



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 3, 2016 – 06:03 PM EDT

PDB ID : 4XHQ
Title : Re-refinement the crystal structure of Dscam1 isoform 1.34, N-terminal four Ig domains
Authors : Chen, Q.; Yu, Y.
Deposited on : 2015-01-06
Resolution : 1.95 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20027939
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) : rb-20027939

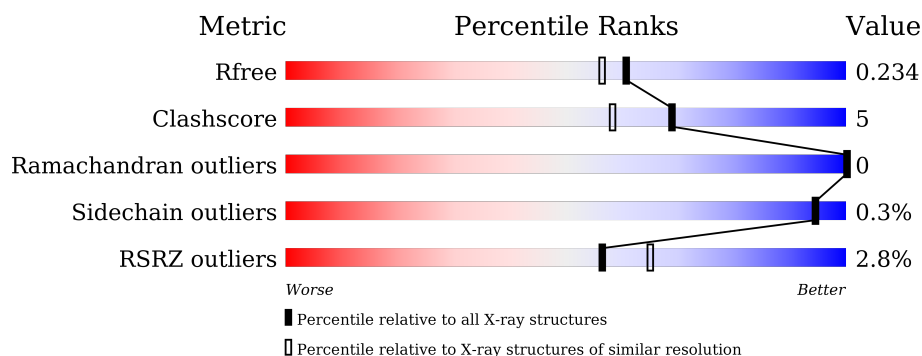
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.95 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	2910 (1.96-1.92)
Clashscore	102246	3095 (1.96-1.92)
Ramachandran outliers	100387	3062 (1.96-1.92)
Sidechain outliers	100360	3062 (1.96-1.92)
RSRZ outliers	91569	2915 (1.96-1.92)

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 $\leq 5\%$. 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	388	<div> <div>3%</div> <div>89%</div> <div>10%</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
2	NAG	A	404	-	-	-	X
5	GOL	A	410	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	NA	A	415	-	-	-	X
7	NA	A	416	-	-	-	X
7	NA	A	417	-	-	-	X
8	SO4	A	420	-	X	-	-

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 3533 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 Dscam.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	388	Total	C	N	O	S	0	0	0
			3014	1903	524	575	12			

There are 20 discrepancies between the modelled and reference sequences:

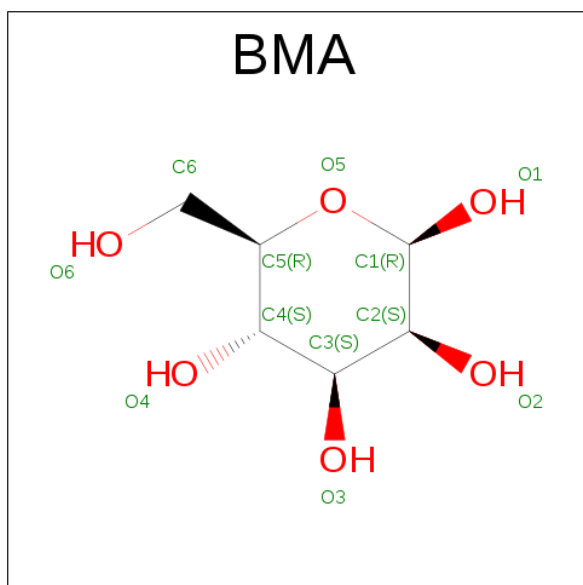
Chain	Residue	Modelled	Actual	Comment	Reference
A	207	ILE	VAL	variant	UNP Q9NBA1
A	211	ALA	PRO	variant	UNP Q9NBA1
A	213	ARG	LYS	variant	UNP Q9NBA1
A	214	THR	ILE	variant	UNP Q9NBA1
A	215	PRO	ASN	variant	UNP Q9NBA1
A	216	ALA	THR	variant	UNP Q9NBA1
A	218	VAL	THR	variant	UNP Q9NBA1
A	219	GLN	TYR	variant	UNP Q9NBA1
A	222	LEU	ASN	variant	UNP Q9NBA1
A	223	GLU	ILE	variant	UNP Q9NBA1
A	224	LEU	VAL	variant	UNP Q9NBA1
A	225	MET	GLU	variant	UNP Q9NBA1
A	226	VAL	SER	variant	UNP Q9NBA1
A	227	ALA	MET	variant	UNP Q9NBA1
A	228	HIS	ALA	variant	UNP Q9NBA1
A	229	THR	SER	variant	UNP Q9NBA1
A	230	ILE	THR	variant	UNP Q9NBA1
A	231	SER	ALA	variant	UNP Q9NBA1
A	232	LEU	ILE	variant	UNP Q9NBA1
A	239	PHE	TYR	variant	UNP Q9NBA1

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C₈H₁₅NO₆).



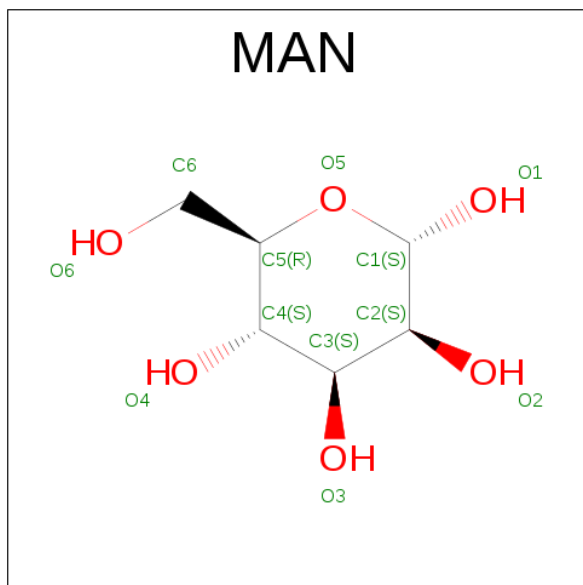
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 3 is BETA-D-MANNOSE (three-letter code: BMA) (formula: $C_6H_{12}O_6$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		
3	A	1	Total	C	O	0	0
			11	6	5		

- Molecule 4 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula: C₆H₁₂O₆).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			11	6	5		
4	A	1	Total	C	O	0	0
			11	6	5		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		
5	A	1	Total	C	O	0	0
			6	3	3		

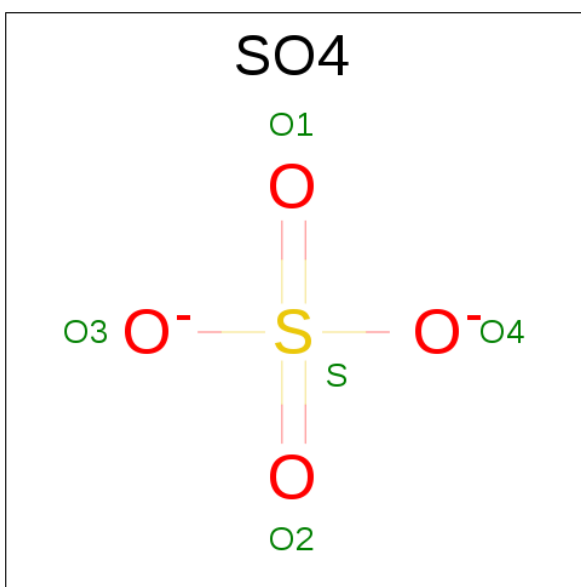
- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	4	Total	Cl	0	0
			4	4		

- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	4	Total	Na	0	0
			4	4		

- Molecule 8 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	O	S	0	0
			5	4	1		
8	A	1	Total	O	S	0	0
			5	4	1		

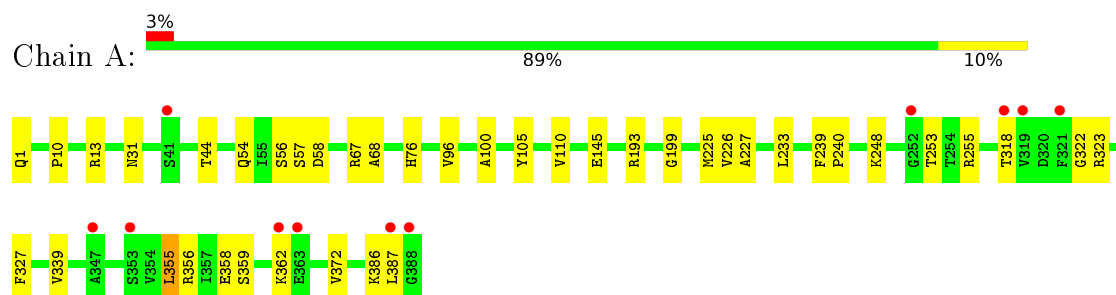
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	383	Total	O	0	0
			383	383		

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: Dscam



4 Data and refinement statistics

Property	Value	Source
Space group	P 42 2 2	Depositor
Cell constants a, b, c, α , β , γ	99.17Å 99.17Å 163.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.83 – 1.95 29.71 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.6 (19.83-1.95) 98.6 (29.71-1.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.47 (at 1.95Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.4_1496)	Depositor
R, R_{free}	0.203 , 0.233 0.203 , 0.234	Depositor DCC
R_{free} test set	3003 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	35.0	Xtriage
Anisotropy	0.377	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 47.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3533	wwPDB-VP
Average B, all atoms (Å ²)	45.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.36% 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, 0.333 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: GOL, BMA, NAG, CL, NA, SO4, MAN

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	0.84	1/3079 (0.0%)	0.79	2/4178 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	105	TYR	CD2-CE2	5.03	1.46	1.39

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	193	ARG	NE-CZ-NH2	-7.36	116.62	120.30
1	A	355	LEU	CA-CB-CG	5.71	128.42	115.30

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	3014	0	2993	27	0
2	A	56	0	48	3	0
3	A	22	0	18	0	0
4	A	22	0	20	0	0
5	A	18	0	23	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	A	4	0	0	1	0
7	A	4	0	0	0	0
8	A	10	0	0	0	0
9	A	383	0	0	10	2
All	All	3533	0	3102	29	2

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.

The worst 5 of 29 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:13:ARG:NH2	9:A:501:HOH:O	2.02	0.90
1:A:67:ARG:HA	2:A:401:NAG:H83	1.65	0.79
2:A:401:NAG:H81	9:A:807:HOH:O	1.91	0.70
1:A:1:GLN:NE2	9:A:504:HOH:O	2.24	0.70
1:A:356:ARG:O	9:A:502:HOH:O	2.15	0.65

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
9:A:776:HOH:O	9:A:776:HOH:O[2_765]	2.06	0.14
9:A:684:HOH:O	9:A:684:HOH:O[6_565]	2.15	0.05

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	386/388 (100%)	376 (97%)	10 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	332/332 (100%)	331 (100%)	1 (0%)	94 94

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	58	ASP

Some 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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 21 ligands modelled in this entry, 8 are monoatomic - leaving 13 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	$\# Z > 2$
2	NAG	A	401	1,2	14,14,15	0.69	0	15,19,21	1.76	3 (20%)
2	NAG	A	402	3,2	14,14,15	0.66	1 (7%)	15,19,21	0.51	0
3	BMA	A	403	2	11,11,12	1.22	2 (18%)	15,15,17	1.23	1 (6%)
2	NAG	A	404	1,2	14,14,15	0.41	0	15,19,21	1.32	2 (13%)
2	NAG	A	405	3,2	14,14,15	0.40	0	15,19,21	0.58	0
3	BMA	A	406	2,4	11,11,12	1.11	0	15,15,17	1.14	1 (6%)
4	MAN	A	407	3	11,11,12	1.15	1 (9%)	15,15,17	0.89	1 (6%)
4	MAN	A	408	3	11,11,12	1.19	2 (18%)	15,15,17	0.89	1 (6%)
5	GOL	A	409	-	5,5,5	0.65	0	5,5,5	1.89	1 (20%)
5	GOL	A	410	-	5,5,5	0.35	0	5,5,5	0.81	0
5	GOL	A	413	-	5,5,5	0.40	0	5,5,5	0.79	0
8	SO4	A	419	-	4,4,4	0.19	0	6,6,6	0.15	0
8	SO4	A	420	-	4,4,4	0.09	0	6,6,6	4.28	4 (66%)

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	401	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	402	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	403	2	-	0/2/19/22	0/1/1/1
2	NAG	A	404	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	405	3,2	-	0/6/23/26	0/1/1/1
3	BMA	A	406	2,4	-	0/2/19/22	0/1/1/1
4	MAN	A	407	3	-	0/2/19/22	0/1/1/1
4	MAN	A	408	3	-	0/2/19/22	0/1/1/1
5	GOL	A	409	-	-	0/4/4/4	0/0/0/0
5	GOL	A	410	-	-	0/4/4/4	0/0/0/0
5	GOL	A	413	-	-	0/4/4/4	0/0/0/0
8	SO4	A	419	-	-	0/0/0/0	0/0/0/0
8	SO4	A	420	-	-	0/0/0/0	0/0/0/0

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	407	MAN	O5-C5	2.04	1.47	1.43
3	A	403	BMA	C2-C3	2.07	1.55	1.52
2	A	402	NAG	C1-C2	2.12	1.55	1.52
4	A	408	MAN	C2-C3	2.22	1.55	1.52
4	A	408	MAN	O5-C5	2.29	1.48	1.43

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	A	420	SO4	O4-S-O2	-2.56	87.98	110.02
8	A	420	SO4	O3-S-O1	-2.55	88.03	110.02
4	A	407	MAN	O2-C2-C3	-2.01	106.13	110.19
4	A	408	MAN	O2-C2-C3	-2.01	106.14	110.19
2	A	404	NAG	C3-C4-C5	2.45	114.60	110.23

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	401	NAG	3	0
5	A	413	GOL	1	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 [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	388/388 (100%)	0.10	11 (2%) 56 65	25, 39, 69, 97	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	321	PHE	5.5
1	A	388	GLY	5.1
1	A	362	LYS	4.9
1	A	363	GLU	3.1
1	A	319	VAL	2.9

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 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(Å ²)	Q<0.9
7	NA	A	416	1/1	0.80	0.44	16.88	68,68,68,68	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
7	NA	A	415	1/1	0.71	0.23	6.32	67,67,67,67	0
5	GOL	A	410	6/6	0.80	0.22	5.30	41,47,54,57	0
2	NAG	A	404	14/15	0.91	0.22	5.20	51,61,69,69	0
7	NA	A	417	1/1	0.82	0.36	4.60	65,65,65,65	0
5	GOL	A	409	6/6	0.94	0.15	1.10	34,37,45,45	0
7	NA	A	418	1/1	0.92	0.11	0.98	58,58,58,58	0
2	NAG	A	401	14/15	0.91	0.11	0.35	25,38,48,65	0
6	CL	A	411	1/1	0.89	0.12	-0.93	80,80,80,80	0
4	MAN	A	408	11/12	0.94	0.31	-	60,68,76,80	0
6	CL	A	412	1/1	0.81	0.08	-	82,82,82,82	0
3	BMA	A	403	11/12	0.77	0.34	-	97,116,125,126	0
5	GOL	A	413	6/6	0.86	0.16	-	46,59,71,74	0
3	BMA	A	406	11/12	0.92	0.19	-	52,66,76,78	0
2	NAG	A	402	14/15	0.87	0.15	-	48,72,101,111	0
6	CL	A	421	1/1	0.89	0.05	-	77,77,77,77	0
2	NAG	A	405	14/15	0.90	0.20	-	56,63,70,78	0
8	SO4	A	420	5/5	0.60	0.41	-	151,151,158,204	5
8	SO4	A	419	5/5	0.94	0.27	-	84,96,114,140	0
4	MAN	A	407	11/12	0.82	0.29	-	54,81,100,105	0
6	CL	A	414	1/1	0.94	0.05	-	78,78,78,78	0

6.5 Other polymers [i](#)

There are no such residues in this entry.