



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 04:45 AM GMT

PDB ID : 2O1S
Title : 1-deoxy-D-xylulose 5-phosphate synthase (DXS) from Escherichia coli
Authors : Xiang, S.; Usunow, G.; Lange, G.; Busch, M.; Tong, L.
Deposited on : 2006-11-29
Resolution : 2.40 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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

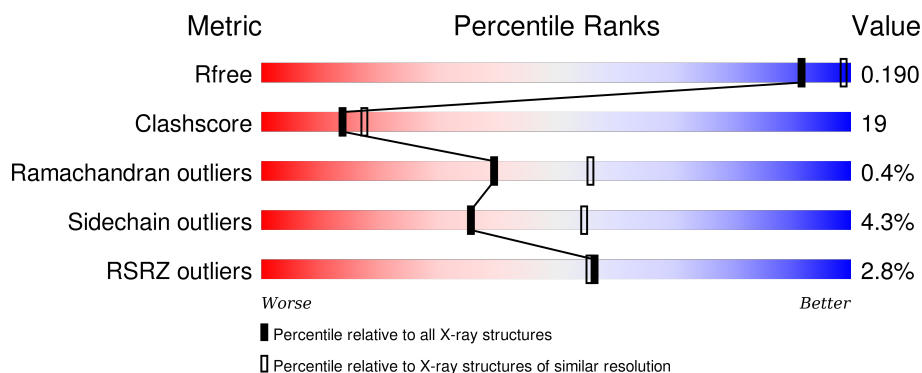
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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}	91344	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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	621	 62% 22% • 14%
1	B	621	 52% 26% • 21%
1	C	621	 62% 22% • 14%
1	D	621	 50% 27% • 22%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 16515 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 1-deoxy-D-xylulose-5-phosphate synthase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	536	Total	C	N	O	S	Se	0	0	0
			4122	2627	708	764	5	18			
1	B	493	Total	C	N	O	S	Se	0	0	0
			3800	2432	643	702	5	18			
1	C	536	Total	C	N	O	S	Se	0	0	0
			4123	2628	708	764	5	18			
1	D	487	Total	C	N	O	S	Se	0	0	0
			3734	2384	631	696	5	18			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	159	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	164	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	174	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	183	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	236	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	269	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	281	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	340	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	346	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	352	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	446	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	449	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	460	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	515	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	532	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	546	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	562	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	574	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	600	MSE	MET	MODIFIED RESIDUE	UNP P77488
A	611	MSE	MET	MODIFIED RESIDUE	UNP P77488

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	621	LEU	-	CLONING ARTIFACT	UNP P77488
B	1	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	159	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	164	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	174	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	183	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	236	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	269	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	281	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	340	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	346	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	352	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	446	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	449	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	460	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	515	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	532	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	546	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	562	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	574	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	600	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	611	MSE	MET	MODIFIED RESIDUE	UNP P77488
B	621	LEU	-	CLONING ARTIFACT	UNP P77488
C	1	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	159	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	164	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	174	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	183	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	236	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	269	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	281	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	340	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	346	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	352	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	446	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	449	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	460	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	515	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	532	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	546	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	562	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	574	MSE	MET	MODIFIED RESIDUE	UNP P77488

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	600	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	611	MSE	MET	MODIFIED RESIDUE	UNP P77488
C	621	LEU	-	CLONING ARTIFACT	UNP P77488
D	1	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	159	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	164	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	174	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	183	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	236	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	269	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	281	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	340	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	346	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	352	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	446	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	449	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	460	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	515	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	532	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	546	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	562	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	574	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	600	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	611	MSE	MET	MODIFIED RESIDUE	UNP P77488
D	621	LEU	-	CLONING ARTIFACT	UNP P77488

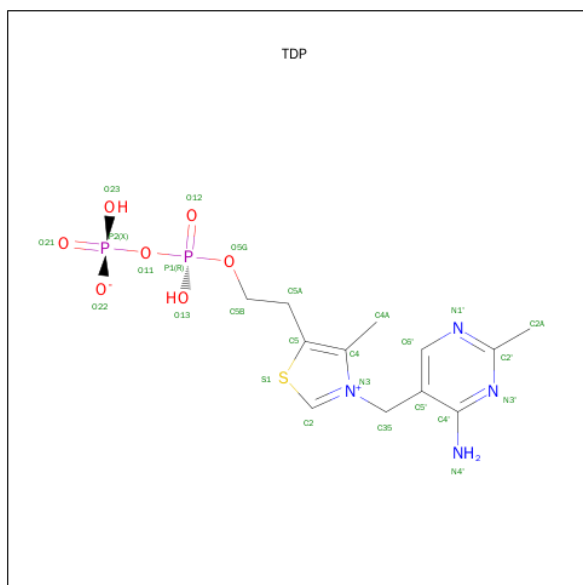
- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

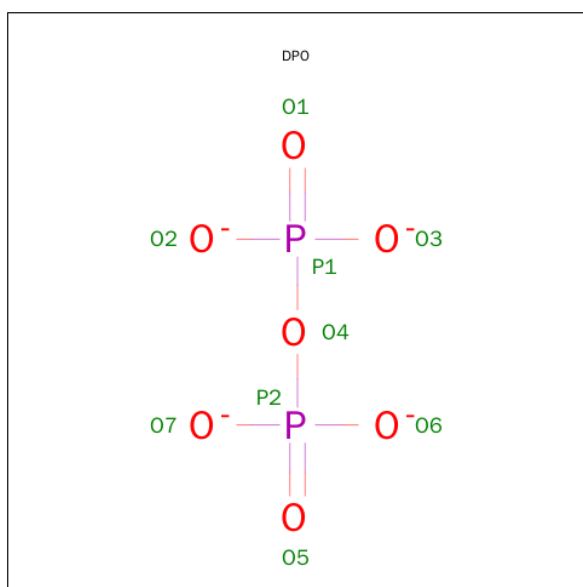
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total K 1 1	0	0
3	A	1	Total K 1 1	0	0
3	D	1	Total K 1 1	0	0
3	C	1	Total K 1 1	0	0

- Molecule 4 is THIAMIN DIPHOSPHATE (three-letter code: TDP) (formula: $C_{12}H_{18}N_4O_7P_2S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	S	0	0
			26	12	4	7	2	1		
4	C	1	Total	C	N	O	P	S	0	0
			26	12	4	7	2	1		

- Molecule 5 is DIPHOSPHATE (three-letter code: DPO) (formula: O_7P_2).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	O	P	0	0
			9	7	2		
5	D	1	Total	O	P	0	0
			9	7	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	210	Total	O	0	0
			210	210		
6	B	134	Total	O	0	0
			134	134		
6	C	186	Total	O	0	0
			186	186		
6	D	128	Total	O	0	0
			128	128		

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ($\text{RSRZ} > 2$). 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.

Chain A: 62% 22% 14%

The sequence logo for Chain A displays the conservation of amino acids across 100 positions. The y-axis represents the log-odds of an amino acid being at a position given the background frequency. The x-axis shows positions 1 to 100. A color scale at the top indicates conservation levels: green (62%), yellow (22%), and grey (14%).

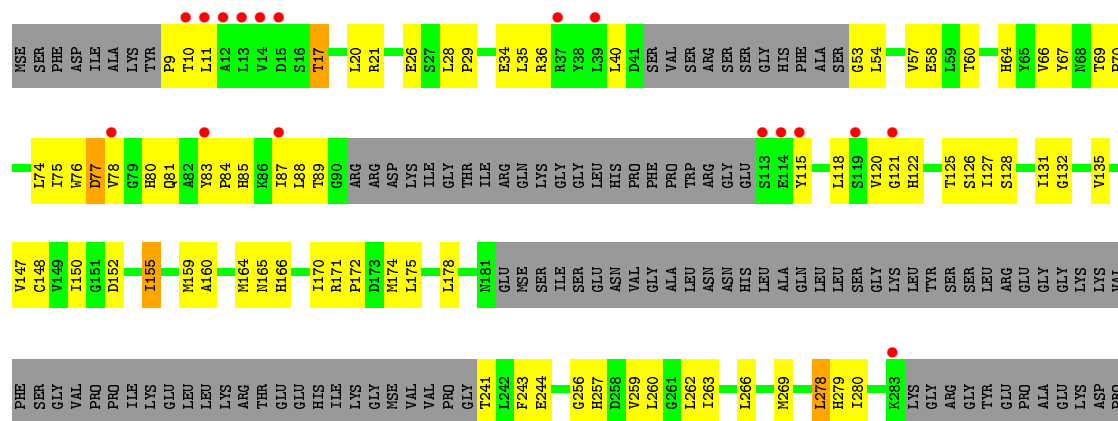
Key features of the logo include:

- Highly Conserved Regions (Green):** Positions 1-10, 15-20, 25-30, 35-40, 45-50, 55-60, 65-70, 75-80, 85-90, and 95-100 show high conservation, with many positions having a log-odds value above 2.0.
- Conserved Regions (Yellow):** Positions 11-14, 21-24, 31-34, 41-44, 51-54, 61-64, 71-74, 81-84, and 91-94 show moderate conservation, with log-odds values generally between 1.0 and 2.0.
- Non-conserved Regions (Grey):** Positions 12-13, 22-23, 32-33, 42-43, 52-53, 62-63, 72-73, 82-83, and 92-93 show low conservation, with log-odds values generally below 1.0.

Chain B:

Sequence logo for Chain B. The y-axis represents information content in bits. The x-axis lists amino acids. A color scale at the top indicates conservation levels: 4% (red), 52% (green), 26% (yellow), and 21% (grey).

Position	Amino Acid	Information Content (bits)
1	LEU	0.15
2	LYS	0.15
3	ARG	0.15
4	THR	0.15
5	GLU	0.15
6	GLY	0.15
7	HIS	0.15
8	ILE	0.15
9	LYS	0.15
10	GLY	0.15
11	MSE	0.15
12	VAL	0.15
13	PRO	0.15
14	GLY	0.15
15	THR	0.15
16	L242	0.15
17	F243	0.15
18	E244	0.15
19	E245	0.15
20	L246	0.15
21	G247	0.15
22	F248	0.15
23	H257	0.15
24	D258	0.15
25	V259	0.15
26	L262	0.15
27	T265	0.15
28	L266	0.15
29	K267	0.15
30	N268	0.15
31	M269	0.15
32	L278	0.15
33	M281	0.15
34	T282	0.15
35	K283	0.15
36	LNS	0.15
37	GLY	0.15
38	ARG	0.15
39	GLY	0.15
40	TYR	0.15
41	GLU	0.15
42	PRO	0.15
43	L286	0.15
44	GLU	0.15
45	LNS	0.15
46	ASP	0.15
47	PRO	0.15
48	IIE	0.15
49	THR	0.15
50	PHE	0.15
51	HIS	0.15
52	ALA	0.15
53	VAL	0.15
54	PRO	0.15
55	L287	0.15
56	GLY	0.15
57	LEU	0.15
58	GLY	0.15
59	LNS	0.15
60	GLY	0.15
61	LNS	0.15
62	TYR	0.15
63	GLY	0.15
64	LNS	0.15
65	GLY	0.15
66	LNS	0.15
67	GLY	0.15
68	LNS	0.15
69	GLY	0.15
70	LNS	0.15
71	GLY	0.15
72	LNS	0.15
73	GLY	0.15
74	LNS	0.15
75	GLY	0.15
76	LNS	0.15
77	GLY	0.15
78	LNS	0.15
79	GLY	0.15
80	LNS	0.15
81	GLY	0.15
82	LNS	0.15
83	GLY	0.15
84	LNS	0.15
85	GLY	0.15
86	LNS	0.15
87	GLY	0.15
88	LNS	0.15
89	GLY	0.15
90	LNS	0.15
91	GLY	0.15
92	LNS	0.15
93	GLY	0.15
94	LNS	0.15
95	GLY	0.15
96	LNS	0.15
97	GLY	0.15
98	LNS	0.15
99	GLY	0.15
100	LNS	0.15





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	86.79 Å 171.15 Å 94.80 Å 90.00° 107.25° 90.00°	Depositor
Resolution (Å)	30.00 – 2.40 29.77 – 2.41	Depositor EDS
% Data completeness (in resolution range)	89.0 (30.00-2.40) 89.7 (29.77-2.41)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.30 (at 2.42 Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.191 , 0.234 0.192 , 0.190	Depositor DCC
R_{free} test set	4603 reflections (5.03%)	DCC
Wilson B-factor (Å ²)	15.5	Xtriage
Anisotropy	0.312	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 42.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 91600 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	16515	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: DPO, K, TDP, 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	0.33	0/4191	0.61	1/5644 (0.0%)
1	B	0.33	0/3861	0.60	0/5203
1	C	0.33	0/4191	0.61	0/5644
1	D	0.31	0/3788	0.58	0/5101
All	All	0.32	0/16031	0.60	1/21592 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	339	LEU	CA-CB-CG	5.12	127.08	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	A	4122	0	4133	132	0
1	B	3800	0	3813	171	0
1	C	4123	0	4136	145	0
1	D	3734	0	3757	175	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	26	0	16	2	0
4	C	26	0	16	3	0
5	B	9	0	0	0	0
5	D	9	0	0	0	0
6	A	210	0	0	7	0
6	B	134	0	0	6	0
6	C	186	0	0	4	0
6	D	128	0	0	4	0
All	All	16515	0	15871	588	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 19.

The worst 5 of 588 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)
4:A:3001:TDP:C2	4:A:3001:TDP:H2	0.97	1.50
4:C:3002:TDP:C2	4:C:3002:TDP:H2	0.97	1.47
1:B:350:SER:HB2	1:B:352:MSE:HE2	1.34	1.10
1:B:503:GLY:H	1:B:528:THR:HG22	1.19	1.08
1:D:423:ILE:HD11	1:D:597:GLN:HA	1.39	1.04

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	530/621 (85%)	501 (94%)	26 (5%)	3 (1%)	30	43
1	B	483/621 (78%)	463 (96%)	19 (4%)	1 (0%)	52	69
1	C	530/621 (85%)	507 (96%)	19 (4%)	4 (1%)	24	35
1	D	477/621 (77%)	458 (96%)	18 (4%)	1 (0%)	52	69
All	All	2020/2484 (81%)	1929 (96%)	82 (4%)	9 (0%)	39	56

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	252	GLY
1	A	48	GLY
1	C	505	LYS
1	D	77	ASP
1	C	48	GLY

5.3.2 Protein sidechains ⓘ

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	433/485 (89%)	413 (95%)	20 (5%)	33	51
1	B	401/485 (83%)	385 (96%)	16 (4%)	38	58
1	C	433/485 (89%)	409 (94%)	24 (6%)	27	42
1	D	395/485 (81%)	383 (97%)	12 (3%)	48	70
All	All	1662/1940 (86%)	1590 (96%)	72 (4%)	35	55

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	441	ARG
1	C	26	GLU
1	D	429	GLN
1	B	461	LEU
1	C	3	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 58 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	432	GLN
1	C	276	GLN
1	D	431	HIS
1	B	480	ASN
1	B	576	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 8 are monoatomic - leaving 4 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	TDP	A	3001	2	21,27,27	2.03	4 (19%)	31,40,40	1.76	8 (25%)
5	DPO	B	4001	2	8,8,8	3.83	4 (50%)	12,13,13	1.53	4 (33%)
4	TDP	C	3002	2	21,27,27	2.02	4 (19%)	31,40,40	1.74	8 (25%)
5	DPO	D	4002	2	8,8,8	3.84	4 (50%)	12,13,13	1.55	4 (33%)

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	TDP	A	3001	2	-	0/16/17/17	0/2/2/2
5	DPO	B	4001	2	-	0/6/6/6	0/0/0/0
4	TDP	C	3002	2	-	0/16/17/17	0/2/2/2
5	DPO	D	4002	2	-	0/6/6/6	0/0/0/0

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	3001	TDP	C4-N3	-6.35	1.34	1.39
4	C	3002	TDP	C4-N3	-6.21	1.34	1.39
4	A	3001	TDP	P1-O12	2.77	1.61	1.51
4	C	3002	TDP	P1-O12	2.78	1.61	1.51
4	A	3001	TDP	P2-O22	3.26	1.61	1.50

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	3002	TDP	C4A-C4-C5	-3.21	121.68	128.90
4	A	3001	TDP	C4A-C4-C5	-3.19	121.74	128.90
4	C	3002	TDP	N1'-C2'-N3'	-3.09	119.89	125.60
5	D	4002	DPO	O2-P1-O1	-3.05	105.69	112.76
5	B	4001	DPO	O2-P1-O1	-3.02	105.74	112.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	3001	TDP	2	0
4	C	3002	TDP	3	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	518/621 (83%)	-0.46	2 (0%) 93 93	3, 14, 31, 43	0
1	B	475/621 (76%)	-0.20	23 (4%) 34 35	3, 17, 53, 62	0
1	C	518/621 (83%)	-0.43	7 (1%) 78 77	4, 13, 30, 42	0
1	D	469/621 (75%)	-0.14	23 (4%) 33 34	5, 19, 45, 58	0
All	All	1980/2484 (79%)	-0.31	55 (2%) 56 55	3, 15, 40, 62	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	15	ASP	4.7
1	B	8	TYR	4.1
1	B	111	GLY	3.9
1	B	109	TRP	3.8
1	B	9	PRO	3.8

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(\AA^2)	Q<0.9
5	DPO	D	4002	9/9	0.91	0.19	0.07	50,51,54,54	0
5	DPO	B	4001	9/9	0.92	0.17	-0.24	50,51,55,56	0
4	TDP	A	3001	26/26	0.98	0.10	-0.37	5,13,16,16	0
4	TDP	C	3002	26/26	0.98	0.10	-0.52	4,13,16,18	0
3	K	D	2004	1/1	0.95	0.09	-0.63	34,34,34,34	0
3	K	C	2003	1/1	0.99	0.09	-0.91	24,24,24,24	0
2	MG	A	1001	1/1	0.98	0.07	-1.09	9,9,9,9	0
2	MG	D	1004	1/1	0.74	0.10	-1.15	61,61,61,61	0
3	K	B	2002	1/1	0.96	0.07	-1.37	14,14,14,14	0
3	K	A	2001	1/1	0.99	0.08	-1.44	23,23,23,23	0
2	MG	C	1003	1/1	0.75	0.15	-	26,26,26,26	0
2	MG	B	1002	1/1	0.87	0.12	-	48,48,48,48	0

6.5 Other polymers [i](#)

There are no such residues in this entry.