



wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 02:51 AM GMT

PDB ID : 2A9G
Title : Structure of C406A arginine deiminase in complex with L-arginine
Authors : Galkin, A.; Lu, X.; Dunaway-Mariano, D.; Herzberg, O.
Deposited on : 2005-07-11
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

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

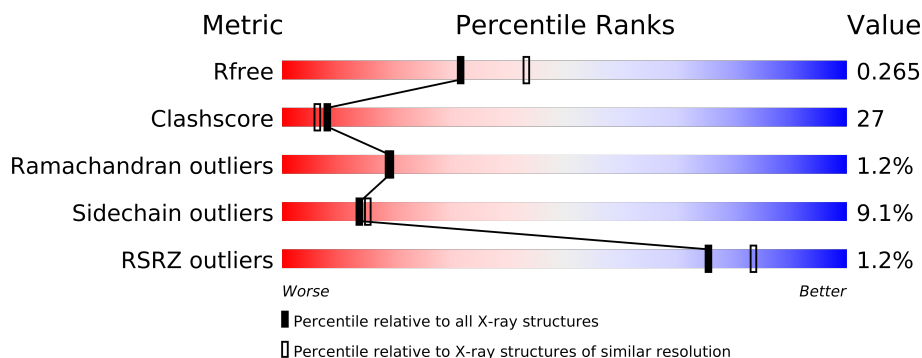
The following versions of software and data (see [references](#)) were used in the production of this report:

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))
R_{free}	66092	2929 (2.30-2.30)
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 ≥ 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	418	
1	B	418	
1	C	418	
1	D	418	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13747 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.

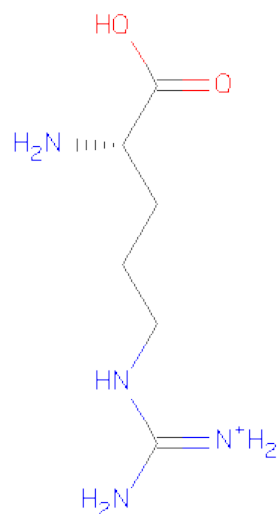
- Molecule 1 is a protein called Arginine deiminase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	405	Total	C	N	O	S	0	0	0
			3163	2001	552	594	16			
1	B	405	Total	C	N	O	S	0	0	0
			3164	2002	552	594	16			
1	C	402	Total	C	N	O	S	0	0	0
			3141	1989	546	590	16			
1	D	406	Total	C	N	O	S	0	0	0
			3167	2003	553	595	16			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	406	ALA	CYS	ENGINEERED	UNP P13981
B	406	ALA	CYS	ENGINEERED	UNP P13981
C	406	ALA	CYS	ENGINEERED	UNP P13981
D	406	ALA	CYS	ENGINEERED	UNP P13981

- Molecule 2 is ARGININE (three-letter code: ARG) (formula: $C_6H_{15}N_4O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			12	6	4	2		
2	B	1	Total	C	N	O	0	0
			12	6	4	2		
2	C	1	Total	C	N	O	0	0
			12	6	4	2		
2	D	1	Total	C	N	O	0	0
			12	6	4	2		

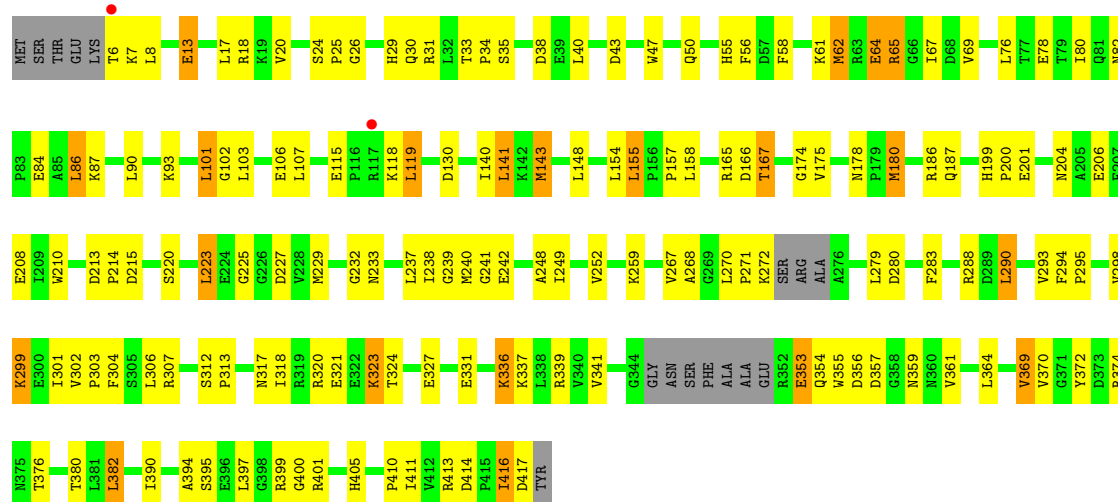
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	223	Total	O	0	0
			223	223		
3	B	277	Total	O	0	0
			277	277		
3	C	309	Total	O	0	0
			309	309		
3	D	255	Total	O	0	0
			255	255		



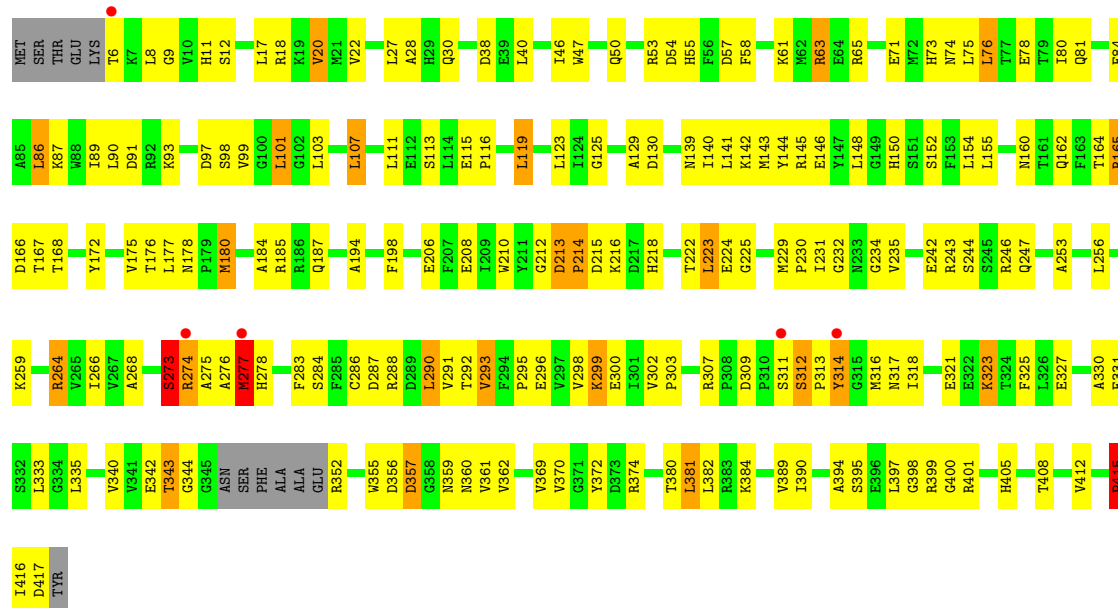
• Molecule 1: Arginine deiminase

Chain C:



• Molecule 1: Arginine deiminase

Chain D:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	91.10Å 120.70Å 151.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.30 19.85 – 2.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (20.00-2.30) 99.2 (19.85-2.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.99 (at 2.30Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.199 , 0.267 0.199 , 0.265	Depositor DCC
R_{free} test set	2001 reflections (2.71%)	DCC
Wilson B-factor (Å ²)	29.3	Xtriage
Anisotropy	0.239	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 51.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	0 of 73863 reflections	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	13747	wwPDB-VP
Average B, all atoms (Å ²)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.75% 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	0.64	0/3230	0.89	2/4381 (0.0%)
1	B	0.68	0/3230	0.94	4/4379 (0.1%)
1	C	0.68	0/3207	0.89	1/4349 (0.0%)
1	D	0.64	0/3234	0.90	2/4386 (0.0%)
All	All	0.66	0/12901	0.90	9/17495 (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	A	0	1

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	273	SER	N-CA-C	7.70	131.79	111.00
1	D	214	PRO	CA-N-CD	-6.46	102.46	111.50
1	A	20	VAL	CB-CA-C	-6.33	99.38	111.40
1	D	415	PRO	CA-N-CD	-5.89	103.25	111.50
1	B	273	SER	CA-C-N	-5.87	104.30	117.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	147	TYR	Sidechain

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	3163	0	3151	203	0
1	B	3164	0	3152	147	0
1	C	3141	0	3127	168	0
1	D	3167	0	3154	184	0
2	A	12	0	12	0	0
2	B	12	0	12	2	0
2	C	12	0	12	0	0
2	D	12	0	12	0	0
3	A	223	0	0	54	0
3	B	277	0	0	31	0
3	C	309	0	0	46	0
3	D	255	0	0	44	0
All	All	13747	0	12632	679	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 27.

The worst 5 of 679 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:14:ALA:O	1:B:366:PRO:HD3	1.34	1.22
1:D:276:ALA:HA	3:D:4622:HOH:O	1.42	1.19
1:B:242:GLU:OE2	1:B:273:SER:HB3	1.42	1.19
1:A:247:GLN:HG2	3:D:4733:HOH:O	1.42	1.16
1:B:73:HIS:HB3	1:B:117:ARG:HH12	1.04	1.14

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	401/418 (96%)	366 (91%)	27 (7%)	8 (2%)	11	8
1	B	399/418 (96%)	367 (92%)	30 (8%)	2 (0%)	38	45
1	C	396/418 (95%)	368 (93%)	26 (7%)	2 (0%)	38	45
1	D	402/418 (96%)	363 (90%)	32 (8%)	7 (2%)	14	11
All	All	1598/1672 (96%)	1464 (92%)	115 (7%)	19 (1%)	19	19

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	242	GLU
1	B	352	ARG
1	A	273	SER
1	A	353	GLU
1	B	351	GLU

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	344/354 (97%)	314 (91%)	30 (9%)	15	17
1	B	344/354 (97%)	312 (91%)	32 (9%)	13	14
1	C	342/354 (97%)	313 (92%)	29 (8%)	15	18
1	D	344/354 (97%)	310 (90%)	34 (10%)	11	13
All	All	1374/1416 (97%)	1249 (91%)	125 (9%)	14	15

5 of 125 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	356	ASP
1	C	101	LEU
1	D	312	SER
1	B	369	VAL
1	C	40	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	187	GLN
1	C	199	HIS
1	D	29	HIS
1	C	55	HIS
1	C	354	GLN

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 ⓘ

4 ligands 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	$\# Z > 2$
2	ARG	A	1500	-	11,11,11	0.53	0	13,13,13	0.96	1 (7%)
2	ARG	B	2500	-	11,11,11	0.50	0	13,13,13	0.79	0
2	ARG	C	3500	-	11,11,11	0.81	0	13,13,13	1.28	1 (7%)
2	ARG	D	4500	-	11,11,11	0.63	0	13,13,13	0.51	0

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
2	ARG	A	1500	-	-	0/11/11/11	0/0/0/0
2	ARG	B	2500	-	-	0/11/11/11	0/0/0/0
2	ARG	C	3500	-	-	0/11/11/11	0/0/0/0
2	ARG	D	4500	-	-	0/11/11/11	0/0/0/0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	3500	ARG	C-CA-N	-3.79	103.08	109.36
2	A	1500	ARG	OXT-C-O	-2.00	119.54	124.07

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, 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	405/418 (96%)	-0.29	6 (1%) 70 78	21, 38, 63, 78	0
1	B	405/418 (96%)	-0.45	6 (1%) 70 78	18, 30, 54, 78	0
1	C	402/418 (96%)	-0.52	2 (0%) 88 94	18, 30, 54, 74	0
1	D	406/418 (97%)	-0.36	5 (1%) 75 83	20, 36, 62, 79	0
All	All	1618/1672 (96%)	-0.40	19 (1%) 75 83	18, 33, 60, 79	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	350	ALA	5.4
1	D	274	ARG	4.8
1	A	275	ALA	4.2
1	B	349	ALA	4.1
1	A	6	THR	4.1

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

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, 95th 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(\AA^2)	Q<0.9
2	ARG	D	4500	12/12	0.10	0.20	16,28,37,38	0
2	ARG	A	1500	12/12	0.09	0.03	16,28,39,46	0
2	ARG	C	3500	12/12	0.09	-0.21	16,21,29,33	0
2	ARG	B	2500	12/12	0.07	-1.38	16,23,26,32	0

6.5 Other polymers ⓘ

There are no such residues in this entry.