wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2025 – 03:18 pm GMT

PDB ID	:	9GRH
Title	:	Crystal structure of the C-terminal phosphatase domain from Saccharomyces cerevisiae Vip1 (apo)
Authors	:	Raia, P.; Lee, K.; Hothorn, M.
Deposited on	:	2024-09-11
Resolution	:	3.20 Å(reported)

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

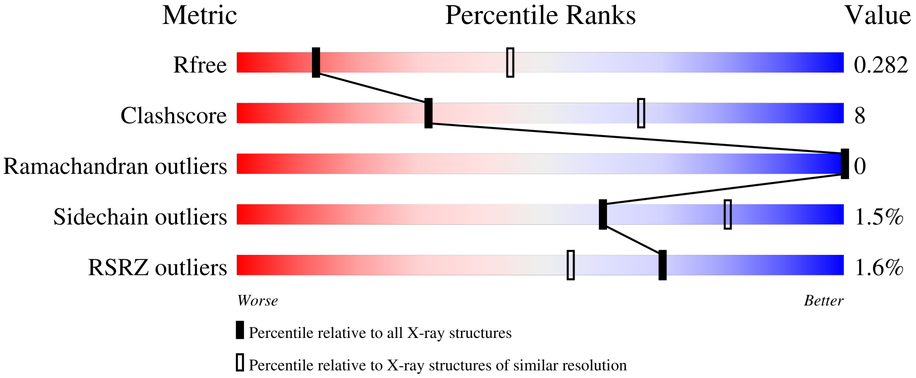
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

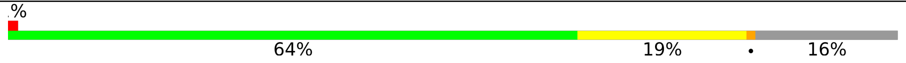
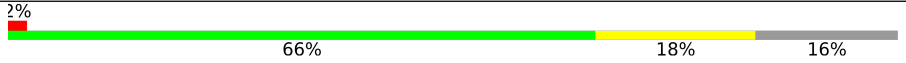
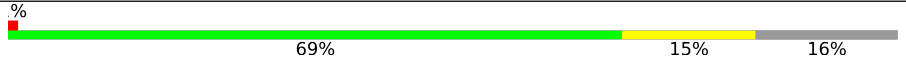
The reported resolution of this entry is 3.20 Å.

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1370 (3.20-3.20)
Clashscore	180529	1497 (3.20-3.20)
Ramachandran outliers	177936	1479 (3.20-3.20)
Sidechain outliers	177891	1478 (3.20-3.20)
RSRZ outliers	164620	1371 (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	574	
1	B	574	
1	C	574	
1	D	574	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 31731 atoms, of which 16021 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 Inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	481	Total	C	H	N	O	S	0	0	0
			7920	2518	3996	672	717	17			
1	B	481	Total	C	H	N	O	S	0	0	0
			7944	2523	4013	674	717	17			
1	C	480	Total	C	H	N	O	S	0	0	0
			7933	2519	4010	673	714	17			
1	D	481	Total	C	H	N	O	S	0	0	0
			7930	2520	4002	673	718	17			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	534	GLY	-	expression tag	UNP Q06685
A	535	ALA	-	expression tag	UNP Q06685
B	534	GLY	-	expression tag	UNP Q06685
B	535	ALA	-	expression tag	UNP Q06685
C	534	GLY	-	expression tag	UNP Q06685
C	535	ALA	-	expression tag	UNP Q06685
D	534	GLY	-	expression tag	UNP Q06685
D	535	ALA	-	expression tag	UNP Q06685

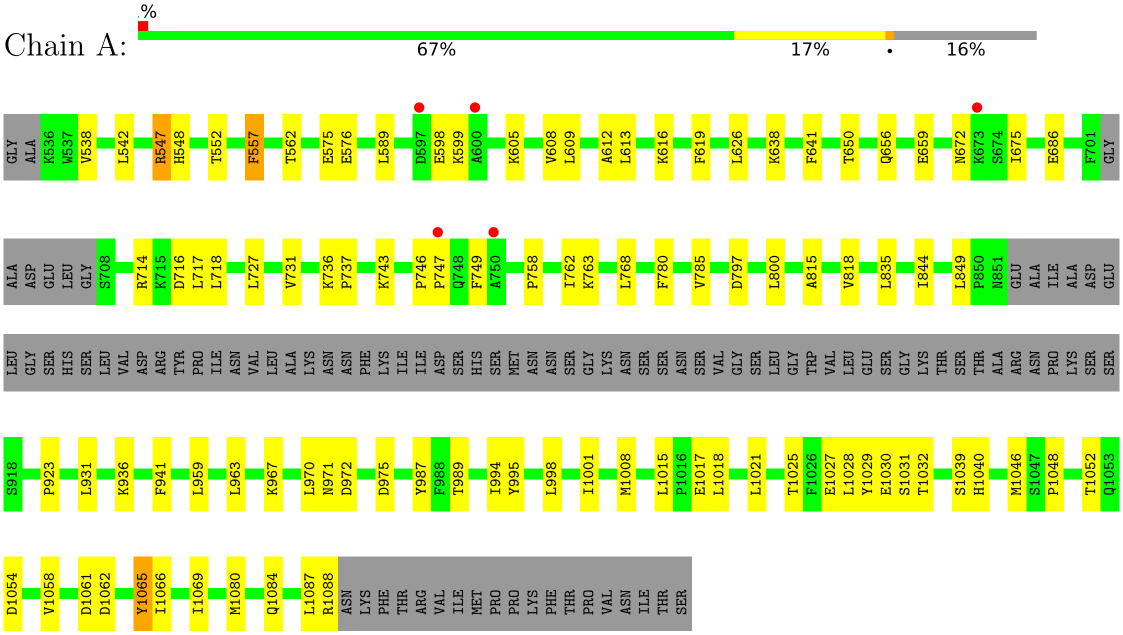
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

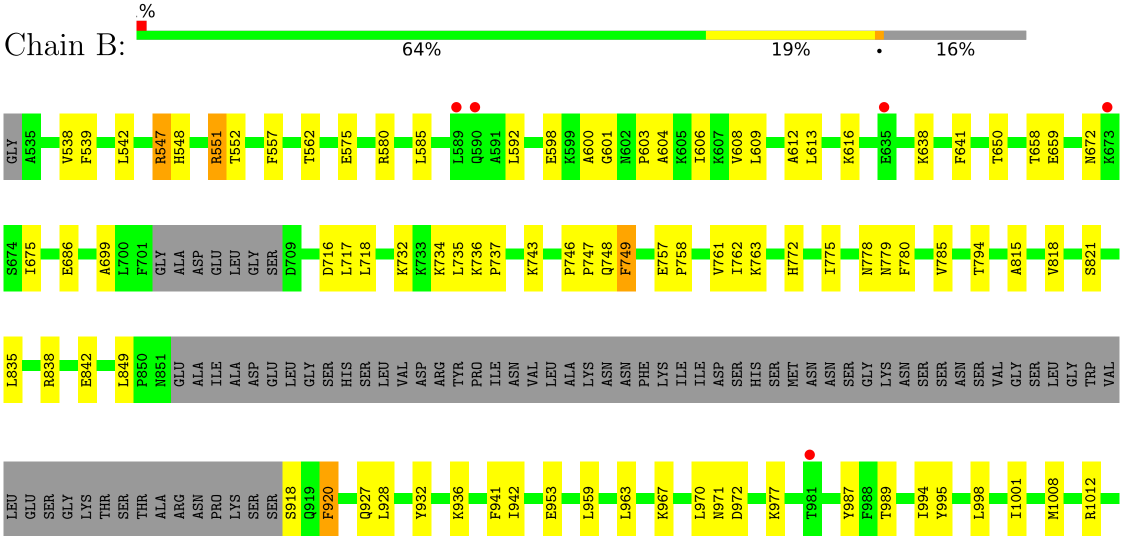
3 Residue-property plots ⓘ

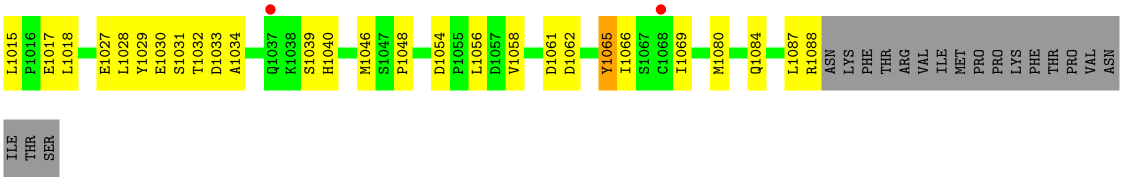
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

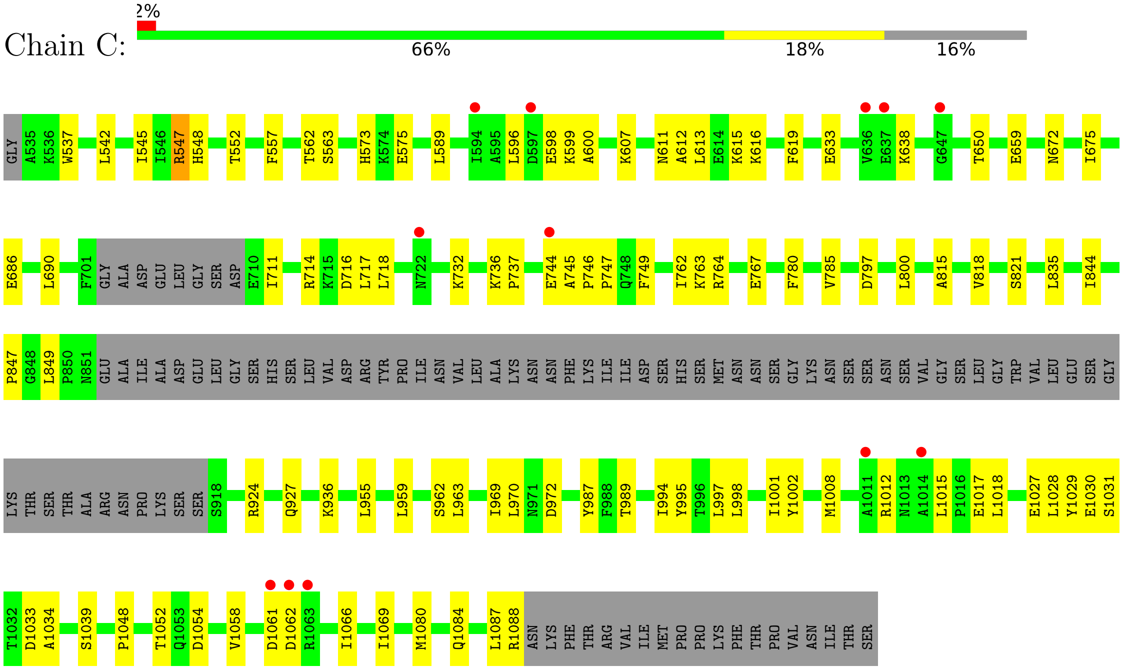


- Molecule 1: Inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase

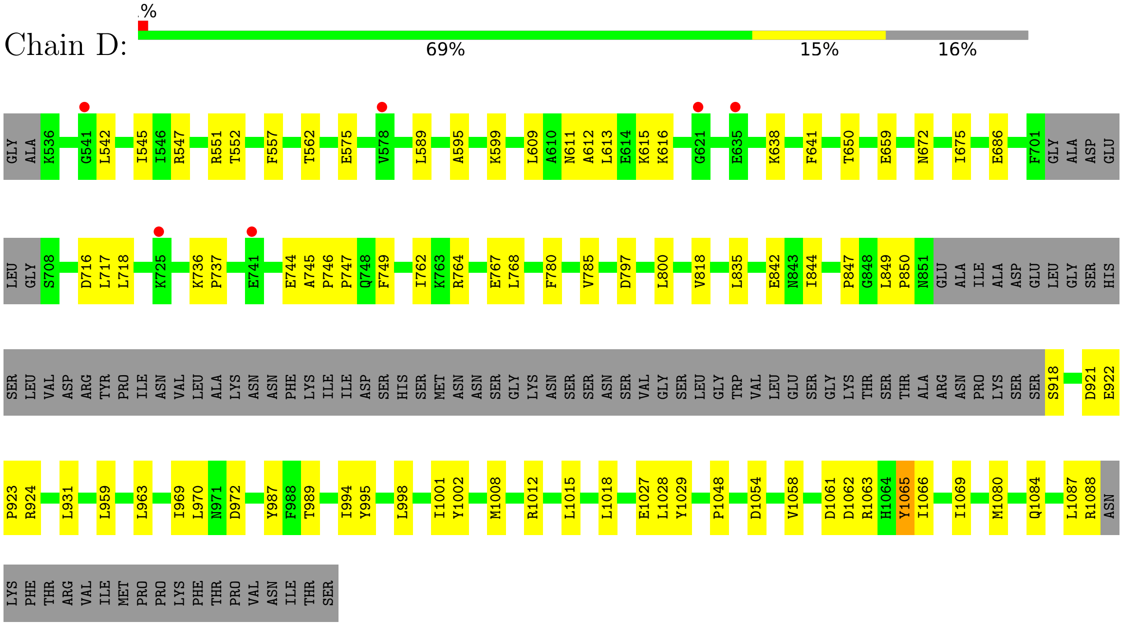




- Molecule 1: Inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase



- Molecule 1: Inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	84.47Å 194.86Å 84.44Å 90.00° 112.16° 90.00°	Depositor
Resolution (Å)	48.71 – 3.20 48.71 – 3.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.71-3.20) 99.9 (48.71-3.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.24 (at 3.07Å)	Xtriage
Refinement program	PHENIX (1.21.1_5286: ???)	Depositor
R, R_{free}	0.218 , 0.276 0.227 , 0.282	Depositor DCC
R_{free} test set	1981 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å^2)	93.0	Xtriage
Anisotropy	0.228	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 257.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.420 for l,-k,h	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	31731	wwPDB-VP
Average B, all atoms (Å^2)	136.0	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 26.10 % 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 2.8043e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[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	0.24	0/4007	0.45	0/5399
1	B	0.24	0/4014	0.45	0/5407
1	C	0.24	0/4006	0.45	0/5396
1	D	0.24	0/4011	0.45	0/5404
All	All	0.24	0/16038	0.45	0/21606

There are no bond length outliers.

There are no bond angle outliers.

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	3924	3996	3995	62	1
1	B	3931	4013	4012	72	2
1	C	3923	4010	4008	66	1
1	D	3928	4002	4001	55	2
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
All	All	15710	16021	16016	253	3

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 8.

The worst 5 of 253 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:659:GLU:OE1	1:D:1065:TYR:OH	1.96	0.83
1:A:659:GLU:OE1	1:A:1065:TYR:OH	1.95	0.82
1:B:1027:GLU:OE1	1:B:1029:TYR:OH	1.97	0.82
1:A:972:ASP:OD2	1:A:987:TYR:OH	1.99	0.81
1:B:972:ASP:OD2	1:B:987:TYR:OH	1.98	0.81

All (3) 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:927:GLN:HE21	1:D:923:PRO:O[2_646]	1.57	0.03
1:B:778:ASN:ND2	1:D:850:PRO:O[2_646]	2.18	0.02
1:A:923:PRO:O	1:C:927:GLN:HE21[2_545]	1.60	0.00

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	475/574 (83%)	463 (98%)	12 (2%)	0	100	100
1	B	475/574 (83%)	459 (97%)	16 (3%)	0	100	100
1	C	474/574 (83%)	459 (97%)	15 (3%)	0	100	100
1	D	475/574 (83%)	459 (97%)	16 (3%)	0	100	100
All	All	1899/2296 (83%)	1840 (97%)	59 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	437/518 (84%)	429 (98%)	8 (2%)	54	77
1	B	438/518 (85%)	430 (98%)	8 (2%)	54	77
1	C	437/518 (84%)	433 (99%)	4 (1%)	75	89
1	D	438/518 (85%)	432 (99%)	6 (1%)	62	82
All	All	1750/2072 (84%)	1724 (98%)	26 (2%)	60	81

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	920	PHE
1	C	557	PHE
1	D	749	PHE
1	C	547	ARG
1	C	686	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 (i)

Of 4 ligands modelled in this entry, 4 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	481/574 (83%)	-0.05	5 (1%)	79	66	67, 125, 256, 401	0
1	B	481/574 (83%)	-0.10	7 (1%)	71	56	67, 129, 216, 366	0
1	C	480/574 (83%)	-0.01	12 (2%)	58	42	72, 126, 229, 404	0
1	D	481/574 (83%)	-0.04	6 (1%)	76	61	72, 123, 200, 301	0
All	All	1923/2296 (83%)	-0.05	30 (1%)	70	55	67, 126, 226, 404	0

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	637	GLU	4.4
1	A	673	LYS	3.9
1	C	1063	ARG	3.9
1	B	590	GLN	3.7
1	B	1037	GLN	3.5

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	A	2000	1/1	0.96	0.10	343,343,343,343	0
2	ZN	B	2000	1/1	0.99	0.04	112,112,112,112	0
2	ZN	C	2000	1/1	0.99	0.04	129,129,129,129	0
2	ZN	D	2000	1/1	0.99	0.06	155,155,155,155	0

6.5 Other polymers

There are no such residues in this entry.