



Full wwPDB X-ray Structure Validation Report

Feb 28, 2014 – 01:52 PM GMT

PDB ID : 2GCJ
Title : Crystal Structure of the Pob3 middle domain
Authors : VanDemark, A.P.
Deposited on : 2006-03-14
Resolution : 2.55 Å(reported)

This is a full wwPDB validation report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

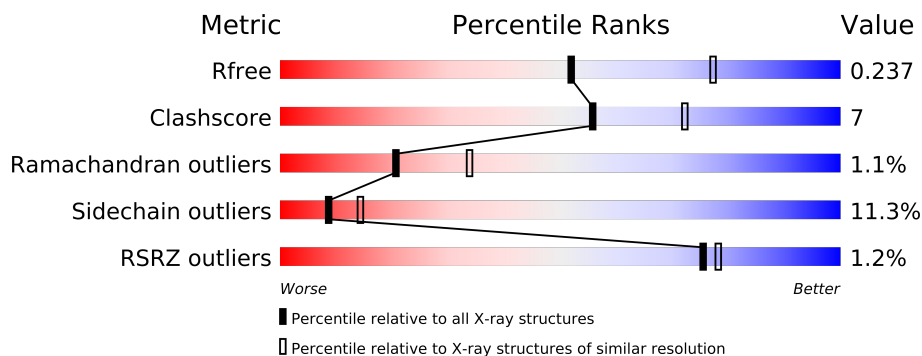
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.55 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}	66092	3413 (2.58-2.50)
Clashscore	79885	4284 (2.58-2.50)
Ramachandran outliers	78287	4193 (2.58-2.50)
Sidechain outliers	78261	4195 (2.58-2.50)
RSRZ outliers	66119	3414 (2.58-2.50)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	261	
1	B	261	
1	C	261	
1	D	261	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7649 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Hypothetical 63.0 kDa protein in DAK1-ORC1 intergenic region.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	229	Total	C	N	O	S	0	0	0
			1888	1210	321	354	3			
1	B	230	Total	C	N	O	S	3	0	0
			1895	1215	322	355	3			
1	C	231	Total	C	N	O	S	5	0	0
			1904	1220	323	358	3			
1	D	229	Total	C	N	O	S	0	0	0
			1882	1207	318	354	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	218	GLY	-	CLONING ARTIFACT	UNP Q04636
A	219	HIS	-	CLONING ARTIFACT	UNP Q04636
A	308	LYS	GLN	ENGINEERED	UNP Q04636
B	218	GLY	-	CLONING ARTIFACT	UNP Q04636
B	219	HIS	-	CLONING ARTIFACT	UNP Q04636
B	308	LYS	GLN	ENGINEERED	UNP Q04636
C	218	GLY	-	CLONING ARTIFACT	UNP Q04636
C	219	HIS	-	CLONING ARTIFACT	UNP Q04636
C	308	LYS	GLN	ENGINEERED	UNP Q04636
D	218	GLY	-	CLONING ARTIFACT	UNP Q04636
D	219	HIS	-	CLONING ARTIFACT	UNP Q04636
D	308	LYS	GLN	ENGINEERED	UNP Q04636

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	21	Total	O	0	0
			21	21		
2	B	25	Total	O	0	0
			25	25		

Continued on next page...

Continued from previous page...

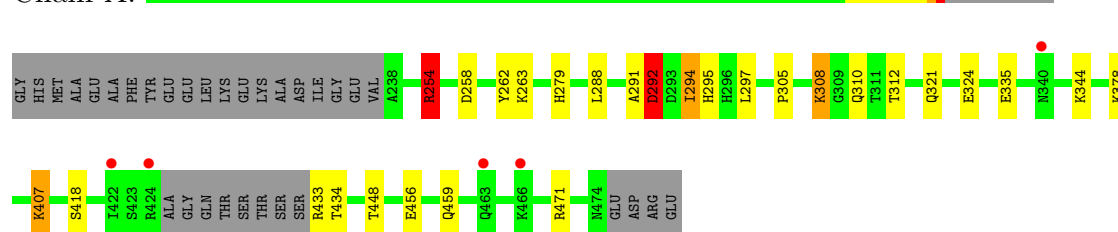
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	18	Total	O	0	0
			18	18		
2	D	16	Total	O	0	0
			16	16		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

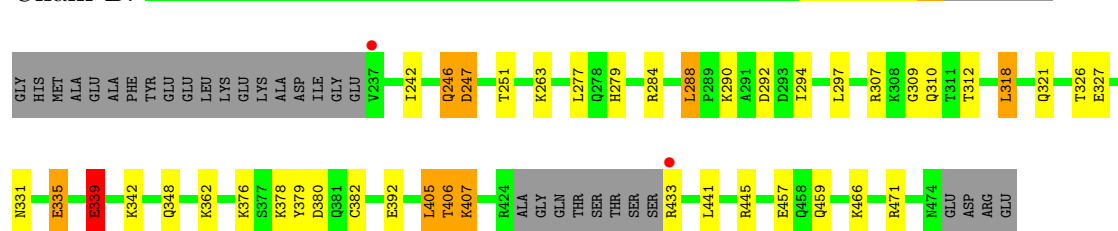
- Molecule 1: Hypothetical 63.0 kDa protein in DAK1-ORC1 intergenic region

Chain A:



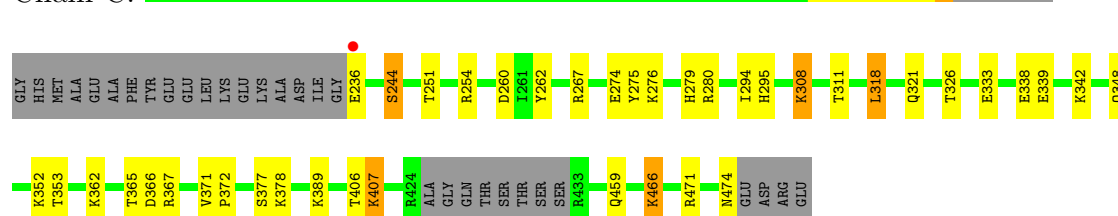
- Molecule 1: Hypothetical 63.0 kDa protein in DAK1-ORC1 intergenic region

Chain B:



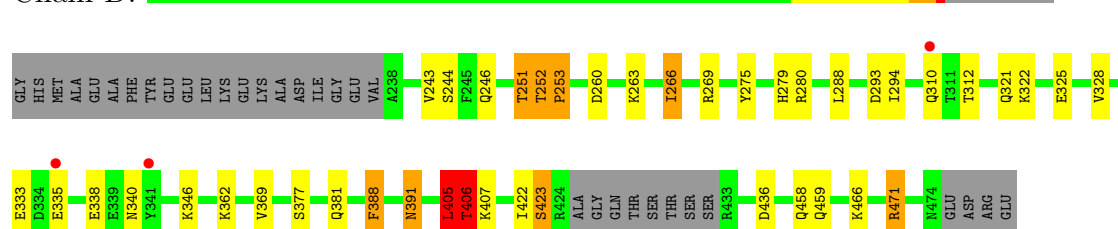
- Molecule 1: Hypothetical 63.0 kDa protein in DAK1-ORC1 intergenic region

Chain C:



- Molecule 1: Hypothetical 63.0 kDa protein in DAK1-ORC1 intergenic region

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	57.87Å 157.53Å 57.79Å 90.00° 89.77° 90.00°	Depositor
Resolution (Å)	78.81 – 2.55 32.30 – 2.49	Depositor EDS
% Data completeness (in resolution range)	91.7 (78.81-2.55) 91.1 (32.30-2.49)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.08 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.215 , 0.303 0.206 , 0.237	Depositor DCC
R_{free} test set	1638 reflections (5.25%)	DCC
Wilson B-factor (Å ²)	45.9	Xtriage
Anisotropy	0.360	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 7.9	EDS
Estimated twinning fraction	0.000 for l,k,-h 0.460 for -h,-k,l 0.000 for -l,-k,-h	Xtriage
L-test for twinning	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	2 of 32801 reflections (0.006%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7649	wwPDB-VP
Average B, all atoms (Å ²)	38.0	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 38.28 % 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 3.7782e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

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	0.57	0/1926	0.70	1/2599 (0.0%)
1	B	0.63	0/1933	0.74	0/2609
1	C	0.62	0/1942	0.75	0/2621
1	D	0.57	0/1920	0.70	0/2592
All	All	0.60	0/7721	0.72	1/10421 (0.0%)

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	1
1	B	0	3
1	C	0	1
1	D	0	4
All	All	0	9

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	254	ARG	NE-CZ-NH1	5.29	122.94	120.30

There are no chirality outliers.

All (9) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	407	LYS	Peptide
1	B	339	GLU	Peptide
1	B	382	CYS	Peptide
1	B	405	LEU	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	C	407	LYS	Peptide
1	D	251	THR	Peptide
1	D	252	THR	Peptide
1	D	388	PHE	Peptide
1	D	405	LEU	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1888	0	0	8	0
1	B	1895	0	0	12	0
1	C	1904	0	0	17	0
1	D	1882	0	0	17	0
2	A	21	0	0	0	0
2	B	25	0	0	1	0
2	C	18	0	0	0	0
2	D	16	0	0	0	0
All	All	7649	0	0	50	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 7.

All (50) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:406:THR:CG2	1:D:407:LYS:N	2.30	0.94
1:C:294:ILE:CD1	1:C:295:HIS:CE1	2.51	0.93
1:B:339:GLU:OE2	1:B:339:GLU:CA	2.30	0.78
1:D:406:THR:O	1:D:407:LYS:CG	2.39	0.70
1:A:292:ASP:OD1	1:A:294:ILE:CG2	2.46	0.64
1:C:294:ILE:CG1	1:C:295:HIS:ND1	2.63	0.62
1:C:262:TYR:O	1:C:279:HIS:CE1	2.54	0.61
1:D:423:SER:O	1:D:436:ASP:N	2.35	0.60
1:B:445:ARG:CD	2:B:43:HOH:O	2.55	0.53
1:A:378:LYS:CD	1:A:456:GLU:OE1	2.57	0.53

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:388:PHE:O	1:D:391:ASN:ND2	2.42	0.53
1:D:252:THR:CG2	1:D:253:PRO:N	2.72	0.52
1:A:294:ILE:CG1	1:A:295:HIS:N	2.74	0.51
1:A:308:LYS:CG	1:A:308:LYS:O	2.59	0.50
1:D:362:LYS:CE	1:D:369:VAL:CG2	2.90	0.50
1:C:275:TYR:OH	1:C:308:LYS:CB	2.60	0.50
1:C:276:LYS:NZ	1:D:325:GLU:OE1	2.46	0.49
1:D:422:ILE:CG2	1:D:458:GLN:OE1	2.60	0.49
1:D:252:THR:CG2	1:D:253:PRO:CD	2.90	0.49
1:B:335:GLU:O	1:B:339:GLU:N	2.47	0.48
1:B:242:ILE:O	1:B:331:ASN:N	2.47	0.48
1:B:290:LYS:O	1:B:292:ASP:N	2.46	0.48
1:D:471:ARG:NH1	1:D:471:ARG:CG	2.77	0.47
1:D:405:LEU:O	1:D:406:THR:CB	2.62	0.46
1:B:318:LEU:CD1	1:B:318:LEU:N	2.78	0.46
1:D:391:ASN:ND2	1:D:391:ASN:N	2.63	0.46
1:D:260:ASP:OD2	1:D:269:ARG:CD	2.63	0.46
1:B:247:ASP:OD1	1:C:367:ARG:NH2	2.49	0.46
1:A:263:LYS:O	1:A:279:HIS:CE1	2.68	0.46
1:B:246:GLN:O	1:C:280:ARG:NH2	2.49	0.46
1:D:266:ILE:CD1	1:D:279:HIS:CD2	3.00	0.45
1:B:263:LYS:O	1:B:279:HIS:CD2	2.69	0.45
1:D:377:SER:N	1:D:381:GLN:O	2.50	0.45
1:C:244:SER:OG	1:C:260:ASP:OD1	2.35	0.44
1:B:406:THR:CG2	1:B:407:LYS:N	2.81	0.44
1:C:294:ILE:CG1	1:C:295:HIS:CE1	2.99	0.43
1:C:274:GLU:OE2	1:D:346:LYS:NZ	2.52	0.43
1:A:254:ARG:NH1	1:A:254:ARG:CG	2.81	0.43
1:A:288:LEU:O	1:A:297:LEU:N	2.52	0.43
1:C:466:LYS:C	1:C:466:LYS:CD	2.87	0.43
1:B:378:LYS:NZ	1:B:379:TYR:OH	2.51	0.43
1:C:371:VAL:CB	1:C:372:PRO:CD	2.96	0.43
1:C:318:LEU:N	1:C:318:LEU:CD1	2.82	0.42
1:D:263:LYS:O	1:D:279:HIS:CE1	2.73	0.42
1:C:406:THR:CG2	1:C:406:THR:O	2.67	0.42
1:C:294:ILE:CD1	1:C:295:HIS:ND1	2.82	0.42
1:C:262:TYR:OH	1:C:267:ARG:NH1	2.53	0.42
1:C:352:LYS:O	1:C:353:THR:C	2.59	0.41
1:B:288:LEU:O	1:B:297:LEU:N	2.54	0.40
1:A:262:TYR:O	1:A:279:HIS:CE1	2.74	0.40

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	225/261 (86%)	197 (88%)	26 (12%)	2 (1%)	25	41
1	B	226/261 (87%)	203 (90%)	21 (9%)	2 (1%)	25	41
1	C	227/261 (87%)	212 (93%)	13 (6%)	2 (1%)	25	41
1	D	225/261 (86%)	210 (93%)	11 (5%)	4 (2%)	13	20
All	All	903/1044 (86%)	822 (91%)	71 (8%)	10 (1%)	21	34

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	406	THR
1	D	322	LYS
1	D	406	THR
1	A	292	ASP
1	C	366	ASP
1	C	377	SER
1	D	243	VAL
1	D	253	PRO
1	A	291	ALA
1	B	309	GLY

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	211/237 (89%)	192 (91%)	19 (9%)	14	24
1	B	212/237 (90%)	182 (86%)	30 (14%)	5	8
1	C	213/237 (90%)	190 (89%)	23 (11%)	9	16

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	210/237 (89%)	186 (89%)	24 (11%)	8	14
All	All	846/948 (89%)	750 (89%)	96 (11%)	9	14

All (96) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	254	ARG
1	A	258	ASP
1	A	292	ASP
1	A	294	ILE
1	A	305	PRO
1	A	308	LYS
1	A	310	GLN
1	A	312	THR
1	A	321	GLN
1	A	324	GLU
1	A	335	GLU
1	A	344	LYS
1	A	407	LYS
1	A	418	SER
1	A	433	ARG
1	A	434	THR
1	A	448	THR
1	A	459	GLN
1	A	471	ARG
1	B	246	GLN
1	B	247	ASP
1	B	251	THR
1	B	277	LEU
1	B	284	ARG
1	B	288	LEU
1	B	294	ILE
1	B	307	ARG
1	B	310	GLN
1	B	312	THR
1	B	318	LEU
1	B	321	GLN
1	B	326	THR
1	B	327	GLU
1	B	335	GLU
1	B	339	GLU
1	B	342	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	348	GLN
1	B	362	LYS
1	B	376	LYS
1	B	380	ASP
1	B	392	GLU
1	B	405	LEU
1	B	407	LYS
1	B	433	ARG
1	B	441	LEU
1	B	457	GLU
1	B	459	GLN
1	B	466	LYS
1	B	471	ARG
1	C	236	GLU
1	C	244	SER
1	C	251	THR
1	C	254	ARG
1	C	308	LYS
1	C	311	THR
1	C	318	LEU
1	C	321	GLN
1	C	326	THR
1	C	333	GLU
1	C	338	GLU
1	C	339	GLU
1	C	342	LYS
1	C	348	GLN
1	C	362	LYS
1	C	365	THR
1	C	378	LYS
1	C	389	LYS
1	C	407	LYS
1	C	459	GLN
1	C	466	LYS
1	C	471	ARG
1	C	474	ASN
1	D	244	SER
1	D	246	GLN
1	D	251	THR
1	D	266	ILE
1	D	275	TYR
1	D	280	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	288	LEU
1	D	293	ASP
1	D	294	ILE
1	D	310	GLN
1	D	312	THR
1	D	321	GLN
1	D	328	VAL
1	D	333	GLU
1	D	335	GLU
1	D	338	GLU
1	D	340	ASN
1	D	391	ASN
1	D	405	LEU
1	D	406	THR
1	D	423	SER
1	D	459	GLN
1	D	466	LYS
1	D	471	ARG

Some 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 chains 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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	229/261 (87%)	0.22	5 (2%) 59 62	15, 41, 75, 105	2 (0%)
1	B	230/261 (88%)	0.03	2 (0%) 81 83	9, 31, 65, 94	1 (0%)
1	C	231/261 (88%)	-0.02	1 (0%) 90 92	7, 30, 65, 100	2 (0%)
1	D	229/261 (87%)	0.13	3 (1%) 74 77	15, 41, 77, 113	0
All	All	919/1044 (88%)	0.09	11 (1%) 75 78	7, 36, 73, 113	5 (0%)

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	237	VAL	3.7
1	A	422	ILE	3.0
1	A	466	LYS	2.9
1	A	340	ASN	2.8
1	D	310	GLN	2.4
1	C	236	GLU	2.3
1	D	341	TYR	2.2
1	A	463	GLN	2.2
1	B	433	ARG	2.1
1	D	335	GLU	2.1
1	A	424	ARG	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 carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.