



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2017 – 08:37 AM EDT

PDB ID : 5MF1  
Title : Crystal structure of a C-terminally truncated trimeric ectodomain of the Chlamydomonas reinhardtii gamete fusion protein HAP2  
Authors : Fedry, J.; Rey, F.A.; Krey, T.  
Deposited on : unknown  
Resolution : 3.30 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20029824  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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-20029824

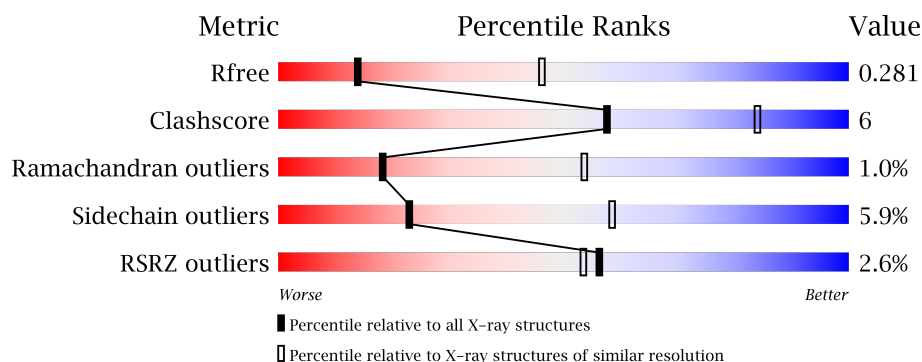
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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))
$R_{free}$	100719	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)

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	611	<div> <div> <div>0.1%</div> <div>59%</div> <div>13%</div> <div>27%</div> </div> </div>
1	B	611	<div> <div>2%</div> <div>60%</div> <div>11%</div> <div>28%</div> </div>
1	C	611	<div> <div>3%</div> <div>60%</div> <div>11%</div> <div>28%</div> </div>

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
3	NGA	C	1002	-	-	-	X

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10032 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 Fusion protein HAP2/GCS1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	446	Total	C	N	O	S	0	0	0
			3349	2104	567	655	23			
1	B	441	Total	C	N	O	S	0	0	0
			3302	2073	560	646	23			
1	C	442	Total	C	N	O	S	0	0	0
			3297	2069	559	646	23			

There are 123 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	21	ARG	-	cloning artifact	UNP A4GRC6
A	22	SER	-	cloning artifact	UNP A4GRC6
A	593	GLY	-	expression tag	UNP A4GRC6
A	594	PRO	-	expression tag	UNP A4GRC6
A	595	PHE	-	expression tag	UNP A4GRC6
A	596	GLU	-	expression tag	UNP A4GRC6
A	597	ASP	-	expression tag	UNP A4GRC6
A	598	ASP	-	expression tag	UNP A4GRC6
A	599	ASP	-	expression tag	UNP A4GRC6
A	600	ASP	-	expression tag	UNP A4GRC6
A	601	LYS	-	expression tag	UNP A4GRC6
A	602	ALA	-	expression tag	UNP A4GRC6
A	603	GLY	-	expression tag	UNP A4GRC6
A	604	TRP	-	expression tag	UNP A4GRC6
A	605	SER	-	expression tag	UNP A4GRC6
A	606	HIS	-	expression tag	UNP A4GRC6
A	607	PRO	-	expression tag	UNP A4GRC6
A	608	GLN	-	expression tag	UNP A4GRC6
A	609	PHE	-	expression tag	UNP A4GRC6
A	610	GLU	-	expression tag	UNP A4GRC6
A	611	LYS	-	expression tag	UNP A4GRC6
A	612	GLY	-	expression tag	UNP A4GRC6
A	613	GLY	-	expression tag	UNP A4GRC6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	614	GLY	-	expression tag	UNP A4GRC6
A	615	SER	-	expression tag	UNP A4GRC6
A	616	GLY	-	expression tag	UNP A4GRC6
A	617	GLY	-	expression tag	UNP A4GRC6
A	618	GLY	-	expression tag	UNP A4GRC6
A	619	SER	-	expression tag	UNP A4GRC6
A	620	GLY	-	expression tag	UNP A4GRC6
A	621	GLY	-	expression tag	UNP A4GRC6
A	622	GLY	-	expression tag	UNP A4GRC6
A	623	SER	-	expression tag	UNP A4GRC6
A	624	TRP	-	expression tag	UNP A4GRC6
A	625	SER	-	expression tag	UNP A4GRC6
A	626	HIS	-	expression tag	UNP A4GRC6
A	627	PRO	-	expression tag	UNP A4GRC6
A	628	GLN	-	expression tag	UNP A4GRC6
A	629	PHE	-	expression tag	UNP A4GRC6
A	630	GLU	-	expression tag	UNP A4GRC6
A	631	LYS	-	expression tag	UNP A4GRC6
B	21	ARG	-	cloning artifact	UNP A4GRC6
B	22	SER	-	cloning artifact	UNP A4GRC6
B	593	GLY	-	expression tag	UNP A4GRC6
B	594	PRO	-	expression tag	UNP A4GRC6
B	595	PHE	-	expression tag	UNP A4GRC6
B	596	GLU	-	expression tag	UNP A4GRC6
B	597	ASP	-	expression tag	UNP A4GRC6
B	598	ASP	-	expression tag	UNP A4GRC6
B	599	ASP	-	expression tag	UNP A4GRC6
B	600	ASP	-	expression tag	UNP A4GRC6
B	601	LYS	-	expression tag	UNP A4GRC6
B	602	ALA	-	expression tag	UNP A4GRC6
B	603	GLY	-	expression tag	UNP A4GRC6
B	604	TRP	-	expression tag	UNP A4GRC6
B	605	SER	-	expression tag	UNP A4GRC6
B	606	HIS	-	expression tag	UNP A4GRC6
B	607	PRO	-	expression tag	UNP A4GRC6
B	608	GLN	-	expression tag	UNP A4GRC6
B	609	PHE	-	expression tag	UNP A4GRC6
B	610	GLU	-	expression tag	UNP A4GRC6
B	611	LYS	-	expression tag	UNP A4GRC6
B	612	GLY	-	expression tag	UNP A4GRC6
B	613	GLY	-	expression tag	UNP A4GRC6
B	614	GLY	-	expression tag	UNP A4GRC6

*Continued on next page...*

*Continued from previous page...*

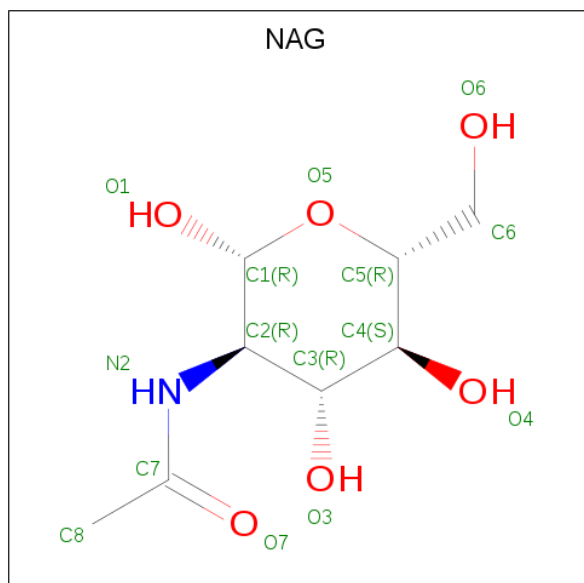
Chain	Residue	Modelled	Actual	Comment	Reference
B	615	SER	-	expression tag	UNP A4GRC6
B	616	GLY	-	expression tag	UNP A4GRC6
B	617	GLY	-	expression tag	UNP A4GRC6
B	618	GLY	-	expression tag	UNP A4GRC6
B	619	SER	-	expression tag	UNP A4GRC6
B	620	GLY	-	expression tag	UNP A4GRC6
B	621	GLY	-	expression tag	UNP A4GRC6
B	622	GLY	-	expression tag	UNP A4GRC6
B	623	SER	-	expression tag	UNP A4GRC6
B	624	TRP	-	expression tag	UNP A4GRC6
B	625	SER	-	expression tag	UNP A4GRC6
B	626	HIS	-	expression tag	UNP A4GRC6
B	627	PRO	-	expression tag	UNP A4GRC6
B	628	GLN	-	expression tag	UNP A4GRC6
B	629	PHE	-	expression tag	UNP A4GRC6
B	630	GLU	-	expression tag	UNP A4GRC6
B	631	LYS	-	expression tag	UNP A4GRC6
C	21	ARG	-	cloning artifact	UNP A4GRC6
C	22	SER	-	cloning artifact	UNP A4GRC6
C	593	GLY	-	expression tag	UNP A4GRC6
C	594	PRO	-	expression tag	UNP A4GRC6
C	595	PHE	-	expression tag	UNP A4GRC6
C	596	GLU	-	expression tag	UNP A4GRC6
C	597	ASP	-	expression tag	UNP A4GRC6
C	598	ASP	-	expression tag	UNP A4GRC6
C	599	ASP	-	expression tag	UNP A4GRC6
C	600	ASP	-	expression tag	UNP A4GRC6
C	601	LYS	-	expression tag	UNP A4GRC6
C	602	ALA	-	expression tag	UNP A4GRC6
C	603	GLY	-	expression tag	UNP A4GRC6
C	604	TRP	-	expression tag	UNP A4GRC6
C	605	SER	-	expression tag	UNP A4GRC6
C	606	HIS	-	expression tag	UNP A4GRC6
C	607	PRO	-	expression tag	UNP A4GRC6
C	608	GLN	-	expression tag	UNP A4GRC6
C	609	PHE	-	expression tag	UNP A4GRC6
C	610	GLU	-	expression tag	UNP A4GRC6
C	611	LYS	-	expression tag	UNP A4GRC6
C	612	GLY	-	expression tag	UNP A4GRC6
C	613	GLY	-	expression tag	UNP A4GRC6
C	614	GLY	-	expression tag	UNP A4GRC6
C	615	SER	-	expression tag	UNP A4GRC6

*Continued on next page...*

Continued from previous page...

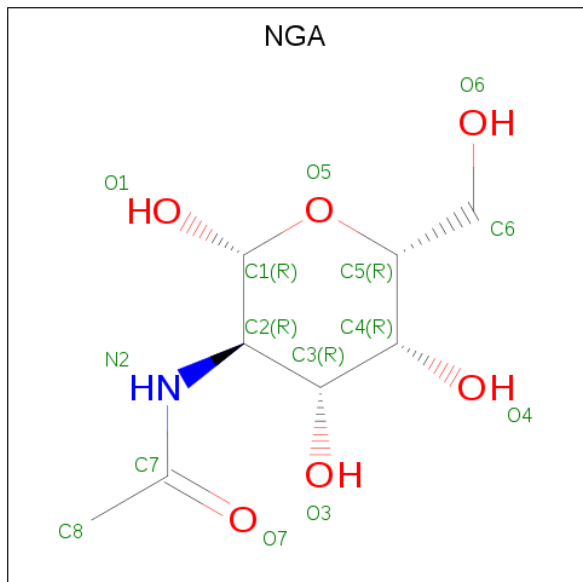
Chain	Residue	Modelled	Actual	Comment	Reference
C	616	GLY	-	expression tag	UNP A4GRC6
C	617	GLY	-	expression tag	UNP A4GRC6
C	618	GLY	-	expression tag	UNP A4GRC6
C	619	SER	-	expression tag	UNP A4GRC6
C	620	GLY	-	expression tag	UNP A4GRC6
C	621	GLY	-	expression tag	UNP A4GRC6
C	622	GLY	-	expression tag	UNP A4GRC6
C	623	SER	-	expression tag	UNP A4GRC6
C	624	TRP	-	expression tag	UNP A4GRC6
C	625	SER	-	expression tag	UNP A4GRC6
C	626	HIS	-	expression tag	UNP A4GRC6
C	627	PRO	-	expression tag	UNP A4GRC6
C	628	GLN	-	expression tag	UNP A4GRC6
C	629	PHE	-	expression tag	UNP A4GRC6
C	630	GLU	-	expression tag	UNP A4GRC6
C	631	LYS	-	expression tag	UNP A4GRC6

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 3 is N-ACETYL-D-GALACTOSAMINE (three-letter code: NGA) (formula:  $C_8H_{15}NO_6$ ).



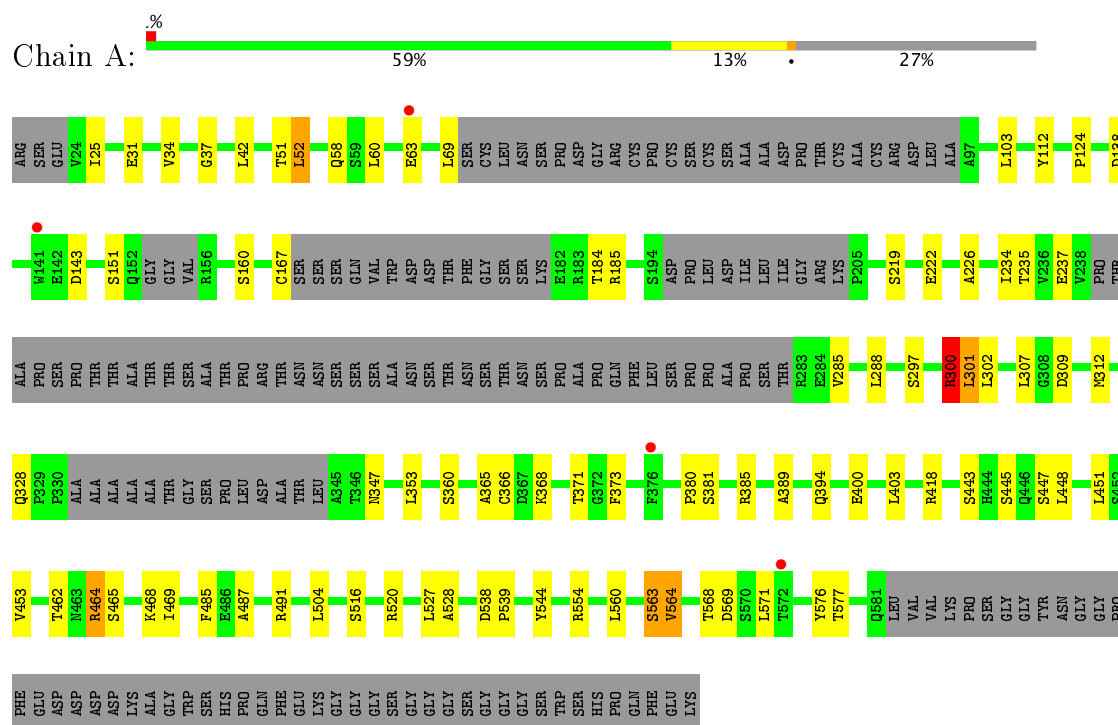
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	C	1	Total	C	N	O	0	0
			14	8	1	5		



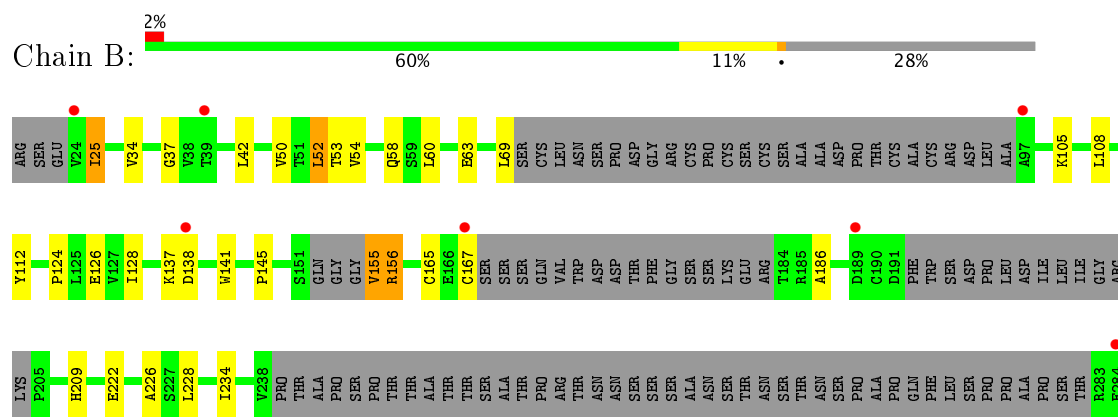
### 3 Residue-property plots

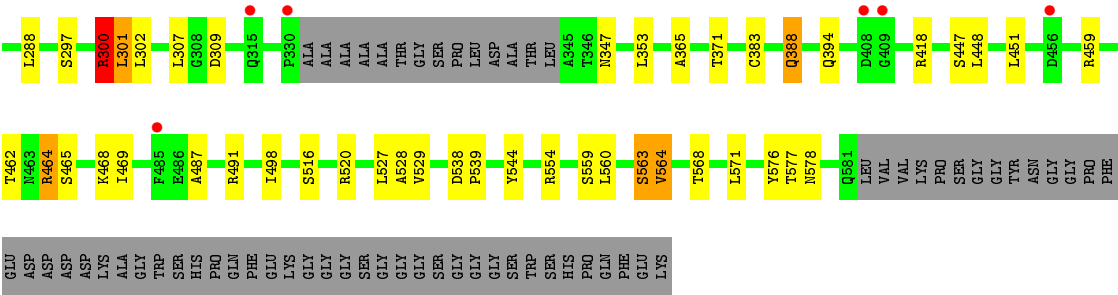
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Fusion protein HAP2/GCS1

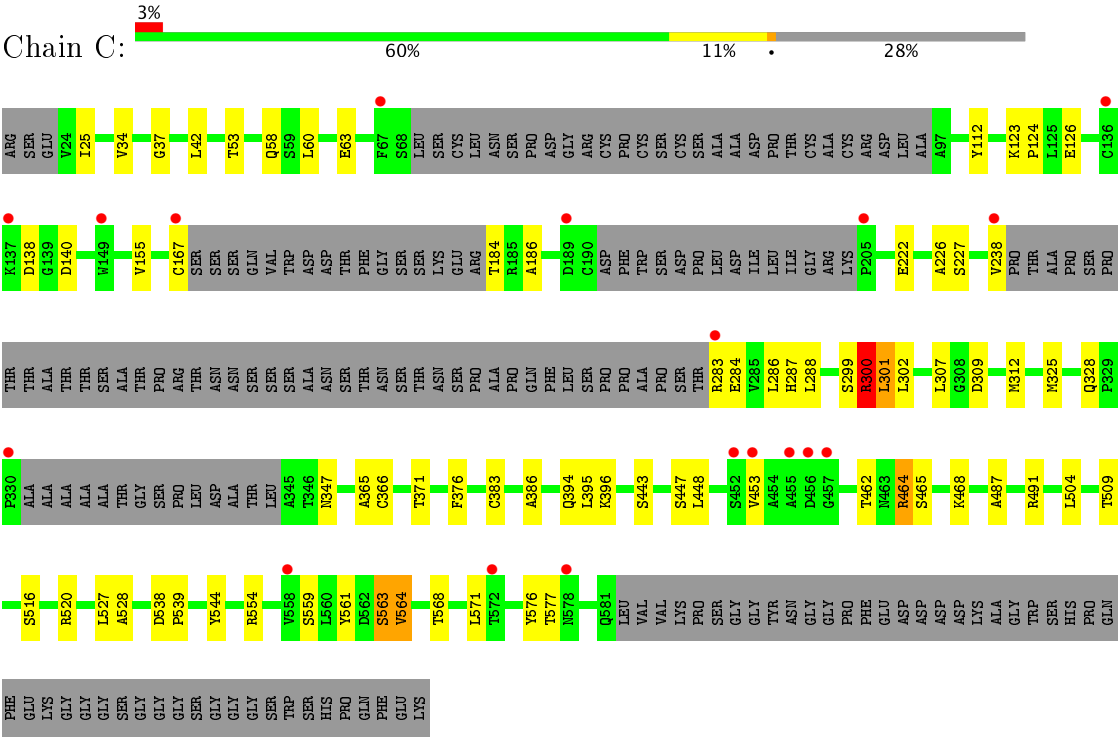


- Molecule 1: Fusion protein HAP2/GCS1





● Molecule 1: Fusion protein HAP2/GCS1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.97Å 114.06Å 137.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.62 – 3.30 48.62 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (48.62-3.30) 99.6 (48.62-3.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 3.33Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.219 , 0.266 0.237 , 0.281	Depositor DCC
$R_{free}$ test set	1259 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	87.0	Xtriage
Anisotropy	0.305	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 54.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	10032	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.27% 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: NAG, NGA

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	0.41	0/3412	0.67	1/4643 (0.0%)
1	B	0.41	0/3362	0.67	1/4575 (0.0%)
1	C	0.41	0/3358	0.67	1/4571 (0.0%)
All	All	0.41	0/10132	0.67	3/13789 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	563	SER	C-N-CA	5.73	136.02	121.70
1	A	563	SER	C-N-CA	5.65	135.82	121.70
1	B	563	SER	C-N-CA	5.62	135.75	121.70

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	3349	0	3272	42	0
1	B	3302	0	3241	44	0
1	C	3297	0	3230	35	0
2	A	14	0	13	0	0
2	B	14	0	13	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	14	0	13	0	0
3	A	14	0	13	0	0
3	B	14	0	13	0	0
3	C	14	0	13	0	0
All	All	10032	0	9821	111	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 6.

All (111) 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:468:LYS:HA	1:A:568:THR:HG21	1.63	0.80
1:B:554:ARG:HD2	1:B:576:TYR:HE1	1.45	0.79
1:C:554:ARG:HD2	1:C:576:TYR:HE1	1.52	0.73
1:A:237:GLU:HG3	1:A:285:VAL:HG22	1.71	0.72
1:A:381:SER:HB3	1:A:385:ARG:HG3	1.72	0.72
1:B:554:ARG:HD2	1:B:576:TYR:CE1	2.25	0.71
1:A:365:ALA:O	1:A:371:THR:HG22	1.91	0.71
1:A:302:LEU:HD13	1:A:453:VAL:HG22	1.73	0.70
1:B:468:LYS:HA	1:B:568:THR:HG21	1.74	0.69
1:C:468:LYS:HA	1:C:568:THR:HG21	1.75	0.69
1:C:309:ASP:HB3	1:C:447:SER:HA	1.77	0.66
1:B:25:ILE:HD11	1:B:53:THR:HG23	1.78	0.66
1:C:554:ARG:HD2	1:C:576:TYR:CE1	2.30	0.65
1:A:309:ASP:HB3	1:A:447:SER:HA	1.77	0.65
1:B:309:ASP:HB3	1:B:447:SER:HA	1.77	0.65
1:B:124:PRO:HD2	1:B:394:GLN:HG3	1.81	0.63
1:A:563:SER:O	1:A:564:VAL:HG22	2.01	0.61
1:C:302:LEU:HD13	1:C:453:VAL:HG22	1.84	0.60
1:B:365:ALA:O	1:B:371:THR:HG22	2.01	0.60
1:A:554:ARG:HD2	1:A:576:TYR:HE1	1.66	0.59
1:B:288:LEU:HD22	1:B:297:SER:HB3	1.84	0.59
1:B:563:SER:O	1:B:564:VAL:HG22	2.03	0.59
1:B:37:GLY:HA3	1:C:465:SER:HA	1.84	0.58
1:C:563:SER:O	1:C:564:VAL:HG22	2.03	0.58
1:C:124:PRO:HD2	1:C:394:GLN:HG3	1.88	0.56
1:A:554:ARG:HD2	1:A:576:TYR:CE1	2.41	0.56
1:B:52:LEU:HD21	1:B:451:LEU:HD13	1.88	0.55
1:A:464:ARG:HB2	1:C:34:VAL:HG13	1.90	0.54
1:A:288:LEU:HD22	1:A:297:SER:HB3	1.89	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:34:VAL:HG13	1:B:464:ARG:HB2	1.89	0.53
1:B:138:ASP:OD1	1:B:145:PRO:HB2	2.08	0.53
1:C:487:ALA:HB2	1:C:576:TYR:HD2	1.74	0.52
1:B:498:ILE:HD11	1:B:529:VAL:HG21	1.92	0.52
1:C:238:VAL:HB	1:C:283:ARG:HB2	1.92	0.52
1:B:108:LEU:HD11	1:B:228:LEU:HB3	1.93	0.51
1:B:487:ALA:HB2	1:B:576:TYR:HD2	1.76	0.51
1:B:34:VAL:HG13	1:C:464:ARG:HB2	1.93	0.50
1:C:123:LYS:HB2	1:C:396:LYS:HB3	1.93	0.50
1:C:365:ALA:O	1:C:371:THR:HG22	2.12	0.50
1:A:368:LYS:O	1:A:371:THR:HG23	2.12	0.49
1:A:487:ALA:HB2	1:A:576:TYR:HD2	1.76	0.49
1:A:31:GLU:HB3	1:B:459:ARG:HG3	1.94	0.49
1:A:380:PRO:HB2	1:A:385:ARG:HH12	1.76	0.49
1:B:165:CYS:HB2	1:B:209:HIS:HE2	1.77	0.49
1:B:300:ARG:HG2	1:B:301:LEU:H	1.78	0.48
1:C:184:THR:HA	1:C:386:ALA:HA	1.95	0.48
1:A:469:ILE:HG13	1:A:560:LEU:HD12	1.96	0.48
1:B:128:ILE:HG12	1:B:209:HIS:ND1	2.28	0.47
1:B:37:GLY:HA2	1:C:504:LEU:HD12	1.95	0.47
1:C:300:ARG:HG2	1:C:301:LEU:H	1.78	0.47
1:B:108:LEU:HD21	1:B:228:LEU:HD22	1.97	0.47
1:C:538:ASP:HB3	1:C:539:PRO:HD3	1.96	0.47
1:A:37:GLY:HA3	1:B:465:SER:HA	1.96	0.46
1:A:185:ARG:HH22	1:A:389:ALA:HB3	1.80	0.46
1:B:112:TYR:CE2	1:B:226:ALA:HB2	2.50	0.46
1:B:520:ARG:HB3	1:B:544:TYR:HB2	1.97	0.46
1:B:538:ASP:HB3	1:B:539:PRO:HD3	1.97	0.46
1:B:234:ILE:HB	1:B:288:LEU:HB2	1.98	0.46
1:C:520:ARG:HB3	1:C:544:TYR:HB2	1.96	0.46
1:A:307:LEU:HD12	1:A:448:LEU:HB3	1.97	0.45
1:A:300:ARG:HG2	1:A:301:LEU:H	1.80	0.45
1:A:520:ARG:HB3	1:A:544:TYR:HB2	1.97	0.45
1:C:126:GLU:HG3	1:C:376:PHE:HZ	1.81	0.45
1:A:538:ASP:HB3	1:A:539:PRO:HD3	1.98	0.45
1:B:112:TYR:CZ	1:B:226:ALA:HB2	2.51	0.45
1:B:60:LEU:HB2	1:B:528:ALA:HB1	1.99	0.45
1:A:112:TYR:CE2	1:A:226:ALA:HB2	2.51	0.45
1:C:307:LEU:HD12	1:C:448:LEU:HB3	1.98	0.45
1:A:52:LEU:HD21	1:A:451:LEU:HD13	1.99	0.44
1:A:160:SER:HA	1:A:373:PHE:HB3	1.98	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:307:LEU:HD12	1:B:448:LEU:HB3	2.00	0.44
1:A:60:LEU:HB2	1:A:528:ALA:HB1	1.98	0.44
1:A:112:TYR:CZ	1:A:226:ALA:HB2	2.52	0.44
1:C:112:TYR:CZ	1:C:226:ALA:HB2	2.52	0.44
1:A:124:PRO:HD2	1:A:394:GLN:HG3	1.99	0.44
1:B:137:LYS:O	1:B:145:PRO:HA	2.18	0.44
1:B:186:ALA:HB2	1:B:383:CYS:HB3	1.99	0.44
1:C:112:TYR:CE2	1:C:226:ALA:HB2	2.52	0.44
1:C:325:MET:HG3	1:C:395:LEU:HD21	2.00	0.44
1:A:380:PRO:HG2	1:A:385:ARG:HH22	1.83	0.43
1:A:51:THR:HG21	1:B:25:ILE:HG23	1.99	0.43
1:C:186:ALA:HB2	1:C:383:CYS:HB3	1.99	0.43
1:A:234:ILE:HB	1:A:288:LEU:HB2	1.99	0.43
1:C:312:MET:HE3	1:C:443:SER:HB3	2.01	0.43
1:A:465:SER:HA	1:C:37:GLY:HA3	2.01	0.43
1:C:60:LEU:HB2	1:C:528:ALA:HB1	2.00	0.43
1:B:353:LEU:HB3	1:B:418:ARG:HH21	1.84	0.43
1:C:309:ASP:HB3	1:C:447:SER:CA	2.46	0.43
1:A:400:GLU:HA	1:A:403:LEU:HD12	2.01	0.42
1:B:155:VAL:HB	1:B:156:ARG:H	1.78	0.42
1:A:485:PHE:O	1:A:576:TYR:HA	2.20	0.42
1:B:302:LEU:HD11	1:B:451:LEU:HD23	2.01	0.42
1:C:509:THR:HB	1:C:561:TYR:HB2	2.01	0.42
1:C:283:ARG:HB3	1:C:284:GLU:H	1.60	0.41
1:B:63:GLU:HB3	1:B:462:THR:HG23	2.02	0.41
1:C:63:GLU:HB3	1:C:462:THR:HG23	2.01	0.41
1:A:312:MET:HE3	1:A:443:SER:HB3	2.02	0.41
1:A:368:LYS:H	1:A:371:THR:CG2	2.34	0.41
1:B:554:ARG:CD	1:B:576:TYR:HE1	2.26	0.41
1:A:504:LEU:HD12	1:C:37:GLY:HA2	2.01	0.41
1:B:50:VAL:HG12	1:B:52:LEU:HD23	2.03	0.41
1:A:103:LEU:HG	1:A:234:ILE:HG12	2.02	0.41
1:A:353:LEU:HB3	1:A:418:ARG:HH21	1.86	0.41
1:B:554:ARG:HA	1:B:554:ARG:HE	1.86	0.41
1:A:360:SER:HB2	1:A:368:LYS:HB2	2.03	0.41
1:B:126:GLU:HB3	1:B:388:GLN:NE2	2.35	0.41
1:C:288:LEU:HD21	1:C:299:SER:HB2	2.02	0.41
1:A:63:GLU:HB3	1:A:462:THR:HG23	2.03	0.40
1:B:54:VAL:HG12	1:B:105:LYS:HB2	2.03	0.40
1:C:554:ARG:HA	1:C:554:ARG:HE	1.86	0.40
1:B:469:ILE:HG13	1:B:560:LEU:HD12	2.02	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	432/611 (71%)	397 (92%)	31 (7%)	4 (1%)	20	55
1	B	427/611 (70%)	389 (91%)	32 (8%)	6 (1%)	13	46
1	C	430/611 (70%)	394 (92%)	33 (8%)	3 (1%)	25	60
All	All	1289/1833 (70%)	1180 (92%)	96 (7%)	13 (1%)	18	53

All (13) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	300	ARG
1	A	564	VAL
1	B	300	ARG
1	B	564	VAL
1	C	300	ARG
1	C	564	VAL
1	A	184	THR
1	B	156	ARG
1	B	141	TRP
1	B	388	GLN
1	A	25	ILE
1	B	25	ILE
1	C	25	ILE

### 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	368/497 (74%)	344 (94%)	24 (6%)	20	55
1	B	364/497 (73%)	346 (95%)	18 (5%)	29	65
1	C	362/497 (73%)	339 (94%)	23 (6%)	20	55
All	All	1094/1491 (73%)	1029 (94%)	65 (6%)	23	58

All (65) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	42	LEU
1	A	52	LEU
1	A	58	GLN
1	A	69	LEU
1	A	138	ASP
1	A	143	ASP
1	A	151	SER
1	A	167	CYS
1	A	219	SER
1	A	222	GLU
1	A	235	THR
1	A	300	ARG
1	A	301	LEU
1	A	328	GLN
1	A	347	ASN
1	A	366	CYS
1	A	445	SER
1	A	464	ARG
1	A	491	ARG
1	A	516	SER
1	A	527	LEU
1	A	569	ASP
1	A	571	LEU
1	A	577	THR
1	B	42	LEU
1	B	52	LEU
1	B	58	GLN
1	B	69	LEU
1	B	155	VAL
1	B	167	CYS
1	B	222	GLU
1	B	300	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	301	LEU
1	B	347	ASN
1	B	464	ARG
1	B	491	ARG
1	B	516	SER
1	B	527	LEU
1	B	559	SER
1	B	571	LEU
1	B	577	THR
1	B	578	ASN
1	C	42	LEU
1	C	53	THR
1	C	58	GLN
1	C	138	ASP
1	C	140	ASP
1	C	155	VAL
1	C	167	CYS
1	C	222	GLU
1	C	227	SER
1	C	286	LEU
1	C	287	HIS
1	C	300	ARG
1	C	301	LEU
1	C	328	GLN
1	C	347	ASN
1	C	366	CYS
1	C	464	ARG
1	C	491	ARG
1	C	516	SER
1	C	527	LEU
1	C	559	SER
1	C	571	LEU
1	C	577	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	347	ASN
1	B	347	ASN
1	B	446	GLN
1	C	216	GLN
1	C	347	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 ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

6 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	NAG	A	1001	1	14,14,15	0.29	0	15,19,21	0.50	0
3	NGA	A	1002	1	14,14,15	0.34	0	15,19,21	1.38	1 (6%)
2	NAG	B	1001	1	14,14,15	0.30	0	15,19,21	0.63	0
3	NGA	B	1002	1	14,14,15	0.43	0	15,19,21	1.64	1 (6%)
2	NAG	C	1001	1	14,14,15	0.31	0	15,19,21	0.90	2 (13%)
3	NGA	C	1002	1	14,14,15	0.41	0	15,19,21	1.80	2 (13%)

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	1001	1	-	0/6/23/26	0/1/1/1
3	NGA	A	1002	1	-	0/6/23/26	0/1/1/1
2	NAG	B	1001	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NGA	B	1002	1	-	0/6/23/26	0/1/1/1
2	NAG	C	1001	1	-	0/6/23/26	0/1/1/1
3	NGA	C	1002	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1002	NGA	O5-C1-C2	-6.03	103.09	111.47
3	B	1002	NGA	O5-C1-C2	-5.70	103.54	111.47
3	A	1002	NGA	O5-C1-C2	-4.62	105.05	111.47
2	C	1001	NAG	O5-C1-C2	-2.38	108.16	111.47
3	C	1002	NGA	C1-O5-C5	2.17	115.16	112.17
2	C	1001	NAG	C1-O5-C5	2.26	115.28	112.17

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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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	446/611 (72%)	0.08	4 (0%) 84 83	60, 89, 119, 146	0
1	B	441/611 (72%)	0.11	13 (2%) 52 50	53, 83, 118, 169	0
1	C	442/611 (72%)	0.20	18 (4%) 38 35	56, 84, 119, 142	0
All	All	1329/1833 (72%)	0.13	35 (2%) 56 53	53, 85, 119, 169	0

All (35) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	455	ALA	4.8
1	C	572	THR	4.1
1	B	189	ASP	3.8
1	C	167	CYS	3.2
1	C	189	ASP	3.1
1	C	457	GLY	2.9
1	C	578	ASN	2.7
1	C	137	LYS	2.7
1	C	205	PRO	2.6
1	C	452	SER	2.6
1	C	456	ASP	2.6
1	B	138	ASP	2.6
1	C	238	VAL	2.6
1	C	67	PHE	2.5
1	A	141	TRP	2.5
1	A	63	GLU	2.5
1	B	24	VAL	2.5
1	A	376	PHE	2.4
1	B	330	PRO	2.4
1	B	167	CYS	2.4
1	C	283	ARG	2.3
1	B	485	PHE	2.3
1	C	453	VAL	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	136	CYS	2.2
1	B	97	ALA	2.2
1	A	572	THR	2.2
1	B	456	ASP	2.2
1	B	284	GLU	2.2
1	C	558	VAL	2.2
1	B	315	GLN	2.1
1	B	39	THR	2.1
1	C	149	TRP	2.1
1	C	330	PRO	2.0
1	B	409	GLY	2.0
1	B	408	ASP	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 carbohydrates 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. 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( $\text{\AA}^2$ )	Q<0.9
3	NGA	C	1002	14/15	0.69	0.31	2.87	92,100,101,102	0
3	NGA	B	1002	14/15	0.57	0.32	1.26	118,120,123,124	0
3	NGA	A	1002	14/15	0.74	0.26	0.56	116,121,124,125	0
2	NAG	B	1001	14/15	0.91	0.17	-0.82	91,92,96,98	0
2	NAG	A	1001	14/15	0.88	0.20	-0.95	95,101,104,104	0
2	NAG	C	1001	14/15	0.85	0.18	-1.01	80,85,88,89	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.