



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 12:03 AM GMT

PDB ID : 4DK1
Title : Crystal Structure of MacA-MexA chimeric protein, containing the Pseudomonas aeruginosa MexA alpha-hairpin domain.
Authors : Xu, Y.; Ha, N.C.
Deposited on : 2012-02-03
Resolution : 3.50 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

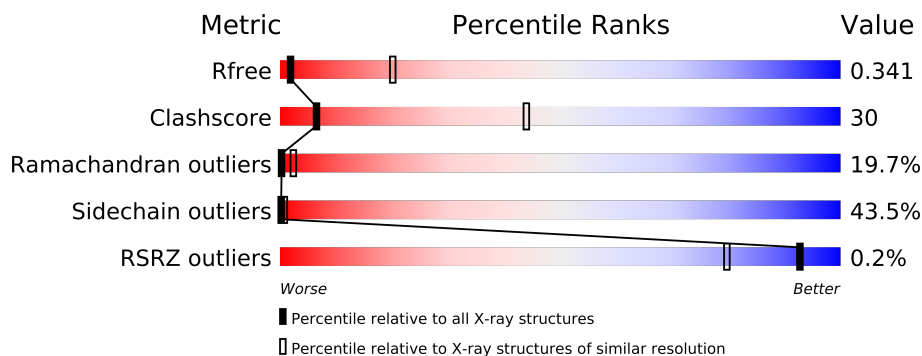
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.50 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	1243 (3.70-3.30)
Clashscore	79885	1039 (3.66-3.34)
Ramachandran outliers	78287	1000 (3.66-3.34)
Sidechain outliers	78261	1000 (3.66-3.34)
RSRZ outliers	66119	1243 (3.70-3.30)

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 ≥ 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	341	
1	B	341	
1	C	341	
1	D	341	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 8568 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 Putative MacA, Multidrug resistance protein mexA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	276	Total	C	N	O	Se	4	0	0
			2108	1316	355	435	2			
1	B	285	Total	C	N	O	Se	4	0	0
			2191	1369	369	451	2			
1	C	278	Total	C	N	O	Se	4	0	0
			2135	1336	358	439	2			
1	D	278	Total	C	N	O	Se	4	0	0
			2134	1334	359	439	2			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	26	GLY	-	EXPRESSION TAG	UNP Q2EHL9
A	27	ALA	-	EXPRESSION TAG	UNP Q2EHL9
A	28	MSE	-	EXPRESSION TAG	UNP Q2EHL9
A	29	ASP	-	EXPRESSION TAG	UNP Q2EHL9
A	68	ILE	VAL	SEE REMARK 999	UNP Q2EHL9
A	85	LEU	ILE	SEE REMARK 999	UNP Q2EHL9
B	26	GLY	-	EXPRESSION TAG	UNP Q2EHL9
B	27	ALA	-	EXPRESSION TAG	UNP Q2EHL9
B	28	MSE	-	EXPRESSION TAG	UNP Q2EHL9
B	29	ASP	-	EXPRESSION TAG	UNP Q2EHL9
B	68	ILE	VAL	SEE REMARK 999	UNP Q2EHL9
B	85	LEU	ILE	SEE REMARK 999	UNP Q2EHL9
C	26	GLY	-	EXPRESSION TAG	UNP Q2EHL9
C	27	ALA	-	EXPRESSION TAG	UNP Q2EHL9
C	28	MSE	-	EXPRESSION TAG	UNP Q2EHL9
C	29	ASP	-	EXPRESSION TAG	UNP Q2EHL9
C	68	ILE	VAL	SEE REMARK 999	UNP Q2EHL9
C	85	LEU	ILE	SEE REMARK 999	UNP Q2EHL9
D	26	GLY	-	EXPRESSION TAG	UNP Q2EHL9
D	27	ALA	-	EXPRESSION TAG	UNP Q2EHL9
D	28	MSE	-	EXPRESSION TAG	UNP Q2EHL9

Continued on next page...

Continued from previous page...

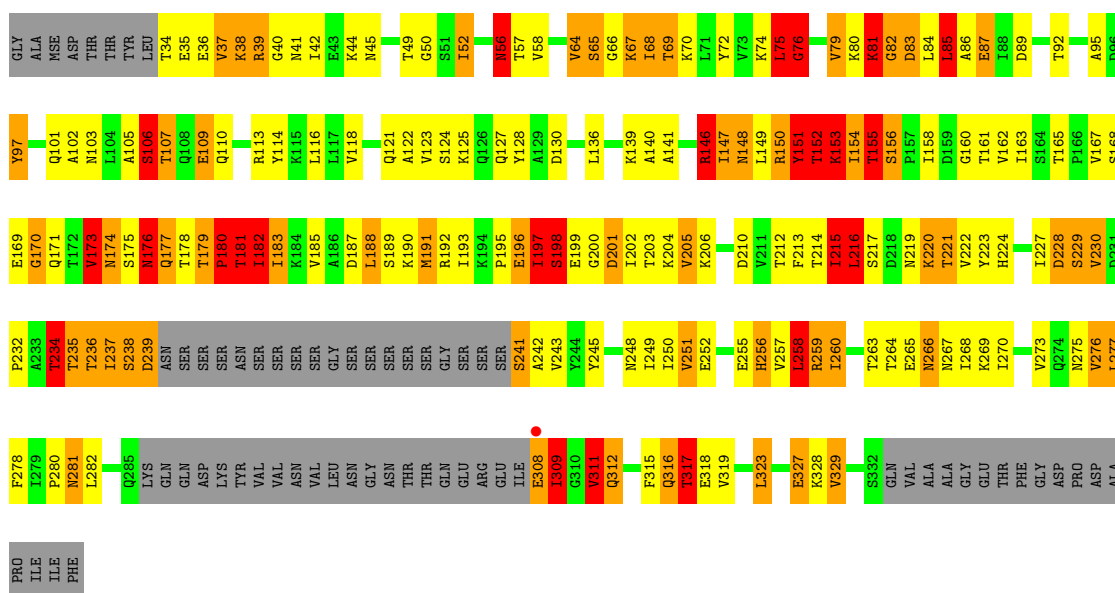
Chain	Residue	Modelled	Actual	Comment	Reference
D	29	ASP	-	EXPRESSION TAG	UNP Q2EHL9
D	68	ILE	VAL	SEE REMARK 999	UNP Q2EHL9
D	85	LEU	ILE	SEE REMARK 999	UNP Q2EHL9

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.

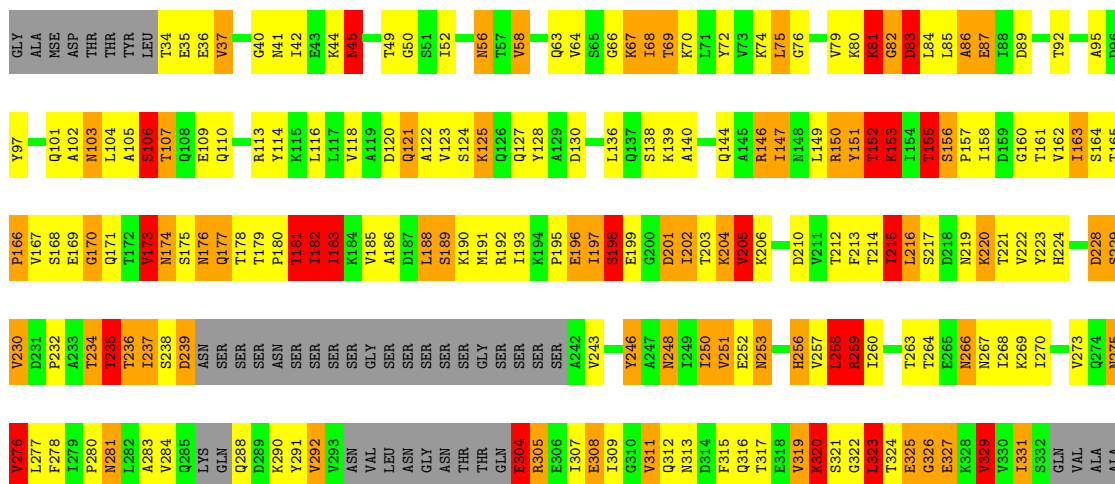
- Molecule 1: Putative MacA, Multidrug resistance protein mexA

Chain A:



- Molecule 1: Putative MacA, Multidrug resistance protein mexA

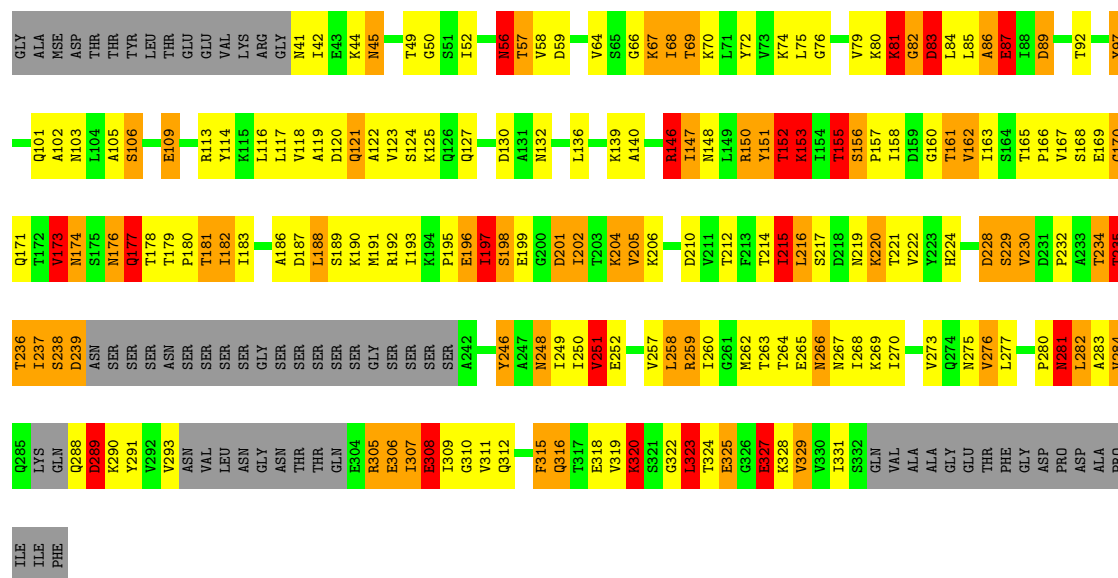
Chain B:



GLY
GLU
THR
PHE
GLY
ASP
PRO
ASP
ALA
PRO
ILE
ILE
PHE

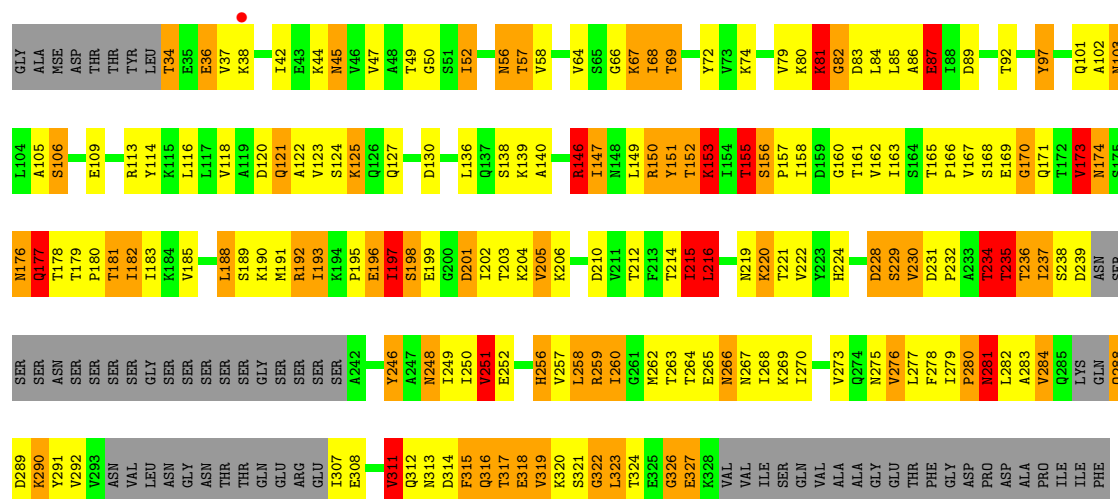
• Molecule 1: Putative MacA, Multidrug resistance protein mexA

Chain C:



• Molecule 1: Putative MacA, Multidrug resistance protein mexA

Chain D:



4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α , β , γ	151.71Å 151.71Å 216.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.95 – 3.50 19.95 – 3.50	Depositor EDS
% Data completeness (in resolution range)	91.3 (19.95-3.50) 91.4 (19.95-3.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.80 (at 3.52Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, R_{free}	0.314 , 0.368 0.289 , 0.341	Depositor DCC
R_{free} test set	1706 reflections (5.07%)	DCC
Wilson B-factor (Å ²)	64.4	Xtriage
Anisotropy	0.394	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.13 , -20.0	EDS
Estimated twinning fraction	0.279 for -h,-k,l	Xtriage
L-test for twinning	$\langle L \rangle = 0.30$, $\langle L^2 \rangle = 0.14$	Xtriage
Outliers	4 of 33690 reflections (0.012%)	Xtriage
F_o, F_c correlation	0.81	EDS
Total number of atoms	8568	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.75% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	1.23	11/2127 (0.5%)	1.47	15/2884 (0.5%)
1	B	1.22	7/2210 (0.3%)	1.44	20/2994 (0.7%)
1	C	0.97	1/2154 (0.0%)	1.35	13/2920 (0.4%)
1	D	1.00	2/2153 (0.1%)	1.35	14/2917 (0.5%)
All	All	1.11	21/8644 (0.2%)	1.41	62/11715 (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	A	0	23
1	B	0	25
1	C	0	22
1	D	0	24
All	All	0	94

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	196	GLU	CG-CD	9.57	1.66	1.51
1	A	196	GLU	CG-CD	8.69	1.65	1.51
1	B	45	ASN	CG-ND2	-6.84	1.15	1.32
1	A	196	GLU	CB-CG	6.58	1.64	1.52
1	B	121	GLN	CD-NE2	-6.55	1.16	1.32

The worst 5 of 62 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	146	ARG	NE-CZ-NH2	16.43	128.51	120.30
1	C	146	ARG	NE-CZ-NH1	16.19	128.40	120.30
1	D	146	ARG	NE-CZ-NH1	16.14	128.37	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	146	ARG	NE-CZ-NH2	14.79	127.70	120.30
1	D	146	ARG	NE-CZ-NH2	-14.74	112.93	120.30

There are no chirality outliers.

5 of 94 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	37	VAL	Peptide
1	A	39	ARG	Peptide
1	A	56	ASN	Peptide
1	A	75	LEU	Peptide
1	A	85	LEU	Mainchain

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	2108	0	0	69	2
1	B	2191	0	0	69	2
1	C	2135	0	0	63	1
1	D	2134	0	0	62	1
All	All	8568	0	0	255	3

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 30.

The worst 5 of 255 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:67:LYS:NZ	1:C:68:ILE:O	1.89	1.05
1:A:67:LYS:NZ	1:A:68:ILE:O	1.89	1.05
1:D:67:LYS:NZ	1:D:68:ILE:O	1.89	1.05
1:A:195:PRO:O	1:A:197:ILE:N	1.97	0.97
1:B:67:LYS:NZ	1:B:68:ILE:O	1.99	0.95

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:176:ASN:O	1:D:174:ASN:ND2[3_565]	1.97	0.23
1:A:199:GLU:O	1:B:217:SER:OG[2_665]	2.11	0.09
1:A:176:ASN:O	1:B:174:ASN:ND2[2_665]	2.12	0.08

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	270/341 (79%)	163 (60%)	51 (19%)	56 (21%)	0	2
1	B	277/341 (81%)	170 (61%)	56 (20%)	51 (18%)	0	3
1	C	270/341 (79%)	164 (61%)	49 (18%)	57 (21%)	0	2
1	D	270/341 (79%)	161 (60%)	59 (22%)	50 (18%)	0	3
All	All	1087/1364 (80%)	658 (60%)	215 (20%)	214 (20%)	0	2

5 of 214 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	38	LYS
1	A	75	LEU
1	A	82	GLY
1	A	121	GLN
1	A	122	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	236/288 (82%)	132 (56%)	104 (44%)	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	245/288 (85%)	134 (55%)	111 (45%)	0	0
1	C	239/288 (83%)	141 (59%)	98 (41%)	0	1
1	D	238/288 (83%)	134 (56%)	104 (44%)	0	1
All	All	958/1152 (83%)	541 (56%)	417 (44%)	0	1

5 of 417 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	258	LEU
1	C	101	GLN
1	D	235	THR
1	B	268	ILE
1	B	323	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 ⓘ

There are no ligands in this entry.

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, 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	276/341 (80%)	-0.38	1 (0%) 90 71	7, 60, 178, 245	1 (0%)
1	B	285/341 (83%)	-0.35	0 100 100	7, 63, 166, 285	1 (0%)
1	C	278/341 (81%)	-0.37	0 100 100	27, 98, 200, 277	1 (0%)
1	D	278/341 (81%)	-0.36	1 (0%) 90 71	27, 93, 190, 275	1 (0%)
All	All	1117/1364 (81%)	-0.36	2 (0%) 93 80	7, 81, 185, 285	4 (0%)

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	308	GLU	2.1
1	D	38	LYS	2.1

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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.