



# Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 09:35 am BST

PDB ID : 4NG3  
Title : Crystal structure of 5-carboxyvanillate decarboxylase from *Sphingomonas paucimobilis* complexed with 4-Hydroxy-3-methoxy-5-nitrobenzoic acid  
Authors : Fedorov, A.A.; Fedorov, E.V.; Vladimirova, A.; Raushel, F.M.; Almo, S.C.  
Deposited on : 2013-11-01  
Resolution : 1.75 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

<https://www.wwpdb.org/validation/2017/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	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
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)	:	2.11

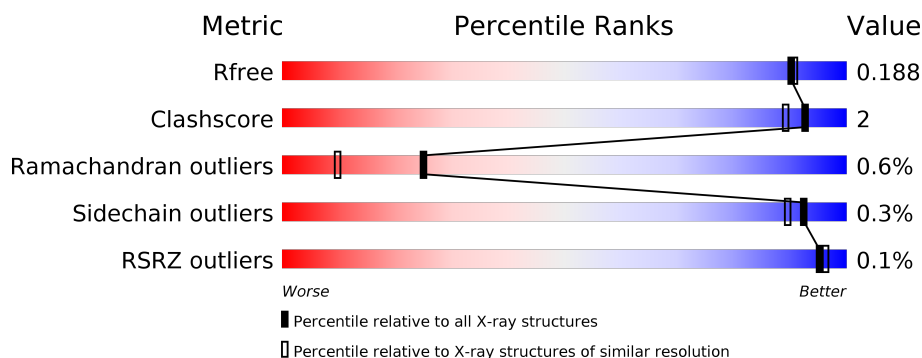
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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.



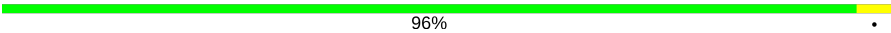
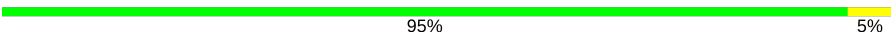
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

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 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  $\leq 5\%$ . 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	335	<div><div></div><div>96%</div><div></div></div>
1	B	335	<div><div></div><div>97%</div><div></div></div>
1	C	335	<div><div></div><div>96%</div><div></div></div>
1	D	335	<div><div></div><div>94%</div><div>6%</div></div>
1	E	335	<div><div></div><div>94%</div><div>6%</div></div>
1	F	335	<div><div></div><div>96%</div><div></div></div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	335	 96% .
1	H	335	 95% 5%

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 23979 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 5-carboxyvanillate decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	335	Total	C	N	O	S	0	1	0
			2615	1660	460	482	13			
1	B	335	Total	C	N	O	S	0	3	0
			2637	1672	468	484	13			
1	C	335	Total	C	N	O	S	0	2	0
			2623	1664	461	485	13			
1	D	335	Total	C	N	O	S	0	2	0
			2631	1669	467	482	13			
1	E	335	Total	C	N	O	S	0	4	0
			2640	1674	464	488	14			
1	F	335	Total	C	N	O	S	0	1	0
			2615	1660	460	482	13			
1	G	335	Total	C	N	O	S	0	1	0
			2615	1660	460	482	13			
1	H	335	Total	C	N	O	S	0	2	0
			2631	1669	467	482	13			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	EXPRESSION TAG	UNP Q8RJ47
A	1	LEU	-	EXPRESSION TAG	UNP Q8RJ47
B	0	SER	-	EXPRESSION TAG	UNP Q8RJ47
B	1	LEU	-	EXPRESSION TAG	UNP Q8RJ47
C	0	SER	-	EXPRESSION TAG	UNP Q8RJ47
C	1	LEU	-	EXPRESSION TAG	UNP Q8RJ47
D	0	SER	-	EXPRESSION TAG	UNP Q8RJ47
D	1	LEU	-	EXPRESSION TAG	UNP Q8RJ47
E	0	SER	-	EXPRESSION TAG	UNP Q8RJ47
E	1	LEU	-	EXPRESSION TAG	UNP Q8RJ47
F	0	SER	-	EXPRESSION TAG	UNP Q8RJ47
F	1	LEU	-	EXPRESSION TAG	UNP Q8RJ47
G	0	SER	-	EXPRESSION TAG	UNP Q8RJ47

*Continued on next page...*

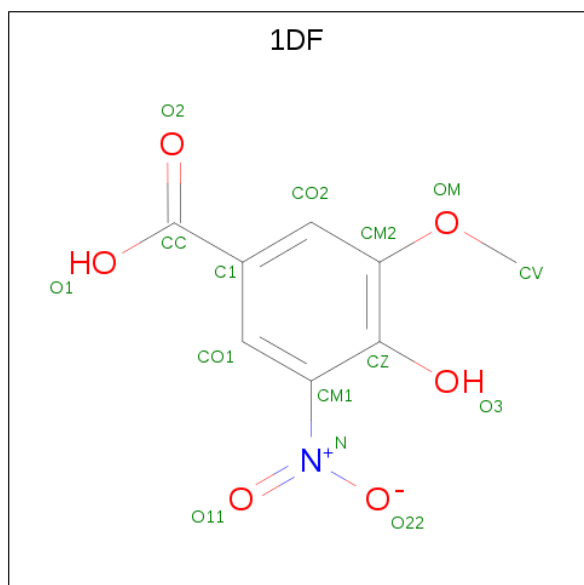
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	1	LEU	-	EXPRESSION TAG	UNP Q8RJ47
H	0	SER	-	EXPRESSION TAG	UNP Q8RJ47
H	1	LEU	-	EXPRESSION TAG	UNP Q8RJ47

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

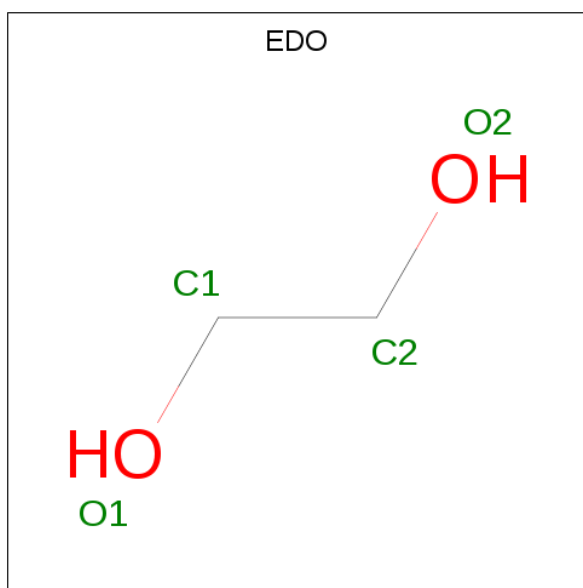
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	1	Total Mn 1 1	0	0
2	D	1	Total Mn 1 1	0	0
2	E	1	Total Mn 1 1	0	0
2	H	1	Total Mn 1 1	0	0
2	B	1	Total Mn 1 1	0	0
2	C	1	Total Mn 1 1	0	0
2	A	1	Total Mn 1 1	0	0
2	F	1	Total Mn 1 1	0	0

- Molecule 3 is 4-hydroxy-3-methoxy-5-nitrobenzoic acid (three-letter code: 1DF) (formula: C<sub>8</sub>H<sub>7</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			15	8	1	6		
3	B	1	Total	C	N	O	0	0
			15	8	1	6		
3	C	1	Total	C	N	O	0	0
			15	8	1	6		
3	D	1	Total	C	N	O	0	0
			15	8	1	6		
3	E	1	Total	C	N	O	0	0
			15	8	1	6		
3	F	1	Total	C	N	O	0	0
			15	8	1	6		
3	G	1	Total	C	N	O	0	0
			15	8	1	6		
3	H	1	Total	C	N	O	0	0
			15	8	1	6		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



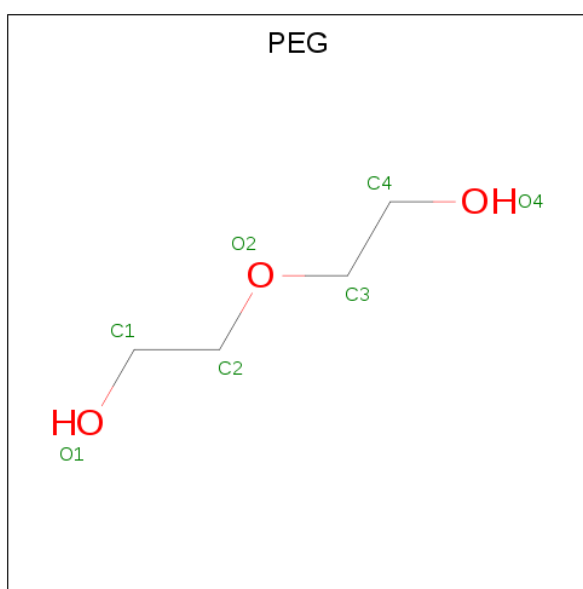
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	C	1	Total	C	O	0	0
			4	2	2		
4	D	1	Total	C	O	0	0
			4	2	2		
4	E	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	E	1	Total	C	O	0	0
			4	2	2		
4	G	1	Total	C	O	0	0
			4	2	2		
4	H	1	Total	C	O	0	0
			4	2	2		
4	H	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			7	4	3		
5	E	1	Total	C	O	0	0
			7	4	3		
5	E	1	Total	C	O	0	0
			7	4	3		
5	F	1	Total	C	O	0	0
			7	4	3		
5	H	1	Total	C	O	0	0
			7	4	3		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	341	Total 341	O 341	0	0
6	B	366	Total 366	O 366	0	0
6	C	343	Total 343	O 343	0	0
6	D	360	Total 360	O 360	0	0
6	E	351	Total 351	O 351	0	0
6	F	341	Total 341	O 341	0	0
6	G	345	Total 345	O 345	0	0
6	H	330	Total 330	O 330	0	0



### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: 5-carboxyvanillate decarboxylase

Chain A:  96%



- Molecule 1: 5-carboxyvanillate decarboxylase

Chain B:  97%



- Molecule 1: 5-carboxyvanillate decarboxylase

Chain C:  96%



- Molecule 1: 5-carboxyvanillate decarboxylase

Chain D:  94% 6%



- Molecule 1: 5-carboxyvanillate decarboxylase

Chain E:  94% 6%



- Molecule 1: 5-carboxyvanillate decarboxylase

Chain F:  96% .



- Molecule 1: 5-carboxyvanillate decarboxylase

Chain G:  96% .



- Molecule 1: 5-carboxyvanillate decarboxylase

Chain H:  95% 5% .



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	80.54Å 100.40Å 100.02Å 66.64° 67.96° 88.24°	Depositor
Resolution (Å)	48.71 – 1.75 48.71 – 1.75	Depositor EDS
% Data completeness (in resolution range)	96.7 (48.71-1.75) 96.7 (48.71-1.75)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.05 (at 1.75Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, $R_{free}$	0.152 , 0.188 0.151 , 0.188	Depositor DCC
$R_{free}$ test set	12939 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.7	Xtriage
Anisotropy	0.140	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 42.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	23979	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% of the height of the origin peak. No significant pseudotranslation is detected.*

<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, 0.333 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: 1DF, PEG, EDO, MN

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	0.35	0/2678	0.54	0/3641
1	B	0.35	0/2700	0.53	0/3669
1	C	0.34	0/2686	0.53	0/3652
1	D	0.35	0/2694	0.54	0/3661
1	E	0.36	0/2703	0.54	0/3674
1	F	0.34	0/2678	0.53	0/3641
1	G	0.35	0/2678	0.54	0/3641
1	H	0.33	0/2694	0.53	0/3661
All	All	0.35	0/21511	0.54	0/29240

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	2615	0	2581	8	0
1	B	2637	0	2605	5	0
1	C	2623	0	2584	8	0
1	D	2631	0	2601	12	0
1	E	2640	0	2599	11	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	2615	0	2581	8	0
1	G	2615	0	2581	8	0
1	H	2631	0	2601	10	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
2	E	1	0	0	0	0
2	F	1	0	0	0	0
2	G	1	0	0	0	0
2	H	1	0	0	0	0
3	A	15	0	5	0	0
3	B	15	0	5	0	0
3	C	15	0	5	0	0
3	D	15	0	5	0	0
3	E	15	0	5	0	0
3	F	15	0	5	0	0
3	G	15	0	5	0	0
3	H	15	0	5	0	0
4	A	4	0	6	0	0
4	C	4	0	6	0	0
4	D	4	0	6	0	0
4	E	8	0	12	1	0
4	G	4	0	6	0	0
4	H	8	0	12	0	0
5	B	7	0	10	0	0
5	E	14	0	20	1	0
5	F	7	0	10	0	0
5	H	7	0	10	0	0
6	A	341	0	0	2	0
6	B	366	0	0	1	0
6	C	343	0	0	4	0
6	D	360	0	0	2	0
6	E	351	0	0	2	0
6	F	341	0	0	3	0
6	G	345	0	0	3	0
6	H	330	0	0	3	0
All	All	23979	0	20871	64	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 2.

All (64) 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:F:263:ARG:NH2	6:F:782:HOH:O	2.21	0.71
1:E:257:LYS:NZ	6:E:840:HOH:O	2.26	0.69
1:E:263:ARG:NH2	6:E:846:HOH:O	2.27	0.68
1:D:167:GLY:O	1:D:332[B]:ARG:NH2	2.31	0.63
1:C:280:ARG:NH2	6:C:690:HOH:O	2.25	0.63
1:A:13:GLN:NE2	6:A:689:HOH:O	2.31	0.62
1:D:36:ARG:NH2	6:D:583:HOH:O	2.33	0.61
1:D:280:ARG:NH2	1:E:284:GLU:OE1	2.33	0.61
1:D:51:ARG:NH2	6:D:742:HOH:O	2.33	0.60
1:C:126:ARG:NH1	6:C:691:HOH:O	2.35	0.59
1:G:126:ARG:NH1	6:G:745:HOH:O	2.34	0.59
1:A:284:GLU:OE1	1:H:280:ARG:NH2	2.37	0.58
1:H:20:ARG:NH2	6:H:730:HOH:O	2.35	0.57
1:D:86:GLU:OE2	1:D:90[A]:ARG:NH1	2.38	0.57
1:G:177:PRO:HB3	1:G:181:ILE:HG13	1.86	0.57
1:F:51:ARG:NH2	6:F:704:HOH:O	2.35	0.56
1:D:284:GLU:OE2	1:E:280:ARG:NH2	2.33	0.55
1:E:3:LEU:HD23	1:E:323:ALA:HB1	1.89	0.55
1:B:3:LEU:HD23	1:B:323:ALA:HB1	1.90	0.54
1:B:280:ARG:NH1	1:G:284:GLU:OE1	2.38	0.53
1:A:177:PRO:HB3	1:A:181:ILE:HG13	1.90	0.53
1:E:177:PRO:HB3	1:E:181:ILE:HG13	1.92	0.52
1:B:177:PRO:HB3	1:B:181:ILE:HG13	1.92	0.51
1:E:213:LEU:HA	4:E:405:EDO:H12	1.92	0.50
1:H:284:GLU:OE1	6:H:826:HOH:O	2.20	0.49
1:G:269:THR:HA	1:G:292:MET:O	2.14	0.47
1:E:45:ARG:HB3	5:E:404:PEG:H41	1.95	0.47
1:H:3:LEU:HD23	1:H:323:ALA:HB1	1.98	0.46
1:B:0:SER:HB2	6:C:801:HOH:O	2.16	0.46
1:A:0:SER:HB2	6:B:711:HOH:O	2.16	0.45
1:H:253:LYS:HE2	6:H:691:HOH:O	2.16	0.45
1:E:262:PHE:O	1:E:266:PHE:HB2	2.17	0.44
1:C:262:PHE:O	1:C:266:PHE:HB2	2.18	0.44
1:G:93:ARG:NH1	6:G:634:HOH:O	2.43	0.44
1:F:269:THR:HA	1:F:292:MET:O	2.17	0.44
1:H:269:THR:HA	1:H:292:MET:O	2.19	0.43
6:C:703:HOH:O	1:D:219:ARG:HD3	2.18	0.43
1:D:296:ASP:HB3	1:D:300:GLN:HG3	2.01	0.43
1:A:269:THR:HA	1:A:292:MET:O	2.19	0.43
1:C:280:ARG:HH22	1:F:284:GLU:CD	2.21	0.43
1:B:6:THR:HA	1:B:72:LEU:HB3	2.00	0.43
1:D:269:THR:HA	1:D:292:MET:O	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:30:LEU:HD22	1:C:182:ASP:HB2	2.00	0.43
1:A:138:ASN:HB3	1:A:331:PHE:O	2.19	0.42
1:D:3:LEU:HD23	1:D:323:ALA:HB1	2.01	0.42
1:F:6:THR:HA	1:F:72:LEU:HB3	2.02	0.42
1:D:177:PRO:HB3	1:D:181:ILE:HG13	2.01	0.42
1:G:329:ARG:NH2	6:G:786:HOH:O	2.42	0.42
1:C:269:THR:HA	1:C:292:MET:O	2.19	0.42
1:F:62:GLU:OE1	6:F:729:HOH:O	2.21	0.42
1:G:262:PHE:O	1:G:266:PHE:HB2	2.19	0.42
1:E:269:THR:HA	1:E:292:MET:O	2.19	0.41
1:A:296:ASP:HB3	1:A:300:GLN:HG3	2.03	0.41
1:A:280:ARG:NH2	6:A:661:HOH:O	2.31	0.41
1:E:116:PHE:O	1:E:118:PRO:HD3	2.20	0.41
1:C:274:GLU:HB2	1:F:236:TRP:CZ2	2.56	0.41
1:F:296:ASP:HB3	1:F:300:GLN:HG3	2.03	0.41
1:H:30:LEU:HD22	1:H:182:ASP:HB2	2.02	0.41
1:C:177:PRO:HB3	1:C:181:ILE:HG13	2.02	0.40
1:H:303:ALA:HB3	1:H:304:PRO:HD3	2.04	0.40
1:H:9:ALA:HB1	1:H:76:THR:HB	2.02	0.40
1:G:177:PRO:HA	1:G:200:GLU:HB2	2.03	0.40
1:H:142:ILE:HA	1:H:142:ILE:HD12	2.00	0.40
1:D:226:HIS:HA	1:D:271:SER:O	2.22	0.40

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	334/335 (100%)	327 (98%)	6 (2%)	1 (0%)	41	22
1	B	336/335 (100%)	326 (97%)	8 (2%)	2 (1%)	25	10
1	C	335/335 (100%)	326 (97%)	7 (2%)	2 (1%)	25	10

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	335/335 (100%)	326 (97%)	7 (2%)	2 (1%)	25	10
1	E	337/335 (101%)	330 (98%)	5 (2%)	2 (1%)	25	10
1	F	334/335 (100%)	326 (98%)	5 (2%)	3 (1%)	17	5
1	G	334/335 (100%)	327 (98%)	5 (2%)	2 (1%)	25	10
1	H	335/335 (100%)	326 (97%)	7 (2%)	2 (1%)	25	10
All	All	2680/2680 (100%)	2614 (98%)	50 (2%)	16 (1%)	25	10

All (16) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	144	SER
1	F	144	SER
1	A	324	GLY
1	B	324	GLY
1	C	324	GLY
1	D	324	GLY
1	E	144	SER
1	E	324	GLY
1	F	324	GLY
1	G	144	SER
1	G	324	GLY
1	H	144	SER
1	H	324	GLY
1	C	144	SER
1	D	144	SER
1	F	195	TRP

### 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	271/270 (100%)	270 (100%)	1 (0%)	91	87
1	B	273/270 (101%)	272 (100%)	1 (0%)	91	87

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	272/270 (101%)	271 (100%)	1 (0%)	91	87
1	D	272/270 (101%)	272 (100%)	0	100	100
1	E	274/270 (102%)	273 (100%)	1 (0%)	91	87
1	F	271/270 (100%)	271 (100%)	0	100	100
1	G	271/270 (100%)	270 (100%)	1 (0%)	91	87
1	H	272/270 (101%)	271 (100%)	1 (0%)	91	87
All	All	2176/2160 (101%)	2170 (100%)	6 (0%)	92	89

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	171	TYR
1	B	171	TYR
1	C	171	TYR
1	E	171	TYR
1	G	171	TYR
1	H	171	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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 29 ligands modelled in this entry, 8 are monoatomic - leaving 21 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	EDO	H	405	-	3,3,3	0.51	0	2,2,2	0.22	0
3	1DF	C	402	2	12,15,15	1.91	2 (16%)	13,21,21	1.73	3 (23%)
3	1DF	B	402	2	12,15,15	1.84	2 (16%)	13,21,21	1.85	4 (30%)
4	EDO	H	404	-	3,3,3	0.44	0	2,2,2	0.43	0
3	1DF	D	402	2	12,15,15	2.00	2 (16%)	13,21,21	1.82	4 (30%)
3	1DF	G	402	2	12,15,15	2.03	3 (25%)	13,21,21	1.41	2 (15%)
3	1DF	F	402	2	12,15,15	1.97	2 (16%)	13,21,21	1.73	3 (23%)
4	EDO	E	405	-	3,3,3	0.51	0	2,2,2	0.26	0
5	PEG	E	403	-	6,6,6	0.49	0	5,5,5	0.22	0
5	PEG	F	403	-	6,6,6	0.44	0	5,5,5	0.28	0
4	EDO	A	403	-	3,3,3	0.50	0	2,2,2	0.20	0
5	PEG	H	403	-	6,6,6	0.46	0	5,5,5	0.37	0
4	EDO	E	406	-	3,3,3	0.45	0	2,2,2	0.41	0
4	EDO	G	403	-	3,3,3	0.46	0	2,2,2	0.24	0
4	EDO	C	403	-	3,3,3	0.45	0	2,2,2	0.29	0
3	1DF	H	402	2	12,15,15	1.77	1 (8%)	13,21,21	1.37	1 (7%)
3	1DF	E	402	2	12,15,15	2.04	2 (16%)	13,21,21	1.46	1 (7%)
4	EDO	D	403	-	3,3,3	0.48	0	2,2,2	0.23	0
5	PEG	B	403	-	6,6,6	0.45	0	5,5,5	0.24	0
5	PEG	E	404	-	6,6,6	0.42	0	5,5,5	0.37	0
3	1DF	A	402	2	12,15,15	1.75	1 (8%)	13,21,21	1.71	4 (30%)

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	EDO	H	405	-	-	0/1/1/1	-
3	1DF	C	402	2	-	0/4/10/10	0/1/1/1
3	1DF	B	402	2	-	2/4/10/10	0/1/1/1
4	EDO	H	404	-	-	1/1/1/1	-
3	1DF	D	402	2	-	2/4/10/10	0/1/1/1

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	1DF	G	402	2	-	0/4/10/10	0/1/1/1
3	1DF	F	402	2	-	2/4/10/10	0/1/1/1
4	EDO	E	405	-	-	1/1/1/1	-
5	PEG	E	403	-	-	3/4/4/4	-
5	PEG	F	403	-	-	1/4/4/4	-
4	EDO	A	403	-	-	0/1/1/1	-
5	PEG	H	403	-	-	1/4/4/4	-
4	EDO	E	406	-	-	0/1/1/1	-
4	EDO	G	403	-	-	0/1/1/1	-
4	EDO	C	403	-	-	0/1/1/1	-
3	1DF	H	402	2	-	0/4/10/10	0/1/1/1
3	1DF	E	402	2	-	0/4/10/10	0/1/1/1
4	EDO	D	403	-	-	0/1/1/1	-
5	PEG	B	403	-	-	0/4/4/4	-
5	PEG	E	404	-	-	2/4/4/4	-
3	1DF	A	402	2	-	0/4/10/10	0/1/1/1

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	402	1DF	O11-N	5.98	1.32	1.22
3	D	402	1DF	O11-N	5.94	1.32	1.22
3	E	402	1DF	O11-N	5.85	1.32	1.22
3	F	402	1DF	O11-N	5.75	1.32	1.22
3	C	402	1DF	O11-N	5.65	1.32	1.22
3	A	402	1DF	O11-N	5.28	1.31	1.22
3	H	402	1DF	O11-N	5.07	1.31	1.22
3	B	402	1DF	O11-N	4.85	1.31	1.22
3	B	402	1DF	C1-CC	2.85	1.50	1.47
3	D	402	1DF	C1-CC	2.49	1.49	1.47
3	E	402	1DF	C1-CC	2.40	1.49	1.47
3	C	402	1DF	C1-CC	2.38	1.49	1.47
3	F	402	1DF	C1-CC	2.37	1.49	1.47
3	G	402	1DF	C1-CC	2.14	1.49	1.47
3	G	402	1DF	CM2-CZ	-2.12	1.37	1.40

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	402	1DF	OM-CM2-CZ	4.68	119.26	114.54
3	F	402	1DF	OM-CM2-CZ	4.52	119.10	114.54

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	402	1DF	OM-CM2-CZ	4.16	118.74	114.54
3	D	402	1DF	OM-CM2-CZ	3.91	118.49	114.54
3	E	402	1DF	OM-CM2-CZ	3.66	118.23	114.54
3	H	402	1DF	OM-CM2-CZ	3.34	117.91	114.54
3	A	402	1DF	CV-OM-CM2	2.93	121.95	117.53
3	G	402	1DF	OM-CM2-CZ	2.77	117.34	114.54
3	A	402	1DF	OM-CM2-CZ	2.60	117.17	114.54
3	A	402	1DF	CO2-C1-CC	-2.54	117.02	120.36
3	C	402	1DF	OM-CM2-CO2	-2.45	119.90	124.12
3	B	402	1DF	OM-CM2-CO2	-2.39	120.01	124.12
3	A	402	1DF	CO1-CM1-CZ	-2.36	119.39	122.51
3	B	402	1DF	CO1-C1-CC	2.21	123.27	120.36
3	B	402	1DF	CO2-C1-CC	-2.21	117.45	120.36
3	D	402	1DF	CV-OM-CM2	2.19	120.84	117.53
3	G	402	1DF	O3-CZ-CM2	-2.19	114.50	119.24
3	F	402	1DF	OM-CM2-CO2	-2.17	120.39	124.12
3	D	402	1DF	CO2-C1-CC	-2.13	117.56	120.36
3	C	402	1DF	CO1-CM1-CZ	-2.12	119.71	122.51
3	D	402	1DF	CO1-CM1-CZ	-2.07	119.77	122.51
3	F	402	1DF	O11-N-CM1	2.07	122.57	119.03

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	404	PEG	O2-C3-C4-O4
3	B	402	1DF	CZ-CM2-OM-CV
5	E	403	PEG	C4-C3-O2-C2
5	E	403	PEG	O1-C1-C2-O2
4	E	405	EDO	O1-C1-C2-O2
3	B	402	1DF	CO2-CM2-OM-CV
3	D	402	1DF	CZ-CM2-OM-CV
3	F	402	1DF	CZ-CM2-OM-CV
3	D	402	1DF	CO2-CM2-OM-CV
5	H	403	PEG	O1-C1-C2-O2
5	F	403	PEG	O2-C3-C4-O4
5	E	404	PEG	C1-C2-O2-C3
5	E	403	PEG	C1-C2-O2-C3
3	F	402	1DF	CO2-CM2-OM-CV
4	H	404	EDO	O1-C1-C2-O2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	405	EDO	1	0
5	E	404	PEG	1	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	335/335 (100%)	-0.56	1 (0%) 94 95	6, 11, 24, 36	0
1	B	335/335 (100%)	-0.56	0 100 100	6, 11, 22, 43	0
1	C	335/335 (100%)	-0.58	1 (0%) 94 95	6, 11, 25, 40	0
1	D	335/335 (100%)	-0.56	0 100 100	6, 10, 21, 40	0
1	E	335/335 (100%)	-0.56	1 (0%) 94 95	6, 10, 20, 37	0
1	F	335/335 (100%)	-0.57	0 100 100	6, 11, 22, 37	0
1	G	335/335 (100%)	-0.53	0 100 100	6, 11, 22, 39	0
1	H	335/335 (100%)	-0.53	0 100 100	7, 12, 24, 37	0
All	All	2680/2680 (100%)	-0.55	3 (0%) 95 96	6, 11, 23, 43	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	334	THR	2.7
1	A	0	SER	2.3
1	C	334	THR	2.1

### 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 ⓘ

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<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	B-factors(Å <sup>2</sup> )	Q<0.9
5	PEG	E	404	7/7	0.83	0.30	22,38,47,48	0
4	EDO	E	406	4/4	0.85	0.18	33,37,38,39	0
4	EDO	E	405	4/4	0.85	0.14	31,33,34,34	0
4	EDO	H	405	4/4	0.88	0.14	27,28,28,30	0
5	PEG	H	403	7/7	0.91	0.11	13,27,40,45	0
5	PEG	E	403	7/7	0.92	0.10	14,27,34,37	0
4	EDO	H	404	4/4	0.93	0.11	15,22,28,39	0
3	1DF	G	402	15/15	0.94	0.08	6,12,22,24	0
5	PEG	F	403	7/7	0.94	0.12	14,30,34,37	0
4	EDO	D	403	4/4	0.94	0.10	15,20,27,29	0
4	EDO	G	403	4/4	0.94	0.12	18,18,29,32	0
4	EDO	A	403	4/4	0.94	0.09	20,24,26,26	0
5	PEG	B	403	7/7	0.95	0.09	15,27,42,44	0
3	1DF	E	402	15/15	0.95	0.08	7,11,20,20	0
3	1DF	F	402	15/15	0.95	0.08	9,15,22,22	0
3	1DF	D	402	15/15	0.96	0.06	8,14,23,26	0
3	1DF	B	402	15/15	0.96	0.07	9,16,22,23	0
3	1DF	C	402	15/15	0.96	0.07	8,14,20,22	0
4	EDO	C	403	4/4	0.96	0.08	15,20,26,30	0
3	1DF	H	402	15/15	0.96	0.07	10,16,23,25	0
3	1DF	A	402	15/15	0.96	0.08	8,13,18,22	0
2	MN	G	401	1/1	0.99	0.04	9,9,9,9	1
2	MN	F	401	1/1	1.00	0.03	10,10,10,10	1
2	MN	C	401	1/1	1.00	0.04	9,9,9,9	1
2	MN	B	401	1/1	1.00	0.04	8,8,8,8	1
2	MN	H	401	1/1	1.00	0.03	8,8,8,8	1
2	MN	A	401	1/1	1.00	0.04	7,7,7,7	1
2	MN	E	401	1/1	1.00	0.05	8,8,8,8	1
2	MN	D	401	1/1	1.00	0.05	9,9,9,9	1

## 6.5 Other polymers ⓘ

There are no such residues in this entry.