

wwPDB X-ray Structure Validation Summary Report (i)

Jun 23, 2024 – 03:32 AM EDT

PDB ID	:	4WID
Title	:	Crystal structure of the immediate-early 1 protein (IE1) at 2.31 angstrom
		(tetragonal form after crystal dehydration)
Authors	:	Klingl, S.; Scherer, M.; Sevvana, M.; Muller, Y.A.; Stamminger, T.
Deposited on	:	2014-09-25
Resolution	:	2.31 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:

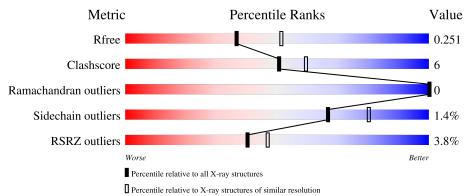
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 2.31 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R_{free}	130704	5974(2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	$5855\ (2.34-2.30)$

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain		
1	А	365	4%	12% ••	
2	В	365	4% 79%	16% · ·	



2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5833 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 RhUL123.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
1	А	353	Total 2867	C 1823	N 482	0 544	S 18	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	31	GLY	-	expression tag	UNP Q2FAE9
А	32	PRO	-	expression tag	UNP Q2FAE9
А	33	LEU	-	expression tag	UNP Q2FAE9
А	34	GLY	-	expression tag	UNP Q2FAE9
А	35	SER	-	expression tag	UNP Q2FAE9

• Molecule 2 is a protein called RhUL123.

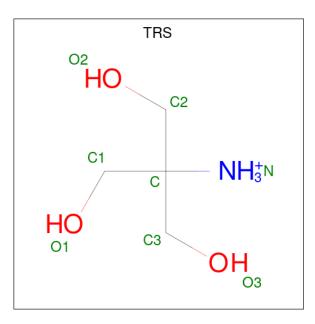
Mo	l Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
2	В	349	Total 2827	C 1803	N 476	O 530	S 18	0	0	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
В	31	GLY	-	expression tag	UNP Q2FAE9
В	32	PRO	-	expression tag	UNP Q2FAE9
В	33	LEU	-	expression tag	UNP Q2FAE9
В	34	GLY	-	expression tag	UNP Q2FAE9
В	35	SER	-	expression tag	UNP Q2FAE9

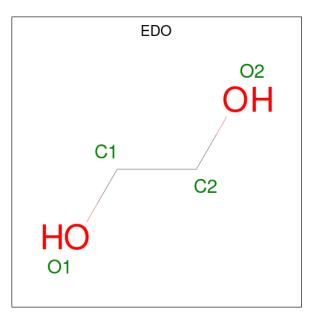
• Molecule 3 is 2-AMINO-2-HYDROXYMETHYL-PROPANE-1,3-DIOL (three-letter code: TRS) (formula: $C_4H_{12}NO_3$).





Μ	lol	Chain	Residues	A	ton	ns		ZeroOcc	AltConf
	3	А	1	Total 8	С 4	N 1	O 3	0	0

• Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 4 & 2 & 2 \end{array}$	0	0
4	В	1	$\begin{array}{ccc} \text{Total} & \text{C} & \text{O} \\ 4 & 2 & 2 \end{array}$	0	0

• Molecule 5 is water.

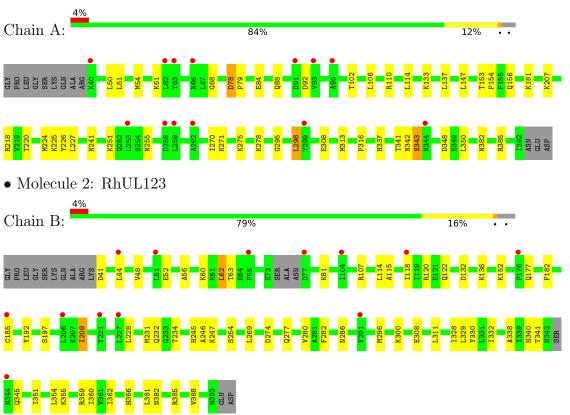


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	А	70	Total O 70 70	0	0
5	В	53	Total O 53 53	0	0



3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.



• Molecule 1: RhUL123



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43	Depositor
Cell constants	56.49Å 56.49 Å 276.91 Å	Depositor
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.18 - 2.31	Depositor
Resolution (A)	48.18 - 2.31	EDS
% Data completeness	99.2 (48.18-2.31)	Depositor
(in resolution range)	99.3 (48.18 - 2.31)	EDS
R _{merge}	0.06	Depositor
R _{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$2.06 (at 2.32 \text{\AA})$	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
P. P.	0.197 , 0.250	Depositor
R, R_{free}	0.200 , 0.251	DCC
R_{free} test set	1923 reflections (5.10%)	wwPDB-VP
Wilson B-factor $(Å^2)$	44.8	Xtriage
Anisotropy	0.351	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.33, 50.7	EDS
L-test for twinning ²	$< L >=0.46, < L^2>=0.28$	Xtriage
Estimated twinning fraction	0.079 for h,-k,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5833	wwPDB-VP
Average B, all atoms $(Å^2)$	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.94% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.



¹Intensities estimated from amplitudes.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: TRS, MLZ, MLY, EDO

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		
	Chain	RMSZ	# Z > 5	RMSZ	# Z > 5	
1	А	0.41	0/2844	0.52	1/3828~(0.0%)	
2	В	0.41	0/2824	0.51	0/3799	
All	All	0.41	0/5668	0.51	1/7627~(0.0%)	

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	A	298	LEU	CA-CB-CG	5.14	127.12	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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	А	2867	0	2911	32	0
2	В	2827	0	2874	40	0
3	А	8	0	12	0	0
4	В	8	0	12	0	0
5	А	70	0	0	3	0
5	В	53	0	0	3	0
All	All	5833	0	5809	66	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 6.

The worst 5 of 66 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)
2:B:81:LYS:O	2:B:107:ARG:NH2	2.16	0.78
1:A:348:ASP:OD1	5:A:501:HOH:O	2.11	0.68
2:B:63:THR:O	2:B:120:ARG:NH2	2.26	0.66
1:A:110:ARG:O	1:A:114:LEU:HG	2.00	0.62
1:A:255:ASN:O	2:B:177:GLN:NE2	2.31	0.61

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 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	Perce	Percentiles	
1	А	345/365~(94%)	338~(98%)	7~(2%)	0	100	100	
2	В	339/365~(93%)	334 (98%)	5 (2%)	0	100	100	
All	All	684/730~(94%)	672~(98%)	12 (2%)	0	100	100	

There are no Ramachandran outliers to report.

5.3.2 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 Outlier		Percentiles		
1	А	316/326~(97%)	312~(99%)	4 (1%)	69 81		
2	В	311/328~(95%)	306~(98%)	5(2%)	62 77		
All	All	627/654~(96%)	618~(99%)	9 (1%)	67 80		

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	В	254	SER
2	В	366	HIS
1	А	343	SER
2	В	52	GLU
2	В	62	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	В	345	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

10 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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| > 2 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link Bond lengths				Bond angles		
10101	Type	Chain	nes		Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z >2
1	MLZ	А	207	1	8,9,10	0.66	0	4,9,11	1.32	1 (25%)
1	MLY	А	181	1	9,10,11	0.68	0	6,11,13	0.58	0
1	MLY	А	199	1	9,10,11	0.59	0	6,11,13	0.60	0
2	MLY	В	355	2	9,10,11	0.62	0	$6,\!11,\!13$	0.39	0
2	MLY	В	179	2	9,10,11	0.70	0	6,11,13	0.33	0



Mol	Turne	Chain	Dec	Tinle	Link Bond lengths			Bond angles		
IVIOI	Type	Chain	Res		Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z >2
1	MLZ	А	61	1	8,9,10	0.58	0	4,9,11	1.46	1 (25%)
1	MLY	А	209	1	9,10,11	0.67	0	6,11,13	0.33	0
2	MLZ	В	247	2	8,9,10	0.70	0	4,9,11	1.45	1 (25%)
1	MLZ	А	251	1	8,9,10	0.66	0	4,9,11	1.55	1 (25%)
2	MLY	В	225	2	9,10,11	0.68	0	$6,\!11,\!13$	0.55	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	MLZ	А	207	1	-	4/7/8/10	-
1	MLY	А	181	1	-	1/8/9/11	-
1	MLY	А	199	1	-	0/8/9/11	-
2	MLY	В	355	2	-	1/8/9/11	-
2	MLY	В	179	2	_	0/8/9/11	-
1	MLZ	А	61	1	-	2/7/8/10	-
1	MLY	А	209	1	-	0/8/9/11	-
2	MLZ	В	247	2	-	4/7/8/10	-
1	MLZ	А	251	1	-	1/7/8/10	-
2	MLY	В	225	2	-	2/8/9/11	_

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Ζ	$\mathbf{Observed}(^{o})$	$Ideal(^{o})$
1	А	251	MLZ	CM-NZ-CE	2.93	120.41	111.95
2	В	247	MLZ	CM-NZ-CE	2.88	120.27	111.95
1	А	61	MLZ	CM-NZ-CE	2.72	119.80	111.95
1	А	207	MLZ	CM-NZ-CE	2.58	119.41	111.95

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	А	207	MLZ	N-CA-CB-CG
1	А	207	MLZ	C-CA-CB-CG
2	В	225	MLY	C-CA-CB-CG
2	В	247	MLZ	C-CA-CB-CG

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	А	181	MLY	CD-CE-NZ-CH1

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	А	181	MLY	2	0
2	В	355	MLY	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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| > 2 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	e Chain	Res	Dog	Dog	Dog	Dec	Dec	Dec	Dec	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z >2	Counts	RMSZ	# Z >2							
4	EDO	В	401	-	3,3,3	0.45	0	2,2,2	0.53	0							
3	TRS	А	401	-	7,7,7	0.27	0	9,9,9	0.27	0							
4	EDO	В	402	-	3,3,3	0.42	0	2,2,2	0.45	0							

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	EDO	В	401	-	-	1/1/1/1	-
3	TRS	А	401	-	-	6/9/9/9	-
4	EDO	В	402	-	-	0/1/1/1	-



There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	А	401	TRS	N-C-C2-O2
3	А	401	TRS	C3-C-C1-O1
3	А	401	TRS	N-C-C1-O1
3	А	401	TRS	C2-C-C1-O1
3	А	401	TRS	C2-C-C3-O3

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.



6 Fit of model and data (i)

6.1 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, 95^{th} 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	$\langle RSRZ \rangle$	#RSRZ>2	$OWAB(Å^2)$	Q<0.9
1	А	347/365~(95%)	0.23	13 (3%) 41 48	33, 54, 98, 139	0
2	В	345/365~(94%)	0.41	13 (3%) 40 47	38, 62, 94, 117	0
All	All	692/730~(94%)	0.32	26 (3%) 40 47	33, 59, 97, 139	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	91	ASP	4.4
1	А	63	THR	3.8
1	А	258	ASP	3.8
2	В	44	LEU	3.5
1	А	253	LEU	3.3

6.2 Non-standard residues in protein, DNA, RNA chains (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, 95^{th} 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
1	MLZ	А	251	10/11	0.90	0.15	80,85,87,91	0
2	MLY	В	179	11/12	0.91	0.30	70,77,79,81	0
1	MLZ	А	61	10/11	0.94	0.19	76,78,83,84	0
2	MLY	В	225	11/12	0.94	0.16	$56,\!64,\!68,\!69$	0
1	MLY	А	209	11/12	0.95	0.14	46,51,61,66	0
1	MLZ	А	207	10/11	0.95	0.17	$46,\!49,\!55,\!57$	0
2	MLY	В	355	11/12	0.95	0.13	68,71,77,77	0
1	MLY	А	181	11/12	0.96	0.13	37,42,59,60	0
2	MLZ	В	247	10/11	0.97	0.20	40,51,64,64	0
1	MLY	А	199	11/12	0.99	0.18	44,47,61,62	0



6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 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, 95^{th} 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
3	TRS	А	401	8/8	0.75	0.23	$101,\!102,\!104,\!105$	0
4	EDO	В	402	4/4	0.92	0.16	74,75,75,76	0
4	EDO	В	401	4/4	0.95	0.24	75,76,76,77	0

6.5 Other polymers (i)

There are no such residues in this entry.

