



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 05:16 PM UTC

PDB ID : 6MDZ / pdb_00006mdz
Title : Human Argonaute2-miR-122 bound to a target RNA with two central mismatches (bu2)
Authors : Sheu-Gruttadauria, J.; MacRae, I.J.
Deposited on : 2018-09-05
Resolution : 3.40 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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
Mogul : 2022.3.0, CSD as543be (2022)
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

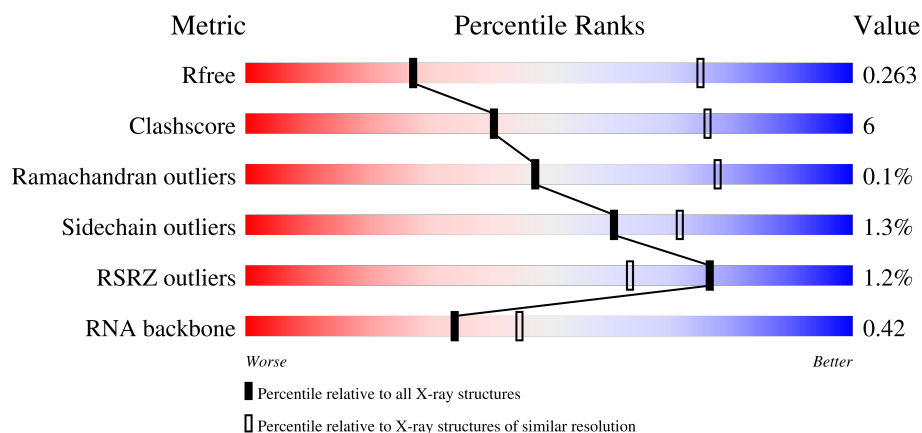
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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	1001 (3.44-3.36)
Clashscore	190562	1022 (3.44-3.36)
Ramachandran outliers	187476	1012 (3.44-3.36)
Sidechain outliers	187428	1012 (3.44-3.36)
RSRZ outliers	180081	1001 (3.44-3.36)
RNA backbone	3983	1157 (3.80-3.00)

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 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 $\leq 5\%$. 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	859	% 77% 16% 6%
1	B	859	% 76% 16% 7%
2	C	21	5% 48% 24% 24% 5%
2	D	21	5% 19% 33% 14% 33%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	23	 4% 74% 26%
3	F	23	 4% 61% 9% 26%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 14382 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 Protein argonaute-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	804	Total	C	N	O	S	0	0	0
			6446	4106	1158	1141	41			
1	B	796	Total	C	N	O	S	0	0	0
			6388	4073	1148	1128	39			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	387	ASP	SER	engineered mutation	UNP Q9UKV8
A	669	ALA	ASP	engineered mutation	UNP Q9UKV8
A	824	ALA	SER	engineered mutation	UNP Q9UKV8
A	828	ASP	SER	engineered mutation	UNP Q9UKV8
A	831	ASP	SER	engineered mutation	UNP Q9UKV8
A	834	ALA	SER	engineered mutation	UNP Q9UKV8
B	387	ASP	SER	engineered mutation	UNP Q9UKV8
B	669	ALA	ASP	engineered mutation	UNP Q9UKV8
B	824	ALA	SER	engineered mutation	UNP Q9UKV8
B	828	ASP	SER	engineered mutation	UNP Q9UKV8
B	831	ASP	SER	engineered mutation	UNP Q9UKV8
B	834	ALA	SER	engineered mutation	UNP Q9UKV8

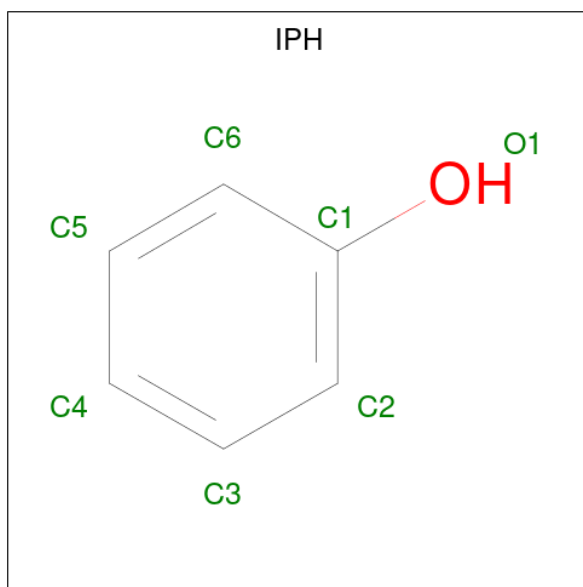
- Molecule 2 is a RNA chain called RNA (5'-R(P*UP*GP*GP*AP*GP*UP*GP*UP*GP*AP*CP*AP*AP*UP*GP*GP*UP*GP*UP*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	20	Total	C	N	O	P	0	0	0
			433	192	77	144	20			
2	D	14	Total	C	N	O	P	0	0	0
			308	136	58	100	14			

- Molecule 3 is a RNA chain called RNA (5'-R(P*AP*AP*AP*CP*AP*CP*CP*AP*UP*UP*UP*CP*CP*AP*CP*AP*CP*UP*CP*CP*AP*AP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	23	Total	C	N	O	P	0	0	0
			480	217	85	155	23			
3	F	17	Total	C	N	O	P	0	0	1
			299	131	50	102	16			

- Molecule 4 is PHENOL (CCD ID: IPH) (formula: C₆H₆O).

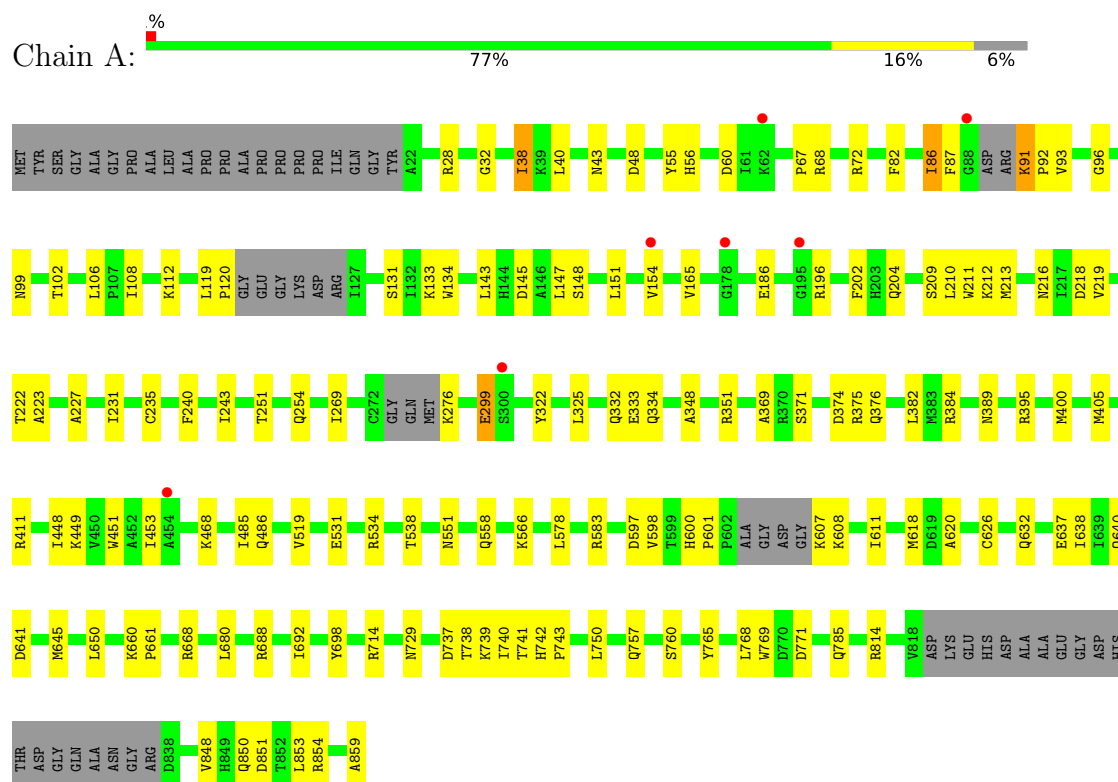


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			7	6	1		
4	A	1	Total	C	O	0	0
			7	6	1		
4	B	1	Total	C	O	0	0
			7	6	1		
4	B	1	Total	C	O	0	0
			7	6	1		

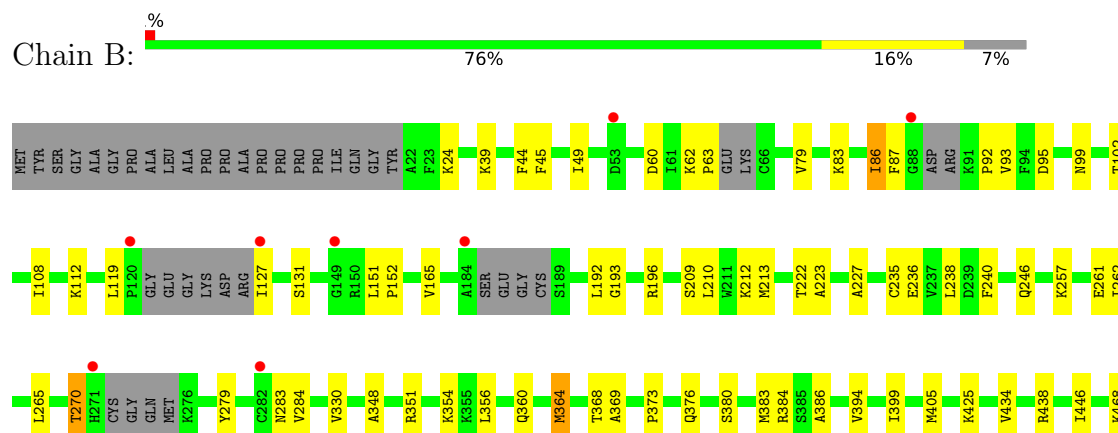
3 Residue-property plots

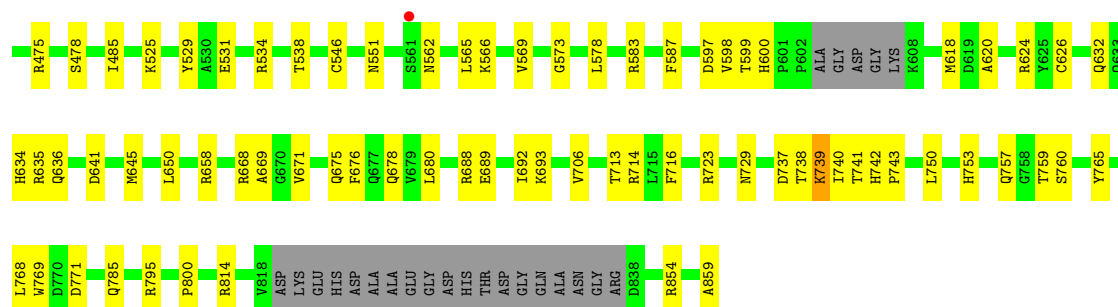
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Protein argonaute-2

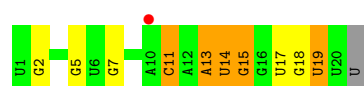


• Molecule 1: Protein argonaute-2

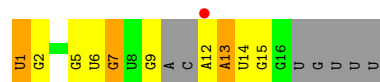
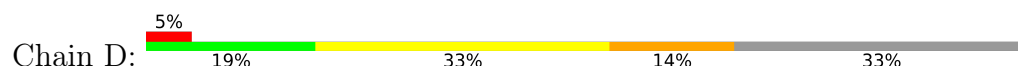




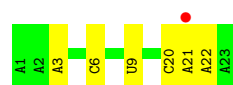
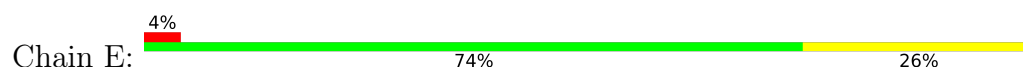
- Molecule 2: RNA (5'-R(P*UP*GP*GP*AP*GP*UP*GP*UP*GP*AP*CP*AP*AP*UP*GP*GP*UP*GP*UP*U)-3')



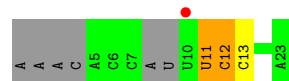
- Molecule 2: RNA (5'-R(P*UP*GP*GP*AP*GP*UP*GP*UP*GP*AP*CP*AP*AP*UP*GP*GP*UP*GP*UP*U)-3')



- Molecule 3: RNA (5'-R(P*AP*AP*AP*CP*AP*CP*CP*AP*UP*UP*UP*CP*CP*AP*CP*A*P*CP*UP*CP*CP*AP*AP*A)-3')



- Molecule 3: RNA (5'-R(P*AP*AP*AP*CP*AP*CP*CP*AP*UP*UP*UP*CP*CP*AP*CP*A*P*CP*UP*CP*CP*AP*AP*A)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	107.61Å 138.29Å 152.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.29 – 3.40 38.29 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.3 (38.29-3.40) 99.3 (38.29-3.40)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.00 (at 3.40Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.220 , 0.264 0.222 , 0.263	Depositor DCC
R_{free} test set	1589 reflections (4.12%)	wwPDB-VP
Wilson B-factor (Å ²)	69.3	Xtriage
Anisotropy	0.233	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 36.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	14382	wwPDB-VP
Average B, all atoms (Å ²)	63.0	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 24.72 % 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 3.6319e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: IPH

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	0.27	0/6598	0.71	2/8930 (0.0%)
1	B	0.28	0/6538	0.72	2/8849 (0.0%)
2	C	0.08	0/484	0.22	0/752
2	D	0.11	0/344	0.37	1/532 (0.2%)
3	E	0.13	0/535	0.29	0/828
3	F	0.15	0/329	0.33	0/506
All	All	0.26	0/14828	0.68	5/20397 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1	U	O3'-P-O5'	-5.59	95.62	104.00
1	B	24	LYS	CA-C-N	5.58	123.73	119.66
1	B	24	LYS	C-N-CA	5.58	123.73	119.66
1	A	96	GLY	N-CA-C	-5.28	108.06	114.92
1	A	299	GLU	CB-CA-C	-5.07	110.33	117.23

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	6446	0	6513	87	1
1	B	6388	0	6455	90	1
2	C	433	0	214	9	0
2	D	308	0	152	11	0
3	E	480	0	250	3	0
3	F	299	0	153	3	0
4	A	14	0	12	2	0
4	B	14	0	12	2	0
All	All	14382	0	13761	182	1

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 6.

The worst 5 of 182 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:B:531:GLU:OE2	1:B:534:ARG:NH1	1.98	0.96
1:B:566:LYS:NZ	1:B:859:ALA:OXT	2.02	0.92
1:A:351:ARG:NH2	2:C:11:C:OP1	2.01	0.92
1:B:583:ARG:NH1	1:B:620:ALA:O	2.19	0.75
1:B:618:MET:HE3	1:B:626:CYS:HB2	1.70	0.74

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	Interatomic distance (Å)	Clash overlap (Å)
1:A:299:GLU:OE2	1:B:438:ARG:NH2[4_455]	2.16	0.04

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	792/859 (92%)	747 (94%)	45 (6%)	0	100	100
1	B	780/859 (91%)	742 (95%)	37 (5%)	1 (0%)	48	78
All	All	1572/1718 (92%)	1489 (95%)	82 (5%)	1 (0%)	48	78

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	739	LYS

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	713/749 (95%)	702 (98%)	11 (2%)	57	69
1	B	706/749 (94%)	698 (99%)	8 (1%)	65	74
All	All	1419/1498 (95%)	1400 (99%)	19 (1%)	61	71

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	270	THR
1	B	468	LYS
1	B	525	LYS
1	B	368	THR
1	A	468	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	640	GLN
1	B	652	GLN
1	A	568	ASN
1	A	553	GLN
1	B	729	ASN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	19/21 (90%)	6 (31%)	2 (10%)
2	D	12/21 (57%)	4 (33%)	0
3	E	22/23 (95%)	3 (13%)	0
3	F	12/23 (52%)	1 (8%)	1 (8%)
All	All	65/88 (73%)	14 (21%)	3 (4%)

5 of 14 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	11	C
2	C	14	U
2	C	15	G
2	C	17	U
2	C	18	G

All (3) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	C	13	A
2	C	14	U
3	F	11	U

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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

4 ligands are modelled in this entry.

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	$\# Z > 2$
4	IPH	A	901	-	7,7,7	0.34	0	8,8,8	0.31	0
4	IPH	B	902	-	7,7,7	0.37	0	8,8,8	0.27	0
4	IPH	B	901	-	7,7,7	0.38	0	8,8,8	0.26	0
4	IPH	A	902	-	7,7,7	0.37	0	8,8,8	0.28	0

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
4	IPH	A	901	-	-	-	0/1/1/1
4	IPH	B	902	-	-	-	0/1/1/1
4	IPH	B	901	-	-	-	0/1/1/1
4	IPH	A	902	-	-	-	0/1/1/1

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.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	901	IPH	2	0
4	B	902	IPH	1	0
4	B	901	IPH	1	0

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	804/859 (93%)	0.01	7 (0%) 81 68	16, 55, 103, 137	0
1	B	796/859 (92%)	0.13	9 (1%) 78 65	16, 58, 120, 163	0
2	C	20/21 (95%)	0.27	1 (5%) 34 25	33, 113, 132, 133	0
2	D	14/21 (66%)	0.44	1 (7%) 22 18	29, 59, 176, 180	0
3	E	23/23 (100%)	0.33	1 (4%) 40 29	47, 99, 140, 148	0
3	F	17/23 (73%)	0.34	1 (5%) 28 22	47, 69, 162, 171	0
All	All	1674/1806 (92%)	0.08	20 (1%) 76 63	16, 57, 116, 180	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	454	ALA	4.6
1	A	300	SER	4.0
1	A	88	GLY	3.4
1	A	195	GLY	2.8
1	B	184	ALA	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 oligosaccharides 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
4	IPH	B	901	7/7	0.81	0.34	68,72,81,89	0
4	IPH	B	902	7/7	0.87	0.14	43,49,55,67	0
4	IPH	A	901	7/7	0.91	0.18	59,60,71,77	0
4	IPH	A	902	7/7	0.93	0.12	39,42,47,48	0

6.5 Other polymers [i](#)

There are no such residues in this entry.