



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:37 AM GMT

PDB ID : 2HXG
Title : Crystal Structure of Mn²⁺ bound ECAI
Authors : Manjasetty, B.A.; Burley, S.K.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2006-08-03
Resolution : 2.80 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

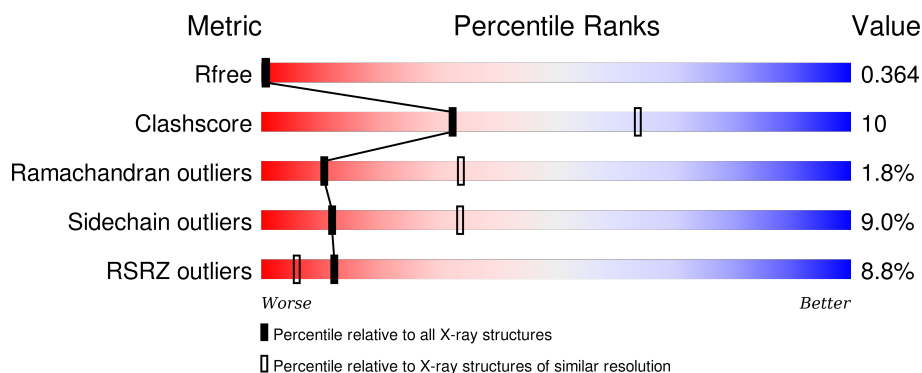
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	500	<div> <div>3%</div> <div>77%</div> <div>19%</div> <div>.</div> </div>
1	B	500	<div> <div>4%</div> <div>75%</div> <div>21%</div> <div>.</div> </div>
1	C	500	<div> <div>19%</div> <div>73%</div> <div>23%</div> <div>.</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 11565 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 L-arabinose isomerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	498	Total	C	N	O	S	0	0	0
			3829	2431	669	705	24			
1	B	498	Total	C	N	O	S	0	0	0
			3895	2474	680	717	24			
1	C	498	Total	C	N	O	S	0	0	0
			3735	2353	650	708	24			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	72	PRO	ARG	SEE REMARK 999	UNP Q8FL89
B	72	PRO	ARG	SEE REMARK 999	UNP Q8FL89
C	72	PRO	ARG	SEE REMARK 999	UNP Q8FL89

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Mn	0	0
			1	1		
2	A	1	Total	Mn	0	0
			1	1		
2	C	1	Total	Mn	0	0
			1	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	41	Total	O	0	0
			41	41		
3	B	37	Total	O	0	0
			37	37		

Continued on next page...

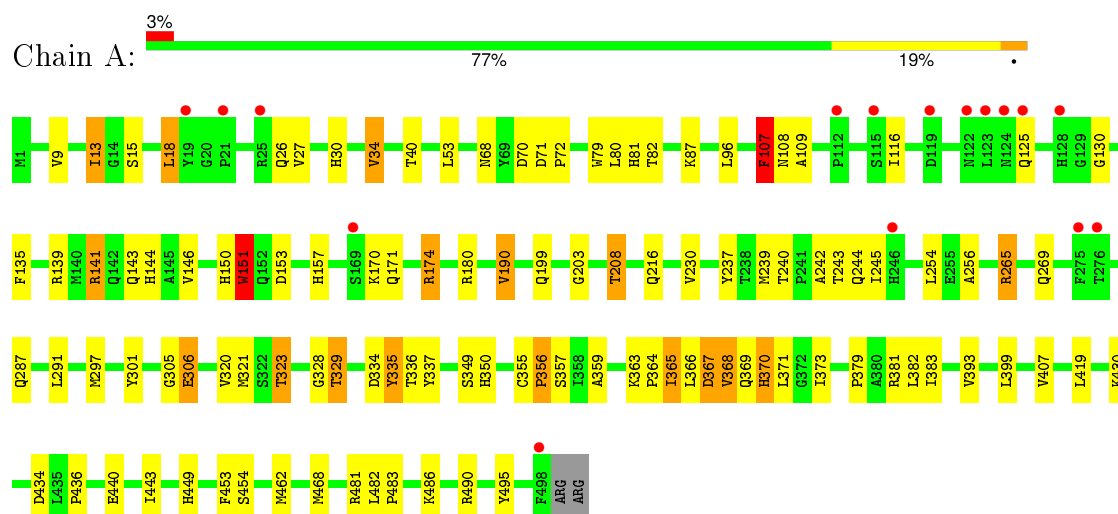
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	C	25	Total	O	0	0
			25	25		

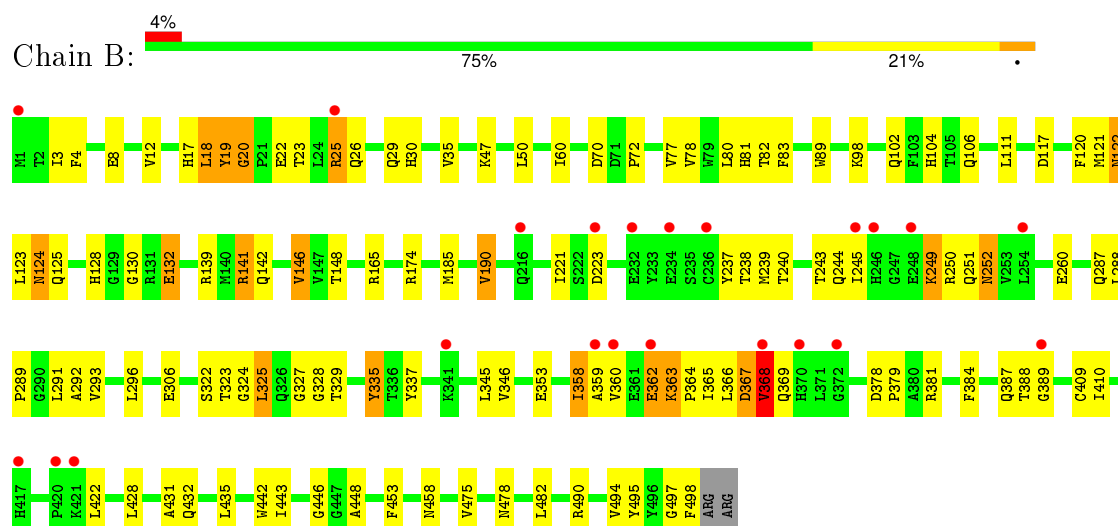
3 Residue-property plots

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 ($\text{RSRZ} > 2$). 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.

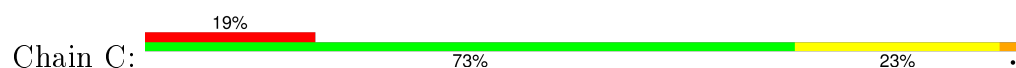
- Molecule 1: L-arabinose isomerase



- Molecule 1: L-arabinose isomerase



- Molecule 1: L-arabinose isomerase



ARG	G374 G375 G376 D377 D378	P289 G290 L291 Q294 Q299 G300 Y301 G302 F303 A304 G305 E306 K310 R316 V320 T323 G324 L325 Q326 G327 G328 T329 S330 D334 F339 N343 D344 L345 S349 H350 M351 L352 E353 V354 C355 I358 A359 V360 E361 E362 K363 P364 V368 Q369 H370 L371 G372 I373	P289 G290 L291 Q294 Q299 G300 Y301 G302 F303 A304 G305 E306 K310 R316 V320 T323 G324 L325 Q326 G327 G328 T329 S330 D334 F339 N343 D344 L345 S349 H350 M351 L352 E353 V354 C355 I358 A359 V360 E361 E362 K363 P364 V368 Q369 H370 L371 G372 I373	L221 S222 D223 V226 N227 A228 L229 Y233 N234 E234 S235 C236 Y237 T238 M239 T240 P241 A242 T243 Q244 I245 H246 G247 E248 K249 R250 Q251 N252 V253 L254 E255 A256 A257 L261 G262 M263 K264 R265 E268 Q269 G270 G271 F272 H273 A274 F275 T276 T277 T278 F279 E280 D281 L282 K286 Q287 L288	L221 S222 D223 V226 N227 A228 L229 Y233 N234 E234 S235 C236 Y237 T238 M239 T240 P241 A242 T243 Q244 I245 H246 G247 E248 K249 R250 Q251 N252 V253 L254 E255 A256 A257 L261 G262 M263 K264 R265 E268 Q269 G270 G271 F272 H273 A274 F275 T276 T277 T278 F279 E280 D281 L282 K286 Q287 L288	L111 F112 W113 D114 S115 I116 D117 M121 N122 L123 N124 Q125 G130 R131 E132 F135 Q143 H144 H150 R151 Q152 D153 K154 H157 V168 K177 V178 C179 R180 M184 M185 R186 E187 V188 T191 D192 K195 V206 N207 T208 W209 D213 Q106 Q216 N219 S220	L111 F112 W113 D114 S115 I116 D117 M121 N122 L123 N124 Q125 G130 R131 E132 F135 Q143 H144 H150 R151 Q152 D153 K154 H157 V168 K177 V178 C179 R180 M184 M185 R186 E187 V188 T191 D192 K195 V206 N207 T208 W209 D213 Q106 Q216 N219 S220	V4 T2 F4 Y7 F11 V12 I13 G14 H17 E22 R25 T28 A31 V35 T40 E41 A42 K43 L44 P45 G46 K47 D70 D71 P72 V77 V78 W79 L80 H81 T82 F83 A86 K98 P99 L100 L101 Q102 F103 H104 T105 Q106 F107 N108 A109 A110	V4 T2 F4 Y7 F11 V12 I13 G14 H17 E22 R25 T28 A31 V35 T40 E41 A42 K43 L44 P45 G46 K47 D70 D71 P72 V77 V78 W79 L80 H81 T82 F83 A86 K98 P99 L100 L101 Q102 F103 H104 T105 Q106 F107 N108 A109 A110
-----	--------------------------------------	--	--	--	--	--	--	---	---

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	116.87Å 116.87Å 215.08Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.80 19.89 – 2.80	Depositor EDS
% Data completeness (in resolution range)	92.3 (20.00-2.80) 92.3 (19.89-2.80)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.13 (at 2.79Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.229 , 0.288 0.318 , 0.364	Depositor DCC
R_{free} test set	1988 reflections (5.34%)	DCC
Wilson B-factor (Å ²)	53.8	Xtriage
Anisotropy	0.683	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 38.3	EDS
Estimated twinning fraction	0.009 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.36$	Xtriage
Outliers	0 of 39226 reflections	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	11565	wwPDB-VP
Average B, all atoms (Å ²)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.26% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MN

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	0.46	0/3922	0.63	0/5325
1	B	0.47	0/3991	0.63	0/5424
1	C	0.42	0/3826	0.59	0/5203
All	All	0.45	0/11739	0.62	0/15952

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
1	B	0	3
1	C	0	3
All	All	0	7

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	107	PHE	Peptide
1	B	327	GLY	Peptide
1	B	362	GLU	Peptide
1	B	368	VAL	Peptide
1	C	46	CYS	Peptide

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	3829	0	3686	88	0
1	B	3895	0	3783	87	0
1	C	3735	0	3440	87	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	41	0	0	3	0
3	B	37	0	0	2	0
3	C	25	0	0	0	0
All	All	11565	0	10909	235	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 10.

The worst 5 of 235 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:C:180:ARG:HD2	1:C:195:LYS:HZ2	1.32	0.94
1:B:360:VAL:HG22	1:B:387:GLN:HG2	1.52	0.90
1:C:109:ALA:HB2	1:C:150:HIS:CG	2.12	0.85
1:A:364:PRO:HD2	1:A:383:ILE:O	1.80	0.79
1:C:180:ARG:HD2	1:C:195:LYS:NZ	1.97	0.79

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	496/500 (99%)	453 (91%)	36 (7%)	7 (1%)	14	42
1	B	496/500 (99%)	451 (91%)	41 (8%)	4 (1%)	24	58
1	C	496/500 (99%)	418 (84%)	62 (12%)	16 (3%)	5	17
All	All	1488/1500 (99%)	1322 (89%)	139 (9%)	27 (2%)	11	34

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	151	TRP
1	A	368	VAL
1	A	370	HIS
1	B	245	ILE
1	C	287	GLN

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	394/419 (94%)	361 (92%)	33 (8%)	14	37
1	B	410/419 (98%)	372 (91%)	38 (9%)	11	32
1	C	369/419 (88%)	335 (91%)	34 (9%)	11	32
All	All	1173/1257 (93%)	1068 (91%)	105 (9%)	12	34

5 of 105 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	139	ARG
1	B	251	GLN
1	C	422	LEU
1	B	141	ARG
1	B	221	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	252	ASN
1	B	298	GLN
1	C	273	HIS
1	B	269	GLN
1	B	283	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](#)

Of 3 ligands modelled in this entry, 3 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 [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	498/500 (99%)	0.41	16 (3%)	51 39	42, 49, 55, 65	0
1	B	498/500 (99%)	0.44	22 (4%)	38 26	39, 47, 55, 71	0
1	C	498/500 (99%)	1.00	93 (18%)	2 1	36, 46, 53, 64	0
All	All	1494/1500 (99%)	0.62	131 (8%)	12 6	36, 47, 54, 71	0

The worst 5 of 131 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	246	HIS	7.5
1	C	251	GLN	6.7
1	B	360	VAL	6.6
1	C	416	PRO	6.2
1	C	237	TYR	5.8

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	MN	C	501	1/1	0.84	0.14	-2.63	71,71,71,71	0
2	MN	A	501	1/1	0.97	0.10	-2.70	35,35,35,35	0
2	MN	B	501	1/1	0.93	0.06	-3.40	50,50,50,50	0

6.5 Other polymers [i](#)

There are no such residues in this entry.