



# Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 01:29 AM GMT

PDB ID : 3D4L  
Title : Human dipeptidyl peptidase IV/CD26 in complex with a novel inhibitor  
Authors : Scapin, G.  
Deposited on : 2008-05-14  
Resolution : 2.00 Å(reported)

This is a full wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

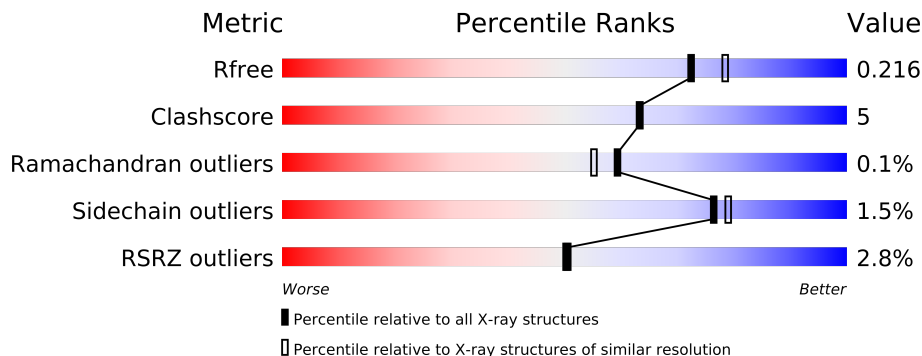
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.00 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))
$R_{free}$	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	728	
1	B	728	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	NAG	A	1520	-	X
3	NAG	B	2150	-	X
3	NAG	B	2321	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 13340 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Dipeptidyl peptidase 4 soluble form.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	728	Total	C	N	O	S	0	3	0
			5987	3840	987	1134	26			
1	B	728	Total	C	N	O	S	0	2	0
			5982	3839	986	1131	26			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	THR	SER	ENGINEERED	UNP P27487
B	39	THR	SER	ENGINEERED	UNP P27487

- Molecule 2 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	2	Total	C	N	O	0	0
			28	16	2	10		
2	A	2	Total	C	N	O	0	0
			28	16	2	10		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	THR	SER	ENGINEERED	UNP P27487
A	39	THR	SER	ENGINEERED	UNP P27487

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	2	Total	C	N	O	0	0
			28	16	2	10		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	THR	SER	ENGINEERED	UNP P27487

- Molecule 5 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	2	Total	C	N	O	0	0
			28	16	2	10		
5	A	2	Total	C	N	O	0	0
			28	16	2	10		
5	B	2	Total	C	N	O	0	0
			28	16	2	10		
5	B	2	Total	C	N	O	0	0
			28	16	2	10		
5	B	2	Total	C	N	O	0	0
			28	16	2	10		

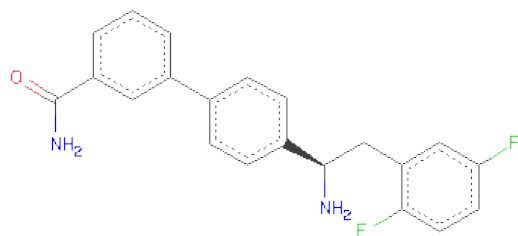
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	39	THR	SER	ENGINEERED	UNP P27487
A	39	THR	SER	ENGINEERED	UNP P27487
B	39	THR	SER	ENGINEERED	UNP P27487
B	39	THR	SER	ENGINEERED	UNP P27487
B	39	THR	SER	ENGINEERED	UNP P27487
B	39	THR	SER	ENGINEERED	UNP P27487

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Na	0	0
			1	1		

- Molecule 7 is 4'-[(1R)-1-AMINO-2-(2,5-DIFLUOROPHENYL)ETHYL]BIPHENYL-3-CARBOXAMIDE (three-letter code: 605) (formula: C<sub>21</sub>H<sub>18</sub>F<sub>2</sub>N<sub>2</sub>O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	A	1	Total	C	F	N	O	0	0
			26	21	2	2	1		
7	B	1	Total	C	F	N	O	0	0
			26	21	2	2	1		

- Molecule 8 is water.

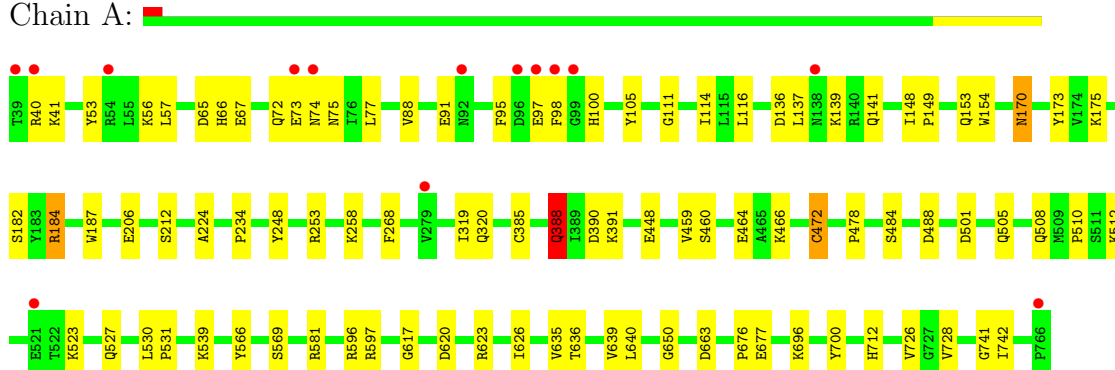
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	500	Total	O	0	0
			500	500		
8	B	482	Total	O	0	0
			482	482		

### 3 Residue-property plots

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

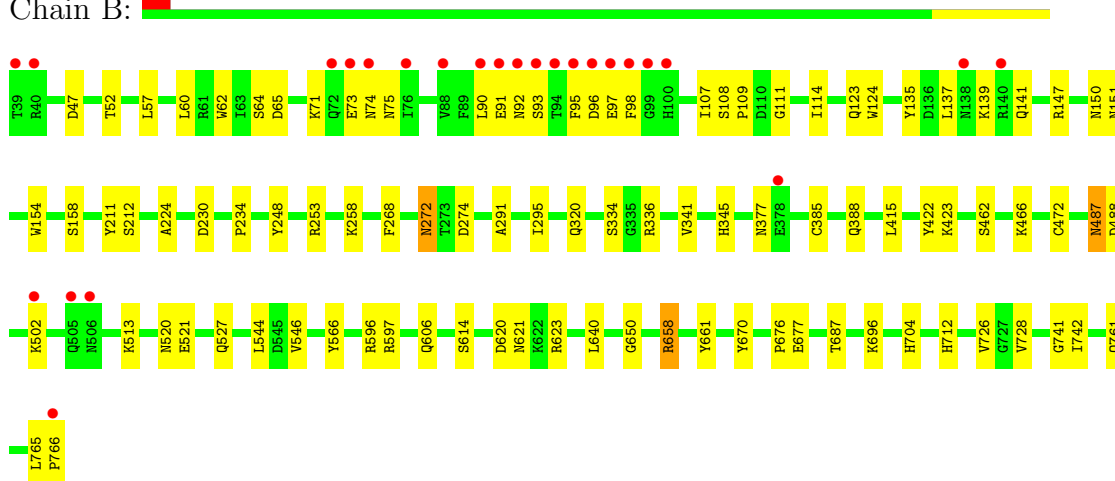
- Molecule 1: Dipeptidyl peptidase 4 soluble form

Chain A:



- Molecule 1: Dipeptidyl peptidase 4 soluble form

Chain B:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.44Å 125.75Å 137.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00 46.35 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.8 (30.00-2.00) 98.5 (46.35-2.00)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.96 (at 2.00Å)	Xtriage
Refinement program	CNX	Depositor
R, $R_{free}$	0.192 , 0.217 0.191 , 0.216	Depositor DCC
$R_{free}$ test set	6944 reflections (5.35%)	DCC
Wilson B-factor (Å <sup>2</sup> )	30.8	Xtriage
Anisotropy	0.396	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 44.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	1 of 136796 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13340	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% of the height of the origin peak. No significant pseudotranslation is detected.*

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



## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 605, NA, NAG, NDG

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	0.50	0/6162	0.69	2/8380 (0.0%)
1	B	0.48	0/6155	0.68	1/8371 (0.0%)
All	All	0.49	0/12317	0.69	3/16751 (0.0%)

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
1	A	0	1
1	B	0	1
All	All	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	388	GLN	N-CA-C	-5.37	96.51	111.00
1	A	319	ILE	N-CA-C	-5.27	96.77	111.00
1	B	388	GLN	N-CA-C	-5.10	97.23	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	700	TYR	Sidechain
1	B	211	TYR	Sidechain

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	5987	0	5690	64	0
1	B	5982	0	5693	67	0
2	A	56	0	50	2	0
3	A	42	0	39	5	0
3	B	42	0	39	1	0
4	A	28	0	25	2	0
5	A	56	0	50	1	0
5	B	112	0	100	0	0
6	A	1	0	0	0	0
7	A	26	0	18	0	0
7	B	26	0	18	0	0
8	A	500	0	0	1	0
8	B	482	0	0	7	0
All	All	13340	0	11722	132	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 5.

All (132) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:253:ARG:HH21	1:B:253:ARG:HH21	1.21	0.85
1:A:581:ARG:CZ	3:A:1520:NAG:H62	2.07	0.83
1:A:75:ASN:HD21	3:A:1092:NAG:HN2	1.23	0.82
1:A:66[B]:HIS:CA	1:A:67:GLU:N	2.45	0.80
1:B:147:ARG:HE	3:B:2150:NAG:H83	1.46	0.79
1:A:596:ARG:O	1:A:597:ARG:HD2	1.89	0.72
1:A:696:LYS:HG2	1:A:728:VAL:HG22	1.71	0.72
1:A:258:LYS:NZ	1:A:712:HIS:HD2	1.88	0.71
1:B:258:LYS:HZ3	1:B:712:HIS:HD2	1.37	0.70
1:A:173:TYR:CE2	1:A:184:ARG:HG2	2.27	0.70
1:B:272:ASN:C	1:B:272:ASN:HD22	1.96	0.69
1:B:91:GLU:HG3	1:B:93:SER:H	1.57	0.68
1:B:640:LEU:HD11	1:B:650:GLY:HA3	1.74	0.68
1:B:658:ARG:HG2	1:B:661:TYR:CE2	2.31	0.66

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:A:1230:NDG:O7	2:A:1230:NDG:H3	1.95	0.65
1:B:258:LYS:NZ	1:B:712:HIS:HD2	1.95	0.63
1:B:596:ARG:O	1:B:597:ARG:HD2	1.98	0.63
1:B:676:PRO:HG2	1:B:677:GLU:OE2	1.98	0.63
1:B:65:ASP:OD2	1:B:466:LYS:HB2	1.99	0.63
1:A:139:LYS:HG3	1:A:141:GLN:HB3	1.82	0.62
1:B:726:VAL:HG23	1:B:728:VAL:HG23	1.82	0.61
1:B:597:ARG:HG2	8:B:2772:HOH:O	2.02	0.59
1:A:175:LYS:HG3	1:A:182:SER:HB3	1.85	0.59
1:B:139:LYS:O	1:B:141:GLN:HG3	2.02	0.59
1:A:149:PRO:HA	4:A:1150:NDG:H8C1	1.84	0.58
1:B:502:LYS:C	1:B:502:LYS:HD3	2.23	0.58
1:B:272:ASN:HD21	1:B:274:ASP:HB2	1.69	0.58
1:B:502:LYS:O	1:B:502:LYS:HD3	2.02	0.58
1:A:676:PRO:HG2	1:A:677:GLU:OE2	2.04	0.58
1:A:581:ARG:NH1	3:A:1520:NAG:H62	2.18	0.57
1:B:95:PHE:HB3	1:B:98:PHE:HB2	1.86	0.57
1:A:640:LEU:HD11	1:A:650:GLY:HA3	1.85	0.57
1:A:40:ARG:HG3	1:A:508:GLN:HG3	1.85	0.57
1:B:114:ILE:CG2	1:B:135:TYR:HB3	2.35	0.57
1:B:258:LYS:HZ3	1:B:712:HIS:CD2	2.19	0.56
1:B:114:ILE:HG23	1:B:135:TYR:HB3	1.87	0.56
1:A:170:ASN:N	1:A:170:ASN:HD22	2.03	0.56
1:A:77:LEU:HD23	1:A:88:VAL:HA	1.88	0.55
1:B:154:TRP:CE2	1:B:212:SER:HB3	2.42	0.55
1:B:614:SER:HB2	1:B:621:ASN:OD1	2.07	0.54
1:B:620:ASP:OD2	1:B:623:ARG:HD3	2.08	0.54
1:A:136:ASP:HB3	1:A:139:LYS:HG2	1.89	0.54
1:A:258:LYS:HZ3	1:A:712:HIS:HD2	1.56	0.54
1:B:75:ASN:ND2	1:B:92:ASN:H	2.06	0.54
4:A:1150:NDG:O3	4:A:1151:NAG:N2	2.42	0.53
1:A:620:ASP:OD2	1:A:623:ARG:HD3	2.08	0.53
1:A:153:GLN:HE22	1:A:170:ASN:ND2	2.07	0.52
1:B:658:ARG:HG3	1:B:658:ARG:O	2.09	0.52
1:A:65[A]:ASP:OD2	1:A:466:LYS:HB2	2.09	0.52
1:A:111:GLY:O	1:A:137:LEU:HD12	2.10	0.52
1:B:341:VAL:HG12	8:B:2773:HOH:O	2.08	0.52
1:B:93:SER:HA	1:B:96:ASP:OD1	2.10	0.51
1:B:60:LEU:HD12	1:B:60:LEU:C	2.32	0.50
1:A:258:LYS:HZ1	1:A:712:HIS:HD2	1.58	0.50
1:B:658:ARG:HG2	1:B:661:TYR:CD2	2.47	0.50
1:B:422:TYR:CE2	1:B:423:LYS:HE2	2.47	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:A:1085:NAG:HO3	2:A:1086:NDG:H6	1.60	0.49
1:B:272:ASN:ND2	1:B:274:ASP:H	2.10	0.49
5:A:1321:NAG:H61	5:A:1322:NAG:O5	2.12	0.49
1:B:544:LEU:HG	1:B:546[A]:VAL:CG2	2.41	0.49
1:A:73:GLU:HB3	3:A:1092:NAG:H3	1.94	0.49
1:B:696:LYS:CG	1:B:728:VAL:HG22	2.43	0.49
1:B:487:ASN:HD22	1:B:487:ASN:C	2.15	0.48
1:B:272:ASN:HD22	1:B:274:ASP:H	1.62	0.48
1:B:334:SER:O	1:B:336:ARG:HG2	2.13	0.48
1:B:109:PRO:HG2	1:B:158:SER:O	2.13	0.48
1:B:741:GLY:O	1:B:742:ILE:C	2.52	0.48
1:A:73:GLU:O	1:A:74:ASN:HB2	2.14	0.48
1:B:150:ASN:O	1:B:151:ASN:HB2	2.14	0.48
1:A:510:PRO:HD3	1:A:569:SER:HB2	1.96	0.48
1:B:73:GLU:O	1:B:74:ASN:HB2	2.13	0.47
1:B:71:LYS:NZ	8:B:2408:HOH:O	2.46	0.47
1:B:520:ASN:O	1:B:521:GLU:HB2	2.15	0.47
1:A:95:PHE:HB3	1:A:98:PHE:HB2	1.97	0.47
1:A:539:LYS:HD3	1:A:617:GLY:O	2.15	0.47
1:A:184:ARG:HH11	1:A:187:TRP:HA	1.81	0.46
1:A:248:TYR:CZ	1:B:234:PRO:HB2	2.51	0.46
1:A:56:LYS:C	1:A:57:LEU:HD12	2.35	0.46
1:A:88:VAL:HG11	1:A:91:GLU:OE2	2.17	0.45
1:A:136:ASP:CG	1:A:139:LYS:HG2	2.37	0.45
1:A:390:ASP:O	1:A:391:LYS:HD2	2.16	0.45
1:A:459:VAL:HG22	1:A:460:SER:N	2.32	0.45
1:B:107:ILE:HG22	1:B:108:SER:O	2.16	0.45
1:B:415:LEU:HD23	1:B:415:LEU:C	2.36	0.45
1:A:726:VAL:HG23	1:A:728:VAL:HG23	1.98	0.45
1:B:696:LYS:HG3	1:B:728:VAL:HG22	1.99	0.45
1:B:345:HIS:HD2	8:B:2452:HOH:O	2.00	0.45
1:A:472:CYS:O	1:A:478:PRO:HA	2.17	0.45
1:A:57:LEU:HB3	8:A:1607:HOH:O	2.17	0.44
1:A:105:TYR:HB2	1:A:114:ILE:HD11	1.98	0.44
1:A:136:ASP:CB	1:A:139:LYS:HG2	2.47	0.44
1:B:513:LYS:O	1:B:527:GLN:HA	2.18	0.44
1:A:258:LYS:HZ3	1:A:712:HIS:CD2	2.35	0.44
1:A:65[A]:ASP:CG	1:A:464:GLU:HB2	2.38	0.44
1:A:741:GLY:O	1:A:742:ILE:C	2.57	0.43
1:B:60:LEU:HD12	1:B:60:LEU:O	2.18	0.43
1:B:253:ARG:HD3	8:B:2794:HOH:O	2.18	0.43
1:A:484[A]:SER:O	1:A:488:ASP:HA	2.19	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:95:PHE:CE1	1:A:116:LEU:HD11	2.53	0.43
1:A:224:ALA:HB1	1:A:268:PHE:CZ	2.54	0.43
1:B:47:ASP:HA	1:B:52:THR:HG23	2.00	0.43
1:B:224:ALA:HB1	1:B:268:PHE:CZ	2.53	0.43
1:A:41:LYS:HE2	1:A:53:TYR:OH	2.18	0.43
1:B:765:LEU:HA	1:B:766:PRO:HD3	1.90	0.42
1:B:62:TRP:CG	1:B:462:SER:HA	2.54	0.42
1:A:206:GLU:OE2	1:A:663:ASP:OD2	2.37	0.42
1:A:388:GLN:HB2	1:A:391:LYS:HB2	2.00	0.42
1:A:388:GLN:CB	1:A:391:LYS:HB2	2.48	0.42
1:B:64:SER:HB2	8:B:2506:HOH:O	2.19	0.42
1:B:291:ALA:O	1:B:295:ILE:HG23	2.19	0.42
1:B:123:GLN:HG2	1:B:124:TRP:N	2.35	0.42
1:A:98:PHE:CE2	1:A:100:HIS:HB2	2.55	0.42
1:A:523:LYS:H	1:A:523:LYS:HG2	1.59	0.42
1:B:272:ASN:C	1:B:272:ASN:ND2	2.69	0.41
1:A:626:ILE:HG23	1:A:636:THR:HG23	2.03	0.41
1:B:111:GLY:O	1:B:137:LEU:HD12	2.20	0.41
1:A:66[B]:HIS:CD2	1:A:67:GLU:HG3	2.56	0.41
1:A:148:ILE:HA	1:A:149:PRO:HD3	1.89	0.41
1:A:635:VAL:O	1:A:639:VAL:HG23	2.20	0.41
1:B:377:ASN:HB2	8:B:2684:HOH:O	2.21	0.41
1:A:530:LEU:HA	1:A:531:PRO:HD3	1.98	0.41
1:A:234:PRO:HB2	1:B:248:TYR:CZ	2.56	0.41
1:A:512:LYS:HE3	1:A:527:GLN:CD	2.41	0.41
1:A:581:ARG:NH2	3:A:1520:NAG:H62	2.35	0.40
1:B:658:ARG:HB2	1:B:687:THR:HG22	2.03	0.40
1:A:72:GLN:O	1:A:73:GLU:HB2	2.21	0.40
1:B:90:LEU:HD12	1:B:90:LEU:HA	1.91	0.40
1:B:546[A]:VAL:HG22	1:B:606:GLN:NE2	2.35	0.40
1:B:487:ASN:HD22	1:B:488:ASP:N	2.19	0.40
1:A:154:TRP:CE2	1:A:212:SER:HB2	2.56	0.40
1:B:596:ARG:N	1:B:670:TYR:O	2.52	0.40
1:A:501:ASP:O	1:A:505:GLN:HG3	2.21	0.40

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	729/728 (100%)	705 (97%)	23 (3%)	1 (0%)	59	55
1	B	728/728 (100%)	700 (96%)	27 (4%)	1 (0%)	59	55
All	All	1457/1456 (100%)	1405 (96%)	50 (3%)	2 (0%)	59	55

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	320	GLN
1	B	320	GLN

### 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	656/653 (100%)	648 (99%)	8 (1%)	82	84
1	B	655/653 (100%)	644 (98%)	11 (2%)	73	75
All	All	1311/1306 (100%)	1292 (99%)	19 (1%)	76	81

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	97	GLU
1	A	170	ASN
1	A	184	ARG
1	A	385	CYS
1	A	388	GLN
1	A	448	GLU
1	A	472	CYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	566	TYR
1	B	57	LEU
1	B	97	GLU
1	B	230	ASP
1	B	272	ASN
1	B	385	CYS
1	B	472	CYS
1	B	487	ASN
1	B	566	TYR
1	B	658	ARG
1	B	704	HIS
1	B	761	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (20) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	74	ASN
1	A	75	ASN
1	A	123	GLN
1	A	169	ASN
1	A	170	ASN
1	A	435	GLN
1	A	505	GLN
1	A	572	ASN
1	A	712	HIS
1	B	51	ASN
1	B	138	ASN
1	B	141	GLN
1	B	169	ASN
1	B	272	ASN
1	B	345	HIS
1	B	487	ASN
1	B	572	ASN
1	B	606	GLN
1	B	712	HIS
1	B	761	GLN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

18 carbohydrates 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	$\# Z  > 2$
2	NAG	A	1085	1,2	12,14,15	0.52	0	15,19,21	1.08	1 (6%)
2	NDG	A	1086	2	12,14,15	0.61	0	15,19,21	0.74	0
4	NDG	A	1150	1,4	12,14,15	0.46	0	15,19,21	0.91	1 (6%)
4	NAG	A	1151	4	12,14,15	0.59	0	15,19,21	0.79	0
5	NAG	A	1219	1,5	12,14,15	0.53	0	15,19,21	0.78	0
5	NAG	A	1220	5	12,14,15	0.49	0	15,19,21	0.67	0
2	NAG	A	1229	1,2	12,14,15	0.48	0	15,19,21	0.89	0
2	NDG	A	1230	2	12,14,15	0.74	0	15,19,21	1.23	1 (6%)
5	NAG	A	1321	1,5	12,14,15	0.65	0	15,19,21	0.92	0
5	NAG	A	1322	5	12,14,15	0.62	0	15,19,21	0.68	0
5	NAG	B	2085	1,5	12,14,15	0.60	0	15,19,21	0.67	0
5	NAG	B	2086	5	12,14,15	0.75	0	15,19,21	1.15	2 (13%)
5	NAG	B	2219	1,5	12,14,15	0.48	0	15,19,21	0.76	0
5	NAG	B	2220	5	12,14,15	0.53	0	15,19,21	0.69	0
5	NAG	B	2229	1,5	12,14,15	0.46	0	15,19,21	1.03	1 (6%)
5	NAG	B	2230	5	12,14,15	0.50	0	15,19,21	0.71	0
5	NAG	B	2281	1,5	12,14,15	0.49	0	15,19,21	0.80	1 (6%)
5	NAG	B	2282	5	12,14,15	0.55	0	15,19,21	0.69	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
2	NAG	A	1085	1,2	-	0/6/23/26	0/1/1/1
2	NDG	A	1086	2	-	0/6/23/26	0/1/1/1
4	NDG	A	1150	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1151	4	-	0/6/23/26	0/1/1/1
5	NAG	A	1219	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1220	5	-	0/6/23/26	0/1/1/1
2	NAG	A	1229	1,2	-	0/6/23/26	0/1/1/1
2	NDG	A	1230	2	-	0/6/23/26	0/1/1/1
5	NAG	A	1321	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1322	5	-	0/6/23/26	0/1/1/1
5	NAG	B	2085	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	2086	5	-	0/6/23/26	0/1/1/1
5	NAG	B	2219	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	2220	5	-	0/6/23/26	0/1/1/1
5	NAG	B	2229	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	2230	5	-	0/6/23/26	0/1/1/1
5	NAG	B	2281	1,5	-	0/6/23/26	0/1/1/1
5	NAG	B	2282	5	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1230	NDG	C3-C2-N2	-3.15	106.96	111.76
5	B	2229	NAG	C3-C2-N2	-2.97	107.24	111.76
5	B	2086	NAG	C3-C2-N2	-2.55	107.88	111.76
5	B	2281	NAG	C2-N2-C7	-2.26	119.30	123.09
4	A	1150	NDG	C3-C2-N2	-2.24	108.34	111.76
5	B	2086	NAG	C4-C3-C2	2.14	116.56	111.32
2	A	1085	NAG	C4-C3-C2	-2.04	106.33	111.32

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.6 Ligand geometry ⓘ

Of 9 ligands modelled in this entry, 1 is monoatomic - leaving 8 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
3	NAG	A	1092	1	12,14,15	0.49	0	15,19,21	0.62	0
3	NAG	A	1281	1	12,14,15	0.45	0	15,19,21	0.85	1 (6%)
3	NAG	A	1520	1	12,14,15	0.60	0	15,19,21	0.79	0
7	605	A	1521	-	28,28,28	3.43	20 (71%)	39,39,39	1.59	8 (20%)
7	605	B	1521	-	28,28,28	3.49	19 (67%)	39,39,39	1.48	10 (25%)
3	NAG	B	2092	1	12,14,15	0.55	0	15,19,21	0.82	0
3	NAG	B	2150	1	12,14,15	0.53	0	15,19,21	0.70	0
3	NAG	B	2321	1	12,14,15	0.49	0	15,19,21	0.85	1 (6%)

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
3	NAG	A	1092	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1281	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1520	1	-	0/6/23/26	0/1/1/1
7	605	A	1521	-	-	0/16/16/16	0/3/3/3
7	605	B	1521	-	-	0/16/16/16	0/3/3/3
3	NAG	B	2092	1	-	0/6/23/26	0/1/1/1
3	NAG	B	2150	1	-	0/6/23/26	0/1/1/1
3	NAG	B	2321	1	-	0/6/23/26	0/1/1/1

All (39) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	1521	605	C29-C30	6.27	1.50	1.39
7	A	1521	605	C21-C14	6.11	1.49	1.39
7	A	1521	605	C29-C30	5.88	1.49	1.39
7	B	1521	605	C21-C14	5.77	1.48	1.39
7	A	1521	605	C20-C19	5.46	1.51	1.39
7	B	1521	605	C17-C14	5.32	1.48	1.39
7	A	1521	605	C6-C5	5.04	1.46	1.36
7	B	1521	605	C20-C19	5.03	1.50	1.39
7	B	1521	605	C6-C5	4.97	1.46	1.36
7	A	1521	605	C17-C14	4.85	1.47	1.39

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	1521	605	C27-C26	4.72	1.49	1.39
7	B	1521	605	C27-C26	4.57	1.49	1.39
7	B	1521	605	C31-C30	4.41	1.46	1.39
7	A	1521	605	C4-C5	4.19	1.45	1.37
7	A	1521	605	C31-C30	4.16	1.45	1.39
7	B	1521	605	C4-C3	4.15	1.47	1.39
7	A	1521	605	C4-C3	4.03	1.46	1.39
7	B	1521	605	C31-C26	3.93	1.47	1.39
7	B	1521	605	C21-C20	3.68	1.45	1.38
7	A	1521	605	C6-C1	3.68	1.45	1.38
7	B	1521	605	C1-C2	3.55	1.46	1.38
7	B	1521	605	C28-C29	3.44	1.47	1.39
7	A	1521	605	C3-C2	3.38	1.44	1.37
7	B	1521	605	C18-C19	3.35	1.46	1.39
7	B	1521	605	C6-C1	3.32	1.45	1.38
7	B	1521	605	C28-C27	3.30	1.46	1.39
7	B	1521	605	C3-C2	3.30	1.44	1.37
7	A	1521	605	C18-C19	3.28	1.46	1.39
7	A	1521	605	C31-C26	3.23	1.45	1.39
7	B	1521	605	C4-C5	3.22	1.43	1.37
7	A	1521	605	C21-C20	3.15	1.44	1.38
7	A	1521	605	C28-C29	3.11	1.46	1.39
7	A	1521	605	C28-C27	3.04	1.46	1.39
7	B	1521	605	F34-C5	-2.87	1.29	1.36
7	A	1521	605	F34-C5	-2.83	1.30	1.36
7	B	1521	605	F35-C2	-2.48	1.29	1.35
7	A	1521	605	C14-C11	-2.41	1.48	1.52
7	A	1521	605	C1-C2	2.32	1.43	1.38
7	A	1521	605	F35-C2	-2.30	1.29	1.35

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1521	605	C17-C14-C11	-4.28	114.63	120.70
7	A	1521	605	C21-C14-C11	3.77	126.05	120.70
7	B	1521	605	C17-C14-C11	-3.44	115.83	120.70
7	B	1521	605	C21-C14-C11	3.22	125.27	120.70
7	A	1521	605	C3-C10-C11	-3.19	106.62	113.87
7	B	1521	605	C3-C10-C11	-3.14	106.74	113.87
7	A	1521	605	C4-C3-C2	3.04	119.10	116.48
7	B	1521	605	O42-C40-N41	-2.56	118.90	122.59
7	B	1521	605	C4-C3-C2	2.54	118.66	116.48

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1521	605	C31-C26-C19	-2.37	116.72	120.89
7	A	1521	605	C18-C19-C26	-2.35	117.27	121.41
7	A	1521	605	C27-C26-C19	2.34	125.52	121.41
7	B	1521	605	C31-C26-C19	-2.33	116.80	120.89
7	B	1521	605	C27-C26-C19	2.32	125.50	121.41
7	A	1521	605	O42-C40-C30	2.32	122.19	119.58
7	B	1521	605	O42-C40-C30	2.20	122.06	119.58
7	B	1521	605	C26-C31-C30	2.12	123.35	121.25
7	B	1521	605	C18-C19-C26	-2.08	117.74	121.41
3	A	1281	NAG	C2-N2-C7	-2.08	119.59	123.09
3	B	2321	NAG	C2-N2-C7	-2.05	119.64	123.09

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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	728/728 (100%)	-0.04	14 (1%) 64 64	21, 29, 48, 59	0
1	B	728/728 (100%)	-0.02	25 (3%) 43 43	21, 31, 49, 59	0
All	All	1456/1456 (100%)	-0.03	39 (2%) 50 52	21, 30, 49, 59	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	39	THR	6.3
1	A	39	THR	6.2
1	B	766	PRO	5.1
1	B	95	PHE	4.6
1	A	766	PRO	4.6
1	A	74	ASN	4.5
1	B	98	PHE	4.4
1	B	97	GLU	4.4
1	B	96	ASP	4.2
1	B	88	VAL	4.1
1	A	73	GLU	3.9
1	B	91	GLU	3.7
1	B	92	ASN	3.6
1	B	73	GLU	3.6
1	A	97	GLU	3.3
1	B	90	LEU	3.3
1	B	94	THR	3.2
1	A	279	VAL	3.1
1	B	100	HIS	3.0
1	B	40	ARG	3.0
1	B	93	SER	3.0
1	B	99	GLY	3.0
1	A	40	ARG	2.9
1	B	505	GLN	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	74	ASN	2.7
1	A	99	GLY	2.6
1	A	96	ASP	2.6
1	B	502	LYS	2.6
1	B	72	GLN	2.5
1	A	521	GLU	2.5
1	B	76	ILE	2.4
1	A	138	ASN	2.4
1	A	98	PHE	2.4
1	B	506	ASN	2.4
1	A	54	ARG	2.1
1	A	92	ASN	2.1
1	B	140	ARG	2.1
1	B	378	GLU	2.0
1	B	138	ASN	2.0

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

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

## 6.3 Carbohydrates ⓘ

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	NDG	A	1230	14/15	0.34	12.81	53,56,58,59	0
5	NAG	A	1321	14/15	0.27	11.05	43,48,49,54	0
5	NAG	B	2282	14/15	0.25	7.24	50,52,54,54	0
4	NDG	A	1150	14/15	0.39	5.09	55,58,60,61	0
5	NAG	A	1322	14/15	0.40	4.56	56,58,60,60	0
5	NAG	B	2230	14/15	0.25	2.27	48,51,52,53	0
5	NAG	B	2220	14/15	0.17	1.79	48,49,51,51	0
5	NAG	B	2219	14/15	0.11	1.65	41,44,46,46	0
5	NAG	A	1220	14/15	0.25	1.60	52,54,57,57	0
5	NAG	A	1219	14/15	0.14	1.54	44,47,49,49	0
2	NAG	A	1229	14/15	0.15	1.32	39,43,47,49	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	NAG	B	2085	14/15	0.20	1.06	50,53,55,57	0
2	NAG	A	1085	14/15	0.14	0.96	46,50,53,56	0
5	NAG	B	2281	14/15	0.11	0.96	40,41,46,48	0
5	NAG	B	2229	14/15	0.10	-0.19	33,35,40,45	0
4	NAG	A	1151	14/15	0.39	-	59,61,62,63	0
2	NDG	A	1086	14/15	0.33	-	58,60,61,61	0
5	NAG	B	2086	14/15	0.52	-	59,60,61,61	0

## 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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	B	2150	14/15	0.39	14.38	53,55,57,57	0
3	NAG	A	1520	14/15	0.38	6.29	59,60,60,60	0
3	NAG	B	2321	14/15	0.31	5.85	42,46,48,49	0
7	605	B	1521	26/26	0.19	1.78	21,26,32,35	0
3	NAG	A	1092	14/15	0.48	1.53	61,62,62,62	0
3	NAG	B	2092	14/15	0.42	0.93	60,60,61,62	0
7	605	A	1521	26/26	0.17	0.33	17,22,29,32	0
3	NAG	A	1281	14/15	0.18	0.16	47,49,50,51	0
6	NA	A	1	1/1	0.07	-2.49	35,35,35,35	0

## 6.5 Other polymers

There are no such residues in this entry.