Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2024 – 01:25 am GMT

PDB ID	:	8PYR
Title	:	Crystal structure of the dual T-loop phosphorylated Cdk7/CycH/Mat1 complex
Authors	:	Anand, K.; Duster, R.; Geyer, M.
Deposited on	:	2023-07-26
Resolution	:	2.15 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 ⓘ 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 ⓘ) were used in the production of this report:

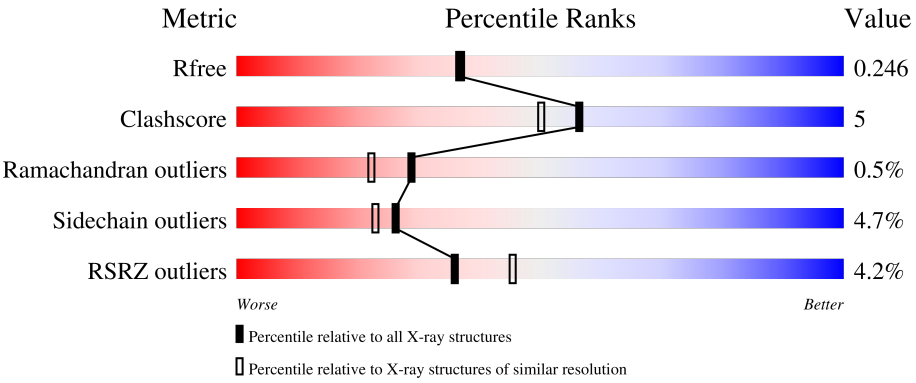
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.15 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.



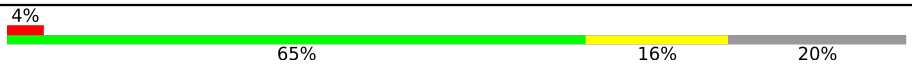
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	346	3% 65% 12% 23%
1	E	346	2% 62% 12% · 25%
2	B	323	5% 76% 12% · 11%
2	F	323	6% 73% 15% · 11%
3	C	82	% 63% 12% · 21%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	82	
4	D	115	
4	H	115	

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
5	EDO	E	401	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 12032 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 Cyclin-dependent kinase 7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	266	Total	C	N	O	P	S	0	0	0
			2104	1356	352	383	2	11			
1	E	260	Total	C	N	O	P	S	0	0	0
			2072	1337	348	374	2	11			

- Molecule 2 is a protein called Cyclin-H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	287	Total	C	N	O	S	4	1	0
			2346	1502	402	423	19			
2	F	287	Total	C	N	O	S	0	0	0
			2338	1497	401	422	18			

- Molecule 3 is a protein called CDK-activating kinase assembly factor MAT1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	65	Total	C	N	O	S	0	0	0
			512	329	87	94	2			
3	G	66	Total	C	N	O	S	5	0	0
			518	332	88	96	2			

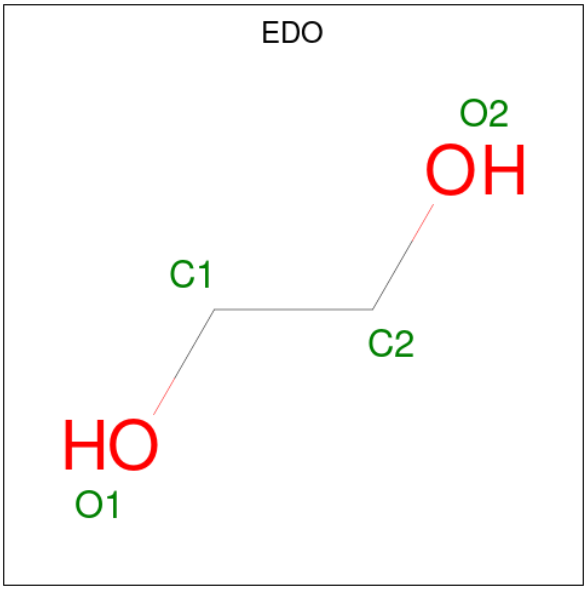
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	228	GLY	-	expression tag	UNP P51948
C	229	SER	-	expression tag	UNP P51948
G	228	GLY	-	expression tag	UNP P51948
G	229	SER	-	expression tag	UNP P51948

- Molecule 4 is a protein called Nanobody (VHH-RD7-04).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	114	Total	C	N	O	S	0	0	0
			867	541	152	170	4			
4	H	114	Total	C	N	O	S	0	0	0
			867	541	152	170	4			

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		
5	E	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	94	Total	O	0	0
			94	94		
6	B	51	Total	O	0	0
			51	51		

Continued on next page...

Continued from previous page...

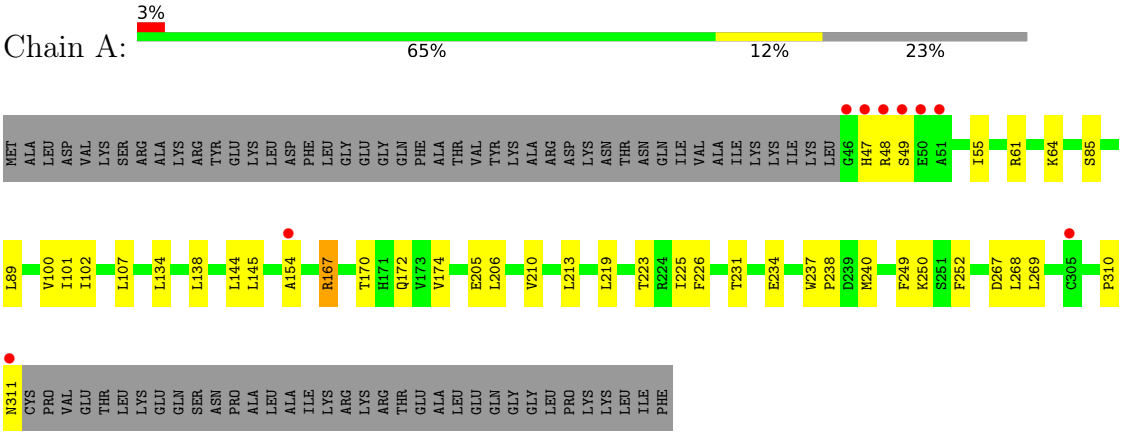
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	15	Total O 15 15	0	0
6	D	24	Total O 24 24	0	0
6	E	94	Total O 94 94	0	0
6	F	60	Total O 60 60	0	0
6	G	21	Total O 21 21	0	0
6	H	25	Total O 25 25	0	0

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.

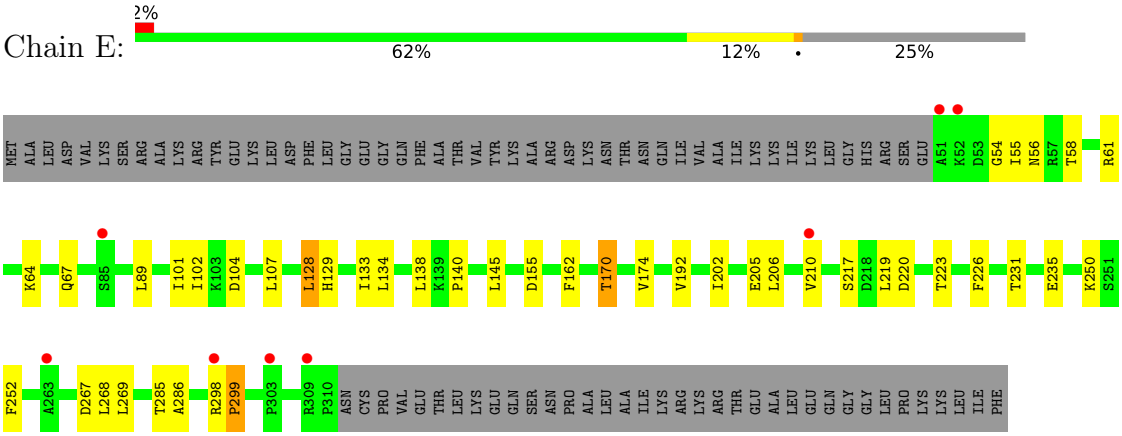
• Molecule 1: Cyclin-dependent kinase 7

Chain A:



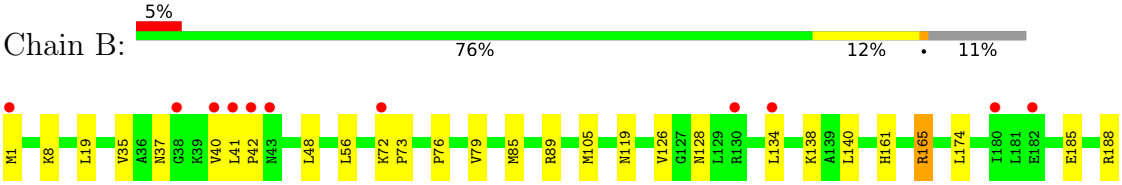
• Molecule 1: Cyclin-dependent kinase 7

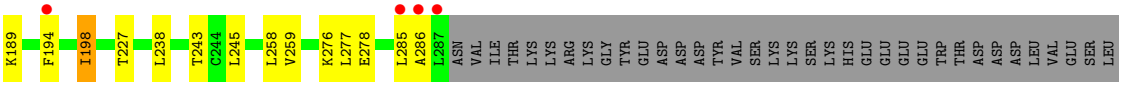
Chain E:



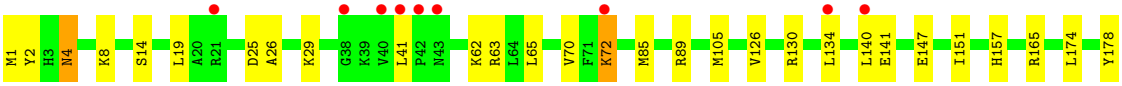
• Molecule 2: Cyclin-H

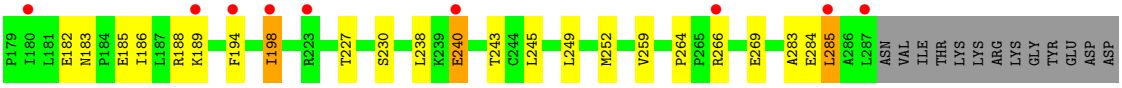
Chain B:





• Molecule 2: Cyclin-H

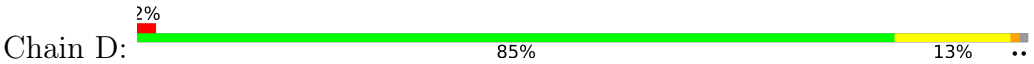


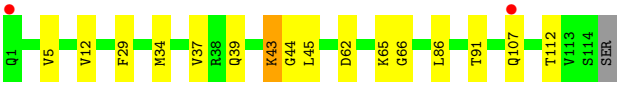


• Molecule 3: CDK-activating kinase assembly factor MAT1

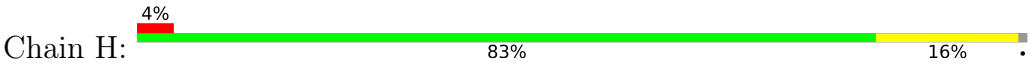
• Molecule 3: CDK-activating kinase assembly factor MAT1

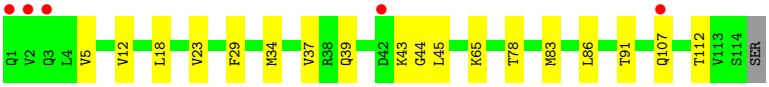
• Molecule 4: Nanobody (VHH-RD7-04)





• Molecule 4: Nanobody (VHH-RD7-04)





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	119.86Å 77.87Å 121.87Å 90.00° 119.31° 90.00°	Depositor
Resolution (Å)	48.05 – 2.15 47.49 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.0 (48.05-2.15) 99.0 (47.49-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.19 (at 2.16Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.220 , 0.241 0.222 , 0.246	Depositor DCC
R_{free} test set	2014 reflections (1.91%)	wwPDB-VP
Wilson B-factor (Å^2)	48.9	Xtriage
Anisotropy	0.278	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 40.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.000 for -h-l,k,h 0.000 for l,k,-h-l 0.009 for h,-k,-h-l 0.001 for -h-l,-k,l 0.004 for l,-k,h	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	12032	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 4.08% 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: TPO, EDO, SEP

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	0.25	0/2136	0.47	0/2901
1	E	0.24	0/2104	0.47	0/2857
2	B	0.24	0/2396	0.46	0/3234
2	F	0.25	0/2388	0.46	0/3224
3	C	0.25	0/528	0.47	0/720
3	G	0.25	0/534	0.48	0/728
4	D	0.30	0/883	0.55	0/1194
4	H	0.27	0/883	0.52	0/1194
All	All	0.25	0/11852	0.48	0/16052

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	B	0	1
2	F	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	165	ARG	Sidechain
2	F	165	ARG	Sidechain

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	2104	0	2084	20	0
1	E	2072	0	2069	23	0
2	B	2346	0	2385	19	0
2	F	2338	0	2375	30	0
3	C	512	0	485	7	0
3	G	518	0	490	6	0
4	D	867	0	849	9	0
4	H	867	0	849	13	0
5	A	8	0	12	0	0
5	B	4	0	6	0	0
5	E	12	0	18	4	0
6	A	94	0	0	1	0
6	B	51	0	0	0	0
6	C	15	0	0	0	0
6	D	24	0	0	0	0
6	E	94	0	0	0	0
6	F	60	0	0	0	0
6	G	21	0	0	0	0
6	H	25	0	0	0	0
All	All	12032	0	11622	117	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 5.

All (117) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:226:PHE:HB3	1:E:231:THR:HG22	1.65	0.77
1:E:192:VAL:HG11	5:E:401:EDO:H12	1.69	0.74
1:A:226:PHE:HB3	1:A:231:THR:HG22	1.73	0.70
1:A:167:ARG:NH2	6:A:501:HOH:O	2.28	0.67
1:A:101:ILE:HG12	1:A:310:PRO:HG3	1.77	0.67
2:B:72:LYS:HB2	2:B:73:PRO:HD3	1.78	0.65
1:E:61:ARG:HA	1:E:64:LYS:HE2	1.79	0.64
1:A:61:ARG:HA	1:A:64:LYS:HE2	1.79	0.64

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:43:LYS:HE2	4:D:44:GLY:H	1.63	0.63
2:F:4:ASN:H	2:F:4:ASN:HD22	1.46	0.62
3:G:264:GLU:HA	4:H:65:LYS:HG2	1.83	0.60
2:B:48:LEU:HD23	2:B:276:LYS:HE3	1.84	0.59
3:G:275:VAL:HG21	3:G:293:CYS:HB2	1.87	0.57
4:H:43:LYS:HD3	4:H:44:GLY:H	1.69	0.57
4:H:12:VAL:HG11	4:H:86:LEU:HD13	1.87	0.56
1:A:237:TRP:CG	1:A:240:MET:HB2	2.40	0.56
4:D:39:GLN:HB2	4:D:45:LEU:HD23	1.88	0.56
2:B:41:LEU:H	2:B:41:LEU:HD23	1.70	0.55
2:B:185:GLU:OE2	2:B:188:ARG:NE	2.28	0.55
4:H:39:GLN:HB2	4:H:45:LEU:HD23	1.88	0.55
1:E:285:THR:HB	5:E:401:EDO:H22	1.88	0.55
2:F:178:TYR:HH	2:F:230:SER:HG	1.53	0.54
4:D:12:VAL:HG11	4:D:86:LEU:HD13	1.89	0.53
1:A:47:HIS:O	2:B:119:ASN:HB3	2.08	0.53
2:B:238:LEU:HB3	2:B:245:LEU:HB2	1.90	0.52
4:H:5:VAL:HA	4:H:107:GLN:HE22	1.74	0.52
2:B:85:MET:HE3	2:B:89:ARG:HE	1.75	0.51
4:D:5:VAL:HA	4:D:107:GLN:HE22	1.75	0.51
1:A:231:THR:HG23	1:A:252:PHE:H	1.75	0.51
2:B:126:VAL:HG11	2:B:140:LEU:HB2	1.91	0.51
1:E:231:THR:HG23	1:E:252:PHE:H	1.75	0.51
1:E:56:ASN:HD21	1:E:58:THR:HB	1.76	0.51
4:H:23:VAL:HG22	4:H:78:THR:HG22	1.94	0.50
4:H:91:THR:HG23	4:H:112:THR:HA	1.93	0.50
1:A:213:LEU:HD13	1:A:225:ILE:HG12	1.92	0.50
3:G:258:PRO:HG2	3:G:301:PHE:CG	2.47	0.50
2:B:8:LYS:HE2	3:C:306:TRP:CE2	2.47	0.50
2:B:35:VAL:HG21	2:B:42:PRO:HA	1.94	0.49
1:E:298:ARG:HB3	1:E:299:PRO:HD3	1.94	0.49
2:B:1[B]:MET:HG3	2:B:161:HIS:NE2	2.26	0.49
3:C:275:VAL:HG21	3:C:293:CYS:HB2	1.94	0.49
2:F:249:LEU:HA	2:F:252:MET:HE2	1.93	0.49
2:F:266:ARG:HB2	2:F:269:GLU:HG2	1.94	0.49
1:E:102:ILE:HD13	1:E:205:GLU:HG2	1.95	0.48
1:A:234:GLU:HG2	1:A:238:PRO:HA	1.96	0.48
2:B:198:ILE:HG22	2:B:259:VAL:HG21	1.96	0.48
2:B:56:LEU:HD13	2:B:277:LEU:HD22	1.96	0.48
3:C:257:GLY:HA3	3:C:301:PHE:O	2.14	0.47
1:E:129:HIS:ND1	5:E:401:EDO:H21	2.29	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:62:LYS:HZ1	2:F:264:PRO:HB3	1.79	0.47
2:B:165:ARG:HD2	3:C:299:ASP:OD2	2.14	0.47
1:E:55:ILE:HD12	2:F:151:ILE:HD12	1.95	0.47
1:E:67:GLN:HG2	2:F:157:HIS:HD2	1.79	0.47
1:A:231:THR:HG21	1:A:250:LYS:O	2.15	0.47
2:F:41:LEU:H	2:F:41:LEU:HD23	1.80	0.47
2:F:72:LYS:HA	2:F:72:LYS:HD3	1.69	0.47
1:E:101:ILE:HG23	1:E:107:LEU:HD23	1.95	0.47
2:F:194:PHE:O	2:F:198:ILE:HG23	2.15	0.47
2:F:198:ILE:HG22	2:F:259:VAL:HG21	1.97	0.47
4:H:37:VAL:HG12	4:H:45:LEU:HD22	1.97	0.47
4:D:37:VAL:HG12	4:D:45:LEU:HD22	1.96	0.47
2:F:238:LEU:HB3	2:F:245:LEU:HB2	1.97	0.47
3:C:268:ARG:NH1	4:D:66:GLY:O	2.47	0.46
4:D:62:ASP:HA	4:D:65:LYS:HE3	1.98	0.46
4:H:43:LYS:HD3	4:H:44:GLY:N	2.30	0.46
4:D:91:THR:HG23	4:D:112:THR:HA	1.97	0.46
2:F:4:ASN:HD22	2:F:4:ASN:N	2.08	0.46
1:A:219:LEU:O	1:A:223:THR:HG23	2.16	0.46
4:H:83:MET:HB3	4:H:86:LEU:HD21	1.98	0.46
2:F:126:VAL:HG11	2:F:140:LEU:HB2	1.97	0.45
1:E:231:THR:HG21	1:E:250:LYS:O	2.17	0.45
1:A:107:LEU:HD12	1:A:107:LEU:HA	1.84	0.45
2:B:76:PRO:HG2	2:B:79:VAL:HG23	1.99	0.45
3:C:297:LEU:HD12	3:C:297:LEU:HA	1.86	0.45
2:F:25:ASP:O	2:F:29:LYS:HG3	2.16	0.45
1:E:128:LEU:HD22	1:E:133:ILE:HB	1.99	0.45
2:F:185:GLU:OE2	2:F:188:ARG:NE	2.32	0.45
4:H:45:LEU:HD23	4:H:45:LEU:HA	1.81	0.45
1:A:101:ILE:HG23	1:A:107:LEU:HD23	1.98	0.44
2:F:240:GLU:H	2:F:240:GLU:HG2	1.50	0.44
1:E:162:PHE:O	3:G:295:ARG:NH1	2.50	0.44
3:C:290:SER:O	3:C:294:HIS:HB2	2.18	0.44
2:F:26:ALA:HA	2:F:29:LYS:HD2	1.99	0.44
1:A:144:LEU:HG	1:A:154:ALA:HB2	2.00	0.44
1:E:104:ASP:HB3	1:E:107:LEU:HB2	2.00	0.44
1:E:286:ALA:H	5:E:401:EDO:H22	1.83	0.44
2:B:194:PHE:O	2:B:198:ILE:HG23	2.18	0.43
1:E:217:SER:H	1:E:220:ASP:HB3	1.83	0.43
2:F:8:LYS:HE2	3:G:306:TRP:CE2	2.53	0.43
1:E:54:GLY:HA3	2:F:147:GLU:OE1	2.19	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:72:LYS:HB2	2:B:72:LYS:HE3	1.82	0.43
4:H:29:PHE:CE1	4:H:34:MET:HG3	2.53	0.43
2:F:4:ASN:H	2:F:4:ASN:ND2	2.13	0.43
1:E:170:TPO:H	1:E:170:TPO:P	2.42	0.43
1:A:223:THR:HG22	1:A:249:PHE:HE2	1.84	0.43
2:B:37:ASN:ND2	2:B:286:ALA:O	2.52	0.43
2:F:283:ALA:O	2:F:285:LEU:N	2.51	0.43
2:F:70:VAL:HG11	2:F:130:ARG:NH2	2.34	0.42
2:F:183:ASN:O	2:F:186:ILE:HG12	2.18	0.42
2:B:72:LYS:N	2:B:128:ASN:OD1	2.33	0.42
1:E:219:LEU:O	1:E:223:THR:HG23	2.20	0.42
4:H:18:LEU:HB2	4:H:86:LEU:HD11	2.02	0.42
1:A:102:ILE:HD13	1:A:205:GLU:HG2	2.02	0.42
4:D:29:PHE:CE1	4:D:34:MET:HG3	2.55	0.41
2:F:2:TYR:CZ	2:F:8:LYS:HB2	2.56	0.41
2:F:182:GLU:H	2:F:182:GLU:CD	2.20	0.41
3:G:306:TRP:CZ3	3:G:308:PRO:HA	2.56	0.41
1:A:231:THR:CG2	1:A:252:PHE:H	2.34	0.41
1:E:267:ASP:OD1	1:E:268:LEU:N	2.53	0.41
2:F:85:MET:HE2	2:F:89:ARG:HH21	1.85	0.41
2:F:4:ASN:N	2:F:4:ASN:ND2	2.69	0.41
1:E:140:PRO:HD3	1:E:202:ILE:HG12	2.03	0.41
1:A:237:TRP:CD2	1:A:240:MET:HB2	2.56	0.41
1:A:267:ASP:OD1	1:A:268:LEU:N	2.54	0.40
1:A:55:ILE:HG12	1:A:85:SER:HA	2.03	0.40
2:F:14:SER:H	2:F:14:SER:HG	1.66	0.40
2:F:1:MET:O	2:F:4:ASN:ND2	2.54	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	262/346 (76%)	250 (95%)	9 (3%)	3 (1%)	14 8
1	E	256/346 (74%)	245 (96%)	8 (3%)	3 (1%)	13 7
2	B	285/323 (88%)	281 (99%)	4 (1%)	0	100 100
2	F	285/323 (88%)	280 (98%)	4 (1%)	1 (0%)	34 29
3	C	63/82 (77%)	63 (100%)	0	0	100 100
3	G	64/82 (78%)	64 (100%)	0	0	100 100
4	D	112/115 (97%)	111 (99%)	1 (1%)	0	100 100
4	H	112/115 (97%)	111 (99%)	1 (1%)	0	100 100
All	All	1439/1732 (83%)	1405 (98%)	27 (2%)	7 (0%)	29 22

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	48	ARG
1	A	49	SER
1	E	155	ASP
1	A	174	VAL
1	E	174	VAL
1	E	299	PRO
2	F	284	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	225/297 (76%)	214 (95%)	11 (5%)	25 21
1	E	224/297 (75%)	215 (96%)	9 (4%)	31 29
2	B	262/296 (88%)	249 (95%)	13 (5%)	24 21
2	F	261/296 (88%)	246 (94%)	15 (6%)	20 16
3	C	52/66 (79%)	46 (88%)	6 (12%)	5 2
3	G	53/66 (80%)	48 (91%)	5 (9%)	8 4
4	D	94/95 (99%)	93 (99%)	1 (1%)	73 78

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	H	94/95 (99%)	94 (100%)	0	100 100
All	All	1265/1508 (84%)	1205 (95%)	60 (5%)	26 23

All (60) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	89	LEU
1	A	100	VAL
1	A	134	LEU
1	A	138	LEU
1	A	145	LEU
1	A	167	ARG
1	A	172	GLN
1	A	206	LEU
1	A	210	VAL
1	A	269	LEU
1	A	311	ASN
2	B	19	LEU
2	B	40	VAL
2	B	105	MET
2	B	134	LEU
2	B	138	LYS
2	B	174	LEU
2	B	189	LYS
2	B	198	ILE
2	B	227	THR
2	B	243	THR
2	B	258	LEU
2	B	278	GLU
2	B	285	LEU
3	C	264	GLU
3	C	268	ARG
3	C	283	LEU
3	C	294	HIS
3	C	297	LEU
3	C	304	LEU
4	D	43	LYS
1	E	89	LEU
1	E	128	LEU
1	E	134	LEU
1	E	138	LEU
1	E	145	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	E	206	LEU
1	E	210	VAL
1	E	235	GLU
1	E	269	LEU
2	F	4	ASN
2	F	19	LEU
2	F	63	ARG
2	F	65	LEU
2	F	72	LYS
2	F	105	MET
2	F	134	LEU
2	F	141	GLU
2	F	174	LEU
2	F	189	LYS
2	F	198	ILE
2	F	227	THR
2	F	240	GLU
2	F	243	THR
2	F	285	LEU
3	G	268	ARG
3	G	283	LEU
3	G	294	HIS
3	G	297	LEU
3	G	304	LEU

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

Mol	Chain	Res	Type
4	D	99	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

4 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	SEP	A	164	1	8,9,10	0.87	0	8,12,14	0.79	0
1	SEP	E	164	1	8,9,10	0.89	0	8,12,14	0.69	0
1	TPO	A	170	1	8,10,11	1.56	1 (12%)	10,14,16	1.18	1 (10%)
1	TPO	E	170	1	8,10,11	1.49	1 (12%)	10,14,16	1.13	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
1	SEP	A	164	1	-	1/5/8/10	-
1	SEP	E	164	1	-	2/5/8/10	-
1	TPO	A	170	1	-	6/9/11/13	-
1	TPO	E	170	1	-	6/9/11/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	170	TPO	P-OG1	4.10	1.67	1.59
1	E	170	TPO	P-OG1	3.86	1.66	1.59

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	170	TPO	O3P-P-OG1	2.05	115.20	105.99

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	164	SEP	N-CA-CB-OG
1	A	170	TPO	N-CA-CB-CG2
1	A	170	TPO	N-CA-CB-OG1
1	A	170	TPO	C-CA-CB-CG2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
1	A	170	TPO	CG2-CB-OG1-P
1	E	164	SEP	N-CA-CB-OG
1	E	170	TPO	N-CA-CB-CG2
1	E	170	TPO	N-CA-CB-OG1
1	E	170	TPO	C-CA-CB-CG2
1	E	170	TPO	O-C-CA-CB
1	E	170	TPO	CB-OG1-P-O1P
1	E	170	TPO	CG2-CB-OG1-P
1	A	170	TPO	CB-OG1-P-O2P
1	E	164	SEP	CA-CB-OG-P
1	A	170	TPO	O-C-CA-CB

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	E	170	TPO	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

6 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
5	EDO	E	401	-	3,3,3	0.47	0	2,2,2	0.42	0
5	EDO	E	403	-	3,3,3	0.48	0	2,2,2	0.40	0
5	EDO	B	401	-	3,3,3	0.47	0	2,2,2	0.50	0
5	EDO	E	402	-	3,3,3	0.47	0	2,2,2	0.47	0
5	EDO	A	401	-	3,3,3	0.46	0	2,2,2	0.51	0
5	EDO	A	402	-	3,3,3	0.47	0	2,2,2	0.54	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	EDO	E	401	-	-	0/1/1/1	-
5	EDO	E	403	-	-	0/1/1/1	-
5	EDO	B	401	-	-	0/1/1/1	-
5	EDO	E	402	-	-	1/1/1/1	-
5	EDO	A	401	-	-	0/1/1/1	-
5	EDO	A	402	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	402	EDO	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	401	EDO	4	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	264/346 (76%)	0.33	9 (3%)	45	53	36, 47, 74, 164	0
1	E	258/346 (74%)	0.17	8 (3%)	49	58	36, 47, 71, 106	0
2	B	287/323 (88%)	0.28	15 (5%)	27	35	40, 59, 89, 115	1 (0%)
2	F	287/323 (88%)	0.29	18 (6%)	20	27	40, 59, 86, 109	0
3	C	65/82 (79%)	0.11	1 (1%)	73	79	42, 50, 79, 88	0
3	G	65/82 (79%)	0.33	3 (4%)	32	42	40, 51, 75, 100	0
4	D	114/115 (99%)	0.12	2 (1%)	68	75	40, 56, 77, 94	0
4	H	114/115 (99%)	0.21	5 (4%)	34	43	39, 55, 77, 117	0
All	All	1454/1732 (83%)	0.25	61 (4%)	36	45	36, 53, 84, 164	1 (0%)

All (61) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	46	GLY	11.8
1	A	49	SER	9.1
1	A	47	HIS	8.2
1	E	51	ALA	5.5
1	E	263	ALA	4.8
2	B	194	PHE	4.8
4	D	1	GLN	4.7
1	A	48	ARG	4.6
2	F	72	LYS	4.5
2	B	286	ALA	4.5
1	A	51	ALA	4.0
2	B	41	LEU	3.9
4	H	42	ASP	3.8
2	B	1[A]	MET	3.7
2	F	42	PRO	3.7
2	B	285	LEU	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	38	GLY	3.7
2	B	40	VAL	3.5
2	B	72	LYS	3.4
2	B	130	ARG	3.4
3	C	244	ALA	3.3
2	F	194	PHE	3.3
2	F	134	LEU	3.3
4	H	1	GLN	3.3
1	A	311	ASN	3.2
2	F	285	LEU	3.2
2	F	38	GLY	3.1
2	F	41	LEU	3.1
2	B	287	LEU	2.9
1	A	154	ALA	2.9
2	B	134	LEU	2.9
2	F	43	ASN	2.8
2	F	287	LEU	2.7
1	E	303	PRO	2.7
1	A	305	CYS	2.6
2	F	140	LEU	2.6
1	E	52	LYS	2.5
1	A	50	GLU	2.5
2	F	240	GLU	2.5
4	H	107	GLN	2.5
4	D	107	GLN	2.4
1	E	85	SER	2.4
3	G	309	SER	2.4
2	F	198	ILE	2.3
2	B	43	ASN	2.3
1	E	298	ARG	2.3
2	F	180	ILE	2.3
3	G	245	LEU	2.2
1	E	210	VAL	2.2
2	F	223	ARG	2.2
2	F	189	LYS	2.2
2	F	40	VAL	2.2
4	H	2	VAL	2.1
2	F	266	ARG	2.1
2	B	182	GLU	2.1
4	H	3	GLN	2.1
2	B	42	PRO	2.1
1	E	309	ARG	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	21	ARG	2.0
3	G	249	GLN	2.0
2	B	180	ILE	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
1	TPO	A	170	11/12	0.94	0.16	46,50,66,67	0
1	SEP	E	164	10/11	0.97	0.12	46,53,57,60	0
1	TPO	E	170	11/12	0.97	0.13	39,50,62,65	0
1	SEP	A	164	10/11	0.99	0.11	48,51,56,58	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
5	EDO	E	401	4/4	0.58	0.30	48,53,62,64	0
5	EDO	E	403	4/4	0.82	0.20	50,53,54,58	0
5	EDO	A	402	4/4	0.92	0.17	57,63,66,72	0
5	EDO	B	401	4/4	0.96	0.13	50,50,51,54	0
5	EDO	A	401	4/4	0.97	0.16	39,41,43,46	0
5	EDO	E	402	4/4	0.98	0.12	45,46,52,53	0

6.5 Other polymers (i)

There are no such residues in this entry.