wwPDB X-ray Structure Validation Summary Report (i)

PDFCREATIONDATE INFOmissingINFO

PDB ID	:	2WBY
Title	:	Crystal structure of human insulin-degrading enzyme in complex with insulin
Authors	:	Manolopoulou, M.; Guo, Q.; Malito, E.; Schilling, A.B.; Tang, W.J.
Deposited on	:	2009-03-06
Resolution	:	2.60 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

PERCENTILES INFOmissingINFO

1 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 16926 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 INSULIN-DEGRADING ENZYME.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	958	7813	5032	1313	1445	23	0	0	1
1	B	955	7787	5018	1307	1440	22	0	0	1

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	110	LEU	CYS	engineered mutation	UNP P14735
A	111	GLN	GLU	engineered mutation	UNP P14735
A	171	SER	CYS	engineered mutation	UNP P14735
A	178	ALA	CYS	engineered mutation	UNP P14735
A	257	VAL	CYS	engineered mutation	UNP P14735
A	414	LEU	CYS	engineered mutation	UNP P14735
A	573	ASN	CYS	engineered mutation	UNP P14735
A	590	SER	CYS	engineered mutation	UNP P14735
A	789	SER	CYS	engineered mutation	UNP P14735
A	812	ALA	CYS	engineered mutation	UNP P14735
A	819	ALA	CYS	engineered mutation	UNP P14735
A	904	SER	CYS	engineered mutation	UNP P14735
A	966	ASN	CYS	engineered mutation	UNP P14735
A	974	ALA	CYS	engineered mutation	UNP P14735
B	110	LEU	CYS	engineered mutation	UNP P14735
B	111	GLN	GLU	engineered mutation	UNP P14735
B	171	SER	CYS	engineered mutation	UNP P14735
B	178	ALA	CYS	engineered mutation	UNP P14735
B	257	VAL	CYS	engineered mutation	UNP P14735
B	414	LEU	CYS	engineered mutation	UNP P14735
B	573	ASN	CYS	engineered mutation	UNP P14735
B	590	SER	CYS	engineered mutation	UNP P14735
B	789	SER	CYS	engineered mutation	UNP P14735
B	812	ALA	CYS	engineered mutation	UNP P14735
B	819	ALA	CYS	engineered mutation	UNP P14735

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	904	SER	CYS	engineered mutation	UNP P14735
B	966	ASN	CYS	engineered mutation	UNP P14735
B	974	ALA	CYS	engineered mutation	UNP P14735

- Molecule 2 is a protein called INSULIN A CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	20	155	95	23	33	4	0	0	0
2	E	20	155	95	23	33	4	0	0	0

- Molecule 3 is a protein called INSULIN B CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	18	143	94	24	24	1	0	0	0
3	F	19	149	97	25	25	2	0	0	0

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

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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	A	408	408	408	0	0
5	B	295	295	295	0	0
5	C	6	6	6	0	0
5	D	4	4	4	0	0
5	E	5	5	5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	F	4	4	4	0	0

SEQUENCE-PLOTS INFOmissingINFO

2 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	262.32Å 262.32Å 90.61Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	32.08 – 2.60 31.46 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.9 (32.08-2.60) 99.9 (31.46-2.60)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.39 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.164 , 0.218 0.168 , 0.220	Depositor DCC
R_{free} test set	5438 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	31.3	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 50.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.019 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	16926	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.82% 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.

3 Model quality (i)

3.1 Standard geometry (i)

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

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	1.21	22/8007 (0.3%)	1.09	34/10833 (0.3%)
1	B	1.19	21/7982 (0.3%)	1.04	31/10801 (0.3%)
2	C	1.25	0/156	1.11	0/209
2	E	1.22	0/156	1.04	0/209
3	D	1.55	3/146 (2.1%)	1.11	0/198
3	F	1.42	1/152 (0.7%)	1.03	0/206
All	All	1.20	47/16599 (0.3%)	1.06	65/22456 (0.3%)

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
1	A	0	1

The worst 5 of 47 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	189	GLU	CD-OE2	17.00	1.44	1.25
1	A	189	GLU	CD-OE1	15.73	1.43	1.25
1	A	189	GLU	CG-CD	8.99	1.65	1.51
1	A	577	GLU	CG-CD	8.46	1.64	1.51
1	B	508	GLU	CG-CD	8.38	1.64	1.51

The worst 5 of 65 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	189	GLU	CG-CD-OE2	-14.71	88.88	118.30
1	A	189	GLU	CG-CD-OE1	13.60	145.50	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	189	GLU	CG-CD-OE1	-11.84	94.62	118.30
1	A	460	ARG	NE-CZ-NH2	-10.47	115.07	120.30
1	B	189	GLU	CG-CD-OE2	9.24	136.78	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	979	ASN	Peptide

3.2 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	7813	0	7744	146	0
1	B	7787	0	7715	161	0
2	C	155	0	144	4	0
2	E	155	0	143	7	0
3	D	143	0	142	13	0
3	F	149	0	146	11	0
4	A	1	0	0	2	0
4	B	1	0	0	3	0
5	A	408	0	0	24	0
5	B	295	0	0	14	0
5	C	6	0	0	0	0
5	D	4	0	0	1	0
5	E	5	0	0	0	0
5	F	4	0	0	0	0
All	All	16926	0	16034	316	0

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 10.

The worst 5 of 316 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:491:ARG:HG3	1:A:491:ARG:HH11	1.02	1.17
4:B:3012:ZN:ZN	3:F:1:PHE:N	1.11	1.14
4:A:3012:ZN:ZN	3:D:1:PHE:N	1.17	1.06
1:B:189:GLU:OE2	3:F:1:PHE:N	1.90	1.03
1:B:119:LYS:HE3	1:B:171:SER:HB2	1.44	0.98

There are no symmetry-related clashes.

3.3 Torsion angles (i)

3.3.1 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	954/990 (96%)	918 (96%)	36 (4%)	0	100	100
1	B	951/990 (96%)	921 (97%)	29 (3%)	1 (0%)	51	75
2	C	18/20 (90%)	16 (89%)	2 (11%)	0	100	100
2	E	18/20 (90%)	17 (94%)	1 (6%)	0	100	100
3	D	16/19 (84%)	12 (75%)	4 (25%)	0	100	100
3	F	17/19 (90%)	17 (100%)	0	0	100	100
All	All	1974/2058 (96%)	1901 (96%)	72 (4%)	1 (0%)	51	75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1010	PRO

3.3.2 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	848/879 (96%)	765 (90%)	83 (10%)	8	15
1	B	845/879 (96%)	761 (90%)	84 (10%)	8	15
2	C	19/19 (100%)	18 (95%)	1 (5%)	22	45
2	E	19/19 (100%)	18 (95%)	1 (5%)	22	45
3	D	16/17 (94%)	14 (88%)	2 (12%)	4	8
3	F	17/17 (100%)	15 (88%)	2 (12%)	5	9
All	All	1764/1830 (96%)	1591 (90%)	173 (10%)	8	15

5 of 173 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	460	ARG
1	B	758	LEU
1	B	524	LEU
1	B	643	LYS
1	B	823	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 62 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	805	ASN
1	B	805	ASN
1	B	184	ASN
1	B	788	ASN
1	B	957	HIS

3.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

3.5 Carbohydrates (i)

There are no monosaccharides in this entry.

3.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 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.

No monomer is involved in short contacts.

3.7 Other polymers (i)

There are no such residues in this entry.

3.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

4 Fit of model and data (i)

4.1 Protein, DNA and RNA chains (i)

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	958/990 (96%)	-0.73	4 (0%)	92	91	13, 24, 40, 71	0
1	B	955/990 (96%)	-0.66	5 (0%)	91	89	15, 28, 44, 78	0
2	C	20/20 (100%)	1.16	6 (30%)	0	0	22, 52, 60, 61	0
2	E	20/20 (100%)	1.35	8 (40%)	0	0	21, 46, 63, 67	0
3	D	18/19 (94%)	1.59	8 (44%)	0	0	26, 52, 56, 57	0
3	F	19/19 (100%)	1.55	7 (36%)	0	0	32, 49, 63, 64	0
All	All	1990/2058 (96%)	-0.61	38 (1%)	66	62	13, 26, 47, 78	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	16	TYR	6.6
3	F	16	TYR	5.4
2	E	20	CYS	5.4
2	E	19	TYR	4.6
2	C	19	TYR	4.3

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

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

4.3 Carbohydrates (i)

There are no monosaccharides in this entry.

4.4 Ligands (i)

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. 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	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	ZN	B	3012	1/1	0.97	0.21	2,2,2,2	0
4	ZN	A	3012	1/1	0.99	0.23	2,2,2,2	0

4.5 Other polymers (i)

There are no such residues in this entry.