wwPDB X-ray Structure Validation Summary Report (i)

Oct 19, 2026 – 01:14 pm BST

PDB ID	:	9T01 / pdb_00009t01
Title	:	Crystal structure of SarF, the ComR paralog from S.salivarius, in complex with DNA and the synthetic peptide sBI7.
Authors	:	Nessler, S.; Gallay Li de la Sierra, I.
Deposited on	:	2025-10-16
Resolution	:	3.49 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 (i) 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 (1)) were used in the production of this report:

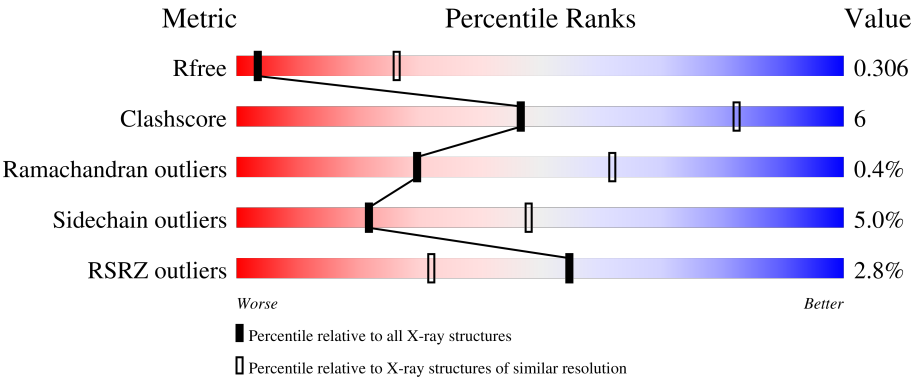
MolProbity	:	4-5-2 with Phenix2.2
Xtriage (Phenix)	:	2.2
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.51

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 3.49 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.



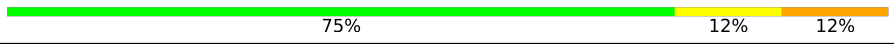
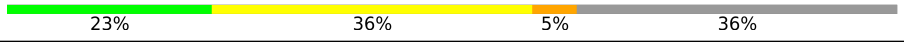
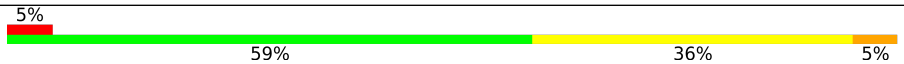
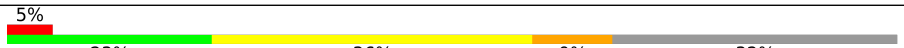
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	1085 (3.54-3.46)
Clashscore	190562	1140 (3.54-3.46)
Ramachandran outliers	187476	1113 (3.54-3.46)
Sidechain outliers	187428	1114 (3.54-3.46)
RSRZ outliers	180081	1084 (3.54-3.46)

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	316	80% 15%
1	B	316	83% 12%
1	C	316	6% 80% 11% 8%
1	D	316	81% 14%
1	E	316	6% 79% 13% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	316	
2	G	8	
2	H	8	
2	I	8	
2	J	8	
3	L	22	
3	N	22	
3	P	22	
4	K	22	
4	M	22	
4	O	22	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 17398 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 Helix-turn-helix transcriptional regulator.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	303	Total	C	N	O	S	0	0	0
			2518	1615	405	487	11			
1	B	303	Total	C	N	O	S	0	0	0
			2518	1615	405	487	11			
1	C	291	Total	C	N	O	S	0	0	0
			2424	1554	392	467	11			
1	D	303	Total	C	N	O	S	0	0	0
			2518	1615	405	487	11			
1	E	291	Total	C	N	O	S	0	0	0
			2424	1554	392	467	11			
1	F	303	Total	C	N	O	S	0	0	0
			2518	1615	405	487	11			

- Molecule 2 is a protein called synthetic peptide sBI7.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	8	Total	C	N	O	0	0	0
			69	51	9	9			
2	J	8	Total	C	N	O	0	0	0
			69	51	9	9			
2	H	8	Total	C	N	O	0	0	0
			69	51	9	9			
2	I	8	Total	C	N	O	0	0	0
			69	51	9	9			

- Molecule 3 is a DNA chain called DNA (5'-D(P*AP*AP*AP*GP*TP*GP*AP*CP*AP*TP *AP*GP*AP*TP*GP*TP*CP*CP*TP*TP*TP*T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	16	Total	C	N	O	P	0	0	0
			335	159	66	94	16			
3	L	22	Total	C	N	O	P	0	0	0
			453	217	80	134	22			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	14	Total	C	N	O	P	0	0	0
			285	138	45	88	14			

- Molecule 4 is a DNA chain called DNA (5'-D(P*AP*AP*AP*AP*GP*GP*AP*CP*AP*TP *CP*TP*AP*TP*GP*TP*CP*AP*CP*TP*TP*T)-3').

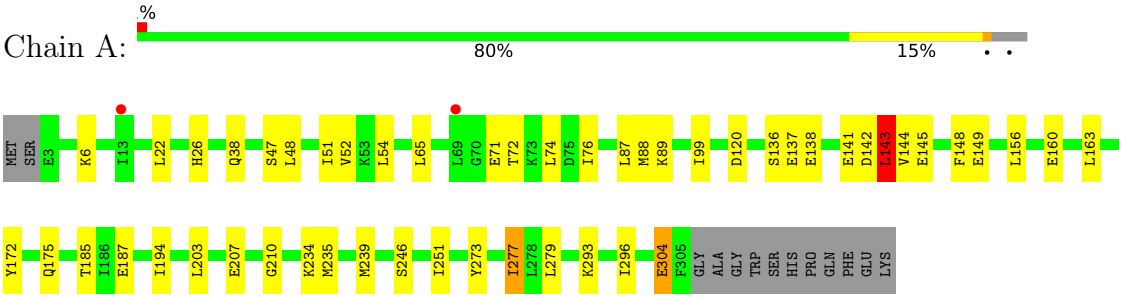
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	M	15	Total	C	N	O	P	0	0	0
			312	148	62	87	15			
4	O	18	Total	C	N	O	P	0	0	0
			366	176	61	111	18			
4	K	22	Total	C	N	O	P	0	0	0
			451	216	81	132	22			

3 Residue-property plots (i)

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.

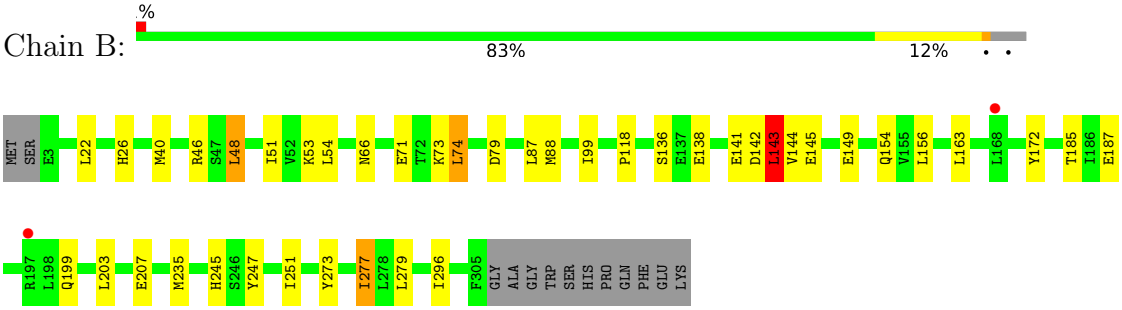
• Molecule 1: Helix-turn-helix transcriptional regulator

Chain A:



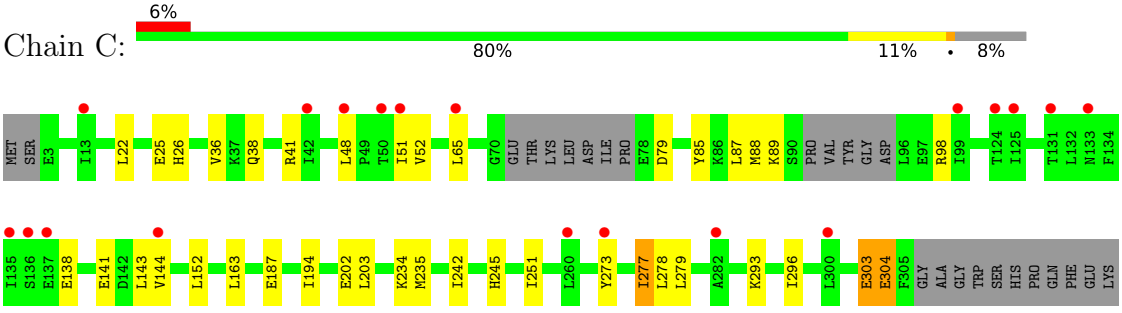
• Molecule 1: Helix-turn-helix transcriptional regulator

Chain B:



• Molecule 1: Helix-turn-helix transcriptional regulator

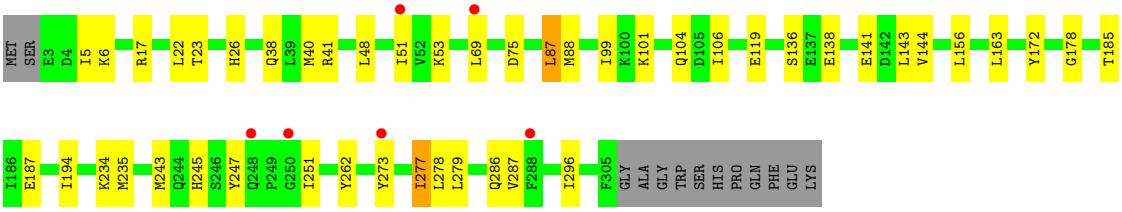
Chain C:



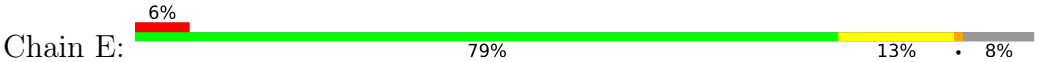
• Molecule 1: Helix-turn-helix transcriptional regulator

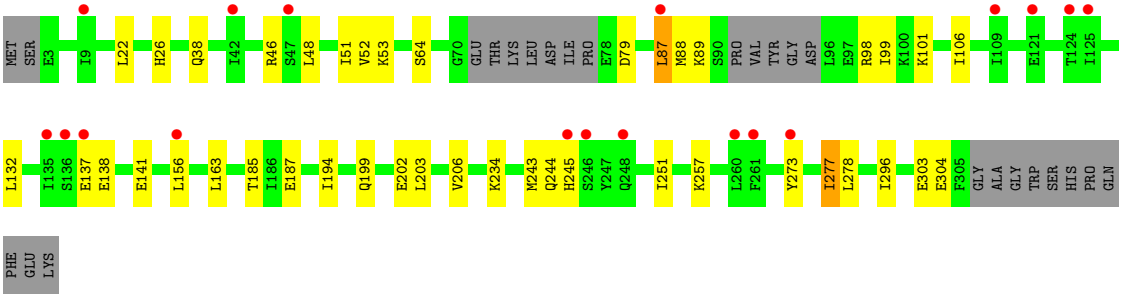
Chain D:



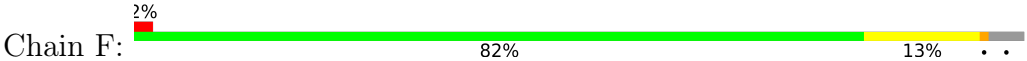


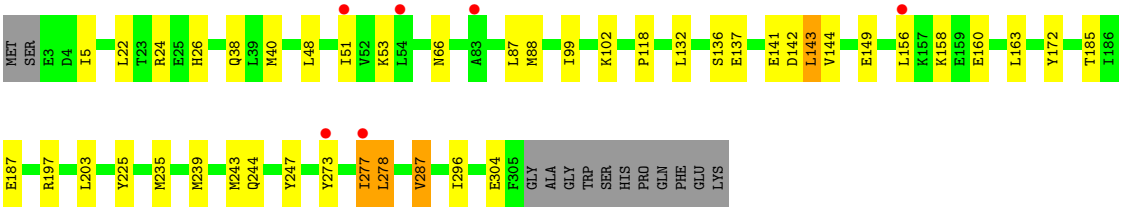
• Molecule 1: Helix-turn-helix transcriptional regulator





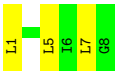
• Molecule 1: Helix-turn-helix transcriptional regulator





• Molecule 2: synthetic peptide sBI7





• Molecule 2: synthetic peptide sBI7





• Molecule 2: synthetic peptide sBI7





• Molecule 2: synthetic peptide sBI7

Chain I:

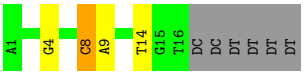




• Molecule 3: DNA (5'-D(P*AP*AP*AP*GP*TP*GP*AP*CP*AP*TP*AP*GP*AP*TP*GP*TP*CP*CP*TP*TP*TP*T)-3')

Chain P:





• Molecule 3: DNA (5'-D(P*AP*AP*AP*GP*TP*GP*AP*CP*AP*TP*AP*GP*AP*TP*GP*TP*CP*CP*TP*TP*TP*T)-3')

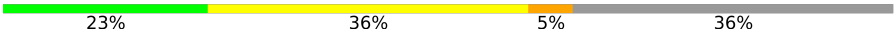
Chain L:

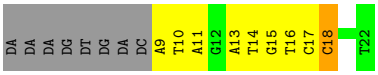




• Molecule 3: DNA (5'-D(P*AP*AP*AP*GP*TP*GP*AP*CP*AP*TP*AP*GP*AP*TP*GP*TP*CP*CP*TP*TP*TP*T)-3')

Chain N:

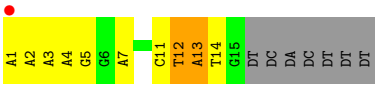




• Molecule 4: DNA (5'-D(P*AP*AP*AP*AP*GP*GP*AP*CP*AP*TP*CP*TP*AP*TP*GP*TP*CP*AP*CP*TP*TP*T)-3')

Chain M:





• Molecule 4: DNA (5'-D(P*AP*AP*AP*AP*GP*GP*AP*CP*AP*TP*CP*TP*AP*TP*GP*TP*CP*AP*CP*TP*TP*T)-3')

Chain O:





• Molecule 4: DNA (5'-D(P*AP*AP*AP*AP*GP*GP*AP*CP*AP*TP*CP*TP*AP*TP*GP*TP*CP*AP*CP*TP*TP*T)-3')

Chain K:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	91.43Å 112.02Å 160.85Å 90.00° 106.14° 90.00°	Depositor
Resolution (Å)	47.22 – 3.49 47.22 – 3.49	Depositor EDS
% Data completeness (in resolution range)	99.4 (47.22-3.49) 99.3 (47.22-3.49)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 3.48Å)	Xtriage
Refinement program	BUSTER 2.10.4 (10-JUL-2024)	Depositor
R, R_{free}	0.281 , 0.316 0.283 , 0.306	Depositor DCC
R_{free} test set	1981 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	144.3	Xtriage
Anisotropy	0.135	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 125.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Estimated twinning fraction	0.064 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	17398	wwPDB-VP
Average B, all atoms (Å^2)	172.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 37.15 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.4815e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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	0.54	0/2556	0.98	0/3435
1	B	0.53	0/2556	0.99	0/3435
1	C	0.53	0/2457	0.99	2/3295 (0.1%)
1	D	0.52	0/2556	0.98	0/3435
1	E	0.52	0/2457	0.98	0/3295
1	F	0.53	0/2556	1.01	2/3435 (0.1%)
2	G	0.77	0/72	0.99	0/97
2	H	0.53	0/72	0.99	0/97
2	I	0.59	0/72	1.02	0/97
2	J	0.76	0/72	1.26	0/97
3	L	0.86	0/507	0.90	1/779 (0.1%)
3	N	0.94	0/317	0.94	1/487 (0.2%)
3	P	0.81	0/377	0.87	1/579 (0.2%)
4	K	0.94	0/505	0.98	2/775 (0.3%)
4	M	1.10	2/351 (0.6%)	1.03	2/538 (0.4%)
4	O	0.90	0/408	0.89	0/627
All	All	0.60	2/17891 (0.0%)	0.98	11/24503 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	M	13	DA	C5'-C4'	5.83	1.55	1.51
4	M	12	DT	C3'-O3'	5.45	1.51	1.44

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	M	13	DA	O5'-C5'-C4'	6.91	114.93	109.40
4	K	22	DT	O5'-C5'-C4'	6.46	114.57	109.40
4	K	7	DA	O5'-C5'-C4'	5.94	114.15	109.40
3	P	8	DC	O4'-C4'-C3'	-5.67	102.60	106.00
3	L	8	DC	O4'-C4'-C3'	-5.40	102.76	106.00

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	2518	0	2554	35	0
1	B	2518	0	2554	35	0
1	C	2424	0	2461	27	0
1	D	2518	0	2554	41	0
1	E	2424	0	2461	26	0
1	F	2518	0	2554	33	0
2	G	69	0	75	1	0
2	H	69	0	75	1	0
2	I	69	0	75	1	0
2	J	69	0	75	1	0
3	L	453	0	251	16	0
3	N	285	0	162	14	0
3	P	335	0	181	10	0
4	K	451	0	250	13	0
4	M	312	0	169	19	0
4	O	366	0	206	6	0
All	All	17398	0	16657	191	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 6.

The worst 5 of 191 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:C:38:GLN:HG2	4:M:13:DA:H3'	1.49	0.93
1:D:41:ARG:NH1	3:N:15:DG:H22	1.69	0.91
1:A:239:MET:HE1	1:B:247:TYR:HB2	1.55	0.89
1:E:52:VAL:HG22	1:F:5:ILE:HD11	1.55	0.89
1:C:293:LYS:HD3	1:C:304:GLU:HG3	1.57	0.86

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	301/316 (95%)	285 (95%)	14 (5%)	2 (1%)	18 51
1	B	301/316 (95%)	289 (96%)	11 (4%)	1 (0%)	36 67
1	C	285/316 (90%)	273 (96%)	12 (4%)	0	100 100
1	D	301/316 (95%)	284 (94%)	15 (5%)	2 (1%)	18 51
1	E	285/316 (90%)	274 (96%)	11 (4%)	0	100 100
1	F	301/316 (95%)	286 (95%)	13 (4%)	2 (1%)	18 51
2	G	6/8 (75%)	6 (100%)	0	0	100 100
2	H	6/8 (75%)	6 (100%)	0	0	100 100
2	I	6/8 (75%)	5 (83%)	1 (17%)	0	100 100
2	J	6/8 (75%)	6 (100%)	0	0	100 100
All	All	1798/1928 (93%)	1714 (95%)	77 (4%)	7 (0%)	30 62

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	75	ASP
1	F	243	MET
1	A	76	ILE
1	A	143	LEU
1	B	143	LEU

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	280/290 (97%)	265 (95%)	15 (5%)	20 46
1	B	280/290 (97%)	266 (95%)	14 (5%)	22 48
1	C	269/290 (93%)	257 (96%)	12 (4%)	24 50
1	D	280/290 (97%)	272 (97%)	8 (3%)	37 60
1	E	269/290 (93%)	255 (95%)	14 (5%)	21 47
1	F	280/290 (97%)	266 (95%)	14 (5%)	22 48
2	G	7/7 (100%)	5 (71%)	2 (29%)	0 2
2	H	7/7 (100%)	6 (86%)	1 (14%)	3 18
2	I	7/7 (100%)	5 (71%)	2 (29%)	0 2
2	J	7/7 (100%)	5 (71%)	2 (29%)	0 2
All	All	1686/1768 (95%)	1602 (95%)	84 (5%)	22 48

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	187	GLU
1	F	187	GLU
1	E	244	GLN
1	F	66	ASN
1	F	287	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	55	GLN
1	D	104	GLN
1	F	55	GLN
1	D	286	GLN
1	E	55	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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

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, 95^{th} 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	303/316 (95%)	-0.18	2 (0%) 84 63	106, 138, 166, 175	0
1	B	303/316 (95%)	-0.20	2 (0%) 84 63	107, 146, 166, 174	0
1	C	291/316 (92%)	0.21	19 (6%) 25 15	119, 240, 258, 263	0
1	D	303/316 (95%)	-0.08	6 (1%) 65 39	120, 157, 226, 234	0
1	E	291/316 (92%)	0.23	18 (6%) 26 16	113, 219, 264, 271	0
1	F	303/316 (95%)	0.04	6 (1%) 65 39	105, 142, 223, 230	0
2	G	8/8 (100%)	-0.15	0 100 100	111, 112, 121, 125	0
2	H	8/8 (100%)	-0.17	0 100 100	109, 111, 117, 120	0
2	I	8/8 (100%)	-0.36	0 100 100	112, 114, 122, 126	0
2	J	8/8 (100%)	-0.31	0 100 100	112, 114, 117, 120	0
3	L	22/22 (100%)	-0.24	0 100 100	170, 178, 199, 200	0
3	N	14/22 (63%)	-0.37	0 100 100	149, 188, 214, 216	0
3	P	16/22 (72%)	-0.42	0 100 100	154, 188, 211, 214	0
4	K	22/22 (100%)	0.02	1 (4%) 38 20	169, 177, 200, 202	0
4	M	15/22 (68%)	-0.13	1 (6%) 24 15	151, 182, 220, 222	0
4	O	18/22 (81%)	-0.27	0 100 100	154, 182, 234, 234	0
All	All	1933/2060 (93%)	-0.02	55 (2%) 55 32	105, 154, 253, 271	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	42	ILE	4.7
1	E	260	LEU	4.5
1	F	51	ILE	3.7
1	C	136	SER	3.7
1	D	51	ILE	3.7

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 oligosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.