



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 01:00 pm BST

PDB ID : 4Q31  
Title : The crystal structure of cystathione gamma lyase (CalE6) from *Micromonospora echinospora*  
Authors : Tan, K.; Bigelow, L.; Jedrzejczak, R.; Babnigg, G.; Bingman, C.A.; Yen-namalli, R.M.; Singh, S.; Kharel, M.K.; Thorson, J.S.; Phillips Jr., G.N.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG); Enzyme Discovery for Natural Product Biosynthesis (NatPro)  
Deposited on : 2014-04-10  
Resolution : 2.10 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

<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

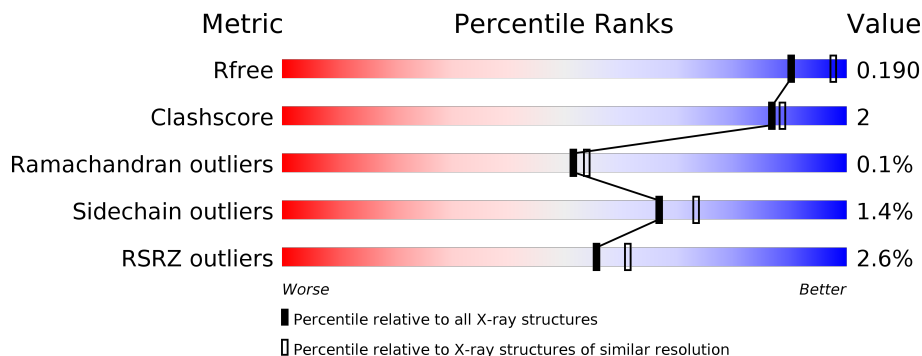
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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	384	 2% 92% 5%
1	B	384	 2% 93% 5%
1	C	384	 4% 91% 8%
1	D	384	 4% 92% 6%
1	E	384	 4% 92% 6%
1	F	384	 2% 92% 6%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	384	 2% 93% 5%
1	H	384	 2% 95%

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	GOL	D	407	-	-	X	-

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 24719 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 cystathione gamma lyase CalE6.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S	Se			
1	A	376	2844	1778	526	530	1	5	4	0	1	0
1	B	378	2851	1781	528	532	1	5	4	0	1	0
1	C	379	2827	1769	523	526	1	5	3	0	0	0
1	D	376	2817	1764	519	524	1	5	4	0	1	0
1	E	376	2813	1764	512	528	1	5	3	0	0	0
1	F	377	2812	1760	518	525	1	5	3	0	0	0
1	G	378	2814	1761	516	528	1	5	3	0	0	0
1	H	378	2828	1770	520	529	1	5	3	0	1	0

There are 32 discrepancies between the modelled and reference sequences:

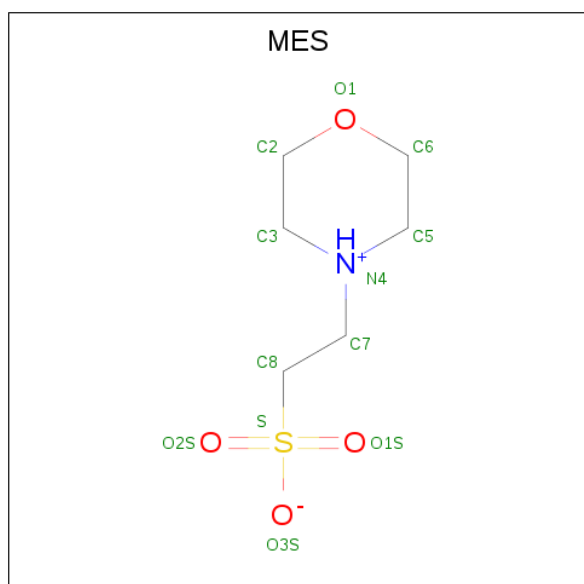
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q8KNG3
A	-1	ASN	-	EXPRESSION TAG	UNP Q8KNG3
A	0	ALA	-	EXPRESSION TAG	UNP Q8KNG3
A	7	GLY	ASP	engineered mutation	UNP Q8KNG3
B	-2	SER	-	EXPRESSION TAG	UNP Q8KNG3
B	-1	ASN	-	EXPRESSION TAG	UNP Q8KNG3
B	0	ALA	-	EXPRESSION TAG	UNP Q8KNG3
B	7	GLY	ASP	engineered mutation	UNP Q8KNG3
C	-2	SER	-	EXPRESSION TAG	UNP Q8KNG3
C	-1	ASN	-	EXPRESSION TAG	UNP Q8KNG3
C	0	ALA	-	EXPRESSION TAG	UNP Q8KNG3
C	7	GLY	ASP	engineered mutation	UNP Q8KNG3
D	-2	SER	-	EXPRESSION TAG	UNP Q8KNG3

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	ASN	-	EXPRESSION TAG	UNP Q8KNG3
D	0	ALA	-	EXPRESSION TAG	UNP Q8KNG3
D	7	GLY	ASP	engineered mutation	UNP Q8KNG3
E	-2	SER	-	EXPRESSION TAG	UNP Q8KNG3
E	-1	ASN	-	EXPRESSION TAG	UNP Q8KNG3
E	0	ALA	-	EXPRESSION TAG	UNP Q8KNG3
E	7	GLY	ASP	engineered mutation	UNP Q8KNG3
F	-2	SER	-	EXPRESSION TAG	UNP Q8KNG3
F	-1	ASN	-	EXPRESSION TAG	UNP Q8KNG3
F	0	ALA	-	EXPRESSION TAG	UNP Q8KNG3
F	7	GLY	ASP	engineered mutation	UNP Q8KNG3
G	-2	SER	-	EXPRESSION TAG	UNP Q8KNG3
G	-1	ASN	-	EXPRESSION TAG	UNP Q8KNG3
G	0	ALA	-	EXPRESSION TAG	UNP Q8KNG3
G	7	GLY	ASP	engineered mutation	UNP Q8KNG3
H	-2	SER	-	EXPRESSION TAG	UNP Q8KNG3
H	-1	ASN	-	EXPRESSION TAG	UNP Q8KNG3
H	0	ALA	-	EXPRESSION TAG	UNP Q8KNG3
H	7	GLY	ASP	engineered mutation	UNP Q8KNG3

- Molecule 2 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>S).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	
			Total	C	N	O			S
2	A	1	12	6	1	4	1	0	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
2	C	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
2	D	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
2	E	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
2	F	1	Total	C	N	O	S	0	0
			11	6	1	3	1		
2	G	1	Total	C	N	O	S	0	0
			12	6	1	4	1		
2	H	1	Total	C	N	O	S	0	0
			12	6	1	4	1		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	A	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0
4	C	1	Total C O 6 3 3	0	0
4	D	1	Total C O 6 3 3	0	0
4	D	1	Total C O 6 3 3	0	0
4	D	1	Total C O 6 3 3	0	0

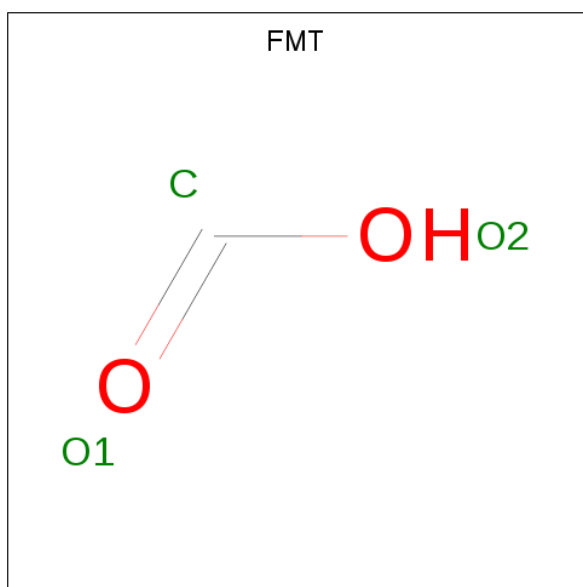
*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total 6	C 3	O 3	0	0
4	E	1	Total 6	C 3	O 3	0	0
4	E	1	Total 6	C 3	O 3	0	0
4	E	1	Total 6	C 3	O 3	0	0
4	E	1	Total 6	C 3	O 3	0	0
4	F	1	Total 6	C 3	O 3	0	0
4	F	1	Total 6	C 3	O 3	0	0
4	F	1	Total 6	C 3	O 3	0	0
4	G	1	Total 6	C 3	O 3	0	0
4	G	1	Total 6	C 3	O 3	0	0
4	G	1	Total 6	C 3	O 3	0	0
4	G	1	Total 6	C 3	O 3	0	0
4	G	1	Total 6	C 3	O 3	0	0
4	H	1	Total 6	C 3	O 3	0	0

- Molecule 5 is FORMIC ACID (three-letter code: FMT) (formula: CH<sub>2</sub>O<sub>2</sub>).





Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 3 1 2	0	0
5	C	1	Total C O 3 1 2	0	0
5	F	1	Total C O 3 1 2	0	0
5	G	1	Total C O 3 1 2	0	0
5	H	1	Total C O 3 1 2	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	243	Total O 243 243	0	0
6	B	228	Total O 228 228	0	0
6	C	230	Total O 230 230	0	0
6	D	215	Total O 215 215	0	0
6	E	250	Total O 250 250	0	0
6	F	204	Total O 204 204	0	0
6	G	215	Total O 215 215	0	0

*Continued on next page...*

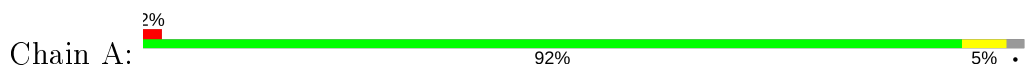
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	H	238	Total 238	O 238	0	0

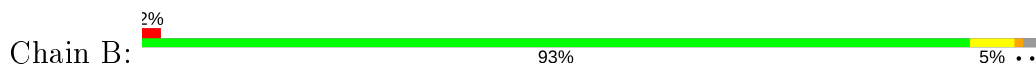
### 3 Residue-property plots [i](#)

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.

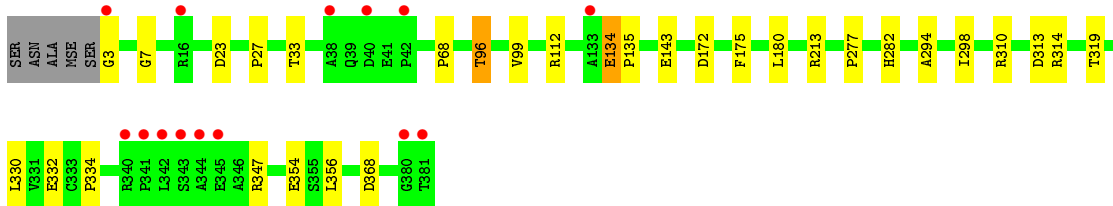
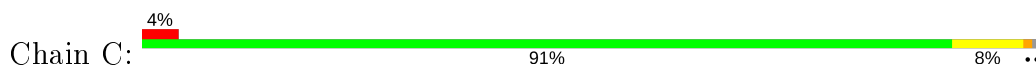
- Molecule 1: cystathione gamma lyase CalE6



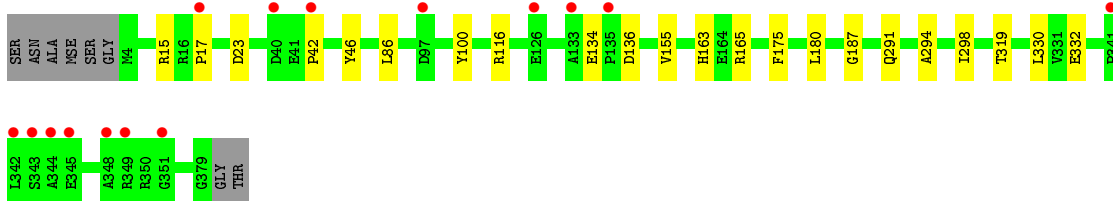
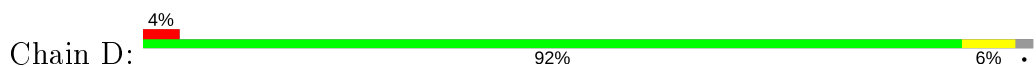
- Molecule 1: cystathione gamma lyase CalE6



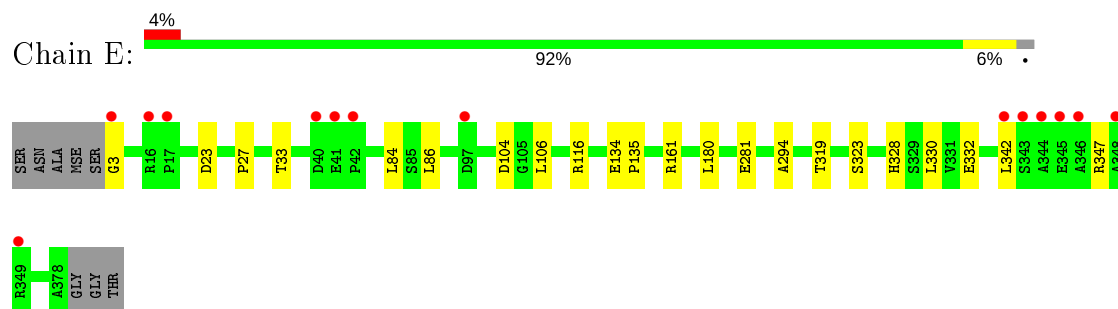
- Molecule 1: cystathione gamma lyase CalE6



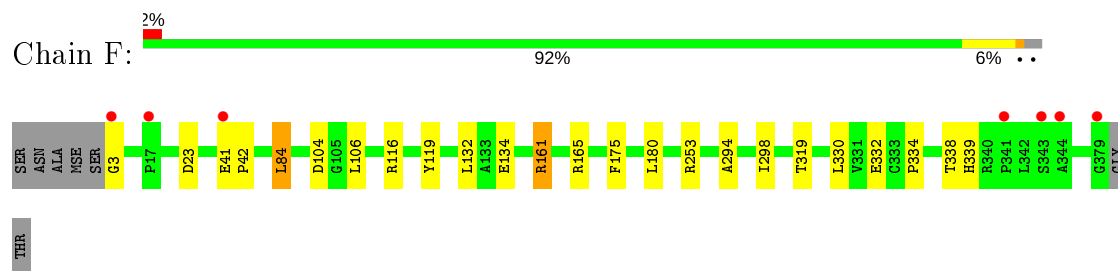
- Molecule 1: cystathione gamma lyase CalE6



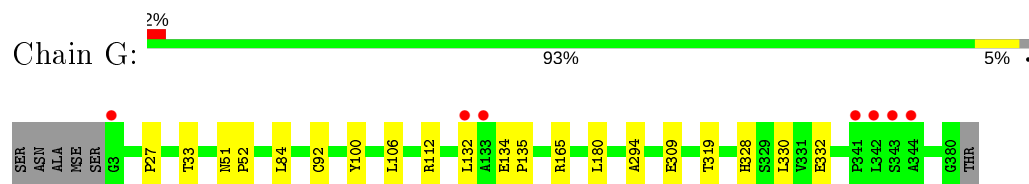
- Molecule 1: cystathione gamma lyase CalE6



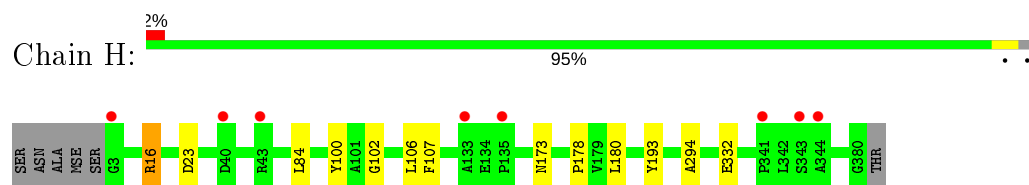
- Molecule 1: cystathione gamma lyase CalE6



- Molecule 1: cystathione gamma lyase CalE6



- Molecule 1: cystathione gamma lyase CalE6



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	146.85Å 146.98Å 349.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.97 – 2.10 33.97 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.9 (33.97-2.10) 99.9 (33.97-2.10)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.01 (at 2.10Å)	Xtrriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, $R_{free}$	0.152 , 0.191 0.152 , 0.190	Depositor DCC
$R_{free}$ test set	11002 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.8	Xtrriage
Anisotropy	0.216	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 52.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.055 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	24719	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.26% 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: GOL, FMT, LLP, MES, CL

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	0.36	0/2876	0.51	0/3916
1	B	0.37	0/2882	0.52	0/3922
1	C	0.37	0/2855	0.52	0/3890
1	D	0.35	0/2849	0.51	0/3884
1	E	0.37	0/2841	0.50	0/3871
1	F	0.36	0/2840	0.51	0/3871
1	G	0.35	0/2842	0.51	0/3874
1	H	0.36	0/2859	0.51	0/3894
All	All	0.36	0/22844	0.51	0/31122

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	2844	0	2841	9	0
1	B	2851	0	2846	14	0
1	C	2827	0	2810	20	0
1	D	2817	0	2798	14	0
1	E	2813	0	2789	11	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	2812	0	2786	14	0
1	G	2814	0	2786	10	0
1	H	2828	0	2811	8	0
2	A	12	0	12	0	0
2	B	12	0	12	1	0
2	C	12	0	12	1	0
2	D	12	0	12	1	0
2	E	12	0	12	0	0
2	F	11	0	12	0	0
2	G	12	0	12	2	0
2	H	12	0	12	1	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
3	C	2	0	0	1	0
3	D	2	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
3	G	1	0	0	0	0
3	H	2	0	0	0	0
4	A	24	0	32	1	0
4	B	36	0	48	1	0
4	C	6	0	8	0	0
4	D	24	0	32	5	0
4	E	24	0	32	1	0
4	F	18	0	24	1	0
4	G	30	0	40	1	0
4	H	6	0	8	0	0
5	A	3	0	1	0	0
5	C	3	0	1	0	0
5	F	3	0	1	0	0
5	G	3	0	1	0	0
5	H	3	0	1	0	0
6	A	243	0	0	0	0
6	B	228	0	0	2	0
6	C	230	0	0	2	0
6	D	215	0	0	0	0
6	E	250	0	0	2	0
6	F	204	0	0	1	0
6	G	215	0	0	0	0
6	H	238	0	0	1	0
All	All	24719	0	22792	96	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.

The worst 5 of 96 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:E:86:LEU:HB3	4:E:405:GOL:H11	1.75	0.67
1:F:134:GLU:O	1:F:165:ARG:NH1	2.28	0.67
1:C:3:GLY:N	6:C:644:HOH:O	2.29	0.65
1:F:338:THR:HG23	1:F:339:HIS:ND1	2.12	0.64
1:F:253:ARG:HD2	4:F:404:GOL:H11	1.83	0.61

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	373/384 (97%)	366 (98%)	7 (2%)	0	100	100
1	B	376/384 (98%)	364 (97%)	12 (3%)	0	100	100
1	C	376/384 (98%)	367 (98%)	9 (2%)	0	100	100
1	D	373/384 (97%)	363 (97%)	9 (2%)	1 (0%)	41	41
1	E	373/384 (97%)	366 (98%)	6 (2%)	1 (0%)	41	41
1	F	374/384 (97%)	367 (98%)	7 (2%)	0	100	100
1	G	375/384 (98%)	364 (97%)	11 (3%)	0	100	100
1	H	376/384 (98%)	365 (97%)	11 (3%)	0	100	100
All	All	2996/3072 (98%)	2922 (98%)	72 (2%)	2 (0%)	51	54

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	323	SER
1	D	42	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	292/293 (100%)	286 (98%)	6 (2%)	53	59
1	B	292/293 (100%)	287 (98%)	5 (2%)	60	67
1	C	286/293 (98%)	283 (99%)	3 (1%)	76	82
1	D	286/293 (98%)	284 (99%)	2 (1%)	84	88
1	E	286/293 (98%)	280 (98%)	6 (2%)	53	59
1	F	284/293 (97%)	280 (99%)	4 (1%)	67	73
1	G	285/293 (97%)	281 (99%)	4 (1%)	67	73
1	H	287/293 (98%)	284 (99%)	3 (1%)	76	82
All	All	2298/2344 (98%)	2265 (99%)	33 (1%)	67	73

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	23	ASP
1	E	104	ASP
1	H	16	ARG
1	D	332	GLU
1	E	23	ASP

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](#)

8 non-standard protein/DNA/RNA residues 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
1	LLP	E	197	1	23,24,25	3.17	7 (30%)	25,32,34	1.28	4 (16%)
1	LLP	F	197	1	23,24,25	3.29	7 (30%)	25,32,34	1.34	4 (16%)
1	LLP	G	197	1	23,24,25	3.21	7 (30%)	25,32,34	1.32	6 (24%)
1	LLP	H	197	1	23,24,25	3.21	7 (30%)	25,32,34	1.49	6 (24%)
1	LLP	A	197	1	23,24,25	3.23	7 (30%)	25,32,34	1.38	5 (20%)
1	LLP	B	197	1	23,24,25	3.25	7 (30%)	25,32,34	1.31	5 (20%)
1	LLP	C	197	1	23,24,25	3.35	7 (30%)	25,32,34	1.38	3 (12%)
1	LLP	D	197	1	23,24,25	3.22	7 (30%)	25,32,34	1.31	5 (20%)

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
1	LLP	E	197	1	-	3/16/17/19	0/1/1/1
1	LLP	F	197	1	-	3/16/17/19	0/1/1/1
1	LLP	G	197	1	-	2/16/17/19	0/1/1/1
1	LLP	H	197	1	-	2/16/17/19	0/1/1/1
1	LLP	A	197	1	-	3/16/17/19	0/1/1/1
1	LLP	B	197	1	-	4/16/17/19	0/1/1/1
1	LLP	C	197	1	-	3/16/17/19	0/1/1/1
1	LLP	D	197	1	-	4/16/17/19	0/1/1/1

The worst 5 of 56 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	197	LLP	C4-C5	-6.82	1.33	1.42
1	C	197	LLP	C4-C3	6.67	1.50	1.40
1	D	197	LLP	C2-N1	6.49	1.46	1.33
1	F	197	LLP	C2-N1	6.47	1.46	1.33
1	F	197	LLP	C4-C5	-6.46	1.33	1.42

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	197	LLP	OP3-P-OP4	3.33	115.58	106.73
1	H	197	LLP	OP2-P-OP4	3.20	115.25	106.73
1	C	197	LLP	OP3-P-OP4	3.12	115.04	106.73
1	H	197	LLP	OP3-P-OP4	3.09	114.96	106.73
1	F	197	LLP	OP3-P-OP4	3.03	114.80	106.73

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	E	197	LLP	O-C-CA-CB
1	F	197	LLP	C-CA-CB-CG
1	F	197	LLP	O-C-CA-CB
1	G	197	LLP	O-C-CA-CB
1	H	197	LLP	O-C-CA-CB

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 53 ligands modelled in this entry, 12 are monoatomic - leaving 41 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	GOL	D	404	-	5,5,5	0.38	0	5,5,5	0.35	0
2	MES	B	401	-	12,12,12	2.13	1 (8%)	14,16,16	2.30	7 (50%)
4	GOL	G	406	-	5,5,5	0.34	0	5,5,5	0.26	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GOL	B	408	-	5,5,5	0.36	0	5,5,5	0.34	0
2	MES	D	401	-	12,12,12	2.14	1 (8%)	14,16,16	1.94	4 (28%)
4	GOL	B	404	-	5,5,5	0.35	0	5,5,5	0.43	0
2	MES	E	401	-	12,12,12	2.09	1 (8%)	14,16,16	1.98	3 (21%)
4	GOL	A	404	-	5,5,5	0.32	0	5,5,5	0.25	0
4	GOL	B	403	-	5,5,5	0.37	0	5,5,5	0.20	0
5	FMT	C	405	-	0,2,2	0.00	-	0,1,1	0.00	-
4	GOL	E	406	-	5,5,5	0.41	0	5,5,5	0.14	0
4	GOL	A	406	-	5,5,5	0.31	0	5,5,5	0.40	0
4	GOL	G	405	-	5,5,5	0.42	0	5,5,5	0.17	0
4	GOL	G	404	-	5,5,5	0.34	0	5,5,5	0.30	0
5	FMT	H	405	-	0,2,2	0.00	-	0,1,1	0.00	-
4	GOL	F	405	-	5,5,5	0.34	0	5,5,5	0.37	0
2	MES	G	401	-	12,12,12	2.20	1 (8%)	14,16,16	2.07	3 (21%)
4	GOL	E	404	-	5,5,5	0.39	0	5,5,5	0.23	0
4	GOL	G	407	-	5,5,5	0.32	0	5,5,5	0.36	0
4	GOL	A	407	-	5,5,5	0.44	0	5,5,5	0.49	0
2	MES	F	401	-	8,11,12	0.40	0	11,13,16	1.97	2 (18%)
4	GOL	C	404	-	5,5,5	0.32	0	5,5,5	0.40	0
5	FMT	G	408	-	0,2,2	0.00	-	0,1,1	0.00	-
4	GOL	H	404	-	5,5,5	0.37	0	5,5,5	0.33	0
4	GOL	A	405	-	5,5,5	0.32	0	5,5,5	0.22	0
4	GOL	B	405	-	5,5,5	0.40	0	5,5,5	0.27	0
5	FMT	F	406	-	0,2,2	0.00	-	0,1,1	0.00	-
4	GOL	D	407	-	5,5,5	0.42	0	5,5,5	0.45	0
4	GOL	E	403	-	5,5,5	0.38	0	5,5,5	0.29	0
4	GOL	B	407	-	5,5,5	0.37	0	5,5,5	0.35	0
4	GOL	E	405	-	5,5,5	0.42	0	5,5,5	0.31	0
5	FMT	A	408	-	0,2,2	0.00	-	0,1,1	0.00	-
4	GOL	D	405	-	5,5,5	0.36	0	5,5,5	0.34	0
4	GOL	D	406	-	5,5,5	0.35	0	5,5,5	0.31	0
2	MES	H	401	-	12,12,12	2.21	1 (8%)	14,16,16	1.89	3 (21%)
2	MES	A	401	-	12,12,12	1.99	1 (8%)	14,16,16	2.10	3 (21%)
4	GOL	B	406	-	5,5,5	0.34	0	5,5,5	0.73	0
2	MES	C	401	-	12,12,12	1.97	1 (8%)	14,16,16	2.03	4 (28%)
4	GOL	G	403	-	5,5,5	0.39	0	5,5,5	0.37	0
4	GOL	F	404	-	5,5,5	0.44	0	5,5,5	0.39	0
4	GOL	F	403	-	5,5,5	0.37	0	5,5,5	0.63	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
4	GOL	D	404	-	-	0/4/4/4	-
2	MES	B	401	-	-	1/6/14/14	0/1/1/1
4	GOL	G	406	-	-	2/4/4/4	-
4	GOL	B	408	-	-	2/4/4/4	-
4	GOL	H	404	-	-	2/4/4/4	-
4	GOL	B	404	-	-	2/4/4/4	-
2	MES	E	401	-	-	5/6/14/14	0/1/1/1
4	GOL	A	404	-	-	2/4/4/4	-
4	GOL	B	403	-	-	2/4/4/4	-
4	GOL	E	406	-	-	4/4/4/4	-
4	GOL	A	406	-	-	2/4/4/4	-
4	GOL	G	405	-	-	2/4/4/4	-
4	GOL	G	404	-	-	4/4/4/4	-
4	GOL	F	405	-	-	2/4/4/4	-
2	MES	G	401	-	-	1/6/14/14	0/1/1/1
4	GOL	E	404	-	-	4/4/4/4	-
4	GOL	G	407	-	-	2/4/4/4	-
4	GOL	A	407	-	-	4/4/4/4	-
2	MES	F	401	-	-	4/5/13/14	0/1/1/1
4	GOL	C	404	-	-	2/4/4/4	-
2	MES	D	401	-	-	1/6/14/14	0/1/1/1
4	GOL	A	405	-	-	2/4/4/4	-
4	GOL	B	405	-	-	4/4/4/4	-
4	GOL	D	407	-	-	0/4/4/4	-
4	GOL	E	403	-	-	2/4/4/4	-
4	GOL	B	407	-	-	2/4/4/4	-
4	GOL	E	405	-	-	3/4/4/4	-
4	GOL	D	405	-	-	2/4/4/4	-
4	GOL	D	406	-	-	2/4/4/4	-
2	MES	H	401	-	-	1/6/14/14	0/1/1/1
2	MES	A	401	-	-	4/6/14/14	0/1/1/1
4	GOL	B	406	-	-	2/4/4/4	-
2	MES	C	401	-	-	0/6/14/14	0/1/1/1
4	GOL	G	403	-	-	3/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	F	404	-	-	4/4/4/4	-
4	GOL	F	403	-	-	0/4/4/4	-

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	401	MES	C8-S	-7.37	1.67	1.77
2	G	401	MES	C8-S	-7.34	1.67	1.77
2	D	401	MES	C8-S	-7.00	1.67	1.77
2	B	401	MES	C8-S	-6.94	1.67	1.77
2	E	401	MES	C8-S	-6.82	1.67	1.77

The worst 5 of 29 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	401	MES	C5-N4-C3	5.38	120.93	108.83
2	G	401	MES	O1S-S-C8	5.25	113.23	106.92
2	F	401	MES	C5-N4-C3	5.22	120.57	108.83
2	B	401	MES	C5-N4-C3	4.73	119.47	108.83
2	A	401	MES	C5-N4-C3	4.60	119.17	108.83

There are no chirality outliers.

5 of 81 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	401	MES	C8-C7-N4-C3
4	B	404	GOL	C1-C2-C3-O3
2	E	401	MES	C8-C7-N4-C3
2	E	401	MES	C7-C8-S-O1S
2	E	401	MES	C7-C8-S-O3S

There are no ring outliers.

12 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	401	MES	1	0
2	D	401	MES	1	0
4	B	403	GOL	1	0
4	G	405	GOL	1	0
2	G	401	MES	2	0
4	A	407	GOL	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	407	GOL	4	0
4	E	405	GOL	1	0
4	D	406	GOL	1	0
2	H	401	MES	1	0
2	C	401	MES	1	0
4	F	404	GOL	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	372/384 (96%)	-0.41	6 (1%) 72 75	13, 24, 57, 90	0
1	B	374/384 (97%)	-0.42	6 (1%) 72 75	13, 24, 56, 72	0
1	C	375/384 (97%)	-0.36	14 (3%) 41 48	13, 24, 58, 81	0
1	D	372/384 (96%)	-0.36	15 (4%) 38 44	13, 23, 62, 82	0
1	E	372/384 (96%)	-0.42	14 (3%) 40 46	14, 23, 56, 88	0
1	F	373/384 (97%)	-0.32	7 (1%) 66 71	13, 24, 57, 74	0
1	G	374/384 (97%)	-0.40	7 (1%) 66 71	13, 26, 57, 71	0
1	H	374/384 (97%)	-0.41	8 (2%) 63 68	13, 24, 57, 76	0
All	All	2986/3072 (97%)	-0.39	77 (2%) 56 61	13, 24, 58, 90	0

The worst 5 of 77 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	3	GLY	4.5
1	H	341	PRO	4.4
1	F	344	ALA	4.4
1	G	3	GLY	4.2
1	D	40	ASP	4.1

### 6.2 Non-standard residues in protein, DNA, RNA chains [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. 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
1	LLP	E	197	24/25	0.98	0.14	13,16,20,21	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
1	LLP	F	197	24/25	0.98	0.15	15,18,25,29	0
1	LLP	B	197	24/25	0.98	0.13	10,17,22,23	0
1	LLP	C	197	24/25	0.98	0.16	13,18,22,25	0
1	LLP	A	197	24/25	0.99	0.13	13,18,21,26	0
1	LLP	G	197	24/25	0.99	0.16	13,19,22,23	0
1	LLP	H	197	24/25	0.99	0.13	13,18,22,23	0
1	LLP	D	197	24/25	0.99	0.12	12,19,22,22	0

### 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. 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( $\text{\AA}^2$ )	Q<0.9
4	GOL	G	404	6/6	0.61	0.25	48,56,63,64	0
5	FMT	A	408	3/3	0.75	0.13	57,57,60,61	0
4	GOL	B	407	6/6	0.77	0.20	36,58,61,63	0
4	GOL	G	407	6/6	0.81	0.25	61,64,65,66	0
4	GOL	D	405	6/6	0.82	0.18	55,58,65,66	0
4	GOL	B	408	6/6	0.83	0.20	71,74,75,76	0
4	GOL	B	403	6/6	0.84	0.23	55,59,60,62	0
4	GOL	F	404	6/6	0.84	0.21	41,50,53,54	0
4	GOL	D	407	6/6	0.85	0.20	49,55,57,57	0
4	GOL	A	407	6/6	0.85	0.19	39,51,53,54	0
4	GOL	A	405	6/6	0.86	0.23	33,53,57,58	0
5	FMT	F	406	3/3	0.86	0.15	29,29,39,44	0
4	GOL	B	405	6/6	0.87	0.26	46,51,56,57	0
4	GOL	E	404	6/6	0.87	0.17	48,55,57,57	0
4	GOL	G	406	6/6	0.88	0.17	65,67,67,69	0
4	GOL	E	406	6/6	0.88	0.19	44,50,54,60	0
4	GOL	G	405	6/6	0.88	0.14	39,47,50,50	0
4	GOL	F	405	6/6	0.89	0.19	44,50,51,52	0
5	FMT	G	408	3/3	0.89	0.07	43,43,46,46	0
4	GOL	F	403	6/6	0.89	0.23	42,56,57,61	0
4	GOL	A	406	6/6	0.90	0.15	45,47,52,52	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	A	402	1/1	0.90	0.09	66,66,66,66	0
3	CL	D	402	1/1	0.90	0.10	61,61,61,61	0
4	GOL	B	406	6/6	0.91	0.33	37,59,63,64	0
5	FMT	H	405	3/3	0.91	0.09	47,47,47,48	0
4	GOL	E	405	6/6	0.91	0.16	54,58,60,63	0
4	GOL	D	404	6/6	0.92	0.19	41,47,48,50	0
3	CL	A	403	1/1	0.92	0.11	53,53,53,53	0
5	FMT	C	405	3/3	0.92	0.07	51,51,52,53	0
4	GOL	A	404	6/6	0.92	0.17	43,51,53,57	0
3	CL	H	403	1/1	0.93	0.08	52,52,52,52	0
4	GOL	E	403	6/6	0.93	0.22	42,43,45,51	0
3	CL	G	402	1/1	0.93	0.06	66,66,66,66	0
4	GOL	D	406	6/6	0.93	0.14	46,56,61,64	0
3	CL	C	402	1/1	0.94	0.07	64,64,64,64	0
4	GOL	G	403	6/6	0.94	0.15	39,47,50,56	0
4	GOL	B	404	6/6	0.94	0.20	43,45,46,48	0
3	CL	C	403	1/1	0.94	0.15	66,66,66,66	0
4	GOL	C	404	6/6	0.95	0.16	39,46,49,51	0
3	CL	E	402	1/1	0.95	0.06	56,56,56,56	0
2	MES	C	401	12/12	0.95	0.10	31,37,41,41	0
4	GOL	H	404	6/6	0.95	0.12	40,49,53,54	0
3	CL	D	403	1/1	0.95	0.05	64,64,64,64	0
3	CL	H	402	1/1	0.95	0.07	64,64,64,64	0
3	CL	F	402	1/1	0.95	0.05	60,60,60,60	0
2	MES	F	401	11/12	0.96	0.11	28,44,53,54	0
3	CL	B	402	1/1	0.96	0.05	67,67,67,67	0
2	MES	B	401	12/12	0.97	0.08	29,32,36,39	0
2	MES	D	401	12/12	0.97	0.09	31,38,45,45	0
2	MES	G	401	12/12	0.97	0.09	32,38,39,41	0
2	MES	A	401	12/12	0.97	0.09	25,28,32,35	0
2	MES	E	401	12/12	0.97	0.09	31,39,50,52	0
2	MES	H	401	12/12	0.99	0.08	26,29,38,43	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.