Primary analysis of hedgehog –the key protein of embryonic development
胚胎发育关键蛋白hedgehog初探

G04组：王志东，陈家楠，刘自珍，刘明
报告人：王志东
2019.1.12

Outline

1.背景
2.序列比对，进化树分析
3.序列预测保守结构
4.Molecular crosstalk about hedgehogs
5.蛋白同源建模
6.总结

Background

Holoprosencephaly (HPE) is a cephalic disorder in which the prosencephalon fails to develop into two hemispheres


Mutations in the gene encoding the Sonic hedgehog protein, which is involved in the development of the central nervous system (CNS), can cause holoprosencephaly

Background

Hedgehog (Hh) was first identified by genetic screens in *Drosophila melanogaster*. It earned its name from the appearance of embryos with null alleles of *hh*, which display a lawn of disorganized, hair-like bristles reminiscent of hedgehog spines. Thereafter three mammalian counterparts, sonic hedgehog (SHH), Indian hedgehog (IHH) and desert hedgehog (DHH) were found.

a) SHH activity reproduces the actions of the zone of polarizing activity in the limb bud and of the notochord and floor plate in the neural tube

b) IHH regulates bone and cartilage development and is partly redundant with SHH

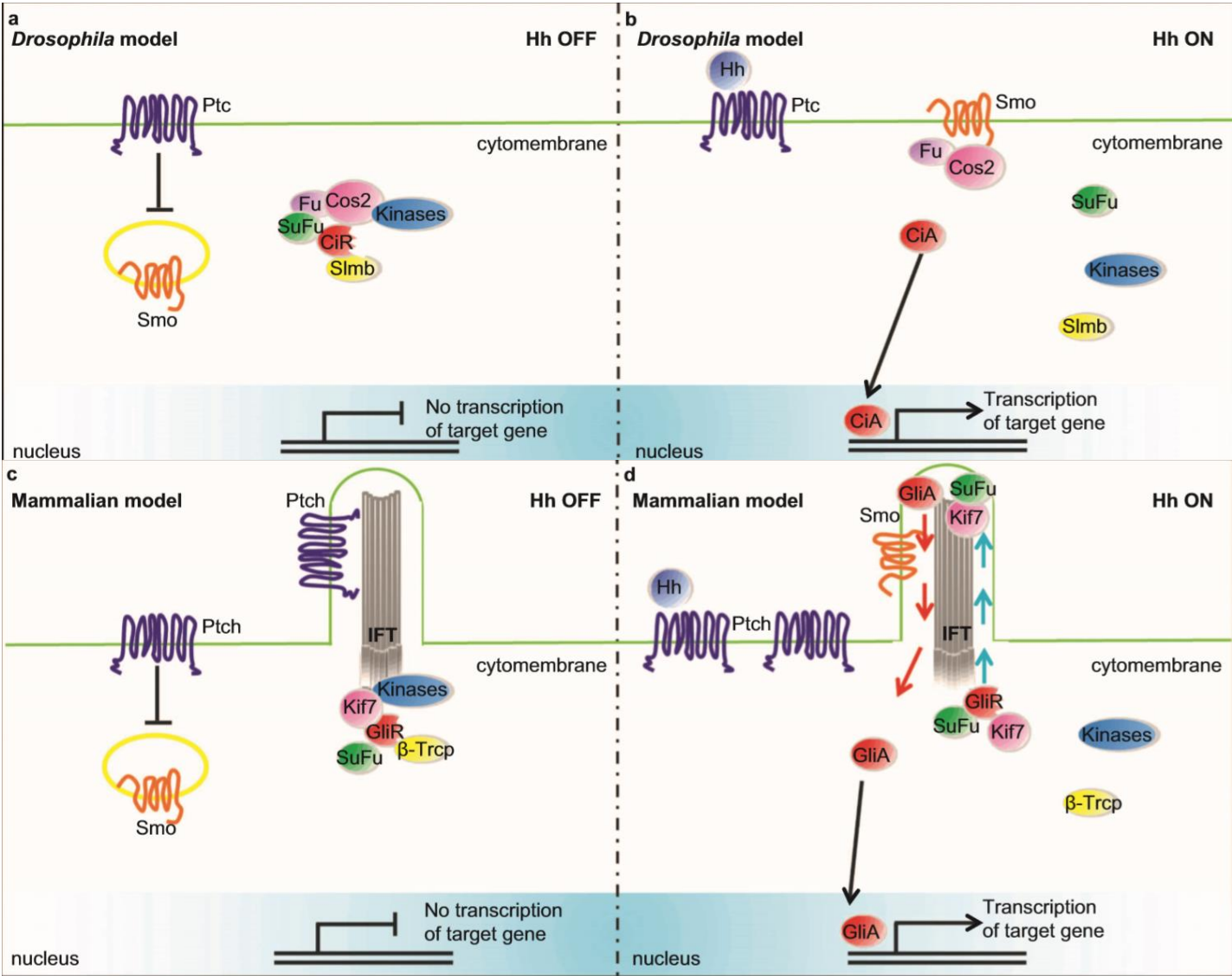c) DHH is essential for germ cell development in the testis and peripheral nerve sheath formation

Background
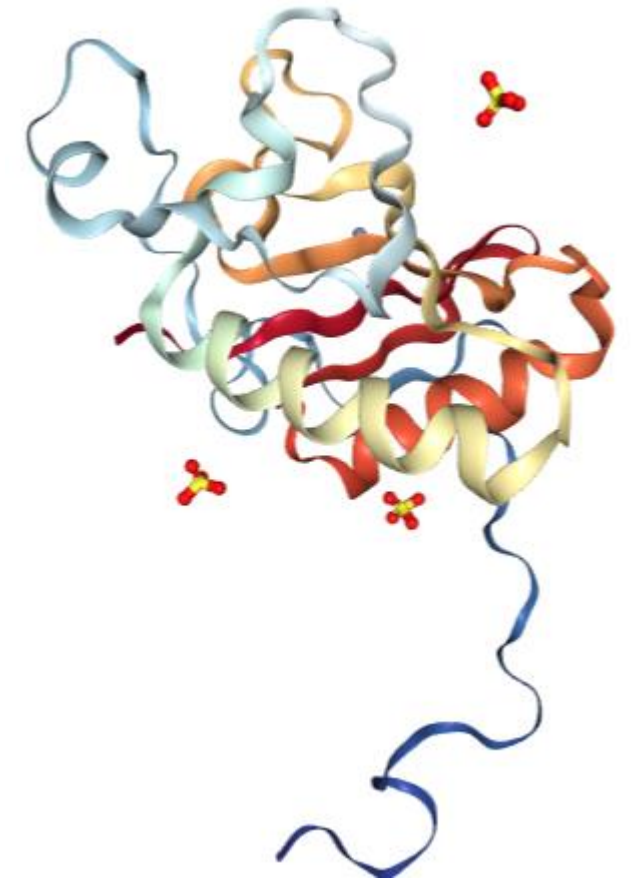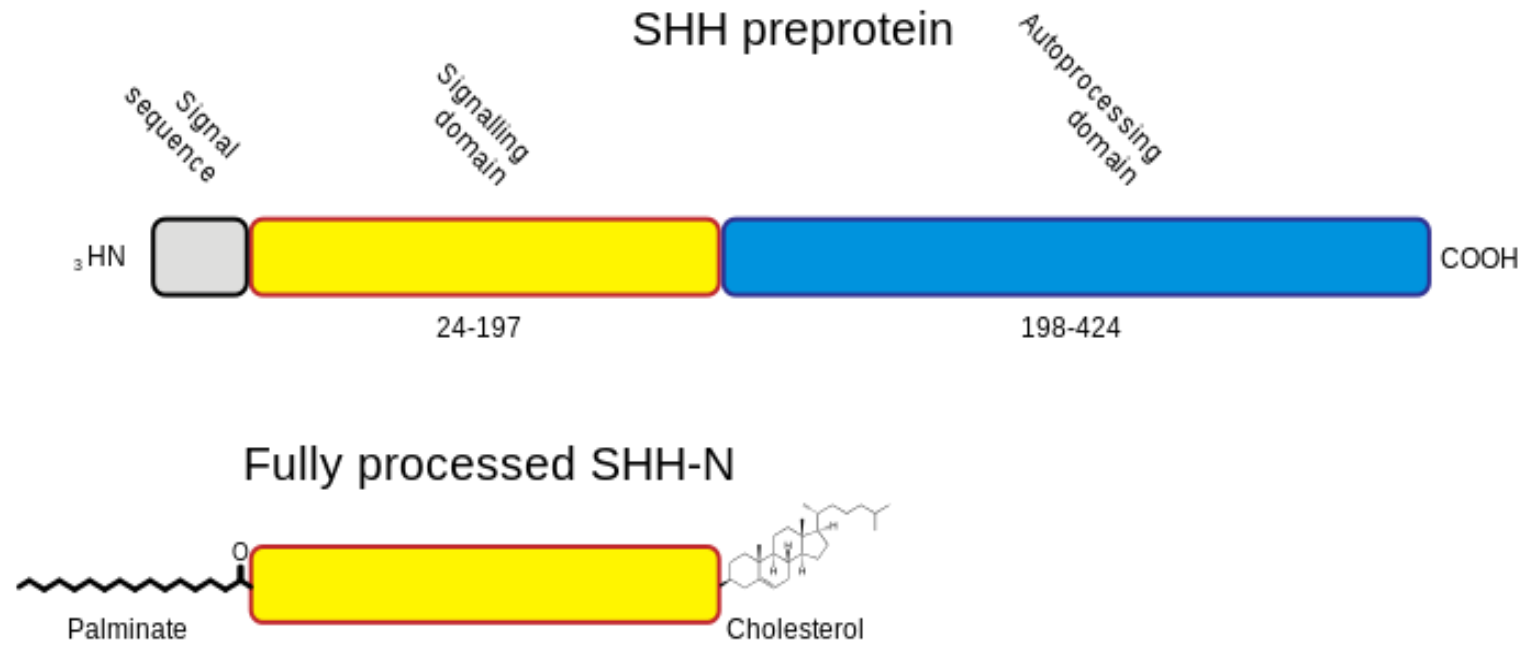
The whole sequence of Shh_human and structure of Shh-N

Hall T M. Nature, 1995, 378(6553):212

Background

The Hedgehog signaling pathway in Drosophila and vertebrates

Yang J, Andre P. *Int J Oral Sci*. 2015;7(2):73-9.

Background

Uniprot search for Shh homologous proteins in three organism

The structures of DHH_Mouse IHH_Mouse SHH_DANRE DHH_DANRE IHH_DANRE are not available in PDB

Alignment and phylogeny analysis

SHH

物种	登录号	Max score	Total score	比例
斑马鱼/人	Q92008/Q15465	567	567	64%
小鼠/斑马鱼	Q62226/Q92008	597	597	68%
小鼠/人	Q62226/Q15465	722	722	87%

DHH

物种	登录号	Max score	Total score	比例
斑马鱼/人	B3DJ46/O43323	386	402	58%
小鼠/斑马鱼	Q61488/B3DJ46	379	394	57%
小鼠/人	Q61488/O43323	767	767	96%

IHH

物种	登录号	Max score	Total score	比例
斑马鱼/人	Q98862/Q14623	449	449	63%
小鼠/斑马鱼	P97812/Q98862	468	468	63%
小鼠/人	P97812/Q14623	778	778	95%

The same protein in different organism alignment indicates that human and mouse are more closely related compared with zebrafish

Alignment and phylogeny analysis

Human

蛋白	Dhh	Ihh	Shh
Dhh	/	60%	56%
Ihh	60%	/	58%
Shh	56%	58%	/

Mouse

蛋白	Dhh	Ihh	Shh
Dhh	/	59%	60%
Ihh	59%	/	60%
Shh	60%	60%	/

Zebrafish

蛋白	Dhh	Ihh	Shh
Dhh	/	51%	49%
Ihh	51%	/	59%
Shh	49%	59%	/

In the same organism different hedgehog alignment indicates that different hedgehog is not as much similar as expected

Alignment and phylogeny analysis

Generally , Consistent with the result of alignment, but the confidence of IHH_DANRE is relatively low

Prediction of structure and function
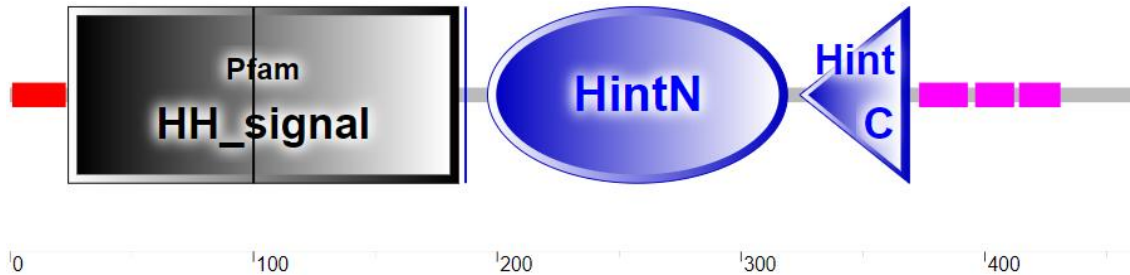
Shh_human domains

Confidently predicted domains, repeats, motifs and features:

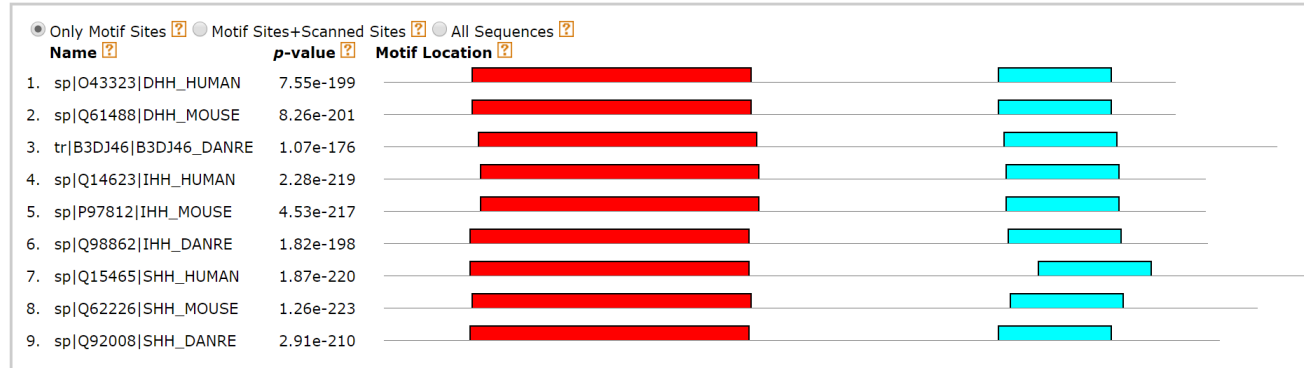| Name	Start	End	E-value
signal peptide	1	23	N/A
Pfam:HH_signal	24	184	2.7e-92
HintN	196	319	4.35e-32
HintC	324	369	0.00000293
low complexity	373	393	N/A
low complexity	396	412	N/A
low complexity	414	431	N/A

Result from SMART

Prediction of structure and function

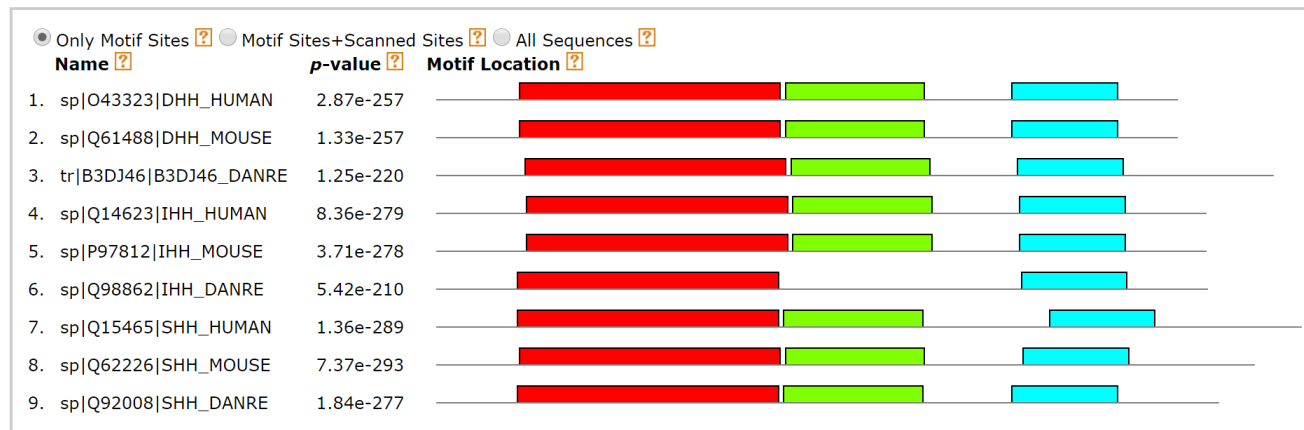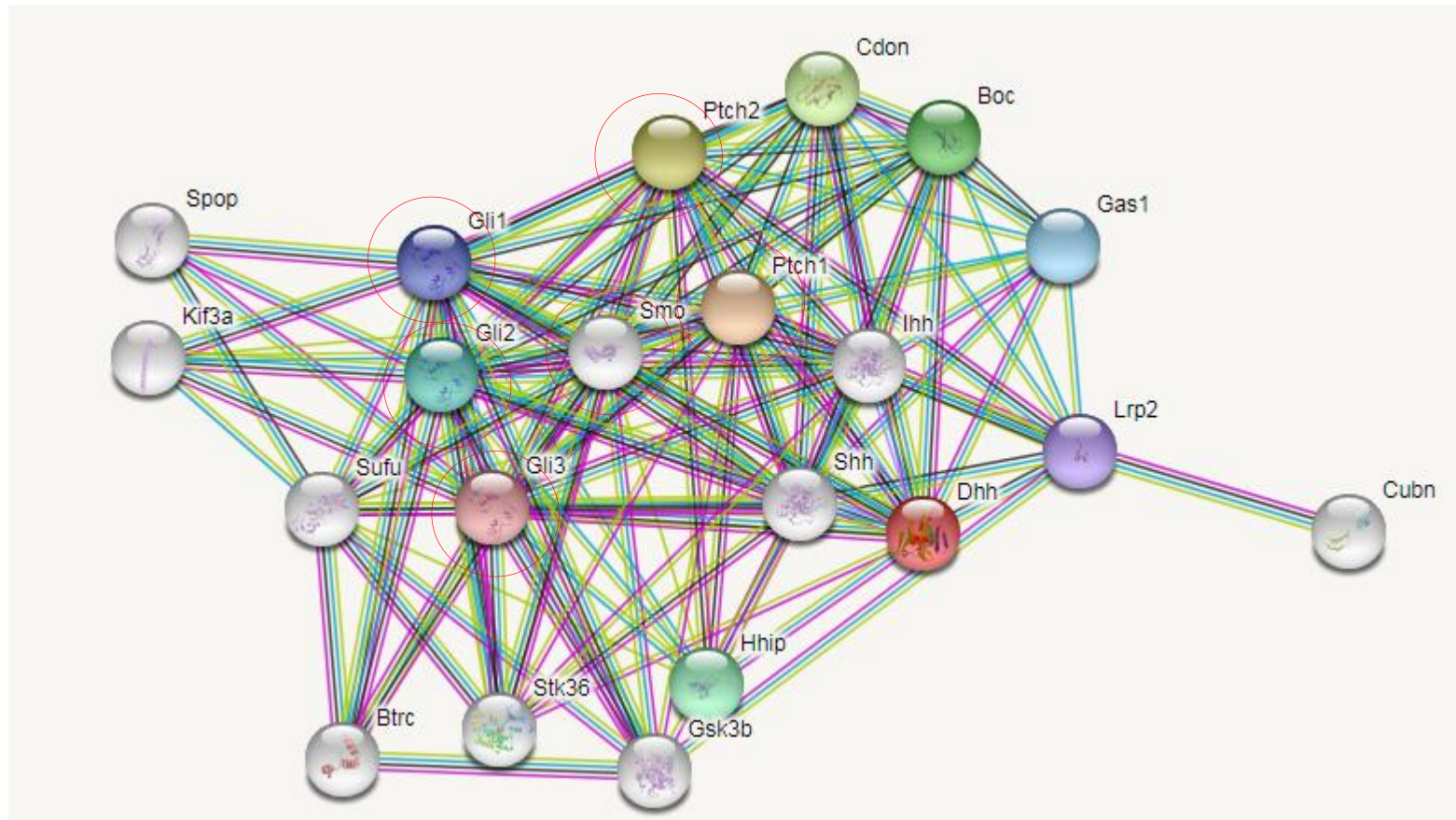## 2motifs

3motifs
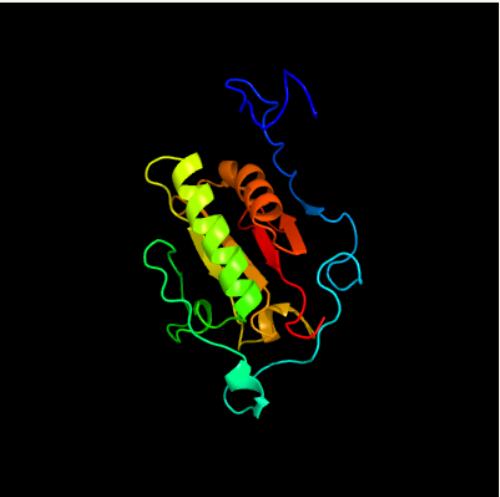

Result from MEME

Molecular crosstalk about hedgehogs

The most related proteins are Ptch1, Ptch2, Smo, Gli1.2.3

Result form STRING search

Protein modeling

a.DHH_mouse

Top model

Image coloured by rainbow N → C terminus
Model dimensions (Å): **X**:56.746 **Y**:51.673 **Z**:35.715

Model (left) based on template c3m1nB

Top template information

PDB header:signaling protein
Chain: B: PDB Molecule:sonic hedgehog protein;
PDBTitle: crystal structure of human sonic hedgehog n-terminal domain

Confidence and coverage

Confidence: **100.0%**   Coverage: **42%**

166 residues (42% of your sequence) have been modelled with 100.0% confidence by the single highest scoring template.

Additional confident templates have been detected (see Domain analysis) which cover other regions of your sequence.

322 residues (81%) could be modelled at >90% confidence using multiple-templates.

You may wish to try resubmitting your sequence in "intensive" mode to model more of your sequence.

3D viewing

Interactive 3D view in JSmol

For other options to view your downloaded structure offline see the FAQ

Good

Bad

Analyses

Quality Function

ProQ2 quality assessment
Clashes
Rotamers
Ramachandran analysis
Alignment confidence
Disorder

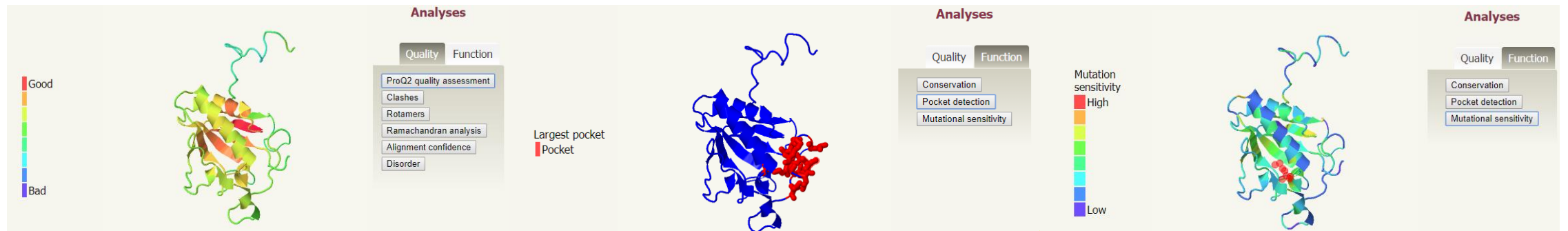Largest pocket
Pocket

Analyses

Quality Function

Conservation
Pocket detection
Mutational sensitivity

Mutation sensitivity
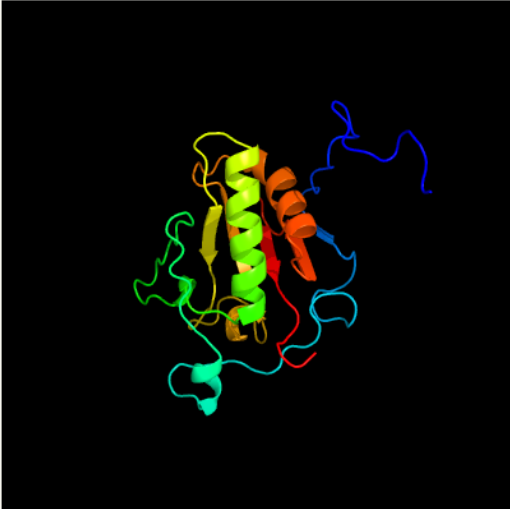High

Low

Analyses

Quality Function

Conservation
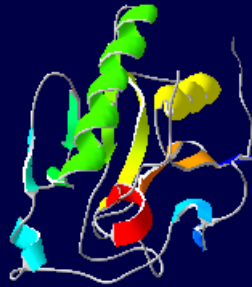Pocket detection
Mutational sensitivity

Protein modeling

b.SHH_DANRE

Top model

Model (left) based on template c3m1nB

Top template information

PDB header: signaling protein
Chain: B: PDB Molecule: sonic hedgehog protein;
PDBTitle: crystal structure of human sonic hedgehog n-terminal domain

Confidence and coverage

Confidence: **100.0%**   Coverage: **40%**

166 residues (40% of your sequence) have been modelled with 100.0% confidence by the single highest scoring template.

Additional confident templates have been detected (see Domain analysis) which cover other regions of your sequence.

320 residues (77%) could be modelled at >90% confidence using multiple-templates.
You may wish to try resubmitting your sequence in "intensive" mode to model more of your sequence.

3D viewing

Interactive 3D view in JSmol
For other options to view your downloaded structure offline see the FAQ

Image coloured by rainbow N → C terminus
Model dimensions (Å): **X:**56.746 **Y:**51.673 **Z:**34.745

Protein modeling
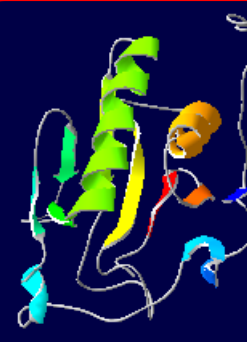
Summary

a) The evolutionary relationship between human and mouse is more close in HH family compared with zebrafish

b) Structure analysis indicate that the hedgehog contains a conservative signal domain and a auto-processing C-terminal domain(hint family)

c) Interaction with hedgehog proteins are ptch1.2, smo, Gli1.2.3, Hhip.

d) The predicted structures are highly similar with SHH_human N-terminal signal domain.