



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 10:44 AM GMT

PDB ID : 1UPM  
Title : ACTIVATED SPINACH RUBISCO COMPLEXED WITH 2-CARBOXYARABINITOL 2 BISPHOSPHAT AND CA<sup>2+</sup>.  
Authors : Karkehabadi, S.; Taylor, T.C.; Andersson, I.  
Deposited on : 2003-10-08  
Resolution : 2.30 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

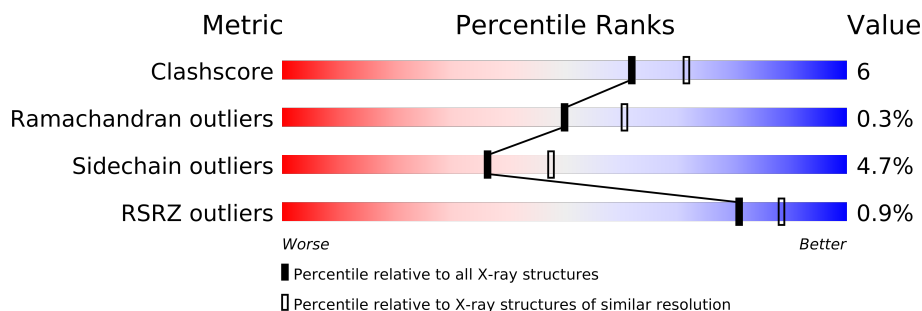
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.30 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))
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (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  $\geq 3$ , 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	B	475	
1	E	475	
1	H	475	
1	K	475	
1	L	475	
1	O	475	
1	R	475	
1	V	475	
2	C	123	
2	F	123	
2	I	123	
2	M	123	
2	P	123	
2	S	123	
2	T	123	
2	W	123	

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 40436 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 RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	467	Total	C	N	O	S	0	0	0
			3652	2316	640	678	18			
1	E	467	Total	C	N	O	S	0	0	0
			3652	2316	640	678	18			
1	H	467	Total	C	N	O	S	0	0	0
			3652	2316	640	678	18			
1	K	467	Total	C	N	O	S	0	0	0
			3652	2316	640	678	18			
1	L	467	Total	C	N	O	S	0	0	0
			3652	2316	640	678	18			
1	O	467	Total	C	N	O	S	0	0	0
			3652	2316	640	678	18			
1	R	467	Total	C	N	O	S	0	0	0
			3652	2316	640	678	18			
1	V	467	Total	C	N	O	S	0	0	0
			3652	2316	640	678	18			

- Molecule 2 is a protein called RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	123	Total	C	N	O	S	0	0	0
			1032	673	167	185	7			
2	F	123	Total	C	N	O	S	0	0	0
			1032	673	167	185	7			
2	I	123	Total	C	N	O	S	0	0	0
			1032	673	167	185	7			
2	M	123	Total	C	N	O	S	0	0	0
			1032	673	167	185	7			
2	P	123	Total	C	N	O	S	0	0	0
			1032	673	167	185	7			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	S	123	Total	C	N	O	S	0	0	0
			1032	673	167	185	7			
2	T	123	Total	C	N	O	S	0	0	0
			1032	673	167	185	7			
2	W	123	Total	C	N	O	S	0	0	0
			1032	673	167	185	7			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	2	GLN	LYS	CONFLICT	UNP Q43832
C	6	ILE	THR	CONFLICT	UNP Q43832
C	7	LEU	GLN	CONFLICT	UNP Q43832
C	9	LEU	MET	CONFLICT	UNP Q43832
C	11	LYS	ARG	CONFLICT	UNP Q43832
C	109	GLU	GLN	CONFLICT	UNP Q43832
C	113	ILE	VAL	CONFLICT	UNP Q43832
F	2	GLN	LYS	CONFLICT	UNP Q43832
F	6	ILE	THR	CONFLICT	UNP Q43832
F	7	LEU	GLN	CONFLICT	UNP Q43832
F	9	LEU	MET	CONFLICT	UNP Q43832
F	11	LYS	ARG	CONFLICT	UNP Q43832
F	109	GLU	GLN	CONFLICT	UNP Q43832
F	113	ILE	VAL	CONFLICT	UNP Q43832
I	2	GLN	LYS	CONFLICT	UNP Q43832
I	6	ILE	THR	CONFLICT	UNP Q43832
I	7	LEU	GLN	CONFLICT	UNP Q43832
I	9	LEU	MET	CONFLICT	UNP Q43832
I	11	LYS	ARG	CONFLICT	UNP Q43832
I	109	GLU	GLN	CONFLICT	UNP Q43832
I	113	ILE	VAL	CONFLICT	UNP Q43832
M	2	GLN	LYS	CONFLICT	UNP Q43832
M	6	ILE	THR	CONFLICT	UNP Q43832
M	7	LEU	GLN	CONFLICT	UNP Q43832
M	9	LEU	MET	CONFLICT	UNP Q43832
M	11	LYS	ARG	CONFLICT	UNP Q43832
M	109	GLU	GLN	CONFLICT	UNP Q43832
M	113	ILE	VAL	CONFLICT	UNP Q43832
P	2	GLN	LYS	CONFLICT	UNP Q43832
P	6	ILE	THR	CONFLICT	UNP Q43832
P	7	LEU	GLN	CONFLICT	UNP Q43832
P	9	LEU	MET	CONFLICT	UNP Q43832

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
P	11	LYS	ARG	CONFLICT	UNP Q43832
P	109	GLU	GLN	CONFLICT	UNP Q43832
P	113	ILE	VAL	CONFLICT	UNP Q43832
S	2	GLN	LYS	CONFLICT	UNP Q43832
S	6	ILE	THR	CONFLICT	UNP Q43832
S	7	LEU	GLN	CONFLICT	UNP Q43832
S	9	LEU	MET	CONFLICT	UNP Q43832
S	11	LYS	ARG	CONFLICT	UNP Q43832
S	109	GLU	GLN	CONFLICT	UNP Q43832
S	113	ILE	VAL	CONFLICT	UNP Q43832
T	2	GLN	LYS	CONFLICT	UNP Q43832
T	6	ILE	THR	CONFLICT	UNP Q43832
T	7	LEU	GLN	CONFLICT	UNP Q43832
T	9	LEU	MET	CONFLICT	UNP Q43832
T	11	LYS	ARG	CONFLICT	UNP Q43832
T	109	GLU	GLN	CONFLICT	UNP Q43832
T	113	ILE	VAL	CONFLICT	UNP Q43832
W	2	GLN	LYS	CONFLICT	UNP Q43832
W	6	ILE	THR	CONFLICT	UNP Q43832
W	7	LEU	GLN	CONFLICT	UNP Q43832
W	9	LEU	MET	CONFLICT	UNP Q43832
W	11	LYS	ARG	CONFLICT	UNP Q43832
W	109	GLU	GLN	CONFLICT	UNP Q43832
W	113	ILE	VAL	CONFLICT	UNP Q43832

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

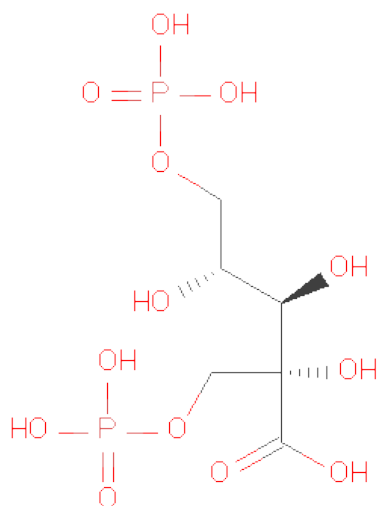
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	K	1	Total Ca 1 1	0	0
3	E	1	Total Ca 1 1	0	0
3	H	1	Total Ca 1 1	0	0
3	B	1	Total Ca 1 1	0	0
3	V	1	Total Ca 1 1	0	0
3	O	1	Total Ca 1 1	0	0
3	R	1	Total Ca 1 1	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	L	1	Total	Ca	0	0
			1	1		

- Molecule 4 is 2-CARBOXYARABINITOL-1,5-DIPHOSPHATE (three-letter code: CAP) (formula:  $C_6H_{14}O_{13}P_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	O	P	0	0
			21	6	13	2		
4	E	1	Total	C	O	P	0	0
			21	6	13	2		
4	H	1	Total	C	O	P	0	0
			21	6	13	2		
4	K	1	Total	C	O	P	0	0
			21	6	13	2		
4	L	1	Total	C	O	P	0	0
			21	6	13	2		
4	O	1	Total	C	O	P	0	0
			21	6	13	2		
4	R	1	Total	C	O	P	0	0
			21	6	13	2		
4	V	1	Total	C	O	P	0	0
			21	6	13	2		

- Molecule 5 is water.

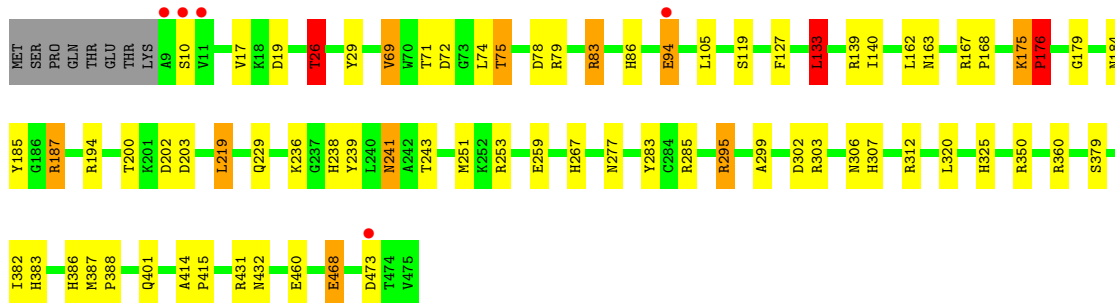
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	263	Total 263	O 263	0	0
5	C	58	Total 58	O 58	0	0
5	E	297	Total 297	O 297	0	0
5	F	100	Total 100	O 100	0	0
5	H	246	Total 246	O 246	0	0
5	I	42	Total 42	O 42	0	0
5	K	278	Total 278	O 278	0	0
5	L	308	Total 308	O 308	0	0
5	M	91	Total 91	O 91	0	0
5	O	261	Total 261	O 261	0	0
5	P	54	Total 54	O 54	0	0
5	R	290	Total 290	O 290	0	0
5	S	93	Total 93	O 93	0	0
5	T	88	Total 88	O 88	0	0
5	V	274	Total 274	O 274	0	0
5	W	45	Total 45	O 45	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.

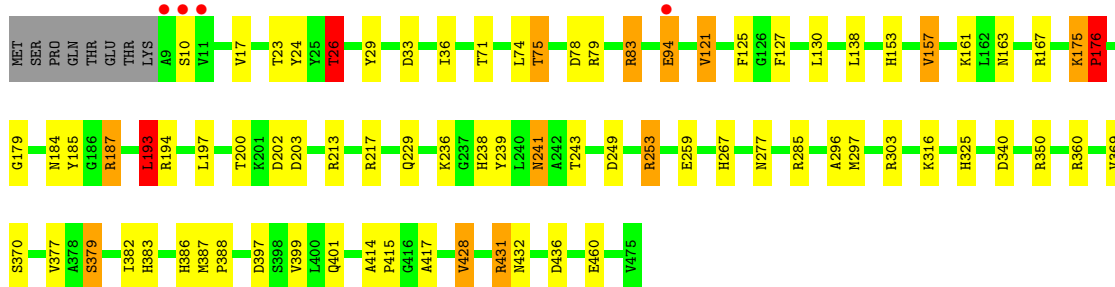
#### • Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN

Chain B: 



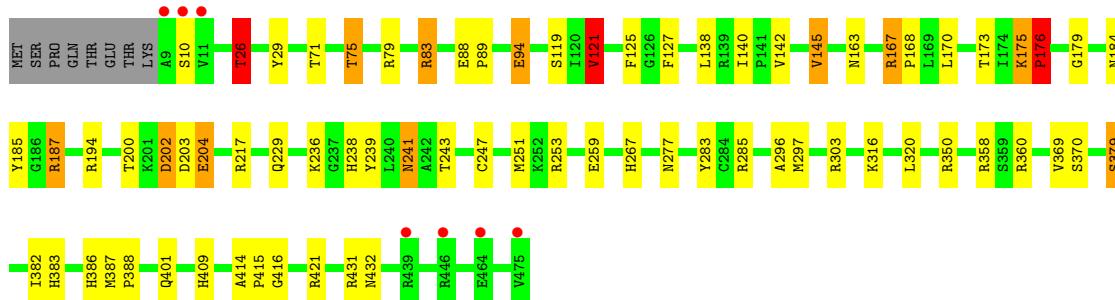
#### • Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN

Chain E: 



#### • Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN

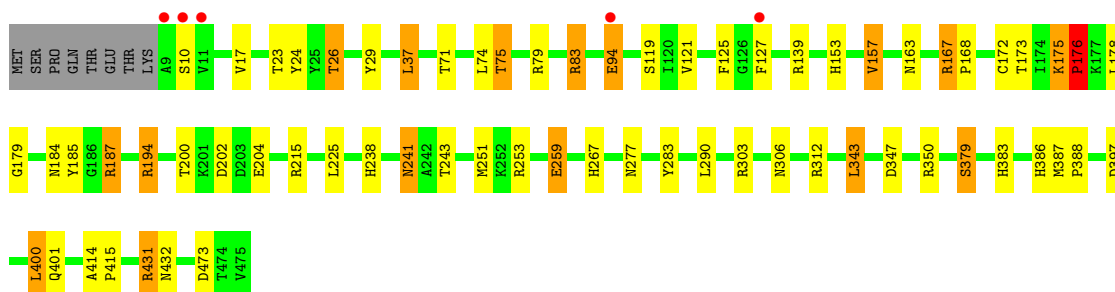
Chain H: 



#### • Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN

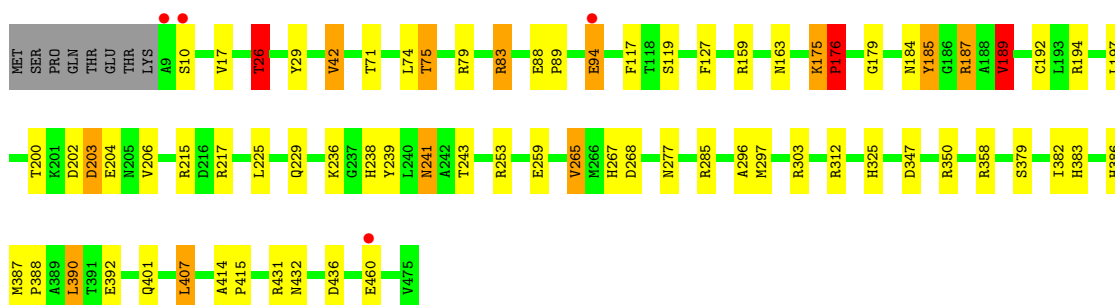


## Chain K:



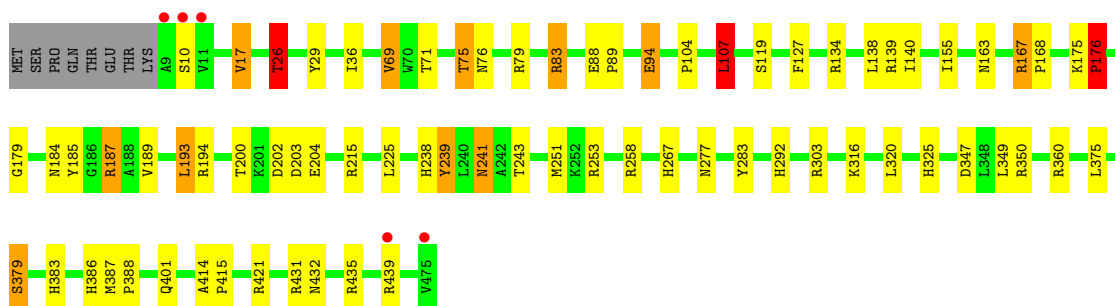
## • Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN

## Chain L:



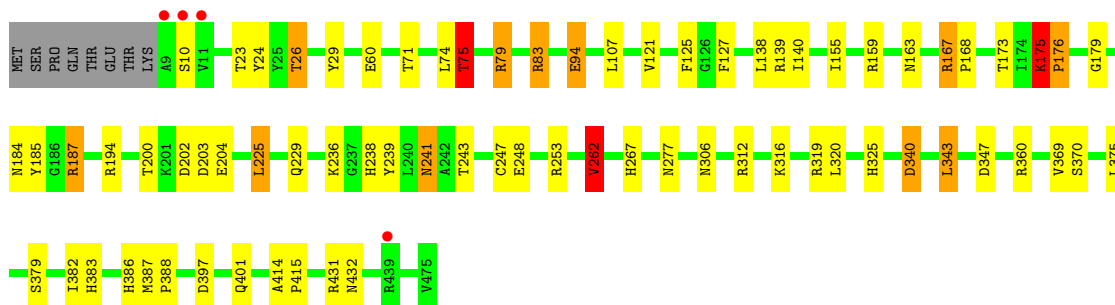
## • Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN

## Chain O:



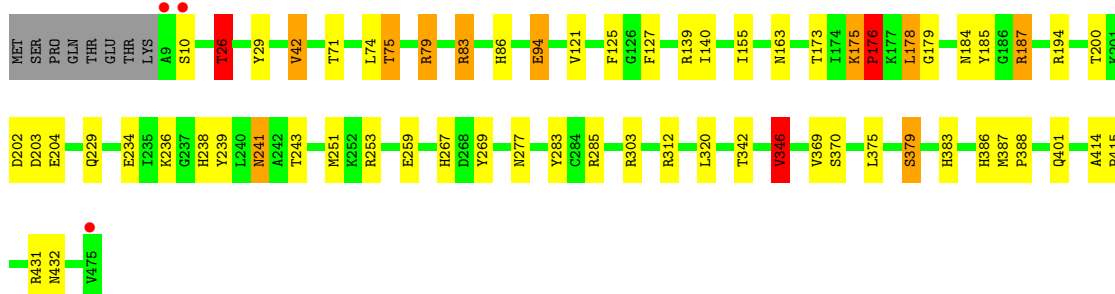
## • Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN

## Chain R:



## • Molecule 1: RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN

Chain V: 



- Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN

Chain C: 



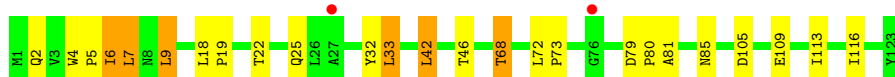
- Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN

Chain F: 



- Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN

Chain I: 



- Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN

Chain M: 



- Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN

Chain P: 



- Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN

Chain S: 



- Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN

Chain T: 



- Molecule 2: RIBULOSE BISPHOSPHATE CARBOXYLASE SMALL CHAIN

Chain W: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	219.60Å 220.94Å 116.98Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 98.33 – 2.30	Depositor EDS
% Data completeness (in resolution range)	90.0 (20.00-2.30) 90.2 (98.33-2.30)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.12 (at 2.29Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.230 , 0.200 0.179 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	24.2	Xtriage
Anisotropy	0.832	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 29.5	EDS
Estimated twinning fraction	0.007 for k,h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 226924 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	40436	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.84% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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	B	0.53	0/3729	1.18	26/5057 (0.5%)
1	E	0.53	0/3729	1.18	29/5057 (0.6%)
1	H	0.52	0/3729	1.15	24/5057 (0.5%)
1	K	0.55	0/3729	1.22	22/5057 (0.4%)
1	L	0.55	0/3729	1.23	33/5057 (0.7%)
1	O	0.55	0/3729	1.21	25/5057 (0.5%)
1	R	0.56	1/3729 (0.0%)	1.22	31/5057 (0.6%)
1	V	0.52	0/3729	1.16	26/5057 (0.5%)
2	C	2.18	1/1067 (0.1%)	1.02	7/1453 (0.5%)
2	F	0.46	0/1067	1.03	4/1453 (0.3%)
2	I	0.44	0/1067	1.03	4/1453 (0.3%)
2	M	0.47	0/1067	1.02	2/1453 (0.1%)
2	P	0.42	0/1067	0.99	3/1453 (0.2%)
2	S	0.45	0/1067	1.06	6/1453 (0.4%)
2	T	0.44	0/1067	1.06	5/1453 (0.3%)
2	W	0.43	0/1067	0.99	3/1453 (0.2%)
All	All	0.63	2/38368 (0.0%)	1.16	250/52080 (0.5%)

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	E	0	1
1	H	0	1
1	K	0	1
1	L	0	1
1	O	0	1
1	R	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
1	V	0	1
All	All	0	8

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	MET	N-CA	69.94	2.86	1.46
1	R	176	PRO	N-CD	8.49	1.59	1.47

The worst 5 of 250 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	R	176	PRO	CA-N-CD	-16.55	88.33	111.50
1	O	303	ARG	NE-CZ-NH1	12.19	126.40	120.30
1	L	187	ARG	NE-CZ-NH2	-12.03	114.28	120.30
1	O	187	ARG	NE-CZ-NH2	-10.51	115.05	120.30
1	L	285	ARG	NE-CZ-NH1	9.71	125.15	120.30

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	175	LYS	Peptide
1	E	175	LYS	Peptide
1	H	175	LYS	Peptide
1	K	175	LYS	Peptide
1	L	175	LYS	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	B	3652	0	3560	40	1
1	E	3652	0	3560	44	0
1	H	3652	0	3560	40	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	K	3652	0	3560	37	0
1	L	3652	0	3560	33	0
1	O	3652	0	3560	39	1
1	R	3652	0	3560	41	0
1	V	3652	0	3560	37	0
2	C	1032	0	990	25	0
2	F	1032	0	990	24	0
2	I	1032	0	990	26	0
2	M	1032	0	990	24	0
2	P	1032	0	990	19	0
2	S	1032	0	990	22	0
2	T	1032	0	990	20	0
2	W	1032	0	990	19	0
3	B	1	0	0	0	0
3	E	1	0	0	0	0
3	H	1	0	0	0	0
3	K	1	0	0	0	0
3	L	1	0	0	0	0
3	O	1	0	0	0	0
3	R	1	0	0	0	0
3	V	1	0	0	0	0
4	B	21	0	7	0	0
4	E	21	0	8	0	0
4	H	21	0	8	0	0
4	K	21	0	8	0	0
4	L	21	0	7	0	0
4	O	21	0	7	0	0
4	R	21	0	8	0	0
4	V	21	0	8	0	0
5	B	263	0	0	6	0
5	C	58	0	0	2	0
5	E	297	0	0	4	0
5	F	100	0	0	3	0
5	H	246	0	0	0	0
5	I	42	0	0	2	0
5	K	278	0	0	3	0
5	L	308	0	0	2	0
5	M	91	0	0	2	0
5	O	261	0	0	1	0
5	P	54	0	0	0	0
5	R	290	0	0	5	0
5	S	93	0	0	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	T	88	0	0	2	0
5	V	274	0	0	2	0
5	W	45	0	0	1	0
All	All	40436	0	36461	420	1

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 6.

The worst 5 of 420 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:K:473:ASP:OD2	5:K:2273:HOH:O	1.71	1.08
2:C:22:THR:H	2:C:25:GLN:HE21	1.14	0.95
1:V:267:HIS:HD2	1:V:277:ASN:HD22	1.14	0.94
1:K:267:HIS:HD2	1:K:277:ASN:HD22	1.15	0.93
1:O:267:HIS:HD2	1:O:277:ASN:HD22	1.12	0.92

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:468:GLU:OE1	1:O:439:ARG:NH1[4_456]	2.13	0.07

## 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	B	464/475 (98%)	440 (95%)	22 (5%)	2 (0%)	43	52
1	E	464/475 (98%)	438 (94%)	24 (5%)	2 (0%)	43	52
1	H	464/475 (98%)	445 (96%)	17 (4%)	2 (0%)	43	52
1	K	464/475 (98%)	443 (96%)	18 (4%)	3 (1%)	33	39
1	L	464/475 (98%)	444 (96%)	18 (4%)	2 (0%)	43	52
1	O	464/475 (98%)	442 (95%)	20 (4%)	2 (0%)	43	52

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	R	464/475 (98%)	442 (95%)	21 (4%)	1 (0%)	56	68
1	V	464/475 (98%)	441 (95%)	21 (4%)	2 (0%)	43	52
2	C	121/123 (98%)	115 (95%)	6 (5%)	0	100	100
2	F	121/123 (98%)	114 (94%)	7 (6%)	0	100	100
2	I	121/123 (98%)	114 (94%)	7 (6%)	0	100	100
2	M	121/123 (98%)	116 (96%)	5 (4%)	0	100	100
2	P	121/123 (98%)	115 (95%)	6 (5%)	0	100	100
2	S	121/123 (98%)	115 (95%)	6 (5%)	0	100	100
2	T	121/123 (98%)	116 (96%)	5 (4%)	0	100	100
2	W	121/123 (98%)	114 (94%)	7 (6%)	0	100	100
All	All	4680/4784 (98%)	4454 (95%)	210 (4%)	16 (0%)	50	60

5 of 16 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	10	SER
1	B	176	PRO
1	E	10	SER
1	E	176	PRO
1	H	10	SER

### 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	B	377/386 (98%)	360 (96%)	17 (4%)	38	50
1	E	377/386 (98%)	362 (96%)	15 (4%)	42	56
1	H	377/386 (98%)	362 (96%)	15 (4%)	42	56
1	K	377/386 (98%)	361 (96%)	16 (4%)	40	53
1	L	377/386 (98%)	358 (95%)	19 (5%)	34	45
1	O	377/386 (98%)	359 (95%)	18 (5%)	35	46
1	R	377/386 (98%)	364 (97%)	13 (3%)	49	64

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	V	377/386 (98%)	364 (97%)	13 (3%)	49	64
2	C	112/112 (100%)	104 (93%)	8 (7%)	21	26
2	F	112/112 (100%)	105 (94%)	7 (6%)	25	32
2	I	112/112 (100%)	105 (94%)	7 (6%)	25	32
2	M	112/112 (100%)	104 (93%)	8 (7%)	21	26
2	P	112/112 (100%)	107 (96%)	5 (4%)	38	50
2	S	112/112 (100%)	105 (94%)	7 (6%)	25	32
2	T	112/112 (100%)	104 (93%)	8 (7%)	21	26
2	W	112/112 (100%)	105 (94%)	7 (6%)	25	32
All	All	3912/3984 (98%)	3729 (95%)	183 (5%)	36	47

5 of 183 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	343	LEU
1	L	390	LEU
1	V	163	ASN
1	L	17	VAL
1	L	127	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 145 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	304	GLN
1	L	304	GLN
1	V	267	HIS
1	K	401	GLN
1	L	184	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

8 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
1	KCX	B	201	1,3	11,11,12	5.42	3 (27%)	10,12,14	2.34	3 (30%)
1	KCX	E	201	1,3	11,11,12	5.96	2 (18%)	10,12,14	2.26	3 (30%)
1	KCX	H	201	1,3	11,11,12	5.90	2 (18%)	10,12,14	2.85	4 (40%)
1	KCX	K	201	1,3	11,11,12	6.01	3 (27%)	10,12,14	2.05	2 (20%)
1	KCX	L	201	1,3	11,11,12	6.27	3 (27%)	10,12,14	1.82	2 (20%)
1	KCX	O	201	1,3	11,11,12	5.57	2 (18%)	10,12,14	2.61	3 (30%)
1	KCX	R	201	1,3	11,11,12	6.02	3 (27%)	10,12,14	1.54	2 (20%)
1	KCX	V	201	1,3	11,11,12	5.61	3 (27%)	10,12,14	2.95	4 (40%)

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
1	KCX	B	201	1,3	-	0/8/10/12	0/0/0/0
1	KCX	E	201	1,3	-	0/8/10/12	0/0/0/0
1	KCX	H	201	1,3	-	0/8/10/12	0/0/0/0
1	KCX	K	201	1,3	-	0/8/10/12	0/0/0/0
1	KCX	L	201	1,3	-	0/8/10/12	0/0/0/0
1	KCX	O	201	1,3	-	0/8/10/12	0/0/0/0
1	KCX	R	201	1,3	-	0/8/10/12	0/0/0/0
1	KCX	V	201	1,3	-	0/8/10/12	0/0/0/0

The worst 5 of 21 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	201	KCX	O-C	20.30	1.25	1.11
1	K	201	KCX	O-C	19.57	1.24	1.11
1	R	201	KCX	O-C	19.50	1.24	1.11
1	E	201	KCX	O-C	19.43	1.24	1.11
1	H	201	KCX	O-C	19.17	1.24	1.11

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	V	201	KCX	OQ2-CX-OQ1	-7.40	112.94	122.17
1	H	201	KCX	OQ2-CX-OQ1	-7.33	113.02	122.17
1	O	201	KCX	OQ2-CX-OQ1	-6.99	113.45	122.17
1	B	201	KCX	OQ2-CX-OQ1	-6.31	114.30	122.17
1	E	201	KCX	OQ2-CX-OQ1	-5.82	114.90	122.17

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 16 ligands modelled in this entry, 8 are monoatomic - leaving 8 for Mogul analysis.

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	$\# Z  > 2$
4	CAP	B	477	3	20,20,20	1.60	4 (20%)	31,31,31	1.97	12 (38%)
4	CAP	E	477	3	20,20,20	1.54	5 (25%)	31,31,31	1.96	11 (35%)
4	CAP	H	477	3	20,20,20	1.51	3 (15%)	31,31,31	1.88	9 (29%)
4	CAP	K	477	3	20,20,20	1.65	7 (35%)	31,31,31	1.99	12 (38%)
4	CAP	L	477	3	20,20,20	1.76	7 (35%)	31,31,31	1.90	11 (35%)
4	CAP	O	477	3	20,20,20	1.61	5 (25%)	31,31,31	1.91	9 (29%)
4	CAP	R	477	3	20,20,20	1.51	4 (20%)	31,31,31	1.91	10 (32%)
4	CAP	V	477	3	20,20,20	1.49	4 (20%)	31,31,31	2.02	10 (32%)

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
4	CAP	B	477	3	-	0/29/29/29	0/0/0/0
4	CAP	E	477	3	-	0/29/29/29	0/0/0/0
4	CAP	H	477	3	-	0/29/29/29	0/0/0/0
4	CAP	K	477	3	-	0/29/29/29	0/0/0/0
4	CAP	L	477	3	-	0/29/29/29	0/0/0/0
4	CAP	O	477	3	-	0/29/29/29	0/0/0/0
4	CAP	R	477	3	-	0/29/29/29	0/0/0/0
4	CAP	V	477	3	-	0/29/29/29	0/0/0/0

The worst 5 of 39 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	H	477	CAP	O4-C4	-4.06	1.34	1.43
4	L	477	CAP	O4-C4	-3.91	1.34	1.43
4	V	477	CAP	O4-C4	-3.75	1.34	1.43
4	R	477	CAP	O4-C4	-3.69	1.34	1.43
4	B	477	CAP	O4-C4	-3.61	1.35	1.43

The worst 5 of 84 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	477	CAP	O3-C3-C4	-4.45	98.48	108.78
4	L	477	CAP	P2-O5-C5	4.38	130.86	118.19
4	K	477	CAP	O6P-P2-O5P	4.15	123.78	107.61
4	H	477	CAP	O6-C-C2	4.11	129.32	122.57
4	H	477	CAP	O7-C-O6	-3.97	111.26	123.76

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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, 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	B	467/475 (98%)	-0.59	5 (1%) 77 85	13, 22, 42, 105	0
1	E	467/475 (98%)	-0.53	4 (0%) 81 88	13, 22, 42, 105	0
1	H	467/475 (98%)	-0.54	7 (1%) 70 78	14, 22, 42, 105	0
1	K	467/475 (98%)	-0.46	5 (1%) 77 85	13, 22, 42, 105	0
1	L	467/475 (98%)	-0.43	4 (0%) 81 88	13, 22, 42, 105	0
1	O	467/475 (98%)	-0.54	5 (1%) 77 85	13, 22, 42, 105	0
1	R	467/475 (98%)	-0.43	4 (0%) 81 88	13, 22, 42, 105	0
1	V	467/475 (98%)	-0.58	3 (0%) 86 92	13, 22, 42, 105	0
2	C	123/123 (100%)	-0.24	0 100 100	19, 32, 51, 66	0
2	F	123/123 (100%)	0.01	1 (0%) 83 90	18, 32, 51, 66	0
2	I	123/123 (100%)	-0.05	2 (1%) 68 77	19, 32, 51, 66	0
2	M	123/123 (100%)	-0.26	0 100 100	18, 32, 51, 66	0
2	P	123/123 (100%)	-0.31	0 100 100	19, 32, 51, 66	0
2	S	123/123 (100%)	0.12	1 (0%) 83 90	18, 32, 51, 66	0
2	T	123/123 (100%)	-0.20	1 (0%) 83 90	18, 32, 51, 66	0
2	W	123/123 (100%)	-0.38	0 100 100	19, 32, 51, 66	0
All	All	4720/4784 (98%)	-0.44	42 (0%) 81 88	13, 23, 48, 105	0

The worst 5 of 42 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	L	9	ALA	10.7
1	V	9	ALA	9.4
1	E	9	ALA	9.4
1	K	9	ALA	8.4
1	H	9	ALA	7.7

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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, 95<sup>th</sup> 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	KCX	E	201	12/13	0.11	0.90	17,18,23,23	0
1	KCX	K	201	12/13	0.11	0.64	17,18,23,24	0
1	KCX	O	201	12/13	0.10	0.53	17,18,24,24	0
1	KCX	H	201	12/13	0.09	0.31	17,18,24,24	0
1	KCX	R	201	12/13	0.10	0.23	17,18,24,24	0
1	KCX	L	201	12/13	0.10	0.01	17,18,23,24	0
1	KCX	B	201	12/13	0.08	-0.15	17,18,24,24	0
1	KCX	V	201	12/13	0.07	-0.92	17,18,23,24	0

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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, 95<sup>th</sup> 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	CA	V	476	1/1	0.12	1.74	42,42,42,42	0
4	CAP	L	477	21/21	0.09	-0.44	16,21,26,26	0
4	CAP	E	477	21/21	0.08	-0.51	16,21,26,26	0
4	CAP	K	477	21/21	0.08	-0.54	16,21,26,27	0
3	CA	L	476	1/1	0.10	-0.56	42,42,42,42	0
4	CAP	B	477	21/21	0.07	-0.70	17,22,26,27	0
3	CA	R	476	1/1	0.09	-0.77	41,41,41,41	0
4	CAP	V	477	21/21	0.07	-0.78	16,21,26,27	0
4	CAP	H	477	21/21	0.07	-0.83	17,21,26,26	0
4	CAP	R	477	21/21	0.07	-0.89	16,21,26,26	0
3	CA	K	476	1/1	0.08	-0.99	42,42,42,42	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CA	E	476	1/1	0.08	-1.22	42,42,42,42	0
4	CAP	O	477	21/21	0.07	-1.24	16,21,26,27	0
3	CA	O	476	1/1	0.07	-1.86	43,43,43,43	0
3	CA	B	476	1/1	0.04	-2.14	42,42,42,42	0
3	CA	H	476	1/1	0.04	-2.29	42,42,42,42	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.