



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 15, 2017 – 12:22 am GMT

PDB ID : 4NE9  
Title : PCSK9 in complex with LDLR peptide  
Authors : Liu, S.  
Deposited on : 2013-10-28  
Resolution : 2.60 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 : trunk28620  
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) : recalc28949

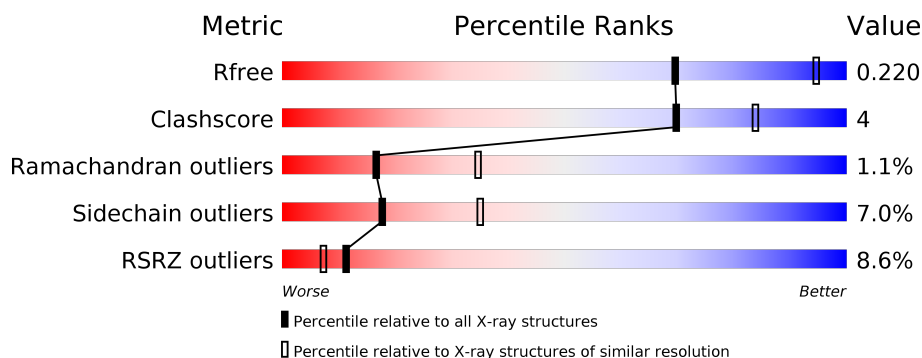
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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	540	<div> <div>7%</div> <div> <div></div> <div>77%</div> <div>12%</div> <div>9%</div> </div> </div>
1	B	540	<div> <div>11%</div> <div> <div></div> <div>76%</div> <div>13%</div> <div>9%</div> </div> </div>
2	C	152	<div> <div></div> <div> <div>52%</div> <div>7%</div> <div>39%</div> </div> </div>
2	P	152	<div> <div></div> <div> <div>54%</div> <div>7%</div> <div>39%</div> </div> </div>
3	D	26	<div> <div>4%</div> <div> <div></div> <div>92%</div> <div>8%</div> </div> </div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 8999 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 Proprotein convertase subtilisin/kexin type 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	491	Total	C	N	O	S	0	1	0
			3646	2246	676	692	32			
1	B	490	Total	C	N	O	S	0	1	0
			3623	2230	669	692	32			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	474	ILE	VAL	engineered mutation	UNP Q8NBP7
A	670	GLU	GLY	engineered mutation	UNP Q8NBP7
B	474	ILE	VAL	engineered mutation	UNP Q8NBP7
B	670	GLU	GLY	engineered mutation	UNP Q8NBP7

- Molecule 2 is a protein called Proprotein convertase subtilisin/kexin type 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	92	Total	C	N	O	S	0	0	0
			740	474	133	131	2			
2	P	92	Total	C	N	O	S	0	0	0
			740	474	133	131	2			

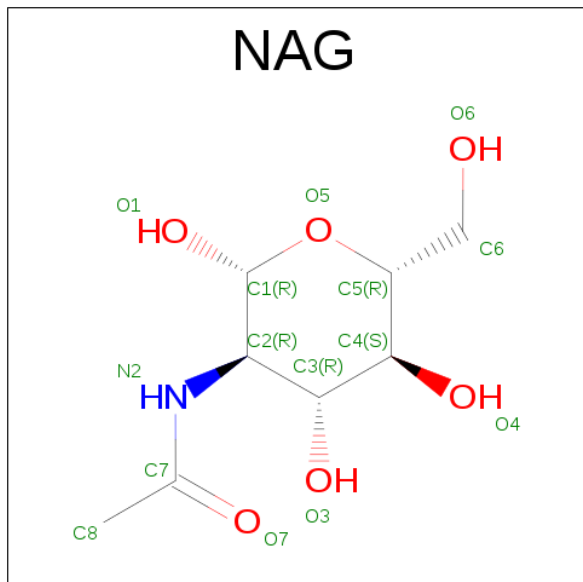
- Molecule 3 is a protein called Low-density lipoprotein receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	26	Total	C	N	O	S	0	0	0
			189	111	33	41	4			

- 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
			24	14	1	9		

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 6 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	1	Total	Ca	0	0
			1	1		

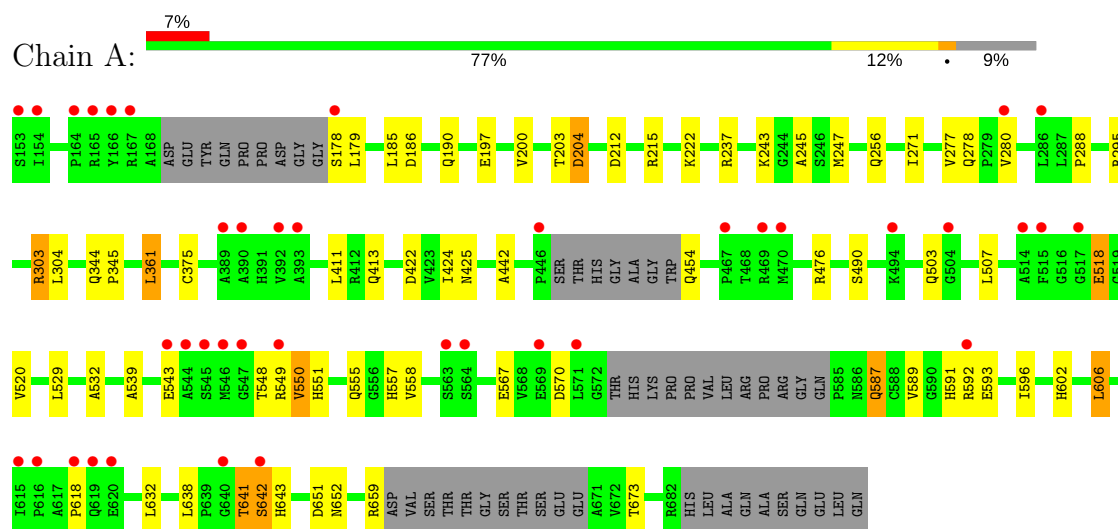
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	9	Total	O	0	0
			9	9		
7	B	5	Total	O	0	0
			5	5		
7	C	4	Total	O	0	0
			4	4		
7	P	4	Total	O	0	0
			4	4		

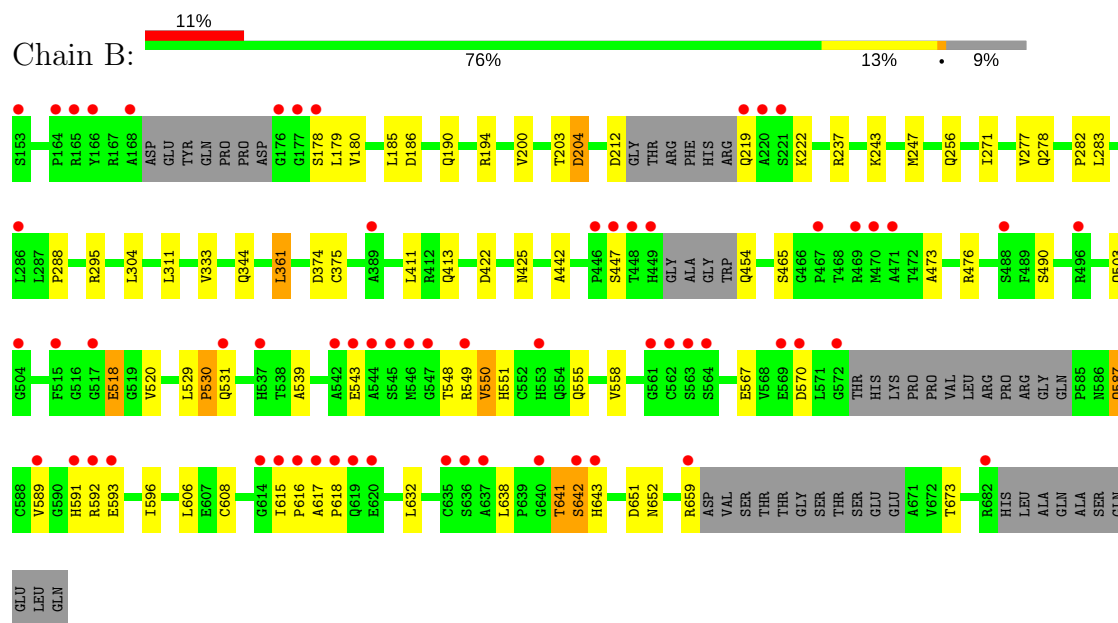
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Proprotein convertase subtilisin/kexin type 9



- Molecule 1: Proprotein convertase subtilisin/kexin type 9



- Molecule 2: Proprotein convertase subtilisin/kexin type 9



MET	GLY	THR	VAL	SER	SER	ARG	ARG	SER	TRP	TRP	PRO	LEU	PRO	LEU	LEU	LEU	LEU	LEU	LEU	GLY	PRO	ALA	GLY	ALA	ALA	ARG	ALA	GLN	GLU	ASP	GLU	ASP	GLY	ASP	TYR	GLU	GLU	LEU	VAL	LEU	ALA	LEU	ARG	SER	GLU	ASP	GLY	ALA	ALA	ALA	PRO	GLU	HIS	GLY	THR
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



- Molecule 2: Proprotein convertase subtilisin/kexin type 9



MET	GLY	THR	VAL	SER	SER	ARG	ARG	SER	TRP	TRP	PRO	LEU	PRO	LEU	LEU	LEU	LEU	LEU	LEU	GLY	PRO	ALA	GLY	ALA	ALA	ARG	ALA	GLN	GLU	ASP	GLU	ASP	GLY	ASP	TYR	GLU	GLU	LEU	VAL	LEU	ALA	LEU	ARG	SER	GLU	ASP	GLY	ALA	ALA	ALA	PRO	GLU	HIS	GLY	THR
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----



- Molecule 3: Low-density lipoprotein receptor



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	126.40Å 131.39Å 134.63Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.02 – 2.60 29.96 – 2.60	Depositor EDS
% Data completeness (in resolution range)	95.2 (30.02-2.60) 95.2 (29.96-2.60)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.53 (at 2.61Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, $R_{free}$	0.190 , 0.211 0.198 , 0.220	Depositor DCC
$R_{free}$ test set	3329 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	57.7	Xtriage
Anisotropy	0.335	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 43.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.016 for -h,l,k	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8999	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.85% 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: CA, NAG, FUC

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	0.49	0/3718	0.69	0/5048
1	B	0.51	0/3693	0.70	0/5014
2	C	0.51	0/757	0.73	0/1023
2	P	0.54	0/757	0.71	0/1023
3	D	0.56	0/190	0.70	0/255
All	All	0.51	0/9115	0.70	0/12363

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3646	0	3559	34	0
1	B	3623	0	3531	36	0
2	C	740	0	750	4	0
2	P	740	0	750	6	0
3	D	189	0	168	1	0
4	A	24	0	22	0	0
5	B	14	0	13	0	0
6	D	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	9	0	0	0	0
7	B	5	0	0	0	0
7	C	4	0	0	0	0
7	P	4	0	0	0	0
All	All	8999	0	8793	74	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 4.

All (74) 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:558:VAL:HG21	1:A:606:LEU:HD13	1.38	1.04
1:A:203:THR:O	1:A:204:ASP:HB2	1.57	1.02
1:B:203:THR:O	1:B:204:ASP:HB2	1.57	1.02
1:A:558:VAL:CG2	1:A:606:LEU:HD13	1.91	1.00
1:B:344:GLN:HE22	1:B:425:ASN:H	1.12	0.88
1:A:344:GLN:HE22	1:A:425:ASN:H	1.16	0.87
1:A:186:ASP:OD1	1:A:288:PRO:HG2	1.93	0.68
1:B:278:GLN:HE22	2:P:88:LEU:H	1.41	0.67
1:A:548:THR:HG22	1:A:596:ILE:HG22	1.77	0.67
1:A:641:THR:HG22	1:A:643:HIS:CE1	2.30	0.67
1:B:186:ASP:OD1	1:B:288:PRO:HG2	1.95	0.66
1:B:548:THR:HG22	1:B:596:ILE:HG22	1.79	0.65
1:B:641:THR:HG22	1:B:643:HIS:CE1	2.31	0.64
1:A:518:GLU:CD	1:A:518:GLU:H	2.02	0.63
1:A:212:ASP:H	1:A:256:GLN:HE22	1.46	0.62
1:B:212:ASP:H	1:B:256:GLN:HE22	1.46	0.62
1:B:518:GLU:H	1:B:518:GLU:CD	2.03	0.62
1:B:539:ALA:HB2	1:B:550:VAL:HG13	1.83	0.60
1:A:203:THR:O	1:A:204:ASP:CB	2.38	0.59
1:A:539:ALA:HB2	1:A:550:VAL:HG13	1.85	0.58
1:B:277:VAL:HG13	2:P:117:GLY:HA2	1.85	0.58
2:C:101:GLN:HE21	2:C:133:LEU:HD11	1.72	0.55
1:A:632:LEU:H	1:A:652:ASN:ND2	2.03	0.55
1:A:185:LEU:HD11	1:A:271:ILE:HD11	1.88	0.54
1:B:311:LEU:HB2	1:B:333:VAL:HG12	1.89	0.54
1:B:549:ARG:HG2	1:B:589:VAL:HG22	1.89	0.54
1:A:529:LEU:HG	1:A:532:ALA:HB2	1.89	0.54
1:B:185:LEU:HD11	1:B:271:ILE:HD11	1.89	0.54
1:B:632:LEU:H	1:B:652:ASN:ND2	2.06	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:549:ARG:HG2	1:A:589:VAL:HG22	1.90	0.54
1:A:476:ARG:HH21	1:A:503:GLN:NE2	2.07	0.53
1:A:237:ARG:O	1:A:243:LYS:HD2	2.09	0.52
1:A:200:VAL:HG22	1:A:247:MET:HB2	1.92	0.52
1:A:295:ARG:HH11	2:P:65:HIS:CE1	2.28	0.52
1:B:212:ASP:H	1:B:256:GLN:NE2	2.07	0.51
1:B:638:LEU:HB2	1:B:673:THR:HB	1.93	0.51
1:B:476:ARG:HH21	1:B:503:GLN:NE2	2.09	0.51
1:B:203:THR:O	1:B:204:ASP:CB	2.38	0.50
1:B:550:VAL:HG22	1:B:596:ILE:HG23	1.94	0.49
1:B:278:GLN:NE2	2:P:88:LEU:H	2.09	0.49
1:A:638:LEU:HB2	1:A:673:THR:HB	1.94	0.49
1:A:212:ASP:H	1:A:256:GLN:NE2	2.09	0.48
1:B:374:ASP:OD2	3:D:15:VAL:HG21	2.12	0.48
1:B:490:SER:HB2	1:B:520:VAL:HG12	1.94	0.48
1:A:557:HIS:CE1	1:A:602:HIS:HB2	2.49	0.48
1:B:615:ILE:HG13	1:B:616:PRO:HD2	1.95	0.48
1:B:361:LEU:HD22	1:B:442:ALA:HB2	1.95	0.47
1:B:200:VAL:HG22	1:B:247:MET:HB2	1.97	0.47
1:A:490:SER:HB2	1:A:520:VAL:HG12	1.94	0.47
1:A:361:LEU:HD22	1:A:442:ALA:HB2	1.96	0.47
1:A:550:VAL:HG22	1:A:596:ILE:HG23	1.96	0.47
2:C:79:VAL:HG22	2:C:123:LEU:HD13	1.97	0.46
1:B:551:HIS:HB3	1:B:587:GLN:HB2	1.97	0.46
1:A:551:HIS:HB3	1:A:587:GLN:HB2	1.97	0.46
1:A:303:ARG:HB2	1:A:303:ARG:HE	1.48	0.44
1:A:539:ALA:CB	1:A:550:VAL:HG13	2.48	0.44
1:A:277:VAL:HG12	1:A:278:GLN:HG2	1.99	0.43
2:C:100:ALA:O	2:C:104:ARG:HG2	2.18	0.43
1:B:558:VAL:HG21	1:B:608:CYS:SG	2.58	0.43
1:B:529:LEU:HA	1:B:530:PRO:HD2	1.66	0.43
1:B:237:ARG:O	1:B:243:LYS:HD2	2.18	0.43
1:A:591:HIS:CD2	1:A:593:GLU:H	2.37	0.42
1:B:591:HIS:CD2	1:B:593:GLU:H	2.38	0.42
1:A:632:LEU:H	1:A:652:ASN:HD22	1.67	0.42
1:B:277:VAL:HG13	2:P:117:GLY:CA	2.49	0.42
1:B:465:SER:HB3	1:B:473:ALA:HB2	2.01	0.41
1:B:539:ALA:CB	1:B:550:VAL:HG13	2.48	0.41
2:P:69:LYS:HD3	2:P:72:TRP:CE2	2.56	0.41
1:B:591:HIS:HD2	1:B:593:GLU:H	1.70	0.40
1:A:345:PRO:HD3	1:A:424:ILE:HG23	2.03	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:591:HIS:HD2	1:A:593:GLU:H	1.69	0.40
1:B:180:VAL:HG22	1:B:282:PRO:HG2	2.02	0.40
1:B:295:ARG:HH11	2:C:65:HIS:CE1	2.38	0.40
1:A:304:LEU:HD12	1:A:304:LEU:HA	1.93	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	482/540 (89%)	458 (95%)	18 (4%)	6 (1%)	15	32
1	B	479/540 (89%)	453 (95%)	19 (4%)	7 (2%)	12	24
2	C	90/152 (59%)	86 (96%)	4 (4%)	0	100	100
2	P	90/152 (59%)	87 (97%)	3 (3%)	0	100	100
3	D	24/26 (92%)	23 (96%)	1 (4%)	0	100	100
All	All	1165/1410 (83%)	1107 (95%)	45 (4%)	13 (1%)	17	35

All (13) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	245	ALA
1	B	530	PRO
1	A	543	GLU
1	A	642	SER
1	B	204	ASP
1	B	447	SER
1	B	543	GLU
1	A	204	ASP
1	B	642	SER
1	A	618	PRO

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	618	PRO
1	A	280	VAL
1	B	617	ALA

### 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	391/431 (91%)	364 (93%)	27 (7%)	18	36
1	B	389/431 (90%)	361 (93%)	28 (7%)	17	33
2	C	79/127 (62%)	71 (90%)	8 (10%)	9	16
2	P	79/127 (62%)	74 (94%)	5 (6%)	21	42
3	D	22/22 (100%)	21 (96%)	1 (4%)	32	59
All	All	960/1138 (84%)	891 (93%)	69 (7%)	18	33

All (69) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	178	SER
1	A	179	LEU
1	A	190	GLN
1	A	197	GLU
1	A	215	ARG
1	A	222	LYS
1	A	303	ARG
1	A	361	LEU
1	A	375	CYS
1	A	411	LEU
1	A	413	GLN
1	A	422	ASP
1	A	454	GLN
1	A	507	LEU
1	A	518	GLU
1	A	550	VAL

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	555	GLN
1	A	567	GLU
1	A	570	ASP
1	A	587	GLN
1	A	592	ARG
1	A	606	LEU
1	A	641	THR
1	A	642	SER
1	A	651[A]	ASP
1	A	651[B]	ASP
1	A	659	ARG
1	B	178	SER
1	B	179	LEU
1	B	190	GLN
1	B	194	ARG
1	B	219	GLN
1	B	222	LYS
1	B	283	LEU
1	B	304	LEU
1	B	361	LEU
1	B	375	CYS
1	B	411	LEU
1	B	413	GLN
1	B	422	ASP
1	B	454	GLN
1	B	518	GLU
1	B	531	GLN
1	B	550	VAL
1	B	555	GLN
1	B	567	GLU
1	B	570	ASP
1	B	587	GLN
1	B	592	ARG
1	B	606	LEU
1	B	641	THR
1	B	642	SER
1	B	651[A]	ASP
1	B	651[B]	ASP
1	B	659	ARG
2	C	66	ARG
2	C	70	ASP
2	C	96	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	C	97	ARG
2	C	104	ARG
2	C	108	LEU
2	C	118	LEU
2	C	133	LEU
3	D	2	THR
2	P	66	ARG
2	P	70	ASP
2	P	104	ARG
2	P	108	LEU
2	P	118	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (28) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	256	GLN
1	A	278	GLN
1	A	298	ASN
1	A	344	GLN
1	A	413	GLN
1	A	464	HIS
1	A	513	ASN
1	A	587	GLN
1	A	591	HIS
1	A	602	HIS
1	A	643	HIS
1	A	652	ASN
1	B	256	GLN
1	B	278	GLN
1	B	298	ASN
1	B	344	GLN
1	B	413	GLN
1	B	464	HIS
1	B	513	ASN
1	B	537	HIS
1	B	554	GLN
1	B	587	GLN
1	B	591	HIS
1	B	643	HIS
1	B	652	ASN
2	C	65	HIS
2	C	101	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	P	65	HIS

### 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 ⓘ

2 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	$\# Z  > 2$
4	NAG	A	901	1,4	14,14,15	0.32	0	15,19,21	0.58	0
4	FUC	A	902	4	9,10,11	0.45	0	13,14,16	0.96	1 (7%)

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
4	NAG	A	901	1,4	-	0/6/23/26	0/1/1/1
4	FUC	A	902	4	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	902	FUC	C1-O5-C5	2.90	118.80	112.39

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.6 Ligand geometry

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	$\# Z  > 2$
5	NAG	B	901	1	14,14,15	0.32	0	15,19,21	0.79	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
5	NAG	B	901	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	901	NAG	C1-O5-C5	2.65	115.83	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	491/540 (90%)	0.25	40 (8%) 13 8	34, 61, 127, 158	0
1	B	490/540 (90%)	0.48	62 (12%) 4 2	36, 65, 139, 189	0
2	C	92/152 (60%)	-0.33	0 100 100	41, 60, 87, 105	0
2	P	92/152 (60%)	-0.48	0 100 100	35, 47, 70, 87	0
3	D	26/26 (100%)	0.26	1 (3%) 41 33	50, 61, 79, 92	0
All	All	1191/1410 (84%)	0.24	103 (8%) 11 7	34, 61, 127, 189	0

All (103) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	178	SER	12.6
1	A	178	SER	12.2
1	B	618	PRO	10.8
1	B	448	THR	8.3
1	B	168	ALA	8.1
1	B	219	GLN	8.0
1	A	618	PRO	7.2
1	B	640	GLY	7.0
1	B	220	ALA	6.9
1	B	545	SER	6.6
1	B	546	MET	6.6
1	B	447	SER	6.5
1	B	592	ARG	6.4
1	A	515	PHE	5.8
1	A	153	SER	5.8
3	D	1	GLY	5.7
1	A	167	ARG	5.3
1	A	165	ARG	5.3
1	A	546	MET	5.2
1	B	221	SER	5.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	640	GLY	4.8
1	B	569	GLU	4.7
1	B	615	ILE	4.6
1	B	165	ARG	4.6
1	A	544	ALA	4.6
1	B	515	PHE	4.5
1	B	543	GLU	4.5
1	B	177	GLY	4.5
1	A	545	SER	4.3
1	B	547	GLY	4.3
1	B	153	SER	4.0
1	B	176	GLY	4.0
1	B	643	HIS	4.0
1	B	544	ALA	3.9
1	B	562	CYS	3.9
1	A	642	SER	3.9
1	B	620	GLU	3.9
1	B	589	VAL	3.8
1	A	469	ARG	3.7
1	B	616	PRO	3.7
1	B	504	GLY	3.6
1	B	542	ALA	3.6
1	B	593	GLU	3.6
1	A	280	VAL	3.5
1	B	470	MET	3.4
1	B	617	ALA	3.4
1	B	619	GLN	3.4
1	A	389	ALA	3.3
1	B	166	TYR	3.2
1	A	549	ARG	3.2
1	B	659	ARG	3.2
1	B	469	ARG	3.1
1	B	682	ARG	3.1
1	B	549	ARG	3.1
1	B	614	GLY	3.1
1	B	517	GLY	3.1
1	A	166	TYR	3.0
1	B	642	SER	2.9
1	B	164	PRO	2.9
1	B	564	SER	2.9
1	B	471	ALA	2.9
1	A	446	PRO	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	390	ALA	2.8
1	A	563	SER	2.8
1	B	570	ASP	2.8
1	B	537	HIS	2.7
1	A	569	GLU	2.7
1	A	154	ILE	2.7
1	A	514	ALA	2.6
1	A	543	GLU	2.6
1	B	561	GLY	2.6
1	B	572	GLY	2.6
1	A	616	PRO	2.6
1	A	286	LEU	2.6
1	A	564	SER	2.6
1	A	164	PRO	2.5
1	B	467	PRO	2.5
1	B	553	HIS	2.5
1	A	620	GLU	2.5
1	A	517	GLY	2.4
1	B	591	HIS	2.4
1	A	571	LEU	2.3
1	A	504	GLY	2.3
1	B	496	ARG	2.3
1	A	392	VAL	2.3
1	A	470	MET	2.3
1	A	592	ARG	2.3
1	A	619	GLN	2.2
1	B	286	LEU	2.2
1	B	449	HIS	2.2
1	A	467	PRO	2.1
1	B	389	ALA	2.1
1	B	531	GLN	2.1
1	A	393	ALA	2.1
1	B	635	CYS	2.1
1	B	446	PRO	2.1
1	A	494	LYS	2.1
1	B	488	SER	2.1
1	B	563	SER	2.1
1	B	636	SER	2.1
1	A	547	GLY	2.1
1	A	615	ILE	2.0
1	B	637	ALA	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](#)

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
4	NAG	A	901	14/15	0.83	0.39	-	116,133,142,144	0
4	FUC	A	902	10/11	0.57	0.46	-	139,146,149,151	0

## 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(Å <sup>2</sup> )	Q<0.9
6	CA	D	101	1/1	0.98	0.07	-2.76	74,74,74,74	0
5	NAG	B	901	14/15	0.80	0.30	-	123,134,138,141	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.