WT	MYYFLAMLAATDLVLSTSTIPKMLAIFWFNLKEISFDACLTQMFFIHSFTGMESGVLLAM	
MT	$\verb"MYYYQAMQAATDQTQSTSTTPKMQAIFWFNLKEISFDACQTQMYYTHSYTGMESGTQQAM"$	
121		180
WT	AFDRYVAICYPLRYTTILTNKVIGKIGMAVVLRAVLLVIPFPFLLKRLPFCGTNIIPHTY * ** *** **** ** *	
MT	AYDRYVAICYPLRYTTILTNKTTGKTGMATTQRATQQTTPYPYQLKRLPFCGTNIIPHTY	
181		240
WT MT		
241		300
WT	STCGSHICVILAFYTPAFFSFLTHRFGHHIPPYIHILLANLYLLVPPMLNPIIYGVKTKQ	
MT	${\tt STCGSHTCTTQAYYTPAYYSFLTHRFGHHIPPYTHTQQANQYQQTPPMQNPTTYGTKTKQ}$	
301		360
WT	IRERVLKIFFKKK	
МТ	IRERVLKIFFKKK	

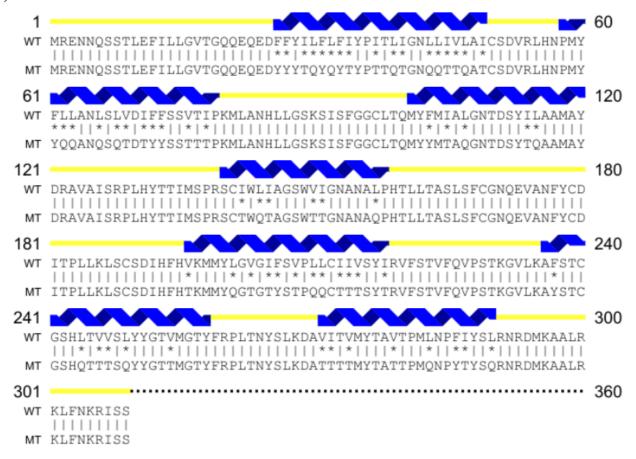
c) TAAR9 vs TAAR9^{QTY}

1		60
WT	MVNNFSQAEAVELCYKNVNESCIKTPYSPGPRSILYAVLGFGAVLAAFGNLLVMIAILHF	
MT	MVNNFSQAEAVELCYKNVNESCIKTPYSPGPRSILYATQGYGATQAAYGNQQTMTATQHF	
61		120
WT	KQLHTPTNFLIASLACADFLVGVTVMPFSTVRSVESCWYFGDSYCKFHTCFDTSFCFASL	
MT	${\tt KQLHTPTNYQTASQACADYQTGTTTMPYSTVRSVESCWYFGDSYCKYHTCYDTSYCYASQ}$	
121		180
WT	FHLCCISVDRYIAVTDPLTYPTKFTVSVSGICIVLSWFFSVTYSFSIFYTGANEEGIEEL * * * *	
MT	YHQCCTSTDRYIAVTDPLTYPTKFTVSTSGTCTTQSWYYSTTYSYSTYYTGANEEGIEEL	
181		240
WT	VVALTCVGGCQAPLNQNWVLLCFLLFFIPNVAMVFIYSKIFLVAKHQARKIESTASQAQS	
MT	VVALTCVGGCQAPLNQNWTQQCYQQYYTPNTAMTYTYSKIFLVAKHQARKIESTASQAQS	
241		
		300
WT	SSESYKERVAKRERKAAKTLGIAMAAFLVSWLPYLVDAVIDAYMNFITPPYVYEILVWCV	300
		300
	SSESYKERVAKRERKAAKTLGIAMAAFLVSWLPYLVDAVIDAYMNFITPPYVYEILVWCV *** *** ** *** * SSESYKERVAKRERKAAKTQGTAMAAYQTSWQPYQTDATTDAYMNFITPPYVYETQTWCT	360
мт 301	SSESYKERVAKRERKAAKTLGIAMAAFLVSWLPYLVDAVIDAYMNFITPPYVYEILVWCV *** *** ** *** * SSESYKERVAKRERKAAKTQGTAMAAYQTSWQPYQTDATTDAYMNFITPPYVYETQTWCT	

d) OR51E1 vs OR51E1^{QTY}

1		60
	MMVDPNGNESSATYFILIGLPGLEEAQFWLAFPLCSLYLIAVLGNLTIIYIVRTEHSLHE	
61 wt	PMYIFLCMLSGIDILISTSSMPKMLAIFWFNSTTIQFDACLLQMFAIHSLSGMESTVLLA	120
	PMYTYQCMQSGTDTQTSTSSMPKMLAIFWFNSTTIQFDACQQQMYATHSQSGMESTTQQA	
121 wt	MAFDRYVAICHPLRHATVLTLPRVTKIGVAAVVRGAALMAPLPVFIKQLPFCRSNILSHS	180
МТ	MAYDRYVAICHPLRHATVLTLPRVTKTGTAATTRGAAQMAPQPTYTKQLPFCRSNILSHS	
101		240
	YCLHQDVMKLACDDIRVNVVYGLIVIISAIGLDSLLISFSYLLILKTVLGLTREAQAKAF ** ***** ****	240
wт мт 241	YCLHQDVMKLACDDIRVNVVYGLIVIISAIGLDSLLISFSYLLILKTVLGLTREAQAKAF ** ***** * *****	300
wт мт 241 wт мт	YCLHQDVMKLACDDIRVNVVYGLIVIISAIGLDSLLISFSYLLILKTVLGLTREAQAKAF	300
wт мт 241 wт мт 301 wт	YCLHQDVMKLACDDIRVNVVYGLIVIISAIGLDSLLISFSYLLILKTVLGLTREAQAKAF ** *****************	

e) OR1A1 vs OR1A1 QTY



f) OR1A2 vs OR1A2^{QTY}

1		60
WT	MKKENQSFNLDFILLGVTSQQEQNNVFFVIFLCIYPITLTGNLLIILAICADIRLHNPMY	
MT	MKKENQSFNLDFILLGVTSQQEQNNTYYTTYQCTYPTTQTGNQQTTQATCADIRLHNPMY	
61		120
WT	FLLANLSLVDIIFSSVTIPKVLANHLLGSKFISFGGCLMQMYFMIALAKADSYTLAAMAY *** * ** **	
MT	YQQANQSQTDTTYSSTTTPKVLANHLLGSKFISFGGCLMQMYYMTAQAKADSYTQAAMAY	
121		180
WT	DRAVAISCPLHYTTIMSPRSCILLIAGSWVIGNTSALPHTLLTASLSFCGNQEVANFYCD	
МТ	DRAVAISCPLHYTTIMSPRSCTQQTAGSWTTGNTSAQPHTLLTASLSFCGNQEVANFYCD	
181		240
WT	<pre>IMPLLKLSCSDVHFNVKMMYLGVGVFSLPLLCIIVSYVQVFSTVFQVPSTKSLFKAFCTC </pre>	240
MT	IMPLLKLSCSDVHFNVKMMYLGVGVFSLPLLCIIVSYVQVFSTVFQVPSTKSLFKAFCTC	
wт мт 241	IMPLLKLSCSDVHFNVKMMYLGVGVFSLPLLCIIVSYVQVFSTVFQVPSTKSLFKAFCTC	300
wт мт 241 wт	IMPLLKLSCSDVHFNVKMMYLGVGVFSLPLLCIIVSYVQVFSTVFQVPSTKSLFKAFCTC	
wт мт 241 wт	<pre>IMPLLKLSCSDVHFNVKMMYLGVGVFSLPLLCIIVSYVQVFSTVFQVPSTKSLFKAFCTC </pre>	
wт мт 241 wт мт	IMPLLKLSCSDVHFNVKMMYLGVGVFSLPLLCIIVSYVQVFSTVFQVPSTKSLFKAFCTC	300