wwPDB X-ray Structure Validation Summary Report (i)

Mar 9, 2024 – 05:25 PM EST

PDB ID	:	3GEB
Title	:	Crystal Structure of edeya2
Authors	:	Kim, S.J.
Deposited on	:	2009-02-25
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 (i)) were used in the production of this report:

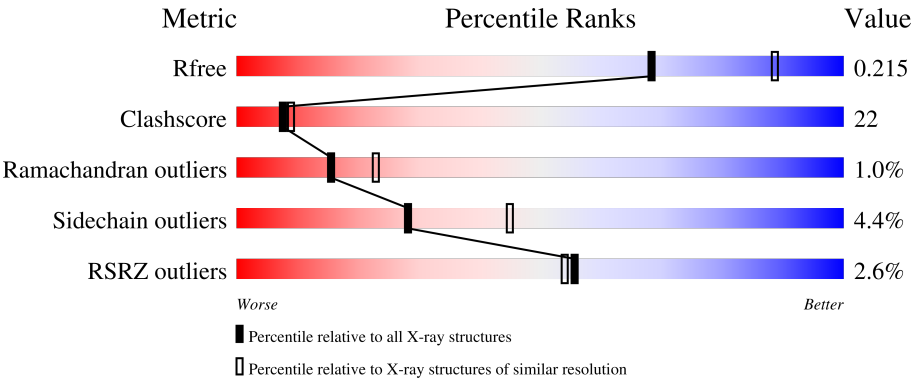
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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)
RSRZ outliers	127900	3811 (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% 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	274	59% 32% • 6%
1	B	274	4% 46% 43% • 8%
1	C	274	3% 51% 41% • 6%
1	D	274	2% 58% 34% • 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8355 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 Eyes absent homolog 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	258	Total	C	N	O	S	0	0	0
			2078	1322	355	389	12			
1	B	253	Total	C	N	O	S	0	0	0
			2049	1305	350	383	11			
1	C	258	Total	C	N	O	S	0	0	0
			2078	1322	355	389	12			
1	D	256	Total	C	N	O	S	0	0	0
			2064	1314	353	386	11			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	265	SER	-	expression tag	UNP O00167
A	266	HIS	-	expression tag	UNP O00167
A	267	MET	-	expression tag	UNP O00167
B	265	SER	-	expression tag	UNP O00167
B	266	HIS	-	expression tag	UNP O00167
B	267	MET	-	expression tag	UNP O00167
C	265	SER	-	expression tag	UNP O00167
C	266	HIS	-	expression tag	UNP O00167
C	267	MET	-	expression tag	UNP O00167
D	265	SER	-	expression tag	UNP O00167
D	266	HIS	-	expression tag	UNP O00167
D	267	MET	-	expression tag	UNP O00167

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	1	Total Mg 1 1	0	0
2	D	1	Total Mg 1 1	0	0

- Molecule 3 is water.

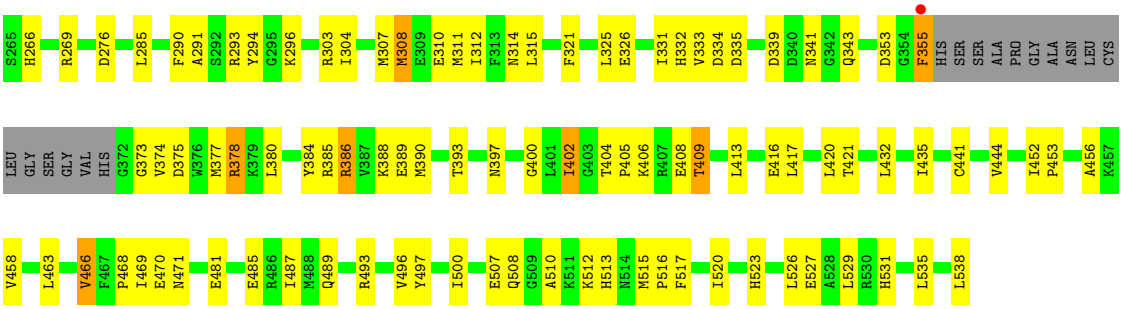
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	28	Total O 28 28	0	0
3	B	12	Total O 12 12	0	0
3	C	13	Total O 13 13	0	0
3	D	29	Total O 29 29	0	0

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 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.

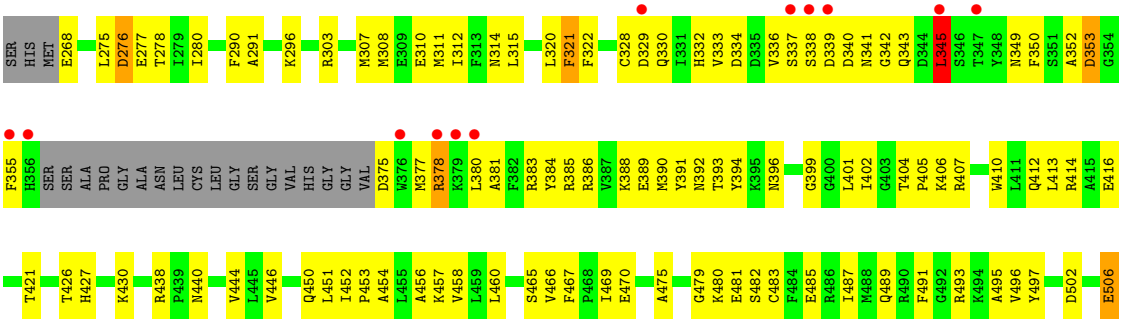
• Molecule 1: Eyes absent homolog 2

Chain A:



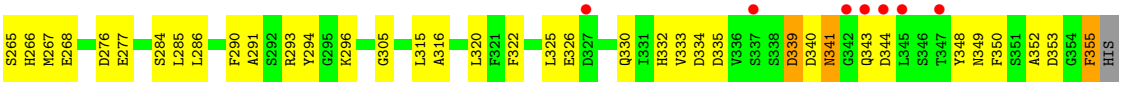
• Molecule 1: Eyes absent homolog 2

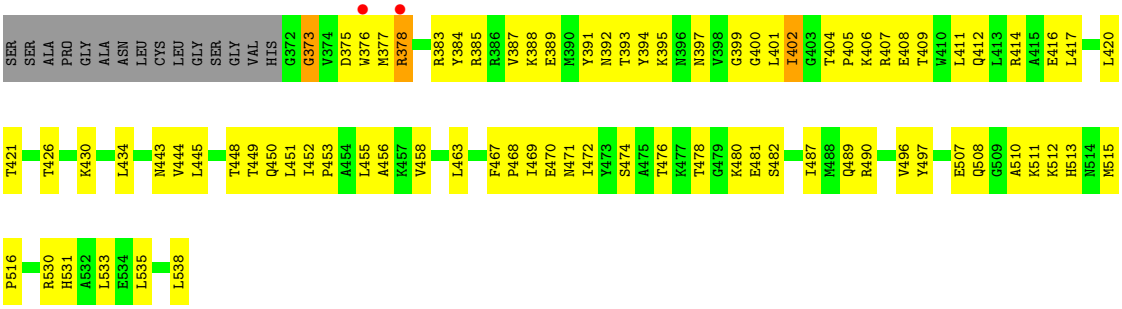
Chain B:



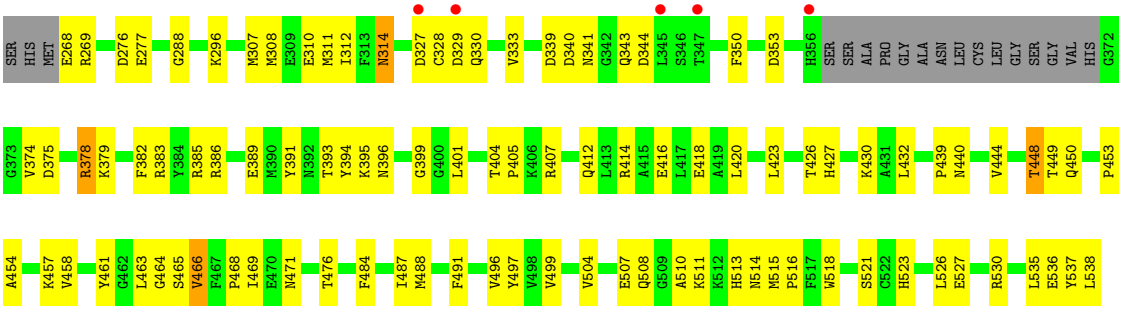
• Molecule 1: Eyes absent homolog 2

Chain C:





• Molecule 1: Eyes absent homolog 2



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, α, β, γ	183.61Å 183.61Å 120.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.40 67.79 – 2.40	Depositor EDS
% Data completeness (in resolution range)	98.7 (50.00-2.40) 99.2 (67.79-2.40)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	3.36 (at 2.40Å)	Xtriage
Refinement program	CNS 0.9	Depositor
R, R_{free}	0.171 , 0.218 0.171 , 0.215	Depositor DCC
R_{free} test set	3694 reflections (4.78%)	wwPDB-VP
Wilson B-factor (Å^2)	37.6	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 28.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.40$, $< L^2 > = 0.23$	Xtriage
Estimated twinning fraction	0.356 for -k,-h,-l	Xtriage
Reported twinning fraction	0.650 for h,k,l 0.350 for h,-k,-l	Depositor
Outliers	0 of 77276 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8355	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.57% of the height of the origin peak. No significant pseudotranslation is detected.*

[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

5.1 Standard geometry

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

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	0.43	0/2121	0.63	0/2867
1	B	0.39	0/2092	0.59	0/2829
1	C	0.40	0/2121	0.59	0/2867
1	D	0.42	0/2107	0.61	0/2849
All	All	0.41	0/8441	0.61	0/11412

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2078	0	2041	71	0
1	B	2049	0	2012	119	0
1	C	2078	0	2041	100	0
1	D	2064	0	2027	79	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	28	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	12	0	0	0	0
3	C	13	0	0	0	0
3	D	29	0	0	2	0
All	All	8355	0	8121	360	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 22.

The worst 5 of 360 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:402:ILE:HG23	1:A:406:LYS:HB2	1.53	0.87
1:B:509:GLY:HA2	1:B:512:LYS:HD2	1.57	0.87
1:B:516:PRO:HB3	1:C:538:LEU:HD23	1.57	0.85
1:D:427:HIS:HD2	1:D:523:HIS:HE1	1.20	0.85
1:D:427:HIS:HD2	1:D:523:HIS:CE1	1.95	0.84

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	254/274 (93%)	240 (94%)	11 (4%)	3 (1%)	13 19
1	B	249/274 (91%)	224 (90%)	21 (8%)	4 (2%)	9 13
1	C	254/274 (93%)	230 (91%)	21 (8%)	3 (1%)	13 19
1	D	252/274 (92%)	235 (93%)	17 (7%)	0	100 100
All	All	1009/1096 (92%)	929 (92%)	70 (7%)	10 (1%)	15 23

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	321	PHE
1	B	337	SER
1	C	341	ASN
1	B	345	LEU
1	C	353	ASP

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	226/237 (95%)	216 (96%)	10 (4%)	28	45
1	B	223/237 (94%)	213 (96%)	10 (4%)	27	44
1	C	226/237 (95%)	215 (95%)	11 (5%)	25	40
1	D	224/237 (94%)	215 (96%)	9 (4%)	31	49
All	All	899/948 (95%)	859 (96%)	40 (4%)	28	45

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	470	GLU
1	D	386	ARG
1	C	489	GLN
1	D	329	ASP
1	D	466	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 28 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	266	HIS
1	D	523	HIS
1	C	412	GLN
1	D	489	GLN
1	C	397	ASN

5.3.3 RNA

There are no RNA molecules 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 monosaccharides in this entry.

5.6 Ligand geometry

Of 4 ligands modelled in this entry, 4 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

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, 95^{th} 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	258/274 (94%)	0.05	1 (0%)	92	91	18, 31, 55, 77	0
1	B	253/274 (92%)	0.40	12 (4%)	31	30	23, 47, 79, 80	0
1	C	258/274 (94%)	0.25	9 (3%)	44	43	21, 40, 77, 80	0
1	D	256/274 (93%)	0.11	5 (1%)	65	63	19, 36, 63, 76	0
All	All	1025/1096 (93%)	0.20	27 (2%)	56	54	18, 37, 74, 80	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	329	ASP	5.3
1	C	345	LEU	4.9
1	C	327	ASP	3.7
1	B	356	HIS	3.4
1	B	337	SER	3.4

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 monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
2	MG	A	701	1/1	0.87	0.21	44,44,44,44	0
2	MG	B	702	1/1	0.87	0.09	49,49,49,49	0
2	MG	D	704	1/1	0.87	0.15	46,46,46,46	0
2	MG	C	703	1/1	0.93	0.18	29,29,29,29	0

6.5 Other polymers (i)

There are no such residues in this entry.