

# Full wwPDB X-ray Structure Validation Report (i)

#### Jun 23, 2024 – 03:22 AM EDT

se variant.

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 (1)) were used in the production of this report:

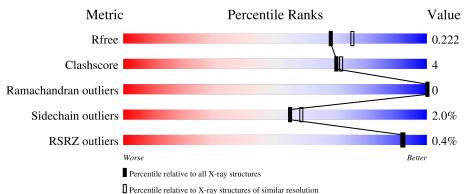
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as $543$ be (2022)
Xtriage (Phenix)	:	1.20.1
$\mathrm{EDS}$	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal 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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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	512	89%	9%	•



#### 6QSQ

# 2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4086 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 Alkaline phosphatase.

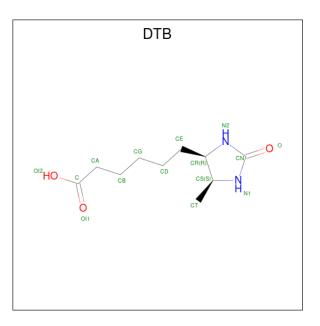
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	А	503	Total 3827	C 2403	N 648	O 763	S 13	0	2	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	41	VAL	ILE	conflict	UNP A0A2N7KUA1
А	79	SER	GLY	conflict	UNP A0A2N7KUA1
А	336	LEU	ARG	engineered mutation	UNP A0A2N7KUA1
А	503	LEU	-	expression tag	UNP A0A2N7KUA1
А	504	ALA	-	expression tag	UNP A0A2N7KUA1
А	505	TRP	-	expression tag	UNP A0A2N7KUA1
А	506	SER	-	expression tag	UNP A0A2N7KUA1
А	507	HIS	-	expression tag	UNP A0A2N7KUA1
А	508	PRO	-	expression tag	UNP A0A2N7KUA1
А	509	GLN	-	expression tag	UNP A0A2N7KUA1
А	510	PHE	-	expression tag	UNP A0A2N7KUA1
А	511	GLU	-	expression tag	UNP A0A2N7KUA1
А	512	LYS	-	expression tag	UNP A0A2N7KUA1

• Molecule 2 is 6-(5-METHYL-2-OXO-IMIDAZOLIDIN-4-YL)-HEXANOIC ACID (three-letter code: DTB) (formula:  $C_{10}H_{18}N_2O_3$ ).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	А	1	Total 15	C 10	N 2	O 3	0	0

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

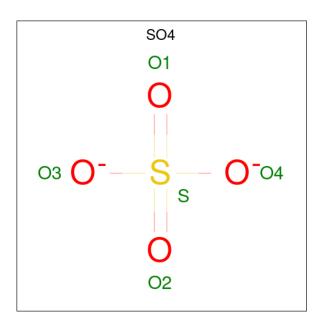
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	А	2	Total Zn 2 2	0	0

• Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

• Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).





Mol	Chain	Residues	Aton	ns	ZeroOcc	AltConf
5	A	1	Total ( 5	O S 4 1	0	0

• Molecule 6 is water.

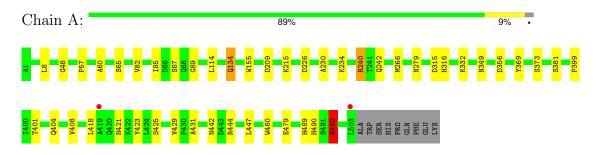
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	А	236	Total         O           236         236	0	0



# 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: Alkaline phosphatase





## 4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants	98.08Å 118.60Å 83.92Å	Depositor
a, b, c, $\alpha$ , $\beta$ , $\gamma$	$90.00^{\circ}$ $90.00^{\circ}$ $90.00^{\circ}$	Depositor
Resolution (Å)	48.43 - 2.00	Depositor
Resolution (A)	48.43 - 2.00	EDS
% Data completeness	99.9 (48.43-2.00)	Depositor
(in resolution range)	99.9 (48.43-2.00)	EDS
R <sub>merge</sub>	0.06	Depositor
R <sub>sym</sub>	(Not available)	Depositor
$< I/\sigma(I) > 1$	$3.30 (at 2.00 \text{\AA})$	Xtriage
Refinement program	REFMAC	Depositor
D D.	0.169 , $0.218$	Depositor
$R, R_{free}$	0.179 , $0.222$	DCC
$R_{free}$ test set	1610 reflections $(4.82\%)$	wwPDB-VP
Wilson B-factor $(Å^2)$	28.3	Xtriage
Anisotropy	0.151	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.34 , $48.5$	EDS
L-test for twinning <sup>2</sup>	$ < L >=0.48, < L^2>=0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	4086	wwPDB-VP
Average B, all atoms $(Å^2)$	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 4.29% of the height of the origin peak. No significant pseudotranslation is detected.

<sup>&</sup>lt;sup>2</sup>Theoretical 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.



<sup>&</sup>lt;sup>1</sup>Intensities 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: SO4, ZN, DTB, MG

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	Bo	nd lengths	Bo	nd angles
	Chain	RMSZ	# Z  > 5	RMSZ	# Z  > 5
1	А	0.90	4/3906~(0.1%)	0.89	8/5307~(0.2%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	А	479	GLU	CD-OE1	6.11	1.32	1.25
1	А	492	GLU	CD-OE2	-6.11	1.19	1.25
1	А	492	GLU	CB-CG	-5.86	1.41	1.52
1	А	492	GLU	CG-CD	-5.50	1.43	1.51

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
1	А	240	ARG	NE-CZ-NH1	7.00	123.80	120.30
1	А	492	GLU	OE1-CD-OE2	6.81	131.47	123.30
1	А	356	ASP	CB-CG-OD1	5.86	123.57	118.30
1	А	492	GLU	N-CA-CB	5.46	120.44	110.60
1	А	492	GLU	CB-CA-C	-5.33	99.73	110.40
1	А	240	ARG	NE-CZ-NH2	-5.25	117.67	120.30
1	А	87	SER	N-CA-CB	-5.23	102.65	110.50
1	А	356	ASP	CB-CG-OD2	-5.08	113.73	118.30

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



Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	А	3827	0	3606	27	0
2	А	15	0	17	0	0
3	А	2	0	0	0	0
4	А	1	0	0	0	0
5	А	5	0	0	1	0
6	А	236	0	0	2	2
All	All	4086	0	3623	27	2

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.

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 4.

All (27) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom 1	A + 2	Interatomic	Clash
Atom-1	Atom-2	distance $(\text{\AA})$	overlap (Å)
1:A:65:SER:OG	5:A:605:SO4:S	2.22	0.98
1:A:226:ASP:H	1:A:242:GLN:HE22	1.35	0.71
1:A:48:GLY:O	6:A:701:HOH:O	2.12	0.66
1:A:134:GLN:HE21	1:A:134:GLN:H	1.49	0.59
1:A:429:VAL:HG12	1:A:442:ASN:OD1	2.03	0.57
1:A:431:ALA:O	1:A:447:LEU:HD21	2.07	0.55
1:A:421:HIS:CD2	1:A:423:TYR:H	2.24	0.54
1:A:369:TYR:O	1:A:373:SER:HB2	2.07	0.54
1:A:57:PRO:HG2	1:A:60:ALA:HB3	1.89	0.53
1:A:408:VAL:O	1:A:444:ARG:NH1	2.38	0.52
1:A:401:THR:OG1	1:A:404:GLN:HG3	2.12	0.49
1:A:421:HIS:HD2	1:A:423:TYR:H	1.59	0.49
1:A:226:ASP:N	1:A:242:GLN:HE22	2.08	0.48
1:A:134:GLN:HE21	1:A:134:GLN:N	2.11	0.47
1:A:85:ILE:HD12	1:A:89:GLY:HA2	1.97	0.47
1:A:234:LYS:HB3	1:A:240:ARG:HG2	1.97	0.47
1:A:489:HIS:O	1:A:492:GLU:HG2	2.15	0.46
1:A:349:ASN:HD22	1:A:349:ASN:HA	1.61	0.44
1:A:226:ASP:H	1:A:242:GLN:NE2	2.08	0.43
1:A:114:LEU:HD23	1:A:155:TRP:CZ2	2.53	0.43
1:A:65:SER:HB3	6:A:711:HOH:O	2.19	0.42
1:A:230:ALA:O	1:A:234:LYS:HG3	2.20	0.41
1:A:315:ASP:CG	1:A:316:HIS:CD2	2.94	0.41
1:A:8:LEU:O	1:A:266:MET:HA	2.21	0.41
1:A:418:LEU:O	1:A:425:SER:HA	2.21	0.41

Continued on next page...



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:82:VAL:HB	1:A:85:ILE:HD11	2.02	0.40
1:A:279:ASN:HB3	1:A:460:TRP:CE2	2.57	0.40

Continued from previous page...

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:793:HOH:O	6:A:793:HOH:O[4_555]	0.68	1.52
6:A:817:HOH:O	6:A:817:HOH:O[4_555]	1.09	1.11

### 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		Allowed	Outliers	Percentiles
1	А	503/512~(98%)	491 (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 side chain 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	Analysed Rotameric Outl		Percentiles
1	А	393/425~(92%)	385~(98%)	8 (2%)	55 58

All (8) residues with a non-rotameric sidechain are listed below:



Mol	Chain	Res	Type
1	А	134	GLN
1	А	209	ASP
1	А	215	LYS
1	А	332	LYS
1	А	381	SER
1	А	399	PRO
1	А	490	HIS
1	А	492	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	А	43	GLN
1	А	134	GLN
1	А	242	GLN
1	А	279	ASN
1	А	334	GLN
1	А	349	ASN
1	А	421	HIS
1	А	433	ASN
1	А	446	ASN
1	А	457	ASN
1	А	467	HIS
1	А	500	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)

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

## 5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry (i)

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 for Mogul analysis.



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 Li		Link	Bo	ond leng	ths	В	ond ang	les		
	Type	Unam	nes	nes		Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z >2
5	SO4	А	605	3	4,4,4	1.11	0	$6,\!6,\!6$	1.97	1 (16%)	
2	DTB	А	601	-	15,15,15	0.76	0	16,19,19	1.13	2 (12%)	

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
2	DTB	А	601	-	-	6/8/20/20	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
5	А	605	SO4	O4-S-O2	-4.19	87.67	109.56
2	А	601	DTB	CG-CD-CE	-2.47	104.30	113.62
2	А	601	DTB	CD-CE-CR	2.20	118.14	113.94

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	А	601	DTB	CD-CE-CR-N2
2	А	601	DTB	C-CA-CB-CG
2	А	601	DTB	CA-CB-CG-CD
2	А	601	DTB	CG-CD-CE-CR
2	А	601	DTB	OI1-C-CA-CB
2	А	601	DTB	OI2-C-CA-CB

There are no ring outliers.

1 monomer is involved in 1 short contact:



Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	А	605	SO4	1	0

## 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	<RSRZ $>$	#RSRZ>2		$\mathbf{OWAB}(\mathbf{A}^2)$	Q<0.9
1	А	503/512~(98%)	-0.23	2 (0%) 92	92	17, 29, 47, 67	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	А	419	ALA	2.7
1	А	503	LEU	2.4

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

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

### 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	$\mathbf{B} ext{-factors}(\mathbf{A}^2)$	Q<0.9
2	DTB	А	601	15/15	0.79	0.22	48,50,54,54	0
5	SO4	А	605	5/5	0.98	0.08	21,22,26,29	0
4	MG	А	604	1/1	0.99	0.08	24,24,24,24	0
3	ZN	А	602	1/1	0.99	0.10	23,23,23,23	0
3	ZN	А	603	1/1	1.00	0.08	21,21,21,21	0



## 6.5 Other polymers (i)

There are no such residues in this entry.

