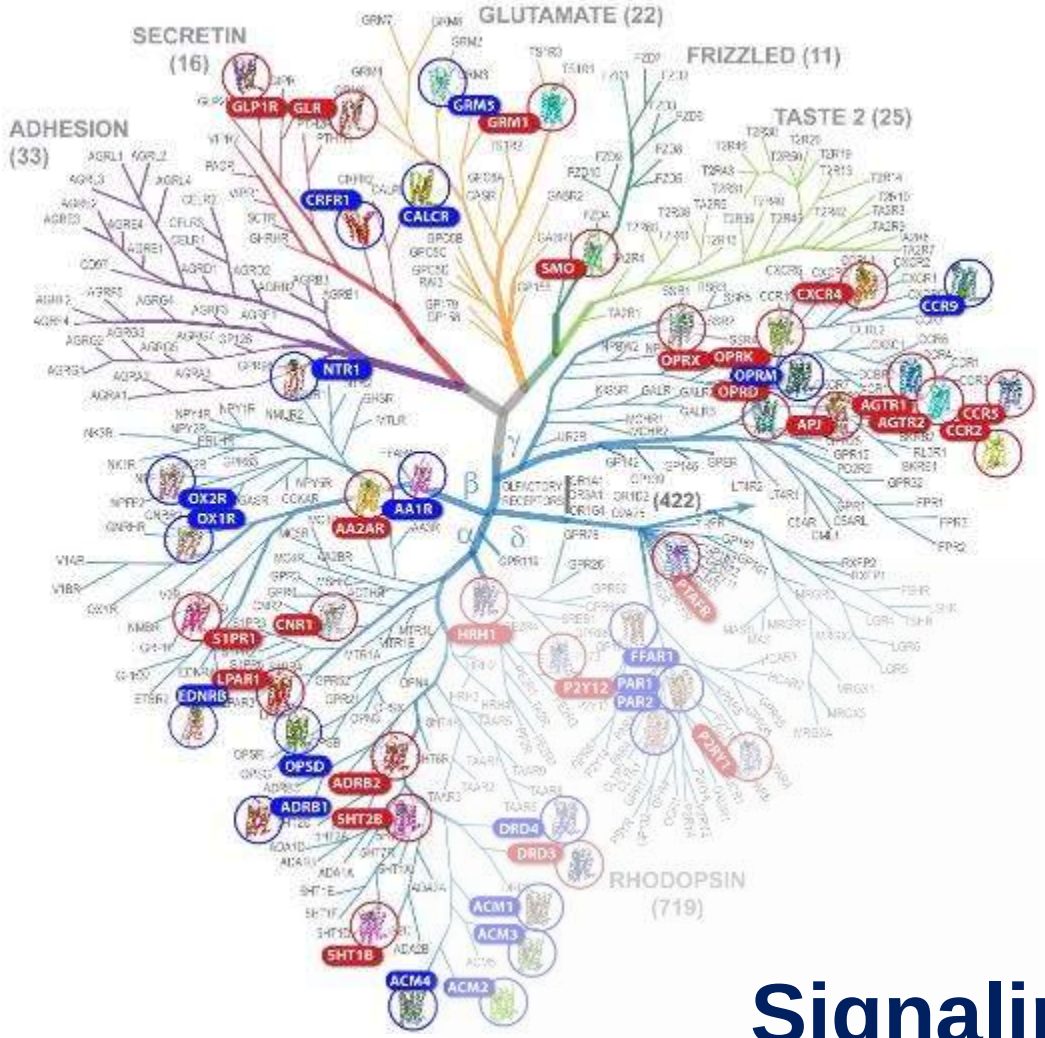
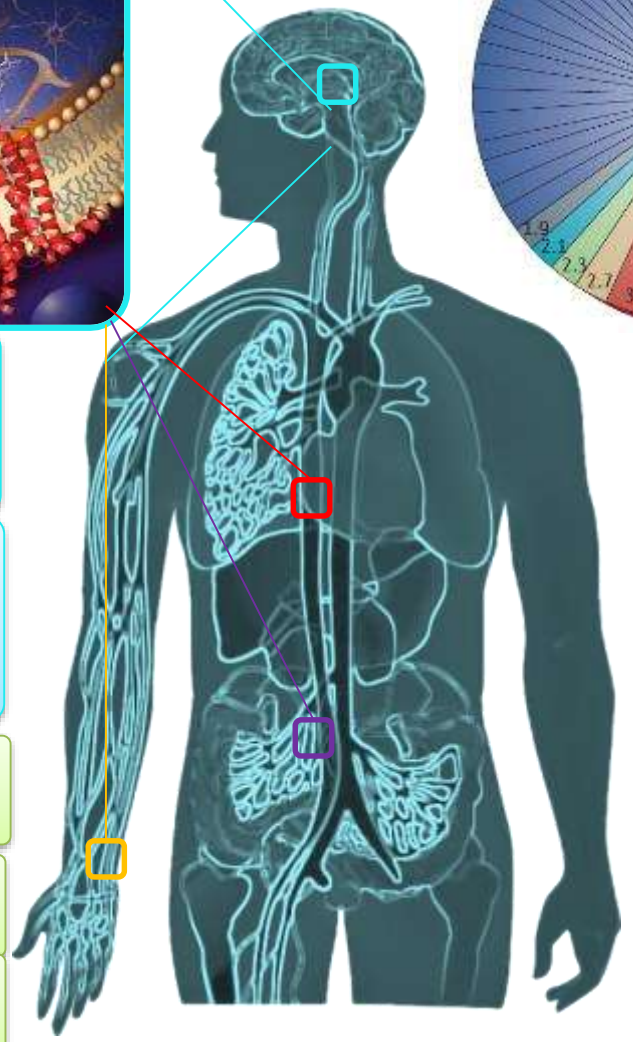


Computational Modeling of GPCRs: Towards Understanding Receptor Signaling for Biology and Drug Discovery

Vsevolod (Seva) Katritch

Barcelona, Jul 13th, 2018

GPCRs in Human Biology & Pharmacology



Nat Rev Drug Disc **5**:993 (2006)

>30% of all drugs

GPCR Targets:
>120 established
> 800 potential

800 GPCR are receptors for:

Neurotransmitters
(**adrenaline, dopamine**, **histamine, acetylcholine**, **adenosine, serotonin**, glutamate, anandamide, GABA)

Hormones & Neuropeptides
(**opioid, neurotensin**, **glucagon**, **CRF**, galanin, orexin, oxytocin, neuromedin, melanocortins, somatostatin, ghrelin, TRH, TSH, GnRH, , PTH, THS, LH… >30 total)

Immune system
(chemokine, sphingosine 1 phosphate)

Development
(**Frizzled**, Adhesion)

Sensory
Light, Taste, Olfactory (388)

Disorders Targeted Clinically

Psychiatric, Learning & Memory, Mood, Sleep, Drug Addiction, Stress, Anxiety, Pain, Social Behavior…

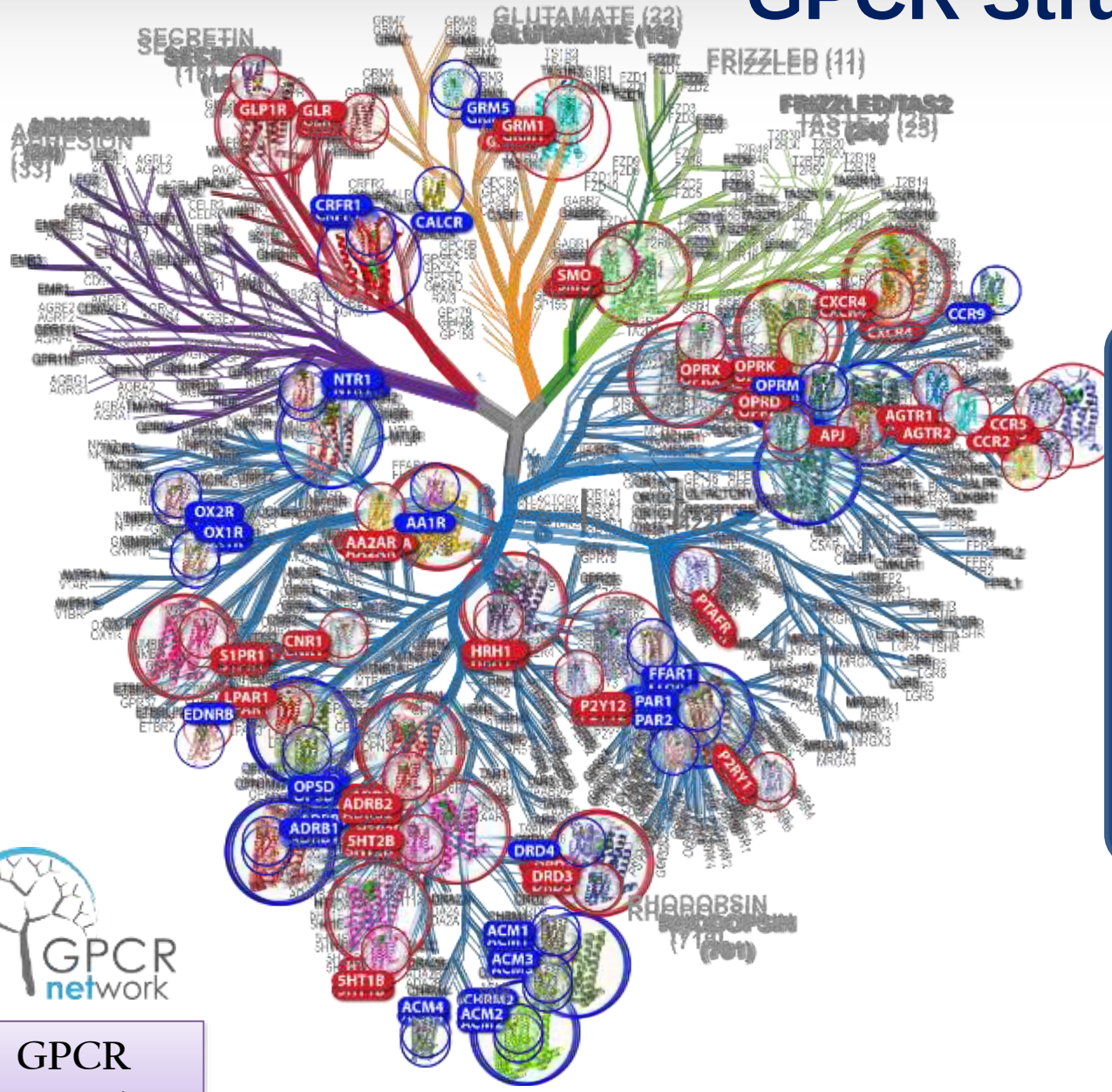
Cardiovascular, Endocrine, Obesity, Immune, HIV, Reproductive …

Neurodegenerative and Autoimmune Disease …

Brain Development, Regenerative Med.

Cancer

GPCR Structure and Function



GPCR Consortium

Thousands of Ligands - Chemical Diversity

>800 Different Human Receptors
(largest family in human genome)
Share 7TM Fold But Diverse Structural Features

Dozens of Effectors

We use structure and modeling to learn:

- Molecular Recognition
- Signaling Mechanisms

and to design:

New tool compounds
New receptor properties

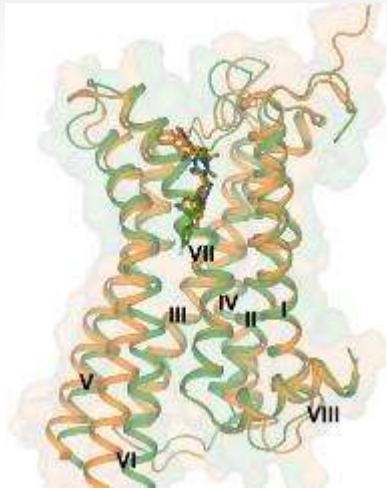
Katritch et al 2013 *Annu Rev Pharm. Tox.* ***53,*** *531-556*
Stevens et al. 2013, *Nat Rev Drug Discov*, **12**: 25–34

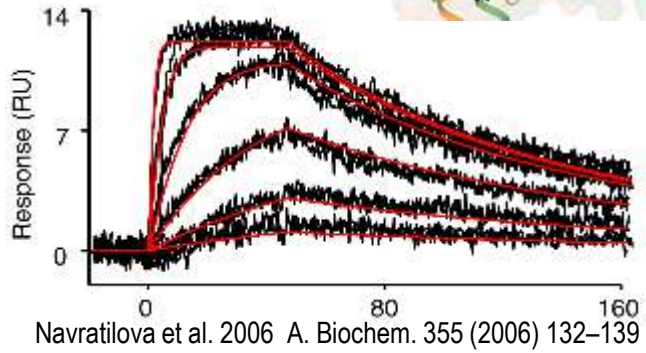
Outline

- Rational prediction of stabilizing mutations: CompoMug
- New insights into GPCR function and allosteric mechanisms
- Structure-Based ligand discovery for GPCRs

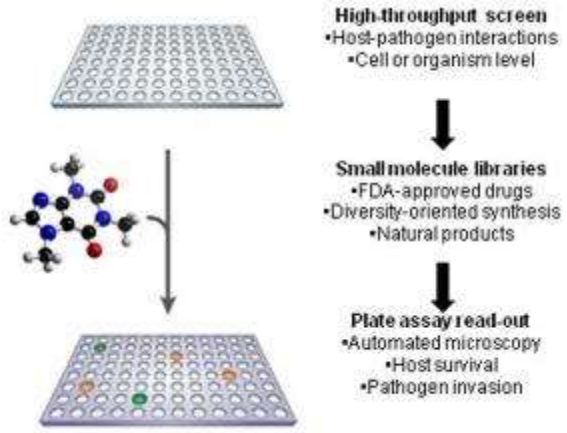
Why Stabilizing GPCRs?

- Crystallography:
  - Synergistic with fusions and truncations
  - Reduces heterogeneity
  - Allows stabilization of active or inactive states
  - Allows co-crystallization with low-affinity ligands
- Biophysical characterization
  - SPR
  - NMR
- Drug discovery
  - More robust assays for HTS
  - NMR-based ligand screening
  - Ligand soaking for large-scale SBDD

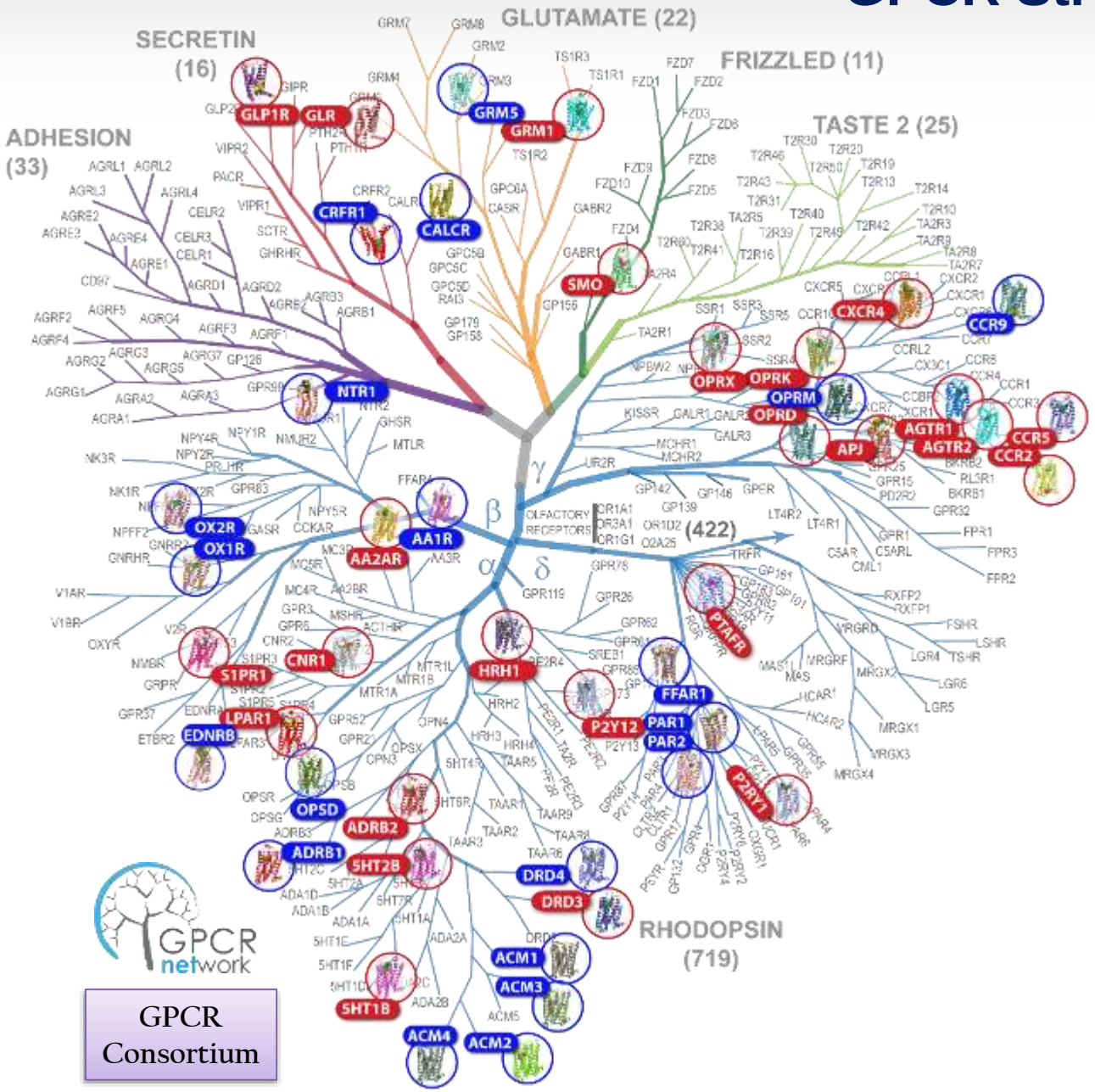




Navratilova et al. 2006 A. Biochem. 355 (2006) 132–139



GPCR Structure and Function



GPCR Consortium

- > 250 structures
- > 55 unique GPCRs

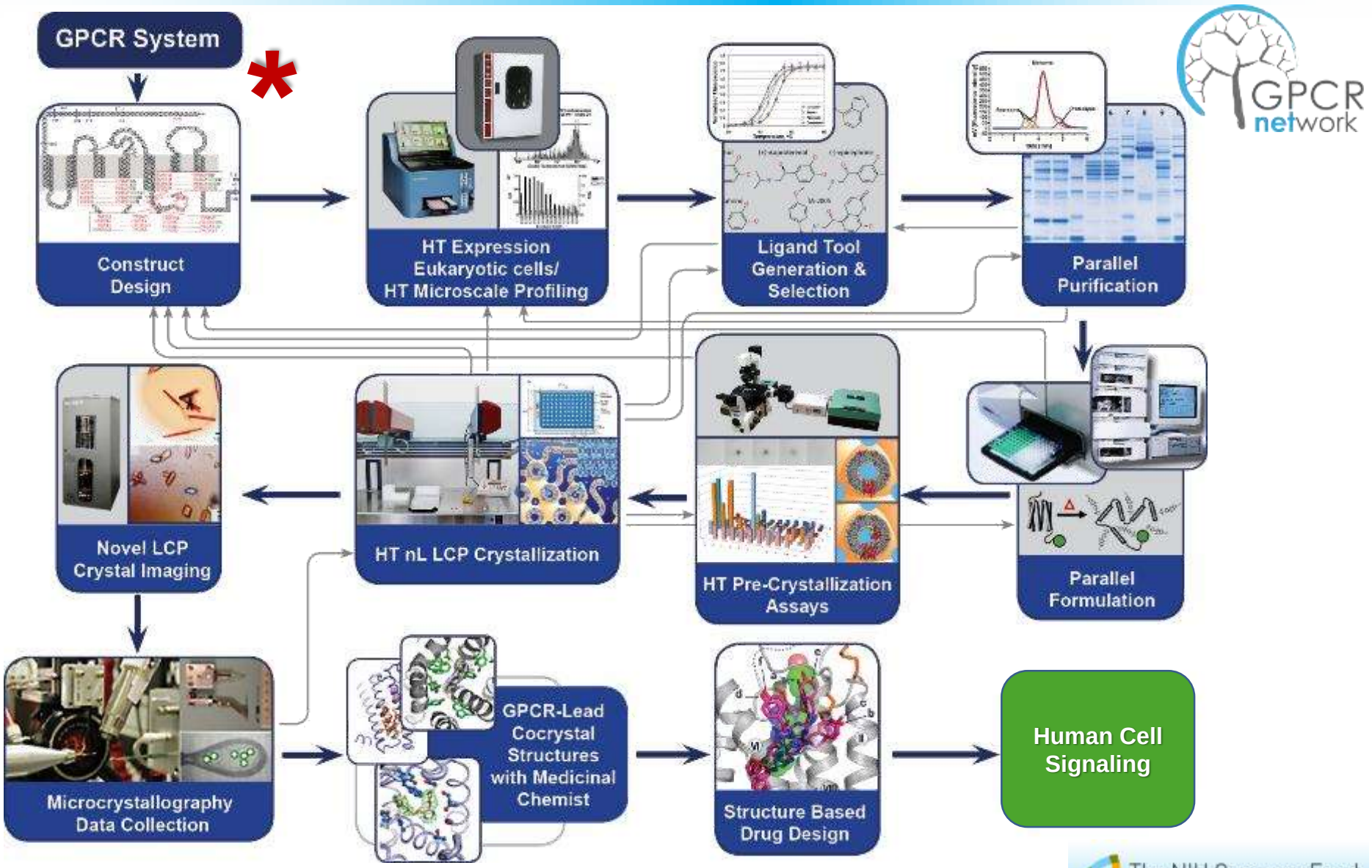
All required some protein engineering:

- 47 with fusion proteins
- 36 with mutations
- 25 with both
- All have truncations

Katritch et al 2013 *Annu Rev Pharm. Tox.* ***53,*** *531-556*
Stevens et al. 2013, *Nat Rev Drug Discov*, **12**: 25–34

GPCR Network Structure-Function Pipeline

Multistage iterative process

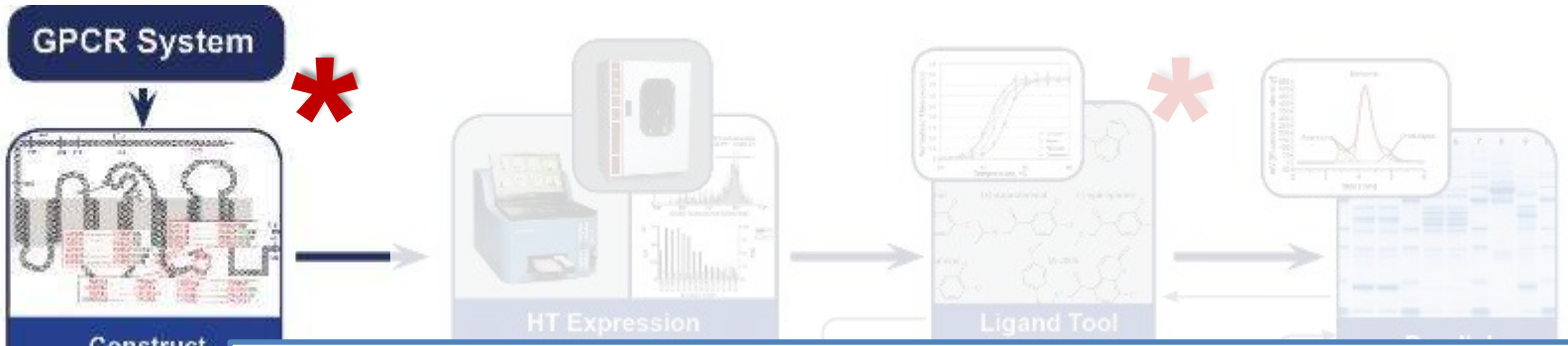


Stevens et al. 2013, *Nat Rev Drug Discov*, **12**: 25–34

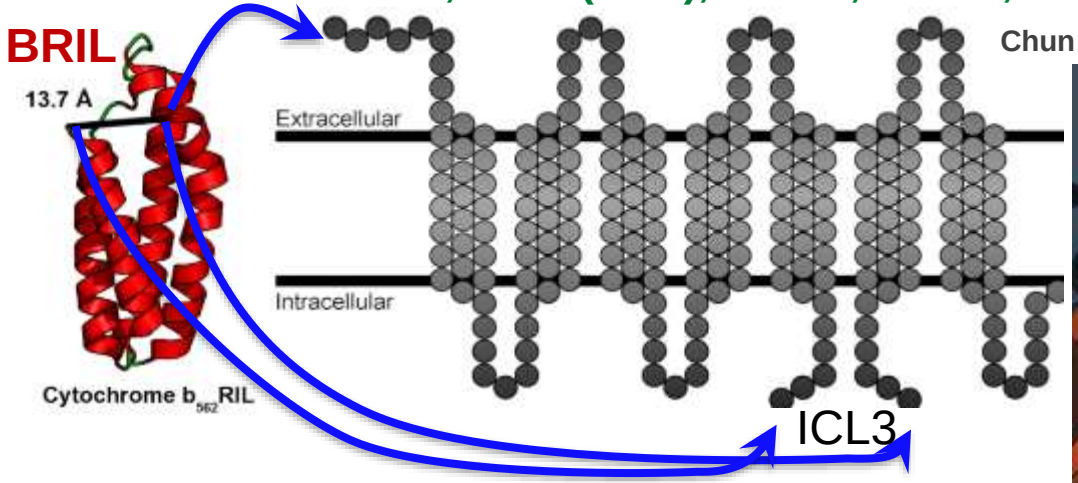
The NIH Common Fund

GPCR Network Structure-Function Pipeline

New Rationally Designed Protein Fusions

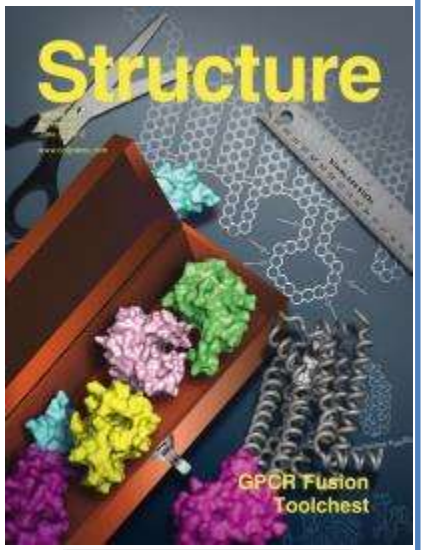


N-term: **NOP, δ-OR (1.8Å), GCGR, SMOH, $mGluR_1$, AT_1 …**



A_{2A} (1.8Å), $5HT_{1B}$, $5HT_{2B}$, $P2Y_{12}$, LPA_1, …

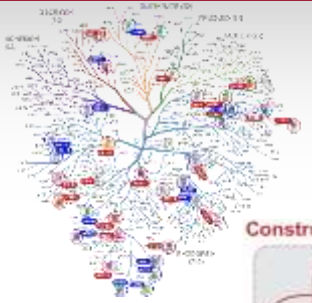
Chun et al. 2012, *Structure* 20: 967

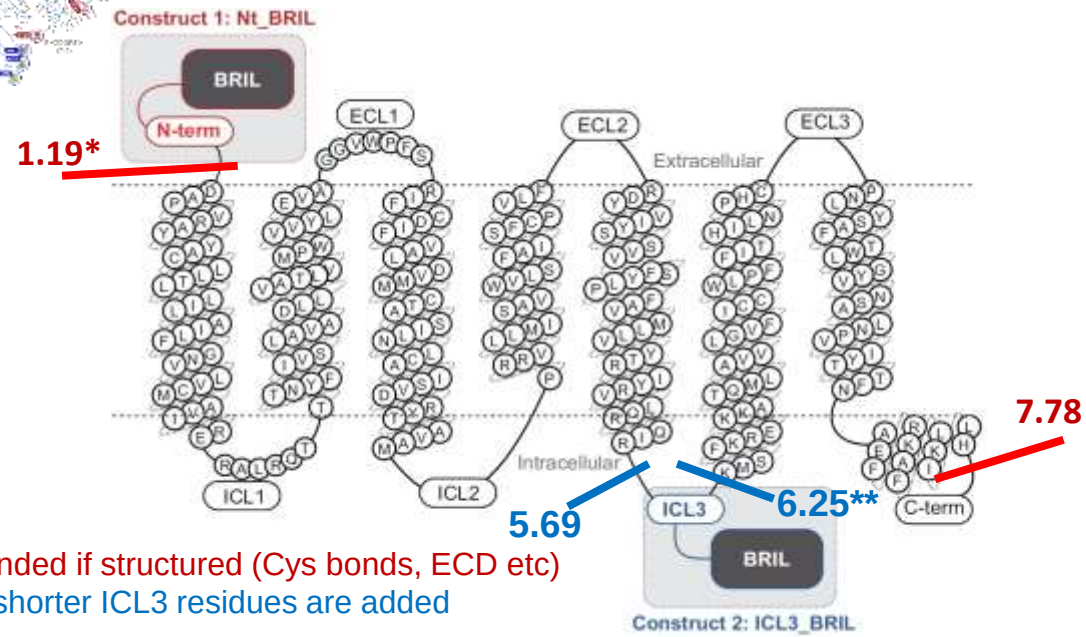


- Higher success and rational design predictability
- Improved resolutions – as high as **1.8 Å !**
- **Less invasive fusion (especially N-terminal)**

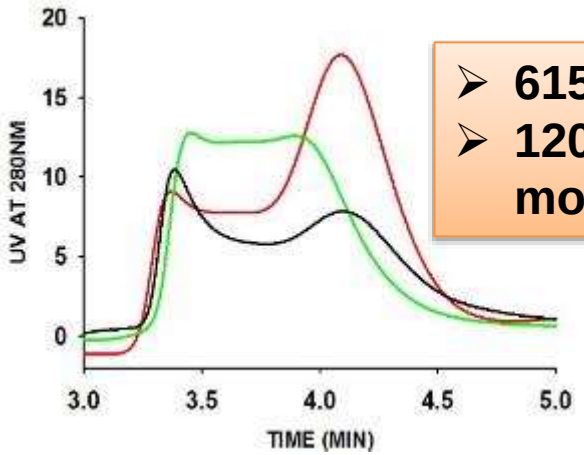
Rubredoxin
Flavodoxin

Initial Constructs: GPCR-826 project



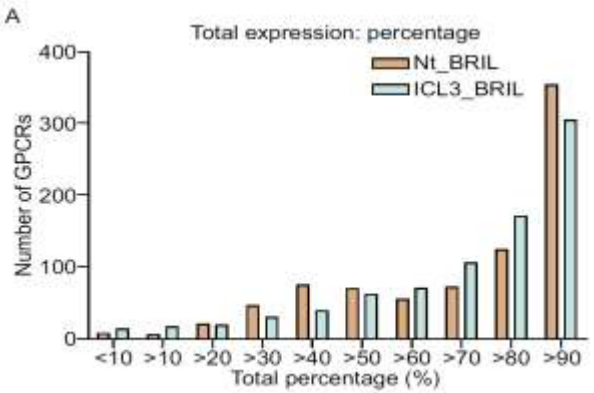


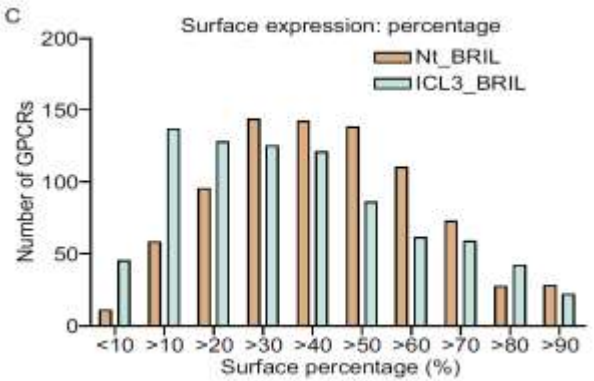
* Extended if structured (Cys bonds, ECD etc)
**For shorter ICL3 residues are added



- **615 GPCRs purified**
- **120 have strong monomer fractions**

Surface Expression >30%:
- 80% (718) of Nt_BRIL
- 60% (480) of ICL3_BRIL



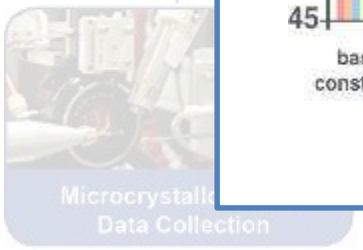


Xuechen Lv et al, Protein & cell (2016)

GPCR Network Structure-Function Process

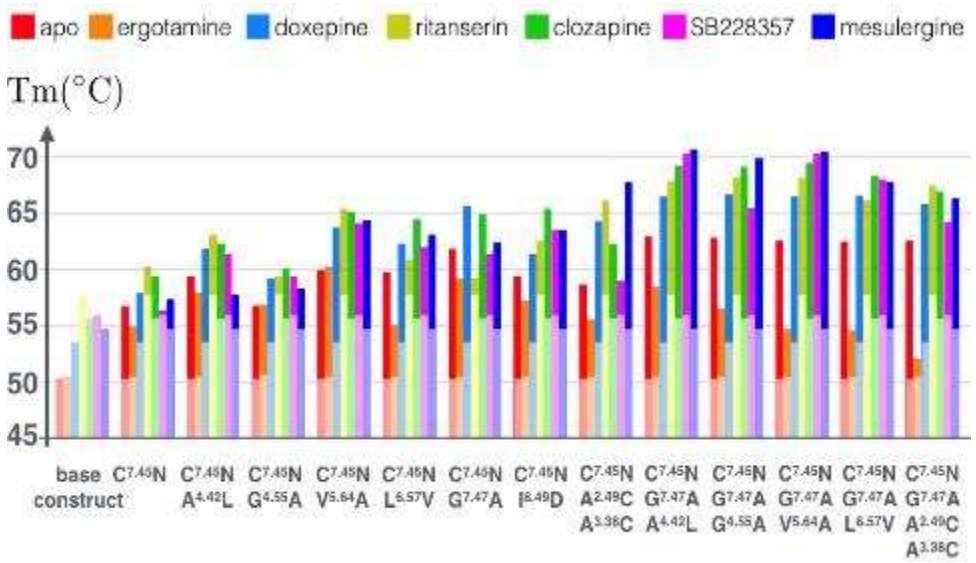
CompoMug: Rationally Designed Mutations

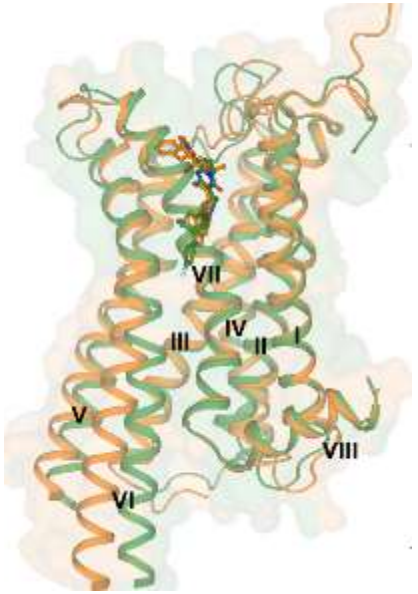
GPCR Consortium



Rationally designed mutations for GPCR crystallization

Peng et al., 2018, *Cell* 172, 719–730; **Popov et al, 2018, *in revision***





The NIH Common Fund

Stabilizing mutations: Experimental approaches

Alanine/Leucine scanning

Number of mutants: 300-2000

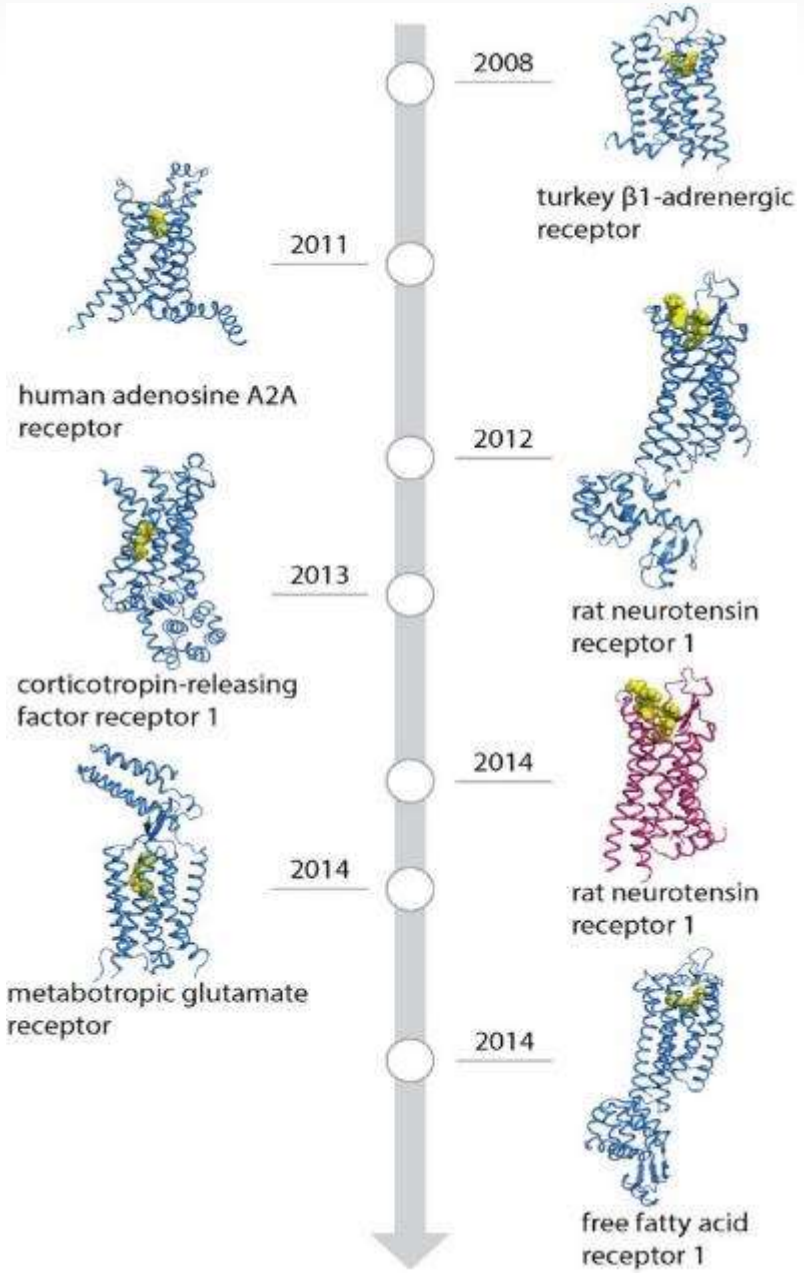
- β1AR (Serrano-Vega et al., 2008, Warne et al., 2009)
- Adenosine A2A receptor (Magnani et al., 2008, Lebon et al., 2011, Robertson et al., 2011, Dore et al., 2011)
- Neurotensin NTSR1 (Shibata et al., 2009, 2013)
- $mGLUR_5$ (Dore et al., 2014)
- CRFR1 (Hollenstein et al., 2013)
- FFAR1 (Hirozane et al., 2014)

Directed evolution

Number of mutants: > 1,000,000

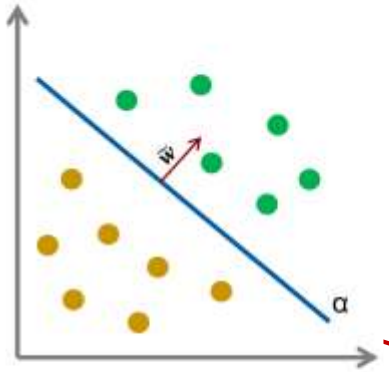
- NTSR1 (Sarkar et al., 2008, Dodevsky et al., 2011)
- α1AR, α1BR (Dodevsky, Plückthun, 2011)
- Tachykinin receptor NK1 (Dodevsky, Plückthun, 2011)

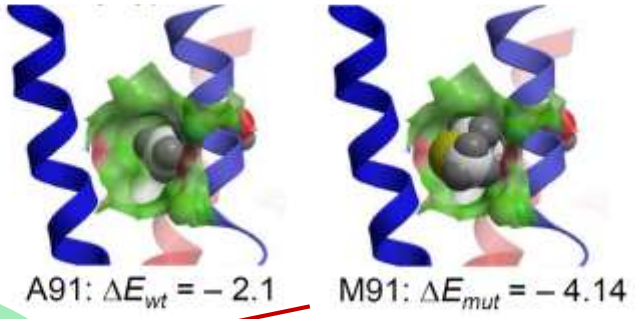
One might hope that in the future it might be possible to design thermostabilizing mutations, computationally predict them or transfer them from other receptors …

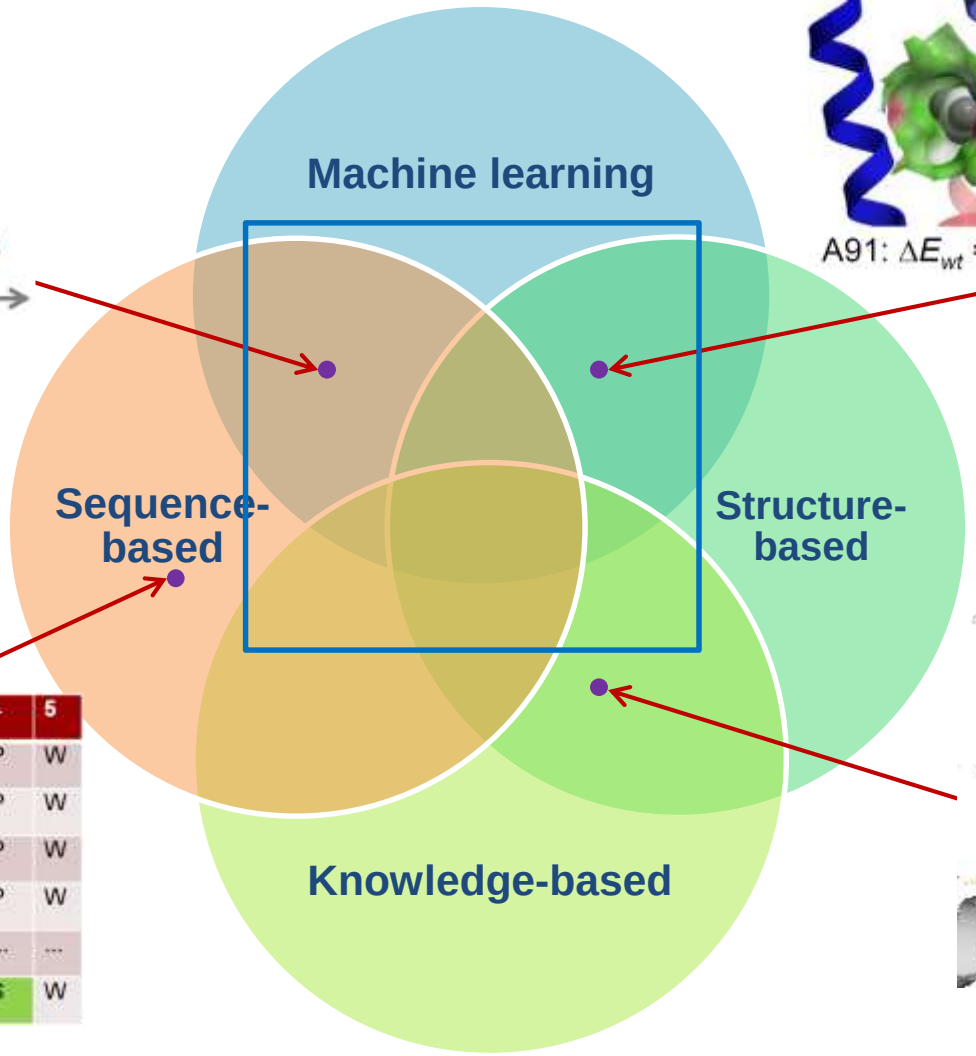


Heydenreich et al., 2015, Frontiers in Pharmacology

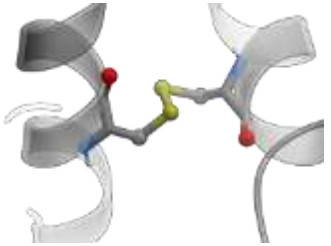
ComPOMuG: Computational Predictions of Optimizing Mutations in GPCRs

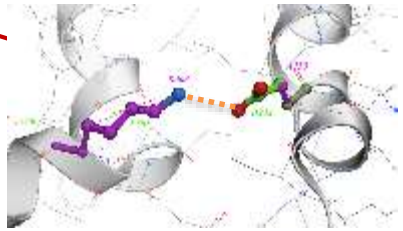






GPCR ID	1	2	3	4	5
GPCR 1	L	V	L	P	W
GPCR 2	L	L	L	P	W
GPCR 3	L	A	L	P	W
GPCR 4	L	T	L	P	W
...	...	....	...	...	...
Target GPCR	L	S	L	S	W





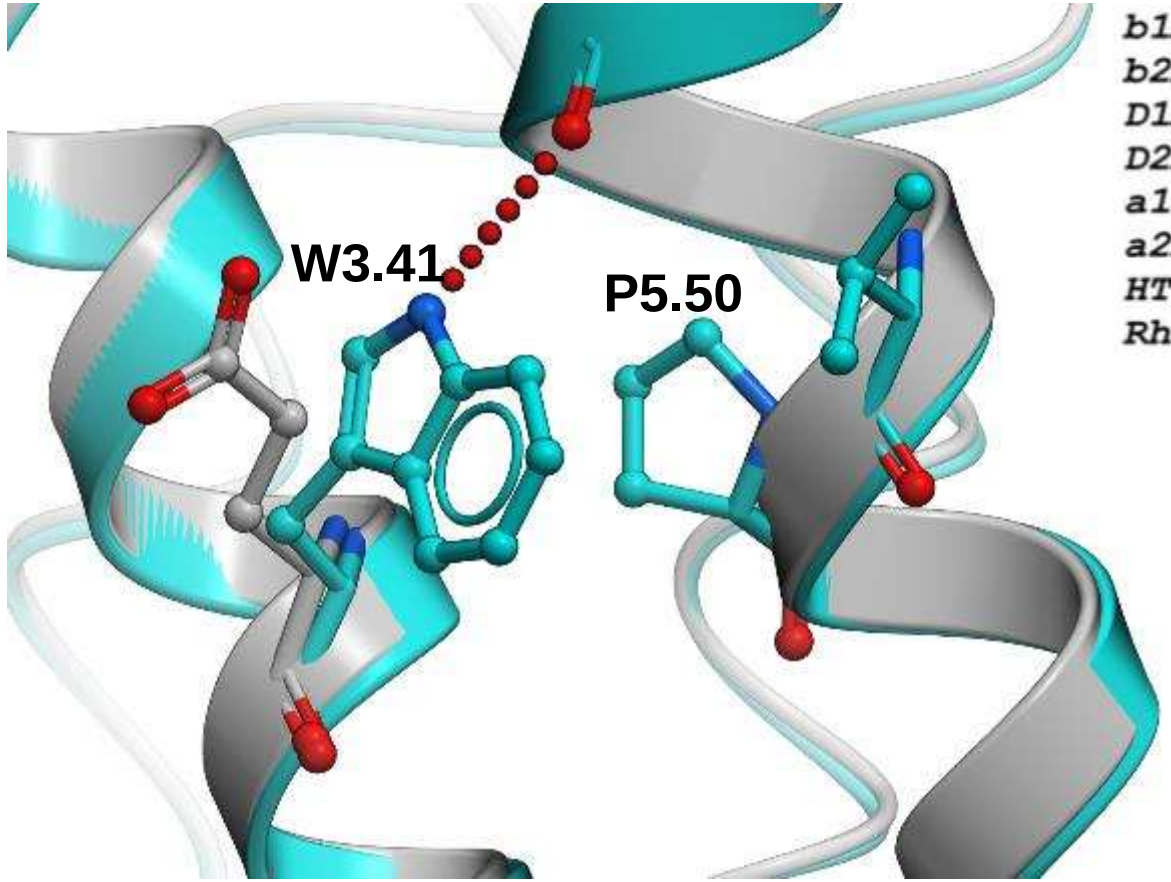
Popov et al, 2018, *eLife 10.7554/eLife.34729*

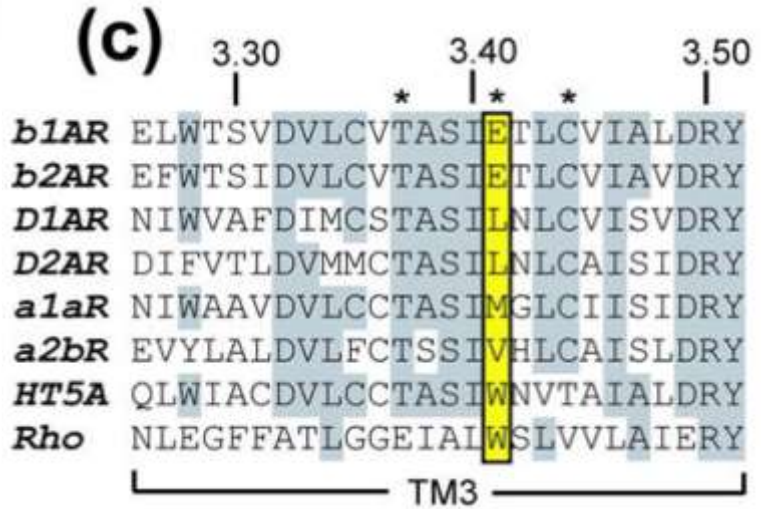
Knowledge Based (Class A only)

- 2 or more structures with this mutant
- Known transferable position
- Most of them destroy/modify a conserved **functional** site

Position	Mutation	Role	Receptor (PDB ID)
3.41	X->W	stabilization of TM3 - TM5 interface	5HT2B (4IB4), 5HT1B (4IAR), ADRB1 (5A8E), ADRB2 (3NY8), CXCR4 (3ODU), DRD3 (3PBL)
2.50	D->N	Sodium pocket	AA2AR (5WF5)
3.39	S->A	Sodium pocket	AA2AR (5WF5)
7.49	D->N	Sodium pocket	P2RY1 (4XNV), P2Y12 (4PXZ)
3.40	I->V, A	P-I-F microswitch motif	ADRB1 (4BVN), APJ (5VBL)
3.49	D,G->A	DRY motif	FFAR1 (5TZR), NTR1 (4XES)
5.58	Y-> A	Conserved activation microswitch	FFAR1 (5TZR), ADRB1 (4BVN)
6.37	L->A	Interferes with DRY motif function	AA2AR (5IU4), NTR1 (4GRV)

X3.41W – “Roth” Mutation





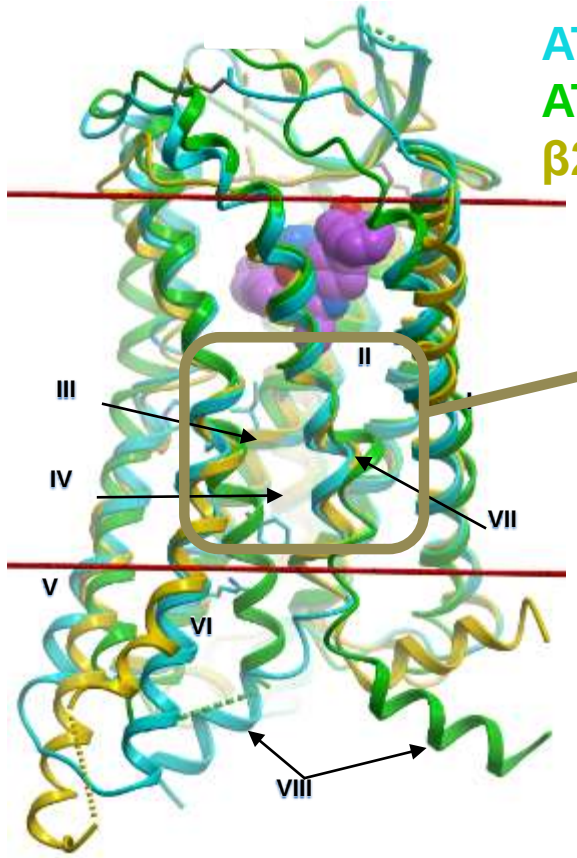
- Stabilizing or neutral in most Class A GPCRs
- Helped to solve >6 receptors, including 2 in active-like state

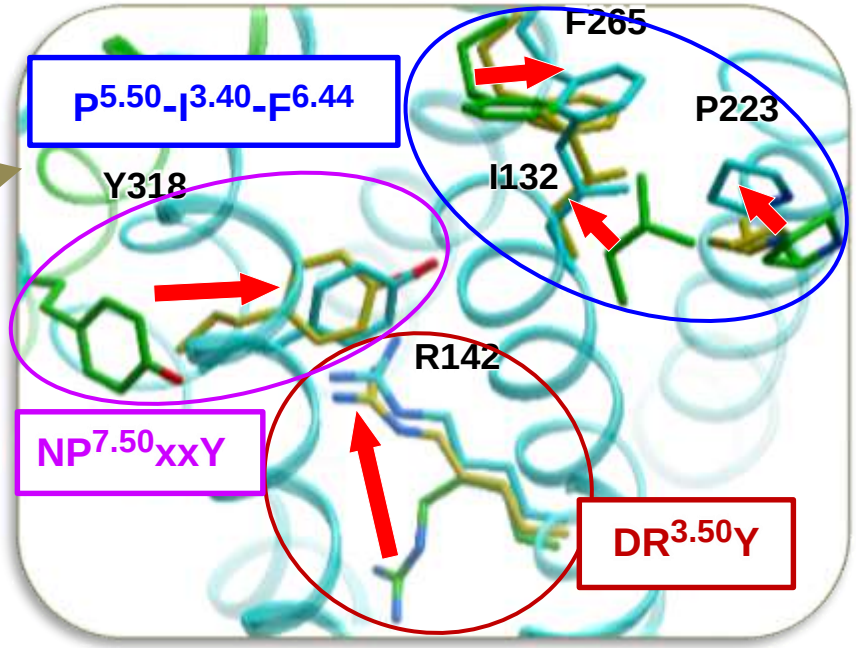
Roth CB et al, (2008) J Mol Biol. 376:1305-19.

Activation Related Changes in Microswitches

Removing switches can decouple from agonist and limit natural motions of receptor

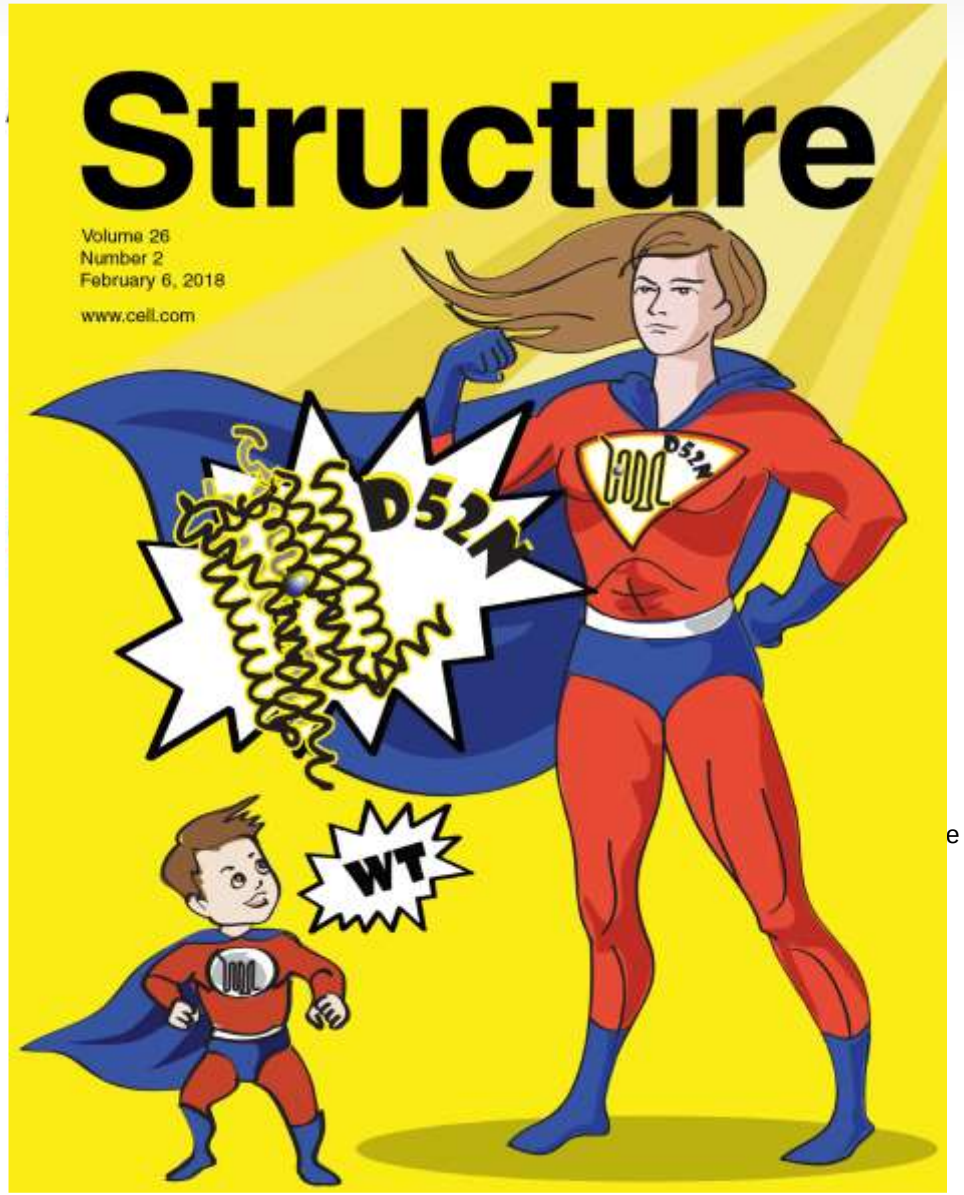
3.40	I->V, A	P-I-F microswitch motif
3.49	D,G->A	DRY motif
6.37	L->A	Interferes with DRY motif function



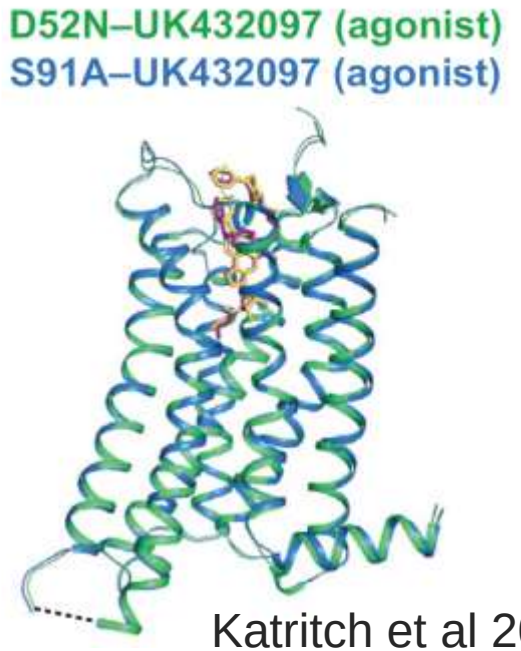


Zhang H et al, *Nature*. 2017;544:327-32.

Stabilizing mutations in conserved Na+ site



	APO T_M ^0C	NECA T_M	Theophylline T_M	ZM241385 Tm
WT	**47**	**53**	**53**	**62**
D2.50A	47	58	54	60
S3.39A	50	61	55	63
N7.49A	54	64	56	64
D2.50N	**58**	**64**	**59**	**66**



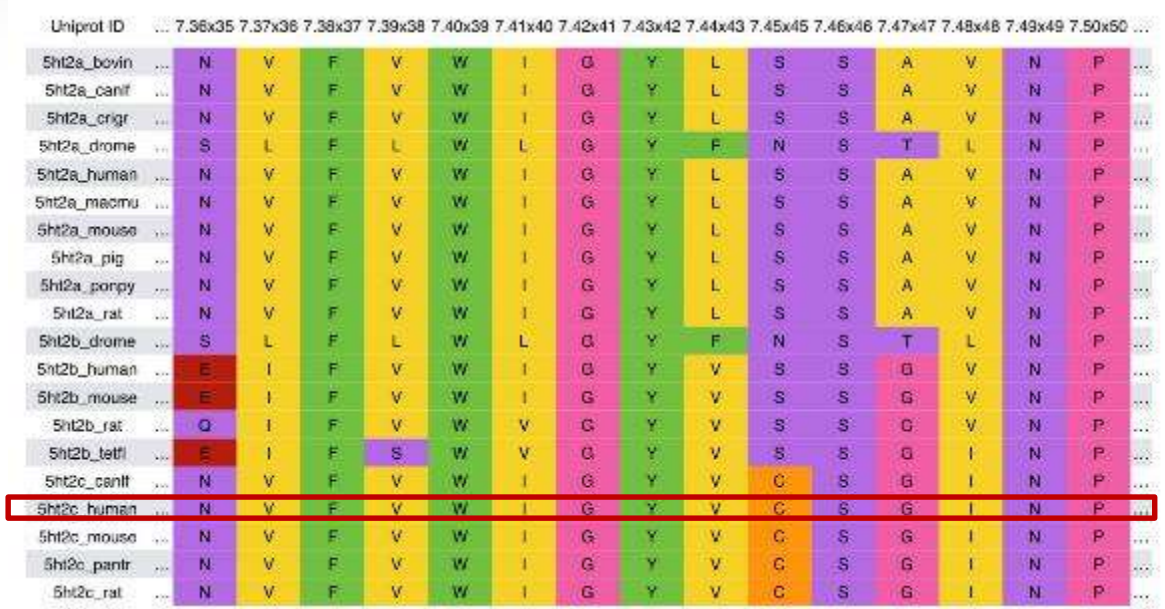
Katritch et al 2014, *TiBS, **39**, 233*
White et al 2018, Structure **26**, 259

Sequence-Based

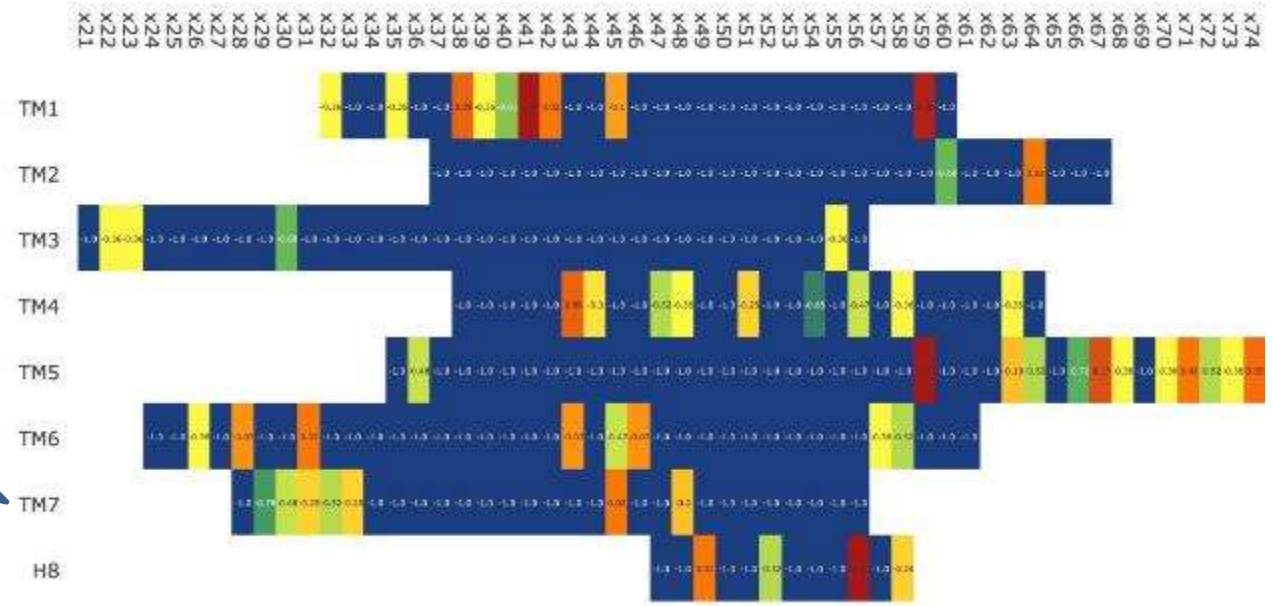
- Goal: identify (and replace) “outliers”: residues which are rarely observed at certain position
- Some positions are highly variable – take into account relative conservation
- Results depend on the GPCR set – use 3 levels of sequence clustering:
  - GPCR Class or branch (SI ≥ 25%)
  - GPCR family (SI ≥ 30%)
  - GPCR subfamily (SI ≥ 35%)

GPCR ID	1	2	3	4	5
GPCR 1	L	V	L	P	W
GPCR 2	L	L	L	P	W
GPCR 3	L	A	L	P	W
GPCR 4	L	T	L	P	W
...	...	...	...	...	...
Target GPCR	L	S	L	S	W

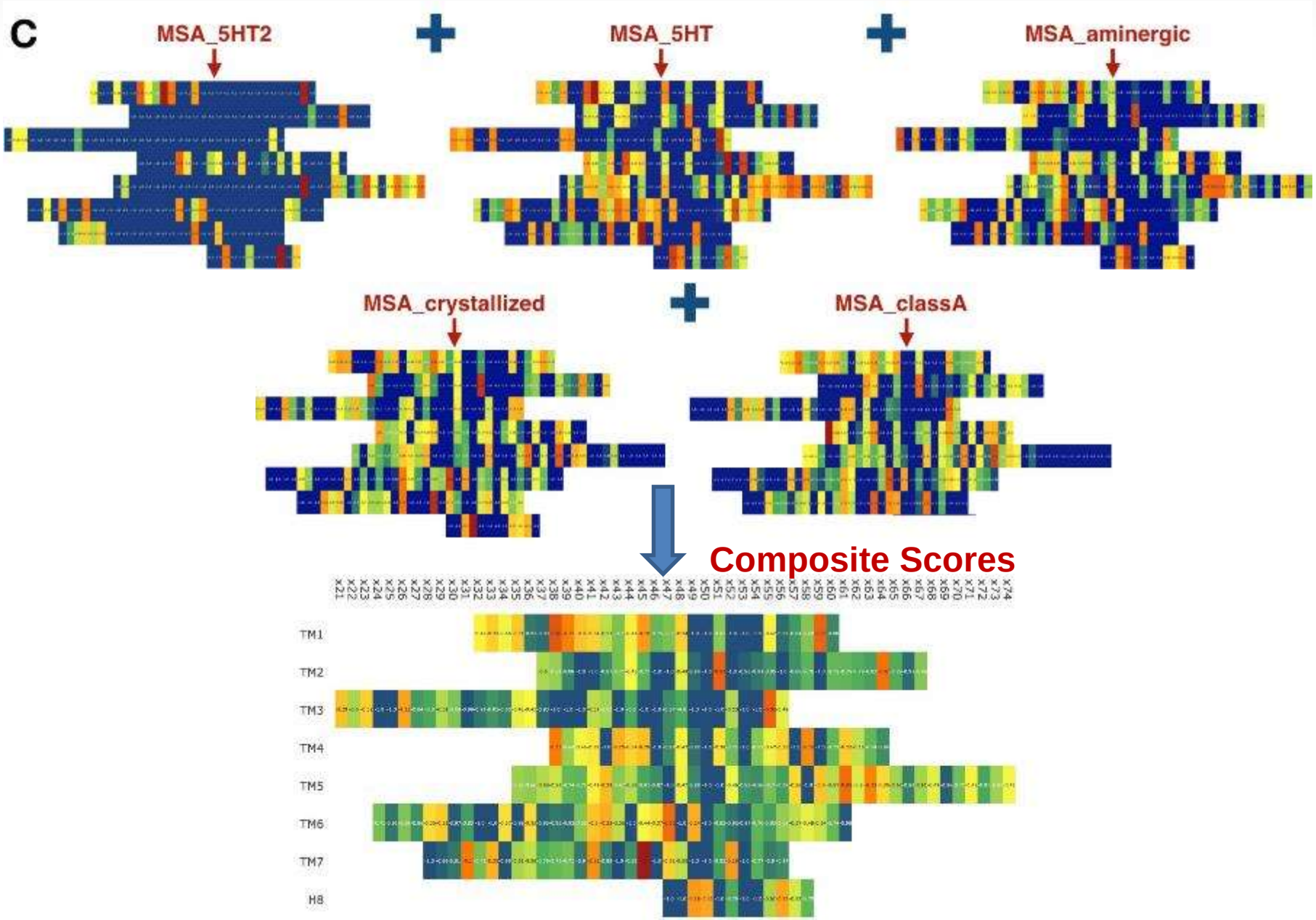
Sequence-Based: Building Scoring Matrix



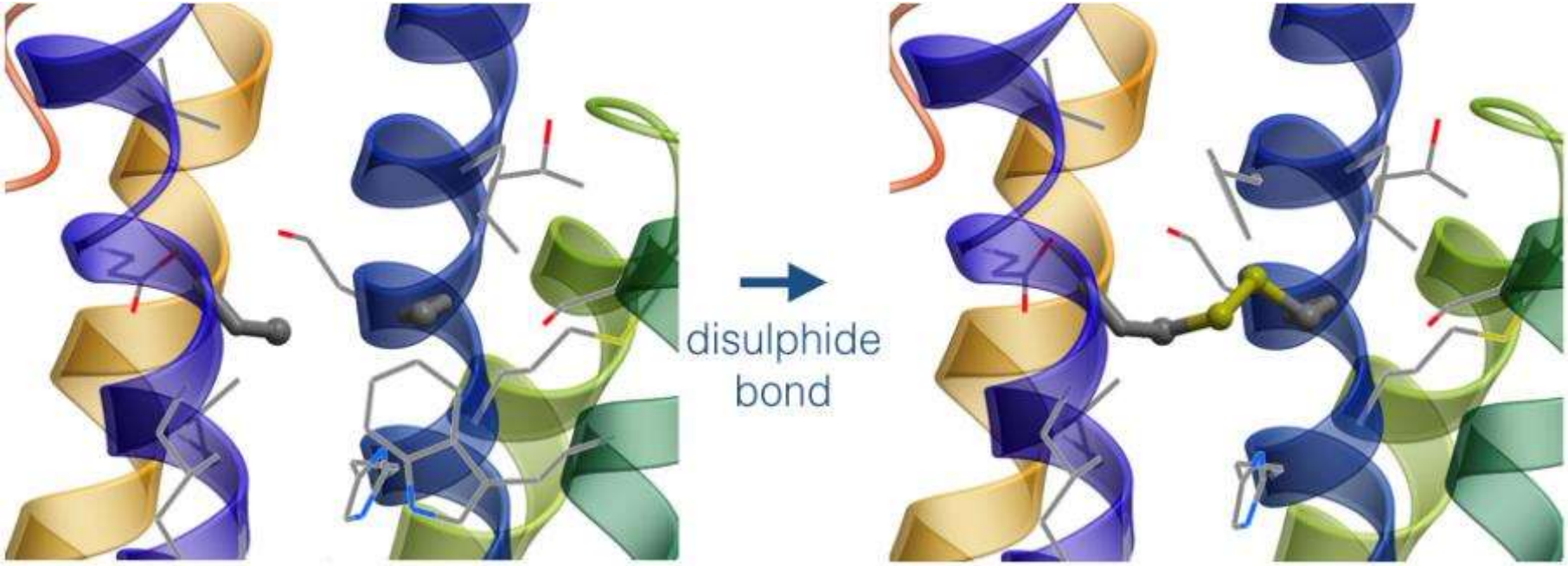
Uniprot ID	...	7.36x35	7.37x36	7.38x37	7.39x38	7.40x39	7.41x40	7.42x41	7.43x42	7.44x43	7.45x45	7.46x46	7.47x47	7.48x48	7.49x49	7.50x50	...
5ht2a_bovin	...	N	V	F	V	W	I	G	Y	L	S	S	A	V	N	P	...
5ht2a_canf	...	N	V	F	V	W	I	G	Y	L	S	S	A	V	N	P	...
5ht2a_crigr	...	N	V	F	V	W	I	G	Y	L	S	S	A	V	N	P	...
5ht2a_drome	...	S	L	F	L	W	L	G	Y	F	N	S	T	L	N	P	...
5ht2a_human	...	N	V	F	V	W	I	G	Y	L	S	S	A	V	N	P	...
5ht2a_macmu	...	N	V	F	V	W	I	G	Y	L	S	S	A	V	N	P	...
5ht2a_mouse	...	N	V	F	V	W	I	G	Y	L	S	S	A	V	N	P	...
5ht2a_pig	...	N	V	F	V	W	I	G	Y	L	S	S	A	V	N	P	...
5ht2a_ponpy	...	N	V	F	V	W	I	G	Y	L	S	S	A	V	N	P	...
5ht2a_rat	...	N	V	F	V	W	I	G	Y	L	S	S	A	V	N	P	...
5ht2b_drome	...	S	L	F	L	W	L	G	Y	F	N	S	T	L	N	P	...
5ht2b_human	...	E	I	F	V	W	I	G	Y	V	S	S	G	V	N	P	...
5ht2b_mouse	...	E	I	F	V	W	I	G	Y	V	S	S	G	V	N	P	...
5ht2b_rat	...	Q	I	F	V	W	V	G	Y	V	S	S	G	V	N	P	...
5ht2b_tetfl	...	E	I	F	S	W	V	G	Y	V	S	S	G	I	N	P	...
5ht2c_canlf	...	N	V	F	V	W	I	G	Y	V	C	S	G	I	N	P	...
5ht2c_human	...	N	V	F	V	W	I	G	Y	V	C	S	G	I	N	P	...
5ht2c_mouse	...	N	V	F	V	W	I	G	Y	V	C	S	G	I	N	P	...
5ht2c_pantr	...	N	V	F	V	W	I	G	Y	V	C	S	G	I	N	P	...
5ht2c_rat	...	N	V	F	V	W	I	G	Y	V	C	S	G	I	N	P	...

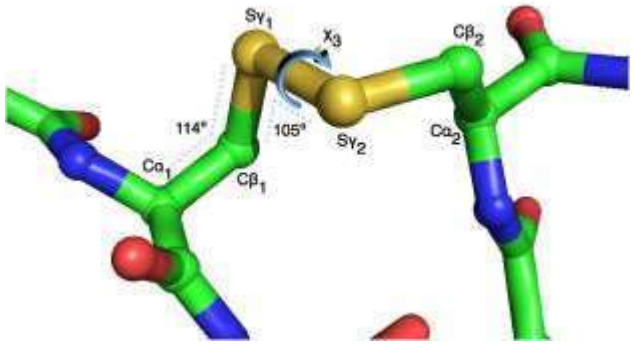


Sequence-Based: Building Scoring Matrix



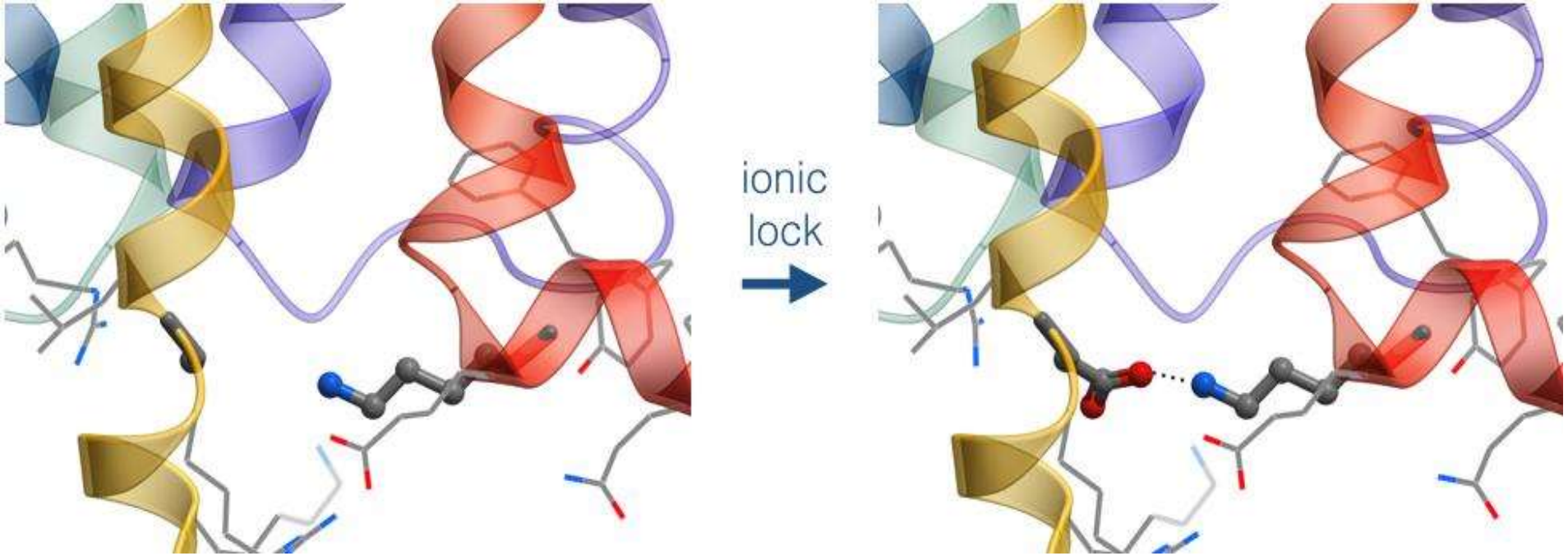
Structure-based: Design of Disulfide Bridges





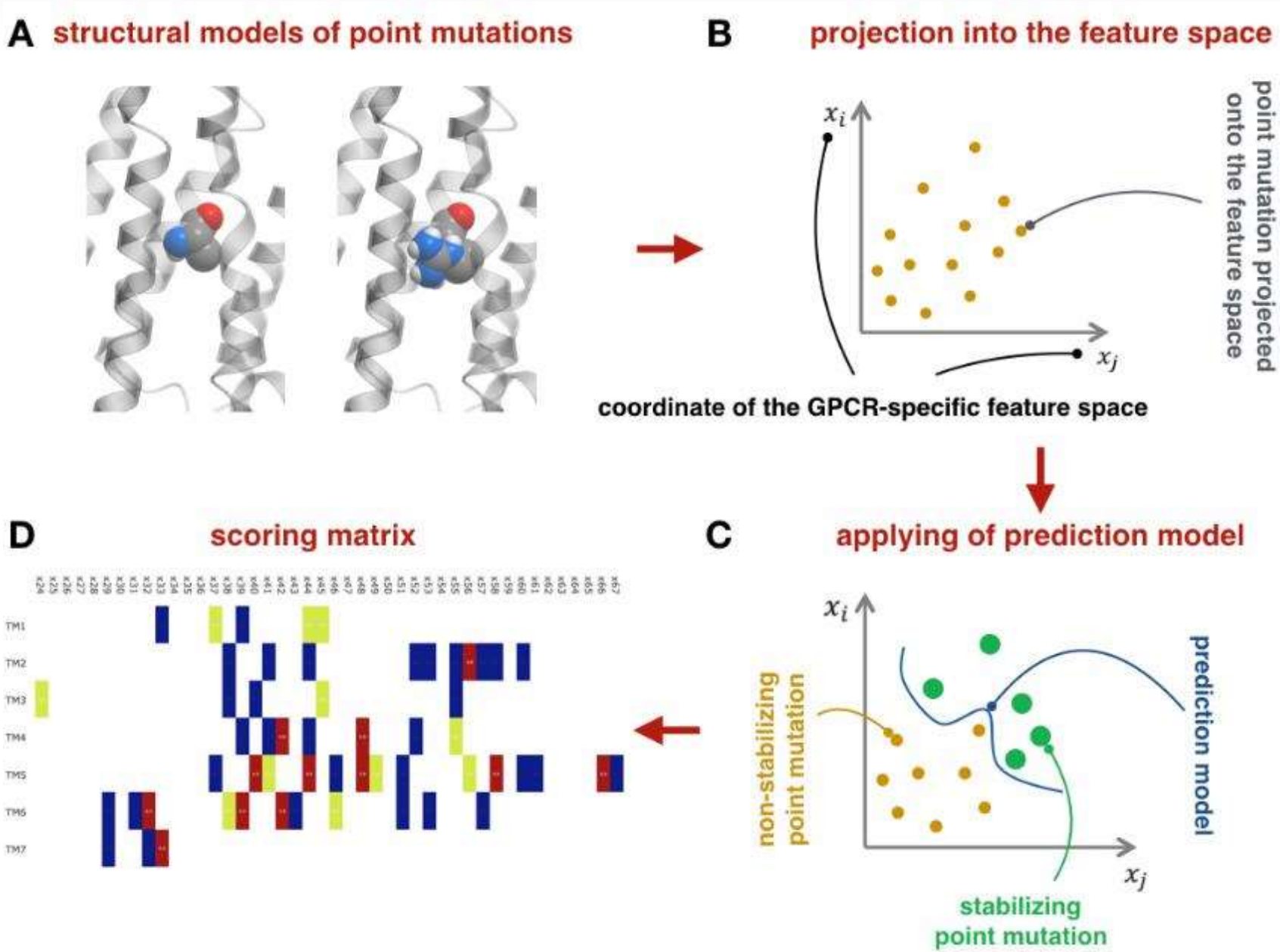
- Find a pair of residues with permissive geometry
- Full energy optimization and scoring energy strain

Structure-based: Design of ionic locks

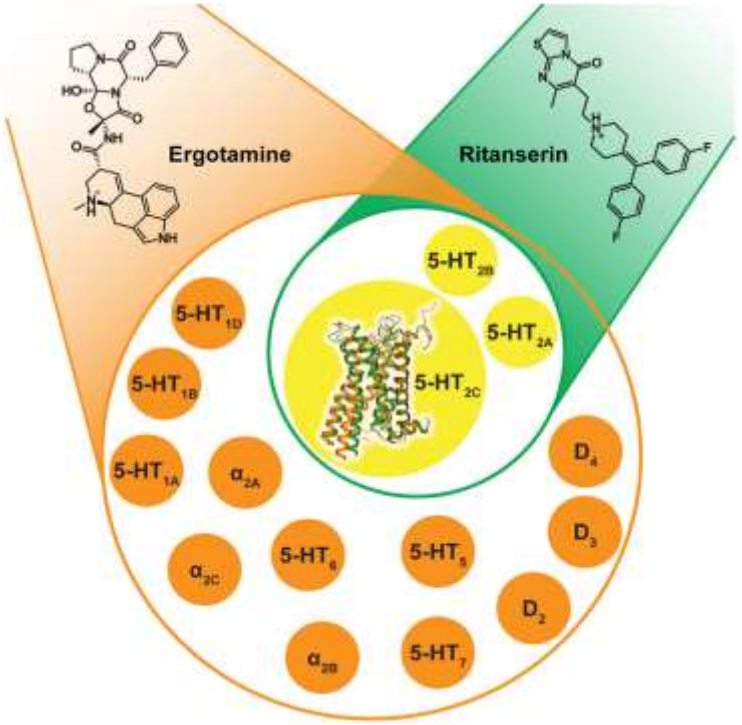


- Find a pair of residues with permissive geometry
- Full energy optimization and scoring energy strain

Machine Learning



Prospective Application to Mutant Discovery: $5HT_{2C}$ structures



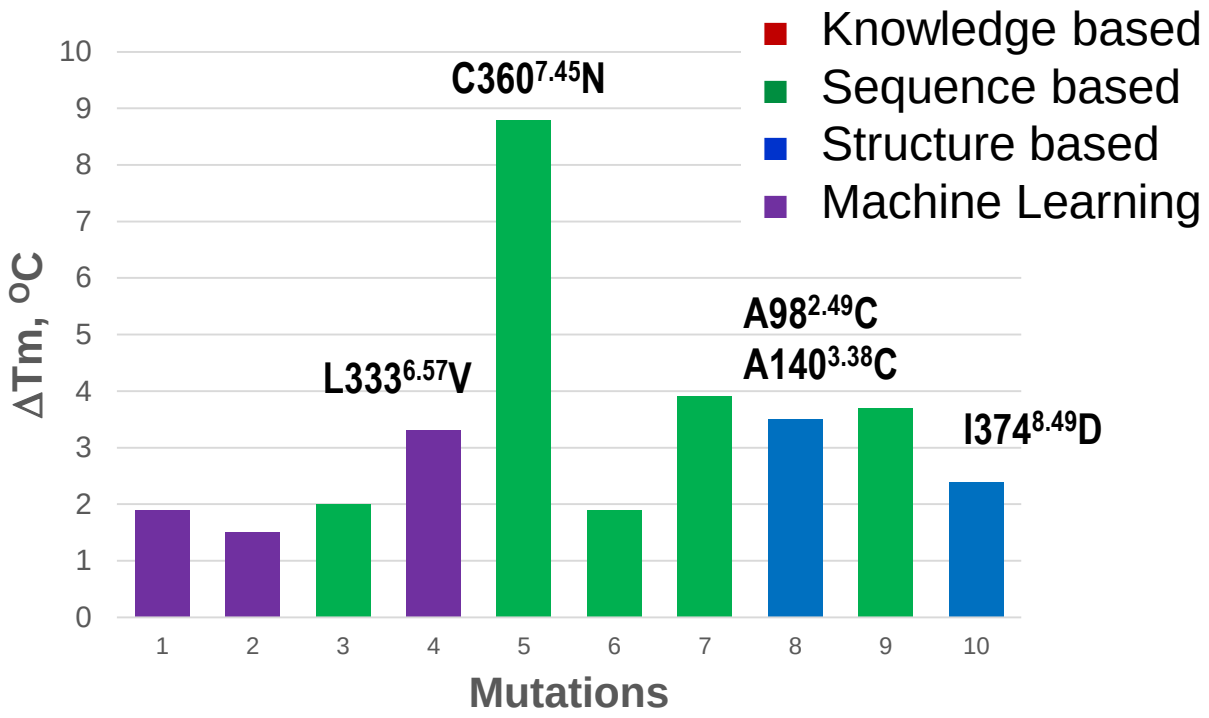
6BQG
6BQH

Peng Y et al. (2018) 5-HT2C Receptor Structures Reveal the Structural Basis of GPCR Polypharmacology. Cell 172(4):719-30

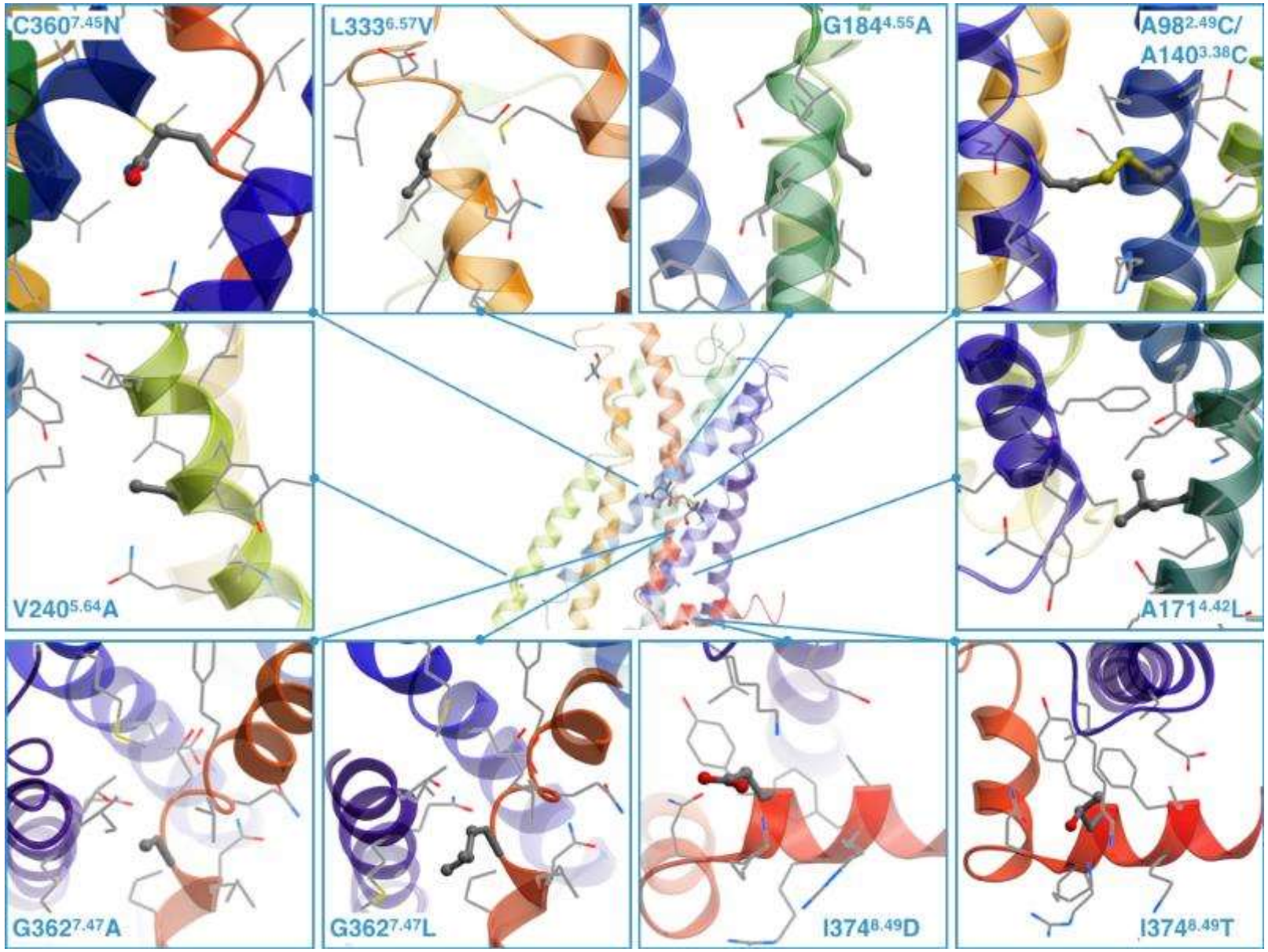
CompoMug prospective screening: $5HT_{2C}$

Mutation	CompoMug Module	aSEC* quality	Tm (°C) ±SEM	ΔTm (C)
WT			50.4 ± 0.8	0.0
$I62^{1.41}V$	Seq-based	~		-0.7
$G69^{1.48}A$	Seq-based	-		-1.4
$D99^{2.50}N$	Knowledge	-		-
$H85^{12.51}E$	Struct-based	N/A		-
$G103^{2.54}A$	Seq-based	-		-4.4
$Y125^{3.23}K$	Seq-based	-		-2.0
$Y125^{3.23}V$	Seq-based	~		-0.7
$M143^{3.41}W$	Knowledge	-		0.6
$R157^{3.55}T$	ML & Seq-based	-		-1.8
$R157^{3.55}Q$	Seq-based	-		-2.0
$T169^{4.40}K$	Seq-based	+		0.2
$A171^{4.42}L$	**ML**	**~**	**52.3 ± 1.2**	**1.9**
$I172^{4.43}A$	Seq-based	-		1.1
$I172^{4.43}F$	Seq-based	~		0.6
$G184^{4.55}A$	**ML**	**+**	**51.9 ± 0.1**	**1.5**
$N203^{ECL2}D$	Struct-based	-		-2.6
$F220^{5.45}I$	ML	~		0.0
$F224^{5.48}Y$	ML &Seq-based	-		-3.3
$C235^{5.59}F$	Seq-based	~		0.1
$L236^{5.60}R$	ML &Seq-based	N/A		-
$V240^{5.64}A$	**Seq-based**	**+**	**52.4 ± 0.5**	**2.0**
$V240^{5.64}S$	Seq-based	+		0.3
$G314^{6.38}A$	ML	-		-4.0
$L333^{6.57}V$	**ML &Seq-based**	**+**	**53.7 ± 0.6**	**3.3**
$K348^{7.32}A$	Seq-based	-		-4.4
$C360^{7.45}N$	**Seq-based**	**+**	**59.2 ± 0.5**	**8.8**
$G362^{7.47}L$	**Seq-based**	**+**	**52.3 ± 0.7**	**1.9**
$G362^{7.47}A$	**Seq-based**	**+**	**54.3 ± 0.7**	**3.9**
$L370^{7.55}D$	Struct-based	-		-2.3
$K373^{8.48}E$	Struct-based	-		-0.4
$I374^{8.49}D$	**Struct-based**	**+**	**53.9 ± 0.8**	**3.5**
$I374^{8.49}T$	**Seq-based**	**+**	**54.1 ± 0.9**	**3.7**
$Y375^{8.50}F$	Seq-based	-		-2.4
$N381^{8.56}R$	Sequence-based	~		0.6
$T67^{1.46}C/G103^{2.54}C$	Struct-based	-		-
$V74^{1.53}C/A96^{2.47}C$	Struct-based	-		-
$A87^{2.38}C/A171^{4.42}C$	Struct-based	~		-
$A98^{2.49}C/A140^{3.38}C$	**Struct-based**	**~**	**52.8 ± 1.0**	**2.4**
$T369^{7.54}C/Y375^{8.50}C$	Struct-based	N/A		-

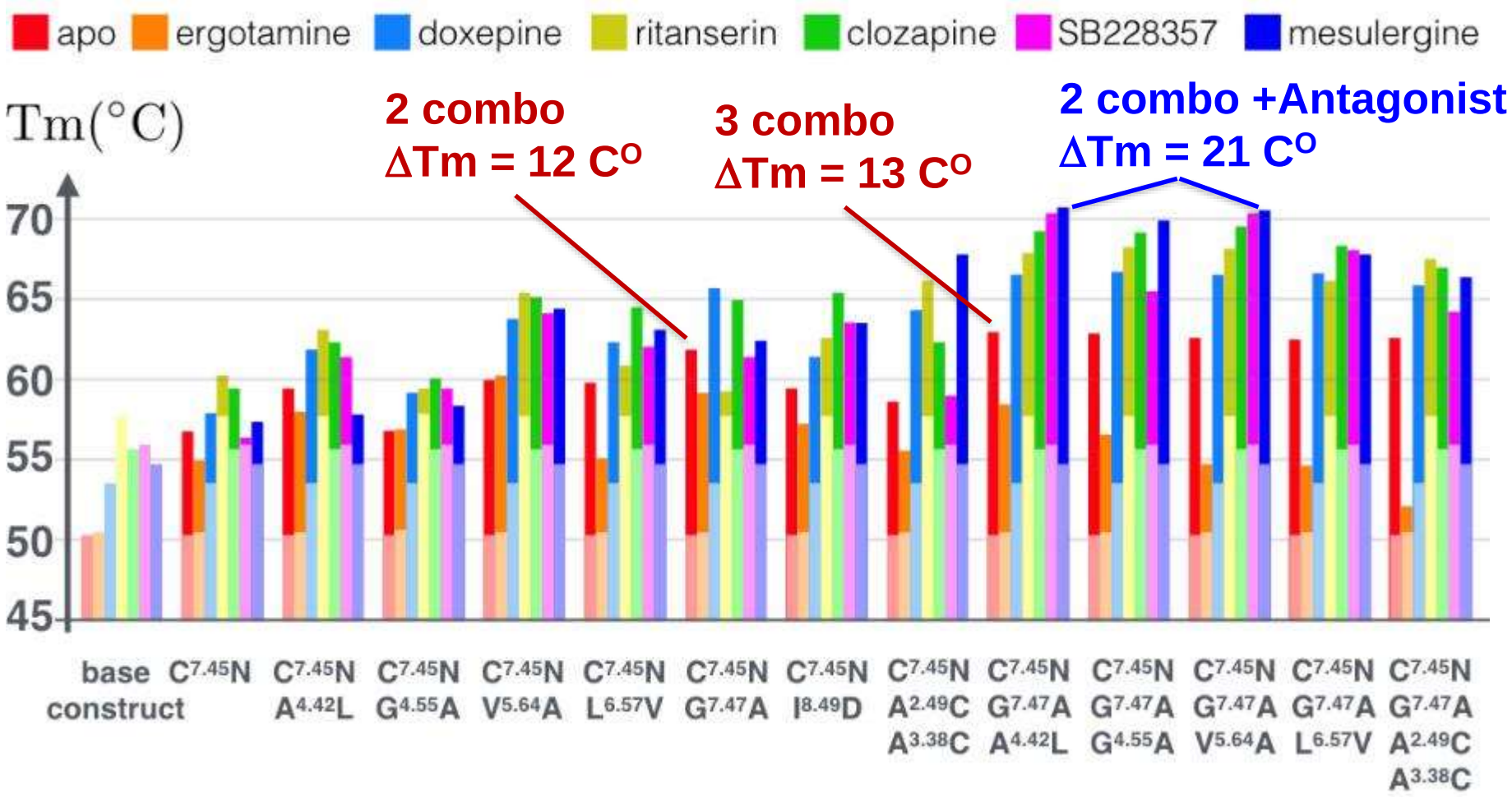
- 40 mutations tested
- 10 improved both Tm & SEC **(~25%)**
- Best shift ΔTm **~ 9°C**
- **3 components contributed, but not Knowledge-based**



$5HT_{2C}$: Mutations in the structural model

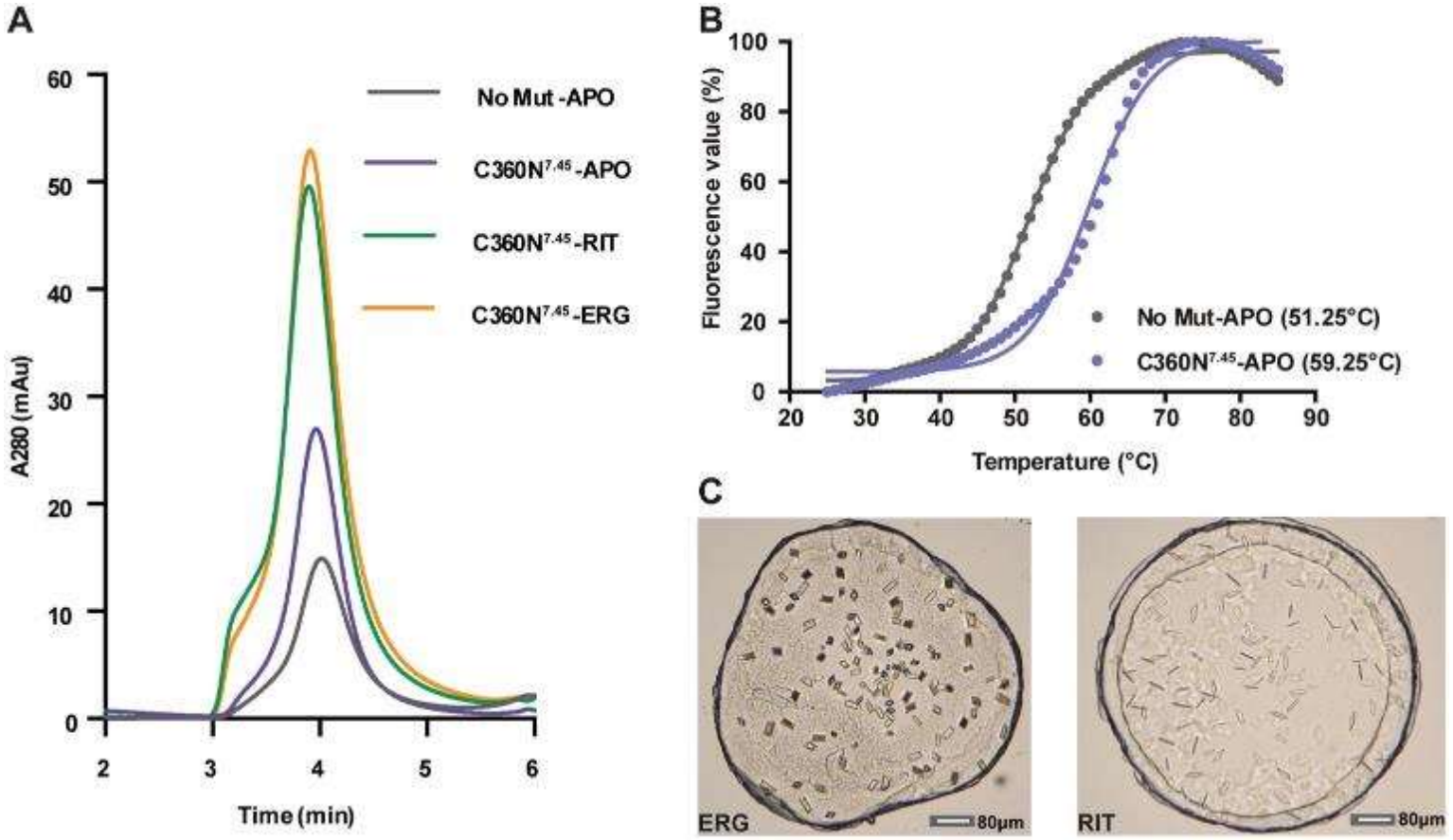


$5HT_{2C}$: Combined mutations in complex with ligands



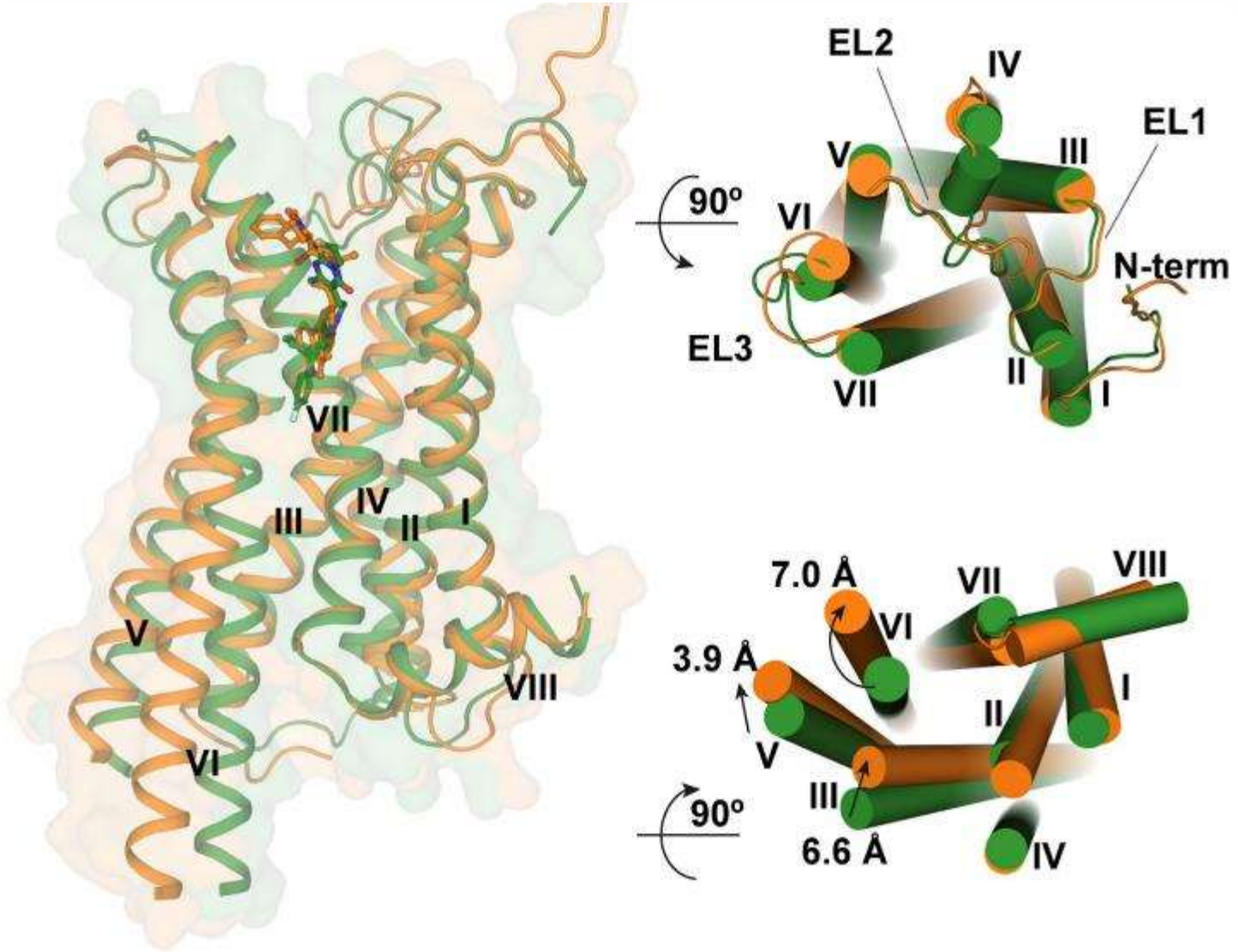
5-HT_{2C}: Crystals and structures

- Combo mutations yield first 5-HT_{2C} crystals with antagonists
- Single $C360^{7.45}N$ allowed crystallization of both agonist and antagonists
- Structures for agonist Ergotamine (3.0 Å) and antagonist Ritanserin (2.7 Å)



Peng Y et al. (2018) Cell 172(4):719-30

5-HT_{2C} inactive and active-like structures



Peng Y et al. (2018) Cell 172(4):719-30

More CompoMug predictions tested

Target	# tested single mutants	Hit rate	Best ΔTm, C°	Best ΔTm, C° combined mutations	Crystallized /Solved	Comment
5HT2C	40	25%	9	13	**Yes/Yes**	1 mutant in structure
Target #2	40	25%	6	7	**Yes/Yes**	2+4 mutants in structure
Target #3	60	17%	10			
Target #4	36	30%	5	9	**Yes/Yes**	5 mutations in structure
Target #5	40	20.0%	4			
Target #6	60	12%	4			
Target #7	40	10.0%	3	4		
Target #8	60	7%	4			
Target #9	60	25%	3	16	**Yes/Yes**	non Class A; 4 mutants
Target #10	60	11%	3			
EP3	30	15%	**0**		**Yes/Yes**	Only Improved Diffraction
Taste /Class C	40-60	0%				

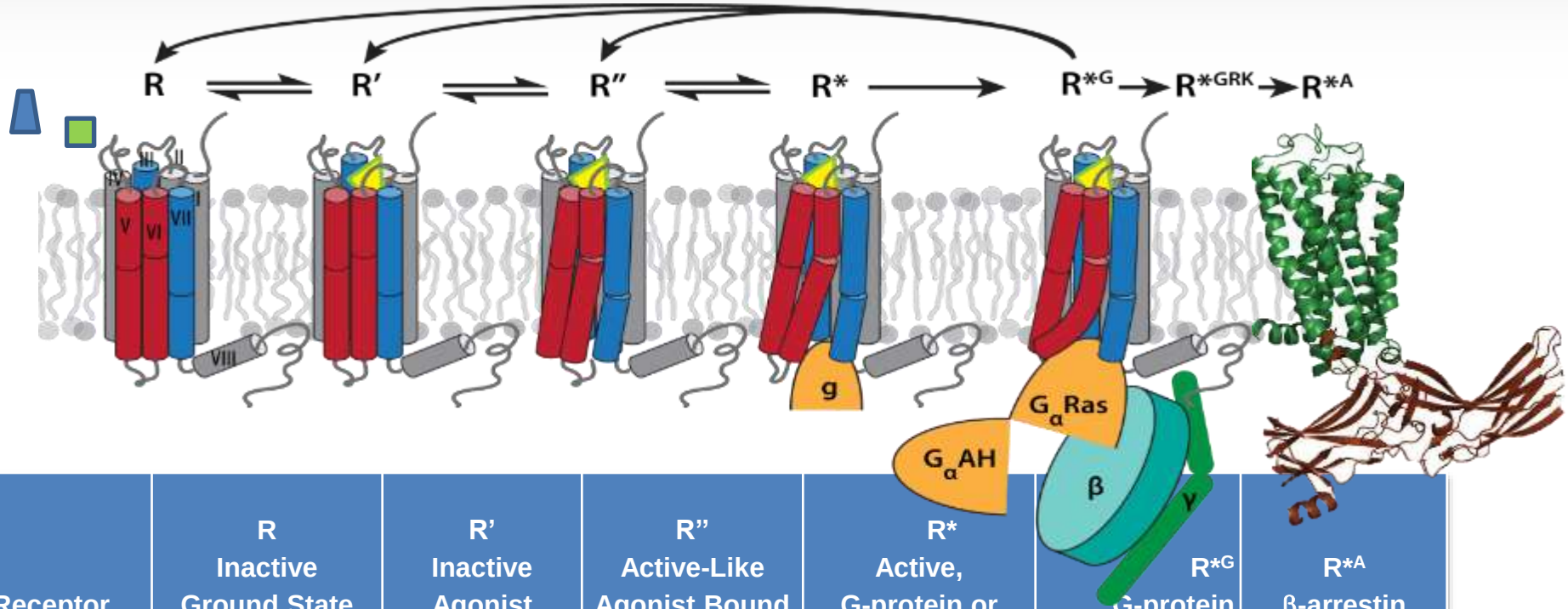
CompoMug: Summary & Outlook

- **CompoMug** already shows 15-30% "hit rate" for most GPCRs targets, and up to 16C° combined ΔTm
- Complementary to other engineering approaches
- All 4 modules make important contributions:
  - **Sequence-based** module most universally applicable
  - **Energy-based predictions can improve with each GPCR structure (better homology models)**
  - **Machine learning continuously can improve with each new stability dataset obtained**
- Utility for both **SBDD** and for **assay development!**

Outline

- Rational prediction of stabilizing mutations: CompoMug
- **New insights into GPCR function and allosteric mechanisms**
- Structure-Based ligand discovery for GPCRs

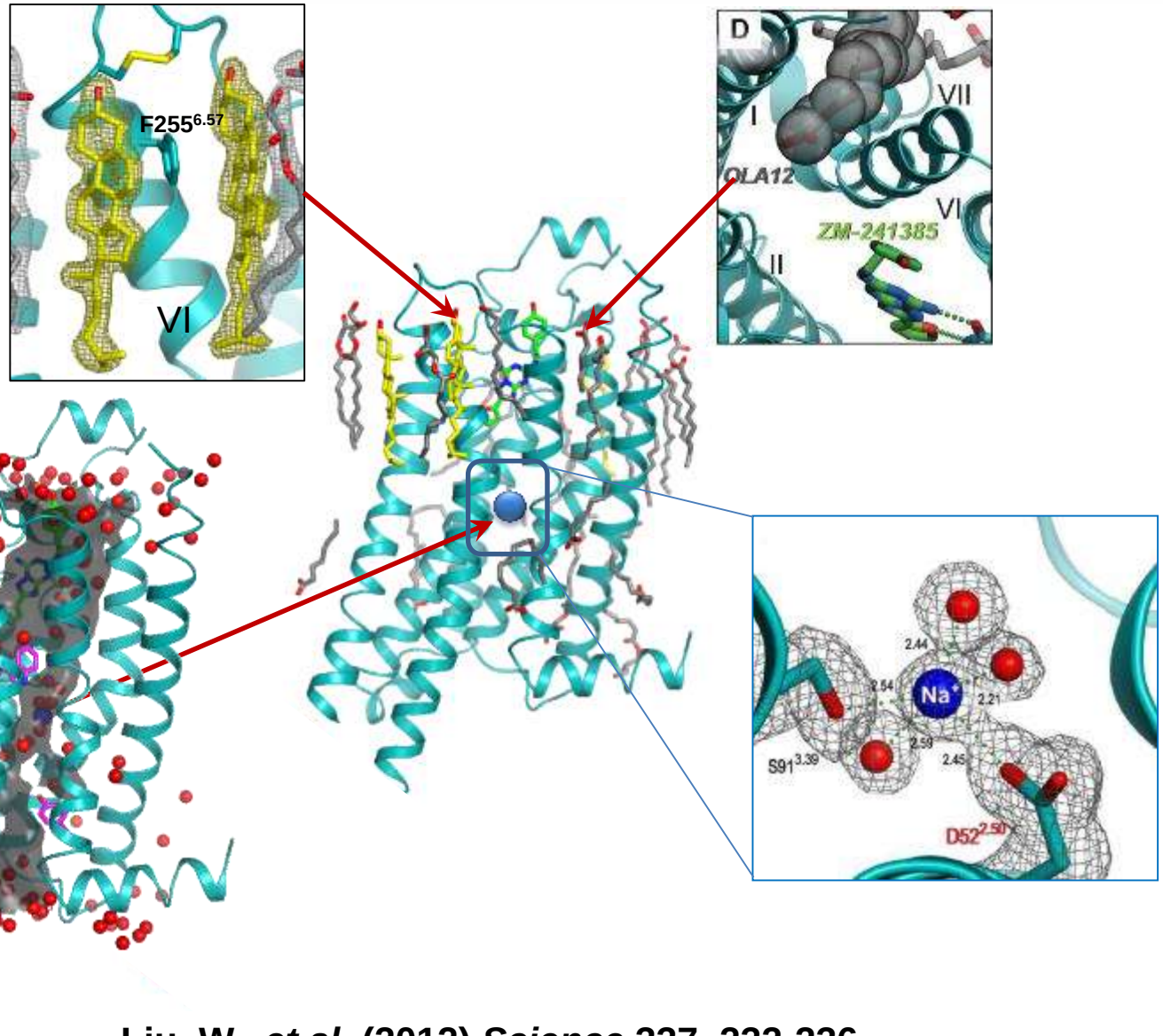
Dynamic Mechanisms of GPCR action on atomic level



Receptor	R Inactive Ground State	R' Inactive Agonist Bound	R'' Active-Like Agonist Bound	R* Active, G-protein or Mimic Binding	R*G G-protein Signaling	R*A β-arrestin Complex
Adenosine A_{2A}	3EML,3REY, 3RFM, 3PWH, **4EIY**	--	**3QAK**,2YDO,2YDV, 4UHR, 4UG2	--	5G53	
β_1-Adrenergic	2VT4,2YCW,2YCX, 2YCY,2YCZ	2Y00 -2Y04	--	--	--	
β_2-Adrenergic	2RH1, 2R4R, 2R4S,3D4S	3PDS	--	3P0G[c] (3.5)	**3SN6**	
Rhodopsin	1F88,1U19	2G87, 2HPY	3CAP[f],	3DQB,2X72,3PQR, 3PXO	--	4ZWJ
Muscarinic	3UON			4MGS,4MGT		
Opioid	4DKL			5C1M, κ-OR		
CB1	5U09		5XRA			
P2Y12	4NTJ	4PXZ				
NTSR1, ETB_4 5HT1B, 5HT2B		5GLH	4GRV, 4IAQ,4IAR, 4IB4			

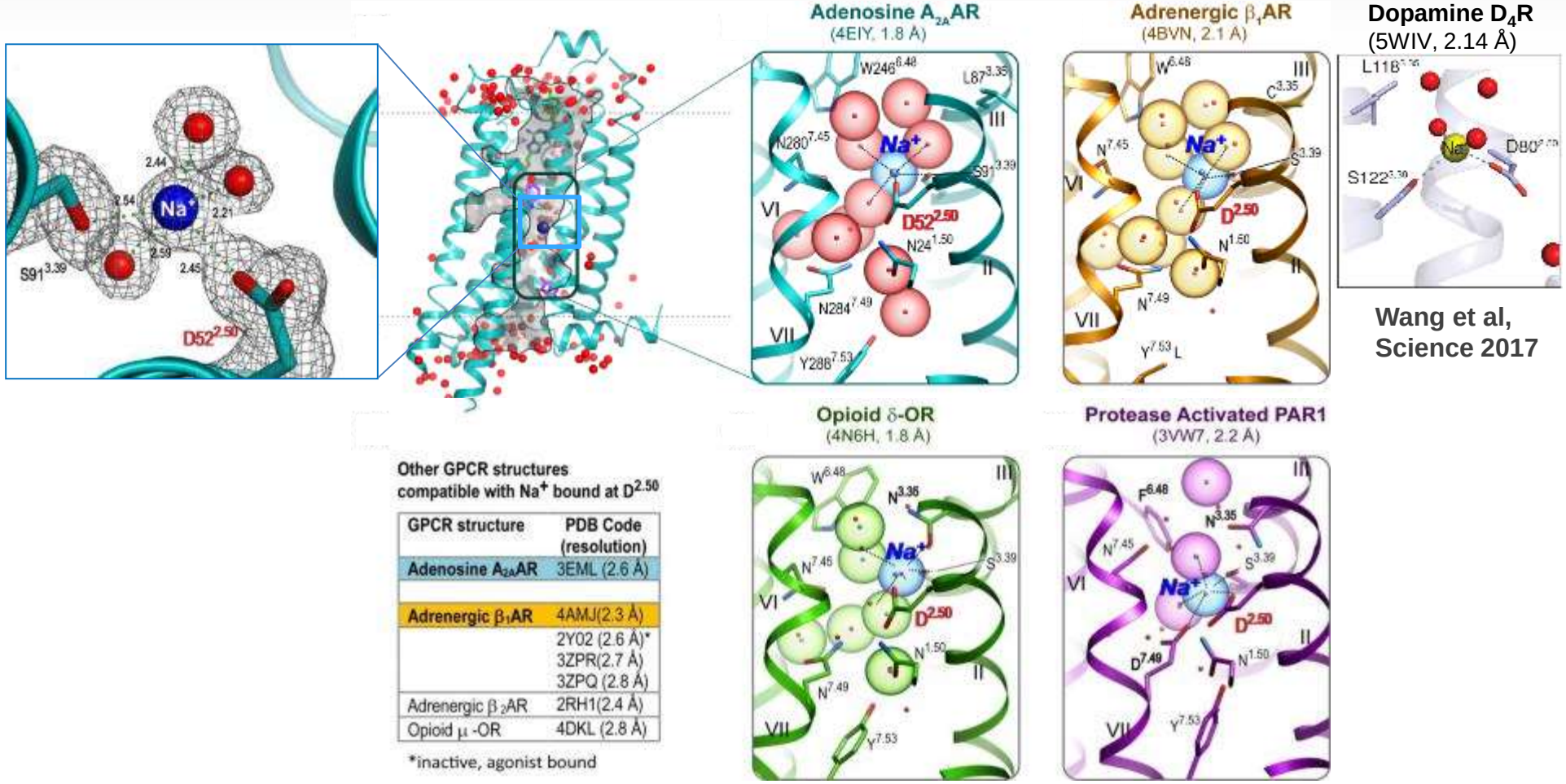
Updated from Katritch, Stevens, Cherezov, *Annual Rev. Pharmacology & Toxicology,* 2013

Endogenous allosteric molecules in 1.8 Å Structure of human $A_{2A}AR$



Liu, W., *et al.* (2012) *Science* 337, 232-236

Allosteric Na^+ site in class A GPCR structures



Other GPCR structures compatible with Na^+ bound at $D^{2.50}$

GPCR structure	PDB Code (resolution)
Adenosine A_{2A}AR	3EML (2.6 Å)
Adrenergic β_1AR	4AMJ(2.3 Å)
	2Y02 (2.6 Å)*
	3ZPR(2.7 Å)
	3ZPQ (2.8 Å)
Adrenergic β_2AR	2RH1(2.4 Å)
Opioid μ -OR	4DKL (2.8 Å)

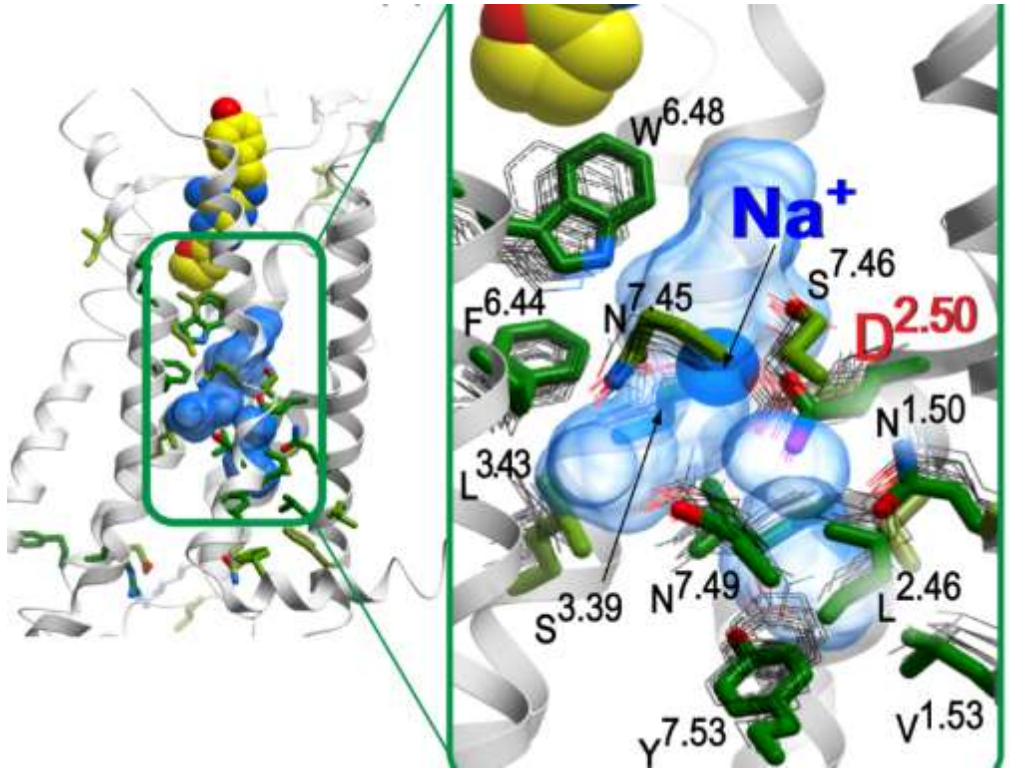
*inactive, agonist bound

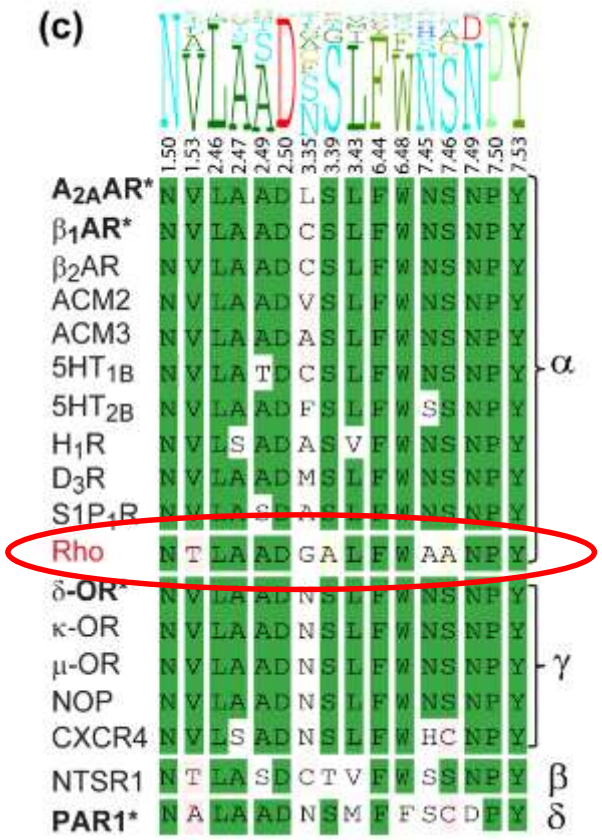
- Na+ crystallographically observed in α, δ, γ branches of Class A
- The pocket highly conserved in most class A GPCRs

Liu et al. 2012 *Science 337, 232-236;* Fenalti et al. 2014, Nature 506, 191-196.; Katritch et al. (2014) *TiBS* 39:233-44
Miller-Gallacher et al, *PLoS One* 2014, 9, e92727, 1. Zhang et all, Nature 2012, 492, 387-392.

Na^+ Pocket Conservation in Class A GPCRs

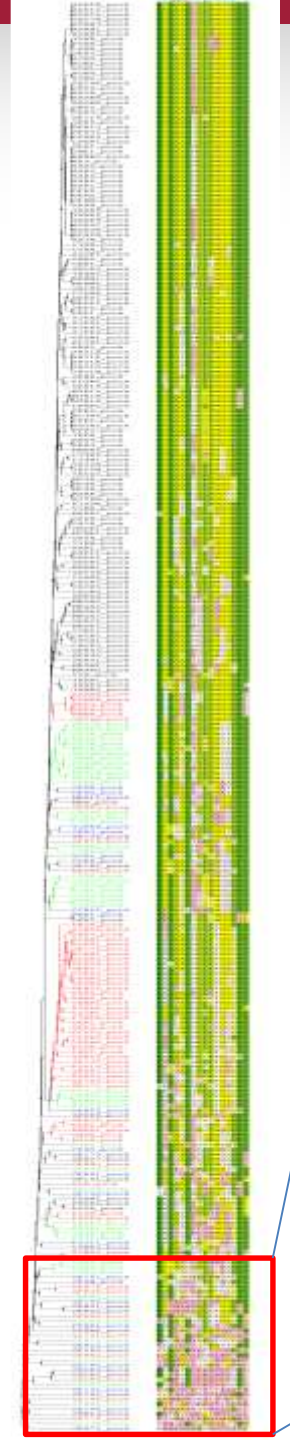
- Largest conserved 3D cluster (15 residues of 34 most conserved residues)
- Includes NPxxY in helix VII and FxxCW in helix VI, but not DRY in helix III
- D2.50 conserved in 95% of GPCRs





Gutierrez-de-Teran et al, Structure 2013, 21, 2175-218
Katritch et al. (2014) *TiBS* 39:233-44

Only 5% of Class A GPCRs lack D2.50



	Uniprot ID	Description*	Protein name	#2.50	Other Acid
1	VNRL4_HUMAN	PSEUDOGENE	Putative vomeronasal receptor-like protein 4	H	
2	CCBP2_HUMAN	DECOY RECEPTOR D6	Chemokine-binding protein 2	N	
3	CCRL2_HUMAN	Non-signaling	C-C chemokine receptor-like 2	N	
4	GNRR2_HUMAN	NON_FUNCTIONAL	Putative gonadotropin-releasing hormone II receptor	F	
5	GPR26_HUMAN	Orphan, Constitutively active	GPCR 26		
6	GPR78_HUMAN	Orphan, Constitutively active	GPCR 78		
7	GP161_HUMAN	Orphan, negative regulator of Shh pathway	GPCR 161	N	
8	LGR4_HUMAN	ORPHAN, binds Frizzled and LRPs	Leucine-rich repeat-containing GPCR 4	N	
9	LGR5_HUMAN	ORPHAN, binds Frizzled and LRPs	Leucine-rich repeat-containing GPCR 5	N	
10	LGR6_HUMAN	ORPHAN, binds Frizzled and LRPs	Leucine-rich repeat-containing GPCR 6	N	
11	NTR2_HUMAN	Low_affinity to NT, constitutively active [19, 20]	Neurotensin receptor type 2	G	
12	GP141_HUMAN	Probable GPCR	Probable G protein coupled receptor 141	H	
13	GP146_HUMAN	Probable GPCR	Probable GPCR 146	A	
14	GP148_HUMAN	Probable GPCR	Probable GPCR 148	Y	
15	GP150_HUMAN	Probable GPCR	Probable GPCR 150	F	
16	GP153_HUMAN	Probable GPCR	Probable GPCR 153	V	
17	GP162_HUMAN	Probable GPCR	Probable GPCR 162	V	
18	GP176_HUMAN	Probable GPCR	Probable GPCR 176 (HB-954)	I	
19	GPR21_HUMAN	Probable GPCR	Probable GPCR 21	A	
20	GPR22_HUMAN	Probable GPCR	Probable GPCR 22	V	
21	GPR33_HUMAN	Probable GPCR	Probable GPCR 33	Y	
22	GPR52_HUMAN	Probable GPCR	Probable GPCR 52	A	
23	GPR62_HUMAN	Probable GPCR	Probable GPCR 62	A	
24	GPR75_HUMAN	Probable GPCR	Probable GPCR 75	L	
25	GPR82_HUMAN	Probable GPCR	Probable GPCR 82	N	
26	O10J6_HUMAN	Putative olfactory	Putative olfactory receptor 10J6	K	
27	OR2G2_HUMAN	Olfactory	Olfactory receptor 2G2	Y	E3.39
28	OR4F6_HUMAN	Olfactory	Olfactory receptor 4F6	N	E3.39
29	OR5BH_HUMAN	Olfactory	Olfactory receptor 5B17	G	E3.39
30	OR8J1_HUMAN	Olfactory	Olfactory receptor 8J1	N	E3.39
31	OR8J2_HUMAN	Olfactory	Olfactory receptor 8J2	N	E3.39
32	OR8J3_HUMAN	Olfactory	Olfactory receptor 8J3	N	E3.39
33	GP143_HUMAN	Signals	GPCR 143 (Ocular albinism type 1)	A	D3.39
34	GPBAR_HUMAN	Signals	G protein coupled bile acid receptor 1	G	D7.50
35	GNRHR_HUMAN	Signals	Gonadotropin-releasing hormone receptor	N	D7.49
36	OPSB_HUMAN	Opsin, no Na^+ binding	Blue-sensitive opsin	G	

Orphans, Constitutive, Non-signaling

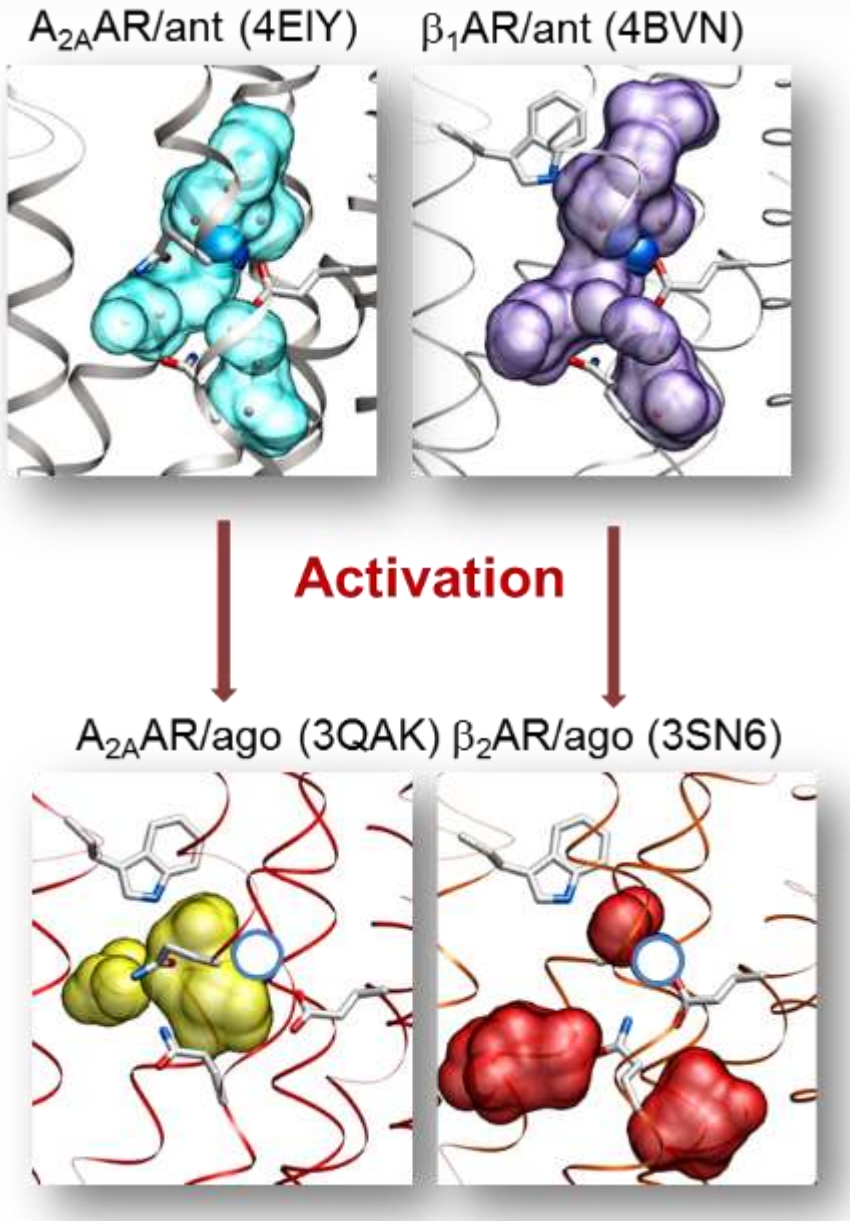
Probable/ Putative

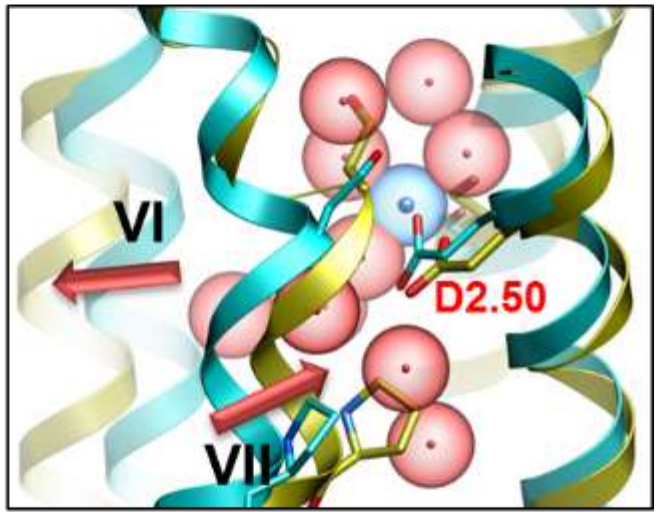
Have alternative acidic side chain In the pocket

BlueOpsin

Katritch et al. (2014) *TiBS* 39(5): p. 233-244

Collapse of the Na+ pocket in GPCR Activation



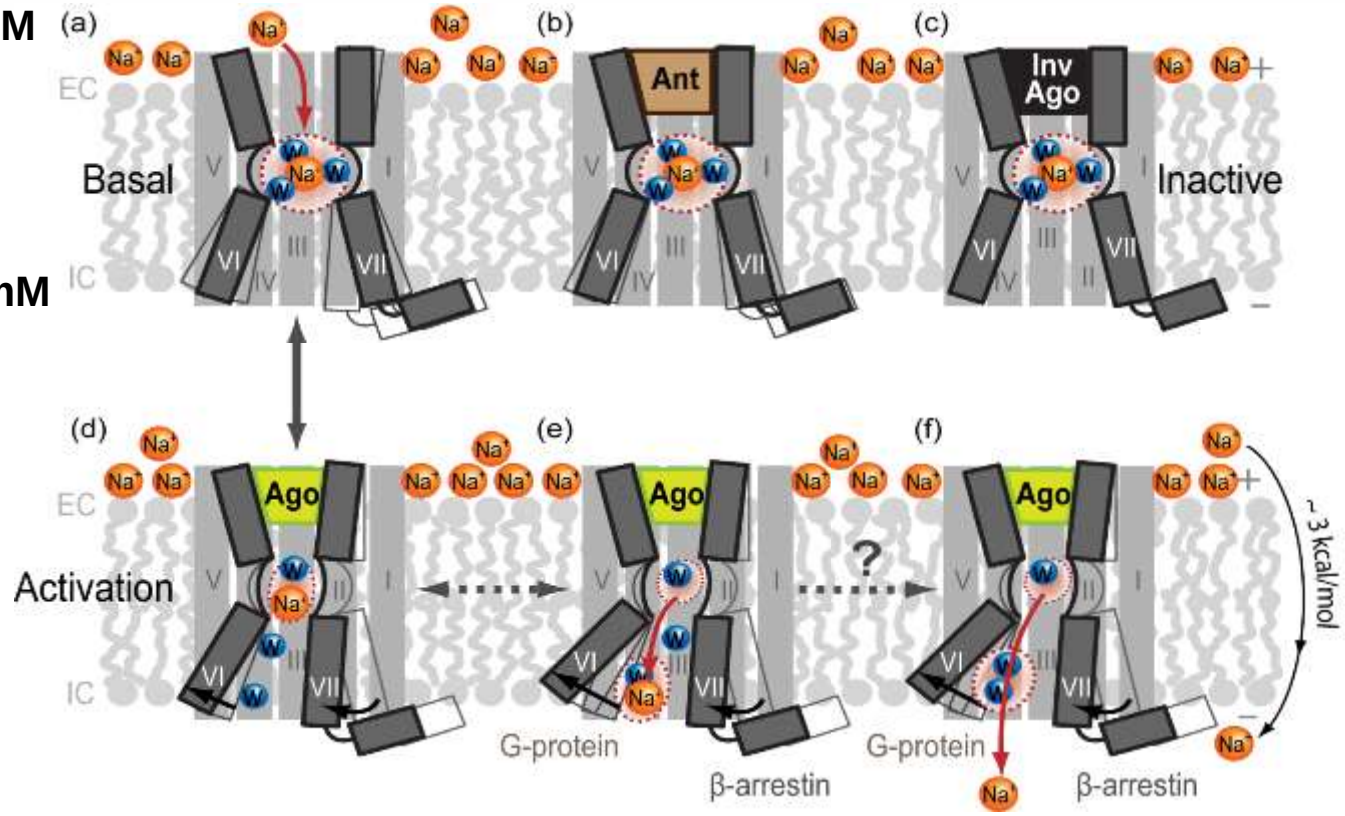


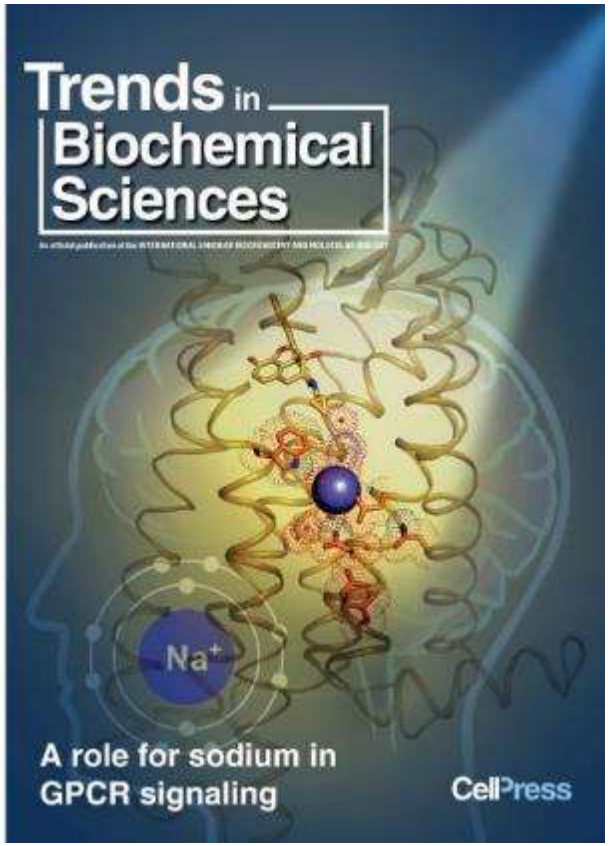
Katritch et al. (April 2014) *TiBS*

Na^+ Plays a Key Role in Activation Mechanism

[Na] ~140 mM

[Na] ~12 mM

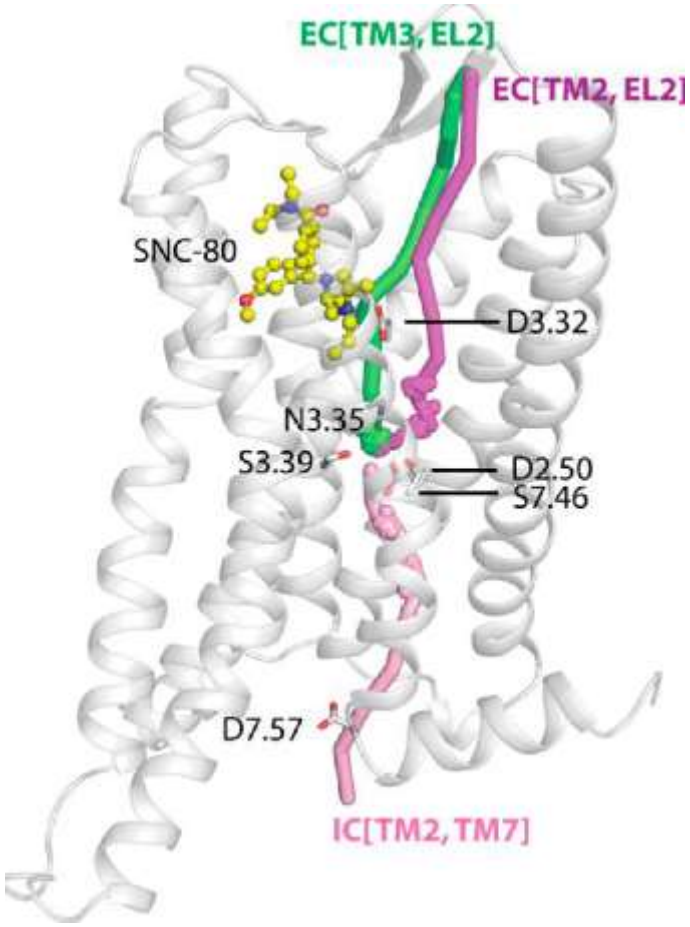




Allosteric Sodium in GPCRs:
One of the ten Science Signaling "Breakthroughs of 2014"

Fenalti et al. 2014, Nature 506, 191-196
Katritch et al. (2014) *TiBS* 39:233-44

But where does it go upon activation?



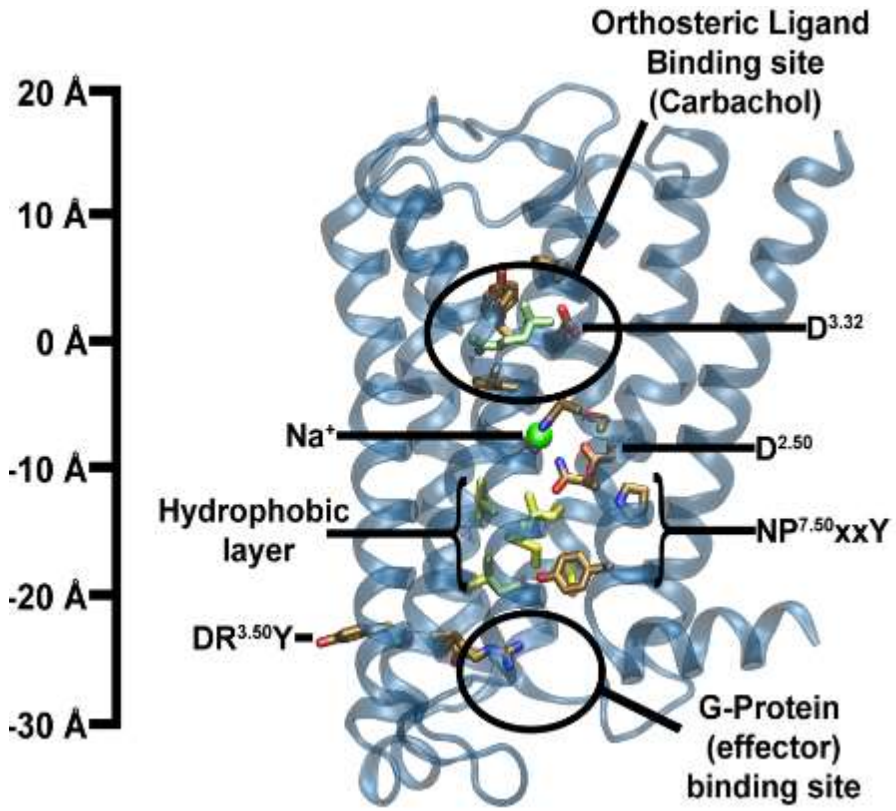
- "In MD simulations of activation transition, Na+ does not leave the allosteric pocket but rather kept its coordination with residues D2.50, N3.35, and S3.39"
- Used Accelerated MD, but only 15% follow the intracellular path

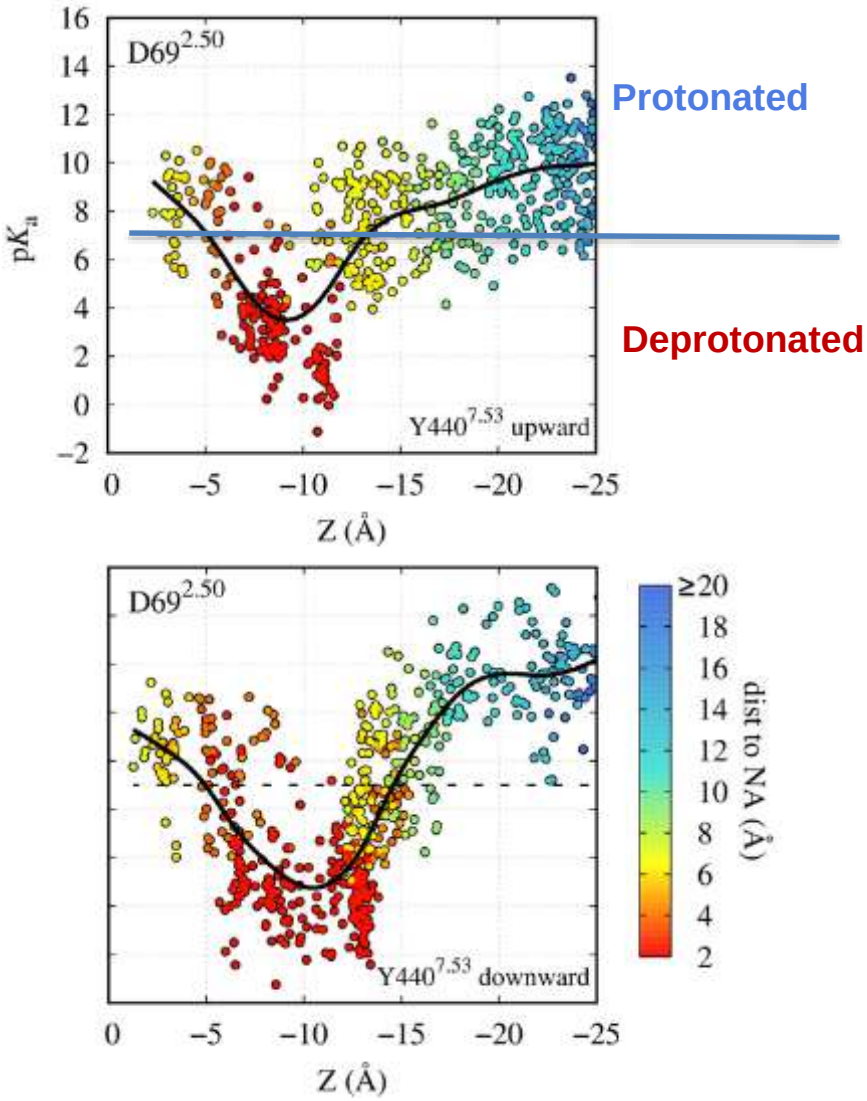
These MD simulations did not take into account:
1. Electrostatic and Na+ concentration gradient
2. Possibility of D2.50 protonation

Shang et al *Biochemistry* ***2014, 53, 5140-5149.***

Is protonation of D2.50 involved in Sodium transition?

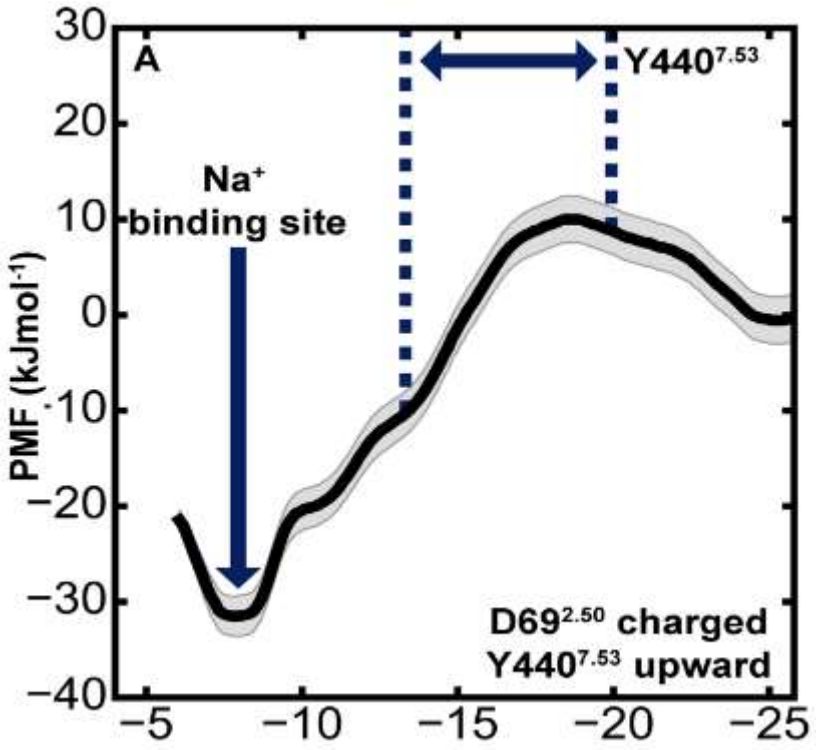
Calculated *pKa* of D2.50 strongly depends on direct contact with Na+

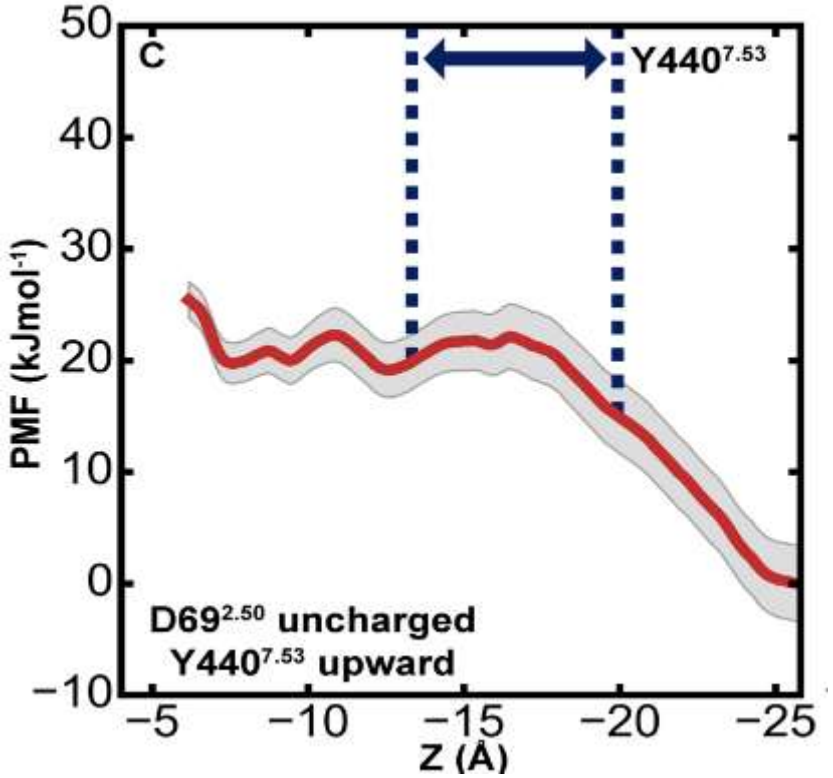




Vickery et al, (2018) *Structure* **26**:171-80

Protonation Changes Na+ Potential along the path

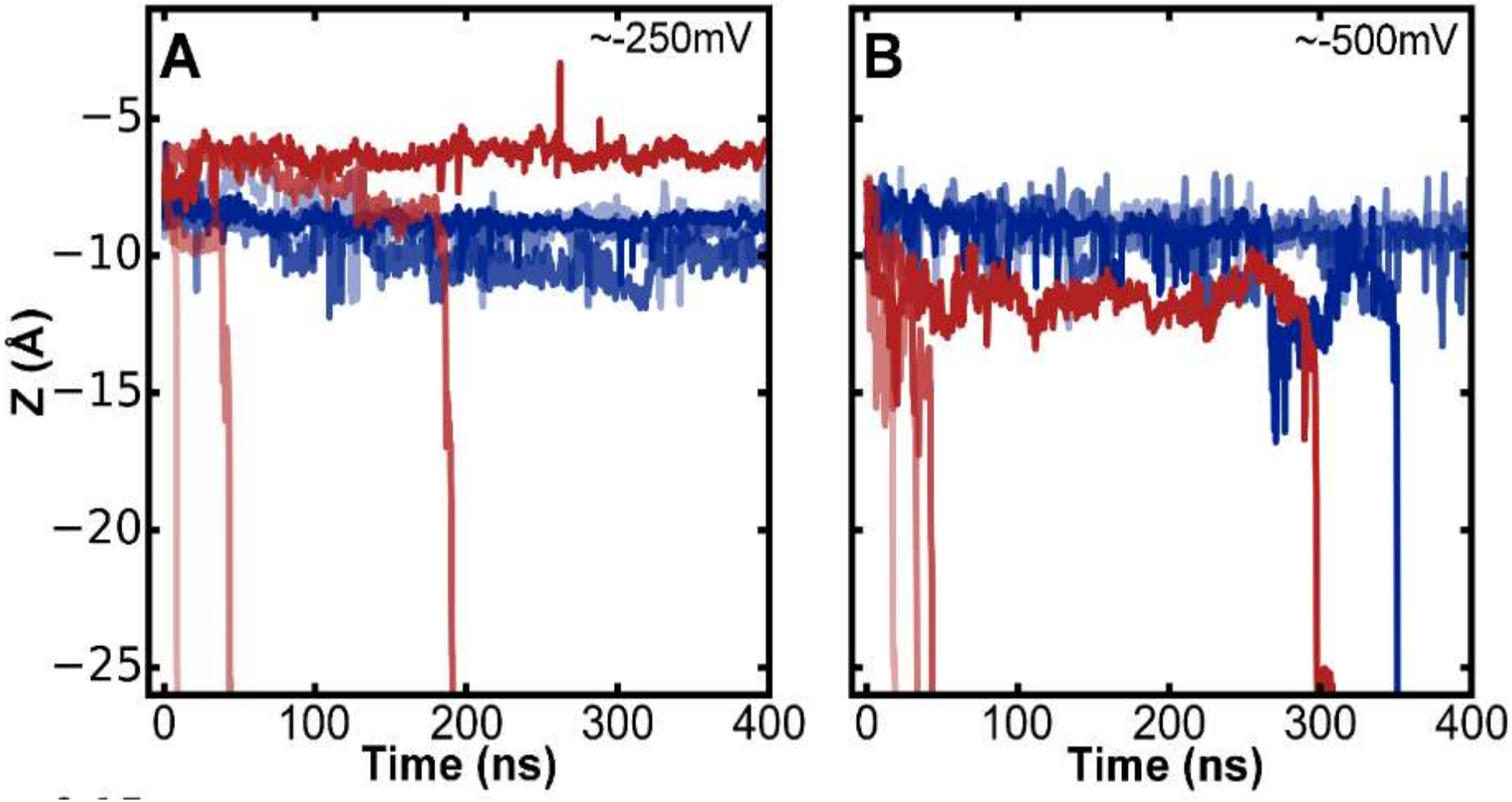




Vickery et al, (2018) *Structure* **26**:171-80

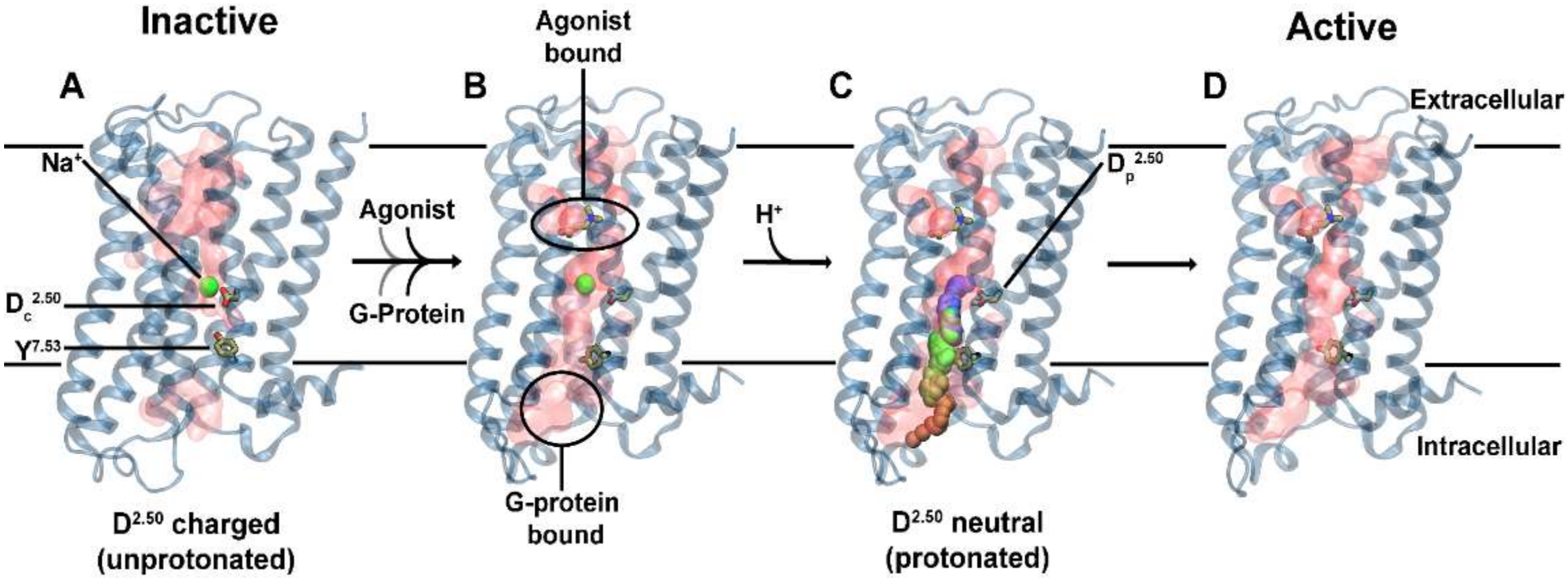
Na^+ ion egress to the intracellular side.

D2.50 protonated
D2.50 charged



Vickery et al, (2018) *Structure* **26**:171-80

Updated Na^+ mechanism



Vickery et al, (2018) *Structure* **26**:171-80

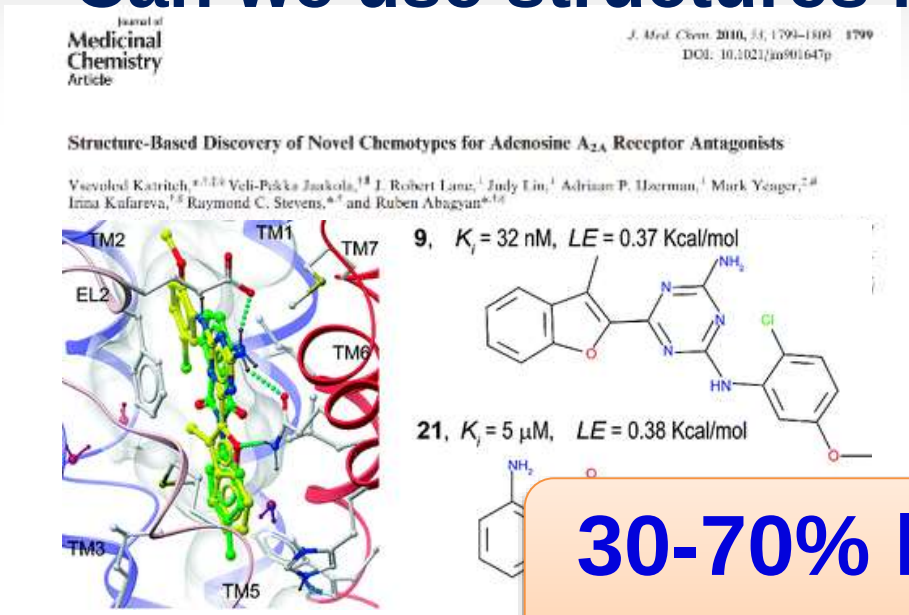
Summary & Outlook : Conserved Na^+ site in GPCRs

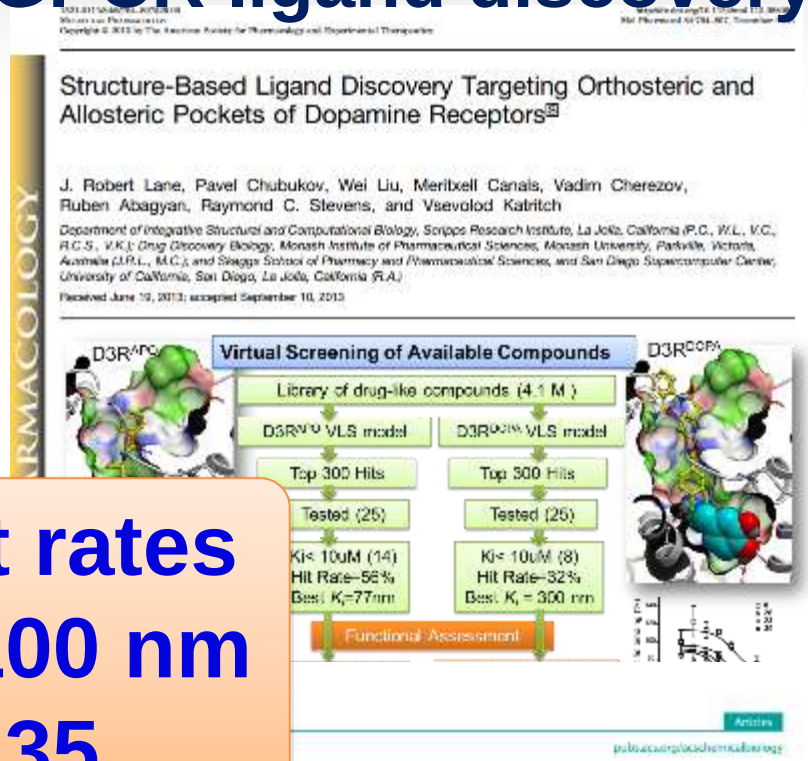
- **Sodium site conserved** in most class A GPCRs
- Sodium coordinating residues and Na^+ itself are involved in GPCR **activation mechanism**
- Sodium can travel through the GPCR “channel” **along the voltage and concentration gradient** (one Na^+ at a time!)
- Sodium transfer may be **coupled with D2.50 protonation**
- Sodium site can be exploited in **ligand discovery** and **receptor stability design**

Outline

- Rational prediction of stabilizing mutations: CompoMug
- New insights into GPCR function and allosteric mechanisms
- **Structure-Based ligand discovery for GPCRs**

Can we use structures in GPCR ligand discovery?

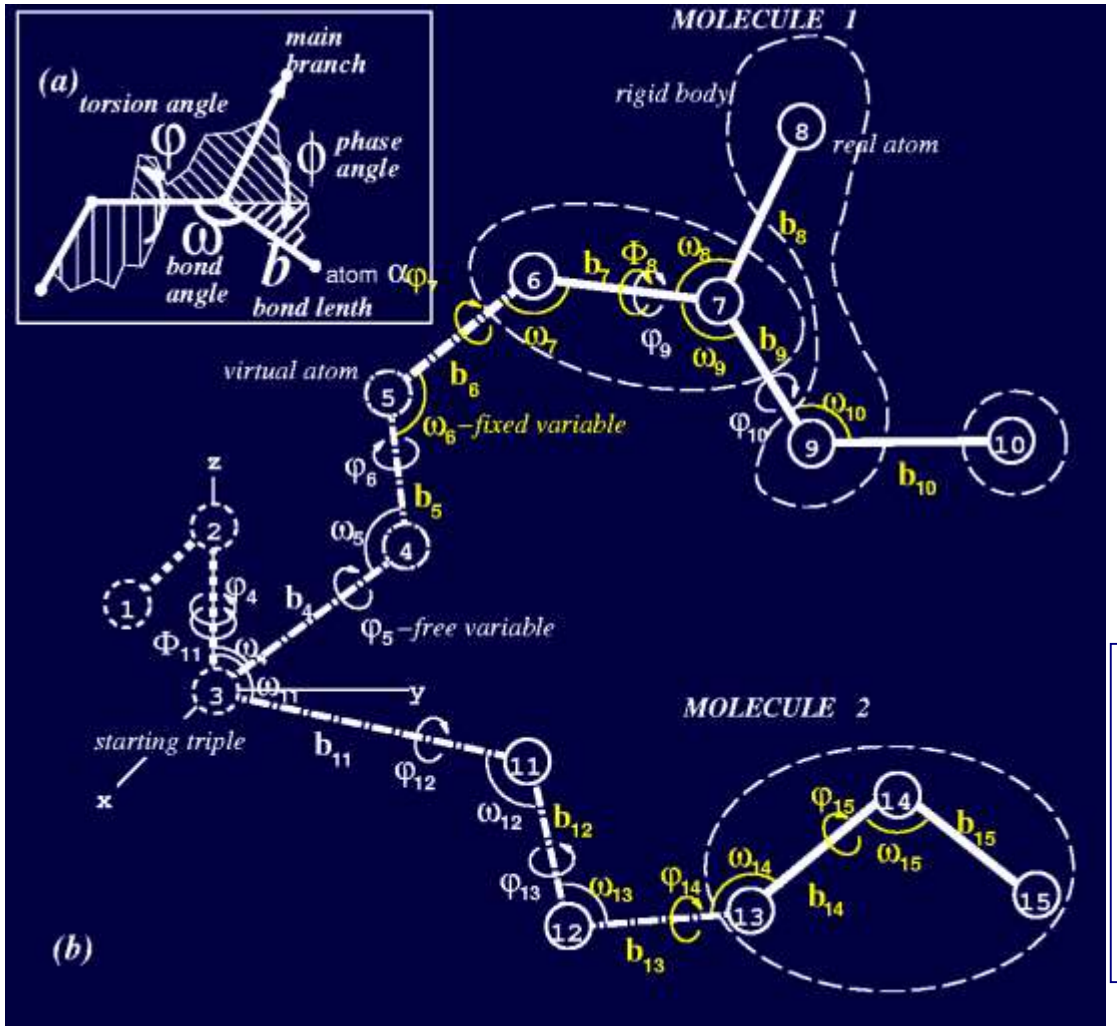




30-70% hit rates
Best K_i < 100 nm
LE > 0.35

Core Modeling Technology: Internal Coordinate Mechanics (ICM)

$$E(\alpha) = \Delta E_{FF} + \Delta E_{EN} + \alpha_1 N_{at} + \alpha_2 \Delta E_{HB} + \alpha_3 \Delta E_{SE} + \alpha_4 \Delta E_{EL} + \alpha_5 \Delta E_{SO}$$



Internal coordinates:

- **+ Eliminates "high frequency" modes**
- **+ Efficient & fast global energy optimization algorithm**
- **+ Large radius of convergence**
- **+ Accurate force field and free energy func (solvation/entropy)**
- **+ Easy to perform sampling of both ligand and receptor conformations**

ICM References:

- **Abagyan et al. (1994)**

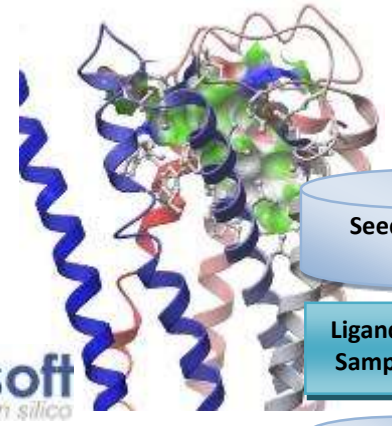
"ICM - a new method for protein modeling.."
***J. Comp. Chem.* 15, 488-506**

- **Katritch et al (2003).**

ICFF: New Internal Coordinate Force Field
J *Comp. Chem.* 24:254-65

Ligand-Guided Model Optimization

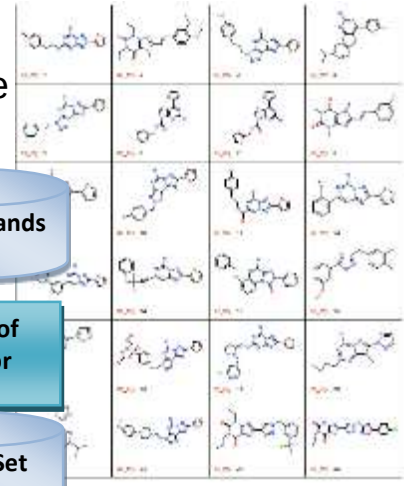
- Improve accuracy and reliability of docking
- Improve VLS performance, especially for lower resolution structures or homology models
- Reduce ligand bias, of shift bias towards desirable ligand/scaffold
- Develop VLS models for different functional states

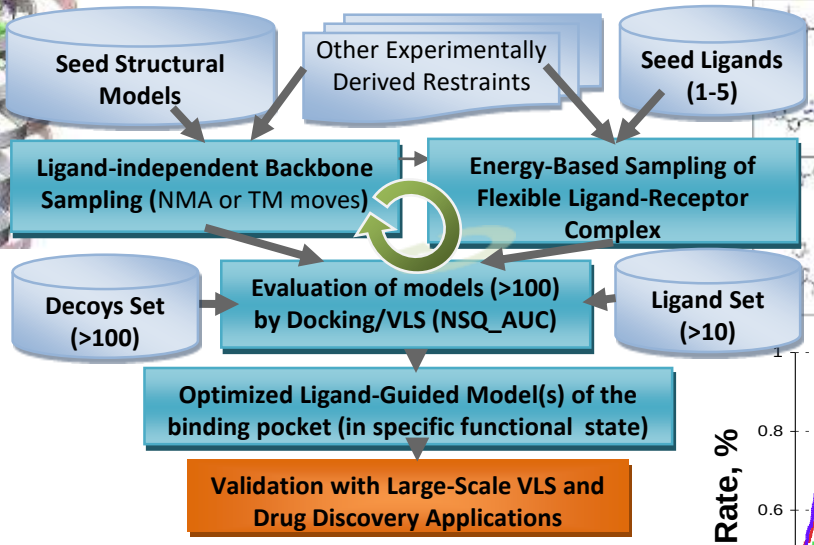


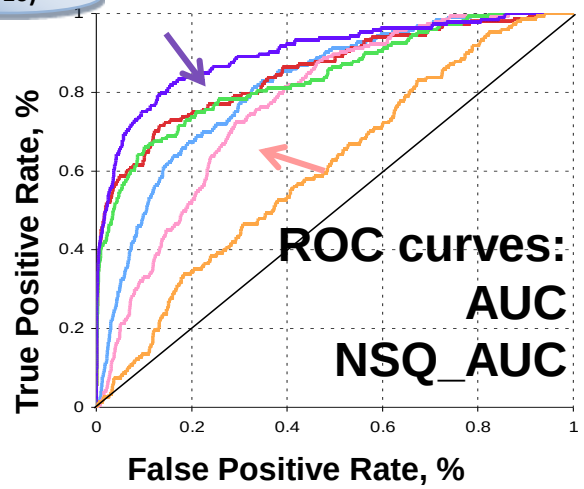
molsoft
molecules in silico

Fast & Reliable molecular modeling platform

LiBERO Algorithm
(Ligand guided Backbone Ensemble Receptor Optimization)







Katritch et al. *Proteins* 2010
Katritch et al. *Methods* 2012

Similar Algorithms:
Pala *et al., J. Chem Inf Model 53, 821-835 (2013) (with Glide)*
Carlsson *et al., Nature Chem Bio 7, 769-778 (2011) (with DOCK).*

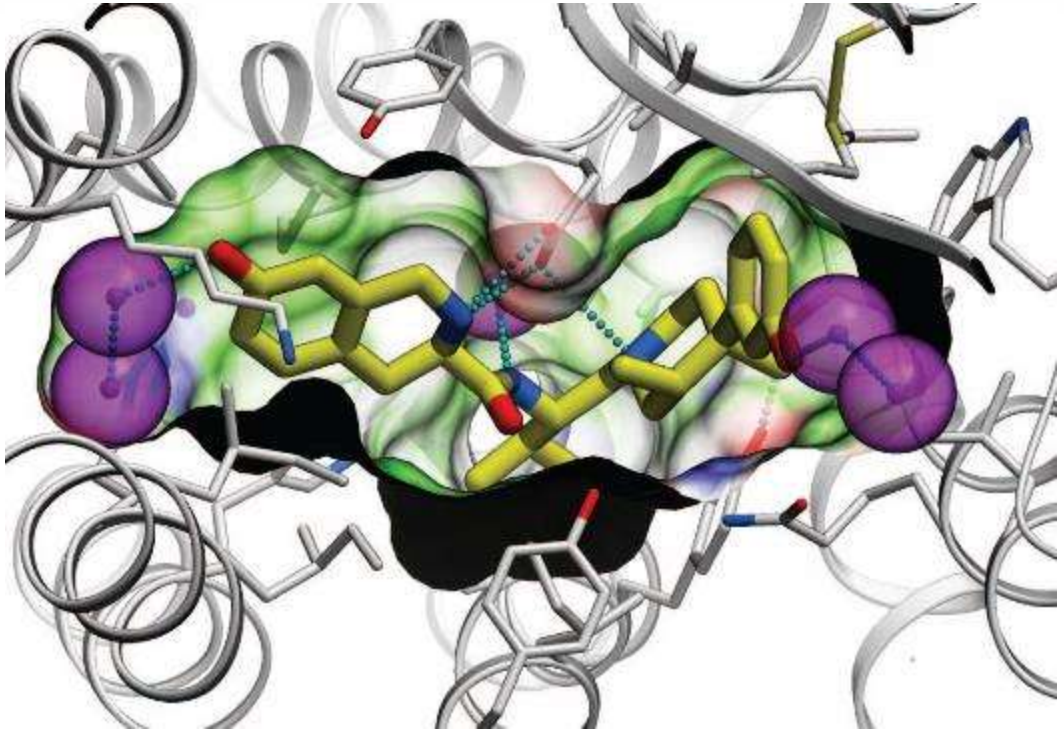
κ-Opioid Receptor as a Therapeutic Target

- κ-OR full agonists (e.g. SalA) lead to hallucination and dysphoria

- κ-OR antagonists: potential antidepressants, anxiolytics and anti-addiction drugs
- JDTic: high affinity selective antagonist, (but cardiac side effects in clinic)

- G-protein biased κ-OR agonists (no β-arrestin signaling): non-addictive analgesia

- **New κ-OR ligand chemotypes needed for both antagonist and biased agonist functional profiles**



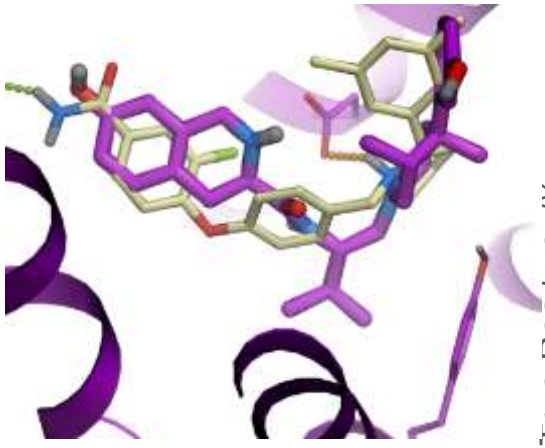
κ-OR structure with JDTic (4DJH)
Wu et al 2012, Nature

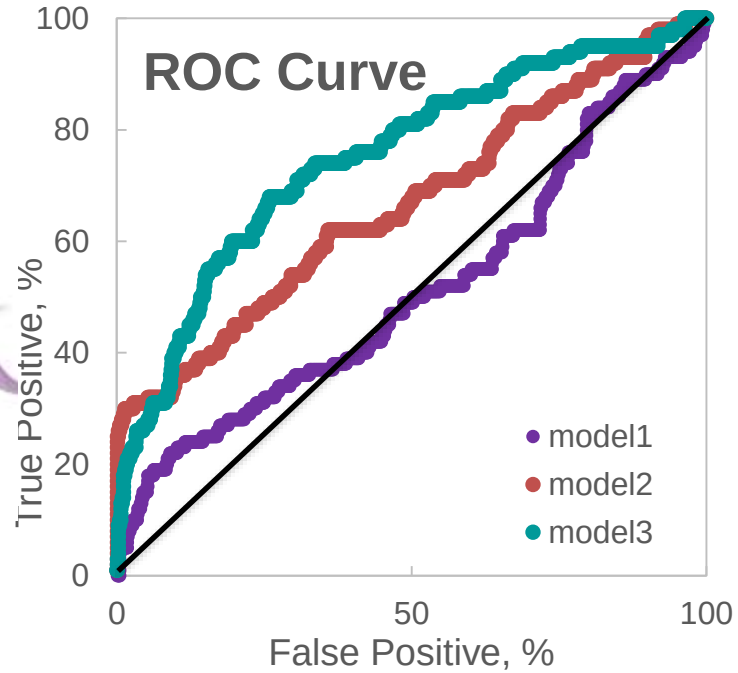
Hard target for VLS – few previously published VLS hits all >**μM** range

Buda et al Neuropsychopharmacology 2015.
Carroll and Carlezon (2015) JMC 56: 2178–2195.

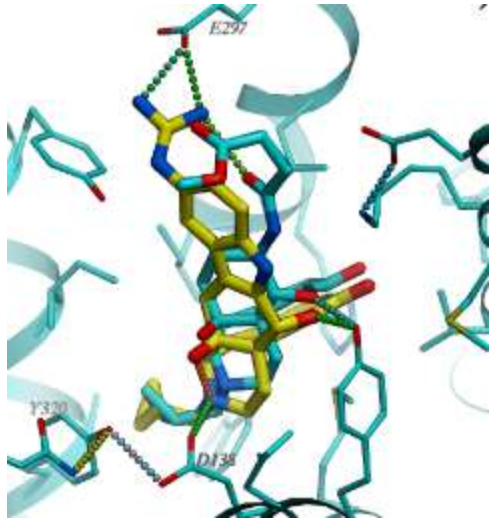
Virtual Ligand Screening for κ-OR: Model Optimization

Model 1
JDTic-bound



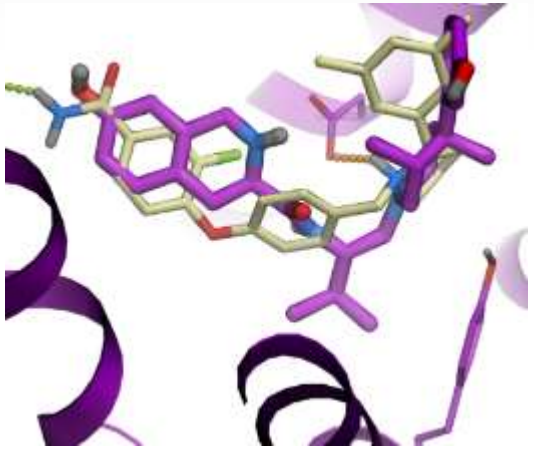


Model 3
Morphinan-bound



Virtual Ligand Screening for κ-OR

New Chemotype Discovery



Three Ligand-Optimized VLS models

Virtual Fragment-like and Lead-like Library (**5M** compounds)

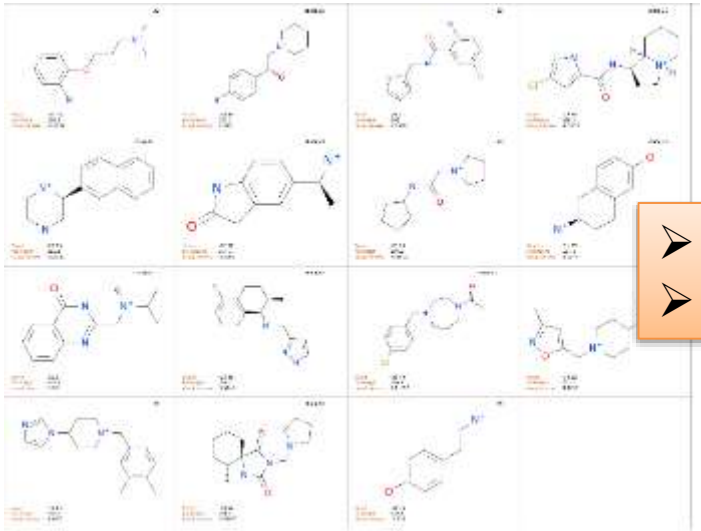
Novel Candidate hits with ΔΔG<30 (**~200**)

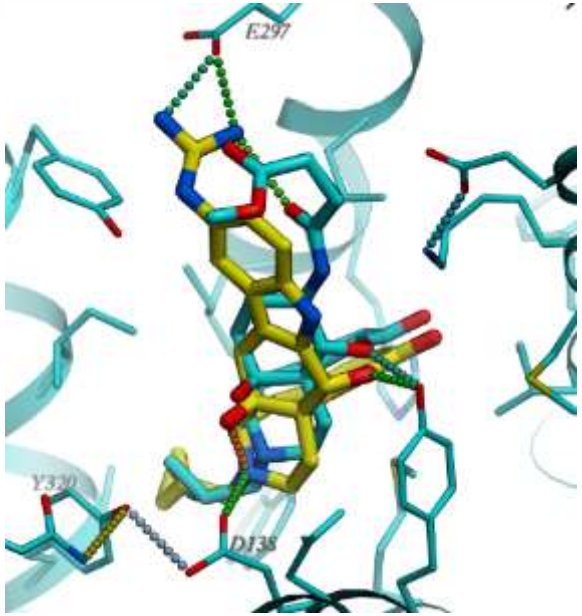
Ordered and tested (**43**)

Ki(κ-OR) < 30μM (**27**)

Ki(κ-OR) < 10μM (**14**)

- Best Ki <200 nm, Ligand Efficiency **LE >0.45**
- Agonists and antagonists, best EC_{50}=260 nm



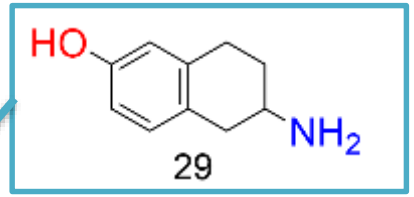


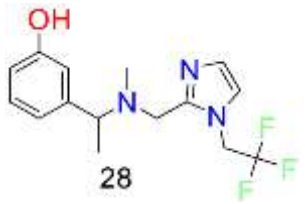
> 6 New Distinct Chemotypes of κ-OR Ligands

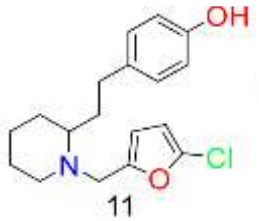
Comp#	Model[a]	Inhibit%[b]	±	SEM[c]	pK_i	±	SEM	K_i(μM)	LE[d]	Tanimoto[e]
28	1	96.9	±	1.2	6.82	±	0.16	0.2	0.43	0.42
11	1	95.3	±	2.5	6.30	±	0.12	0.5	0.40	0.19
21	3	77.5	±	6.8	5.95	±	0.26	1.1	0.31	0.29
43	2	76.5	±	4.7	5.64	±	0.16	2.3	0.32	0.39
40	3	77.5	±	5.6	5.59	±	0.22	2.6	0.36	0.40
6	1	80.1	±	4.2	5.58	±	0.17	2.6	0.39	0.39
17	2	53.0	±	7.8	5.52	±	0.12	3.0	0.32	0.52
16	2	60.7	±	14.7	5.52	±	0.17	3.0	0.37	0.38
20	2	49.1	±	16.4	5.21	±	0.14	6.2	0.30	0.42
35	2	49.8	±	8.4	5.19	±	0.17	6.4	0.36	0.51
3	1	70.2	±	3.9	5.18	±	0.18	6.6	0.36	0.26
8	1	73.7	±	4.0	5.09	±	0.22	8.1	0.29	0.34
29	2	48.6	±	9.6	5.08	±	0.12	8.3	0.59	0.15
22	2	47.3	±	7.7	5.01	±	0.13	9.8	0.28	0.39

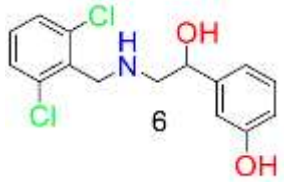
"Similar" to JDTic core

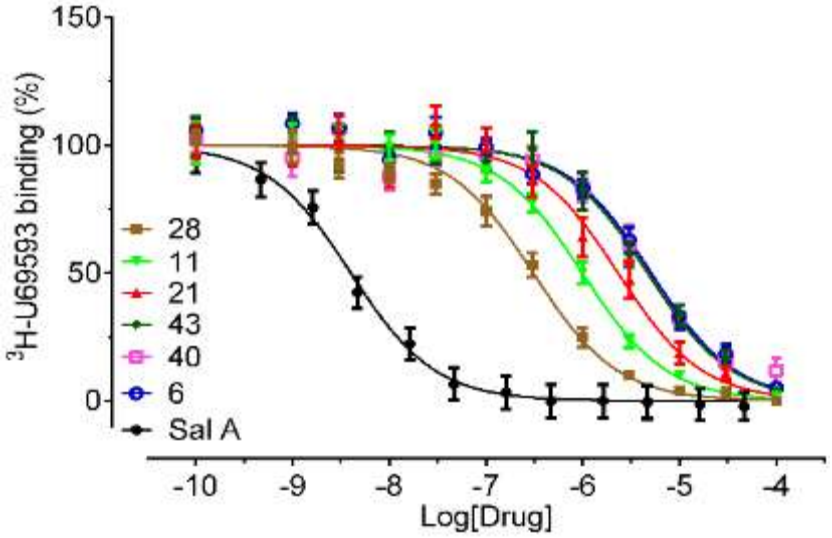
tetrahydronaphthalene vs.
tetrahydroisoquinoline



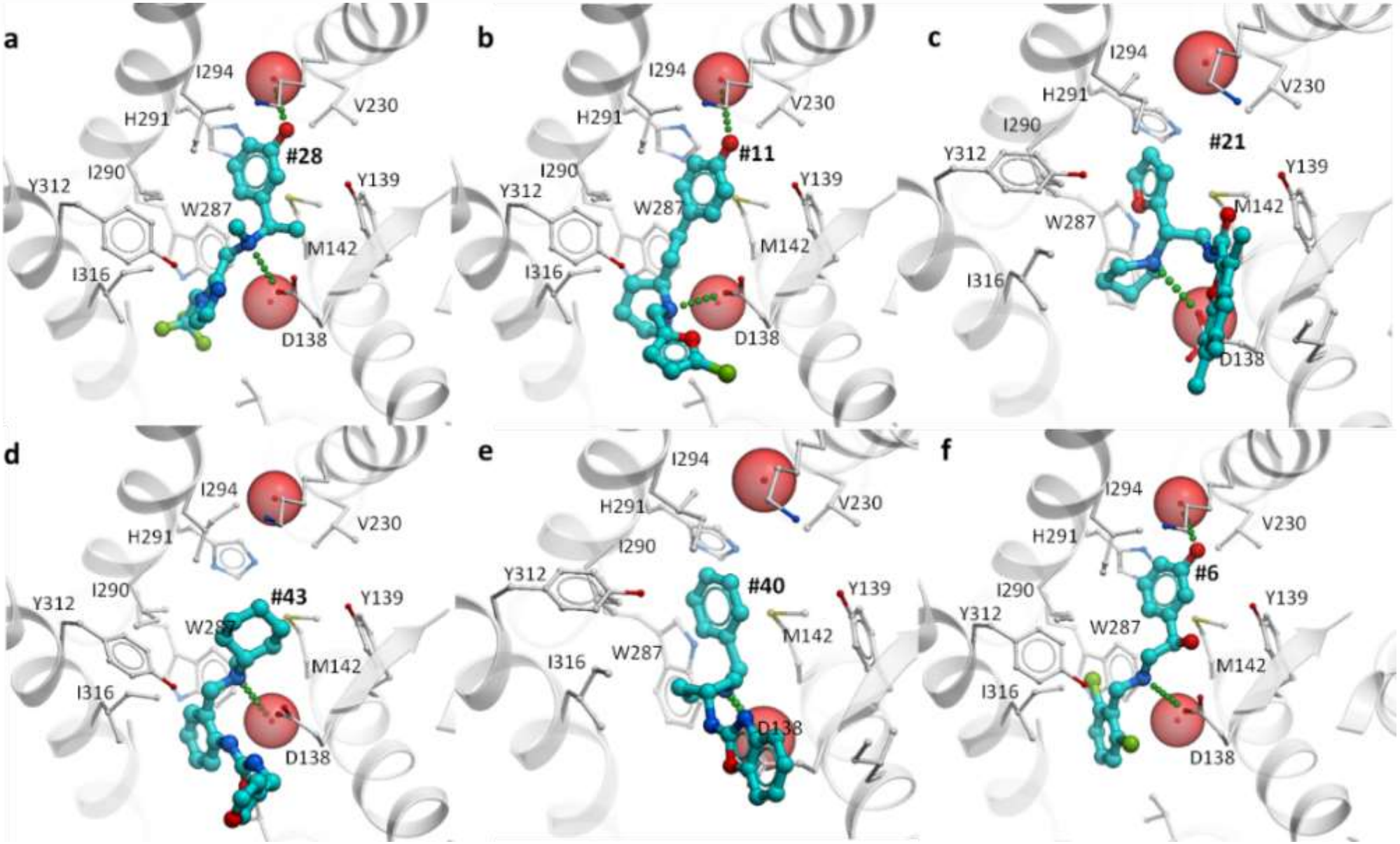




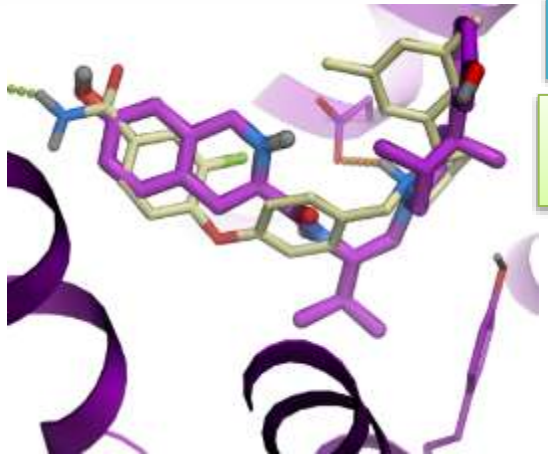




> 6 New Distinct Chemotypes of κ-OR Ligands



Virtual Ligand Screening for κ-OR
First Round of Optimization



Three Ligand-Optimized VLS models

↓

Virtual Fragment-like and Lead-like Library (**5M** compounds)

↓

Novel Candidate hits with ΔΔG<30 (**~200**)

↓

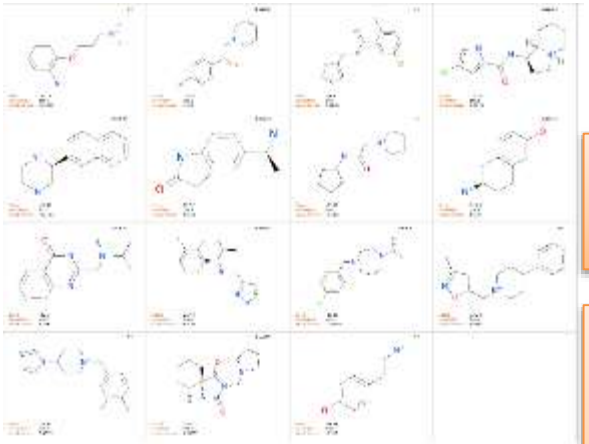
Ordered and tested (**43**)

↓

Ki(κ-OR) < 30μM (**27**)

↓

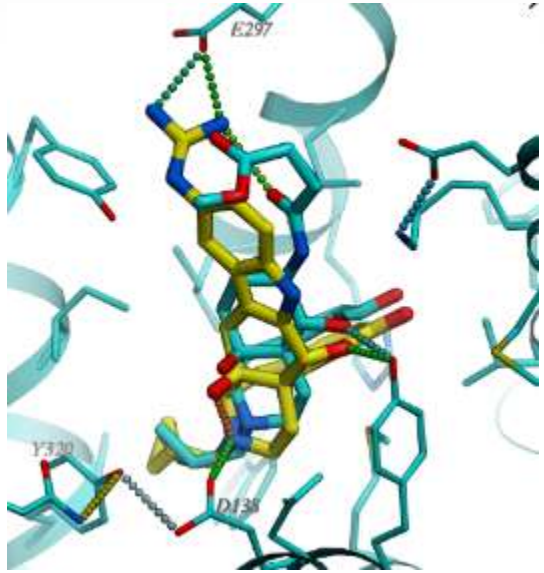
Ki(κ-OR) < 10μM (**14**)

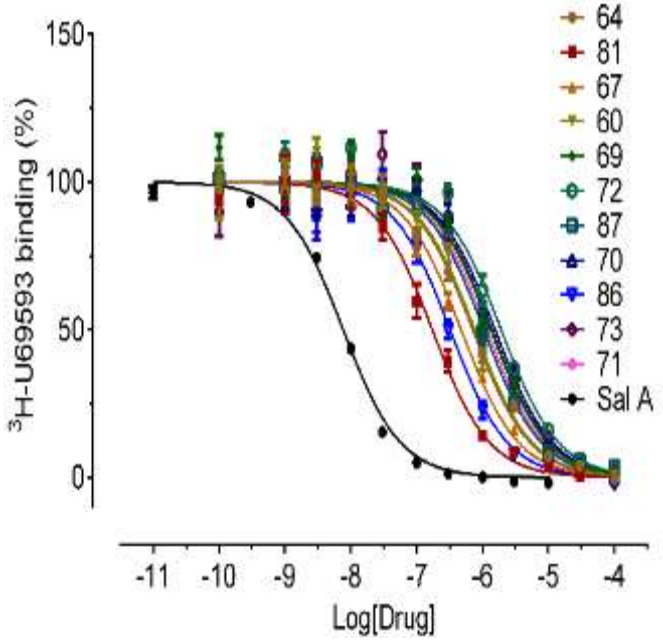


- Best Ki <200 nm, Ligand Efficiency **LE >0.45**
- Agonists and antagonists, best EC_{50}=260 nm

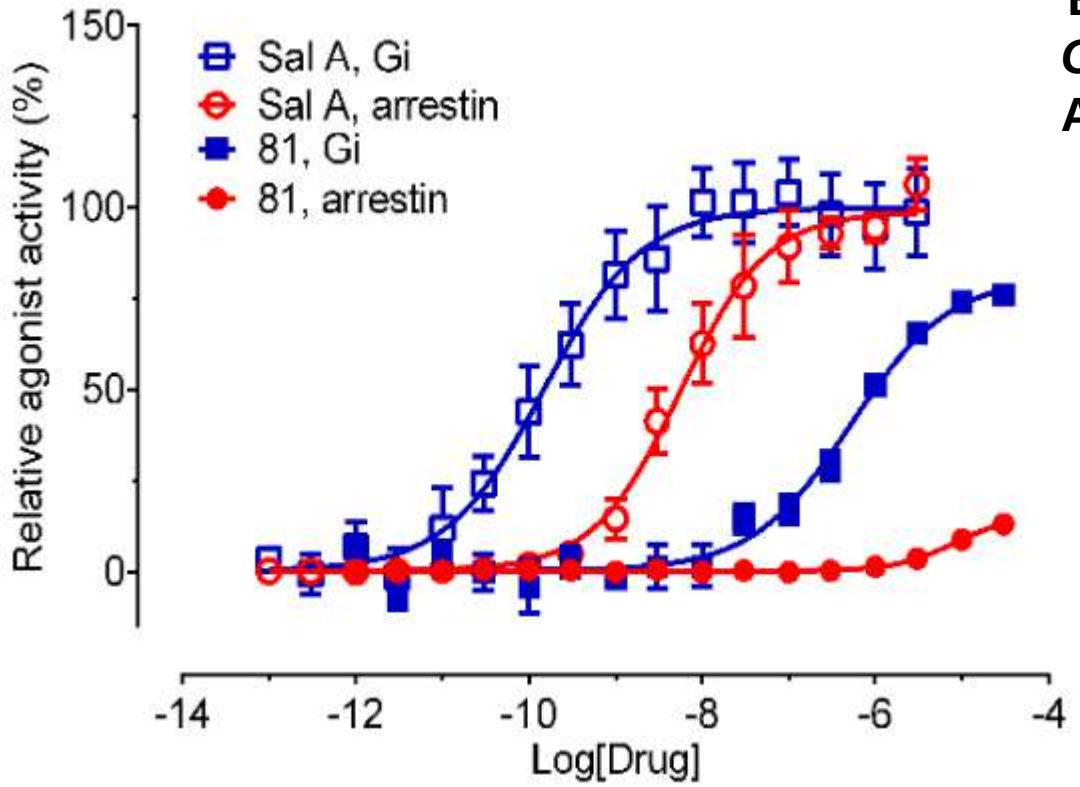
Analog by catalog for the top leads

- 30 compounds tested (in 6 scaffolds)
- **10** sub-micromolar compounds!
- **Best Ki < 90 nm!**
- **Functional Profiles >**





Discovery of a new Biased Agonist



Comp #81:

Binding Affinity K_i **= 160 nM**

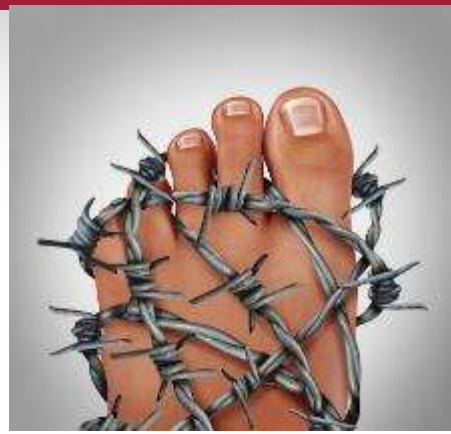
G_i pathway: EC_{50} **= 530 nM**

Arrestin : EC_{50} **> 10000 nM**

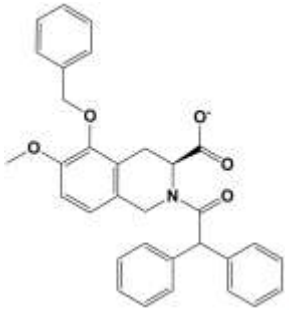
Bias factor: $\Delta\Delta\log(\tau/K_A) > 6.0$
(G_i vs arrestin pathways)

Ongoing optimizations of best scaffolds

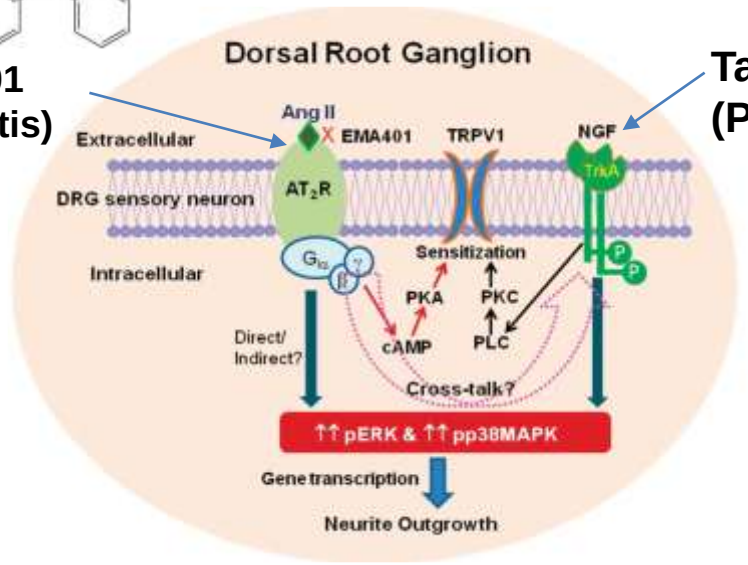
Angiotensin AT_2R antagonists for neuropathic pain relief



- AT_2R regulates nerve sensitization and neurite outgrowth in neuropathic pain
- Blocking AT_2R in PNS can reduce the pain
- **EMA401** is in Phase II trials for post-herpetic neuropathic pain (Novartis)

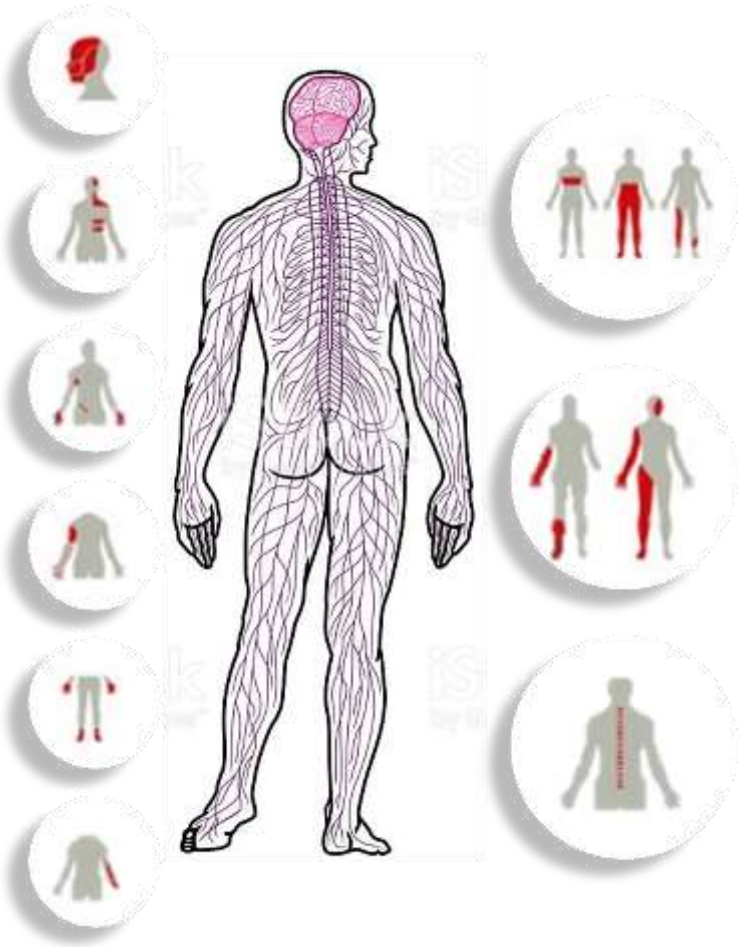


EMA401 (Novartis)



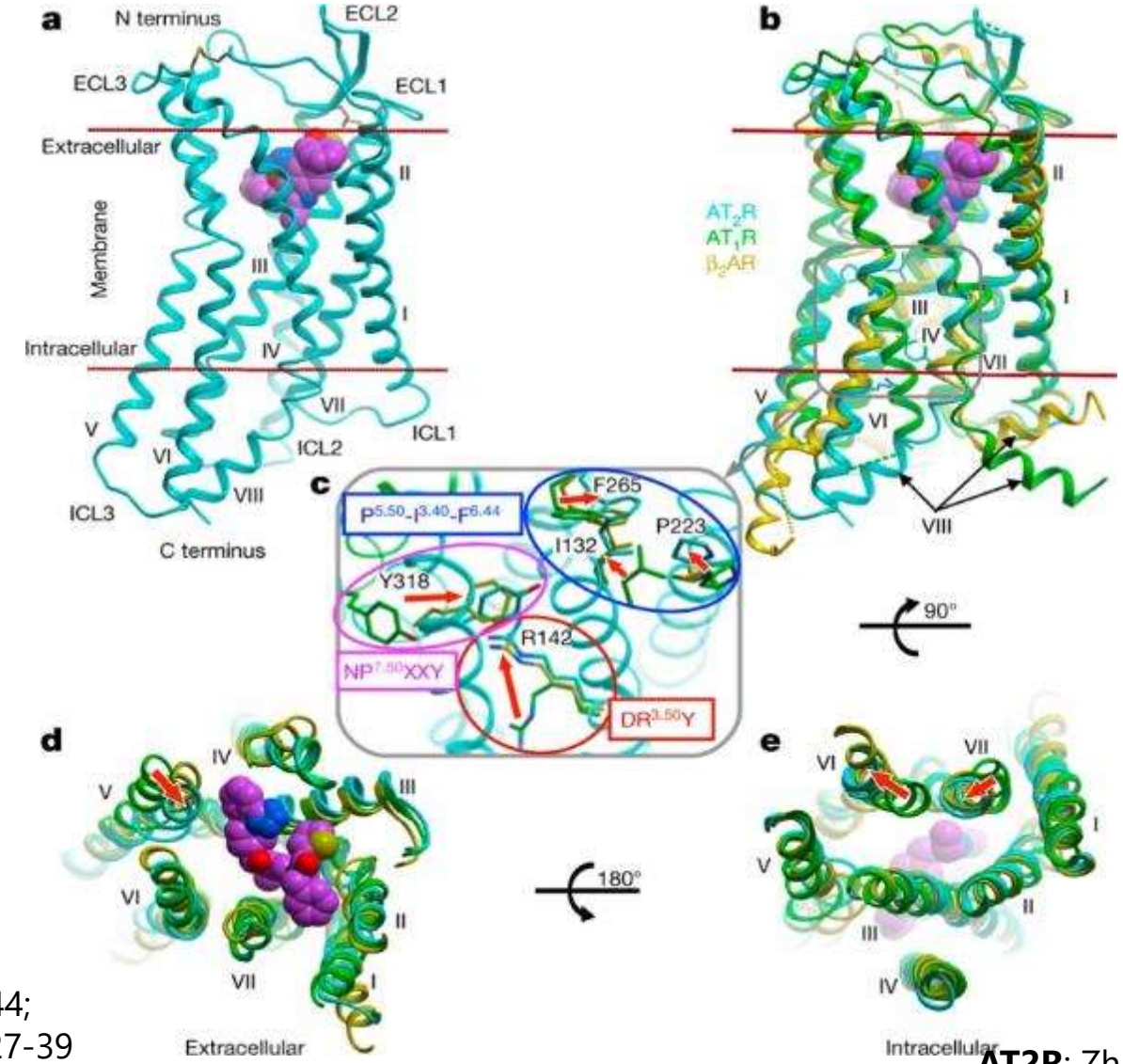
Tanezumab (Pfizer)

Smith et al, *Pain* 2016



Also AT_2R **agonists** (C21) confer neuroprotection in CNS (stroke)

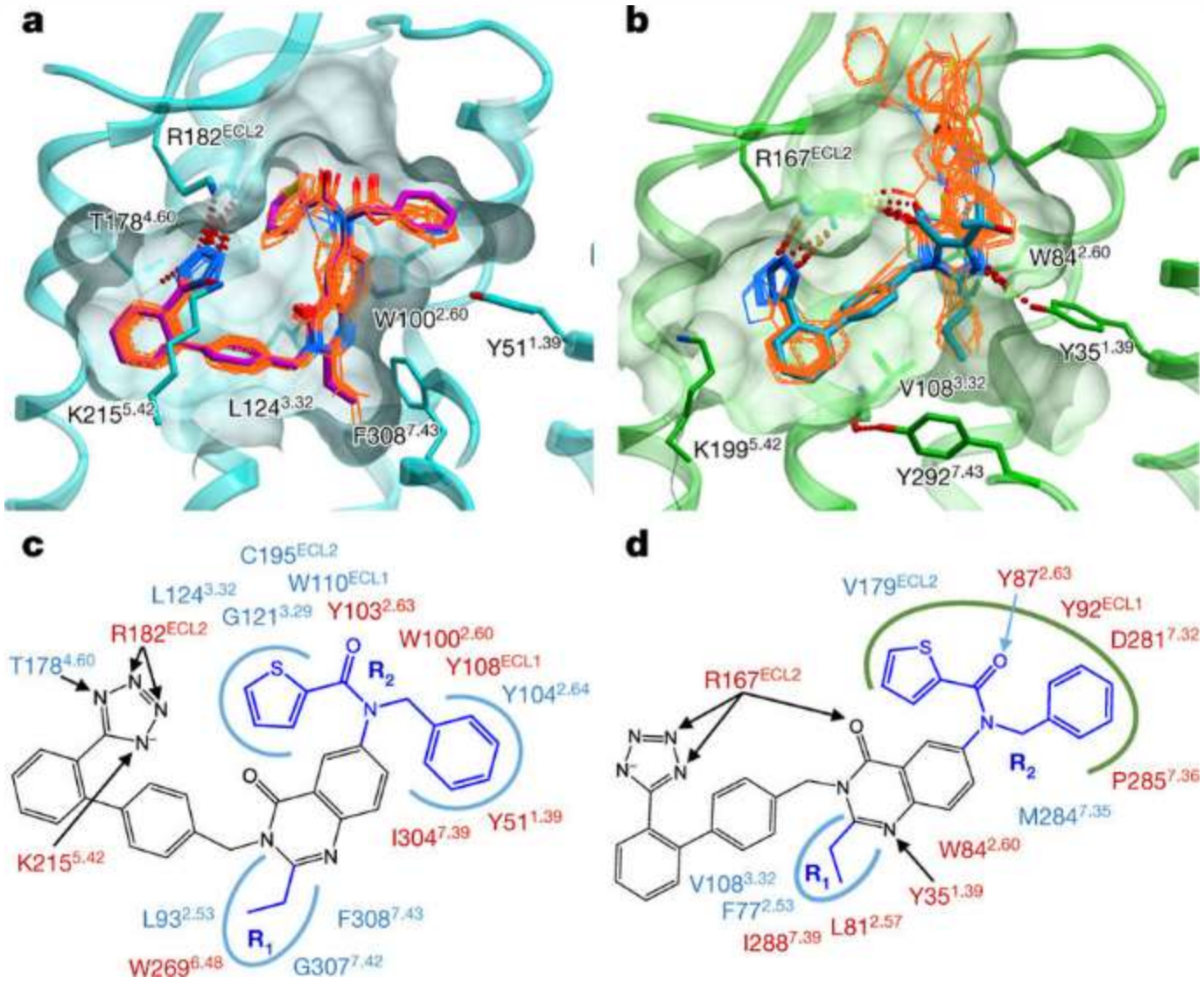
Crystal Structures of AT_1R and AT_2R give new insights into function …



AT1R: Zhang H, *Cell*. 2015;161(4):833-44;
Zhang H, *JBC* 2015;290(49):29127-39

AT2R: Zhang H, *Nature*. 2017;544(7650):327-32;

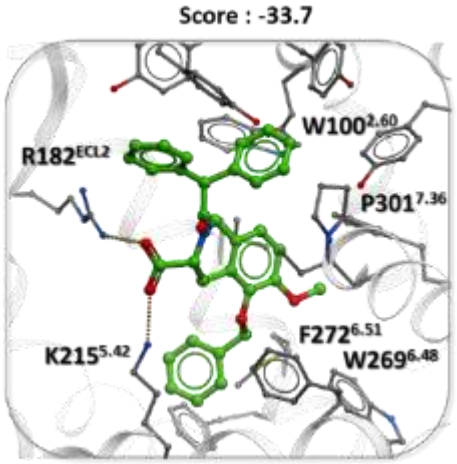
… and AT_1R vs. AT_2R selectivity



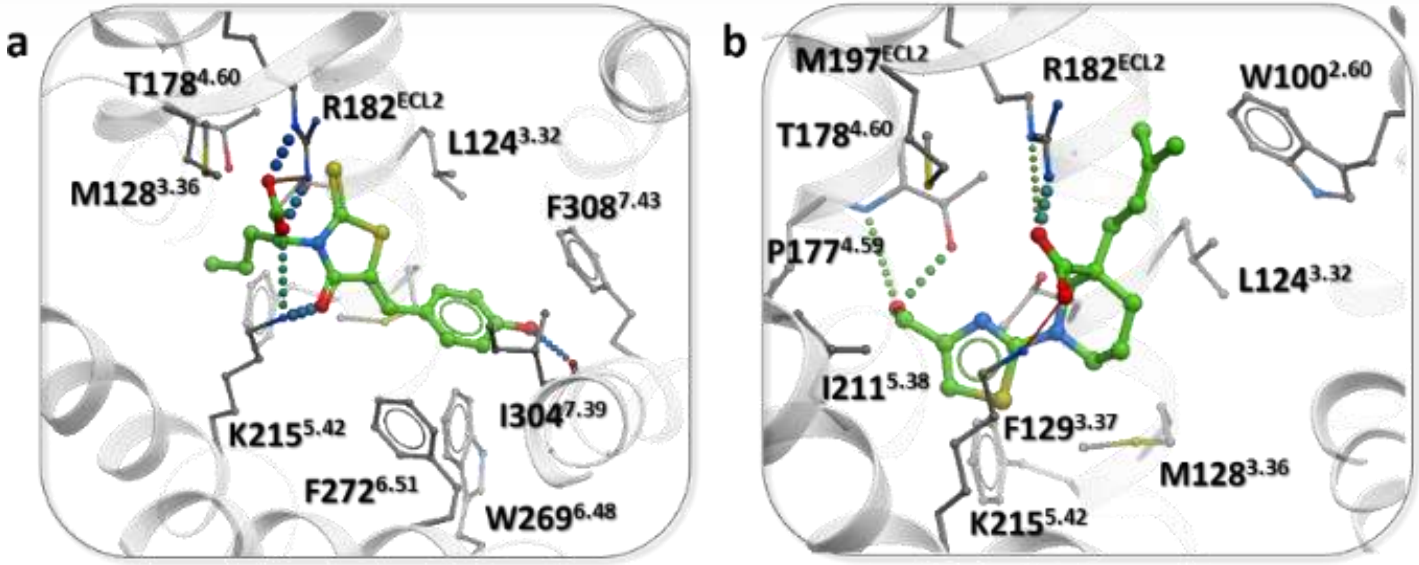
AT2R: Zhang H, *Nature*. 2017;544(7650):327-32;

Lead-like AT_2R selective hits from initial VLS

Model tested and refined with EMA401



- **1M** • Lead Like Compounds
- **5K** • Top rank analyzed for selectivity
- **52** • Purchased and tested
- **6** • Initial Validated Hits



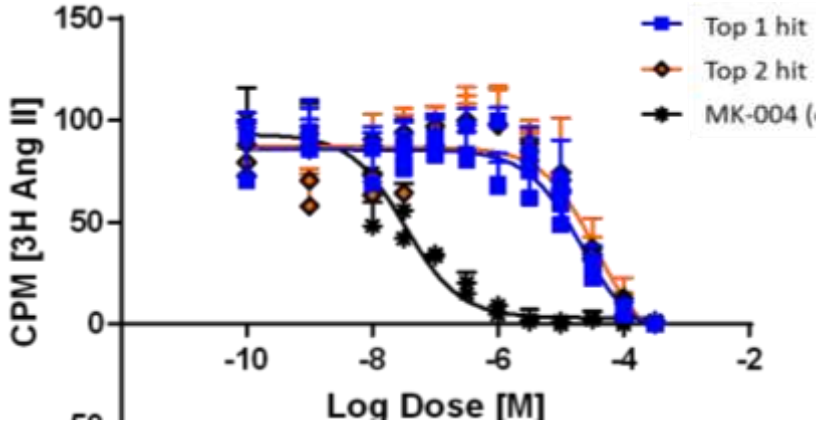


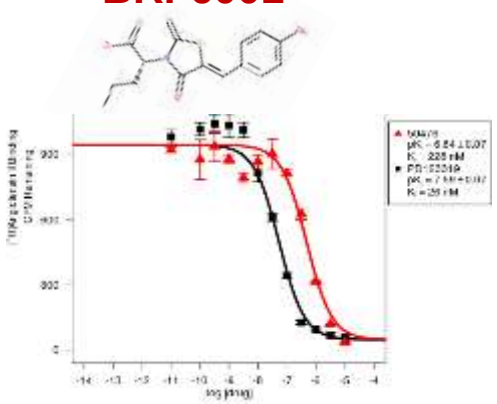
Fig.10 Radioligand competition binding assay for the MK-4 and top two hit compounds.

Zarzycka, Sadybekov et al, manuscript in preparation

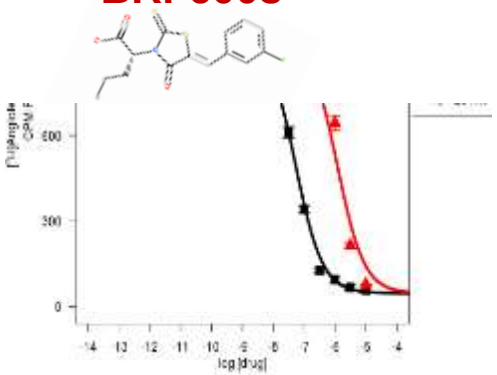
New SAR Results

Confidential

BRI-6001



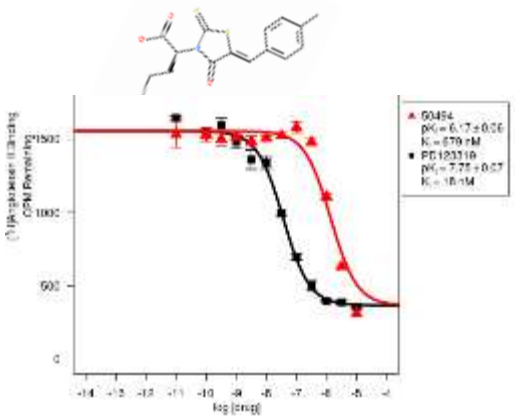
BRI-6003



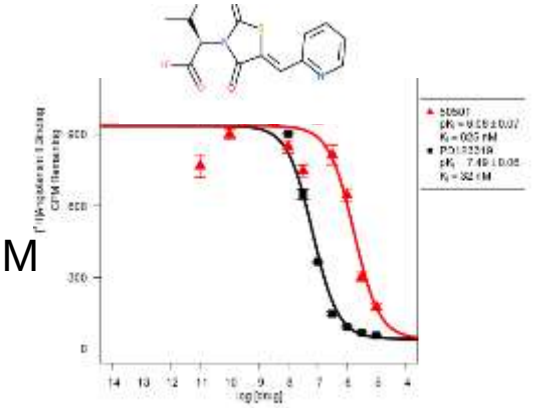
Compound # (Name)	*pKi*±SEM	Ki, nM	LE
50476 (BRI-6001)	**6.64±0.07**	**228**	**0.42**
50486 (BRI-6103)	**6.24±0.07**	**581**	**0.40**
50491 (BRI-6108)	5.65±0.07	2254	0.36
50492 (BRI-6109)	5.43±0.08	3739	0.35
50494 (BRI-6111)	**6.17±0.06**	**680**	**0.39**
50501 (BRI-6118)	**6.08±0.07**	**825**	**0.41**
50502 (BRI-6119)	5.25±0.09	5681	0.32
50477 (BRI-6002)	5.67±0.08	2158	0.38
50533 (EMA401)	**7.60±0.04**	**25**	**0.28**
50534 (C21)	7.44±0.04	36	0.33

Ligand Efficiency, LE = $1.4*pK_i/N_{atom}$

BRI-6011



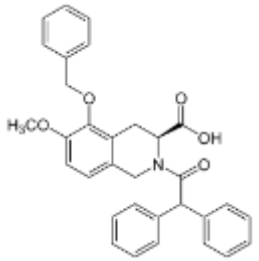
BRI-6018



- Four sub-micromolar hits
- Very high Ligand Efficiency (**LE**)
- AT_2R selective: no AT_1R binding at 30 uM
- Amenable for lead optimization

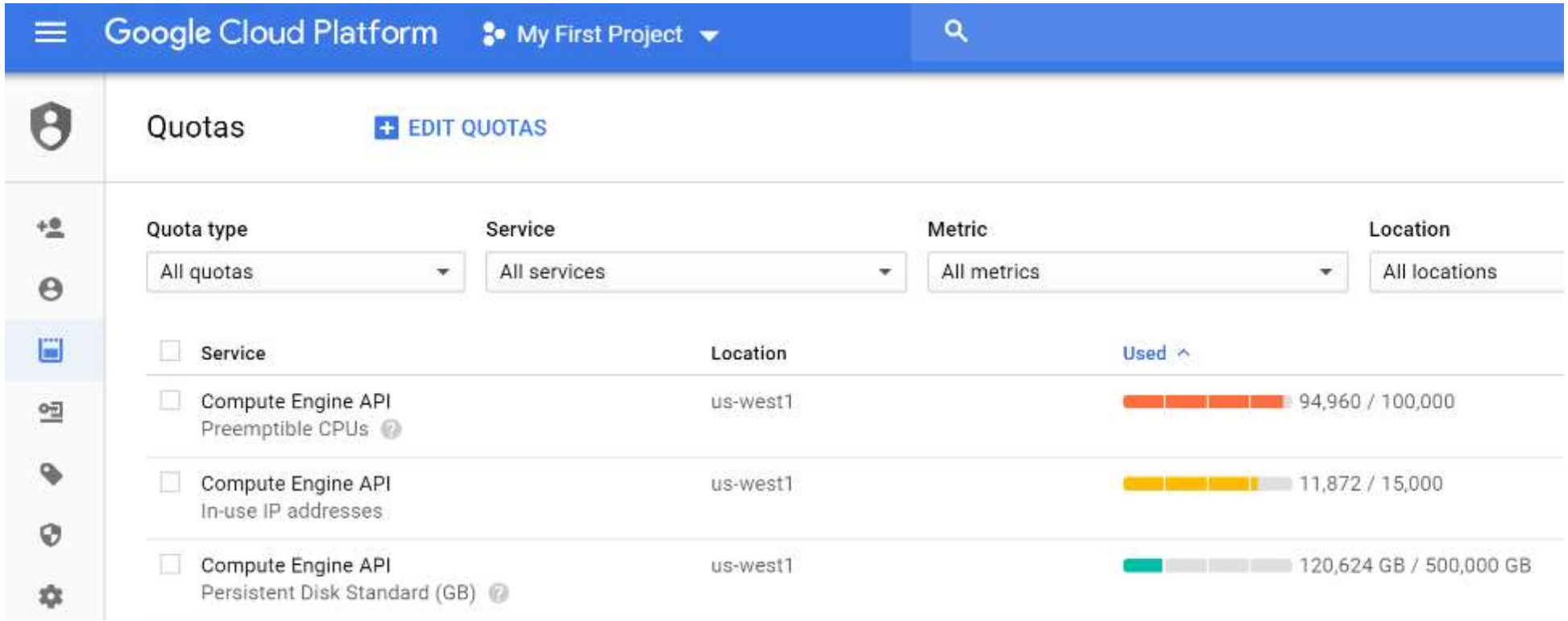
Control: PD123319 (**EMA401**)

N_{atom} =38



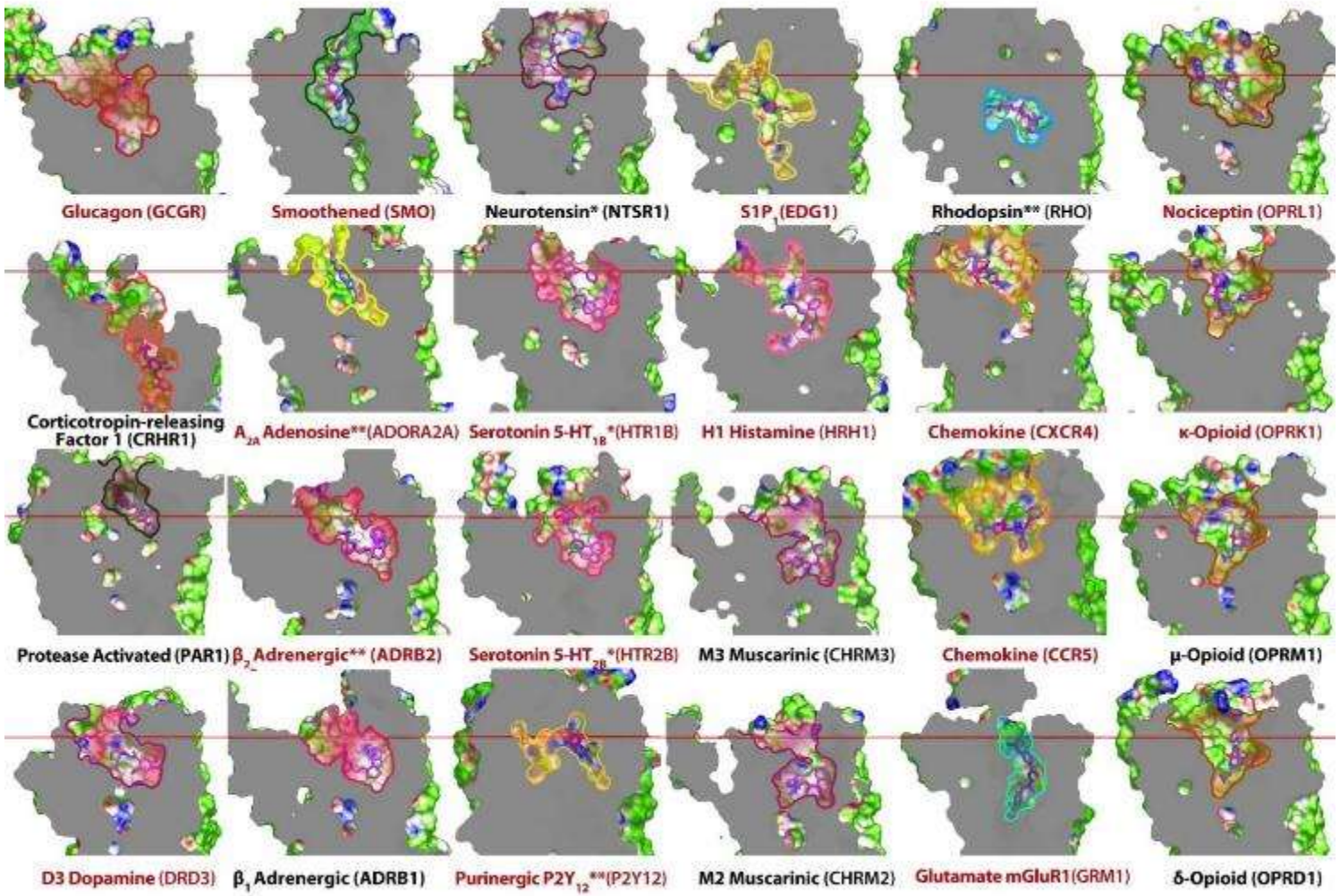
Google Cloud screening of 500M compounds (NIH Commons Pilot Project)

- Used new ZINC library of “Wait-OK” compounds (4-6 weeks, 80% success)
- Google Cloud screening used ~100,000 CPUs at once
- Completed in 24 hours
- Would take ~5 month on our servers



Other Hot GPCR Targets for Ligand Discovery

>50 GPCRs with structures solved and ~10 more coming this year…



* Solved in agonist-bound partially activated state

** Both agonist-bound and antagonist-bound structures known

Summary & Outlook : Structure-based VLS

- Structure-based VLS allows effective and fast discovery of novel chemotypes for GPCR targets
- High hit rates (20-40%), sub-uM affinities, high Ligand Efficiencies
- Subtype- and functional selectivity can be improved via SAR and lead optimization
- Targeting new allosteric pockets (e.g. sodium site)

USC University of Southern California

The Bridge @ USC

Thanks!

GPCR Consortium

NIH National Institute on Drug Abuse — The Science of Drug Abuse & Addiction

R33 DA038858; P01 DA035764
R21 EY027620 (BRAIN Initiative)

MIPT

Ray Stevens & Vadim Cherezov

Computational Group


Nilkanth Patel


Anastasiia Sadybekov


Barabara Zarzycka, PhD.


Saheem Zaidi, PhD


Jessica Grandner, PhD.


Petr Popov, PhD (+MIPT)


Arman Sadybekov, PhD


Natalie Binczewski

Collaborators:

- Bryan Roth (*UNC*)
- *James Zhi-Jie (iHuman)*
- Beili Wu & Quang Zhao *(SIMM)*
- Uli Zachariae (U. Dundee)
- Ken Jacobson *(NIH/NIDDK)*
- *Patrick Giguere (Ottawa U)*
- *Rob Lane (Monash U)*
- Ivy Carroll (*RTI*)
- *Irene Coin (Leipzig U)*
- *Dirk Trauner (NYU)*
- *Ruben Abagyan (UCSD/Molsoft)*

Google Cloud

molsoft molecules in silico