wwPDB X-ray Structure Validation Summary Report (i)

Oct 3, 2021 – 01:51 AM EDT

PDB ID	:	3HLS
Title	:	Crystal structure of the signaling helix coiled-coil doimain of the BETA-1 subunit of the soluble guanylyl cyclase
Authors	:	Ma, X.; van den Akker, F.
Deposited on	:	2009-05-28
Resolution	:	2.15 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 following versions of software and data (see references (i)) were used in the production of this report:

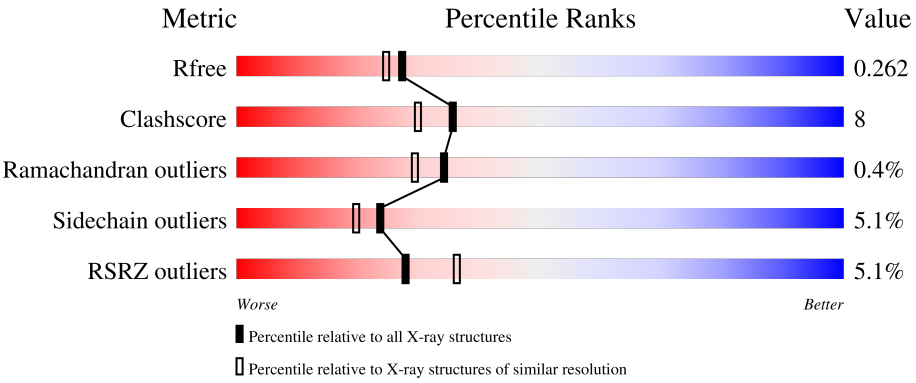
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.15 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.



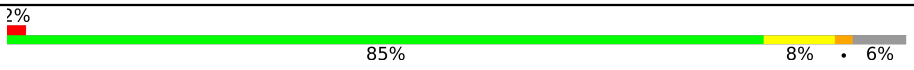
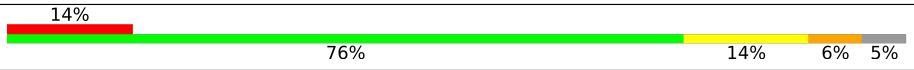
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	66	3% 86% 11%
1	B	66	3% 88% 8% 5%
1	C	66	9% 80% 15% 5%
1	D	66	82% 11% 5%
1	E	66	74% 15% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	66	
1	G	66	
1	H	66	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4612 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 Guanylate cyclase soluble subunit beta-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	65	Total	C	N	O	Se	0	0	0
			531	333	92	104	2			
1	B	63	Total	C	N	O	Se	0	0	0
			521	328	90	101	2			
1	C	63	Total	C	N	O	Se	0	0	0
			517	325	88	102	2			
1	D	63	Total	C	N	O	Se	0	0	0
			517	325	88	102	2			
1	E	64	Total	C	N	O	Se	0	0	0
			521	327	89	103	2			
1	F	62	Total	C	N	O	Se	0	0	0
			511	322	87	100	2			
1	G	63	Total	C	N	O	Se	0	0	0
			517	325	88	102	2			
1	H	63	Total	C	N	O	Se	0	0	0
			517	325	88	102	2			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	344	GLY	-	expression tag	UNP P20595
A	345	SER	-	expression tag	UNP P20595
A	346	HIS	-	expression tag	UNP P20595
A	347	MSE	-	expression tag	UNP P20595
A	371	MSE	ILE	engineered mutation	UNP P20595
B	344	GLY	-	expression tag	UNP P20595
B	345	SER	-	expression tag	UNP P20595
B	346	HIS	-	expression tag	UNP P20595
B	347	MSE	-	expression tag	UNP P20595
B	371	MSE	ILE	engineered mutation	UNP P20595
C	344	GLY	-	expression tag	UNP P20595
C	345	SER	-	expression tag	UNP P20595
C	346	HIS	-	expression tag	UNP P20595

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	347	MSE	-	expression tag	UNP P20595
C	371	MSE	ILE	engineered mutation	UNP P20595
D	344	GLY	-	expression tag	UNP P20595
D	345	SER	-	expression tag	UNP P20595
D	346	HIS	-	expression tag	UNP P20595
D	347	MSE	-	expression tag	UNP P20595
D	371	MSE	ILE	engineered mutation	UNP P20595
E	344	GLY	-	expression tag	UNP P20595
E	345	SER	-	expression tag	UNP P20595
E	346	HIS	-	expression tag	UNP P20595
E	347	MSE	-	expression tag	UNP P20595
E	371	MSE	ILE	engineered mutation	UNP P20595
F	344	GLY	-	expression tag	UNP P20595
F	345	SER	-	expression tag	UNP P20595
F	346	HIS	-	expression tag	UNP P20595
F	347	MSE	-	expression tag	UNP P20595
F	371	MSE	ILE	engineered mutation	UNP P20595
G	344	GLY	-	expression tag	UNP P20595
G	345	SER	-	expression tag	UNP P20595
G	346	HIS	-	expression tag	UNP P20595
G	347	MSE	-	expression tag	UNP P20595
G	371	MSE	ILE	engineered mutation	UNP P20595
H	344	GLY	-	expression tag	UNP P20595
H	345	SER	-	expression tag	UNP P20595
H	346	HIS	-	expression tag	UNP P20595
H	347	MSE	-	expression tag	UNP P20595
H	371	MSE	ILE	engineered mutation	UNP P20595

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	56	Total O 56 56	0	0
2	B	79	Total O 79 79	0	0
2	C	55	Total O 55 55	0	0
2	D	56	Total O 56 56	0	0
2	E	58	Total O 58 58	0	0
2	F	56	Total O 56 56	0	0

Continued on next page...

Continued from previous page...

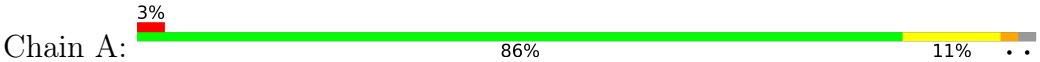
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	G	58	Total O 58 58	0	0
2	H	42	Total O 42 42	0	0

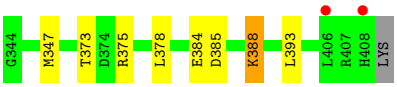
3 Residue-property plots (i)

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: Guanylate cyclase soluble subunit beta-1

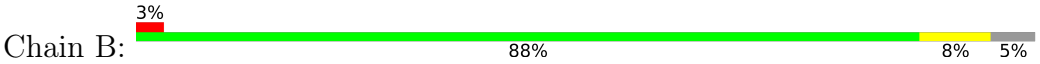
Chain A:





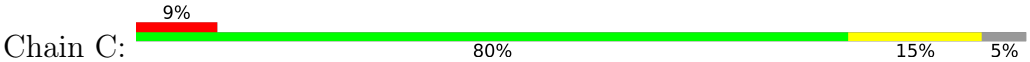
• Molecule 1: Guanylate cyclase soluble subunit beta-1

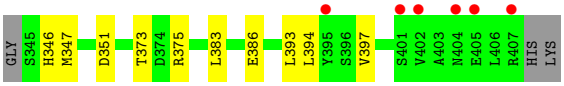
Chain B:



• Molecule 1: Guanylate cyclase soluble subunit beta-1

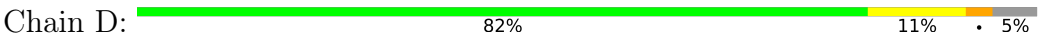
Chain C:

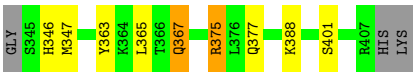




• Molecule 1: Guanylate cyclase soluble subunit beta-1

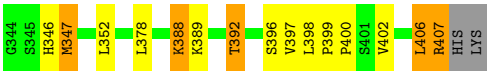
Chain D:





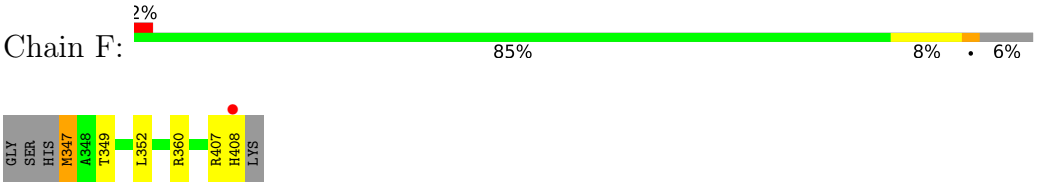
• Molecule 1: Guanylate cyclase soluble subunit beta-1

Chain E:



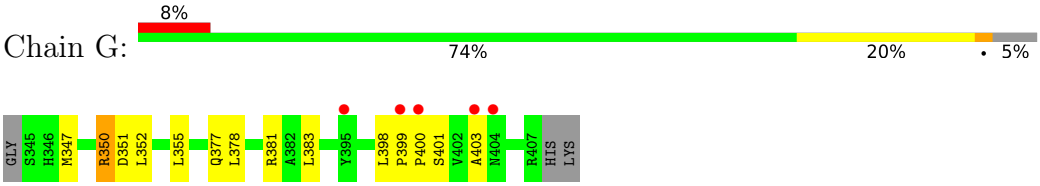
• Molecule 1: Guanylate cyclase soluble subunit beta-1

Chain F:



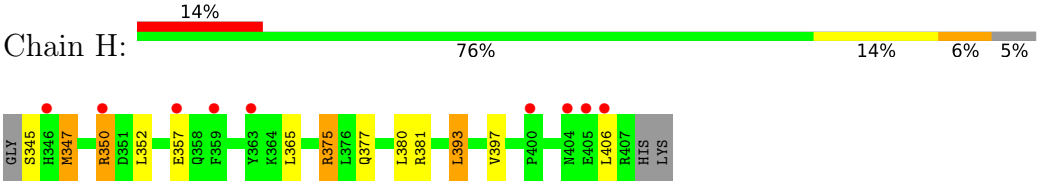
• Molecule 1: Guanylate cyclase soluble subunit beta-1

Chain G:



• Molecule 1: Guanylate cyclase soluble subunit beta-1

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	152.04Å 65.81Å 98.63Å 90.00° 129.95° 90.00°	Depositor
Resolution (Å)	39.81 – 2.15 39.81 – 2.15	Depositor EDS
% Data completeness (in resolution range)	98.9 (39.81-2.15) 98.9 (39.81-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	2.20 (at 2.16Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.213 , 0.261 0.213 , 0.262	Depositor DCC
R_{free} test set	2023 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	32.2	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 46.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.039 for -h-2*l,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4612	wwPDB-VP
Average B, all atoms (Å^2)	40.0	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 34.65 % 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 6.5966e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

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	0.48	0/535	0.62	0/717
1	B	0.54	0/525	0.59	0/704
1	C	0.47	0/520	0.60	0/697
1	D	0.51	0/520	0.67	0/697
1	E	0.51	0/524	0.62	0/702
1	F	0.51	0/515	0.62	0/692
1	G	0.49	0/520	0.64	0/697
1	H	0.58	0/520	0.70	0/697
All	All	0.51	0/4179	0.64	0/5603

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	531	0	542	8	0
1	B	521	0	534	6	0
1	C	517	0	532	12	0
1	D	517	0	532	13	0
1	E	521	0	535	16	0
1	F	511	0	527	7	0
1	G	517	0	532	15	0
1	H	517	0	532	13	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	56	0	0	3	0
2	B	79	0	0	1	0
2	C	55	0	0	1	0
2	D	56	0	0	0	0
2	E	58	0	0	2	0
2	F	56	0	0	0	0
2	G	58	0	0	2	0
2	H	42	0	0	0	0
All	All	4612	0	4266	68	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 8.

The worst 5 of 68 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:G:350:ARG:HH11	1:G:350:ARG:HG2	1.02	1.08
1:E:388:LYS:HE3	2:E:84:HOH:O	1.55	1.05
1:E:347:MSE:HE3	1:E:352:LEU:HD23	1.43	0.98
1:H:375:ARG:HG2	1:H:375:ARG:HH11	1.30	0.96
1:A:347:MSE:CE	1:C:346:HIS:HB2	2.00	0.91

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	63/66 (96%)	62 (98%)	1 (2%)	0	100	100
1	B	61/66 (92%)	61 (100%)	0	0	100	100
1	C	61/66 (92%)	59 (97%)	2 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	D	61/66 (92%)	60 (98%)	1 (2%)	0	100 100
1	E	62/66 (94%)	61 (98%)	0	1 (2%)	9 4
1	F	60/66 (91%)	60 (100%)	0	0	100 100
1	G	61/66 (92%)	60 (98%)	0	1 (2%)	9 4
1	H	61/66 (92%)	61 (100%)	0	0	100 100
All	All	490/528 (93%)	484 (99%)	4 (1%)	2 (0%)	34 29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	347	MSE
1	G	400	PRO

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	60/59 (102%)	57 (95%)	3 (5%)	24 21
1	B	59/59 (100%)	58 (98%)	1 (2%)	60 65
1	C	59/59 (100%)	59 (100%)	0	100 100
1	D	59/59 (100%)	56 (95%)	3 (5%)	24 20
1	E	59/59 (100%)	54 (92%)	5 (8%)	10 6
1	F	58/59 (98%)	55 (95%)	3 (5%)	23 19
1	G	59/59 (100%)	57 (97%)	2 (3%)	37 35
1	H	59/59 (100%)	52 (88%)	7 (12%)	5 2
All	All	472/472 (100%)	448 (95%)	24 (5%)	24 20

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	360	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	H	347	MSE
1	G	378	LEU
1	H	350	ARG
1	D	401	SER

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

Mol	Chain	Res	Type
1	A	377	GLN
1	B	377	GLN
1	E	346	HIS
1	G	346	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

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	63/66 (95%)	0.13	2 (3%) 47 56	20, 33, 54, 89	0
1	B	61/66 (92%)	0.07	2 (3%) 46 55	20, 29, 49, 74	0
1	C	61/66 (92%)	0.26	6 (9%) 7 11	21, 35, 86, 101	0
1	D	61/66 (92%)	0.22	0 100 100	24, 36, 76, 81	0
1	E	62/66 (93%)	0.11	0 100 100	22, 35, 53, 58	0
1	F	60/66 (90%)	0.01	1 (1%) 70 76	23, 33, 49, 61	0
1	G	61/66 (92%)	0.42	5 (8%) 11 16	25, 39, 77, 87	0
1	H	61/66 (92%)	0.51	9 (14%) 2 3	21, 38, 87, 91	0
All	All	490/528 (92%)	0.22	25 (5%) 28 36	20, 35, 76, 101	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	395	TYR	5.3
1	G	404	ASN	4.3
1	A	408	HIS	4.1
1	C	395	TYR	3.9
1	B	408	HIS	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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.