Full wwPDB X-ray Structure Validation Report (i)

Dec 14, 2023 – 04:59 am GMT

PDB ID	:	3ZU1
Title	:	Structure of LysB29(Nepsilon omega-carboxyheptadecanoyl) des(B30) Human Insulin
Authors	:	Schluckebier, G.; Whittingham, J.
Deposited on	:	2011-07-13
Resolution	:	1.60 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

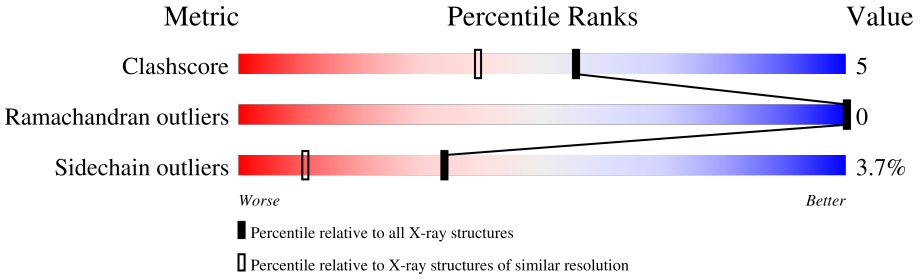
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.60 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	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 1811 atoms, of which 927 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.

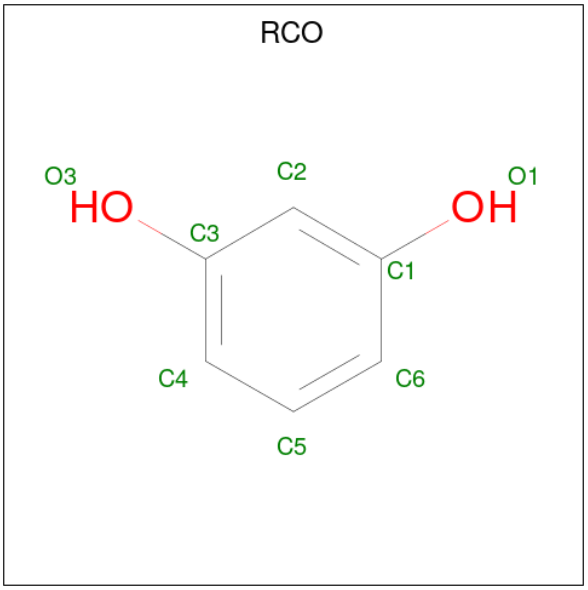
- Molecule 1 is a protein called INSULIN A CHAIN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	21	294	91	142	25	32	4	12	0	0
1	C	21	312	99	149	25	35	4	13	0	0

- Molecule 2 is a protein called INSULIN B CHAIN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	28	464	154	233	39	36	2	10	3	0
2	D	28	481	164	239	38	38	2	11	4	0

- Molecule 3 is RESORCINOL (three-letter code: RCO) (formula: $C_6H_6O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
3	B	1	14	6	6	2	1	0
3	D	1	14	6	6	2	1	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
4	B	1	1	1	0	0
4	D	1	1	1	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
5	B	1	1	1	0	0
5	D	1	1	1	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	H	O		
6	A	19	57	38	19	0	0
6	B	21	63	42	21	0	0
6	C	23	69	46	23	0	0
6	D	13	39	26	13	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	H 3	Depositor
Cell constants a, b, c, α, β, γ	79.06Å 79.06Å 41.02Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	21.88 – 1.60	Depositor
% Data completeness (in resolution range)	99.2 (21.88-1.60)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.42 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.165 , 0.214	Depositor
Wilson B-factor (Å^2)	27.3	Xtriage
Anisotropy	0.362	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.028 for h,-h-k,-l	Xtriage
Total number of atoms	1811	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.39% 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 (i)

4.1 Standard geometry (i)

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

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	0.94	0/152	0.74	0/204
1	C	0.95	0/164	0.93	0/220
2	B	1.17	0/240	1.05	1/327 (0.3%)
2	D	1.25	1/258 (0.4%)	1.07	0/350
All	All	1.12	1/814 (0.1%)	0.98	1/1101 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	16	TYR	CD2-CE2	-5.72	1.30	1.39

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	16	TYR	CB-CG-CD2	6.74	125.05	121.00

There are no chirality outliers.

There are no planarity outliers.

4.2 Too-close contacts (i)

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	152	142	138	1	0
1	C	163	149	149	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	231	233	227	3	0
2	D	242	239	245	4	0
3	B	8	6	5	0	0
3	D	8	6	6	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
5	B	1	0	0	1	0
5	D	1	0	0	0	0
6	A	19	38	0	0	1
6	B	21	42	0	3	0
6	C	23	46	0	0	0
6	D	13	26	0	1	1
All	All	884	927	770	8	1

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 5.

All (8) 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)
5:B:1031:CL:CL	6:B:2021:HOH:O	2.20	0.94
1:C:2:ILE:HG12	2:D:11[A]:LEU:HD21	1.58	0.85
1:C:21:ASN:HD21	2:D:25:PHE:HD1	1.50	0.60
2:D:1:PHE:HB2	6:D:2001:HOH:O	2.01	0.60
1:A:13:LEU:HD22	2:B:18[A]:VAL:CG1	2.43	0.48
2:B:3:ASN:ND2	6:B:2001:HOH:O	2.31	0.46
1:C:13:LEU:HG	2:D:18:VAL:CG2	2.49	0.42
2:B:28:PRO:C	6:B:2020:HOH:O	2.60	0.40

All (1) 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:2013:HOH:H1	6:D:2010:HOH:H2[3_555]	1.30	0.30

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	19/21 (90%)	19 (100%)	0	0	100	100
1	C	19/21 (90%)	19 (100%)	0	0	100	100
2	B	29/30 (97%)	29 (100%)	0	0	100	100
2	D	30/30 (100%)	30 (100%)	0	0	100	100
All	All	97/102 (95%)	97 (100%)	0	0	100	100

There are no Ramachandran outliers to report.

4.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	18/20 (90%)	18 (100%)	0	100	100
1	C	20/20 (100%)	19 (95%)	1 (5%)	24	6
2	B	24/26 (92%)	24 (100%)	0	100	100
2	D	27/26 (104%)	25 (93%)	2 (7%)	13	2
All	All	89/92 (97%)	86 (97%)	3 (3%)	34	13

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	4	GLU
2	D	1	PHE
2	D	27	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	21	ASN
2	D	4	GLN

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 6 ligands modelled in this entry, 4 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 (i)

5.1 Protein, DNA and RNA chains (i)

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

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

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

5.3 Carbohydrates (i)

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

5.4 Ligands (i)

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

5.5 Other polymers (i)

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