wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 12:21 PM GMT

PDB ID	:	3R05
Title	:	Structure of neurexin 1 alpha (domains LNS1-LNS6), with splice insert SS3
Authors	:	Venugopal, V.; Rudenko, G.
Deposited on	:	2011-03-07
Resolution	:	2.95 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

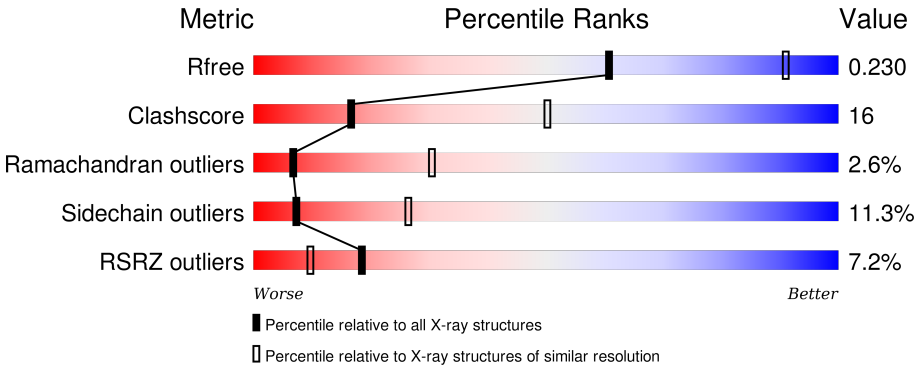
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.95 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	2184 (3.00-2.92)
Clashscore	102246	2552 (3.00-2.92)
Ramachandran outliers	100387	2468 (3.00-2.92)
Sidechain outliers	100360	2471 (3.00-2.92)
RSRZ outliers	91569	2201 (3.00-2.92)

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 <=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	1254	6% 50% 24% 5% 20%
1	B	1254	5% 50% 24% 6% 20%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 15522 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 Neurexin-1-alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1008	Total	C	N	O	S	0	0	0
			7733	4857	1334	1499	43			
1	B	1008	Total	C	N	O	S	0	0	0
			7733	4857	1334	1499	43			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1340	ALA	-	EXPRESSION TAG	UNP Q28146
A	1341	SER	-	EXPRESSION TAG	UNP Q28146
A	1342	THR	-	EXPRESSION TAG	UNP Q28146
A	1343	SER	-	EXPRESSION TAG	UNP Q28146
A	1344	HIS	-	EXPRESSION TAG	UNP Q28146
A	1345	HIS	-	EXPRESSION TAG	UNP Q28146
A	1346	HIS	-	EXPRESSION TAG	UNP Q28146
A	1347	HIS	-	EXPRESSION TAG	UNP Q28146
A	1348	HIS	-	EXPRESSION TAG	UNP Q28146
A	1349	HIS	-	EXPRESSION TAG	UNP Q28146
B	1340	ALA	-	EXPRESSION TAG	UNP Q28146
B	1341	SER	-	EXPRESSION TAG	UNP Q28146
B	1342	THR	-	EXPRESSION TAG	UNP Q28146
B	1343	SER	-	EXPRESSION TAG	UNP Q28146
B	1344	HIS	-	EXPRESSION TAG	UNP Q28146
B	1345	HIS	-	EXPRESSION TAG	UNP Q28146
B	1346	HIS	-	EXPRESSION TAG	UNP Q28146
B	1347	HIS	-	EXPRESSION TAG	UNP Q28146
B	1348	HIS	-	EXPRESSION TAG	UNP Q28146
B	1349	HIS	-	EXPRESSION TAG	UNP Q28146

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

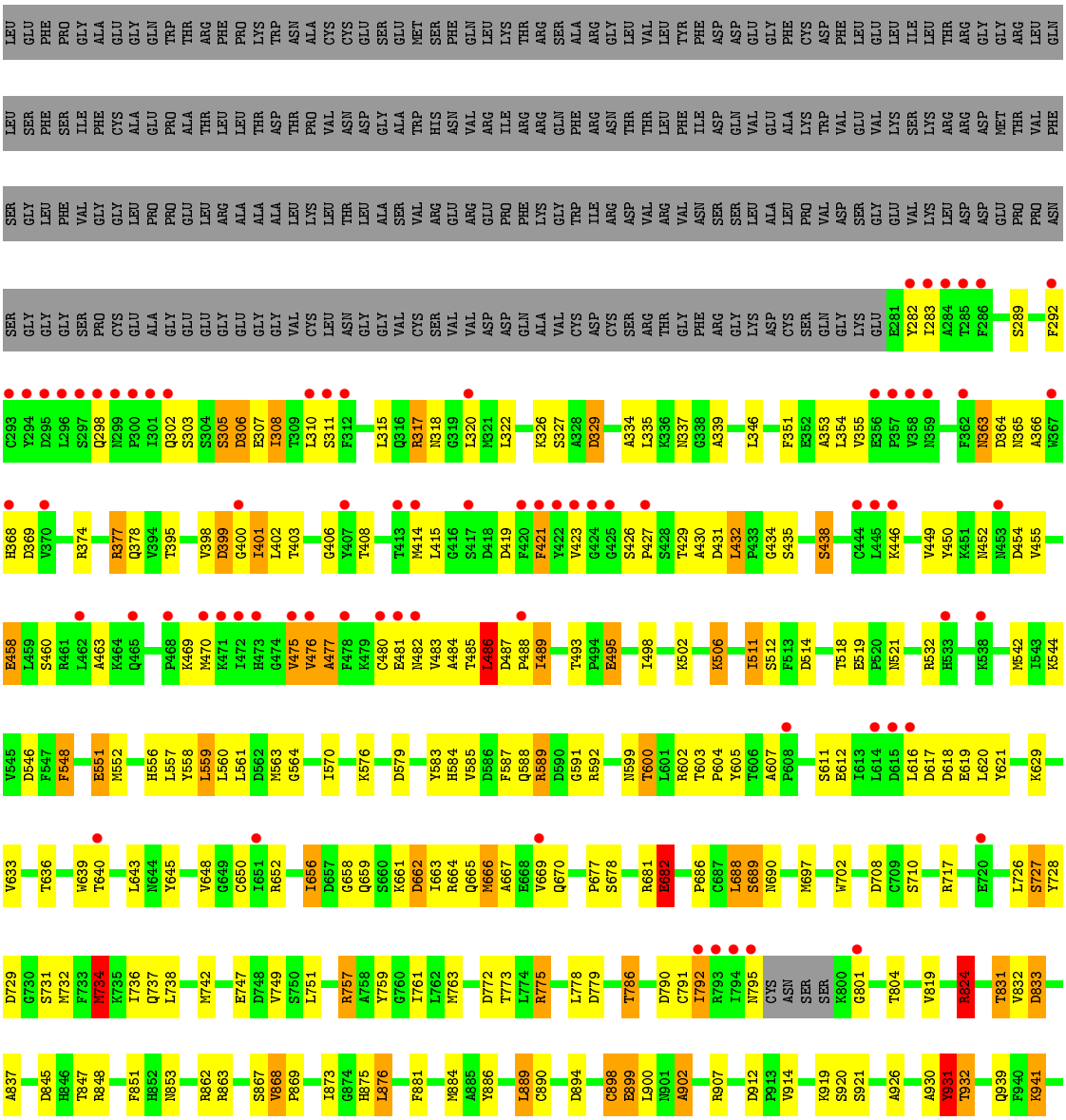
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	2	Total	C	N	O	0	0
			28	16	2	10		
2	B	2	Total	C	N	O	0	0
			28	16	2	10		

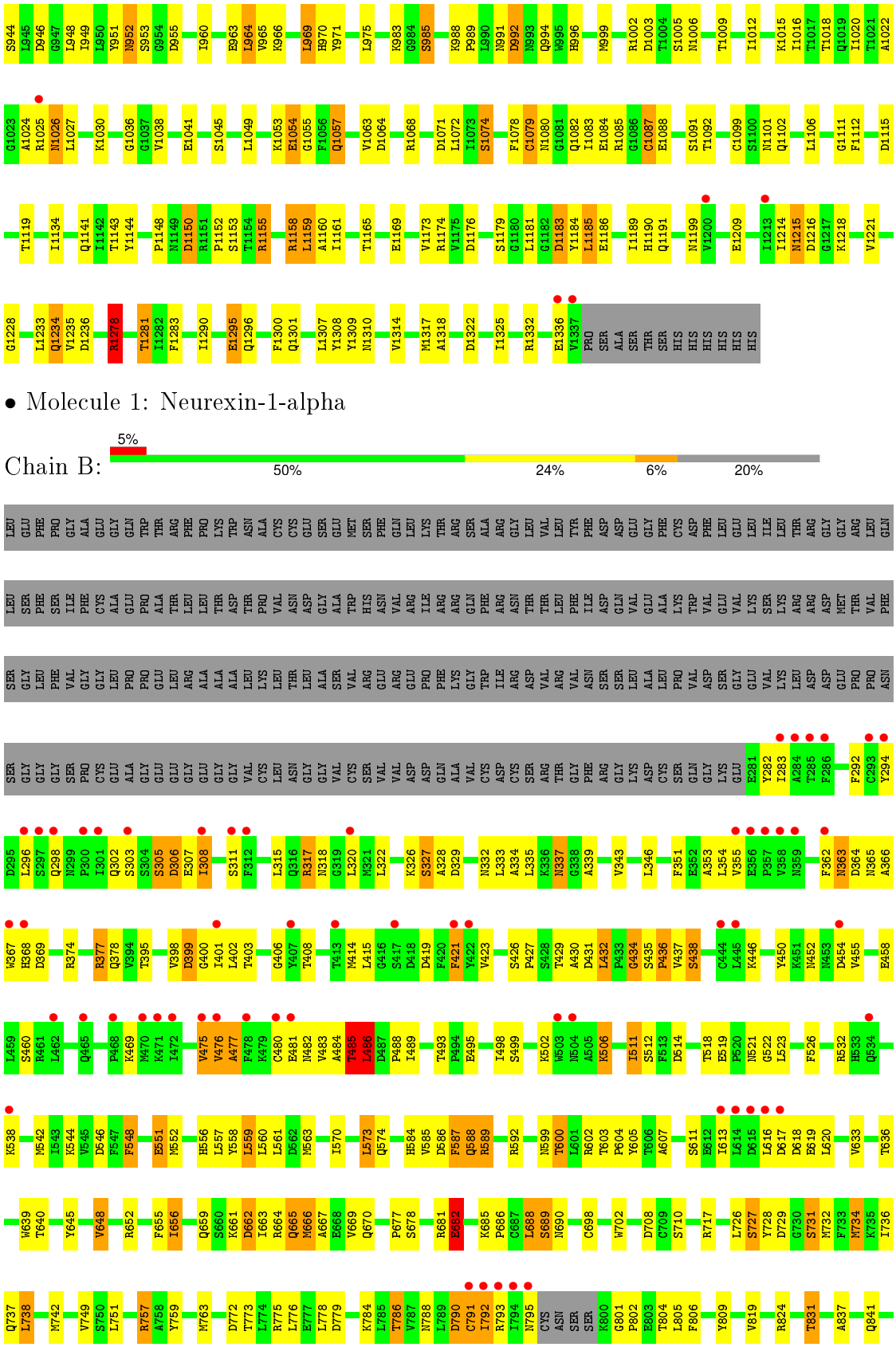
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Neurexin-1-alpha

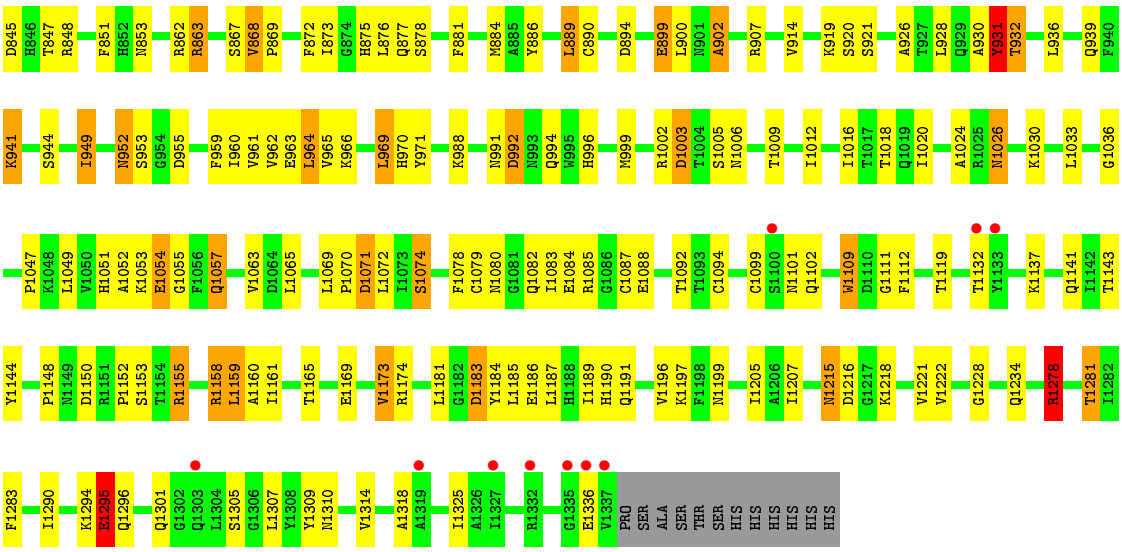
Chain A:





- Molecule 1: Neurexin-1-alpha

Chain B:


D845 H846 T847 R848 F851 H852 N853 R862 R863 S867 V868 P869 F872 I873 G874 H875 L876 Q877 S878 F881 M884 A885 Y886 L889 C890 D894 E899 L900 N901 A902 R907 V914 K919 S920 S921 A926 T927 L928 Q929 A930 Y931 T932 L936 Q939 F940
K941 S944 I949 N952 S953 G954 D955 F959 I960 V961 V962 E963 L964 V965 K966 L969 H970 Y971 K988 N991 D992 N993 Q994 W995 H996 M999 R1002 D1003 T1004 S1005 N1006 T1009 I1012 I1016 T1017 T1018 Q1019 I1020 A1024 R1025 N1026 K1030 L1033 G1036
P1047 K1048 L1049 V1050 H1051 A1052 K1053 E1054 G1055 F1056 Q1057 V1063 D1064 L1065 L1069 P1070 D1071 L1072 I1073 S1074 F1078 C1079 N1080 G1081 Q1082 I1083 E1084 R1085 G1086 C1087 E1088 T1092 T1093 C1094 C1099 S1100 N1101 Q1102 W1109 D1110 G1111 F1112 T1119 T1132 Y1133 K1137 Q1141 I1142 T1143
Y1144 P1148 N1149 D1150 R1151 P1152 S1153 T1154 R1155 R1158 L1159 A1160 I1161 T1165 E1169 V1173 R1174 L1181 G1182 D1183 Y1184 L1185 E1186 L1187 H1188 I1189 H1190 Q1191 V1196 K1197 F1198 N1199 I1205 A1206 I1207 N1215 D1216 G1217 K1218 V1221 V1222 G1228 Q1234 R1278 T1281 I1282
F1283 I1290 K1294 E1295 Q1296 Q1301 G1302 Q1303 L1304 S1305 G1306 L1307 Y1308 Y1309 N1310 V1314 A1318 A1319 I1325 A1326 I1327 R1332 G1335 E1336 V1337 PRO SER ALA SER THR SER HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	61.03Å 114.57Å 160.13Å 89.60° 89.98° 87.96°	Depositor
Resolution (Å)	44.29 – 2.95 44.29 – 2.94	Depositor EDS
% Data completeness (in resolution range)	98.2 (44.29-2.95) 97.6 (44.29-2.94)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.32 (at 2.96Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.208 , 0.235 0.202 , 0.230	Depositor DCC
R_{free} test set	4655 reflections (5.48%)	DCC
Wilson B-factor (Å^2)	62.7	Xtriage
Anisotropy	0.812	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 47.5	EDS
Estimated twinning fraction	0.125 for h,-k,-l 0.004 for -h,k,-l 0.010 for -h,-k,l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	0 of 89621 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	15522	wwPDB-VP
Average B, all atoms (Å^2)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	1.19	32/7886 (0.4%)	1.12	25/10691 (0.2%)
1	B	1.19	29/7886 (0.4%)	1.13	28/10691 (0.3%)
All	All	1.19	61/15772 (0.4%)	1.13	53/21382 (0.2%)

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	3
1	B	0	3
All	All	0	6

The worst 5 of 61 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1057	GLN	CG-CD	14.07	1.83	1.51
1	B	1057	GLN	CG-CD	13.73	1.82	1.51
1	A	898	CYS	CB-SG	10.65	2.00	1.82
1	B	868	VAL	CB-CG2	-9.30	1.33	1.52
1	B	963	GLU	CD-OE2	9.26	1.35	1.25

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1071	ASP	CB-CG-OD1	10.07	127.37	118.30
1	A	824	ARG	NE-CZ-NH1	9.87	125.23	120.30
1	B	1053	LYS	CD-CE-NZ	-9.41	90.05	111.70
1	A	941	LYS	CD-CE-NZ	-8.72	91.65	111.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	824	ARG	NE-CZ-NH2	-8.72	115.94	120.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1215	ASN	Peptide
1	A	484	ALA	Peptide
1	A	747	GLU	Peptide
1	B	484	ALA	Peptide
1	B	790	ASP	Peptide

5.2 Too-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 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	7733	0	7482	253	0
1	B	7733	0	7482	260	0
2	A	28	0	25	0	0
2	B	28	0	25	4	0
All	All	15522	0	15014	503	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 16.

The worst 5 of 503 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:B:1057:GLN:CG	1:B:1057:GLN:CD	1.82	1.46
1:A:1057:GLN:CG	1:A:1057:GLN:CD	1.83	1.44
1:B:519:GLU:HG3	1:B:648:VAL:HG22	1.40	1.02
1:A:988:LYS:HG3	1:B:731:SER:HB3	1.41	1.02
1:A:727:SER:HB2	1:A:875:HIS:CD2	1.94	1.01

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	1004/1254 (80%)	881 (88%)	97 (10%)	26 (3%)	7 30
1	B	1004/1254 (80%)	881 (88%)	96 (10%)	27 (3%)	6 29
All	All	2008/2508 (80%)	1762 (88%)	193 (10%)	53 (3%)	7 30

5 of 53 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	327	SER
1	A	401	ILE
1	A	476	VAL
1	A	477	ALA
1	A	506	LYS

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	839/1068 (79%)	742 (88%)	97 (12%)	7 25
1	B	839/1068 (79%)	747 (89%)	92 (11%)	8 27
All	All	1678/2136 (79%)	1489 (89%)	189 (11%)	7 26

5 of 189 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1176	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	377	ARG
1	B	1155	ARG
1	A	1185	LEU
1	A	1295	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	378	GLN
1	B	452	ASN
1	B	1008	HIS
1	B	365	ASN
1	B	1141	GLN

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

4 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	2000	1,2	14,14,15	0.83	0	15,19,21	2.94	9 (60%)
2	NAG	A	2001	2	14,14,15	0.60	0	15,19,21	1.61	3 (20%)
2	NAG	B	2000	1,2	14,14,15	0.80	0	15,19,21	3.04	8 (53%)
2	NAG	B	2001	2	14,14,15	0.81	1 (7%)	15,19,21	1.50	4 (26%)

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	2000	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	2001	2	-	0/6/23/26	0/1/1/1
2	NAG	B	2000	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	2001	2	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2001	NAG	C1-C2	2.14	1.55	1.52

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	2000	NAG	O7-C7-C8	-4.29	114.19	122.06
2	B	2000	NAG	O7-C7-C8	-3.61	115.43	122.06
2	B	2000	NAG	O4-C4-C3	-3.09	103.38	110.34
2	A	2001	NAG	C1-O5-C5	-2.99	108.45	112.25
2	A	2000	NAG	O4-C4-C3	-2.98	103.63	110.34

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2000	NAG	4	0

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

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^{th} 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	1008/1254 (80%)	0.21	78 (7%) 16 8	35, 91, 151, 196	0
1	B	1008/1254 (80%)	0.22	68 (6%) 21 11	34, 92, 152, 201	0
All	All	2016/2508 (80%)	0.22	146 (7%) 18 10	34, 91, 152, 201	0

The worst 5 of 146 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	472	ILE	6.0
1	A	359	ASN	5.9
1	B	358	VAL	5.7
1	A	480	CYS	5.7
1	B	445	LEU	5.5

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^{th} 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^2)	Q<0.9
2	NAG	B	2000	14/15	0.85	0.18	-0.15	87,101,111,114	0
2	NAG	A	2000	14/15	0.86	0.16	-0.35	89,102,110,114	0
2	NAG	A	2001	14/15	0.84	0.23	-	110,124,129,133	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
2	NAG	B	2001	14/15	0.82	0.33	-	113,126,132,134	0

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.