W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 15, 2017 – 12:30 am GMT

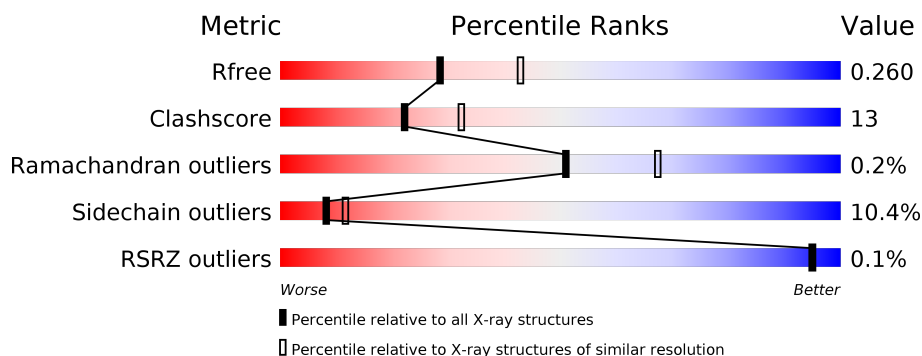|		
PDB ID	:	2NNN
Title	:	Crystal structure of probable transcriptional regulator from Pseudomonas aeruginosa
Authors	:	Chang, C.; Evdokimova, E.; Altamentova, S.; Savchenko, A.; Edwards, A.M.; Joachimiak, A.; Midwest Center for Structural Genomics (MCSG)
Deposited on	:	2006-10-24
Resolution	:	2.40 Å(reported)

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 2.40 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		0.260
Clashscore		13
Ramachandran outliers		0.2%
Sidechain outliers		10.4%
RSRZ outliers		0.1%

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}	100719	3166 (2.40-2.40)
Clashscore	112137	3674 (2.40-2.40)
Ramachandran outliers	110173	3616 (2.40-2.40)
Sidechain outliers	110143	3617 (2.40-2.40)
RSRZ outliers	101464	3195 (2.40-2.40)

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 <=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	140	72% · · · 20% · ·
1	B	140	74% · · · 19% · ·
1	C	140	71% · · · 21% · 6%
1	D	140	74% · · · 19% · 5%
1	E	140	66% · · · 23% 6% 6%
1	F	140	69% · · · 22% · 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	140	59% 31% 6%
1	H	140	75% 19% 5%
1	I	140	72% 21% 5%
1	J	140	71% 21% 5%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 10506 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 Probable transcriptional regulator.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	134	Total	C	N	O	S	Se	0	0	0
			1009	624	194	189	1	1			
1	B	134	Total	C	N	O	S	Se	0	0	0
			1009	624	194	189	1	1			
1	C	132	Total	C	N	O	S	Se	0	0	0
			995	616	192	185	1	1			
1	D	133	Total	C	N	O	S	Se	0	0	0
			1004	621	193	188	1	1			
1	E	132	Total	C	N	O	S	Se	0	0	0
			996	617	192	185	1	1			
1	F	133	Total	C	N	O	S	Se	0	0	0
			1004	621	193	188	1	1			
1	G	132	Total	C	N	O	S	Se	0	0	0
			995	616	192	185	1	1			
1	H	133	Total	C	N	O	S	Se	0	0	0
			1004	621	193	188	1	1			
1	I	133	Total	C	N	O	S	Se	0	0	0
			1004	621	193	188	1	1			
1	J	133	Total	C	N	O	S	Se	0	0	0
			1004	621	193	188	1	1			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
A	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
B	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
B	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
C	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
C	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
D	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
D	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
E	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
F	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
F	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
G	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
G	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
H	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
H	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
I	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
I	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
J	1	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8
J	64	MSE	MET	MODIFIED RESIDUE	UNP Q9I3B8

- Molecule 2 is water.

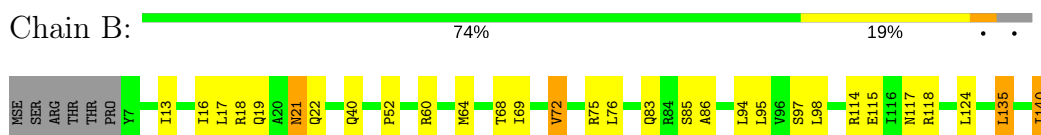| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	67	Total 67	O 67	0	0
2	B	62	Total 62	O 62	0	0
2	C	36	Total 36	O 36	0	0
2	D	63	Total 63	O 63	0	0
2	E	45	Total 45	O 45	0	0
2	F	38	Total 38	O 38	0	0
2	G	42	Total 42	O 42	0	0
2	H	51	Total 51	O 51	0	0
2	I	33	Total 33	O 33	0	0
2	J	45	Total 45	O 45	0	0

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 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.
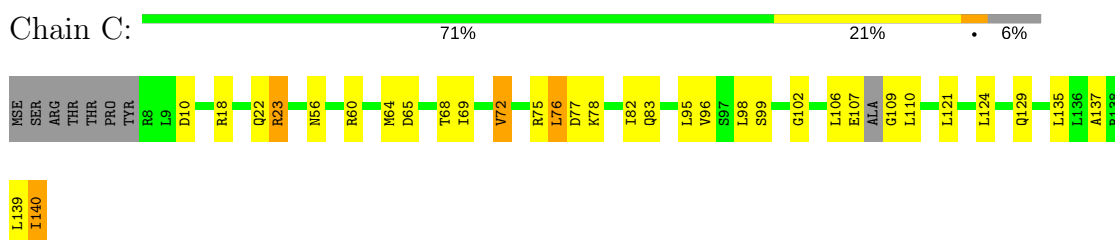
- Molecule 1: Probable transcriptional regulator
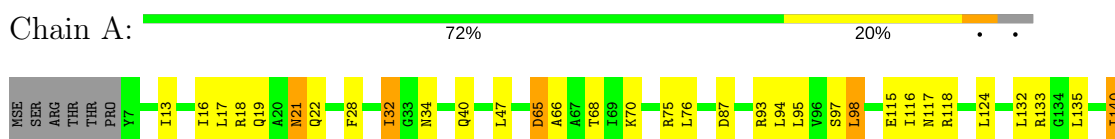
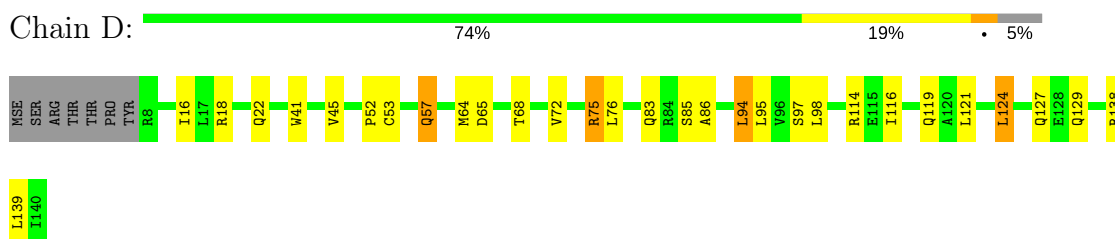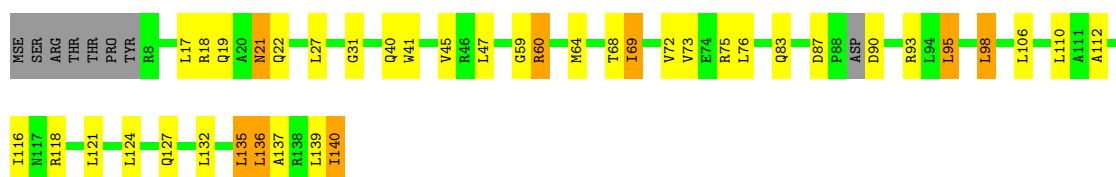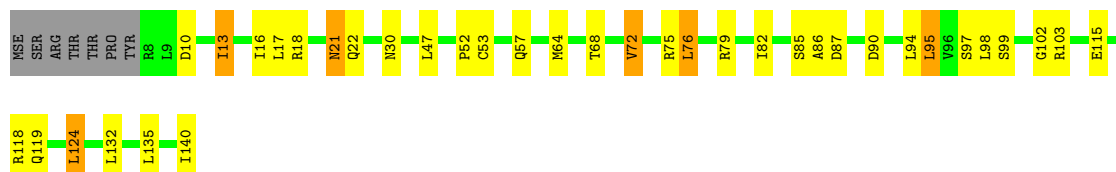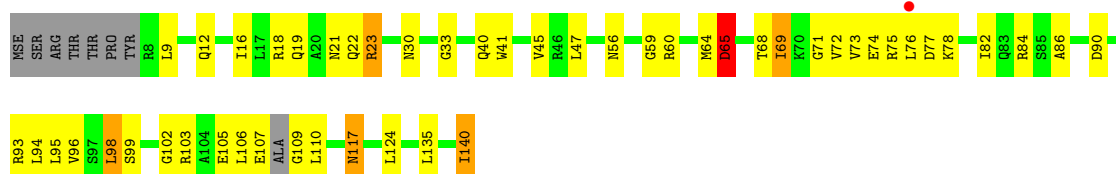Chain A:

- Molecule 1: Probable transcriptional regulator

Chain B:

- Molecule 1: Probable transcriptional regulator

Chain C:

- Molecule 1: Probable transcriptional regulator

Chain D:

- Molecule 1: Probable transcriptional regulator

Chain E:

MSE SER ARG THR PRO TYR R8 L17 R18 Q19 A20 N21 Q22 L27 G31 Q40 W41 V45 R46 L47 G59 R60 M64 T68 I69 V72 V73 E74 R75 L76 Q83 D87 P88 ASP D90 R93 L94 L95 L98 L106 L110 A111 A112

I116 N117 R118 L121 L124 Q127 L132 L135 L136 A137 R138 L139 I140

• Molecule 1: Probable transcriptional regulator

Chain F:　69%　22%　•　5%

MSE SER ARG THR PRO TYR R8 L9 D10 I13 I16 L17 R18 N21 Q22 N30 L47 P52 C53 Q57 M64 T68 V72 R75 L76 R79 I82 S85 A86 D87 D90 L94 L95 V96 S97 L98 S99 G102 R103 E115

R118 Q119 L124 L132 L135 I140

• Molecule 1: Probable transcriptional regulator

Chain G:　%　59%　31%　•　•　6%

MSE SER ARG THR PRO TYR R8 L9 Q12 I16 L17 R18 Q19 A20 N21 Q22 R23 N30 G33 Q40 W41 V45 R46 L47 N56 G59 R60 M64 D65 T68 I69 K70 G71 V72 V73 E74 R75 L76 D77 K78 I82 Q83 R84 S85 A86 D90

R93 L94 L95 V96 S97 L98 S99 G102 R103 A104 E105 L106 E107 ALA G109 L110 N117 L124 L135 I140

• Molecule 1: Probable transcriptional regulator

Chain H:　75%　19%　•　5%

MSE SER ARG THR PRO TYR R8 Q12 I13 I16 L17 R18 Q19 Q22 Q40 Q57 R60 M64 T68 G71 V72 R75 L76 D90 G91 R92 R93 L94 L95 V96 S97 L98 E115 R118 A122 P123 L139 I140

• Molecule 1: Probable transcriptional regulator

Chain I:　72%　21%　•　5%

MSE SER ARG THR PRO TYR R8 L9 D10 D11 Q12 L17 R18 Q19 A20 N21 Q22 F28 G33 N34 G35 L36 E49 M64 D65 T68 V72 V73 E74 R75 L76 D77 K78 R79 D87 D90 L94 L98 S99 G102 E105 I116 N117 R118

L132 L135 I140
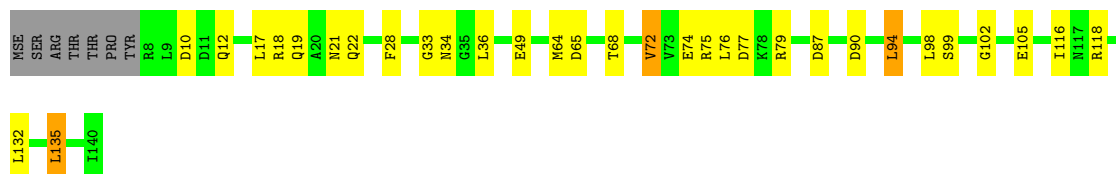
• Molecule 1: Probable transcriptional regulator

Chain J:　71%　21%　•　5%

MSE SER ARG THR PRO TYR R8 I16 L17 R18 N21 Q22 A26 N30 G31 I32 Q40 W41 V45 G59 R60 I69 V72 R75 L76 D87 P88 D89 D90 L94 L95 L98 S99 G102 A112 I116 N117 L124 Q127 L132
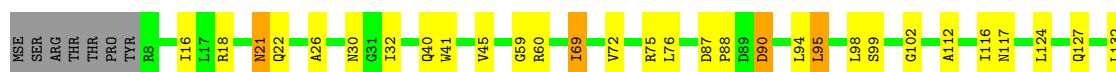
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	89.37Å 89.37Å 217.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	34.14 – 2.40 34.13 – 2.40	Depositor EDS
% Data completeness (in resolution range)	98.5 (34.14-2.40) 98.6 (34.13-2.40)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.09 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.218 , 0.262 0.217 , 0.260	Depositor DCC
R_{free} test set	3761 reflections (5.29%)	DCC
Wilson B-factor (Å^2)	36.9	Xtriage
Anisotropy	0.729	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 37.7	EDS
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.010 for -h,-k,l 0.480 for h,-h-k,-l 0.014 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	10506	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 13.59% 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)

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	0.49	0/1018	0.69	0/1374
1	B	0.49	0/1018	0.68	0/1374
1	C	0.44	0/1003	0.67	0/1352
1	D	0.48	0/1013	0.66	0/1367
1	E	0.43	0/1004	0.62	0/1353
1	F	0.45	0/1013	0.66	0/1367
1	G	0.46	0/1003	0.68	0/1352
1	H	0.48	0/1013	0.64	0/1367
1	I	0.45	0/1013	0.66	0/1367
1	J	0.43	0/1013	0.66	0/1367
All	All	0.46	0/10111	0.66	0/13640

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	1009	0	1035	34	0
1	B	1009	0	1035	37	0
1	C	995	0	1023	25	0
1	D	1004	0	1033	23	0
1	E	996	0	1028	44	0
1	F	1004	0	1033	42	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	995	0	1023	48	0
1	H	1004	0	1033	30	0
1	I	1004	0	1033	25	0
1	J	1004	0	1033	22	0
2	A	67	0	0	3	0
2	B	62	0	0	2	0
2	C	36	0	0	0	0
2	D	63	0	0	3	0
2	E	45	0	0	1	0
2	F	38	0	0	3	0
2	G	42	0	0	5	0
2	H	51	0	0	4	0
2	I	33	0	0	0	0
2	J	45	0	0	0	0
All	All	10506	0	10309	269	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 13.

The worst 5 of 269 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:64:MSE:HE2	1:D:68:THR:HG22	1.19	1.15
1:B:64:MSE:HE2	1:B:68:THR:HG22	1.37	1.07
1:F:64:MSE:HE2	1:F:68:THR:HG22	1.29	1.06
1:C:64:MSE:HE2	1:C:68:THR:HG22	1.36	1.05
1:H:64:MSE:CE	1:H:68:THR:HG22	1.86	1.04

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	132/140 (94%)	129 (98%)	3 (2%)	0	100	100
1	B	132/140 (94%)	130 (98%)	2 (2%)	0	100	100
1	C	128/140 (91%)	125 (98%)	3 (2%)	0	100	100
1	D	131/140 (94%)	131 (100%)	0	0	100	100
1	E	128/140 (91%)	127 (99%)	1 (1%)	0	100	100
1	F	131/140 (94%)	131 (100%)	0	0	100	100
1	G	128/140 (91%)	126 (98%)	0	2 (2%)	11	15
1	H	131/140 (94%)	130 (99%)	1 (1%)	0	100	100
1	I	131/140 (94%)	128 (98%)	3 (2%)	0	100	100
1	J	131/140 (94%)	130 (99%)	0	1 (1%)	22	33
All	All	1303/1400 (93%)	1287 (99%)	13 (1%)	3 (0%)	51	67

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	J	90	ASP
1	G	65	ASP
1	G	90	ASP

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	100/106 (94%)	90 (90%)	10 (10%)	9	13
1	B	100/106 (94%)	90 (90%)	10 (10%)	9	13
1	C	99/106 (93%)	89 (90%)	10 (10%)	9	12
1	D	100/106 (94%)	91 (91%)	9 (9%)	11	16
1	E	99/106 (93%)	86 (87%)	13 (13%)	5	5
1	F	100/106 (94%)	91 (91%)	9 (9%)	11	16
1	G	99/106 (93%)	86 (87%)	13 (13%)	5	5
1	H	100/106 (94%)	94 (94%)	6 (6%)	22	35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	100/106 (94%)	89 (89%)	11 (11%)	7	10
1	J	100/106 (94%)	87 (87%)	13 (13%)	5	5
All	All	997/1060 (94%)	893 (90%)	104 (10%)	8	12

5 of 104 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	124	LEU
1	F	103	ARG
1	J	76	LEU
1	E	127	GLN
1	F	10	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	19	GLN
1	E	117	ASN
1	J	22	GLN
1	E	21	ASN
1	E	40	GLN

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 carbohydrates in this entry.

5.6　Ligand geometry ⓘ

There are no ligands in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

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	133/140 (95%)	-0.16	0	100	100	28, 42, 52, 57	0
1	B	133/140 (95%)	-0.13	0	100	100	28, 42, 52, 57	0
1	C	131/140 (93%)	-0.02	0	100	100	37, 49, 62, 63	0
1	D	132/140 (94%)	-0.13	0	100	100	28, 42, 58, 62	0
1	E	131/140 (93%)	-0.06	0	100	100	42, 53, 62, 64	0
1	F	132/140 (94%)	0.02	0	100	100	35, 46, 60, 64	0
1	G	131/140 (93%)	0.03	1 (0%)	86	84	37, 49, 61, 65	0
1	H	132/140 (94%)	-0.20	0	100	100	31, 42, 60, 65	0
1	I	132/140 (94%)	-0.04	0	100	100	35, 47, 61, 65	0
1	J	132/140 (94%)	-0.02	0	100	100	41, 54, 62, 65	0
All	All	1319/1400 (94%)	-0.07	1 (0%)	95	95	28, 47, 61, 65	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	76	LEU	2.0

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 carbohydrates 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.