



wwPDB X-ray Structure Validation Summary Report ⓘ

May 25, 2020 – 07:16 am BST

PDB ID : 2ZDS
Title : Crystal Structure of SCO6571 from *Streptomyces coelicolor* A3(2)
Authors : Begum, P.; Gao, Y.G.; Sakai, N.; Yao, M.; Watanabe, N.; Tanaka, I.
Deposited on : 2007-11-27
Resolution : 2.30 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

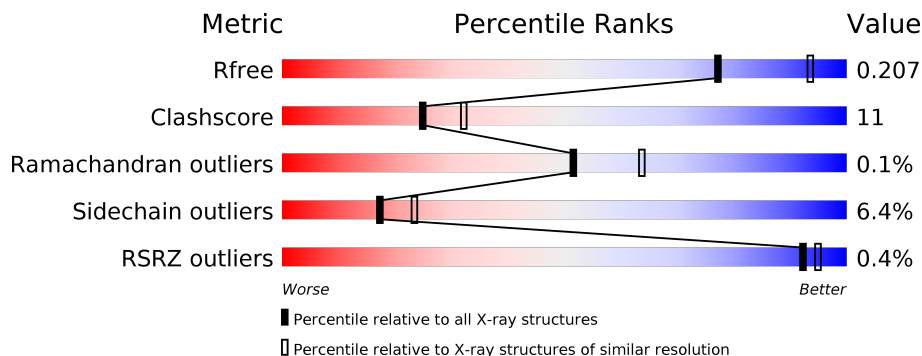
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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 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 $\leq 5\%$. 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	340	
1	B	340	
1	C	340	
1	D	340	
1	E	340	
1	F	340	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 16757 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 Putative DNA-binding protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	321	2584	1645	461	469	5	4	0	0	0
1	B	322	2590	1648	462	471	5	4	0	0	0
1	C	322	2590	1648	462	471	5	4	0	0	0
1	D	322	2590	1648	462	471	5	4	0	0	0
1	E	321	2584	1645	461	469	5	4	0	0	0
1	F	322	2590	1648	462	471	5	4	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	332	LEU	-	EXPRESSION TAG	UNP O69946
A	333	GLU	-	EXPRESSION TAG	UNP O69946
A	334	HIS	-	EXPRESSION TAG	UNP O69946
A	335	HIS	-	EXPRESSION TAG	UNP O69946
A	336	HIS	-	EXPRESSION TAG	UNP O69946
A	337	HIS	-	EXPRESSION TAG	UNP O69946
A	338	HIS	-	EXPRESSION TAG	UNP O69946
A	339	HIS	-	EXPRESSION TAG	UNP O69946
B	332	LEU	-	EXPRESSION TAG	UNP O69946
B	333	GLU	-	EXPRESSION TAG	UNP O69946
B	334	HIS	-	EXPRESSION TAG	UNP O69946
B	335	HIS	-	EXPRESSION TAG	UNP O69946
B	336	HIS	-	EXPRESSION TAG	UNP O69946
B	337	HIS	-	EXPRESSION TAG	UNP O69946
B	338	HIS	-	EXPRESSION TAG	UNP O69946
B	339	HIS	-	EXPRESSION TAG	UNP O69946
C	332	LEU	-	EXPRESSION TAG	UNP O69946

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	333	GLU	-	EXPRESSION TAG	UNP O69946
C	334	HIS	-	EXPRESSION TAG	UNP O69946
C	335	HIS	-	EXPRESSION TAG	UNP O69946
C	336	HIS	-	EXPRESSION TAG	UNP O69946
C	337	HIS	-	EXPRESSION TAG	UNP O69946
C	338	HIS	-	EXPRESSION TAG	UNP O69946
C	339	HIS	-	EXPRESSION TAG	UNP O69946
D	332	LEU	-	EXPRESSION TAG	UNP O69946
D	333	GLU	-	EXPRESSION TAG	UNP O69946
D	334	HIS	-	EXPRESSION TAG	UNP O69946
D	335	HIS	-	EXPRESSION TAG	UNP O69946
D	336	HIS	-	EXPRESSION TAG	UNP O69946
D	337	HIS	-	EXPRESSION TAG	UNP O69946
D	338	HIS	-	EXPRESSION TAG	UNP O69946
D	339	HIS	-	EXPRESSION TAG	UNP O69946
E	332	LEU	-	EXPRESSION TAG	UNP O69946
E	333	GLU	-	EXPRESSION TAG	UNP O69946
E	334	HIS	-	EXPRESSION TAG	UNP O69946
E	335	HIS	-	EXPRESSION TAG	UNP O69946
E	336	HIS	-	EXPRESSION TAG	UNP O69946
E	337	HIS	-	EXPRESSION TAG	UNP O69946
E	338	HIS	-	EXPRESSION TAG	UNP O69946
E	339	HIS	-	EXPRESSION TAG	UNP O69946
F	332	LEU	-	EXPRESSION TAG	UNP O69946
F	333	GLU	-	EXPRESSION TAG	UNP O69946
F	334	HIS	-	EXPRESSION TAG	UNP O69946
F	335	HIS	-	EXPRESSION TAG	UNP O69946
F	336	HIS	-	EXPRESSION TAG	UNP O69946
F	337	HIS	-	EXPRESSION TAG	UNP O69946
F	338	HIS	-	EXPRESSION TAG	UNP O69946
F	339	HIS	-	EXPRESSION TAG	UNP O69946

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	153	Total O 153 153	0	0
2	B	205	Total O 205 205	0	0
2	C	206	Total O 206 206	0	0
2	D	217	Total O 217 217	0	0

Continued on next page...

Continued from previous page...

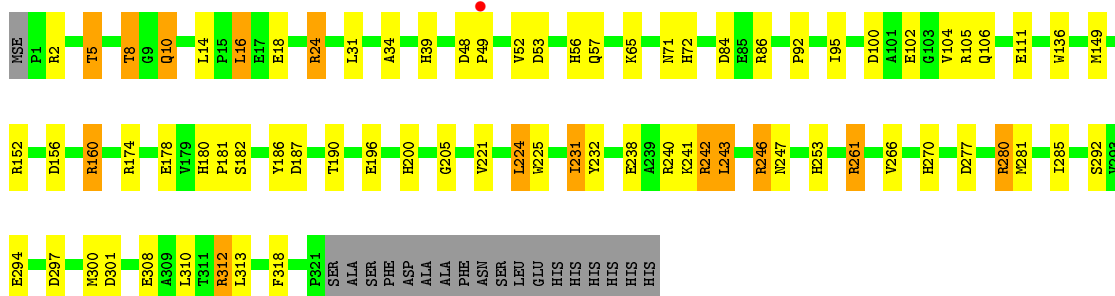
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	E	229	Total 229	O 229	0	0
2	F	219	Total 219	O 219	0	0

3 Residue-property plots [i](#)


These plots are drawn for all protein, RNA and DNA 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.

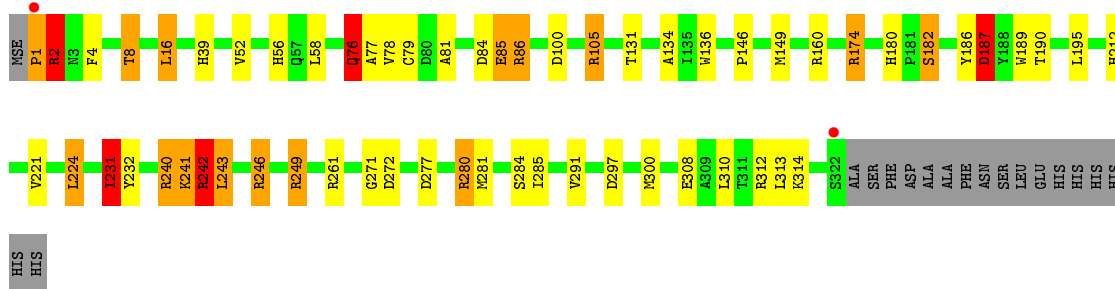
- Molecule 1: Putative DNA-binding protein

Chain A: 



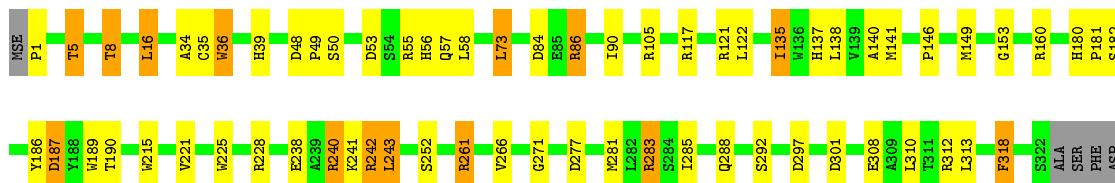
- Molecule 1: Putative DNA-binding protein

Chain B: 



- Molecule 1: Putative DNA-binding protein

Chain C: 



ALA
ALA
PHE
ASN
SER
SER
LEU
GLU
HIS
HIS
HIS
HIS
HIS

- Molecule 1: Putative DNA-binding protein

Chain D: 77% 14% 5%

MSE P1 T5 T8 C9 Q10 L16 R24 D25 A34 H39 D43 V52 H56 Q57 L58 K65 C66 W67 N71 H72 E85 P92 A93 R94 I95 R105 F130 W136 H137 P146 M149 R152 R174 E178 H179 H180 P181 S182

D187 T190 G205 H212 W215 V221 R228 Y232 E238 A239 R240 R241 L243 N247 H253 R261 V266 H270 G271 D277 M281 L282 R283 I285 S292 D297 E308 R312 L313 F318 S322 ALA SER PHE H180 P181 ASP ALA

ALA
PHE
ASN
SER
SER
LEU
GLU
HIS
HIS
HIS
HIS

- Molecule 1: Putative DNA-binding protein

Chain E: 76% 14% 6%

MSE P1 R2 T5 T8 R21 A34 C35 W36 H39 S70 N71 H72 L73 D84 E85 R86 I90 L91 P92 I95 R105 K113 R121 W136 P143 P146 M149 D156 D301 R160 V167 E178 S182 D187 T190 H192

E196 R201 G205 H212 W215 D219 P220 V221 L224 R230 T231 Y232 R240 K241 R242 L243 R246 R261 G271 E276 D277 R280 M281 L282 R283 S284 S292 D297 M300 D301 E308 A309 L310 T311 L313 P321 SER ALA SER PHE

ASP
ALA
PHE
ASN
SER
SER
LEU
GLU
HIS
HIS
HIS
HIS

- Molecule 1: Putative DNA-binding protein

Chain F: 75% 15% 5%

MSE P1 L6 F7 T8 C9 Q10 W11 A12 L16 E17 E18 R21 R24 A34 H39 V52 H56 A68 H72 I82 I83 D84 E85 R86 I90 D98 T1128 T1131 W136 M149 R152 G153 Y154 R160 R174

H177 E178 P181 S182 D187 Y188 R193 H212 W215 D219 W225 E226 F227 R228 I231 Y232 C236 R237 E238 A239 R240 K241 R242 L243 R246 R261 D264 F265 V266 G271 W275 E276 R280 M281 I285 W295 E296 D297 L303 E308

R312 L313 P320 P321 S322 ALA SER PHE ASP ALA ALA PHE ASN ASN LEU GLU HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	84.48Å 171.59Å 184.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 41.91 – 2.29	Depositor EDS
% Data completeness (in resolution range)	99.6 (20.00-2.30) 99.6 (41.91-2.29)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.63 (at 2.29Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.167 , 0.210 0.164 , 0.207	Depositor DCC
R_{free} test set	6032 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	16.5	Xtrriage
Anisotropy	0.130	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 33.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	16757	wwPDB-VP
Average B, all atoms (Å ²)	13.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.28% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.90	1/2663 (0.0%)	1.13	20/3618 (0.6%)
1	B	0.97	5/2669 (0.2%)	1.13	29/3626 (0.8%)
1	C	0.94	1/2669 (0.0%)	1.12	24/3626 (0.7%)
1	D	0.96	1/2669 (0.0%)	1.03	15/3626 (0.4%)
1	E	0.99	1/2663 (0.0%)	1.15	26/3618 (0.7%)
1	F	0.98	3/2669 (0.1%)	1.25	35/3626 (1.0%)
All	All	0.96	12/16002 (0.1%)	1.14	149/21740 (0.7%)

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	B	0	1
1	F	0	1
All	All	0	2

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	76	GLN	CB-CG	9.44	1.78	1.52
1	F	187	ASP	CG-OD1	8.98	1.46	1.25
1	D	187	ASP	CB-CG	-7.95	1.35	1.51
1	C	187	ASP	CG-OD1	7.62	1.42	1.25
1	A	187	ASP	CG-OD1	7.00	1.41	1.25

The worst 5 of 149 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	160	ARG	NE-CZ-NH2	-17.40	111.60	120.30
1	A	280	ARG	NE-CZ-NH2	-17.39	111.60	120.30
1	E	240	ARG	NE-CZ-NH2	-17.13	111.73	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	240	ARG	NE-CZ-NH2	-17.05	111.77	120.30
1	F	187	ASP	CB-CG-OD2	-16.47	103.47	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	1	PRO	Peptide
1	F	320	PRO	Peptide

5.2 Too-close contacts [\(i\)](#)

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	2584	0	2423	69	0
1	B	2590	0	2428	51	0
1	C	2590	0	2428	51	0
1	D	2590	0	2428	58	0
1	E	2584	0	2423	61	0
1	F	2590	0	2427	58	0
2	A	153	0	0	4	0
2	B	205	0	0	6	0
2	C	206	0	0	7	0
2	D	217	0	0	8	0
2	E	229	0	0	8	0
2	F	219	0	0	7	0
All	All	16757	0	14557	326	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 11.

The worst 5 of 326 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:E:281:MSE:CE	1:E:284:SER:HB2	1.29	1.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:76:GLN:CG	1:B:76:GLN:CB	1.78	1.59
1:A:281:MSE:CE	1:A:285:ILE:HG23	1.46	1.44
1:B:281:MSE:HE2	1:B:285:ILE:CG2	1.49	1.43
1:F:281:MSE:CE	1:F:285:ILE:HG23	1.49	1.42

There are no symmetry-related clashes.

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	319/340 (94%)	310 (97%)	9 (3%)	0	100	100
1	B	320/340 (94%)	314 (98%)	6 (2%)	0	100	100
1	C	320/340 (94%)	310 (97%)	10 (3%)	0	100	100
1	D	320/340 (94%)	310 (97%)	10 (3%)	0	100	100
1	E	319/340 (94%)	314 (98%)	5 (2%)	0	100	100
1	F	320/340 (94%)	311 (97%)	8 (2%)	1 (0%)	41	50
All	All	1918/2040 (94%)	1869 (97%)	48 (2%)	1 (0%)	51	64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	320	PRO

5.3.2 Protein sidechains [i](#)

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	260/271 (96%)	241 (93%)	19 (7%)	14	18
1	B	261/271 (96%)	244 (94%)	17 (6%)	17	23
1	C	261/271 (96%)	244 (94%)	17 (6%)	17	23
1	D	261/271 (96%)	244 (94%)	17 (6%)	17	23
1	E	260/271 (96%)	244 (94%)	16 (6%)	18	25
1	F	261/271 (96%)	247 (95%)	14 (5%)	22	30
All	All	1564/1626 (96%)	1464 (94%)	100 (6%)	17	23

5 of 100 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	135	ILE
1	D	8	THR
1	F	174	ARG
1	C	242	ARG
1	C	310	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	288	GLN
1	D	76	GLN
1	F	76	GLN
1	D	10	GLN
1	D	71	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	317/340 (93%)	-0.59	1 (0%) 94 96	6, 16, 29, 38	0
1	B	318/340 (93%)	-0.65	2 (0%) 89 92	6, 12, 24, 41	0
1	C	318/340 (93%)	-0.77	0 100 100	5, 12, 24, 42	0
1	D	318/340 (93%)	-0.75	2 (0%) 89 92	5, 11, 22, 46	0
1	E	317/340 (93%)	-0.79	0 100 100	4, 10, 20, 32	0
1	F	318/340 (93%)	-0.71	2 (0%) 89 92	5, 10, 22, 50	0
All	All	1906/2040 (93%)	-0.71	7 (0%) 92 95	4, 11, 24, 50	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	322	SER	3.5
1	B	1	PRO	3.4
1	A	49	PRO	2.9
1	F	320	PRO	2.8
1	F	322	SER	2.3

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.