Exp 12 Structure Validation

BY YUEJIAN MO

11510511

May 31, 2018

1 Introduction

In this final dry lab experiment, we try to confirm the validation of protein structure from others. Model validation can be compared by geometry, dihedral angels from Ramachandran plot, sterics and packing from all atom clash and B-factor in existing datebase.

Here, I just mark B-factor. Firstly, the Deby-Waller factor(DWF), named after Pepter and Ivar Waller, is used in condensed matter physics to describe the attenuation of x-ran scattering or coherent neutron scattering caused by thermal motion. In the context of protein structures, the term B-factor is used. The B-factor is defined as

$$B = 8\pi^2 < u^2 >$$

where $< u^2 >$ is the mean squared displacement. It is measured in units of ångström^2, indicated the relative vibraional motion of different parts of the structure. Each ATOM record in PDB file contain a B-factor for that atom.

2 Methods

1. Analyze the two Hemoglobin structure (4HHB and 2HHB) in PDB. Find which one is better and why.
2. Analyze the 2HHB structure in PyMOL considering the effects of B-factor. Find out which region is more flexible and which region is more stable and reliable?
3. Investigate the Murthy's faked structure (1BEF) using MolProbity sever. Find out why this structure is faked.

3 Results

3.1 Model quality check of 4HHB and 2HHB within PDB database

Generaly, 2HHB show more better global quanlity than 4HHB(Figure 1,2).(I wonder why here are large difference from same paper!)

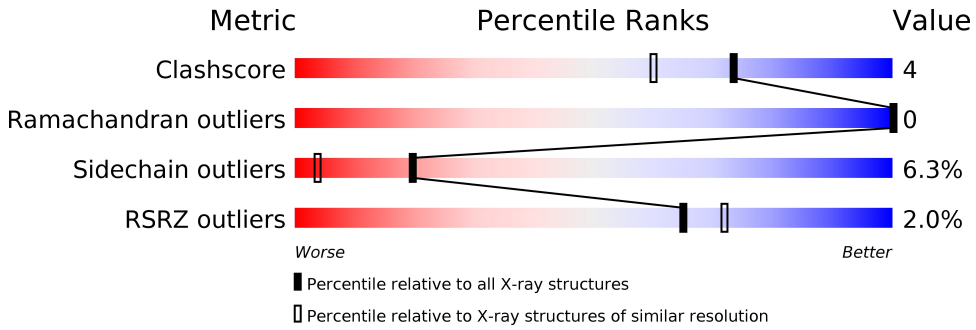

Figure 1. The 2HHB wwPDB validation from RCSB

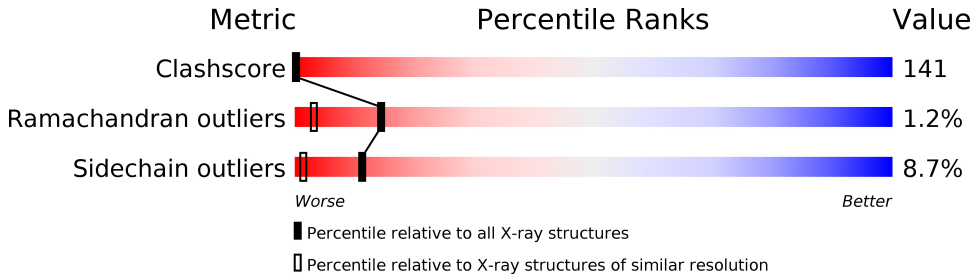

Figure 2. The 4HHB wwPDB validation from RCSB

Then I use the align for 2hhb and 4hhb, find that the RMSD is 0.159 angstroms in 568 atoms. But I find 4hhb was calculate by MolProbity without Xtriange and EDS paramters.

3.2 Comparing by using b-factor

The red color indicates higher b-factor, suggested flexible. Othewise, blue color indicates lower b-factor, suggested stable. Most red or organal colored region present in loop, turn and terminal where almost in the surface of globin.(Figure 3) It mean loop, turn and terminal, especially localated at surface, are flexible. The helix in the core are stable and reliable.

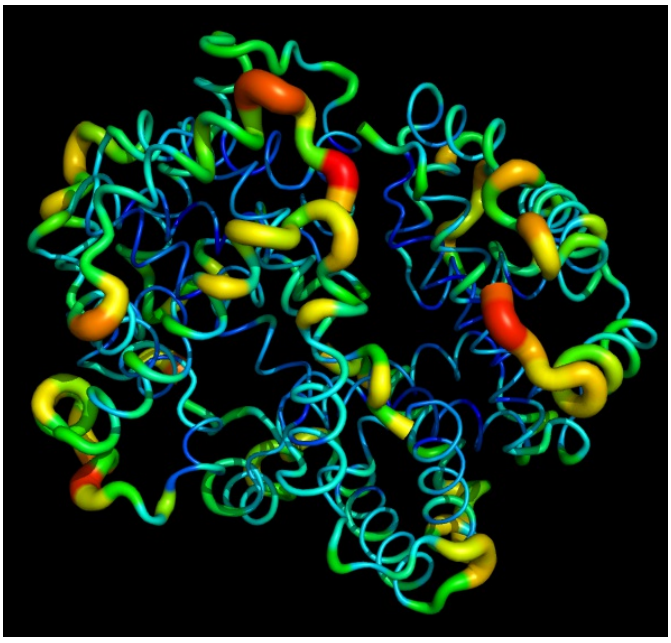

Figure 3. B-factor analysis of 2HHB. Red Color show higher b-factor, which suggest the loop, terminal and trun are more flexible than helix.

3.3 Analysis 1BEF

All atom contacts and protein geometry suggested 1BEF is faked.

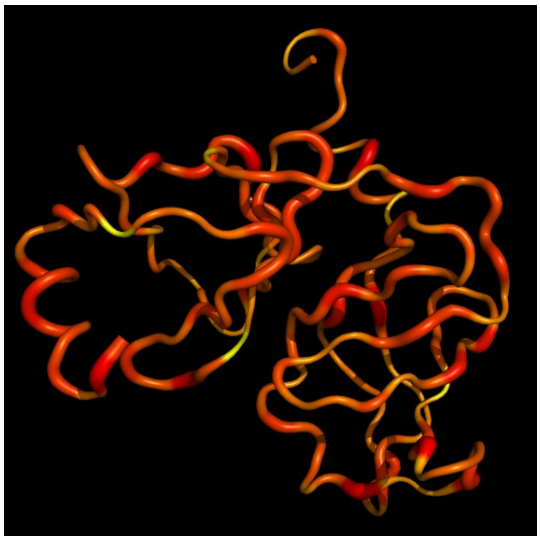

Figure 4. B-factor analysis of 1BEF. Most region are red, which suggests this structure may be faked.

All-Atom Contacts	Clashscore, all atoms:	87.8		0th percentile* (N=576, 2.10Å ± 0.25Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.			
Protein Geometry	Poor rotamers	26	18.57%	Goal: <0.3%
	Favored rotamers	105	75.00%	Goal: >98%
	Ramachandran outliers	4	2.29%	Goal: <0.05%
	Ramachandran favored	153	87.43%	Goal: >98%
	MolProbity score^	3.99		1st percentile* (N=11758, 2.10Å ± 0.25Å)
	Cβ deviations >0.25Å	1	0.66%	Goal: 0
	Bad bonds:	0 / 1367	0.00%	Goal: 0%
	Bad angles:	0 / 1850	0.00%	Goal: <0.1%
Peptide Omegas	Cis Prolines:	0 / 11	0.00%	Expected: ≤1 per chain, or ≤5%

In the two column results, the left column gives the raw count, right column gives the percentage.

* 100th percentile is the best among structures of comparable resolution; 0th percentile is the worst. For clashscore the comparative set of structures was selected in 2004, for MolProbity score in 2006.

^ MolProbity score combines the clashscore, rotamer, and Ramachandran evaluations into a single score, normalized to be on the same scale as X-ray resolution.

Figure 5. The 1bef analysis result from MolProb.

4 Conclusion

The story of Murthy's faked protein structure, awares us to care the validation of protein form public datebase. Although here are not unique index to profile the validation of protein structure, we analysis the validation from difference factor.

5 Reference

- Debye–Waller_factor— Wikipidia