



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:30 AM GMT

PDB ID : 2X7J  
Title : STRUCTURE OF THE MENAQUINONE BIOSYNTHESIS PROTEIN  
MEND FROM BACILLUS SUBTILIS  
Authors : Dawson, A.; Chen, M.; Fyfe, P.K.; Guo, Z.; Hunter, W.N.  
Deposited on : 2010-03-01  
Resolution : 2.35 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](mailto: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

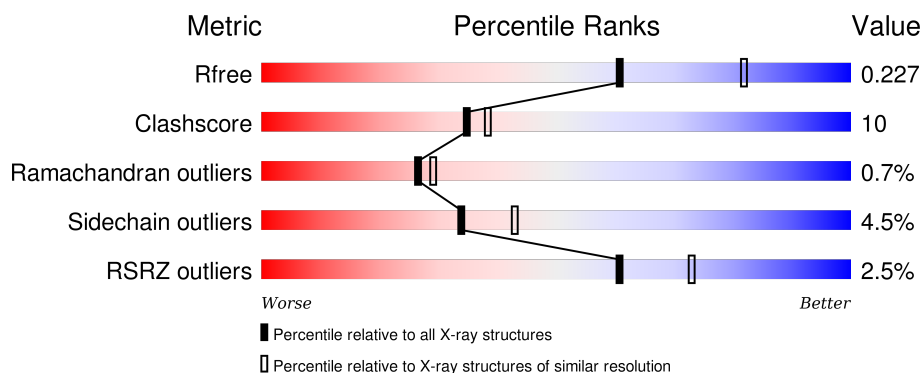
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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	1352 (2.38-2.34)
Clashscore	102246	1456 (2.38-2.34)
Ramachandran outliers	100387	1435 (2.38-2.34)
Sidechain outliers	100360	1436 (2.38-2.34)
RSRZ outliers	91569	1358 (2.38-2.34)

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  $\geq 3$ , 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	604	<div> <div>0%</div> <div> <div></div> <div>79%</div> <div>15%</div> <div>• •</div> </div> </div>
1	B	604	<div> <div>4%</div> <div> <div></div> <div>76%</div> <div>17%</div> <div>• • •</div> </div> </div>
1	C	604	<div> <div>3%</div> <div> <div></div> <div>77%</div> <div>17%</div> <div>• •</div> </div> </div>
1	D	604	<div> <div>2%</div> <div> <div></div> <div>75%</div> <div>19%</div> <div>• •</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	EDO	A	603	-	-	X	X
4	EDO	B	603	-	-	X	-
4	EDO	C	603	-	-	X	X
4	EDO	D	603	-	-	X	X
6	SO4	A	1582	-	-	X	X
6	SO4	B	1582	-	-	X	X
6	SO4	C	1580	-	-	X	X
6	SO4	D	1582	-	-	X	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 19356 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 2-SUCCINYL-5-ENOLPYRUVYL-6-HYDROXY-3-CYCLOHEXENE -1-CARBOXYLATE SYNTHASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	579	Total	C	N	O	S	0	6	0
			4544	2885	789	851	19			
1	B	579	Total	C	N	O	S	0	2	0
			4520	2868	784	849	19			
1	C	577	Total	C	N	O	S	0	2	0
			4502	2857	782	844	19			
1	D	579	Total	C	N	O	S	0	6	0
			4548	2887	793	849	19			

There are 96 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-23	MET	-	EXPRESSION TAG	UNP P23970
A	-22	GLY	-	EXPRESSION TAG	UNP P23970
A	-21	SER	-	EXPRESSION TAG	UNP P23970
A	-20	SER	-	EXPRESSION TAG	UNP P23970
A	-19	HIS	-	EXPRESSION TAG	UNP P23970
A	-18	HIS	-	EXPRESSION TAG	UNP P23970
A	-17	HIS	-	EXPRESSION TAG	UNP P23970
A	-16	HIS	-	EXPRESSION TAG	UNP P23970
A	-15	HIS	-	EXPRESSION TAG	UNP P23970
A	-14	HIS	-	EXPRESSION TAG	UNP P23970
A	-13	SER	-	EXPRESSION TAG	UNP P23970
A	-12	SER	-	EXPRESSION TAG	UNP P23970
A	-11	GLY	-	EXPRESSION TAG	UNP P23970
A	-10	GLU	-	EXPRESSION TAG	UNP P23970
A	-9	ASN	-	EXPRESSION TAG	UNP P23970
A	-8	LEU	-	EXPRESSION TAG	UNP P23970
A	-7	TYR	-	EXPRESSION TAG	UNP P23970
A	-6	PHE	-	EXPRESSION TAG	UNP P23970
A	-5	GLN	-	EXPRESSION TAG	UNP P23970
A	-4	GLY	-	EXPRESSION TAG	UNP P23970

*Continued on next page...*

*Continued from previous page...*

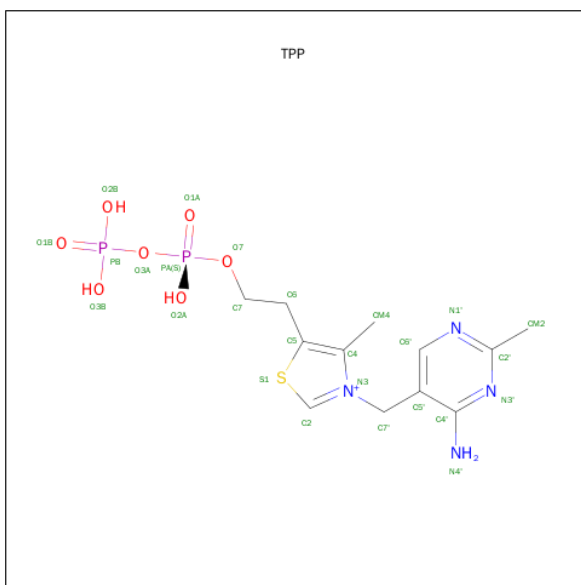
Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	HIS	-	EXPRESSION TAG	UNP P23970
A	-2	MET	-	EXPRESSION TAG	UNP P23970
A	-1	LEU	-	EXPRESSION TAG	UNP P23970
A	0	GLU	-	EXPRESSION TAG	UNP P23970
B	-23	MET	-	EXPRESSION TAG	UNP P23970
B	-22	GLY	-	EXPRESSION TAG	UNP P23970
B	-21	SER	-	EXPRESSION TAG	UNP P23970
B	-20	SER	-	EXPRESSION TAG	UNP P23970
B	-19	HIS	-	EXPRESSION TAG	UNP P23970
B	-18	HIS	-	EXPRESSION TAG	UNP P23970
B	-17	HIS	-	EXPRESSION TAG	UNP P23970
B	-16	HIS	-	EXPRESSION TAG	UNP P23970
B	-15	HIS	-	EXPRESSION TAG	UNP P23970
B	-14	HIS	-	EXPRESSION TAG	UNP P23970
B	-13	SER	-	EXPRESSION TAG	UNP P23970
B	-12	SER	-	EXPRESSION TAG	UNP P23970
B	-11	GLY	-	EXPRESSION TAG	UNP P23970
B	-10	GLU	-	EXPRESSION TAG	UNP P23970
B	-9	ASN	-	EXPRESSION TAG	UNP P23970
B	-8	LEU	-	EXPRESSION TAG	UNP P23970
B	-7	TYR	-	EXPRESSION TAG	UNP P23970
B	-6	PHE	-	EXPRESSION TAG	UNP P23970
B	-5	GLN	-	EXPRESSION TAG	UNP P23970
B	-4	GLY	-	EXPRESSION TAG	UNP P23970
B	-3	HIS	-	EXPRESSION TAG	UNP P23970
B	-2	MET	-	EXPRESSION TAG	UNP P23970
B	-1	LEU	-	EXPRESSION TAG	UNP P23970
B	0	GLU	-	EXPRESSION TAG	UNP P23970
C	-23	MET	-	EXPRESSION TAG	UNP P23970
C	-22	GLY	-	EXPRESSION TAG	UNP P23970
C	-21	SER	-	EXPRESSION TAG	UNP P23970
C	-20	SER	-	EXPRESSION TAG	UNP P23970
C	-19	HIS	-	EXPRESSION TAG	UNP P23970
C	-18	HIS	-	EXPRESSION TAG	UNP P23970
C	-17	HIS	-	EXPRESSION TAG	UNP P23970
C	-16	HIS	-	EXPRESSION TAG	UNP P23970
C	-15	HIS	-	EXPRESSION TAG	UNP P23970
C	-14	HIS	-	EXPRESSION TAG	UNP P23970
C	-13	SER	-	EXPRESSION TAG	UNP P23970
C	-12	SER	-	EXPRESSION TAG	UNP P23970
C	-11	GLY	-	EXPRESSION TAG	UNP P23970
C	-10	GLU	-	EXPRESSION TAG	UNP P23970

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	-9	ASN	-	EXPRESSION TAG	UNP P23970
C	-8	LEU	-	EXPRESSION TAG	UNP P23970
C	-7	TYR	-	EXPRESSION TAG	UNP P23970
C	-6	PHE	-	EXPRESSION TAG	UNP P23970
C	-5	GLN	-	EXPRESSION TAG	UNP P23970
C	-4	GLY	-	EXPRESSION TAG	UNP P23970
C	-3	HIS	-	EXPRESSION TAG	UNP P23970
C	-2	MET	-	EXPRESSION TAG	UNP P23970
C	-1	LEU	-	EXPRESSION TAG	UNP P23970
C	0	GLU	-	EXPRESSION TAG	UNP P23970
D	-23	MET	-	EXPRESSION TAG	UNP P23970
D	-22	GLY	-	EXPRESSION TAG	UNP P23970
D	-21	SER	-	EXPRESSION TAG	UNP P23970
D	-20	SER	-	EXPRESSION TAG	UNP P23970
D	-19	HIS	-	EXPRESSION TAG	UNP P23970
D	-18	HIS	-	EXPRESSION TAG	UNP P23970
D	-17	HIS	-	EXPRESSION TAG	UNP P23970
D	-16	HIS	-	EXPRESSION TAG	UNP P23970
D	-15	HIS	-	EXPRESSION TAG	UNP P23970
D	-14	HIS	-	EXPRESSION TAG	UNP P23970
D	-13	SER	-	EXPRESSION TAG	UNP P23970
D	-12	SER	-	EXPRESSION TAG	UNP P23970
D	-11	GLY	-	EXPRESSION TAG	UNP P23970
D	-10	GLU	-	EXPRESSION TAG	UNP P23970
D	-9	ASN	-	EXPRESSION TAG	UNP P23970
D	-8	LEU	-	EXPRESSION TAG	UNP P23970
D	-7	TYR	-	EXPRESSION TAG	UNP P23970
D	-6	PHE	-	EXPRESSION TAG	UNP P23970
D	-5	GLN	-	EXPRESSION TAG	UNP P23970
D	-4	GLY	-	EXPRESSION TAG	UNP P23970
D	-3	HIS	-	EXPRESSION TAG	UNP P23970
D	-2	MET	-	EXPRESSION TAG	UNP P23970
D	-1	LEU	-	EXPRESSION TAG	UNP P23970
D	0	GLU	-	EXPRESSION TAG	UNP P23970

- Molecule 2 is THIAMINE DIPHOSPHATE (three-letter code: TPP) (formula: C<sub>12</sub>H<sub>19</sub>N<sub>4</sub>O<sub>7</sub>P<sub>2</sub>S).



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

- Molecule 3 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

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

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $\text{C}_2\text{H}_6\text{O}_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	C	1	Total	C	O	0	0
			4	2	2		
4	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Na	0	0
			1	1		
5	A	1	Total	Na	0	0
			1	1		
5	D	1	Total	Na	0	0
			1	1		
5	C	1	Total	Na	0	0
			1	1		

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).





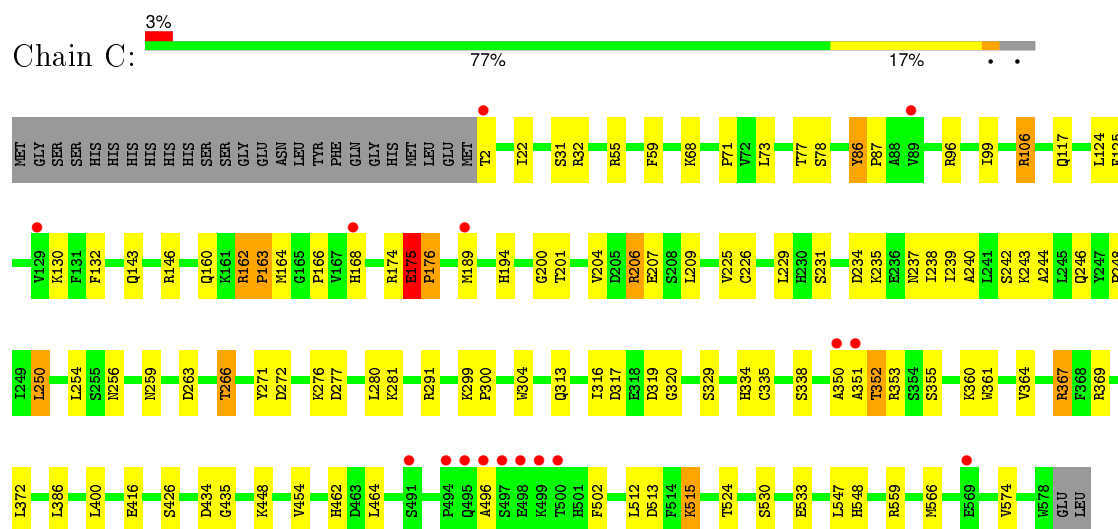
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	S	0	0
			5	4	1		
6	A	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		
6	B	1	Total	O	S	0	0
			5	4	1		
6	C	1	Total	O	S	0	0
			5	4	1		
6	C	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 7 is water.

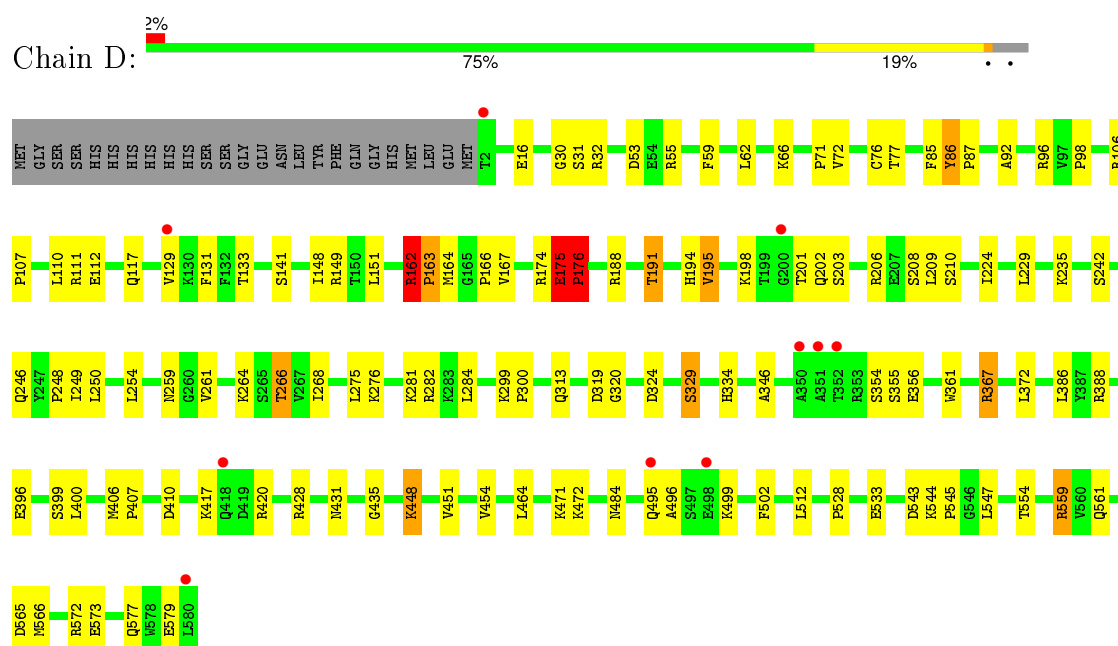
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	285	Total	O	0	0
			285	285		
7	B	239	Total	O	0	0
			239	239		
7	C	256	Total	O	0	0
			256	256		
7	D	294	Total	O	0	0
			294	294		



• Molecule 1: 2-SUCCINYL-5-ENOLPYRUVYL-6-HYDROXY-3-CYCLOHEXENE -1-CARBOXYLATE SYNTHASE



• Molecule 1: 2-SUCCINYL-5-ENOLPYRUVYL-6-HYDROXY-3-CYCLOHEXENE -1-CARBOXYLATE SYNTHASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	100.18Å 152.99Å 158.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	62.14 – 2.35 62.15 – 2.35	Depositor EDS
% Data completeness (in resolution range)	100.0 (62.14-2.35) 100.0 (62.15-2.35)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.39 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.172 , 0.228 0.173 , 0.227	Depositor DCC
$R_{free}$ test set	5118 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.3	Xtriage
Anisotropy	0.082	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 43.0	EDS
Estimated twinning fraction	0.010 for -h,l,k	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	4 of 101872 reflections (0.004%)	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	19356	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	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 26.04 % 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 2.8341e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, TPP, MN, EDO, SO4

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	0.48	0/4671	0.64	3/6342 (0.0%)
1	B	0.45	0/4635	0.61	3/6294 (0.0%)
1	C	0.45	0/4617	0.60	2/6271 (0.0%)
1	D	0.48	0/4675	0.64	3/6346 (0.0%)
All	All	0.46	0/18598	0.62	11/25253 (0.0%)

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	A	0	2
1	B	0	2
1	C	0	2
1	D	0	2
All	All	0	8

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	175	GLU	C-N-CD	-11.48	95.34	120.60
1	C	162	ARG	C-N-CD	-11.36	95.60	120.60
1	A	175	GLU	C-N-CD	-11.26	95.83	120.60
1	A	162	ARG	C-N-CD	-10.23	98.08	120.60
1	B	162	ARG	C-N-CD	-10.04	98.52	120.60

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	162	ARG	Peptide
1	A	175	GLU	Peptide
1	B	162	ARG	Peptide
1	B	175	GLU	Peptide
1	C	162	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	4544	0	4503	74	0
1	B	4520	0	4465	91	0
1	C	4502	0	4448	82	0
1	D	4548	0	4513	111	0
2	A	26	0	16	2	0
2	B	26	0	16	0	0
2	C	26	0	16	1	0
2	D	26	0	16	2	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	4	0	6	13	0
4	B	4	0	6	6	0
4	C	4	0	6	8	0
4	D	4	0	6	8	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
6	A	10	0	0	5	0
6	B	10	0	0	6	0
6	C	10	0	0	6	0
6	D	10	0	0	9	0
7	A	285	0	0	11	0
7	B	239	0	0	9	0
7	C	256	0	0	10	0
7	D	294	0	0	17	0
All	All	19356	0	18017	359	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 10.

The worst 5 of 359 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:C:31:SER:H	4:C:603:EDO:H11	1.04	1.20
1:D:31:SER:H	4:D:603:EDO:H22	1.00	1.17
1:D:163:PRO:HD3	1:D:319:ASP:HB3	1.34	1.08
1:C:206:ARG:HH11	1:C:206:ARG:HG3	0.89	1.06
1:B:163:PRO:CD	1:B:319:ASP:HB3	1.85	1.06

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	583/604 (96%)	555 (95%)	24 (4%)	4 (1%)	26	29
1	B	579/604 (96%)	555 (96%)	17 (3%)	7 (1%)	16	15
1	C	577/604 (96%)	552 (96%)	22 (4%)	3 (0%)	34	39
1	D	583/604 (96%)	564 (97%)	17 (3%)	2 (0%)	46	55
All	All	2322/2416 (96%)	2226 (96%)	80 (3%)	16 (1%)	26	29

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	163	PRO
1	A	176	PRO
1	A	498	GLU
1	B	163	PRO
1	B	176	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	495/511 (97%)	467 (94%)	28 (6%)	25	30
1	B	491/511 (96%)	466 (95%)	25 (5%)	29	36
1	C	489/511 (96%)	468 (96%)	21 (4%)	35	45
1	D	495/511 (97%)	473 (96%)	22 (4%)	35	44
All	All	1970/2044 (96%)	1874 (95%)	96 (5%)	34	39

5 of 96 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	367	ARG
1	C	86[A]	TYR
1	D	448	LYS
1	B	386	LEU
1	B	498	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	431	ASN
1	C	117	GLN
1	D	127	ASN
1	B	334	HIS
1	D	117	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 24 ligands modelled in this entry, 8 are monoatomic - leaving 16 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
6	SO4	A	1582	-	4,4,4	0.51	0	6,6,6	0.28	0
6	SO4	A	1583	-	4,4,4	0.34	0	6,6,6	0.12	0
2	TPP	A	601	3	20,27,27	1.47	2 (10%)	31,40,40	2.09	10 (32%)
4	EDO	A	603	-	3,3,3	0.46	0	2,2,2	0.44	0
6	SO4	B	1582	-	4,4,4	0.56	0	6,6,6	0.37	0
6	SO4	B	1583	-	4,4,4	0.14	0	6,6,6	0.14	0
2	TPP	B	601	3	20,27,27	1.49	3 (15%)	31,40,40	1.85	8 (25%)
4	EDO	B	603	-	3,3,3	0.54	0	2,2,2	0.54	0
6	SO4	C	1580	-	4,4,4	0.70	0	6,6,6	0.44	0
6	SO4	C	1581	-	4,4,4	0.26	0	6,6,6	0.18	0
2	TPP	C	601	3	20,27,27	1.32	4 (20%)	31,40,40	1.93	9 (29%)
4	EDO	C	603	-	3,3,3	0.54	0	2,2,2	0.42	0
6	SO4	D	1582	-	4,4,4	0.57	0	6,6,6	0.13	0
6	SO4	D	1583	-	4,4,4	0.37	0	6,6,6	0.27	0
2	TPP	D	601	3	20,27,27	1.56	4 (20%)	31,40,40	1.88	11 (35%)
4	EDO	D	603	-	3,3,3	0.56	0	2,2,2	0.36	0

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
6	SO4	A	1582	-	-	0/0/0/0	0/0/0/0
6	SO4	A	1583	-	-	0/0/0/0	0/0/0/0
2	TPP	A	601	3	-	0/16/17/17	0/2/2/2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	EDO	A	603	-	-	0/1/1/1	0/0/0/0
6	SO4	B	1582	-	-	0/0/0/0	0/0/0/0
6	SO4	B	1583	-	-	0/0/0/0	0/0/0/0
2	TPP	B	601	3	-	0/16/17/17	0/2/2/2
4	EDO	B	603	-	-	0/1/1/1	0/0/0/0
6	SO4	C	1580	-	-	0/0/0/0	0/0/0/0
6	SO4	C	1581	-	-	0/0/0/0	0/0/0/0
2	TPP	C	601	3	-	0/16/17/17	0/2/2/2
4	EDO	C	603	-	-	0/1/1/1	0/0/0/0
6	SO4	D	1582	-	-	0/0/0/0	0/0/0/0
6	SO4	D	1583	-	-	0/0/0/0	0/0/0/0
2	TPP	D	601	3	-	0/16/17/17	0/2/2/2
4	EDO	D	603	-	-	0/1/1/1	0/0/0/0

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	601	TPP	C4-N3	-4.21	1.36	1.39
2	B	601	TPP	C4-N3	-3.25	1.36	1.39
2	D	601	TPP	C4-N3	-2.92	1.37	1.39
2	C	601	TPP	C4-N3	-2.52	1.37	1.39
2	D	601	TPP	PB-O3B	-2.03	1.47	1.54

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	601	TPP	CM4-C4-C5	-4.11	119.66	128.90
2	D	601	TPP	CM4-C4-C5	-3.66	120.67	128.90
2	A	601	TPP	PA-O3A-PB	-3.48	121.01	132.67
2	C	601	TPP	CM4-C4-C5	-3.44	121.17	128.90
2	A	601	TPP	CM4-C4-C5	-3.40	121.27	128.90

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

12 monomers are involved in 64 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	1582	SO4	5	0
2	A	601	TPP	2	0
4	A	603	EDO	13	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	1582	SO4	6	0
4	B	603	EDO	6	0
6	C	1580	SO4	5	0
6	C	1581	SO4	1	0
2	C	601	TPP	1	0
4	C	603	EDO	8	0
6	D	1582	SO4	9	0
2	D	601	TPP	2	0
4	D	603	EDO	8	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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	579/604 (95%)	-0.06	6 (1%) 84 92	15, 29, 46, 61	1 (0%)
1	B	579/604 (95%)	0.22	25 (4%) 39 53	16, 34, 57, 67	3 (0%)
1	C	577/604 (95%)	0.13	16 (2%) 56 69	15, 33, 55, 68	6 (1%)
1	D	579/604 (95%)	-0.16	10 (1%) 73 83	17, 28, 44, 60	2 (0%)
All	All	2314/2416 (95%)	0.03	57 (2%) 61 73	15, 30, 52, 68	12 (0%)

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	2	THR	6.7
1	B	352	THR	5.1
1	C	351	ALA	4.8
1	D	2	THR	4.8
1	B	2	THR	4.6

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
6	SO4	A	1582	5/5	0.96	0.18	9.10	10,10,11,11	5
6	SO4	D	1582	5/5	0.98	0.28	8.07	16,16,17,17	5
4	EDO	A	603	4/4	0.94	0.20	6.88	25,27,28,29	0
4	EDO	D	603	4/4	0.90	0.20	4.78	26,27,29,29	0
6	SO4	B	1582	5/5	0.94	0.22	4.39	13,14,15,15	5
6	SO4	C	1580	5/5	0.98	0.20	4.19	10,12,12,12	5
4	EDO	C	603	4/4	0.93	0.18	3.03	27,27,27,28	0
5	NA	D	1581	1/1	0.96	0.14	1.81	33,33,33,33	0
4	EDO	B	603	4/4	0.94	0.15	1.58	23,24,25,25	0
2	TPP	C	601	26/26	0.98	0.10	-0.45	22,23,24,24	0
5	NA	A	1581	1/1	0.91	0.10	-0.45	35,35,35,35	0
3	MN	D	602	1/1	1.00	0.09	-0.59	23,23,23,23	0
2	TPP	A	601	26/26	0.98	0.10	-0.64	20,23,25,25	0
2	TPP	B	601	26/26	0.98	0.10	-0.70	21,24,26,27	0
2	TPP	D	601	26/26	0.99	0.09	-0.76	19,21,23,24	0
5	NA	C	1579	1/1	0.96	0.10	-0.99	36,36,36,36	0
5	NA	B	1581	1/1	0.96	0.10	-1.49	33,33,33,33	0
3	MN	A	602	1/1	0.99	0.07	-2.13	24,24,24,24	0
3	MN	B	602	1/1	0.99	0.08	-2.24	29,29,29,29	0
3	MN	C	602	1/1	1.00	0.04	-3.89	24,24,24,24	0
6	SO4	A	1583	5/5	0.97	0.14	-	24,24,24,24	5
6	SO4	D	1583	5/5	0.97	0.24	-	21,22,22,22	5
6	SO4	C	1581	5/5	0.94	0.15	-	26,27,27,27	5
6	SO4	B	1583	5/5	0.93	0.13	-	31,31,31,31	5

## 6.5 Other polymers ⓘ

There are no such residues in this entry.