



# wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 03:27 PM GMT

PDB ID : 4E2F  
Title : Crystal Structure of E. coli Aspartate Transcarbamoylase K164E/E239K Mutant in an intermediate state  
Authors : Guo, W.; Kantrowitz, E.R.  
Deposited on : 2012-03-08  
Resolution : 2.80 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>

---

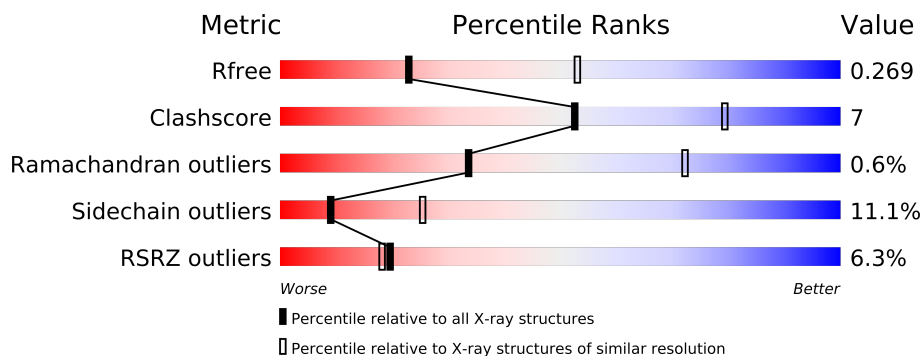
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.80 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))
$R_{free}$	66092	1799 (2.80-2.80)
Clashscore	79885	2295 (2.80-2.80)
Ramachandran outliers	78287	2252 (2.80-2.80)
Sidechain outliers	78261	2254 (2.80-2.80)
RSRZ outliers	66119	1802 (2.80-2.80)

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	A	310	
1	C	310	
1	E	310	
1	G	310	
1	I	310	
1	K	310	
2	B	153	
2	D	153	
2	F	153	
2	H	153	
2	J	153	
2	L	153	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 21764 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 Aspartate carbamoyltransferase catalytic chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	I	310	Total	C	N	O	S	0	0	0
			2415	1527	423	456	9			
1	K	310	Total	C	N	O	S	0	0	0
			2415	1527	423	456	9			
1	G	310	Total	C	N	O	S	0	0	0
			2415	1527	423	456	9			
1	C	310	Total	C	N	O	S	0	0	0
			2415	1527	423	456	9			
1	A	310	Total	C	N	O	S	0	0	0
			2415	1527	423	456	9			
1	E	310	Total	C	N	O	S	0	0	0
			2415	1527	423	456	9			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	164	GLU	LYS	ENGINEERED MUTATION	UNP P0A786
I	239	LYS	GLU	ENGINEERED MUTATION	UNP P0A786
K	164	GLU	LYS	ENGINEERED MUTATION	UNP P0A786
K	239	LYS	GLU	ENGINEERED MUTATION	UNP P0A786
G	164	GLU	LYS	ENGINEERED MUTATION	UNP P0A786
G	239	LYS	GLU	ENGINEERED MUTATION	UNP P0A786
C	164	GLU	LYS	ENGINEERED MUTATION	UNP P0A786
C	239	LYS	GLU	ENGINEERED MUTATION	UNP P0A786
A	164	GLU	LYS	ENGINEERED MUTATION	UNP P0A786
A	239	LYS	GLU	ENGINEERED MUTATION	UNP P0A786
E	164	GLU	LYS	ENGINEERED MUTATION	UNP P0A786
E	239	LYS	GLU	ENGINEERED MUTATION	UNP P0A786

- Molecule 2 is a protein called Aspartate carbamoyltransferase regulatory chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	144	Total	C	N	O	S	0	0	0
			1127	707	199	216	5			
2	B	144	Total	C	N	O	S	0	0	0
			1127	707	199	216	5			
2	J	144	Total	C	N	O	S	0	0	0
			1127	707	199	216	5			
2	L	144	Total	C	N	O	S	0	0	0
			1127	707	199	216	5			
2	H	144	Total	C	N	O	S	0	0	0
			1127	707	199	216	5			
2	F	144	Total	C	N	O	S	0	0	0
			1127	707	199	216	5			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	J	1	Total	Zn	0	0
			1	1		
3	D	1	Total	Zn	0	0
			1	1		
3	H	1	Total	Zn	0	0
			1	1		
3	B	1	Total	Zn	0	0
			1	1		
3	L	1	Total	Zn	0	0
			1	1		
3	F	1	Total	Zn	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	I	49	Total	O	0	0
			49	49		
4	K	44	Total	O	0	0
			44	44		
4	G	41	Total	O	0	0
			41	41		
4	C	52	Total	O	0	0
			52	52		
4	A	78	Total	O	0	0
			78	78		
4	D	32	Total	O	0	0
			32	32		

*Continued on next page...*

*Continued from previous page...*

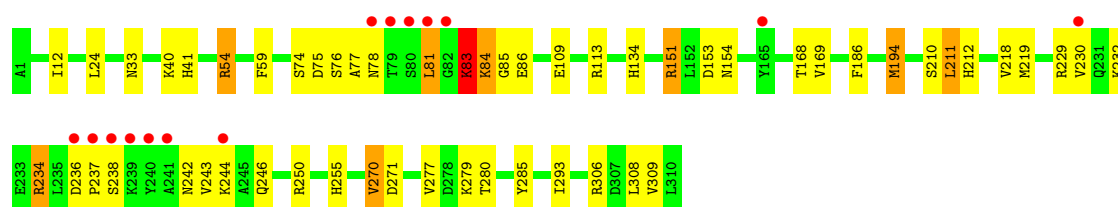
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	35	Total 35	O 35	0	0
4	E	71	Total 71	O 71	0	0
4	J	21	Total 21	O 21	0	0
4	L	26	Total 26	O 26	0	0
4	H	16	Total 16	O 16	0	0
4	F	41	Total 41	O 41	0	0

### 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.

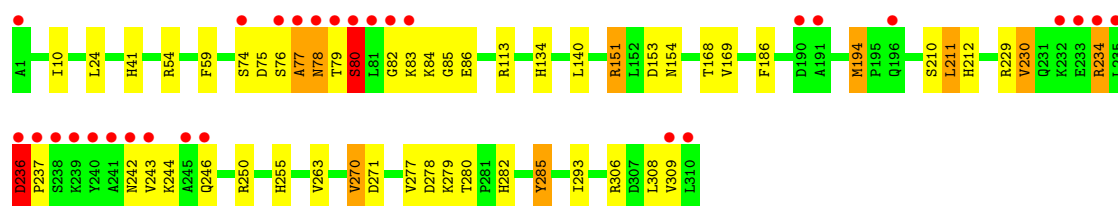
- Molecule 1: Aspartate carbamoyltransferase catalytic chain

Chain I: 



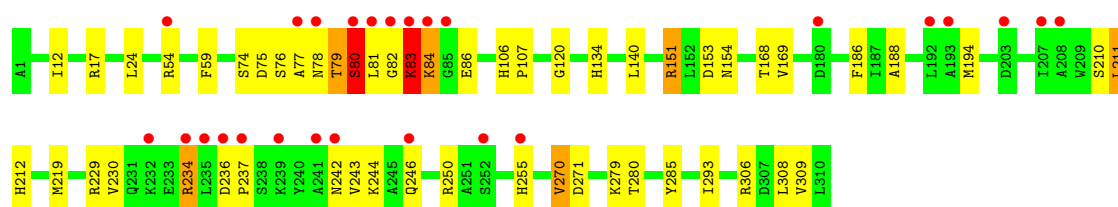
- Molecule 1: Aspartate carbamoyltransferase catalytic chain

Chain K: 



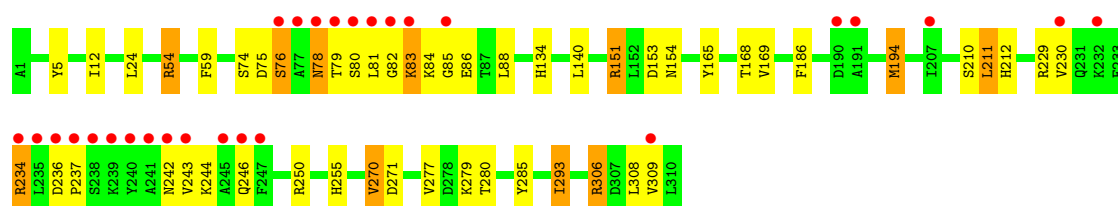
- Molecule 1: Aspartate carbamoyltransferase catalytic chain

Chain G: 

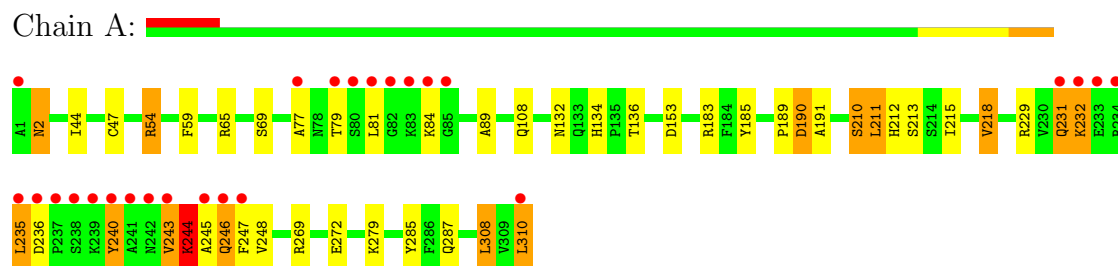


- Molecule 1: Aspartate carbamoyltransferase catalytic chain

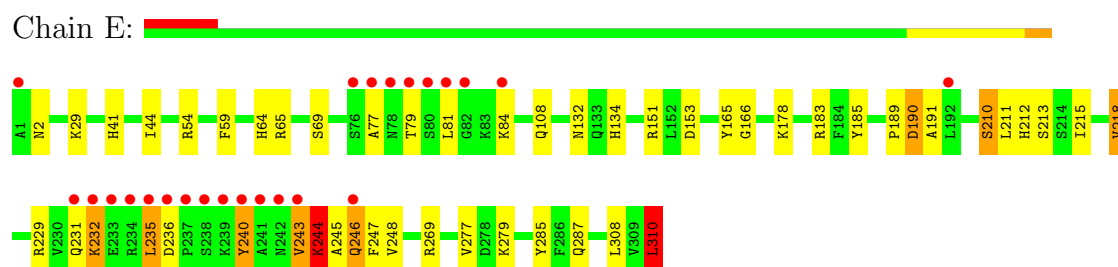
Chain C: 



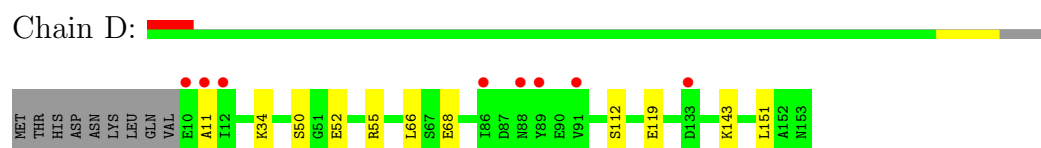
- Molecule 1: Aspartate carbamoyltransferase catalytic chain



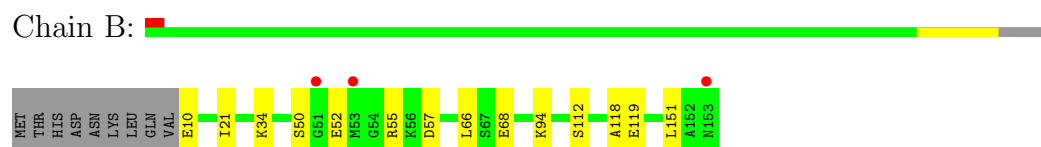
- Molecule 1: Aspartate carbamoyltransferase catalytic chain



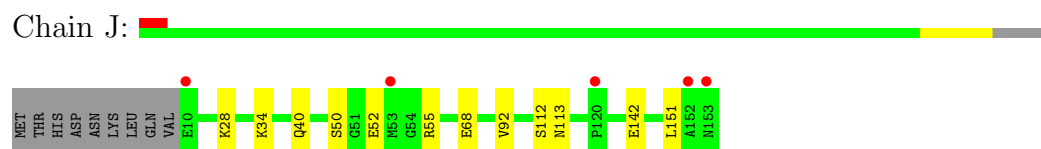
- Molecule 2: Aspartate carbamoyltransferase regulatory chain



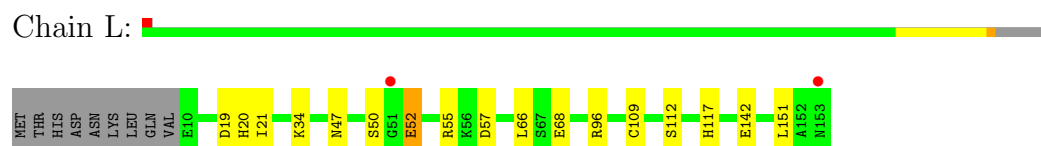
- Molecule 2: Aspartate carbamoyltransferase regulatory chain



- Molecule 2: Aspartate carbamoyltransferase regulatory chain



- Molecule 2: Aspartate carbamoyltransferase regulatory chain



- Molecule 2: Aspartate carbamoyltransferase regulatory chain





● Molecule 2: Aspartate carbamoyltransferase regulatory chain

Chain F:





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	124.09Å 144.81Å 203.36Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.89 – 2.80 29.89 – 2.80	Depositor EDS
% Data completeness (in resolution range)	92.4 (29.89-2.80) 92.4 (29.89-2.80)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.84 (at 2.80Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, $R_{free}$	0.214 , 0.274 0.207 , 0.269	Depositor DCC
$R_{free}$ test set	4205 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	75.6	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 49.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.53$ , $\langle L^2 \rangle = 0.37$	Xtriage
Outliers	0 of 83807 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	21764	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.54% 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: ZN

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	0.73	1/2461 (0.0%)	0.86	3/3339 (0.1%)
1	C	0.59	0/2461	0.70	2/3339 (0.1%)
1	E	0.76	0/2461	0.89	3/3339 (0.1%)
1	G	0.63	3/2461 (0.1%)	1.02	13/3339 (0.4%)
1	I	0.60	0/2461	0.72	2/3339 (0.1%)
1	K	0.60	0/2461	0.79	6/3339 (0.2%)
2	B	0.61	0/1144	0.67	0/1546
2	D	0.65	0/1144	0.69	0/1546
2	F	0.65	0/1144	0.69	0/1546
2	H	0.64	0/1144	0.71	0/1546
2	J	0.64	0/1144	0.72	1/1546 (0.1%)
2	L	0.62	0/1144	0.70	0/1546
All	All	0.65	4/21630 (0.0%)	0.79	30/29310 (0.1%)

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	C	0	1
2	H	0	2
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	G	83	LYS	C-O	-8.77	1.06	1.23
1	G	82	GLY	CA-C	6.28	1.61	1.51
1	G	234	ARG	CZ-NH1	6.09	1.41	1.33
1	A	47	CYS	CB-SG	-5.86	1.72	1.81

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	234	ARG	NE-CZ-NH2	-30.54	105.03	120.30
1	G	234	ARG	NE-CZ-NH1	15.91	128.25	120.30
1	K	236	ASP	CB-CG-OD1	-14.09	105.62	118.30
1	G	17	ARG	NE-CZ-NH2	10.36	125.48	120.30
1	G	17	ARG	NE-CZ-NH1	-10.35	115.13	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	82	GLY	Peptide
2	H	11	ALA	Peptide
2	H	12	ILE	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	A	2415	0	0	29	0
1	C	2415	0	0	19	0
1	E	2415	0	0	31	0
1	G	2415	0	0	17	0
1	I	2415	0	0	21	0
1	K	2415	0	0	18	0
2	B	1127	0	0	5	0
2	D	1127	0	0	3	0
2	F	1127	0	0	1	0
2	H	1127	0	0	5	0
2	J	1127	0	0	4	0
2	L	1127	0	0	8	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
3	H	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	J	1	0	0	0	0
3	L	1	0	0	0	0
4	A	78	0	0	5	0
4	B	35	0	0	3	0
4	C	52	0	0	3	0
4	D	32	0	0	1	0
4	E	71	0	0	4	0
4	F	41	0	0	1	0
4	G	41	0	0	4	0
4	H	16	0	0	2	0
4	I	49	0	0	3	0
4	J	21	0	0	1	0
4	K	44	0	0	4	0
4	L	26	0	0	4	0
All	All	21764	0	0	150	0

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 7.

The worst 5 of 150 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:K:236:ASP:OD1	1:E:165:TYR:OH	1.64	1.13
1:E:189:PRO:CB	1:E:247:PHE:CZ	2.44	1.01
1:A:189:PRO:CB	1:A:247:PHE:CZ	2.47	0.96
1:E:240:TYR:CE1	1:E:244:LYS:NZ	2.34	0.95
1:A:240:TYR:CE1	1:A:244:LYS:NZ	2.42	0.88

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	308/310 (99%)	281 (91%)	24 (8%)	3 (1%)	22 60

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	308/310 (99%)	281 (91%)	25 (8%)	2 (1%)	33	72
1	E	308/310 (99%)	285 (92%)	19 (6%)	4 (1%)	18	51
1	G	308/310 (99%)	284 (92%)	21 (7%)	3 (1%)	22	60
1	I	308/310 (99%)	281 (91%)	27 (9%)	0	100	100
1	K	308/310 (99%)	285 (92%)	20 (6%)	3 (1%)	22	60
2	B	142/153 (93%)	137 (96%)	5 (4%)	0	100	100
2	D	142/153 (93%)	136 (96%)	6 (4%)	0	100	100
2	F	142/153 (93%)	136 (96%)	6 (4%)	0	100	100
2	H	142/153 (93%)	131 (92%)	11 (8%)	0	100	100
2	J	142/153 (93%)	135 (95%)	7 (5%)	0	100	100
2	L	142/153 (93%)	135 (95%)	7 (5%)	0	100	100
All	All	2700/2778 (97%)	2507 (93%)	178 (7%)	15 (1%)	33	72

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	K	78	ASN
1	C	80	SER
1	A	190	ASP
1	E	190	ASP
1	K	77	ALA

### 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	261/261 (100%)	230 (88%)	31 (12%)	8	22
1	C	261/261 (100%)	223 (85%)	38 (15%)	5	13
1	E	261/261 (100%)	230 (88%)	31 (12%)	8	22
1	G	261/261 (100%)	223 (85%)	38 (15%)	5	13
1	I	261/261 (100%)	224 (86%)	37 (14%)	5	14
1	K	261/261 (100%)	223 (85%)	38 (15%)	5	13

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	128/137 (93%)	119 (93%)	9 (7%)	21	52
2	D	128/137 (93%)	120 (94%)	8 (6%)	25	59
2	F	128/137 (93%)	120 (94%)	8 (6%)	25	59
2	H	128/137 (93%)	121 (94%)	7 (6%)	30	65
2	J	128/137 (93%)	121 (94%)	7 (6%)	30	65
2	L	128/137 (93%)	120 (94%)	8 (6%)	25	59
All	All	2334/2388 (98%)	2074 (89%)	260 (11%)	9	25

5 of 260 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	74	SER
1	C	280	THR
2	L	55	ARG
1	C	83	LYS
1	C	211	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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	A	310/310 (100%)	0.02	26 (8%) 11 9	42, 61, 202, 311	0
1	C	310/310 (100%)	0.15	28 (9%) 10 8	51, 79, 210, 318	0
1	E	310/310 (100%)	0.14	24 (7%) 13 12	41, 60, 201, 335	0
1	G	310/310 (100%)	0.05	26 (8%) 11 9	55, 79, 205, 316	0
1	I	310/310 (100%)	-0.20	14 (4%) 32 33	43, 70, 204, 320	0
1	K	310/310 (100%)	0.24	29 (9%) 9 7	51, 78, 203, 307	0
2	B	144/153 (94%)	-0.48	3 (2%) 60 61	51, 71, 121, 189	0
2	D	144/153 (94%)	-0.12	8 (5%) 24 23	57, 78, 136, 190	0
2	F	144/153 (94%)	-0.45	2 (1%) 72 72	52, 72, 125, 189	0
2	H	144/153 (94%)	-0.14	4 (2%) 50 52	63, 84, 128, 189	0
2	J	144/153 (94%)	-0.16	5 (3%) 42 42	50, 74, 124, 188	0
2	L	144/153 (94%)	-0.38	2 (1%) 72 72	55, 73, 127, 190	0
All	All	2724/2778 (98%)	-0.05	171 (6%) 19 18	41, 74, 189, 335	0

The worst 5 of 171 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	236	ASP	15.0
1	E	241	ALA	12.6
1	E	235	LEU	12.6
1	E	237	PRO	11.3
1	K	242	ASN	11.2

### 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 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	ZN	H	201	1/1	0.14	-0.22	98,98,98,98	0
3	ZN	F	201	1/1	0.11	-0.70	58,58,58,58	0
3	ZN	J	201	1/1	0.10	-0.74	75,75,75,75	0
3	ZN	B	201	1/1	0.07	-1.47	59,59,59,59	0
3	ZN	D	201	1/1	0.08	-1.62	59,59,59,59	0
3	ZN	L	201	1/1	0.07	-1.69	57,57,57,57	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.