wwPDB X-ray Structure Validation Summary Report (i)

Oct 24, 2023 – 06:51 AM EDT

PDB ID	:	2Z2E
Title	:	Crystal Structure of Canine Milk Lysozyme Stabilized against Non-enzymatic Deamidation
Authors	:	Nonaka, Y.; Akieda, D.; Watanabe, N.; Tanaka, I.; Kamiya, M.; Aizawa, T.; Nitta, K.; Demura, M.; Kawano, K.
Deposited on	:	2007-05-21
Resolution	:	2.01 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:

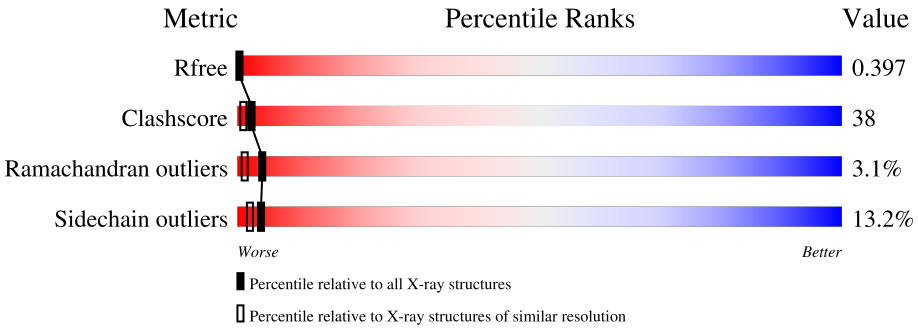
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.01 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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	A	129	39% 45% 14% .
1	B	129	37% 40% 19% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 2255 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.

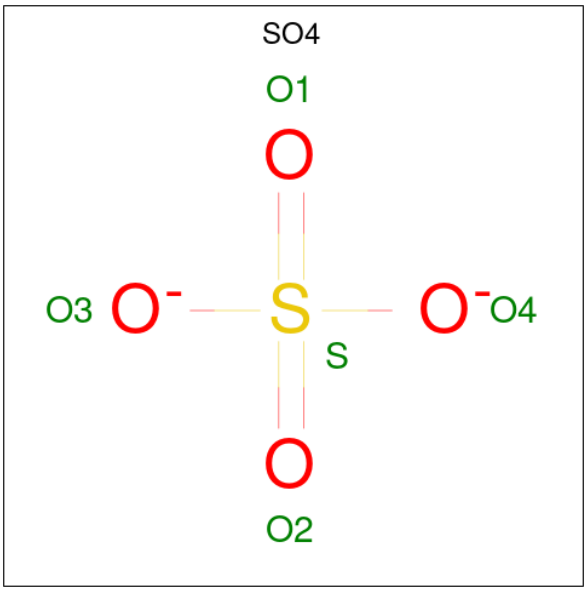
- Molecule 1 is a protein called Lysozyme C, milk isozyme.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	129	Total	C	N	O	S	0	0	0
			1015	635	177	190	13			
1	B	129	Total	C	N	O	S	0	0	0
			1015	635	177	190	13			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	44	GLN	ASN	engineered mutation	UNP P81708
A	47	GLN	ASN	engineered mutation	UNP P81708
A	49	GLN	ASN	engineered mutation	UNP P81708
A	68	GLN	ASN	engineered mutation	UNP P81708
A	103	GLN	ASN	engineered mutation	UNP P81708
B	44	GLN	ASN	engineered mutation	UNP P81708
B	47	GLN	ASN	engineered mutation	UNP P81708
B	49	GLN	ASN	engineered mutation	UNP P81708
B	68	GLN	ASN	engineered mutation	UNP P81708
B	103	GLN	ASN	engineered mutation	UNP P81708

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is water.

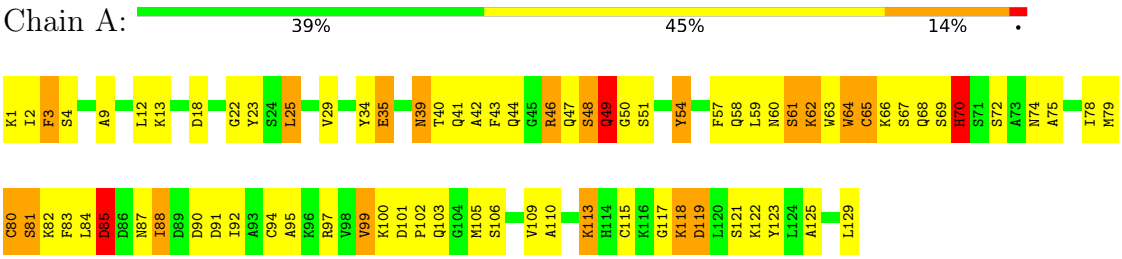
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	103	Total O 103 103	0	0
3	B	112	Total O 112 112	0	0

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.

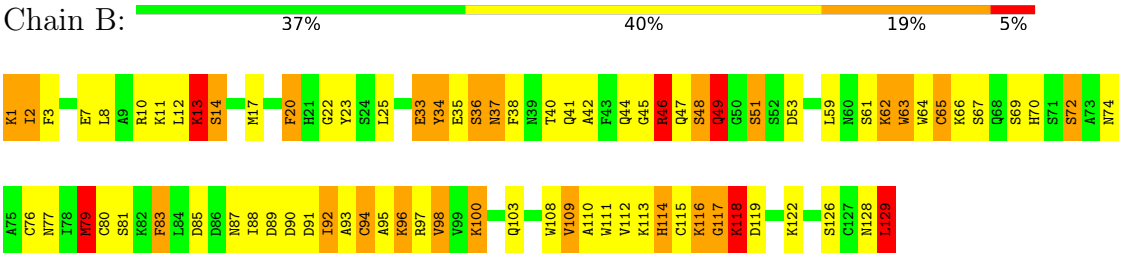
• Molecule 1: Lysozyme C, milk isozyme

Chain A:



• Molecule 1: Lysozyme C, milk isozyme

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	31.22Å 31.22Å 198.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	27.04 – 2.01 27.03 – 2.01	Depositor EDS
% Data completeness (in resolution range)	95.5 (27.04-2.01) 95.5 (27.03-2.01)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	8.58 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.273 , 0.413 0.261 , 0.397	Depositor DCC
R_{free} test set	682 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	17.6	Xtriage
Anisotropy	0.033	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 25.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.068 for -h,-k,l 0.085 for h,-h-k,-l 0.477 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	2255	wwPDB-VP
Average B, all atoms (Å^2)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.30% 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: 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	2.15	22/1038 (2.1%)	1.49	10/1391 (0.7%)
1	B	2.07	28/1038 (2.7%)	1.50	11/1391 (0.8%)
All	All	2.11	50/2076 (2.4%)	1.49	21/2782 (0.8%)

The worst 5 of 50 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	80	CYS	CB-SG	-12.06	1.61	1.82
1	A	129	LEU	C-OXT	-11.72	1.01	1.23
1	A	123	TYR	CD2-CE2	-10.91	1.23	1.39
1	A	99	VAL	CB-CG2	-10.09	1.31	1.52
1	B	129	LEU	C-OXT	-8.94	1.06	1.23

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	46	ARG	NE-CZ-NH1	9.91	125.25	120.30
1	A	85	ASP	CB-CG-OD2	-7.47	111.58	118.30
1	A	49	GLN	N-CA-C	-6.60	93.19	111.00
1	B	46	ARG	CG-CD-NE	6.34	125.12	111.80
1	B	62	LYS	CD-CE-NZ	6.20	125.96	111.70

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	1015	0	967	67	1
1	B	1015	0	967	87	1
2	A	5	0	0	1	0
2	B	5	0	0	0	0
3	A	103	0	0	11	3
3	B	112	0	0	20	2
All	All	2255	0	1934	153	4

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 38.

The worst 5 of 153 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:B:96:LYS:O	1:B:100:LYS:CD	1.72	1.37
1:A:67:SER:HB3	1:A:70:HIS:ND1	1.38	1.33
1:B:97:ARG:HG2	3:B:1100:HOH:O	1.35	1.26
1:B:8:LEU:HB2	3:B:1046:HOH:O	1.33	1.23
1:B:96:LYS:O	1:B:100:LYS:HD2	1.10	1.23

All (4) 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)
1:A:48:SER:O	3:A:1046:HOH:O[1_545]	1.20	1.00
1:B:46:ARG:NH1	3:B:1063:HOH:O[1_545]	1.71	0.49
3:A:1017:HOH:O	3:B:1030:HOH:O[1_565]	1.92	0.28
3:A:1028:HOH:O	3:A:1051:HOH:O[1_545]	2.17	0.03

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	127/129 (98%)	112 (88%)	12 (9%)	3 (2%)	6 2
1	B	127/129 (98%)	108 (85%)	14 (11%)	5 (4%)	3 1
All	All	254/258 (98%)	220 (87%)	26 (10%)	8 (3%)	4 1

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	48	SER
1	B	80	CYS
1	B	98	VAL
1	A	70	HIS
1	A	125	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	110/110 (100%)	96 (87%)	14 (13%)	4 2
1	B	110/110 (100%)	95 (86%)	15 (14%)	3 2
All	All	220/220 (100%)	191 (87%)	29 (13%)	4 2

5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	2	ILE
1	B	122	LYS
1	B	37	ASN
1	B	100	LYS
1	B	14	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	60	ASN
1	B	114	HIS
1	B	21	HIS
1	B	44	GLN
1	B	47	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)

2 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	B	1001	-	4,4,4	0.14	0	6,6,6	1.02	1 (16%)
2	SO4	A	1002	-	4,4,4	0.14	0	6,6,6	0.05	0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1001	SO4	O3-S-O1	2.13	120.42	109.31

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1002	SO4	1	0

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

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

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

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

6.4 Ligands

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

6.5 Other polymers

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