wwPDB X-ray Structure Validation Summary Report (i)

Sep 4, 2023 – 06:43 PM EDT

PDB ID	:	3UN1
Title	:	Crystal structure of an oxidoreductase from Sinorhizobium meliloti 1021
Authors	:	Agarwal, R.; Chamala, S.; Evans, B.; Foti, R.; Gizzi, A.; Hillerich, B.; Kar, A.; LaFleur, J.; Seidel, R.; Villigas, G.; Zencheck, W.; Almo, S.C.; Swaminathan, S.; New York Structural Genomics Research Consortium (NYSGRC)
Deposited on	:	2011-11-15
Resolution	:	2.45 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:

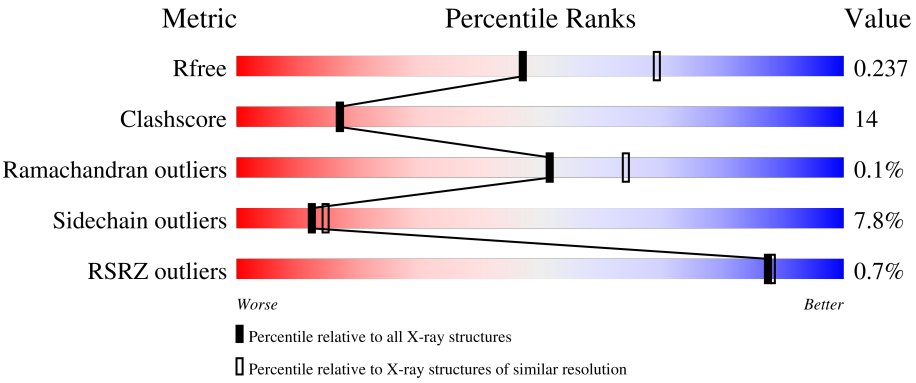
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.45 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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	260	71% 17% 7%
1	B	260	69% 18% 10%
1	C	260	70% 18% 10%
1	D	260	68% 19% 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7202 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 Probable oxidoreductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	241	Total	C	N	O	Se	0	0	0
			1801	1124	332	336	9			
1	B	235	Total	C	N	O	Se	0	0	0
			1743	1088	321	327	7			
1	C	234	Total	C	N	O	Se	0	0	0
			1746	1089	320	330	7			
1	D	236	Total	C	N	O	Se	0	0	0
			1744	1087	323	328	6			

There are 92 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MSE	-	expression tag	UNP Q9EXU6
A	-22	HIS	-	expression tag	UNP Q9EXU6
A	-21	HIS	-	expression tag	UNP Q9EXU6
A	-20	HIS	-	expression tag	UNP Q9EXU6
A	-19	HIS	-	expression tag	UNP Q9EXU6
A	-18	HIS	-	expression tag	UNP Q9EXU6
A	-17	HIS	-	expression tag	UNP Q9EXU6
A	-16	SER	-	expression tag	UNP Q9EXU6
A	-15	SER	-	expression tag	UNP Q9EXU6
A	-14	GLY	-	expression tag	UNP Q9EXU6
A	-13	VAL	-	expression tag	UNP Q9EXU6
A	-12	ASP	-	expression tag	UNP Q9EXU6
A	-11	LEU	-	expression tag	UNP Q9EXU6
A	-10	GLY	-	expression tag	UNP Q9EXU6
A	-9	THR	-	expression tag	UNP Q9EXU6
A	-8	GLU	-	expression tag	UNP Q9EXU6
A	-7	ASN	-	expression tag	UNP Q9EXU6
A	-6	LEU	-	expression tag	UNP Q9EXU6
A	-5	TYR	-	expression tag	UNP Q9EXU6
A	-4	PHE	-	expression tag	UNP Q9EXU6
A	-3	GLN	-	expression tag	UNP Q9EXU6

Continued on next page...

Continued from previous page...

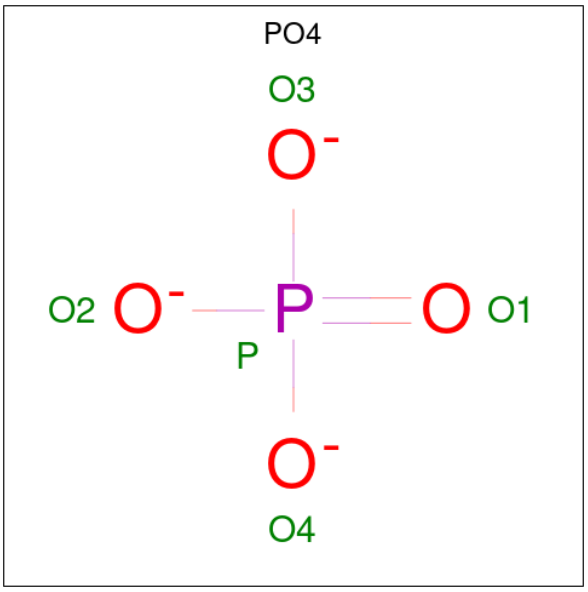
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q9EXU6
A	-1	MSE	-	expression tag	UNP Q9EXU6
B	-23	MSE	-	expression tag	UNP Q9EXU6
B	-22	HIS	-	expression tag	UNP Q9EXU6
B	-21	HIS	-	expression tag	UNP Q9EXU6
B	-20	HIS	-	expression tag	UNP Q9EXU6
B	-19	HIS	-	expression tag	UNP Q9EXU6
B	-18	HIS	-	expression tag	UNP Q9EXU6
B	-17	HIS	-	expression tag	UNP Q9EXU6
B	-16	SER	-	expression tag	UNP Q9EXU6
B	-15	SER	-	expression tag	UNP Q9EXU6
B	-14	GLY	-	expression tag	UNP Q9EXU6
B	-13	VAL	-	expression tag	UNP Q9EXU6
B	-12	ASP	-	expression tag	UNP Q9EXU6
B	-11	LEU	-	expression tag	UNP Q9EXU6
B	-10	GLY	-	expression tag	UNP Q9EXU6
B	-9	THR	-	expression tag	UNP Q9EXU6
B	-8	GLU	-	expression tag	UNP Q9EXU6
B	-7	ASN	-	expression tag	UNP Q9EXU6
B	-6	LEU	-	expression tag	UNP Q9EXU6
B	-5	TYR	-	expression tag	UNP Q9EXU6
B	-4	PHE	-	expression tag	UNP Q9EXU6
B	-3	GLN	-	expression tag	UNP Q9EXU6
B	-2	SER	-	expression tag	UNP Q9EXU6
B	-1	MSE	-	expression tag	UNP Q9EXU6
C	-23	MSE	-	expression tag	UNP Q9EXU6
C	-22	HIS	-	expression tag	UNP Q9EXU6
C	-21	HIS	-	expression tag	UNP Q9EXU6
C	-20	HIS	-	expression tag	UNP Q9EXU6
C	-19	HIS	-	expression tag	UNP Q9EXU6
C	-18	HIS	-	expression tag	UNP Q9EXU6
C	-17	HIS	-	expression tag	UNP Q9EXU6
C	-16	SER	-	expression tag	UNP Q9EXU6
C	-15	SER	-	expression tag	UNP Q9EXU6
C	-14	GLY	-	expression tag	UNP Q9EXU6
C	-13	VAL	-	expression tag	UNP Q9EXU6
C	-12	ASP	-	expression tag	UNP Q9EXU6
C	-11	LEU	-	expression tag	UNP Q9EXU6
C	-10	GLY	-	expression tag	UNP Q9EXU6
C	-9	THR	-	expression tag	UNP Q9EXU6
C	-8	GLU	-	expression tag	UNP Q9EXU6
C	-7	ASN	-	expression tag	UNP Q9EXU6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-6	LEU	-	expression tag	UNP Q9EXU6
C	-5	TYR	-	expression tag	UNP Q9EXU6
C	-4	PHE	-	expression tag	UNP Q9EXU6
C	-3	GLN	-	expression tag	UNP Q9EXU6
C	-2	SER	-	expression tag	UNP Q9EXU6
C	-1	MSE	-	expression tag	UNP Q9EXU6
D	-23	MSE	-	expression tag	UNP Q9EXU6
D	-22	HIS	-	expression tag	UNP Q9EXU6
D	-21	HIS	-	expression tag	UNP Q9EXU6
D	-20	HIS	-	expression tag	UNP Q9EXU6
D	-19	HIS	-	expression tag	UNP Q9EXU6
D	-18	HIS	-	expression tag	UNP Q9EXU6
D	-17	HIS	-	expression tag	UNP Q9EXU6
D	-16	SER	-	expression tag	UNP Q9EXU6
D	-15	SER	-	expression tag	UNP Q9EXU6
D	-14	GLY	-	expression tag	UNP Q9EXU6
D	-13	VAL	-	expression tag	UNP Q9EXU6
D	-12	ASP	-	expression tag	UNP Q9EXU6
D	-11	LEU	-	expression tag	UNP Q9EXU6
D	-10	GLY	-	expression tag	UNP Q9EXU6
D	-9	THR	-	expression tag	UNP Q9EXU6
D	-8	GLU	-	expression tag	UNP Q9EXU6
D	-7	ASN	-	expression tag	UNP Q9EXU6
D	-6	LEU	-	expression tag	UNP Q9EXU6
D	-5	TYR	-	expression tag	UNP Q9EXU6
D	-4	PHE	-	expression tag	UNP Q9EXU6
D	-3	GLN	-	expression tag	UNP Q9EXU6
D	-2	SER	-	expression tag	UNP Q9EXU6
D	-1	MSE	-	expression tag	UNP Q9EXU6

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total O P 5 4 1	0	0
2	D	1	Total O P 5 4 1	0	0

- Molecule 3 is water.

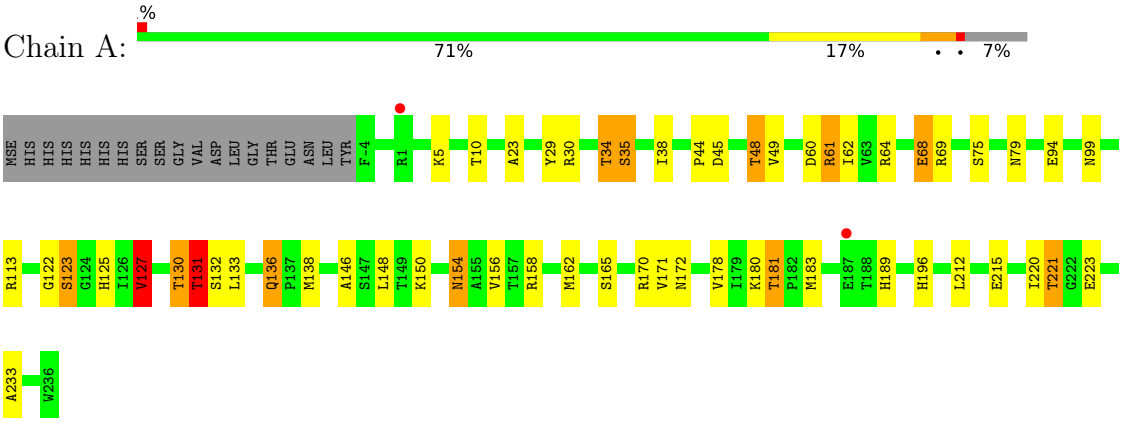
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	41	Total O 41 41	0	0
3	B	48	Total O 48 48	0	0
3	C	37	Total O 37 37	0	0
3	D	32	Total O 32 32	0	0

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.

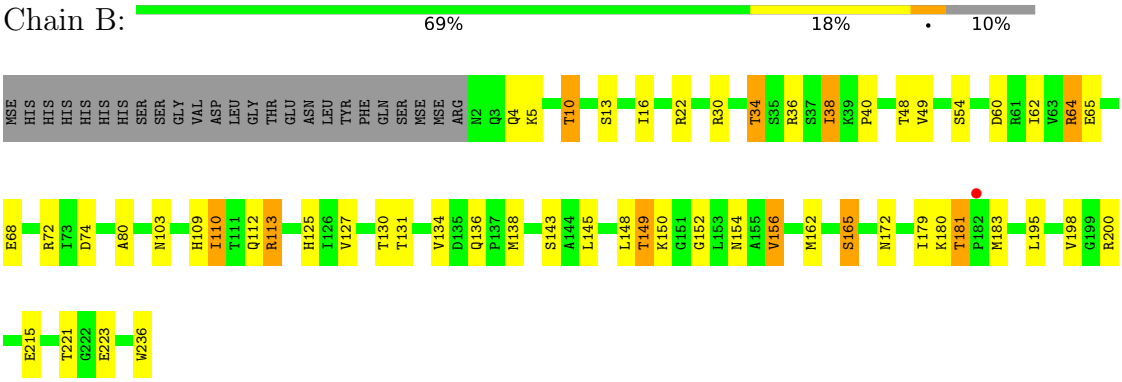
• Molecule 1: Probable oxidoreductase

Chain A:



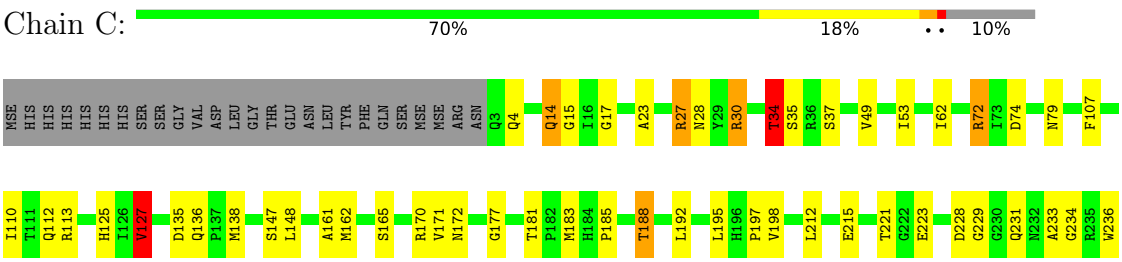
• Molecule 1: Probable oxidoreductase

Chain B:



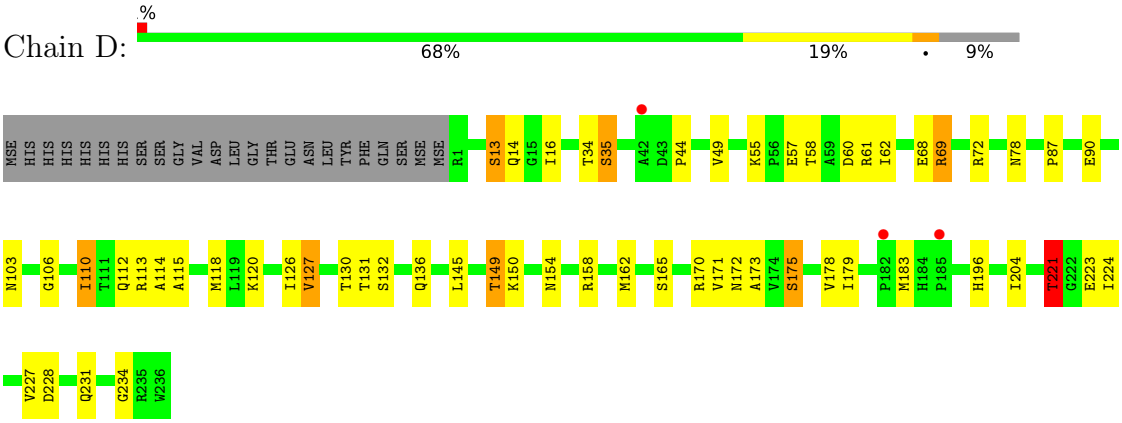
• Molecule 1: Probable oxidoreductase

Chain C:



• Molecule 1: Probable oxidoreductase

Chain D:


%
68%
19%
9%
MSE HIS HIS HIS HIS HIS HIS SER SER GLY VAL ASP LEU GLY THR GLU ASN LEU TYR PHE GLN SER MSE MSE R1 S13 Q14 G15 I16 T34 S35 A42 D43 P44 V49 K55 P56 E57 T58 A59 D60 R61 I62 E68 R69 R72 N78 P87 E90
N103 G106 I110 T111 Q112 R113 A114 A115 M118 L119 K120 I126 V127 T130 T131 S132 Q136 L145 T149 K150 N154 R158 M162 S165 R170 V171 N172 A173 V174 S175 V178 I179 P182 M183 H184 P185 H196 I204 T221 G222 E223 I224
V227 D228 Q231 G234 R235 W236

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	84.22Å 84.22Å 249.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.38 – 2.45 47.38 – 2.45	Depositor EDS
% Data completeness (in resolution range)	99.7 (47.38-2.45) 99.7 (47.38-2.45)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.30 (at 2.45Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.185 , 0.241 0.182 , 0.237	Depositor DCC
R_{free} test set	1939 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	38.1	Xtriage
Anisotropy	0.025	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 37.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.025 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7202	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.30% 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.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.95	3/1824 (0.2%)	0.94	3/2457 (0.1%)
1	B	1.01	1/1767 (0.1%)	0.96	1/2386 (0.0%)
1	C	0.98	1/1770 (0.1%)	0.95	3/2390 (0.1%)
1	D	0.89	0/1768	0.91	2/2388 (0.1%)
All	All	0.96	5/7129 (0.1%)	0.94	9/9621 (0.1%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	68	GLU	CG-CD	5.56	1.60	1.51
1	C	127	VAL	CB-CG2	-5.48	1.41	1.52
1	A	132	SER	CB-OG	5.20	1.49	1.42
1	B	156	VAL	CB-CG2	-5.20	1.42	1.52
1	A	94	GLU	CG-CD	5.11	1.59	1.51

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	61	ARG	NE-CZ-NH2	-5.92	117.34	120.30
1	D	221	THR	CB-CA-C	-5.89	95.70	111.60
1	A	131	THR	CB-CA-C	-5.63	96.39	111.60
1	D	127	VAL	CB-CA-C	-5.53	100.90	111.40
1	B	110	ILE	CB-CA-C	-5.49	100.62	111.60

There are no chirality outliers.

There are no planarity outliers.

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	1801	0	1793	55	0
1	B	1743	0	1734	51	0
1	C	1746	0	1739	46	0
1	D	1744	0	1729	61	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
3	A	41	0	0	3	0
3	B	48	0	0	2	0
3	C	37	0	0	2	0
3	D	32	0	0	3	0
All	All	7202	0	6995	197	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 14.

The worst 5 of 197 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:D:115:ALA:HA	1:D:118:MSE:CE	1.75	1.16
1:B:152:GLY:O	1:B:156:VAL:HG23	1.50	1.07
1:D:115:ALA:HA	1:D:118:MSE:HE2	1.33	1.04
1:C:138:MSE:HE1	1:C:195:LEU:HD22	1.39	1.01
1:B:103:ASN:HD22	1:B:149:THR:HG21	1.29	0.95

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	239/260 (92%)	227 (95%)	12 (5%)	0	100 100
1	B	233/260 (90%)	222 (95%)	11 (5%)	0	100 100
1	C	232/260 (89%)	226 (97%)	6 (3%)	0	100 100
1	D	234/260 (90%)	227 (97%)	6 (3%)	1 (0%)	34 41
All	All	938/1040 (90%)	902 (96%)	35 (4%)	1 (0%)	51 64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	183	MSE

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	188/200 (94%)	175 (93%)	13 (7%)	15 18
1	B	182/200 (91%)	164 (90%)	18 (10%)	8 7
1	C	184/200 (92%)	169 (92%)	15 (8%)	11 13
1	D	181/200 (90%)	170 (94%)	11 (6%)	18 24
All	All	735/800 (92%)	678 (92%)	57 (8%)	12 15

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	165	SER
1	D	175	SER
1	C	35	SER
1	D	149	THR
1	D	69	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	112	GLN
1	D	125	HIS
1	D	196	HIS
1	B	136	GLN
1	B	125	HIS

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)

2 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	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	PO4	C	237	-	4,4,4	1.04	0	6,6,6	0.84	0
2	PO4	D	237	-	4,4,4	0.65	0	6,6,6	1.47	1 (16%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	237	PO4	O3-P-O2	3.05	117.77	107.97

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

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	232/260 (89%)	-0.37	2 (0%) 84 85	20, 31, 55, 67	0
1	B	228/260 (87%)	-0.44	1 (0%) 92 93	18, 31, 52, 73	0
1	C	227/260 (87%)	-0.47	0 100 100	19, 30, 46, 57	0
1	D	229/260 (88%)	-0.28	3 (1%) 77 76	20, 35, 60, 71	0
All	All	916/1040 (88%)	-0.39	6 (0%) 87 88	18, 32, 55, 73	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	42	ALA	5.1
1	B	182	PRO	3.1
1	D	182	PRO	2.9
1	A	187	GLU	2.8
1	D	185	PRO	2.5

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	PO4	C	237	5/5	0.95	0.23	57,59,60,60	0
2	PO4	D	237	5/5	0.95	0.25	46,49,50,55	0

6.5 Other polymers (i)

There are no such residues in this entry.