实用生物信息技术期末总结交流报告会

分子对接应用实例
Application Examples of Molecular Docking

2020年12月12日

报告人：孙皓然 刘孝贺
中国农业科学院研究生院
2020级博士班
孙：solomoncat@163.com
刘：liuxiaoheaye@163.com

小组成员

编号	姓名	研究所	导师	研究方向
G5A	孙皓然	特产所	李光玉	小型食肉动物营养与健康
G5B	刘理想	特产所	许保增	卵母细胞减数分裂相关调控
G5C	邵　静	特产所	许保增	特种经济动物遗传育种与繁殖
G5D	刘孝贺	植保所	张永军	农业昆虫与害虫防治

目录

Part1 分子对接原理简介

1 分子对接的概念
2 分子对接的原理
3 分子对接方法的分类
4 分子对接常用工具

Part1 分子对接原理简介

1 分子对接的概念

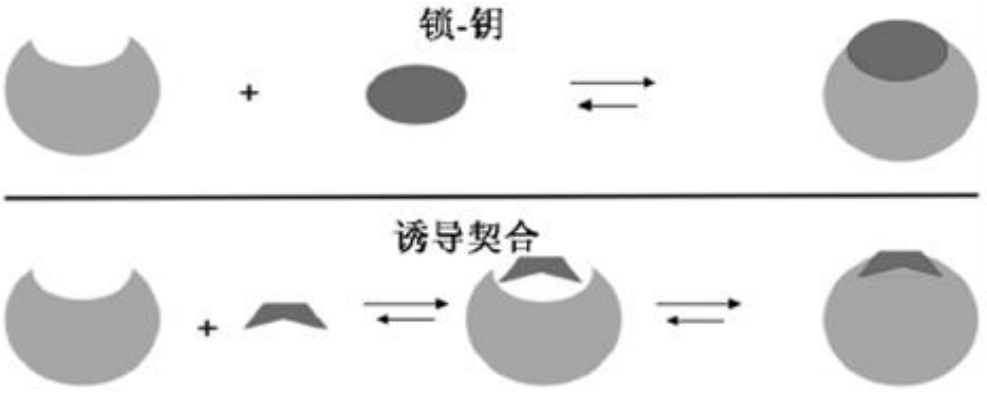
分子对接（Molecular Docking）是依据配体与受体作用的“锁-钥原理”（lock & key principle），模拟配体与受体相互作用的一种分子模拟方法。

配体或受体可以是大分子（蛋白质、DAN、RNA等），也可以是小分子（药物分子、气味分子等）。

Part1 分子对接原理简介

2 分子对接的原理——理论基础

锁-钥模型
1890年E.Fischer 提出



诱导契合模型
1958年D.E.Koshland提出

Part1 分子对接原理简介

2 分子对接的原理——对接过程考虑的因素

形状互补

亲疏水性

表面电荷分布

Part1 分子对接原理简介

3 分子对接方法的分类

方法	研究对象	适用范围	特点
刚性对接	研究体系的构象不发生变化	蛋白质和蛋白质以及蛋白质和核酸	计算较为粗糙，原理相对简单
半柔性对接	允许在一定范围内变化	小分子（柔性）和大分子（刚性）	药物分子筛选常用方法
柔性对接	对接过程中基于分子力学和分子动力学的分子对接方法	精确考察分子之间的识别情况	计算量非常大，耗时较多，得到的对接精度较高

Part1 分子对接原理简介

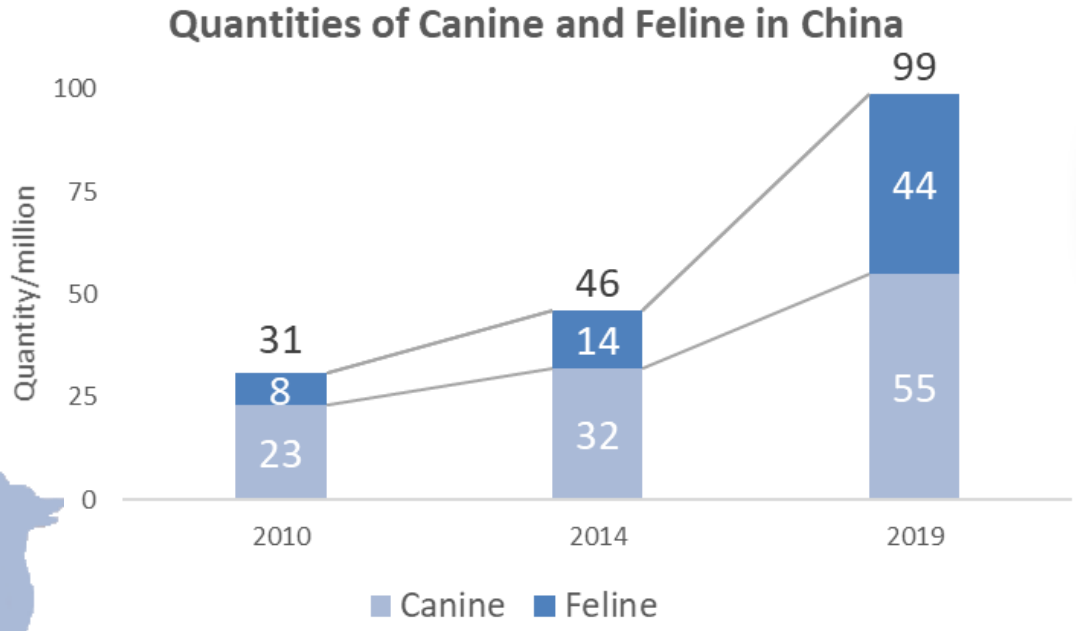
4 分子对接常用工具

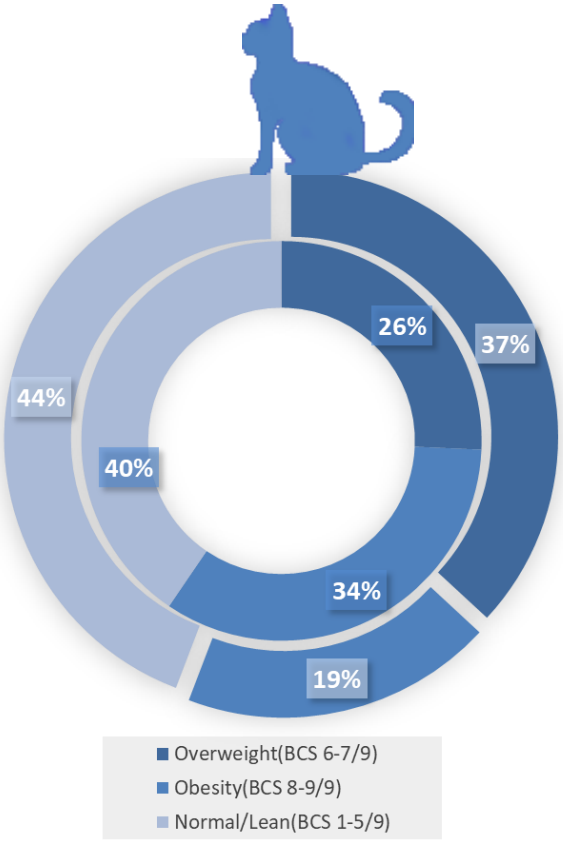
名称	构象搜索方法	结合评价方法	速度
Flex X (Sybyl)	片段生长法	半经验自由能	快
LigandFit(Cerius2)	蒙地卡罗模拟	半经验自由能	快
Glide (薛定谔软件)	系统搜索	半经验自由能	一般
Gold	遗传算法	半经验自由能	快
Affinity （InsightII)	蒙地卡罗/MM/MD	分子力场	慢
AutoDock	遗传算法	半经验自由能	一般
Dock	片段生长法	分子力场	快
ICM-Dock	随机全局优化	半经验自由能	快
Fred (openeye)	系统搜索	半经验自由能	快

Part2 蛋白质-蛋白质分子对接实例演示
——瘦素与瘦素受体的分子对接初探

1 研究背景
2 瘦素与瘦素受体的结构
3 瘦素与瘦素受体的分子对接
4 分子对接结果评价

1 研究背景

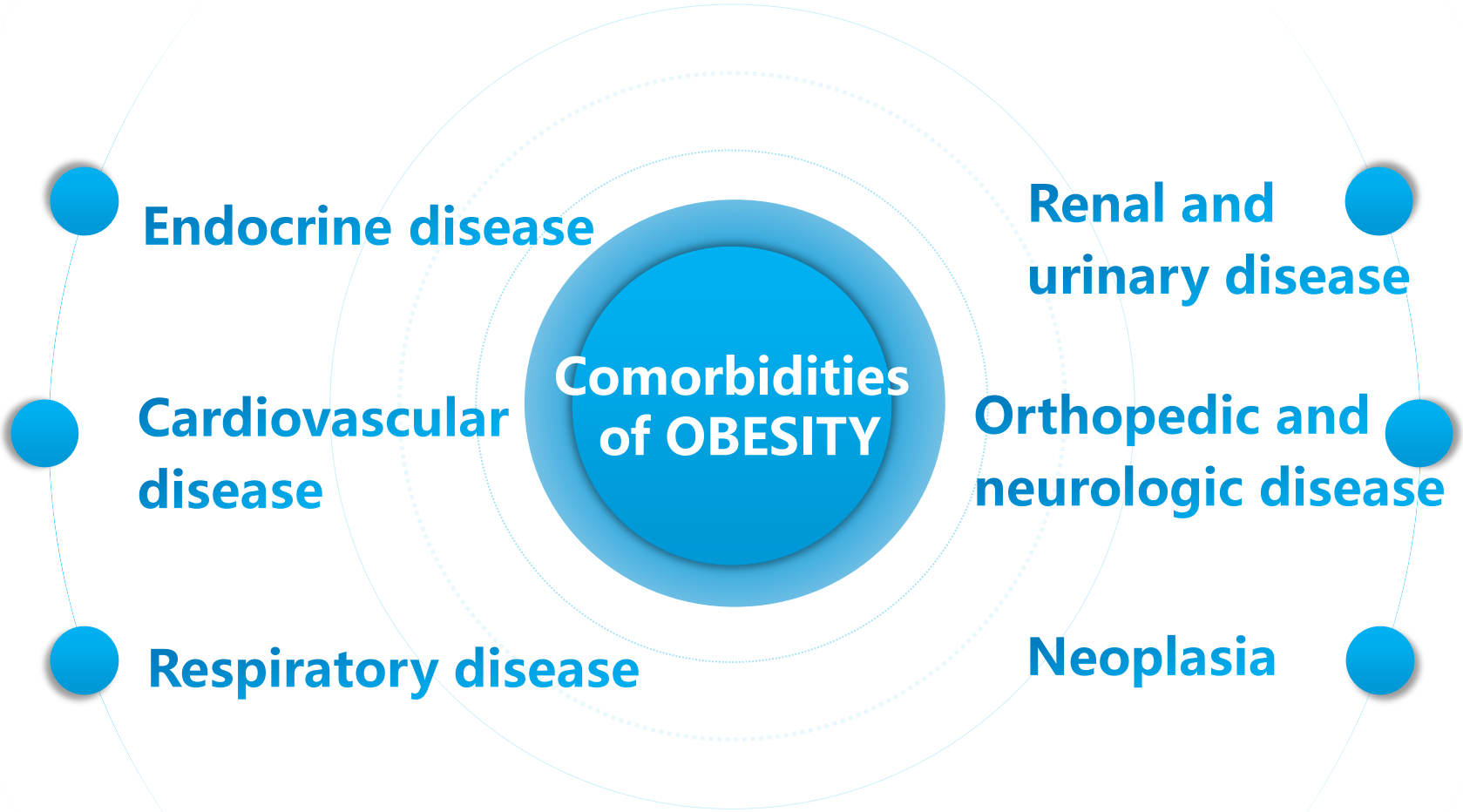




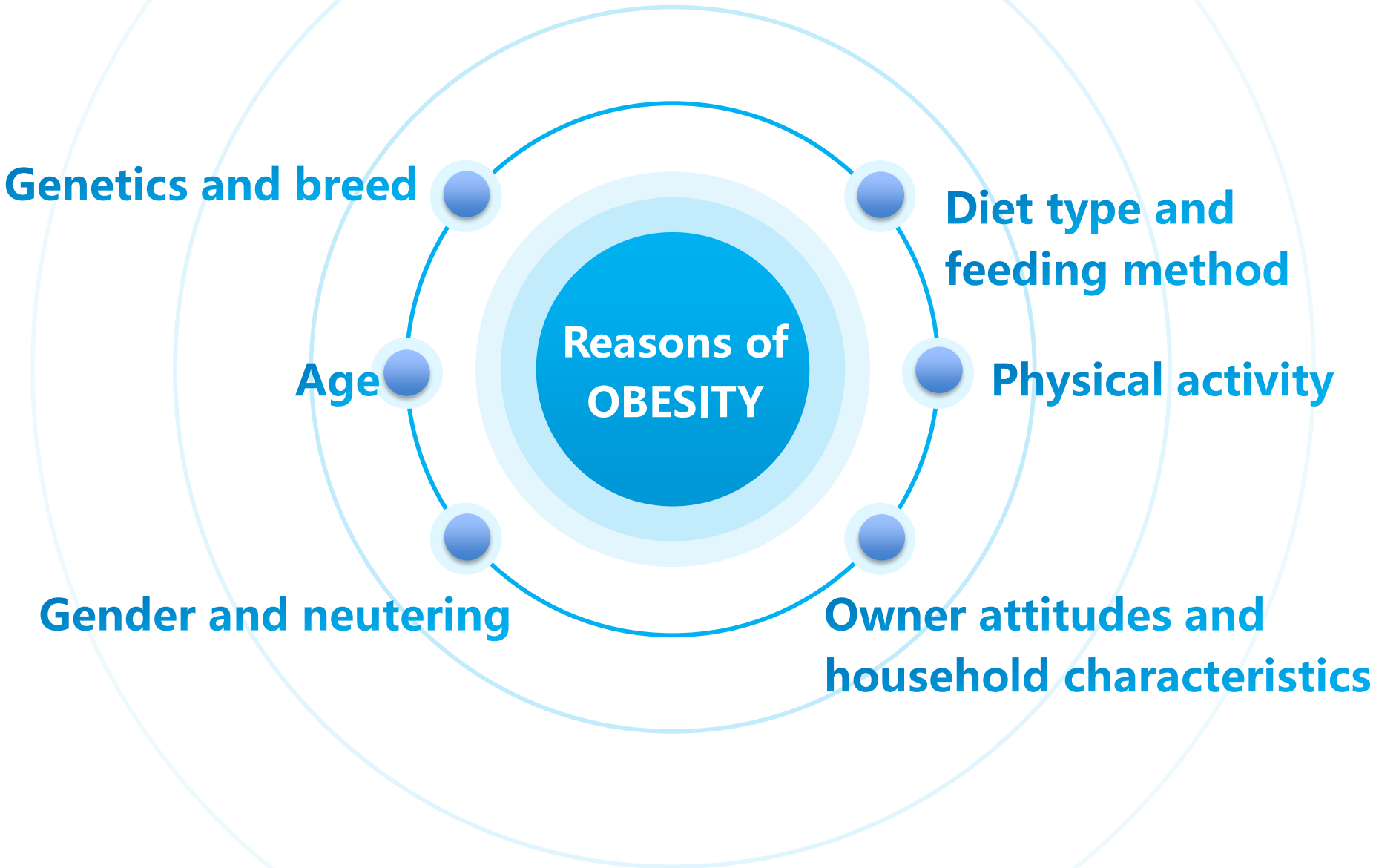
https://petobesityprevention.org/2018

White Paper on Pet Industry in China 2010-2019

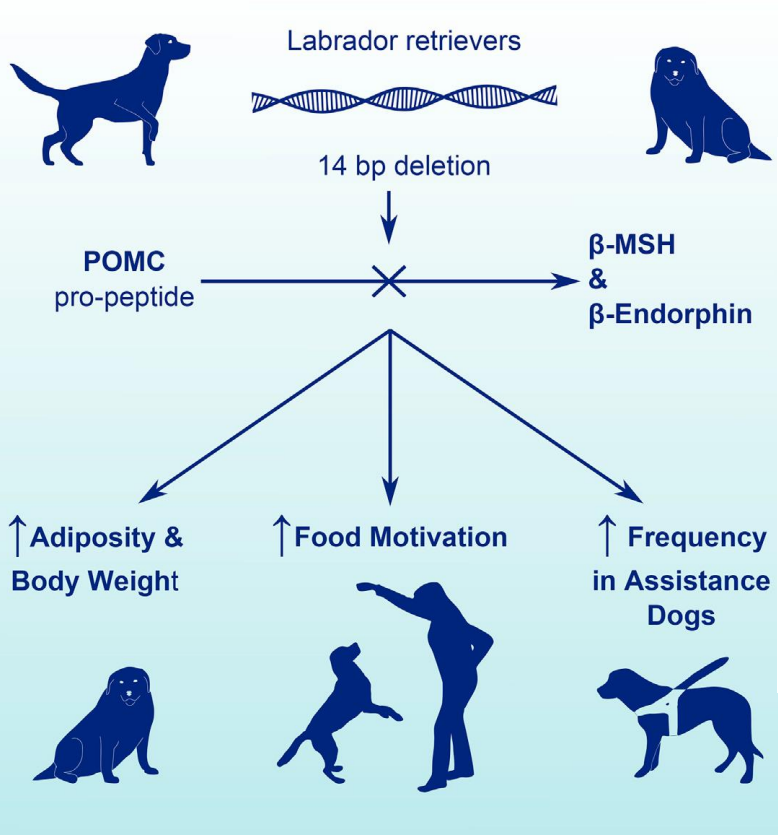
1 研究背景

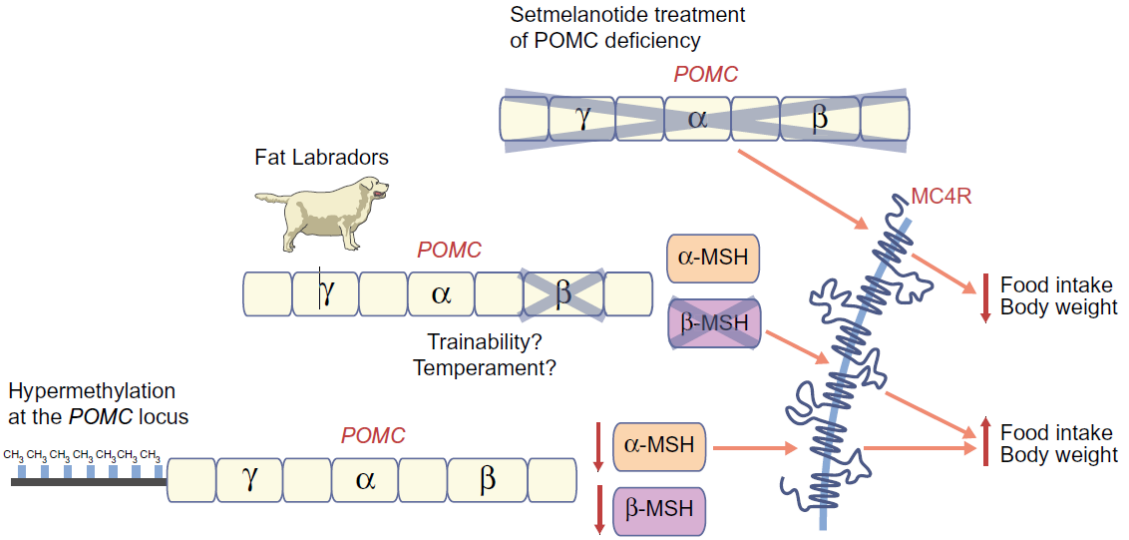

Comorbidities of OBESITY
Endocrine disease
Renal and urinary disease
Cardiovascular disease
Orthopedic and neurologic disease
Respiratory disease
Neoplasia

1 研究背景


Reasons of OBESITY
Genetics and breed
Diet type and feeding method
Age
Physical activity
Gender and neutering
Owner attitudes and household characteristics

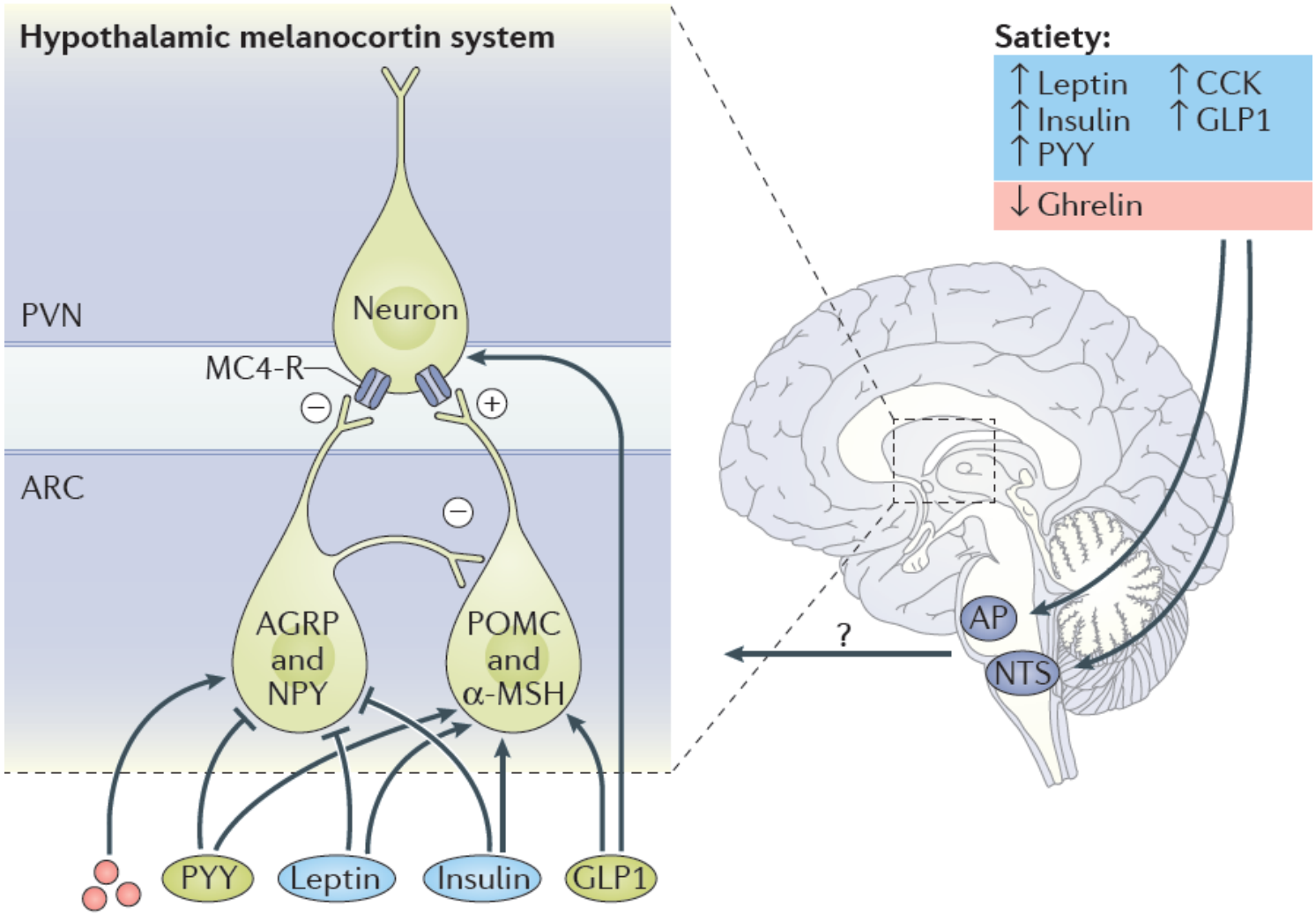
1 研究背景





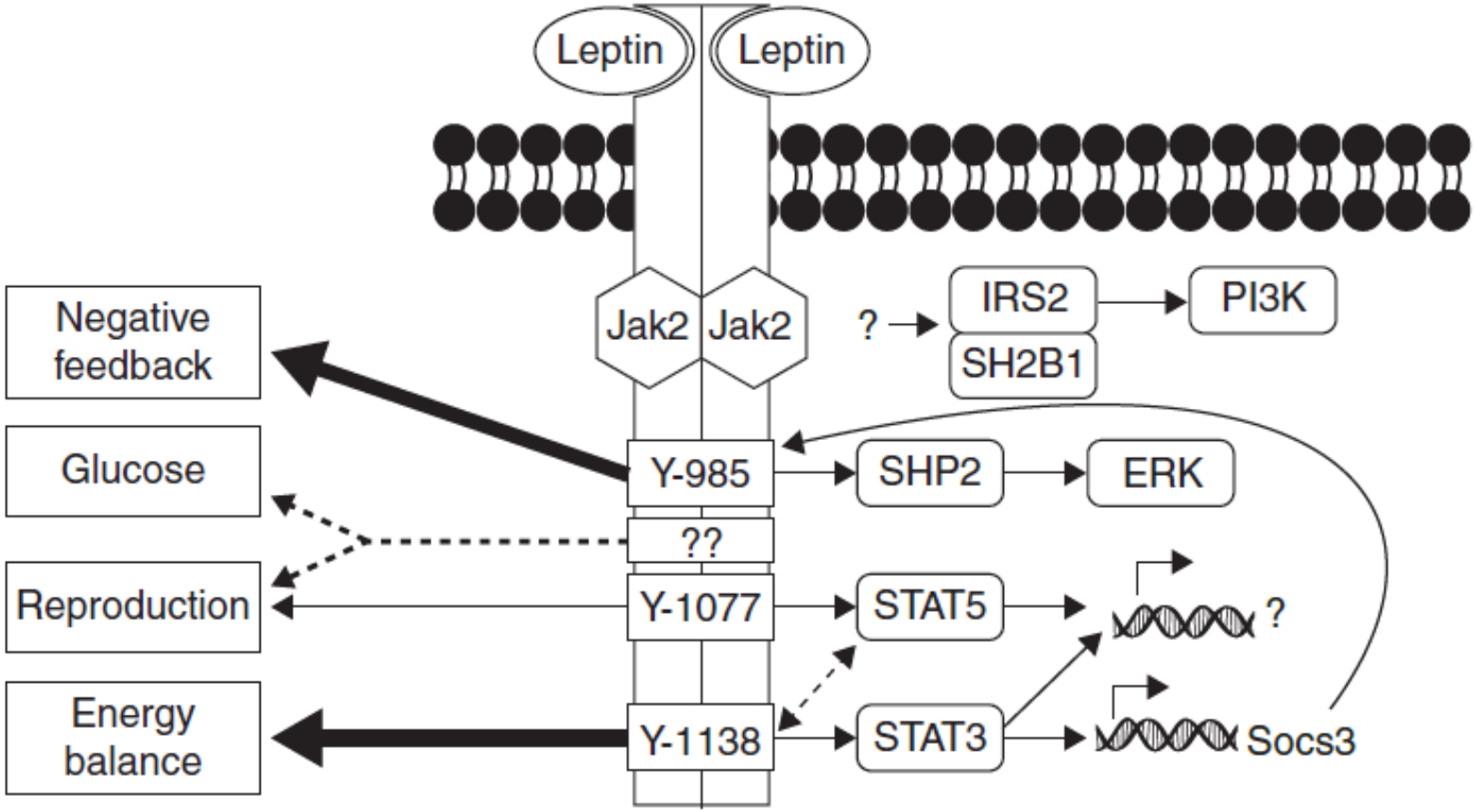
Yeo GSH, 2017 .Diabetologia.
Raffan E, *et al.*, 2016 .Cell Metab

1 研究背景



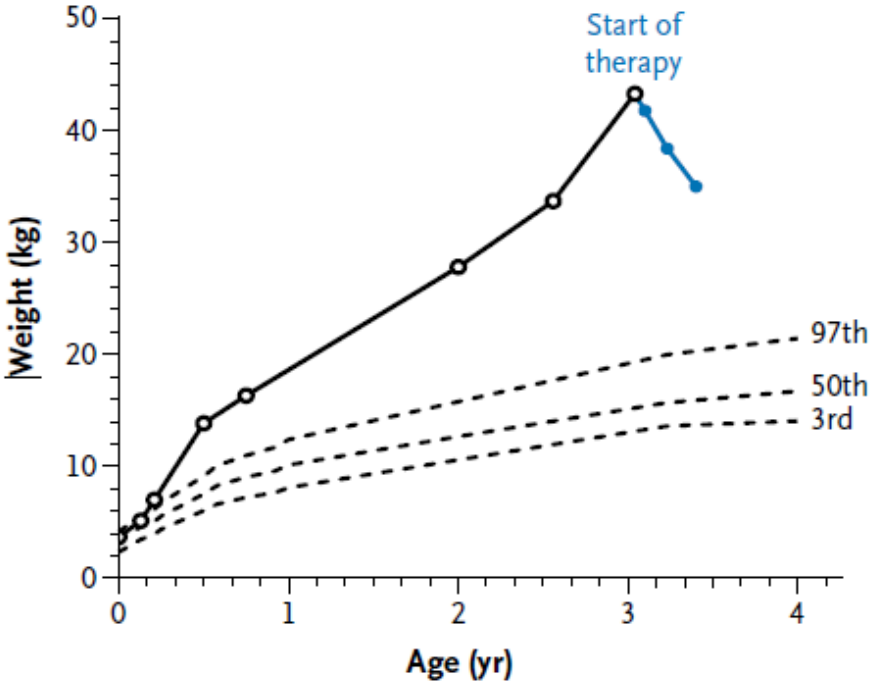
Kim KS, *et al.*, 2018 .Nat Rev Neurosci.

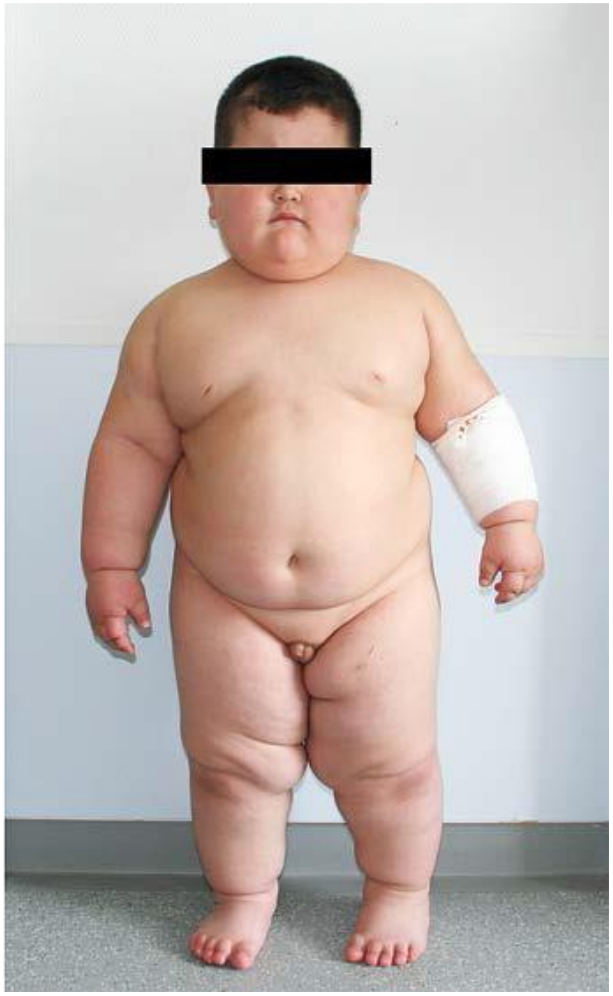
1 研究背景



Margaret B Allison, 2014. J Endocrinol

1 研究背景





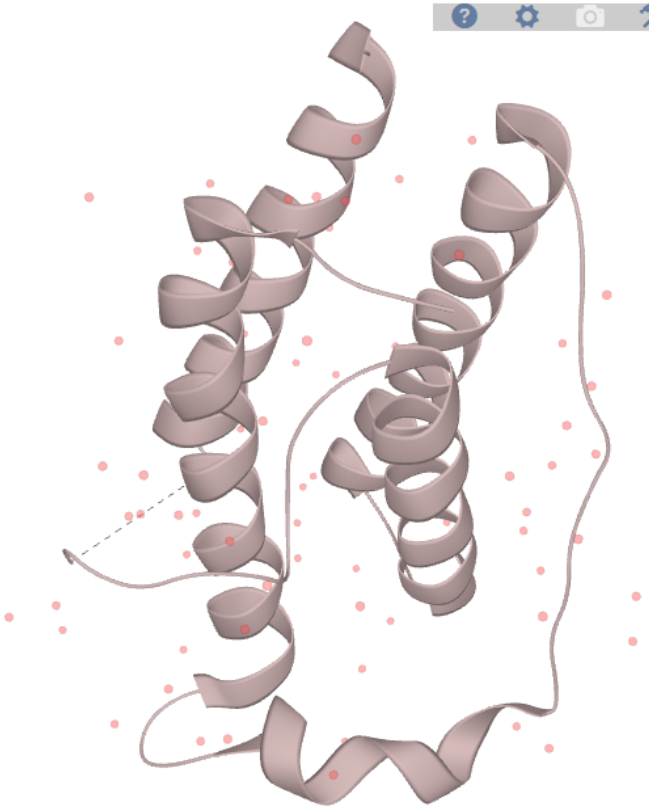
Wabitsch , *et al.*, 2015. N Engl J Med

1 研究背景

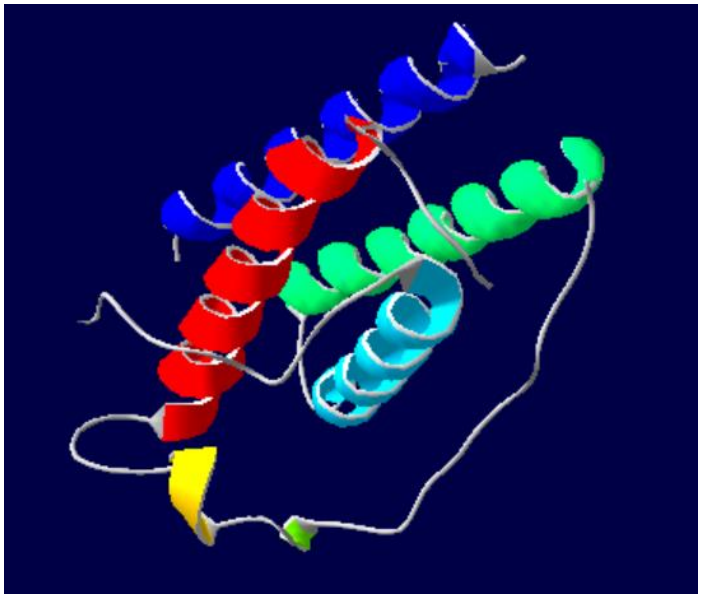
几种有代表性的瘦素氨基酸突变

突变	特点	文献
p. D100Y	可在外周检测到瘦素，使用重组瘦素后症状缓解	Wabitsch , *et al*., 2015. N Engl J Med
p. C117F	外周检测瘦素不足，该突变导致二硫键破坏	Yupanqui-Lozno,*et al*., 2019.Genes
p. N103K	部分报道可在外周检测到瘦素，使用重组瘦素后症状缓解	[1]Mazen*et al*., 2009. Mol Genet Metab [2]Shabana,*et al*., 2016. Biol Res [3]Wabitsch,*et al*., 2015. J Clin Endocrinol Metab
p. H118L	外周检测情况未报道，中国汉族来源	Zhao, *et al*., 2014. Biomed Res Int.

2 瘦素与瘦素受体的结构



PDB Entry	Method	Resolution	Chain	Positions	Links
1AX8	X-ray	2.40 Å	A	22-167	PDBe RCSB ... PDBj PDBsum



Natural variant

Feature key	Position(s)	Description	Actions	Graphical view	Length
Natural variant (VAR_004196)	49	Missing .			1
Natural variant (VAR_004197)	94	V → M 2 Publications Corresponds to variant dbSNP:rs17151919	Ensembl, ClinVar.		1
Natural variant (VAR_075144)	100	D → Y n LEPD; no effect on secretion; does not bind or activate LEPR. 1 Publication Corresponds to variant dbSNP:rs724159998	Ensembl, ClinVar.		1
Natural variant (VAR_008094)	105	R → W in LEPD. 1 Publication Corresponds to variant dbSNP:rs104894023	Ensembl, ClinVar.		1
Natural variant (VAR_011955)	110	V → M. Corresponds to variant dbSNP:rs1800564	Ensembl.		1

2 瘦素与瘦素受体的结构

PDB Entry	Method	Resolution	Chain	Positions	Links
3V6O	X-ray	1.95 Å	A/B	428-633	PDBe RCSB ... PDBj PDBsum
6E2P	X-ray	2.83 Å	C/D	863-933	PDBe RCSB ... PDBj PDBsum

Region

Feature key	Position(s)	Description	Actions	Graphical view	Length
Region	467 – 484	Leptin-binding 1 Publication	Add BLAST		18
Region	893 – 898	Required for JAK2 activation By similarity			6
Region	898 – 906	Required for STAT3 phosphorylation By similarity			9

Byron Carpenter. Structure 2012

2 瘦素与瘦素受体的结构

Leptin receptor

Chains: A, B
Length: 206 amino acids
Theoretical weight: 23.38 KDa
Source organism: *Homo sapiens*
Expression system: *Escherichia coli*
UniProt:
- Canonical: P48357 (Residues: 428-633; Coverage: 18%)

Gene names: DB, LEPR, OBR
Sequence domains: Obesity receptor immunoglobulin like domain
Structure domains: Immunoglobulins

Molecule details ›

Monoclonal antibody 9F8 fab fragment Heavy chain

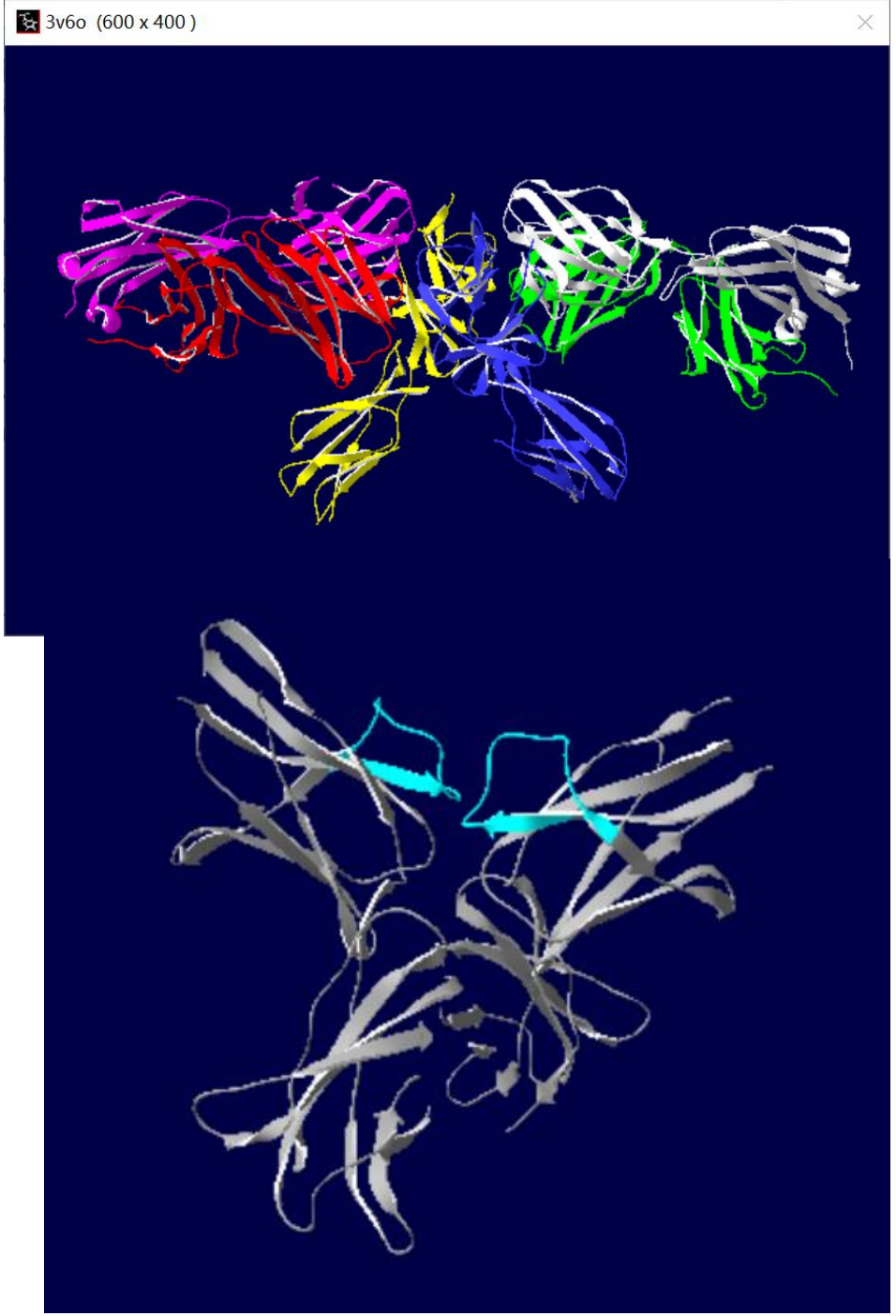
Chains: C, D
Length: 221 amino acids
Theoretical weight: 23.59 KDa
Source organism: *Mus musculus*
Structure domains: Immunoglobulins

Molecule details ›

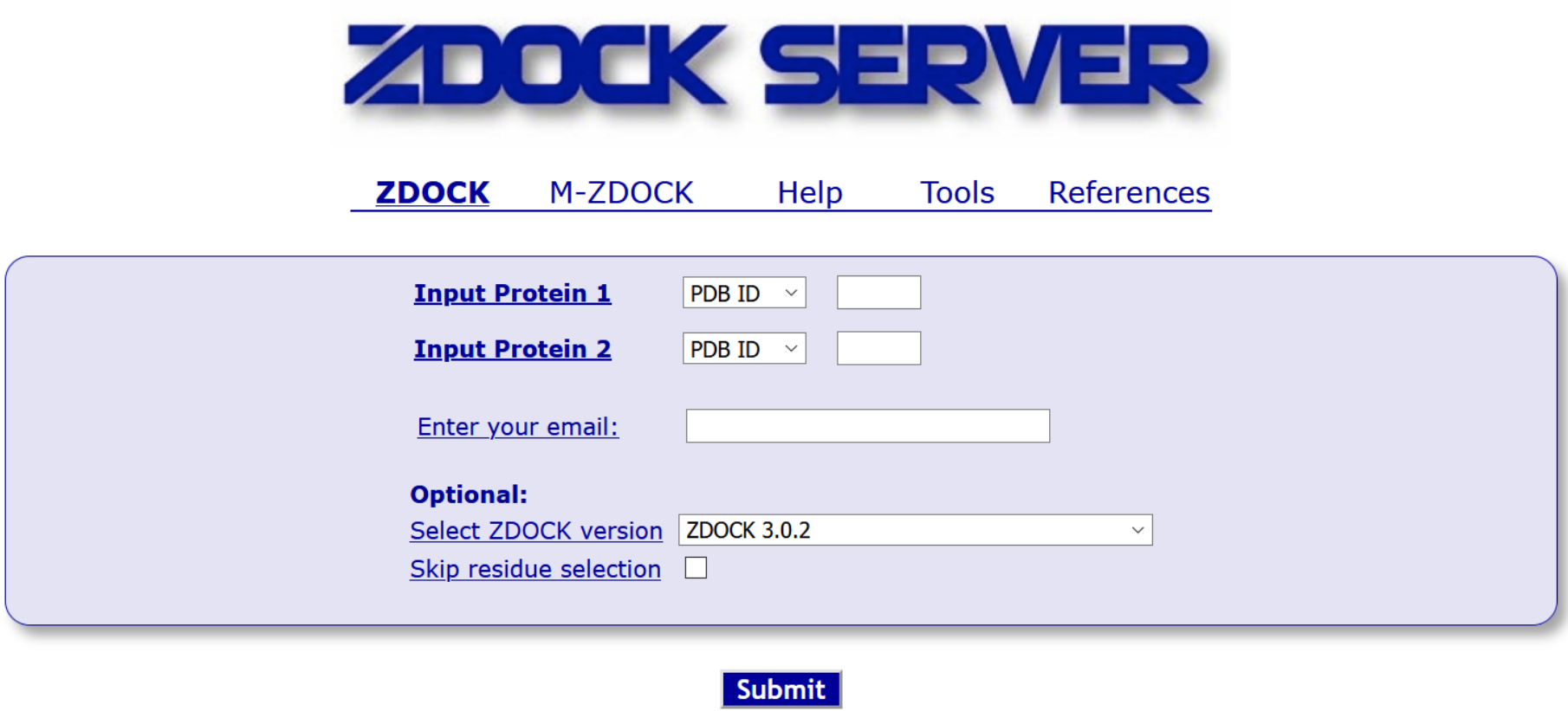
Monoclonal antibody 9F8 fab fragment Light chain
Chains: E, F

Chains: E, F
Length: 215 amino acids
Theoretical weight: 23.87 KDa
Source organism: *Mus musculus*
Structure domains: Immunoglobulins

Molecule details ›

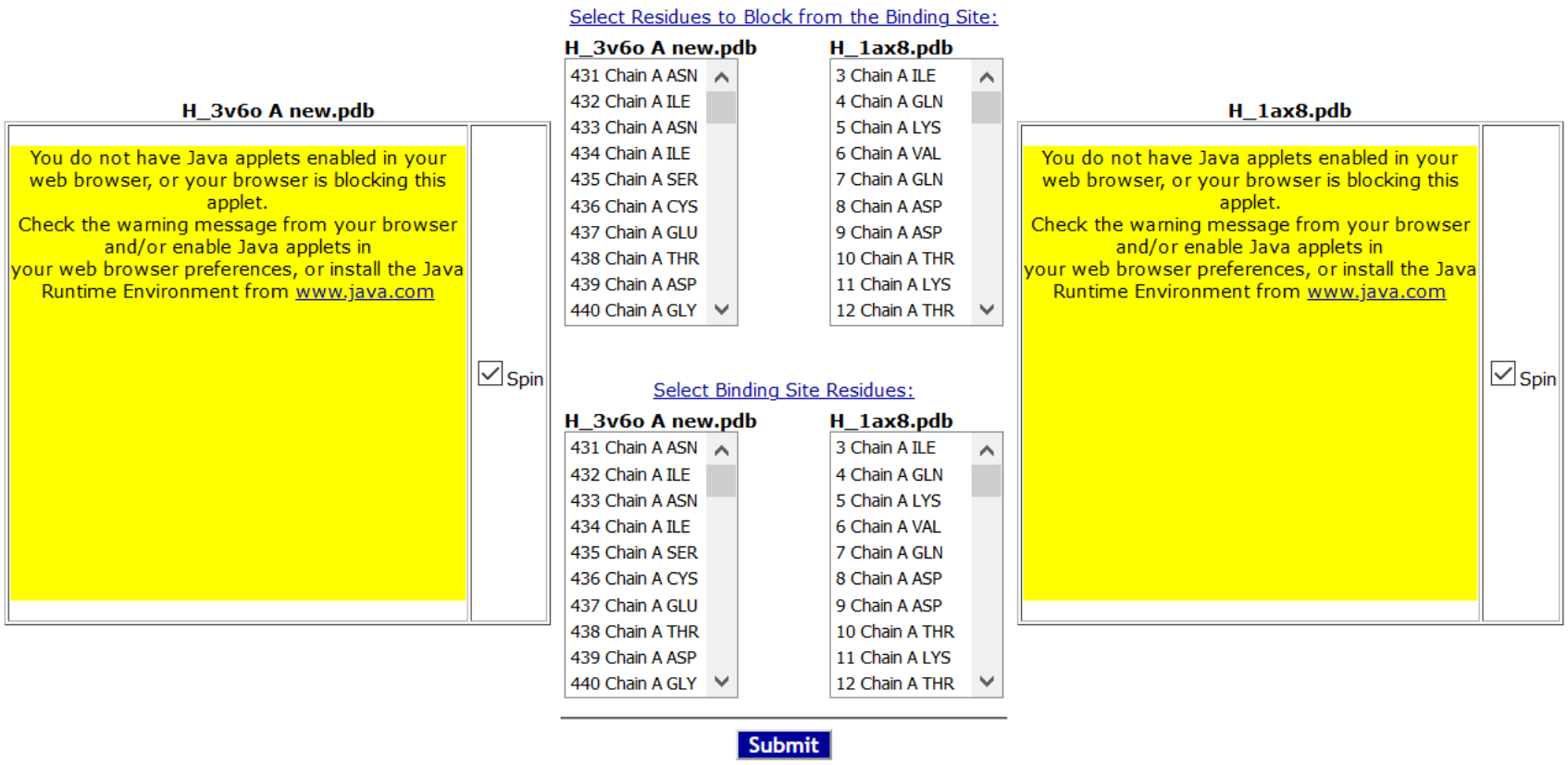


3 瘦素与瘦素受体的分子对接

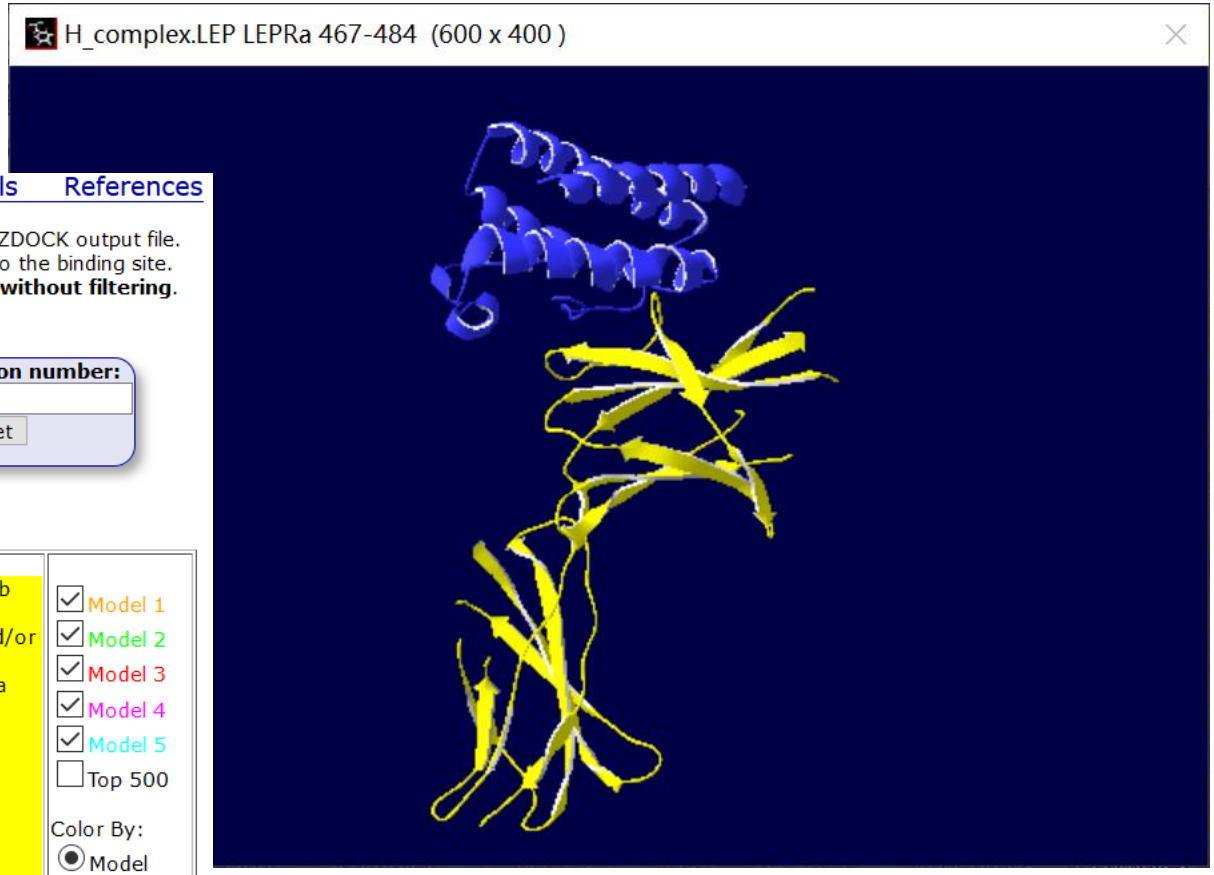


http://zdock.umassmed.edu/

3 瘦素与瘦素受体的分子对接



3 瘦素与瘦素受体的分子对接



ZDOCK M-ZDOCK Help Tools References

Contact filtering removed ALL predictions from the ZDOCK output file.
You may want to select fewer residues to force into the binding site.
For your reference, below is the ZDOCK output file **without filtering**.

Download Files

ZDOCK Output
Receptor PDB
Ligand PDB
Top 10 Predictions

Get prediction number:

Get

Top 5 Predictions

You do not have Java applets enabled in your web browser, or your browser is blocking this applet.
Check the warning message from your browser and/or enable Java applets in
your web browser preferences, or install the Java Runtime Environment from www.java.com

- [x] Model 1
- [x] Model 2
- [x] Model 3
- [x] Model 4
- [x] Model 5
- [ ] Top 500

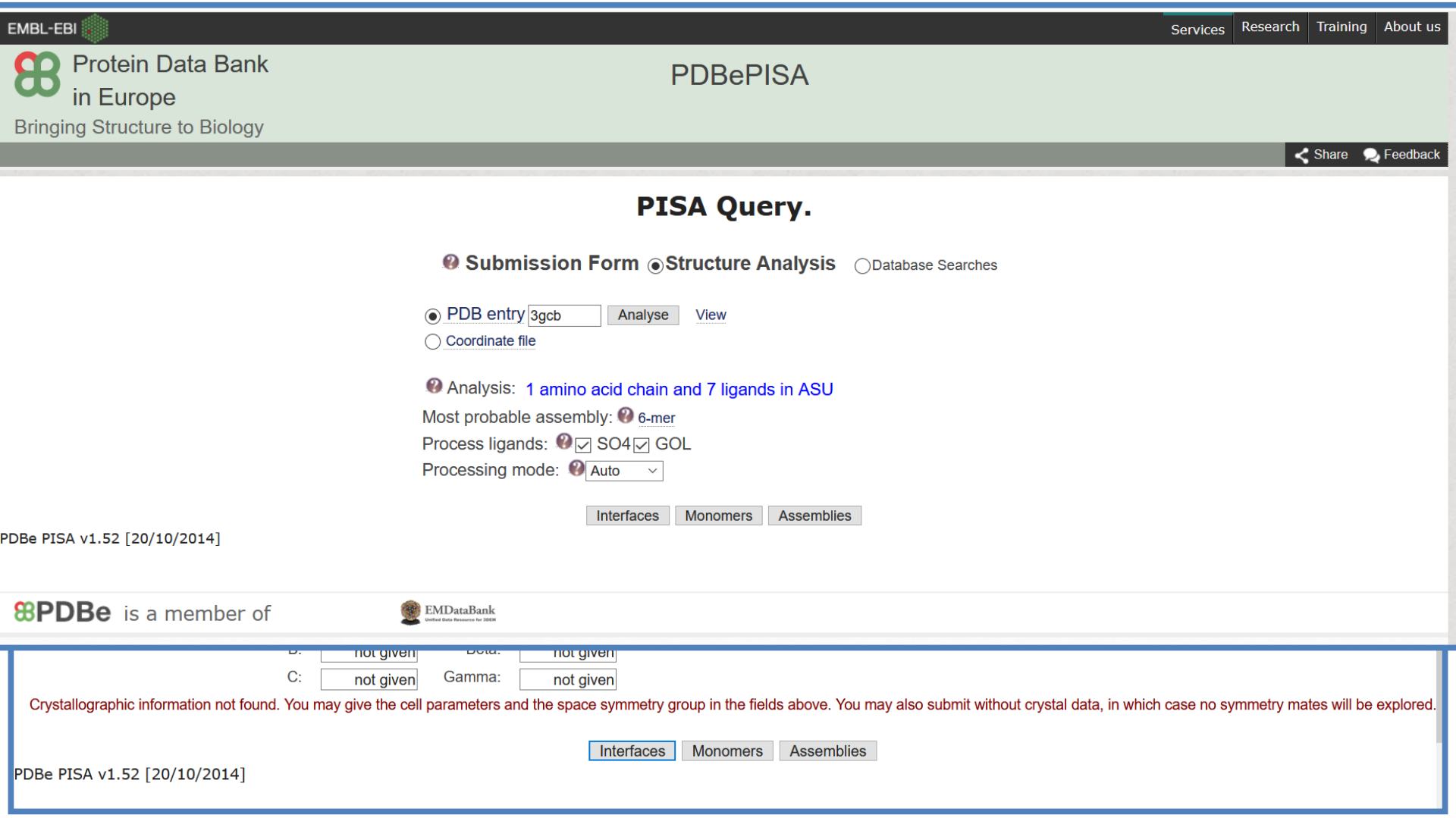
Color By:

- (•) Model
- ( ) Chain

- [x] Spin
- [ ] Stereo

- [x] Cartoon
- [ ] Wireframe
- [ ] Spacefill

4 分子对接结果评价



https://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver

4 分子对接结果评价

PISA Interface List.

Session Map (id=669-MI-52G)

Start | Interfaces | Interface Search | Monomers | Assemblies

Interfaces in H_complex.LEP LEPRa 467-484.pdb

Interfaces XML View Details Download Search

## NN	«»	Structure 1 Range	$^{i}N_{at}$	$^{i}N_{res}$	Surface Å^2	×	Structure 2 Range	$^{i}N_{at}$	$^{i}N_{res}$	Surface Å^2	interface area, Å^2	$\Delta^{i}G$ kcal/mol	$\Delta^{i}G$ P-value	N_{HB}	N_{SB}	N_{DS}	CSS
1	◉	A	68	19	11080	◊	B	81	21	7074	725.0	−11.2	0.177	3	0	0	0.000

View Details Download Search

4 分子对接结果评价

Hydrogen bonds

XML

No disulfide bonds found

No covalent bonds found

No salt bridges found

##	Structure 1	Dist. [Å]	Structure 2
1	A:HIS 467[NE2]	2.60	B:PHE 41[O]
2	A:SER 470[N]	3.65	B:PHE 41[O]
3	A:SER 478[OG]	2.30	B:ASP 135[OD1]

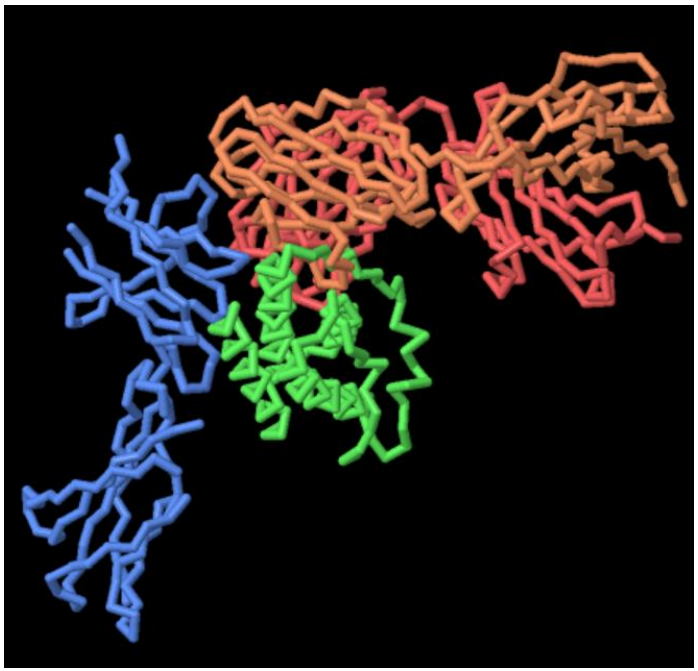
A:TYR 466		31.95	4.63 ||	0.03
A:HIS 467	H	39.26	31.93 ||||||||	0.89
A:ARG 468		103.46	39.99 ||||	-0.18
A:SER 469		31.97	27.08 ||||||||	-0.07
A:SER 470	H	114.34	58.65 ||||||	-0.05
A:LEU 471		111.51	61.27 ||||||	0.97
A:TYR 472		124.77	6.17 |	-0.04
A:CYS 473		49.69	0.00	0.00
A:SER 474		85.80	45.99 ||||||	0.04
A:ASP 475		149.34	36.20 |||	0.02
A:ILE 476		135.86	104.30 ||||||||	1.50
A:PRO 477		52.63	1.22 |	-0.01
A:SER 478	H	51.48	49.01 |||||||||	0.45
A:ILE 479		66.79	53.90 ||||||||	0.34
A:HIS 480		50.47	34.47 |||||||	-0.31
A:PRO 481		141.75	73.82 ||||||	1.18
A:ILE 482		142.26	81.99 ||||||	1.29
A:SER 483		21.98	3.44 ||	0.06
A:GLU 484		105.33	21.76 |||	-0.16

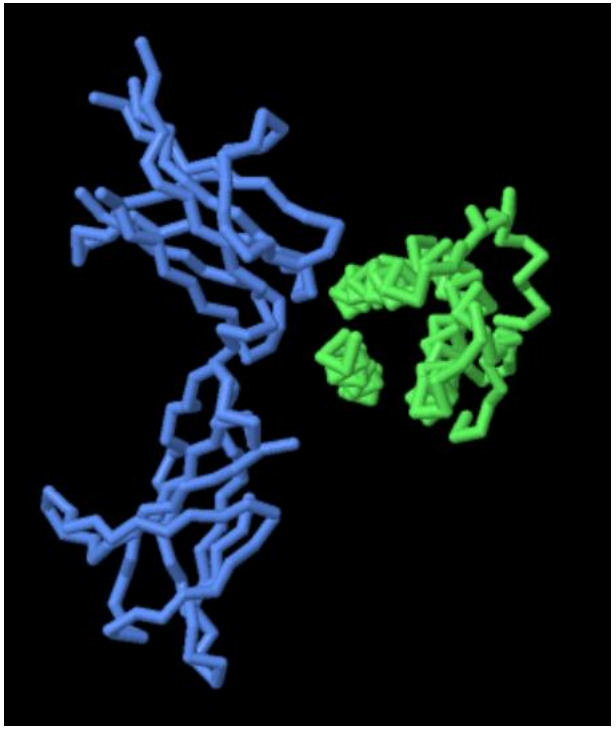
B:LEU 39		79.78	11.24 ||	0.18
B:ASP 40		130.38	37.14 |||	-0.18
B:PHE 41	H	188.97	161.78 ||||||||	2.18
B:ILE 42		29.15	24.25 ||||||||	0.26
B:PRO 43		95.51	83.93 ||||||||	1.34
B:GLY 44		13.58	5.20 ||||	-0.06
B:LEU 45		18.98	0.00	0.00
B:HIS 46		93.87	26.74 |||	0.36
B:PRO 47		18.95	15.50 ||||||||	0.02
B:ILE 48		56.85	49.70 ||||||||	0.80
B:LEU 49		91.42	24.93 |||	0.40
B:THR 50		33.27	0.00	0.00
B:LEU 51		0.00	0.00	0.00
B:SER 52		26.85	0.00	0.00
B:LYS 53		72.20	5.40 |	-0.20
B:MET 54		0.17	0.00	0.00
B:ASP 55		2.27	0.00	0.00
B:GLN 56		43.40	0.00	0.00
B:THR 57		0.00	0.00	0.00
B:LEU 58		0.00	0.00	0.00
B:ALA 59		11.78	0.00	0.00
B:VAL 60		20.19	0.00	0.00
B:TYR 61		0.00	0.00	0.00

B:ARG 128		34.53	26.44 |||||||	-0.73
B:LEU 129		0.00	0.00	0.00
B:GLN 130		31.57	6.17 ||	-0.10
B:GLY 131		21.03	12.42 |||||	0.09
B:SER 132		0.00	0.00	0.00
B:LEU 133		0.00	0.00	0.00
B:GLN 134		85.73	62.55 |||||||	-0.08
B:ASP 135	H	39.26	35.74 ||||||||	-0.27
B:MET 136		1.51	0.00	0.00
B:LEU 137		44.88	7.36 ||	0.12
B:TRP 138		142.95	88.25 ||||||	1.13
B:GLN 139		24.68	15.55 ||||||	-0.18

存在的问题：

1. 瘦素受体结构不明确，对后续结果影响较大



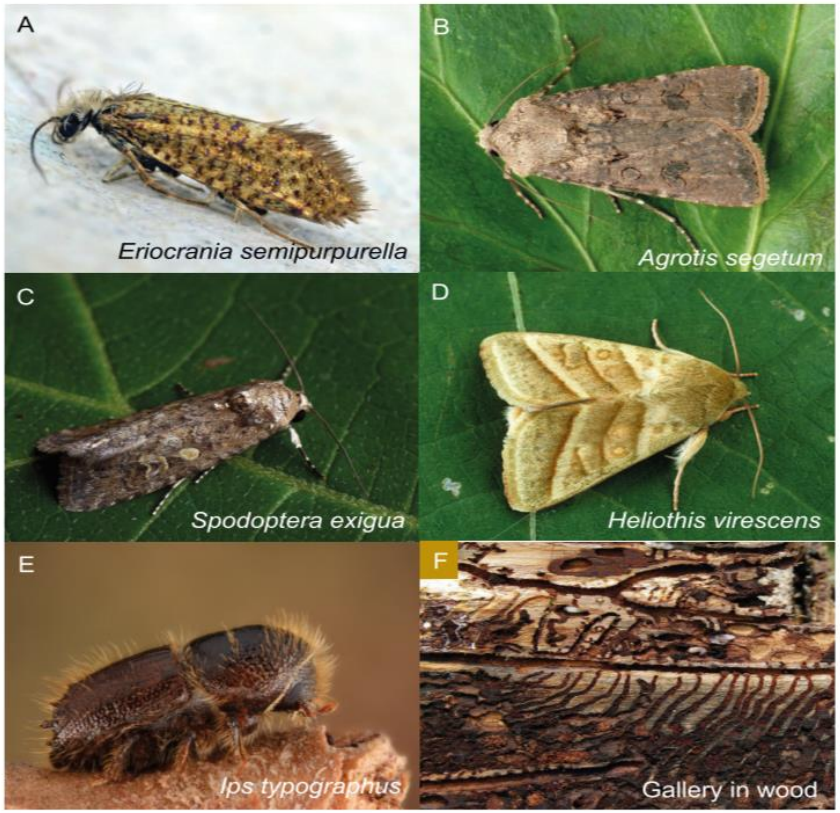


https://pdb101.rcsb.org/motm/149

Part3 蛋白质分子-化学小分子物质对接示例演示

- 背景介绍
- 化学感受蛋白
- 气味受体
- 研究计划及面对的问题

背景介绍



- **Insects use their remarkable olfactory system to detect and discriminate different chemicals signals from the surrounding environment to locate the food sources.**

- **To find suitable oviposition sites and mating partners, and to avoid predators**

昆虫嗅觉系统

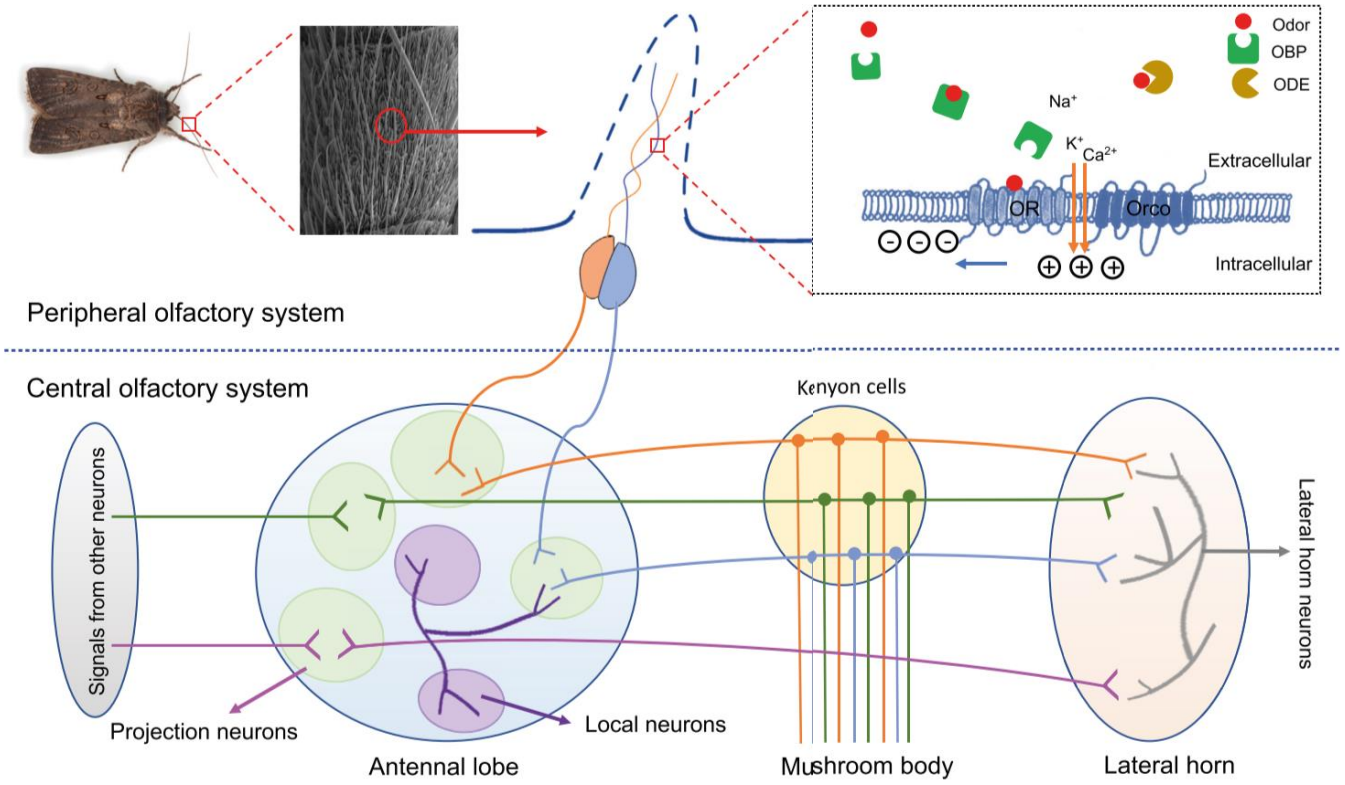
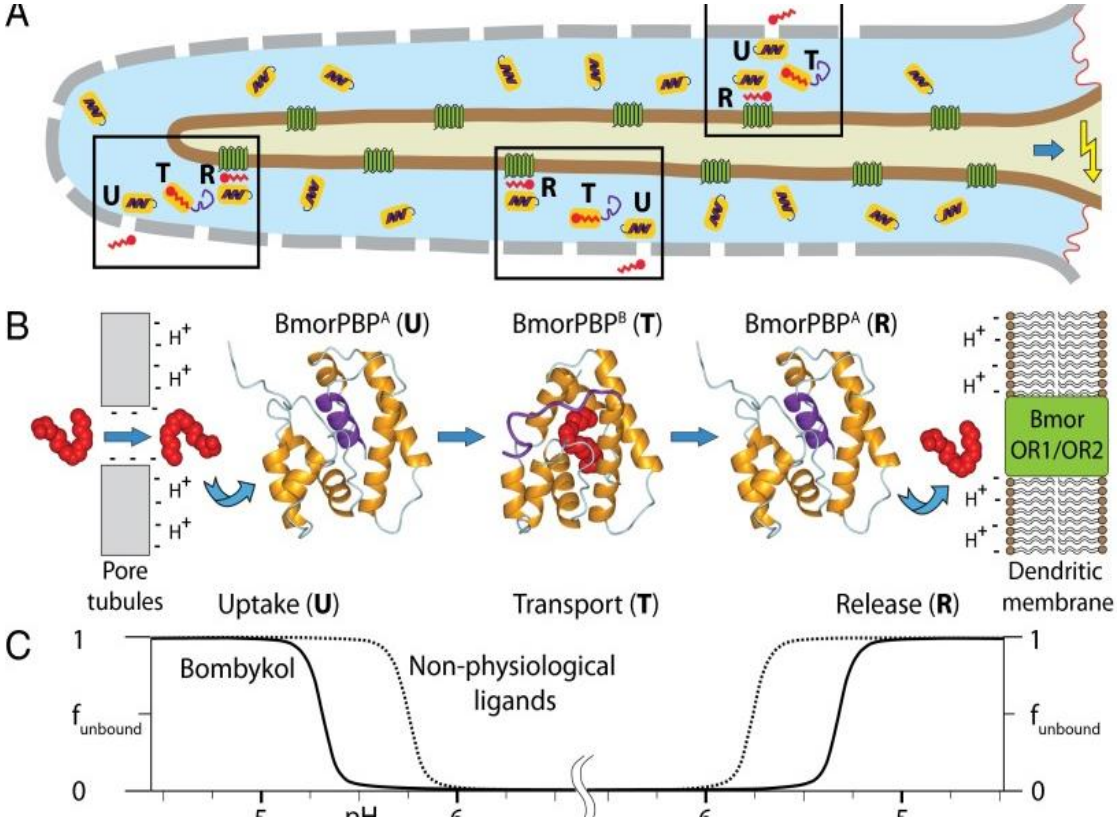


Figure 1. Schematic representation of the insect olfactory system

Hou, 2020. Lund University

昆虫外周神经信号传递过程



①Odour chemical molecules enter into sensilla through small pores

②They are captured by olfactory binding proteins(OBPs) in lymph

③ Then transported to the odorant receptors(ORs) expressed on the dendritic membrane of OSNs

④The interaction of receptors and ligands elicit the transduction of chemical signals to electric signals.

⑤The chemical molecules will be afterward inactivated by odorant degrading enzymes(ODEs) rapidly

外周神经信号传递中相关蛋白

- Chemosensory Proteins (CSPs): 化学感受蛋白
- Odorant Binding Proteins (OBPs): 气味结合蛋白

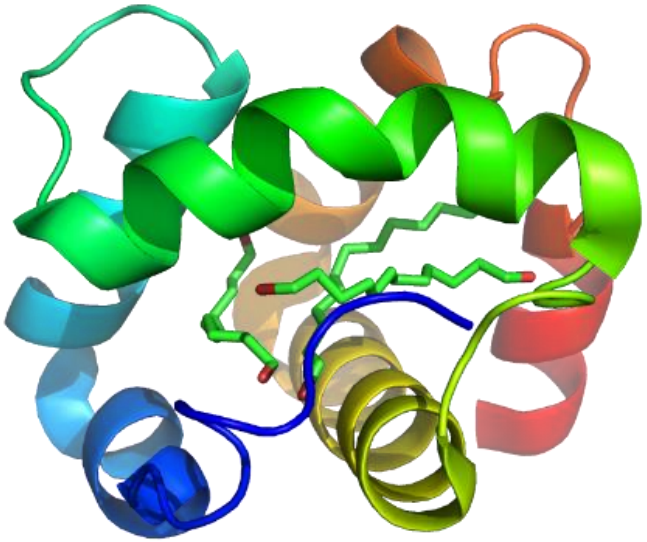
载体蛋白

- Odorant Receptors (ORs) : 嗅觉受体
- Ionotropic Receptor (IRs):离子型受体
- Sensory Neuron Membrane Protein (SNMP): 昆虫感觉神经元膜蛋白

膜蛋白

- Odorant Degradation Enzyme (ODEs):气味降解酶

化学感受蛋白结构特征

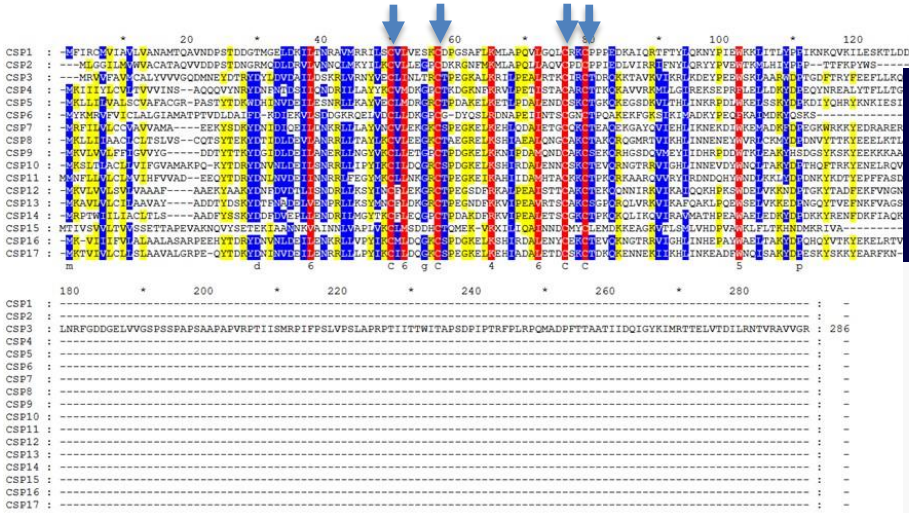


- MbraA6CSP
- Access Number: 1n8u
- Resolution:1.14Å
- Method: X-ray

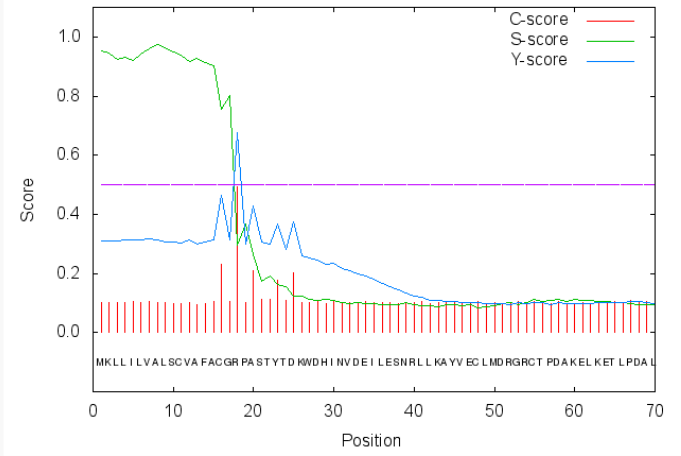
- CSPs是分泌蛋白，有15-25个氨基酸残基组成的信号肽
- 由110-130个氨基酸残基组成
- 4个保守的半胱氨酸残基
- 6个α-螺旋组成的一个球状蛋白

Campanacci *et al*, 2003. PNAS

化学感受蛋白蛋白序列特征

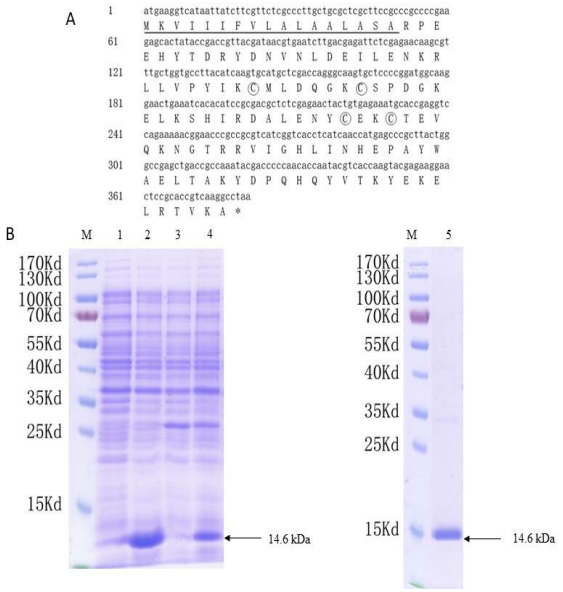


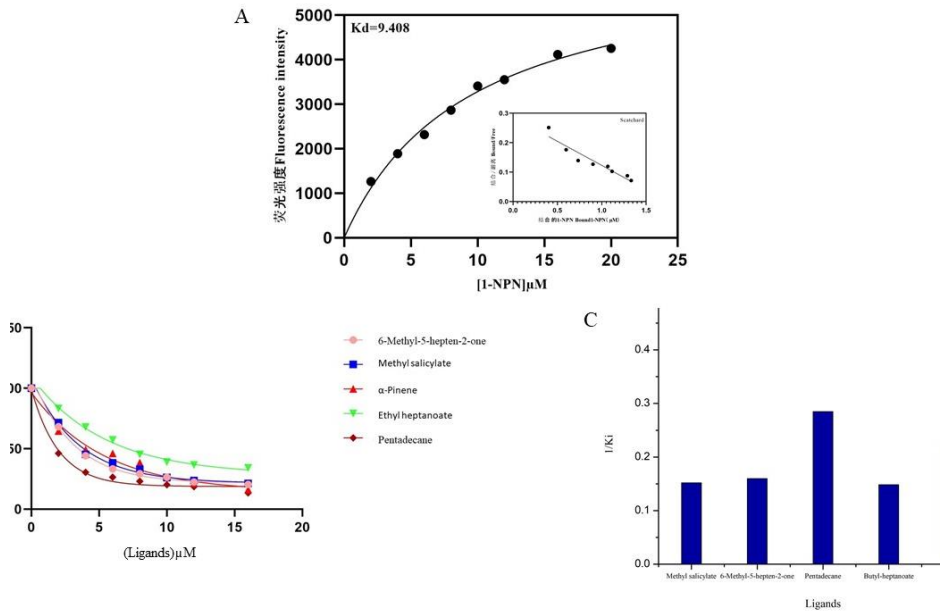
SignalP-4.1 prediction (euk networks): Sequence




```
# Measure  Position  Value    Cutoff   signal peptide?
  max. C    18       0.493
  max. Y    18       0.676
  max. S     8       0.973
  mean S     1-17    0.919
         D   1-17    0.807   0.450   YES
Name=Sequence   SP='YES' Cleavage site between pos. 17 and 18: ACG-RP D=0.807 D-cutoff=0.450 Networks=SignalP-
# data
# gnuplot script
```

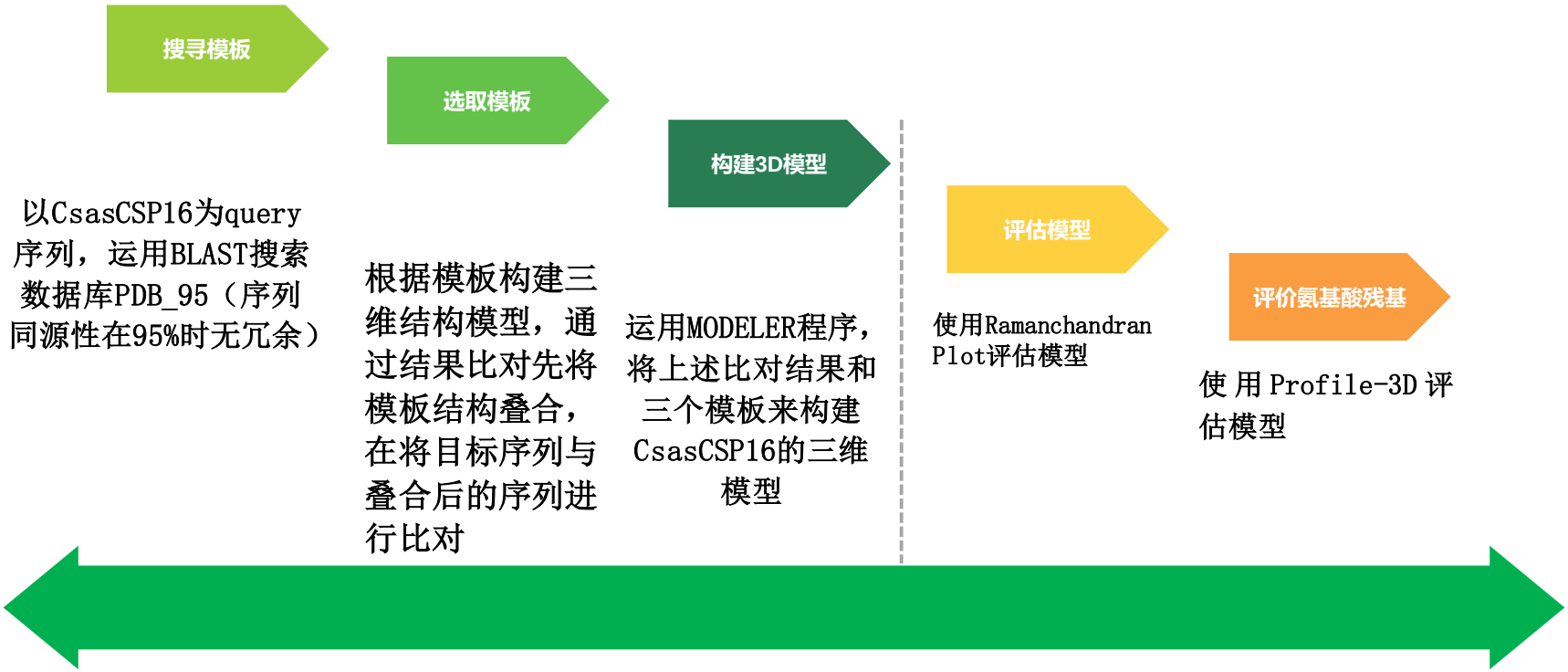
化学感受蛋白体外表达与荧光竞争结合



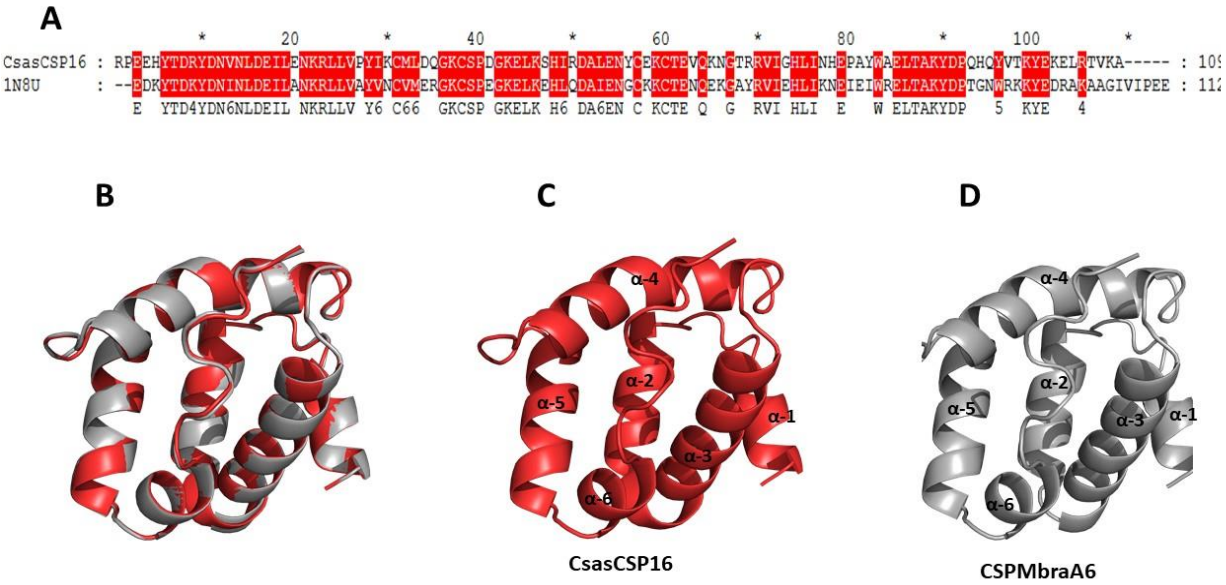


该蛋白与水杨酸甲酯、庚酸丁酯、十五烷、alpha-蒎烯和6甲基-5-庚烯-2-酮的结合亲和性较高

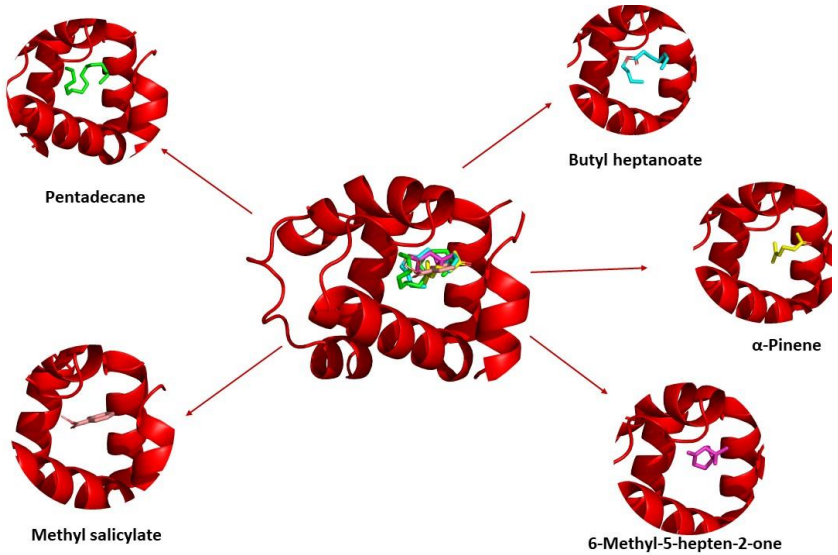
化学感受蛋白三维模型构建

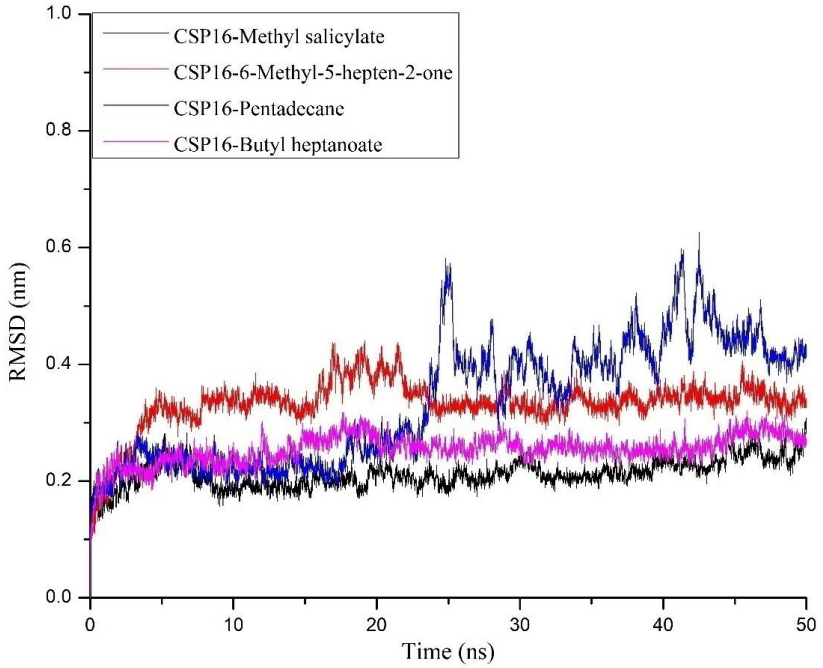


化学感受蛋白三维模型构建



化学感受蛋白分子对接与动力学模拟





MM/PBSA方法获得复合物相互作用力

CsasCSP16与配体分子的结合自由能和组分

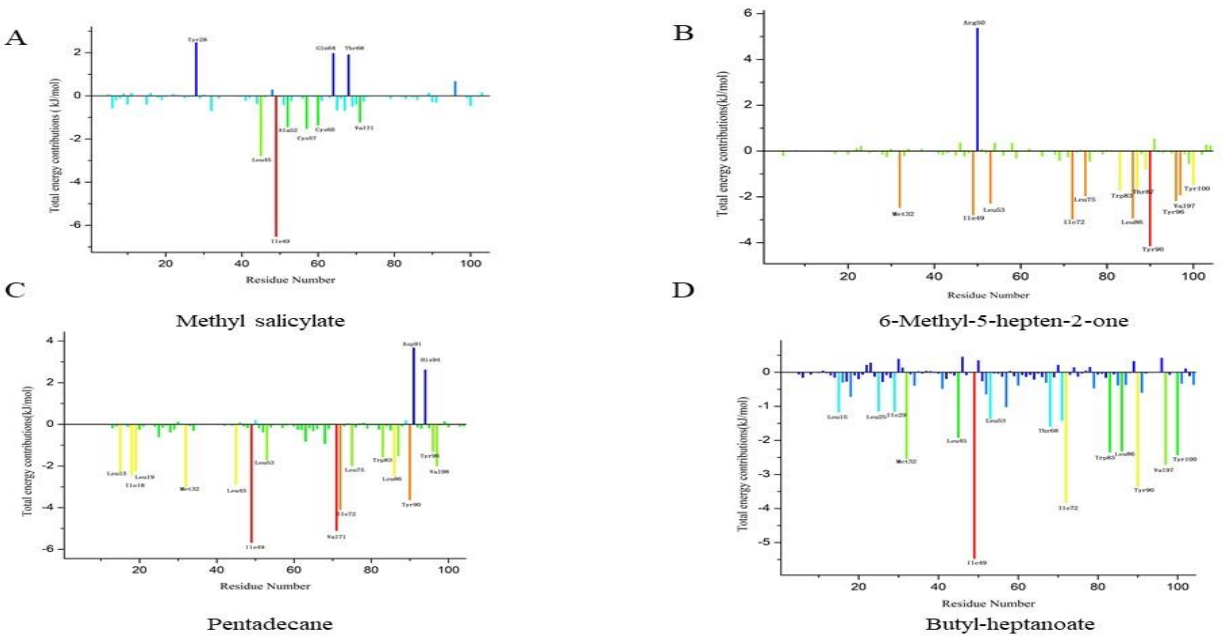
配体 Ligand name	范德华力 (kJ/mol) Val der Waals	静电作用力 （kJ/mol) Electrostatic force	极性溶剂化能 （kJ/mol） Polar solvation energy	非极性溶剂化能 （kJ/mol） Nonpolar solvation energy	结合能（kJ/mol） Binding energy
水杨酸甲酯 Methyl salicylate	-82.466 （ 0.682）	-18.084 （0.850）	61.336 （ 0.750）	-11.103 （0.055）	-50.264 （0.702）
6-甲基-5-庚烯-2酮 6-Methyl-5-hepten-2-one	-93.570 （ 0.927）	-16.158 （1.325）	55.604 （0.973）	-11.428 （ 0.061）	-65.551 （ 1.080 ）
十五烷 Pentadecane	-164.748 （ 0.845 ）	-0.969 （0.069 ）	49.259 （0.848 ）	-19.537 （ 0.089 ）	-136.035 （ 1.047）
庚酸丁酯 Butyl heptanoate	-109.800 （ 0.787 ）	-1.567 （0.211 ）	33.365 （0.782 ）	-15.787 （ 0.079 ）	-93.805 （1.191 ）

MM/PBSA获得单个氨基酸贡献的结合自由能

主要残基对结合自由能的贡献

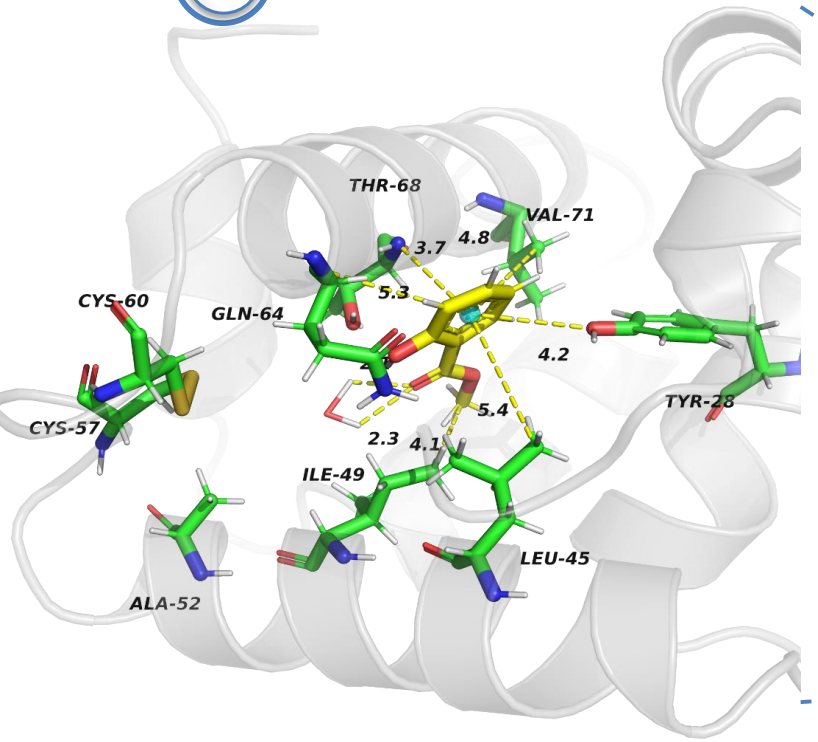
气味分子 Volatiles ligands	相互作用能 Interaction energy/(kJ/mol)			
	范围 Range(>2.0)	范围 Range(-1.0 ~ -3.0)	范围 Range (-3.0 ~ -5.0)	范围 Range(< -5.0)
水杨酸甲酯 Methyl salicylate	Tyr28, Gln64, Thr68	Leu45, Ala52, Cys57, Cys60, Val71	-	Ile49
6-甲基-5-庚烯-2酮 6-Methyl-5-hepten-2-one	Arg50	Met32, Ile49, Leu53, Ile72, Leu75, Trp83, Leu86, Thr87 Tyr96, Val 97, Tyr100	Thr90	-
十五烷 Pentadecane	Asp91, His94	Leu15, Ile18, Leu19, Met32, Leu45, Leu53, Leu75, Trp83, Leu86, Thr87, Tyr96, Val97	Ile72, Thr90	Ile49, Val71
庚酸丁酯 Butyl heptanoate	-	Leu15, Leu25, Ile29, Met32, Leu45, Leu53, Cys57 Thr68, Val71, Trp83, Leu86, Val97, Tyr100	Ile72, Thr90	Ile49

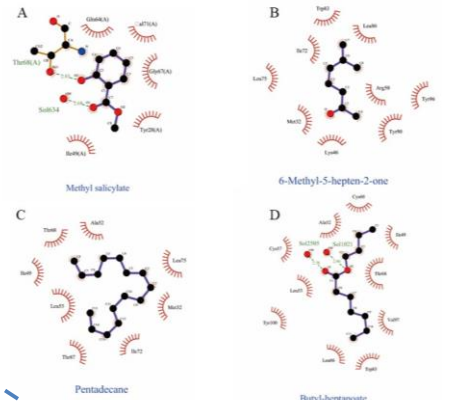
MM/PBSA获得单个氨基酸贡献的结合自由能

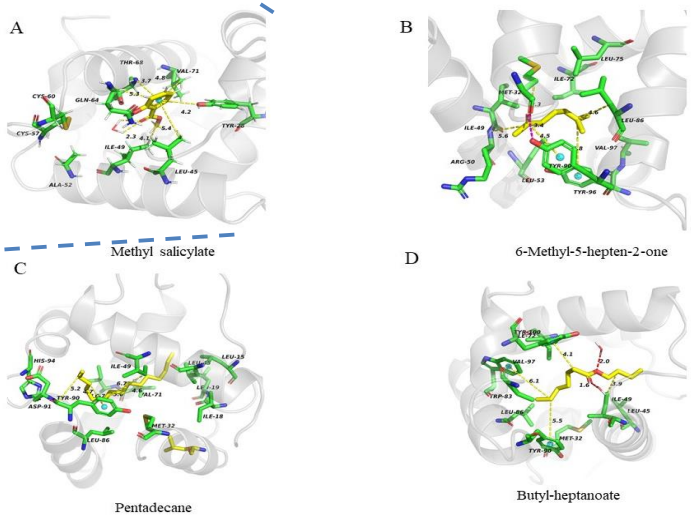


异亮氨酸49（Ile49）、缬氨酸71（Val71）、异亮氨酸72 （Ile72）苏氨酸90（Thr90）是关键氨基酸残基

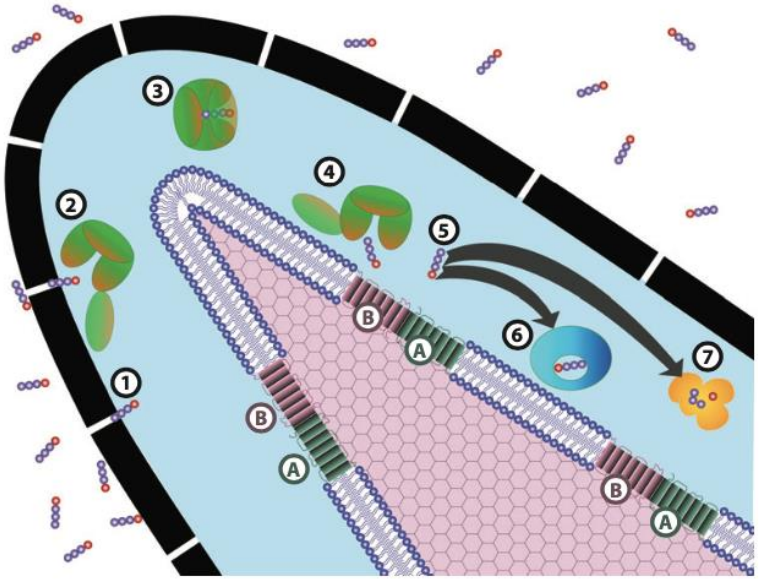
配体分子与相互作用的氨基酸残基的距离及位置

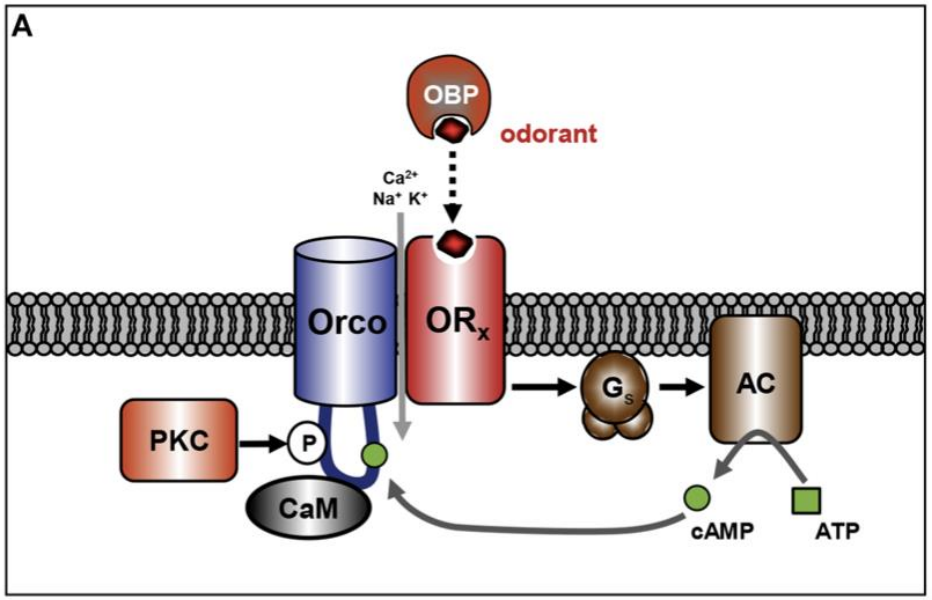






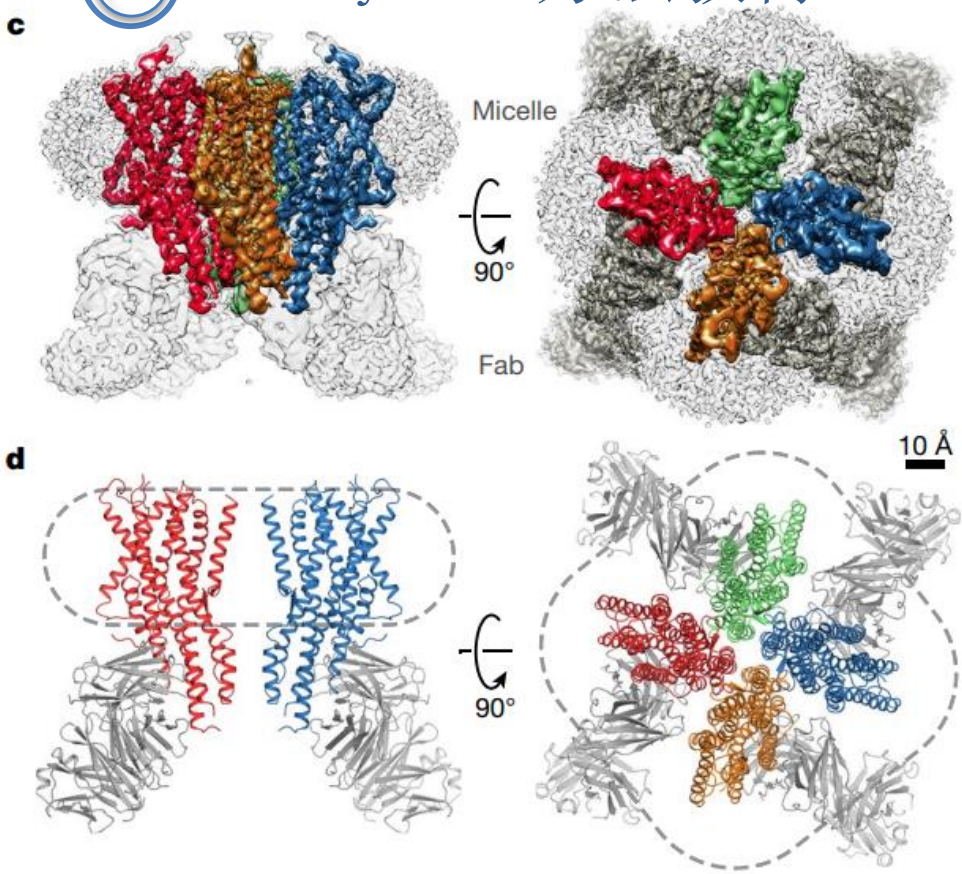
背景介绍—嗅觉受体（博士期间拟开展课题）

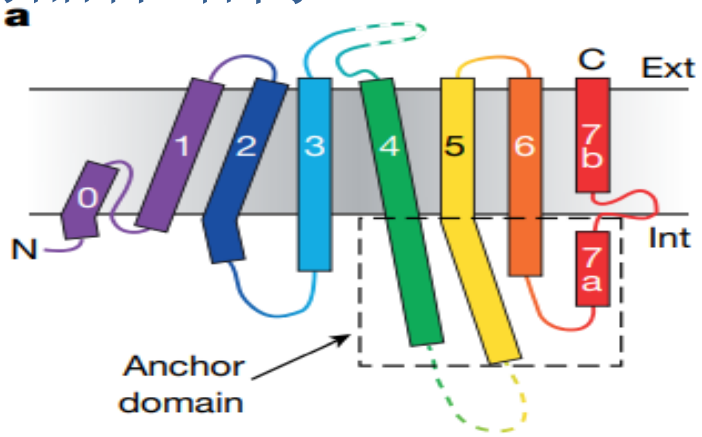


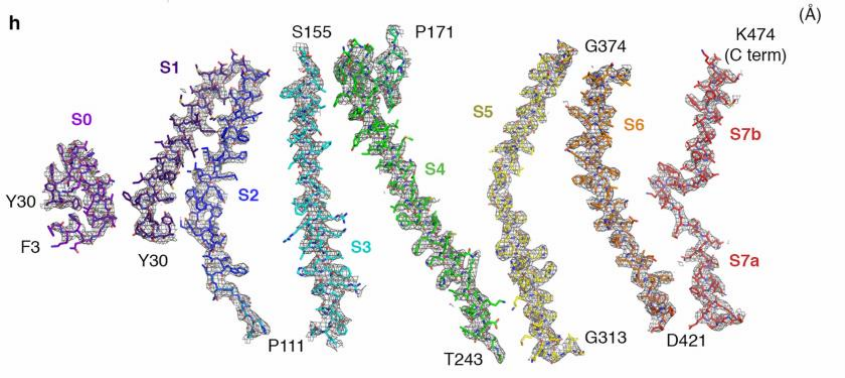


Leal *et al*., 2013. Annu Rev Entomol
Fleischer *et al*., 2018. Cell Mol Life Sci

Cryo-EM方法获得ORCO的晶体结构

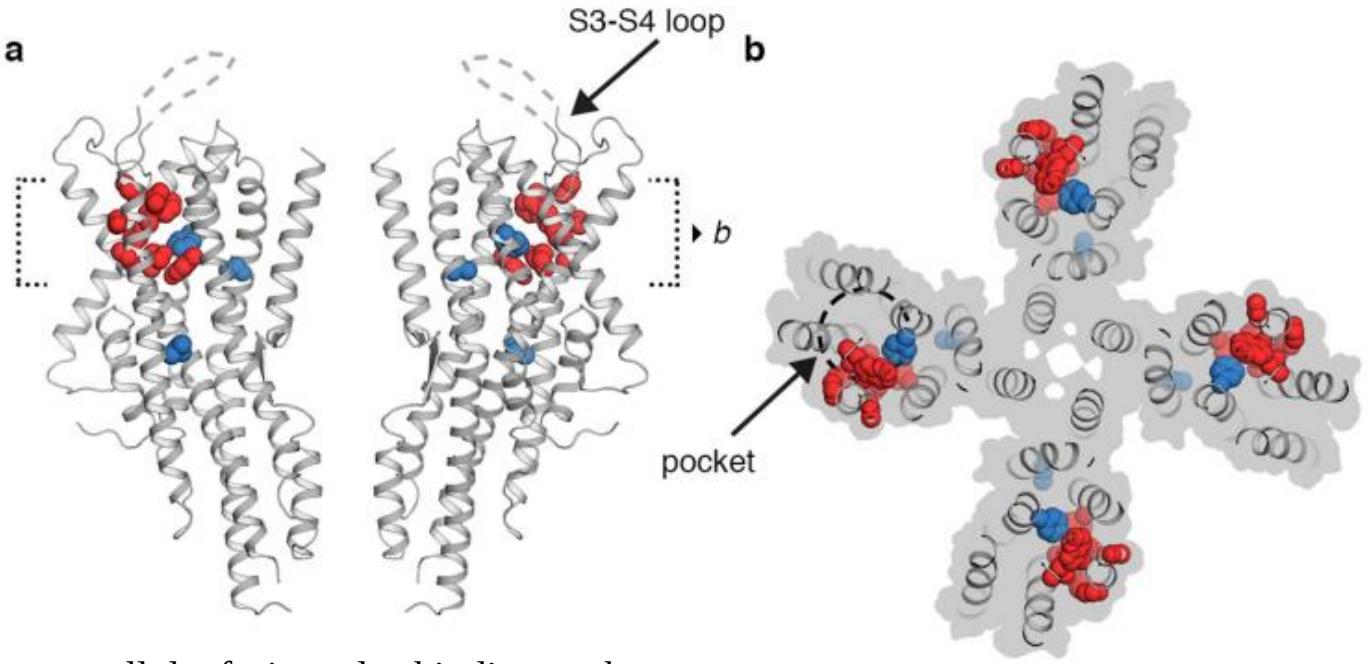






Butterwick *et al*.,2018.Nature

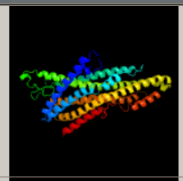
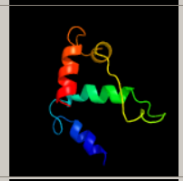
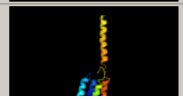
研究进展-预测的与配体分子的结合口袋



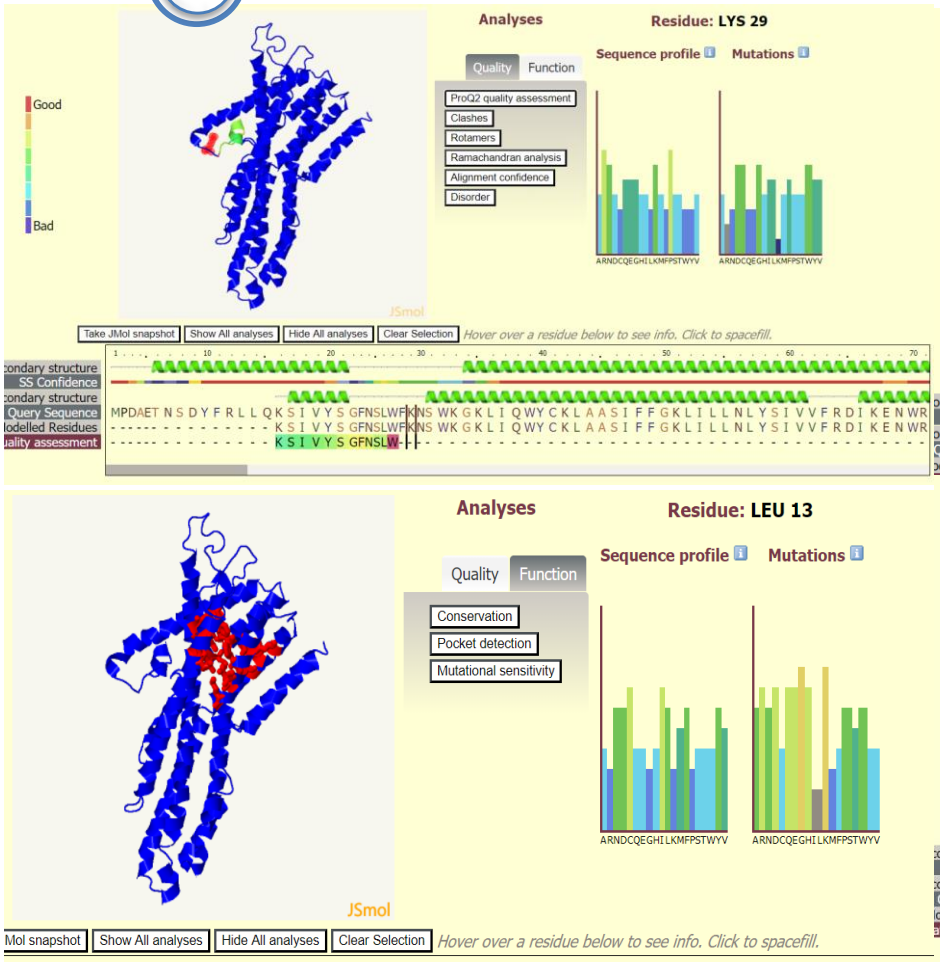
- Potential extracellular facing odor-binding pocket
- Side view of Orco highlighting the location

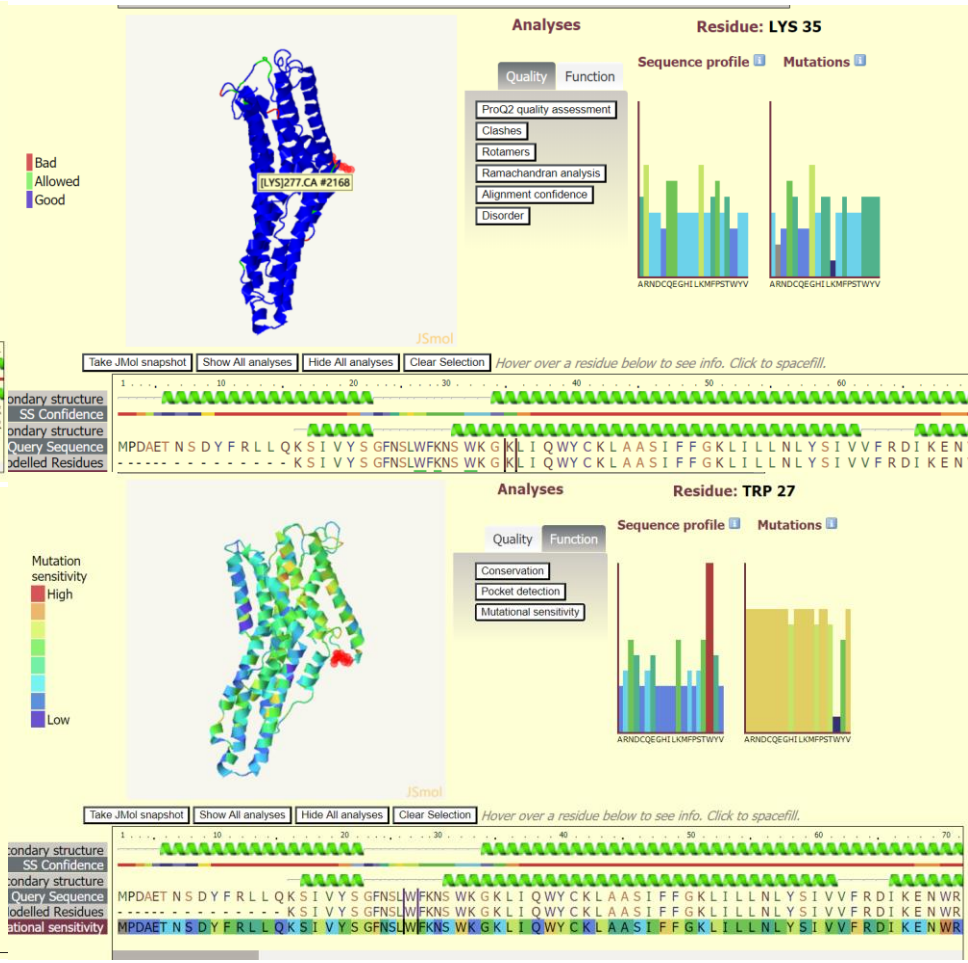
Butterwick *et al.*,2018.Nature

Phyre2-HvarOR20三维模型预测

#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	c6c70D	Alignment		100.0	12	**PDB header:**membrane protein **Chain:** D: **PDB Molecule:**odorant receptor; **PDBTitle:** cryo-em structure of orco Run Investigator
2	d2nrac1	Alignment		44.7	12	**Fold:**DNA/RNA-binding 3-helical bundle **Superfamily:**"Winged helix" DNA-binding domain **Family:**Replication initiation protein Run Investigator
3	d1hkqa	Alignment		19.5	13	**Fold:**DNA/RNA-binding 3-helical bundle **Superfamily:**"Winged helix" DNA-binding domain **Family:**Replication initiation protein Run Investigator
4	c1y66D	Alignment		14.3	26	**PDB header:**de novo protein **Chain:** D: **PDB Molecule:**engrailed homeodomain; **PDBTitle:** dioxane contributes to the altered conformation and2 oligomerization state of a designed engrailed homeodomain3 variant Run Investigator
5	c4aq9E	Alignment		11.7	12	**PDB header:**membrane protein **Chain:** E: **PDB Molecule:**acetylcholine receptor gamma subunit; **PDBTitle:** gating movement in acetylcholine receptor analysed by time-resolved2 electron cryo-microscopy (open class)

HvarOR20三维模型评估





研究计划及面对的问题

- 目前关于OR受体的晶体结构少
  - 多个蛋白建模主要用同一个模板
- Orco与OR特异性识别气味分子的机制需要进一步研究
  - 昆虫特异性与灵敏性是植保工作者关心的主要方面
- 构建的三维模型评价标准
  - AlphaFold2
- 分子对接与动力学模拟的结果分析及理论基础
  - 对分子对接结果的评价维度
  - 分子对接过程中活性位点的寻找
  - 动力学模拟的理论基础（力场、能量最小化、NPT平衡）

参考文献

Butterwick, J.A., Del Marmol, J., Kim, K.H., Kahlson, M.A., Rogow, J.A., Walz, T., and Ruta, V. (2018). Cryo-EM structure of the insect olfactory receptor Orco. Nature *560*, 447-452.

Campanacci, V., Lartigue, A., Hällberg, B.M., Jones, T.A., Giudici-Orticoni, M.-T., Tegoni, M., and Cambillau, C. (2003). Moth chemosensory protein exhibits drastic conformational changes and cooperativity on ligand binding. Proceedings of the National Academy of Sciences *100*, 5069.

Fleischer, J., Pregitzer, P., Breer, H., and Krieger, J. (2018). Access to the odor world: olfactory receptors and their role for signal transduction in insects. Cell Mol Life Sci *75*, 485-508.

Hou Xiaoqing. 2020. From sequence to function:Comparative studies of insect olfactory receptor. Lund University.

Leal, W.S. (2013). Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. Annu Rev Entomol *58*, 373-391.

参考文献

Wabitsch, Martin et al. “Biologically inactive leptin and early-onset extreme obesity.” *The New England journal of medicine* vol. 372,1 (2015): 48-54. doi:10.1056/NEJMoa1406653

Yupanqui-Lozno, Hernan et al. “Congenital Leptin Deficiency and Leptin Gene Missense Mutation Found in Two Colombian Sisters with Severe Obesity.” *Genes* vol. 10,5 342. 7 May. 2019, doi:10.3390/genes10050342

Mazen, I et al. “A novel homozygous missense mutation of the leptin gene (N103K) in an obese Egyptian pati ent.” *Molecular genetics and metabolism* vol. 97,4 (2009): 305-8. doi:10.1016/j.ymgme.2009.04.002

Shabana, and Shahida Hasnain. “The p. N103K mutation of leptin (LEP) gene and severe early onset obesity i n Pakistan.” *Biological research* vol. 49 23. 13 Apr. 2016, doi:10.1186/s40659-016-0082-7

Wabitsch, Martin et al. “Severe Early-Onset Obesity Due to Bioinactive Leptin Caused by a p.N103K Mutation in the Leptin Gene.” *The Journal of clinical endocrinology and metabolism* vol. 100,9 (2015): 3227-30. doi:10.1210/jc.2015-2263

Zhao, Yue et al. “A novel mutation in leptin gene is associated with severe obesity in Chinese individuals.” *BioMed research international* vol. 2014 (2014): 912052. doi:10.1155/2014/912052

Carpenter, Byron et al. “Structure of the human obesity receptor leptin-binding domain reveals the mechanism of leptin antagonism by a monoclonal antibody.” *Structure (London, England : 1993)* vol. 20,3 (2012): 487-97. doi:10.1016/j.str.2012.01.019

参考文献

Cline, Martha G., and Maryanne Murphy. *Obesity in the Dog and Cat.* 1st ed. CRC Press, 2019.

de Clercq, Nicolien C et al. “Gut Microbiota in Obesity and Undernutrition.” *Advances in nutrition (Bethesda, Md.)* vol. 7,6 1080-1089. 15 Nov. 2016, doi:10.3945/an.116.012914

Ernie Ward. 2018. U.S. “Pet Obesity Rates Plateau and Nutritional Confusion Grows.” Association for Pet Obesity Prevention. Accessed March 12, 2019. https://petobesityprevention.org/2018

Kim, Ki-Suk et al. “Signalling from the periphery to the brain that regulates energy homeostasis.” *Nature reviews. Neuroscience* vol. 19,4 (2018): 185-196. doi:10.1038/nrn.2018.8

Raffan, Eleanor et al. “A Deletion in the Canine POMC Gene Is Associated with Weight and Appetite in Obesity-Prone Labrador Retriever Dogs.” *Cell metabolism* vol. 23,5 (2016): 893-900. doi:10.1016/j.cmet.2016.04.012

Torres-Fuentes, Cristina et al. “The microbiota-gut-brain axis in obesity.” *The lancet. Gastroenterology & hepatology* vol. 2,10 (2017): 747-756. doi:10.1016/S2468-1253(17)30147-4

Yeo, Giles S H. “Genetics of obesity: can an old dog teach us new tricks?.” *Diabetologia* vol. 60,5 (2017): 778-783. doi:10.1007/s00125-016-4187-x

Allison, Margaret B, and Martin G Myers Jr. “20 years of leptin: connecting leptin signaling to biological function.” *The Journal of endocrinology* vol. 223,1 (2014): T25-35. doi:10.1530/JOE-14-0404

Thanks FOR YOUR ATTENTION