



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 04:43 AM GMT

PDB ID : 2NY1
Title : HIV-1 gp120 Envelope Glycoprotein (I109C, T257S, S334A, S375W, Q428C)
Complexed with CD4 and Antibody 17b
Authors : Zhou, T.; Xu, L.; Dey, B.; Hessel, A.J.; Van Ryk, D.; Xiang, S.H.; Yang, X;
Zhang, M.Y.; Zwick, M.B.; Arthos, J.; Burton, D.R.; Dimitrov, D.S.; Sodroski,
J.; Wyatt, R.; Nabel, G.J.; Kwong, P.D.
Deposited on : 2006-11-20
Resolution : 1.99 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
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 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

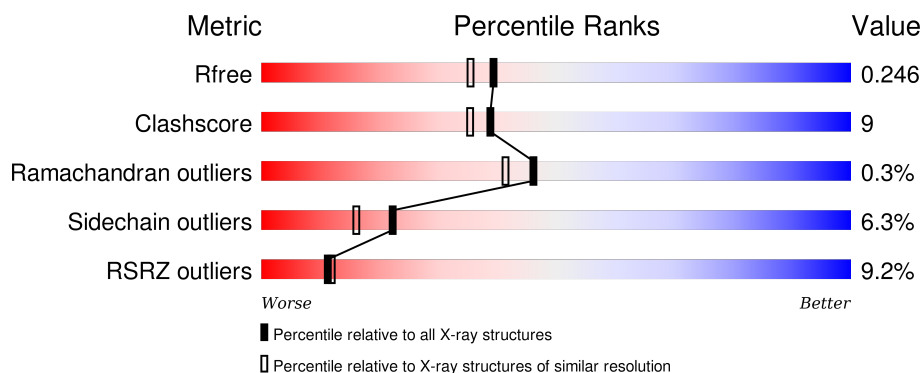
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.99 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (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 ≥ 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	317	<div> <div>15%</div> <div>75%</div> <div>20%</div> <div>• •</div> </div>
2	B	184	<div> <div>14%</div> <div>79%</div> <div>17%</div> <div>• •</div> </div>
3	C	214	<div> <div>3%</div> <div>85%</div> <div>14%</div> <div>•</div> </div>
4	D	229	<div> <div>3%</div> <div>83%</div> <div>14%</div> <div>• •</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
5	NAG	A	730	-	-	-	X
5	NAG	A	734	-	-	-	X
6	SUC	D	4000	-	-	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7852 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 ENVELOPE GLYCOPROTEIN GP120.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	305	Total	C	N	O	S	0	0	0
			2362	1482	411	447	22			

- Molecule 2 is a protein called T-cell surface glycoprotein CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	181	Total	C	N	O	S	0	0	0
			1412	885	247	276	4			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1000	MET	-	INITIATING METHIONINE	UNP P01730

- Molecule 3 is a protein called ANTIBODY 17B, LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	214	Total	C	N	O	S	0	0	0
			1646	1028	282	331	5			

- Molecule 4 is a protein called ANTIBODY 17B, HEAVY CHAIN.

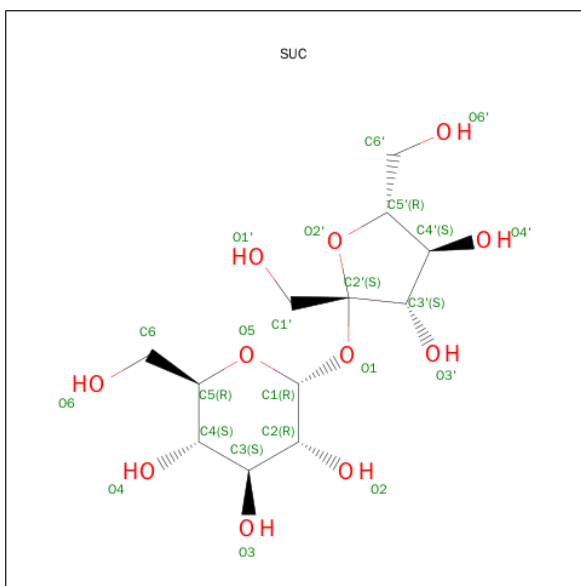
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	223	Total	C	N	O	S	0	0	0
			1681	1064	282	330	5			

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C₈H₁₅NO₆).



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

- Molecule 6 is SUGAR (SUCROSE) (three-letter code: SUC) (formula: $C_{12}H_{22}O_{11}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	D	1	Total	C	O	0	0
			23	12	11		

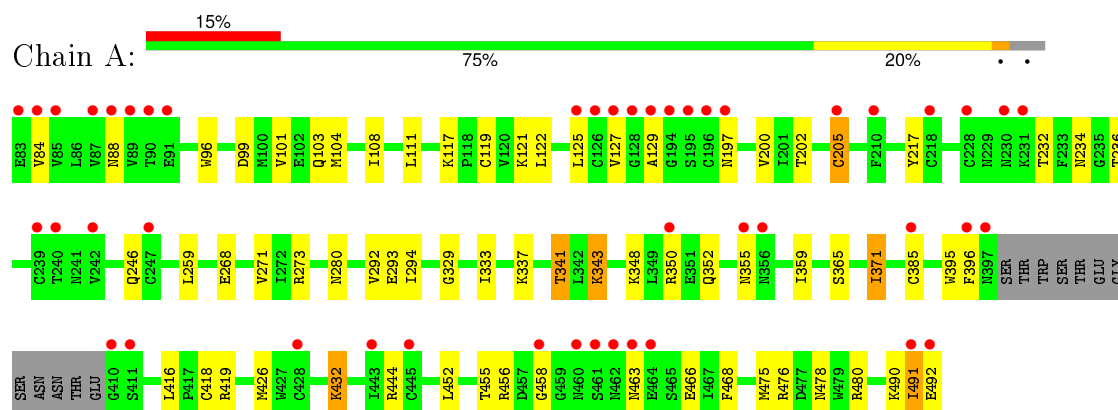
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	154	Total	O	0	0
			154	154		
7	B	86	Total	O	0	0
			86	86		
7	C	131	Total	O	0	0
			131	131		
7	D	161	Total	O	0	0
			161	161		

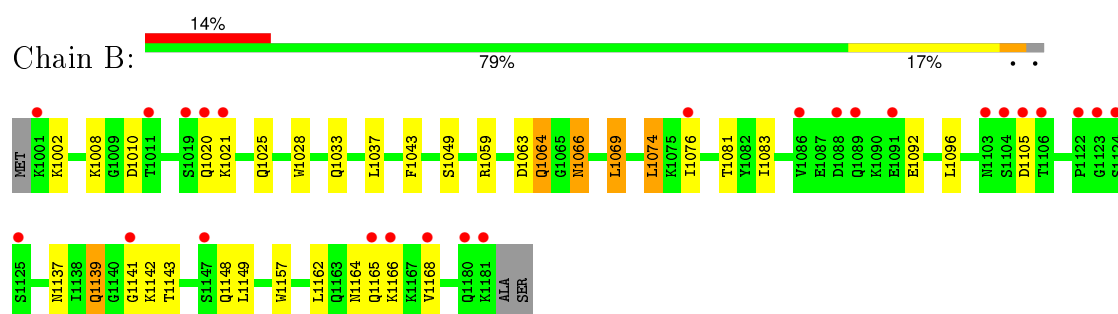
3 Residue-property plots

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 ($\text{RSRZ} > 2$). 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.

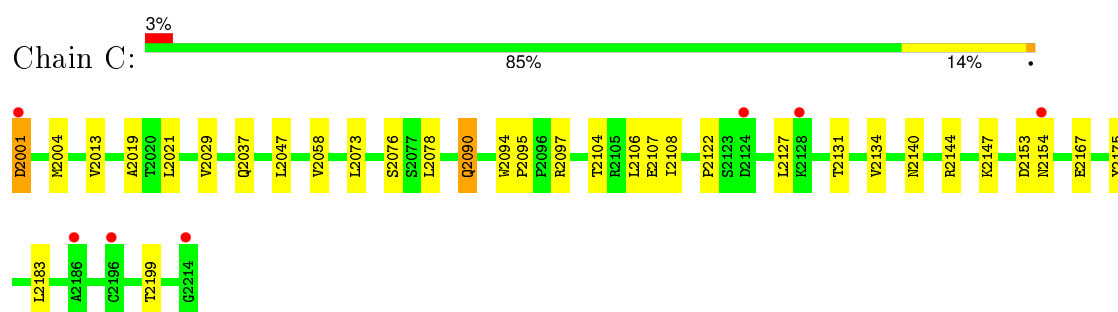
• Molecule 1: ENVELOPE GLYCOPROTEIN GP120



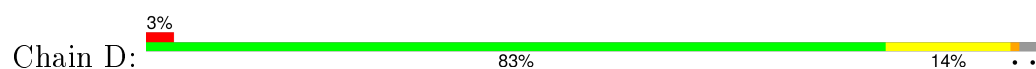
• Molecule 2: T-cell surface glycoprotein CD4

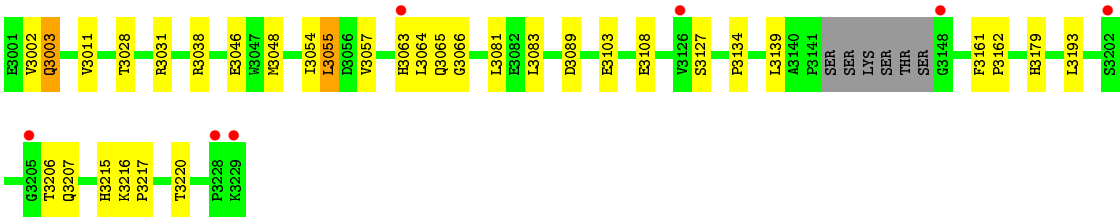


• Molecule 3: ANTIBODY 17B, LIGHT CHAIN



• Molecule 4: ANTIBODY 17B, HEAVY CHAIN





4 Data and refinement statistics

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	71.11Å 88.16Å 197.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.40 – 1.99 42.38 – 2.00	Depositor EDS
% Data completeness (in resolution range)	81.0 (42.40-1.99) 81.5 (42.38-2.00)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.18 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.209 , 0.250 0.209 , 0.246	Depositor DCC
R_{free} test set	3536 reflections (5.35%)	DCC
Wilson B-factor (Å ²)	24.6	Xtriage
Anisotropy	0.205	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 46.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 69592 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7852	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.20% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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

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	0.32	0/2410	0.49	0/3267
2	B	0.29	0/1432	0.51	0/1930
3	C	0.32	0/1683	0.53	0/2288
4	D	0.35	0/1720	0.54	0/2343
All	All	0.32	0/7245	0.51	0/9828

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	2362	0	2294	59	0
2	B	1412	0	1441	19	0
3	C	1646	0	1590	29	0
4	D	1681	0	1647	26	0
5	A	196	0	182	6	0
6	D	23	0	22	3	0
7	A	154	0	0	12	0
7	B	86	0	0	8	0
7	C	131	0	0	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	D	161	0	0	6	0
All	All	7852	0	7176	130	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 9.

All (130) 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)
3:C:2144:ARG:NH1	7:C:355:HOH:O	1.79	1.09
1:A:476:ARG:HB3	1:A:480:ARG:HH12	1.28	0.97
2:B:1025:GLN:HG2	7:B:442:HOH:O	1.66	0.95
3:C:2167:GLU:O	7:C:205:HOH:O	1.86	0.93
2:B:1066:ASN:ND2	7:B:74:HOH:O	2.02	0.89
4:D:3215:HIS:HB3	4:D:3220:THR:HG22	1.55	0.88
4:D:3215:HIS:HB3	4:D:3220:THR:CG2	2.06	0.85
4:D:3046:GLU:OE2	4:D:3063:HIS:NE2	2.15	0.79
1:A:491:ILE:O	1:A:492:GLU:HG3	1.82	0.78
4:D:3065:GLN:NE2	7:D:8:HOH:O	2.16	0.77
4:D:3048:MET:HG2	7:D:435:HOH:O	1.81	0.77
7:A:1104:HOH:O	4:D:3108:GLU:HB2	1.83	0.77
1:A:476:ARG:HB3	1:A:480:ARG:NH1	1.99	0.76
1:A:280:ASN:HD22	1:A:458:GLY:HA3	1.51	0.75
1:A:395:TRP:CZ3	5:A:839:NAG:O7	2.40	0.75
1:A:491:ILE:HG13	1:A:492:GLU:H	1.52	0.73
1:A:280:ASN:ND2	1:A:458:GLY:HA3	2.03	0.73
1:A:395:TRP:HZ3	5:A:839:NAG:O7	1.72	0.71
1:A:456:ARG:HB2	7:A:1040:HOH:O	1.92	0.69
1:A:292:VAL:HG21	1:A:341:THR:HG21	1.73	0.69
3:C:2144:ARG:NE	7:C:500:HOH:O	1.97	0.69
2:B:1148:GLN:OE1	7:B:103:HOH:O	2.13	0.67
4:D:3134:PRO:HD2	4:D:3220:THR:HG21	1.74	0.67
1:A:456:ARG:HD3	7:A:1088:HOH:O	1.95	0.67
3:C:2013:VAL:HG11	3:C:2078:LEU:CD1	2.25	0.67
3:C:2013:VAL:HG11	3:C:2078:LEU:HD12	1.76	0.66
3:C:2107:GLU:HG2	3:C:2108:ILE:N	2.13	0.64
2:B:1033:GLN:NE2	7:B:494:HOH:O	2.08	0.64
2:B:1164:ASN:O	2:B:1165:GLN:HB2	1.99	0.62
1:A:104:MET:O	1:A:108:ILE:HG12	1.99	0.62
1:A:337:LYS:O	1:A:341:THR:HG23	2.02	0.60
1:A:292:VAL:CG2	1:A:341:THR:HG21	2.32	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:1066:ASN:HB2	7:B:396:HOH:O	2.03	0.59
1:A:294:ILE:HD12	1:A:333:ILE:HD11	1.85	0.58
5:A:588:NAG:C3	5:A:588:NAG:O7	2.51	0.58
4:D:3063:HIS:CE1	6:D:4000:SUC:O6'	2.57	0.57
1:A:121:LYS:HB2	1:A:426:MET:HE3	1.85	0.57
1:A:119:CYS:N	1:A:205:CYS:SG	2.76	0.57
3:C:2029:VAL:HG11	3:C:2090:GLN:CG	2.35	0.57
3:C:2078:LEU:HD21	3:C:2106:LEU:HD21	1.88	0.56
4:D:3046:GLU:OE2	4:D:3063:HIS:CE1	2.59	0.56
4:D:3065:GLN:HG3	7:D:174:HOH:O	2.06	0.56
3:C:2029:VAL:HG11	3:C:2090:GLN:HG3	1.87	0.56
1:A:350:ARG:NH2	1:A:355:ASN:O	2.40	0.55
1:A:491:ILE:HG13	1:A:492:GLU:N	2.21	0.54
1:A:280:ASN:HD22	1:A:458:GLY:CA	2.19	0.54
3:C:2019:ALA:HB2	3:C:2078:LEU:HD11	1.88	0.54
1:A:432:LYS:HE2	7:A:1089:HOH:O	2.07	0.54
1:A:350:ARG:HD3	1:A:396:PHE:CE1	2.42	0.54
1:A:350:ARG:HD3	1:A:396:PHE:CD1	2.43	0.54
1:A:491:ILE:CG1	1:A:492:GLU:H	2.19	0.53
2:B:1157:TRP:HB3	7:B:440:HOH:O	2.09	0.53
4:D:3065:GLN:CG	7:D:174:HOH:O	2.55	0.53
5:A:588:NAG:O3	5:A:588:NAG:O7	2.26	0.53
3:C:2037:GLN:HB2	3:C:2047:LEU:HD11	1.91	0.53
1:A:293:GLU:OE1	5:A:795:NAG:H3	2.08	0.52
1:A:259:LEU:HD23	1:A:452:LEU:HD22	1.92	0.52
4:D:3002:VAL:C	4:D:3003:GLN:NE2	2.63	0.52
2:B:1083:ILE:HG13	2:B:1092:GLU:HG3	1.92	0.52
3:C:2021:LEU:HD22	3:C:2104:THR:HG21	1.91	0.52
1:A:329:GLY:C	1:A:418:CYS:HB2	2.31	0.51
4:D:3002:VAL:O	4:D:3003:GLN:NE2	2.44	0.50
3:C:2175:TYR:HA	7:C:205:HOH:O	2.11	0.50
3:C:2004:MET:SD	3:C:2090:GLN:HB2	2.52	0.50
3:C:2013:VAL:CG1	3:C:2078:LEU:CD1	2.90	0.49
3:C:2094:TRP:HA	3:C:2095:PRO:C	2.33	0.49
1:A:122:LEU:HD21	4:D:3057:VAL:HG21	1.95	0.49
1:A:259:LEU:CD2	1:A:452:LEU:HD22	2.43	0.49
1:A:125:LEU:HB3	1:A:197:ASN:O	2.13	0.48
1:A:129:ALA:HB3	7:A:1073:HOH:O	2.13	0.48
4:D:3011:VAL:HG21	4:D:3162:PRO:HG3	1.95	0.48
1:A:480:ARG:CZ	7:A:1024:HOH:O	2.62	0.48
3:C:2140:ASN:HD21	4:D:3179:HIS:CE1	2.32	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:3066:GLY:HA3	7:D:504:HOH:O	2.14	0.47
3:C:2097:ARG:HD2	7:C:460:HOH:O	2.13	0.47
4:D:3216:LYS:N	4:D:3217:PRO:CD	2.78	0.47
2:B:1010:ASP:O	2:B:1074:LEU:HD22	2.15	0.47
6:D:4000:SUC:O1'	7:D:39:HOH:O	2.20	0.47
1:A:456:ARG:HG2	7:A:1088:HOH:O	2.14	0.47
3:C:2147:LYS:HB3	3:C:2199:THR:HB	1.97	0.47
1:A:104:MET:HE3	1:A:217:TYR:CE1	2.50	0.47
3:C:2001:ASP:OD2	3:C:2001:ASP:N	2.48	0.47
1:A:117:LYS:HD3	7:A:1020:HOH:O	2.14	0.46
1:A:352:GLN:NE2	7:A:976:HOH:O	2.05	0.46
4:D:3063:HIS:CE1	6:D:4000:SUC:HO6'	2.34	0.46
1:A:478:ASN:H	1:A:478:ASN:HD22	1.63	0.46
3:C:2122:PRO:HD3	3:C:2134:VAL:HG22	1.98	0.46
2:B:1008:LYS:HD2	2:B:1076:ILE:HD11	1.98	0.46
1:A:259:LEU:HD22	1:A:452:LEU:CD2	2.46	0.45
3:C:2013:VAL:CG1	3:C:2078:LEU:HD12	2.45	0.45
3:C:2153:ASP:O	3:C:2154:ASN:CB	2.64	0.45
4:D:3193:LEU:HD12	4:D:3193:LEU:C	2.38	0.44
1:A:419:ARG:NH2	4:D:3103:GLU:OE2	2.36	0.44
1:A:96:TRP:CD1	1:A:236:THR:HG21	2.52	0.44
2:B:1141:GLY:C	7:B:493:HOH:O	2.56	0.44
2:B:1020:GLN:HG2	7:B:130:HOH:O	2.18	0.44
2:B:1081:THR:CG2	2:B:1092:GLU:HG2	2.48	0.44
1:A:234:ASN:CG	5:A:734:NAG:N2	2.71	0.44
1:A:232:THR:HG22	1:A:232:THR:O	2.17	0.44
1:A:456:ARG:HB3	1:A:468:PHE:CE2	2.53	0.44
2:B:1166:LYS:N	2:B:1166:LYS:HD2	2.33	0.44
4:D:3028:THR:HB	4:D:3031:ARG:HD2	2.00	0.44
1:A:466:GLU:HB3	7:A:1040:HOH:O	2.18	0.44
1:A:271:VAL:HG11	1:A:273:ARG:NH1	2.32	0.44
4:D:3206:THR:OG1	4:D:3207:GLN:N	2.51	0.44
1:A:121:LYS:HE2	7:A:1093:HOH:O	2.16	0.43
3:C:2153:ASP:O	3:C:2154:ASN:HB3	2.18	0.43
2:B:1021:LYS:HE3	2:B:1064:GLN:HE21	1.84	0.43
1:A:348:LYS:HA	1:A:348:LYS:HD3	1.88	0.43
3:C:2107:GLU:HG3	3:C:2175:TYR:OH	2.19	0.43
1:A:121:LYS:HB2	1:A:426:MET:CE	2.48	0.43
3:C:2144:ARG:NH2	7:C:500:HOH:O	2.51	0.43
1:A:202:THR:HG22	3:C:2095:PRO:HG3	2.00	0.43
1:A:122:LEU:HD22	1:A:200:VAL:HG22	1.99	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:490:LYS:NZ	7:A:1066:HOH:O	2.52	0.43
4:D:3054:ILE:HG23	4:D:3055:LEU:HD13	2.00	0.42
1:A:104:MET:CE	1:A:217:TYR:CE1	3.02	0.42
1:A:268:GLU:OE1	1:A:268:GLU:HA	2.19	0.42
2:B:1028:TRP:CE2	2:B:1069:LEU:HB2	2.56	0.41
2:B:1139:GLN:HB2	2:B:1139:GLN:HE21	1.70	0.41
1:A:343:LYS:HE2	1:A:343:LYS:HB3	1.76	0.41
1:A:101:VAL:HG11	1:A:480:ARG:NH1	2.35	0.41
1:A:371:ILE:HG12	2:B:1043:PHE:HB3	2.02	0.41
1:A:259:LEU:CD2	1:A:452:LEU:CD2	2.99	0.41
3:C:2140:ASN:HD21	4:D:3179:HIS:HE1	1.68	0.41
1:A:96:TRP:HD1	1:A:236:THR:HG21	1.86	0.41
4:D:3127:SER:HB3	4:D:3161:PHE:CZ	2.56	0.41
2:B:1164:ASN:O	2:B:1166:LYS:HD2	2.22	0.40
3:C:2047:LEU:HA	3:C:2058:VAL:HG21	2.03	0.40
1:A:478:ASN:N	1:A:478:ASN:HD22	2.19	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	301/317 (95%)	283 (94%)	16 (5%)	2 (1%)	26	19
2	B	179/184 (97%)	171 (96%)	7 (4%)	1 (1%)	30	22
3	C	212/214 (99%)	204 (96%)	8 (4%)	0	100	100
4	D	219/229 (96%)	213 (97%)	6 (3%)	0	100	100
All	All	911/944 (96%)	871 (96%)	37 (4%)	3 (0%)	46	41

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	127	VAL
2	B	1105	ASP
1	A	491	ILE

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	269/280 (96%)	250 (93%)	19 (7%)	18	12
2	B	164/166 (99%)	147 (90%)	17 (10%)	9	4
3	C	184/184 (100%)	177 (96%)	7 (4%)	40	36
4	D	187/193 (97%)	179 (96%)	8 (4%)	35	30
All	All	804/823 (98%)	753 (94%)	51 (6%)	22	16

All (51) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	84	VAL
1	A	88	ASN
1	A	99	ASP
1	A	103	GLN
1	A	111	LEU
1	A	205	CYS
1	A	246	GLN
1	A	341	THR
1	A	343	LYS
1	A	359	ILE
1	A	365	SER
1	A	371	ILE
1	A	385	CYS
1	A	416	LEU
1	A	432	LYS
1	A	444	ARG
1	A	455	THR
1	A	463	ASN
1	A	475	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	1002	LYS
2	B	1037	LEU
2	B	1049	SER
2	B	1059	ARG
2	B	1063	ASP
2	B	1064	GLN
2	B	1066	ASN
2	B	1069	LEU
2	B	1074	LEU
2	B	1096	LEU
2	B	1137	ASN
2	B	1139	GLN
2	B	1142	LYS
2	B	1143	THR
2	B	1149	LEU
2	B	1162	LEU
2	B	1168	VAL
3	C	2001	ASP
3	C	2073	LEU
3	C	2076	SER
3	C	2090	GLN
3	C	2127	LEU
3	C	2131	THR
3	C	2183	LEU
4	D	3003	GLN
4	D	3038	ARG
4	D	3055	LEU
4	D	3064	LEU
4	D	3081	LEU
4	D	3083	LEU
4	D	3089	ASP
4	D	3139	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (15) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	103	GLN
1	A	355	ASN
1	A	377	ASN
1	A	442	GLN
1	A	478	ASN
2	B	1040	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	1064	GLN
2	B	1073	ASN
2	B	1148	GLN
3	C	2140	ASN
3	C	2149	GLN
4	D	3003	GLN
4	D	3065	GLN
4	D	3179	HIS
4	D	3207	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

15 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	$\# Z > 2$
5	NAG	A	588	1	14,14,15	0.42	0	15,19,21	0.79	0
5	NAG	A	697	1	14,14,15	0.47	0	15,19,21	0.95	1 (6%)
5	NAG	A	730	1	14,14,15	0.56	0	15,19,21	0.82	0
5	NAG	A	734	1	14,14,15	0.42	0	15,19,21	2.62	5 (33%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	NAG	A	741	1	14,14,15	0.56	0	15,19,21	0.90	0
5	NAG	A	762	1	14,14,15	0.53	0	15,19,21	0.95	0
5	NAG	A	776	1	14,14,15	0.53	0	15,19,21	0.82	0
5	NAG	A	789	1	14,14,15	0.55	0	15,19,21	0.73	0
5	NAG	A	795	1	14,14,15	0.53	0	15,19,21	1.22	0
5	NAG	A	839	1	14,14,15	0.64	0	15,19,21	1.00	1 (6%)
5	NAG	A	886	1	14,14,15	0.71	0	15,19,21	0.65	0
5	NAG	A	892	1	14,14,15	0.46	0	15,19,21	0.80	0
5	NAG	A	948	1	14,14,15	0.48	0	15,19,21	0.81	0
5	NAG	A	963	1	14,14,15	0.56	0	15,19,21	0.98	1 (6%)
6	SUC	D	4000	-	24,24,24	0.40	0	36,36,36	0.65	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
5	NAG	A	588	1	-	0/6/23/26	0/1/1/1
5	NAG	A	697	1	-	0/6/23/26	0/1/1/1
5	NAG	A	730	1	-	0/6/23/26	0/1/1/1
5	NAG	A	734	1	-	0/6/23/26	0/1/1/1
5	NAG	A	741	1	-	0/6/23/26	0/1/1/1
5	NAG	A	762	1	-	0/6/23/26	0/1/1/1
5	NAG	A	776	1	-	0/6/23/26	0/1/1/1
5	NAG	A	789	1	-	0/6/23/26	0/1/1/1
5	NAG	A	795	1	-	0/6/23/26	0/1/1/1
5	NAG	A	839	1	-	0/6/23/26	0/1/1/1
5	NAG	A	886	1	-	0/6/23/26	0/1/1/1
5	NAG	A	892	1	-	1/6/23/26	0/1/1/1
5	NAG	A	948	1	-	0/6/23/26	0/1/1/1
5	NAG	A	963	1	-	0/6/23/26	0/1/1/1
6	SUC	D	4000	-	-	0/12/51/51	0/2/2/2

There are no bond length outliers.

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	734	NAG	C6-C5-C4	-5.24	100.09	113.02
5	A	734	NAG	O5-C5-C6	-4.93	96.67	107.35
5	A	734	NAG	C4-C3-C2	-3.07	106.46	111.23

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	697	NAG	C2-N2-C7	-2.31	120.06	123.04
5	A	963	NAG	C1-O5-C5	2.81	115.81	112.25
5	A	839	NAG	C4-C3-C2	3.10	116.05	111.23
5	A	734	NAG	O6-C6-C5	3.17	121.82	111.33
5	A	734	NAG	C1-O5-C5	4.86	118.41	112.25

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	892	NAG	O7-C7-N2-C2

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	588	NAG	2	0
5	A	734	NAG	1	0
5	A	795	NAG	1	0
5	A	839	NAG	2	0
6	D	4000	SUC	3	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 ⓘ

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, 95th 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 ²)	Q<0.9
1	A	305/317 (96%)	0.99	46 (15%) 3 3	14, 31, 70, 95	0
2	B	181/184 (98%)	0.82	25 (13%) 4 4	17, 35, 62, 76	0
3	C	214/214 (100%)	0.29	7 (3%) 50 51	12, 28, 61, 74	0
4	D	223/229 (97%)	0.44	7 (3%) 52 53	11, 23, 61, 89	0
All	All	923/944 (97%)	0.66	85 (9%) 11 12	11, 30, 65, 95	0

All (85) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	1106	THR	8.8
1	A	195	SER	8.2
1	A	410	GLY	7.7
1	A	492	GLU	7.3
1	A	129	ALA	6.5
1	A	128	GLY	6.4
4	D	3148	GLY	5.9
1	A	87	VAL	5.8
1	A	127	VAL	5.8
1	A	210	PHE	5.7
1	A	126	CYS	5.1
2	B	1001	LYS	5.1
1	A	84	VAL	4.6
1	A	197	ASN	4.6
1	A	205	CYS	4.5
1	A	228	CYS	4.5
1	A	385	CYS	4.2
2	B	1125	SER	4.2
2	B	1103	ASN	4.0
1	A	464	GLU	4.0
1	A	194	GLY	3.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	C	2001	ASP	3.9
2	B	1181	LYS	3.8
1	A	411	SER	3.6
1	A	196	CYS	3.6
1	A	88	ASN	3.6
4	D	3205	GLY	3.5
1	A	428	CYS	3.3
1	A	445	CYS	3.3
2	B	1019	SER	3.2
1	A	89	VAL	3.2
2	B	1104	SER	3.2
1	A	461	SER	3.1
1	A	125	LEU	3.1
1	A	356	ASN	3.1
3	C	2196	CYS	3.1
2	B	1141	GLY	3.0
1	A	231	LYS	3.0
1	A	240	THR	3.0
2	B	1123	GLY	3.0
1	A	491	ILE	3.0
2	B	1166	LYS	2.9
4	D	3229	LYS	2.9
1	A	462	ASN	2.9
1	A	90	THR	2.7
1	A	397	ASN	2.7
1	A	239	CYS	2.7
2	B	1180	GLN	2.6
1	A	91	GLU	2.6
2	B	1091	GLU	2.6
2	B	1021	LYS	2.5
1	A	218	CYS	2.5
2	B	1124	SER	2.5
2	B	1122	PRO	2.5
3	C	2186	ALA	2.4
3	C	2128	LYS	2.4
3	C	2214	GLY	2.4
1	A	247	CYS	2.4
4	D	3202	SER	2.4
3	C	2154	ASN	2.4
4	D	3228	PRO	2.4
4	D	3063	HIS	2.3
1	A	230	ASN	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	85	VAL	2.3
1	A	463	ASN	2.3
2	B	1086	VAL	2.3
2	B	1147	SER	2.2
1	A	443	ILE	2.2
1	A	83	GLU	2.2
1	A	350	ARG	2.2
2	B	1105	ASP	2.2
1	A	458	GLY	2.1
2	B	1165	GLN	2.1
1	A	396	PHE	2.1
3	C	2124	ASP	2.1
2	B	1168	VAL	2.1
2	B	1089	GLN	2.1
2	B	1020	GLN	2.1
1	A	242	VAL	2.1
2	B	1011	THR	2.0
1	A	355	ASN	2.0
2	B	1076	ILE	2.0
2	B	1088	ASP	2.0
1	A	460	ASN	2.0
4	D	3126	VAL	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, 95th 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(\AA^2)	Q<0.9
5	NAG	A	730	14/15	0.65	0.53	4.38	55,57,57,58	0
6	SUC	D	4000	23/23	0.80	0.27	3.44	46,47,48,48	0
5	NAG	A	734	14/15	0.61	0.29	2.78	49,51,51,52	0
5	NAG	A	789	14/15	0.92	0.21	1.62	35,37,39,39	0
5	NAG	A	948	14/15	0.81	0.20	0.72	39,40,42,42	0
5	NAG	A	886	14/15	0.86	0.15	0.12	39,41,41,42	0
5	NAG	A	795	14/15	0.91	0.14	0.07	28,29,30,31	0
5	NAG	A	776	14/15	0.89	0.14	-0.46	38,40,41,41	0
5	NAG	A	762	14/15	0.91	0.14	-0.60	30,33,34,36	0
5	NAG	A	588	14/15	0.54	0.53	-	62,63,64,64	0
5	NAG	A	892	14/15	0.71	0.39	-	44,46,48,48	0
5	NAG	A	839	14/15	0.57	0.52	-	47,49,51,51	0
5	NAG	A	963	14/15	0.40	0.56	-	60,62,62,62	0
5	NAG	A	741	14/15	0.70	0.39	-	50,52,53,54	0
5	NAG	A	697	14/15	0.66	0.49	-	52,54,54,55	0

6.5 Other polymers [i](#)

There are no such residues in this entry.