



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 01:42 am GMT

PDB ID : 3I4J
Title : Crystal structure of Aminotransferase, class III from *Deinococcus radiodurans*
Authors : Fedorov, A.A.; Fedorov, E.V.; Toro, R.; Sauder, J.M.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2009-07-01
Resolution : 1.70 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
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) : recalc28949

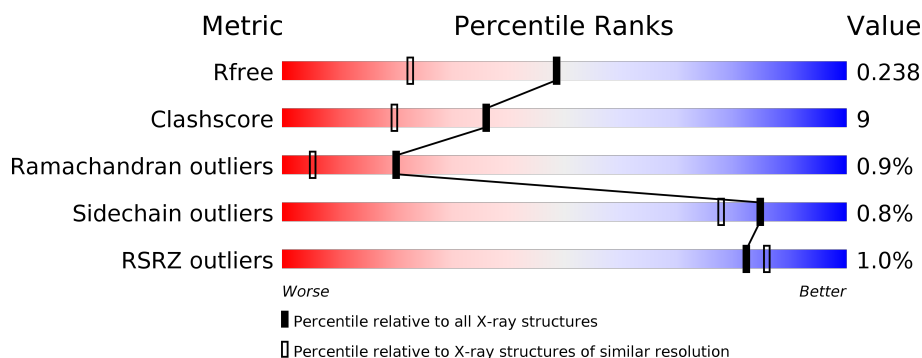
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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}	100719	3453 (1.70-1.70)
Clashscore	112137	3876 (1.70-1.70)
Ramachandran outliers	110173	3815 (1.70-1.70)
Sidechain outliers	110143	3815 (1.70-1.70)
RSRZ outliers	101464	3491 (1.70-1.70)

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	430	<div> <div>%</div> <div> <div></div> <div>72%</div> <div>16%</div> <div>12%</div> </div> </div>
1	B	430	<div> <div></div> <div>73%</div> <div>16%</div> <div>10%</div> </div>
1	C	430	<div> <div>%</div> <div> <div></div> <div>72%</div> <div>17%</div> <div>10%</div> </div> </div>
1	D	430	<div> <div>%</div> <div> <div></div> <div>70%</div> <div>18%</div> <div>12%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11873 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 Aminotransferase, class III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	379	Total	C	N	O	S	0	0	0
			2804	1779	498	517	10			
1	B	386	Total	C	N	O	S	0	0	0
			2875	1827	508	529	11			
1	C	386	Total	C	N	O	S	0	0	0
			2875	1827	508	529	11			
1	D	379	Total	C	N	O	S	0	0	0
			2804	1779	498	517	10			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	416	GLY	GLU	engineered	UNP Q9RZ32
B	416	GLY	GLU	engineered	UNP Q9RZ32
C	416	GLY	GLU	engineered	UNP Q9RZ32
D	416	GLY	GLU	engineered	UNP Q9RZ32

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		

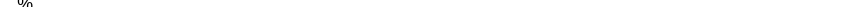
- Molecule 3 is water.

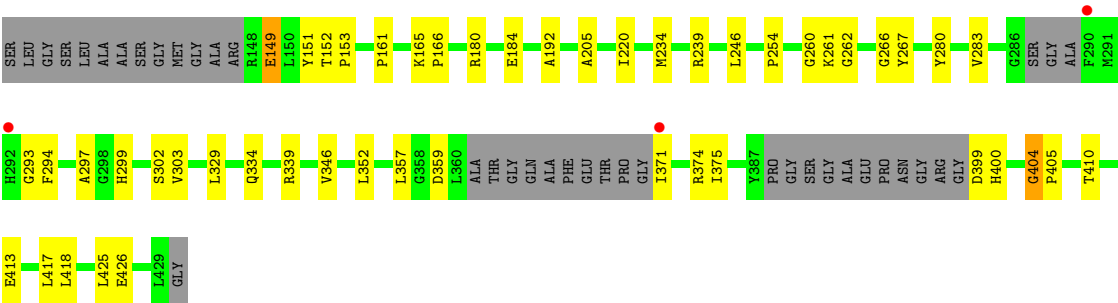
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	104	Total	O	0	0
			104	104		
3	B	143	Total	O	0	0
			143	143		
3	C	139	Total	O	0	0
			139	139		
3	D	109	Total	O	0	0
			109	109		

These plots are drawn for all protein, RNA and DNA 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 ($\text{RSRZ} > 2$). 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.

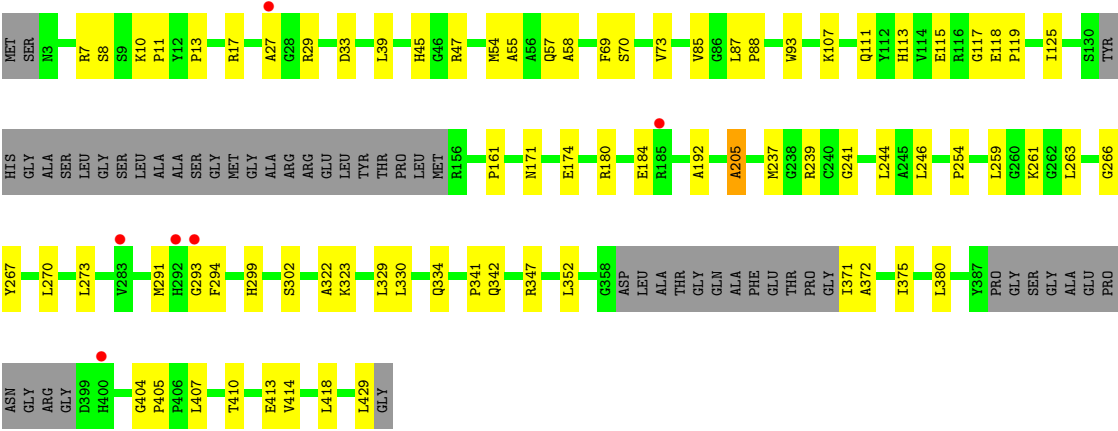
- [illegible]

- Chain B:
-
- 73% 16% 10%
- MET N3 V4 F5 S8 S9 K10 P13 D33 L39 H45 G46 A47 A48 E52 R53 M54 A55 A56 Q57 R60 F69 S70 V73 L74 R91 F92 W93 A94 K107 Q111 Y112 H113 W114 E115 R127 V128 P129 S130 TYR HIS GLY ALA SER LEU GLY
- SER LEU ALA ALA SER GLY MET GLY ARG R148 Y151 P152 P153 K165 P166 E177 R180 A192 S203 D204 A205 I220 M234 R239 L246 P254 G260 K261 G262 G266 Y267 Y269 V283 G286 SER GLY ALA F290 M291 H292 G293 F294
- A297 G298 H299 S302 V303 R315 L352 L360 ALA THR GLY GLN ALA PHE GLU THR PRO PRO GLY I371 I375 L380 Y387 PRO GLY SER GLY ALA GLU PRO ASN ARG GLY G399 H400 L401 G404 P405 T410 E413 L417 L421 L425 V428 L429 GLY

- Chain C: 



• Molecule 1: Aminotransferase, class III



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	64.75Å 193.12Å 64.77Å 90.00° 105.41° 90.00°	Depositor
Resolution (Å)	24.89 – 1.70 32.85 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.1 (24.89-1.70) 98.7 (32.85-1.70)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.47 (at 1.70Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.218 , 0.239 0.218 , 0.238	Depositor DCC
R_{free} test set	8346 reflections (5.03%)	DCC
Wilson B-factor (Å ²)	23.7	Xtriage
Anisotropy	0.311	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 29.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.479 for l,-k,h	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	11873	wwPDB-VP
Average B, all atoms (Å ²)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.04% 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, 0.333 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: SO4

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	0.31	0/2860	0.59	1/3877 (0.0%)
1	B	0.31	0/2932	0.60	1/3974 (0.0%)
1	C	0.31	0/2932	0.61	1/3974 (0.0%)
1	D	0.30	0/2860	0.59	1/3877 (0.0%)
All	All	0.31	0/11584	0.60	4/15702 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	70	SER	N-CA-C	-5.30	96.70	111.00
1	D	70	SER	N-CA-C	-5.22	96.90	111.00
1	B	70	SER	N-CA-C	-5.15	97.09	111.00
1	C	70	SER	N-CA-C	-5.12	97.17	111.00

There are no chirality outliers.

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	2804	0	2809	59	0
1	B	2875	0	2883	53	0
1	C	2875	0	2883	53	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	2804	0	2809	59	0
2	A	5	0	0	0	0
2	B	5	0	0	0	0
2	C	5	0	0	0	0
2	D	5	0	0	0	0
3	A	104	0	0	4	0
3	B	143	0	0	4	0
3	C	139	0	0	4	0
3	D	109	0	0	4	0
All	All	11873	0	11384	207	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 207 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:180:ARG:HH21	1:A:220:ILE:HD11	1.32	0.93
1:A:57:GLN:OE1	1:A:73:VAL:HG11	1.73	0.88
1:B:57:GLN:OE1	1:B:73:VAL:HG11	1.74	0.86
1:A:342:GLN:H	1:A:342:GLN:HE21	1.26	0.84
1:D:57:GLN:OE1	1:D:73:VAL:HG11	1.76	0.83

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	371/430 (86%)	357 (96%)	11 (3%)	3 (1%)	22 7
1	B	376/430 (87%)	361 (96%)	11 (3%)	4 (1%)	17 4

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	376/430 (87%)	361 (96%)	11 (3%)	4 (1%)	17	4
1	D	371/430 (86%)	355 (96%)	13 (4%)	3 (1%)	22	7
All	All	1494/1720 (87%)	1434 (96%)	46 (3%)	14 (1%)	20	5

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	261	LYS
1	A	261	LYS
1	B	261	LYS
1	B	293	GLY
1	C	293	GLY

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	275/308 (89%)	273 (99%)	2 (1%)	87	81
1	B	284/308 (92%)	281 (99%)	3 (1%)	78	68
1	C	284/308 (92%)	280 (99%)	4 (1%)	71	58
1	D	275/308 (89%)	275 (100%)	0	100	100
All	All	1118/1232 (91%)	1109 (99%)	9 (1%)	85	78

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	315	ARG
1	C	359	ASP
1	C	129	PRO
1	B	5	PHE
1	C	5	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	337	GLN
1	C	113	HIS
1	D	113	HIS
1	B	299	HIS
1	D	299	HIS

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	SO4	A	431	-	4,4,4	0.31	0	6,6,6	0.21	0
2	SO4	B	431	-	4,4,4	0.33	0	6,6,6	0.16	0
2	SO4	C	431	-	4,4,4	0.25	0	6,6,6	0.14	0
2	SO4	D	431	-	4,4,4	0.24	0	6,6,6	0.19	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	SO4	A	431	-	-	0/0/0/0	0/0/0/0
2	SO4	B	431	-	-	0/0/0/0	0/0/0/0
2	SO4	C	431	-	-	0/0/0/0	0/0/0/0
2	SO4	D	431	-	-	0/0/0/0	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.

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, 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	379/430 (88%)	-0.29	4 (1%) 80 84	17, 25, 40, 50	0
1	B	386/430 (89%)	-0.27	2 (0%) 90 93	17, 25, 41, 59	0
1	C	386/430 (89%)	-0.28	3 (0%) 86 88	17, 24, 40, 59	0
1	D	379/430 (88%)	-0.27	6 (1%) 72 77	17, 25, 41, 48	0
All	All	1530/1720 (88%)	-0.27	15 (0%) 82 86	17, 25, 41, 59	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	292	HIS	5.0
1	B	292	HIS	4.4
1	D	27	ALA	3.3
1	A	185	ARG	3.2
1	A	292	HIS	3.0

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	SO4	C	431	5/5	0.98	0.10	1.50	39,41,42,43	0
2	SO4	B	431	5/5	0.93	0.11	0.91	39,39,41,41	0
2	SO4	A	431	5/5	0.93	0.11	0.78	40,40,41,42	0
2	SO4	D	431	5/5	0.96	0.12	0.73	39,42,43,44	0

6.5 Other polymers [i](#)

There are no such residues in this entry.