10.1071/CH19461_AC

Australian Journal of Chemistry 2020, 73(2&3), 202-211

Supplementary Material

Phote-HrTH (*Phormia terraenovae* hypertrehalosaemic hormone), the Metabolic Hormone of the Fruit Fly: Solution Structure and Receptor Binding Model

Ibrahim A. Abdulganiyyu,[A] *Marc-Antoine Sani,*[C] *Frances Separovic,*[C] *Heather Marco,*[B] *and Graham E. Jackson*[A,D]

[A]Department of Chemistry, University of Cape Town, Private Bag, Rondebosch, Cape Town, 7701, South Africa.

[B]Biological Sciences, University of Cape Town, Private Bag, Rondebosch, Cape Town, 7701, South Africa.

[C]School of Chemistry, Bio21 Institute, University of Melbourne, Melbourne, Vic. 3010, Australia.

[D]Corresponding author. Email: graham.jackson@uct.ac.za

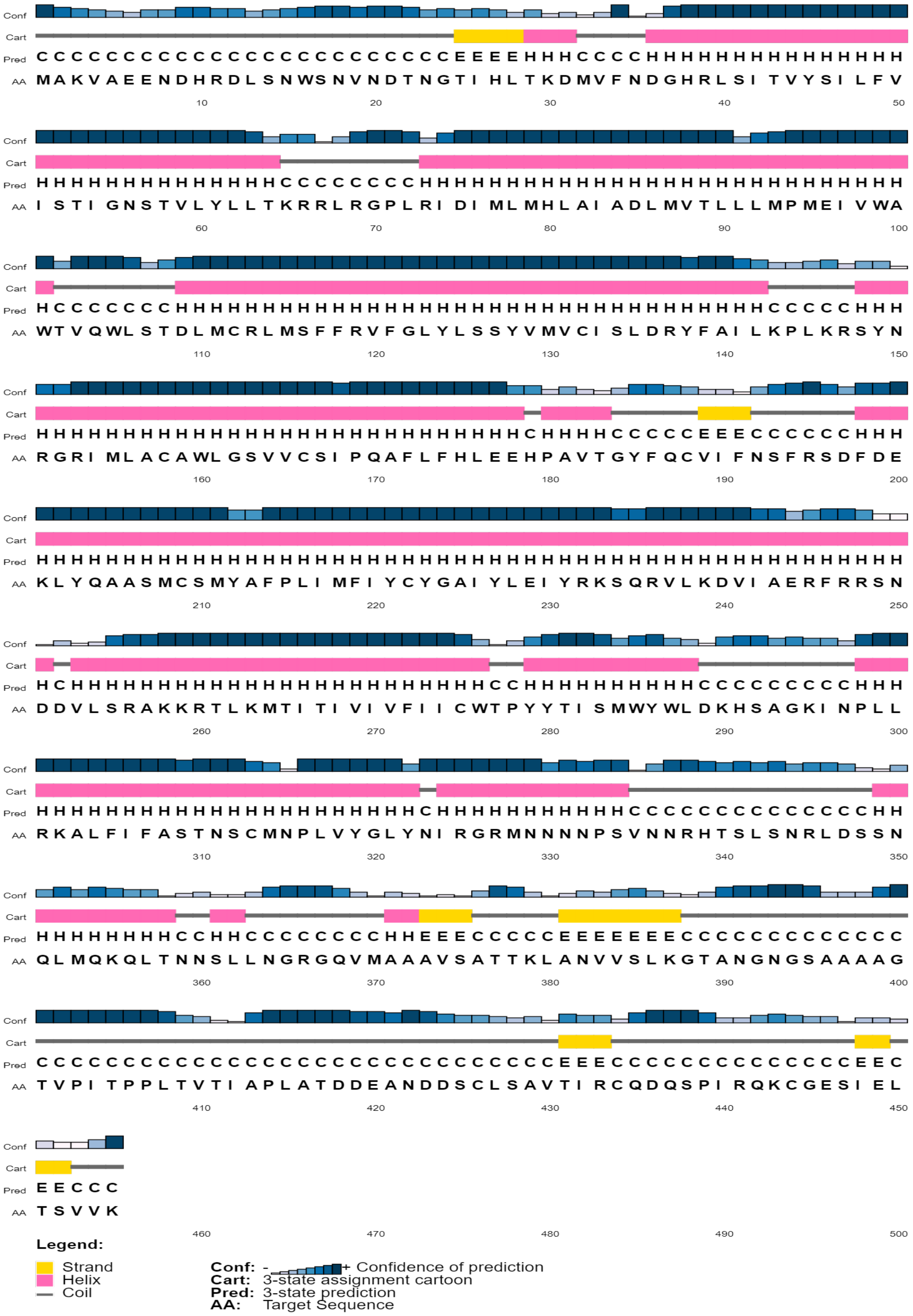


Figure S1. PSIPRED predicted secondary structure of drome-AKHR. H= helix (in yellow), S = strand (in pink) and C =coil (ash)

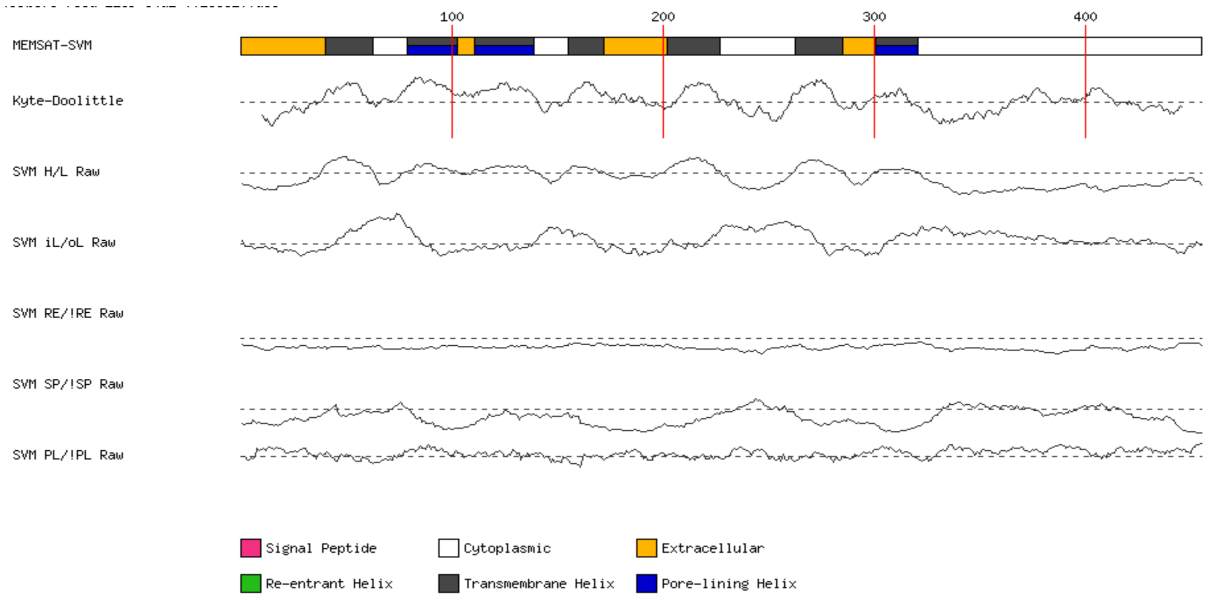


Figure S2. Schematic diagram of the MEMSAT-SVM predictions for the query sequence of Drome-AKHR. Traces indicate the RAW outputs for the prediction SVMs. Dashed lines indicate the prediction threshold. PL: Pore lining residue SP: Signal peptide residue RE: Re-entrant helix residue iL/oL & H/L: Helix prediction.

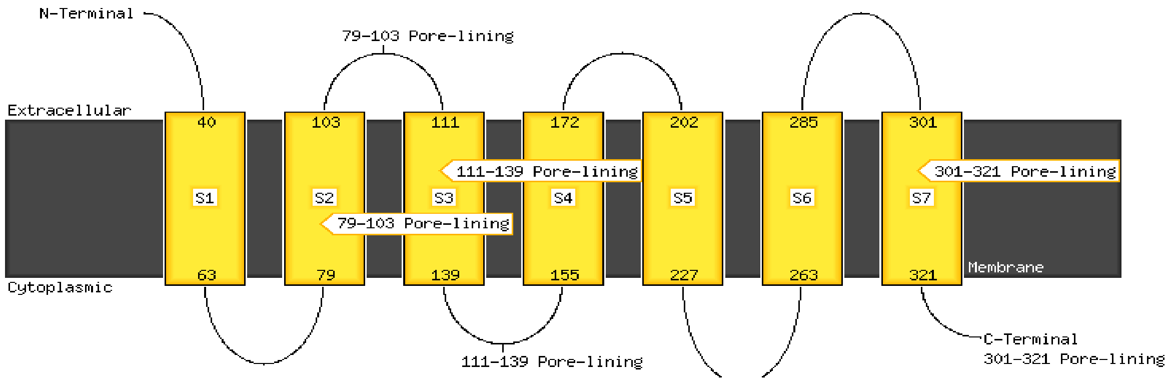


Figure S3. Prediction of transmembrane helices and the topology analysis of Drome-AKHR. The helices are represented in yellow and labelled S1-S7, the membrane (black) and the loops (thin black line) starting from the N-terminus (the extracellular region) and terminating at the C-terminus (the intracellular region)

```
Chain              1   5    10   15   20   25   30   35   40   45   50   55   60   65   70   75   80   85
Consensus           RLSITVYSILFVISTIGNSTVLYLLTKRRLRGPLRIDIMLMHLAIADLMVTLLLMPMEIVWAWTVQWLSTDLMCRLMSFFRVFG
. 1: 2rodofrf.pdb.A HRLSITVYSILFVISTIGNSTVLYLLTKRRLRGPLRIDIMLMHLAIADLMVTLLLMPMEIVWAWTVQWLSTDLMCRLMSFFRVFG
.. 2: 2beta2frf.pdb.A RLSITVYSILFVISTIGNSTVLYLLTKRRLRGPLRIDIMLMHLAIADLMVTLLLMPMEIVWAWTVQWLSTDLMCRLMSFFRVFG

Chain              86  90   95   100  105  110  115  120  125  130  135  140  145  150  155  160  165  170
Consensus          LYLSSYVMVCISLDRYFAILKPLKRSYNRGRIMLACAWLGSVVCSIPQAFLFHLEEHPAVTGYFQCVIFNSFRSDFDEKLYQAAS
. 1: 2rodofrf.pdb.A LYLSSYVMVCISLDRYFAILKPLKRSYNRGRIMLACAWLGSVVCSIPQAFLFHLEEHPAVTGYFQCVIFNSFRSDFDEKLYQAAS
.. 2: 2beta2frf.pdb.A LYLSSYVMVCISLDRYFAILKPLKRSYNRGRIMLACAWLGSVVCSIPQAFLFHLEEHPAVTGYFQCVIFNSFRSDFDEKLYQAAS

Chain              171 175  180  185  190  195  200  205  210  215  220  225  230  235  240  245  250  255
Consensus          MCSMYAFPLIMFIYCYGAIYLEIYRKSQRVLKDVIAERFRRSNDDVLSRAKKRTLKMTITIVIVFIICWTPYYTISMWYWLDKHS
. 1: 2rodofrf.pdb.A MCSMYAFPLIMFIYCYGAIYLEIYRKSQRVLKDVIAERFRRSNDDVLSRAKKRTLKMTITIVIVFIICWTPYYTISMWYWLDKHS
.. 2: 2beta2frf.pdb.A MCSMYAFPLIMFIYCYGAIYLEIYRKSQRVLKDVIAERFRRSNDDVLSRAKKRTLKMTITIVIVFIICWTPYYTISMWYWLDKHS

Chain              256 260  265  270  275  280  285  290 293
Consensus          AGKINPLLRKALFIFASTNSCMNPLVYGLYNIRGRMNn
. 1: 2rodofrf.pdb.A AGKINPLLRKALFIFASTNSCMNPLVYGLYNIRGRMNN
.. 2: 2beta2frf.pdb.A AGKINPLLRKALFIFASTNSCMNPLVYGLYNIRGRMN
```

Figure S4. Schematic diagram representing an alignment of both models constructed from both Rhodopsin and beta2 androgenic receptor templates of Drome-AKHR, the *Drosophila melanogaster* adipokinetic hormone receptor. Red indicates residues in the seven transmembrane helices, blue shows the extra-cellular regions, while the intra-cellular regions are represented in purple, yellow denotes cystine ionic lock, and green shows the highly conserved residues.

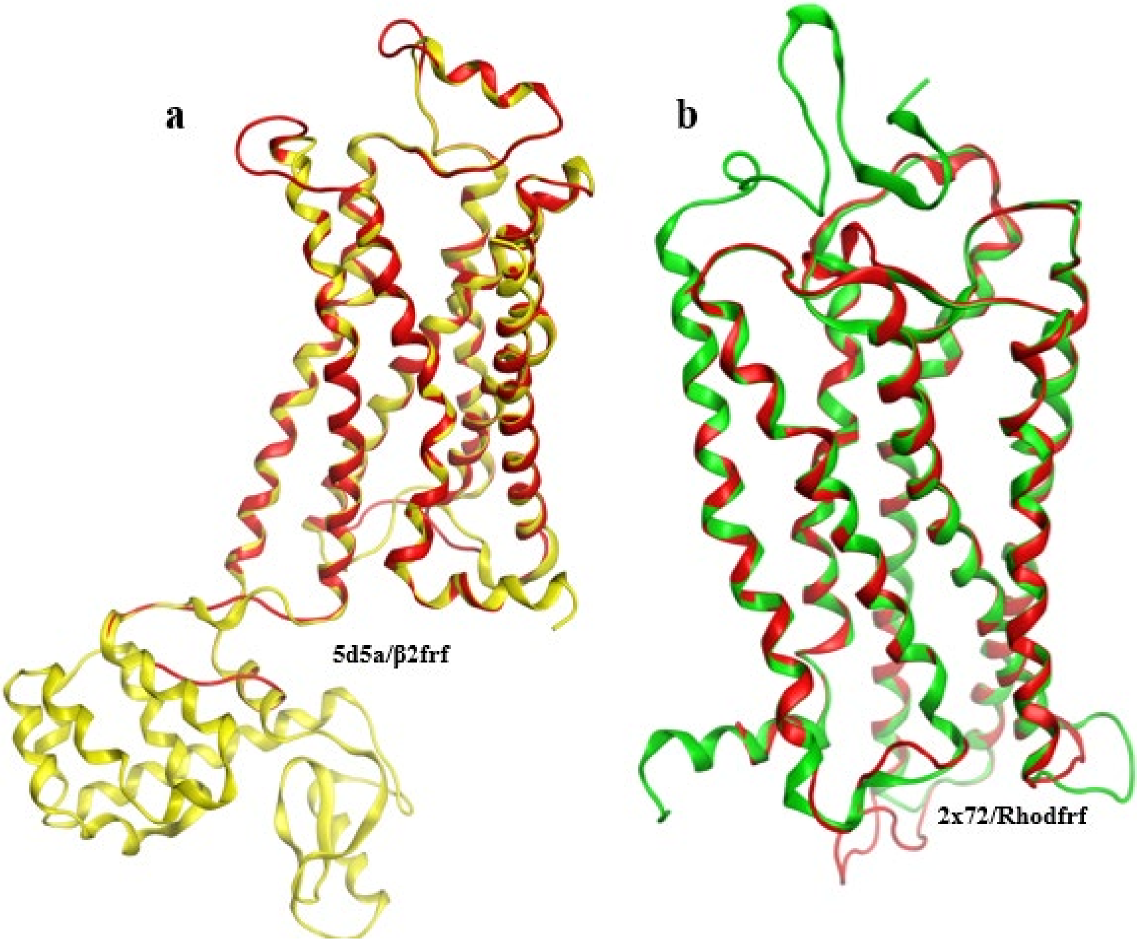
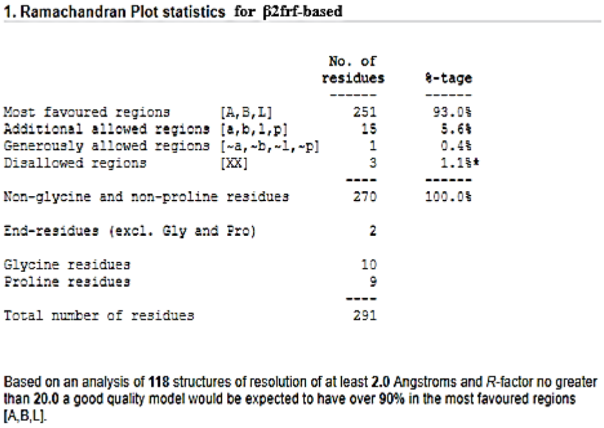


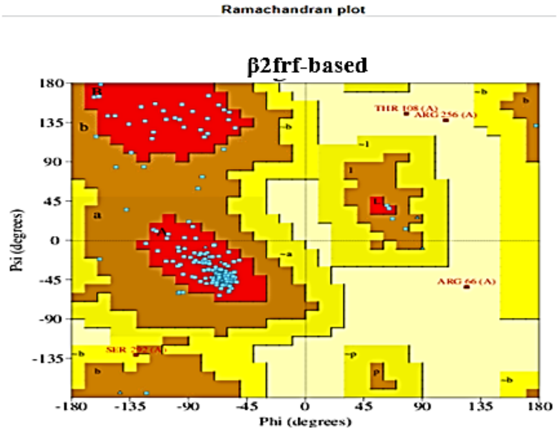
Figure S5. a) An overlay of β2frf-based-AKHR (the constructed model, in red) and 5d5a (the template used for constructing the model, in yellow). b) shows an overlay of Rhodfrf-based-AKHR (the constructed model, in red) and 2x72 (the template used for constructing the model, in green).



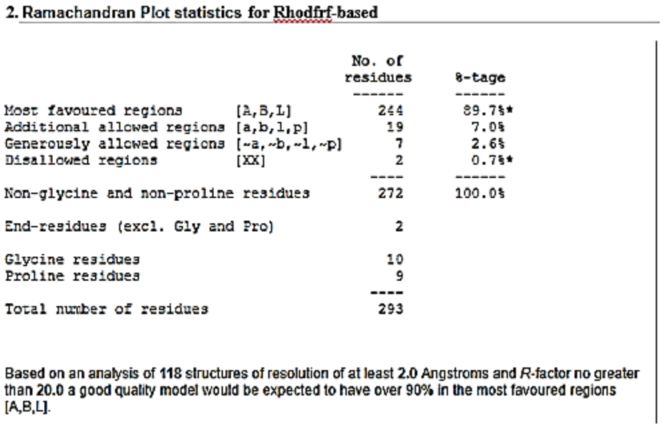
1. Ramachandran Plot statistics for β2frf-based

```
                                            No. of
                                           residues     %-tage
                                            ------      ------
Most favoured regions      [A,B,L]            251        93.0%
Additional allowed regions [a,b,l,p]           15         5.6%
Generously allowed regions [~a,~b,~l,~p]        1         0.4%
Disallowed regions         [XX]                 3         1.1%*
                                             ----       ------
Non-glycine and non-proline residues          270       100.0%

End-residues (excl. Gly and Pro)                2

Glycine residues                               10
Proline residues                                9
                                             ----
Total number of residues                      291
```

Based on an analysis of **118** structures of resolution of at least **2.0** Angstroms and *R*-factor no greater than **20.0** a good quality model would be expected to have over **90%** in the most favoured regions [A,B,L].



b)



2. Ramachandran Plot statistics for Rhodfrf-based

```
                                               No. of
                                              residues     %-tage
                                               ------      ------
Most favoured regions         [A,B,L]            244        89.7%*
Additional allowed regions    [a,b,l,p]           19         7.0%
Generously allowed regions    [~a,~b,~l,~p]        7         2.6%
Disallowed regions            [XX]                 2         0.7%*
                                                ----       ------
Non-glycine and non-proline residues             272       100.0%

End-residues (excl. Gly and Pro)                   2

Glycine residues                                  10
Proline residues                                   9
                                                ----
Total number of residues                         293
```

Based on an analysis of **118** structures of resolution of at least **2.0** Angstroms and *R*-factor no greater than **20.0** a good quality model would be expected to have over **90%** in the most favoured regions [A,B,L].

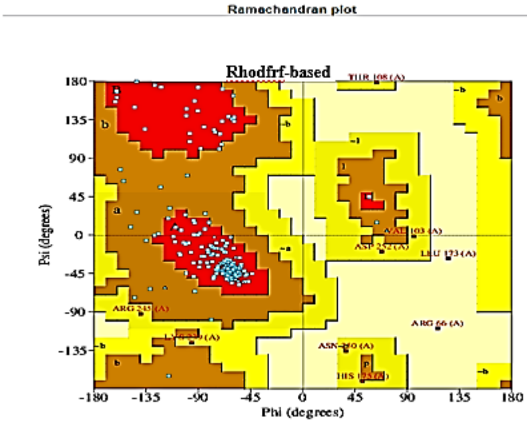


Figure S6. Ramachandran plots of a) β2frf-based and b) Rhodfrf-based models

Table S1. Comparison of helical sequences of Drome-AKHR with β2AR and rhodopsin.

Helix	% sequence identity with		% sequence similarity with	
	β2AR (β2frf)	rhodopsin (rhodfrf)	β2AR(β2frf)	rhodopsin (rhodfrf)
1	16	20	60	61
2	25	16	58	52
3	29	30	68	60
4	28	40	76	80
5	23	17	65	60
6	34	23	83	74
7	43	45	88	88