



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 24, 2024 – 12:19 PM EDT

PDB ID : 6C08
Title : Zebrafish SLC38A9 with arginine bound in the cytosol open state
Authors : Lei, H.-T.; Gonen, T.
Deposited on : 2017-12-28
Resolution : 3.17 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 ⓘ 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 : 2.37.1
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

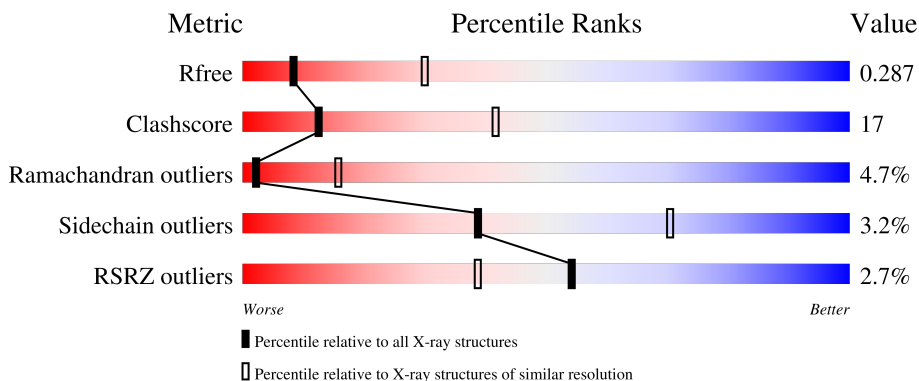
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.17 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.



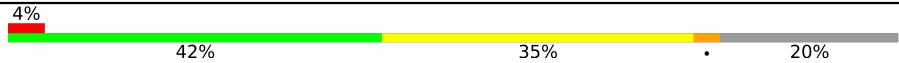
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1467 (3.20-3.16)
Clashscore	141614	1599 (3.20-3.16)
Ramachandran outliers	138981	1574 (3.20-3.16)
Sidechain outliers	138945	1573 (3.20-3.16)
RSRZ outliers	127900	1423 (3.20-3.16)

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 $\leq 5\%$. 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	218	 79% 18%
1	D	218	 77% 20%
2	B	215	 78% 21%
2	E	215	 79% 21%
3	C	479	 41% 31% 24%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	479	 <p>A horizontal bar chart representing the quality of chain. The bar is divided into four segments: a small red segment at the beginning labeled '4%', followed by a green segment labeled '42%', a yellow segment labeled '35%', and a grey segment at the end labeled '20%'. A small black dot is visible on the yellow segment.</p>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 12341 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 antibody Fab Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	214	1561	983	258	316	4	0	0	0
1	D	214	1561	983	258	316	4	0	0	0

- Molecule 2 is a protein called antibody Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	215	1642	1025	274	339	4	0	0	0
2	E	215	1642	1025	274	339	4	0	0	0

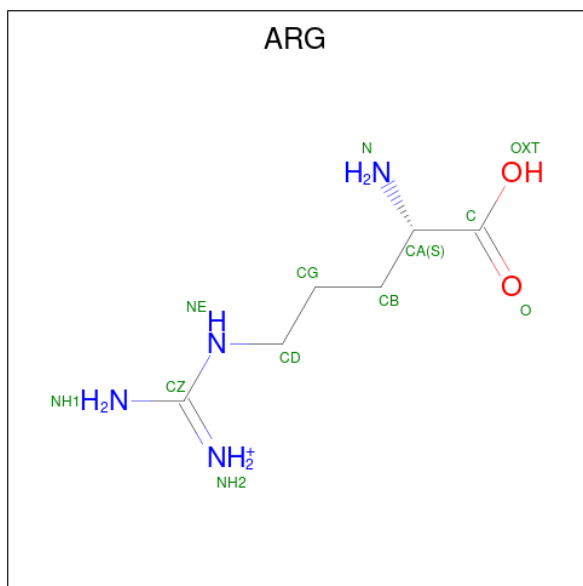
- Molecule 3 is a protein called Sodium-coupled neutral amino acid transporter 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	364	2866	1942	439	466	19	0	0	0
3	F	385	3057	2066	471	501	19	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	227	GLN	ASN	engineered mutation	UNP Q08BA4
C	235	GLN	ASN	engineered mutation	UNP Q08BA4
C	252	GLN	ASN	engineered mutation	UNP Q08BA4
C	263	GLN	ASN	engineered mutation	UNP Q08BA4
F	227	GLN	ASN	engineered mutation	UNP Q08BA4
F	235	GLN	ASN	engineered mutation	UNP Q08BA4
F	252	GLN	ASN	engineered mutation	UNP Q08BA4
F	263	GLN	ASN	engineered mutation	UNP Q08BA4

- Molecule 4 is ARGinine (three-letter code: ARG) (formula: $C_6H_{15}N_4O_2$).

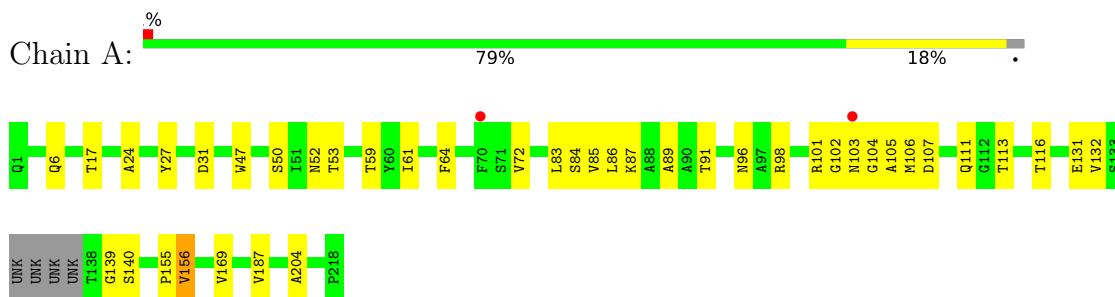


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	F	1	12	6	4	2	0	0

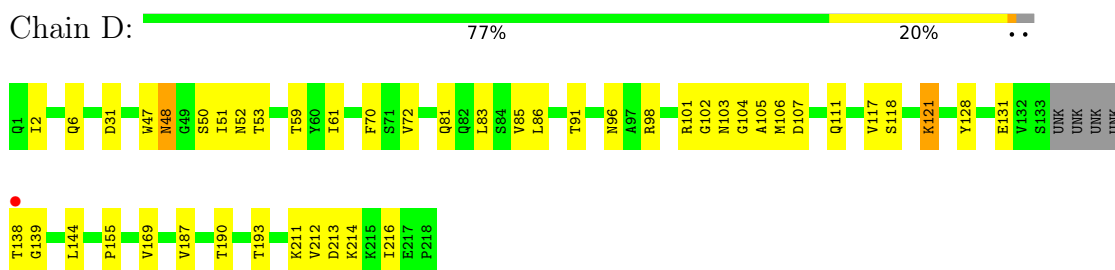
3 Residue-property plots

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.

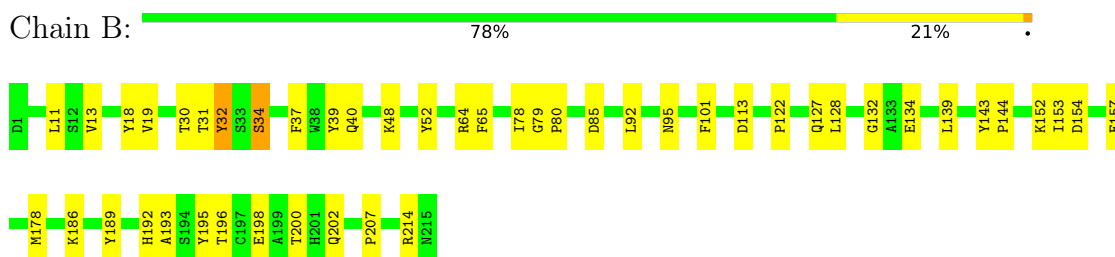
- Molecule 1: antibody Fab Heavy Chain



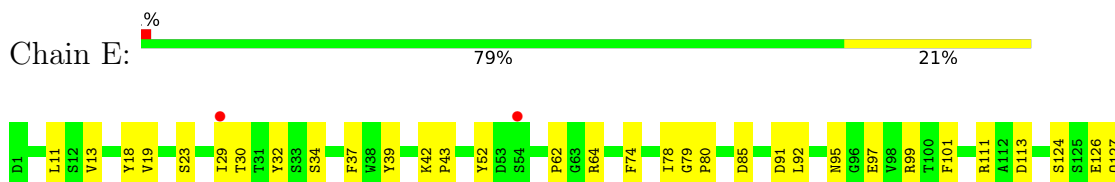
- Molecule 1: antibody Fab Heavy Chain



- Molecule 2: antibody Fab light chain

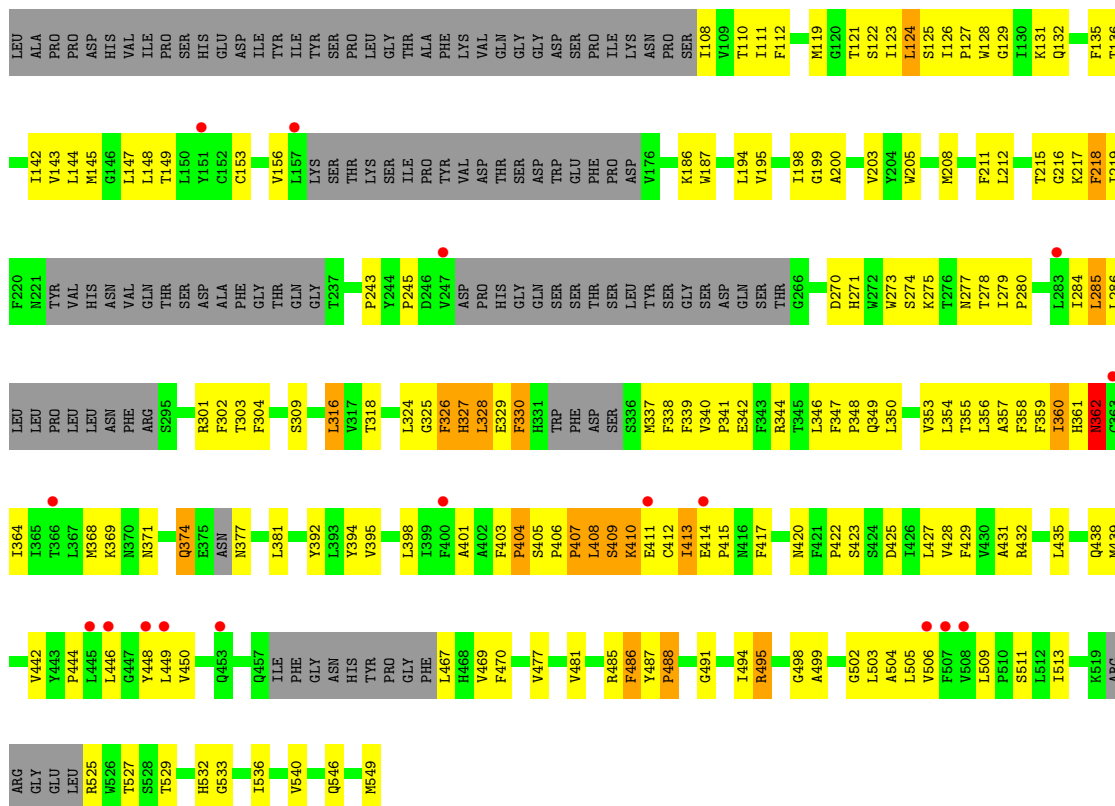


- Molecule 2: antibody Fab light chain

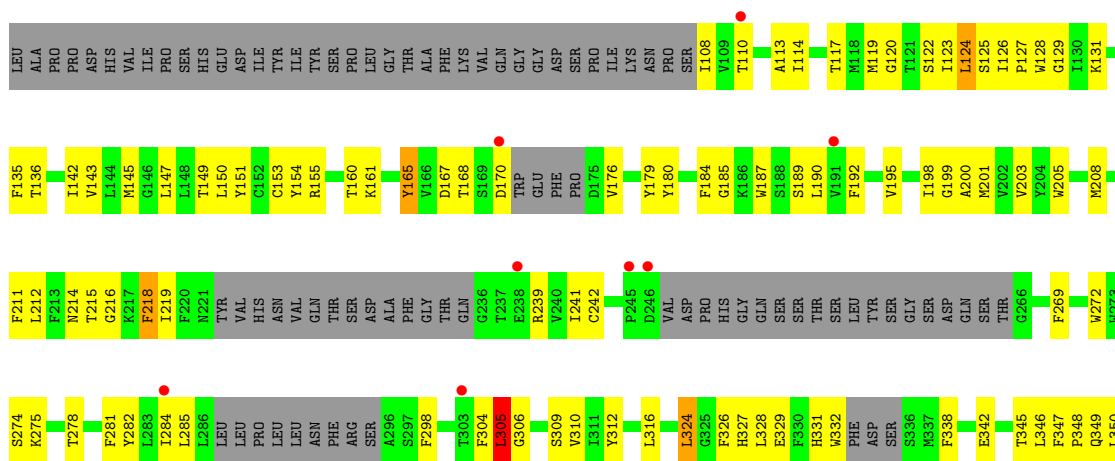


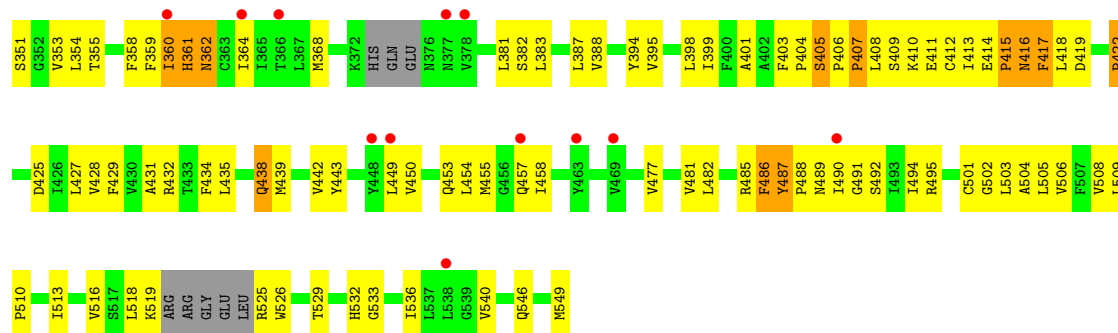


• Molecule 3: Sodium-coupled neutral amino acid transporter 9



• Molecule 3: Sodium-coupled neutral amino acid transporter 9





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	136.61Å 82.81Å 158.92Å 90.00° 100.02° 90.00°	Depositor
Resolution (Å)	156.50 – 3.17 156.50 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.8 (156.50-3.17) 93.8 (156.50-3.10)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.02 (at 3.13Å)	Xtrriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.267 , 0.285 0.266 , 0.287	Depositor DCC
R_{free} test set	2000 reflections (3.14%)	wwPDB-VP
Wilson B-factor (Å ²)	105.2	Xtrriage
Anisotropy	0.042	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 68.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	12341	wwPDB-VP
Average B, all atoms (Å ²)	113.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% 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](#)

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	0/1597	0.51	0/2178
1	D	0.30	0/1597	0.51	0/2178
2	B	0.29	0/1679	0.50	0/2285
2	E	0.30	0/1679	0.51	0/2285
3	C	0.32	0/2942	0.70	0/3998
3	F	0.33	0/3145	0.70	0/4275
All	All	0.31	0/12639	0.61	0/17199

There are no bond length outliers.

There are no bond angle outliers.

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	1561	0	1524	31	0
1	D	1561	0	1524	34	0
2	B	1642	0	1561	31	0
2	E	1642	0	1561	31	0
3	C	2866	0	2923	143	0
3	F	3057	0	3090	153	0
4	F	12	0	12	3	0
All	All	12341	0	12195	411	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 17.

The worst 5 of 411 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)
3:F:122:SER:HG	4:F:601:ARG:N	1.49	1.10
2:E:13:VAL:HG21	2:E:19:VAL:HG12	1.33	1.09
3:C:491:GLY:O	3:C:495:ARG:HD3	1.62	0.98
2:B:13:VAL:HG21	2:B:19:VAL:HG12	1.43	0.98
3:F:489:ASN:HB3	3:F:492:SER:HB2	1.51	0.92

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	210/218 (96%)	196 (93%)	8 (4%)	6 (3%)	4	26
1	D	210/218 (96%)	196 (93%)	10 (5%)	4 (2%)	8	37
2	B	213/215 (99%)	197 (92%)	12 (6%)	4 (2%)	8	37
2	E	213/215 (99%)	198 (93%)	13 (6%)	2 (1%)	17	54
3	C	346/479 (72%)	290 (84%)	28 (8%)	28 (8%)	1	5
3	F	369/479 (77%)	305 (83%)	34 (9%)	30 (8%)	1	5
All	All	1561/1824 (86%)	1382 (88%)	105 (7%)	74 (5%)	2	16

5 of 74 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	103	ASN
1	A	140	SER
1	A	155	PRO
3	C	122	SER
3	C	326	PHE

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	174/176 (99%)	172 (99%)	2 (1%)	73 88
1	D	174/176 (99%)	169 (97%)	5 (3%)	42 72
2	B	184/185 (100%)	182 (99%)	2 (1%)	73 88
2	E	184/185 (100%)	177 (96%)	7 (4%)	33 66
3	C	311/424 (73%)	298 (96%)	13 (4%)	30 63
3	F	332/424 (78%)	318 (96%)	14 (4%)	30 63
All	All	1359/1570 (87%)	1316 (97%)	43 (3%)	39 70

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	158	ARG
3	F	359	PHE
3	F	165	TYR
3	F	269	PHE
3	F	416	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	438	GLN
3	F	489	ASN
2	E	202	GLN
3	F	132	GLN
3	F	214	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](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

1 ligand is 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	ARG	F	601	-	10,11,11	0.73	1 (10%)	11,13,13	1.21	2 (18%)

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	ARG	F	601	-	-	2/11/11/11	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	601	ARG	OXT-C	-2.03	1.23	1.30

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	601	ARG	OXT-C-O	-3.20	116.81	124.09
4	F	601	ARG	OXT-C-CA	2.28	121.15	113.38

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	F	601	ARG	O-C-CA-N
4	F	601	ARG	OXT-C-CA-N

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	601	ARG	3	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	214/218 (98%)	0.18	2 (0%) 84 75	61, 86, 111, 138	0
1	D	214/218 (98%)	0.22	1 (0%) 91 86	63, 85, 105, 145	0
2	B	215/215 (100%)	0.07	0 100 100	65, 87, 123, 138	0
2	E	215/215 (100%)	0.18	3 (1%) 75 63	65, 86, 122, 141	0
3	C	364/479 (75%)	0.16	17 (4%) 31 18	92, 142, 172, 189	0
3	F	385/479 (80%)	0.24	20 (5%) 27 15	94, 144, 176, 200	0
All	All	1607/1824 (88%)	0.18	43 (2%) 54 39	61, 107, 166, 200	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	449	LEU	11.7
3	F	378	VAL	6.5
2	E	215	ASN	5.6
3	F	449	LEU	5.0
3	F	490	ILE	4.7

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
4	ARG	F	601	12/12	0.78	0.23	133,144,157,163	0

6.5 Other polymers [i](#)

There are no such residues in this entry.