Full wwPDB X-ray Structure Validation Report (i)

Jan 8, 2024 – 05:51 am GMT

PDB ID	:	5MT9
Title	:	Human insulin in complex with serotonin and arginine
Authors	:	Brzozowski, A.M.; Turkenburg, J.P.; Jiracek, J.; Zakova, L.
Deposited on	:	2017-01-07
Resolution	:	1.88 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i)) were used in the production of this report:

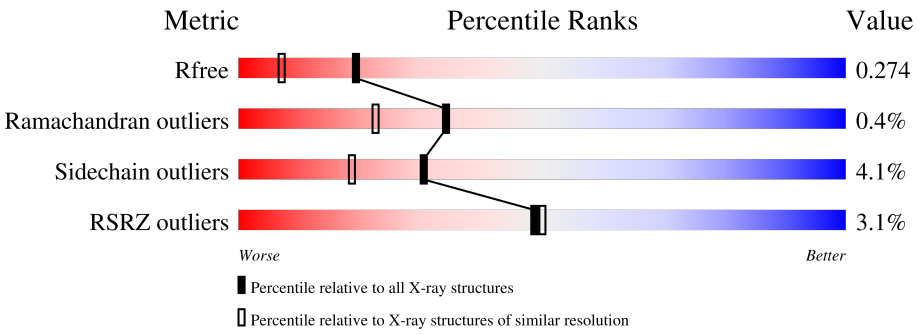
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.88 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	9470 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

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	21	14%, 86%, 10%, 5%
1	C	21	5%, 86%, 10%, 5%
1	E	21	14%, 100%
1	G	21	5%, 100%
1	I	21	5%, 95%, 5%
1	K	21	5%, 81%, 14%, 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	M	21	95% 5%
1	O	21	100%
1	Q	21	5% 95% 5%
1	S	21	100%
1	U	21	90% 5% 5%
1	W	21	95% 5%
1	Y	21	5% 95% 5%
1	a	21	100%
1	c	21	5% 100%
1	e	21	95% 5%
2	B	30	83% 17%
2	D	30	3% 93% 7%
2	F	30	83% 17%
2	H	30	93%
2	J	30	3% 83% 7% 7%
2	L	30	77% 17%
2	N	30	83% 17%
2	P	30	3% 90% 7%
2	R	30	3% 73% 7% 20%
2	T	30	7% 90% 7%
2	V	30	80% 20%
2	X	30	7% 83% 10% 7%
2	Z	30	80% 20%
2	b	30	90% 7%
2	d	30	80% 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	f	30	7% 83% 10%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	ARG	B	103	-	-	-	X
5	ARG	F	103	-	-	-	X

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 6235 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 Insulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	20	147	89	24	30	4	0	0	0
1	C	20	152	91	24	33	4	0	0	0
1	E	21	155	95	24	32	4	0	0	0
1	G	21	163	99	25	35	4	0	0	0
1	I	21	161	98	25	34	4	0	0	0
1	K	20	144	87	23	30	4	0	0	0
1	M	21	158	97	25	32	4	0	0	0
1	O	21	156	93	25	34	4	0	0	0
1	Q	20	140	88	21	27	4	0	0	0
1	S	21	159	97	25	33	4	0	0	0
1	U	20	146	91	22	29	4	0	0	0
1	W	21	159	97	25	33	4	0	0	0
1	Y	20	145	91	22	28	4	0	0	0
1	a	21	163	99	25	35	4	0	0	0
1	c	21	153	93	24	32	4	0	0	0
1	e	21	163	99	25	35	4	0	0	0

- Molecule 2 is a protein called Insulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	25	186	124	27	33	2	0	0	0
2	D	28	211	138	37	34	2	0	0	0
2	F	25	190	125	30	33	2	0	0	0
2	H	29	230	154	39	35	2	0	1	0
2	J	28	214	143	36	33	2	0	0	0
2	L	25	193	127	31	33	2	0	0	0
2	N	25	195	128	32	33	2	0	0	0
2	P	28	217	141	40	34	2	0	1	0
2	R	24	182	121	29	30	2	0	0	0
2	T	29	219	142	38	37	2	0	0	0
2	V	24	175	114	29	30	2	0	0	0
2	X	28	215	140	37	36	2	0	0	0
2	Z	24	175	114	29	30	2	0	0	0
2	b	29	219	142	38	37	2	0	0	0
2	d	24	179	116	29	32	2	0	0	0
2	f	29	214	137	38	37	2	0	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	B	1	1	1	0	0
3	D	1	1	1	0	0
3	F	1	1	1	0	0

Continued on next page...

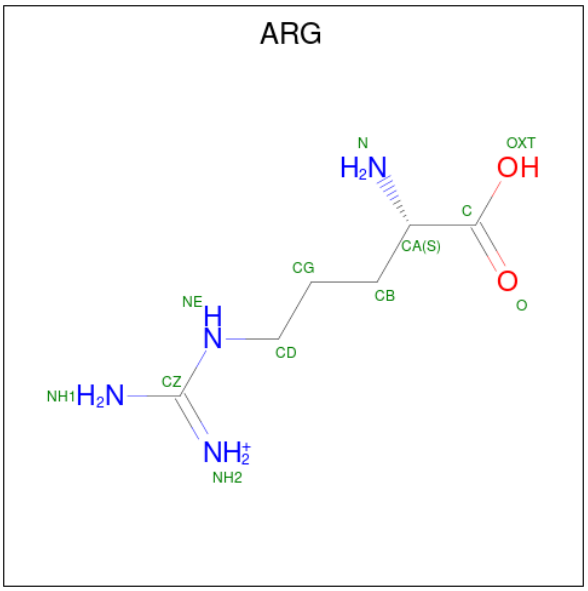
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	H	1	Total Zn 1 1	0	0
3	R	1	Total Zn 1 1	0	0
3	T	1	Total Zn 1 1	0	0
3	d	1	Total Zn 1 1	0	0
3	f	1	Total Zn 1 1	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

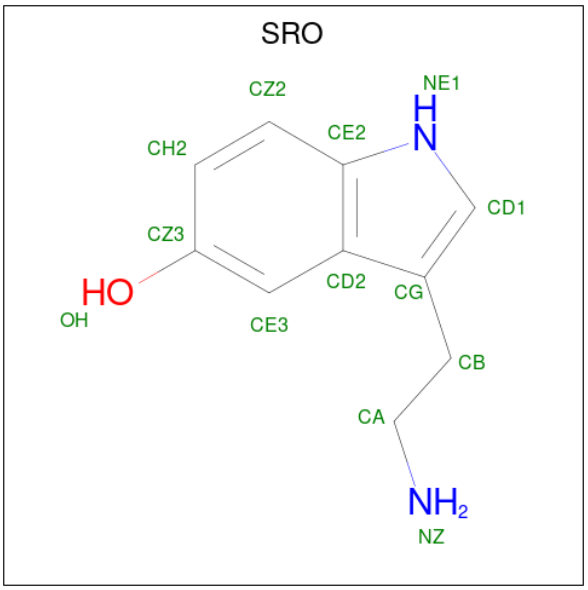
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Cl 1 1	0	0
4	D	1	Total Cl 1 1	0	0
4	F	1	Total Cl 1 1	0	0
4	P	1	Total Cl 1 1	0	0
4	T	1	Total Cl 1 1	0	0
4	Z	1	Total Cl 1 1	0	0
4	d	1	Total Cl 1 1	0	0
4	f	1	Total Cl 1 1	0	0

- Molecule 5 is ARGININE (three-letter code: ARG) (formula: $C_6H_{15}N_4O_2$).



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

- Molecule 6 is SEROTONIN (three-letter code: SRO) (formula: $C_{10}H_{12}N_2O$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	C	1	13	10	2	1	0	0
6	E	1	13	10	2	1	0	0
6	G	1	13	10	2	1	0	0
6	I	1	13	10	2	1	0	0
6	K	1	13	10	2	1	0	0
6	M	1	13	10	2	1	0	0
6	O	1	13	10	2	1	0	0
6	Q	1	13	10	2	1	0	0
6	R	1	13	10	2	1	0	0
6	S	1	13	10	2	1	0	0
6	S	1	13	10	2	1	0	0
6	W	1	13	10	2	1	0	0
6	Y	1	13	10	2	1	0	0
6	a	1	13	10	2	1	0	0
6	c	1	13	10	2	1	0	0
6	e	1	13	10	2	1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	2	2	2	0	0
7	B	1	1	1	0	0
7	C	3	3	3	0	0
7	D	9	9	9	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	E	5	Total O 5 5	0	0
7	F	6	Total O 6 6	0	0
7	G	17	Total O 17 17	0	0
7	H	19	Total O 19 19	0	0
7	I	11	Total O 11 11	0	0
7	J	18	Total O 18 18	0	0
7	K	2	Total O 2 2	0	0
7	L	6	Total O 6 6	0	0
7	M	7	Total O 7 7	0	0
7	N	8	Total O 8 8	0	0
7	O	15	Total O 15 15	0	0
7	P	13	Total O 13 13	0	0
7	Q	2	Total O 2 2	0	0
7	R	8	Total O 8 8	0	0
7	S	1	Total O 1 1	0	0
7	T	15	Total O 15 15	0	0
7	V	6	Total O 6 6	0	0
7	W	9	Total O 9 9	0	0
7	X	12	Total O 12 12	0	0
7	Z	11	Total O 11 11	0	0
7	a	15	Total O 15 15	0	0

Continued on next page...

Continued from previous page...

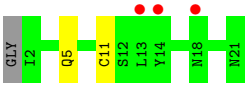
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	b	19	19	19	0	0
			Total	O		
7	c	10	10	10	0	0
			Total	O		
7	d	14	14	14	0	0
			Total	O		
7	e	9	9	9	0	0
			Total	O		
7	f	12	12	12	0	0

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.

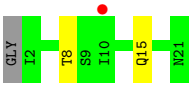
- Molecule 1: Insulin

Chain A:



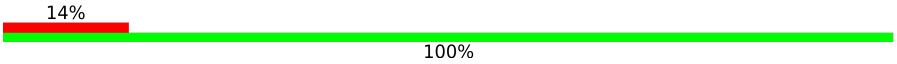
- Molecule 1: Insulin

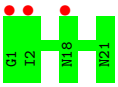
Chain C:



- Molecule 1: Insulin

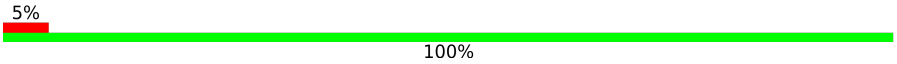
Chain E:





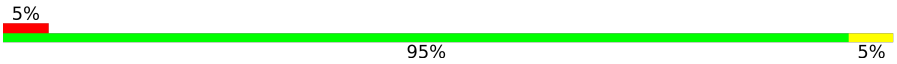
- Molecule 1: Insulin

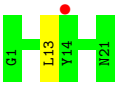
Chain G:



- Molecule 1: Insulin

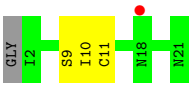
Chain I:



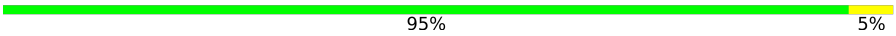


- Molecule 1: Insulin

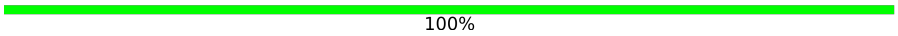
Chain K:



• Molecule 1: Insulin

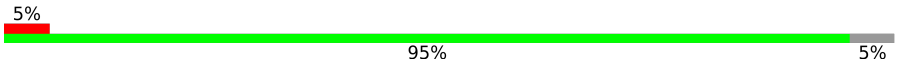
Chain M: 

• Molecule 1: Insulin

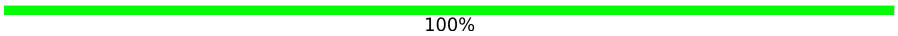
Chain O: 

There are no outlier residues recorded for this chain.

• Molecule 1: Insulin

Chain Q: 

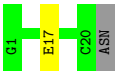
• Molecule 1: Insulin

Chain S: 

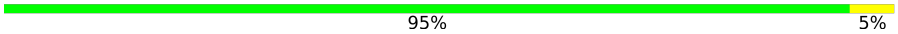
There are no outlier residues recorded for this chain.

• Molecule 1: Insulin

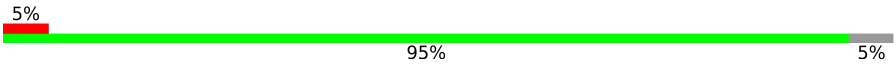
Chain U:

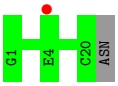


• Molecule 1: Insulin

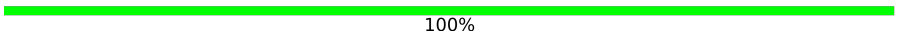
Chain W: 

• Molecule 1: Insulin

Chain Y: 

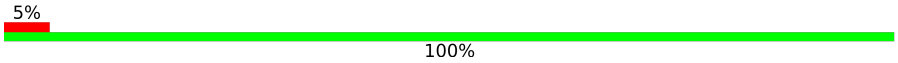


• Molecule 1: Insulin

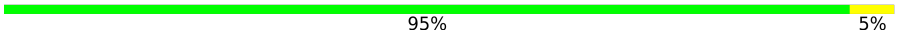
Chain a: 

There are no outlier residues recorded for this chain.

• Molecule 1: Insulin

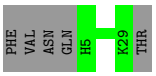
Chain c: 

• Molecule 1: Insulin

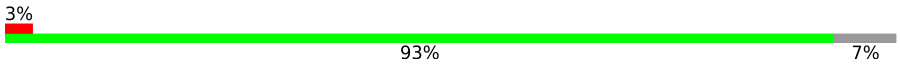
Chain e: 

• Molecule 2: Insulin

Chain B:

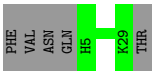


• Molecule 2: Insulin

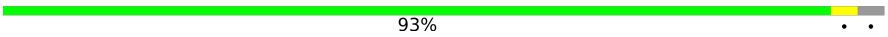
Chain D: 

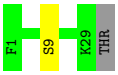
• Molecule 2: Insulin

Chain F:



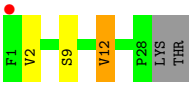
• Molecule 2: Insulin

Chain H: 



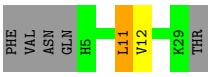
- Molecule 2: Insulin

Chain J:



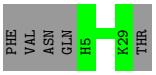
- Molecule 2: Insulin

Chain L:



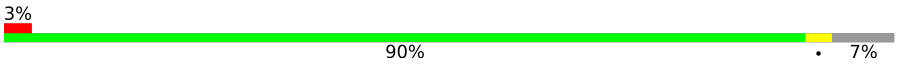
- Molecule 2: Insulin

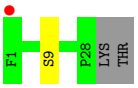
Chain N:



- Molecule 2: Insulin

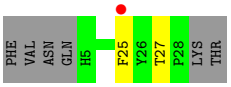
Chain P:





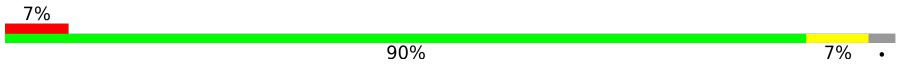
- Molecule 2: Insulin

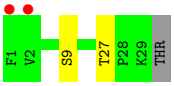
Chain R:



- Molecule 2: Insulin

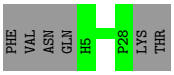
Chain T:





- Molecule 2: Insulin

Chain V:

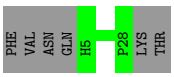


- Molecule 2: Insulin

Chain X:

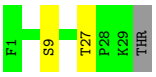
• Molecule 2: Insulin

Chain Z:



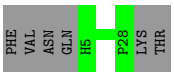
• Molecule 2: Insulin

Chain b:



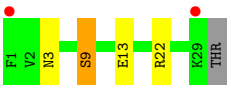
• Molecule 2: Insulin

Chain d:



• Molecule 2: Insulin

Chain f:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	159.31Å 159.31Å 76.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.99 – 1.88 45.99 – 1.88	Depositor EDS
% Data completeness (in resolution range)	98.3 (45.99-1.88) 98.4 (45.99-1.88)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.89 (at 1.88Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.203 , 0.269 0.212 , 0.274	Depositor DCC
R_{free} test set	2867 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	31.4	Xtriage
Anisotropy	0.260	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 48.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.52$, $< L^2 > = 0.35$	Xtriage
Estimated twinning fraction	0.024 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6235	wwPDB-VP
Average B, all atoms (Å^2)	41.0	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 72.02 % 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 2.3711e-06. 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CL, ZN, SRO

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	0.68	0/148	0.76	0/200
1	C	0.78	0/152	0.80	0/204
1	E	0.81	0/156	0.71	0/210
1	G	1.11	0/164	0.81	0/220
1	I	0.88	0/162	0.79	0/217
1	K	0.72	0/145	0.66	0/195
1	M	0.88	0/159	0.75	0/214
1	O	0.96	0/156	0.85	0/209
1	Q	0.79	0/141	0.84	0/191
1	S	0.73	0/160	0.76	0/215
1	U	0.65	0/147	0.82	0/199
1	W	0.93	0/160	0.78	0/215
1	Y	0.72	0/146	0.76	0/198
1	a	0.89	0/164	0.88	0/220
1	c	0.75	0/154	0.74	0/206
1	e	0.87	0/164	0.78	0/220
2	B	0.84	0/191	0.91	0/259
2	D	0.87	0/217	0.74	0/295
2	F	0.99	0/195	1.05	0/264
2	H	0.92	1/240 (0.4%)	0.84	0/325
2	J	0.99	1/221 (0.5%)	1.05	1/300 (0.3%)
2	L	0.89	0/198	1.04	1/266 (0.4%)
2	N	0.99	0/201	0.96	0/272
2	P	1.03	1/226 (0.4%)	0.82	0/305
2	R	0.95	0/187	0.82	0/253
2	T	0.92	1/225 (0.4%)	0.86	0/305
2	V	0.90	0/179	0.98	0/243
2	X	0.99	1/221 (0.5%)	0.86	0/300
2	Z	0.84	0/179	0.81	0/243
2	b	1.00	1/225 (0.4%)	0.82	0/305
2	d	1.07	0/183	0.96	0/248
2	f	0.99	1/219 (0.5%)	1.03	2/298 (0.7%)

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#|Z| >5	RMSZ	#|Z| >5
All	All	0.90	7/5785 (0.1%)	0.86	4/7814 (0.1%)

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	J	9	SER	CB-OG	-7.11	1.33	1.42
2	b	9	SER	CB-OG	-6.86	1.33	1.42
2	f	9	SER	CB-OG	-6.06	1.34	1.42
2	P	9	SER	CB-OG	-5.60	1.34	1.42
2	T	9	SER	CB-OG	-5.35	1.35	1.42
2	X	9	SER	CB-OG	-5.24	1.35	1.42
2	H	9	SER	CB-OG	-5.07	1.35	1.42

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	f	22	ARG	NE-CZ-NH2	-7.79	116.41	120.30
2	f	22	ARG	NE-CZ-NH1	6.77	123.69	120.30
2	J	12	VAL	CG1-CB-CG2	6.50	121.30	110.90
2	L	11	LEU	CB-CG-CD2	5.11	119.69	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	18/21 (86%)	17 (94%)	1 (6%)	0	100	100
1	C	18/21 (86%)	17 (94%)	1 (6%)	0	100	100
1	E	19/21 (90%)	18 (95%)	1 (5%)	0	100	100
1	G	19/21 (90%)	19 (100%)	0	0	100	100
1	I	19/21 (90%)	19 (100%)	0	0	100	100
1	K	18/21 (86%)	15 (83%)	1 (6%)	2 (11%)	0	0
1	M	19/21 (90%)	19 (100%)	0	0	100	100
1	O	19/21 (90%)	19 (100%)	0	0	100	100
1	Q	18/21 (86%)	18 (100%)	0	0	100	100
1	S	19/21 (90%)	19 (100%)	0	0	100	100
1	U	18/21 (86%)	18 (100%)	0	0	100	100
1	W	19/21 (90%)	19 (100%)	0	0	100	100
1	Y	18/21 (86%)	17 (94%)	1 (6%)	0	100	100
1	a	19/21 (90%)	19 (100%)	0	0	100	100
1	c	19/21 (90%)	19 (100%)	0	0	100	100
1	e	19/21 (90%)	19 (100%)	0	0	100	100
2	B	23/30 (77%)	22 (96%)	1 (4%)	0	100	100
2	D	26/30 (87%)	24 (92%)	2 (8%)	0	100	100
2	F	23/30 (77%)	21 (91%)	2 (9%)	0	100	100
2	H	28/30 (93%)	28 (100%)	0	0	100	100
2	J	26/30 (87%)	24 (92%)	1 (4%)	1 (4%)	3	0
2	L	23/30 (77%)	23 (100%)	0	0	100	100
2	N	23/30 (77%)	23 (100%)	0	0	100	100
2	P	27/30 (90%)	27 (100%)	0	0	100	100
2	R	22/30 (73%)	22 (100%)	0	0	100	100
2	T	27/30 (90%)	27 (100%)	0	0	100	100
2	V	22/30 (73%)	22 (100%)	0	0	100	100
2	X	26/30 (87%)	26 (100%)	0	0	100	100
2	Z	22/30 (73%)	22 (100%)	0	0	100	100
2	b	27/30 (90%)	27 (100%)	0	0	100	100
2	d	22/30 (73%)	22 (100%)	0	0	100	100
2	f	27/30 (90%)	27 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	692/816 (85%)	678 (98%)	11 (2%)	3 (0%)	34 22

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	K	10	ILE
1	K	9	SER
2	J	2	VAL

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	16/20 (80%)	14 (88%)	2 (12%)	4 1
1	C	19/20 (95%)	17 (90%)	2 (10%)	7 2
1	E	18/20 (90%)	18 (100%)	0	100 100
1	G	20/20 (100%)	20 (100%)	0	100 100
1	I	19/20 (95%)	18 (95%)	1 (5%)	22 11
1	K	15/20 (75%)	14 (93%)	1 (7%)	16 6
1	M	18/20 (90%)	17 (94%)	1 (6%)	21 10
1	O	19/20 (95%)	19 (100%)	0	100 100
1	Q	15/20 (75%)	15 (100%)	0	100 100
1	S	19/20 (95%)	19 (100%)	0	100 100
1	U	17/20 (85%)	16 (94%)	1 (6%)	19 9
1	W	19/20 (95%)	18 (95%)	1 (5%)	22 11
1	Y	16/20 (80%)	16 (100%)	0	100 100
1	a	20/20 (100%)	20 (100%)	0	100 100
1	c	18/20 (90%)	18 (100%)	0	100 100
1	e	20/20 (100%)	19 (95%)	1 (5%)	24 13
2	B	18/26 (69%)	18 (100%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	21/26 (81%)	21 (100%)	0	100	100
2	F	19/26 (73%)	19 (100%)	0	100	100
2	H	24/26 (92%)	24 (100%)	0	100	100
2	J	21/26 (81%)	20 (95%)	1 (5%)	25	14
2	L	20/26 (77%)	18 (90%)	2 (10%)	7	2
2	N	20/26 (77%)	20 (100%)	0	100	100
2	P	22/26 (85%)	22 (100%)	0	100	100
2	R	18/26 (69%)	16 (89%)	2 (11%)	6	1
2	T	22/26 (85%)	21 (96%)	1 (4%)	27	16
2	V	17/26 (65%)	17 (100%)	0	100	100
2	X	22/26 (85%)	20 (91%)	2 (9%)	9	3
2	Z	17/26 (65%)	17 (100%)	0	100	100
2	b	22/26 (85%)	21 (96%)	1 (4%)	27	16
2	d	18/26 (69%)	18 (100%)	0	100	100
2	f	21/26 (81%)	18 (86%)	3 (14%)	3	1
All	All	610/736 (83%)	588 (96%)	22 (4%)	30	23

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	11	CYS
1	C	8	THR
1	C	15	GLN
1	I	13	LEU
2	J	12	VAL
1	K	11	CYS
2	L	11	LEU
2	L	12	VAL
1	M	8	THR
2	R	25	PHE
2	R	27	THR
2	T	27	THR
1	U	17	GLU
1	W	21	ASN
2	X	13	GLU
2	X	25	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	b	27	THR
1	e	9	SER
2	f	3	ASN
2	f	9	SER
2	f	13	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	18	ASN
1	C	18	ASN
1	C	21	ASN
2	H	5	HIS
1	I	5	GLN
1	I	18	ASN
1	K	5	GLN
1	K	15	GLN
1	S	5	GLN
1	W	5	GLN
1	W	21	ASN
1	a	21	ASN
2	b	5	HIS
1	e	21	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)

Of 36 ligands modelled in this entry, 16 are monoatomic - leaving 20 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	$\#\|Z\| > 2$
6	SRO	S	102	-	12,14,14	1.22	1 (8%)	12,19,19	1.20	1 (8%)
6	SRO	R	102	-	12,14,14	1.15	0	12,19,19	1.95	3 (25%)
6	SRO	O	101	-	12,14,14	1.37	3 (25%)	12,19,19	1.25	1 (8%)
5	ARG	B	103	-	10,11,11	0.85	0	11,13,13	1.54	2 (18%)
6	SRO	Y	101	-	12,14,14	1.28	2 (16%)	12,19,19	1.24	2 (16%)
6	SRO	Q	101	-	12,14,14	1.39	2 (16%)	12,19,19	1.73	3 (25%)
6	SRO	W	101	-	12,14,14	1.39	2 (16%)	12,19,19	1.23	1 (8%)
5	ARG	L	101	-	10,11,11	0.71	0	11,13,13	1.57	2 (18%)
6	SRO	E	101	-	12,14,14	1.22	1 (8%)	12,19,19	1.76	4 (33%)
6	SRO	K	101	-	12,14,14	1.34	3 (25%)	12,19,19	1.43	3 (25%)
6	SRO	M	101	-	12,14,14	1.07	1 (8%)	12,19,19	1.69	3 (25%)
6	SRO	c	101	-	12,14,14	1.19	2 (16%)	12,19,19	1.43	2 (16%)
6	SRO	I	101	-	12,14,14	1.09	1 (8%)	12,19,19	1.17	1 (8%)
6	SRO	G	101	-	12,14,14	1.10	0	12,19,19	1.59	4 (33%)
5	ARG	N	101	-	10,11,11	0.71	1 (10%)	11,13,13	1.29	2 (18%)
5	ARG	F	103	-	10,11,11	0.84	0	11,13,13	1.81	2 (18%)
6	SRO	S	101	-	12,14,14	1.41	2 (16%)	12,19,19	1.81	3 (25%)
6	SRO	C	101	-	12,14,14	1.11	0	12,19,19	1.21	1 (8%)
6	SRO	a	101	-	12,14,14	1.35	2 (16%)	12,19,19	1.31	2 (16%)
6	SRO	e	101	-	12,14,14	1.57	4 (33%)	12,19,19	1.14	1 (8%)

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
6	SRO	S	102	-	-	0/3/3/3	0/2/2/2
6	SRO	R	102	-	-	2/3/3/3	0/2/2/2
6	SRO	O	101	-	-	0/3/3/3	0/2/2/2

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	ARG	B	103	-	-	6/11/11/11	-
6	SRO	Y	101	-	-	1/3/3/3	0/2/2/2
6	SRO	Q	101	-	-	0/3/3/3	0/2/2/2
6	SRO	W	101	-	-	0/3/3/3	0/2/2/2
5	ARG	L	101	-	-	7/11/11/11	-
6	SRO	E	101	-	-	1/3/3/3	0/2/2/2
6	SRO	K	101	-	-	1/3/3/3	0/2/2/2
6	SRO	M	101	-	-	2/3/3/3	0/2/2/2
6	SRO	c	101	-	-	0/3/3/3	0/2/2/2
6	SRO	I	101	-	-	0/3/3/3	0/2/2/2
6	SRO	G	101	-	-	0/3/3/3	0/2/2/2
5	ARG	N	101	-	-	5/11/11/11	-
5	ARG	F	103	-	-	7/11/11/11	-
6	SRO	S	101	-	-	1/3/3/3	0/2/2/2
6	SRO	C	101	-	-	0/3/3/3	0/2/2/2
6	SRO	a	101	-	-	1/3/3/3	0/2/2/2
6	SRO	e	101	-	-	1/3/3/3	0/2/2/2

All (27) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	K	101	SRO	CZ2-CH2	2.82	1.42	1.36
6	E	101	SRO	CZ2-CH2	2.82	1.42	1.36
6	W	101	SRO	CE3-CD2	-2.73	1.36	1.42
6	e	101	SRO	CZ2-CH2	2.73	1.42	1.36
6	e	101	SRO	CD2-CE2	2.49	1.49	1.42
6	O	101	SRO	CZ2-CH2	2.43	1.41	1.36
6	Y	101	SRO	CD2-CE2	2.42	1.49	1.42
6	e	101	SRO	CD1-CG	2.42	1.43	1.37
6	O	101	SRO	CD2-CE2	2.36	1.49	1.42
6	a	101	SRO	CZ2-CE2	-2.33	1.37	1.41
6	Q	101	SRO	CZ2-CH2	2.28	1.41	1.36
6	Q	101	SRO	CD2-CE2	2.28	1.48	1.42
6	W	101	SRO	CZ2-CH2	2.28	1.41	1.36
6	K	101	SRO	CE3-CZ3	2.22	1.41	1.37
6	M	101	SRO	CZ2-CE2	-2.22	1.38	1.41
6	S	101	SRO	CH2-CZ3	2.22	1.43	1.38
6	Y	101	SRO	CE3-CZ3	2.20	1.41	1.37
6	c	101	SRO	CZ2-CH2	2.16	1.41	1.36
6	e	101	SRO	CH2-CZ3	2.12	1.43	1.38
6	K	101	SRO	CD2-CE2	2.09	1.48	1.42

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	I	101	SRO	CD2-CE2	2.09	1.48	1.42
6	S	102	SRO	CD2-CE2	2.09	1.48	1.42
6	a	101	SRO	CH2-CZ3	2.09	1.42	1.38
6	S	101	SRO	CD2-CE2	2.08	1.48	1.42
6	O	101	SRO	CE3-CD2	-2.03	1.38	1.42
6	c	101	SRO	CD2-CE2	2.02	1.48	1.42
5	N	101	ARG	OXT-C	-2.01	1.24	1.30

All (43) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	R	102	SRO	CZ3-CE3-CD2	-4.29	117.64	120.64
6	S	101	SRO	CZ3-CE3-CD2	-4.29	117.64	120.64
6	M	101	SRO	CZ3-CE3-CD2	-4.26	117.66	120.64
5	F	103	ARG	OXT-C-O	-4.04	114.91	124.09
5	B	103	ARG	OXT-C-O	-3.78	115.50	124.09
6	Q	101	SRO	CZ3-CE3-CD2	-3.67	118.08	120.64
6	E	101	SRO	CE3-CD2-CE2	3.47	123.01	118.26
6	R	102	SRO	CA-CB-CG	-3.41	105.16	112.93
6	E	101	SRO	CZ3-CE3-CD2	-3.31	118.33	120.64
5	L	101	ARG	OXT-C-O	-3.27	116.66	124.09
5	L	101	ARG	OXT-C-CA	3.24	124.43	113.38
6	c	101	SRO	CZ3-CE3-CD2	-3.15	118.43	120.64
6	S	102	SRO	CZ3-CE3-CD2	-2.87	118.64	120.64
6	Q	101	SRO	CE3-CD2-CE2	2.83	122.13	118.26
5	N	101	ARG	OXT-C-O	-2.82	117.69	124.09
6	Y	101	SRO	CZ3-CE3-CD2	-2.81	118.68	120.64
5	N	101	ARG	OXT-C-CA	2.74	122.71	113.38
6	R	102	SRO	CE3-CD2-CE2	2.69	121.94	118.26
6	K	101	SRO	CE3-CD2-CE2	2.67	121.92	118.26
5	B	103	ARG	OXT-C-CA	2.64	122.37	113.38
6	a	101	SRO	CE3-CD2-CE2	2.57	121.77	118.26
6	c	101	SRO	CE3-CD2-CE2	2.50	121.68	118.26
6	S	101	SRO	CH2-CZ3-CE3	2.47	123.55	120.39
6	I	101	SRO	CZ3-CE3-CD2	-2.46	118.92	120.64
6	M	101	SRO	CH2-CZ3-CE3	2.44	123.51	120.39
6	G	101	SRO	CE3-CD2-CE2	2.42	121.57	118.26
6	G	101	SRO	CZ2-CE2-NE1	2.39	137.41	130.80
5	F	103	ARG	CB-CA-C	2.35	115.89	110.30
6	a	101	SRO	CZ2-CE2-CD2	-2.33	116.50	120.76
6	K	101	SRO	CZ2-CE2-NE1	2.25	137.02	130.80
6	E	101	SRO	CZ2-CE2-CD2	-2.24	116.66	120.76

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	K	101	SRO	CZ2-CE2-CD2	-2.16	116.80	120.76
6	G	101	SRO	CA-CB-CG	-2.16	108.01	112.93
6	G	101	SRO	CZ2-CE2-CD2	-2.16	116.81	120.76
6	O	101	SRO	CE3-CD2-CE2	2.14	121.19	118.26
6	S	101	SRO	CA-CB-CG	-2.12	108.09	112.93
6	Q	101	SRO	CZ2-CH2-CZ3	2.12	122.59	120.15
6	e	101	SRO	CE3-CD2-CE2	2.11	121.15	118.26
6	E	101	SRO	CZ2-CE2-NE1	2.11	136.65	130.80
6	M	101	SRO	OH-CZ3-CH2	-2.07	114.12	120.02
6	W	101	SRO	CZ2-CE2-NE1	2.04	136.44	130.80
6	C	101	SRO	CB-CG-CD1	-2.03	122.91	127.19
6	Y	101	SRO	CH2-CZ3-CE3	2.02	122.97	120.39

There are no chirality outliers.

All (35) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	103	ARG	N-CA-CB-CG
5	B	103	ARG	C-CA-CB-CG
5	F	103	ARG	O-C-CA-N
5	F	103	ARG	N-CA-CB-CG
5	F	103	ARG	C-CA-CB-CG
5	L	101	ARG	O-C-CA-N
5	L	101	ARG	N-CA-CB-CG
5	L	101	ARG	C-CA-CB-CG
5	N	101	ARG	O-C-CA-N
6	M	101	SRO	NZ-CA-CB-CG
6	R	102	SRO	CA-CB-CG-CD2
6	S	101	SRO	NZ-CA-CB-CG
6	Y	101	SRO	NZ-CA-CB-CG
6	e	101	SRO	NZ-CA-CB-CG
5	F	103	ARG	NE-CD-CG-CB
5	L	101	ARG	OXT-C-CA-N
5	F	103	ARG	OXT-C-CA-N
5	N	101	ARG	OXT-C-CA-N
5	B	103	ARG	CA-CB-CG-CD
5	N	101	ARG	NE-CD-CG-CB
6	M	101	SRO	CA-CB-CG-CD2
6	a	101	SRO	NZ-CA-CB-CG
5	F	103	ARG	OXT-C-CA-CB
6	R	102	SRO	CA-CB-CG-CD1
5	F	103	ARG	O-C-CA-CB

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	L	101	ARG	O-C-CA-CB
5	L	101	ARG	CG-CD-NE-CZ
6	E	101	SRO	NZ-CA-CB-CG
6	K	101	SRO	NZ-CA-CB-CG
5	N	101	ARG	O-C-CA-CB
5	L	101	ARG	OXT-C-CA-CB
5	N	101	ARG	OXT-C-CA-CB
5	B	103	ARG	O-C-CA-CB
5	B	103	ARG	OXT-C-CA-CB
5	B	103	ARG	OXT-C-CA-N

There are no ring outliers.

No monomer is involved in short contacts.

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	20/21 (95%)	1.23	3 (15%)	2	2	59, 69, 83, 85	0
1	C	20/21 (95%)	0.45	1 (5%)	28	30	43, 57, 72, 75	0
1	E	21/21 (100%)	0.98	3 (14%)	2	2	38, 51, 73, 84	0
1	G	21/21 (100%)	0.03	1 (4%)	30	32	25, 29, 39, 56	0
1	I	21/21 (100%)	0.12	1 (4%)	30	32	27, 38, 53, 61	0
1	K	20/21 (95%)	0.51	1 (5%)	28	30	44, 57, 75, 75	0
1	M	21/21 (100%)	0.32	0	100	100	35, 48, 59, 66	0
1	O	21/21 (100%)	-0.42	0	100	100	24, 33, 44, 52	0
1	Q	20/21 (95%)	0.10	1 (5%)	28	30	44, 54, 62, 63	0
1	S	21/21 (100%)	0.00	0	100	100	37, 48, 58, 62	0
1	U	20/21 (95%)	0.33	0	100	100	48, 57, 67, 68	0
1	W	21/21 (100%)	-0.32	0	100	100	27, 37, 43, 60	0
1	Y	20/21 (95%)	0.46	1 (5%)	28	30	43, 55, 70, 74	0
1	a	21/21 (100%)	-0.10	0	100	100	27, 35, 45, 50	0
1	c	21/21 (100%)	0.07	1 (4%)	30	32	34, 42, 55, 59	0
1	e	21/21 (100%)	-0.18	0	100	100	24, 31, 45, 57	0
2	B	25/30 (83%)	0.13	0	100	100	32, 43, 61, 77	0
2	D	28/30 (93%)	0.13	1 (3%)	42	44	27, 43, 64, 72	0
2	F	25/30 (83%)	-0.20	0	100	100	20, 30, 52, 68	0
2	H	29/30 (96%)	-0.15	0	100	100	19, 26, 40, 68	0
2	J	28/30 (93%)	-0.02	1 (3%)	42	44	20, 28, 43, 56	0
2	L	25/30 (83%)	-0.20	0	100	100	24, 35, 59, 72	0
2	N	25/30 (83%)	-0.17	0	100	100	23, 31, 61, 65	0
2	P	28/30 (93%)	-0.03	1 (3%)	42	44	19, 29, 49, 56	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
2	R	24/30 (80%)	0.10	1 (4%) 36 37	29, 39, 57, 63	0
2	T	29/30 (96%)	0.29	2 (6%) 16 18	23, 39, 64, 76	0
2	V	24/30 (80%)	0.08	0 100 100	30, 42, 55, 62	0
2	X	28/30 (93%)	0.29	2 (7%) 16 17	23, 36, 62, 70	0
2	Z	24/30 (80%)	-0.13	0 100 100	21, 33, 52, 58	0
2	b	29/30 (96%)	-0.11	0 100 100	21, 32, 50, 62	0
2	d	24/30 (80%)	-0.14	0 100 100	22, 32, 50, 53	0
2	f	29/30 (96%)	0.13	2 (6%) 16 18	21, 28, 49, 70	0
All	All	754/816 (92%)	0.10	23 (3%) 49 50	19, 40, 67, 85	0

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	1	GLY	5.2
2	J	1	PHE	4.2
2	D	1	PHE	4.0
2	X	25	PHE	3.8
2	R	25	PHE	3.8
1	A	14	TYR	3.7
1	G	14	TYR	3.5
1	A	18	ASN	3.3
2	X	1	PHE	3.2
2	T	2	VAL	2.8
2	P	1	PHE	2.7
1	C	10	ILE	2.7
2	f	1	PHE	2.7
2	T	1	PHE	2.7
2	f	29	LYS	2.6
1	K	18	ASN	2.4
1	E	2	ILE	2.4
1	A	13	LEU	2.3
1	E	18	ASN	2.3
1	I	14	TYR	2.3
1	Y	4	GLU	2.1
1	Q	1	GLY	2.1
1	c	21	ASN	2.0

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
5	ARG	L	101	12/12	0.32	0.40	32,41,46,61	11
5	ARG	F	103	12/12	0.41	0.43	65,87,99,100	0
6	SRO	R	102	13/13	0.59	0.31	53,56,59,59	13
5	ARG	B	103	12/12	0.64	0.49	59,86,90,94	0
5	ARG	N	101	12/12	0.68	0.31	37,46,47,49	12
4	CL	Z	101	1/1	0.72	0.09	89,89,89,89	0
6	SRO	M	101	13/13	0.80	0.30	29,32,41,41	13
6	SRO	Y	101	13/13	0.82	0.18	52,59,68,76	0
6	SRO	S	102	13/13	0.83	0.22	61,64,73,75	0
4	CL	F	102	1/1	0.84	0.07	66,66,66,66	0
6	SRO	K	101	13/13	0.87	0.21	41,45,48,48	0
6	SRO	Q	101	13/13	0.88	0.16	39,41,46,54	0
6	SRO	C	101	13/13	0.92	0.11	40,42,48,50	0
6	SRO	S	101	13/13	0.92	0.10	40,43,47,51	0
6	SRO	c	101	13/13	0.92	0.10	30,35,39,40	0
6	SRO	E	101	13/13	0.93	0.16	35,40,46,59	0
6	SRO	a	101	13/13	0.93	0.09	27,33,35,36	0
6	SRO	O	101	13/13	0.93	0.11	22,26,28,30	0
6	SRO	G	101	13/13	0.96	0.09	25,28,30,30	0
6	SRO	W	101	13/13	0.96	0.09	24,29,31,32	0
6	SRO	I	101	13/13	0.96	0.10	29,33,36,39	0
4	CL	B	102	1/1	0.96	0.07	58,58,58,58	1
4	CL	D	102	1/1	0.96	0.15	30,30,30,30	1
6	SRO	e	101	13/13	0.96	0.09	26,29,32,33	0
4	CL	P	101	1/1	0.97	0.11	21,21,21,21	0
4	CL	T	102	1/1	0.98	0.11	22,22,22,22	0
3	ZN	R	101	1/1	0.99	0.10	35,35,35,35	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	CL	f	102	1/1	0.99	0.14	21,21,21,21	1
3	ZN	T	101	1/1	1.00	0.10	20,20,20,20	0
3	ZN	d	101	1/1	1.00	0.11	28,28,28,28	1
4	CL	d	102	1/1	1.00	0.09	61,61,61,61	1
3	ZN	f	101	1/1	1.00	0.11	19,19,19,19	1
3	ZN	D	101	1/1	1.00	0.12	25,25,25,25	1
3	ZN	F	101	1/1	1.00	0.07	32,32,32,32	0
3	ZN	H	101	1/1	1.00	0.12	17,17,17,17	0
3	ZN	B	101	1/1	1.00	0.07	39,39,39,39	1

6.5 Other polymers (i)

There are no such residues in this entry.