Full wwPDB X-ray Structure Validation Report ⓘ

Feb 16, 2023 – 02:23 PM EST

PDB ID	:	8EIL
Title	:	C-Terminal Domain of BrxL from Acinetobacter BREX type I phage restriction system
Authors	:	Doyle, L.A.; Stoddard, B.L.; Kaiser, B.
Deposited on	:	2022-09-15
Resolution	:	2.25 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ 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 ⓘ) were used in the production of this report:

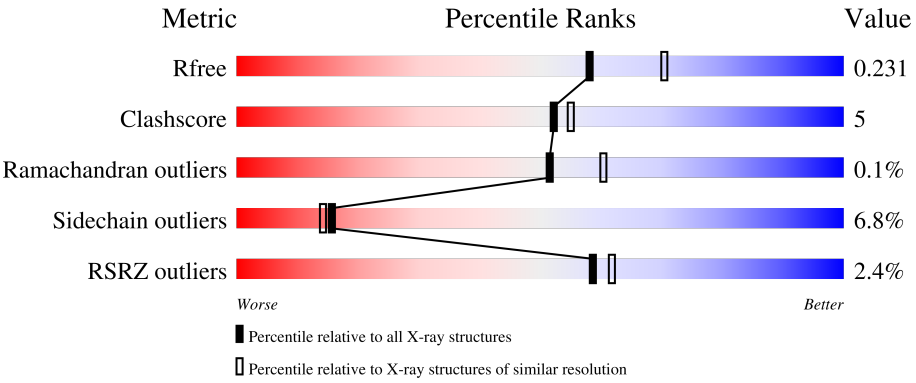
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.32.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.1

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 2.25 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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	186	80% 17%
1	B	186	82% 13%
1	C	186	3% 78% 16%
1	D	186	6% 82% 12% 5%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5208 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.

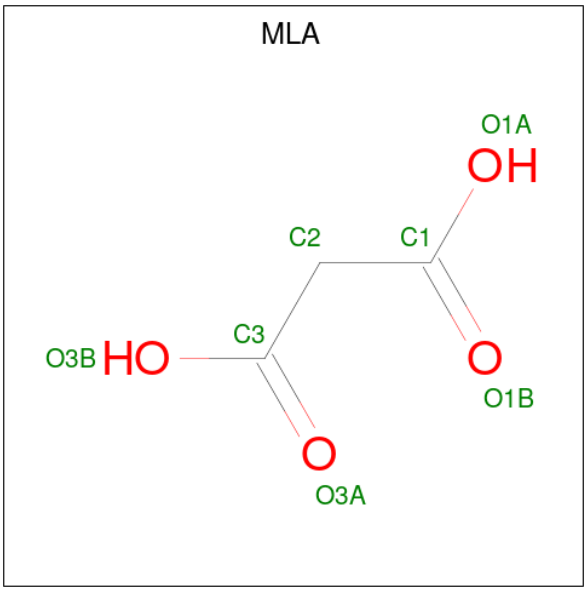
- Molecule 1 is a protein called Protease Lon-related BREX system protein BrxL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	181	Total	C	N	O	S	0	0	0
			1290	833	208	243	6			
1	B	180	Total	C	N	O	S	0	0	0
			1300	838	209	246	7			
1	C	178	Total	C	N	O	S	0	0	0
			1243	799	206	233	5			
1	D	177	Total	C	N	O	S	0	0	0
			1233	792	205	231	5			

There are 16 discrepancies between the modelled and reference sequences:

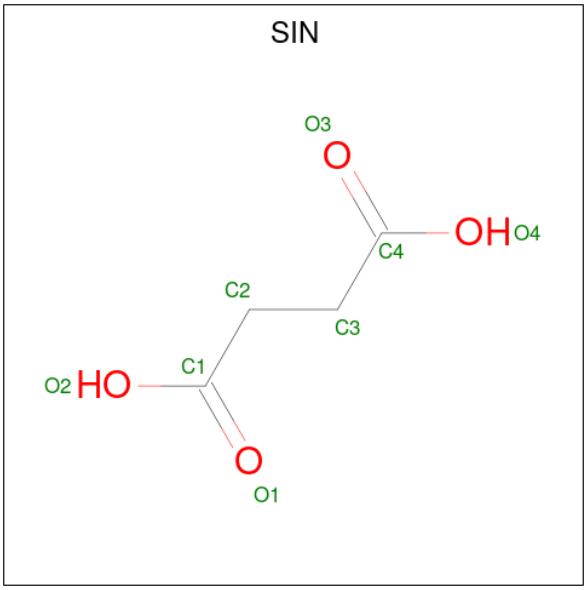
Chain	Residue	Modelled	Actual	Comment	Reference
A	494	GLY	-	expression tag	UNP A0A3R9EDI8
A	495	SER	-	expression tag	UNP A0A3R9EDI8
A	496	HIS	-	expression tag	UNP A0A3R9EDI8
A	497	MET	-	expression tag	UNP A0A3R9EDI8
B	494	GLY	-	expression tag	UNP A0A3R9EDI8
B	495	SER	-	expression tag	UNP A0A3R9EDI8
B	496	HIS	-	expression tag	UNP A0A3R9EDI8
B	497	MET	-	expression tag	UNP A0A3R9EDI8
C	494	GLY	-	expression tag	UNP A0A3R9EDI8
C	495	SER	-	expression tag	UNP A0A3R9EDI8
C	496	HIS	-	expression tag	UNP A0A3R9EDI8
C	497	MET	-	expression tag	UNP A0A3R9EDI8
D	494	GLY	-	expression tag	UNP A0A3R9EDI8
D	495	SER	-	expression tag	UNP A0A3R9EDI8
D	496	HIS	-	expression tag	UNP A0A3R9EDI8
D	497	MET	-	expression tag	UNP A0A3R9EDI8

- Molecule 2 is MALONIC ACID (three-letter code: MLA) (formula: $C_3H_4O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 7 3 4	0	0
2	B	1	Total C O 7 3 4	0	0

- Molecule 3 is SUCCINIC ACID (three-letter code: SIN) (formula: $C_4H_6O_4$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	1	Total C O 8 4 4	0	0

- Molecule 4 is water.

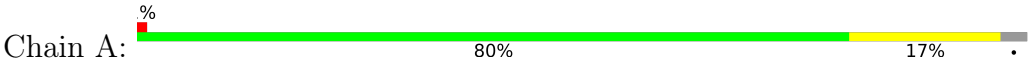
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	40	Total O 40 40	0	0
4	B	35	Total O 35 35	0	0
4	C	24	Total O 24 24	0	0
4	D	21	Total O 21 21	0	0

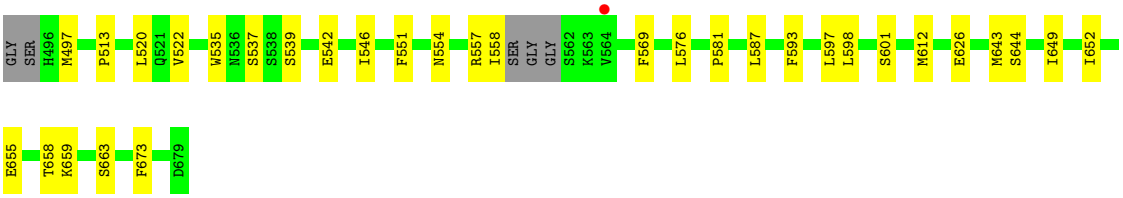
3 Residue-property plots (i)

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: Protease Lon-related BREX system protein BrxL

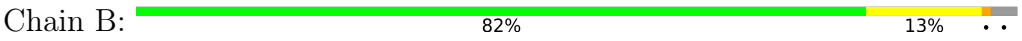
Chain A:





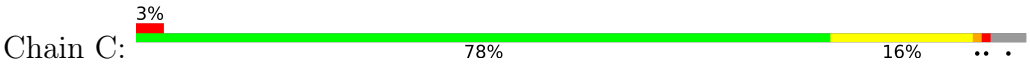
• Molecule 1: Protease Lon-related BREX system protein BrxL

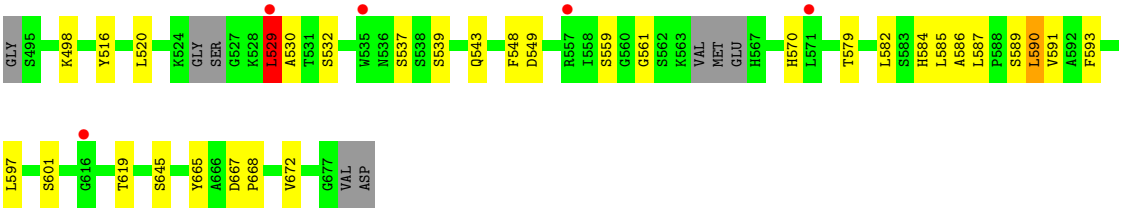
Chain B:



• Molecule 1: Protease Lon-related BREX system protein BrxL

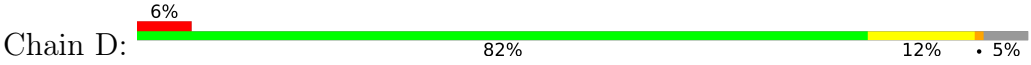
Chain C:

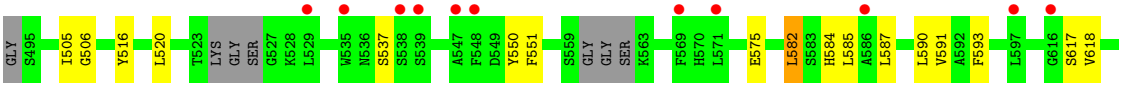




• Molecule 1: Protease Lon-related BREX system protein BrxL

Chain D:






V630
M643
S644
S645
Y665
A666
D667
P668
V672
G677
VAL
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	F 2 3	Depositor
Cell constants a, b, c, α, β, γ	213.34Å 213.34Å 213.34Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.94 – 2.25 48.94 – 2.25	Depositor EDS
% Data completeness (in resolution range)	92.6 (48.94-2.25) 92.1 (48.94-2.25)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.13 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.18_3855	Depositor
R, R_{free}	0.216 , 0.230 0.216 , 0.231	Depositor DCC
R_{free} test set	2024 reflections (5.33%)	wwPDB-VP
Wilson B-factor (Å^2)	45.3	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 54.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.499 for k,h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5208	wwPDB-VP
Average B, all atoms (Å^2)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% 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

5.1 Standard geometry

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

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	0.27	0/1317	0.47	0/1794
1	B	0.27	0/1327	0.48	0/1807
1	C	0.27	0/1270	0.51	1/1731 (0.1%)
1	D	0.27	0/1260	0.46	0/1719
All	All	0.27	0/5174	0.48	1/7051 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	529	LEU	CA-CB-CG	5.83	128.72	115.30

There are no chirality outliers.

There are no planarity outliers.

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	1290	0	1226	15	0
1	B	1300	0	1252	13	0
1	C	1243	0	1144	14	0
1	D	1233	0	1119	14	0
2	A	7	0	2	0	0
2	B	7	0	2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	8	0	4	0	0
4	A	40	0	0	0	0
4	B	35	0	0	1	0
4	C	24	0	0	0	0
4	D	21	0	0	0	0
All	All	5208	0	4749	54	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 5.

All (54) 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:D:520:LEU:HD22	1:D:593:PHE:HA	1.67	0.74
1:C:520:LEU:HD22	1:C:593:PHE:HA	1.70	0.73
1:D:516:TYR:HD1	1:D:575:GLU:HA	1.59	0.67
1:C:516:TYR:OH	1:C:579:THR:OG1	2.14	0.65
1:D:582:LEU:HD22	1:D:584:HIS:H	1.62	0.65
1:A:497:MET:O	1:A:601:SER:OG	2.15	0.64
1:A:520:LEU:HD11	1:A:569:PHE:HB3	1.87	0.57
1:B:600:ARG:HB2	1:B:676:LEU:HD22	1.86	0.55
1:B:505:ILE:HD12	1:B:607:VAL:HG11	1.88	0.55
1:A:520:LEU:HD22	1:A:593:PHE:HA	1.90	0.52
1:C:529:LEU:HD23	1:C:548:PHE:CG	2.45	0.52
1:C:585:LEU:O	1:C:589:SER:OG	2.22	0.52
1:B:649:ILE:HA	1:B:652:ILE:HD12	1.92	0.51
1:C:582:LEU:HD23	1:C:585:LEU:HG	1.92	0.51
1:A:649:ILE:HA	1:A:652:ILE:HD12	1.93	0.50
1:B:520:LEU:HD11	1:B:569:PHE:HB3	1.93	0.50
1:C:587:LEU:HD21	1:C:665:TYR:CE1	2.46	0.50
1:B:557:ARG:O	1:B:558:ILE:HG13	2.11	0.50
1:D:506:GLY:HA3	1:D:582:LEU:HD21	1.95	0.49
1:D:587:LEU:HD21	1:D:665:TYR:CE2	2.49	0.48
1:D:591:VAL:HA	1:D:672:VAL:HG22	1.96	0.48
1:A:554:ASN:HB3	1:A:557:ARG:HD3	1.94	0.48
1:B:513:PRO:HG3	1:B:626:GLU:HB2	1.96	0.48
1:B:497:MET:O	1:B:601:SER:HB3	2.13	0.48
1:C:543:GLN:NE2	1:C:586:ALA:H	2.12	0.48
1:C:520:LEU:HD13	1:C:593:PHE:CD1	2.50	0.47
1:C:530:ALA:HB3	1:C:570:HIS:HD2	1.80	0.47
1:D:516:TYR:CE2	1:D:582:LEU:HG	2.50	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:590:LEU:HD22	1:C:668:PRO:HB3	1.96	0.46
1:A:542:GLU:O	1:A:546:ILE:HG13	2.16	0.46
1:D:516:TYR:CD1	1:D:575:GLU:HA	2.46	0.45
1:A:535:TRP:CE2	1:A:581:PRO:HA	2.52	0.45
1:A:513:PRO:HG3	1:A:626:GLU:HB3	1.99	0.44
1:A:658:THR:HG21	1:D:643:MET:HG3	1.99	0.44
1:C:520:LEU:HD13	1:C:593:PHE:CE1	2.52	0.44
1:D:550:TYR:CZ	1:D:668:PRO:HG2	2.53	0.44
1:A:558:ILE:HG22	1:A:598:LEU:HD11	2.00	0.44
1:B:520:LEU:HD22	1:B:593:PHE:HA	2.00	0.44
1:A:576:LEU:HD23	1:A:576:LEU:HA	1.86	0.44
1:B:535:TRP:CE2	1:B:581:PRO:HA	2.53	0.43
1:A:551:PHE:CZ	1:A:597:LEU:HD12	2.53	0.43
1:D:520:LEU:HD13	1:D:593:PHE:CE1	2.52	0.43
1:B:641:LEU:O	1:B:664:PHE:HA	2.18	0.43
1:B:543:GLN:NE2	4:B:801:HOH:O	2.42	0.43
1:B:576:LEU:HD23	1:B:576:LEU:HA	1.90	0.43
1:C:591:VAL:HA	1:C:672:VAL:HG22	1.99	0.43
1:D:505:ILE:HD13	1:D:630:VAL:HG12	2.00	0.43
1:A:655:GLU:O	1:A:659:LYS:HG2	2.19	0.42
1:A:522:VAL:HG22	1:A:569:PHE:CD1	2.54	0.42
1:A:673:PHE:CD2	1:C:498:LYS:HE2	2.54	0.42
1:B:522:VAL:HG22	1:B:569:PHE:CD1	2.55	0.41
1:C:582:LEU:HG	1:C:584:HIS:H	1.86	0.41
1:D:551:PHE:CD2	1:D:593:PHE:HB3	2.56	0.41
1:D:618:VAL:HG13	1:D:668:PRO:HD3	2.04	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	177/186 (95%)	176 (99%)	1 (1%)	0	100 100
1	B	176/186 (95%)	173 (98%)	3 (2%)	0	100 100
1	C	172/186 (92%)	167 (97%)	4 (2%)	1 (1%)	25 25
1	D	171/186 (92%)	168 (98%)	3 (2%)	0	100 100
All	All	696/744 (94%)	684 (98%)	11 (2%)	1 (0%)	51 60

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	561	GLY

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	126/150 (84%)	119 (94%)	7 (6%)	21 21
1	B	132/150 (88%)	125 (95%)	7 (5%)	22 23
1	C	116/150 (77%)	104 (90%)	12 (10%)	7 5
1	D	113/150 (75%)	106 (94%)	7 (6%)	18 17
All	All	487/600 (81%)	454 (93%)	33 (7%)	16 14

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	537	SER
1	A	539	SER
1	A	587	LEU
1	A	612	MET
1	A	643	MET
1	A	644	SER
1	A	663	SER
1	B	497	MET
1	B	537	SER
1	B	587	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	601	SER
1	B	617	SER
1	B	623	SER
1	B	643	MET
1	C	529	LEU
1	C	532	SER
1	C	537	SER
1	C	539	SER
1	C	549	ASP
1	C	559	SER
1	C	590	LEU
1	C	597	LEU
1	C	601	SER
1	C	619	THR
1	C	645	SER
1	C	667	ASP
1	D	537	SER
1	D	582	LEU
1	D	585	LEU
1	D	590	LEU
1	D	617	SER
1	D	645	SER
1	D	667	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	543	GLN
1	C	570	HIS

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)

3 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	$\#\|Z\| > 2$
2	MLA	B	701	-	6,6,6	1.40	0	7,7,7	1.12	0
2	MLA	A	701	-	6,6,6	1.35	0	7,7,7	1.28	0
3	SIN	D	701	-	7,7,7	1.01	0	8,8,8	1.78	2 (25%)

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	MLA	B	701	-	-	2/4/4/4	-
2	MLA	A	701	-	-	2/4/4/4	-
3	SIN	D	701	-	-	5/5/5/5	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	701	SIN	C2-C3-C4	-2.49	108.25	113.60
3	D	701	SIN	C3-C2-C1	-2.22	108.82	113.60

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	D	701	SIN	C1-C2-C3-C4
2	B	701	MLA	O1A-C1-C2-C3
2	B	701	MLA	O1B-C1-C2-C3
2	A	701	MLA	O1B-C1-C2-C3
2	A	701	MLA	O1A-C1-C2-C3
3	D	701	SIN	C2-C3-C4-O4
3	D	701	SIN	C2-C3-C4-O3
3	D	701	SIN	O2-C1-C2-C3
3	D	701	SIN	O1-C1-C2-C3

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	181/186 (97%)	-0.11	1 (0%) 89 89	32, 53, 82, 98	0
1	B	180/186 (96%)	-0.15	0 100 100	34, 53, 84, 95	0
1	C	178/186 (95%)	0.12	5 (2%) 53 55	34, 71, 101, 115	0
1	D	177/186 (95%)	0.17	11 (6%) 20 22	35, 71, 98, 121	0
All	All	716/744 (96%)	0.00	17 (2%) 59 62	32, 61, 95, 121	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	564	VAL	5.2
1	D	616	GLY	3.9
1	C	616	GLY	3.8
1	C	571	LEU	3.7
1	D	535	TRP	3.5
1	D	547	ALA	3.4
1	D	529	LEU	3.3
1	C	529	LEU	3.2
1	C	535	TRP	3.1
1	D	569	PHE	3.0
1	D	538	SER	2.7
1	D	586	ALA	2.7
1	D	548	PHE	2.6
1	C	557	ARG	2.5
1	D	597	LEU	2.4
1	D	571	LEU	2.2
1	D	539	SER	2.1

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 (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, 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	B-factors(Å^2)	Q<0.9
2	MLA	A	701	7/7	0.89	0.12	57,62,73,74	0
2	MLA	B	701	7/7	0.91	0.11	59,61,71,72	0
3	SIN	D	701	8/8	0.93	0.14	55,67,68,71	0

6.5 Other polymers (i)

There are no such residues in this entry.