wwPDB X-ray Structure Validation Summary Report (i)

Jul 29, 2021 – 12:10 PM EDT

PDB ID	:	1ZVW
Title	:	The Crystal Structure of TrpD (Rv2192c) from Mycobacterium tuberculosis in Complex with PRPP and Magnesium
Authors	:	Lee, C.E.; Lott, J.S.; Baker, E.N.; Arcus, V.L.; Javid-Majd, F.; Goodfellow, C.; Hung, L.-W.; TB Structural Genomics Consortium (TBSGC)
Deposited on	:	2005-06-02
Resolution	:	2.30 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 following versions of software and data (see references (i)) were used in the production of this report:

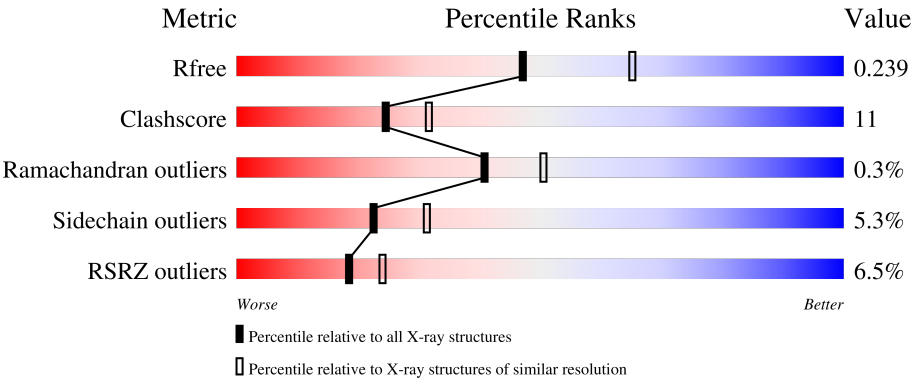
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.22
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.22

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	A	378	6% 70% 20% · 8%
1	B	378	6% 69% 17% · 12%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 5186 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 Anthranilate phosphoribosyltransferase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	346	Total	C	N	O	S	Se	0	0	0
			2493	1560	459	465	2	7			
1	B	333	Total	C	N	O	S	Se	0	0	0
			2383	1496	435	443	2	7			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP P66992
A	49	MSE	MET	modified residue	UNP P66992
A	53	MSE	MET	modified residue	UNP P66992
A	69	MSE	MET	modified residue	UNP P66992
A	71	MSE	MET	modified residue	UNP P66992
A	86	MSE	MET	modified residue	UNP P66992
A	121	MSE	MET	modified residue	UNP P66992
A	230	MSE	MET	modified residue	UNP P66992
A	371	LEU	-	cloning artifact	UNP P66992
A	372	GLU	-	cloning artifact	UNP P66992
A	373	HIS	-	cloning artifact	UNP P66992
A	374	HIS	-	cloning artifact	UNP P66992
A	375	HIS	-	cloning artifact	UNP P66992
A	376	HIS	-	cloning artifact	UNP P66992
A	377	HIS	-	cloning artifact	UNP P66992
A	378	HIS	-	cloning artifact	UNP P66992
B	1	MSE	MET	modified residue	UNP P66992
B	49	MSE	MET	modified residue	UNP P66992
B	53	MSE	MET	modified residue	UNP P66992
B	69	MSE	MET	modified residue	UNP P66992
B	71	MSE	MET	modified residue	UNP P66992
B	86	MSE	MET	modified residue	UNP P66992
B	121	MSE	MET	modified residue	UNP P66992
B	230	MSE	MET	modified residue	UNP P66992
B	371	LEU	-	cloning artifact	UNP P66992

Continued on next page...

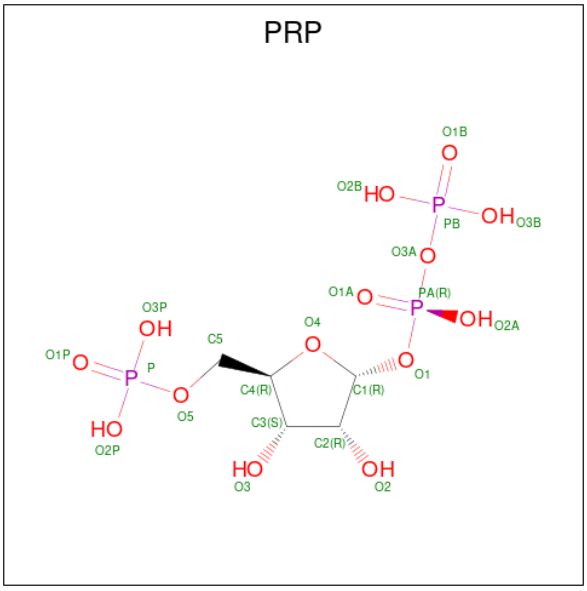
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	372	GLU	-	cloning artifact	UNP P66992
B	373	HIS	-	cloning artifact	UNP P66992
B	374	HIS	-	cloning artifact	UNP P66992
B	375	HIS	-	cloning artifact	UNP P66992
B	376	HIS	-	cloning artifact	UNP P66992
B	377	HIS	-	cloning artifact	UNP P66992
B	378	HIS	-	cloning artifact	UNP P66992

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

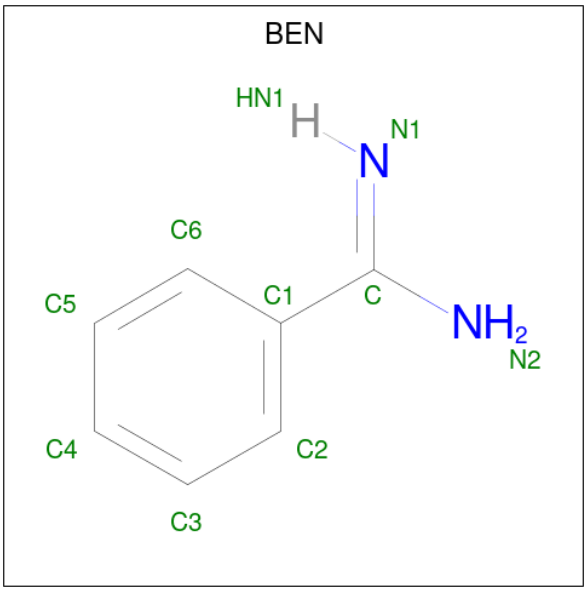
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Mg 2 2	0	0

- Molecule 3 is 1-O-pyrophosphono-5-O-phosphono-alpha-D-ribofuranose (three-letter code: PRP) (formula: $C_5H_{13}O_{14}P_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O P 22 5 14 3	0	0

- Molecule 4 is BENZAMIDINE (three-letter code: BEN) (formula: $C_7H_8N_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N 9 7 2	0	0

- Molecule 5 is water.

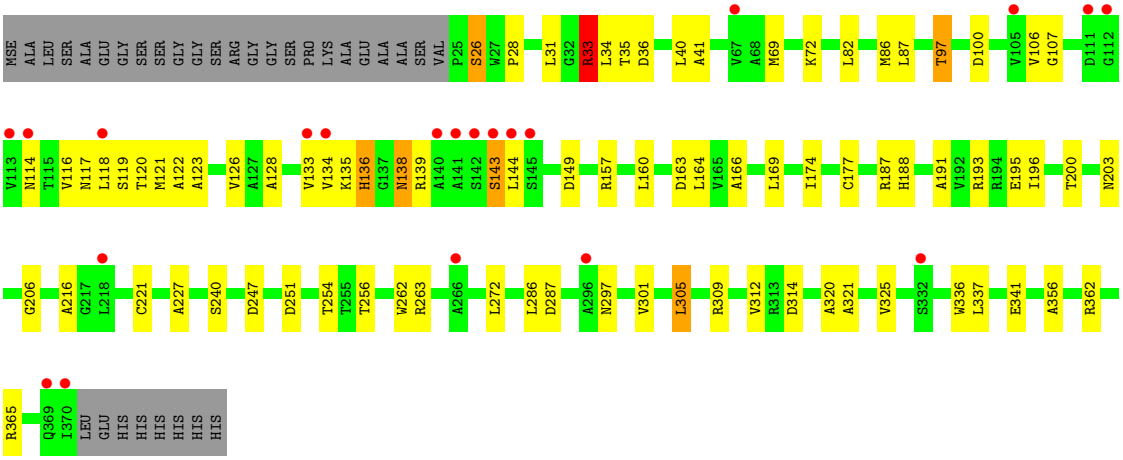
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	142	Total O 142 142	0	0
5	B	135	Total O 135 135	0	0

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.

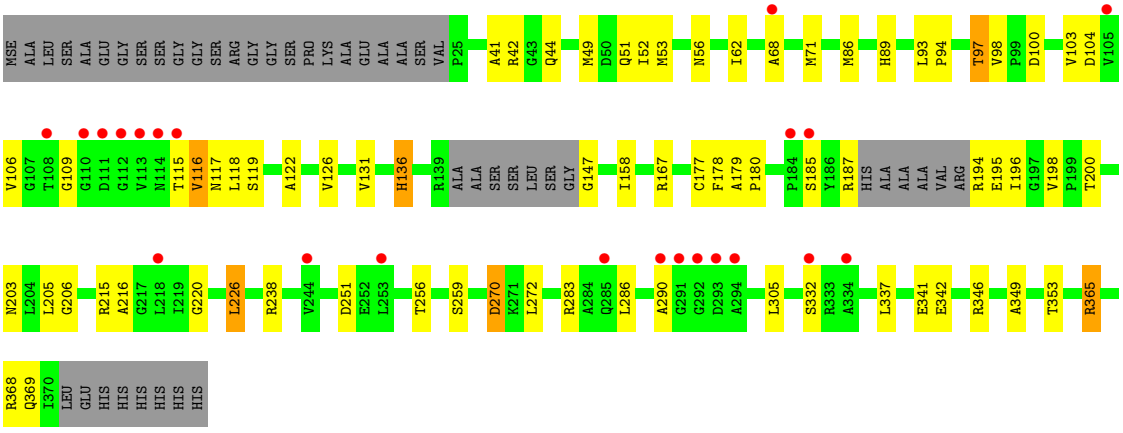
• Molecule 1: Anthranilate phosphoribosyltransferase

Chain A:



• Molecule 1: Anthranilate phosphoribosyltransferase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	80.42Å 82.08Å 118.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.30 24.92 – 2.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (25.00-2.30) 99.9 (24.92-2.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.10 (at 2.31Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.194 , 0.255 0.194 , 0.239	Depositor DCC
R_{free} test set	1784 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	38.6	Xtriage
Anisotropy	0.016	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 46.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.020 for k,h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5186	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 20.62 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 8.3603e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, BEN, PRP

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	1.12	2/2534 (0.1%)	1.05	11/3448 (0.3%)
1	B	1.11	2/2420 (0.1%)	1.03	8/3293 (0.2%)
All	All	1.11	4/4954 (0.1%)	1.04	19/6741 (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	B	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	262	TRP	CB-CG	6.04	1.61	1.50
1	A	320	ALA	CA-CB	5.32	1.63	1.52
1	B	270	ASP	CB-CG	-5.25	1.40	1.51
1	B	259	SER	CA-CB	5.05	1.60	1.52

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	238	ARG	NE-CZ-NH2	-10.36	115.12	120.30
1	B	368	ARG	NE-CZ-NH1	9.80	125.20	120.30
1	B	368	ARG	NE-CZ-NH2	-8.94	115.83	120.30
1	A	362	ARG	NE-CZ-NH2	-8.12	116.24	120.30
1	B	238	ARG	NE-CZ-NH1	7.91	124.26	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	194	ARG	Peptide

5.2 Too-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 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	2493	0	2479	58	0
1	B	2383	0	2356	48	0
2	A	2	0	0	0	0
3	A	22	0	8	6	0
4	A	9	0	7	0	0
5	A	142	0	0	4	0
5	B	135	0	0	4	0
All	All	5186	0	4850	106	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 11.

The worst 5 of 106 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:B:86:MSE:HE1	1:B:205:LEU:HD12	1.18	1.13
1:A:69:MSE:HA	1:A:69:MSE:HE2	1.30	1.13
1:A:157:ARG:HG2	1:A:160:LEU:HD21	1.39	1.03
1:B:51:GLN:HE22	1:B:56:ASN:ND2	1.60	0.98
1:A:116:VAL:HG13	1:A:121:MSE:HE1	1.45	0.97

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	344/378 (91%)	333 (97%)	11 (3%)	0	100	100
1	B	327/378 (86%)	315 (96%)	10 (3%)	2 (1%)	25	31
All	All	671/756 (89%)	648 (97%)	21 (3%)	2 (0%)	41	50

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	185	SER
1	B	332	SER

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	241/257 (94%)	226 (94%)	15 (6%)	18	25
1	B	228/257 (89%)	218 (96%)	10 (4%)	28	39
All	All	469/514 (91%)	444 (95%)	25 (5%)	22	31

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	305	LEU
1	B	116	VAL
1	B	305	LEU
1	B	97	THR
1	B	136	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	56	ASN
1	B	136	HIS
1	B	203	ASN
1	A	136	HIS
1	A	114	ASN

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 4 ligands modelled in this entry, 2 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	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	BEN	A	5001	-	9,9,9	1.45	1 (11%)	7,11,11	1.77	2 (28%)
3	PRP	A	3966	2	19,22,22	0.71	0	33,35,35	1.28	4 (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
4	BEN	A	5001	-	-	4/4/4/4	0/1/1/1
3	PRP	A	3966	2	-	2/16/33/33	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	5001	BEN	C1-C	2.85	1.52	1.47

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	5001	BEN	C1-C-N2	3.45	123.24	118.05
3	A	3966	PRP	C1-C2-C3	2.80	105.84	102.30
3	A	3966	PRP	O3A-PA-O1	-2.63	97.19	102.48
3	A	3966	PRP	PA-O1-C1	2.35	128.81	119.74
3	A	3966	PRP	O3P-P-O2P	2.27	116.33	107.64

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	3966	PRP	PA-O3A-PB-O1B
4	A	5001	BEN	N2-C-C1-C2
4	A	5001	BEN	N2-C-C1-C6
4	A	5001	BEN	N1-C-C1-C2
4	A	5001	BEN	N1-C-C1-C6

There are no ring outliers.

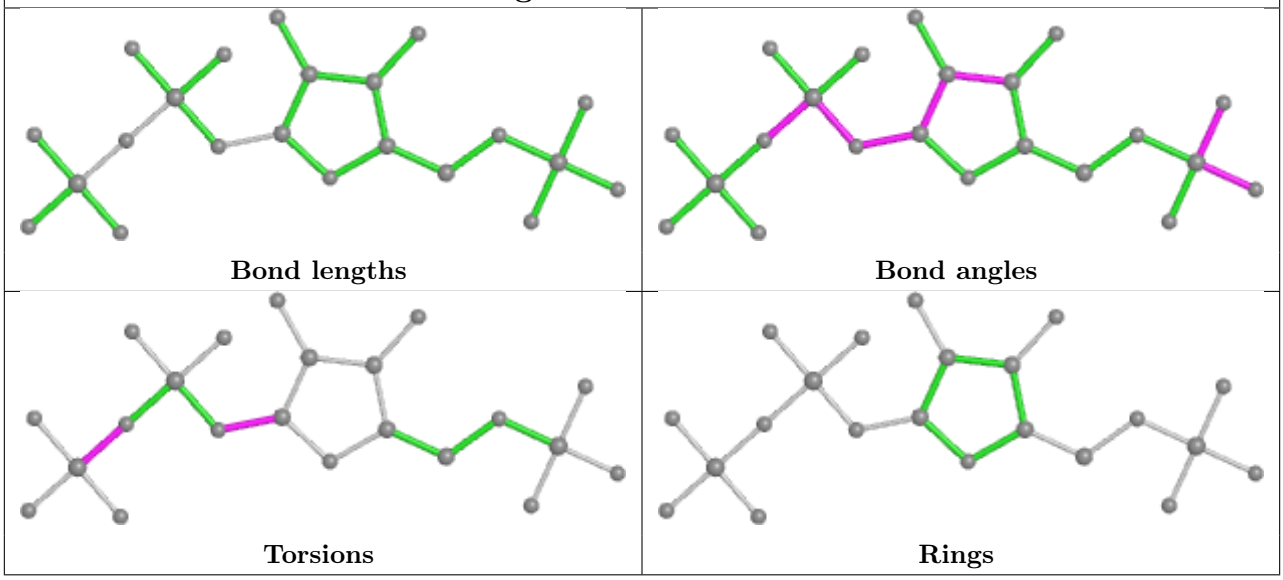
1 monomer is involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	3966	PRP	6	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and

any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

Ligand PRP A 3966



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

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, 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			OWAB(Å^2)	Q<0.9
1	A	339/378 (89%)	0.00	21 (6%)	20	26	22, 36, 63, 78	0
1	B	326/378 (86%)	0.09	22 (6%)	17	23	24, 37, 62, 81	0
All	All	665/756 (87%)	0.05	43 (6%)	18	24	22, 37, 62, 81	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	112	GLY	8.3
1	B	113	VAL	7.7
1	A	143	SER	7.5
1	B	110	GLY	6.3
1	A	144	LEU	5.9

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 monosaccharides in this entry.

6.4 Ligands

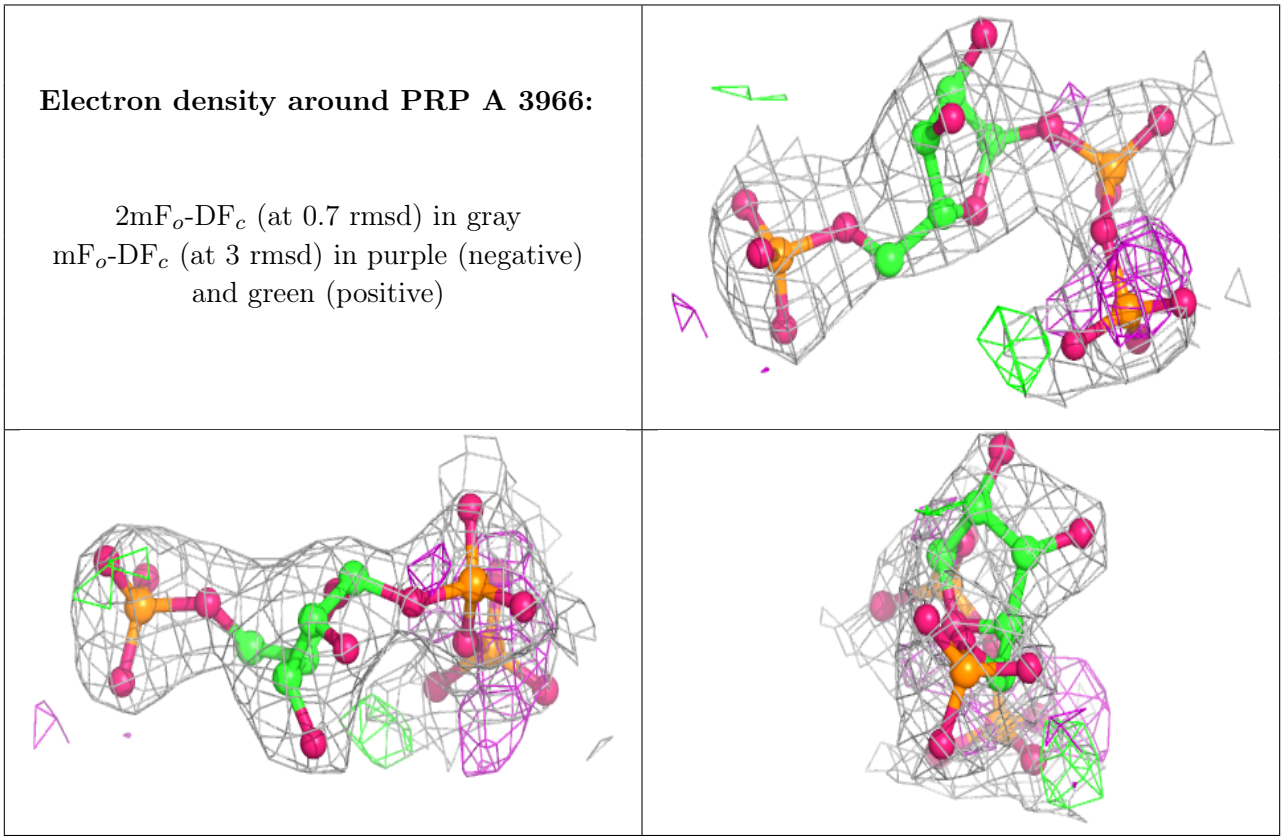
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
2	MG	A	4001	1/1	0.86	0.07	50,50,50,50	0
4	BEN	A	5001	9/9	0.87	0.17	48,51,54,54	0
2	MG	A	6001	1/1	0.90	0.06	55,55,55,55	0
3	PRP	A	3966	22/22	0.91	0.15	66,69,70,70	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around PRP A 3966:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)



6.5 Other polymers (i)

There are no such residues in this entry.