



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:38 PM GMT

PDB ID : 1UIK
Title : Crystal structure of soybean beta-conglycinin alpha prime homotrimer
Authors : Maruyama, Y.; Maruyama, N.; Mikami, B.; Utsumi, S.
Deposited on : 2003-07-16
Resolution : 2.30 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

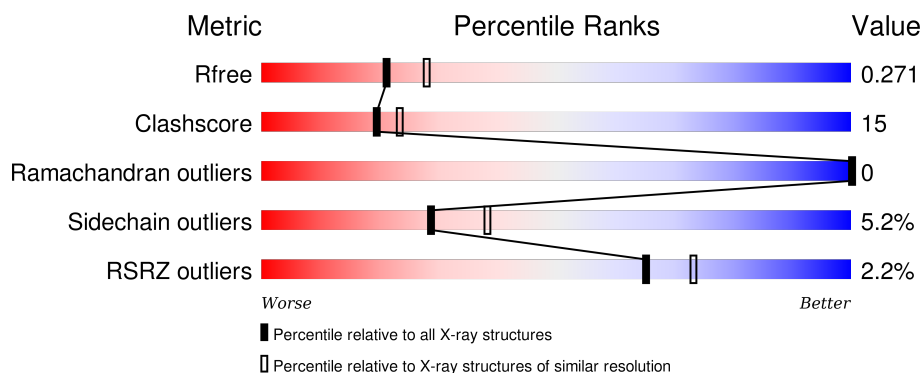
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	418	<div> <div>3%</div> <div>64%</div> <div>21%</div> <div>•</div> <div>13%</div> </div>
1	B	418	<div> <div>%</div> <div>63%</div> <div>21%</div> <div>•</div> <div>13%</div> </div>
1	C	418	<div> <div>2%</div> <div>65%</div> <div>19%</div> <div>•</div> <div>13%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9348 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 alpha prime subunit of beta-conglycinin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	364	Total	C	N	O	S	0	0	0
			2921	1857	504	558	2			
1	B	363	Total	C	N	O	S	0	1	0
			2919	1855	503	559	2			
1	C	364	Total	C	N	O	S	0	0	0
			2921	1857	504	558	2			

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

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

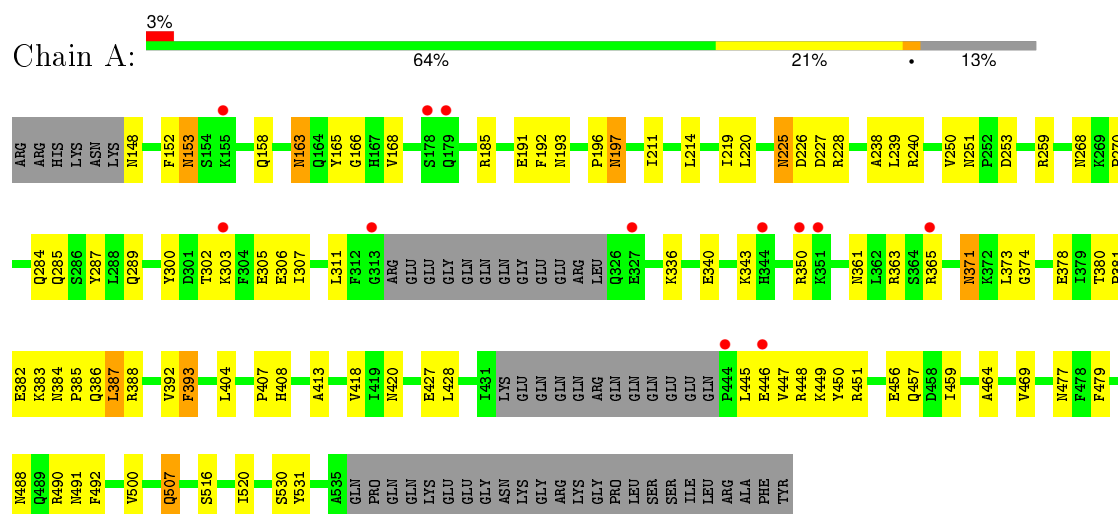
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	188	Total	O	0	0
			188	188		
3	B	196	Total	O	0	0
			196	196		
3	C	200	Total	O	0	0
			200	200		

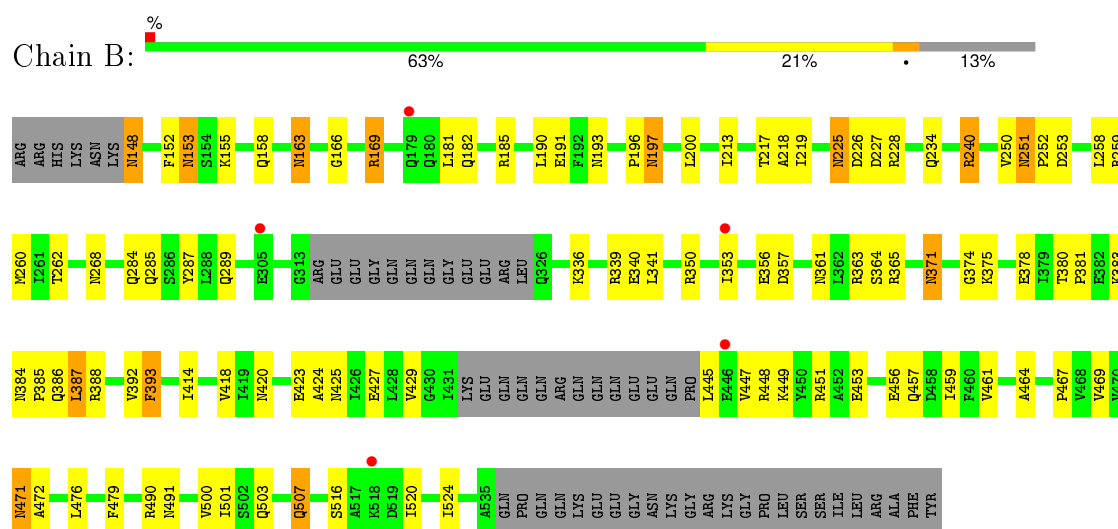
3 Residue-property plots

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 ($RSRZ > 2$). 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: alpha prime subunit of beta-conglycinin



- Molecule 1: alpha prime subunit of beta-conglycinin



- Molecule 1: alpha prime subunit of beta-conglycinin



N491	F492	Q503	I504	P505	S506	Q507	W508	Q509	S516	I520	I524	A535	GLN	PRO	GLN	GLN	LYS	GLU	GLU	GLY	ASN	LYS	GLY	ARG	LYS	GLY	PRO	LEU	SER	SER	ILE	LEU	LEU	ARG	ALA	PHE	TYR	K372	L373	E378	I379	T380	P381	E382	K383	N384	P385	Q386	L387	R388	V392	F393	H408	L417	E427	L428	I431	LYS	GLU	GLN	GLN	GLN	ARG	GLN	GLN	GLN	LEU	Q326	I330	K336	E340	K343	H344	A345	K346	S347	S348	S349	R350	I353	K358	N361	L362	R363	S364	R365	N371	P266																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
ARG	ARG	HIS	LYS	ASN	LYS	N143	N153	Q158	T159	N163	G166	R169	S178	Q179	Q180	R185	D186	E191	F192	N193	P196	N197	L200	I211	V212	I213	L214	N215	T217	A218	I219	N225	D226	A238	V250	N251	N257	L258	R259	W260	I261	P266																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
V267	N268	Q284	Q285	S286	Y287	I295	D301	T302	K303	F304	E305	E306	K309	G313	ARG	GLU	GLY	GLN	GLN	GLN	GLY	GLU	GLU	ARG	LEU	Q326	I330	K336	E340	K343	H344	A345	K346	S347	S348	S349	R350	I353	K358	N361	L362	R363	S364	R365	N371	P266																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
K372	L373	E378	I379	T380	P381	E382	K383	N384	P385	Q386	L387	R388	V392	F393	H408	L417	E427	L428	I431	LYS	GLU	GLN	GLN	GLN	ARG	GLN	GLN	GLU	GLN	P444	L445	E446	V447	Y450	R451	E456	Q457	D458	I459	I462	P463	A464	V468	V469	Q489	R490																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
M491	F492	D503	I504	P505	S506	Q507	W508	D509	S516	T520	T524	A535	GLN	PRO	GLN	GLN	LYS	GLU	GLU	GLY	ASN	LYS	GLY	ARG	LYS	GLY	PRO	LEU	SER	SER	ILE	LEU	LEU	ARG	ALA	PHE	TYR																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																</

4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	164.96Å 164.96Å 110.05Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	6.00 – 2.30 29.99 – 2.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (6.00-2.30) 88.5 (29.99-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.21 (at 2.00Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.221 , 0.269 0.223 , 0.271	Depositor DCC
R_{free} test set	3319 reflections (5.01%)	DCC
Wilson B-factor (Å ²)	26.6	Xtriage
Anisotropy	0.605	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 55.8	EDS
Estimated twinning fraction	0.000 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.54$, $\langle L^2 \rangle = 0.38$	Xtriage
Outliers	7 of 102805 reflections (0.007%)	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9348	wwPDB-VP
Average B, all atoms (Å ²)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 33.62 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.7719e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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: 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	0.36	0/2977	0.61	0/4026
1	B	0.38	0/2978	0.63	0/4027
1	C	0.36	0/2977	0.61	0/4026
All	All	0.37	0/8932	0.62	0/12079

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 [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	2921	0	2901	87	0
1	B	2919	0	2895	105	0
1	C	2921	0	2901	87	0
2	A	1	0	0	0	1
2	B	1	0	0	0	1
2	C	1	0	0	0	0
3	A	188	0	0	6	0
3	B	196	0	0	7	0
3	C	200	0	0	3	0
All	All	9348	0	8697	261	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 15.

The worst 5 of 261 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:507:GLN:HE21	1:A:507:GLN:H	1.02	0.99
1:A:507:GLN:NE2	1:A:507:GLN:H	1.66	0.94
1:B:507:GLN:HE21	1:B:507:GLN:H	1.05	0.93
1:C:193:ASN:HD21	1:C:259:ARG:HE	1.20	0.88
1:B:507:GLN:H	1:B:507:GLN:NE2	1.71	0.88

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)
2:A:1001:MG:MG	2:B:1002:MG:MG[6_666]	0.06	2.14

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	358/418 (86%)	344 (96%)	14 (4%)	0	100	100
1	B	358/418 (86%)	346 (97%)	12 (3%)	0	100	100
1	C	358/418 (86%)	346 (97%)	12 (3%)	0	100	100
All	All	1074/1254 (86%)	1036 (96%)	38 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	328/376 (87%)	315 (96%)	13 (4%)	38	52
1	B	328/376 (87%)	309 (94%)	19 (6%)	25	33
1	C	328/376 (87%)	309 (94%)	19 (6%)	25	33
All	All	984/1128 (87%)	933 (95%)	51 (5%)	29	38

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	251	ASN
1	B	387	LEU
1	C	387	LEU
1	B	341	LEU
1	B	393	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 92 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	257	ASN
1	B	471	ASN
1	C	484	ASN
1	B	268	ASN
1	B	361	ASN

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	364/418 (87%)	-0.09	12 (3%) 50 59	20, 32, 47, 57	0
1	B	363/418 (86%)	-0.22	5 (1%) 78 83	18, 31, 46, 55	0
1	C	364/418 (87%)	-0.18	7 (1%) 70 76	17, 31, 46, 55	0
All	All	1091/1254 (87%)	-0.16	24 (2%) 65 73	17, 31, 47, 57	0

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	350	ARG	3.8
1	A	444	PRO	3.7
1	B	353	ILE	3.4
1	A	155	LYS	3.1
1	C	179	GLN	3.1

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	MG	B	1002	1/1	0.99	0.06	-2.18	49,49,49,49	0
2	MG	A	1001	1/1	0.98	0.06	-2.31	47,47,47,47	0
2	MG	C	1003	1/1	0.96	0.15	-	34,34,34,34	1

6.5 Other polymers [i](#)

There are no such residues in this entry.