Full wwPDB X-ray Structure Validation Report ⓘ

Apr 15, 2026 – 12:32 AM UTC

PDB ID	:	9NX3 / pdb_00009nx3
Title	:	Crystal structure of GP232 tail fiber recognition domain from mycobacteriophage Bxz-1
Authors	:	Di, D.; Tsai, J.H.; Krieger, I.V.; Sacchettini, J.C.
Deposited on	:	2025-03-25
Resolution	:	1.90 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 ⓘ 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 ⓘ) were used in the production of this report:

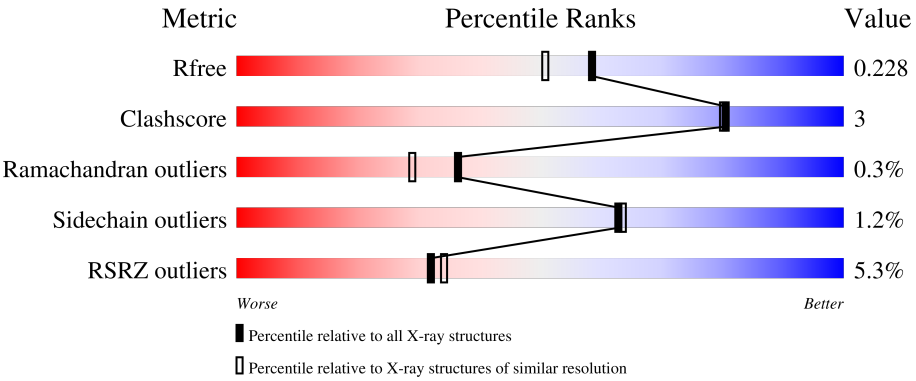
MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 1.90 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.



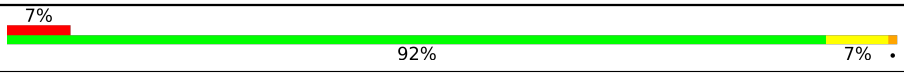
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	180053	7789 (1.90-1.90)
Clashscore	190562	8410 (1.90-1.90)
Ramachandran outliers	187476	8333 (1.90-1.90)
Sidechain outliers	187428	8333 (1.90-1.90)
RSRZ outliers	180081	7790 (1.90-1.90)

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	242	92% 6%
1	B	242	2% 93% 6%
1	C	242	% 95%
1	D	242	2% 95%
1	E	242	19% 86% 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	242	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 11966 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 Capsid decoration protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	237	Total	C	N	O	S	0	0	0
			1685	1043	301	340	1			
1	B	240	Total	C	N	O	S	0	0	0
			1700	1052	304	343	1			
1	C	237	Total	C	N	O	S	0	0	0
			1685	1043	301	340	1			
1	D	239	Total	C	N	O	S	0	0	0
			1709	1058	308	342	1			
1	E	237	Total	C	N	O	S	0	0	0
			1689	1046	302	340	1			
1	F	242	Total	C	N	O	S	0	0	0
			1734	1073	315	345	1			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	HIS	-	expression tag	UNP A0A411CKL9
A	4	HIS	-	expression tag	UNP A0A411CKL9
A	5	HIS	-	expression tag	UNP A0A411CKL9
A	6	HIS	-	expression tag	UNP A0A411CKL9
A	7	HIS	-	expression tag	UNP A0A411CKL9
B	3	HIS	-	expression tag	UNP A0A411CKL9
B	4	HIS	-	expression tag	UNP A0A411CKL9
B	5	HIS	-	expression tag	UNP A0A411CKL9
B	6	HIS	-	expression tag	UNP A0A411CKL9
B	7	HIS	-	expression tag	UNP A0A411CKL9
C	3	HIS	-	expression tag	UNP A0A411CKL9
C	4	HIS	-	expression tag	UNP A0A411CKL9
C	5	HIS	-	expression tag	UNP A0A411CKL9
C	6	HIS	-	expression tag	UNP A0A411CKL9
C	7	HIS	-	expression tag	UNP A0A411CKL9
D	3	HIS	-	expression tag	UNP A0A411CKL9
D	4	HIS	-	expression tag	UNP A0A411CKL9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	5	HIS	-	expression tag	UNP A0A411CKL9
D	6	HIS	-	expression tag	UNP A0A411CKL9
D	7	HIS	-	expression tag	UNP A0A411CKL9
E	3	HIS	-	expression tag	UNP A0A411CKL9
E	4	HIS	-	expression tag	UNP A0A411CKL9
E	5	HIS	-	expression tag	UNP A0A411CKL9
E	6	HIS	-	expression tag	UNP A0A411CKL9
E	7	HIS	-	expression tag	UNP A0A411CKL9
F	3	HIS	-	expression tag	UNP A0A411CKL9
F	4	HIS	-	expression tag	UNP A0A411CKL9
F	5	HIS	-	expression tag	UNP A0A411CKL9
F	6	HIS	-	expression tag	UNP A0A411CKL9
F	7	HIS	-	expression tag	UNP A0A411CKL9

- Molecule 2 is water.

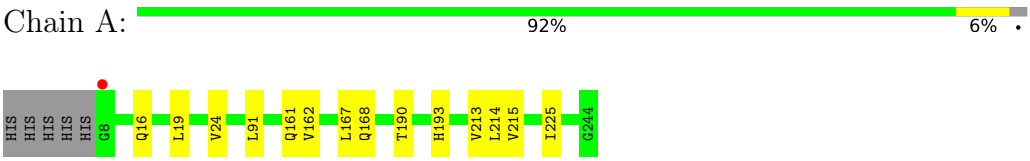
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	367	Total O 367 367	0	0
2	B	257	Total O 257 257	0	0
2	C	332	Total O 332 332	0	0
2	D	301	Total O 301 301	0	0
2	E	203	Total O 203 203	0	0
2	F	304	Total O 304 304	0	0

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.

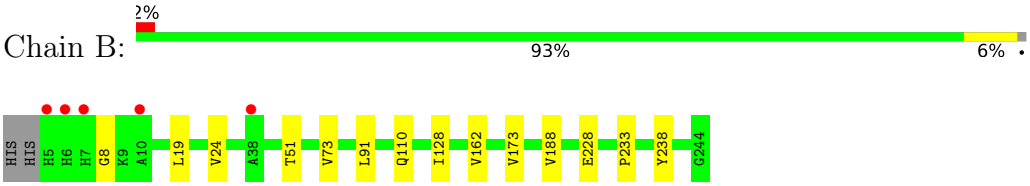
• Molecule 1: Capsid decoration protein

Chain A:



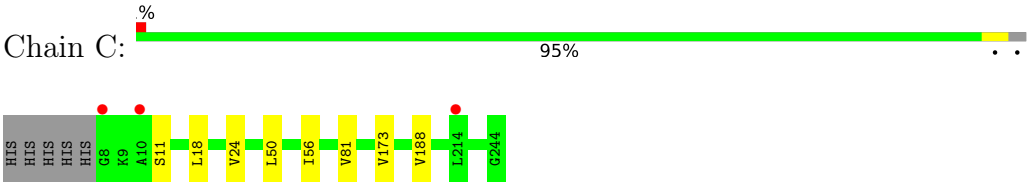
• Molecule 1: Capsid decoration protein

Chain B:



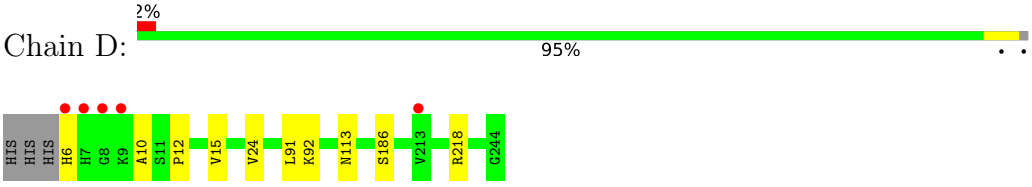
• Molecule 1: Capsid decoration protein

Chain C:



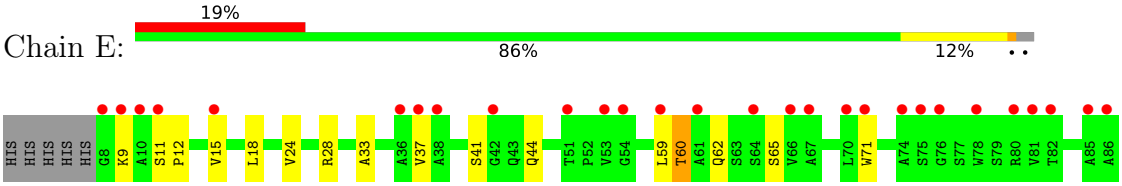
• Molecule 1: Capsid decoration protein

Chain D:

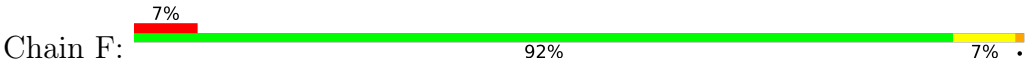


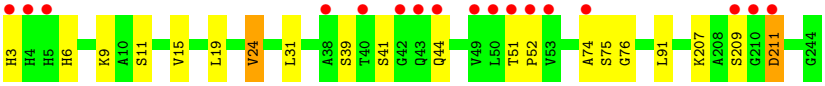
• Molecule 1: Capsid decoration protein

Chain E:



• Molecule 1: Capsid decoration protein





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	56.84Å 88.39Å 96.33Å 96.16° 102.08° 107.98°	Depositor
Resolution (Å)	41.31 – 1.90 41.31 – 1.90	Depositor EDS
% Data completeness (in resolution range)	93.6 (41.31-1.90) 93.6 (41.31-1.90)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.99 (at 1.89Å)	Xtriage
Refinement program	REFMAC 1.20.1_4487	Depositor
R, R_{free}	0.182 , 0.220 0.194 , 0.228	Depositor DCC
R_{free} test set	6367 reflections (4.71%)	wwPDB-VP
Wilson B-factor (Å^2)	24.6	Xtriage
Anisotropy	0.397	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 41.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	11966	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.56% 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

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	0.33	0/1712	0.52	0/2343
1	B	0.27	0/1727	0.46	0/2364
1	C	0.32	0/1712	0.50	0/2343
1	D	0.31	0/1738	0.51	0/2377
1	E	0.26	0/1716	0.47	0/2347
1	F	0.30	0/1765	0.48	1/2414 (0.0%)
All	All	0.30	0/10370	0.49	1/14188 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	F	52	PRO	CA-N-CD	-6.46	102.96	112.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	1685	0	1657	11	0
1	B	1700	0	1663	9	0
1	C	1685	0	1657	4	0
1	D	1709	0	1682	9	0
1	E	1689	0	1668	18	0
1	F	1734	0	1698	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	367	0	0	1	0
2	B	257	0	0	1	0
2	C	332	0	0	0	0
2	D	301	0	0	1	0
2	E	203	0	0	1	0
2	F	304	0	0	3	0
All	All	11966	0	10025	52	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 3.

All (52) 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:24:VAL:HG22	1:C:24:VAL:HG11	1.68	0.73
1:B:24:VAL:HG23	2:B:440:HOH:O	1.88	0.73
1:E:9:LYS:HD2	1:F:9:LYS:HB2	1.71	0.73
1:B:24:VAL:HG22	1:B:91:LEU:HD12	1.78	0.64
1:A:16:GLN:HG3	1:C:18:LEU:HD21	1.81	0.62
1:A:162:VAL:HG12	1:A:167:LEU:HD23	1.82	0.60
1:A:161:GLN:OE1	1:A:168:GLN:NE2	2.25	0.58
1:E:37:VAL:HB	1:E:62:GLN:HE22	1.69	0.58
1:F:39:SER:O	1:F:44:GLN:NE2	2.37	0.57
1:B:188:VAL:HG22	1:B:228:GLU:HG3	1.88	0.56
1:F:24:VAL:HG13	1:F:91:LEU:HD12	1.85	0.56
1:E:229:PHE:O	1:E:231:SER:N	2.40	0.55
1:E:33:ALA:O	1:E:60:THR:HB	2.08	0.54
1:F:74:ALA:HB3	2:F:313:HOH:O	2.06	0.54
1:D:24:VAL:HG23	1:D:91:LEU:HD12	1.90	0.54
1:A:19:LEU:HD12	1:B:19:LEU:HD21	1.91	0.53
1:E:59:LEU:HD13	1:E:71:TRP:CD1	2.45	0.52
1:D:24:VAL:CG2	1:D:91:LEU:HD12	2.39	0.52
1:D:6:HIS:N	2:D:307:HOH:O	2.43	0.52
1:E:91:LEU:HD11	1:F:19:LEU:HD22	1.92	0.52
1:A:24:VAL:HG13	2:A:532:HOH:O	2.09	0.51
1:E:24:VAL:HG22	1:E:91:LEU:HD12	1.92	0.50
1:E:28:ARG:HD2	1:E:99:THR:OG1	2.12	0.49
1:F:31:LEU:HD12	1:F:51:THR:HG22	1.93	0.49
1:A:24:VAL:HG12	1:A:91:LEU:HD12	1.94	0.49
1:D:24:VAL:HB	1:F:24:VAL:HG21	1.95	0.49
1:A:213:VAL:HG12	1:A:215:VAL:HG13	1.94	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:E:41:SER:O	1:E:44:GLN:NE2	2.46	0.48
1:E:65:SER:HB2	1:E:122:ALA:HB1	1.95	0.48
1:E:213:VAL:HG12	1:E:215:VAL:HG13	1.94	0.48
1:A:24:VAL:CG1	1:A:91:LEU:HD12	2.45	0.47
1:F:209:SER:O	1:F:211:ASP:N	2.45	0.47
1:D:92:LYS:HG3	1:D:113:ASN:O	2.16	0.46
1:B:233:PRO:HB3	1:B:238:TYR:HD2	1.82	0.45
1:E:162:VAL:HG22	1:E:167:LEU:HD23	1.99	0.44
1:D:218:ARG:HD2	1:E:209:SER:O	2.18	0.43
1:E:24:VAL:CG2	1:E:91:LEU:HD12	2.48	0.43
1:B:8:GLY:HA2	1:C:11:SER:HB2	2.00	0.43
1:E:24:VAL:HG23	2:E:437:HOH:O	2.17	0.43
1:F:207:LYS:NZ	2:F:317:HOH:O	2.52	0.43
1:F:41:SER:HA	1:F:76:GLY:O	2.18	0.43
1:E:233:PRO:HB3	1:E:238:TYR:CD2	2.54	0.43
1:F:3:HIS:N	2:F:318:HOH:O	2.52	0.43
1:B:51:THR:HB	1:B:73:VAL:HG21	2.01	0.42
1:A:19:LEU:HD11	1:B:19:LEU:HD11	1.99	0.42
1:B:110:GLN:HB2	1:B:128:ILE:HD13	2.01	0.42
1:D:15:VAL:HG21	1:F:15:VAL:HG22	2.02	0.41
1:D:12:PRO:HG2	1:F:6:HIS:O	2.21	0.41
1:A:193:HIS:CD2	1:A:225:ILE:HG23	2.56	0.41
1:D:10:ALA:HB2	1:E:12:PRO:HD3	2.02	0.41
1:E:11:SER:O	1:E:15:VAL:HG23	2.19	0.41
1:C:56:ILE:HD12	1:C:81:VAL:HG12	2.02	0.41

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	235/242 (97%)	232 (99%)	3 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	238/242 (98%)	233 (98%)	5 (2%)	0	100	100
1	C	235/242 (97%)	231 (98%)	4 (2%)	0	100	100
1	D	237/242 (98%)	233 (98%)	4 (2%)	0	100	100
1	E	235/242 (97%)	225 (96%)	7 (3%)	3 (1%)	9	3
1	F	240/242 (99%)	233 (97%)	6 (2%)	1 (0%)	30	22
All	All	1420/1452 (98%)	1387 (98%)	29 (2%)	4 (0%)	36	29

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	230	GLU
1	F	211	ASP
1	E	231	SER
1	E	185	GLY

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	179/185 (97%)	177 (99%)	2 (1%)	65	67
1	B	179/185 (97%)	177 (99%)	2 (1%)	65	67
1	C	179/185 (97%)	176 (98%)	3 (2%)	53	52
1	D	182/185 (98%)	181 (100%)	1 (0%)	81	84
1	E	180/185 (97%)	178 (99%)	2 (1%)	65	67
1	F	184/185 (100%)	181 (98%)	3 (2%)	55	54
All	All	1083/1110 (98%)	1070 (99%)	13 (1%)	63	63

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	190	THR
1	A	214	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	162	VAL
1	B	173	VAL
1	C	50	LEU
1	C	173	VAL
1	C	188	VAL
1	D	186	SER
1	E	18	LEU
1	E	60	THR
1	F	11	SER
1	F	24	VAL
1	F	75	SER

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	44	GLN
1	B	161	GLN
1	B	168	GLN
1	C	106	ASN
1	C	161	GLN
1	C	168	GLN
1	D	161	GLN
1	D	168	GLN
1	E	44	GLN
1	E	161	GLN
1	F	3	HIS
1	F	43	GLN
1	F	161	GLN
1	F	168	GLN

5.3.3 RNA

There are no RNA molecules 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 (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	237/242 (97%)	-0.33	1 (0%) 88 90	15, 24, 35, 65	0
1	B	240/242 (99%)	0.23	5 (2%) 63 67	18, 34, 58, 77	0
1	C	237/242 (97%)	-0.08	3 (1%) 75 78	18, 26, 39, 66	0
1	D	239/242 (98%)	0.10	5 (2%) 63 67	16, 28, 44, 93	0
1	E	237/242 (97%)	0.87	45 (18%) 3 3	17, 39, 69, 83	0
1	F	242/242 (100%)	0.21	17 (7%) 22 24	17, 27, 56, 61	0
All	All	1432/1452 (98%)	0.17	76 (5%) 32 34	15, 28, 57, 93	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	210	GLY	5.8
1	D	7	HIS	4.9
1	B	6	HIS	4.2
1	C	8	GLY	4.1
1	B	5	HIS	3.9
1	F	53	VAL	3.9
1	B	7	HIS	3.8
1	F	209	SER	3.8
1	F	74	ALA	3.7
1	D	6	HIS	3.6
1	E	119	GLY	3.6
1	E	64	SER	3.5
1	F	42	GLY	3.5
1	E	61	ALA	3.4
1	E	42	GLY	3.4
1	E	74	ALA	3.4
1	E	87	GLY	3.3
1	F	49	VAL	3.3
1	E	75	SER	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	54	GLY	3.2
1	E	85	ALA	3.2
1	E	229	PHE	3.1
1	E	67	ALA	3.1
1	F	3	HIS	3.0
1	D	213	VAL	3.0
1	F	52	PRO	3.0
1	E	10	ALA	2.9
1	F	211	ASP	2.9
1	D	8	GLY	2.9
1	E	86	ALA	2.8
1	E	53	VAL	2.8
1	E	235	SER	2.8
1	A	8	GLY	2.7
1	E	76	GLY	2.7
1	B	38	ALA	2.6
1	E	36	ALA	2.6
1	F	5	HIS	2.5
1	E	213	VAL	2.5
1	E	9	LYS	2.5
1	E	122	ALA	2.5
1	E	232	ALA	2.5
1	E	8	GLY	2.4
1	E	71	TRP	2.4
1	E	78	TRP	2.4
1	E	186	SER	2.4
1	F	50	LEU	2.4
1	F	51	THR	2.4
1	C	214	LEU	2.4
1	E	70	LEU	2.3
1	E	11	SER	2.3
1	E	66	VAL	2.3
1	E	82	THR	2.3
1	E	118	VAL	2.3
1	E	238	TYR	2.3
1	E	233	PRO	2.2
1	E	51	THR	2.2
1	E	185	GLY	2.2
1	E	234	ALA	2.2
1	E	117	VAL	2.2
1	F	38	ALA	2.2
1	E	81	VAL	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	236	GLY	2.1
1	E	59	LEU	2.1
1	D	9	LYS	2.1
1	C	10	ALA	2.1
1	E	121	ASN	2.1
1	E	120	THR	2.1
1	E	37	VAL	2.1
1	F	4	HIS	2.1
1	B	10	ALA	2.1
1	E	80	ARG	2.1
1	F	40	THR	2.1
1	F	43	GLN	2.1
1	E	38	ALA	2.1
1	F	44	GLN	2.0
1	E	15	VAL	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 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.