GPCRdb Homology Models - "Less Model & More Crystal"

GPCRdb.org

G. Pándy-Szekeres[1,2], S. Mordalski[2], A. Bojarski[2], D.E. Gloriam[1]

[1]Department of Drug Design and Pharmacology, University of Copenhagen, Denmark;

[2]Department of Medicinal Chemistry, Polish Academy of Sciences, Krakow, Poland

UNIVERSITY OF COPENHAGEN

SUMMARY

- **Knowledge-based approach**
- **Multiple backbone templates**
- **Multiple rotamer templates**
- **Alternative loop templates**
- **Modeling structural distortions in the helices**
- **MODELLER [1] for regions without a template**
- **Automated model building pipeline**
- **Models automatically updated with new Xtal data**

MODEL BUILDING STEPS

1. Main template selection – The frame

- Same class (A, B, C, F) as target
- Same activation state as target
- Highest overall sequence similarity
- If multiple choices, one with best resolution
- GPCRdb [2] generic numbers [3] are used for the sequence similarity alignment
- Only 7TM and H8 – **Initial model**

2. Helix start and end adjustments

- Distortions in the crystal packing or fused proteins can
- cause too short or too long TM1-7 and H8 starts and ends.
- GPCRdb has manually annotated segment ends for all structure templates
- When helix start or end is too long → residues are removed from the initial model
- When too short → an alternative template is used to model the missing residues
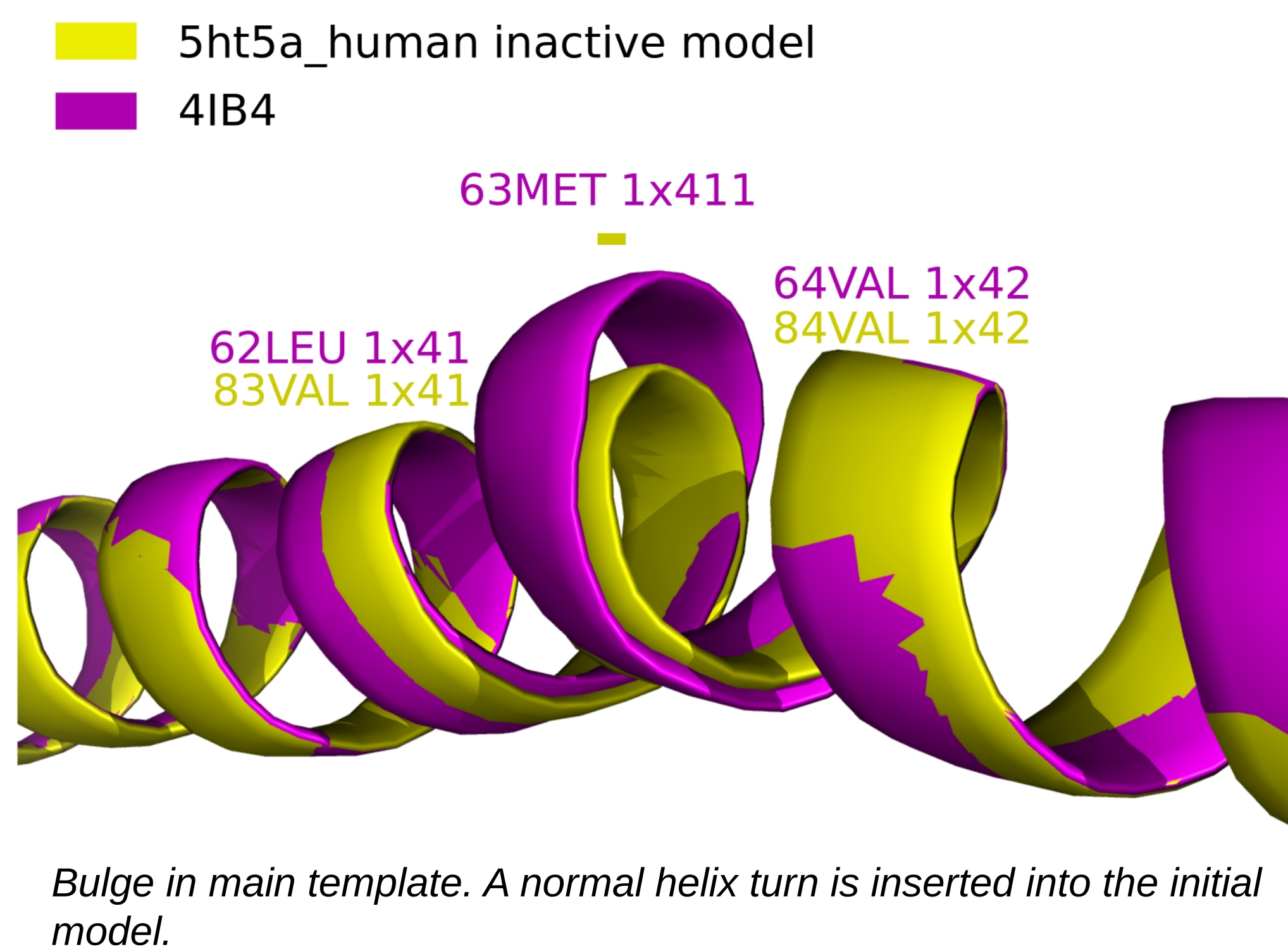
3. Loop backbone modeling

- Trying loops from main template first, if not suitable, an alternative template is found
- Selection: 1. length, 2. overall sequence similarity and 3. resolution
- ECL2 – **3-part modeling** – ECL2_1, ECL2_mid (CYS 45x50, X 45x51, X 45x52), ECL2_2
- Long ICL3 cut to 10-10 residues on each end

Validation, comparison with SWISS-MODEL

- RMSD calculations comparing to target structure
- 3 recently determined receptor structures
- High sequence similarities
- Good GPCRdb scores in 7TM region
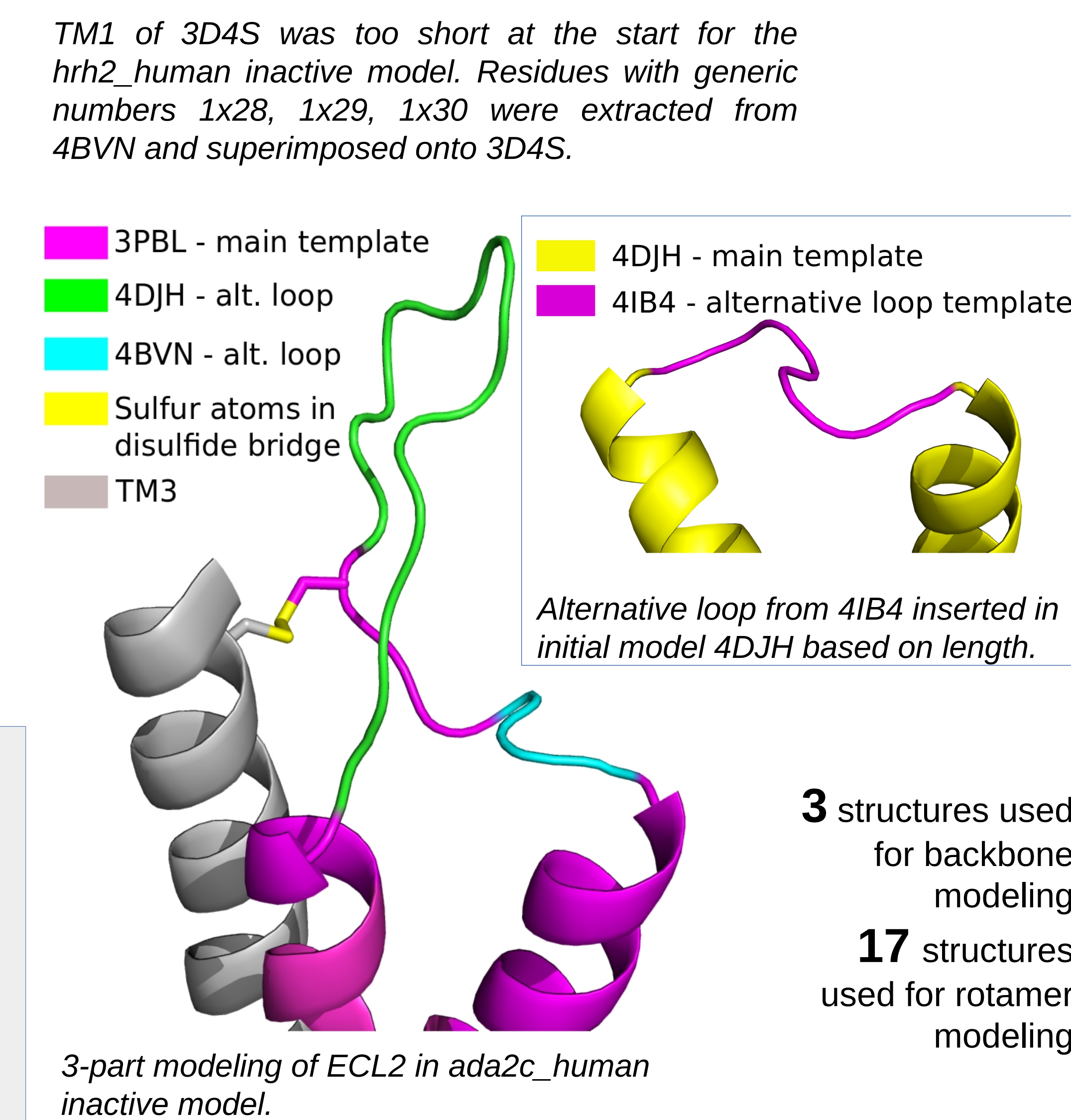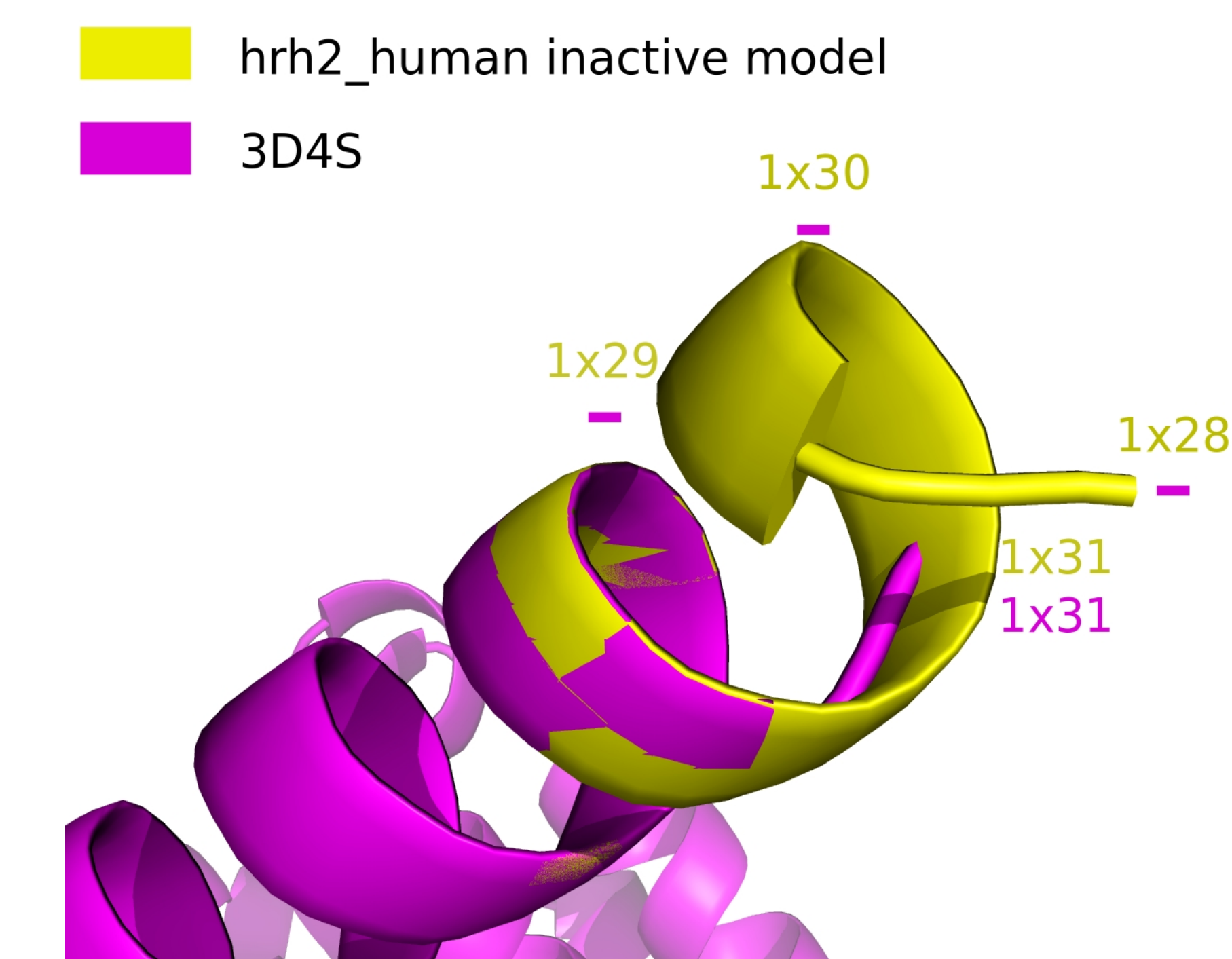- Loop and binding pocket modeling needs improvement

5ht5a_human inactive model
4IB4

63MET 1x411
64VAL 1x42
84VAL 1x42
62LEU 1x41
83VAL 1x41

Bulge in main template. A normal helix turn is inserted into the initial model.

5ht1b_human
5ht2b_human
ccr5_human
cxcr4_human
ntr1_rat
oprd_mouse
oprd_human

Alignment for receptors

hrh2_human inactive model
3D4S

1x30
1x29
1x28
1x31
1x31

TM1 of 3D4S was too short at the start for the hrh2_human inactive model. Residues with generic numbers 1x28, 1x29, 1x30 were extracted from 4BVN and superimposed onto 3D4S.

3PBL - main template
4DJH - alt. loop
4BVN - alt. loop
Sulfur atoms in disulfide bridge
TM3

4DJH - main template
4IB4 - alternative loop template

Alternative loop from 4IB4 inserted in initial model 4DJH based on length.

3-part modeling of ECL2 in ada2c_human inactive model.

3 structures used for backbone modeling

17 structures used for rotamer modeling

4. Modeling local structural distortions in the helix – bulges and constrictions

- GPCRdb generic numbers can show local structural distortions in the helix
- Bulge – **5 residues** instead of normal 4 in a helix turn
- Constriction – **3 residues** instead of normal 4 in a helix turn
- A bulged or constricted turn can be switched to a normal turn
- Alternative templates are found based on generic numbers

5. Modeling the backbone of N- and C-termini

- N-term – last five residues modeled with TM1 template
- C-term – first five residues modeled with H8 (if not present, TM7) template
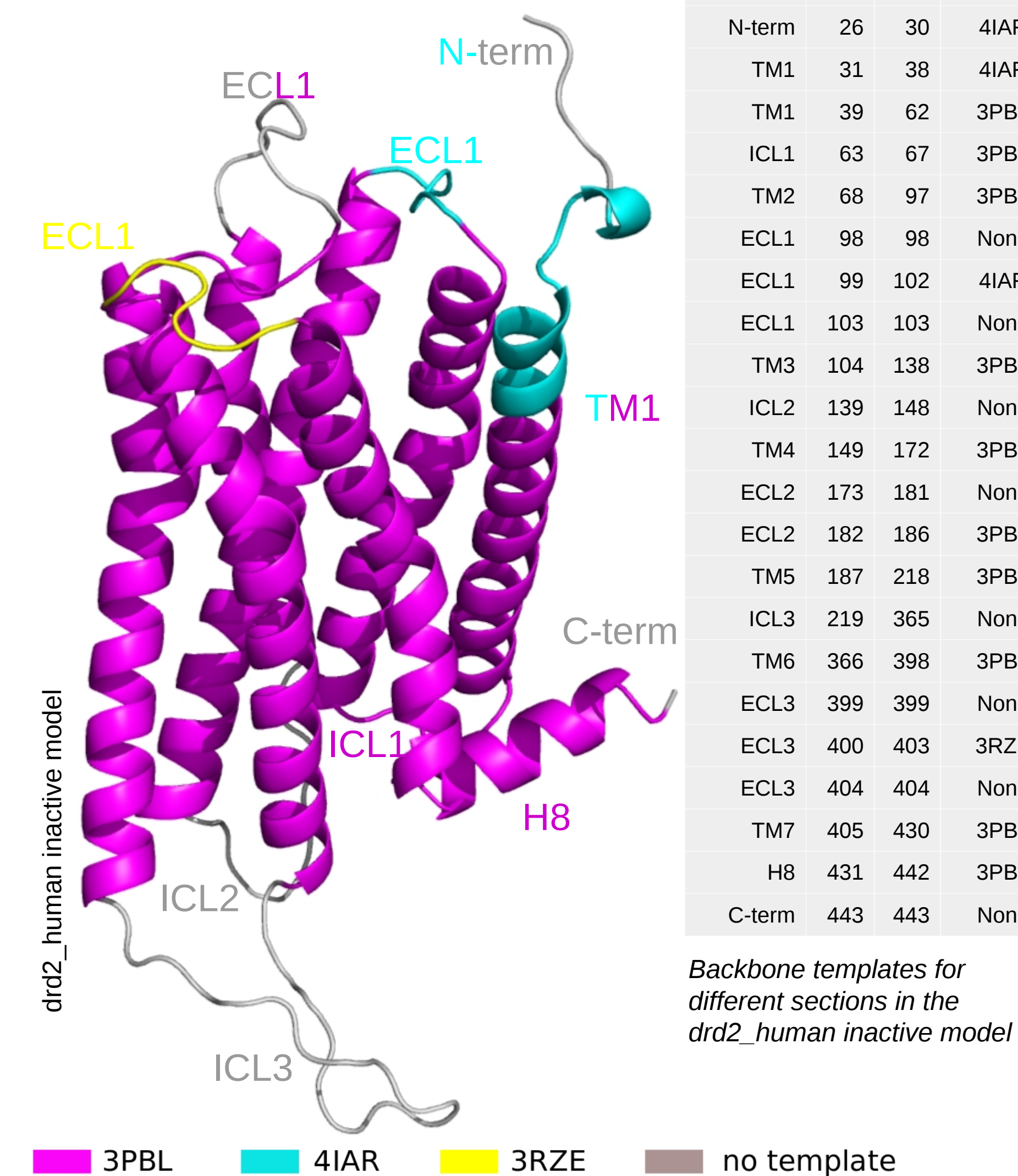- Up to ten residues are modeled freely with MODELLER

6. Modeling the rotamers – in-house rotamer library

- When residue mismatch between target and main template → alternative rotamer template based on GPCRdb generic number
- In-house rotamer library
- Selection: 1. overall sequence similarity, 2. same activation state, 3. highest resolution

7. Remaining parts - MODELLER

- When no suitable template – free modeling
- Residues are mutated to ALA

THE FINAL MODELS

N-term
ECL1
ECL1
ECL1
TM1
C-term
ICL1
H8
ICL2
ICL3

drd2_human inactive model

3PBL *4IAR* *3RZE* *no template*

Segment	Start num	End num	Backbone Template
N-term	1	25	None
N-term	26	30	4IAR
TM1	31	38	4IAR
TM1	39	62	3PBL
ICL1	63	67	3PBL
TM2	68	97	3PBL
ECL1	98	98	None
ECL1	99	102	4IAR
ECL1	103	103	None
TM3	104	138	3PBL
ICL2	139	148	None
TM4	149	172	3PBL
ECL2	173	181	None
ECL2	182	186	3PBL
TM5	187	218	3PBL
ICL3	219	365	None
TM6	366	398	3PBL
ECL3	399	399	None
ECL3	400	403	3RZE
ECL3	404	404	None
TM7	405	430	3PBL
H8	431	442	3PBL
C-term	443	443	None

Backbone templates for different sections in the drd2_human inactive model

Receptor (PDB)	Model*	Main template	Overall all (Å)	Overall backbone (Å)	7TM all (Å)	7TM backbone (Å)	Binding pocket (Å)
ox1r_human (4ZJC)	**GPCRdb**	4S0V	3.4	2.8	1.3	0.5	2.2
	SwissModel_repo	-	-	-	-	-	-
	SwissModel	4S0V	2.7	2.0	1.5	0.5	0.6
acm1_human (5CXV)	**GPCRdb**	4U15	1.8	1.3	1.6	1.1	0.7
	SwissModel_repo	3UON	2.3	1.7	2.1	1.5	0.9
	SwissModel	4U14	2.6	2.1	2.3	1.7	0.8
acm4_human (5DSG)	**GPCRdb**	3UON	1.8	1.4	1.5	1.1	1.2
	SwissModel_repo	4U14	2.0	1.5	2.0	1.5	0.8
	SwissModel	3UON	1.8	1.3	1.5	1.1	1.2

*SwissModel_repo – Model form the SwissModel repository
SwissModel – Model built by SwissModel using the same (or the closest) main template as in the GPCRdb model

REFERENCES

[1] B. Webb, A. Sali.: Comparative Protein Structure Modeling Using Modeller. Current Protocols in Bioinformatics, John Wiley & Sons, Inc., 5.6.1-5.6.32, 2014.
[2] Isberg V, Mordalski S, Munk C, Rataj K, Harpsøe K, Hauser AS, Vroling B, Bojarski AJ, Vriend G, Gloriam DE. GPCRdb: an information system for G protein-coupled receptors. Nucleic Acids Res. 2016 Jan 4;44(D1):D356-64.
[3] Isberg V, et al.; Generic GPCR Residue Numbers - Aligning Topology Map Minding The Gaps; *Trends Pharmacol Sci, (2015) 36:22-31*

CONTACT Web: www.gpcrdb.org Email: info@gpcrdb.org @gpcrdb