



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 28, 2020 – 06:06 PM BST

PDB ID : 6T9H  
Title : C171S mutant of Linalool Dehydratase Isomerase  
Authors : Cuetos, A.; Zukic, E.; Danesh-Azari, H.R.; Grogan, G.  
Deposited on : 2019-10-28  
Resolution : 2.58 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](mailto: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  
Xtriage (Phenix) : 1.13  
EDS : 2.14.6  
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) : 2.14.6

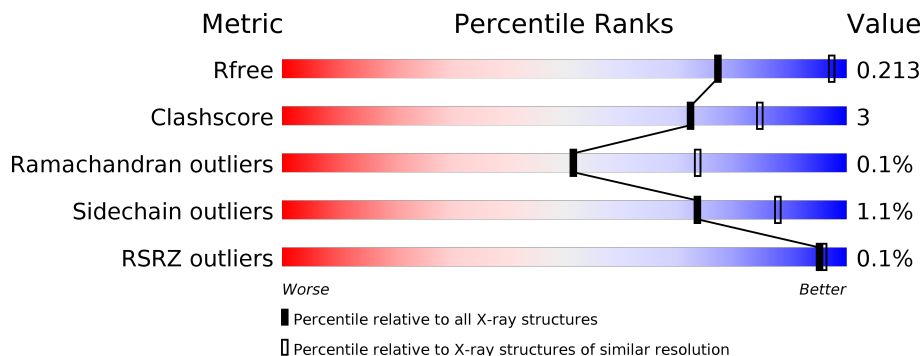
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.58 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	3676 (2.60-2.56)
Clashscore	141614	4049 (2.60-2.56)
Ramachandran outliers	138981	3979 (2.60-2.56)
Sidechain outliers	138945	3979 (2.60-2.56)
RSRZ outliers	127900	3614 (2.60-2.56)

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  $\geq 3$ , 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	372	 90% 8% •
1	B	372	 90% 8% •
1	C	372	 91% 7% •
1	D	372	 90% 8% •
1	E	372	 89% 7% ••
1	S	372	 90% 8% •

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	T	372	 92% 5%
1	U	372	 90% 7%
1	V	372	 88% 9%
1	W	372	 89% 9%

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 29358 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 Linalool dehydratase-isomerase protein LDI.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	363	2864	1856	476	520	12	0	0	0
1	B	363	2862	1855	474	521	12	0	0	0
1	C	363	2853	1844	476	521	12	0	0	0
1	D	363	2863	1851	479	521	12	0	0	0
1	E	362	2855	1848	475	520	12	0	0	0
1	S	363	2856	1847	478	519	12	0	0	0
1	T	362	2831	1836	471	512	12	0	0	0
1	U	362	2807	1820	460	515	12	0	0	0
1	V	361	2842	1836	477	517	12	0	0	0
1	W	362	2849	1844	475	518	12	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP W8X534
A	171	SER	CYS	engineered mutation	UNP W8X534
B	1	MET	-	initiating methionine	UNP W8X534
B	171	SER	CYS	engineered mutation	UNP W8X534
C	1	MET	-	initiating methionine	UNP W8X534
C	171	SER	CYS	engineered mutation	UNP W8X534
D	1	MET	-	initiating methionine	UNP W8X534
D	171	SER	CYS	engineered mutation	UNP W8X534
E	1	MET	-	initiating methionine	UNP W8X534

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	171	SER	CYS	engineered mutation	UNP W8X534
S	1	MET	-	initiating methionine	UNP W8X534
S	171	SER	CYS	engineered mutation	UNP W8X534
T	1	MET	-	initiating methionine	UNP W8X534
T	171	SER	CYS	engineered mutation	UNP W8X534
U	1	MET	-	initiating methionine	UNP W8X534
U	171	SER	CYS	engineered mutation	UNP W8X534
V	1	MET	-	initiating methionine	UNP W8X534
V	171	SER	CYS	engineered mutation	UNP W8X534
W	1	MET	-	initiating methionine	UNP W8X534
W	171	SER	CYS	engineered mutation	UNP W8X534

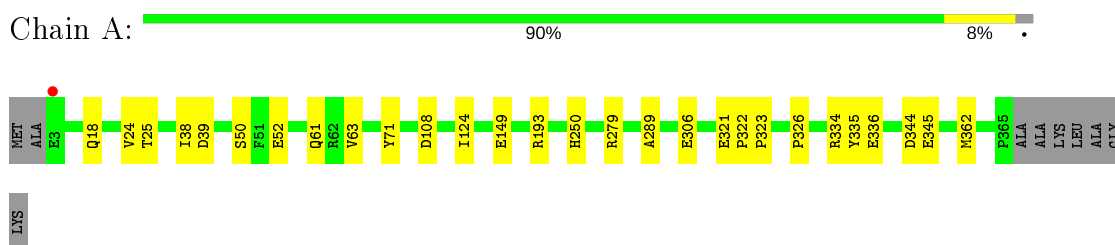
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	108	Total O 108 108	0	0
2	B	92	Total O 92 92	0	0
2	C	87	Total O 87 87	0	0
2	D	102	Total O 102 102	0	0
2	E	102	Total O 102 102	0	0
2	S	84	Total O 84 84	0	0
2	T	66	Total O 66 66	0	0
2	U	59	Total O 59 59	0	0
2	V	94	Total O 94 94	0	0
2	W	82	Total O 82 82	0	0

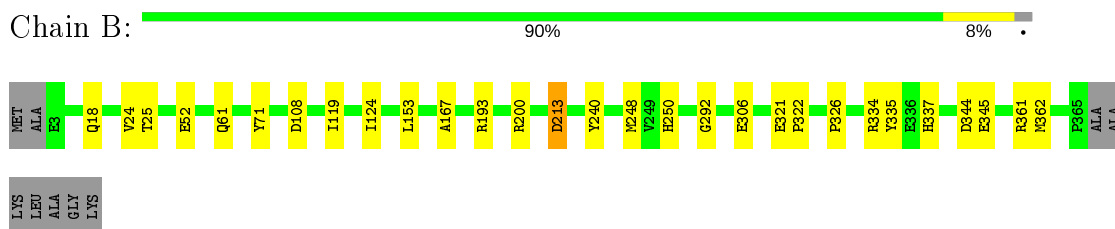
### 3 Residue-property plots [i](#)

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.

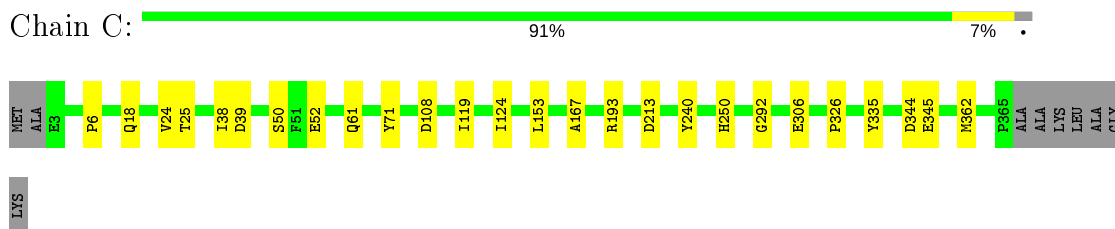
- Molecule 1: Linalool dehydratase-isomerase protein LDI



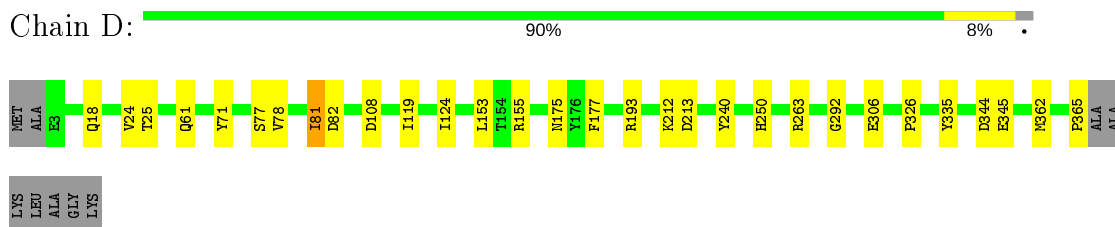
- Molecule 1: Linalool dehydratase-isomerase protein LDI



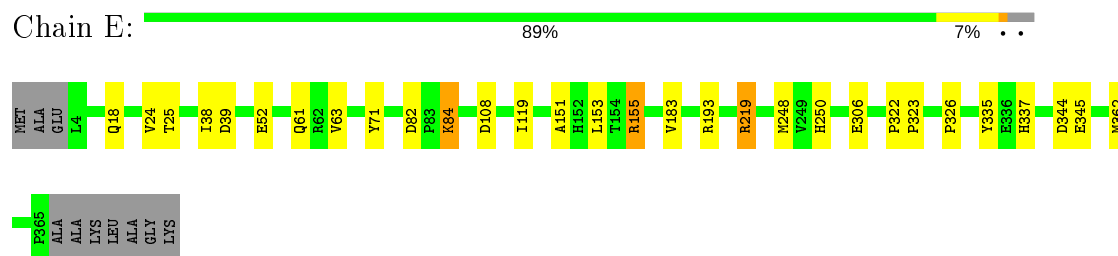
- Molecule 1: Linalool dehydratase-isomerase protein LDI



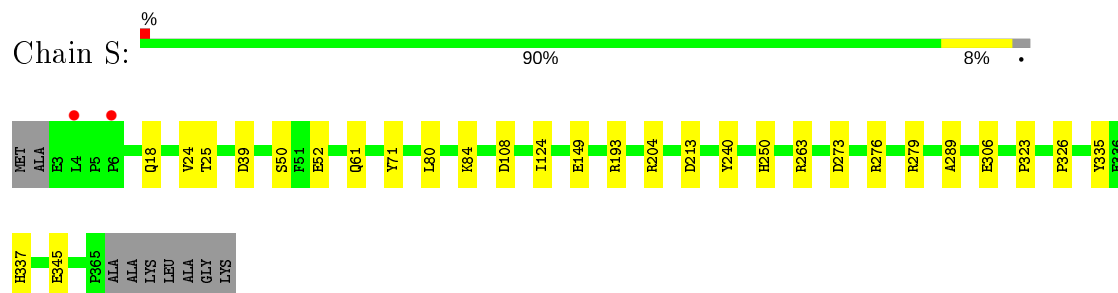
- Molecule 1: Linalool dehydratase-isomerase protein LDI



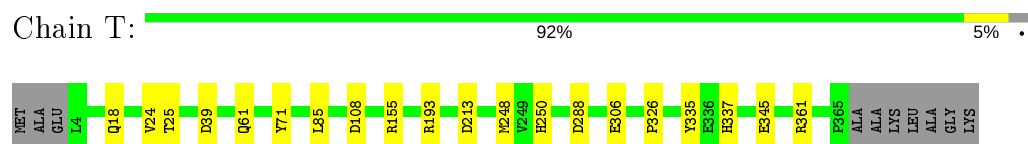
- Molecule 1: Linalool dehydratase-isomerase protein LDI



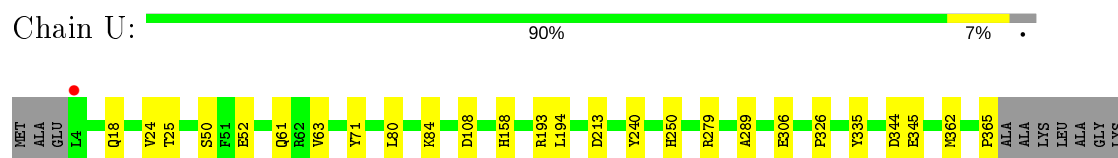
- Molecule 1: Linalool dehydratase-isomerase protein LDI



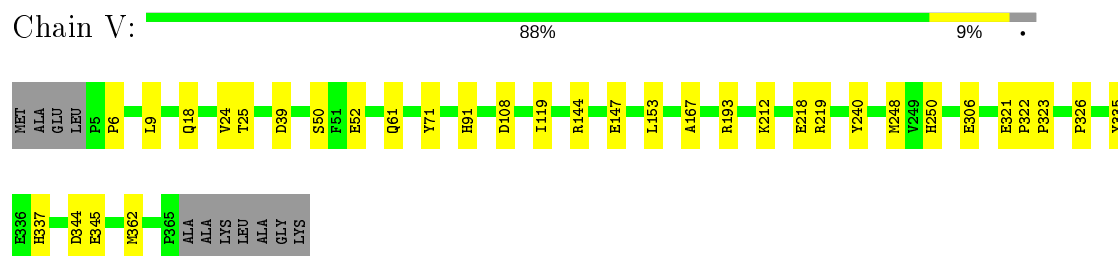
- Molecule 1: Linalool dehydratase-isomerase protein LDI



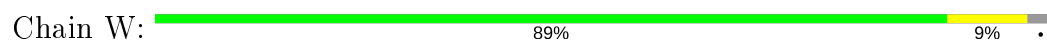
- Molecule 1: Linalool dehydratase-isomerase protein LDI

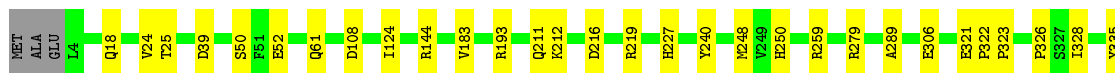


- Molecule 1: Linalool dehydratase-isomerase protein LDI



- Molecule 1: Linalool dehydratase-isomerase protein LDI







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	87.77Å 109.47Å 232.78Å 90.00° 99.53° 90.00°	Depositor
Resolution (Å)	67.99 – 2.58 67.90 – 2.58	Depositor EDS
% Data completeness (in resolution range)	99.2 (67.99-2.58) 99.3 (67.90-2.58)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.13 (at 2.58Å)	Xtrriage
Refinement program	REFMAC 5.8.0257	Depositor
R, $R_{free}$	0.179 , 0.212 0.183 , 0.213	Depositor DCC
$R_{free}$ test set	6674 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.8	Xtrriage
Anisotropy	0.357	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 30.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.003 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	29358	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 35.31 % 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 5.9536e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

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	0.70	0/2950	0.78	0/4017
1	B	0.68	0/2948	0.81	0/4015
1	C	0.69	0/2939	0.78	0/4008
1	D	0.69	0/2949	0.80	0/4018
1	E	0.69	0/2940	0.80	2/4004 (0.0%)
1	S	0.69	0/2942	0.78	3/4010 (0.1%)
1	T	0.68	0/2917	0.79	3/3978 (0.1%)
1	U	0.68	0/2892	0.75	0/3948
1	V	0.69	0/2928	0.79	1/3990 (0.0%)
1	W	0.68	0/2935	0.78	0/4000
All	All	0.69	0/29340	0.79	9/39988 (0.0%)

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	219	ARG	CG-CD-NE	8.80	130.28	111.80
1	T	155	ARG	CG-CD-NE	-7.83	95.35	111.80
1	S	193	ARG	NE-CZ-NH2	7.81	124.21	120.30
1	S	193	ARG	NE-CZ-NH1	-6.07	117.26	120.30
1	E	219	ARG	NE-CZ-NH1	5.57	123.08	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	2864	0	2745	18	0
1	B	2862	0	2740	17	0
1	C	2853	0	2708	19	0
1	D	2863	0	2737	27	0
1	E	2855	0	2738	28	0
1	S	2856	0	2724	16	0
1	T	2831	0	2696	12	0
1	U	2807	0	2650	16	0
1	V	2842	0	2709	21	0
1	W	2849	0	2722	24	0
2	A	108	0	0	4	0
2	B	92	0	0	4	0
2	C	87	0	0	4	0
2	D	102	0	0	4	0
2	E	102	0	0	4	0
2	S	84	0	0	3	0
2	T	66	0	0	3	0
2	U	59	0	0	5	0
2	V	94	0	0	5	0
2	W	82	0	0	2	0
All	All	29358	0	27169	187	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 3.

The worst 5 of 187 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:V:6:PRO:HB2	2:V:467:HOH:O	1.48	1.11
1:E:183:VAL:HA	1:E:248:MET:CE	2.03	0.87
1:W:183:VAL:HA	1:W:248:MET:CE	2.04	0.87
1:D:365:PRO:C	2:D:473:HOH:O	2.13	0.86
1:D:212:LYS:CB	2:D:493:HOH:O	2.24	0.84

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	361/372 (97%)	352 (98%)	9 (2%)	0	100	100
1	B	361/372 (97%)	353 (98%)	8 (2%)	0	100	100
1	C	361/372 (97%)	353 (98%)	8 (2%)	0	100	100
1	D	361/372 (97%)	353 (98%)	8 (2%)	0	100	100
1	E	360/372 (97%)	352 (98%)	8 (2%)	0	100	100
1	S	361/372 (97%)	353 (98%)	8 (2%)	0	100	100
1	T	360/372 (97%)	353 (98%)	7 (2%)	0	100	100
1	U	360/372 (97%)	353 (98%)	7 (2%)	0	100	100
1	V	359/372 (96%)	351 (98%)	7 (2%)	1 (0%)	41	62
1	W	360/372 (97%)	352 (98%)	7 (2%)	1 (0%)	41	62
All	All	3604/3720 (97%)	3525 (98%)	77 (2%)	2 (0%)	51	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	W	212	LYS
1	V	212	LYS

### 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	288/303 (95%)	285 (99%)	3 (1%)	76	89

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	288/303 (95%)	282 (98%)	6 (2%)	53	75
1	C	286/303 (94%)	285 (100%)	1 (0%)	92	97
1	D	289/303 (95%)	286 (99%)	3 (1%)	76	89
1	E	288/303 (95%)	284 (99%)	4 (1%)	67	84
1	S	287/303 (95%)	284 (99%)	3 (1%)	76	89
1	T	282/303 (93%)	281 (100%)	1 (0%)	91	97
1	U	279/303 (92%)	275 (99%)	4 (1%)	67	84
1	V	286/303 (94%)	282 (99%)	4 (1%)	67	84
1	W	287/303 (95%)	285 (99%)	2 (1%)	84	93
All	All	2860/3030 (94%)	2829 (99%)	31 (1%)	73	88

5 of 31 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	84	LYS
1	S	84	LYS
1	V	344	ASP
1	E	155	ARG
1	S	213	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 67 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	310	GLN
1	S	227	HIS
1	W	91	HIS
1	S	18	GLN
1	S	91	HIS

### 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 monosaccharides 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	363/372 (97%)	-0.31	1 (0%) 94   94	27, 38, 57, 102	0
1	B	363/372 (97%)	-0.28	0 100   100	29, 41, 62, 96	0
1	C	363/372 (97%)	-0.26	0 100   100	33, 45, 63, 104	0
1	D	363/372 (97%)	-0.30	0 100   100	27, 39, 58, 95	0
1	E	362/372 (97%)	-0.28	0 100   100	28, 40, 62, 83	0
1	S	363/372 (97%)	-0.23	2 (0%) 89   89	30, 42, 63, 98	0
1	T	362/372 (97%)	-0.29	0 100   100	33, 45, 63, 88	0
1	U	362/372 (97%)	-0.19	1 (0%) 94   94	34, 54, 75, 110	0
1	V	361/372 (97%)	-0.28	0 100   100	29, 42, 59, 86	0
1	W	362/372 (97%)	-0.26	0 100   100	29, 43, 61, 80	0
All	All	3624/3720 (97%)	-0.27	4 (0%) 95   96	27, 43, 65, 110	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	3	GLU	3.0
1	S	6	PRO	2.5
1	S	4	LEU	2.2
1	U	4	LEU	2.1

### 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 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.