



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:23 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 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

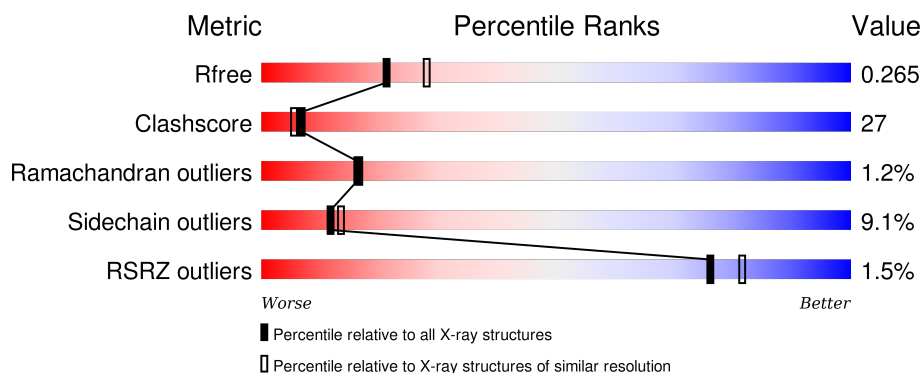
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

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}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (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. 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 $\leq 5\%$. 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
1	A	418	<div> <div>2%</div> <div>49% 42% 5% . .</div> </div>
1	B	418	<div> <div>2%</div> <div>61% 32% . . .</div> </div>
1	C	418	<div> <div>59% 32% 5% .</div> </div>
1	D	418	<div> <div>2%</div> <div>52% 39% 6% . .</div> </div>

2 Entry composition

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

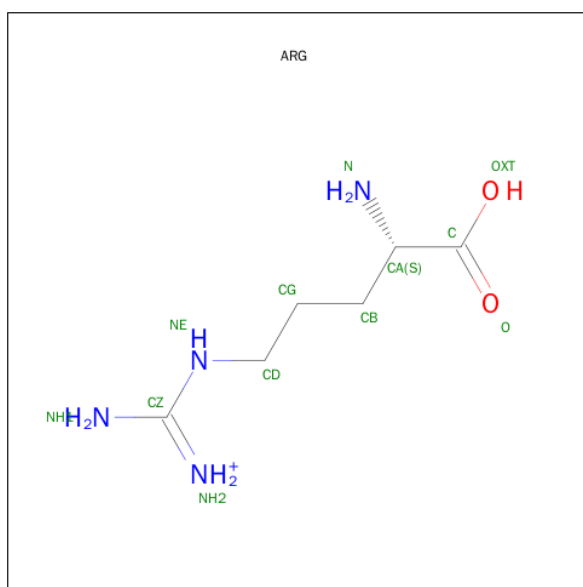
- Molecule 1 is a protein called Arginine deiminase.

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: C₆H₁₅N₄O₂).



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		

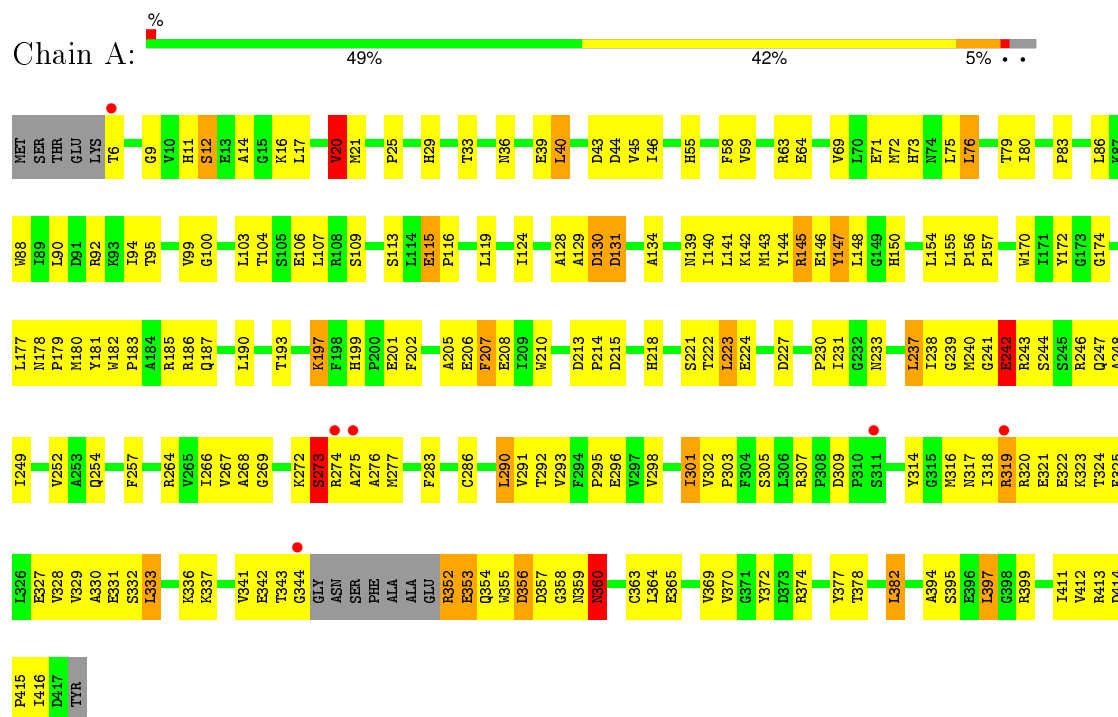
- Molecule 3 is water.

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		

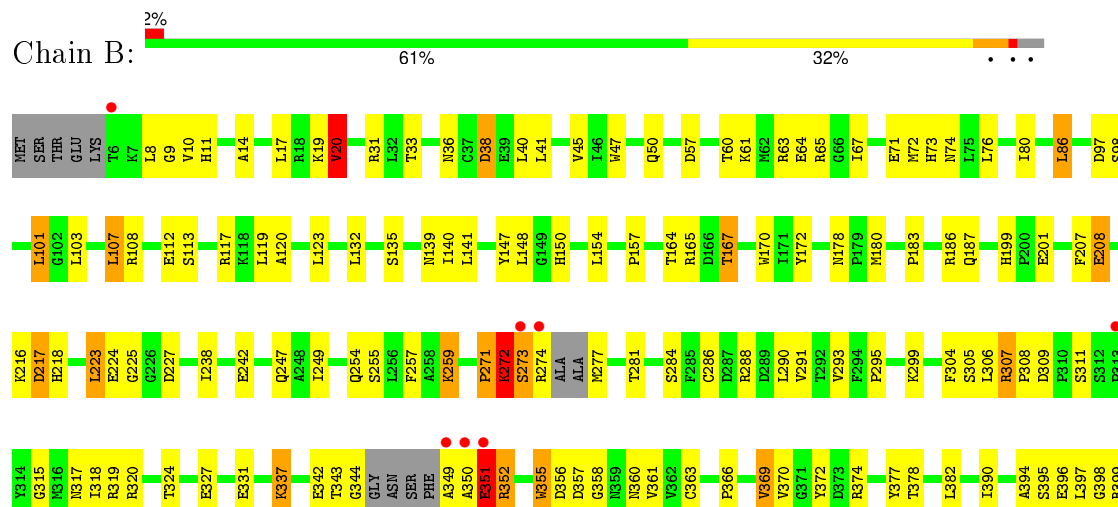
3 Residue-property plots [i](#)

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: Arginine deiminase



• Molecule 1: Arginine deiminase



G400
R401
G402
G403
G404
H405
H406
H407
H408
G409
P410
V411
V412
R413
D414
P415
I416
ASP
TYR

• Molecule 1: Arginine deiminase

Chain C: 59% 32% 5%

MET SER THR LYS L7 K7 L3
 E13
 L17
 R18
 K19
 V20
 S24
 P25
 G26
 E29
 Q30
 R31
 T32
 T33
 P34
 S35
 D38
 E39
 L40
 D43
 V47
 Q50
 H55
 P56
 D57
 F58
 K61
 K62
 R63
 E64
 R65
 G66
 L67
 D68
 V69
 L76
 T77
 E78
 T79
 R80
 Q81
 N82
 P83
 E84
 R85
 L86
 K87
 L90
 R93
 L101
 G102
 L103
 E106
 L107
 E115
 P116
 R117
 L119
 D130
 T140
 L141
 M143
 L148
 L154
 L155
 P157
 L158
 R165
 D166
 T167
 F174
 V175
 M178
 P179
 M180
 R186
 Q187
 H199
 P200
 E201
 N204
 A205
 E206
 F207
 E208
 L209
 W210
 D213
 P214
 D215
 S220
 L223
 E224
 G225
 G226
 D227
 V228
 M229
 G232
 N233
 L237
 I238
 G239
 M240
 G241
 E242
 A248
 I249
 V252
 K259
 V267
 A268
 G269
 L270
 P271
 K272
 SER
 ARG
 ALA
 A276
 L279
 D280
 F283
 R288
 D289
 L290
 V293
 F294
 P295
 V298
 K299
 E300
 I301
 V302
 P303
 F304
 S305
 L306
 R307
 S312
 P313
 R317
 R318
 R319
 G400
 R401
 E321
 E322
 K323
 T324
 E327
 E331
 K336
 K337
 L338
 R339
 V340
 V341
 G344
 GLY
 ASN
 SER
 PHE
 ALA
 ALA
 GLU
 R352
 E353
 Q354
 K355
 D356
 D357
 G358
 R359
 K360
 V361
 L364
 V369
 V370
 G371
 I372
 I373
 R374
 N375
 T376
 T380
 L381
 L382
 I390
 A394
 S395
 E396
 L397
 G398
 R399
 R401
 H405
 P410
 I411
 V412
 R413
 D414
 P415
 I416
 D417
 TYR

• Molecule 1: Arginine deiminase

Chain D: 2% 52% 39% 6%

MET SER THR LYS L7 K7 L8
 G9
 H11
 S12
 L17
 R18
 K19
 V20
 M21
 V22
 L27
 A28
 H29
 Q30
 D38
 E39
 L40
 I46
 W47
 Q50
 R53
 D54
 H55
 F56
 D57
 F58
 K61
 K62
 R63
 E64
 R65
 E71
 W72
 H73
 N74
 L75
 L76
 T77
 E78
 T79
 I80
 Q81
 E84
 A85
 L86
 K87
 W88
 L89
 L90
 D91
 R92
 R93
 D97
 H98
 S98
 M100
 V99
 G100
 L101
 G102
 L103
 L107
 L111
 E112
 S113
 H29
 L114
 P116
 L119
 L123
 D213
 P214
 T124
 G125
 A129
 D130
 N139
 T140
 L141
 K142
 M143
 Y144
 R145
 E146
 T147
 L148
 G149
 H150
 S151
 S152
 F153
 L154
 L155
 N160
 T161
 Q162
 T164
 R165
 D166
 T167
 T168
 Y172
 V175
 T176
 L177
 M178
 M180
 A184
 R185
 L186
 Q187
 A194
 F198
 E206
 F207
 E208
 E209
 W210
 Y211
 G212
 D213
 P214
 D215
 K216
 D217
 H218
 T222
 L223
 E224
 G225
 M229
 P230
 I231
 G232
 G234
 V235
 E242
 R243
 S244
 S245
 R246
 Q247
 A253
 L256
 K259
 R264
 V265
 I266
 A268
 S273
 A275
 A276
 W277
 H278
 F283
 S284
 F285
 C286
 D287
 R288
 D289
 L290
 V291
 T292
 V293
 F294
 P295
 E296
 V297
 V298
 K299
 E300
 I301
 V302
 P303
 R307
 P308
 D309
 P310
 S311
 S312
 P313
 Y314
 G315
 M316
 N317
 I318
 E321
 E322
 K323
 T324
 F325
 L326
 E327
 A330
 E331
 S332
 L333
 G334
 L335
 V340
 V341
 E342
 T343
 G344
 G345
 ASN
 SER
 PHE
 ALA
 ALA
 GLU
 R352
 W355
 D356
 D357
 G358
 N359
 N360
 V361
 V362
 V369
 V370
 G371
 Y372
 B373
 R374
 T380
 L381
 L382
 R383
 K384
 V389
 I390
 A394
 S395
 E396
 L397
 G398
 R399
 G400
 R401
 H405
 T406
 V412
 P415
 I416
 D417
 TYR

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)	0.31 , 56.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.41$, $\langle L^2 \rangle = 0.24$	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.

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

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 Too-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 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	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

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 27.

The worst 5 of 679 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: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%)	9	7
1	B	399/418 (96%)	367 (92%)	30 (8%)	2 (0%)	34	41
1	C	396/418 (95%)	368 (93%)	26 (7%)	2 (0%)	34	41
1	D	402/418 (96%)	363 (90%)	32 (8%)	7 (2%)	11	10
All	All	1598/1672 (96%)	1464 (92%)	115 (7%)	19 (1%)	16	16

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%)	13	15
1	B	344/354 (97%)	312 (91%)	32 (9%)	11	13
1	C	342/354 (97%)	313 (92%)	29 (8%)	13	16
1	D	344/354 (97%)	310 (90%)	34 (10%)	10	11
All	All	1374/1416 (97%)	1249 (91%)	125 (9%)	12	13

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

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
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 [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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

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	-	5,11,11	0.52	0	3,13,13	0.69	0
2	ARG	B	2500	-	5,11,11	0.30	0	3,13,13	0.79	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	ARG	C	3500	-	5,11,11	0.46	0	3,13,13	0.60	0
2	ARG	D	4500	-	5,11,11	0.50	0	3,13,13	0.05	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	-	-	0/5/11/11	0/0/0/0
2	ARG	B	2500	-	-	0/5/11/11	0/0/0/0
2	ARG	C	3500	-	-	0/5/11/11	0/0/0/0
2	ARG	D	4500	-	-	0/5/11/11	0/0/0/0

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.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2500	ARG	2	0

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, 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%)	-0.25	6 (1%) 76 81	21, 38, 63, 78	0
1	B	405/418 (96%)	-0.42	7 (1%) 73 79	18, 30, 54, 78	0
1	C	402/418 (96%)	-0.49	2 (0%) 91 94	18, 30, 54, 74	0
1	D	406/418 (97%)	-0.33	9 (2%) 65 73	20, 36, 62, 79	0
All	All	1618/1672 (96%)	-0.37	24 (1%) 76 81	18, 33, 60, 79	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	350	ALA	5.2
1	A	6	THR	4.9
1	D	274	ARG	4.7
1	B	349	ALA	4.2
1	D	311	SER	4.2

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 carbohydrates 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. 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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	ARG	D	4500	12/12	0.97	0.10	0.55	16,28,37,38	0
2	ARG	A	1500	12/12	0.98	0.10	0.09	16,28,39,46	0
2	ARG	C	3500	12/12	0.97	0.09	-0.18	16,21,29,33	0
2	ARG	B	2500	12/12	0.98	0.07	-1.20	16,23,26,32	0

6.5 Other polymers [i](#)

There are no such residues in this entry.