Full wwPDB X-ray Structure Validation Report (i)

Oct 8, 2023 – 09:04 AM EDT

PDB ID	:	4QYI
Title	:	1.95 Angstrom resolution crystal structure of a hypoxanthine-guanine phosphoribosyltransferase (hpt-2) from Bacillus anthracis str. 'Ames Ancestor' with HEPES molecule in the active site
Authors	:	Halavaty, A.S.; Minasov, G.; Dubrovska, I.; Winsor, J.; Shuvalova, L.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on	:	2014-07-24
Resolution	:	1.95 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

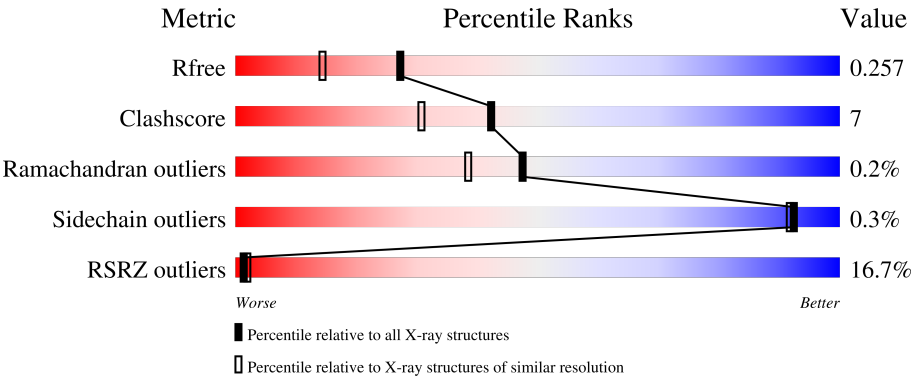
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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.



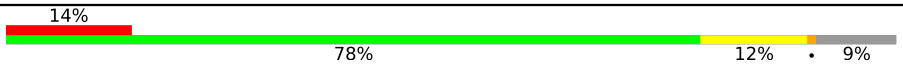
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	186	26% 79% 11% • 9%
2	B	186	6% 81% 11% 8%
2	C	186	7% 81% 11% • 8%
2	D	186	23% 80% 11% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	186	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 7518 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 Hypoxanthine phosphoribosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	170	Total	C	N	O	S	0	4	0
			1390	899	225	263	3			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q81KC7
A	-1	ASN	-	expression tag	UNP Q81KC7
A	0	ALA	-	expression tag	UNP Q81KC7

- Molecule 2 is a protein called Hypoxanthine phosphoribosyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	171	Total	C	N	O	S	0	9	0
			1447	931	235	277	4			
2	C	171	Total	C	N	O	S	0	10	0
			1456	937	233	280	6			
2	D	169	Total	C	N	O	S	0	0	0
			1353	877	218	254	4			
2	E	169	Total	C	N	O	S	0	4	0
			1384	893	222	265	4			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	SER	-	expression tag	UNP Q81KC7
B	-1	ASN	-	expression tag	UNP Q81KC7
B	0	ALA	-	expression tag	UNP Q81KC7
C	-2	SER	-	expression tag	UNP Q81KC7
C	-1	ASN	-	expression tag	UNP Q81KC7
C	0	ALA	-	expression tag	UNP Q81KC7
D	-2	SER	-	expression tag	UNP Q81KC7

Continued on next page...

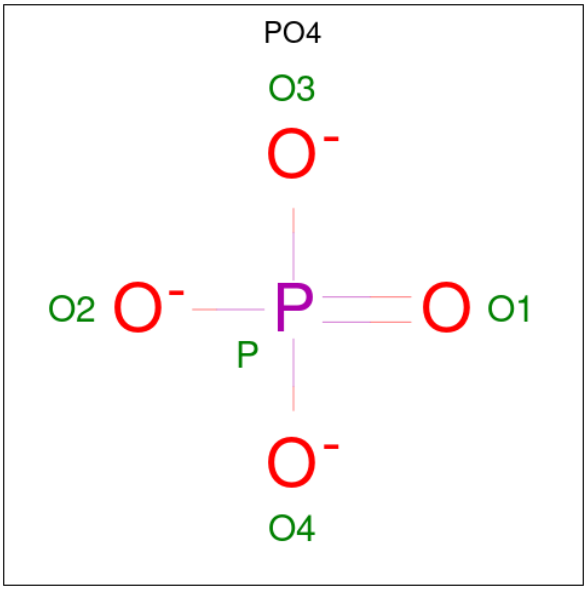
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-1	ASN	-	expression tag	UNP Q81KC7
D	0	ALA	-	expression tag	UNP Q81KC7
E	-2	SER	-	expression tag	UNP Q81KC7
E	-1	ASN	-	expression tag	UNP Q81KC7
E	0	ALA	-	expression tag	UNP Q81KC7

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mg 1 1	0	0
3	B	2	Total Mg 2 2	0	0
3	C	2	Total Mg 2 2	0	0
3	D	1	Total Mg 1 1	0	0
3	E	3	Total Mg 3 3	0	0

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



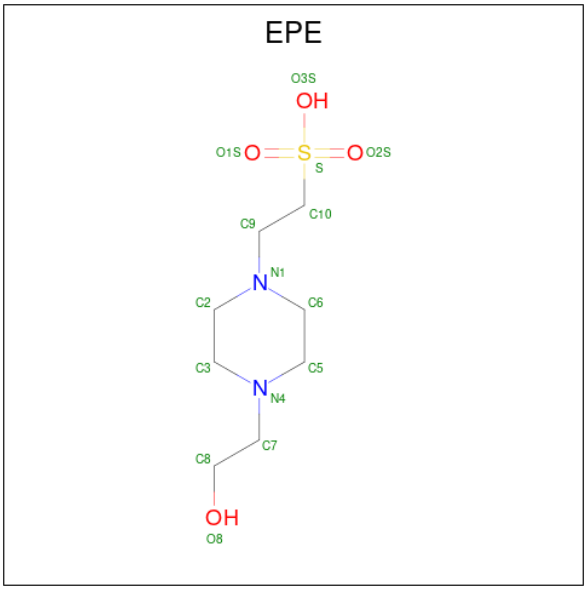
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 5 4 1	0	0

Continued on next page...

Continued from previous page...

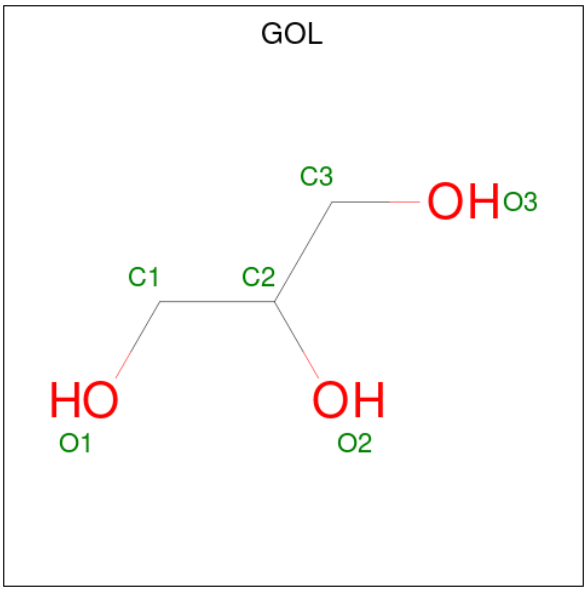
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O P 5 4 1	0	0
4	B	1	Total O P 5 4 1	0	0
4	C	1	Total O P 5 4 1	0	0
4	D	1	Total O P 5 4 1	0	0
4	D	1	Total O P 5 4 1	0	1
4	E	1	Total O P 5 4 1	0	0
4	E	1	Total O P 5 4 1	0	1

- Molecule 5 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total C N O S 15 8 2 4 1	0	0
5	C	1	Total C N O S 15 8 2 4 1	0	0
5	E	1	Total C N O S 15 8 2 4 1	0	1

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	E	1	Total C O 6 3 3	0	0

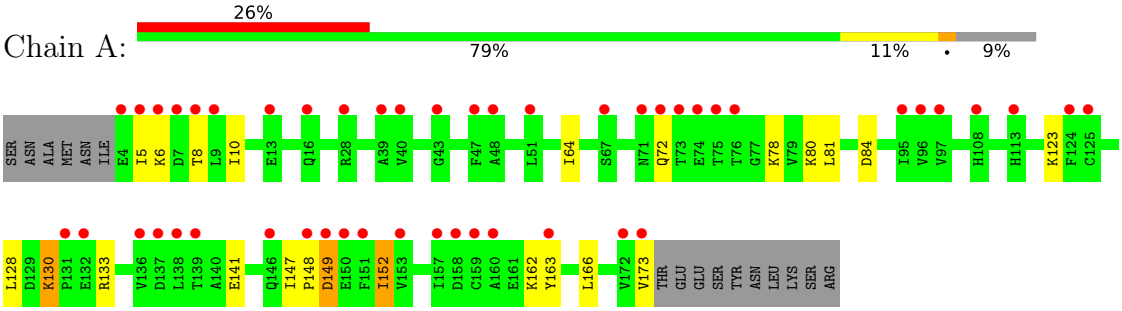
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	34	Total O 35 35	0	1
7	B	122	Total O 126 126	0	9
7	C	101	Total O 109 109	0	11
7	D	26	Total O 26 26	0	1
7	E	87	Total O 92 92	0	8

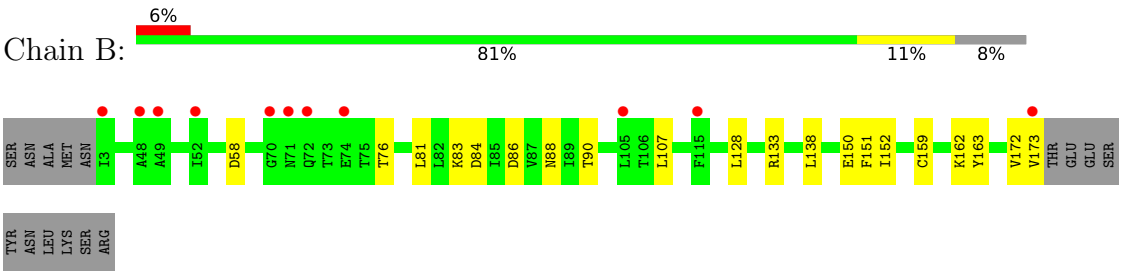
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

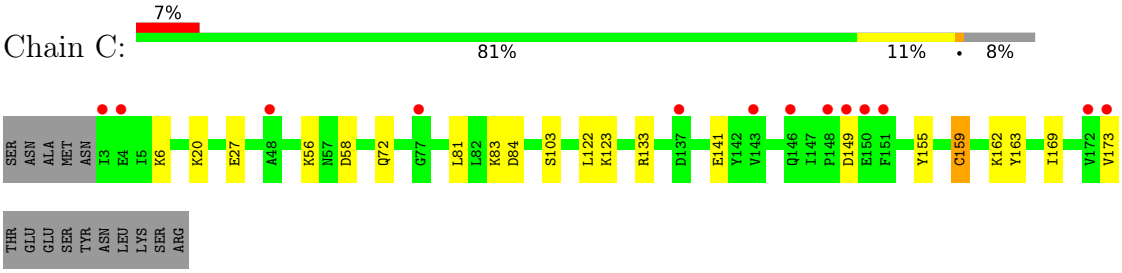
• Molecule 1: Hypoxanthine phosphoribosyltransferase



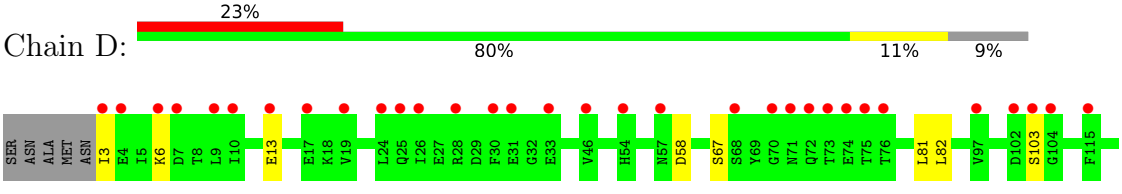
• Molecule 2: Hypoxanthine phosphoribosyltransferase

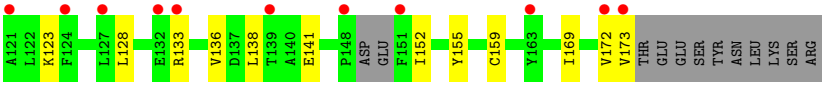


• Molecule 2: Hypoxanthine phosphoribosyltransferase



• Molecule 2: Hypoxanthine phosphoribosyltransferase





• Molecule 2: Hypoxanthine phosphoribosyltransferase

Chain E:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	172.78Å 117.01Å 56.88Å 90.00° 90.96° 90.00°	Depositor
Resolution (Å)	27.99 – 1.95 27.97 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.5 (27.99-1.95) 98.6 (27.97-1.95)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.41 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.221 , 0.252 0.225 , 0.257	Depositor DCC
R_{free} test set	4060 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	34.5	Xtriage
Anisotropy	0.535	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 53.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.018 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7518	wwPDB-VP
Average B, all atoms (Å^2)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.56% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, MG, EPE, PO4, CME

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.48	0/1412	0.80	1/1903 (0.1%)
2	B	0.67	0/1458	0.83	0/1963
2	C	0.69	0/1457	0.81	0/1961
2	D	0.48	0/1363	0.77	0/1836
2	E	0.57	0/1393	0.79	0/1877
All	All	0.59	0/7083	0.80	1/9540 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	130	LYS	CB-CG-CD	5.49	125.87	111.60

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	1390	0	1425	26	0
2	B	1447	0	1468	25	0
2	C	1456	0	1475	17	0
2	D	1353	0	1389	19	0
2	E	1384	0	1412	24	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	1	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	1	0	0	0	0
3	E	3	0	0	0	0
4	A	10	0	0	0	0
4	B	5	0	0	0	0
4	C	5	0	0	0	0
4	D	10	0	0	0	0
4	E	10	0	0	1	0
5	B	15	0	18	1	0
5	C	15	0	18	0	0
5	E	15	0	17	4	0
6	E	6	0	8	0	0
7	A	35	0	0	2	0
7	B	126	0	0	2	0
7	C	109	0	0	4	0
7	D	26	0	0	0	0
7	E	92	0	0	0	0
All	All	7518	0	7230	100	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 7.

All (100) 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)
2:B:133[B]:ARG:HB3	2:B:133[B]:ARG:CZ	1.96	0.95
2:E:68:SER:HB2	2:E:73:THR:OG1	1.73	0.88
2:D:172:VAL:HG22	2:D:173:VAL:N	1.89	0.86
2:B:133[B]:ARG:HB3	2:B:133[B]:ARG:NH1	1.92	0.83
2:D:172:VAL:HG22	2:D:173:VAL:H	1.41	0.83
1:A:10:ILE:HD11	1:A:147:ILE:HD11	1.65	0.79
1:A:128:LEU:HD22	1:A:147:ILE:HD12	1.69	0.73
2:B:150:GLU:OE1	2:B:151:PHE:CE2	2.42	0.73
2:E:106:THR:HB	5:E:207[B]:EPE:O2S	1.89	0.72
2:D:3:ILE:HD13	2:D:159:CME:SG	2.33	0.68
2:E:64:ILE:HD13	2:E:114:PHE:HZ	1.59	0.68
1:A:152:ILE:H	1:A:152:ILE:HD13	1.57	0.67
2:E:162:LYS:HE2	2:E:163:TYR:CZ	2.30	0.67
2:C:133:ARG:HD2	2:C:149:ASP:CG	2.15	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:20:LYS:NZ	7:C:339:HOH:O	2.28	0.66
2:B:58[A]:ASP:OD2	2:C:162:LYS:NZ	2.28	0.65
2:D:172:VAL:CG2	2:D:173:VAL:N	2.59	0.65
2:D:13:GLU:OE1	2:D:13:GLU:N	2.18	0.63
1:A:147:ILE:HG23	1:A:148:PRO:HD2	1.81	0.62
2:D:172:VAL:CG2	2:D:173:VAL:H	2.11	0.62
2:C:162:LYS:HE3	2:C:163:TYR:CZ	2.34	0.62
2:B:133[B]:ARG:CZ	2:B:133[B]:ARG:CB	2.75	0.62
2:D:13:GLU:H	2:D:13:GLU:CD	2.02	0.61
2:C:83[B]:LYS:HE3	7:C:337:HOH:O	2.02	0.58
2:E:102:ASP:N	4:E:206[A]:PO4:O4	2.34	0.58
1:A:80[B]:LYS:NZ	2:B:86:ASP:OD1	2.30	0.58
1:A:162:LYS:HE3	1:A:163:TYR:CZ	2.39	0.57
1:A:130:LYS:HE3	1:A:149:ASP:HA	1.87	0.56
2:B:162:LYS:NZ	2:B:163:TYR:CZ	2.74	0.55
2:D:3:ILE:HG23	2:D:172:VAL:HG21	1.89	0.54
1:A:8:THR:HG21	7:A:333:HOH:O	2.07	0.54
1:A:8:THR:CG2	7:A:333:HOH:O	2.55	0.53
2:B:83[B]:LYS:HE3	7:B:404:HOH:O	2.09	0.53
2:D:6:LYS:HE3	2:D:173:VAL:HG21	1.92	0.52
2:E:2:ASN:O	2:E:174:THR:HA	2.11	0.51
1:A:128:LEU:HD22	1:A:147:ILE:CD1	2.38	0.51
2:E:107:LEU:N	5:E:207[B]:EPE:O1S	2.43	0.51
2:E:40:VAL:HG22	2:E:64:ILE:HD11	1.92	0.51
2:C:81:LEU:HD12	2:D:81:LEU:HD12	1.93	0.51
2:E:64:ILE:HD13	2:E:114:PHE:CZ	2.43	0.51
2:E:106:THR:HB	5:E:207[B]:EPE:S	2.50	0.51
1:A:123[A]:LYS:HD2	1:A:141:GLU:HG2	1.94	0.50
2:E:68:SER:CB	2:E:73:THR:OG1	2.53	0.49
2:E:64:ILE:HG22	2:E:85:ILE:HG23	1.94	0.49
2:C:155:TYR:HB3	2:C:169[A]:ILE:HD11	1.94	0.49
1:A:152:ILE:HD13	1:A:152:ILE:N	2.25	0.49
2:E:123:LYS:HD2	2:E:141:GLU:HG2	1.94	0.49
2:E:107:LEU:HB2	5:E:207[B]:EPE:O1S	2.13	0.49
2:D:123:LYS:HD2	2:D:141:GLU:HG2	1.94	0.48
1:A:133:ARG:HD3	1:A:149:ASP:OD2	2.12	0.48
2:E:155:TYR:HB3	2:E:169[A]:ILE:HD11	1.96	0.48
2:B:159:CME:HE3	2:B:159:CME:HB3	1.67	0.47
1:A:81[B]:LEU:HD21	1:A:84:ASP:CG	2.35	0.47
2:B:150:GLU:HG3	2:B:151:PHE:N	2.30	0.47
2:E:133:ARG:HD3	2:E:149:ASP:OD2	2.15	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:76:THR:C	2:E:78:LYS:H	2.19	0.46
1:A:64:ILE:HD11	1:A:81[B]:LEU:CD1	2.45	0.46
2:B:83[B]:LYS:HD3	2:B:84:ASP:O	2.15	0.46
2:C:122:LEU:C	2:C:123[B]:LYS:HD2	2.35	0.46
2:C:123[A]:LYS:HD2	2:C:141:GLU:HG2	1.96	0.46
2:D:155:TYR:HB3	2:D:169:ILE:HD11	1.97	0.46
2:C:103:SER:HA	2:C:133:ARG:O	2.16	0.46
1:A:78:LYS:HE2	2:B:90:THR:OG1	2.16	0.46
2:E:68:SER:HB3	2:E:79:VAL:HG12	1.97	0.45
2:E:68:SER:CB	2:E:79:VAL:HG12	2.47	0.45
1:A:78:LYS:CE	2:B:90:THR:OG1	2.65	0.45
2:B:107:LEU:HB2	5:B:204:EPE:O1S	2.17	0.44
1:A:147:ILE:HG22	1:A:148:PRO:O	2.18	0.44
1:A:72:GLN:OE1	2:B:88:ASN:ND2	2.50	0.44
2:D:138:LEU:HD12	2:D:138:LEU:HA	1.78	0.44
2:C:72:GLN:HG3	7:C:392:HOH:O	2.16	0.44
2:B:162:LYS:HE2	2:C:58[A]:ASP:OD2	2.17	0.44
2:C:83[B]:LYS:HD3	2:C:84:ASP:O	2.18	0.43
2:E:68:SER:OG	2:E:79:VAL:HG12	2.18	0.43
2:C:133:ARG:HD2	2:C:149:ASP:OD1	2.17	0.43
2:D:13:GLU:N	2:D:13:GLU:CD	2.68	0.43
2:D:3:ILE:HG23	2:D:172:VAL:CG2	2.49	0.43
1:A:78:LYS:NZ	2:B:90:THR:OG1	2.50	0.43
2:E:2:ASN:O	2:E:174:THR:HG23	2.19	0.43
1:A:64:ILE:HD11	1:A:81[B]:LEU:HD11	2.01	0.42
2:D:103:SER:HA	2:D:133:ARG:O	2.19	0.42
2:D:128:LEU:CD1	2:D:152:ILE:CD1	2.97	0.42
2:E:76:THR:C	2:E:78:LYS:N	2.72	0.42
1:A:81[A]:LEU:HD12	2:B:81[A]:LEU:HD22	2.01	0.42
2:B:172:VAL:HG12	2:B:173:VAL:N	2.34	0.42
2:C:159[A]:CME:HE2	7:C:395:HOH:O	2.18	0.42
1:A:6:LYS:HD2	1:A:173:VAL:HG11	2.02	0.42
2:C:6:LYS:HB2	2:C:173:VAL:CG1	2.50	0.42
2:B:150:GLU:HG3	2:B:151:PHE:CG	2.55	0.42
2:B:138:LEU:C	2:B:138:LEU:HD23	2.41	0.41
2:B:58[A]:ASP:OD1	2:B:58[A]:ASP:C	2.59	0.41
1:A:162:LYS:NZ	2:D:58:ASP:OD2	2.54	0.41
2:B:76:THR:O	7:B:398:HOH:O	2.22	0.41
2:C:27:GLU:OE2	2:C:56:LYS:HG2	2.21	0.41
2:D:67:SER:HB3	2:D:82:LEU:HD11	2.03	0.41
2:B:128:LEU:CD1	2:B:152:ILE:CD1	3.00	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:132[A]:GLU:H	2:E:132[A]:GLU:CD	2.23	0.40
2:E:64:ILE:CG2	2:E:85:ILE:HG23	2.51	0.40
1:A:80[A]:LYS:NZ	2:B:83[A]:LYS:NZ	2.69	0.40
1:A:5:ILE:CD1	1:A:166:LEU:HD11	2.52	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	172/186 (92%)	169 (98%)	2 (1%)	1 (1%)	25 14
2	B	177/186 (95%)	173 (98%)	4 (2%)	0	100 100
2	C	177/186 (95%)	174 (98%)	3 (2%)	0	100 100
2	D	164/186 (88%)	161 (98%)	3 (2%)	0	100 100
2	E	168/186 (90%)	165 (98%)	2 (1%)	1 (1%)	25 14
All	All	858/930 (92%)	842 (98%)	14 (2%)	2 (0%)	47 38

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	149	ASP
2	E	149	ASP

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	156/167 (93%)	155 (99%)	1 (1%)	86 85
2	B	161/166 (97%)	161 (100%)	0	100 100
2	C	161/166 (97%)	161 (100%)	0	100 100
2	D	150/166 (90%)	149 (99%)	1 (1%)	84 82
2	E	155/166 (93%)	155 (100%)	0	100 100
All	All	783/831 (94%)	781 (100%)	2 (0%)	92 92

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	152	ILE
2	D	136	VAL

Sometimes 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)

5 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	$\#\|Z\| > 2$
2	CME	C	159[B]	2	8,9,10	0.64	0	5,9,11	0.75	0
2	CME	B	159	2	8,9,10	0.81	0	5,9,11	1.00	0
2	CME	E	159	2	8,9,10	0.70	0	5,9,11	1.24	1 (20%)
2	CME	C	159[A]	2	8,9,10	0.82	0	5,9,11	1.87	2 (40%)
2	CME	D	159	2	8,9,10	0.80	0	5,9,11	1.12	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
2	CME	C	159[B]	2	-	0/5/8/10	-
2	CME	B	159	2	-	2/5/8/10	-
2	CME	E	159	2	-	1/5/8/10	-
2	CME	C	159[A]	2	-	2/5/8/10	-
2	CME	D	159	2	-	2/5/8/10	-

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	159[A]	CME	CB-SG-SD	-3.04	95.95	103.82
2	C	159[A]	CME	CZ-CE-SD	-2.28	105.44	113.37
2	E	159	CME	CB-SG-SD	-2.04	98.53	103.82

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	159	CME	SD-CE-CZ-OH
2	D	159	CME	CE-SD-SG-CB
2	D	159	CME	SD-CE-CZ-OH
2	B	159	CME	CE-SD-SG-CB
2	E	159	CME	SD-CE-CZ-OH
2	C	159[A]	CME	CE-SD-SG-CB
2	C	159[A]	CME	CZ-CE-SD-SG

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	159	CME	1	0
2	C	159[A]	CME	1	0
2	D	159	CME	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 21 ligands modelled in this entry, 9 are monoatomic - leaving 12 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	$\#\|Z\| > 2$
5	EPE	E	207[B]	3	15,15,15	1.94	1 (6%)	18,20,20	1.31	2 (11%)
5	EPE	C	204	-	15,15,15	1.83	1 (6%)	18,20,20	1.44	3 (16%)
4	PO4	B	203	3	4,4,4	1.02	0	6,6,6	0.83	0
4	PO4	A	202	-	4,4,4	0.90	0	6,6,6	0.59	0
4	PO4	C	203	3	4,4,4	1.32	1 (25%)	6,6,6	0.99	0
5	EPE	B	204	3	15,15,15	1.76	1 (6%)	18,20,20	1.28	1 (5%)
4	PO4	D	203[A]	-	4,4,4	0.93	0	6,6,6	0.63	0
6	GOL	E	204	-	5,5,5	0.48	0	5,5,5	0.82	0
4	PO4	A	203	-	4,4,4	1.13	0	6,6,6	0.79	0
4	PO4	E	206[A]	-	4,4,4	0.98	0	6,6,6	0.42	0
4	PO4	E	205	3	4,4,4	0.94	0	6,6,6	0.78	0
4	PO4	D	202	-	4,4,4	0.85	0	6,6,6	0.80	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
5	EPE	E	207[B]	3	-	5/9/19/19	0/1/1/1
5	EPE	B	204	3	-	0/9/19/19	0/1/1/1
6	GOL	E	204	-	-	2/4/4/4	-
5	EPE	C	204	-	-	2/9/19/19	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	207[B]	EPE	C10-S	-7.29	1.67	1.77
5	C	204	EPE	C10-S	-6.73	1.67	1.77
5	B	204	EPE	C10-S	-6.51	1.68	1.77
4	C	203	PO4	P-O3	-2.28	1.47	1.54

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	C	204	EPE	O2S-S-C10	4.09	111.84	106.92
5	B	204	EPE	O2S-S-C10	3.40	111.01	106.92
5	E	207[B]	EPE	O3S-S-C10	2.92	110.49	105.77
5	C	204	EPE	O3S-S-O1S	2.45	117.26	111.27
5	E	207[B]	EPE	C6-N1-C2	2.37	114.15	108.83
5	C	204	EPE	O2S-S-O1S	-2.31	105.97	113.95

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	207[B]	EPE	C10-C9-N1-C2
5	E	207[B]	EPE	C8-C7-N4-C3
6	E	204	GOL	C1-C2-C3-O3
6	E	204	GOL	O2-C2-C3-O3
5	C	204	EPE	N4-C7-C8-O8
5	C	204	EPE	C8-C7-N4-C5
5	E	207[B]	EPE	S-C10-C9-N1
5	E	207[B]	EPE	C10-C9-N1-C6
5	E	207[B]	EPE	C9-C10-S-O3S

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	207[B]	EPE	4	0
5	B	204	EPE	1	0
4	E	206[A]	PO4	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

6.1 Protein, DNA and RNA chains

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^{th} 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^2)	Q<0.9
1	A	170/186 (91%)	1.32	48 (28%) 0 0	42, 63, 105, 117	0
2	B	170/186 (91%)	0.53	11 (6%) 18 27	23, 40, 69, 94	0
2	C	170/186 (91%)	0.62	13 (7%) 13 21	25, 42, 78, 99	0
2	D	168/186 (90%)	1.46	43 (25%) 0 0	43, 66, 103, 115	0
2	E	168/186 (90%)	0.76	26 (15%) 2 3	28, 48, 92, 122	0
All	All	846/930 (90%)	0.94	141 (16%) 1 2	23, 52, 97, 122	0

All (141) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	173	VAL	8.2
2	D	104	GLY	7.0
1	A	148	PRO	6.7
2	C	3	ILE	6.5
2	B	3	ILE	6.4
2	D	3	ILE	6.3
2	D	70	GLY	6.1
1	A	172	VAL	5.8
2	D	6	LYS	5.7
2	B	173	VAL	5.6
2	D	28	ARG	5.5
2	D	133	ARG	5.4
1	A	8	THR	5.3
2	D	75	THR	5.2
2	D	71	ASN	5.1
2	D	132	GLU	5.1
2	C	173	VAL	5.1
2	D	173	VAL	5.0
2	D	73	THR	4.9
1	A	151	PHE	4.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	C	149	ASP	4.7
2	E	67	SER	4.7
2	D	9	LEU	4.6
2	D	148	PRO	4.5
1	A	159	CYS	4.5
1	A	7	ASP	4.4
2	B	72	GLN	4.4
2	B	71	ASN	4.3
1	A	67	SER	4.3
2	D	72	GLN	4.1
1	A	72	GLN	4.1
2	E	68	SER	4.0
2	E	116	MET	3.9
2	E	115	PHE	3.9
2	D	30	PHE	3.9
2	C	148	PRO	3.9
2	D	74	GLU	3.9
1	A	136	VAL	3.9
2	D	172	VAL	3.9
2	C	172	VAL	3.8
2	E	45	PHE	3.8
2	C	4	GLU	3.8
2	D	163	TYR	3.8
2	D	31	GLU	3.8
2	B	74	GLU	3.7
1	A	5	ILE	3.7
1	A	138	LEU	3.5
2	D	115	PHE	3.5
1	A	132	GLU	3.4
1	A	139	THR	3.4
1	A	137	ASP	3.4
2	D	4	GLU	3.3
1	A	4	GLU	3.3
2	E	97	VAL	3.3
2	D	139	THR	3.3
2	D	25	GLN	3.3
2	B	115	PHE	3.2
2	C	137	ASP	3.2
2	D	103	SER	3.2
1	A	75	THR	3.2
1	A	96	VAL	3.1
2	D	19	VAL	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	E	51	LEU	3.1
2	D	68	SER	3.1
2	E	137	ASP	3.0
1	A	97	VAL	3.0
2	D	7	ASP	3.0
2	E	76	THR	3.0
2	D	151	PHE	2.9
2	C	150	GLU	2.9
1	A	76	THR	2.9
2	D	13	GLU	2.9
1	A	131	PRO	2.9
2	C	151	PHE	2.9
2	E	49	ALA	2.8
1	A	158	ASP	2.8
1	A	6	LYS	2.8
2	D	57	ASN	2.8
1	A	40	VAL	2.8
1	A	160	ALA	2.7
2	E	148	PRO	2.7
2	D	24	LEU	2.7
1	A	95	ILE	2.7
2	C	77	GLY	2.7
2	E	55	ILE	2.7
2	E	78	LYS	2.7
2	E	46	VAL	2.6
1	A	13	GLU	2.6
2	B	105	LEU	2.6
1	A	48	ALA	2.6
2	C	146	GLN	2.6
1	A	73	THR	2.6
2	B	70	GLY	2.6
2	D	33	GLU	2.5
2	D	124	PHE	2.5
1	A	74	GLU	2.5
2	E	80	LYS	2.5
1	A	153	VAL	2.5
2	D	54	HIS	2.4
1	A	146	GLN	2.4
2	E	37	VAL	2.4
2	D	17	GLU	2.4
2	E	48	ALA	2.4
2	E	40	VAL	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	E	110	LEU	2.3
1	A	149	ASP	2.3
1	A	28	ARG	2.3
1	A	124	PHE	2.3
2	E	77	GLY	2.3
2	E	41	LEU	2.3
2	E	79	VAL	2.3
2	B	52	ILE	2.3
1	A	51	LEU	2.3
1	A	47	PHE	2.2
1	A	71	ASN	2.2
2	E	38	ILE	2.2
1	A	43	GLY	2.2
2	C	48	ALA	2.2
2	D	26	ILE	2.2
2	E	75	THR	2.2
2	D	102	ASP	2.2
2	C	143	VAL	2.2
2	D	46	VAL	2.2
2	E	63	PHE	2.1
1	A	157	ILE	2.1
1	A	9	LEU	2.1
1	A	39	ALA	2.1
2	D	121	ALA	2.1
1	A	108	HIS	2.1
1	A	150	GLU	2.1
2	D	10	ILE	2.1
2	B	48	ALA	2.1
1	A	163	TYR	2.1
2	B	49	ALA	2.1
1	A	16	GLN	2.1
2	D	76	THR	2.1
2	D	97	VAL	2.1
1	A	125	CYS	2.1
2	D	127	LEU	2.0
2	E	52	ILE	2.0
1	A	113	HIS	2.0

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^{th} 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^2)	Q<0.9
2	CME	C	159[A]	10/11	0.84	0.18	45,53,67,67	10
2	CME	C	159[B]	10/11	0.84	0.18	45,53,70,75	10
2	CME	D	159	10/11	0.87	0.22	61,70,103,104	0
2	CME	E	159	10/11	0.89	0.14	43,47,71,74	0
2	CME	B	159	10/11	0.95	0.19	34,40,64,64	0

6.3 Carbohydrates (i)

There are no monosaccharides 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^{th} 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^2)	Q<0.9
3	MG	C	202	1/1	0.75	0.10	63,63,63,63	0
4	PO4	A	202	5/5	0.79	0.20	65,79,88,98	5
6	GOL	E	204	6/6	0.79	0.19	48,67,74,78	0
4	PO4	D	203[A]	5/5	0.81	0.40	47,54,65,70	5
3	MG	B	202	1/1	0.82	0.07	67,67,67,67	0
3	MG	D	201	1/1	0.85	0.15	66,66,66,66	0
5	EPE	E	207[B]	15/15	0.92	0.22	44,56,67,70	15
4	PO4	A	203	5/5	0.92	0.17	54,63,65,72	5
4	PO4	D	202	5/5	0.93	0.10	56,73,76,78	0
3	MG	E	203	1/1	0.93	0.07	62,62,62,62	0
3	MG	E	202	1/1	0.94	0.07	71,71,71,71	0
3	MG	C	201	1/1	0.94	0.09	35,35,35,35	0
3	MG	A	201	1/1	0.95	0.07	55,55,55,55	0
5	EPE	B	204	15/15	0.95	0.16	44,65,83,87	0
5	EPE	C	204	15/15	0.96	0.15	41,61,85,90	0
4	PO4	B	203	5/5	0.96	0.08	36,46,55,55	0
3	MG	B	201	1/1	0.96	0.08	35,35,35,35	0
3	MG	E	201	1/1	0.97	0.09	52,52,52,52	0
4	PO4	C	203	5/5	0.97	0.08	41,43,52,53	0
4	PO4	E	205	5/5	0.98	0.08	50,56,64,74	0
4	PO4	E	206[A]	5/5	0.98	0.16	43,48,56,64	5

6.5 Other polymers (i)

There are no such residues in this entry.