wwPDB X-ray Structure Validation Summary Report (i)

Feb 6, 2024 – 08:19 PM EST

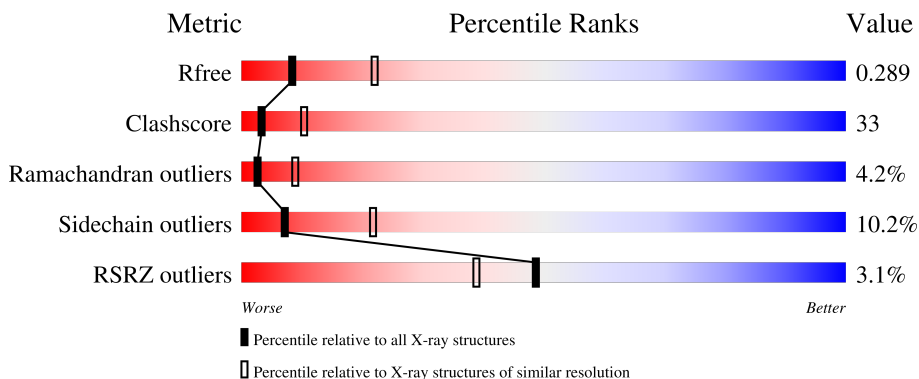|		
PDB ID	:	2B5N
Title	:	Crystal Structure of the DDB1 BPB Domain
Authors	:	Li, T.; Chen, X.; Garbutt, K.C.; Zhou, P.; Zheng, N.
Deposited on	:	2005-09-28
Resolution	:	2.80 Å(reported)

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 2.80 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		0.289
Clashscore		33
Ramachandran outliers		4.2%
Sidechain outliers		10.2%
RSRZ outliers		3.1%

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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	323	3% / 46% / 42% / 7% · ·
1	B	323	3% / 51% / 39% / 8% ··
1	C	323	3% / 48% / 39% / 11% ··
1	D	323	3% / 48% / 45% / 6% ··

WORLDWIDE
wwPDB
PROTEIN DATA BANK

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9893 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 damage-specific DNA binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	312	Total	C	N	O	S	0	0	0
			2421	1535	407	466	13			
1	B	321	Total	C	N	O	S	0	0	0
			2486	1570	422	480	14			
1	C	320	Total	C	N	O	S	0	0	0
			2478	1564	421	479	14			
1	D	320	Total	C	N	O	S	0	0	0
			2482	1568	421	479	14			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	387	GLY	-	cloning artifact	GB 13435359
A	388	SER	-	cloning artifact	GB 13435359
A	389	HIS	-	cloning artifact	GB 13435359
A	390	MET	-	initiating methionine	GB 13435359
B	387	GLY	-	cloning artifact	GB 13435359
B	388	SER	-	cloning artifact	GB 13435359
B	389	HIS	-	cloning artifact	GB 13435359
B	390	MET	-	initiating methionine	GB 13435359
C	387	GLY	-	cloning artifact	GB 13435359
C	388	SER	-	cloning artifact	GB 13435359
C	389	HIS	-	cloning artifact	GB 13435359
C	390	MET	-	initiating methionine	GB 13435359
D	387	GLY	-	cloning artifact	GB 13435359
D	388	SER	-	cloning artifact	GB 13435359
D	389	HIS	-	cloning artifact	GB 13435359
D	390	MET	-	initiating methionine	GB 13435359

- Molecule 2 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C_3H_8O).

IPA

C3

C2

C1

OH

O2

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	B	1	Total 4	C 3	O 1	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	4	Total 4	O 4	0	0
3	B	6	Total 6	O 6	0	0
3	C	6	Total 6	O 6	0	0
3	D	6	Total 6	O 6	0	0

WORLDWIDE
PDB
PROTEIN DATA BANK

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.
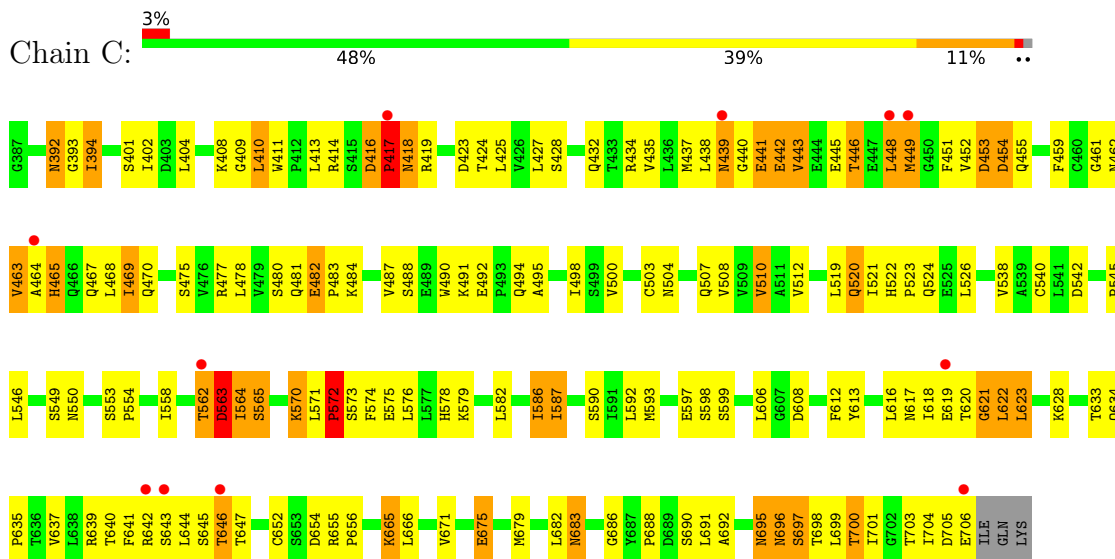
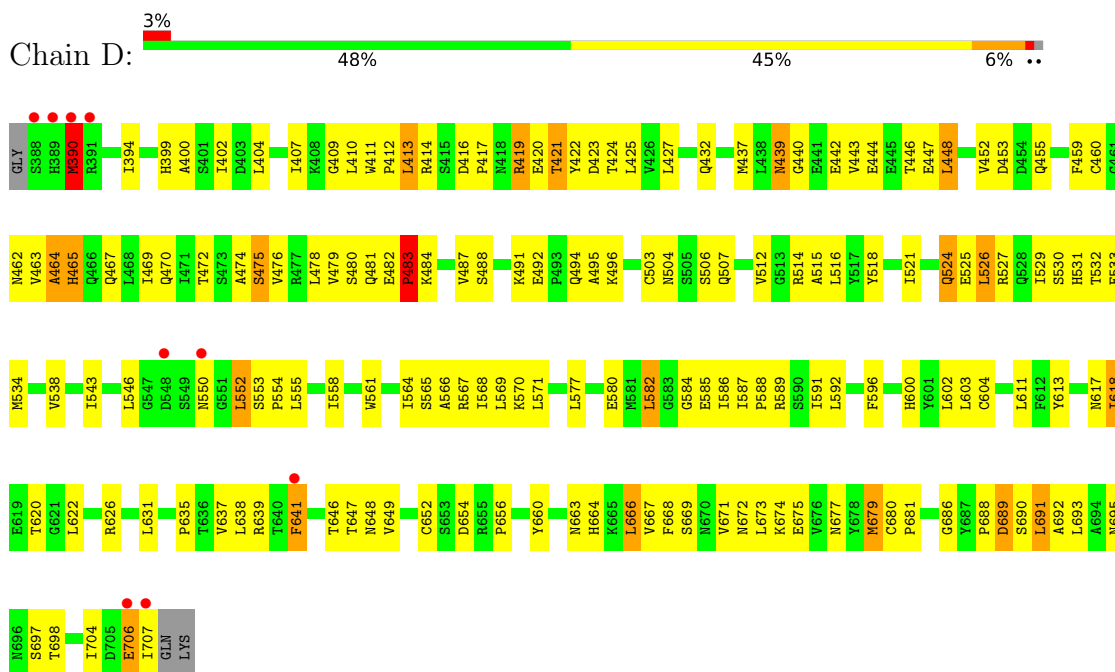• Molecule 1: damage-specific DNA binding protein 1
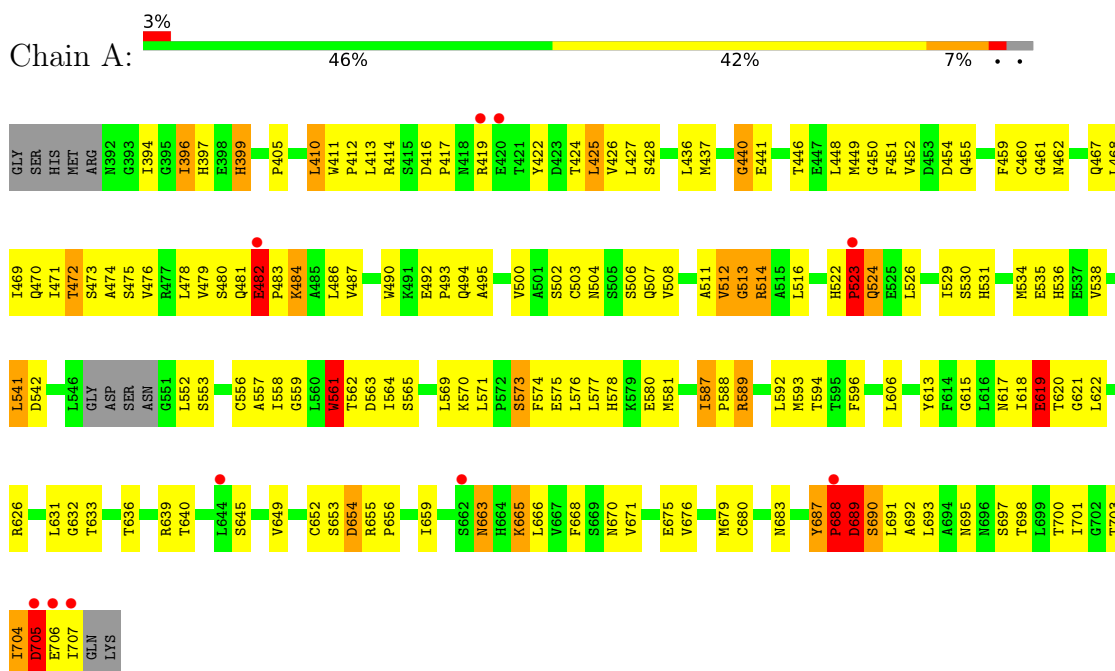
• Molecule 1: damage-specific DNA binding protein 1

• Molecule 1: damage-specific DNA binding protein 1

Chain C:

• Molecule 1: damage-specific DNA binding protein 1

Chain D:
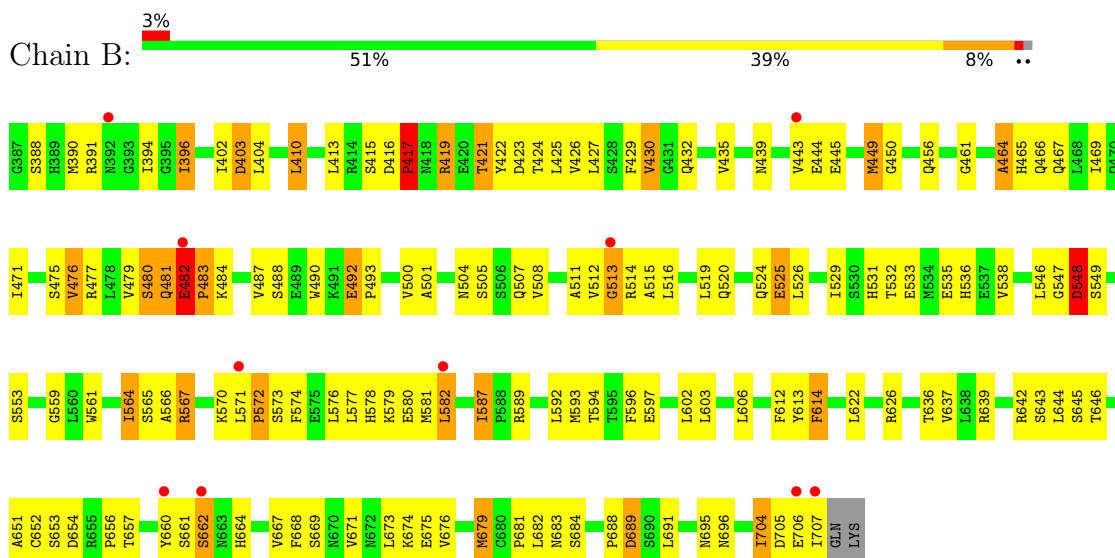
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	101.54Å 73.75Å 136.78Å 90.00° 111.73° 90.00°	Depositor
Resolution (Å)	48.80 – 2.80 48.77 – 2.54	Depositor EDS
% Data completeness (in resolution range)	(Not available) (48.80-2.80) 98.9 (48.77-2.54)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	4.75 (at 2.54Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.236 , 0.296 0.239 , 0.289	Depositor DCC
R_{free} test set	2401 reflections (3.86%)	wwPDB-VP
Wilson B-factor (Å²)	33.2	Xtriage
Anisotropy	0.543	Xtriage
Bulk solvent k_{sol}(e/Å³), B_{sol}(Å²)	0.30 , 1.4	EDS
L-test for twinning [2]	$<|L|>$ = 0.44, $<L^2>$ = 0.26	Xtriage
Estimated twinning fraction	0.164 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	9893	wwPDB-VP
Average B, all atoms (Å²)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% 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 ⓘ

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

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	0.43	0/2467	0.87	10/3354 (0.3%)
1	B	0.38	0/2534	0.81	4/3444 (0.1%)
1	C	0.37	0/2526	0.98	11/3433 (0.3%)
1	D	0.32	0/2530	0.73	0/3439
All	All	0.38	0/10057	0.85	25/13670 (0.2%)

There are no bond length outliers.

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	571	LEU	C-N-CD	-20.05	76.49	120.60
1	C	442	GLU	N-CA-C	-16.93	65.29	111.00
1	C	571	LEU	C-N-CA	13.74	179.69	122.00
1	C	442	GLU	CB-CA-C	12.23	134.87	110.40
1	B	417	PRO	CA-N-CD	-11.36	95.60	111.50

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	2421	0	2421	154	0
1	B	2486	0	2477	144	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2478	0	2466	177	1
1	D	2482	0	2474	175	1
2	B	4	0	8	2	0
3	A	4	0	0	0	0
3	B	6	0	0	2	1
3	C	6	0	0	0	0
3	D	6	0	0	3	0
All	All	9893	0	9846	646	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 33.

The worst 5 of 646 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:D:465:HIS:N	3:D:16:HOH:O	1.64	1.25
1:A:683:ASN:OD1	1:A:688:PRO:O	1.61	1.17
1:A:417:PRO:HG3	1:A:481:GLN:HG2	1.25	1.16
1:A:482:GLU:HB2	1:A:483:PRO:HD3	1.22	1.15
1:A:689:ASP:O	1:A:689:ASP:OD2	1.65	1.15

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:B:388:SER:O	1:D:524:GLN:O[2_544]	1.97	0.23
1:C:524:GLN:O	3:B:9:HOH:O[1_454]	2.08	0.12

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	308/323 (95%)	253 (82%)	40 (13%)	15 (5%)	2	7
1	B	319/323 (99%)	271 (85%)	35 (11%)	13 (4%)	3	9
1	C	318/323 (98%)	267 (84%)	36 (11%)	15 (5%)	2	7
1	D	318/323 (98%)	260 (82%)	48 (15%)	10 (3%)	4	14
All	All	1263/1292 (98%)	1051 (83%)	159 (13%)	53 (4%)	3	9

5 of 53 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	482	GLU
1	A	514	ARG
1	A	561	TRP
1	A	645	SER
1	B	449	MET

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	276/285 (97%)	244 (88%)	32 (12%)	5	17
1	B	283/285 (99%)	254 (90%)	29 (10%)	7	22
1	C	282/285 (99%)	249 (88%)	33 (12%)	5	16
1	D	283/285 (99%)	262 (93%)	21 (7%)	13	37
All	All	1124/1140 (99%)	1009 (90%)	115 (10%)	7	22

5 of 115 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	679	MET
1	D	641	PHE
1	C	449	MET
1	D	622	LEU
1	D	439	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 53 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	462	ASN
1	C	617	ASN
1	D	617	ASN
1	C	466	GLN
1	C	520	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 ⓘ

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	IPA	B	1001	-	3,3,3	0.26	0	3,3,3	1.81	1 (33%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1001	IPA	C3-C2-C1	-2.79	92.38	113.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1001	IPA	2	0

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	312/323 (96%)	-0.02	10 (3%)	47	37	10, 32, 77, 124	0
1	B	321/323 (99%)	-0.10	10 (3%)	49	39	6, 25, 77, 129	0
1	C	320/323 (99%)	0.09	11 (3%)	45	35	4, 27, 75, 130	0
1	D	320/323 (99%)	-0.01	9 (2%)	53	43	8, 30, 85, 141	0
All	All	1273/1292 (98%)	-0.01	40 (3%)	49	39	4, 28, 80, 141	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	388	SER	6.4
1	C	448	LEU	4.9
1	D	389	HIS	4.8
1	D	390	MET	4.6
1	D	391	ARG	4.3

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.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($Å^2$)	Q<0.9
2	IPA	B	1001	4/4	0.78	0.34	20,20,20,20	0

6.5　Other polymers ⓘ

There are no such residues in this entry.