



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:41 AM GMT

PDB ID : 2I4N
Title : Rhodopseudomonas palustris prolyl-tRNA synthetase in complex with CysAMS
Authors : Crepin, T.; Yaremchuk, A.; Tukalo, M.; Cusack, S.
Deposited on : 2006-08-22
Resolution : 2.85 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

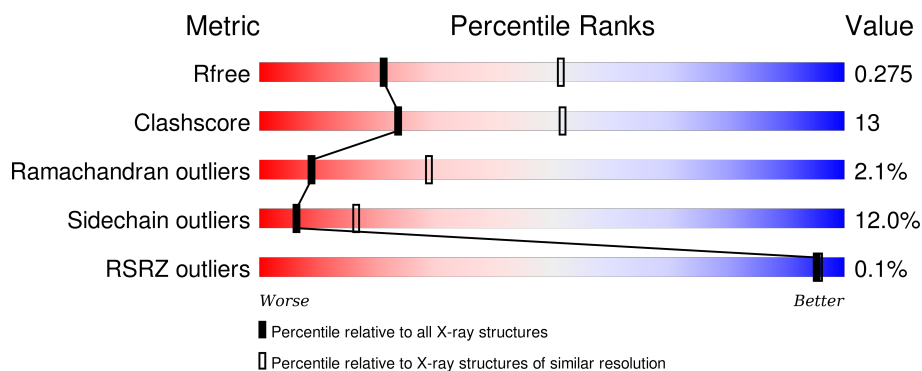
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.85 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}	91344	2228 (2.90-2.82)
Clashscore	102246	2499 (2.90-2.82)
Ramachandran outliers	100387	2439 (2.90-2.82)
Sidechain outliers	100360	2442 (2.90-2.82)
RSRZ outliers	91569	2236 (2.90-2.82)

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. 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	458	
1	B	458	
1	C	458	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10672 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 Proline-tRNA ligase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	441	Total	C	N	O	S	0	0	0
			3495	2211	620	650	14			
1	B	442	Total	C	N	O	S	0	0	0
			3506	2217	624	651	14			
1	C	442	Total	C	N	O	S	0	0	0
			3506	2217	624	651	14			

There are 60 discrepancies between the modelled and reference sequences:

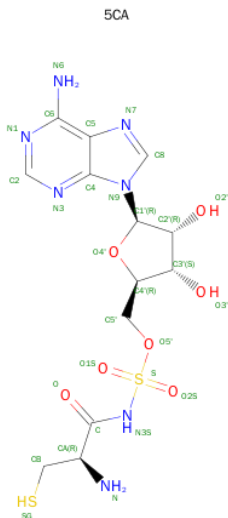
Chain	Residue	Modelled	Actual	Comment	Reference
A	-19	MET	-	EXPRESSION TAG	UNP Q6N5P6
A	-18	GLY	-	EXPRESSION TAG	UNP Q6N5P6
A	-17	SER	-	EXPRESSION TAG	UNP Q6N5P6
A	-16	SER	-	EXPRESSION TAG	UNP Q6N5P6
A	-15	HIS	-	EXPRESSION TAG	UNP Q6N5P6
A	-14	HIS	-	EXPRESSION TAG	UNP Q6N5P6
A	-13	HIS	-	EXPRESSION TAG	UNP Q6N5P6
A	-12	HIS	-	EXPRESSION TAG	UNP Q6N5P6
A	-11	HIS	-	EXPRESSION TAG	UNP Q6N5P6
A	-10	HIS	-	EXPRESSION TAG	UNP Q6N5P6
A	-9	SER	-	EXPRESSION TAG	UNP Q6N5P6
A	-8	SER	-	EXPRESSION TAG	UNP Q6N5P6
A	-7	GLY	-	EXPRESSION TAG	UNP Q6N5P6
A	-6	LEU	-	EXPRESSION TAG	UNP Q6N5P6
A	-5	VAL	-	EXPRESSION TAG	UNP Q6N5P6
A	-4	PRO	-	EXPRESSION TAG	UNP Q6N5P6
A	-3	ARG	-	EXPRESSION TAG	UNP Q6N5P6
A	-2	GLY	-	EXPRESSION TAG	UNP Q6N5P6
A	-1	SER	-	EXPRESSION TAG	UNP Q6N5P6
A	0	HIS	-	EXPRESSION TAG	UNP Q6N5P6
B	-19	MET	-	EXPRESSION TAG	UNP Q6N5P6
B	-18	GLY	-	EXPRESSION TAG	UNP Q6N5P6
B	-17	SER	-	EXPRESSION TAG	UNP Q6N5P6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-16	SER	-	EXPRESSION TAG	UNP Q6N5P6
B	-15	HIS	-	EXPRESSION TAG	UNP Q6N5P6
B	-14	HIS	-	EXPRESSION TAG	UNP Q6N5P6
B	-13	HIS	-	EXPRESSION TAG	UNP Q6N5P6
B	-12	HIS	-	EXPRESSION TAG	UNP Q6N5P6
B	-11	HIS	-	EXPRESSION TAG	UNP Q6N5P6
B	-10	HIS	-	EXPRESSION TAG	UNP Q6N5P6
B	-9	SER	-	EXPRESSION TAG	UNP Q6N5P6
B	-8	SER	-	EXPRESSION TAG	UNP Q6N5P6
B	-7	GLY	-	EXPRESSION TAG	UNP Q6N5P6
B	-6	LEU	-	EXPRESSION TAG	UNP Q6N5P6
B	-5	VAL	-	EXPRESSION TAG	UNP Q6N5P6
B	-4	PRO	-	EXPRESSION TAG	UNP Q6N5P6
B	-3	ARG	-	EXPRESSION TAG	UNP Q6N5P6
B	-2	GLY	-	EXPRESSION TAG	UNP Q6N5P6
B	-1	SER	-	EXPRESSION TAG	UNP Q6N5P6
B	0	HIS	-	EXPRESSION TAG	UNP Q6N5P6
C	-19	MET	-	EXPRESSION TAG	UNP Q6N5P6
C	-18	GLY	-	EXPRESSION TAG	UNP Q6N5P6
C	-17	SER	-	EXPRESSION TAG	UNP Q6N5P6
C	-16	SER	-	EXPRESSION TAG	UNP Q6N5P6
C	-15	HIS	-	EXPRESSION TAG	UNP Q6N5P6
C	-14	HIS	-	EXPRESSION TAG	UNP Q6N5P6
C	-13	HIS	-	EXPRESSION TAG	UNP Q6N5P6
C	-12	HIS	-	EXPRESSION TAG	UNP Q6N5P6
C	-11	HIS	-	EXPRESSION TAG	UNP Q6N5P6
C	-10	HIS	-	EXPRESSION TAG	UNP Q6N5P6
C	-9	SER	-	EXPRESSION TAG	UNP Q6N5P6
C	-8	SER	-	EXPRESSION TAG	UNP Q6N5P6
C	-7	GLY	-	EXPRESSION TAG	UNP Q6N5P6
C	-6	LEU	-	EXPRESSION TAG	UNP Q6N5P6
C	-5	VAL	-	EXPRESSION TAG	UNP Q6N5P6
C	-4	PRO	-	EXPRESSION TAG	UNP Q6N5P6
C	-3	ARG	-	EXPRESSION TAG	UNP Q6N5P6
C	-2	GLY	-	EXPRESSION TAG	UNP Q6N5P6
C	-1	SER	-	EXPRESSION TAG	UNP Q6N5P6
C	0	HIS	-	EXPRESSION TAG	UNP Q6N5P6

- Molecule 2 is 5'-O-(N-(L-CYSTEINYL)-SULFAMOYL)ADENOSINE (three-letter code: 5CA) (formula: C₁₃H₁₉N₇O₇S₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 29	C 13	N 7	O 7	S 2	0	0
2	B	1	Total 29	C 13	N 7	O 7	S 2	0	0
2	C	1	Total 29	C 13	N 7	O 7	S 2	0	0

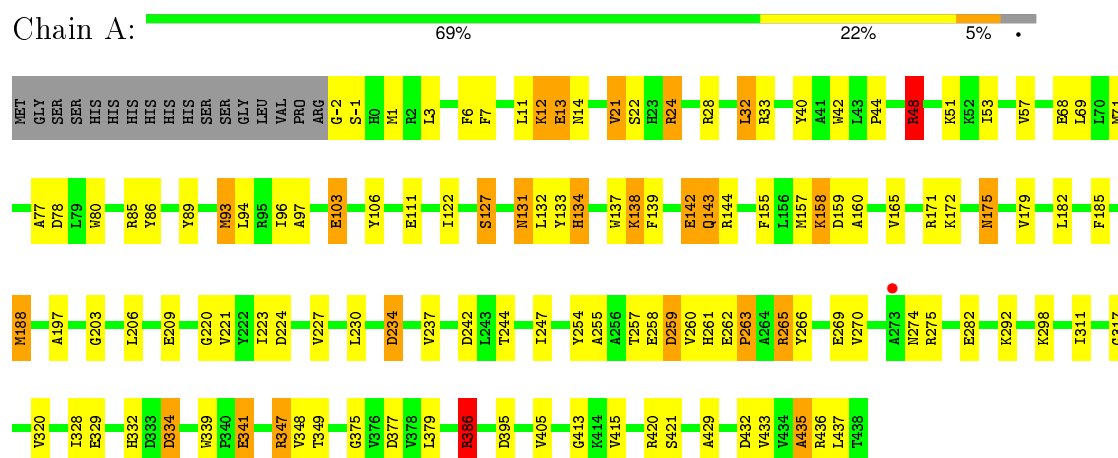
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	28	Total O 28 28	0	0
3	B	30	Total O 30 30	0	0
3	C	20	Total O 20 20	0	0

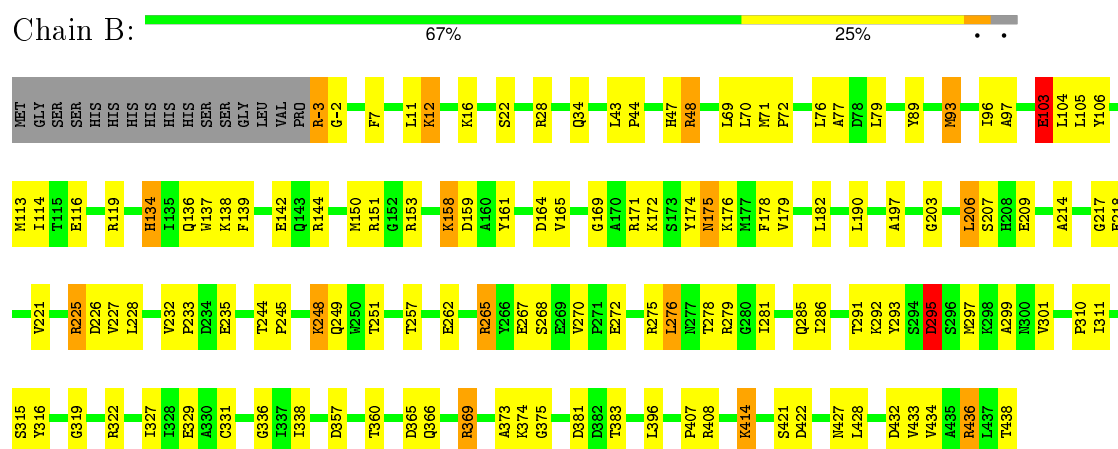
3 Residue-property plots

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.

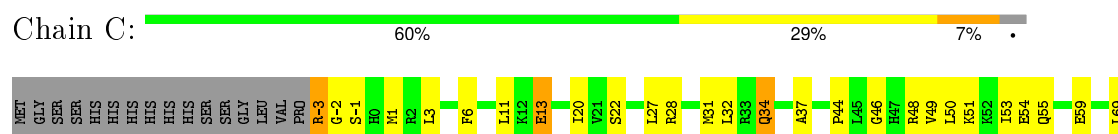
• Molecule 1: Proline-tRNA ligase



• Molecule 1: Proline-tRNA ligase



• Molecule 1: Proline-tRNA ligase



L367	Y368	R369	E370	L371	S372	V376	D377	V378	L379	R386	I399	R408	G409	E416	R419	R420	S421	A424	R425	E426	D432	V433	V434	A435	R436	I437	T438	D159	A160	Y161	V165	D166	Y174	F178	R183	T184	R187	M188	I193	P194	M195	R196	A197	E198	T199	G204	D205	L206	E209	T216	G217	S219	Y222	R225	P233	D234	E235	N236	V237	L243	T244	P245	I246	E247	K248	Q249	S252	E258	D259	V260	H261	Q75	L76	A77	D78	L79	H80	R85	Y86	Y89	E92	M93	L94	R95	R99	H100	E103	L104	L105	M110	E111	E112	M113	I114	T115	E116	S124	Y125	K126	S127	L128	P129	L130	N131	L132	Y133	H134	I135	Q136	W137	K138	E142	Q143	R144	M150	R153	L156	M157	K158
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	217.90Å 107.47Å 110.66Å 90.00° 120.43° 90.00°	Depositor
Resolution (Å)	(Not available) – 2.85 29.84 – 2.70	Depositor EDS
% Data completeness (in resolution range)	94.6 ((Not available)-2.85) 94.6 (29.84-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.01	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.20 (at 2.68Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.211 , 0.277 0.216 , 0.275	Depositor DCC
R_{free} test set	2903 reflections (5.35%)	DCC
Wilson B-factor (Å ²)	63.7	Xtriage
Anisotropy	0.144	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 33.6	EDS
Estimated twinning fraction	0.047 for -h+k-l,-l,-k 0.034 for -h-k-l,l,k 0.064 for -h-2*l,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	0 of 57190 reflections	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	10672	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.69% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: 5CA

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	0.63	0/3570	0.76	2/4828 (0.0%)
1	B	0.63	0/3581	0.75	1/4842 (0.0%)
1	C	0.64	0/3581	0.74	0/4842
All	All	0.64	0/10732	0.75	3/14512 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	48	ARG	NE-CZ-NH1	5.22	122.91	120.30
1	B	79	LEU	CA-CB-CG	5.12	127.08	115.30
1	A	386	ARG	NE-CZ-NH1	5.09	122.84	120.30

There are no chirality outliers.

There are no planarity outliers.

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	3495	0	3466	94	0
1	B	3506	0	3479	89	0
1	C	3506	0	3479	99	0
2	A	29	0	19	1	0
2	B	29	0	19	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	29	0	19	1	0
3	A	28	0	0	3	0
3	B	30	0	0	0	0
3	C	20	0	0	1	0
All	All	10672	0	10481	278	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 13.

The worst 5 of 278 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:C:-3:ARG:HH11	1:C:-3:ARG:HG3	1.10	1.13
1:B:436:ARG:HH11	1:B:436:ARG:HG2	0.95	1.09
1:B:48:ARG:HH11	1:B:48:ARG:HG2	1.18	1.00
1:A:-2:GLY:HA3	1:A:375:GLY:HA2	1.48	0.96
1:B:175:ASN:HD22	1:B:175:ASN:N	1.63	0.96

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	439/458 (96%)	393 (90%)	41 (9%)	5 (1%)	17	47
1	B	440/458 (96%)	392 (89%)	36 (8%)	12 (3%)	6	22
1	C	440/458 (96%)	374 (85%)	55 (12%)	11 (2%)	7	24
All	All	1319/1374 (96%)	1159 (88%)	132 (10%)	28 (2%)	9	29

5 of 28 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	260	VAL
1	A	436	ARG
1	B	12	LYS
1	B	103	GLU
1	C	408	ARG

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	361/376 (96%)	322 (89%)	39 (11%)	8	21
1	B	362/376 (96%)	329 (91%)	33 (9%)	12	31
1	C	362/376 (96%)	304 (84%)	58 (16%)	3	7
All	All	1085/1128 (96%)	955 (88%)	130 (12%)	6	16

5 of 130 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	251	THR
1	C	-3	ARG
1	C	324	LEU
1	B	268	SER
1	B	408	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	300	ASN
1	C	23	HIS
1	C	274	ASN
1	B	285	GLN
1	C	285	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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

3 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z > 2$	Counts	RMSZ	# $ Z > 2$
2	5CA	A	439	-	24,31,31	2.68	4 (16%)	29,46,46	2.59	7 (24%)
2	5CA	B	439	-	24,31,31	2.16	2 (8%)	29,46,46	3.03	10 (34%)
2	5CA	C	439	-	24,31,31	2.43	4 (16%)	29,46,46	2.48	7 (24%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	5CA	A	439	-	-	0/15/37/37	0/3/3/3
2	5CA	B	439	-	-	0/15/37/37	0/3/3/3
2	5CA	C	439	-	-	0/15/37/37	0/3/3/3

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	439	5CA	CB-CA	-2.78	1.49	1.52
2	A	439	5CA	CB-SG	-2.09	1.76	1.81
2	C	439	5CA	C2-N3	2.03	1.35	1.32
2	C	439	5CA	O4'-C1'	2.24	1.44	1.41
2	B	439	5CA	O1S-S	6.22	1.47	1.42

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	439	5CA	N3-C2-N1	-9.36	121.72	128.89
2	A	439	5CA	N3-C2-N1	-9.12	121.91	128.89
2	C	439	5CA	O2S-S-O1S	-8.27	110.86	120.77
2	B	439	5CA	O2S-S-O1S	-7.96	111.24	120.77
2	C	439	5CA	N3-C2-N1	-7.14	123.43	128.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	439	5CA	1	0
2	B	439	5CA	1	0
2	C	439	5CA	1	0

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	441/458 (96%)	-0.26	1 (0%) 95 95	44, 55, 77, 84	0
1	B	442/458 (96%)	-0.33	0 100 100	45, 58, 69, 81	0
1	C	442/458 (96%)	-0.59	0 100 100	33, 63, 86, 101	0
All	All	1325/1374 (96%)	-0.39	1 (0%) 95 95	33, 58, 81, 101	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	273	ALA	2.5

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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
2	5CA	C	439	29/29	0.98	0.12	-0.72	37,43,46,48	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	5CA	A	439	29/29	0.97	0.14	-1.91	46,48,50,51	0
2	5CA	B	439	29/29	0.97	0.13	-1.93	44,50,53,54	0

6.5 Other polymers [i](#)

There are no such residues in this entry.