WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 07:33 AM EDT

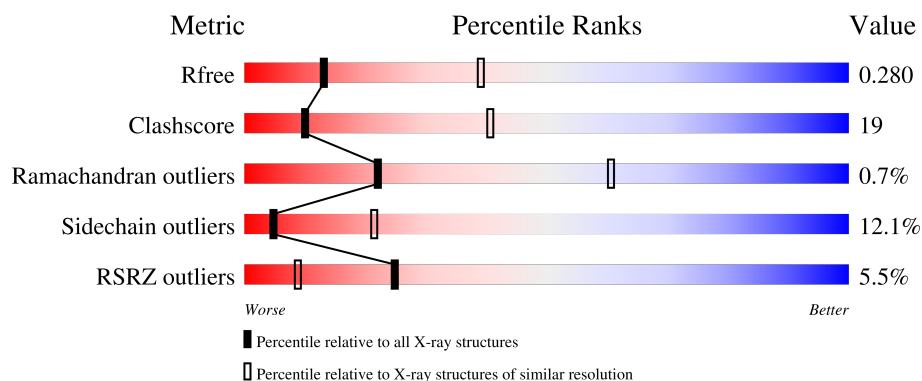|		
PDB ID	:	5JSZ
Title	:	Folate ECF transporter: apo state
Authors	:	Swier, L.J.Y.M.; Guskov, A.; Slotboom, D.J.
Deposited on	:	2016-05-09
Resolution	:	3.00 Å(reported)

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 3.00 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	Percentile Ranks	Value
Rfree		0.280
Clashscore		19
Ramachandran outliers		0.7%
Sidechain outliers		12.1%
RSRZ outliers		5.5%

Worse   *Better*

▮ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	300	4% 54% 35% • 7%
1	E	300	3% 51% 35% 7% • 7%
2	B	287	61% 34% • •
2	F	287	2% 62% 33% • •
3	C	184	9% 44% 40% 5% • 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	184	5% 46% 39% 5% 10%
4	D	265	9% 48% 42% 7%
4	H	265	12% 49% 37% 9%

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 15394 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 Energy-coupling factor transporter ATP-binding protein EcfA1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	280	Total	C	N	O	S	0	0	0
			2147	1354	359	430	4			
1	E	280	Total	C	N	O	S	0	0	0
			2147	1354	359	430	4			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-17	MET	-	initiating methionine	UNP Q1GBJ0
A	-16	HIS	-	expression tag	UNP Q1GBJ0
A	-15	HIS	-	expression tag	UNP Q1GBJ0
A	-14	HIS	-	expression tag	UNP Q1GBJ0
A	-13	HIS	-	expression tag	UNP Q1GBJ0
A	-12	HIS	-	expression tag	UNP Q1GBJ0
A	-11	HIS	-	expression tag	UNP Q1GBJ0
A	-10	HIS	-	expression tag	UNP Q1GBJ0
A	-9	HIS	-	expression tag	UNP Q1GBJ0
A	-8	HIS	-	expression tag	UNP Q1GBJ0
A	-7	HIS	-	expression tag	UNP Q1GBJ0
A	-6	GLY	-	expression tag	UNP Q1GBJ0
A	-5	GLU	-	expression tag	UNP Q1GBJ0
A	-4	ASN	-	expression tag	UNP Q1GBJ0
A	-3	LEU	-	expression tag	UNP Q1GBJ0
A	-2	TYR	-	expression tag	UNP Q1GBJ0
A	-1	PHE	-	expression tag	UNP Q1GBJ0
A	0	GLN	-	expression tag	UNP Q1GBJ0
A	1	GLY	-	expression tag	UNP Q1GBJ0
E	-17	MET	-	initiating methionine	UNP Q1GBJ0
E	-16	HIS	-	expression tag	UNP Q1GBJ0
E	-15	HIS	-	expression tag	UNP Q1GBJ0
E	-14	HIS	-	expression tag	UNP Q1GBJ0
E	-13	HIS	-	expression tag	UNP Q1GBJ0
E	-12	HIS	-	expression tag	UNP Q1GBJ0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	-11	HIS	-	expression tag	UNP Q1GBJ0
E	-10	HIS	-	expression tag	UNP Q1GBJ0
E	-9	HIS	-	expression tag	UNP Q1GBJ0
E	-8	HIS	-	expression tag	UNP Q1GBJ0
E	-7	HIS	-	expression tag	UNP Q1GBJ0
E	-6	GLY	-	expression tag	UNP Q1GBJ0
E	-5	GLU	-	expression tag	UNP Q1GBJ0
E	-4	ASN	-	expression tag	UNP Q1GBJ0
E	-3	LEU	-	expression tag	UNP Q1GBJ0
E	-2	TYR	-	expression tag	UNP Q1GBJ0
E	-1	PHE	-	expression tag	UNP Q1GBJ0
E	0	GLN	-	expression tag	UNP Q1GBJ0
E	1	GLY	-	expression tag	UNP Q1GBJ0

- Molecule 2 is a protein called Energy-coupling factor transporter ATP-binding protein EcfA2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	282	Total 2191	C 1399	N 371	O 412	S 9	0	0	0
2	F	282	Total 2191	C 1399	N 371	O 412	S 9	0	0	0

- Molecule 3 is a protein called Conserved hypothetical membrane protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	165	Total 1276	C 861	N 199	O 206	S 10	0	0	0
3	G	165	Total 1276	C 861	N 199	O 206	S 10	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	177	TRP	-	expression tag	UNP Q1G929
C	178	SER	-	expression tag	UNP Q1G929
C	179	HIS	-	expression tag	UNP Q1G929
C	180	PRO	-	expression tag	UNP Q1G929
C	181	GLN	-	expression tag	UNP Q1G929
C	182	PHE	-	expression tag	UNP Q1G929
C	183	GLU	-	expression tag	UNP Q1G929
C	184	LYS	-	expression tag	UNP Q1G929
G	177	TRP	-	expression tag	UNP Q1G929

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	178	SER	-	expression tag	UNP Q1G929
G	179	HIS	-	expression tag	UNP Q1G929
G	180	PRO	-	expression tag	UNP Q1G929
G	181	GLN	-	expression tag	UNP Q1G929
G	182	PHE	-	expression tag	UNP Q1G929
G	183	GLU	-	expression tag	UNP Q1G929
G	184	LYS	-	expression tag	UNP Q1G929

- Molecule 4 is a protein called Energy-coupling factor transporter transmembrane protein EcfT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	259	Total	C	N	O	S	0	0	0
			2083	1391	333	345	14			
4	H	259	Total	C	N	O	S	0	0	0
			2083	1391	333	345	14			

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: Energy-coupling factor transporter ATP-binding protein EcfA1

- Molecule 1: Energy-coupling factor transporter ATP-binding protein EcfA1

- Molecule 2: Energy-coupling factor transporter ATP-binding protein EcfA2

M1 F5 E6 N7 V8 S13 P14 G15 S16 P17 L18 E19 A20 I21 Q25 L26 N27 F28 E31 E32 G33 K34 L38 M50 K63 I66 A67 T72 K81 K86 V87 S88 L89 A90 F91 Q92 A96 Q97 L98 F99 E100 N101 T102 V103 L104 K105 D106 V107

P111 F114 E120 E123 A124 K127 W128 L129 K130 K131 L134 K135 D136 D137 L138 I139 E140 H141 S142 P143 L146 Q150 M151 R152 R153 L160 A161 Y162 E163 P164 I167 C168 L169 D170 E171 P172 A173 L176 R181 T197 V198 I199 L200 V201 N204

D207 Y211 D214 V215 L218 E219 I224 S228 P229 K230 D235 L239 Q240 R248 R251 K255 L256 E257 G260 L261 K262 L263 L268 T269 M270 P271 E272 L273 A274 D275 A276 I277 S280 L281 K282 GLY GLY HIS GLU

• Molecule 2: Energy-coupling factor transporter ATP-binding protein EcfA2

Chain F:

M1 A2 I3 K4 F5 E6 N7 Y10 V11 Y12 S13 P14 E19 A20 I21 G22 L23 L26 N27 F28 E31 E32 G33 K34 L38 M50 L56 L57 T60 K63 I66 A67 T72 L80 K81 K86 V87 A90 F91 Q92 A96 Q97 L98 F99 E100
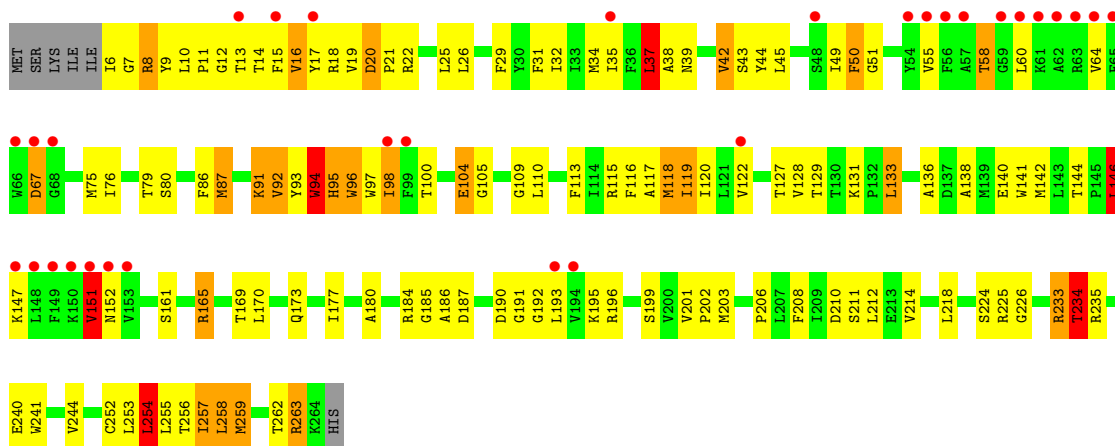
N101 T102 V103 L104 K105 D106 V107 P111 F114 G115 F116 E120 K127 W128 L129 K130 K131 L134 K135 D136 D137 L138 I139 E140 H141 S142 P143 L146 Q150 M151 R152 R153 L156 Y162 E163 P164 E165 I166 I167 C168 L169 D170 E171 P172 A173 L176 R181 M185

K189 T197 V198 I199 L200 V201 N204 M205 D206 D207 Y211 D214 L218 E219 I224 S228 P229 K230 F233 K234 D235 R248 R251 K255 L256 E257 G260 L261 L268 T269 M270 P271 E272 L273 A274 D275 A276 S280 L281 K282 GLY GLY GLU HIS GLU

• Molecule 3: Conserved hypothetical membrane protein

Chain C:

MET LYS SER GLU SER LYS VAL SER S9 E12 E15 L18 L19 A20 M21 V22 I23 K26 K33 V34 G35 V41 G42 F45 I46 Q47 S48 V49 M50 L51 L54 F55 W59 G60 F61 A65 L66 S67 D68 L69 V70 S71 S72 V73 I74 F75 N77 L78 G79

G80 F81 F82 F85 T86 L87 T88 A89 A90 L91 M94 I95 Y96 G97 F98 F99 L100 P104 I105 Q106 I107 W108 R109 V110 I115 C116 V117 T118 V119 I120 C121 N122 I123 N126 W129 M133 Y134 G135 I136 N137 F138 M139 R145 I146 L147 K148 E149 M150 I151 T152 P153 W154

I155 Q156 M157 V158 A159 V160 W161 F162 I163 L164 E165 G166 L167 S168 R169 V170 K171 L172 S173 ARG LYS PHE TRP SER HIS PRO GLN PHE GLU LYS

• Molecule 3: Conserved hypothetical membrane protein

Chain G:

MET LYS SER GLU SER LYS VAL SER S9 R14 E15 L18 L19 A20 M21 V22 I23 K26 K33 V34 G35 N36 Q40 F45 S48 V49 M50 L51 F55 W59 G60 F61 A62 G63 G64 A65 L66 S67 D68 S71 G76 N77 L78 G79 G80 F81 F82 I83 G84

F85 T86 A90 L91 M94 I95 Y96 F99 L100 Q103 P104 I105 Q106 I107 W108 R109 V110 S113 V114 I115 C116 V117 T118 V119 N122 I123 N126 T127 L128 M133 Y134 G135 I136 N137 F138 A141 L142 R145 I146 L147 K148 E149 M150 I151 T152 P153 W154 I155 Q156 M157

V158 A159 V160 W161 F162 I163 L164 E165 G166 L167 S168 R169 V170 K171 L172 S173 ARG LYS PHE TRP SER HIS PRO GLN PHE GLU LYS

• Molecule 4: Energy-coupling factor transporter transmembrane protein EcfT

Chain D:

• Molecule 4: Energy-coupling factor transporter transmembrane protein EcfT

Chain H:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	88.82Å 95.32Å 107.57Å 83.45° 65.75° 61.99°	Depositor
Resolution (Å)	49.00 – 3.00 49.01 – 3.00	Depositor EDS
% Data completeness (in resolution range)	73.8 (49.00-3.00) 73.2 (49.01-3.00)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 3.01Å)	Xtriage
Refinement program	PHENIX (1.10_2155: ???)	Depositor
R, R_{free}	0.233 , 0.277 0.234 , 0.280	Depositor DCC
R_{free} test set	2088 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	89.9	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 49.4	EDS
L-test for twinning[2]	$< |L| > = 0.45, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.327 for h,h-k,h-l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	15394	wwPDB-VP
Average B, all atoms (Å^2)	102.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.42% 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 ⓘ

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	0.28	0/2180	0.52	0/2959
1	E	0.28	0/2180	0.58	2/2959 (0.1%)
2	B	0.27	0/2237	0.48	0/3019
2	F	0.26	0/2237	0.48	1/3019 (0.0%)
3	C	0.28	0/1307	0.55	0/1773
3	G	0.29	0/1307	0.58	0/1773
4	D	0.34	0/2136	0.60	1/2901 (0.0%)
4	H	0.37	0/2136	0.65	4/2901 (0.1%)
All	All	0.30	0/15720	0.56	8/21304 (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	3
1	E	0	3
2	B	0	1
2	F	0	1
3	C	0	1
4	D	0	1
4	H	0	3
All	All	0	13

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	26	LEU	CA-CB-CG	6.32	129.83	115.30
4	H	254	LEU	CB-CG-CD1	-6.10	100.63	111.00
4	H	146	LEU	CA-CB-CG	5.71	128.42	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	12	TYR	C-N-CA	5.56	135.60	121.70
1	E	222	LEU	CA-CB-CG	5.53	128.03	115.30

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	12	THR	Peptide
1	A	18	SER	Peptide
1	A	69	VAL	Peptide
2	B	7	ASN	Peptide
3	C	82	PHE	Peptide

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	2147	0	2159	84	0
1	E	2147	0	2159	83	0
2	B	2191	0	2199	60	0
2	F	2191	0	2199	61	0
3	C	1276	0	1349	68	0
3	G	1276	0	1349	60	0
4	D	2083	0	2185	117	0
4	H	2083	0	2185	111	0
All	All	15394	0	15784	593	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 19.

The worst 5 of 593 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)
4:H:94:TRP:HB2	4:H:95:HIS:CG	2.09	0.87
4:D:94:TRP:HB3	4:D:95:HIS:HB2	1.58	0.86

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:118:THR:HA	3:C:122:ASN:HB3	1.58	0.83
3:G:61:PHE:HB2	3:G:94:MET:HB2	1.61	0.81
4:D:22:ARG:NH2	4:D:240:GLU:O	2.14	0.81

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	278/300 (93%)	250 (90%)	25 (9%)	3 (1%)	14	50
1	E	278/300 (93%)	254 (91%)	22 (8%)	2 (1%)	22	60
2	B	280/287 (98%)	264 (94%)	16 (6%)	0	100	100
2	F	280/287 (98%)	263 (94%)	16 (6%)	1 (0%)	34	72
3	C	163/184 (89%)	143 (88%)	18 (11%)	2 (1%)	13	48
3	G	163/184 (89%)	147 (90%)	15 (9%)	1 (1%)	25	64
4	D	257/265 (97%)	244 (95%)	11 (4%)	2 (1%)	19	57
4	H	257/265 (97%)	241 (94%)	13 (5%)	3 (1%)	13	48
All	All	1956/2072 (94%)	1806 (92%)	136 (7%)	14 (1%)	22	60

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	13	PHE
4	D	153	VAL
4	D	234	THR
2	F	13	SER
4	H	234	THR

5.3.2 Protein sidechains ⓘ

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	240/259 (93%)	220 (92%)	20 (8%)	11	39
1	E	240/259 (93%)	202 (84%)	38 (16%)	2	12
2	B	231/234 (99%)	209 (90%)	22 (10%)	8	32
2	F	231/234 (99%)	209 (90%)	22 (10%)	8	32
3	C	136/155 (88%)	118 (87%)	18 (13%)	4	18
3	G	136/155 (88%)	120 (88%)	16 (12%)	5	22
4	D	227/233 (97%)	199 (88%)	28 (12%)	4	21
4	H	227/233 (97%)	189 (83%)	38 (17%)	2	11
All	All	1668/1762 (95%)	1466 (88%)	202 (12%)	5	21

5 of 202 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	196	THR
2	F	219	GLU
4	H	258	LEU
1	E	217	ASP
2	F	34	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	96	GLN
1	E	96	GLN
3	G	77	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 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(Å²)	Q<0.9
1	A	280/300 (93%)	-0.22	11 (3%)	39	15	41, 81, 147, 261	0
1	E	280/300 (93%)	-0.24	8 (2%)	51	23	40, 83, 154, 279	0
2	B	282/287 (98%)	-0.41	1 (0%)	92	79	44, 76, 134, 209	0
2	F	282/287 (98%)	-0.20	5 (1%)	68	40	39, 74, 147, 225	0
3	C	165/184 (89%)	0.19	17 (10%)	6	2	51, 121, 192, 247	0
3	G	165/184 (89%)	0.09	10 (6%)	21	7	59, 121, 198, 261	0
4	D	259/265 (97%)	0.20	25 (9%)	7	2	54, 110, 203, 255	0
4	H	259/265 (97%)	0.41	31 (11%)	4	1	45, 110, 225, 279	0
All	All	1972/2072 (95%)	-0.05	108 (5%)	25	9	39, 90, 186, 279	0

The worst 5 of 108 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	H	60	LEU	13.5
3	G	173	SER	11.2
4	H	13	THR	11.1
4	H	150	LYS	10.0
3	C	173	SER	9.9

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.