



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:15 PM GMT

PDB ID : 4BZ4  
Title : CorA is a surface-associated copper-binding protein important in *Methylomicrobium album* BG8 copper acquisition  
Authors : Johnson, K.A.; Ve, T.; Pedersen, R.B.; Lillehaug, J.R.; Jensen, H.B.; Helland, R.; Karlsen, O.A.  
Deposited on : 2013-07-24  
Resolution : 1.60 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  
<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

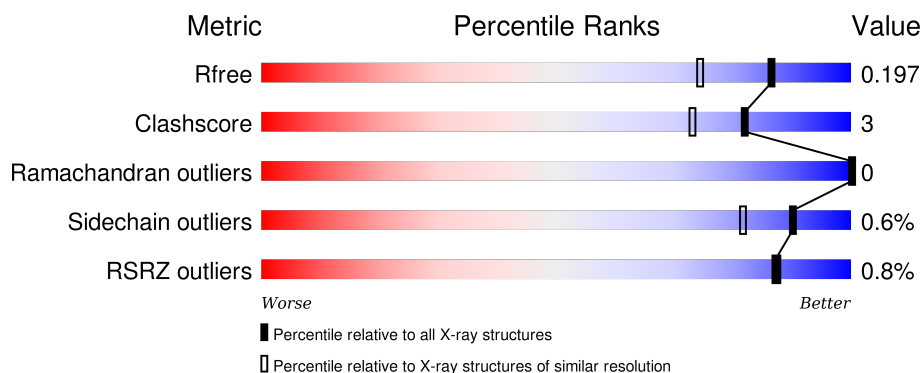
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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	2475 (1.60-1.60)
Clashscore	102246	2732 (1.60-1.60)
Ramachandran outliers	100387	2654 (1.60-1.60)
Sidechain outliers	100360	2653 (1.60-1.60)
RSRZ outliers	91569	2479 (1.60-1.60)

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	231	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">%</span> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>78%</span> <span>8%</span> <span>14%</span> </div> </div>
1	B	231	<div> <div style="width: 100%; height: 10px; background-color: green; position: relative;"> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>79%</span> <span>9%</span> <span>12%</span> </div> </div>
1	C	231	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">%</span> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>80%</span> <span>6%</span> <span>14%</span> </div> </div>
1	D	231	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">%</span> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>78%</span> <span>7%</span> <span>14%</span> </div> </div>
1	E	231	<div> <div style="width: 100%; height: 10px; background-color: green; position: relative;"> </div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> <span>80%</span> <span>7%</span> <span>12%</span> </div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	231	

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	PG4	A	1234	-	-	-	X
4	PG4	B	1234	-	-	-	X
4	PG4	C	1234	-	-	-	X
4	PG4	D	1234	-	-	-	X
4	PG4	E	1234	-	-	-	X
4	PG4	F	1234	-	-	-	X
5	P6G	A	1235	-	-	-	X
5	P6G	A	1236	-	-	-	X
5	P6G	B	1235	-	-	-	X
5	P6G	B	1236	-	-	-	X
5	P6G	C	1235	-	-	-	X
5	P6G	D	1235	-	-	-	X
5	P6G	E	1235	-	-	-	X
5	P6G	E	1236	-	-	-	X
5	P6G	E	1237	-	-	-	X
5	P6G	F	1235	-	-	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10634 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 COPPER-REPRESSIBLE POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	199	Total	C	N	O	S	0	4	0
			1539	986	255	293	5			
1	B	204	Total	C	N	O	S	0	3	0
			1573	1003	263	302	5			
1	C	199	Total	C	N	O	S	0	2	0
			1529	979	255	289	6			
1	D	198	Total	C	N	O	S	0	2	0
			1521	973	253	290	5			
1	E	204	Total	C	N	O	S	0	6	0
			1582	1008	263	306	5			
1	F	199	Total	C	N	O	S	0	6	0
			1542	987	256	294	5			

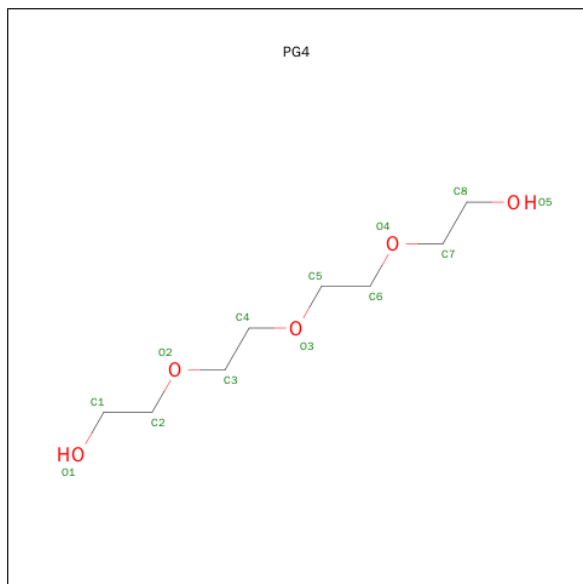
- Molecule 2 is COPPER (I) ION (three-letter code: CU1) (formula: Cu).

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

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

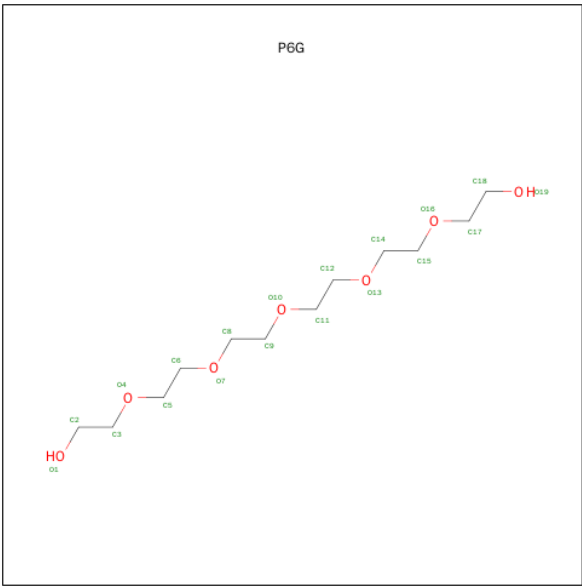
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	1	Total Ca 1 1	0	0
3	E	1	Total Ca 1 1	0	0
3	B	1	Total Ca 1 1	0	0
3	C	1	Total Ca 1 1	0	0
3	A	1	Total Ca 1 1	0	0
3	F	1	Total Ca 1 1	0	0

- Molecule 4 is TETRAETHYLENE GLYCOL (three-letter code: PG4) (formula:  $C_8H_{18}O_5$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	B	1	Total C O 6 4 2	0	0
4	C	1	Total C O 6 4 2	0	0
4	D	1	Total C O 7 4 3	0	0
4	E	1	Total C O 6 4 2	0	0
4	F	1	Total C O 8 5 3	0	0

- Molecule 5 is HEXAETHYLENE GLYCOL (three-letter code: P6G) (formula: C<sub>12</sub>H<sub>26</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			10	6	4		
5	A	1	Total	C	O	0	0
			7	4	3		
5	B	1	Total	C	O	0	0
			11	7	4		
5	B	1	Total	C	O	0	0
			8	5	3		
5	C	1	Total	C	O	0	0
			19	12	7		
5	C	1	Total	C	O	0	0
			15	10	5		
5	D	1	Total	C	O	0	0
			8	5	3		
5	E	1	Total	C	O	0	0
			14	9	5		
5	E	1	Total	C	O	0	0
			13	8	5		
5	E	1	Total	C	O	0	0
			8	5	3		
5	F	1	Total	C	O	0	0
			11	7	4		

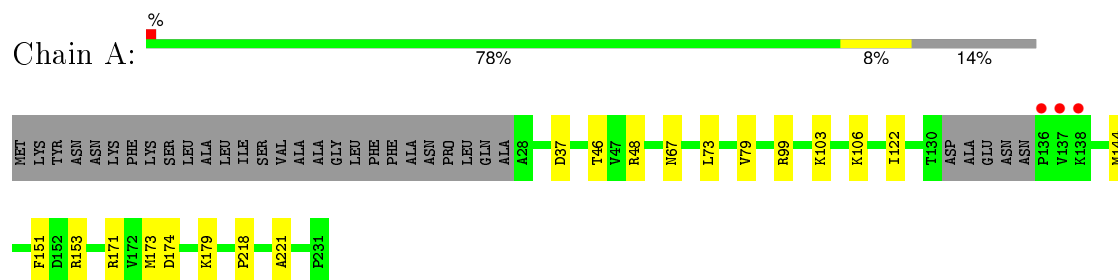
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	206	Total 206	O 206	0	0
6	B	179	Total 179	O 179	0	0
6	C	188	Total 188	O 188	0	0
6	D	200	Total 200	O 200	0	0
6	E	222	Total 222	O 222	0	0
6	F	180	Total 180	O 180	0	0

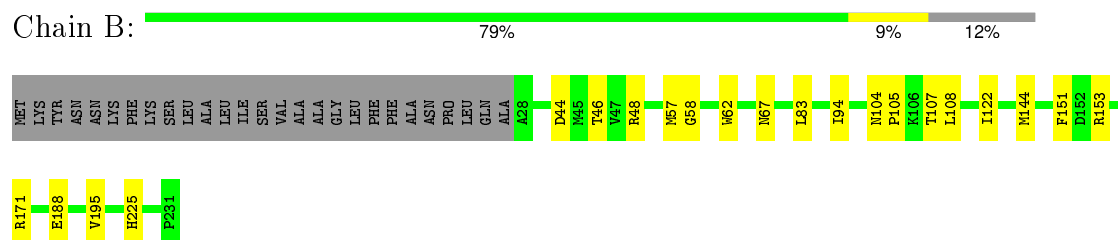
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

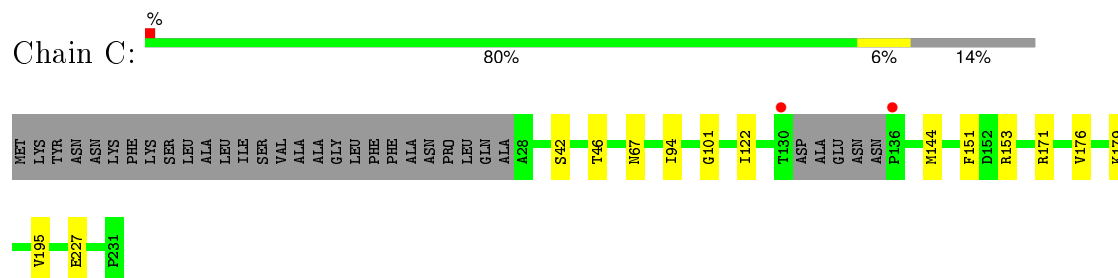
- Molecule 1: COPPER-REPRESSIBLE POLYPEPTIDE



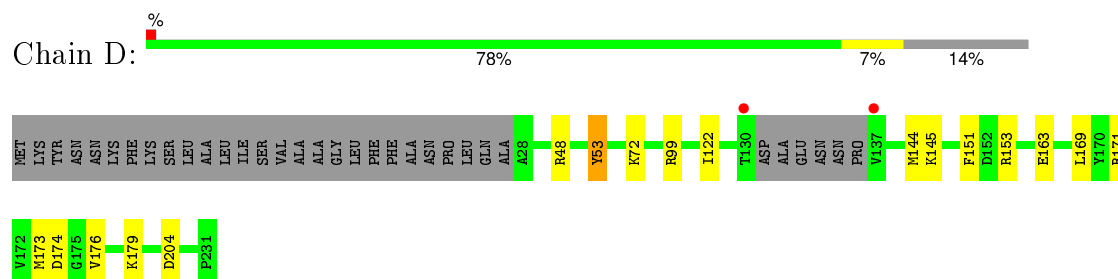
- Molecule 1: COPPER-REPRESSIBLE POLYPEPTIDE



- Molecule 1: COPPER-REPRESSIBLE POLYPEPTIDE

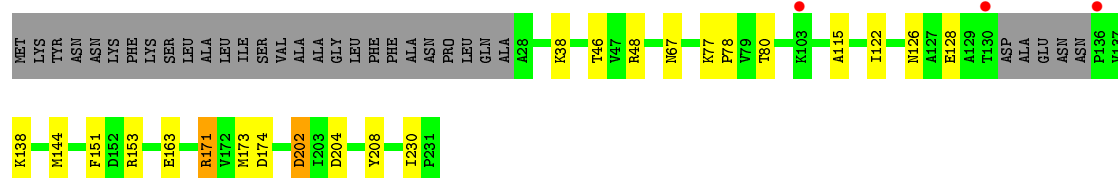
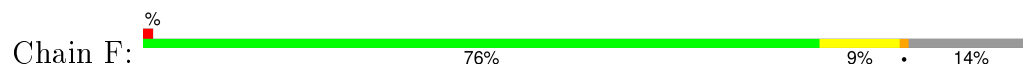


- Molecule 1: COPPER-REPRESSIBLE POLYPEPTIDE





Chain E:  80% 7% • 12%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.84Å 113.17Å 81.49Å 90.00° 104.51° 90.00°	Depositor
Resolution (Å)	30.00 – 1.60 29.87 – 1.60	Depositor EDS
% Data completeness (in resolution range)	98.7 (30.00-1.60) 98.6 (29.87-1.60)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.35 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.158 , 0.188 0.169 , 0.197	Depositor DCC
$R_{free}$ test set	8426 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.8	Xtriage
Anisotropy	0.149	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 50.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 168264 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	10634	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	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 46.47 % 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 1.1404e-04. 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PG4, CA, P6G, KYN, CU1

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	0.97	0/1585	1.03	5/2157 (0.2%)
1	B	1.03	0/1616	1.04	4/2201 (0.2%)
1	C	1.03	0/1567	0.98	0/2132
1	D	1.02	1/1558 (0.1%)	1.06	6/2120 (0.3%)
1	E	1.04	2/1637 (0.1%)	1.09	7/2230 (0.3%)
1	F	1.04	0/1596	1.07	8/2171 (0.4%)
All	All	1.02	3/9559 (0.0%)	1.05	30/13011 (0.2%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	163	GLU	CD-OE1	8.02	1.34	1.25
1	D	163	GLU	CD-OE2	5.69	1.31	1.25
1	E	163	GLU	CD-OE2	-5.52	1.19	1.25

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	173	MET	CG-SD-CE	-9.86	84.43	100.20
1	A	173	MET	CG-SD-CE	-8.78	86.15	100.20
1	B	48	ARG	NE-CZ-NH2	-7.87	116.36	120.30
1	E	153	ARG	NE-CZ-NH2	-7.27	116.66	120.30
1	A	174	ASP	CB-CG-OD2	-7.12	111.89	118.30
1	B	48	ARG	CB-CG-CD	6.99	129.77	111.60
1	E	131	ASP	CB-CG-OD2	-6.86	112.12	118.30
1	D	48	ARG	NE-CZ-NH2	-6.75	116.93	120.30
1	F	202	ASP	CB-CG-OD1	6.60	124.24	118.30
1	B	48	ARG	NE-CZ-NH1	6.45	123.52	120.30
1	A	37	ASP	CB-CG-OD2	-6.40	112.54	118.30
1	E	163	GLU	CG-CD-OE2	-6.10	106.09	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	173	MET	CG-SD-CE	-5.94	90.70	100.20
1	A	37	ASP	CB-CG-OD1	5.78	123.50	118.30
1	E	202	ASP	CB-CG-OD2	-5.76	113.12	118.30
1	E	171	ARG	NE-CZ-NH2	-5.68	117.46	120.30
1	D	204	ASP	CB-CG-OD2	-5.66	113.20	118.30
1	D	174	ASP	CB-CG-OD2	-5.62	113.24	118.30
1	D	99	ARG	NE-CZ-NH2	-5.62	117.49	120.30
1	F	202	ASP	CB-CG-OD2	-5.60	113.26	118.30
1	E	53	TYR	CB-CG-CD2	-5.56	117.67	121.00
1	F	171	ARG	NE-CZ-NH1	5.34	122.97	120.30
1	F	174	ASP	CB-CG-OD2	-5.29	113.54	118.30
1	D	53	TYR	CB-CG-CD1	-5.19	117.89	121.00
1	E	158	ASP	CB-CG-OD1	5.19	122.97	118.30
1	F	204	ASP	CB-CG-OD1	5.12	122.91	118.30
1	B	48	ARG	CD-NE-CZ	5.08	130.71	123.60
1	F	48	ARG	NE-CZ-NH2	-5.05	117.78	120.30
1	A	99	ARG	NE-CZ-NH1	5.05	122.82	120.30
1	F	48	ARG	CG-CD-NE	-5.03	101.23	111.80

There are no chirality outliers.

There are no planarity outliers.

## 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	1539	0	1468	11	0
1	B	1573	0	1486	13	0
1	C	1529	0	1462	7	0
1	D	1521	0	1450	6	0
1	E	1582	0	1499	9	0
1	F	1542	0	1473	11	1
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

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	1	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	4	0	5	0	0
4	B	6	0	6	1	0
4	C	6	0	6	0	0
4	D	7	0	9	0	0
4	E	6	0	6	0	0
4	F	8	0	8	2	0
5	A	17	0	22	0	0
5	B	19	0	22	2	0
5	C	34	0	44	0	0
5	D	8	0	9	0	0
5	E	35	0	42	0	0
5	F	11	0	12	0	0
6	A	206	0	0	4	0
6	B	179	0	0	4	0
6	C	188	0	0	1	0
6	D	200	0	0	2	1
6	E	222	0	0	2	0
6	F	180	0	0	4	0
All	All	10634	0	9029	59	1

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.

All (59) 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:B:188[B]:GLU:OE2	6:B:2152:HOH:O	1.57	1.20
1:B:57:MET:C	1:B:58:GLY:CA	2.30	1.00
1:B:67[A]:ASN:OD1	6:B:2009:HOH:O	1.80	0.97
5:B:1235:P6G:H31	6:B:2176:HOH:O	1.82	0.79
1:F:80[B]:THR:HG23	6:F:2068:HOH:O	1.86	0.75
1:C:101:GLY:O	6:C:2090:HOH:O	2.04	0.73
1:C:46:THR:H	1:C:67[B]:ASN:ND2	1.92	0.67
1:B:44:ASP:OD1	1:B:225[B]:HIS:ND1	2.30	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:163:GLU:OE2	6:F:2136:HOH:O	2.17	0.59
1:B:105:PRO:HA	1:B:108:LEU:HD12	1.85	0.58
1:B:104:ASN:O	1:B:107:THR:HG22	2.04	0.57
1:D:176:VAL:HG11	1:D:179:LYS:HD2	1.87	0.57
1:E:128:GLU:CD	1:E:138:LYS:NZ	2.58	0.56
1:F:115:ALA:HB2	4:F:1234:PG4:H41	1.88	0.55
1:A:67[A]:ASN:OD1	6:A:2008:HOH:O	2.18	0.55
1:D:145:LYS:HE3	6:D:2123:HOH:O	2.06	0.54
4:F:1234:PG4:H42	6:F:2176:HOH:O	2.10	0.52
1:F:38[B]:LYS:NZ	1:F:230:ILE:O	2.43	0.51
1:B:151:PHE:CZ	1:B:171:ARG:HG2	2.46	0.50
5:B:1235:P6G:H32	6:B:2178:HOH:O	2.12	0.50
1:C:42:SER:HB3	1:C:227:GLU:OE1	2.12	0.49
1:E:128:GLU:CD	1:E:138:LYS:HZ3	2.16	0.49
1:B:46:THR:H	1:B:67[B]:ASN:ND2	2.10	0.49
1:E:122:ILE:HB	1:E:144:MET:HB2	1.96	0.48
1:E:227[A]:GLU:OE2	6:E:2085:HOH:O	2.20	0.48
1:B:83:LEU:HD12	1:B:83:LEU:N	2.28	0.48
1:C:151:PHE:CZ	1:C:171:ARG:HG2	2.49	0.48
1:E:151:PHE:CZ	1:E:171:ARG:HG2	2.49	0.47
1:A:122:ILE:HB	1:A:144:MET:HB2	1.96	0.47
1:F:202:ASP:O	1:F:208:TYR:HB2	2.14	0.47
1:D:151:PHE:CZ	1:D:171:ARG:HG2	2.49	0.47
1:B:62:KYN:CE2	4:B:1234:PG4:H31	2.45	0.46
1:A:151:PHE:CZ	1:A:171:ARG:HG2	2.51	0.46
1:F:38[B]:LYS:HD2	6:F:2015:HOH:O	2.15	0.46
1:C:94:ILE:HA	1:C:195:VAL:O	2.16	0.46
1:F:122:ILE:HB	1:F:144:MET:HB2	1.98	0.46
1:D:72:LYS:NZ	6:D:2063:HOH:O	2.49	0.45
1:A:48:ARG:HD3	1:A:218:PRO:O	2.17	0.45
1:C:176:VAL:HG11	1:C:179:LYS:HE3	1.98	0.45
1:F:151:PHE:CZ	1:F:171:ARG:HG2	2.52	0.45
1:F:46:THR:H	1:F:67:ASN:ND2	2.16	0.44
1:D:122:ILE:HB	1:D:144:MET:HB2	2.00	0.44
1:F:128:GLU:HG2	1:F:138:LYS:HD2	1.99	0.44
1:A:103:LYS:HE2	1:A:103:LYS:HA	2.00	0.43
1:F:77:LYS:HA	1:F:78:PRO:HD3	1.91	0.43
1:B:94:ILE:HA	1:B:195:VAL:O	2.19	0.43
1:E:105:PRO:HA	1:E:108:LEU:HD12	2.01	0.42
1:B:122:ILE:HB	1:B:144:MET:HB2	2.01	0.42
1:C:122:ILE:HB	1:C:144:MET:HB2	2.00	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:53:TYR:CE1	1:D:169:LEU:HD11	2.56	0.41
1:B:107:THR:HG21	1:E:204:ASP:OD2	2.20	0.41
1:E:46:THR:H	1:E:67[B]:ASN:ND2	2.18	0.41
1:A:46:THR:H	1:A:67[A]:ASN:ND2	2.18	0.41
1:E:174[A]:ASP:OD2	6:E:2176:HOH:O	2.20	0.41
1:A:46:THR:HB	1:A:221:ALA:HB1	2.02	0.40
1:A:103:LYS:HB2	6:A:2100:HOH:O	2.20	0.40
1:A:73:LEU:HD13	1:A:79:VAL:HG21	2.04	0.40
1:A:179:LYS:HD2	6:A:2164:HOH:O	2.21	0.40
1:A:106:LYS:HE3	6:A:2104:HOH:O	2.22	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:126[B]:ASN:ND2	6:D:2177:HOH:O[2_656]	1.88	0.32

## 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	198/231 (86%)	194 (98%)	4 (2%)	0	100	100
1	B	202/231 (87%)	198 (98%)	4 (2%)	0	100	100
1	C	196/231 (85%)	191 (97%)	5 (3%)	0	100	100
1	D	195/231 (84%)	190 (97%)	5 (3%)	0	100	100
1	E	207/231 (90%)	200 (97%)	7 (3%)	0	100	100
1	F	200/231 (87%)	193 (96%)	7 (4%)	0	100	100
All	All	1198/1386 (86%)	1166 (97%)	32 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	161/183 (88%)	160 (99%)	1 (1%)	90	82
1	B	164/183 (90%)	163 (99%)	1 (1%)	90	82
1	C	159/183 (87%)	158 (99%)	1 (1%)	90	82
1	D	158/183 (86%)	157 (99%)	1 (1%)	90	82
1	E	167/183 (91%)	166 (99%)	1 (1%)	90	82
1	F	163/183 (89%)	162 (99%)	1 (1%)	90	82
All	All	972/1098 (88%)	966 (99%)	6 (1%)	90	82

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

Mol	Chain	Res	Type
1	A	153	ARG
1	B	153	ARG
1	C	153	ARG
1	D	153	ARG
1	E	153	ARG
1	F	153	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	67	ASN
1	F	67	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.



## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

6 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	KYN	A	62	1	13,14,15	1.20	2 (15%)	13,18,20	1.83	4 (30%)
1	KYN	B	62	1	13,14,15	1.48	3 (23%)	13,18,20	1.91	3 (23%)
1	KYN	C	62	1	13,14,15	1.40	2 (15%)	13,18,20	2.46	6 (46%)
1	KYN	D	62	1	13,14,15	1.49	2 (15%)	13,18,20	2.00	4 (30%)
1	KYN	E	62	1	13,14,15	1.20	2 (15%)	13,18,20	1.39	2 (15%)
1	KYN	F	62	1	13,14,15	1.23	1 (7%)	13,18,20	2.52	5 (38%)

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	KYN	A	62	1	-	0/8/10/12	0/1/1/1
1	KYN	B	62	1	-	0/8/10/12	0/1/1/1
1	KYN	C	62	1	-	0/8/10/12	0/1/1/1
1	KYN	D	62	1	-	0/8/10/12	0/1/1/1
1	KYN	E	62	1	-	0/8/10/12	0/1/1/1
1	KYN	F	62	1	-	0/8/10/12	0/1/1/1

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	62	KYN	CB-CA	-2.54	1.48	1.53
1	A	62	KYN	CG-N1	2.04	1.44	1.38
1	B	62	KYN	CE2-CD2	2.08	1.43	1.39
1	E	62	KYN	CG-N1	2.11	1.45	1.38
1	B	62	KYN	O2-C1	2.14	1.26	1.22
1	D	62	KYN	CD2-CG	2.25	1.45	1.41
1	E	62	KYN	CD2-CG	2.41	1.45	1.41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	62	KYN	CB-C1	2.43	1.54	1.51
1	C	62	KYN	O2-C1	2.56	1.26	1.22
1	F	62	KYN	CD2-C1	2.57	1.54	1.48
1	A	62	KYN	CD2-C1	2.77	1.54	1.48
1	B	62	KYN	CG-N1	3.23	1.48	1.38

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	62	KYN	CD1-CG-CD2	-5.58	112.40	118.06
1	C	62	KYN	CA-CB-C1	-4.67	106.83	113.40
1	B	62	KYN	CA-CB-C1	-4.39	107.22	113.40
1	C	62	KYN	O2-C1-CB	-4.26	115.98	120.75
1	D	62	KYN	O2-C1-CB	-4.03	116.23	120.75
1	A	62	KYN	CA-CB-C1	-3.98	107.80	113.40
1	E	62	KYN	O2-C1-CB	-3.47	116.86	120.75
1	D	62	KYN	CA-CB-C1	-3.47	108.52	113.40
1	C	62	KYN	CD1-CG-CD2	-3.41	114.60	118.06
1	F	62	KYN	CE1-CZ-CE2	-3.33	115.31	120.19
1	F	62	KYN	O2-C1-CB	-3.13	117.25	120.75
1	B	62	KYN	CD1-CG-CD2	-2.98	115.03	118.06
1	D	62	KYN	CD1-CG-CD2	-2.73	115.29	118.06
1	A	62	KYN	CD1-CG-CD2	-2.60	115.42	118.06
1	A	62	KYN	O2-C1-CB	-2.50	117.96	120.75
1	E	62	KYN	CA-CB-C1	-2.23	110.26	113.40
1	C	62	KYN	CE1-CZ-CE2	-2.10	117.11	120.19
1	A	62	KYN	CD2-CG-N1	2.02	125.32	122.74
1	C	62	KYN	CD2-CG-N1	2.22	125.58	122.74
1	D	62	KYN	CE2-CD2-CG	2.35	121.76	119.01
1	F	62	KYN	CD1-CG-N1	2.43	125.03	120.15
1	B	62	KYN	CE2-CD2-CG	3.36	122.95	119.01
1	C	62	KYN	CE2-CD2-CG	3.79	123.45	119.01
1	F	62	KYN	CE2-CD2-CG	3.93	123.62	119.01

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	62	KYN	1	0

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 29 ligands modelled in this entry, 12 are monoatomic - leaving 17 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	PG4	A	1234	-	3,3,12	0.56	0	2,2,11	0.18	0
5	P6G	A	1235	-	9,9,18	0.63	0	8,8,17	1.01	0
5	P6G	A	1236	-	6,6,18	0.59	0	5,5,17	0.47	0
4	PG4	B	1234	-	5,5,12	0.67	0	4,4,11	0.64	0
5	P6G	B	1235	-	10,10,18	0.53	0	9,9,17	0.82	0
5	P6G	B	1236	-	7,7,18	0.65	0	6,6,17	0.81	0
4	PG4	C	1234	-	5,5,12	0.79	0	4,4,11	0.50	0
5	P6G	C	1235	-	18,18,18	0.64	0	17,17,17	0.69	0
5	P6G	C	1236	-	14,14,18	0.55	0	13,13,17	0.39	0
4	PG4	D	1234	-	6,6,12	0.39	0	5,5,11	0.97	0
5	P6G	D	1235	-	7,7,18	0.73	0	6,6,17	0.72	0
4	PG4	E	1234	-	5,5,12	0.62	0	4,4,11	0.58	0
5	P6G	E	1235	-	13,13,18	0.58	0	12,12,17	0.90	0
5	P6G	E	1236	-	12,12,18	0.36	0	11,11,17	0.83	0
5	P6G	E	1237	-	7,7,18	0.64	0	6,6,17	0.87	0
4	PG4	F	1234	-	7,7,12	0.69	0	6,6,11	0.45	0
5	P6G	F	1235	-	10,10,18	0.72	0	9,9,17	1.33	1 (11%)

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	PG4	A	1234	-	-	0/1/1/10	0/0/0/0
5	P6G	A	1235	-	-	0/7/7/16	0/0/0/0
5	P6G	A	1236	-	-	0/4/4/16	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	PG4	B	1234	-	-	0/3/3/10	0/0/0/0
5	P6G	B	1235	-	-	0/8/8/16	0/0/0/0
5	P6G	B	1236	-	-	0/5/5/16	0/0/0/0
4	PG4	C	1234	-	-	0/3/3/10	0/0/0/0
5	P6G	C	1235	-	-	0/16/16/16	0/0/0/0
5	P6G	C	1236	-	-	0/12/12/16	0/0/0/0
4	PG4	D	1234	-	-	0/4/4/10	0/0/0/0
5	P6G	D	1235	-	-	0/5/5/16	0/0/0/0
4	PG4	E	1234	-	-	0/3/3/10	0/0/0/0
5	P6G	E	1235	-	-	0/11/11/16	0/0/0/0
5	P6G	E	1236	-	-	0/10/10/16	0/0/0/0
5	P6G	E	1237	-	-	0/5/5/16	0/0/0/0
4	PG4	F	1234	-	-	0/5/5/10	0/0/0/0
5	P6G	F	1235	-	-	0/8/8/16	0/0/0/0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	F	1235	P6G	C11-O10-C9	2.60	124.50	113.31

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	1234	PG4	1	0
5	B	1235	P6G	2	0
4	F	1234	PG4	2	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	198/231 (85%)	-0.34	3 (1%) 76 75	8, 15, 28, 63	0
1	B	203/231 (87%)	-0.41	0 100 100	7, 14, 25, 31	0
1	C	198/231 (85%)	-0.34	2 (1%) 84 84	7, 13, 29, 52	0
1	D	197/231 (85%)	-0.30	2 (1%) 84 84	7, 14, 31, 53	0
1	E	203/231 (87%)	-0.46	0 100 100	6, 11, 22, 33	0
1	F	198/231 (85%)	-0.39	3 (1%) 76 75	6, 11, 26, 52	0
All	All	1197/1386 (86%)	-0.37	10 (0%) 87 87	6, 13, 28, 63	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	136	PRO	4.4
1	F	130	THR	4.0
1	D	137	VAL	3.3
1	C	136	PRO	3.3
1	A	136	PRO	3.3
1	A	138	LYS	2.3
1	D	130	THR	2.3
1	A	137	VAL	2.2
1	C	130	THR	2.2
1	F	103	LYS	2.1

### 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. 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
1	KYN	F	62	14/15	0.96	0.09	-	8,10,16,16	0
1	KYN	A	62	14/15	0.95	0.08	-	11,13,17,17	0
1	KYN	D	62	14/15	0.95	0.08	-	9,11,15,15	0
1	KYN	B	62	14/15	0.95	0.07	-	8,12,15,16	0
1	KYN	C	62	14/15	0.95	0.07	-	10,12,17,17	0
1	KYN	E	62	14/15	0.96	0.07	-	8,11,18,18	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. 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
4	PG4	F	1234	8/13	0.80	0.18	15.71	22,30,35,36	0
5	P6G	F	1235	11/19	0.85	0.14	11.41	27,30,35,42	0
4	PG4	D	1234	7/13	0.90	0.17	9.00	23,27,31,41	0
4	PG4	E	1234	6/13	0.92	0.10	8.66	21,23,25,25	0
4	PG4	B	1234	6/13	0.90	0.12	6.96	28,30,33,34	0
4	PG4	C	1234	6/13	0.85	0.13	6.58	27,28,29,32	0
5	P6G	B	1236	8/19	0.88	0.11	6.56	35,37,44,49	0
5	P6G	B	1235	11/19	0.80	0.15	6.28	30,32,40,40	0
5	P6G	E	1235	14/19	0.88	0.14	5.28	24,29,45,48	0
5	P6G	E	1237	8/19	0.83	0.13	4.74	35,39,40,44	0
5	P6G	E	1236	13/19	0.92	0.11	4.56	25,28,41,47	0
4	PG4	A	1234	4/13	0.90	0.11	3.74	30,31,32,33	0
5	P6G	D	1235	8/19	0.78	0.16	2.72	31,37,41,43	0
5	P6G	C	1235	19/19	0.90	0.10	2.58	20,25,43,44	0
5	P6G	A	1235	10/19	0.91	0.09	2.43	29,33,36,37	0
5	P6G	A	1236	7/19	0.81	0.13	2.41	41,45,46,52	0
5	P6G	C	1236	15/19	0.86	0.12	0.89	28,31,37,38	0
3	CA	B	1233	1/1	0.99	0.06	-0.33	16,16,16,16	0
3	CA	E	1233	1/1	1.00	0.07	-0.40	13,13,13,13	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CA	A	1233	1/1	1.00	0.04	-1.39	14,14,14,14	0
3	CA	D	1233	1/1	1.00	0.04	-1.68	13,13,13,13	0
3	CA	C	1233	1/1	1.00	0.04	-1.74	14,14,14,14	0
3	CA	F	1233	1/1	1.00	0.04	-1.85	11,11,11,11	0
2	CU1	F	1232	1/1	0.99	0.03	-	18,18,18,18	0
2	CU1	A	1232	1/1	0.99	0.03	-	20,20,20,20	0
2	CU1	C	1232	1/1	1.00	0.02	-	19,19,19,19	0
2	CU1	B	1232	1/1	1.00	0.03	-	13,13,13,13	0
2	CU1	E	1232	1/1	0.99	0.04	-	14,14,14,14	0
2	CU1	D	1232	1/1	1.00	0.03	-	18,18,18,18	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.