Full wwPDB X-ray Structure Validation Report ⓘ

Aug 29, 2023 – 09:23 PM EDT

PDB ID	:	3MYW
Title	:	The Bowman-Birk type inhibitor from mung bean in ternary complex with porcine trypsin
Authors	:	Engh, R.A.; Bode, W.; Huber, R.; Lin, G.; Chi, C.
Deposited on	:	2010-05-11
Resolution	:	2.50 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:

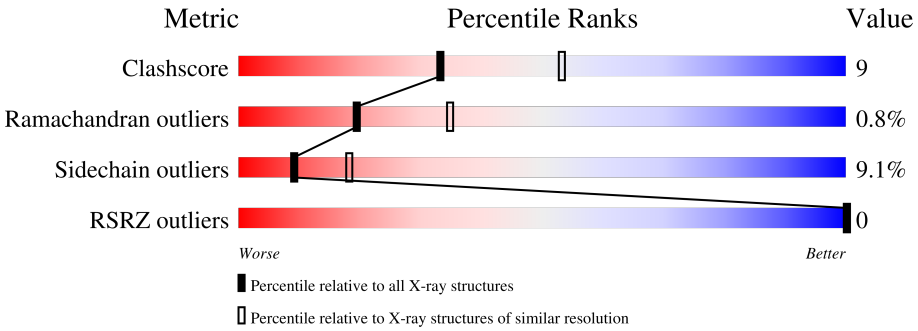
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.50 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))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	223	78% 20% ·
1	B	223	78% 19% ·
2	I	72	40% 28% 6% · 24%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 3953 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 Trypsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	223	Total	C	N	O	S	35	0	0
			1640	1019	287	320	14			
1	B	223	Total	C	N	O	S	0	0	0
			1640	1019	287	320	14			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	VAL	ILE	variant	UNP P00761
A	171	ALA	SER	SEE REMARK 999	UNP P00761
B	27	VAL	ILE	variant	UNP P00761
B	171	ALA	SER	SEE REMARK 999	UNP P00761

- Molecule 2 is a protein called Bowman-Birk type trypsin inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	55	Total	C	N	O	S	40	0	0
			409	240	76	78	15			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	22	LYS	ILE	SEE REMARK 999	UNP P01062
I	25	GLN	GLU	SEE REMARK 999	UNP P01062

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		
3	B	1	Total	Ca	0	0
			1	1		

- Molecule 4 is water.

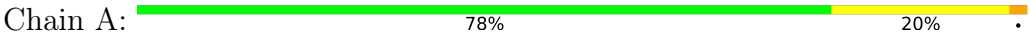
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	125	Total O 125 125	0	0
4	I	12	Total O 12 12	0	0
4	B	125	Total O 125 125	0	0

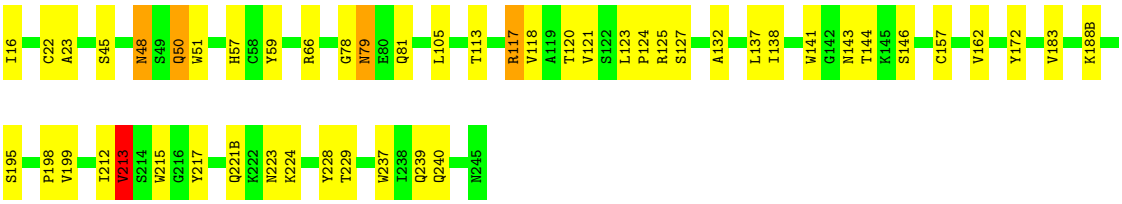
3 Residue-property plots (i)

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: Trypsin

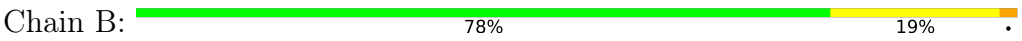
Chain A:





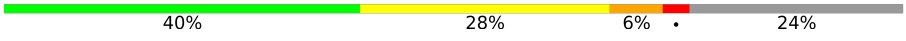
- Molecule 1: Trypsin

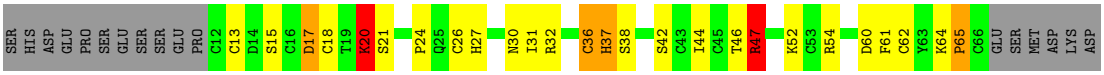
Chain B:



- Molecule 2: Bowman-Birk type trypsin inhibitor

Chain I:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, α, β, γ	62.46Å 62.46Å 160.01Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	8.00 – 2.50 37.98 – 2.40	Depositor EDS
% Data completeness (in resolution range)	(Not available) (8.00-2.50) 72.0 (37.98-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$	-	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.179 , (Not available) 0.170 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	35.3	Xtriage
Anisotropy	0.070	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 71.0	EDS
L-test for twinning[1]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.045 for -h,-k,l 0.499 for h,-h-k,-l 0.045 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	3953	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 6.18% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]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

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

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	1.09	2/1672 (0.1%)	1.29	12/2271 (0.5%)
1	B	1.11	2/1672 (0.1%)	1.30	12/2271 (0.5%)
2	I	1.31	3/416 (0.7%)	1.44	2/558 (0.4%)
All	All	1.13	7/3760 (0.2%)	1.31	26/5100 (0.5%)

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	B	0	1
2	I	0	3
All	All	0	4

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	47	ARG	C-O	8.16	1.38	1.23
1	A	213	VAL	CA-CB	6.46	1.68	1.54
1	B	213	VAL	CA-CB	6.28	1.68	1.54
1	B	237	TRP	CG-CD2	-6.22	1.33	1.43
1	A	237	TRP	CG-CD2	-6.16	1.33	1.43
2	I	20	LYS	C-O	5.81	1.34	1.23
2	I	37	HIS	CA-CB	-5.77	1.41	1.53

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	237	TRP	CD1-CG-CD2	9.37	113.79	106.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	237	TRP	CD1-CG-CD2	9.34	113.77	106.30
1	B	141	TRP	CD1-CG-CD2	9.11	113.59	106.30
1	A	141	TRP	CD1-CG-CD2	9.10	113.58	106.30
1	B	215	TRP	CD1-CG-CD2	8.38	113.01	106.30
1	A	215	TRP	CD1-CG-CD2	8.19	112.85	106.30
1	A	51	TRP	CD1-CG-CD2	7.35	112.18	106.30
1	B	215	TRP	CE2-CD2-CG	-7.33	101.44	107.30
1	B	51	TRP	CD1-CG-CD2	7.30	112.14	106.30
1	A	215	TRP	CE2-CD2-CG	-7.12	101.60	107.30
1	B	141	TRP	CE2-CD2-CG	-6.71	101.93	107.30
1	A	141	TRP	CE2-CD2-CG	-6.65	101.98	107.30
1	A	117	ARG	NE-CZ-NH2	6.25	123.42	120.30
1	A	237	TRP	CE2-CD2-CG	-6.20	102.34	107.30
1	A	51	TRP	CE2-CD2-CG	-6.19	102.35	107.30
1	B	237	TRP	CE2-CD2-CG	-6.13	102.40	107.30
1	B	117	ARG	NE-CZ-NH2	6.11	123.35	120.30
1	B	51	TRP	CE2-CD2-CG	-6.11	102.42	107.30
2	I	36	CYS	N-CA-C	-5.93	94.98	111.00
1	B	141	TRP	CB-CG-CD1	-5.62	119.70	127.00
1	A	141	TRP	CB-CG-CD1	-5.57	119.76	127.00
2	I	47	ARG	NE-CZ-NH2	5.42	123.01	120.30
1	B	215	TRP	CG-CD2-CE3	5.37	138.73	133.90
1	A	215	TRP	CB-CG-CD1	-5.23	120.20	127.00
1	B	215	TRP	CB-CG-CD1	-5.22	120.21	127.00
1	A	215	TRP	CG-CD2-CE3	5.10	138.49	133.90

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	172	TYR	Sidechain
2	I	20	LYS	Mainchain
2	I	47	ARG	Mainchain
2	I	64	LYS	Peptide

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	1640	0	1593	26	1
1	B	1640	0	1593	26	1
2	I	409	0	376	19	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	125	0	0	2	1
4	B	125	0	0	2	1
4	I	12	0	0	0	0
All	All	3953	0	3562	61	2

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 9.

All (61) 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)
2:I:20:LYS:C	1:B:195:SER:OG	1.71	1.28
2:I:20:LYS:C	1:B:195:SER:HG	1.37	1.25
1:A:195:SER:HA	1:A:213:VAL:HG22	1.56	0.87
1:B:195:SER:HA	1:B:213:VAL:HG22	1.58	0.85
2:I:26:CYS:SG	2:I:62:CYS:HB2	2.22	0.79
1:A:195:SER:OG	2:I:47:ARG:C	2.27	0.73
1:A:217:TYR:CE1	2:I:44:ILE:HG22	2.32	0.65
2:I:31:ILE:HD13	2:I:54:ARG:HB2	1.80	0.63
1:A:79:ASN:HB3	1:A:117:ARG:NH2	2.14	0.63
1:B:79:ASN:HB3	1:B:117:ARG:NH2	2.14	0.63
1:A:48:ASN:HD22	1:A:50:GLN:H	1.47	0.61
1:B:48:ASN:HD22	1:B:50:GLN:H	1.47	0.61
1:A:48:ASN:ND2	1:A:50:GLN:H	1.99	0.60
1:B:48:ASN:ND2	1:B:50:GLN:H	1.99	0.60
2:I:24:PRO:O	2:I:60:ASP:HA	2.00	0.60
1:A:138:ILE:HG12	1:A:199:VAL:HG22	1.89	0.55
1:B:138:ILE:HG12	1:B:199:VAL:HG22	1.89	0.55
2:I:20:LYS:HA	1:B:192:GLN:HG3	1.90	0.54
1:A:195:SER:CB	2:I:47:ARG:C	2.78	0.52
1:A:183:VAL:HB	1:A:228:TYR:CE2	2.44	0.52
2:I:17:ASP:HB2	1:B:217:TYR:CE1	2.45	0.52
1:A:81:GLN:NE2	1:A:118:VAL:HG21	2.27	0.50
1:B:81:GLN:NE2	1:B:118:VAL:HG21	2.27	0.49
2:I:30:ASN:ND2	2:I:32:ARG:HH21	2.10	0.49
1:B:79:ASN:HB3	1:B:117:ARG:CZ	2.42	0.49
1:A:79:ASN:HB3	1:A:117:ARG:CZ	2.43	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:212:ILE:HB	1:B:229:THR:HB	1.95	0.48
1:A:212:ILE:HB	1:A:229:THR:HB	1.95	0.48
1:B:45:SER:OG	1:B:198:PRO:HB3	2.14	0.48
1:A:45:SER:OG	1:A:198:PRO:HB3	2.14	0.47
1:A:132:ALA:HA	1:A:162:VAL:HG23	1.95	0.47
1:B:132:ALA:HA	1:B:162:VAL:HG23	1.95	0.47
1:A:16:ILE:N	4:A:300:HOH:O	2.48	0.47
2:I:32:ARG:NH1	2:I:37:HIS:NE2	2.63	0.46
1:B:221(C):GLN:HB2	1:B:224:LYS:HB2	1.98	0.46
1:A:57:HIS:HB2	2:I:52:LYS:HE2	1.98	0.46
1:B:16:ILE:N	4:B:456:HOH:O	2.49	0.45
1:A:48:ASN:HD22	1:A:50:GLN:N	2.13	0.45
1:B:48:ASN:HD22	1:B:50:GLN:N	2.13	0.45
2:I:20:LYS:CA	1:B:195:SER:OG	2.60	0.45
2:I:18:CYS:SG	2:I:62:CYS:HB2	2.56	0.45
1:A:221(B):GLN:HB2	1:A:224:LYS:HB2	1.97	0.44
1:A:162:VAL:HG12	1:A:183:VAL:HG22	1.99	0.44
2:I:30:ASN:HD22	2:I:32:ARG:HH21	1.65	0.44
1:B:124:PRO:HB3	1:B:127:SER:O	2.16	0.44
1:A:124:PRO:HB3	1:A:127:SER:O	2.17	0.44
1:B:172:TYR:HE2	1:B:217:TYR:CD1	2.35	0.44
1:B:162:VAL:HG12	1:B:183:VAL:HG22	1.99	0.44
1:B:88(C):LYS:HE2	4:B:485:HOH:O	2.17	0.43
1:B:137:LEU:HD11	1:B:157:CYS:HB3	2.00	0.42
1:B:183:VAL:HB	1:B:228:TYR:CE2	2.53	0.42
2:I:20:LYS:O	1:B:195:SER:OG	2.08	0.42
1:A:213:VAL:HA	1:A:228:TYR:CD1	2.55	0.42
1:A:137:LEU:HD11	1:A:157:CYS:HB3	2.01	0.42
2:I:13:CYS:SG	2:I:15:SER:O	2.78	0.42
1:A:188(B):LYS:HE2	4:A:337:HOH:O	2.20	0.41
1:A:143:ASN:OD1	1:A:144:THR:N	2.53	0.41
2:I:32:ARG:NH1	2:I:37:HIS:CE1	2.88	0.41
1:A:172:TYR:CE2	1:A:224:LYS:HD2	2.56	0.40
1:A:22:CYS:O	1:A:23:ALA:HB3	2.21	0.40
1:B:22:CYS:O	1:B:23:ALA:HB3	2.21	0.40

All (2) 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	Interatomic distance (Å)	Clash overlap (Å)
1:A:59:TYR:OH	4:B:465:HOH:O[2_665]	2.09	0.11

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:59:TYR:OH	4:A:309:HOH:O[3_574]	2.13	0.07

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	221/223 (99%)	205 (93%)	15 (7%)	1 (0%)	29 48
1	B	221/223 (99%)	207 (94%)	13 (6%)	1 (0%)	29 48
2	I	53/72 (74%)	42 (79%)	9 (17%)	2 (4%)	3 4
All	All	495/518 (96%)	454 (92%)	37 (8%)	4 (1%)	19 35

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	I	61	PHE
2	I	65	PRO
1	A	78	GLY
1	B	78	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	182/182 (100%)	167 (92%)	15 (8%)	11 22
1	B	182/182 (100%)	167 (92%)	15 (8%)	11 22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	I	52/69 (75%)	44 (85%)	8 (15%)	2 5
All	All	416/433 (96%)	378 (91%)	38 (9%)	9 18

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	48	ASN
1	A	50	GLN
1	A	66	ARG
1	A	79	ASN
1	A	105	LEU
1	A	113	THR
1	A	120	THR
1	A	121	VAL
1	A	123	LEU
1	A	125	ARG
1	A	146	SER
1	A	213	VAL
1	A	223	ASN
1	A	239	GLN
1	A	240	GLN
2	I	17	ASP
2	I	21	SER
2	I	27	HIS
2	I	36	CYS
2	I	38	SER
2	I	42	SER
2	I	46	THR
2	I	65	PRO
1	B	48	ASN
1	B	50	GLN
1	B	66	ARG
1	B	79	ASN
1	B	105	LEU
1	B	113	THR
1	B	120	THR
1	B	121	VAL
1	B	123	LEU
1	B	125	ARG
1	B	146	SER
1	B	213	VAL
1	B	223	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	239	GLN
1	B	240	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	30	GLN
1	A	48	ASN
1	A	50	GLN
1	A	101	ASN
1	A	175	GLN
1	A	210	GLN
1	B	30	GLN
1	B	48	ASN
1	B	50	GLN
1	B	97	ASN
1	B	101	ASN
1	B	210	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

There are no monosaccharides 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.

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

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	219/223 (98%)	-0.67	0	100	100	12, 31, 52, 64	0
1	B	223/223 (100%)	-0.67	0	100	100	2, 30, 52, 64	0
2	I	50/72 (69%)	-0.72	0	100	100	13, 25, 45, 52	0
All	All	492/518 (94%)	-0.68	0	100	100	2, 30, 52, 64	0

There are no RSRZ outliers to report.

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 monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} 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
3	CA	B	246	1/1	0.96	0.03	48,48,48,48	0
3	CA	A	299	1/1	1.00	0.06	17,17,17,17	0

6.5 Other polymers (i)

There are no such residues in this entry.