W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 10:15 PM GMT

PDB ID	:	4X4U
Title	:	Crystal structure of the A.fulgidus CCA-adding enzyme in complex with a human MenBeta minihelix ending in CCACC
Authors	:	Kuhn, C.-D.; Joshua-Tor, L.
Deposited on	:	2014-12-03
Resolution	:	2.70 Å(reported)

The following versions of software and data (see references (i)) 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 (i)

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

The reported resolution of this entry is 2.70 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.
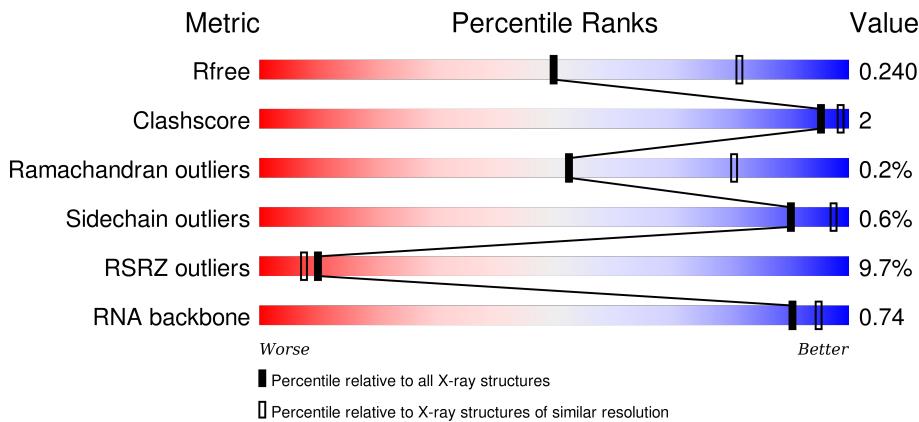

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)
RNA backbone	2183	1069 (3.10-2.30)

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 >=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 <=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	457	8% ... 91% ... 5%
1	C	457	14% ... 91% ...
1	E	457	4% ... 93% ...
1	F	457	3% ... 92% ... 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	B	37	43% / 65% / 14% / 5% / 16%
2	D	37	59% / 49% / 22% / 30%
2	G	37	8% / 32% / 16% / 51%
3	H	2	100% / 100%
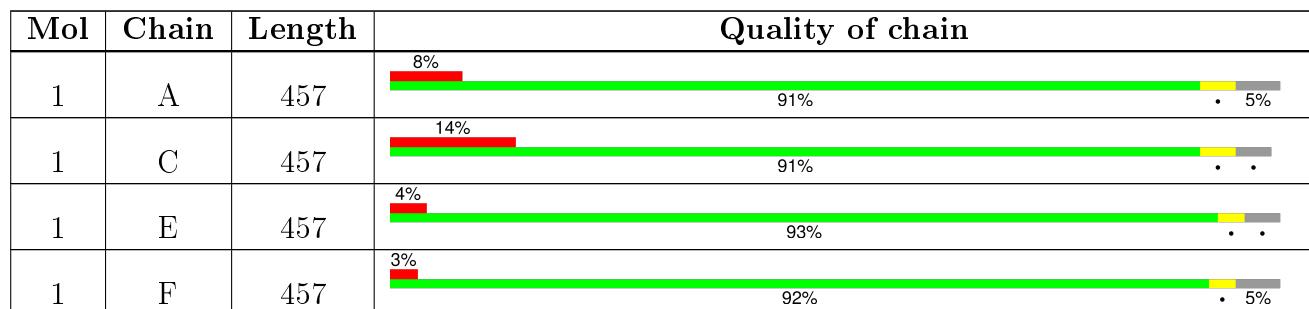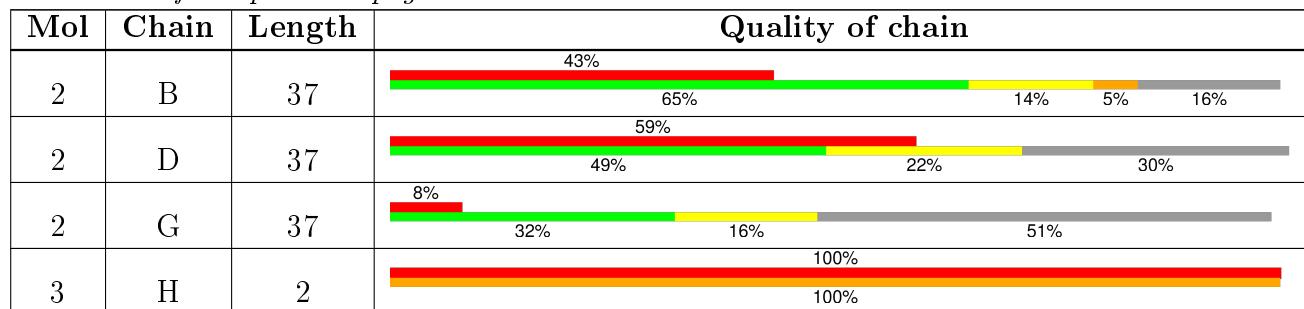
The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	APC	A	501	-	-	-	X
4	APC	F	501	-	-	-	X
5	GOL	A	502	-	-	-	X
5	GOL	A	503	-	-	-	X
5	GOL	A	504	-	-	-	X
5	GOL	H	101	-	-	-	X
6	PEG	C	501	-	-	-	X
6	PEG	E	502	-	-	-	X

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 31814 atoms, of which 15468 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 CCA-adding enzyme.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	435	Total	C	H	N	O	S	0	0	0
			7244	2325	3629	630	647	13			
1	C	437	Total	C	H	N	O	S	0	0	0
			7265	2333	3637	631	651	13			
1	E	437	Total	C	H	N	O	S	0	1	0
			7285	2338	3649	634	651	13			
1	F	436	Total	C	H	N	O	S	0	0	0
			7250	2328	3629	631	650	12			

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	438	SER	-	expression tag	UNP O28126
A	439	ASN	-	expression tag	UNP O28126
A	440	SER	-	expression tag	UNP O28126
A	441	SER	-	expression tag	UNP O28126
A	442	SER	-	expression tag	UNP O28126
A	443	VAL	-	expression tag	UNP O28126
A	444	ASP	-	expression tag	UNP O28126
A	445	LYS	-	expression tag	UNP O28126
A	446	LEU	-	expression tag	UNP O28126
A	447	ALA	-	expression tag	UNP O28126
A	448	ALA	-	expression tag	UNP O28126
A	449	ALA	-	expression tag	UNP O28126
A	450	LEU	-	expression tag	UNP O28126
A	451	GLU	-	expression tag	UNP O28126
A	452	HIS	-	expression tag	UNP O28126
A	453	HIS	-	expression tag	UNP O28126
A	454	HIS	-	expression tag	UNP O28126
A	455	HIS	-	expression tag	UNP O28126
A	456	HIS	-	expression tag	UNP O28126
A	457	HIS	-	expression tag	UNP O28126
C	438	SER	-	expression tag	UNP O28126

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	439	ASN	-	expression tag	UNP O28126
C	440	SER	-	expression tag	UNP O28126
C	441	SER	-	expression tag	UNP O28126
C	442	SER	-	expression tag	UNP O28126
C	443	VAL	-	expression tag	UNP O28126
C	444	ASP	-	expression tag	UNP O28126
C	445	LYS	-	expression tag	UNP O28126
C	446	LEU	-	expression tag	UNP O28126
C	447	ALA	-	expression tag	UNP O28126
C	448	ALA	-	expression tag	UNP O28126
C	449	ALA	-	expression tag	UNP O28126
C	450	LEU	-	expression tag	UNP O28126
C	451	GLU	-	expression tag	UNP O28126
C	452	HIS	-	expression tag	UNP O28126
C	453	HIS	-	expression tag	UNP O28126
C	454	HIS	-	expression tag	UNP O28126
C	455	HIS	-	expression tag	UNP O28126
C	456	HIS	-	expression tag	UNP O28126
C	457	HIS	-	expression tag	UNP O28126
E	438	SER	-	expression tag	UNP O28126
E	439	ASN	-	expression tag	UNP O28126
E	440	SER	-	expression tag	UNP O28126
E	441	SER	-	expression tag	UNP O28126
E	442	SER	-	expression tag	UNP O28126
E	443	VAL	-	expression tag	UNP O28126
E	444	ASP	-	expression tag	UNP O28126
E	445	LYS	-	expression tag	UNP O28126
E	446	LEU	-	expression tag	UNP O28126
E	447	ALA	-	expression tag	UNP O28126
E	448	ALA	-	expression tag	UNP O28126
E	449	ALA	-	expression tag	UNP O28126
E	450	LEU	-	expression tag	UNP O28126
E	451	GLU	-	expression tag	UNP O28126
E	452	HIS	-	expression tag	UNP O28126
E	453	HIS	-	expression tag	UNP O28126
E	454	HIS	-	expression tag	UNP O28126
E	455	HIS	-	expression tag	UNP O28126
E	456	HIS	-	expression tag	UNP O28126
E	457	HIS	-	expression tag	UNP O28126
F	438	SER	-	expression tag	UNP O28126
F	439	ASN	-	expression tag	UNP O28126
F	440	SER	-	expression tag	UNP O28126

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	441	SER	-	expression tag	UNP O28126
F	442	SER	-	expression tag	UNP O28126
F	443	VAL	-	expression tag	UNP O28126
F	444	ASP	-	expression tag	UNP O28126
F	445	LYS	-	expression tag	UNP O28126
F	446	LEU	-	expression tag	UNP O28126
F	447	ALA	-	expression tag	UNP O28126
F	448	ALA	-	expression tag	UNP O28126
F	449	ALA	-	expression tag	UNP O28126
F	450	LEU	-	expression tag	UNP O28126
F	451	GLU	-	expression tag	UNP O28126
F	452	HIS	-	expression tag	UNP O28126
F	453	HIS	-	expression tag	UNP O28126
F	454	HIS	-	expression tag	UNP O28126
F	455	HIS	-	expression tag	UNP O28126
F	456	HIS	-	expression tag	UNP O28126
F	457	HIS	-	expression tag	UNP O28126

- Molecule 2 is a RNA chain called human MenBeta minihelix.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	31	Total	C	H	N	O	P	0	0	0
			997	293	336	116	221	31			
2	D	26	Total	C	H	N	O	P	0	0	0
			836	245	282	96	187	26			
2	G	18	Total	C	H	N	O	P	0	0	0
			577	171	194	68	127	17			

- Molecule 3 is a RNA chain called RNA (5'-D(*GP*G)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	H	2	Total	C	H	N	O	P	0	0	0
			67	20	24	10	12	1			

- Molecule 4 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: APC) (formula: $C_{11}H_{18}N_5O_{12}P_3$).
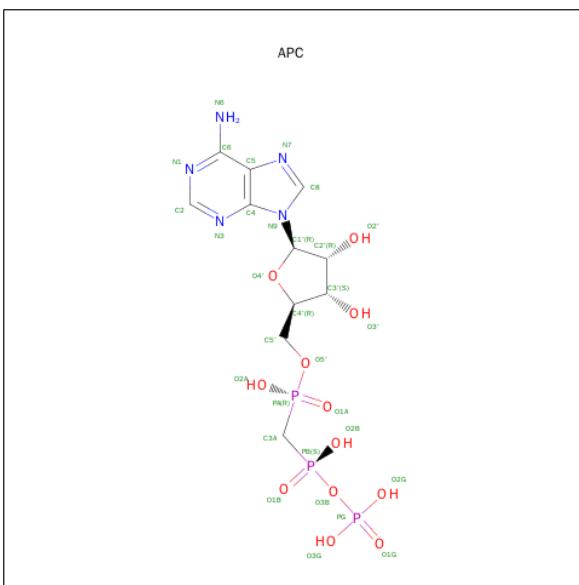
APC

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	P	0	0
			45	11	14	5	12	3		
4	F	1	Total	C	H	N	O	P	0	0
			45	11	14	5	12	3		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

GOL

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			14	3	8	3		
5	A	1	Total	C	H	O	0	0
			14	3	8	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total 14	C 3	H 8	O 3	0	0
5	E	1	Total 14	C 3	H 8	O 3	0	0
5	H	1	Total 14	C 3	H 8	O 3	0	0

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).
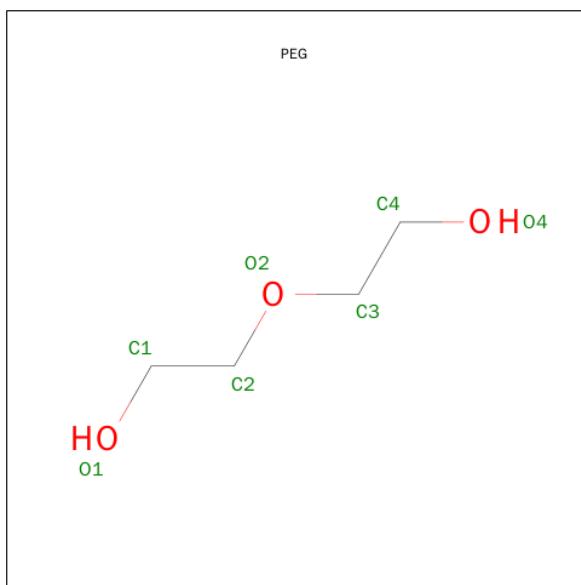
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	C	1	Total 17	C 4	H 10	O 3	0	0
6	E	1	Total 17	C 4	H 10	O 3	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	26	Total 26	O 26	0	0
7	B	1	Total 1	O 1	0	0
7	C	12	Total 12	O 12	0	0
7	E	29	Total 29	O 29	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	F	28	Total 28	O 28	0	0
7	G	3	Total 3	O 3	0	0

3 Residue-property plots ⓘ

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: CCA-adding enzyme
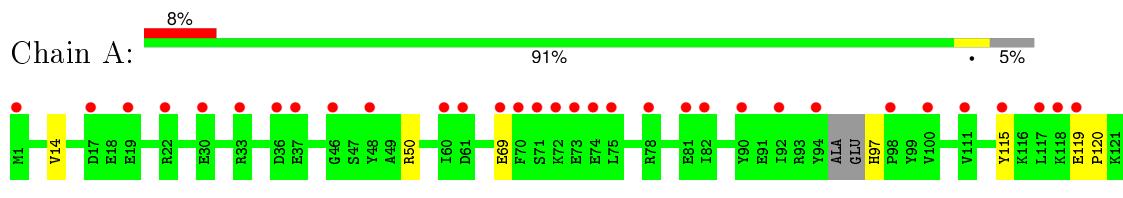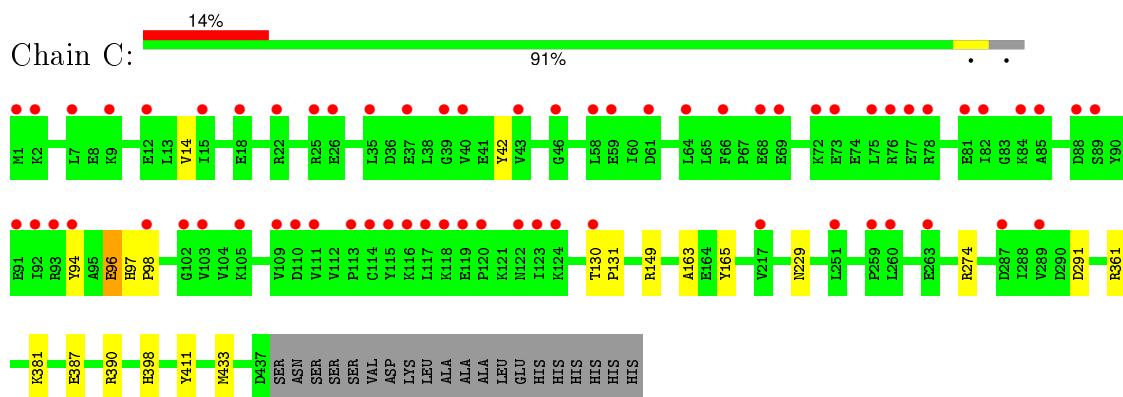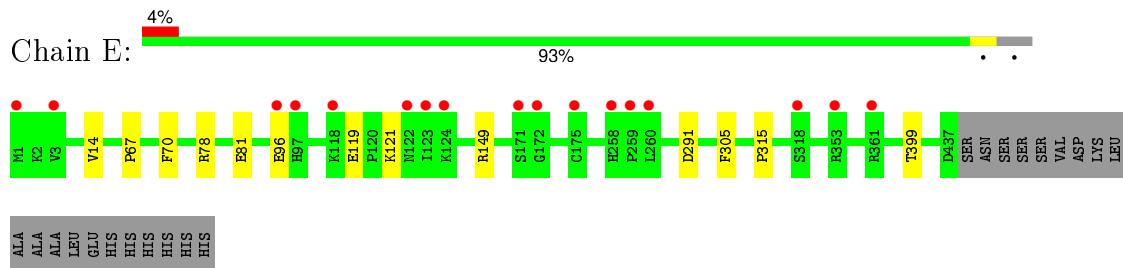
- Molecule 1: CCA-adding enzyme

- Molecule 1: CCA-adding enzyme

- Molecule 1: CCA-adding enzyme
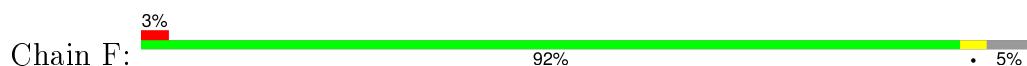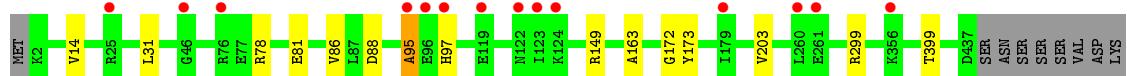
• Molecule 2: human MenBeta minihelix
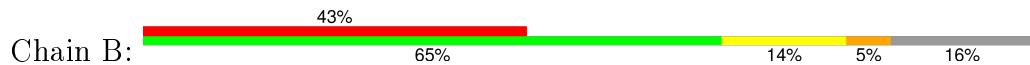

• Molecule 2: human MenBeta minihelix
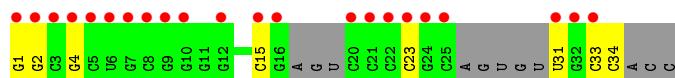
• Molecule 2: human MenBeta minihelix
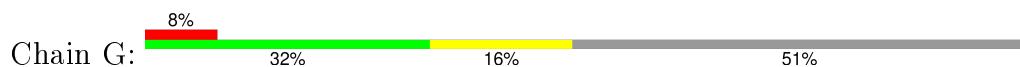
• Molecule 3: RNA (5'-D(*GP*G)-3')

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	114.99Å 84.05Å 135.52Å 90.00° 103.42° 90.00°	Depositor
Resolution (Å)	29.79 – 2.70	Depositor
	29.79 – 2.70	EDS
% Data completeness (in resolution range)	97.9 (29.79-2.70)	Depositor
	97.9 (29.79-2.70)	EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.06 (at 2.68Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.183 , 0.233	Depositor
	0.192 , 0.240	DCC
R_{free} test set	3411 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	65.2	Xtriage
Anisotropy	0.002	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 50.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$<|L|> = 0.49, <L^2> = 0.33$	Xtriage
Outliers	0 of 67888 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	31814	wwPDB-VP
Average B, all atoms (Å^2)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.35% 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, 0.375 respectively for untwinned datasets, and 0.333, 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, PEG, APC

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	0.31	0/3697	0.44	0/4965
1	C	0.29	0/3711	0.42	0/4985
1	E	0.34	0/3723	0.45	0/5002
1	F	0.33	0/3704	0.45	0/4977
2	B	0.44	1/735 (0.1%)	0.72	1/1140 (0.1%)
2	D	0.48	1/614 (0.2%)	0.72	1/949 (0.1%)
2	G	0.33	0/426	0.86	0/661
3	H	0.47	0/48	1.24	1/74 (1.4%)
All	All	0.33	2/16658 (0.0%)	0.49	3/22753 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	G	OP3-P	-10.56	1.48	1.61
2	D	1	G	OP3-P	-10.55	1.48	1.61

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	1	G	OP1-P-O3'	5.21	116.66	105.20
3	H	1	G	OP1-P-O3'	5.13	116.48	105.20
2	B	1	G	OP1-P-O3'	5.08	116.36	105.20

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	3615	3629	3621	12	0
1	C	3628	3637	3629	12	0
1	E	3636	3649	3640	6	0
1	F	3621	3629	3621	10	0
2	B	661	336	338	6	0
2	D	554	282	285	5	0
2	G	383	194	197	3	0
3	H	43	24	24	2	0
4	A	31	14	14	1	0
4	F	31	14	14	4	0
5	A	18	24	24	0	0
5	E	6	8	8	0	0
5	H	6	8	8	1	0
6	C	7	10	9	2	0
6	E	7	10	9	0	0
7	A	26	0	0	0	0
7	B	1	0	0	0	0
7	C	12	0	0	0	0
7	E	29	0	0	0	0
7	F	28	0	0	1	0
7	G	3	0	0	1	0
All	All	16346	15468	15441	50	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (50) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:299:ARG:NH1	1:F:399:THR:O	2.17	0.78
1:A:363:ARG:NH1	1:A:378:GLU:OE2	2.19	0.76
1:F:173:TYR:N	4:F:501:APC:O2A	2.18	0.74
2:B:34:C:OP2	2:B:37:C:N4	2.21	0.73
1:C:361:ARG:NH1	2:D:15:C:OP1	2.24	0.71
2:B:34:C:H3'	2:B:35:A:H5"	1.75	0.68

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:274:ARG:NH2	1:A:433:MET:O	2.27	0.68
2:B:17:A:O2'	2:B:19:U:OP2	2.17	0.62
1:E:291:ASP:OD1	2:G:31:U:O2'	2.20	0.60
1:C:387:GLU:OE2	1:C:390:ARG:NH1	2.36	0.58
2:G:25:C:OP2	7:G:103:HOH:O	2.17	0.58
1:E:119:GLU:OE1	1:E:121:LYS:NZ	2.35	0.58
1:C:274:ARG:NH2	1:C:433:MET:O	2.36	0.58
1:F:78:ARG:NH1	1:F:81:GLU:OE1	2.36	0.58
1:A:69:GLU:N	1:A:69:GLU:OE2	2.40	0.54
1:A:133:HIS:CE1	4:A:501:APC:H2	2.43	0.53
1:E:78:ARG:NH1	1:E:81:GLU:OE1	2.40	0.53
2:D:4:G:H1	2:D:31:U:H3	1.57	0.52
1:E:14:VAL:CG1	1:E:149:ARG:HB3	2.41	0.51
1:A:14:VAL:CG1	1:A:149:ARG:HB3	2.41	0.50
1:C:291:ASP:OD2	2:D:34:C:O2'	2.30	0.49
1:A:358:PHE:O	1:A:363:ARG:NH2	2.45	0.49
1:F:172:GLY:N	4:F:501:APC:O2A	2.46	0.48
1:F:163:ALA:N	4:F:501:APC:O3G	2.41	0.48
1:C:163:ALA:O	1:C:229:ASN:ND2	2.47	0.47
2:G:2:G:H1	2:G:30:U:H3	1.62	0.47
1:A:165:TYR:CE2	2:B:34:C:H1'	2.49	0.47
1:C:97:HIS:HB3	1:C:98:PRO:HD2	1.96	0.47
1:A:50:ARG:NH1	1:A:176:GLU:OE2	2.49	0.46
1:A:129:ARG:HH11	1:A:224:ARG:HD2	1.81	0.46
1:F:95:ALA:O	1:F:97:HIS:N	2.48	0.46
1:F:172:GLY:N	4:F:501:APC:O1B	2.34	0.46
1:C:381:LYS:H	6:C:501:PEG:H41	1.80	0.46
1:A:133:HIS:NE2	1:A:218:ASP:OD2	2.45	0.45
1:A:119:GLU:HB3	1:A:120:PRO:HD2	1.98	0.45
3:H:2:G:H22	5:H:101:GOL:C1	2.30	0.45
1:C:14:VAL:CG1	1:C:149:ARG:HB3	2.47	0.45
1:A:115:TYR:N	1:A:128:ASP:OD1	2.51	0.44
1:C:130:THR:N	1:C:131:PRO:HD2	2.32	0.43
2:B:34:C:C3'	2:B:35:A:H5"	2.48	0.42
3:H:1:G:H3'	3:H:2:G:C5'	2.50	0.42
1:F:31:LEU:HD12	1:F:86:VAL:CG1	2.49	0.41
1:C:398:HIS:HD2	2:D:23:C:H5"	1.85	0.41
1:F:203:VAL:HG23	7:F:626:HOH:O	2.20	0.41
1:E:305:PHE:CE2	1:E:315:PRO:HB2	2.56	0.41
2:B:34:C:H2'	2:B:35:A:C4'	2.50	0.41
1:F:14:VAL:CG1	1:F:149:ARG:HB3	2.51	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:67:PRO:HG2	1:E:70:PHE:CE1	2.56	0.41
1:C:165:TYR:OH	2:D:33:C:O2	2.35	0.41
1:C:381:LYS:H	6:C:501:PEG:C4	2.34	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	431/457 (94%)	409 (95%)	22 (5%)	0	100	100
1	C	435/457 (95%)	411 (94%)	23 (5%)	1 (0%)	52	80
1	E	436/457 (95%)	424 (97%)	11 (2%)	1 (0%)	52	80
1	F	434/457 (95%)	422 (97%)	11 (2%)	1 (0%)	52	80
All	All	1736/1828 (95%)	1666 (96%)	67 (4%)	3 (0%)	52	80

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	96	GLU
1	F	95	ALA
1	E	96	GLU

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	386/404 (96%)	383 (99%)	3 (1%)	86	96
1	C	386/404 (96%)	382 (99%)	4 (1%)	82	94
1	E	388/404 (96%)	387 (100%)	1 (0%)	94	99
1	F	386/404 (96%)	385 (100%)	1 (0%)	94	99
All	All	1546/1616 (96%)	1537 (99%)	9 (1%)	90	97

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	97	HIS
1	A	129	ARG
1	A	362	ASN
1	C	42	TYR
1	C	94	TYR
1	C	96	GLU
1	C	411	TYR
1	E	399	THR
1	F	88	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	97	HIS

5.3.3   RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	29/37 (78%)	3 (10%)	0
2	D	23/37 (62%)	1 (4%)	0
2	G	16/37 (43%)	2 (12%)	0
3	H	1/2 (50%)	1 (100%)	0
All	All	69/113 (61%)	7 (10%)	0

All (7) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	35	A
2	B	36	C
2	B	37	C

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	2	G
2	G	5	C
2	G	28	U
3	H	2	G

There are no RNA pucker outliers to report.

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)

9 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 Counts	Bond lengths RMSZ	Bond lengths #$|Z| > 2$	Bond angles Counts	Bond angles RMSZ	Bond angles #$|Z| > 2$
4	APC	A	501	-	25,33,33	2.25	6 (24%)	30,52,52	2.45	5 (16%)
5	GOL	A	502	-	5,5,5	0.35	0	5,5,5	0.28	0
5	GOL	A	503	-	5,5,5	0.31	0	5,5,5	0.27	0
5	GOL	A	504	-	5,5,5	0.33	0	5,5,5	0.40	0
6	PEG	C	501	-	6,6,6	0.60	0	5,5,5	0.34	0
5	GOL	E	501	-	5,5,5	0.33	0	5,5,5	0.32	0
6	PEG	E	502	-	6,6,6	0.64	0	5,5,5	0.40	0
4	APC	F	501	-	25,33,33	2.28	8 (32%)	30,52,52	3.00	8 (26%)
5	GOL	H	101	-	5,5,5	0.33	0	5,5,5	0.29	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	APC	A	501	-	-	0/15/38/38	0/3/3/3
5	GOL	A	502	-	-	0/4/4/4	0/0/0/0
5	GOL	A	503	-	-	0/4/4/4	0/0/0/0
5	GOL	A	504	-	-	0/4/4/4	0/0/0/0
6	PEG	C	501	-	-	0/4/4/4	0/0/0/0
5	GOL	E	501	-	-	0/4/4/4	0/0/0/0
6	PEG	E	502	-	-	0/4/4/4	0/0/0/0
4	APC	F	501	-	-	0/15/38/38	0/3/3/3
5	GOL	H	101	-	-	0/4/4/4	0/0/0/0

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	501	APC	C2'-C3'	-3.40	1.44	1.53
4	A	501	APC	C2'-C3'	-3.20	1.44	1.53
4	F	501	APC	PA-O2A	-3.03	1.49	1.56
4	A	501	APC	PB-O2B	-2.32	1.50	1.56
4	F	501	APC	C5-N7	-2.08	1.32	1.39
4	F	501	APC	PB-O3B	-2.06	1.56	1.58
4	F	501	APC	C6-N6	2.03	1.41	1.34
4	A	501	APC	C6-N6	2.07	1.41	1.34
4	F	501	APC	C2-N3	2.62	1.36	1.32
4	A	501	APC	O4'-C4'	2.69	1.51	1.45
4	A	501	APC	C2-N3	2.69	1.37	1.32
4	F	501	APC	O4'-C4'	2.97	1.51	1.45
4	F	501	APC	O4'-C1'	8.02	1.51	1.41
4	A	501	APC	O4'-C1'	8.59	1.52	1.41

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	501	APC	N3-C2-N1	-10.92	120.54	128.89
4	A	501	APC	N3-C2-N1	-10.55	120.82	128.89
4	F	501	APC	C4'-O4'-C1'	-8.55	100.32	109.72
4	A	501	APC	C4'-O4'-C1'	-5.60	103.56	109.72
4	F	501	APC	C2'-C1'-N9	-3.24	109.34	114.29
4	A	501	APC	C2'-C1'-N9	-3.14	109.50	114.29
4	F	501	APC	PG-O3B-PB	-2.84	123.15	132.67
4	A	501	APC	C4-C5-N7	-2.70	107.00	109.48

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	F	501	APC	C4-C5-N7	-2.57	107.11	109.48
4	A	501	APC	O5'-PA-C3A	2.03	110.10	104.42
4	F	501	APC	C2'-C3'-C4'	2.31	107.37	102.61
4	F	501	APC	O4'-C1'-N9	3.03	114.44	108.10
4	F	501	APC	O5'-PA-C3A	3.45	114.07	104.42

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	501	APC	1	0
6	C	501	PEG	2	0
4	F	501	APC	4	0
5	H	101	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, 95th 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	435/457 (95%)	0.46	37 (8%)	13	10	34, 70, 159, 219	0
1	C	437/457 (95%)	0.82	65 (14%)	3	2	36, 78, 180, 280	0
1	E	437/457 (95%)	0.22	17 (3%)	43	43	35, 58, 104, 160	0
1	F	436/457 (95%)	0.21	14 (3%)	51	51	40, 60, 102, 160	0
2	B	31/37 (83%)	2.56	16 (51%)	0	0	113, 146, 280, 423	0
2	D	26/37 (70%)	3.12	22 (84%)	0	0	143, 171, 199, 254	0
2	G	18/37 (48%)	0.84	3 (16%)	2	2	75, 88, 157, 159	0
3	H	2/2 (100%)	3.32	2 (100%)	0	0	148, 148, 148, 161	0
All	All	1822/1941 (93%)	0.51	176 (9%)	10	7	34, 66, 162, 423	0

All (176) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	94	TYR	10.4
1	C	92	ILE	10.2
1	C	115	TYR	8.8
1	C	73	GLU	8.2
1	C	68	GLU	8.1
1	C	1	MET	7.6
1	C	77	GLU	7.4
2	B	37	C	7.2
1	A	1	MET	6.9
2	B	7	G	6.8
1	A	78	ARG	6.7
2	B	35	A	6.0
2	B	8	C	5.8
2	B	36	C	5.3
1	E	124	LYS	5.3
2	D	7	G	5.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	9	G	5.1
1	C	122	ASN	5.1
1	E	123	ILE	5.0
2	B	18	G	4.8
1	C	93	ARG	4.7
2	D	20	C	4.7
1	C	81	GLU	4.7
2	D	25	C	4.6
1	C	84	LYS	4.6
2	G	9	G	4.5
1	F	124	LYS	4.4
2	D	10	G	4.4
1	A	61	ASP	4.3
1	C	76	ARG	4.3
1	A	72	LYS	4.3
1	F	260	LEU	4.2
2	D	6	U	4.2
2	B	19	U	4.1
1	C	22	ARG	4.1
1	A	82	ILE	4.0
1	A	74	GLU	4.0
2	B	6	U	4.0
2	D	16	G	3.9
1	F	123	ILE	3.9
1	A	81	GLU	3.8
2	D	24	G	3.8
1	A	98	PRO	3.8
2	B	31	U	3.8
2	D	4	G	3.8
1	C	105	LYS	3.7
1	E	259	PRO	3.7
1	E	258	HIS	3.7
1	E	175	CYS	3.7
1	F	261	GLU	3.6
1	A	33	ARG	3.6
3	H	1	G	3.6
1	C	123	ILE	3.6
1	C	78	ARG	3.6
1	C	72	LYS	3.6
1	E	3	VAL	3.6
1	A	75	LEU	3.5
2	D	21	C	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	94	TYR	3.5
1	C	114	CYS	3.5
2	D	2	G	3.5
1	C	120	PRO	3.4
1	C	75	LEU	3.4
1	C	117	LEU	3.4
2	G	1	G	3.3
1	C	61	ASP	3.3
2	D	31	U	3.2
1	C	260	LEU	3.2
2	B	9	G	3.2
1	C	39	GLY	3.2
2	D	1	G	3.1
1	C	85	ALA	3.1
1	A	70	PHE	3.1
1	C	26	GLU	3.0
1	A	73	GLU	3.0
3	H	2	G	3.0
1	A	122	ASN	3.0
1	C	259	PRO	3.0
1	C	98	PRO	3.0
1	C	119	GLU	2.9
1	C	12	GLU	2.9
1	F	96	GLU	2.9
1	C	116	LYS	2.9
1	C	69	GLU	2.9
2	D	33	C	2.9
1	F	119	GLU	2.9
1	E	97	HIS	2.8
1	E	361	ARG	2.8
1	A	100	VAL	2.8
2	D	32	G	2.8
1	C	2	LYS	2.8
1	A	69	GLU	2.8
2	D	15	C	2.8
1	F	122	ASN	2.8
1	C	66	PHE	2.7
1	A	117	LEU	2.7
1	C	109	VAL	2.7
2	D	8	C	2.7
1	C	58	LEU	2.7
1	C	59	GLU	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	289	VAL	2.7
1	E	122	ASN	2.7
1	C	91	GLU	2.6
1	C	15	ILE	2.6
1	C	263	GLU	2.6
1	F	356	LYS	2.6
1	A	17	ASP	2.6
1	A	90	TYR	2.6
1	A	46	GLY	2.6
2	D	3	C	2.6
1	E	353	ARG	2.5
1	F	25	ARG	2.5
1	A	172	GLY	2.5
1	C	287	ASP	2.5
2	B	11	G	2.5
1	C	102	GLY	2.5
1	C	130	THR	2.5
1	A	22	ARG	2.5
1	A	60	ILE	2.4
1	A	119	GLU	2.4
1	C	18	GLU	2.4
1	C	89	SER	2.4
2	G	23	C	2.4
1	C	82	ILE	2.4
2	B	1	G	2.4
2	D	22	C	2.4
2	B	10	G	2.4
1	C	103	VAL	2.4
2	D	5	C	2.4
1	A	259	PRO	2.4
2	B	17	A	2.4
1	C	64	LEU	2.3
1	C	113	PRO	2.3
2	B	5	C	2.3
1	A	115	TYR	2.3
1	A	71	SER	2.3
1	A	37	GLU	2.3
1	A	324	GLU	2.3
2	B	34	C	2.3
1	C	9	LYS	2.3
1	C	40	VAL	2.3
1	C	46	GLY	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	92	ILE	2.3
1	C	25	ARG	2.2
1	A	48	TYR	2.2
1	C	35	LEU	2.2
1	E	118	LYS	2.2
1	C	43	VAL	2.2
1	C	7	LEU	2.2
1	E	96	GLU	2.2
1	E	260	LEU	2.2
1	F	76	ARG	2.2
1	F	97	HIS	2.2
1	A	223	LYS	2.2
1	A	111	VAL	2.2
1	E	318	SER	2.2
1	F	95	ALA	2.2
1	A	36	ASP	2.2
1	C	217	VAL	2.2
1	C	124	LYS	2.1
1	E	1	MET	2.1
1	C	37	GLU	2.1
1	C	110	ASP	2.1
1	A	19	GLU	2.1
1	F	179	ILE	2.1
1	E	172	GLY	2.1
1	E	171	SER	2.1
1	C	111	VAL	2.0
2	D	23	C	2.0
1	C	118	LYS	2.0
2	D	12	G	2.0
1	F	46	GLY	2.0
1	A	30	GLU	2.0
1	A	118	LYS	2.0
1	C	251	LEU	2.0
1	C	88	ASP	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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. 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, 95th 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(Å^2)	Q<0.9
5	GOL	A	503	6/6	0.71	0.51	7.08	84,101,113,116	0
5	GOL	A	504	6/6	0.77	0.57	6.75	87,106,117,128	0
4	APC	F	501	31/31	0.62	0.63	6.69	46,70,84,91	45
6	PEG	C	501	7/7	0.83	0.37	5.50	89,107,124,127	0
5	GOL	H	101	6/6	0.67	0.50	5.06	64,77,82,85	14
4	APC	A	501	31/31	0.49	0.69	4.09	51,73,86,92	45
6	PEG	E	502	7/7	0.90	0.28	3.87	74,90,93,95	0
5	GOL	A	502	6/6	0.86	0.26	2.23	82,98,111,114	0
5	GOL	E	501	6/6	0.87	0.23	0.85	80,97,114,116	0

6.5 Other polymers (i)

There are no such residues in this entry.