wwPDB X-ray Structure Validation Summary Report (i)

Jun 22, 2024 – 10:16 AM EDT

PDB ID	:	5E6U
Title	:	Structures of leukocyte integrin aLb2: The aI domain, the headpiece, and the pocket for the internal ligand
Authors	:	Sen, M.; Springer, T.A.
Deposited on	:	2015-10-10
Resolution	:	2.50 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

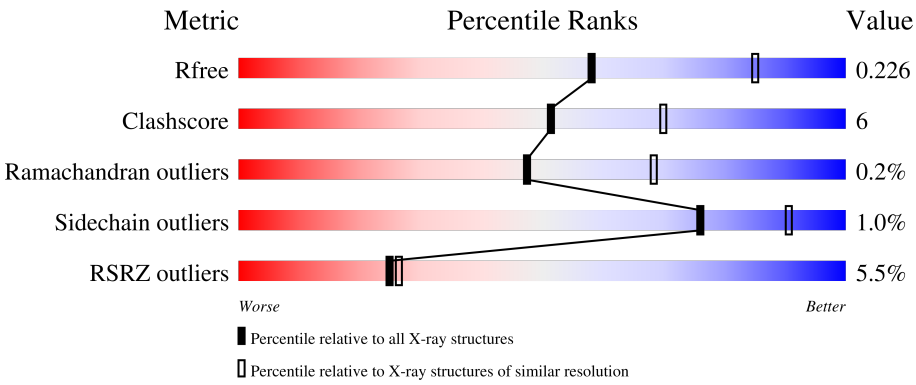
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.50 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}	130704	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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	795	2%, 63%, 11%, 27%
2	B	510	8%, 78%, 11%, 11%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 8784 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 Integrin alpha-L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	583	4525	2862	779	868	16	0	6	0

There are 53 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	189	TRP	ARG	variant	UNP P20701
A	645	ARG	ASN	engineered mutation	UNP P20701
A	701	ARG	ASN	engineered mutation	UNP P20701
A	746	PRO	-	expression tag	UNP P20701
A	747	ALA	-	expression tag	UNP P20701
A	748	ALA	-	expression tag	UNP P20701
A	749	LEU	-	expression tag	UNP P20701
A	750	GLN	-	expression tag	UNP P20701
A	751	THR	-	expression tag	UNP P20701
A	752	LEU	-	expression tag	UNP P20701
A	753	PHE	-	expression tag	UNP P20701
A	754	GLN	-	expression tag	UNP P20701
A	755	GLY	-	expression tag	UNP P20701
A	756	PRO	-	expression tag	UNP P20701
A	757	LEU	-	expression tag	UNP P20701
A	758	GLY	-	expression tag	UNP P20701
A	759	ALA	-	expression tag	UNP P20701
A	760	GLN	-	expression tag	UNP P20701
A	761	GLY	-	expression tag	UNP P20701
A	762	GLU	-	expression tag	UNP P20701
A	763	LYS	-	expression tag	UNP P20701
A	764	GLU	-	expression tag	UNP P20701
A	765	LEU	-	expression tag	UNP P20701
A	766	GLN	-	expression tag	UNP P20701
A	767	ALA	-	expression tag	UNP P20701
A	768	LEU	-	expression tag	UNP P20701
A	769	GLU	-	expression tag	UNP P20701

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	770	LYS	-	expression tag	UNP P20701
A	771	GLU	-	expression tag	UNP P20701
A	772	ASN	-	expression tag	UNP P20701
A	773	ALA	-	expression tag	UNP P20701
A	774	GLN	-	expression tag	UNP P20701
A	775	LEU	-	expression tag	UNP P20701
A	776	GLU	-	expression tag	UNP P20701
A	777	TRP	-	expression tag	UNP P20701
A	778	GLU	-	expression tag	UNP P20701
A	779	LEU	-	expression tag	UNP P20701
A	780	GLN	-	expression tag	UNP P20701
A	781	ALA	-	expression tag	UNP P20701
A	782	LEU	-	expression tag	UNP P20701
A	783	GLU	-	expression tag	UNP P20701
A	784	LYS	-	expression tag	UNP P20701
A	785	GLU	-	expression tag	UNP P20701
A	786	LEU	-	expression tag	UNP P20701
A	787	ALA	-	expression tag	UNP P20701
A	788	GLN	-	expression tag	UNP P20701
A	789	HIS	-	expression tag	UNP P20701
A	790	HIS	-	expression tag	UNP P20701
A	791	HIS	-	expression tag	UNP P20701
A	792	HIS	-	expression tag	UNP P20701
A	793	HIS	-	expression tag	UNP P20701
A	794	HIS	-	expression tag	UNP P20701
A	795	ALA	-	expression tag	UNP P20701

- Molecule 2 is a protein called Integrin beta-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	453	3556	2217	633	677	29	0	6	0

There are 51 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	232	LYS	ASN	engineered mutation	UNP P05107
B	461	PRO	-	expression tag	UNP P05107
B	462	ALA	-	expression tag	UNP P05107
B	463	ALA	-	expression tag	UNP P05107
B	464	LEU	-	expression tag	UNP P05107
B	465	GLN	-	expression tag	UNP P05107

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	466	THR	-	expression tag	UNP P05107
B	467	LEU	-	expression tag	UNP P05107
B	468	PHE	-	expression tag	UNP P05107
B	469	GLN	-	expression tag	UNP P05107
B	470	GLY	-	expression tag	UNP P05107
B	471	PRO	-	expression tag	UNP P05107
B	472	LEU	-	expression tag	UNP P05107
B	473	GLY	-	expression tag	UNP P05107
B	474	ALA	-	expression tag	UNP P05107
B	475	GLN	-	expression tag	UNP P05107
B	476	GLY	-	expression tag	UNP P05107
B	477	LYS	-	expression tag	UNP P05107
B	478	LYS	-	expression tag	UNP P05107
B	479	LYS	-	expression tag	UNP P05107
B	480	LEU	-	expression tag	UNP P05107
B	481	GLN	-	expression tag	UNP P05107
B	482	ALA	-	expression tag	UNP P05107
B	483	LEU	-	expression tag	UNP P05107
B	484	LYS	-	expression tag	UNP P05107
B	485	LYS	-	expression tag	UNP P05107
B	486	LYS	-	expression tag	UNP P05107
B	487	ASN	-	expression tag	UNP P05107
B	488	ALA	-	expression tag	UNP P05107
B	489	GLN	-	expression tag	UNP P05107
B	490	LEU	-	expression tag	UNP P05107
B	491	LYS	-	expression tag	UNP P05107
B	492	TRP	-	expression tag	UNP P05107
B	493	LYS	-	expression tag	UNP P05107
B	494	LEU	-	expression tag	UNP P05107
B	495	GLN	-	expression tag	UNP P05107
B	496	ALA	-	expression tag	UNP P05107
B	497	LEU	-	expression tag	UNP P05107
B	498	LYS	-	expression tag	UNP P05107
B	499	LYS	-	expression tag	UNP P05107
B	500	LYS	-	expression tag	UNP P05107
B	501	LEU	-	expression tag	UNP P05107
B	502	ALA	-	expression tag	UNP P05107
B	503	GLN	-	expression tag	UNP P05107
B	504	HIS	-	expression tag	UNP P05107
B	505	HIS	-	expression tag	UNP P05107
B	506	HIS	-	expression tag	UNP P05107
B	507	HIS	-	expression tag	UNP P05107

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	508	HIS	-	expression tag	UNP P05107
B	509	HIS	-	expression tag	UNP P05107
B	510	ALA	-	expression tag	UNP P05107

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Mg 1 1	0	0
3	B	1	Total Mg 1 1	0	0

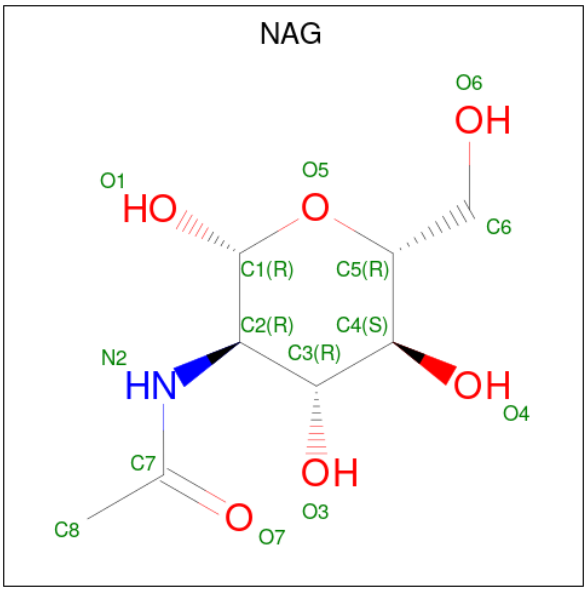
- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	2	Total Ca 2 2	0	0
4	B	2	Total Ca 2 2	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Cl 1 1	0	0

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 7 is water.

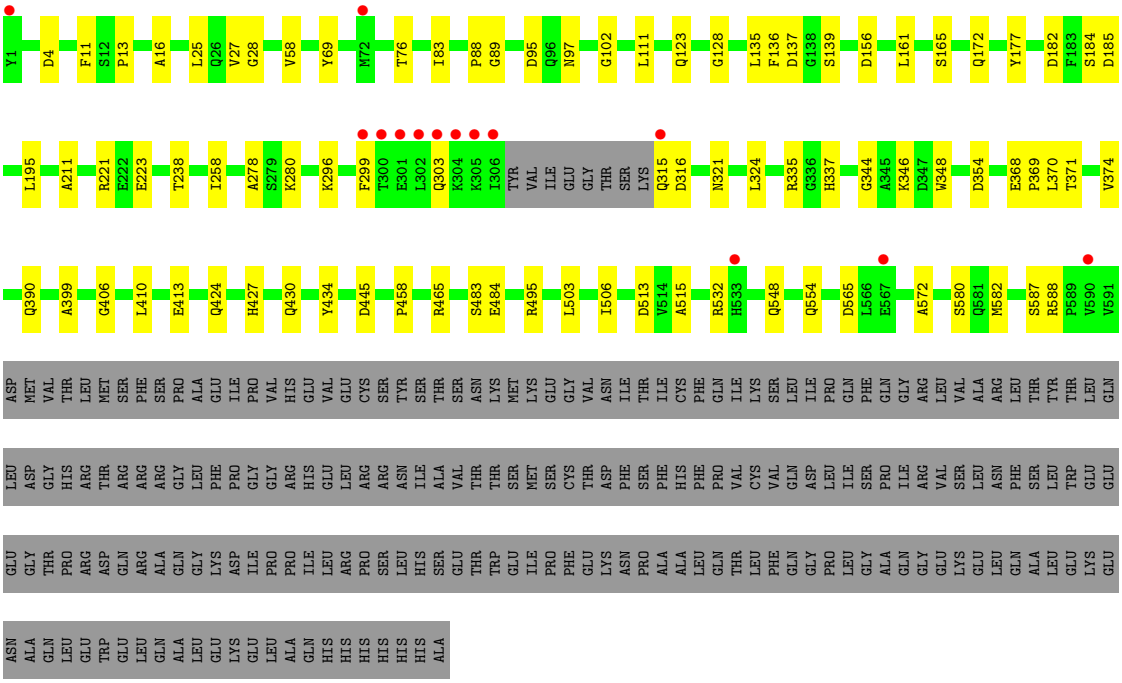
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	403	403	403	0	0
			Total	O		
7	B	251	251	251	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

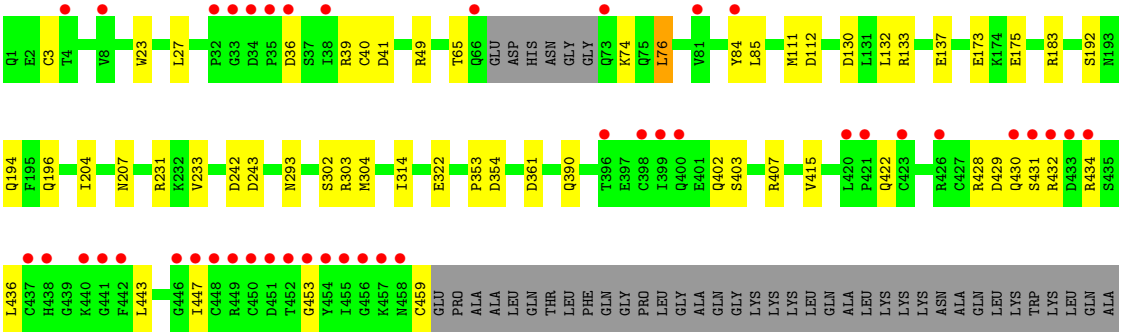
• Molecule 1: Integrin alpha-L

Chain A: 2% 63% 11% 27%



• Molecule 2: Integrin beta-2

Chain B: 8% 78% 11% 11%



LEU
LYS
LYS
LYS
LEU
ALA
GLN
HIS
HIS
HIS
HIS
HIS
HIS
HIS
ALA

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	154.98Å 154.98Å 115.45Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.44 – 2.50 46.44 – 2.50	Depositor EDS
% Data completeness (in resolution range)	99.5 (46.44-2.50) 99.5 (46.44-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.37 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.179 , 0.226 0.179 , 0.226	Depositor DCC
R_{free} test set	1645 reflections (3.03%)	wwPDB-VP
Wilson B-factor (Å^2)	40.1	Xtriage
Anisotropy	0.140	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 47.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.038 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8784	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.26% 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, 0.333 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: MG, CL, NAG, CA

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	0.30	0/4630	0.52	0/6264
2	B	0.29	0/3637	0.51	0/4919
All	All	0.30	0/8267	0.52	0/11183

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
2	B	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	242	ASP	Peptide
2	B	453	GLY	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	4525	0	4414	51	0
2	B	3556	0	3488	40	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
5	A	1	0	0	1	0
6	A	14	0	13	0	0
6	B	28	0	26	0	0
7	A	403	0	0	13	0
7	B	251	0	0	11	0
All	All	8784	0	7941	90	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 90 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)
2:B:173:GLU:OE2	7:B:4001:HOH:O	2.02	0.76
1:A:337[A]:HIS:NE2	1:A:354:ASP:OD1	2.21	0.70
2:B:3:CYS:SG	7:B:4186:HOH:O	2.49	0.69
2:B:361:ASP:HB2	2:B:390:GLN:HB3	1.73	0.69
1:A:4:ASP:OD2	1:A:588:ARG:NH1	2.27	0.68

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	585/795 (74%)	564 (96%)	19 (3%)	2 (0%)	41 61
2	B	455/510 (89%)	436 (96%)	19 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1040/1305 (80%)	1000 (96%)	38 (4%)	2 (0%)	47 68

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	28	GLY
1	A	316	ASP

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	482/666 (72%)	479 (99%)	3 (1%)	86 95
2	B	398/440 (90%)	392 (98%)	6 (2%)	65 85
All	All	880/1106 (80%)	871 (99%)	9 (1%)	76 90

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	403	SER
2	B	459	CYS
2	B	65	THR
2	B	76	LEU
2	B	175	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 7 are monoatomic - leaving 3 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	$\#\|Z\| > 2$
6	NAG	B	3094	2	14,14,15	0.51	0	17,19,21	0.46	0
6	NAG	B	3190	2	14,14,15	0.29	0	17,19,21	0.60	0
6	NAG	A	3064	1	14,14,15	0.34	0	17,19,21	0.47	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
6	NAG	B	3094	2	-	2/6/23/26	0/1/1/1
6	NAG	B	3190	2	-	3/6/23/26	0/1/1/1
6	NAG	A	3064	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	3064	NAG	C4-C5-C6-O6
6	B	3094	NAG	O5-C5-C6-O6
6	A	3064	NAG	O5-C5-C6-O6
6	B	3094	NAG	C4-C5-C6-O6
6	B	3190	NAG	C4-C5-C6-O6

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 (i)

6.1 Protein, DNA and RNA chains (i)

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	583/795 (73%)	-0.19	14 (2%)	59	62	27, 39, 82, 222	0
2	B	453/510 (88%)	0.17	43 (9%)	8	8	26, 49, 148, 350	0
All	All	1036/1305 (79%)	-0.03	57 (5%)	25	26	26, 44, 114, 350	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	431	SER	10.2
2	B	432	ARG	9.2
2	B	455	ILE	7.2
2	B	454	TYR	7.2
2	B	457	LYS	6.9

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 monosaccharides 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. 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	B-factors(Å^2)	Q<0.9
6	NAG	A	3064	14/15	0.83	0.34	114,126,130,130	0
6	NAG	B	3190	14/15	0.87	0.23	93,101,104,105	0
4	CA	A	2002	1/1	0.93	0.07	57,57,57,57	0
6	NAG	B	3094	14/15	0.95	0.14	49,63,71,79	0
4	CA	A	2003	1/1	0.96	0.06	49,49,49,49	0
3	MG	A	2001	1/1	0.96	0.03	25,25,25,25	1
4	CA	B	2002	1/1	0.97	0.09	56,56,56,56	0
3	MG	B	2001	1/1	0.97	0.06	27,27,27,27	1
5	CL	A	2004	1/1	0.98	0.07	61,61,61,61	0
4	CA	B	2003	1/1	1.00	0.16	34,34,34,34	0

6.5 Other polymers (i)

There are no such residues in this entry.