wwPDB X-ray Structure Validation Summary Report (i)

Jun 19, 2024 – 04:35 AM EDT

PDB ID	:	4K2C
Title	:	HSA Ligand Free
Authors	:	Wang, Y.; Luo, Z.; Shi, X.; Huang, M.
Deposited on	:	2013-04-08
Resolution	:	3.23 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:

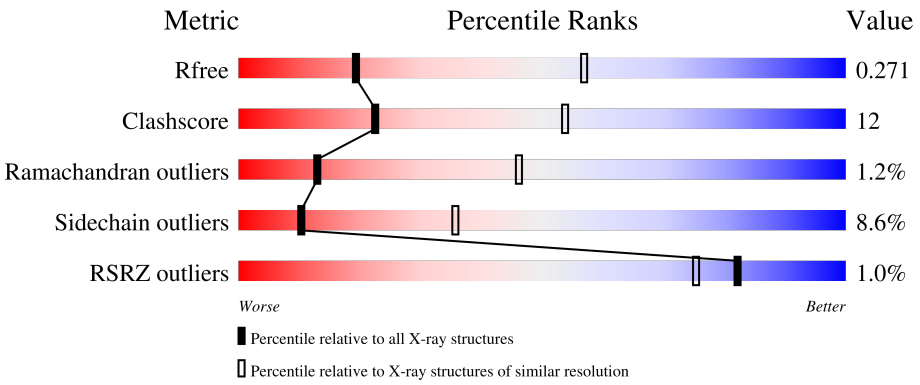
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.23 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	1619 (3.28-3.20)
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)
RSRZ outliers	127900	1567 (3.28-3.20)

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	585	% 69% 26% • •
1	B	585	% 66% 30% • •

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 9182 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 Serum albumin.

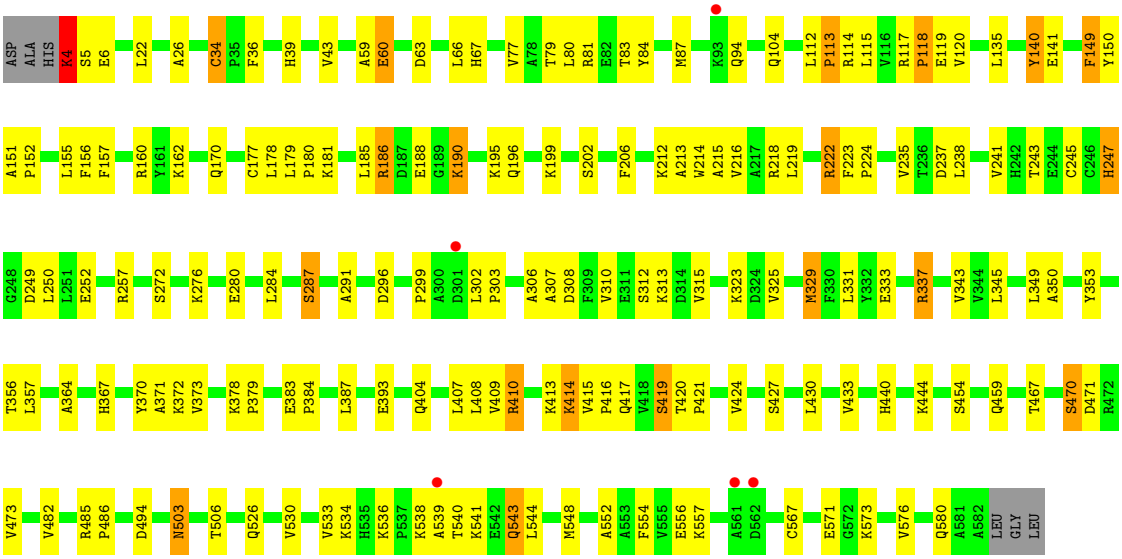
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	579	4591	2898	775	877	41	0	0	0
1	B	579	4591	2898	775	877	41	0	0	0

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.

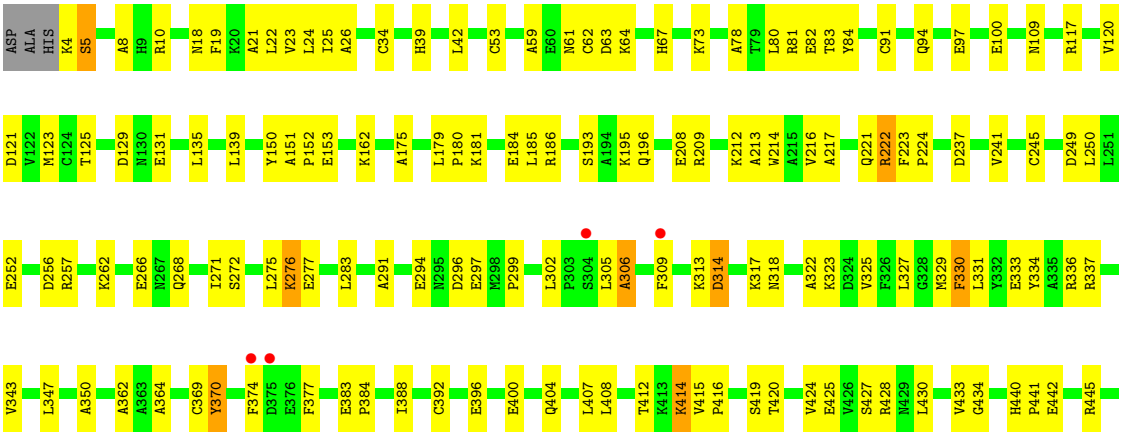
• Molecule 1: Serum albumin

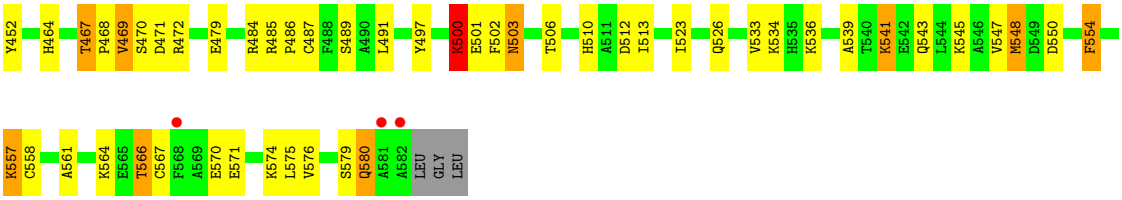
Chain A:



• Molecule 1: Serum albumin

Chain B:




Y452
H464
T467
P468
V469
S470
D471
R472
E479
R484
R485
P486
C487
F488
S489
A490
L491
Y497
K500
E501
F502
N503
T506
H510
A511
D512
I513
I523
Q526
V533
K534
H535
K536
A539
T540
K541
E542
Q543
L544
K545
A546
V547
M548
D549
D550
F554
K557
C558
A561
K564
E565
T566
C567
F568
A569
E570
E571
K574
L575
V576
S579
Q580
A581
A582
LEU
GLY
LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	58.41Å 179.89Å 57.46Å 90.00° 105.86° 90.00°	Depositor
Resolution (Å)	47.70 – 3.23 47.65 – 2.99	Depositor EDS
% Data completeness (in resolution range)	80.0 (47.70-3.23) 82.4 (47.65-2.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.29 (at 3.01Å)	Xtriage
Refinement program	CNS, REFMAC 5.7.0032	Depositor
R, R_{free}	0.213 , 0.265 0.231 , 0.271	Depositor DCC
R_{free} test set	975 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å^2)	90.3	Xtriage
Anisotropy	0.256	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 66.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.035 for l,-k,h	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9182	wwPDB-VP
Average B, all atoms (Å^2)	105.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.07% 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	A	0.57	0/4680	0.76	0/6316
1	B	0.53	0/4680	0.72	0/6316
All	All	0.55	0/9360	0.74	0/12632

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	2
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	4	LYS	Peptide
1	B	4	LYS	Peptide
1	B	500	LYS	Peptide

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	4591	0	4494	95	0
1	B	4591	0	4494	115	0
All	All	9182	0	8988	210	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 12.

The worst 5 of 210 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:80:LEU:HB2	1:A:84:TYR:HE2	1.46	0.81
1:B:151:ALA:HB2	1:B:250:LEU:HD22	1.65	0.78
1:B:302:LEU:HD13	1:B:337:ARG:HG2	1.67	0.77
1:A:299:PRO:HB2	1:A:302:LEU:HD21	1.66	0.76
1:A:222:ARG:NH1	1:A:291:ALA:O	2.20	0.75

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	577/585 (99%)	516 (89%)	55 (10%)	6 (1%)	15	50
1	B	577/585 (99%)	516 (89%)	53 (9%)	8 (1%)	11	43
All	All	1154/1170 (99%)	1032 (89%)	108 (9%)	14 (1%)	13	46

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	79	THR
1	A	364	ALA
1	A	60	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	81	ARG
1	B	362	ALA

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	503/511 (98%)	459 (91%)	44 (9%)	10	35
1	B	503/511 (98%)	460 (92%)	43 (8%)	10	36
All	All	1006/1022 (98%)	919 (91%)	87 (9%)	10	36

5 of 87 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	245	CYS
1	B	419	SER
1	B	277	GLU
1	B	323	LYS
1	B	479	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	503	ASN
1	B	221	GLN
1	B	94	GLN
1	B	196	GLN
1	B	67	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 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)

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	579/585 (98%)	-0.60	5 (0%)	84	78	51, 95, 147, 202	0
1	B	579/585 (98%)	-0.51	7 (1%)	79	70	63, 106, 164, 221	0
All	All	1158/1170 (98%)	-0.55	12 (1%)	82	75	51, 100, 159, 221	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	539	ALA	4.7
1	A	93	LYS	3.1
1	A	301	ASP	3.0
1	B	375	ASP	2.8
1	B	582	ALA	2.7

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 monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.