WORLDWIDE PDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2024 – 09:11 PM EST

PDB ID	:	3MZK
Title	:	Sec13/Sec16 complex, S.cerevisiae
Authors	:	Whittle, J.R.; Schwartz, T.U.
Deposited on	:	2010-05-12
Resolution	:	2.69 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
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:

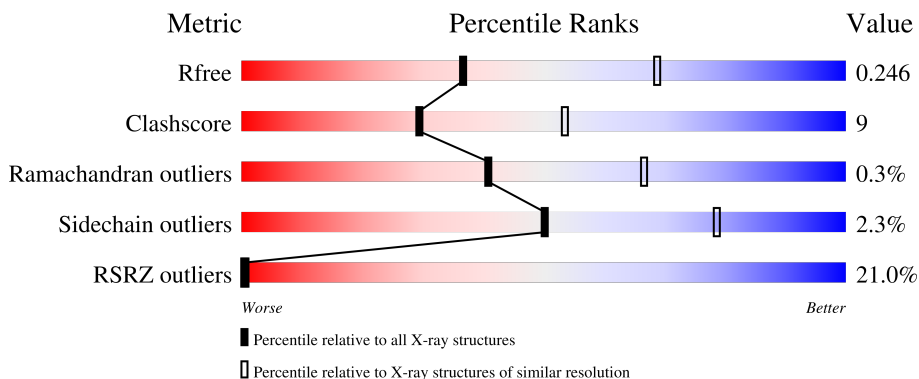|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 2.69 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	Percentile Ranks	Value
Rfree		0.246
Clashscore		9
Ramachandran outliers		0.3%
Sidechain outliers		2.3%
RSRZ outliers		21.0%

Worse — *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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	297	% 80% 19% •
1	D	297	74% 77% 23%
2	B	441	% 74% 13% • 13%
2	C	441	13% 65% 21% 14%

WORLDWIDE wwPDB PROTEIN DATA BANK

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10662 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 Protein transport protein SEC13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	297	Total 2329	C 1481	N 400	O 444	S 4	0	0	0
1	D	297	Total 2322	C 1477	N 398	O 443	S 4	0	0	0

- Molecule 2 is a protein called Protein transport protein SEC16.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	385	Total 2932	C 1904	N 472	O 544	S 12	0	0	0
2	C	381	Total 2904	C 1890	N 467	O 536	S 11	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	980	GLY	-	expression tag	UNP P48415
B	981	PRO	-	expression tag	UNP P48415
B	982	GLY	-	expression tag	UNP P48415
B	983	SER	-	expression tag	UNP P48415
C	980	GLY	-	expression tag	UNP P48415
C	981	PRO	-	expression tag	UNP P48415
C	982	GLY	-	expression tag	UNP P48415
C	983	SER	-	expression tag	UNP P48415

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	116	Total 116	O 116	0	0
3	B	48	Total 48	O 48	0	0

Continued on next page...

Continued from previous page...

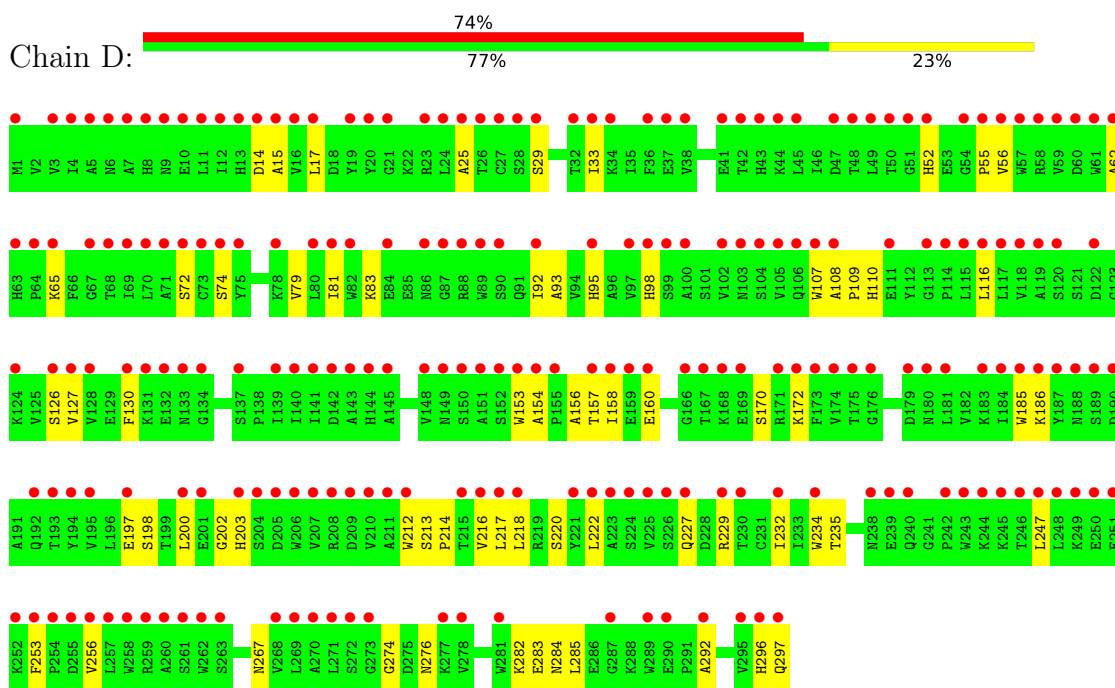| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	10	Total 10	O 10	0	0
3	D	1	Total 1	O 1	0	0

3 Residue-property plots ⓘ

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.
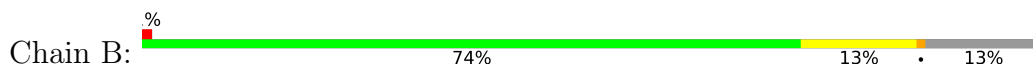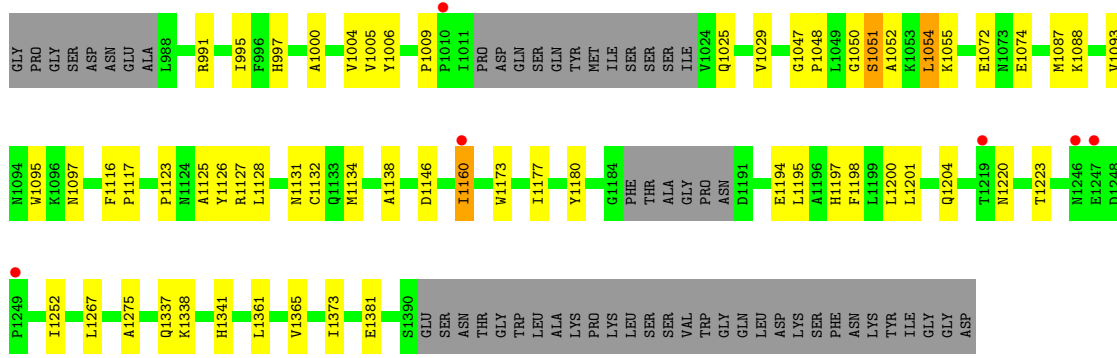
• Molecule 1: Protein transport protein SEC13

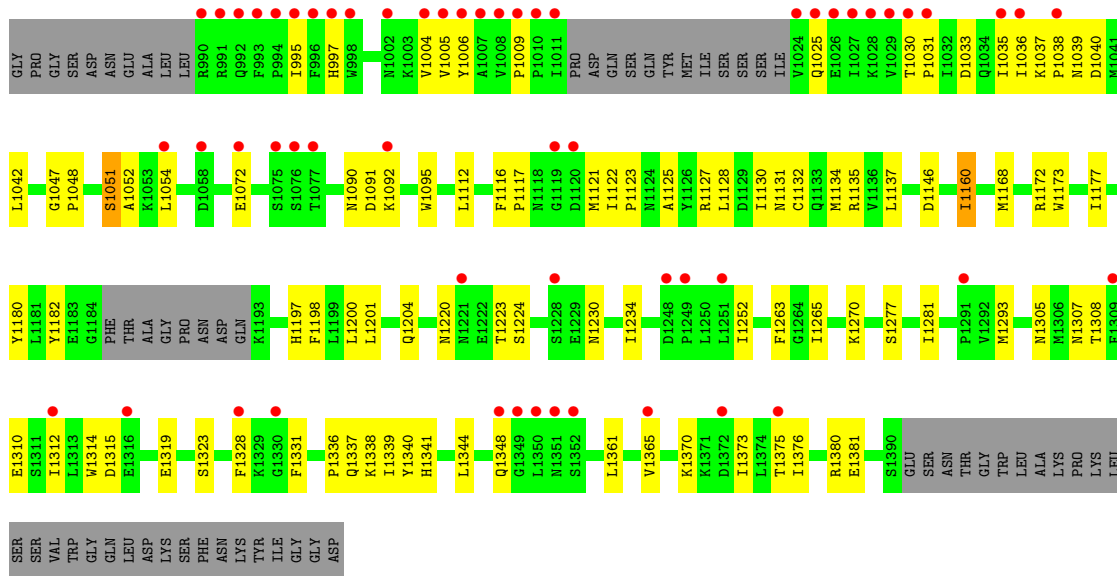
Chain A:

• Molecule 1: Protein transport protein SEC13

Chain D:

• Molecule 2: Protein transport protein SEC16

Chain B:

GLY PRO GLY SER ASP ASN GLU ALA L988 R991 I995 F996 H997 A1000 V1004 V1005 Y1006 P1009 P1010 I1011 PRO ASP GLN SER GLN TYR MET ILE SER SER ILE V1024 Q1025 V1029 G1047 P1048 L1049 G1050 S1051 A1052 K1053 L1054 K1055 E1072 N1073 E1074 M1087 K1088 V1093

N1094 W1095 K1096 N1097 F1116 P1117 P1123 N1124 A1125 Y1126 R1127 L1128 N1131 C1132 Q1133 M1134 A1138 D1146 I1160 W1173 I1177 Y1180 G1184 PHE THR ALA GLY PRO ASN D1191 E1194 L1195 A1196 H1197 F1198 L1199 L1200 L1201 Q1204 T1219 N1220 T1223 N1246 E1247 D1248

P1249 I1252 L1267 A1275 Q1337 K1338 H1341 L1361 V1365 I1373 E1381 S1390 GLU SER ASN THR GLY TRP LEU ALA LYS PRO LYS LEU SER SER VAL TRP GLY GLN LEU ASP LYS SER PHE ASN LYS TYR ILE GLY GLY ASP

• Molecule 2: Protein transport protein SEC16

Chain C:

GLY PRO GLY SER ASP ASN GLU ALA LEU LEU R990 R991 Q992 F993 P994 I995 F996 H997 W998 N1002 K1003 V1004 V1005 Y1006 A1007 V1008 P1009 P1010 I1011 PRO ASP GLN SER GLN TYR MET ILE SER SER ILE V1024 Q1025 E1026 I1027 K1028 V1029 T1030 P1031 I1032 D1033 Q1034 I1035 I1036 K1037 P1038 N1039 D1040 M1041

L1042 G1047 P1048 S1051 A1052 K1053 L1054 D1058 E1072 S1075 S1076 T1077 N1090 D1091 K1092 W1095 L1112 F1116 P1117 N1118 G1119 D1120 M1121 I1122 P1123 N1124 A1125 Y1126 R1127 L1128 D1129 I1130 N1131 C1132 Q1133 M1134 R1135 V1136 L1137 D1146 I1160 M1168 R1172 W1173 I1177

Y1180 L1181 Y1182 E1183 G1184 PHE THR ALA GLY PRO ASN ASP GLN K193 H1197 F1198 L1199 L1200 L1201 Q1204 N1220 N1221 E1222 T1223 S1224 S1228 E1229 N1230 I1234 D1248 P1249 L1250 L1251 I1252 F1263 G1264 I1265 K1270 S1277 I1281 P1291 V1292 M1293 N1305 M1306 N1307 T1308 F1309

E1310 S1311 I1312 L1313 W1314 D1315 E1316 E1319 S1323 F1328 K1329 G1330 F1331 P1336 Q1337 K1338 I1339 Y1340 H1341 L1344 Q1348 G1349 L1350 N1351 S1352 L1361 V1365 K1370 K1371 D1372 I1373 I1374 T1375 I1376 R1380 E1381 S1390 GLU SER ASN THR GLY TRP LEU ALA LYS PRO LYS LEU

SER SER VAL TRP GLY GLN LEU ASP LYS SER PHE ASN LYS TYR ILE GLY GLY ASP

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.73Å 139.01Å 205.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.71 – 2.69 30.71 – 2.69	Depositor EDS
% Data completeness (in resolution range)	92.3 (30.71-2.69) 98.1 (30.71-2.69)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.94 (at 2.68Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.1_357)	Depositor
R, R_{free}	0.198 , 0.243 0.200 , 0.246	Depositor DCC
R_{free} test set	2292 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	56.2	Xtriage
Anisotropy	0.511	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 79.0	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10662	wwPDB-VP
Average B, all atoms (Å^2)	114.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% 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 (i)

5.1   Standard geometry (i)

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	0.34	0/2392	0.53	0/3263
1	D	0.31	0/2385	0.51	0/3255
2	B	0.29	0/2999	0.42	0/4094
2	C	0.24	0/2971	0.41	0/4057
All	All	0.30	0/10747	0.47	0/14669

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2329	0	2254	41	0
1	D	2322	0	2239	65	0
2	B	2932	0	2818	49	0
2	C	2904	0	2791	87	0
3	A	116	0	0	3	0
3	B	48	0	0	5	0
3	C	10	0	0	0	0
3	D	1	0	0	0	0
All	All	10662	0	10102	195	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 9.

The worst 5 of 195 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:1128:LEU:HD22	2:B:1160:ILE:HD12	1.46	0.97
1:D:227:GLN:HA	1:D:256:VAL:HG23	1.49	0.95
2:C:1128:LEU:HD22	2:C:1160:ILE:HD12	1.47	0.94
1:A:227:GLN:HA	1:A:256:VAL:HG23	1.47	0.93
2:C:1381:GLU:HB3	1:D:216:VAL:HG22	1.55	0.89

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	295/297 (99%)	278 (94%)	16 (5%)	1 (0%)	41	66
1	D	295/297 (99%)	284 (96%)	10 (3%)	1 (0%)	41	66
2	B	379/441 (86%)	360 (95%)	18 (5%)	1 (0%)	41	66
2	C	375/441 (85%)	347 (92%)	27 (7%)	1 (0%)	41	66
All	All	1344/1476 (91%)	1269 (94%)	71 (5%)	4 (0%)	41	66

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	1047	GLY
2	C	1047	GLY
1	D	202	GLY
1	A	202	GLY

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	250/252 (99%)	246 (98%)	4 (2%)	62	85
1	D	248/252 (98%)	246 (99%)	2 (1%)	81	93
2	B	306/395 (78%)	297 (97%)	9 (3%)	42	71
2	C	302/395 (76%)	292 (97%)	10 (3%)	38	67
All	All	1106/1294 (86%)	1081 (98%)	25 (2%)	50	78

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	1040	ASP
2	C	1122	ILE
1	D	198	SER
2	C	1072	GLU
2	C	1146	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	1337	GLN
2	C	1341	HIS
1	D	98	HIS
2	C	1348	GLN
2	B	1341	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 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(Å²)	Q<0.9
1	A	297/297 (100%)	-0.40	2 (0%)	87	89	20, 39, 90, 133	0
1	D	297/297 (100%)	4.22	220 (74%)	0	0	122, 243, 326, 392	0
2	B	385/441 (87%)	-0.23	6 (1%)	72	74	33, 72, 123, 174	0
2	C	381/441 (86%)	0.63	57 (14%)	2	1	49, 99, 228, 287	0
All	All	1360/1476 (92%)	0.94	285 (20%)	1	0	20, 84, 276, 392	0

The worst 5 of 285 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	68	THR	19.6
1	D	67	GLY	15.4
1	D	189	SER	15.1
1	D	272	SER	14.0
1	D	253	PHE	13.4

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.