wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 10:04 AM GMT

PDB ID	:	3KM9
Title	:	Structure of complement C5 in complex with the C-terminal beta-grasp domain of SSL7
Authors	:	Laursen, N.S.; Gordon, N.; Hermans, S.; Lorenz, N.; Jackson, N.; Wines, B.; Spillner, E.; Christensen, J.B.; Jensen, M.; Fredslund, F.; Bjerre, M.; Sottrup-Jensen, L.; Fraser, J.D.; Andersen, G.R.
Deposited on	:	2009-11-10
Resolution	:	4.20 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:

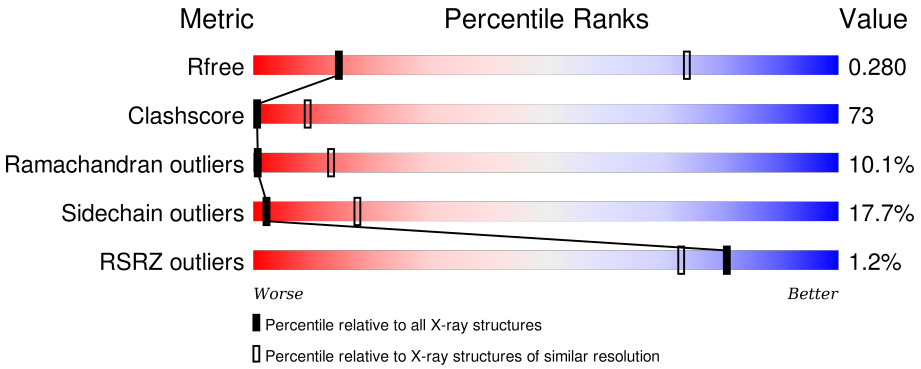
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 4.20 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	1039 (4.80-3.60)
Clashscore	102246	1140 (4.80-3.60)
Ramachandran outliers	100387	1083 (4.80-3.60)
Sidechain outliers	100360	1067 (4.80-3.60)
RSRZ outliers	91569	1042 (4.80-3.60)

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 >=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	1676	% 18% 50% 18% 13%
1	B	1676	% 18% 49% 18% 13%
2	X	103	9% 37% 52% 10%
2	Y	103	5% 38% 52% 9%

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CD	A	1678	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 24809 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 Complement C5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1459	Total	C	N	O	S	0	0	0
			11541	7396	1903	2200	42			
1	B	1459	Total	C	N	O	S	0	0	0
			11541	7396	1903	2200	42			

- Molecule 2 is a protein called Staphylococcal enterotoxin-like toxin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	X	102	Total	C	N	O	S	0	0	0
			819	517	138	163	1			
2	Y	102	Total	C	N	O	S	0	0	0
			819	517	138	163	1			

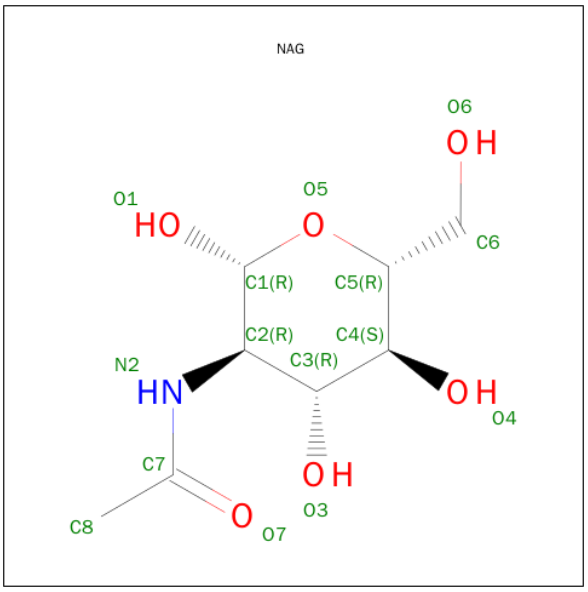
- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Cd	0	0
			2	2		
3	A	3	Total	Cd	0	0
			3	3		

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

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

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



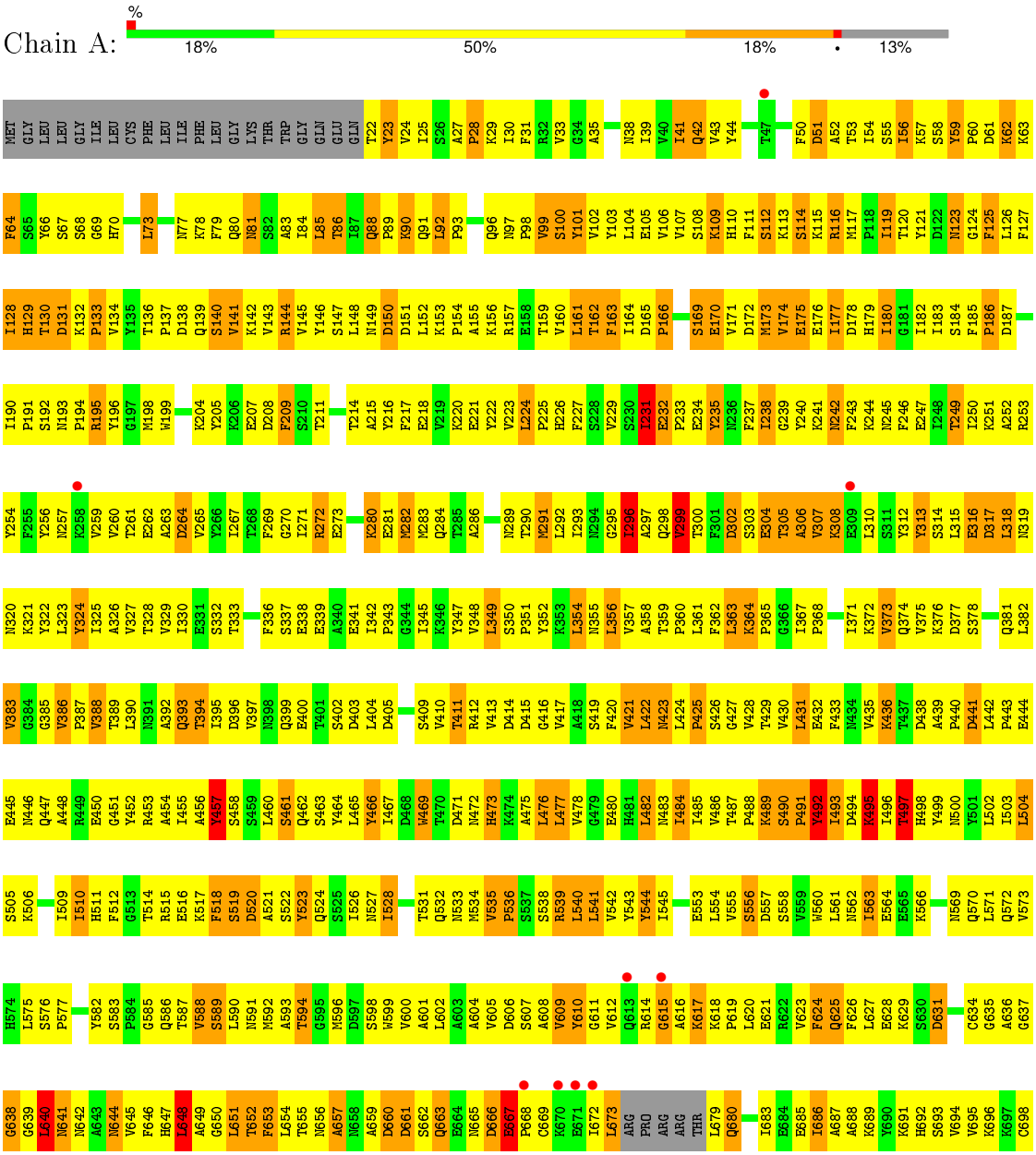
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C N O 14 8 1 5	0	0
5	B	1	Total C N O 14 8 1 5	0	0

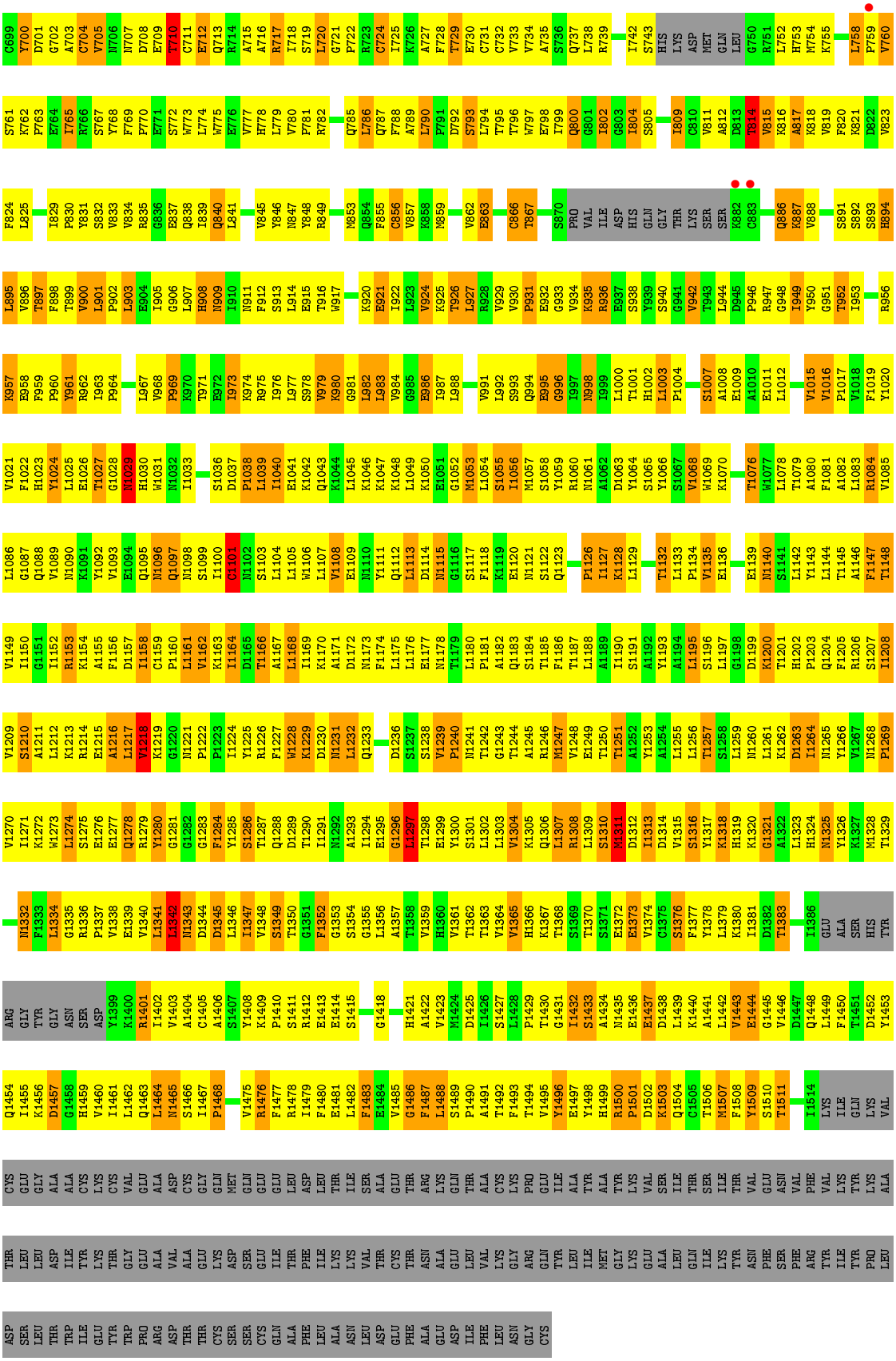
3 Residue-property plots (i)

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: Complement C5

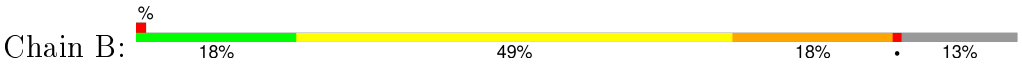
Chain A:

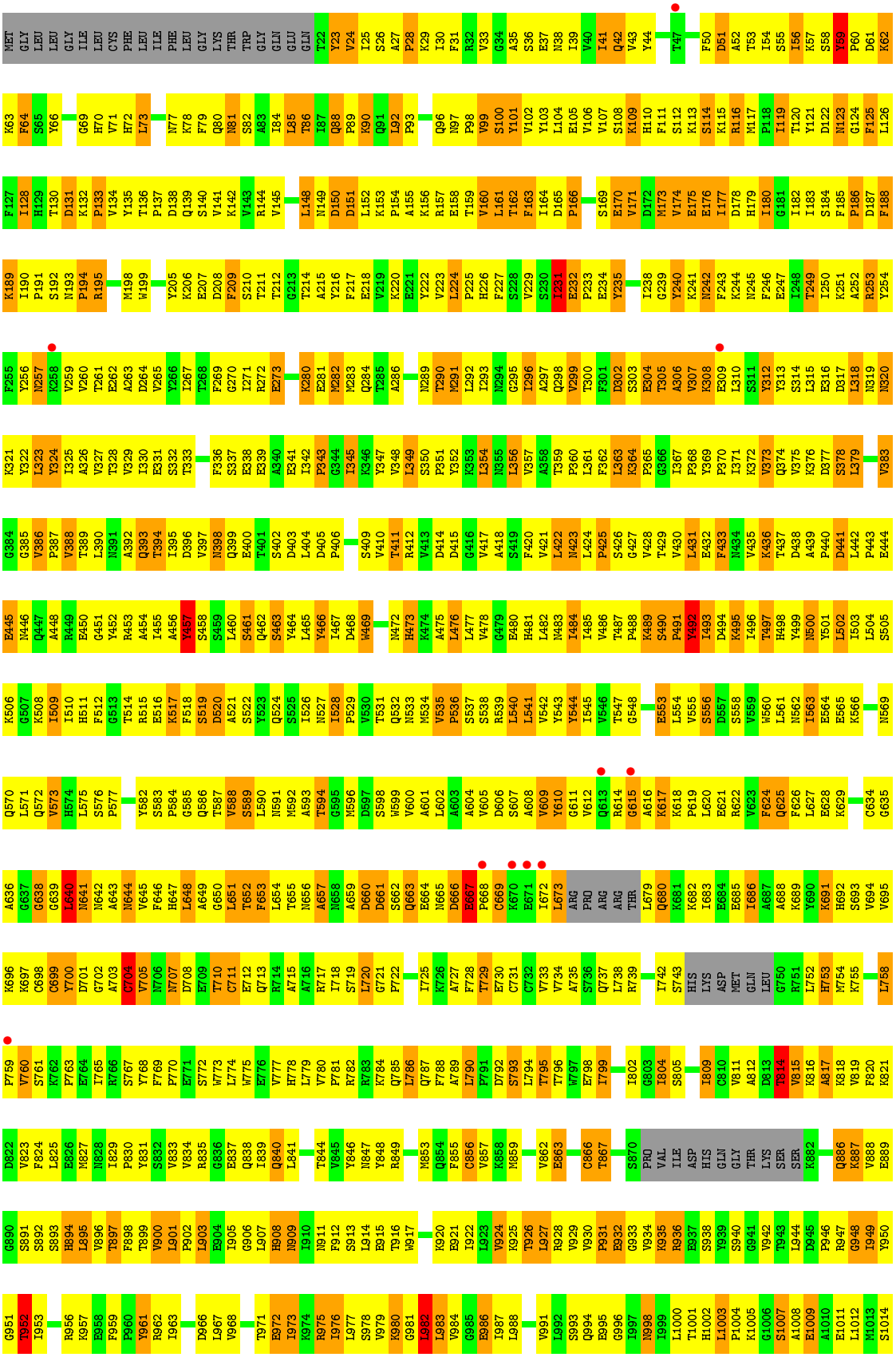




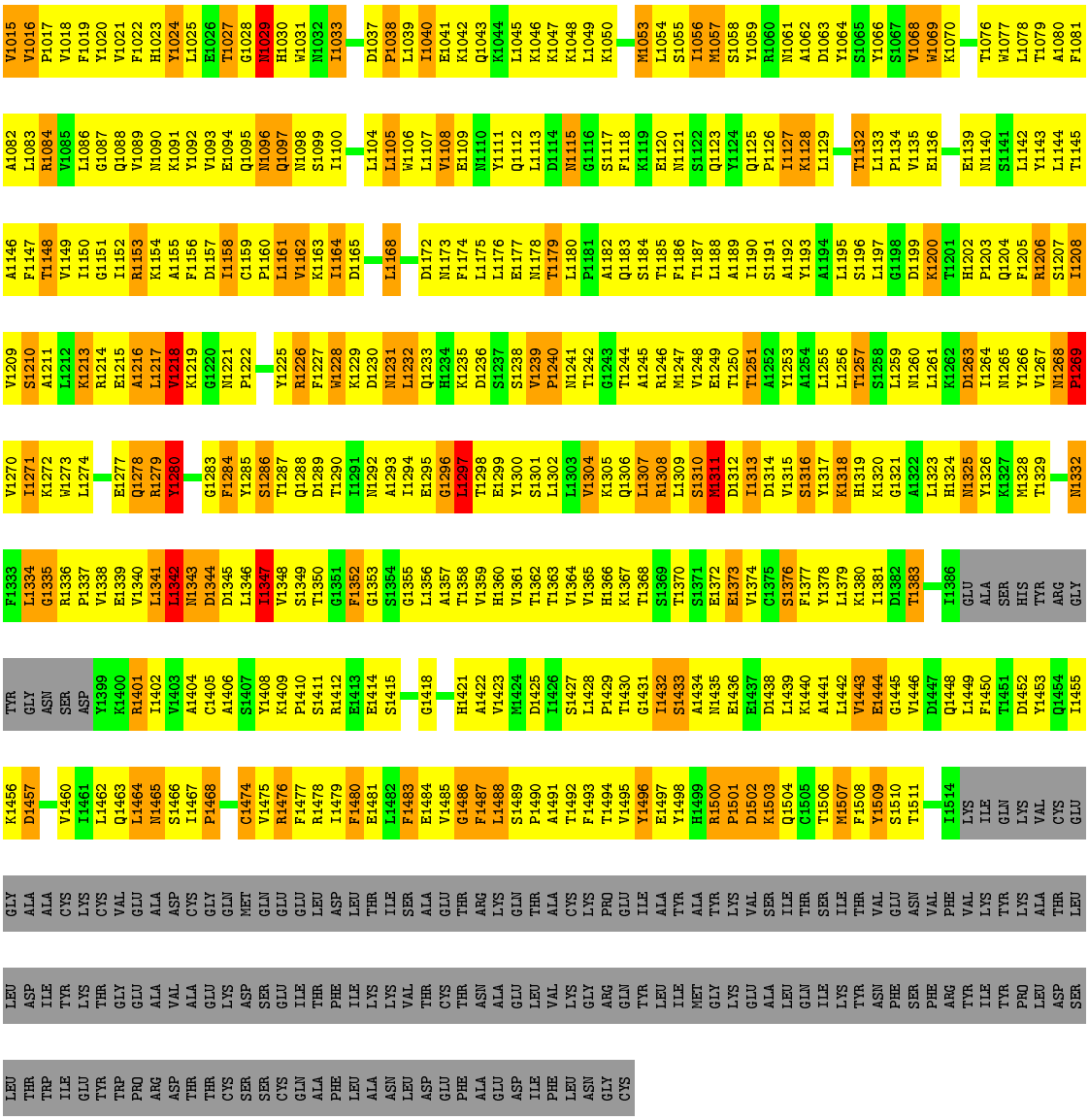
- Molecule 1: Complement C5

Chain B:

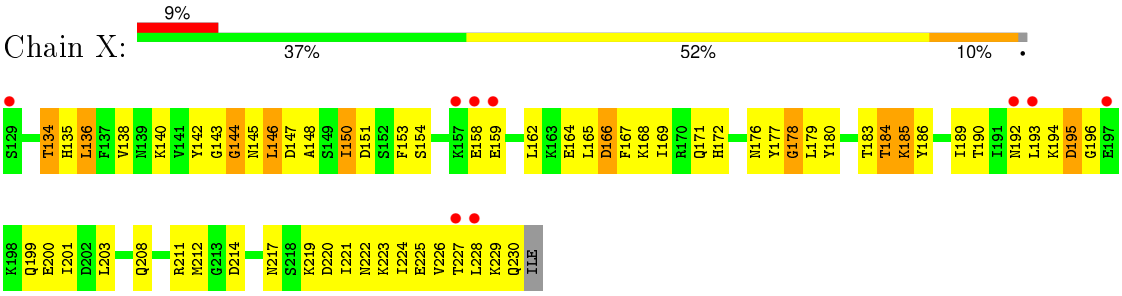




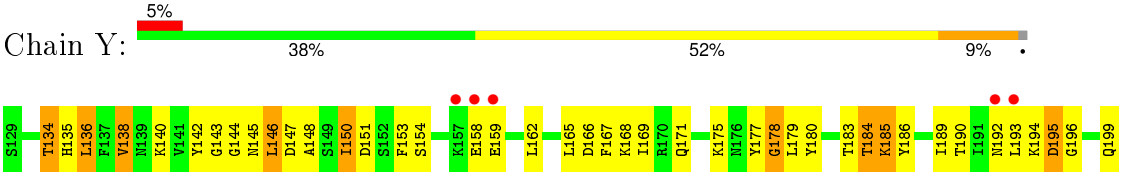
WORLDWIDE PDB PROTEIN DATA BANK

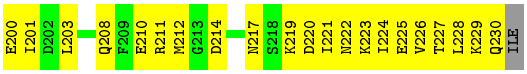


• Molecule 2: Staphylococcal enterotoxin-like toxin



• Molecule 2: Staphylococcal enterotoxin-like toxin




E200
I201
D202
L203
Q208
F209
E210
R211
M212
G213
D214
N217
S218
K219
D220
I221
N222
K223
I224
E225
V226
T227
L228
K229
Q230
TLE

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	144.79Å 144.79Å 245.28Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.75 – 4.20 49.75 – 4.20	Depositor EDS
% Data completeness (in resolution range)	99.6 (49.75-4.20) 99.8 (49.75-4.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	2.76 (at 4.14Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, R_{free}	0.233 , 0.297 0.212 , 0.280	Depositor DCC
R_{free} test set	2030 reflections (4.86%)	DCC
Wilson B-factor (Å^2)	115.5	Xtriage
Anisotropy	0.855	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.19 , 90.3	EDS
Estimated twinning fraction	0.044 for -h,-k,l 0.397 for h,-h-k,-l 0.045 for -k,-h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Outliers	0 of 41960 reflections	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	24809	wwPDB-VP
Average B, all atoms (Å^2)	205.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.16% 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, CD

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	0.53	0/11793	0.77	6/16003 (0.0%)
1	B	0.53	0/11793	0.77	5/16003 (0.0%)
2	X	0.34	0/828	0.54	0/1107
2	Y	0.34	0/828	0.56	1/1107 (0.1%)
All	All	0.52	0/25242	0.75	12/34220 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	2
All	All	0	3

There are no bond length outliers.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	264	ASP	CB-CG-OD1	-8.57	110.58	118.30
1	B	640	LEU	CA-CB-CG	7.23	131.94	115.30
1	B	1374	VAL	CB-CA-C	-6.61	98.85	111.40
1	A	640	LEU	CA-CB-CG	6.51	130.28	115.30
1	A	1195	LEU	CA-CB-CG	-5.71	102.17	115.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	651	LEU	Peptide
1	B	1179	THR	Peptide
1	B	651	LEU	Peptide

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	11541	0	11511	1721	0
1	B	11541	0	11511	1730	0
2	X	819	0	831	85	0
2	Y	819	0	831	83	0
3	A	3	0	0	0	0
3	B	2	0	0	0	0
4	A	28	0	25	2	0
4	B	28	0	25	3	0
5	A	14	0	13	0	0
5	B	14	0	13	0	0
All	All	24809	0	24760	3610	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 73.

The worst 5 of 3610 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:535:VAL:HG23	1:A:536:PRO:HD3	1.21	1.17
1:A:698:CYS:SG	1:A:724:CYS:CB	2.33	1.16
1:A:698:CYS:SG	1:A:724:CYS:HB2	1.86	1.15
1:A:968:VAL:HG12	1:A:1368:THR:HG22	1.19	1.13
1:B:535:VAL:HG23	1:B:536:PRO:HD3	1.29	1.13

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	1449/1676 (86%)	1039 (72%)	255 (18%)	155 (11%)	0 11
1	B	1449/1676 (86%)	1026 (71%)	274 (19%)	149 (10%)	1 12
2	X	100/103 (97%)	86 (86%)	9 (9%)	5 (5%)	3 31
2	Y	100/103 (97%)	84 (84%)	11 (11%)	5 (5%)	3 31
All	All	3098/3558 (87%)	2235 (72%)	549 (18%)	314 (10%)	1 13

5 of 314 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	59	TYR
1	A	86	THR
1	A	97	ASN
1	A	99	VAL
1	A	170	GLU

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	1296/1484 (87%)	1061 (82%)	235 (18%)	2 16
1	B	1296/1484 (87%)	1051 (81%)	245 (19%)	2 14
2	X	93/94 (99%)	87 (94%)	6 (6%)	21 61
2	Y	93/94 (99%)	86 (92%)	7 (8%)	17 56
All	All	2778/3156 (88%)	2285 (82%)	493 (18%)	2 18

5 of 493 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1465	ASN
1	B	184	SER
1	B	1313	ILE
1	A	1496	TYR
1	B	73	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 98 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1463	GLN
1	B	139	GLN
1	B	1366	HIS
1	A	1465	ASN
1	B	77	ASN

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)

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	$\#\|Z\| > 2$
4	NAG	A	2001	1,4	14,14,15	0.50	0	15,19,21	1.07	1 (6%)
4	NAG	A	2002	4	14,14,15	0.46	0	15,19,21	1.09	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	NAG	B	2001	1,4	14,14,15	0.54	0	15,19,21	0.89	0
4	NAG	B	2002	4	14,14,15	0.46	0	15,19,21	1.02	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
4	NAG	A	2001	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	2002	4	-	0/6/23/26	0/1/1/1
4	NAG	B	2001	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	2002	4	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	B	2002	NAG	C1-O5-C5	2.76	115.75	112.25
4	A	2001	NAG	C1-O5-C5	3.13	116.23	112.25
4	A	2002	NAG	C1-O5-C5	3.14	116.23	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	2001	NAG	2	0
4	B	2001	NAG	3	0
4	B	2002	NAG	1	0

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 5 are monoatomic - leaving 2 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	A	1680	1	14,14,15	0.65	0	15,19,21	0.88	1 (6%)
5	NAG	B	1679	1	14,14,15	0.70	1 (7%)	15,19,21	0.97	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	1680	1	-	0/6/23/26	0/1/1/1
5	NAG	B	1679	1	-	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)
5	B	1679	NAG	C1-C2	2.02	1.55	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	1680	NAG	O5-C5-C6	2.02	111.73	107.35

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^{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	1459/1676 (87%)	-0.60	12 (0%) 87 82	81, 190, 311, 455	0
1	B	1459/1676 (87%)	-0.61	10 (0%) 89 84	85, 190, 308, 475	0
2	X	102/103 (99%)	0.08	9 (8%) 12 9	157, 292, 386, 530	0
2	Y	102/103 (99%)	0.07	5 (4%) 33 25	156, 292, 377, 494	0
All	All	3122/3558 (87%)	-0.56	36 (1%) 81 73	81, 194, 328, 530	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	671	GLU	5.0
2	Y	193	LEU	4.8
2	X	193	LEU	4.7
1	B	671	GLU	4.2
1	A	670	LYS	3.7

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
4	NAG	A	2001	14/15	0.85	0.28	-	293,293,293,293	0
4	NAG	A	2002	14/15	0.66	0.46	-	343,343,343,343	0
4	NAG	B	2002	14/15	0.73	0.50	-	363,363,363,363	0
4	NAG	B	2001	14/15	0.79	0.30	-	280,280,280,280	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^{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
3	CD	A	1678	1/1	0.87	0.44	2.30	481,481,481,481	0
3	CD	B	1677	1/1	0.90	0.38	1.04	466,466,466,466	0
5	NAG	A	1680	14/15	0.51	0.36	-	301,301,301,301	0
3	CD	B	1678	1/1	0.39	0.11	-	397,397,397,397	0
3	CD	A	1677	1/1	0.89	0.09	-	229,229,229,229	1
5	NAG	B	1679	14/15	0.59	0.35	-	290,290,290,290	0
3	CD	A	1679	1/1	0.22	0.12	-	402,402,402,402	0

6.5 Other polymers (i)

There are no such residues in this entry.