



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 05:04 AM GMT

PDB ID : 4EX1
Title : Human Insulin
Authors : Favero-Retto, M.P.; Palmieri, L.C.; Lima, L.M.T.R.
Deposited on : 2012-04-29
Resolution : 1.66 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

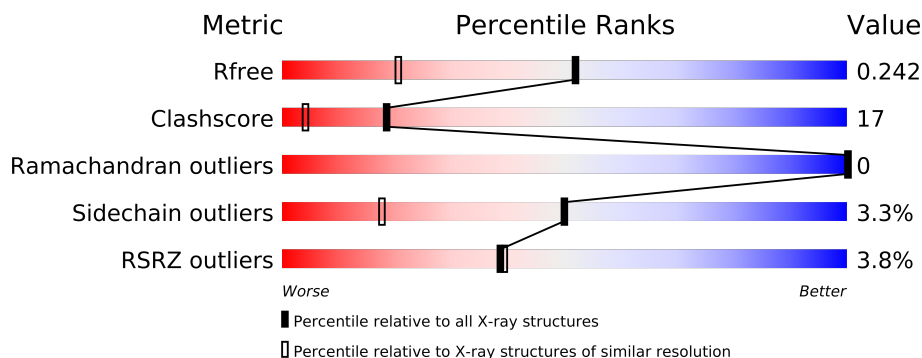
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance





The reported resolution of this entry is 1.66 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))
R_{free}	66092	1404 (1.68-1.64)
Clashscore	79885	1001 (1.66-1.66)
Ramachandran outliers	78287	1581 (1.68-1.64)
Sidechain outliers	78261	1580 (1.68-1.64)
RSRZ outliers	66119	1404 (1.68-1.64)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	21	
1	C	21	
2	B	30	
2	D	30	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 996 atoms, of which 0 are hydrogen and 0 are deuterium.

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
1	A	21	Total	C	N	O	S	0	0	0
			162	99	25	34	4			
1	C	21	Total	C	N	O	S	0	2	0
			176	108	27	37	4			

- Molecule 2 is a protein called Insulin B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	30	Total	C	N	O	S	0	4	0
			275	180	47	46	2			
2	D	30	Total	C	N	O	S	0	3	0
			265	176	43	44	2			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	25	Total 25	O 25	0	0
5	B	33	Total 33	O 33	0	0
5	C	25	Total 25	O 25	0	0
5	D	31	Total 31	O 31	0	0

i

- Molecule 1: Insulin A chain

Diagram illustrating a sequence of nodes: G1, I2, C7, L13, L16, E17, and N21. The nodes are connected sequentially by green lines, indicating a flow or relationship between them.

G1	I2	V3	Y19	C20	N21
----	----	----	-----	-----	-----

```

graph LR
    F1[F1] --- L11[L11]
    L11 --- L17[L17]
    L17 --- R22[R22]
    R22 --- F25[F25]
    F25 --- T30[T30]
    style F1 fill:#00FF00
    style L11 fill:#00FF00
    style L17 fill:#FFFF00
    style R22 fill:#FFFF00
    style F25 fill:#00FF00
    style T30 fill:#FFFF00
  
```

4 Data and refinement statistics

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	80.67Å 80.67Å 33.21Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.33 – 1.66 16.74 – 1.66	Depositor EDS
% Data completeness (in resolution range)	99.8 (40.33-1.66) 99.9 (16.74-1.66)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	6.35 (at 1.65Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.164 , 0.239 0.176 , 0.242	Depositor DCC
R_{free} test set	459 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	14.9	Xtriage
Anisotropy	0.050	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 40.1	EDS
Estimated twinning fraction	0.031 for h,-h-k,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 9564 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	996	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.10% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.79	0/163	0.90	0/219
1	C	0.90	0/177	0.93	0/239
2	B	0.87	0/282	1.06	2/381 (0.5%)
2	D	0.78	0/273	0.89	0/369
All	All	0.84	0/895	0.95	2/1208 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	17[A]	LEU	CA-CB-CG	5.01	126.82	115.30
2	B	17[B]	LEU	CA-CB-CG	5.01	126.82	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	29	LYS	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	162	0	149	7	0
1	C	176	0	163	10	0
2	B	275	0	269	17	0
2	D	265	0	254	6	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
5	A	25	0	0	2	0
5	B	33	0	0	1	0
5	C	25	0	0	0	0
5	D	31	0	0	3	0
All	All	996	0	835	29	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 17.

The worst 5 of 29 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:7:CYS:O	2:B:5:HIS:HD2	1.67	0.76
1:C:2[B]:ILE:HG21	1:C:19:TYR:CE2	2.22	0.75
1:C:2[B]:ILE:HG12	1:C:19:TYR:CZ	2.24	0.71
2:B:29:LYS:O	2:B:30:THR:HB	1.89	0.71
2:B:4:GLN:HE21	2:B:6:LEU:HD21	1.59	0.67

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	21/21 (100%)	19 (90%)	2 (10%)	0	100	100
2	B	32/30 (107%)	32 (100%)	0	0	100	100
2	D	31/30 (103%)	31 (100%)	0	0	100	100
All	All	103/102 (101%)	101 (98%)	2 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	20/20 (100%)	19 (95%)	1 (5%)	34	8
1	C	22/20 (110%)	22 (100%)	0	100	100
2	B	30/26 (115%)	27 (90%)	3 (10%)	11	1
2	D	29/26 (112%)	29 (100%)	0	100	100
All	All	101/92 (110%)	97 (96%)	4 (4%)	50	13

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	13	LEU
2	B	17[A]	LEU
2	B	17[B]	LEU
2	B	30	THR

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

Mol	Chain	Res	Type
2	B	4	GLN
2	B	5	HIS

5.3.3 RNA ⓘ

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 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.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	21/21 (100%)	-0.16	0 100 100	18, 24, 29, 41	0
1	C	21/21 (100%)	-0.19	0 100 100	16, 21, 23, 27	0
2	B	30/30 (100%)	0.16	4 (13%) 4 3	15, 20, 47, 50	0
2	D	30/30 (100%)	0.02	0 100 100	15, 19, 32, 43	0
All	All	102/102 (100%)	-0.02	4 (3%) 38 38	15, 21, 40, 50	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	2	VAL	3.1
2	B	1	PHE	2.9
2	B	4	GLN	2.6
2	B	30	THR	2.1

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
3	ZN	D	101	1/1	0.09	1.64	16,16,16,16	1
3	ZN	B	101	1/1	0.06	-0.58	13,13,13,13	1
4	CL	D	102	1/1	0.18	-	44,44,44,44	1
4	CL	B	102	1/1	0.08	-	32,32,32,32	1

6.5 Other polymers ⓘ

There are no such residues in this entry.