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 ⓘ

Sep 6, 2016 – 12:06 AM EDT

PDB ID	:	4Z53
Title	:	Quinolone(Trovafloxacin)-DNA cleavage complex of topoisomerase IV from S. pneumoniae
Authors	:	Laponogov, I.; Veselkov, D.A.; Pan, X.-S.; Selvarajah, J.; Crevel, I.M.-T.; Fisher, L.M.; Sanderson, M.R.
Deposited on	:	2015-04-02
Resolution	:	3.26 Å(reported)

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

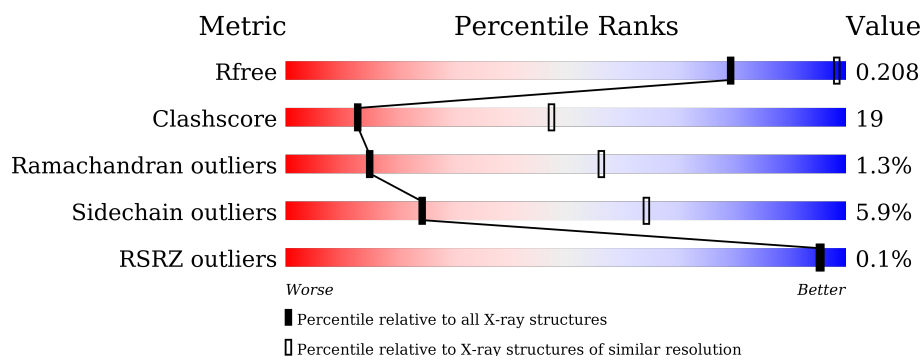|		
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.26 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		0.208
Clashscore		19
Ramachandran outliers		1.3%
Sidechain outliers		5.9%
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}	91344	1624 (3.32-3.20)
Clashscore	102246	1806 (3.32-3.20)
Ramachandran outliers	100387	1773 (3.32-3.20)
Sidechain outliers	100360	1771 (3.32-3.20)
RSRZ outliers	91569	1632 (3.32-3.20)

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	742	67% 22% · 7%
1	B	742	66% 25% · 6%
2	E	7	14% 86%
3	F	11	36% 64%
4	G	7	29% 71%
5	H	11	9% 91%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 11112 atoms, of which 28 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 DNA topoisomerase 4 subunit B,DNA topoisomerase 4 subunit A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	690	Total 5138	C 3258	N 892	O 969	S 19	0	0	0
1	B	695	Total 5148	C 3265	N 887	O 978	S 18	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	403	MET	-	expression tag	UNP Q59961
A	460	ILE	VAL	engineered mutation	UNP Q59961
A	996	ALA	-	linker	UNP Q59961
A	997	THR	-	linker	UNP Q59961
A	998	VAL	-	linker	UNP Q59961
A	999	PHE	-	linker	UNP Q59961
A	1000	HIS	-	linker	UNP Q59961
A	1001	MET	-	linker	UNP Q59961
A	1002	SER	-	linker	UNP Q59961
A	1257	THR	ILE	engineered mutation	UNP P72525
A	1485	THR	-	expression tag	UNP P72525
A	1486	ALA	-	expression tag	UNP P72525
A	1487	LYS	-	expression tag	UNP P72525
A	1488	ALA	-	expression tag	UNP P72525
A	1489	LEU	-	expression tag	UNP P72525
A	1490	GLU	-	expression tag	UNP P72525
A	1491	HIS	-	expression tag	UNP P72525
A	1492	HIS	-	expression tag	UNP P72525
A	1493	HIS	-	expression tag	UNP P72525
A	1494	HIS	-	expression tag	UNP P72525
A	1495	HIS	-	expression tag	UNP P72525
A	1496	HIS	-	expression tag	UNP P72525
B	403	MET	-	expression tag	UNP Q59961
B	460	ILE	VAL	engineered mutation	UNP Q59961

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	996	ALA	-	linker	UNP Q59961
B	997	THR	-	linker	UNP Q59961
B	998	VAL	-	linker	UNP Q59961
B	999	PHE	-	linker	UNP Q59961
B	1000	HIS	-	linker	UNP Q59961
B	1001	MET	-	linker	UNP Q59961
B	1002	SER	-	linker	UNP Q59961
B	1257	THR	ILE	engineered mutation	UNP P72525
B	1485	THR	-	expression tag	UNP P72525
B	1486	ALA	-	expression tag	UNP P72525
B	1487	LYS	-	expression tag	UNP P72525
B	1488	ALA	-	expression tag	UNP P72525
B	1489	LEU	-	expression tag	UNP P72525
B	1490	GLU	-	expression tag	UNP P72525
B	1491	HIS	-	expression tag	UNP P72525
B	1492	HIS	-	expression tag	UNP P72525
B	1493	HIS	-	expression tag	UNP P72525
B	1494	HIS	-	expression tag	UNP P72525
B	1495	HIS	-	expression tag	UNP P72525
B	1496	HIS	-	expression tag	UNP P72525

- Molecule 2 is a DNA chain called E-site DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	7	Total	C	N	O	P	0	0	0
			140	69	27	38	6			

- Molecule 3 is a DNA chain called E-site DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	11	Total	C	N	O	P	0	0	0
			225	108	39	67	11			

- Molecule 4 is a DNA chain called E-site DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	7	Total	C	N	O	P	0	0	0
			139	68	25	40	6			

- Molecule 5 is a DNA chain called E-site DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	H	11	Total 226	C 107	N 43	O 65	P 11	0	0	0

• Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	2	Total 2	Mg 2	0	0
6	A	2	Total 2	Mg 2	0	0

• Molecule 7 is Trovafloxacin (three-letter code: TR6) (formula: $C_{20}H_{15}F_3N_4O_3$).
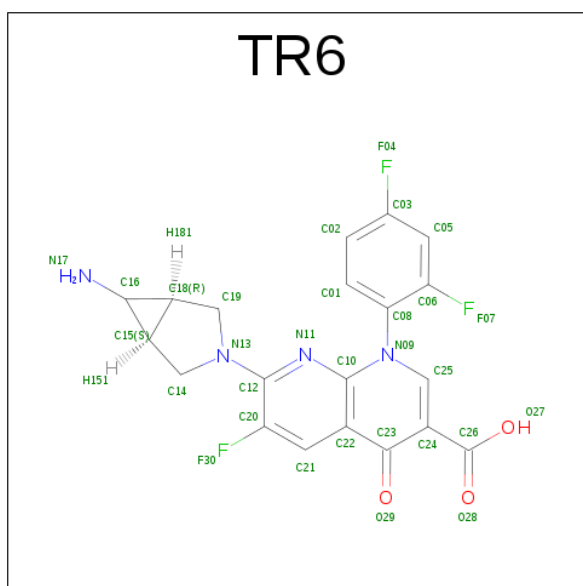
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
7	F	1	Total 44	C 20	F 3	H 14	N 4	O 3	0	0
7	H	1	Total 44	C 20	F 3	H 14	N 4	O 3	0	0

• Molecule 8 is water.

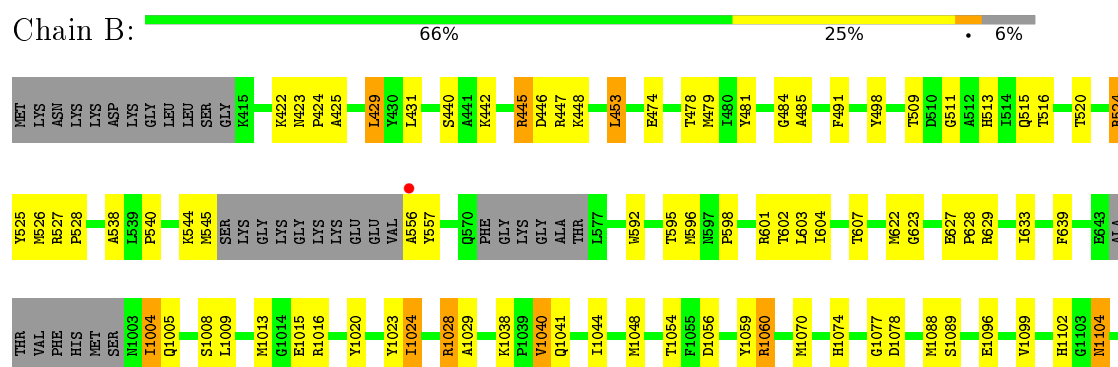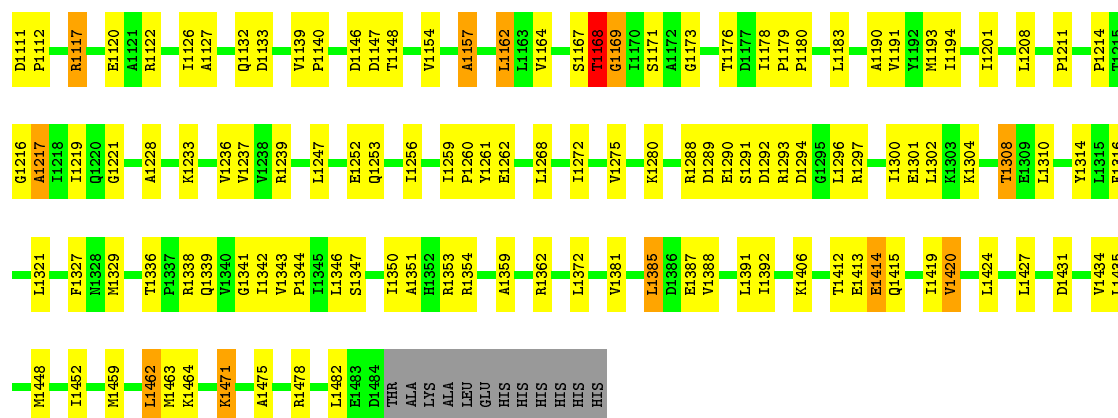| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	3	Total 3	O 3	0	0
8	B	1	Total 1	O 1	0	0

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 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 (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.

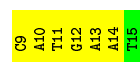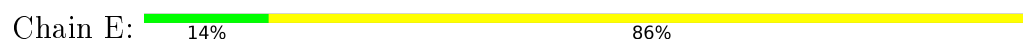- Molecule 1: DNA topoisomerase 4 subunit B,DNA topoisomerase 4 subunit A

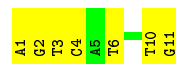- Molecule 1: DNA topoisomerase 4 subunit B,DNA topoisomerase 4 subunit A

D1111 P1112 R1117 E1120 A1121 R1122 I1126 A1127 Q1132 D1133 V1139 P1140 D1146 D1147 T1148 V1154 A1157 L1162 L1163 V1164 S1167 T1168 G1169 I1170 S1171 A1172 G1173 T1176 D1177 I1178 P1179 P1180 L1183 A1190 V1191 Y1192 M1193 I1194 I1201 L1208 P1211 P1214 T1215

G1216 A1217 I1218 I1219 Q1220 G1221 A1228 K1233 V1236 V1237 V1238 R1239 L1247 E1252 Q1253 I1256 I1259 P1260 Y1261 E1262 L1268 L1272 V1275 K1280 R1288 D1289 E1290 S1291 D1292 D1293 D1294 G1295 L1296 R1297 I1300 E1301 L1302 K1303 K1304 T1308 E1309 L1310 Y1314 L1315 F1316

L1321 F1327 N1328 M1329 T1336 P1337 R1338 Q1339 V1340 G1341 I1342 V1343 P1344 L1345 L1346 S1347 I1350 A1351 H1352 R1353 R1354 A1359 R1362 L1372 V1381 L1385 D1386 E1387 V1388 L1391 I1392 K1406 T1412 E1413 E1414 Q1415 I1419 V1420 L1424 L1427 D1431 V1434 L1435

M1448 I1452 M1459 L1462 M1463 K1464 K1471 A1475 R1478 L1482 E1483 D1484 THR ALA LYS ALA LEU GLU HIS HIS HIS HIS HIS HIS

- Molecule 2: E-site DNA

Chain E: 14% 86%

C9 A10 T11 G12 A13 A14 T15

- Molecule 3: E-site DNA

Chain F: 36% 64%

A1 G2 T3 C4 A5 T6 T10 G11

- Molecule 4: E-site DNA

Chain G: 29% 71%

C9 G10 T11 G12 C13 A14 T15

- Molecule 5: E-site DNA

Chain H: 9% 91%

G1 A2 C3 T4 A5 T6 G7 C8 A9 C10 G11

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	157.83Å 157.83Å 210.58Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	73.90 – 3.26 73.90 – 3.01	Depositor EDS
% Data completeness (in resolution range)	99.9 (73.90-3.26) 95.0 (73.90-3.01)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	3.33 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.173 , 0.207 0.171 , 0.208	Depositor DCC
R_{free} test set	2345 reflections (5.02%)	DCC
Wilson B-factor (Å^2)	75.3	Xtriage
Anisotropy	0.583	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 53.4	EDS
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Estimated twinning fraction	0.028 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11112	wwPDB-VP
Average B, all atoms (Å^2)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.17% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, TR6

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	0.23	0/5224	0.42	0/7102
1	B	0.23	0/5236	0.42	0/7123
2	E	0.51	0/157	1.08	0/241
3	F	0.48	0/251	1.22	3/385 (0.8%)
4	G	0.48	0/155	1.12	1/238 (0.4%)
5	H	0.47	0/253	1.04	0/388
All	All	0.26	0/11276	0.51	4/15477 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	6	DT	O4'-C4'-C3'	-7.99	101.21	106.00
3	F	6	DT	O4'-C1'-N1	5.38	111.77	108.00
4	G	11	DT	N3-C4-O4	5.27	123.06	119.90
3	F	10	DT	C1'-O4'-C4'	-5.01	105.09	110.10

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	5138	0	4821	196	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	5148	0	4781	188	0
2	E	140	0	78	9	0
3	F	225	0	126	8	0
4	G	139	0	78	5	0
5	H	226	0	124	13	0
6	A	2	0	0	0	0
6	B	2	0	0	0	0
7	F	30	14	14	1	0
7	H	30	14	14	0	0
8	A	3	0	0	1	0
8	B	1	0	0	0	0
All	All	11084	28	10036	404	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 404 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:B:545:MET:HB3	1:B:556:ALA:CB	1.69	1.21
1:B:545:MET:CB	1:B:556:ALA:HB3	1.71	1.20
1:B:1117:ARG:HH11	1:B:1117:ARG:HG2	1.05	1.17
1:A:1146:ASP:HB3	1:A:1148:THR:HG22	1.27	1.13
1:A:1293:ARG:HB3	1:A:1293:ARG:HH11	1.13	1.12

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	682/742 (92%)	630 (92%)	45 (7%)	7 (1%)	19	62

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	689/742 (93%)	627 (91%)	51 (7%)	11 (2%)	12	52
All	All	1371/1484 (92%)	1257 (92%)	96 (7%)	18 (1%)	15	56

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1307	ASN
1	B	1168	THR
1	B	1157	ALA
1	B	1293	ARG
1	A	1104	ASN

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	480/635 (76%)	450 (94%)	30 (6%)	22	62
1	B	474/635 (75%)	448 (94%)	26 (6%)	27	67
All	All	954/1270 (75%)	898 (94%)	56 (6%)	24	64

5 of 56 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1385	LEU
1	B	445	ARG
1	B	1385	LEU
1	A	1393	ARG
1	A	1420	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1253	GLN

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 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	#$|Z| > 2$
7	TR6	F	101	6	28,34,34	1.68	8 (28%)	30,53,53	1.85	8 (26%)
7	TR6	H	101	6	28,34,34	1.67	7 (25%)	30,53,53	1.79	10 (33%)

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
7	TR6	F	101	6	-	0/8/29/29	0/4/5/5
7	TR6	H	101	6	-	0/8/29/29	0/4/5/5

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	F	101	TR6	C12-C20	-2.80	1.38	1.42

Continued on next page...

wwPDB X-ray Structure Validation Summary Report

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	F	101	TR6	C14-C15	-2.77	1.49	1.53
7	H	101	TR6	C12-C20	-2.69	1.38	1.42
7	H	101	TR6	F04-C03	-2.48	1.30	1.36
7	F	101	TR6	F04-C03	-2.44	1.30	1.36

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	H	101	TR6	C08-N09-C10	-3.89	115.13	120.11
7	F	101	TR6	C21-C20-C12	-3.43	119.27	121.73
7	F	101	TR6	C08-N09-C10	-3.41	115.75	120.11
7	H	101	TR6	C22-C10-N11	-3.05	119.82	123.55
7	H	101	TR6	C21-C20-C12	-3.01	119.58	121.73

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	F	101	TR6	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	690/742 (92%)	-0.57	0	100	100	46, 79, 138, 173	0
1	B	695/742 (93%)	-0.60	1 (0%)	95	95	47, 77, 145, 184	0
2	E	7/7 (100%)	-0.49	0	100	100	79, 84, 108, 130	0
3	F	11/11 (100%)	-0.58	0	100	100	81, 99, 138, 144	0
4	G	7/7 (100%)	-0.41	0	100	100	72, 78, 106, 129	0
5	H	11/11 (100%)	-0.57	0	100	100	81, 97, 131, 140	0
All	All	1421/1520 (93%)	-0.58	1 (0%)	95	95	46, 79, 141, 184	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	556	ALA	2.4

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	TR6	F	101	30/30	0.95	0.20	1.02	84,109,134,154	0
7	TR6	H	101	30/30	0.94	0.19	0.19	95,120,141,165	0
6	MG	B	1502	1/1	0.33	0.29	-	84,84,84,84	0
6	MG	A	1501	1/1	0.92	0.37	-	92,92,92,92	0
6	MG	A	1502	1/1	0.70	0.33	-	90,90,90,90	0
6	MG	B	1501	1/1	0.95	0.45	-	101,101,101,101	0

6.5 Other polymers (i)

There are no such residues in this entry.