wwPDB X-ray Structure Validation Summary Report (i)

Dec 8, 2023 – 03:13 am GMT

PDB ID	:	1V15
Title	:	CRYSTAL STRUCTURE OF THE COLICIN E9, MUTANT HIS103ALA, IN COMPLEX WITH ZN+2 AND DSDNA (RESOLUTION 2.4A)
Authors	:	Mate, M.J.; Kleanthous, C.
Deposited on	:	2004-04-06
Resolution	:	2.40 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:

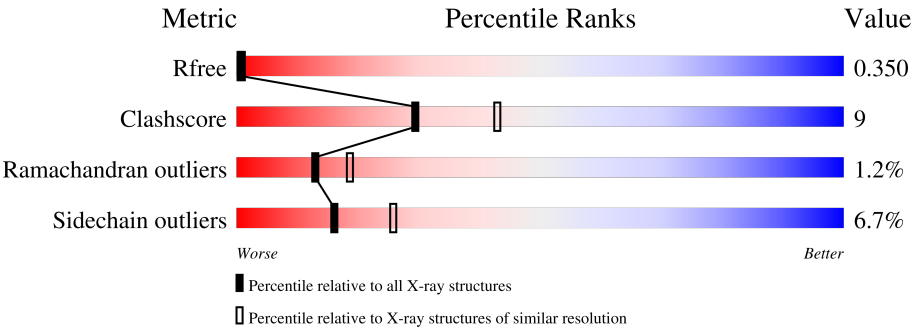
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.40 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.



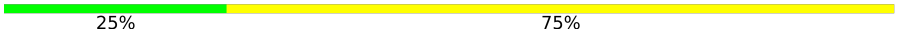
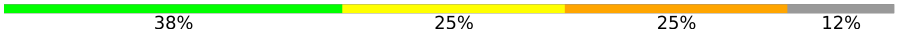
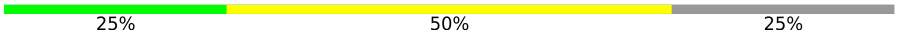
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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%

Mol	Chain	Length	Quality of chain
1	A	134	57% 15% 6% · 21%
1	B	134	81% 14% · ·
1	C	134	66% 26% · ·
1	D	134	72% 23% · ··
2	E	8	62% 12% 12% 12%
2	F	8	25% 25% 50%
2	G	8	62% 12% 25%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	8	
2	I	8	
2	J	8	
2	K	8	
2	L	8	

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5261 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 COLICIN E9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	106	Total	C	N	O	S	15	0	0
			828	518	151	158	1			
1	B	134	Total	C	N	O	S	0	0	0
			1058	658	195	203	2			
1	C	129	Total	C	N	O	S	19	0	0
			1017	634	186	196	1			
1	D	133	Total	C	N	O	S	0	0	0
			1050	653	194	202	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	ALA	HIS	engineered mutation	UNP P09883
B	103	ALA	HIS	engineered mutation	UNP P09883
C	103	ALA	HIS	engineered mutation	UNP P09883
D	103	ALA	HIS	engineered mutation	UNP P09883

- Molecule 2 is a DNA chain called 5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	7	Total	C	N	O	P	0	0	0
			139	67	26	40	6			
2	F	8	Total	C	N	O	P	0	0	0
			161	77	31	46	7			
2	G	6	Total	C	N	O	P	0	0	0
			123	58	23	36	6			
2	H	8	Total	C	N	O	P	0	0	0
			161	77	31	46	7			
2	I	7	Total	C	N	O	P	0	0	0
			139	67	26	40	6			
2	J	8	Total	C	N	O	P	0	0	0
			161	77	31	46	7			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	K	6	Total	C	N	O	P	0	0	0
			123	58	23	36	6			
2	L	8	Total	C	N	O	P	0	0	0
			161	77	31	46	7			

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		
3	B	1	Total	Zn	0	0
			1	1		
3	C	1	Total	Zn	0	0
			1	1		
3	D	1	Total	Zn	0	0
			1	1		
3	E	1	Total	Zn	0	0
			1	1		
3	G	1	Total	Zn	0	0
			1	1		
3	I	1	Total	Zn	0	0
			1	1		
3	K	1	Total	Zn	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	13	Total	O	0	0
			13	13		
4	B	20	Total	O	0	0
			20	20		
4	C	27	Total	O	0	0
			27	27		
4	D	25	Total	O	0	0
			25	25		
4	E	4	Total	O	0	0
			4	4		
4	F	3	Total	O	0	0
			3	3		
4	G	9	Total	O	0	0
			9	9		

Continued on next page...

Continued from previous page...

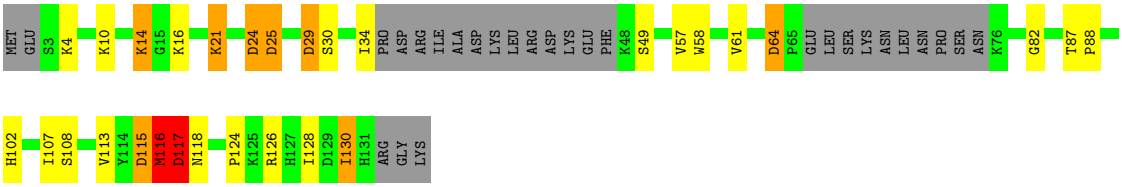
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	H	6	Total O 6 6	0	0
4	I	6	Total O 6 6	0	0
4	J	9	Total O 9 9	0	0
4	K	3	Total O 3 3	0	0
4	L	7	Total O 7 7	0	0

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. 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. 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: COLICIN E9

Chain A:

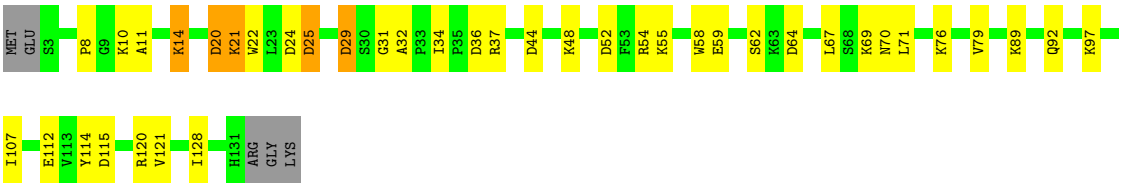


- Molecule 1: COLICIN E9

Chain B:

- Molecule 1: COLICIN E9

Chain C:

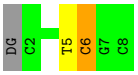


- Molecule 1: COLICIN E9

Chain D:

- Molecule 2: 5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3'

Chain E:

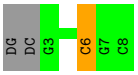


• Molecule 2: 5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3'

Chain F:

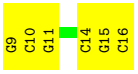
• Molecule 2: 5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3'

Chain G:



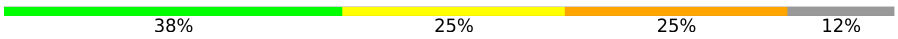
• Molecule 2: 5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3'

Chain H:



• Molecule 2: 5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3'

Chain I:

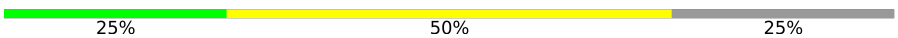


• Molecule 2: 5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3'

Chain J:

• Molecule 2: 5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3'

Chain K:



• Molecule 2: 5'-D(*GP*CP*GP*AP*TP*CP*GP*CP)-3'

Chain L:


G9
C10
G11
A12
T13
C14
G15
C16

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	93.14Å 123.31Å 110.53Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	74.54 – 2.40 74.32 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (74.54-2.40) 99.9 (74.32-2.40)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.73 (at 2.40Å)	Xtriage
Refinement program	REFMAC 5.2.0001	Depositor
R, R_{free}	0.242 , 0.329 0.281 , 0.350	Depositor DCC
R_{free} test set	1286 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å^2)	51.1	Xtriage
Anisotropy	0.513	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 56.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	5261	wwPDB-VP
Average B, all atoms (Å^2)	70.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 47.01 % 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 1.0548e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

[2] Theoretical 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.

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	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	1.94	3/846 (0.4%)	1.57	12/1133 (1.1%)
1	B	0.63	2/1081 (0.2%)	0.86	8/1449 (0.6%)
1	C	0.86	3/1040 (0.3%)	1.64	13/1397 (0.9%)
1	D	0.47	0/1073	0.85	9/1439 (0.6%)
2	E	0.83	0/155	1.62	4/237 (1.7%)
2	F	0.87	0/180	2.27	11/276 (4.0%)
2	G	1.11	0/137	1.94	4/209 (1.9%)
2	H	1.22	0/180	2.63	18/276 (6.5%)
2	I	1.03	0/155	1.99	4/237 (1.7%)
2	J	1.43	2/180 (1.1%)	2.24	11/276 (4.0%)
2	K	0.87	0/137	2.06	6/209 (2.9%)
2	L	0.96	0/180	2.64	14/276 (5.1%)
All	All	1.07	10/5344 (0.2%)	1.58	114/7414 (1.5%)

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	1	0
1	C	1	0
All	All	2	0

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	21	LYS	CA-CB	-46.81	0.51	1.53
1	A	14	LYS	CA-CB	26.01	2.11	1.53
1	C	21	LYS	CA-CB	-20.20	1.09	1.53
1	A	10	LYS	CA-CB	12.69	1.81	1.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	66	GLU	CD-OE2	9.82	1.36	1.25

The worst 5 of 114 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	21	LYS	CB-CA-C	33.32	177.04	110.40
1	A	14	LYS	N-CA-CB	-32.33	52.40	110.60
1	C	21	LYS	N-CA-CB	-24.41	66.66	110.60
1	A	21	LYS	CB-CA-C	22.26	154.92	110.40
1	C	21	LYS	CA-CB-CG	-19.67	70.13	113.40

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	21	LYS	CA
1	C	21	LYS	CA

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	828	0	820	24	0
1	B	1058	0	1058	17	0
1	C	1017	0	1011	19	0
1	D	1050	0	1046	16	0
2	E	139	0	80	1	0
2	F	161	0	91	7	0
2	G	123	0	68	2	0
2	H	161	0	91	1	0
2	I	139	0	80	2	0
2	J	161	0	91	5	0
2	K	123	0	68	0	0
2	L	161	0	91	8	0
3	A	1	0	0	0	0
3	B	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	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	G	1	0	0	0	0
3	I	1	0	0	0	0
3	K	1	0	0	0	0
4	A	13	0	0	0	0
4	B	20	0	0	0	0
4	C	27	0	0	0	0
4	D	25	0	0	0	0
4	E	4	0	0	0	0
4	F	3	0	0	0	0
4	G	9	0	0	0	0
4	H	6	0	0	0	0
4	I	6	0	0	1	0
4	J	9	0	0	0	0
4	K	3	0	0	0	0
4	L	7	0	0	0	0
All	All	5261	0	4595	91	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 9.

The worst 5 of 91 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:A:34:ILE:HD11	1:A:116:MET:CG	1.66	1.23
2:L:15:DG:C5'	2:L:16:DC:H3'	1.81	1.11
2:L:15:DG:H5'	2:L:16:DC:H3'	1.35	1.08
1:A:34:ILE:CD1	1:A:116:MET:HG2	1.89	1.02
1:A:34:ILE:CD1	1:A:116:MET:CG	2.42	0.98

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	100/134 (75%)	84 (84%)	13 (13%)	3 (3%)	4 3
1	B	132/134 (98%)	122 (92%)	9 (7%)	1 (1%)	19 29
1	C	127/134 (95%)	119 (94%)	7 (6%)	1 (1%)	19 29
1	D	131/134 (98%)	123 (94%)	7 (5%)	1 (1%)	19 29
All	All	490/536 (91%)	448 (91%)	36 (7%)	6 (1%)	13 19

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	116	MET
1	D	29	ASP
1	C	29	ASP
1	A	130	ILE
1	A	49	SER

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	90/116 (78%)	81 (90%)	9 (10%)	7 11
1	B	116/116 (100%)	112 (97%)	4 (3%)	37 56
1	C	112/116 (97%)	106 (95%)	6 (5%)	22 36
1	D	115/116 (99%)	105 (91%)	10 (9%)	10 15
All	All	433/464 (93%)	404 (93%)	29 (7%)	16 26

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	21	LYS
1	D	90	ASN
1	C	89	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	78	SER
1	C	34	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	6	ASN
1	C	6	ASN
1	D	90	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 8 ligands modelled in this entry, 8 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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.