

wwPDB X-ray Structure Validation Summary Report (i)

Feb 18, 2024 – 08:31 AM EST

PDB ID : 4E2F

Title : Crystal Structure of E. coli Aspartate Transcarbamoylase K164E/E239K Mu-

tant in an intermediate state

Authors : Guo, W.; Kantrowitz, E.R.

Deposited on : 2012-03-08

Resolution : 2.80 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 (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

MolProbity: 4.02b-467 Xtriage (Phenix): 1.13

EDS: 2.36

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)

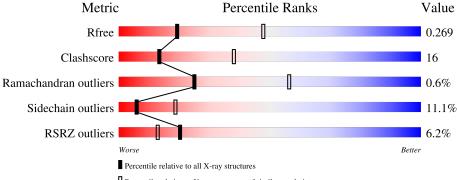
Validation Pipeline (wwPDB-VP) : 2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: X-RAY DIFFRACTION

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.



Percentile relative to X-ray structures of similar resolution

Metric	Whole archive $(\# \mathrm{Entries})$	$\begin{array}{c} {\rm Similar\ resolution} \\ (\#{\rm Entries},{\rm resolution\ range}(\mathring{\rm A})) \end{array}$
R_{free}	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (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 of 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 <=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				
			8%				
1	A	310	69%	24%	6% •		
			9%				
1	С	310	66%	26%	7%		
			8%				
1	E	310	68%	26%	5% •		
			9%				
1	G	310	67%	25%	6% •		
			5%				
1	I	310	67%	26%	6% •		

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quality of chain		
1	K	310	68%	25%	7% •
2	В	153	2%	220/	• 6%
	Ъ	199	71%	22%	• 6%
2	D	153	67%	26%	• 6%
2	F	153	72%	20%	• 6%
2	Н	153	72%	18%	• 6%
2	J	153	68%	25%	• 6%
2	L	153	67%	25%	• 6%



2 Entry composition (i)

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
1	Т	310	Total	С	N	О	S	0	0	0
1	1	310	2415	1527	423	456	9	0	U	
1	K	310	Total	С	N	О	S	0	0	0
1	11	310	2415	1527	423	456	9	0	0	
1	G	310	Total	С	N	О	S	0	0	0
1	G	310	2415	1527	423	456	9	0		
1	С	310	Total	С	N	О	S	0	0	0
1		310	2415	1527	423	456	9	0	0	
1	A	310	Total	С	N	О	S	0	0	0
1	Λ	310	2415	1527	423	456	9	0	0	
1	E	210	Total	С	N	О	S	0	0	0
1	1 E	Ξ 310	2415	1527	423	456	9	0	0	

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
С	164	GLU	LYS	engineered mutation	UNP P0A786
С	239	LYS	GLU	engineered mutation	UNP P0A786
A	164	GLU	LYS	engineered mutation	UNP P0A786
A	239	LYS	GLU	engineered mutation	UNP P0A786
Е	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		At	oms			ZeroOcc	AltConf	Trace
2	D	144	Total	С	N	О	S	0	0	0
2	ע	144	1127	707	199	216	5	0	0	0
2	В	144	Total	С	N	О	S	0	0	0
2	Б	144	1127	707	199	216	5	0	U	0
2	J	144	Total	С	N	О	S	0	0	0
2	J	144	1127	707	199	216	5 5		0	U
2	L	144	Total	С	N	О	S	0	0	0
2	ш	144	1127	707	199	216	5	0		0
2	Н	144	Total	С	N	О	S	0	0	0
2	11	144	1127	707	199	216	5	0	U	0
2	2 F	144	Total	С	N	О	S	0	0	0
	Г	144	1127	707	199	216	5	0	0	U

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	1	Total Zn 1 1	0	0
3	В	1	Total Zn 1 1	0	0
3	J	1	Total Zn 1 1	0	0
3	L	1	Total Zn 1 1	0	0
3	Н	1	Total Zn 1 1	0	0
3	F	1	Total Zn 1 1	0	0

• Molecule 4 is water.

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

Continued on next page...



Continued from previous page...

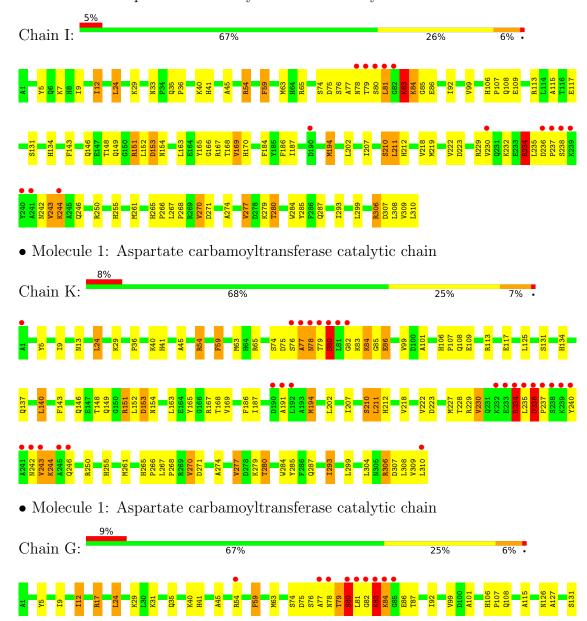
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	В	35	Total O 35 35	0	0
4	Е	71	Total O 71 71	0	0
4	J	21	Total O 21 21	0	0
4	L	26	Total O 26 26	0	0
4	Н	16	Total O 16 16	0	0
4	F	41	Total O 41 41	0	0



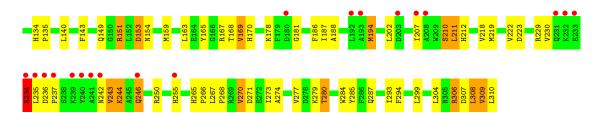
3 Residue-property plots (i)

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.

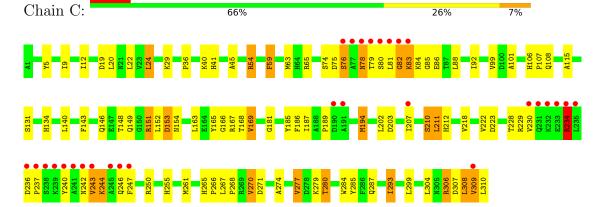
• Molecule 1: Aspartate carbamoyltransferase catalytic chain



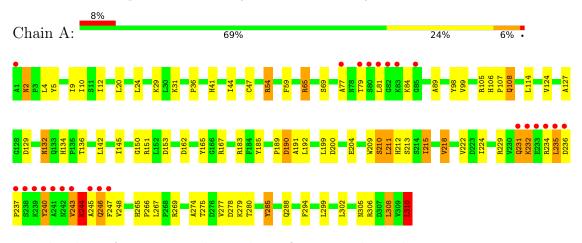




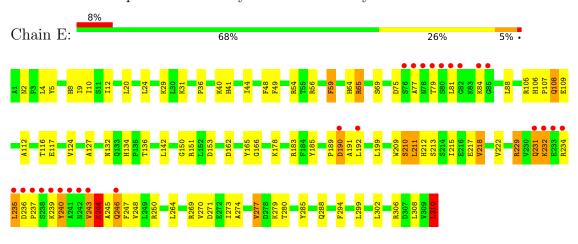
• Molecule 1: Aspartate carbamoyltransferase catalytic chain



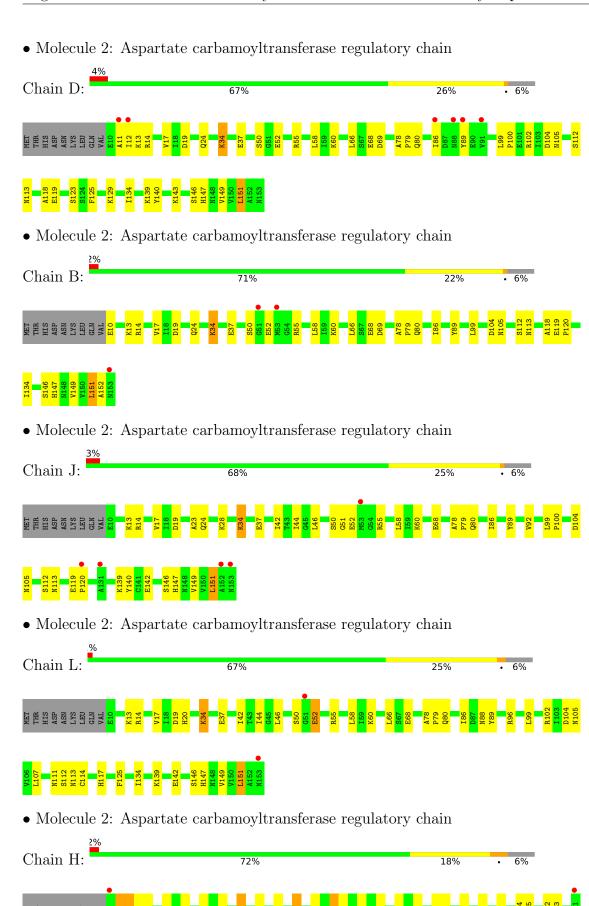
• Molecule 1: Aspartate carbamoyltransferase catalytic chain



• Molecule 1: Aspartate carbamoyltransferase catalytic chain



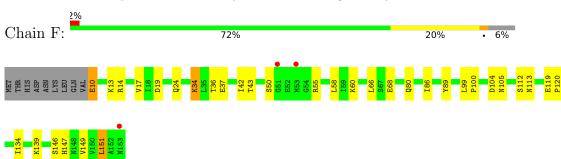








• Molecule 2: Aspartate carbamoyltransferase regulatory chain





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants	124.09Å 144.81Å 203.36Å	Donogiton
a, b, c, α , β , γ	90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.89 - 2.80	Depositor
Resolution (A)	29.89 - 2.80	EDS
% Data completeness	92.4 (29.89-2.80)	Depositor
(in resolution range)	92.4 (29.89-2.80)	EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	1.84 (at 2.80Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
P. P.	0.214 , 0.274	Depositor
R, R_{free}	0.207 , 0.269	DCC
R_{free} test set	4208 reflections $(5.02%)$	wwPDB-VP
Wilson B-factor (Å ²)	75.6	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.29, 65.7	EDS
L-test for twinning ²	$< L >=0.53, < L^2>=0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	21764	wwPDB-VP
Average B, all atoms (Å ²)	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.

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



¹Intensities estimated from amplitudes.

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: 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		Bo	nd lengths	Bond angles		
IVIOI	Chain	RMSZ	# Z > 5	RMSZ	# Z >5	
1	A	0.73	1/2461 (0.0%)	0.86	3/3339 (0.1%)	
1	С	0.59	0/2461	0.70	2/3339~(0.1%)	
1	Е	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	В	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	Н	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	С	0	1
2	Н	0	2
All	All	0	3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	Observed(A)	$\operatorname{Ideal}(ext{\AA})$
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:
110 110100	$\overline{}$	O.	90	OILG	WII SIC	Cathere	COL C	IID CCC	OCIO III .

Mol	Chain	Res	Type	Atoms	Z	$\mathbf{Observed}(^o)$	$Ideal(^{o})$
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	С	82	GLY	Peptide
2	Н	11	ALA	Peptide
2	Н	12	ILE	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	2415	0	2422	112	0
1	С	2415	0	2422	85	0
1	Е	2415	0	2422	114	0
1	G	2415	0	2422	82	0
1	I	2415	0	2422	80	0
1	K	2415	0	2422	84	0
2	В	1127	0	1142	28	0
2	D	1127	0	1143	31	0
2	F	1127	0	1142	26	0
2	Н	1127	0	1142	35	0
2	J	1127	0	1142	27	0
2	L	1127	0	1142	31	0
3	В	1	0	0	0	0
3	D	1	0	0	0	0
3	F	1	0	0	0	0
3	Н	1	0	0	0	0
3	J	1	0	0	0	0
3	L	1	0	0	0	0

Continued on next page...



I 'omtamalod	trom	mmonia	maaa
Continued	11 0116	DICUIUUS	Daue
	.,	10	1

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	78	0	0	7	0
4	В	35	0	0	4	0
4	С	52	0	0	8	0
4	D	32	0	0	7	0
4	Ε	71	0	0	9	0
4	F	41	0	0	1	0
4	G	41	0	0	11	0
4	Н	16	0	0	5	0
4	I	49	0	0	6	0
4	J	21	0	0	2	0
4	K	44	0	0	9	0
4	L	26	0	0	3	0
All	All	21764	0	21385	688	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 16.

The worst 5 of 688 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ (\rm \AA) \end{array}$	$egin{aligned} ext{Clash} \ ext{overlap } (ext{Å}) \end{aligned}$
1:E:189:PRO:HB3	1:E:247:PHE:CZ	1.52	1.43
1:A:189:PRO:HB3	1:A:247:PHE:CZ	1.55	1.41
1:E:189:PRO:CB	1:E:247:PHE:HZ	1.53	1.20
1:A:189:PRO:CB	1:A:247:PHE:HZ	1.56	1.17
1:E:244:LYS:HE2	1:E:247:PHE:CB	1.76	1.16

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	Perce	entiles
1	A	308/310 (99%)	281 (91%)	24 (8%)	3 (1%)	15	44
1	\mathbf{C}	308/310 (99%)	281 (91%)	25 (8%)	2 (1%)	25	56
1	\mathbf{E}	308/310 (99%)	285 (92%)	19 (6%)	4 (1%)	12	36
1	G	308/310 (99%)	284 (92%)	21 (7%)	3 (1%)	15	44
1	I	308/310 (99%)	281 (91%)	27 (9%)	0	100	100
1	K	308/310 (99%)	285 (92%)	20 (6%)	3 (1%)	15	44
2	В	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	Н	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%)	25	56

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	K	78	ASN
1	С	80	SER
1	A	190	ASP
1	Е	190	ASP
1	K	77	ALA

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	261/261 (100%)	230 (88%)	31 (12%)	5 16
1	С	261/261 (100%)	223 (85%)	38 (15%)	3 9
1	Е	261/261 (100%)	230 (88%)	31 (12%)	5 16
1	G	261/261 (100%)	223 (85%)	38 (15%)	3 9

Continued on next page...



Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	261/261 (100%)	224 (86%)	37 (14%)	3 10	
1	K	261/261 (100%)	223 (85%)	38 (15%)	3 9	
2	В	128/137 (93%)	119 (93%)	9 (7%)	15 40	
2	D	128/137 (93%)	120 (94%)	8 (6%)	18 46	
2	F	128/137 (93%)	120 (94%)	8 (6%)	18 46	
2	Н	128/137 (93%)	121 (94%)	7 (6%)	21 52	
2	J	128/137 (93%)	121 (94%)	7 (6%)	21 52	
2	L	128/137 (93%)	120 (94%)	8 (6%)	18 46	
All	All	2334/2388 (98%)	2074 (89%)	260 (11%)	6 19	

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

Mol	Chain	Res	Type
2	J	50	SER
2	L	50	SER
1	G	230	VAL
1	G	210	SER
2	L	112	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	Е	41	HIS
2	Н	113	ASN
1	Е	212	HIS
2	F	113	ASN
2	L	113	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

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.

No monomer is involved in short contacts.

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^{th} 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>	$\# \mathrm{RSRZ}{>}2$	$\mathbf{OWAB}(\mathrm{\AA}^2)$	Q < 0.9
1	A	310/310 (100%)	0.00	24 (7%) 13 7	42, 61, 202, 311	0
1	С	310/310 (100%)	0.11	29 (9%) 8 4	51, 79, 210, 318	0
1	E	310/310 (100%)	0.10	25 (8%) 12 6	41, 60, 201, 335	0
1	G	310/310 (100%)	0.01	28 (9%) 9 5	55, 79, 205, 316	0
1	I	310/310 (100%)	-0.23	14 (4%) 33 23	43, 70, 204, 320	0
1	K	310/310 (100%)	0.17	26 (8%) 11 5	51, 78, 203, 307	0
2	В	144/153 (94%)	-0.44	3 (2%) 63 54	51, 71, 121, 189	0
2	D	144/153 (94%)	-0.14	6 (4%) 36 26	57, 78, 136, 190	0
2	F	144/153 (94%)	-0.40	3 (2%) 63 54	52, 72, 125, 189	0
2	Н	144/153 (94%)	-0.12	3 (2%) 63 54	63, 84, 128, 189	0
2	J	144/153 (94%)	-0.15	5 (3%) 44 34	50, 74, 124, 188	0
2	L	144/153 (94%)	-0.41	2 (1%) 75 70	55, 73, 127, 190	0
All	All	2724/2778 (98%)	-0.07	168 (6%) 20 13	41, 74, 189, 335	0

The worst 5 of 168 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	Ε	236	ASP	14.5
1	Ε	235	LEU	11.7
1	Е	241	ALA	10.8
1	K	191	ALA	10.0
1	E	82	GLY	9.7

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 (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. The B-factors column lists the minimum, median, 95^{th} 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	$\mathbf{B} ext{-}\mathbf{factors}(\mathring{\mathbf{A}}^2)$	Q<0.9
3	ZN	Н	201	1/1	0.98	0.13	98,98,98,98	0
3	ZN	J	201	1/1	0.99	0.10	75,75,75,75	0
3	ZN	L	201	1/1	0.99	0.08	57,57,57,57	0
3	ZN	D	201	1/1	0.99	0.08	59,59,59,59	0
3	ZN	F	201	1/1	0.99	0.11	58,58,58,58	0
3	ZN	В	201	1/1	1.00	0.07	59,59,59,59	0

6.5 Other polymers (i)

There are no such residues in this entry.

