WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

May 16, 2020 – 08:57 am BST

PDB ID	:	4NJ7
Title	:	PB1 Domain of AtARF7 - SeMet Derivative
Authors	:	Korasick, D.A.; Westfall, C.S.; Lee, S.G.; Nanao, M.; Jez, J.M.; Strader, L.C.
Deposited on	:	2013-11-08
Resolution	:	3.00 Å(reported)

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

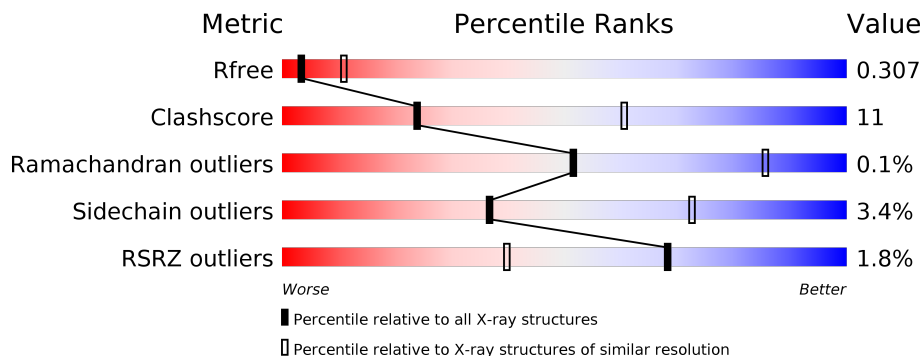|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance ⓘ

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

The reported resolution of this entry is 3.00 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	Percentile Ranks	Value
Rfree		0.307
Clashscore		11
Ramachandran outliers		0.1%
Sidechain outliers		3.4%
RSRZ outliers		1.8%

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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 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	95	62% 16% • 19%
1	B	95	3% / 54% 24% • 18%
1	C	95	73% 9% • 16%
1	D	95	% / 54% 20% 26%
1	E	95	67% 15% • 17%
1	F	95	66% 18% • 15%

WORLDWIDE
PDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	95	3% / 63% / 19% / 18%
1	H	95	% / 63% / 16% / 20%
1	I	95	% / 63% / 17% / 17%
1	J	95	61% / 14% / 22%
1	K	95	3% / 60% / 18% / 21%
1	L	95	3% / 62% / 13% / 25%
1	M	95	3% / 62% / 14% / 24%
1	N	95	% / 54% / 16% / 27%
1	O	95	% / 66% / 16% / 16%
1	P	95	2% / 65% / 16% / 19%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 9822 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 Auxin response factor 7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	77	Total 620	C 393	N 108	O 117	S 1	Se 1	0	0	0
1	B	78	Total 631	C 400	N 112	O 117	S 1	Se 1	0	0	0
1	C	80	Total 650	C 410	N 113	O 124	S 1	Se 2	0	0	0
1	D	70	Total 568	C 363	N 100	O 103	S 1	Se 1	0	0	0
1	E	79	Total 633	C 400	N 109	O 122	S 1	Se 1	0	0	0
1	F	81	Total 653	C 412	N 113	O 125	S 1	Se 2	0	0	0
1	G	78	Total 629	C 399	N 111	O 116	S 1	Se 2	0	0	0
1	H	76	Total 616	C 392	N 108	O 113	S 1	Se 2	0	0	0
1	I	79	Total 638	C 404	N 112	O 119	S 1	Se 2	0	0	0
1	J	74	Total 599	C 382	N 105	O 110	S 1	Se 1	0	0	0
1	K	75	Total 607	C 385	N 107	O 113	S 1	Se 1	0	0	0
1	L	71	Total 567	C 360	N 100	O 105	S 1	Se 1	0	0	0
1	M	72	Total 585	C 374	N 102	O 107	S 1	Se 1	0	0	0
1	N	69	Total 550	C 349	N 96	O 102	S 1	Se 2	0	0	0
1	O	80	Total 650	C 410	N 113	O 124	S 1	Se 2	0	0	0
1	P	77	Total 626	C 397	N 110	O 117	S 1	Se 1	0	0	0

There are 112 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1033	GLY	-	EXPRESSION TAG	UNP P93022
A	1034	SER	-	EXPRESSION TAG	UNP P93022
A	1035	HIS	-	EXPRESSION TAG	UNP P93022
A	1036	MSE	-	EXPRESSION TAG	UNP P93022
A	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
A	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
A	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
B	1033	GLY	-	EXPRESSION TAG	UNP P93022
B	1034	SER	-	EXPRESSION TAG	UNP P93022
B	1035	HIS	-	EXPRESSION TAG	UNP P93022
B	1036	MSE	-	EXPRESSION TAG	UNP P93022
B	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
B	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
B	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
C	1033	GLY	-	EXPRESSION TAG	UNP P93022
C	1034	SER	-	EXPRESSION TAG	UNP P93022
C	1035	HIS	-	EXPRESSION TAG	UNP P93022
C	1036	MSE	-	EXPRESSION TAG	UNP P93022
C	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
C	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
C	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
D	1033	GLY	-	EXPRESSION TAG	UNP P93022
D	1034	SER	-	EXPRESSION TAG	UNP P93022
D	1035	HIS	-	EXPRESSION TAG	UNP P93022
D	1036	MSE	-	EXPRESSION TAG	UNP P93022
D	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
D	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
D	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
E	1033	GLY	-	EXPRESSION TAG	UNP P93022
E	1034	SER	-	EXPRESSION TAG	UNP P93022
E	1035	HIS	-	EXPRESSION TAG	UNP P93022
E	1036	MSE	-	EXPRESSION TAG	UNP P93022
E	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
E	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
E	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
F	1033	GLY	-	EXPRESSION TAG	UNP P93022
F	1034	SER	-	EXPRESSION TAG	UNP P93022
F	1035	HIS	-	EXPRESSION TAG	UNP P93022
F	1036	MSE	-	EXPRESSION TAG	UNP P93022
F	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
F	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
F	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	1033	GLY	-	EXPRESSION TAG	UNP P93022
G	1034	SER	-	EXPRESSION TAG	UNP P93022
G	1035	HIS	-	EXPRESSION TAG	UNP P93022
G	1036	MSE	-	EXPRESSION TAG	UNP P93022
G	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
G	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
G	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
H	1033	GLY	-	EXPRESSION TAG	UNP P93022
H	1034	SER	-	EXPRESSION TAG	UNP P93022
H	1035	HIS	-	EXPRESSION TAG	UNP P93022
H	1036	MSE	-	EXPRESSION TAG	UNP P93022
H	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
H	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
H	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
I	1033	GLY	-	EXPRESSION TAG	UNP P93022
I	1034	SER	-	EXPRESSION TAG	UNP P93022
I	1035	HIS	-	EXPRESSION TAG	UNP P93022
I	1036	MSE	-	EXPRESSION TAG	UNP P93022
I	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
I	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
I	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
J	1033	GLY	-	EXPRESSION TAG	UNP P93022
J	1034	SER	-	EXPRESSION TAG	UNP P93022
J	1035	HIS	-	EXPRESSION TAG	UNP P93022
J	1036	MSE	-	EXPRESSION TAG	UNP P93022
J	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
J	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
J	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
K	1033	GLY	-	EXPRESSION TAG	UNP P93022
K	1034	SER	-	EXPRESSION TAG	UNP P93022
K	1035	HIS	-	EXPRESSION TAG	UNP P93022
K	1036	MSE	-	EXPRESSION TAG	UNP P93022
K	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
K	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
K	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
L	1033	GLY	-	EXPRESSION TAG	UNP P93022
L	1034	SER	-	EXPRESSION TAG	UNP P93022
L	1035	HIS	-	EXPRESSION TAG	UNP P93022
L	1036	MSE	-	EXPRESSION TAG	UNP P93022
L	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
L	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
L	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
M	1033	GLY	-	EXPRESSION TAG	UNP P93022
M	1034	SER	-	EXPRESSION TAG	UNP P93022
M	1035	HIS	-	EXPRESSION TAG	UNP P93022
M	1036	MSE	-	EXPRESSION TAG	UNP P93022
M	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
M	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
M	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
N	1033	GLY	-	EXPRESSION TAG	UNP P93022
N	1034	SER	-	EXPRESSION TAG	UNP P93022
N	1035	HIS	-	EXPRESSION TAG	UNP P93022
N	1036	MSE	-	EXPRESSION TAG	UNP P93022
N	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
N	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
N	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
O	1033	GLY	-	EXPRESSION TAG	UNP P93022
O	1034	SER	-	EXPRESSION TAG	UNP P93022
O	1035	HIS	-	EXPRESSION TAG	UNP P93022
O	1036	MSE	-	EXPRESSION TAG	UNP P93022
O	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
O	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
O	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022
P	1033	GLY	-	EXPRESSION TAG	UNP P93022
P	1034	SER	-	EXPRESSION TAG	UNP P93022
P	1035	HIS	-	EXPRESSION TAG	UNP P93022
P	1036	MSE	-	EXPRESSION TAG	UNP P93022
P	1042	ALA	LYS	ENGINEERED MUTATION	UNP P93022
P	1092	ALA	ASP	ENGINEERED MUTATION	UNP P93022
P	1096	ALA	ASP	ENGINEERED MUTATION	UNP P93022

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Auxin response factor 7
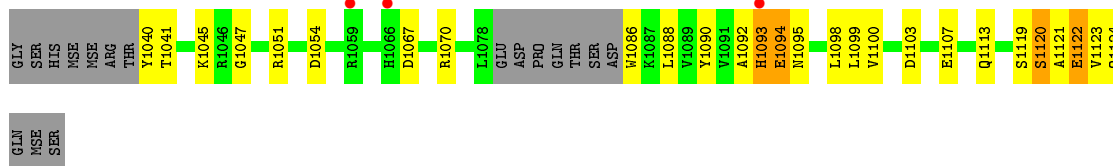
- Molecule 1: Auxin response factor 7

- Molecule 1: Auxin response factor 7

- Molecule 1: Auxin response factor 7

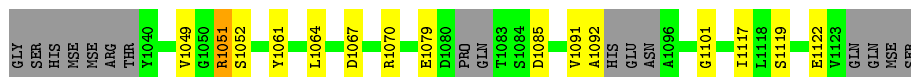- Molecule 1: Auxin response factor 7

- Molecule 1: Auxin response factor 7

Chain F:

- Molecule 1: Auxin response factor 7

Chain G:

- Molecule 1: Auxin response factor 7

Chain H:

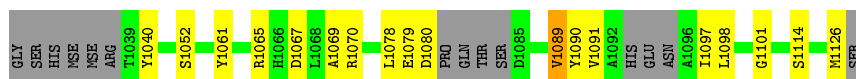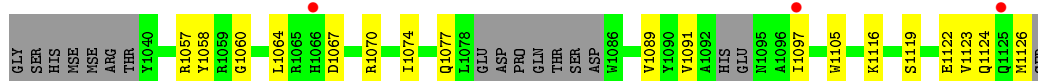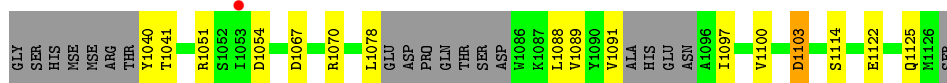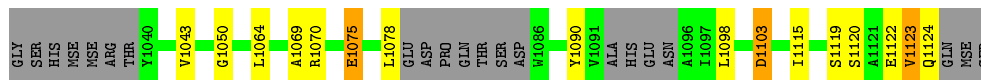- Molecule 1: Auxin response factor 7

Chain I:

- Molecule 1: Auxin response factor 7

Chain J:

- Molecule 1: Auxin response factor 7

Chain K:

- Molecule 1: Auxin response factor 7

Chain L:

• Molecule 1: Auxin response factor 7

Chain M:

• Molecule 1: Auxin response factor 7

Chain N:

• Molecule 1: Auxin response factor 7

Chain O:

• Molecule 1: Auxin response factor 7

Chain P:
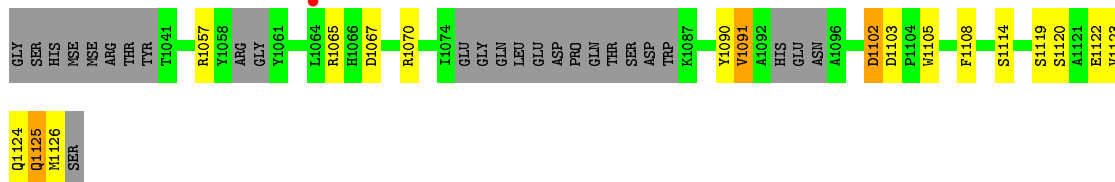
4 Data and refinement statistics (i)

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α, β, γ	150.94Å 150.94Å 184.24Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.60 – 3.00 47.72 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.9 (47.60-3.00) 98.5 (47.72-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.31 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, R_{free}	0.257 , 0.304 0.256 , 0.307	Depositor DCC
R_{free} test set	2029 reflections (6.47%)	wwPDB-VP
Wilson B-factor (Å^2)	68.3	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 51.7	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.035 for -2/3*h-1/3*k+2/3*l,-1/3*h-2/3*k-2/3*l,2/3*h-2/3*k+1/3*l 0.028 for -h,1/3*h-1/3*k+2/3*l,2/3*h+4/3*k+1/3*l 0.035 for -1/3*h+1/3*k-2/3*l,-k,-4/3*h-2/3*k+1/3*l 0.028 for -h,2/3*h+1/3*k-2/3*l,-2/3*h-4/3*k-1/3*l 0.028 for 1/3*h+2/3*k+2/3*l,-k,4/3*h+2/3*k-1/3*l 0.038 for -1/3*h-2/3*k-2/3*l,-2/3*h-1/3*k+2/3*l,-2/3*h+2/3*k-1/3*l 0.037 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	9822	wwPDB-VP
Average B, all atoms (Å^2)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.67% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$ >5	RMSZ	#$|Z|$ >5
1	A	1.24	3/628 (0.5%)	1.09	0/844
1	B	1.17	0/641	1.06	0/863
1	C	1.28	1/657 (0.2%)	1.09	3/880 (0.3%)
1	D	1.20	0/576	1.07	2/774 (0.3%)
1	E	1.32	2/641 (0.3%)	1.15	1/862 (0.1%)
1	F	1.07	1/660 (0.2%)	1.05	0/885
1	G	1.19	1/636 (0.2%)	1.10	2/852 (0.2%)
1	H	1.21	0/623	1.00	0/834
1	I	1.14	0/645	0.98	0/864
1	J	1.08	0/607	1.05	1/815 (0.1%)
1	K	1.17	1/617 (0.2%)	1.04	2/832 (0.2%)
1	L	1.05	0/573	0.98	0/769
1	M	1.07	0/593	1.02	0/796
1	N	1.09	0/553	1.10	3/738 (0.4%)
1	O	1.25	1/657 (0.2%)	1.09	3/880 (0.3%)
1	P	1.13	0/637	1.04	0/859
All	All	1.17	10/9944 (0.1%)	1.06	17/13347 (0.1%)

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	1064	LEU	CG-CD1	5.97	1.74	1.51
1	E	1051	ARG	CG-CD	-5.87	1.37	1.51
1	C	1091	VAL	CB-CG1	-5.75	1.40	1.52
1	A	1042	ALA	CA-CB	-5.72	1.40	1.52
1	G	1091	VAL	CA-CB	5.63	1.66	1.54

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	1051	ARG	NE-CZ-NH1	-5.90	117.35	120.30
1	J	1070	ARG	NE-CZ-NH2	-5.83	117.39	120.30
1	O	1086	TRP	CB-CA-C	-5.67	99.07	110.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	1057	ARG	NE-CZ-NH1	-5.56	117.52	120.30
1	D	1068	LEU	CB-CG-CD2	5.48	120.32	111.00

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	620	0	606	37	0
1	B	631	0	618	55	0
1	C	650	0	632	8	0
1	D	568	0	563	9	0
1	E	633	0	616	8	0
1	F	653	0	636	15	0
1	G	629	0	621	13	0
1	H	616	0	610	11	0
1	I	638	0	627	24	0
1	J	599	0	593	12	0
1	K	607	0	593	14	0
1	L	567	0	567	5	0
1	M	585	0	580	7	0
1	N	550	0	551	15	0
1	O	650	0	632	13	0
1	P	626	0	609	8	0
All	All	9822	0	9654	217	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 11.

The worst 5 of 217 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:A:1059:ARG:HG3	1:B:1095:ASN:CB	1.35	1.51
1:A:1059:ARG:CG	1:B:1095:ASN:HB3	1.57	1.31

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1105:TRP:CB	1:B:1094:GLU:HG2	1.73	1.19
1:B:1120:SER:O	1:B:1123:VAL:HG22	1.46	1.16
1:A:1059:ARG:HG3	1:B:1095:ASN:CG	1.63	1.15

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	71/95 (75%)	69 (97%)	2 (3%)	0	100	100
1	B	74/95 (78%)	71 (96%)	2 (3%)	1 (1%)	11	43
1	C	74/95 (78%)	73 (99%)	1 (1%)	0	100	100
1	D	64/95 (67%)	64 (100%)	0	0	100	100
1	E	73/95 (77%)	72 (99%)	1 (1%)	0	100	100
1	F	75/95 (79%)	74 (99%)	1 (1%)	0	100	100
1	G	72/95 (76%)	71 (99%)	1 (1%)	0	100	100
1	H	70/95 (74%)	68 (97%)	2 (3%)	0	100	100
1	I	73/95 (77%)	69 (94%)	4 (6%)	0	100	100
1	J	68/95 (72%)	66 (97%)	2 (3%)	0	100	100
1	K	71/95 (75%)	71 (100%)	0	0	100	100
1	L	65/95 (68%)	64 (98%)	1 (2%)	0	100	100
1	M	66/95 (70%)	65 (98%)	1 (2%)	0	100	100
1	N	61/95 (64%)	60 (98%)	1 (2%)	0	100	100
1	O	74/95 (78%)	72 (97%)	2 (3%)	0	100	100
1	P	73/95 (77%)	72 (99%)	1 (1%)	0	100	100
All	All	1124/1520 (74%)	1101 (98%)	22 (2%)	1 (0%)	51	85

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	1122	GLU

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	66/79 (84%)	62 (94%)	4 (6%)	18	53
1	B	67/79 (85%)	63 (94%)	4 (6%)	19	53
1	C	70/79 (89%)	70 (100%)	0	100	100
1	D	61/79 (77%)	59 (97%)	2 (3%)	38	73
1	E	68/79 (86%)	66 (97%)	2 (3%)	42	76
1	F	70/79 (89%)	70 (100%)	0	100	100
1	G	67/79 (85%)	67 (100%)	0	100	100
1	H	66/79 (84%)	65 (98%)	1 (2%)	65	87
1	I	68/79 (86%)	64 (94%)	4 (6%)	19	54
1	J	64/79 (81%)	60 (94%)	4 (6%)	18	51
1	K	65/79 (82%)	63 (97%)	2 (3%)	40	75
1	L	61/79 (77%)	58 (95%)	3 (5%)	25	61
1	M	63/79 (80%)	62 (98%)	1 (2%)	62	86
1	N	60/79 (76%)	56 (93%)	4 (7%)	16	49
1	O	70/79 (89%)	68 (97%)	2 (3%)	42	76
1	P	67/79 (85%)	64 (96%)	3 (4%)	27	64
All	All	1053/1264 (83%)	1017 (97%)	36 (3%)	37	72

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	1125	GLN
1	J	1124	GLN
1	P	1057	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	J	1103	ASP
1	K	1093	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1110	ASN
1	B	1095	ASN
1	B	1113	GLN
1	H	1110	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 carbohydrates 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 (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	76/95 (80%)	0.04	0	100	100	43, 71, 121, 148	0
1	B	77/95 (81%)	0.18	3 (3%)	39	15	40, 71, 118, 125	0
1	C	78/95 (82%)	-0.07	0	100	100	41, 69, 120, 169	0
1	D	69/95 (72%)	0.11	1 (1%)	75	49	42, 81, 123, 144	0
1	E	78/95 (82%)	0.07	0	100	100	45, 73, 123, 134	0
1	F	79/95 (83%)	0.23	0	100	100	46, 79, 122, 148	0
1	G	76/95 (80%)	0.17	3 (3%)	39	15	46, 81, 122, 170	0
1	H	74/95 (77%)	0.19	1 (1%)	75	49	50, 77, 137, 154	0
1	I	77/95 (81%)	0.13	1 (1%)	77	51	55, 81, 146, 161	0
1	J	73/95 (76%)	0.06	0	100	100	45, 81, 128, 165	0
1	K	74/95 (77%)	0.27	3 (4%)	37	14	39, 82, 138, 143	0
1	L	70/95 (73%)	0.31	3 (4%)	35	13	52, 91, 144, 228	0
1	M	71/95 (74%)	0.38	3 (4%)	36	14	45, 91, 135, 161	0
1	N	67/95 (70%)	0.26	1 (1%)	73	46	52, 94, 173, 213	0
1	O	78/95 (82%)	-0.01	1 (1%)	77	51	49, 74, 124, 150	0
1	P	76/95 (80%)	0.29	2 (2%)	56	27	49, 79, 128, 149	0
All	All	1193/1520 (78%)	0.16	22 (1%)	68	40	39, 79, 134, 228	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	1057	ARG	3.7
1	B	1066	HIS	3.2
1	K	1059	ARG	2.9
1	L	1068	LEU	2.9
1	B	1093	HIS	2.7

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 carbohydrates 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.