



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 10:27 PM GMT

PDB ID : 4YA5  
Title : Yeast 20S proteasome beta2-H114D mutant in complex with Ac-PAE-ep  
Authors : Huber, E.M.; Groll, M.  
Deposited on : 2015-02-17  
Resolution : 2.50 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

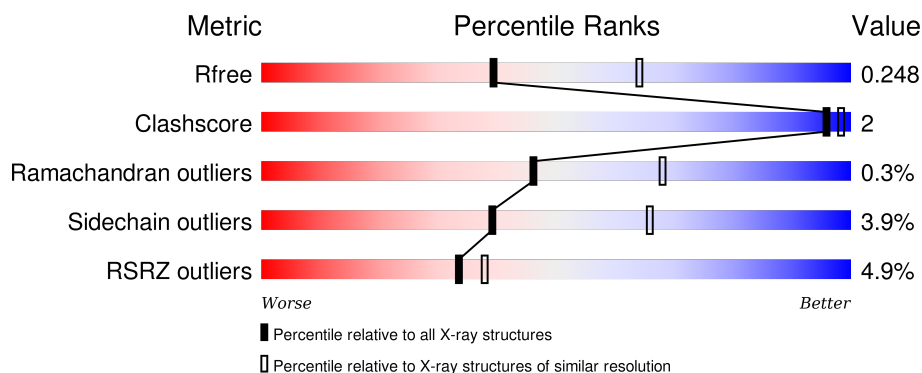
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	250	<div> <div>4%</div> <div>98%</div> <div>.</div> </div>
1	O	250	<div> <div>4%</div> <div>97%</div> <div>.</div> </div>
2	B	258	<div> <div>9%</div> <div>84%</div> <div>10%</div> <div>5%</div> </div>
2	P	258	<div> <div>5%</div> <div>85%</div> <div>9%</div> <div>5%</div> </div>
3	C	254	<div> <div>10%</div> <div>85%</div> <div>9%</div> <div>6%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Q	254	
4	D	260	
4	R	260	
5	E	234	
5	S	234	
6	F	288	
6	T	288	
7	G	252	
7	U	252	
8	H	232	
8	V	232	
9	I	205	
9	W	205	
10	J	198	
10	X	198	
11	K	212	
11	Y	212	
12	L	222	
12	Z	222	
13	M	246	
13	a	246	
14	N	196	
14	b	196	
15	c	5	
15	d	5	

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
16	MG	I	301	-	-	-	X

## 2 Entry composition

There are 18 unique types of molecules in this entry. The entry contains 50389 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 Proteasome subunit alpha type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	250	Total	C	N	O	S	0	0	0
			1915	1219	315	377	4			
1	O	250	Total	C	N	O	S	0	0	0
			1915	1219	315	377	4			

- Molecule 2 is a protein called Proteasome subunit alpha type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	244	Total	C	N	O	S	0	0	0
			1904	1201	321	379	3			
2	P	244	Total	C	N	O	S	0	0	0
			1904	1201	321	379	3			

- Molecule 3 is a protein called Proteasome subunit alpha type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	240	Total	C	N	O	S	0	0	0
			1881	1176	329	372	4			
3	Q	240	Total	C	N	O	S	0	0	0
			1881	1176	329	372	4			

- Molecule 4 is a protein called Proteasome subunit alpha type-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	235	Total	C	N	O	S	0	0	0
			1813	1136	304	366	7			
4	R	235	Total	C	N	O	S	0	0	0
			1813	1136	304	366	7			

- Molecule 5 is a protein called Proteasome subunit alpha type-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	231	Total	C	N	O	S	0	0	0
			1773	1114	307	348	4			
5	S	231	Total	C	N	O	S	0	0	0
			1773	1114	307	348	4			

- Molecule 6 is a protein called Probable proteasome subunit alpha type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	243	Total	C	N	O	S	0	0	0
			1892	1203	329	356	4			
6	T	243	Total	C	N	O	S	0	0	0
			1892	1203	329	356	4			

- Molecule 7 is a protein called Proteasome subunit alpha type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	241	Total	C	N	O	S	0	0	0
			1907	1214	320	365	8			
7	U	241	Total	C	N	O	S	0	0	0
			1907	1214	320	365	8			

- Molecule 8 is a protein called Proteasome subunit beta type-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	226	Total	C	N	O	S	0	0	0
			1717	1080	296	334	7			
8	V	226	Total	C	N	O	S	0	0	0
			1717	1080	296	334	7			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
H	114	ASP	HIS	engineered mutation	UNP P25043
V	114	ASP	HIS	engineered mutation	UNP P25043

- Molecule 9 is a protein called Proteasome subunit beta type-3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	204	Total	C	N	O	S	0	0	0
			1581	1010	258	305	8			
9	W	204	Total	C	N	O	S	0	0	0
			1581	1010	258	305	8			

- Molecule 10 is a protein called Proteasome subunit beta type-4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	195	Total	C	N	O	S	0	0	0
			1561	992	264	299	6			
10	X	195	Total	C	N	O	S	0	0	0
			1561	992	264	299	6			

- Molecule 11 is a protein called Proteasome subunit beta type-5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	212	Total	C	N	O	S	0	0	0
			1644	1045	280	312	7			
11	Y	212	Total	C	N	O	S	0	0	0
			1644	1045	280	312	7			

- Molecule 12 is a protein called Proteasome subunit beta type-6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	4			
12	Z	222	Total	C	N	O	S	0	0	0
			1757	1115	303	335	4			

- Molecule 13 is a protein called Proteasome subunit beta type-7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	M	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			
13	a	233	Total	C	N	O	S	0	0	0
			1824	1154	312	351	7			

- Molecule 14 is a protein called Proteasome subunit beta type-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			
14	b	196	Total	C	N	O	S	0	0	0
			1512	955	250	300	7			

- Molecule 15 is a protein called Ac-PAE-ep.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
15	c	5	Total	C	N	O	0	0	0
			28	18	3	7			
15	d	5	Total	C	N	O	0	0	0
			28	18	3	7			

- Molecule 16 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	G	1	Total	Mg	0	0
			1	1		
16	K	1	Total	Mg	0	0
			1	1		
16	I	1	Total	Mg	0	0
			1	1		
16	V	1	Total	Mg	0	0
			1	1		
16	Z	1	Total	Mg	0	0
			1	1		
16	N	1	Total	Mg	0	0
			1	1		
16	Y	1	Total	Mg	0	0
			1	1		

- Molecule 17 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	G	1	Total	Cl	0	0
			1	1		

- Molecule 18 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	A	37	Total	O	0	0
			37	37		
18	B	27	Total	O	0	0
			27	27		
18	C	25	Total	O	0	0
			25	25		
18	D	27	Total	O	0	0
			27	27		
18	E	17	Total	O	0	0
			17	17		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	F	32	Total 32	O 32	0	0
18	G	52	Total 52	O 52	0	0
18	H	50	Total 50	O 50	0	0
18	I	39	Total 39	O 39	0	0
18	J	42	Total 42	O 42	0	0
18	K	41	Total 41	O 41	0	0
18	L	40	Total 40	O 40	0	0
18	M	50	Total 50	O 50	0	0
18	N	41	Total 41	O 41	0	0
18	O	26	Total 26	O 26	0	0
18	P	26	Total 26	O 26	0	0
18	Q	13	Total 13	O 13	0	0
18	R	18	Total 18	O 18	0	0
18	S	15	Total 15	O 15	0	0
18	T	28	Total 28	O 28	0	0
18	U	40	Total 40	O 40	0	0
18	V	38	Total 38	O 38	0	0
18	W	42	Total 42	O 42	0	0
18	X	37	Total 37	O 37	0	0
18	Y	36	Total 36	O 36	0	0
18	Z	40	Total 40	O 40	0	0

*Continued on next page...*

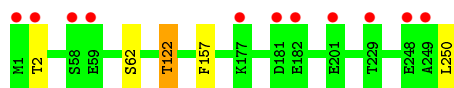
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	a	41	Total 41	O 41	0	0
18	b	38	Total 38	O 38	0	0
18	c	3	Total 3	O 3	0	0
18	d	2	Total 2	O 2	0	0

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

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

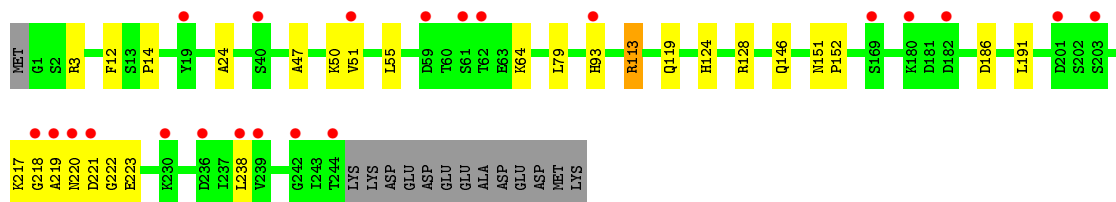
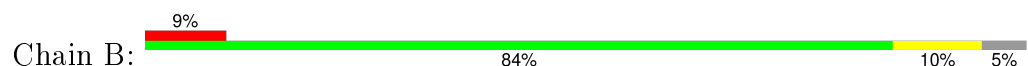
- Molecule 1: Proteasome subunit alpha type-2



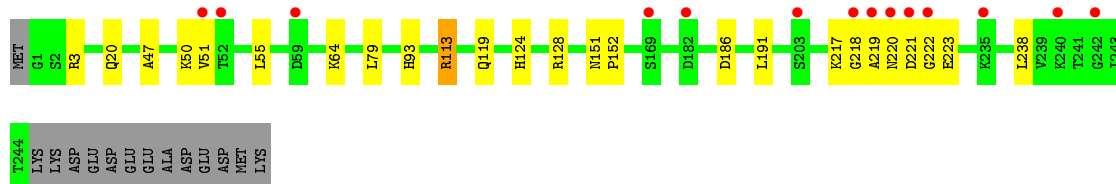
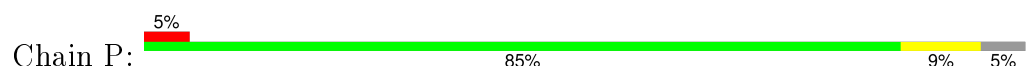
- Molecule 1: Proteasome subunit alpha type-2



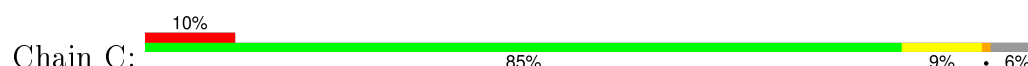
- Molecule 2: Proteasome subunit alpha type-3



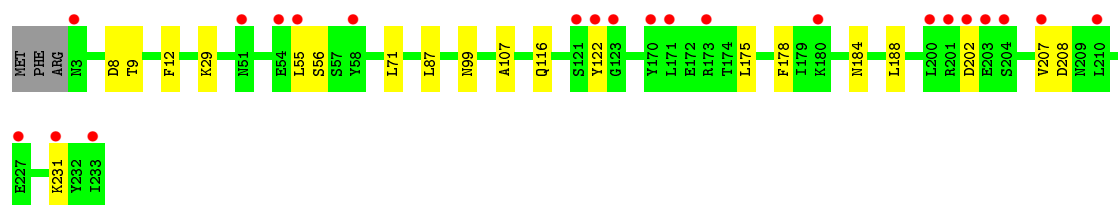
- Molecule 2: Proteasome subunit alpha type-3



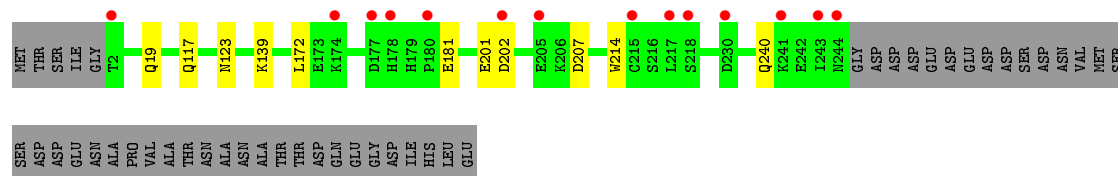
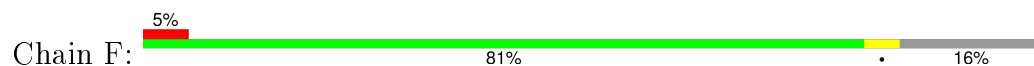
- Molecule 3: Proteasome subunit alpha type-4



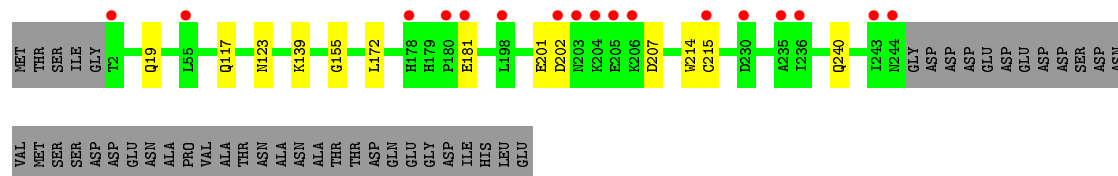




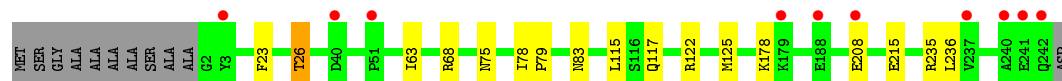
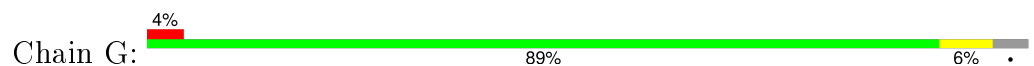
- Molecule 6: Probable proteasome subunit alpha type-7



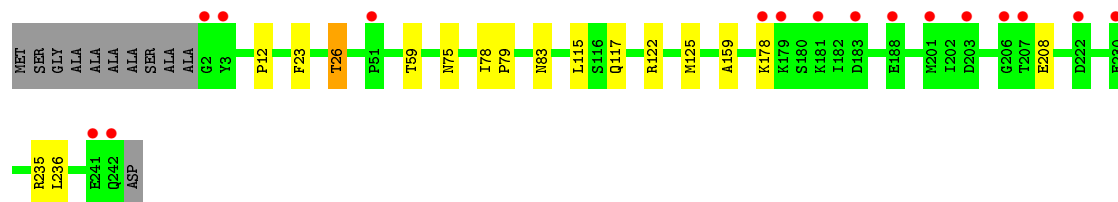
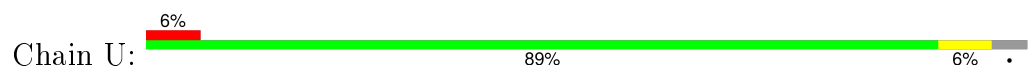
- Molecule 6: Probable proteasome subunit alpha type-7



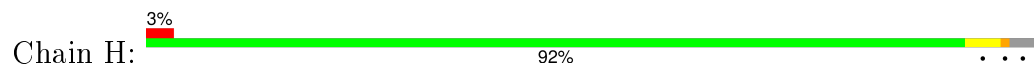
- Molecule 7: Proteasome subunit alpha type-1



- Molecule 7: Proteasome subunit alpha type-1



- Molecule 8: Proteasome subunit beta type-2

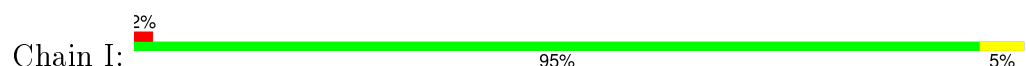




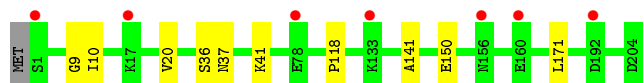
- Molecule 8: Proteasome subunit beta type-2



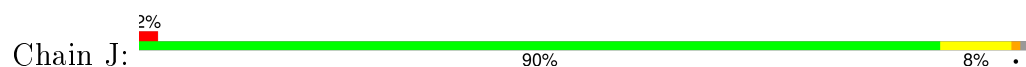
- Molecule 9: Proteasome subunit beta type-3



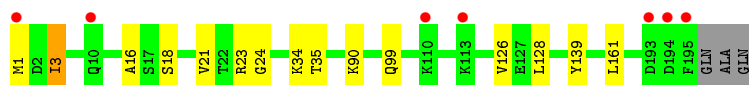
- Molecule 9: Proteasome subunit beta type-3



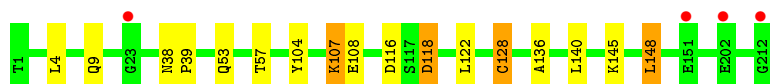
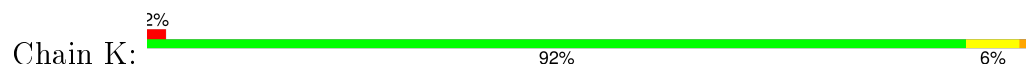
- Molecule 10: Proteasome subunit beta type-4



- Molecule 10: Proteasome subunit beta type-4

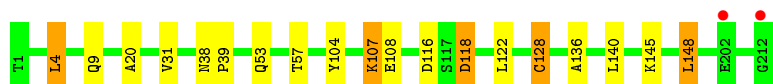


- Molecule 11: Proteasome subunit beta type-5



- Molecule 11: Proteasome subunit beta type-5





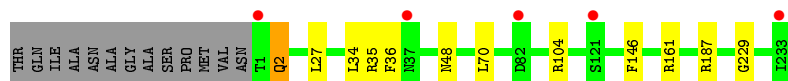
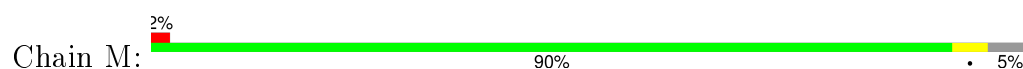
- Molecule 12: Proteasome subunit beta type-6



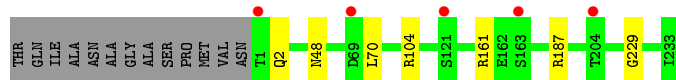
- Molecule 12: Proteasome subunit beta type-6



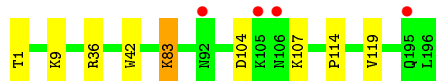
- Molecule 13: Proteasome subunit beta type-7



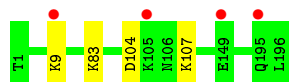
- Molecule 13: Proteasome subunit beta type-7



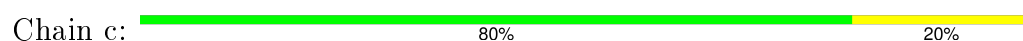
- Molecule 14: Proteasome subunit beta type-1



- Molecule 14: Proteasome subunit beta type-1



- Molecule 15: Ac-PAE-ep





- Molecule 15: Ac-PAE-ep

Chain d:





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	136.71Å 301.32Å 145.93Å 90.00° 113.18° 90.00°	Depositor
Resolution (Å)	15.00 – 2.50 15.00 – 2.50	Depositor EDS
% Data completeness (in resolution range)	95.9 (15.00-2.50) 95.9 (15.00-2.50)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.81 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.226 , 0.242 0.232 , 0.248	Depositor DCC
$R_{free}$ test set	17789 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	44.2	Xtriage
Anisotropy	0.075	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 38.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	47 of 355770 reflections (0.013%)	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	50389	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.58% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: MG, POL, GAU, ACE, CL

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	0.28	0/1952	0.47	0/2642
1	O	0.28	0/1952	0.47	0/2642
2	B	0.29	0/1934	0.50	0/2618
2	P	0.28	0/1934	0.50	0/2618
3	C	0.28	0/1910	0.51	0/2586
3	Q	0.28	0/1910	0.51	0/2586
4	D	0.28	0/1837	0.48	0/2475
4	R	0.28	0/1837	0.49	0/2475
5	E	0.28	0/1800	0.48	0/2433
5	S	0.28	0/1800	0.48	0/2433
6	F	0.28	0/1932	0.46	0/2609
6	T	0.28	0/1932	0.46	0/2609
7	G	0.28	0/1945	0.47	0/2634
7	U	0.28	0/1945	0.47	0/2634
8	H	0.26	0/1747	0.48	0/2369
8	V	0.26	0/1747	0.48	0/2369
9	I	0.29	0/1611	0.50	0/2174
9	W	0.29	0/1611	0.50	0/2174
10	J	0.27	0/1589	0.49	0/2142
10	X	0.27	0/1589	0.49	0/2142
11	K	0.27	0/1681	0.51	0/2274
11	Y	0.27	0/1681	0.51	1/2274 (0.0%)
12	L	0.28	0/1795	0.50	0/2420
12	Z	0.28	0/1795	0.50	0/2420
13	M	0.28	0/1855	0.51	0/2514
13	a	0.29	0/1855	0.51	0/2514
14	N	0.29	0/1541	0.50	1/2087 (0.0%)
14	b	0.27	0/1541	0.49	0/2087
15	c	2.57	1/13 (7.7%)	0.98	0/18
15	d	1.89	1/13 (7.7%)	0.94	0/18
All	All	0.28	2/50284 (0.0%)	0.49	2/67990 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	c	2	PRO	CA-C	-7.03	1.38	1.52
15	d	2	PRO	CA-C	-6.32	1.40	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
14	N	1	THR	N-CA-C	5.16	124.93	111.00
11	Y	4	LEU	CA-CB-CG	5.13	127.09	115.30

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	1915	0	1929	1	0
1	O	1915	0	1929	4	0
2	B	1904	0	1904	16	0
2	P	1904	0	1904	11	0
3	C	1881	0	1895	11	0
3	Q	1881	0	1895	7	0
4	D	1813	0	1797	7	0
4	R	1813	0	1797	7	0
5	E	1773	0	1775	6	0
5	S	1773	0	1775	5	0
6	F	1892	0	1883	1	0
6	T	1892	0	1883	3	0
7	G	1907	0	1901	4	0
7	U	1907	0	1901	5	0
8	H	1717	0	1716	9	0
8	V	1717	0	1716	8	0
9	I	1581	0	1574	6	0
9	W	1581	0	1574	5	0
10	J	1561	0	1569	10	0
10	X	1561	0	1569	9	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
11	K	1644	0	1595	8	0
11	Y	1644	0	1595	9	0
12	L	1757	0	1711	1	0
12	Z	1757	0	1711	1	0
13	M	1824	0	1832	7	0
13	a	1824	0	1832	0	0
14	N	1512	0	1478	5	0
14	b	1512	0	1478	0	0
15	c	28	0	27	0	0
15	d	28	0	27	0	0
16	G	1	0	0	0	0
16	I	1	0	0	0	0
16	K	1	0	0	0	0
16	N	1	0	0	0	0
16	V	1	0	0	0	0
16	Y	1	0	0	0	0
16	Z	1	0	0	0	0
17	G	1	0	0	0	0
18	A	37	0	0	0	0
18	B	27	0	0	2	0
18	C	25	0	0	0	0
18	D	27	0	0	1	0
18	E	17	0	0	0	0
18	F	32	0	0	0	0
18	G	52	0	0	0	0
18	H	50	0	0	0	0
18	I	39	0	0	0	0
18	J	42	0	0	0	0
18	K	41	0	0	0	0
18	L	40	0	0	0	0
18	M	50	0	0	1	0
18	N	41	0	0	1	0
18	O	26	0	0	0	0
18	P	26	0	0	2	0
18	Q	13	0	0	0	0
18	R	18	0	0	0	0
18	S	15	0	0	0	0
18	T	28	0	0	1	0
18	U	40	0	0	0	0
18	V	38	0	0	0	0
18	W	42	0	0	0	0
18	X	37	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
18	Y	36	0	0	0	0
18	Z	40	0	0	1	0
18	a	41	0	0	0	0
18	b	38	0	0	0	0
18	c	3	0	0	0	0
18	d	2	0	0	0	0
All	All	50389	0	49172	133	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.

The worst 5 of 133 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)
8:H:112:SER:OG	8:H:120:ASP:OD1	2.01	0.75
8:V:112:SER:OG	8:V:120:ASP:OD1	2.02	0.74
8:H:113:ILE:HG13	8:H:119:THR:HG22	1.70	0.74
8:V:113:ILE:HG13	8:V:119:THR:HG22	1.71	0.73
11:Y:116:ASP:OD2	11:Y:118:ASP:OD1	2.15	0.66

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	248/250 (99%)	239 (96%)	8 (3%)	1 (0%)	39 61
1	O	248/250 (99%)	238 (96%)	9 (4%)	1 (0%)	39 61
2	B	242/258 (94%)	234 (97%)	4 (2%)	4 (2%)	11 19
2	P	242/258 (94%)	234 (97%)	4 (2%)	4 (2%)	11 19
3	C	238/254 (94%)	233 (98%)	3 (1%)	2 (1%)	24 41

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	Q	238/254 (94%)	233 (98%)	2 (1%)	3 (1%)	15	26
4	D	231/260 (89%)	229 (99%)	2 (1%)	0	100	100
4	R	231/260 (89%)	229 (99%)	2 (1%)	0	100	100
5	E	229/234 (98%)	224 (98%)	5 (2%)	0	100	100
5	S	229/234 (98%)	224 (98%)	5 (2%)	0	100	100
6	F	241/288 (84%)	238 (99%)	3 (1%)	0	100	100
6	T	241/288 (84%)	238 (99%)	3 (1%)	0	100	100
7	G	239/252 (95%)	238 (100%)	1 (0%)	0	100	100
7	U	239/252 (95%)	239 (100%)	0	0	100	100
8	H	224/232 (97%)	219 (98%)	5 (2%)	0	100	100
8	V	224/232 (97%)	218 (97%)	6 (3%)	0	100	100
9	I	202/205 (98%)	196 (97%)	6 (3%)	0	100	100
9	W	202/205 (98%)	196 (97%)	6 (3%)	0	100	100
10	J	193/198 (98%)	189 (98%)	3 (2%)	1 (0%)	34	55
10	X	193/198 (98%)	189 (98%)	3 (2%)	1 (0%)	34	55
11	K	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
11	Y	210/212 (99%)	206 (98%)	4 (2%)	0	100	100
12	L	220/222 (99%)	215 (98%)	5 (2%)	0	100	100
12	Z	220/222 (99%)	215 (98%)	5 (2%)	0	100	100
13	M	231/246 (94%)	223 (96%)	7 (3%)	1 (0%)	39	61
13	a	231/246 (94%)	224 (97%)	6 (3%)	1 (0%)	39	61
14	N	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
14	b	194/196 (99%)	189 (97%)	5 (3%)	0	100	100
15	c	2/5 (40%)	2 (100%)	0	0	100	100
15	d	2/5 (40%)	2 (100%)	0	0	100	100
All	All	6288/6624 (95%)	6148 (98%)	121 (2%)	19 (0%)	46	68

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	51	VAL
2	B	222	GLY
3	C	202	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	P	51	VAL
2	P	222	GLY

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	209/209 (100%)	205 (98%)	4 (2%)	65	87
1	O	209/209 (100%)	205 (98%)	4 (2%)	65	87
2	B	203/216 (94%)	195 (96%)	8 (4%)	39	66
2	P	203/216 (94%)	195 (96%)	8 (4%)	39	66
3	C	212/226 (94%)	200 (94%)	12 (6%)	25	46
3	Q	212/226 (94%)	200 (94%)	12 (6%)	25	46
4	D	194/215 (90%)	182 (94%)	12 (6%)	23	41
4	R	194/215 (90%)	182 (94%)	12 (6%)	23	41
5	E	190/193 (98%)	177 (93%)	13 (7%)	20	36
5	S	190/193 (98%)	177 (93%)	13 (7%)	20	36
6	F	201/239 (84%)	191 (95%)	10 (5%)	30	53
6	T	201/239 (84%)	191 (95%)	10 (5%)	30	53
7	G	206/210 (98%)	195 (95%)	11 (5%)	28	50
7	U	206/210 (98%)	195 (95%)	11 (5%)	28	50
8	H	185/190 (97%)	181 (98%)	4 (2%)	60	84
8	V	185/190 (97%)	181 (98%)	4 (2%)	60	84
9	I	172/173 (99%)	170 (99%)	2 (1%)	78	93
9	W	172/173 (99%)	170 (99%)	2 (1%)	78	93
10	J	173/175 (99%)	168 (97%)	5 (3%)	50	77
10	X	173/175 (99%)	168 (97%)	5 (3%)	50	77
11	K	169/169 (100%)	161 (95%)	8 (5%)	32	56
11	Y	169/169 (100%)	161 (95%)	8 (5%)	32	56

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
12	L	185/185 (100%)	179 (97%)	6 (3%)	46	74
12	Z	185/185 (100%)	179 (97%)	6 (3%)	46	74
13	M	199/208 (96%)	193 (97%)	6 (3%)	48	76
13	a	199/208 (96%)	193 (97%)	6 (3%)	48	76
14	N	162/162 (100%)	158 (98%)	4 (2%)	55	82
14	b	162/162 (100%)	158 (98%)	4 (2%)	55	82
15	c	1/1 (100%)	1 (100%)	0	100	100
15	d	1/1 (100%)	1 (100%)	0	100	100
All	All	5322/5542 (96%)	5112 (96%)	210 (4%)	39	66

5 of 210 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
12	L	150	LEU
3	Q	4	ARG
12	Z	3	ASN
13	M	70	LEU
1	O	122	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 116 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	108	ASN
3	Q	38	ASN
12	Z	158	ASN
13	M	194	ASN
2	P	58	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

2 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	$\# Z  > 2$
15	GAU	c	4	15,14	5,8,8	1.45	1 (20%)	4,9,9	2.03	2 (50%)
15	GAU	d	4	15,14	5,8,8	1.38	1 (20%)	4,9,9	1.82	1 (25%)

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
15	GAU	c	4	15,14	-	0/5/7/7	0/0/0/0
15	GAU	d	4	15,14	-	0/5/7/7	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	c	4	GAU	C-CA	3.04	1.56	1.52
15	d	4	GAU	C-CA	3.07	1.56	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	d	4	GAU	CG-CB-CA	-3.51	105.04	113.80
15	c	4	GAU	OXT-C-CA	-2.68	104.68	111.77
15	c	4	GAU	CB-CA-N	-2.47	102.05	108.92

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	250/250 (100%)	0.14	11 (4%)	38 43	31, 43, 77, 114	0
1	O	250/250 (100%)	0.25	11 (4%)	38 43	36, 52, 95, 125	0
2	B	244/258 (94%)	0.31	22 (9%)	12 12	30, 48, 89, 135	0
2	P	244/258 (94%)	0.41	14 (5%)	27 31	37, 54, 99, 145	0
3	C	240/254 (94%)	0.42	25 (10%)	8 8	31, 54, 115, 140	0
3	Q	240/254 (94%)	0.69	39 (16%)	2 2	36, 65, 144, 169	0
4	D	235/260 (90%)	0.25	12 (5%)	32 36	37, 56, 88, 132	0
4	R	235/260 (90%)	0.32	13 (5%)	29 32	40, 60, 99, 134	0
5	E	231/234 (98%)	0.37	12 (5%)	31 35	40, 58, 96, 144	0
5	S	231/234 (98%)	0.52	22 (9%)	10 11	41, 65, 107, 155	0
6	F	243/288 (84%)	0.19	14 (5%)	26 30	32, 50, 100, 130	0
6	T	243/288 (84%)	0.36	17 (6%)	19 22	36, 61, 120, 149	0
7	G	241/252 (95%)	0.10	10 (4%)	41 46	25, 45, 78, 130	0
7	U	241/252 (95%)	0.21	16 (6%)	22 24	34, 49, 81, 124	0
8	H	226/232 (97%)	0.12	7 (3%)	52 57	25, 41, 72, 133	0
8	V	226/232 (97%)	0.22	9 (3%)	42 47	32, 45, 74, 150	0
9	I	204/205 (99%)	-0.16	5 (2%)	61 65	26, 39, 66, 90	0
9	W	204/205 (99%)	-0.12	7 (3%)	49 54	30, 43, 71, 96	0
10	J	195/198 (98%)	-0.00	4 (2%)	67 71	30, 44, 69, 126	0
10	X	195/198 (98%)	0.07	7 (3%)	46 51	30, 45, 70, 132	0
11	K	212/212 (100%)	0.01	4 (1%)	70 73	28, 44, 67, 89	0
11	Y	212/212 (100%)	0.03	2 (0%)	85 88	33, 45, 69, 88	0
12	L	222/222 (100%)	0.03	5 (2%)	64 67	30, 46, 77, 110	0
12	Z	222/222 (100%)	0.07	6 (2%)	58 62	30, 44, 75, 109	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	M	233/246 (94%)	-0.08	5 (2%) 67 71	28, 43, 65, 83	0
13	a	233/246 (94%)	-0.05	5 (2%) 67 71	28, 42, 62, 79	0
14	N	196/196 (100%)	-0.08	4 (2%) 68 72	28, 39, 65, 94	0
14	b	196/196 (100%)	-0.02	4 (2%) 68 72	28, 40, 65, 99	0
15	c	2/5 (40%)	-0.33	0 100 100	52, 52, 52, 54	0
15	d	2/5 (40%)	-0.41	0 100 100	44, 44, 44, 51	0
All	All	6348/6624 (95%)	0.17	312 (4%) 33 38	25, 48, 93, 169	0

The worst 5 of 312 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	P	219	ALA	9.5
2	B	221	ASP	9.1
3	Q	49	THR	8.6
6	T	243	ILE	7.9
2	P	220	ASN	7.7

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
15	GAU	d	4	9/9	0.89	0.17	-	48,52,53,57	0
15	GAU	c	4	9/9	0.93	0.17	-	56,59,62,63	0

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron

density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
16	MG	I	301	1/1	0.91	0.20	2.10	63,63,63,63	0
16	MG	N	201	1/1	0.96	0.12	-0.90	47,47,47,47	0
16	MG	Z	301	1/1	0.96	0.12	-1.33	50,50,50,50	0
16	MG	G	301	1/1	0.95	0.08	-1.45	43,43,43,43	0
16	MG	K	301	1/1	0.94	0.08	-1.83	42,42,42,42	0
16	MG	V	301	1/1	0.96	0.06	-2.79	48,48,48,48	0
16	MG	Y	301	1/1	0.99	0.04	-3.87	37,37,37,37	0
17	CL	G	302	1/1	0.99	0.10	-	30,30,30,30	0

## 6.5 Other polymers

There are no such residues in this entry.