



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 05:57 AM GMT

PDB ID : 3A1M
Title : A fusion protein of a beta helix region of gene product 5 and the foldon region of bacteriophage T4
Authors : Yokoi, N.; Suzuki, A.; Hikage, T.; Koshiyama, T.; Terauchi, M.; Yutani, K.; Kanamaru, S.; Arisaka, F.; Yamane, T.; Watanabe, Y.; Ueno, T.
Deposited on : 2009-04-11
Resolution : 2.00 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

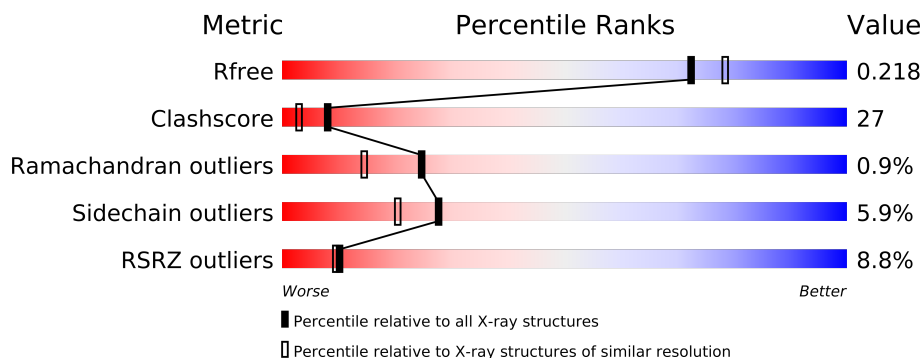
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	66092	4888 (2.00-2.00)
Clashscore	79885	6188 (2.00-2.00)
Ramachandran outliers	78287	6102 (2.00-2.00)
Sidechain outliers	78261	6100 (2.00-2.00)
RSRZ outliers	66119	4890 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	139	
1	B	139	
1	C	139	
1	D	139	
1	E	139	
1	F	139	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4820 atoms, of which 0 are hydrogen and 0 are deuterium.

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 chimera of thrombin cleavage site, Tail-associated lysozyme, Fibrin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	112	Total	C	N	O	S	0	5	0
			856	531	140	183	2			
1	B	112	Total	C	N	O	S	0	5	0
			856	531	140	183	2			
1	C	111	Total	C	N	O	S	0	5	0
			847	526	139	180	2			
1	D	85	Total	C	N	O	S	0	2	0
			626	383	104	137	2			
1	E	85	Total	C	N	O	S	0	2	0
			626	383	104	137	2			
1	F	85	Total	C	N	O	S	0	3	0
			630	386	104	138	2			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	16	LEU	VAL	ENGINEERED	UNP P16009
A	102	SER	-	LINKER	UNP P16009
A	103	VAL	-	LINKER	UNP P16009
A	104	GLU	-	LINKER	UNP P16009
A	132	VAL	-	EXPRESSION TAG	UNP P10104
A	133	GLU	-	EXPRESSION TAG	UNP P10104
A	134	HIS	-	EXPRESSION TAG	UNP P10104
A	135	HIS	-	EXPRESSION TAG	UNP P10104
A	136	HIS	-	EXPRESSION TAG	UNP P10104
A	137	HIS	-	EXPRESSION TAG	UNP P10104
A	138	HIS	-	EXPRESSION TAG	UNP P10104
A	139	HIS	-	EXPRESSION TAG	UNP P10104
B	16	LEU	VAL	ENGINEERED	UNP P16009
B	102	SER	-	LINKER	UNP P16009
B	103	VAL	-	LINKER	UNP P16009
B	104	GLU	-	LINKER	UNP P16009

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	132	VAL	-	EXPRESSION TAG	UNP P10104
B	133	GLU	-	EXPRESSION TAG	UNP P10104
B	134	HIS	-	EXPRESSION TAG	UNP P10104
B	135	HIS	-	EXPRESSION TAG	UNP P10104
B	136	HIS	-	EXPRESSION TAG	UNP P10104
B	137	HIS	-	EXPRESSION TAG	UNP P10104
B	138	HIS	-	EXPRESSION TAG	UNP P10104
B	139	HIS	-	EXPRESSION TAG	UNP P10104
C	16	LEU	VAL	ENGINEERED	UNP P16009
C	102	SER	-	LINKER	UNP P16009
C	103	VAL	-	LINKER	UNP P16009
C	104	GLU	-	LINKER	UNP P16009
C	132	VAL	-	EXPRESSION TAG	UNP P10104
C	133	GLU	-	EXPRESSION TAG	UNP P10104
C	134	HIS	-	EXPRESSION TAG	UNP P10104
C	135	HIS	-	EXPRESSION TAG	UNP P10104
C	136	HIS	-	EXPRESSION TAG	UNP P10104
C	137	HIS	-	EXPRESSION TAG	UNP P10104
C	138	HIS	-	EXPRESSION TAG	UNP P10104
C	139	HIS	-	EXPRESSION TAG	UNP P10104
D	16	LEU	VAL	ENGINEERED	UNP P16009
D	102	SER	-	LINKER	UNP P16009
D	103	VAL	-	LINKER	UNP P16009
D	104	GLU	-	LINKER	UNP P16009
D	132	VAL	-	EXPRESSION TAG	UNP P10104
D	133	GLU	-	EXPRESSION TAG	UNP P10104
D	134	HIS	-	EXPRESSION TAG	UNP P10104
D	135	HIS	-	EXPRESSION TAG	UNP P10104
D	136	HIS	-	EXPRESSION TAG	UNP P10104
D	137	HIS	-	EXPRESSION TAG	UNP P10104
D	138	HIS	-	EXPRESSION TAG	UNP P10104
D	139	HIS	-	EXPRESSION TAG	UNP P10104
E	16	LEU	VAL	ENGINEERED	UNP P16009
E	102	SER	-	LINKER	UNP P16009
E	103	VAL	-	LINKER	UNP P16009
E	104	GLU	-	LINKER	UNP P16009
E	132	VAL	-	EXPRESSION TAG	UNP P10104
E	133	GLU	-	EXPRESSION TAG	UNP P10104
E	134	HIS	-	EXPRESSION TAG	UNP P10104
E	135	HIS	-	EXPRESSION TAG	UNP P10104
E	136	HIS	-	EXPRESSION TAG	UNP P10104
E	137	HIS	-	EXPRESSION TAG	UNP P10104

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	138	HIS	-	EXPRESSION TAG	UNP P10104
E	139	HIS	-	EXPRESSION TAG	UNP P10104
F	16	LEU	VAL	ENGINEERED	UNP P16009
F	102	SER	-	LINKER	UNP P16009
F	103	VAL	-	LINKER	UNP P16009
F	104	GLU	-	LINKER	UNP P16009
F	132	VAL	-	EXPRESSION TAG	UNP P10104
F	133	GLU	-	EXPRESSION TAG	UNP P10104
F	134	HIS	-	EXPRESSION TAG	UNP P10104
F	135	HIS	-	EXPRESSION TAG	UNP P10104
F	136	HIS	-	EXPRESSION TAG	UNP P10104
F	137	HIS	-	EXPRESSION TAG	UNP P10104
F	138	HIS	-	EXPRESSION TAG	UNP P10104
F	139	HIS	-	EXPRESSION TAG	UNP P10104

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total K 1 1	0	0
2	D	1	Total K 1 1	0	0

- Molecule 3 is water.

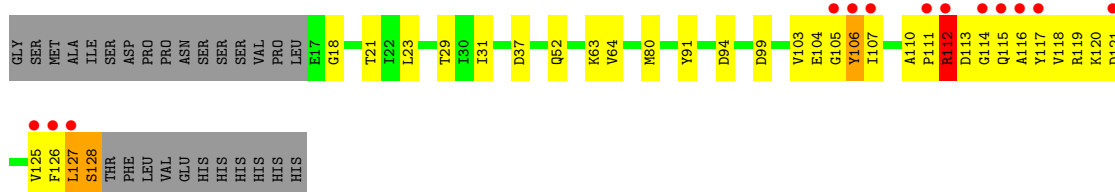
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	83	Total O 83 83	0	0
3	B	73	Total O 73 73	0	0
3	C	86	Total O 86 86	0	0
3	D	49	Total O 49 49	0	0
3	E	41	Total O 41 41	0	0
3	F	45	Total O 45 45	0	0

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 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: chimera of thrombin cleavage site, Tail-associated lysozyme, Fibrin

Chain A: 



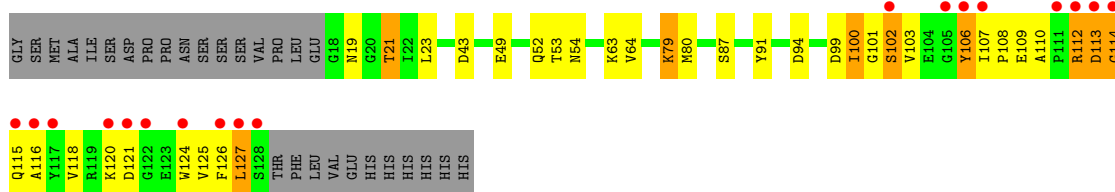
- Molecule 1: chimera of thrombin cleavage site, Tail-associated lysozyme, Fibrin

Chain B: 



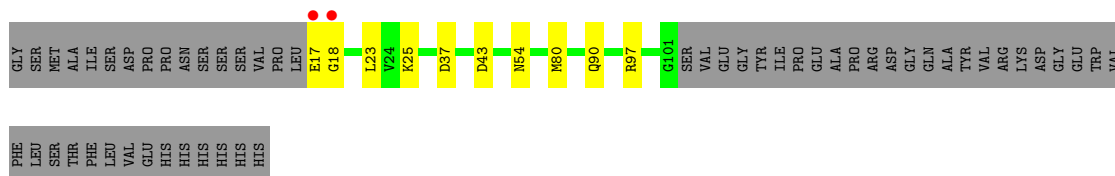
- Molecule 1: chimera of thrombin cleavage site, Tail-associated lysozyme, Fibrin

Chain C: 

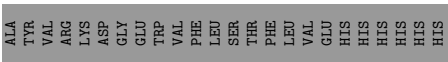
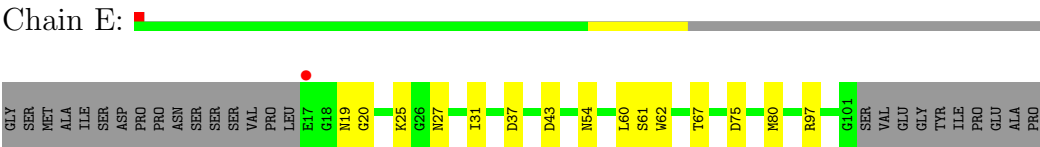


- Molecule 1: chimera of thrombin cleavage site, Tail-associated lysozyme, Fibrin

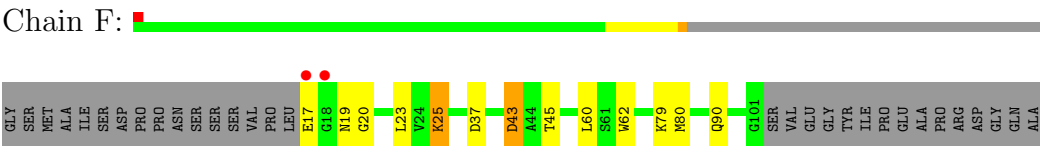
Chain D: 



● Molecule 1: chimera of thrombin cleavage site, Tail-associated lysozyme, Fibrin



● Molecule 1: chimera of thrombin cleavage site, Tail-associated lysozyme, Fibrin



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	98.81Å 57.01Å 136.59Å 90.00° 103.90° 90.00°	Depositor
Resolution (Å)	29.17 – 2.00 29.17 – 2.00	Depositor EDS
% Data completeness (in resolution range)	89.0 (29.17-2.00) 88.9 (29.17-2.00)	Depositor EDS
R_{merge}	0.03	Depositor
R_{sym}	0.03	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.33 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.5.0066	Depositor
R, R_{free}	0.223 , 0.268 0.205 , 0.218	Depositor DCC
R_{free} test set	2253 reflections (5.34%)	DCC
Wilson B-factor (Å ²)	20.6	Xtriage
Anisotropy	0.156	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 39.9	EDS
Estimated twinning fraction	0.480 for 1/2*h-3/2*k,-1/2*h-1/2*k,-1/2*h +1/2*k-l 0.487 for 1/2*h+3/2*k,1/2*h-1/2*k,-1/2*h- 1/2*k-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	5 of 44534 reflections (0.011%)	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	4820	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.81% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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

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	1.25	2/884 (0.2%)	1.05	1/1203 (0.1%)
1	B	1.18	0/884	1.08	3/1203 (0.2%)
1	C	1.17	2/875 (0.2%)	1.06	3/1191 (0.3%)
1	D	1.00	0/638	1.04	2/867 (0.2%)
1	E	0.97	0/638	0.99	2/867 (0.2%)
1	F	0.99	0/645	1.00	2/877 (0.2%)
All	All	1.12	4/4564 (0.1%)	1.04	13/6208 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	91	TYR	CE1-CZ	6.89	1.47	1.38
1	C	64	VAL	CB-CG2	6.38	1.66	1.52
1	C	91	TYR	CE1-CZ	5.89	1.46	1.38
1	A	64	VAL	CB-CG2	5.78	1.65	1.52

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	43	ASP	CB-CG-OD1	8.77	126.19	118.30
1	C	94	ASP	CB-CG-OD1	8.57	126.02	118.30
1	A	94	ASP	CB-CG-OD1	7.55	125.10	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	43	ASP	CB-CG-OD1	7.24	124.82	118.30
1	F	43	ASP	CB-CG-OD2	-6.67	112.30	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	112	ARG	Peptide

5.2 Close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	856	0	826	92	0
1	B	856	0	826	71	0
1	C	847	0	820	96	0
1	D	626	0	606	19	0
1	E	626	0	606	17	0
1	F	630	0	613	20	0
2	A	1	0	0	2	0
2	D	1	0	0	0	0
3	A	83	0	0	1	0
3	B	73	0	0	5	0
3	C	86	0	0	5	0
3	D	49	0	0	6	0
3	E	41	0	0	3	1
3	F	45	0	0	8	0
All	All	4820	0	4297	230	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 27.

The worst 5 of 230 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:120:LYS:HG2	1:A:125:VAL:CG2	1.58	1.33

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:120:LYS:CG	1:A:125:VAL:HG21	1.66	1.24
1:B:127:LEU:CD1	1:C:118:VAL:HG11	1.66	1.23
1:A:120:LYS:CD	1:A:125:VAL:HG21	1.72	1.18
1:A:103:VAL:HG23	1:C:106:TYR:OH	1.47	1.13

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:E:238:HOH:O	3:E:300:HOH:O[2_555]	2.15	0.05

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	115/139 (83%)	110 (96%)	4 (4%)	1 (1%)	25	14
1	B	115/139 (83%)	109 (95%)	4 (4%)	2 (2%)	14	5
1	C	114/139 (82%)	107 (94%)	5 (4%)	2 (2%)	13	5
1	D	85/139 (61%)	85 (100%)	0	0	100	100
1	E	85/139 (61%)	84 (99%)	1 (1%)	0	100	100
1	F	86/139 (62%)	85 (99%)	1 (1%)	0	100	100
All	All	600/834 (72%)	580 (97%)	15 (2%)	5 (1%)	25	17

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	127	LEU
1	B	127	LEU
1	C	102	SER
1	B	114	GLY
1	C	114	GLY

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	95/115 (83%)	91 (96%)	4 (4%)	40	34
1	B	95/115 (83%)	87 (92%)	8 (8%)	16	9
1	C	94/115 (82%)	83 (88%)	11 (12%)	8	4
1	D	70/115 (61%)	69 (99%)	1 (1%)	78	81
1	E	70/115 (61%)	67 (96%)	3 (4%)	40	33
1	F	71/115 (62%)	69 (97%)	2 (3%)	56	54
All	All	495/690 (72%)	466 (94%)	29 (6%)	28	20

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	63	LYS
1	C	100	ILE
1	E	80	MET
1	C	79	LYS
1	C	106	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	19	ASN
1	C	52	GLN
1	E	19	ASN
1	B	115	GLN
1	D	90	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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(Å ²)	Q<0.9
1	A	112/139 (80%)	0.74	13 (11%) 5 5	14, 22, 78, 83	0
1	B	112/139 (80%)	0.62	16 (14%) 3 3	14, 22, 74, 81	0
1	C	111/139 (79%)	0.73	18 (16%) 2 3	13, 22, 74, 78	0
1	D	85/139 (61%)	0.23	2 (2%) 56 56	25, 30, 42, 63	0
1	E	85/139 (61%)	0.16	1 (1%) 75 76	25, 30, 41, 62	0
1	F	85/139 (61%)	0.17	2 (2%) 56 56	24, 30, 42, 65	0
All	All	590/834 (70%)	0.48	52 (8%) 10 9	13, 28, 73, 83	0

The worst 5 of 52 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	116	ALA	9.5
1	C	116	ALA	6.4
1	C	106	TYR	6.3
1	A	114	GLY	5.2
1	C	121	ASP	5.1

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

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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	K	A	140	1/1	0.07	-5.31	39,39,39,39	0
2	K	D	140	1/1	0.03	-7.06	55,55,55,55	0

6.5 Other polymers

There are no such residues in this entry.