



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 16, 2017 – 10:55 PM EDT

PDB ID : 3DZO
Title : Crystal structure of a rhopty kinase from toxoplasma gondii
Authors : Wernimont, A.K.; Lam, A.; Ali, A.; Lin, Y.H.; Ni, S.; Ravichandran, M.; Wasney, G.; Vedadi, M.; Kozieradzki, I.; Schapira, M.; Bochkarev, A.; Wilkstrom, M.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.M.; Sibley, D.; Hui, R.; Qiu, W.; Structural Genomics Consortium (SGC)
Deposited on : unknown
Resolution : 1.80 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20030345

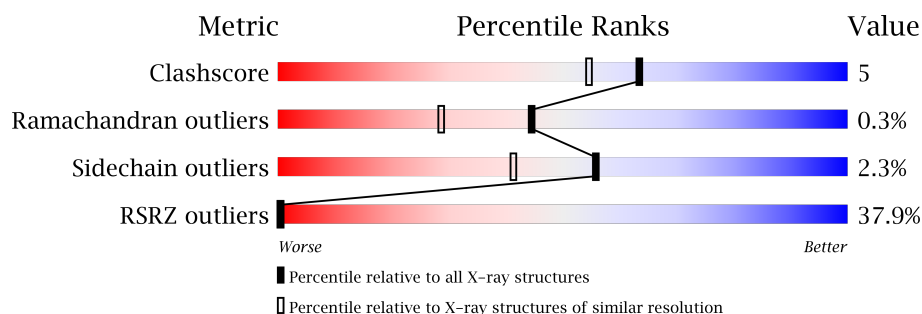
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	112137	5742 (1.80-1.80)
Ramachandran outliers	110173	5676 (1.80-1.80)
Sidechain outliers	110143	5675 (1.80-1.80)
RSRZ outliers	101464	4906 (1.80-1.80)

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. 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 $\leq 5\%$. 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	A	413	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 2828 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 Rhoptyr kinase domain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	348	Total	C	N	O	S	0	7	0
			2709	1755	460	487	7			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	178	MET	-	EXPRESSION TAG	PDB 3DZO
A	179	HIS	-	EXPRESSION TAG	PDB 3DZO
A	180	HIS	-	EXPRESSION TAG	PDB 3DZO
A	181	HIS	-	EXPRESSION TAG	PDB 3DZO
A	182	HIS	-	EXPRESSION TAG	PDB 3DZO
A	183	HIS	-	EXPRESSION TAG	PDB 3DZO
A	184	HIS	-	EXPRESSION TAG	PDB 3DZO
A	185	SER	-	EXPRESSION TAG	PDB 3DZO
A	186	SER	-	EXPRESSION TAG	PDB 3DZO
A	187	GLY	-	EXPRESSION TAG	PDB 3DZO
A	188	ARG	-	EXPRESSION TAG	PDB 3DZO
A	189	GLU	-	EXPRESSION TAG	PDB 3DZO
A	190	ASN	-	EXPRESSION TAG	PDB 3DZO
A	191	LEU	-	EXPRESSION TAG	PDB 3DZO
A	192	TYR	-	EXPRESSION TAG	PDB 3DZO
A	193	PHE	-	EXPRESSION TAG	PDB 3DZO
A	194	GLN	-	EXPRESSION TAG	PDB 3DZO
A	195	GLY	-	EXPRESSION TAG	PDB 3DZO

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	118	Total 118	O 118	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	41.26 Å 44.29 Å 105.30 Å 90.00° 98.90° 90.00°	Depositor
Resolution (Å)	35.00 – 1.80 29.99 – 1.80	Depositor EDS
% Data completeness (in resolution range)	97.6 (35.00-1.80) 93.6 (29.99-1.80)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.04 (at 1.79 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.244 , 0.270 0.237 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	DCC
Wilson B-factor (Å ²)	32.8	Xtriage
Anisotropy	0.505	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 47.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.024 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	2828	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 6.62% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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.

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: 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	Bond lengths		Bond angles	
		RMSZ	$\# Z > 5$	RMSZ	$\# Z > 5$
1	A	0.39	0/2800	0.53	0/3820

There are no bond length outliers.

There are no bond angle outliers.

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	A	2709	0	2584	26	0
2	A	1	0	0	0	0
3	A	118	0	0	0	0
All	All	2828	0	2584	26	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 5.

The worst 5 of 26 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:246:VAL:HB	1:A:334:VAL:HG11	1.31	1.12

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:401:LEU:HD11	1:A:414:LEU:HD22	1.57	0.82
1:A:246:VAL:HB	1:A:334:VAL:CG1	2.11	0.78
1:A:208:GLU:OE1	1:A:212:ARG:HD2	1.84	0.77
1:A:319[A]:LYS:H	1:A:319[A]:LYS:HD2	1.50	0.76

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	Percentiles
1	A	347/413 (84%)	333 (96%)	13 (4%)	1 (0%)	44 29

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	337	PRO

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	Outliers	Percentiles
1	A	273/357 (76%)	266 (97%)	7 (3%)	51 36

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	319[B]	LYS
1	A	470	LEU
1	A	413	TYR
1	A	238	TRP
1	A	437	ARG

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

Mol	Chain	Res	Type
1	A	190	ASN
1	A	546	GLN

5.3.3 RNA ⓘ

There are no RNA molecules 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 1 ligands modelled in this entry, 1 is 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.

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, 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	348/413 (84%)	1.92	132 (37%) 0 0	20, 31, 35, 37	8 (2%)

The worst 5 of 132 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	397[A]	VAL	8.6
1	A	258	VAL	8.3
1	A	235	THR	8.2
1	A	393	LEU	7.5
1	A	469	THR	6.7

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 carbohydrates 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. 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	RSCC	RSR	LLDF	B-factors(Å ²)	Q<0.9
2	MG	A	601	1/1	0.95	0.08	-	47,47,47,47	0

6.5 Other polymers [i](#)

There are no such residues in this entry.