



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 3, 2021 – 01:11 PM GMT

PDB ID : 6YA4  
Title : Crystal structure of PnrA from *S. pneumoniae* in complex with cytidine  
Authors : Batuecas, M.T.; Hermoso, J.A.  
Deposited on : 2020-03-11  
Resolution : 2.20 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.16  
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) : 2.16

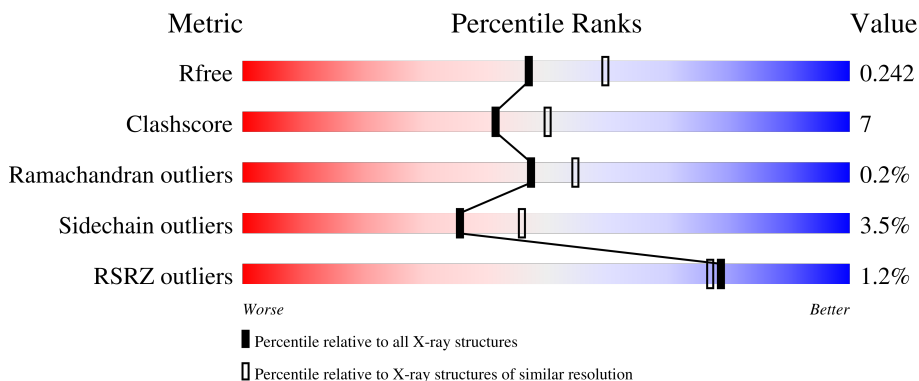
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	332	88% 11% .
1	B	332	84% 13% ..
1	C	332	86% 13% .
1	D	332	90% 9% ..

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	ACT	C	404	-	-	X	-

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 10516 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 Lipoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	C	332	2482	1551	424	506	1	0	0	0
1	A	332	2486	1554	424	507	1	0	1	0
1	B	330	2484	1555	425	503	1	0	2	0
1	D	330	2472	1546	422	503	1	0	0	0

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	2	GLY	-	expression tag	UNP A0A0H2UPF3
C	3	SER	-	expression tag	UNP A0A0H2UPF3
C	4	SER	-	expression tag	UNP A0A0H2UPF3
C	5	HIS	-	expression tag	UNP A0A0H2UPF3
C	6	HIS	-	expression tag	UNP A0A0H2UPF3
C	7	HIS	-	expression tag	UNP A0A0H2UPF3
C	8	HIS	-	expression tag	UNP A0A0H2UPF3
C	9	HIS	-	expression tag	UNP A0A0H2UPF3
C	10	HIS	-	expression tag	UNP A0A0H2UPF3
C	11	MET	-	expression tag	UNP A0A0H2UPF3
C	12	SER	-	expression tag	UNP A0A0H2UPF3
C	13	GLY	-	expression tag	UNP A0A0H2UPF3
C	14	GLU	-	expression tag	UNP A0A0H2UPF3
C	15	ASN	-	expression tag	UNP A0A0H2UPF3
C	16	LEU	-	expression tag	UNP A0A0H2UPF3
C	17	TYR	-	expression tag	UNP A0A0H2UPF3
C	18	PHE	-	expression tag	UNP A0A0H2UPF3
C	19	GLN	-	expression tag	UNP A0A0H2UPF3
C	20	GLY	-	expression tag	UNP A0A0H2UPF3
C	21	ALA	-	expression tag	UNP A0A0H2UPF3
C	22	SER	-	expression tag	UNP A0A0H2UPF3

*Continued on next page...*

*Continued from previous page...*

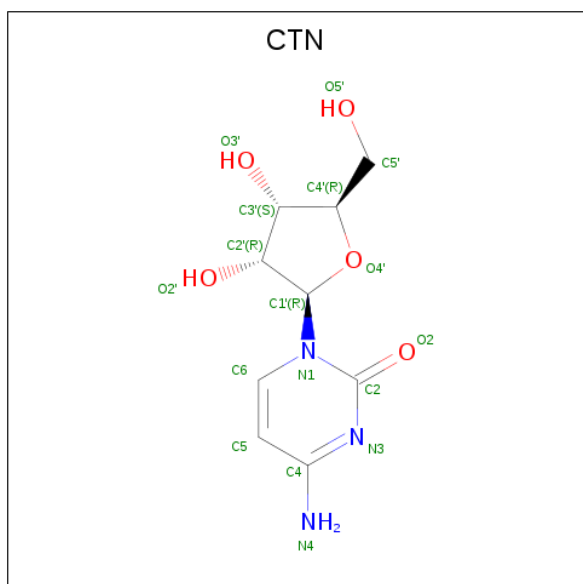
Chain	Residue	Modelled	Actual	Comment	Reference
A	2	GLY	-	expression tag	UNP A0A0H2UPF3
A	3	SER	-	expression tag	UNP A0A0H2UPF3
A	4	SER	-	expression tag	UNP A0A0H2UPF3
A	5	HIS	-	expression tag	UNP A0A0H2UPF3
A	6	HIS	-	expression tag	UNP A0A0H2UPF3
A	7	HIS	-	expression tag	UNP A0A0H2UPF3
A	8	HIS	-	expression tag	UNP A0A0H2UPF3
A	9	HIS	-	expression tag	UNP A0A0H2UPF3
A	10	HIS	-	expression tag	UNP A0A0H2UPF3
A	11	MET	-	expression tag	UNP A0A0H2UPF3
A	12	SER	-	expression tag	UNP A0A0H2UPF3
A	13	GLY	-	expression tag	UNP A0A0H2UPF3
A	14	GLU	-	expression tag	UNP A0A0H2UPF3
A	15	ASN	-	expression tag	UNP A0A0H2UPF3
A	16	LEU	-	expression tag	UNP A0A0H2UPF3
A	17	TYR	-	expression tag	UNP A0A0H2UPF3
A	18	PHE	-	expression tag	UNP A0A0H2UPF3
A	19	GLN	-	expression tag	UNP A0A0H2UPF3
A	20	GLY	-	expression tag	UNP A0A0H2UPF3
A	21	ALA	-	expression tag	UNP A0A0H2UPF3
A	22	SER	-	expression tag	UNP A0A0H2UPF3
B	2	GLY	-	expression tag	UNP A0A0H2UPF3
B	3	SER	-	expression tag	UNP A0A0H2UPF3
B	4	SER	-	expression tag	UNP A0A0H2UPF3
B	5	HIS	-	expression tag	UNP A0A0H2UPF3
B	6	HIS	-	expression tag	UNP A0A0H2UPF3
B	7	HIS	-	expression tag	UNP A0A0H2UPF3
B	8	HIS	-	expression tag	UNP A0A0H2UPF3
B	9	HIS	-	expression tag	UNP A0A0H2UPF3
B	10	HIS	-	expression tag	UNP A0A0H2UPF3
B	11	MET	-	expression tag	UNP A0A0H2UPF3
B	12	SER	-	expression tag	UNP A0A0H2UPF3
B	13	GLY	-	expression tag	UNP A0A0H2UPF3
B	14	GLU	-	expression tag	UNP A0A0H2UPF3
B	15	ASN	-	expression tag	UNP A0A0H2UPF3
B	16	LEU	-	expression tag	UNP A0A0H2UPF3
B	17	TYR	-	expression tag	UNP A0A0H2UPF3
B	18	PHE	-	expression tag	UNP A0A0H2UPF3
B	19	GLN	-	expression tag	UNP A0A0H2UPF3
B	20	GLY	-	expression tag	UNP A0A0H2UPF3
B	21	ALA	-	expression tag	UNP A0A0H2UPF3
B	22	SER	-	expression tag	UNP A0A0H2UPF3

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	2	GLY	-	expression tag	UNP A0A0H2UPF3
D	3	SER	-	expression tag	UNP A0A0H2UPF3
D	4	SER	-	expression tag	UNP A0A0H2UPF3
D	5	HIS	-	expression tag	UNP A0A0H2UPF3
D	6	HIS	-	expression tag	UNP A0A0H2UPF3
D	7	HIS	-	expression tag	UNP A0A0H2UPF3
D	8	HIS	-	expression tag	UNP A0A0H2UPF3
D	9	HIS	-	expression tag	UNP A0A0H2UPF3
D	10	HIS	-	expression tag	UNP A0A0H2UPF3
D	11	MET	-	expression tag	UNP A0A0H2UPF3
D	12	SER	-	expression tag	UNP A0A0H2UPF3
D	13	GLY	-	expression tag	UNP A0A0H2UPF3
D	14	GLU	-	expression tag	UNP A0A0H2UPF3
D	15	ASN	-	expression tag	UNP A0A0H2UPF3
D	16	LEU	-	expression tag	UNP A0A0H2UPF3
D	17	TYR	-	expression tag	UNP A0A0H2UPF3
D	18	PHE	-	expression tag	UNP A0A0H2UPF3
D	19	GLN	-	expression tag	UNP A0A0H2UPF3
D	20	GLY	-	expression tag	UNP A0A0H2UPF3
D	21	ALA	-	expression tag	UNP A0A0H2UPF3
D	22	SER	-	expression tag	UNP A0A0H2UPF3

- Molecule 2 is 4-AMINO-1-BETA-D-RIBOFURANOSYL-2(1H)-PYRIMIDINONE (three-letter code: CTN) (formula: C<sub>9</sub>H<sub>13</sub>N<sub>3</sub>O<sub>5</sub>).

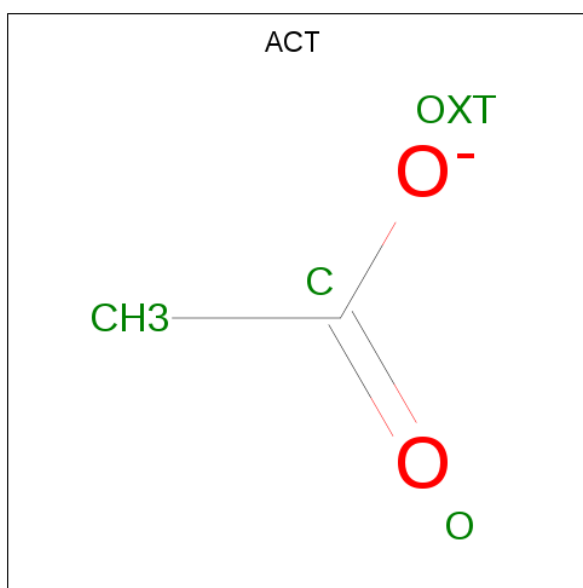


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	C	1	Total	C	N	O	0	0
			17	9	3	5		
2	A	1	Total	C	N	O	0	0
			17	9	3	5		
2	B	1	Total	C	N	O	0	0
			17	9	3	5		
2	D	1	Total	C	N	O	0	0
			17	9	3	5		

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

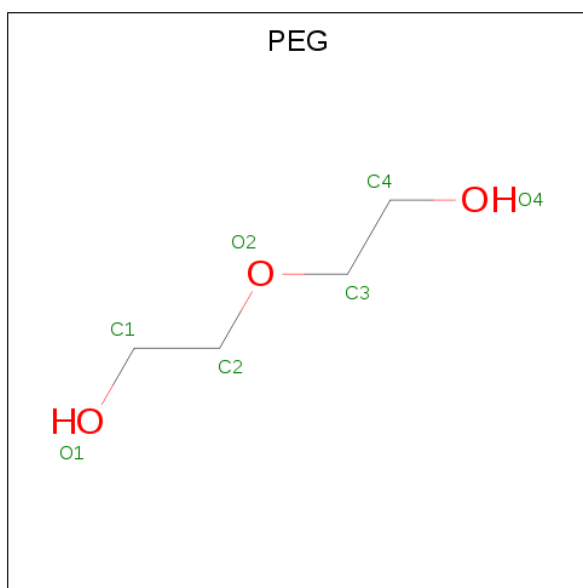
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Ni	0	0
			1	1		
3	A	2	Total	Ni	0	0
			2	2		
3	D	2	Total	Ni	0	0
			2	2		
3	C	2	Total	Ni	0	0
			2	2		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 7 4 3	0	0
5	B	1	Total C O 7 4 3	0	0

- Molecule 6 is water.


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	159	Total O 159 159	0	0
6	A	159	Total O 159 159	0	0
6	B	132	Total O 132 132	0	0
6	D	49	Total O 49 49	0	0

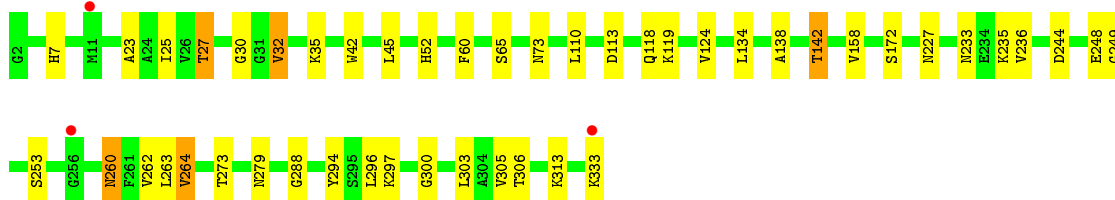


### 3 Residue-property plots [i](#)


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.

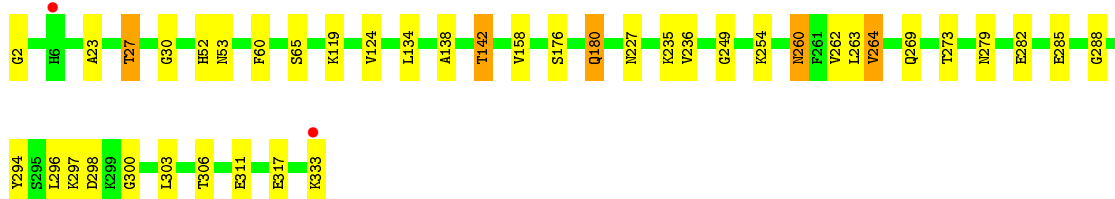
- Molecule 1: Lipoprotein

Chain C:  86% 13%




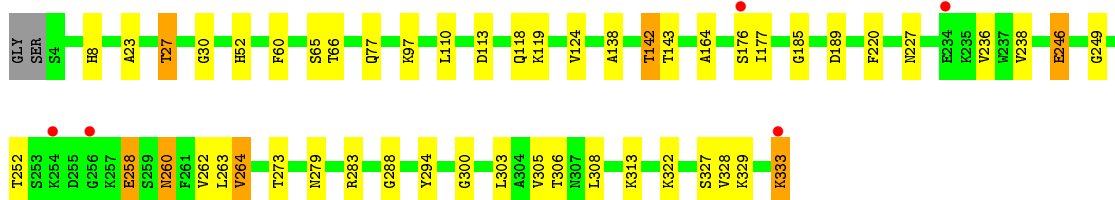
- Molecule 1: Lipoprotein

Chain A:  88% 11%

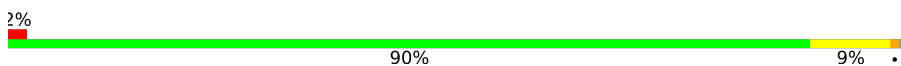


- Molecule 1: Lipoprotein

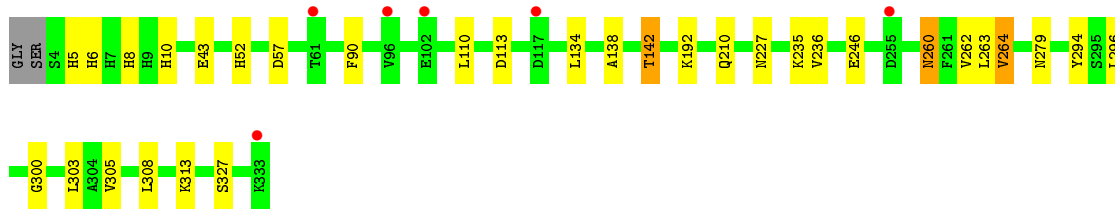
Chain B:  84% 13%



- Molecule 1: Lipoprotein

Chain D:  90% 9%





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.61Å 304.39Å 128.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.07 – 2.20 49.07 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.07-2.20) 99.7 (49.07-2.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.47 (at 2.20Å)	Xtrriage
Refinement program	REFMAC 7.0.078	Depositor
R, $R_{free}$	0.200 , 0.240 0.199 , 0.242	Depositor DCC
$R_{free}$ test set	3921 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.1	Xtrriage
Anisotropy	0.088	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 36.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10516	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 16.97% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: NI, CTN, PEG, ACT

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	0.76	1/2531 (0.0%)	0.83	0/3418
1	B	0.71	0/2532	0.87	2/3419 (0.1%)
1	C	0.75	0/2524	0.84	0/3408
1	D	0.71	0/2514	0.82	0/3395
All	All	0.73	1/10101 (0.0%)	0.84	2/13640 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	2	GLY	C-O	5.80	1.32	1.23

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	283[A]	ARG	NE-CZ-NH2	5.28	122.94	120.30
1	B	283[B]	ARG	NE-CZ-NH2	5.28	122.94	120.30

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	2486	0	2405	29	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	2484	0	2412	41	0
1	C	2482	0	2398	42	0
1	D	2472	0	2390	27	0
2	A	17	0	13	0	0
2	B	17	0	13	0	0
2	C	17	0	13	0	0
2	D	17	0	13	0	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
4	C	4	0	3	2	0
5	A	7	0	10	2	0
5	B	7	0	10	0	0
6	A	159	0	0	7	0
6	B	132	0	0	5	0
6	C	159	0	0	9	0
6	D	49	0	0	5	0
All	All	10516	0	9680	138	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

All (138) 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:A:227:ASN:HD21	1:A:236:VAL:H	1.10	1.00
1:A:282:GLU:OE2	6:A:501:HOH:O	1.84	0.95
1:C:227:ASN:HD21	1:C:236:VAL:H	1.03	0.91
1:C:118:GLN:HG3	6:C:581:HOH:O	1.69	0.90
1:B:164:ALA:HB2	1:B:333:LYS:HB2	1.53	0.88
1:B:227:ASN:HD21	1:B:236:VAL:H	1.23	0.87
1:B:264:VAL:HG13	1:B:303:LEU:HD11	1.54	0.87
1:A:298:ASP:OD2	5:A:404:PEG:H41	1.78	0.83
1:D:305:VAL:HG11	1:D:313:LYS:HG2	1.61	0.83
1:B:97:LYS:HG2	1:B:118:GLN:HE22	1.43	0.83
1:B:305:VAL:HG12	1:B:308:LEU:HB3	1.61	0.82
1:B:305:VAL:CG1	1:B:308:LEU:HB3	2.14	0.77
1:C:32:VAL:CG1	1:C:42:TRP:CD1	2.67	0.77
1:A:52:HIS:CD2	6:A:501:HOH:O	2.38	0.76
1:B:27:THR:HG21	1:B:30:GLY:O	1.86	0.75

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:32:VAL:HG13	1:C:42:TRP:CD1	2.22	0.74
1:C:227:ASN:HD21	1:C:236:VAL:N	1.84	0.72
1:A:27:THR:HG21	1:A:30:GLY:O	1.91	0.71
1:B:142:THR:O	6:B:501:HOH:O	2.09	0.70
1:D:6:HIS:N	6:D:501:HOH:O	2.16	0.70
1:B:305:VAL:HG11	1:B:313:LYS:HG2	1.73	0.70
1:C:313:LYS:NZ	6:C:503:HOH:O	2.26	0.68
1:B:246:GLU:HG3	1:B:306:THR:HG21	1.76	0.68
1:C:27:THR:HG21	1:C:30:GLY:O	1.93	0.68
1:A:297:LYS:HE2	6:A:611:HOH:O	1.93	0.67
1:D:227:ASN:HD21	1:D:236:VAL:H	1.42	0.66
1:B:264:VAL:CG1	1:B:303:LEU:HD11	2.25	0.66
1:C:264:VAL:HG13	1:C:303:LEU:HD11	1.77	0.65
1:B:164:ALA:CB	1:B:333:LYS:HB2	2.25	0.65
1:C:333:LYS:HG3	6:C:536:HOH:O	1.98	0.64
1:C:294:TYR:HB3	1:C:300:GLY:HA3	1.79	0.64
1:B:305:VAL:CG1	1:B:308:LEU:CB	2.76	0.64
1:D:305:VAL:CG1	1:D:313:LYS:HG2	2.28	0.63
1:A:138:ALA:O	1:A:142:THR:HG23	1.99	0.63
1:A:227:ASN:HD21	1:A:236:VAL:N	1.90	0.63
1:B:305:VAL:HG12	1:B:308:LEU:CB	2.30	0.61
1:D:138:ALA:O	1:D:142:THR:HG23	2.00	0.61
1:D:305:VAL:HG12	1:D:305:VAL:O	2.01	0.61
1:B:294:TYR:HB3	1:B:300:GLY:HA3	1.82	0.60
1:A:260:ASN:ND2	1:A:262:VAL:H	2.00	0.60
1:C:172:SER:O	4:C:404:ACT:C	2.50	0.60
1:B:260:ASN:ND2	1:B:262:VAL:H	1.99	0.60
1:A:294:TYR:HB3	1:A:300:GLY:HA3	1.84	0.59
1:C:32:VAL:HG11	1:C:42:TRP:CD1	2.36	0.59
1:A:298:ASP:OD2	5:A:404:PEG:C4	2.50	0.59
1:D:57:ASP:HA	6:D:503:HOH:O	2.02	0.59
1:A:264:VAL:HG13	1:A:303:LEU:HD11	1.84	0.58
1:B:97:LYS:HG2	1:B:118:GLN:NE2	2.14	0.58
1:D:110:LEU:HD21	1:D:113:ASP:HB3	1.85	0.58
1:C:158:VAL:HG23	6:C:613:HOH:O	2.04	0.58
1:A:138:ALA:O	1:A:142:THR:CG2	2.52	0.57
1:C:253:SER:O	6:C:501:HOH:O	2.17	0.57
1:C:260:ASN:ND2	1:C:262:VAL:H	2.03	0.57
1:A:27:THR:HG22	1:A:65:SER:H	1.70	0.56
1:B:322:LYS:HB3	1:B:328:VAL:HG23	1.86	0.56
1:C:73:ASN:ND2	6:C:506:HOH:O	2.38	0.56

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:264:VAL:HG13	1:D:303:LEU:HD11	1.88	0.56
1:B:8:HIS:HD2	6:B:518:HOH:O	1.89	0.56
1:C:110:LEU:HD21	1:C:113:ASP:HB3	1.87	0.55
1:C:124:VAL:HG11	1:C:273:THR:HG21	1.86	0.55
1:D:260:ASN:ND2	1:D:262:VAL:H	2.05	0.55
1:B:260:ASN:HD22	1:B:262:VAL:H	1.54	0.55
1:A:269:GLN:O	1:A:273[B]:THR:HG23	2.07	0.54
1:C:35:LYS:NZ	1:C:248:GLU:OE2	2.38	0.54
1:D:260:ASN:HD22	1:D:262:VAL:H	1.55	0.54
1:D:52:HIS:HE1	1:D:279:ASN:OD1	1.90	0.54
1:C:138:ALA:O	1:C:142:THR:HG23	2.08	0.54
1:D:294:TYR:HB3	1:D:300:GLY:HA3	1.89	0.53
1:B:110:LEU:HD21	1:B:113:ASP:HB3	1.91	0.53
1:B:138:ALA:O	1:B:142:THR:HG23	2.08	0.53
1:A:260:ASN:HD22	1:A:262:VAL:H	1.57	0.52
1:A:23:ALA:O	1:A:60:PHE:HA	2.09	0.52
1:B:227:ASN:ND2	1:B:236:VAL:H	2.00	0.52
1:C:227:ASN:ND2	1:C:235:LYS:HA	2.25	0.52
1:D:8:HIS:HE1	1:D:10:HIS:HE1	1.57	0.51
1:B:97:LYS:NZ	6:B:508:HOH:O	2.43	0.51
1:A:317:GLU:HG2	6:A:650:HOH:O	2.10	0.50
1:C:142:THR:HG21	1:C:263:LEU:HD21	1.93	0.49
1:D:8:HIS:HD2	6:D:540:HOH:O	1.96	0.49
1:B:138:ALA:O	1:B:142:THR:CG2	2.60	0.49
1:A:227:ASN:ND2	1:A:235:LYS:HA	2.28	0.48
1:C:52:HIS:HE1	1:C:279:ASN:OD1	1.96	0.48
1:B:27:THR:HG22	1:B:65:SER:H	1.78	0.48
1:D:142:THR:HG21	1:D:263:LEU:HD21	1.96	0.48
1:C:248:GLU:O	6:C:502:HOH:O	2.19	0.48
1:C:233:ASN:ND2	6:C:512:HOH:O	2.47	0.48
1:D:43:GLU:O	6:D:502:HOH:O	2.20	0.47
1:D:138:ALA:O	1:D:142:THR:CG2	2.62	0.47
1:D:227:ASN:HD21	1:D:236:VAL:HG22	1.80	0.47
1:B:249:GLY:O	1:B:260:ASN:HA	2.15	0.47
1:C:110:LEU:CD2	1:C:113:ASP:HB3	2.45	0.47
1:C:138:ALA:O	1:C:142:THR:CG2	2.63	0.47
1:D:227:ASN:ND2	1:D:236:VAL:H	2.11	0.47
1:B:142:THR:HG21	1:B:263:LEU:HD21	1.96	0.47
1:B:252:THR:HG22	1:B:258:GLU:HG3	1.97	0.47
1:B:249:GLY:HA3	1:B:262:VAL:HG23	1.97	0.46
1:B:124:VAL:HG11	1:B:273:THR:HG21	1.98	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:124:VAL:HG11	1:A:273[A]:THR:HG21	1.97	0.46
1:D:110:LEU:CD2	1:D:113:ASP:HB3	2.45	0.46
1:B:66:THR:O	6:B:502:HOH:O	2.21	0.45
1:C:260:ASN:HD22	1:C:262:VAL:H	1.64	0.45
1:A:158:VAL:HG23	6:A:619:HOH:O	2.17	0.45
1:A:52:HIS:HE1	1:A:279:ASN:OD1	2.00	0.45
1:C:27:THR:HG22	1:C:65:SER:H	1.82	0.45
1:C:23:ALA:O	1:C:60:PHE:HA	2.17	0.45
1:B:52:HIS:HE1	1:B:279:ASN:OD1	1.99	0.45
1:D:305:VAL:CG1	1:D:308:LEU:HB3	2.47	0.45
1:A:142:THR:HG21	1:A:263:LEU:HD21	1.99	0.45
1:A:285:GLU:OE2	6:A:502:HOH:O	2.21	0.45
1:C:249:GLY:O	1:C:260:ASN:HA	2.17	0.44
1:D:134:LEU:HD11	1:D:296:LEU:HD21	1.98	0.44
1:C:244:ASP:HB3	1:A:53:ASN:HD21	1.83	0.44
1:C:134:LEU:HD11	1:C:296:LEU:HD21	2.00	0.44
1:B:185:GLY:HA2	6:B:600:HOH:O	2.17	0.44
1:D:8:HIS:CE1	1:D:10:HIS:CE1	3.06	0.44
1:D:210:GLN:H	1:D:210:GLN:HG3	1.67	0.43
1:C:249:GLY:HA3	1:C:262:VAL:HG23	1.99	0.43
1:C:297:LYS:HG3	6:C:572:HOH:O	2.18	0.43
1:A:249:GLY:O	1:A:260:ASN:HA	2.19	0.43
1:A:119:LYS:O	1:A:288:GLY:HA3	2.19	0.43
1:B:220:PHE:CE1	1:B:238[A]:VAL:HG11	2.54	0.43
1:B:143:THR:CG2	1:B:177:ILE:HD13	2.49	0.42
1:B:119:LYS:O	1:B:288:GLY:HA3	2.20	0.42
1:C:172:SER:O	4:C:404:ACT:OXT	2.37	0.42
1:B:110:LEU:CD2	1:B:113:ASP:HB3	2.49	0.42
1:B:23:ALA:O	1:B:60:PHE:HA	2.20	0.42
1:A:134:LEU:HD11	1:A:296:LEU:HD21	2.02	0.42
1:C:142:THR:CG2	1:C:263:LEU:HD21	2.49	0.41
1:C:32:VAL:CG1	1:C:42:TRP:CG	3.03	0.41
1:C:25:ILE:HG23	1:C:45:LEU:HD22	2.02	0.41
1:D:90:PHE:N	6:D:511:HOH:O	2.52	0.41
1:A:180:GLN:NE2	6:A:517:HOH:O	2.54	0.41
1:C:119:LYS:O	1:C:288:GLY:HA3	2.21	0.40
1:C:32:VAL:HG13	1:C:42:TRP:CG	2.55	0.40
1:B:143:THR:HG21	1:B:177:ILE:HD13	2.04	0.40
1:B:142:THR:CG2	1:B:263:LEU:HD21	2.52	0.40
1:C:296:LEU:HD12	1:C:296:LEU:HA	1.96	0.40
1:D:52:HIS:CE1	1:D:279:ASN:OD1	2.73	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	331/332 (100%)	312 (94%)	18 (5%)	1 (0%)	41	46
1	B	330/332 (99%)	315 (96%)	15 (4%)	0	100	100
1	C	330/332 (99%)	309 (94%)	20 (6%)	1 (0%)	41	46
1	D	328/332 (99%)	309 (94%)	19 (6%)	0	100	100
All	All	1319/1328 (99%)	1245 (94%)	72 (6%)	2 (0%)	47	55

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	306	THR
1	C	306	THR

### 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	256/255 (100%)	247 (96%)	9 (4%)	36	46
1	B	256/255 (100%)	244 (95%)	12 (5%)	26	33
1	C	255/255 (100%)	248 (97%)	7 (3%)	44	57
1	D	254/255 (100%)	246 (97%)	8 (3%)	40	51
All	All	1021/1020 (100%)	985 (96%)	36 (4%)	36	46

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	7	HIS
1	C	27	THR
1	C	32	VAL
1	C	142	THR
1	C	260	ASN
1	C	264	VAL
1	C	305	VAL
1	A	27	THR
1	A	142	THR
1	A	176	SER
1	A	180	GLN
1	A	254	LYS
1	A	260	ASN
1	A	264	VAL
1	A	311	GLU
1	A	333	LYS
1	B	27	THR
1	B	77	GLN
1	B	142	THR
1	B	176	SER
1	B	189	ASP
1	B	246	GLU
1	B	258	GLU
1	B	260	ASN
1	B	264	VAL
1	B	327	SER
1	B	329	LYS
1	B	333	LYS
1	D	5	HIS
1	D	142	THR
1	D	192	LYS
1	D	235	LYS
1	D	246	GLU
1	D	260	ASN
1	D	264	VAL
1	D	327	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (31) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	5	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	C	7	HIS
1	C	9	HIS
1	C	38	ASN
1	C	52	HIS
1	C	64	GLN
1	C	227	ASN
1	C	260	ASN
1	A	5	HIS
1	A	15	ASN
1	A	38	ASN
1	A	52	HIS
1	A	53	ASN
1	A	64	GLN
1	A	227	ASN
1	A	260	ASN
1	B	8	HIS
1	B	38	ASN
1	B	52	HIS
1	B	118	GLN
1	B	227	ASN
1	B	260	ASN
1	D	7	HIS
1	D	8	HIS
1	D	9	HIS
1	D	10	HIS
1	D	38	ASN
1	D	52	HIS
1	D	76	GLN
1	D	227	ASN
1	D	260	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 7 are monoatomic - leaving 7 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	CTN	A	401	-	15,18,18	1.42	1 (6%)	17,26,26	1.59	3 (17%)
5	PEG	A	404	-	6,6,6	0.28	0	5,5,5	0.34	0
2	CTN	D	401	-	15,18,18	0.91	1 (6%)	17,26,26	1.82	3 (17%)
2	CTN	C	401	-	15,18,18	1.45	3 (20%)	17,26,26	1.34	2 (11%)
2	CTN	B	401	-	15,18,18	1.21	2 (13%)	17,26,26	1.59	2 (11%)
5	PEG	B	403	-	6,6,6	0.16	0	5,5,5	0.11	0
4	ACT	C	404	-	1,3,3	4.33	1 (100%)	0,3,3	0.00	-

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CTN	A	401	-	-	0/4/22/22	0/2/2/2
5	PEG	A	404	-	-	3/4/4/4	-
2	CTN	D	401	-	-	1/4/22/22	0/2/2/2
2	CTN	C	401	-	-	0/4/22/22	0/2/2/2
2	CTN	B	401	-	-	0/4/22/22	0/2/2/2
5	PEG	B	403	-	-	4/4/4/4	-

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	404	ACT	CH3-C	4.33	1.54	1.48

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	401	CTN	C6-N1	-3.04	1.32	1.35
2	A	401	CTN	O4'-C1'	3.03	1.45	1.41
2	C	401	CTN	C2'-C1'	-3.03	1.49	1.53
2	B	401	CTN	C2-N3	-2.78	1.32	1.38
2	B	401	CTN	O4'-C1'	2.53	1.44	1.41
2	D	401	CTN	C6-N1	-2.43	1.32	1.35
2	C	401	CTN	C5-C4	-2.22	1.36	1.41

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	401	CTN	N4-C4-N3	4.89	124.22	116.49
2	A	401	CTN	C2-N3-C4	4.80	121.21	116.34
2	B	401	CTN	C2-N3-C4	4.79	121.19	116.34
2	D	401	CTN	C5-C4-N4	-3.27	115.47	121.14
2	C	401	CTN	N4-C4-N3	3.17	121.51	116.49
2	C	401	CTN	C5-C4-N4	-2.96	116.00	121.14
2	B	401	CTN	N4-C4-N3	2.74	120.83	116.49
2	A	401	CTN	N4-C4-N3	2.54	120.51	116.49
2	D	401	CTN	C2-N3-C4	2.17	118.54	116.34
2	A	401	CTN	O2'-C2'-C3'	-2.10	105.03	111.82

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	404	PEG	O1-C1-C2-O2
5	B	403	PEG	O1-C1-C2-O2
5	B	403	PEG	O2-C3-C4-O4
5	A	404	PEG	C4-C3-O2-C2
5	B	403	PEG	C1-C2-O2-C3
5	B	403	PEG	C4-C3-O2-C2
5	A	404	PEG	C1-C2-O2-C3
2	D	401	CTN	O4'-C4'-C5'-O5'

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	404	PEG	2	0
4	C	404	ACT	2	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, 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	332/332 (100%)	-0.47	2 (0%) 89 88	17, 34, 58, 91	0
1	B	330/332 (99%)	-0.23	5 (1%) 73 72	20, 41, 80, 113	0
1	C	332/332 (100%)	-0.45	3 (0%) 84 83	16, 34, 66, 85	0
1	D	330/332 (99%)	0.04	6 (1%) 68 66	30, 58, 93, 105	0
All	All	1324/1328 (99%)	-0.28	16 (1%) 79 77	16, 41, 79, 113	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	333	LYS	5.9
1	A	333	LYS	4.4
1	C	333	LYS	4.3
1	B	333	LYS	4.3
1	C	256	GLY	4.2
1	D	117	ASP	3.2
1	A	6	HIS	3.1
1	B	254	LYS	2.8
1	B	176	SER	2.6
1	D	61	THR	2.5
1	D	96	VAL	2.5
1	B	234	GLU	2.4
1	C	11	MET	2.3
1	B	256	GLY	2.1
1	D	255	ASP	2.1
1	D	102	GLU	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 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<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	B-factors(Å <sup>2</sup> )	Q<0.9
5	PEG	A	404	7/7	0.77	0.18	62,64,68,75	0
5	PEG	B	403	7/7	0.91	0.15	44,46,56,59	0
4	ACT	C	404	4/4	0.91	0.18	57,58,62,63	0
2	CTN	D	401	17/17	0.95	0.08	36,41,45,45	0
2	CTN	B	401	17/17	0.97	0.10	21,24,30,30	0
2	CTN	A	401	17/17	0.98	0.09	18,21,26,26	0
3	NI	A	402	1/1	0.98	0.10	37,37,37,37	0
3	NI	D	403	1/1	0.99	0.14	99,99,99,99	0
3	NI	C	402	1/1	0.99	0.13	21,21,21,21	0
3	NI	B	402	1/1	0.99	0.13	16,16,16,16	0
2	CTN	C	401	17/17	0.99	0.09	18,22,23,25	0
3	NI	A	403	1/1	0.99	0.10	37,37,37,37	0
3	NI	C	403	1/1	1.00	0.13	20,20,20,20	0
3	NI	D	402	1/1	1.00	0.10	25,25,25,25	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.