Full wwPDB X-ray Structure Validation Report (i)

Feb 19, 2025 – 10:02 PM EST

PDB ID	:	9E4S
Title	:	TAD from Carmabin Biosynthetic Pathway - Crystal Form 1
Authors	:	Rankin, M.R.; Smith, J.L.
Deposited on	:	2024-10-25
Resolution	:	2.28 Å(reported)

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

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.3

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 2.28 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.
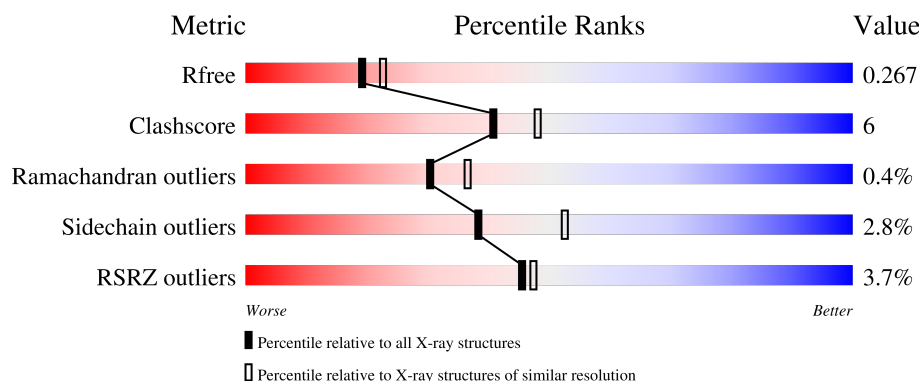
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	8487 (2.30-2.26)
Clashscore	180529	9437 (2.30-2.26)
Ramachandran outliers	177936	9341 (2.30-2.26)
Sidechain outliers	177891	9342 (2.30-2.26)
RSRZ outliers	164620	8487 (2.30-2.26)

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	425	4% / 80% / 17% / ..
1	B	425	3% / 80% / 14% / 6%

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 6550 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 Amino acid adenylation domain protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	420	Total 3306	C 2116	N 549	O 623	S 18	0	0	0
1	B	401	Total 3156	C 2026	N 522	O 590	S 18	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1887	MET	-	initiating methionine	UNP F4Y2B0
A	2304	LEU	-	expression tag	UNP F4Y2B0
A	2305	GLU	-	expression tag	UNP F4Y2B0
A	2306	HIS	-	expression tag	UNP F4Y2B0
A	2307	HIS	-	expression tag	UNP F4Y2B0
A	2308	HIS	-	expression tag	UNP F4Y2B0
A	2309	HIS	-	expression tag	UNP F4Y2B0
A	2310	HIS	-	expression tag	UNP F4Y2B0
A	2311	HIS	-	expression tag	UNP F4Y2B0
B	1887	MET	-	initiating methionine	UNP F4Y2B0
B	2304	LEU	-	expression tag	UNP F4Y2B0
B	2305	GLU	-	expression tag	UNP F4Y2B0
B	2306	HIS	-	expression tag	UNP F4Y2B0
B	2307	HIS	-	expression tag	UNP F4Y2B0
B	2308	HIS	-	expression tag	UNP F4Y2B0
B	2309	HIS	-	expression tag	UNP F4Y2B0
B	2310	HIS	-	expression tag	UNP F4Y2B0
B	2311	HIS	-	expression tag	UNP F4Y2B0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	38	Total 38	O 38	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	50	Total O 50 50		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: Amino acid adenylation domain protein
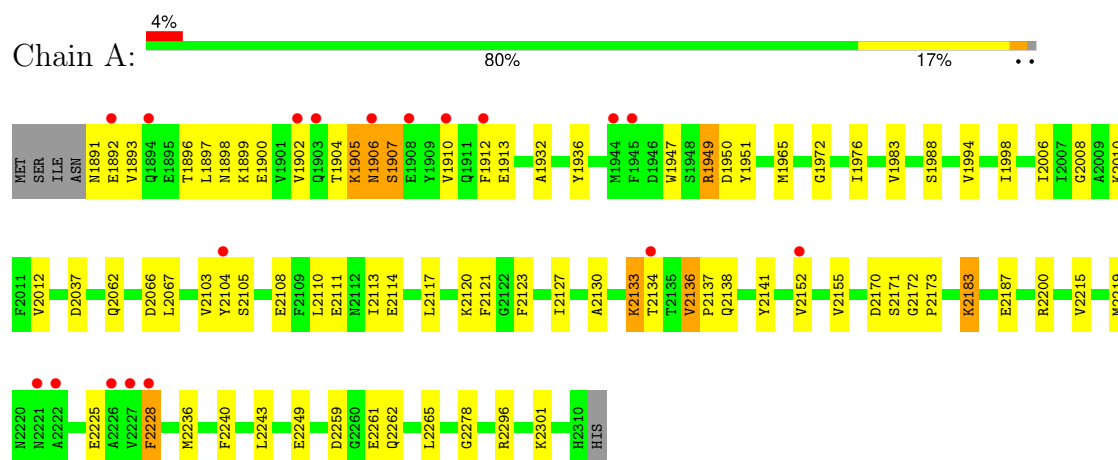
• Molecule 1: Amino acid adenylation domain protein
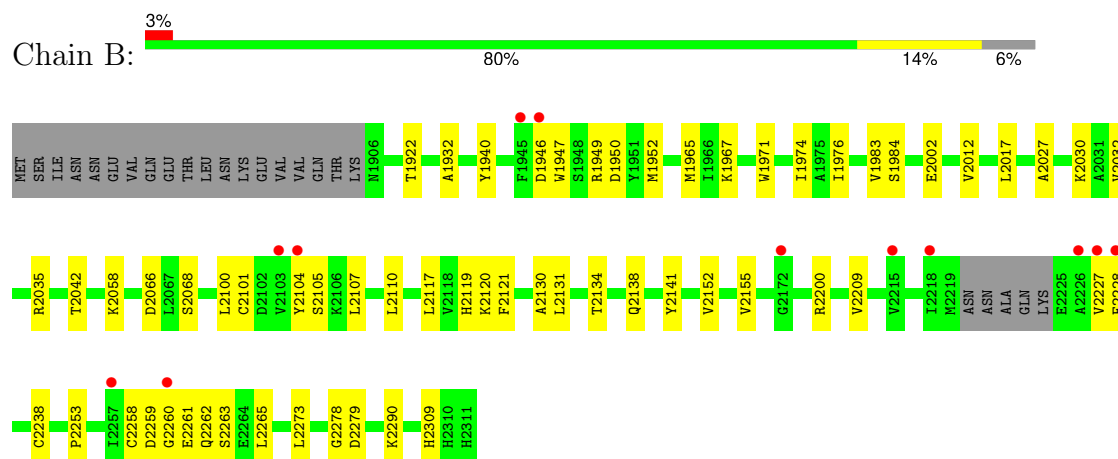
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	58.90Å 58.90Å 479.04Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.03 – 2.28 45.03 – 2.28	Depositor EDS
% Data completeness (in resolution range)	99.8 (45.03-2.28) 99.8 (45.03-2.28)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.58 (at 2.29Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.235 , 0.267 0.235 , 0.267	Depositor DCC
R_{free} test set	2138 reflections (4.65%)	wwPDB-VP
Wilson B-factor (Å^2)	66.2	Xtriage
Anisotropy	0.308	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 40.1	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.039 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6550	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% 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 ⓘ

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	0.30	0/3377	0.49	0/4578
1	B	0.26	0/3227	0.47	0/4373
All	All	0.28	0/6604	0.48	0/8951

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	2172	GLY	Peptide

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	3306	0	3295	49	0
1	B	3156	0	3139	34	0
2	A	38	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	50	0	0	1	0
All	All	6550	0	6434	82	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 6.

All (82) 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:2120:LYS:HD3	1:A:2121:PHE:CZ	1.97	0.99
1:A:2249:GLU:OE1	1:A:2249:GLU:N	2.12	0.82
1:B:2104:TYR:HA	1:B:2131:LEU:HD13	1.71	0.71
1:A:2136:VAL:HG22	1:A:2137:PRO:HD2	1.73	0.70
1:A:2120:LYS:HE2	1:A:2261:GLU:OE2	1.93	0.69
1:A:2120:LYS:HD3	1:A:2121:PHE:CE2	2.28	0.67
1:A:1965:MET:HB2	1:A:1976:ILE:HB	1.79	0.64
1:A:2225:GLU:N	1:A:2225:GLU:OE1	2.31	0.63
1:A:1907:SER:HB3	1:A:1910:VAL:H	1.65	0.61
1:A:1893:VAL:O	1:A:1896:THR:HG22	2.01	0.61
1:B:2027:ALA:HA	1:B:2030:LYS:HE2	1.82	0.60
1:A:2138:GLN:HA	1:A:2141:TYR:HD2	1.67	0.59
1:A:1932:ALA:HB3	1:A:2012:VAL:HG22	1.85	0.58
1:B:2120:LYS:HD2	1:B:2120:LYS:N	2.17	0.58
1:B:2209:VAL:HG21	1:B:2227:VAL:HG11	1.84	0.57
1:B:2066:ASP:OD1	1:B:2068:SER:N	2.34	0.57
1:A:1947:TRP:HB3	1:A:1951:TYR:HD2	1.69	0.57
1:A:1972:GLY:HA3	1:A:2243:LEU:HD11	1.87	0.57
1:B:1965:MET:HB2	1:B:1976:ILE:HB	1.88	0.55
1:B:2152:VAL:HG13	1:B:2155:VAL:HG21	1.87	0.55
1:A:2114:GLU:HA	1:A:2117:LEU:HD12	1.87	0.55
1:A:2120:LYS:CD	1:A:2121:PHE:CZ	2.82	0.55
1:A:1994:VAL:O	1:A:1998:ILE:HG12	2.06	0.54
1:A:2110:LEU:HD23	1:A:2113:ILE:HD11	1.89	0.54
1:A:2152:VAL:HG13	1:A:2155:VAL:HG11	1.90	0.53
1:A:2240:PHE:CE1	1:A:2243:LEU:HD23	2.44	0.53
1:A:1891:ASN:O	1:A:1892:GLU:C	2.48	0.52
1:A:2200:ARG:HB2	1:A:2278:GLY:HA2	1.92	0.52
1:B:2134:THR:O	1:B:2155:VAL:HG13	2.09	0.52
1:B:2262:GLN:OE1	1:B:2262:GLN:N	2.41	0.51
1:A:2183:LYS:HE3	1:A:2187:GLU:OE2	2.09	0.51
1:B:2117:LEU:HD23	1:B:2121:PHE:HD2	1.75	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1967:LYS:HB2	1:B:1974:ILE:HG23	1.93	0.51
1:A:1949:ARG:NH2	2:A:2403:HOH:O	2.43	0.51
1:A:2103:VAL:HG12	1:A:2104:TYR:H	1.77	0.50
1:B:2200:ARG:HB2	1:B:2278:GLY:HA2	1.94	0.50
1:B:2260:GLY:O	1:B:2263:SER:OG	2.26	0.50
1:B:2032:VAL:O	1:B:2035:ARG:HB2	2.12	0.50
1:A:2062:GLN:O	1:A:2296:ARG:NH1	2.37	0.49
1:A:2108:GLU:HA	1:A:2111:GLU:HB2	1.95	0.49
1:A:2008:GLY:O	1:A:2010:LYS:NZ	2.34	0.49
1:B:2119:HIS:HB2	1:B:2120:LYS:HD2	1.95	0.49
1:A:2104:TYR:CD1	1:A:2133:LYS:HG3	2.48	0.48
1:B:2107:LEU:HD23	1:B:2131:LEU:HD21	1.95	0.48
1:B:2101:CYS:HA	1:B:2130:ALA:O	2.14	0.48
1:A:2215:VAL:O	1:A:2219:MET:HG3	2.15	0.47
1:B:2100:LEU:HB3	1:B:2110:LEU:HD22	1.96	0.47
1:A:1904:THR:O	1:A:1905:LYS:C	2.52	0.46
1:B:2058:LYS:NZ	1:B:2279:ASP:OD1	2.45	0.46
1:A:1905:LYS:HE3	1:A:1905:LYS:HB2	1.47	0.46
1:A:1936:TYR:CE2	1:A:1983:VAL:HG22	2.51	0.45
1:B:1940:TYR:OH	1:B:1947:TRP:O	2.20	0.45
1:A:1902:VAL:C	1:A:1904:THR:N	2.69	0.45
1:B:1947:TRP:HB2	1:B:1952:MET:HG3	1.99	0.45
1:A:2006:ILE:HD12	1:B:1922:THR:HA	1.98	0.45
1:A:2066:ASP:OD1	1:A:2067:LEU:N	2.50	0.45
1:B:2138:GLN:HA	1:B:2141:TYR:HD2	1.83	0.44
1:A:1897:LEU:O	1:A:1898:ASN:C	2.55	0.44
1:A:1912:PHE:HE1	1:A:2228:PHE:HB2	1.83	0.44
1:A:2170:ASP:O	1:A:2171:SER:C	2.55	0.44
1:B:1971:TRP:HE1	1:B:2273:LEU:HD12	1.83	0.43
1:A:1912:PHE:CE1	1:A:2228:PHE:HB2	2.54	0.43
1:A:2130:ALA:HB1	1:A:2137:PRO:HG3	2.01	0.43
1:B:2002:GLU:OE1	1:B:2035:ARG:NH2	2.50	0.43
1:B:2258:CYS:SG	1:B:2259:ASP:N	2.90	0.43
1:A:2262:GLN:HG3	1:A:2265:LEU:HD12	2.01	0.43
1:B:2265:LEU:HD23	1:B:2265:LEU:HA	1.84	0.43
1:A:2123:PHE:HD2	1:A:2127:ILE:HD11	1.83	0.43
1:A:1902:VAL:C	1:A:1904:THR:H	2.21	0.43
1:A:2136:VAL:HG13	1:A:2137:PRO:O	2.18	0.43
1:B:2290:LYS:HA	1:B:2290:LYS:HD3	1.68	0.43
1:A:2103:VAL:HG12	1:A:2104:TYR:N	2.33	0.42
1:B:2042:THR:HB	1:B:2253:PRO:HB3	2.02	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1932:ALA:HB3	1:B:2012:VAL:HG22	2.02	0.42
1:B:2066:ASP:OD1	1:B:2066:ASP:C	2.58	0.42
1:A:1899:LYS:O	1:A:1900:GLU:C	2.58	0.42
1:B:1983:VAL:HG21	1:B:2017:LEU:HD12	2.02	0.41
1:A:1902:VAL:O	1:A:1904:THR:N	2.49	0.41
1:A:1906:ASN:ND2	1:A:1906:ASN:H	2.17	0.41
1:A:2037:ASP:OD1	1:A:2037:ASP:N	2.45	0.41
1:B:2238:CYS:HB3	2:B:2444:HOH:O	2.21	0.41
1:B:1949:ARG:NH1	1:B:1984:SER:OG	2.55	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	418/425 (98%)	388 (93%)	27 (6%)	3 (1%)	19	22
1	B	397/425 (93%)	385 (97%)	12 (3%)	0	100	100
All	All	815/850 (96%)	773 (95%)	39 (5%)	3 (0%)	30	36

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1905	LYS
1	A	2259	ASP
1	A	2173	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	367/372 (99%)	353 (96%)	14 (4%)	28	40
1	B	349/372 (94%)	343 (98%)	6 (2%)	56	70
All	All	716/744 (96%)	696 (97%)	20 (3%)	38	52

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1906	ASN
1	A	1907	SER
1	A	1913	GLU
1	A	1949	ARG
1	A	1950	ASP
1	A	1988	SER
1	A	2105	SER
1	A	2133	LYS
1	A	2134	THR
1	A	2136	VAL
1	A	2183	LYS
1	A	2228	PHE
1	A	2236	MET
1	A	2301	LYS
1	B	1946	ASP
1	B	1950	ASP
1	B	2105	SER
1	B	2228	PHE
1	B	2261	GLU
1	B	2309	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1891	ASN
1	A	1906	ASN
1	A	2268	GLN
1	A	2310	HIS
1	B	1995	GLN
1	B	2151	ASN

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 ⓘ

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	420/425 (98%)	0.49	18 (4%)	40	41	49, 77, 118, 142	0
1	B	401/425 (94%)	0.31	12 (2%)	52	54	49, 71, 110, 135	0
All	All	821/850 (96%)	0.40	30 (3%)	45	47	49, 74, 113, 142	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	2227	VAL	3.6
1	B	2227	VAL	3.6
1	A	2228	PHE	3.2
1	B	2226	ALA	3.1
1	B	2215	VAL	2.8
1	A	1906	ASN	2.7
1	B	2103	VAL	2.7
1	A	2222	ALA	2.5
1	A	1945	PHE	2.4
1	A	2104	TYR	2.4
1	B	2260	GLY	2.4
1	B	2228	PHE	2.4
1	B	2104	TYR	2.3
1	A	1944	MET	2.3
1	B	2218	ILE	2.3
1	B	2257	ILE	2.3
1	A	2152	VAL	2.2
1	A	1912	PHE	2.2
1	B	1945	PHE	2.2
1	A	1892	GLU	2.2
1	A	2221	ASN	2.2
1	A	2226	ALA	2.2
1	A	2134	THR	2.2
1	A	1910	VAL	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1894	GLN	2.2
1	B	2172	GLY	2.2
1	A	1903	GLN	2.2
1	A	1902	VAL	2.0
1	A	1908	GLU	2.0
1	B	1946	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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.