wwPDB X-ray Structure Validation Summary Report (i)

Sep 13, 2020 – 07:56 PM BST

PDB ID	:	5FXN
Title	:	Structure of thermolysin solved by SAD from data collected by Direct Data Collection (DDC) using the ESRF RoboDiff goniometer
Authors	:	Bowler, M.W.; Nurizzo, D.
Deposited on	:	2016-03-02
Resolution	:	1.45 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

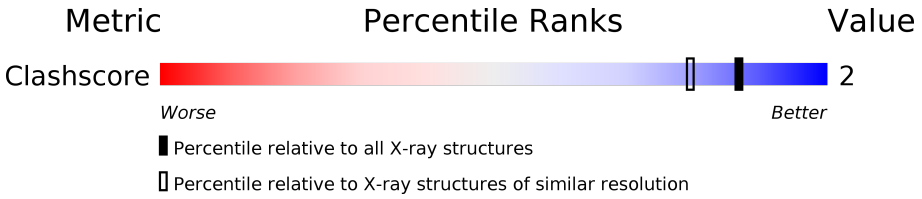
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.4.dev1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.45 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1202 (1.46-1.46)

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 2747 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

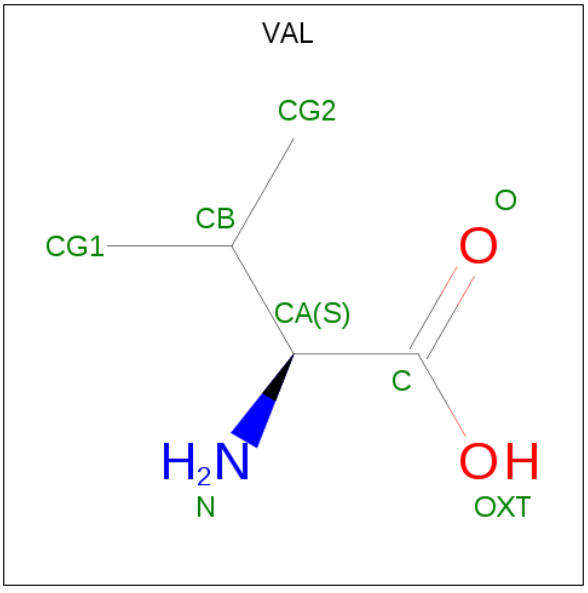
- Molecule 1 is a protein called THERMOLYSIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	314	Total	C	N	O	S	0	5	0
			2442	1537	408	494	3			

There are 2 discrepancies between the modelled and reference sequences:

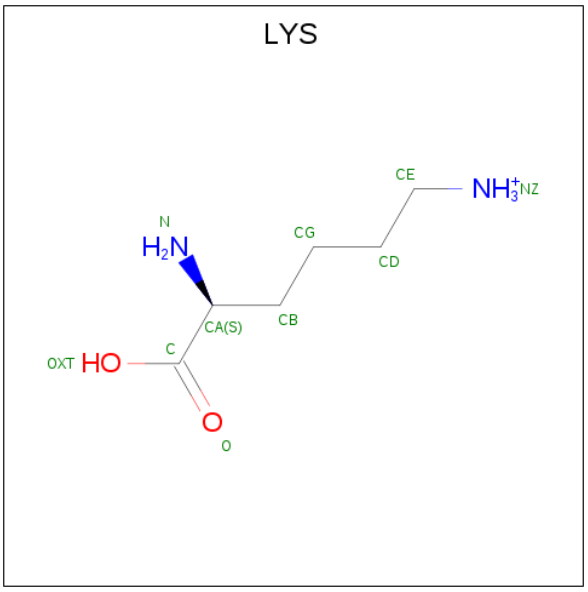
Chain	Residue	Modelled	Actual	Comment	Reference
A	37	ASP	ASN	conflict	UNP P00800
A	119	GLU	GLN	conflict	UNP P00800

- Molecule 2 is VALINE (three-letter code: VAL) (formula: $C_5H_{11}NO_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			7	5	1	1		

- Molecule 3 is LYSINE (three-letter code: LYS) (formula: $C_6H_{15}N_2O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 10 6 2 2	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Zn 1 1	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	4	Total Ca 4 4	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	283	Total O 283 283	0	0

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics (i)

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, α, β, γ	93.10Å 93.10Å 130.30Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	80.63 – 1.45	Depositor
% Data completeness (in resolution range)	90.2 (80.63-1.45)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.25 (at 1.45Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.144 , 0.180	Depositor
Wilson B-factor (Å^2)	18.3	Xtriage
Anisotropy	0.233	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2747	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.85% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

4 Model quality

4.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, CA

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.74	1/2517 (0.0%)	0.78	1/3427 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	166	GLU	CD-OE2	-6.51	1.18	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	150	ASP	CB-CG-OD1	5.24	123.01	118.30

There are no chirality outliers.

There are no planarity outliers.

4.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2442	0	2281	10	0
2	A	7	0	8	0	0
3	A	10	0	13	0	0
4	A	1	0	0	0	0
5	A	4	0	0	0	0
6	A	283	0	0	6	2

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	2747	0	2302	10	2

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

The worst 5 of 10 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:2:THR:N	6:A:2003:HOH:O	2.07	0.88
1:A:89:ASN:HB2	6:A:2116:HOH:O	1.96	0.66
1:A:11:ARG:NH1	6:A:2024:HOH:O	2.16	0.47
1:A:100:ILE:HG23	1:A:120[B]:MET:CE	2.45	0.46
1:A:89:ASN:CB	6:A:2116:HOH:O	2.60	0.45

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:2163:HOH:O	6:A:2163:HOH:O[12_555]	0.52	1.68
6:A:2161:HOH:O	6:A:2161:HOH:O[12_555]	1.06	1.14

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

There are no protein backbone outliers to report in this entry.

4.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

4.3.3 RNA (i)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 5 are monoatomic - leaving 2 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

5.2 Non-standard residues in protein, DNA, RNA chains

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

5.4 Ligands

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers

EDS failed to run properly - this section is therefore empty.