



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 10:53 am BST

PDB ID : 6B4G  
Title : Crystal structure of Chaetomium thermophilum Gle1 CTD-Nup42 GBM complex  
Authors : Lin, D.H.; Correia, A.R.; Cai, S.W.; Huber, F.M.; Jette, C.A.; Hoelz, A.  
Deposited on : 2017-09-26  
Resolution : 2.65 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

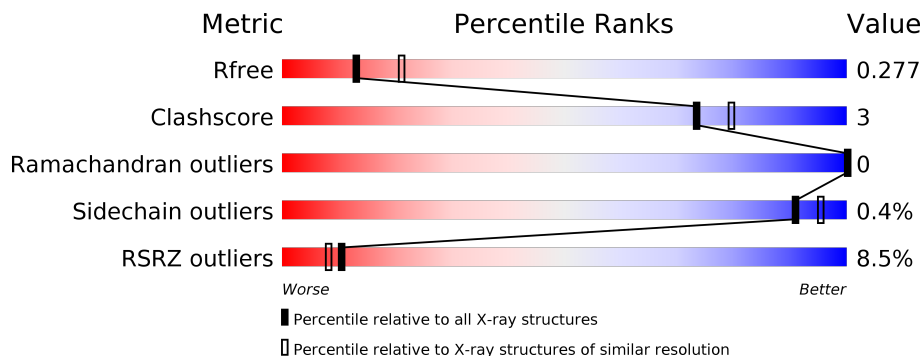
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.65 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	1426 (2.66-2.62)
Clashscore	141614	1472 (2.66-2.62)
Ramachandran outliers	138981	1446 (2.66-2.62)
Sidechain outliers	138945	1446 (2.66-2.62)
RSRZ outliers	127900	1408 (2.66-2.62)

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	318	
1	C	318	
1	E	318	
1	G	318	
2	B	73	
2	D	73	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	F	73	
2	H	73	

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 23147 atoms, of which 11586 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 Nucleoporin GLE1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
1	A	315	5013	1579	2537	442	442	13	0	1	0
1	C	310	4950	1560	2505	437	435	13	0	0	0
1	E	314	4996	1573	2528	441	440	14	0	0	0
1	G	308	4924	1551	2493	435	432	13	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	212	GLY	-	expression tag	UNP G0S7F3
A	213	PRO	-	expression tag	UNP G0S7F3
A	214	HIS	-	expression tag	UNP G0S7F3
A	215	MET	-	expression tag	UNP G0S7F3
C	212	GLY	-	expression tag	UNP G0S7F3
C	213	PRO	-	expression tag	UNP G0S7F3
C	214	HIS	-	expression tag	UNP G0S7F3
C	215	MET	-	expression tag	UNP G0S7F3
E	212	GLY	-	expression tag	UNP G0S7F3
E	213	PRO	-	expression tag	UNP G0S7F3
E	214	HIS	-	expression tag	UNP G0S7F3
E	215	MET	-	expression tag	UNP G0S7F3
G	212	GLY	-	expression tag	UNP G0S7F3
G	213	PRO	-	expression tag	UNP G0S7F3
G	214	HIS	-	expression tag	UNP G0S7F3
G	215	MET	-	expression tag	UNP G0S7F3

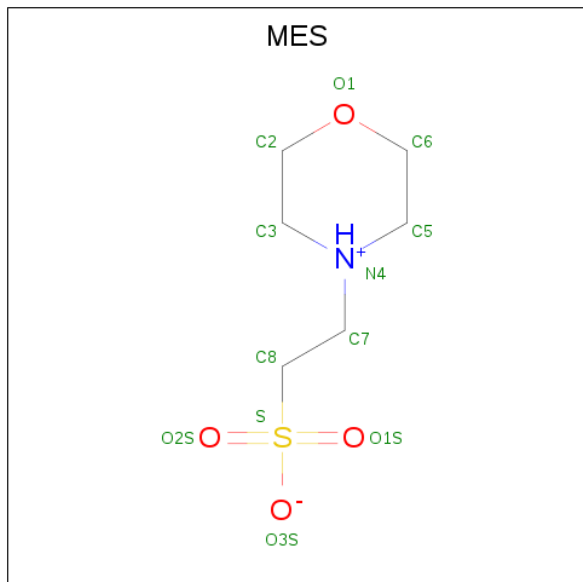
- Molecule 2 is a protein called Nucleoporin AMO1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
2	B	47	Total	C	H	N	O	S	0	0	0
			750	247	357	67	77	2			
2	D	50	Total	C	H	N	O	S	0	0	0
			796	264	379	70	81	2			
2	F	50	Total	C	H	N	O	S	0	0	0
			796	264	379	70	81	2			
2	H	45	Total	C	H	N	O	S	0	0	0
			714	232	343	64	73	2			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	485	GLY	-	expression tag	UNP G0S381
B	486	PRO	-	expression tag	UNP G0S381
B	487	HIS	-	expression tag	UNP G0S381
B	488	MET	-	expression tag	UNP G0S381
B	489	GLY	-	expression tag	UNP G0S381
B	490	SER	-	expression tag	UNP G0S381
B	491	PRO	-	expression tag	UNP G0S381
B	492	GLU	-	expression tag	UNP G0S381
D	485	GLY	-	expression tag	UNP G0S381
D	486	PRO	-	expression tag	UNP G0S381
D	487	HIS	-	expression tag	UNP G0S381
D	488	MET	-	expression tag	UNP G0S381
D	489	GLY	-	expression tag	UNP G0S381
D	490	SER	-	expression tag	UNP G0S381
D	491	PRO	-	expression tag	UNP G0S381
D	492	GLU	-	expression tag	UNP G0S381
F	485	GLY	-	expression tag	UNP G0S381
F	486	PRO	-	expression tag	UNP G0S381
F	487	HIS	-	expression tag	UNP G0S381
F	488	MET	-	expression tag	UNP G0S381
F	489	GLY	-	expression tag	UNP G0S381
F	490	SER	-	expression tag	UNP G0S381
F	491	PRO	-	expression tag	UNP G0S381
F	492	GLU	-	expression tag	UNP G0S381
H	485	GLY	-	expression tag	UNP G0S381
H	486	PRO	-	expression tag	UNP G0S381
H	487	HIS	-	expression tag	UNP G0S381
H	488	MET	-	expression tag	UNP G0S381
H	489	GLY	-	expression tag	UNP G0S381
H	490	SER	-	expression tag	UNP G0S381
H	491	PRO	-	expression tag	UNP G0S381
H	492	GLU	-	expression tag	UNP G0S381

- Molecule 3 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	H	N	O			S
3	A	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
3	A	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
3	C	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
3	E	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		
3	G	1	Total	C	H	N	O	S	0	0
			25	6	13	1	4	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	36	Total	O	0	0
			36	36		
4	B	9	Total	O	0	0
			9	9		
4	D	1	Total	O	0	0
			1	1		
4	H	1	Total	O	0	0
			1	1		
4	C	27	Total	O	0	0
			27	27		

*Continued on next page...*

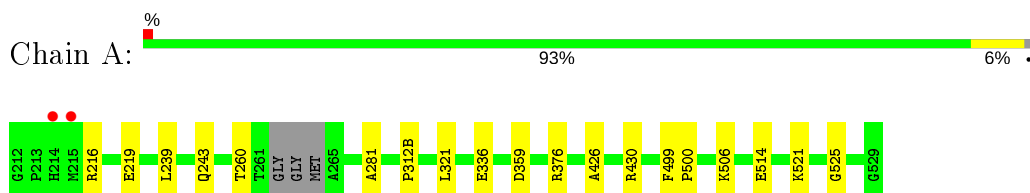
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
4	G	9	Total O 9 9	0	0

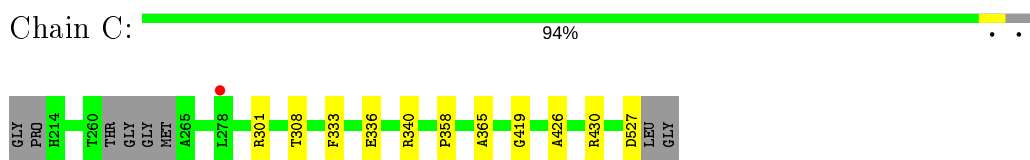
### 3 Residue-property plots [i](#)

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.

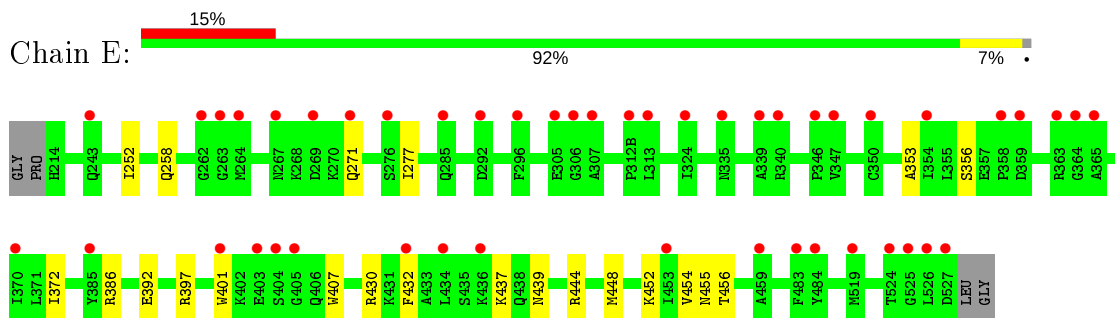
- Molecule 1: Nucleoporin GLE1



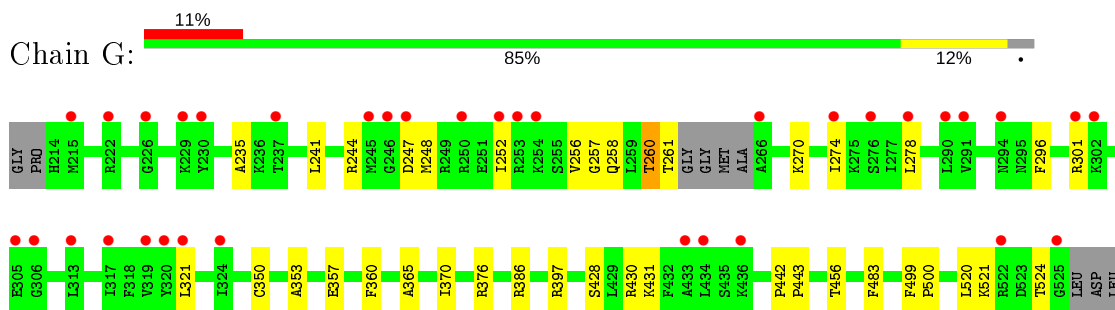
- Molecule 1: Nucleoporin GLE1



- Molecule 1: Nucleoporin GLE1



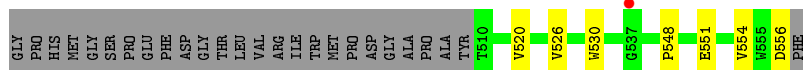
- Molecule 1: Nucleoporin GLE1



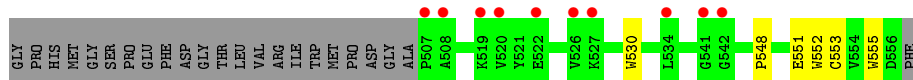


GLY

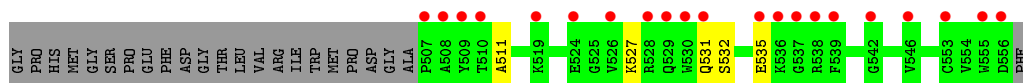
- Molecule 2: Nucleoporin AMO1



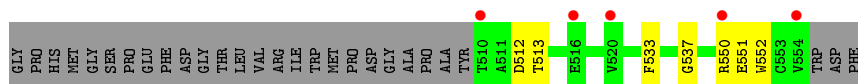
- Molecule 2: Nucleoporin AMO1



- Molecule 2: Nucleoporin AMO1



- Molecule 2: Nucleoporin AMO1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	84.48Å 93.06Å 229.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.86 – 2.65 48.86 – 2.65	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.86-2.65) 99.8 (48.86-2.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.09	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.85 (at 2.65Å)	Xtrriage
Refinement program	PHENIX (dev_2006: ???)	Depositor
R, $R_{free}$	0.240 , 0.277 0.241 , 0.277	Depositor DCC
$R_{free}$ test set	2000 reflections (3.74%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	66.1	Xtrriage
Anisotropy	0.310	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 51.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	23147	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	97.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.16% 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: MES

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	0.27	0/2534	0.46	0/3420
1	C	0.25	0/2494	0.43	0/3366
1	E	0.26	0/2518	0.46	0/3399
1	G	0.30	0/2480	0.47	0/3347
2	B	0.26	0/405	0.45	0/548
2	D	0.28	0/431	0.50	0/584
2	F	0.26	0/431	0.50	0/584
2	H	0.30	0/381	0.48	0/514
All	All	0.27	0/11674	0.46	0/15762

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	2476	2537	2532	11	0
1	C	2445	2505	2505	5	0
1	E	2468	2528	2528	13	0
1	G	2431	2493	2492	29	0
2	B	393	357	357	3	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	417	379	379	3	0
2	F	417	379	379	3	0
2	H	371	343	343	15	0
3	A	24	26	26	0	0
3	C	12	13	13	0	0
3	E	12	13	13	0	0
3	G	12	13	13	3	0
4	A	36	0	0	1	0
4	B	9	0	0	0	0
4	C	27	0	0	0	0
4	D	1	0	0	0	0
4	G	9	0	0	0	0
4	H	1	0	0	0	0
All	All	11561	11586	11580	75	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 3.

The worst 5 of 75 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:A:359:ASP:OD1	1:E:258:GLN:NE2	2.08	0.86
1:G:257:GLY:O	1:G:260:THR:OG1	1.95	0.83
2:H:513:THR:HG21	2:H:550:ARG:NH2	1.95	0.82
2:H:513:THR:CB	2:H:550:ARG:NH2	2.50	0.75
2:H:513:THR:OG1	2:H:550:ARG:NH2	2.20	0.73

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	312/318 (98%)	304 (97%)	8 (3%)	0	100	100
1	C	306/318 (96%)	302 (99%)	4 (1%)	0	100	100
1	E	312/318 (98%)	298 (96%)	14 (4%)	0	100	100
1	G	304/318 (96%)	297 (98%)	7 (2%)	0	100	100
2	B	45/73 (62%)	44 (98%)	1 (2%)	0	100	100
2	D	48/73 (66%)	47 (98%)	1 (2%)	0	100	100
2	F	48/73 (66%)	48 (100%)	0	0	100	100
2	H	43/73 (59%)	43 (100%)	0	0	100	100
All	All	1418/1564 (91%)	1383 (98%)	35 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	261/261 (100%)	261 (100%)	0	100	100
1	C	257/261 (98%)	256 (100%)	1 (0%)	91	95
1	E	259/261 (99%)	258 (100%)	1 (0%)	91	95
1	G	256/261 (98%)	254 (99%)	2 (1%)	81	89
2	B	41/61 (67%)	40 (98%)	1 (2%)	49	67
2	D	43/61 (70%)	43 (100%)	0	100	100
2	F	43/61 (70%)	43 (100%)	0	100	100
2	H	39/61 (64%)	39 (100%)	0	100	100
All	All	1199/1288 (93%)	1194 (100%)	5 (0%)	91	95

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	556	ASP
1	C	527	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	401	TRP
1	G	260	THR
1	G	261	THR

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](#)

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	MES	C	601	-	12,12,12	0.66	0	14,16,16	0.39	0
3	MES	E	601	-	12,12,12	0.81	0	14,16,16	0.35	0
3	MES	G	601	-	12,12,12	0.65	0	14,16,16	0.40	0
3	MES	A	602	-	12,12,12	0.69	0	14,16,16	0.39	0
3	MES	A	601	-	12,12,12	0.74	0	14,16,16	0.56	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
3	MES	C	601	-	-	0/6/14/14	0/1/1/1
3	MES	E	601	-	-	4/6/14/14	0/1/1/1
3	MES	G	601	-	-	1/6/14/14	0/1/1/1
3	MES	A	602	-	-	3/6/14/14	0/1/1/1
3	MES	A	601	-	-	4/6/14/14	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	601	MES	N4-C7-C8-S
3	A	602	MES	C7-C8-S-O1S
3	A	601	MES	N4-C7-C8-S
3	A	601	MES	C7-C8-S-O2S
3	A	602	MES	C7-C8-S-O3S

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	G	601	MES	3	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	315/318 (99%)	0.15	2 (0%) 89 88	35, 59, 92, 129	0
1	C	310/318 (97%)	0.07	1 (0%) 94 95	40, 59, 94, 127	0
1	E	314/318 (98%)	0.99	47 (14%) 2 1	92, 121, 156, 175	0
1	G	308/318 (96%)	0.71	35 (11%) 5 3	50, 94, 141, 154	0
2	B	47/73 (64%)	0.04	1 (2%) 63 60	39, 56, 74, 96	0
2	D	50/73 (68%)	0.99	10 (20%) 1 1	63, 98, 126, 134	0
2	F	50/73 (68%)	2.08	21 (42%) 0 0	122, 153, 170, 182	0
2	H	45/73 (61%)	0.72	5 (11%) 5 3	62, 105, 123, 138	0
All	All	1439/1564 (92%)	0.55	122 (8%) 10 8	35, 81, 145, 182	0

The worst 5 of 122 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	364	GLY	8.4
2	F	553	CYS	8.3
2	F	537	GLY	8.1
1	E	263	GLY	7.4
2	F	530	TRP	6.2

### 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. 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
3	MES	G	601	12/12	0.74	0.26	99,119,136,186	0
3	MES	E	601	12/12	0.84	0.20	114,130,143,143	0
3	MES	A	602	12/12	0.93	0.18	101,114,131,134	0
3	MES	A	601	12/12	0.96	0.21	55,66,78,85	0
3	MES	C	601	12/12	0.98	0.15	68,78,93,94	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.