W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Oct 24, 2024 – 12:17 AM EDT

PDB ID	:	2GBF
Title	:	rat dpp-IV with alkynyl cyanopyrrolidine #1
Authors	:	Longenecker, K.L.; Jakob, C.G.; Fry, E.H.; Wilk, S.
Deposited on	:	2006-03-10
Resolution	:	3.10 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 (i) 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 (i)) were used in the production of this report:

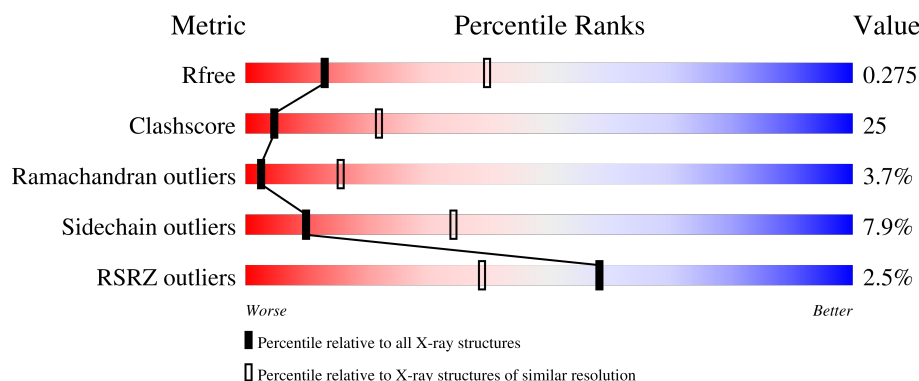|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 3.10 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		0.275
Clashscore		25
Ramachandran outliers		3.7%
Sidechain outliers		7.9%
RSRZ outliers		2.5%

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}	164625	1351 (3.10-3.10)
Clashscore	180529	1454 (3.10-3.10)
Ramachandran outliers	177936	1391 (3.10-3.10)
Sidechain outliers	177891	1391 (3.10-3.10)
RSRZ outliers	164620	1351 (3.10-3.10)

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	730	2% 54% 40% 6% •
1	B	730	3% 54% 39% 6% •

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 11858 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 Dipeptidyl peptidase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	730	Total 5920	C 3789	N 981	O 1124	S 26	0	0	0
1	B	730	Total 5920	C 3789	N 981	O 1124	S 26	0	0	0

- Molecule 2 is (1S)-2-[(2S,5R)-2-(AMINOMETHYL)-5-ETHYNYLPYRROLIDIN-1-YL]-1-CYCLOPENTYL-2-OXOETHANAMINE (three-letter code: AIA) (formula: $C_{14}H_{23}N_3O$).
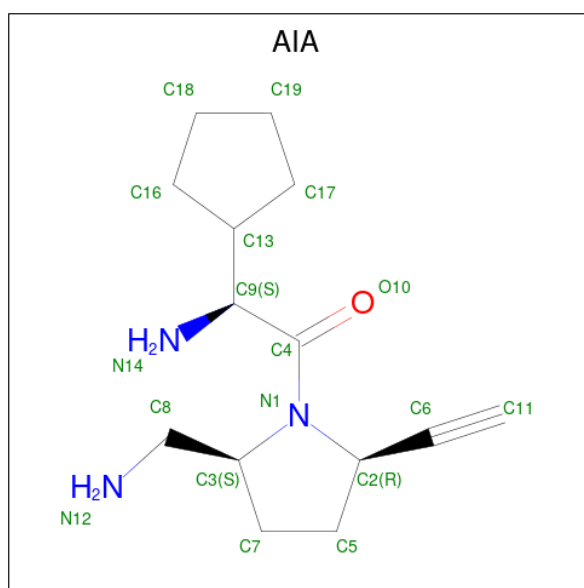

| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total 18	C 14	N 3	O 1	0	0

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: Dipeptidyl peptidase 4

Chain A:

- Molecule 1: Dipeptidyl peptidase 4

Chain B:
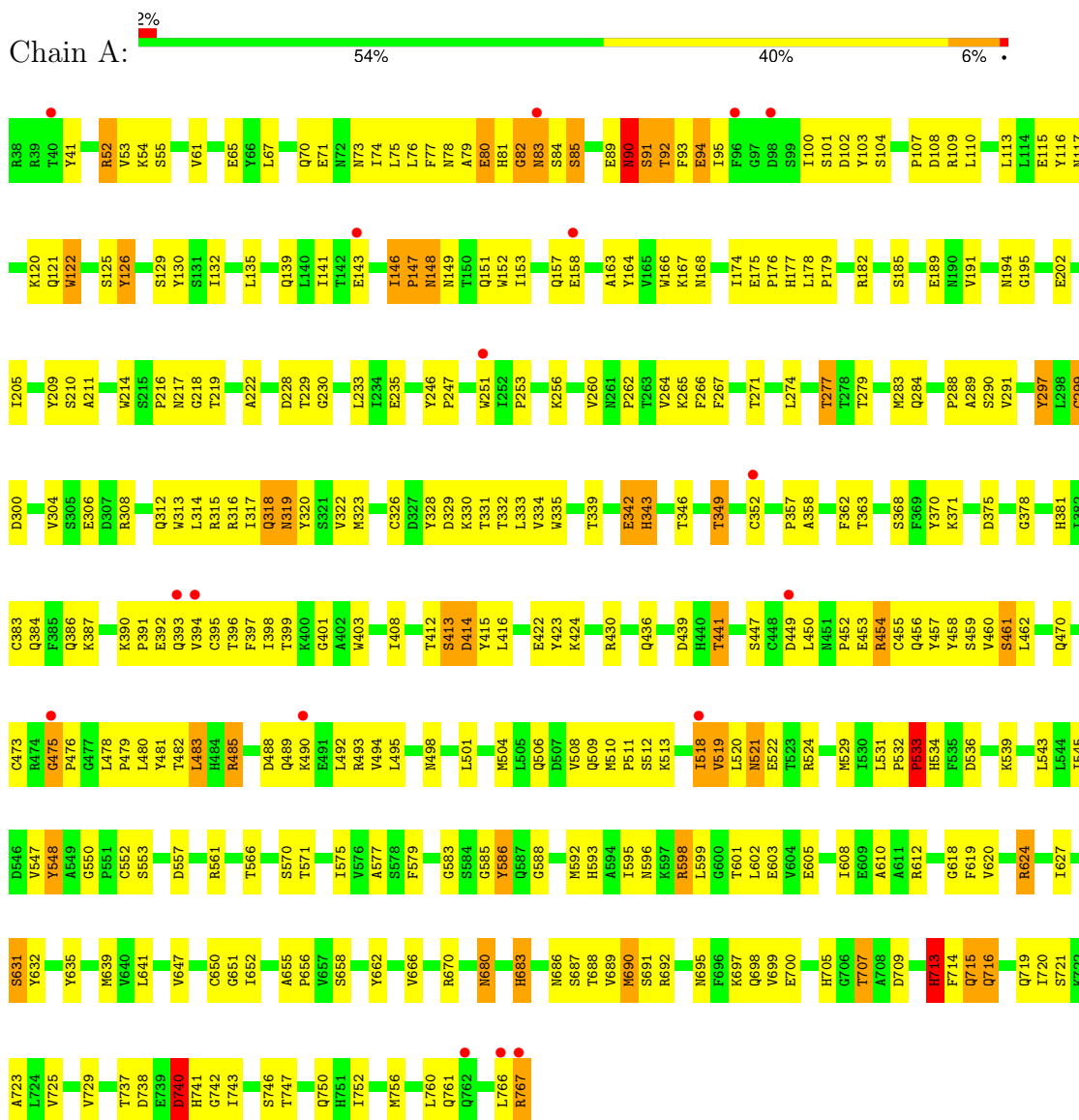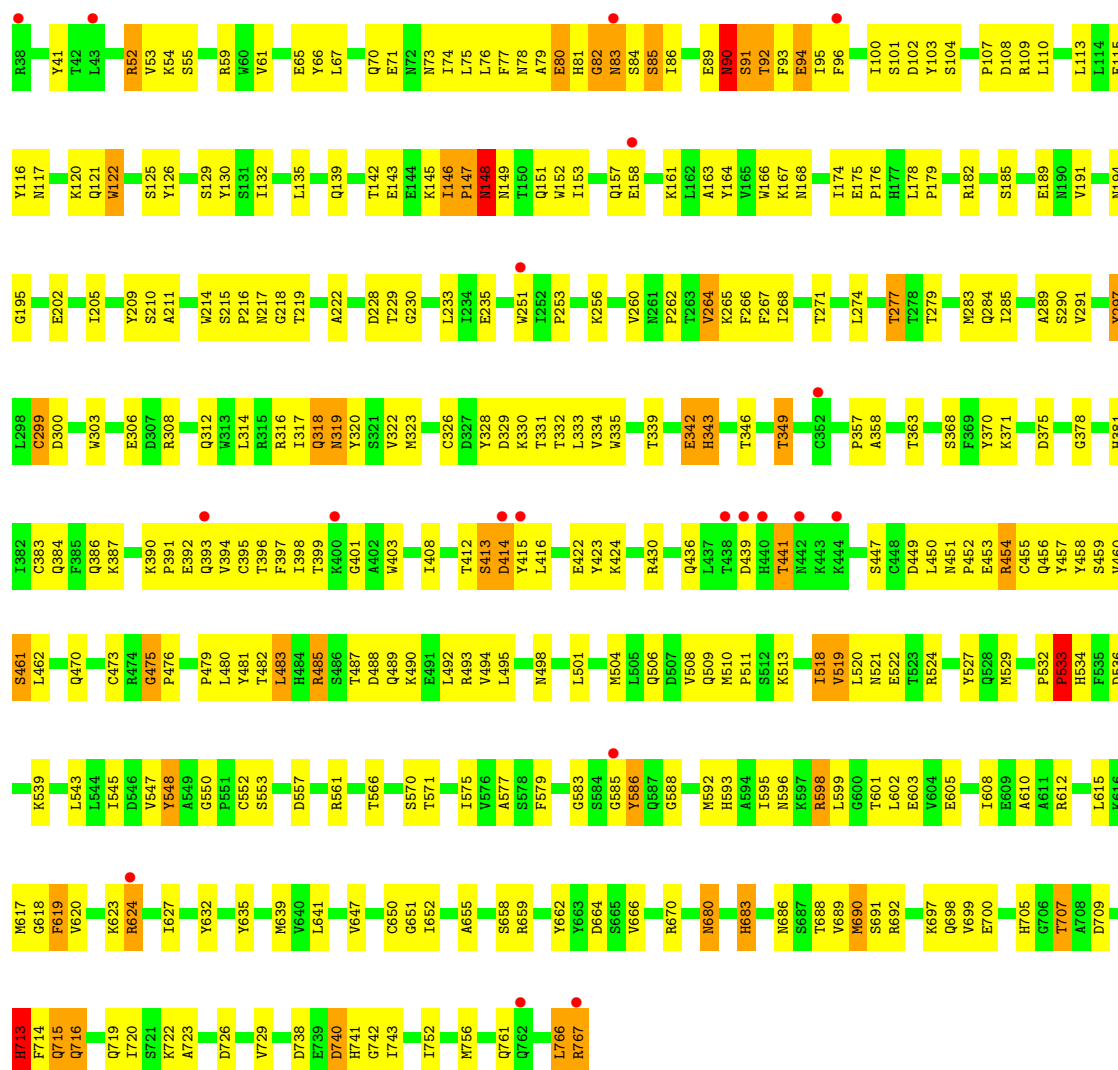
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α, β, γ	208.17Å 208.17Å 208.17Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.10 20.00 – 3.10	Depositor EDS
% Data completeness (in resolution range)	98.8 (20.00-3.10) 98.9 (20.00-3.10)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.57 (at 3.09Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.250 , 0.289 0.237 , 0.275	Depositor DCC
R_{free} test set	2743 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	73.1	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 57.6	EDS
L-test for twinning[2]	$< |L| > = 0.54, < L^2 > = 0.38$	Xtriage
Estimated twinning fraction	0.000 for l,-k,h	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	11858	wwPDB-VP
Average B, all atoms (Å^2)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.78% 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)

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

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	0.46	2/6088 (0.0%)	0.71	2/8278 (0.0%)
1	B	0.43	0/6088	0.70	2/8278 (0.0%)
All	All	0.45	2/12176 (0.0%)	0.71	4/16556 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	631	SER	C-O	8.88	1.40	1.23
1	A	352	CYS	CB-SG	-5.14	1.73	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	713	HIS	N-CA-C	7.29	130.67	111.00
1	A	713	HIS	N-CA-C	7.22	130.50	111.00
1	A	90	ASN	N-CA-C	-5.15	97.10	111.00
1	B	90	ASN	N-CA-C	-5.12	97.18	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	5920	0	5631	294	0
1	B	5920	0	5632	292	0
2	A	18	0	21	0	0
All	All	11858	0	11284	584	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 25.

The worst 5 of 584 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:761:GLN:HB3	1:A:767:ARG:HB3	1.39	1.03
1:B:349:THR:HB	1:B:593:HIS:HD2	1.26	0.99
1:A:460:VAL:HG22	1:A:461:SER:H	1.28	0.99
1:A:349:THR:HB	1:A:593:HIS:HD2	1.25	0.98
1:B:761:GLN:HB3	1:B:767:ARG:HB3	1.42	0.98

There are no symmetry-related clashes.

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	728/730 (100%)	614 (84%)	87 (12%)	27 (4%)	2	16
1	B	728/730 (100%)	615 (84%)	86 (12%)	27 (4%)	2	16
All	All	1456/1460 (100%)	1229 (84%)	173 (12%)	54 (4%)	2	16

5 of 54 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	80	GLU
1	A	143	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	450	LEU
1	A	521	ASN
1	A	533	PRO

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	643/651 (99%)	594 (92%)	49 (8%)	11	36
1	B	643/651 (99%)	591 (92%)	52 (8%)	9	33
All	All	1286/1302 (99%)	1185 (92%)	101 (8%)	10	34

5 of 101 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	175	GLU
1	B	395	CYS
1	B	740	ASP
1	B	260	VAL
1	B	300	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	470	GLN
1	B	698	GLN
1	B	484	HIS
1	B	607	GLN
1	B	715	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 oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

1 ligand is 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	$\#|Z| > 2$
2	AIA	A	768	1	18,19,19	1.18	2 (11%)	17,26,26	1.76	4 (23%)

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
2	AIA	A	768	1	-	2/14/36/36	0/2/2/2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	768	AIA	C8-N12	-2.64	1.35	1.48
2	A	768	AIA	C9-C4	2.61	1.55	1.53

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	768	AIA	C13-C9-C4	4.98	116.56	110.33
2	A	768	AIA	C7-C3-N1	2.43	104.32	101.68

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	768	AIA	C4-C9-N14	-2.42	106.01	110.09
2	A	768	AIA	C7-C3-C8	-2.24	103.82	112.17

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	768	AIA	C7-C3-C8-N12
2	A	768	AIA	N1-C3-C8-N12

There are no ring outliers.

No monomer is involved in short contacts.

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	730/730 (100%)	-0.13	17 (2%)	61	42	32, 58, 91, 165	0
1	B	730/730 (100%)	0.05	20 (2%)	56	36	35, 65, 104, 161	0
All	All	1460/1460 (100%)	-0.04	37 (2%)	58	39	32, 61, 99, 165	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	96	PHE	4.0
1	B	440	HIS	3.8
1	B	38	ARG	3.8
1	B	442	ASN	3.7
1	B	438	THR	3.6

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, 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
2	AIA	A	768	18/18	0.94	0.16	65,65,65,65	0

6.5 Other polymers ⓘ

There are no such residues in this entry.