WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2024 – 06:33 PM EST

PDB ID	:	3QZ2
Title	:	The structure of cysteine-free human insulin degrading enzyme
Authors	:	Guo, Q.; Tang, W.J.
Deposited on	:	2011-03-04
Resolution	:	3.20 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

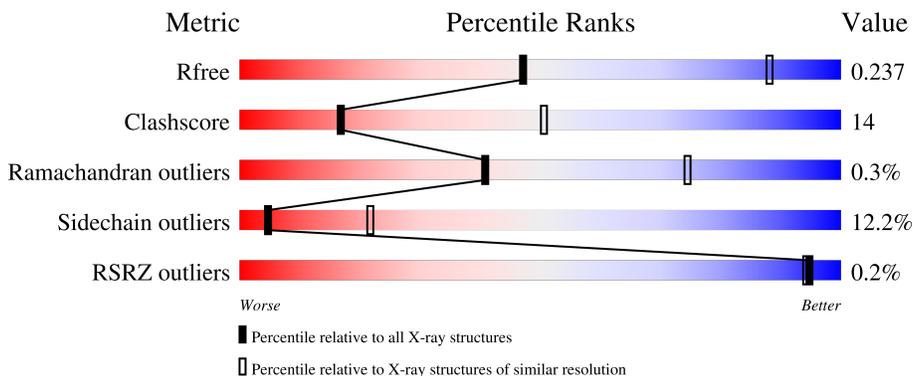|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.20 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		0.237
Clashscore		14
Ramachandran outliers		0.3%
Sidechain outliers		12.2%
RSRZ outliers		0.2%

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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	990	66%　　　26%　　•　•
1	B	990	62%　　　28%　　6%　•

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 15748 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 Insulin-degrading enzyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	953	Total 7790	C 5019	N 1308	O 1441	S 22	0	0	0
1	B	952	Total 7780	C 5013	N 1305	O 1440	S 22	0	0	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	30	MET	-	expression tag	UNP P14735
A	31	HIS	-	expression tag	UNP P14735
A	32	HIS	-	expression tag	UNP P14735
A	33	HIS	-	expression tag	UNP P14735
A	34	HIS	-	expression tag	UNP P14735
A	35	HIS	-	expression tag	UNP P14735
A	36	HIS	-	expression tag	UNP P14735
A	37	ALA	-	expression tag	UNP P14735
A	38	ALA	-	expression tag	UNP P14735
A	39	GLY	-	expression tag	UNP P14735
A	40	ILE	-	expression tag	UNP P14735
A	41	PRO	-	expression tag	UNP P14735
A	110	LEU	CYS	engineered mutation	UNP P14735
A	171	SER	CYS	engineered mutation	UNP P14735
A	178	ALA	CYS	engineered mutation	UNP P14735
A	257	VAL	CYS	engineered mutation	UNP P14735
A	414	LEU	CYS	engineered mutation	UNP P14735
A	573	ASN	CYS	engineered mutation	UNP P14735
A	590	SER	CYS	engineered mutation	UNP P14735
A	789	SER	CYS	engineered mutation	UNP P14735
A	812	ALA	CYS	engineered mutation	UNP P14735
A	819	ALA	CYS	engineered mutation	UNP P14735
A	904	SER	CYS	engineered mutation	UNP P14735
A	966	ASN	CYS	engineered mutation	UNP P14735
A	974	ALA	CYS	engineered mutation	UNP P14735

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	30	MET	-	expression tag	UNP P14735
B	31	HIS	-	expression tag	UNP P14735
B	32	HIS	-	expression tag	UNP P14735
B	33	HIS	-	expression tag	UNP P14735
B	34	HIS	-	expression tag	UNP P14735
B	35	HIS	-	expression tag	UNP P14735
B	36	HIS	-	expression tag	UNP P14735
B	37	ALA	-	expression tag	UNP P14735
B	38	ALA	-	expression tag	UNP P14735
B	39	GLY	-	expression tag	UNP P14735
B	40	ILE	-	expression tag	UNP P14735
B	41	PRO	-	expression tag	UNP P14735
B	110	LEU	CYS	engineered mutation	UNP P14735
B	171	SER	CYS	engineered mutation	UNP P14735
B	178	ALA	CYS	engineered mutation	UNP P14735
B	257	VAL	CYS	engineered mutation	UNP P14735
B	414	LEU	CYS	engineered mutation	UNP P14735
B	573	ASN	CYS	engineered mutation	UNP P14735
B	590	SER	CYS	engineered mutation	UNP P14735
B	789	SER	CYS	engineered mutation	UNP P14735
B	812	ALA	CYS	engineered mutation	UNP P14735
B	819	ALA	CYS	engineered mutation	UNP P14735
B	904	SER	CYS	engineered mutation	UNP P14735
B	966	ASN	CYS	engineered mutation	UNP P14735
B	974	ALA	CYS	engineered mutation	UNP P14735

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	Zn 1	0	0
2	B	1	Total 1	Zn 1	0	0

- Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	94	Total 94	O 94	0	0
3	B	82	Total 82	O 82	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: Insulin-degrading enzyme

Chain A:

- Molecule 1: Insulin-degrading enzyme

Chain B:
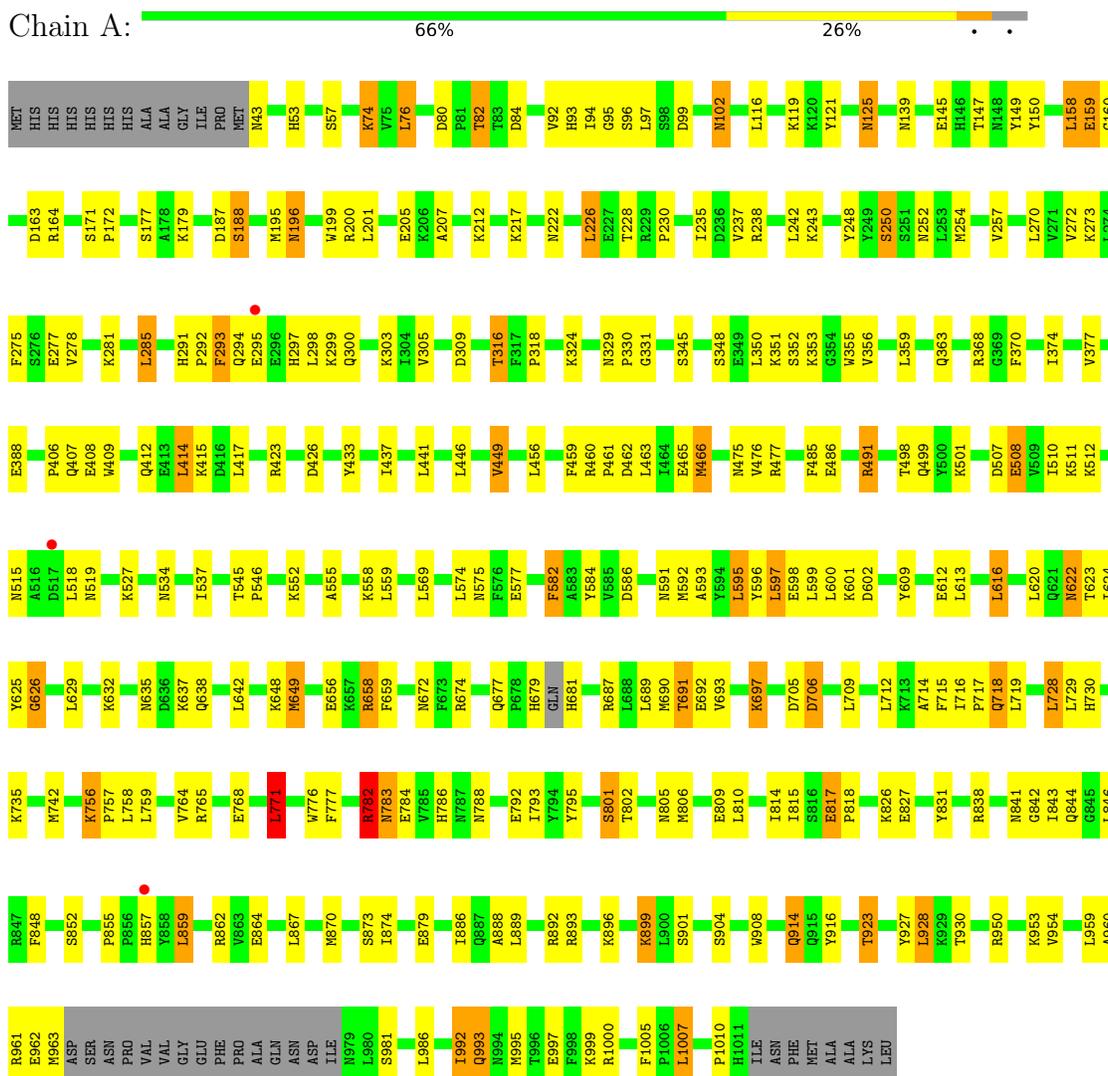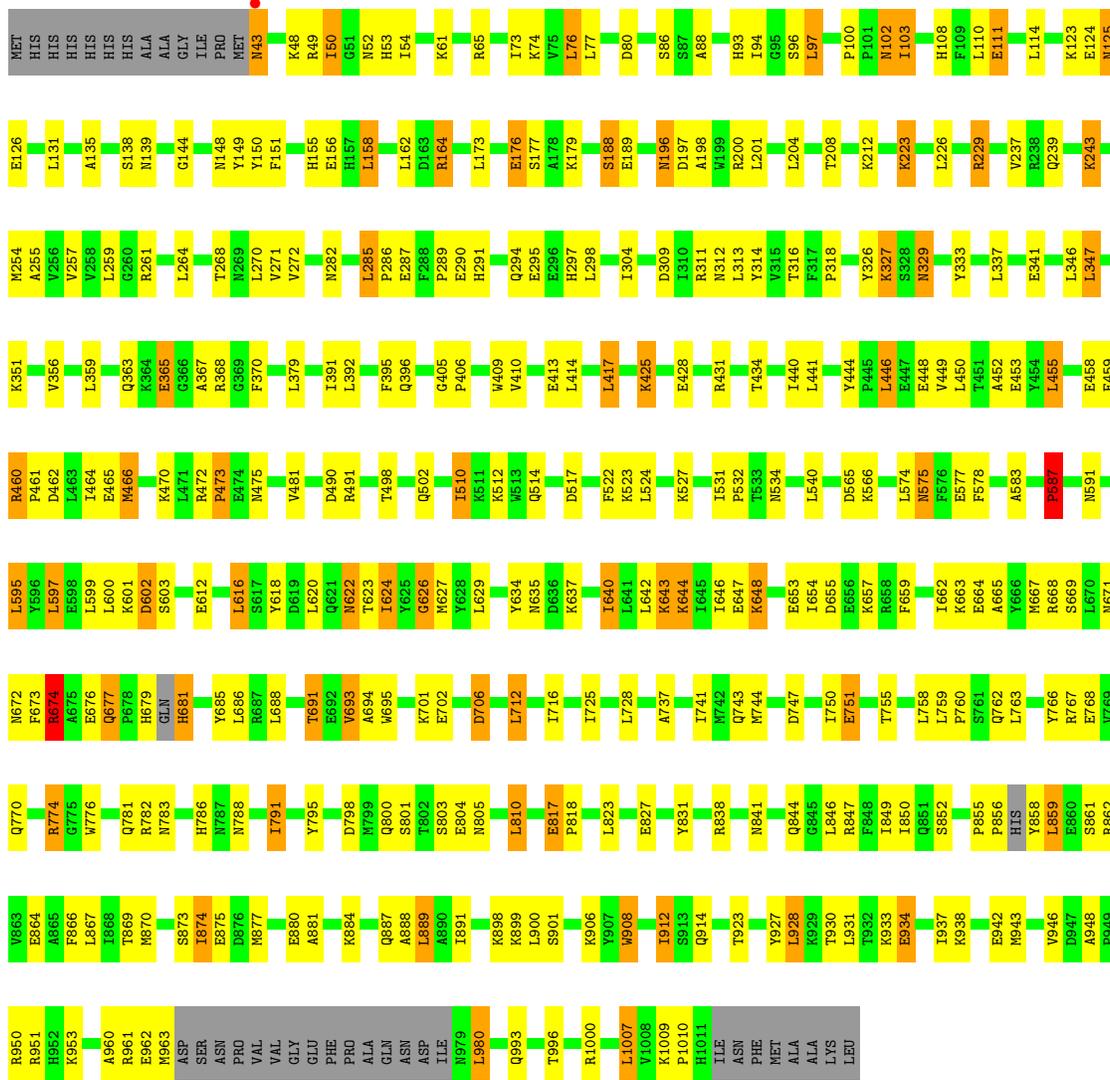
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	262.07Å 262.07Å 90.84Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.20 49.53 – 3.21	Depositor EDS
% Data completeness (in resolution range)	99.6 (50.00-3.20) 99.6 (49.53-3.21)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.18 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.172 , 0.241 0.173 , 0.237	Depositor DCC
R_{free} test set	2967 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	38.4	Xtriage
Anisotropy	0.094	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 36.2	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.018 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	15748	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.07% 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: ZN

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	0.84	4/7984 (0.1%)	0.90	13/10800 (0.1%)
1	B	0.81	5/7972 (0.1%)	0.88	8/10782 (0.1%)
All	All	0.82	9/15956 (0.1%)	0.89	21/21582 (0.1%)

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	A	0	2

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	577	GLU	CG-CD	8.82	1.65	1.51
1	B	453	GLU	CB-CG	6.97	1.65	1.52
1	B	577	GLU	CG-CD	6.67	1.61	1.51
1	B	50	ILE	CG1-CD1	6.24	1.93	1.50
1	A	577	GLU	CB-CG	5.81	1.63	1.52

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	365	GLU	CB-CA-C	-8.45	93.50	110.40
1	A	992	ILE	O-C-N	-8.15	109.66	122.70
1	A	285	LEU	CA-CB-CG	7.45	132.43	115.30
1	A	993	GLN	N-CA-CB	7.26	123.68	110.60
1	A	626	GLY	N-CA-C	-6.86	95.94	113.10

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	992	ILE	Peptide,Mainchain

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	7790	0	7724	208	0
1	B	7780	0	7717	231	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	A	94	0	0	38	0
3	B	82	0	0	20	0
All	All	15748	0	15441	436	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 14.

The worst 5 of 436 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:50:ILE:CD1	1:B:50:ILE:CG1	1.93	1.44
1:B:782:ARG:HD2	3:B:1059:HOH:O	1.46	1.12
1:B:679:HIS:O	1:B:681:HIS:N	1.86	1.09
1:A:491:ARG:HG3	1:A:491:ARG:HH11	1.12	1.08
1:B:856:PRO:O	1:B:858:TYR:N	1.88	1.07

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	947/990 (96%)	881 (93%)	65 (7%)	1 (0%)	51	83
1	B	944/990 (95%)	876 (93%)	63 (7%)	5 (0%)	29	67
All	All	1891/1980 (96%)	1757 (93%)	128 (7%)	6 (0%)	41	74

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	946	VAL
1	A	228	THR
1	B	406	PRO
1	B	405	GLY
1	B	587	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	847/879 (96%)	761 (90%)	86 (10%)	7	29
1	B	846/879 (96%)	725 (86%)	121 (14%)	3	15
All	All	1693/1758 (96%)	1486 (88%)	207 (12%)	5	22

5 of 207 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	329	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	587	PRO
1	B	928	LEU
1	B	414	LEU
1	B	466	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 63 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	828	GLN
1	B	783	ASN
1	B	102	ASN
1	B	743	GLN
1	B	828	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, 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	953/990 (96%)	-0.34	3 (0%)	94	92	11, 25, 40, 60	0
1	B	952/990 (96%)	-0.30	1 (0%)	95	95	14, 27, 41, 61	0
All	All	1905/1980 (96%)	-0.32	4 (0%)	95	94	11, 26, 41, 61	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	857	HIS	2.7
1	B	43	ASN	2.4
1	A	295	GLU	2.2
1	A	517	ASP	2.0

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	ZN	B	2	1/1	0.97	0.12	28,28,28,28	0

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	ZN	A	1	1/1	0.98	0.10	27,27,27,27	0

6.5   Other polymers ⓘ

There are no such residues in this entry.