wwPDB X-ray Structure Validation Summary Report (i)

Aug 20, 2023 – 03:27 AM EDT

PDB ID	:	2HEO
Title	:	General Structure-Based Approach to the Design of Protein Ligands: Application to the Design of Kv1.2 Potassium Channel Blockers.
Authors	:	Magis, C.; Gasparini, S.; Charbonnier, J.B.; Stura, E.; Le Du, M.H.; Menez, A.; Cuniasse, P.
Deposited on	:	2006-06-21
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
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 (1)) were used in the production of this report:

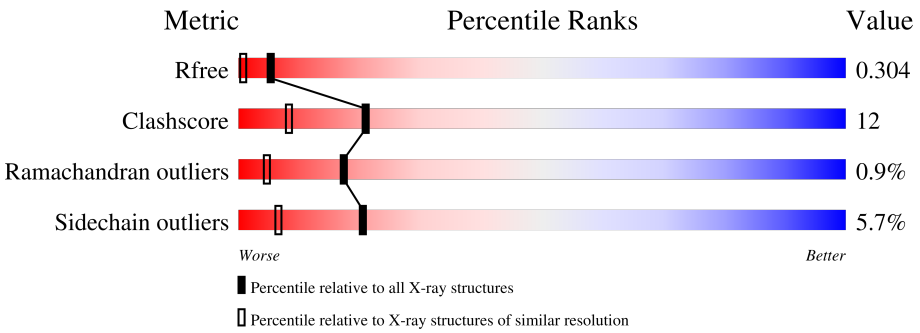
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 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}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (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 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%

Mol	Chain	Length	Quality of chain
1	B	7	57% 29% 14%
1	E	7	71% 14% 14%
2	A	67	66% 19% • 12%
2	D	67	67% 16% • 13%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 1409 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 DNA chain called 5'-D(*TP*CP*GP*CP*GP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	B	6	123	57	24	36	6	0	0	0
1	E	6	123	57	24	36	6	0	0	0

- Molecule 2 is a protein called Z-DNA binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	59	477	305	85	86	1	0	3	0
2	D	58	458	294	78	85	1	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	104	GLY	-	cloning artifact	UNP Q9QY24
A	105	SER	-	cloning artifact	UNP Q9QY24
A	106	HIS	-	cloning artifact	UNP Q9QY24
A	107	MET	-	cloning artifact	UNP Q9QY24
A	130	ALA	LYS	engineered mutation	UNP Q9QY24
A	132	PHE	GLY	engineered mutation	UNP Q9QY24
A	162	SER	GLU	engineered mutation	UNP Q9QY24
A	164	LYS	ALA	engineered mutation	UNP Q9QY24
A	165	TYR	THR	engineered mutation	UNP Q9QY24
D	104	GLY	-	cloning artifact	UNP Q9QY24
D	105	SER	-	cloning artifact	UNP Q9QY24
D	106	HIS	-	cloning artifact	UNP Q9QY24
D	107	MET	-	cloning artifact	UNP Q9QY24
D	130	ALA	LYS	engineered mutation	UNP Q9QY24
D	132	PHE	GLY	engineered mutation	UNP Q9QY24
D	162	SER	GLU	engineered mutation	UNP Q9QY24

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	164	LYS	ALA	engineered mutation	UNP Q9QY24
D	165	TYR	THR	engineered mutation	UNP Q9QY24

- Molecule 3 is water.

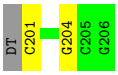
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	B	32	32	32	0	0
			Total	O		
3	E	29	29	29	0	0
			Total	O		
3	A	83	83	83	0	0
			Total	O		
3	D	84	84	84	0	0

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. 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. 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: 5'-D(*TP*CP*GP*CP*GP*CP*G)-3'

Chain B:

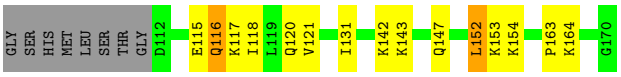


- Molecule 1: 5'-D(*TP*CP*GP*CP*GP*CP*G)-3'

Chain E:

- Molecule 2: Z-DNA binding protein 1

Chain A:



- Molecule 2: Z-DNA binding protein 1

Chain D:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	79.91Å 79.91Å 55.34Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 1.70 43.22 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-1.70) 99.8 (43.22-1.70)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.28 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.241 , 0.305 0.245 , 0.304	Depositor DCC
R_{free} test set	1129 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	24.4	Xtriage
Anisotropy	0.380	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 37.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.21$	Xtriage
Estimated twinning fraction	0.439 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	1409	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.00% 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 (i)

5.1 Standard geometry (i)

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	B	1.16	0/137	2.03	3/209 (1.4%)
1	E	1.32	2/137 (1.5%)	2.18	5/209 (2.4%)
2	A	0.76	1/500 (0.2%)	0.94	2/670 (0.3%)
2	D	0.65	0/466	0.74	0/630
All	All	0.86	3/1240 (0.2%)	1.27	10/1718 (0.6%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	152	LEU	C-N	8.07	1.52	1.34
1	E	204	DG	C2-N3	5.65	1.37	1.32
1	E	205	DC	N1-C6	5.07	1.40	1.37

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	152	LEU	O-C-N	-11.17	104.83	122.70
1	E	204	DG	O4'-C1'-N9	-9.70	101.21	108.00
1	B	204	DG	O4'-C1'-N9	-9.26	101.52	108.00
1	E	201	DC	OP1-P-OP2	8.94	133.00	119.60
2	A	152	LEU	CA-C-N	7.48	133.65	117.20

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	B	123	0	67	1	0
1	E	123	0	67	2	0
2	A	477	0	509	12	0
2	D	458	0	471	13	0
3	A	83	0	0	2	0
3	B	32	0	0	2	1
3	D	84	0	0	2	1
3	E	29	0	0	3	0
All	All	1409	0	1114	28	1

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 12.

The worst 5 of 28 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)
2:D:140:VAL:HG12	2:D:141:PRO:HD2	1.55	0.87
1:E:206:DG:N7	3:E:210:HOH:O	2.16	0.77
2:D:140:VAL:CG1	2:D:141:PRO:HD2	2.14	0.77
2:A:117:LYS:NZ	3:A:215:HOH:O	2.26	0.67
2:D:140:VAL:HG12	2:D:141:PRO:CD	2.27	0.65

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:B:175:HOH:O	3:D:219:HOH:O[6_665]	2.15	0.05

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	
2	A	60/67 (90%)	59 (98%)	1 (2%)	0	100	100
2	D	56/67 (84%)	55 (98%)	0	1 (2%)	8	1
All	All	116/134 (87%)	114 (98%)	1 (1%)	1 (1%)	17	5

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	114	LEU

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	
2	A	56/60 (93%)	52 (93%)	4 (7%)	14	3
2	D	52/60 (87%)	49 (94%)	3 (6%)	20	6
All	All	108/120 (90%)	101 (94%)	7 (6%)	20	4

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	154[B]	LYS
2	D	113	ASN
2	D	137	LYS
2	D	115	GLU
2	A	154[A]	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	116	GLN
2	A	120	GLN
2	A	133	GLN
2	D	133	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands (i)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers (i)

Unable to reproduce the depositors R factor - this section is therefore empty.