

Integrative Structure Validation Report ?

March 27, 2025 - 10:00 AM PDT

The following software was used in the production of this report:

Integrative Modeling Validation Version 2.0

Python-IHM Version 1.8

MolProbity Version 4.5.2

PDB ID	9A0B
PDB-Dev ID	PDBDEV_00000047
Structure Title	Structure determination of the HgcAB complex using metagenome sequence data, coevolution analysis, and ab initio structure calculations
Structure Authors	Cooper CJ; Zheng K; Rush KW; Johs A; Sanders BC; Pavlopoulos GA; Kyrpides NC; Podar M; Ovchinnikov S; Ragsdale SW; Parks JM
Deposited on	2020-04-30

This is a PDB-IHM IM Structure Validation Report for a publicly released PDB-IHM entry.

We welcome your comments at helpdesk@pdb-ihm.org

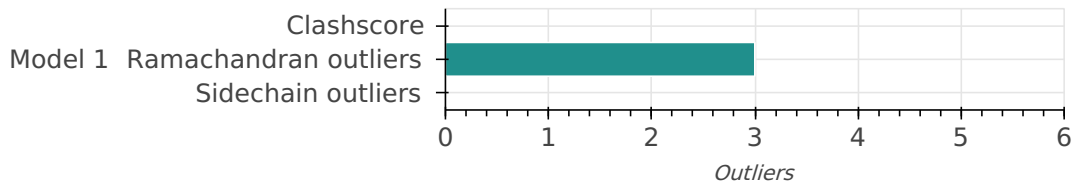
A user guide is available at https://pdb-ihm.org/validation_help.html with specific help available everywhere you see the ? symbol.

List of references used to build this report is available [here](#).

Overall quality ?

This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS datasets. Data quality and fit to model assessments for other datasets and model uncertainty are under development. Number of plots is limited to 256.

Model Quality: MolProbity Analysis



Ensemble information ?

This entry consists of 0 distinct ensemble(s).

Summary ?

This entry consists of 1 model(s). A total of 6 datasets were used to build this entry.

Representation ?

This entry has 1 representation(s).

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
1	1	1	HgcA	A	329	-	1-179, 161-329	100.00 / 100.00	Atomic
		2	HgcB	B	95	-	1-95	100.00 / 100.00	Atomic
		3	COBALAMIN	C [A]	Non-polymeric	-	-	Not available / Not available	Atomic
		4	IRON/SULFUR CLUSTER	D [B] E [B]	Non-polymeric	-	-	Not available / Not available	Atomic

Datasets used for modeling ?

There are 6 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Predicted contacts	Not available	Not available
2	De Novo model	Not available	Not available
3	De Novo model	Not available	Not available
4	De Novo model	Not available	Not available
5	Experimental model	PDB	2FDN
6	Experimental model	PDB	8C7P

Methodology and software ?

This entry is a result of 1 distinct protocol(s).

Step number	Protocol ID	Method name	Method type	Method description	Number of computed models	Multi state modeling	Multi scale modeling
1	1	modeling with restraints from coevolution analysis	ab initio modeling	None	1500	False	False

There are 4 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	HH-Suite	Not available	sequence analysis	https://github.com/soedinglab/hh-suite
2	HMMER	3.1b1	sequence analysis	http://hmmer.org
3	GREMLIN	Not available	contact prediction	http://gremlin.bakerlab.org/
4	ROSETTA	Not available	model building	https://github.com/RosettaCommons

Data quality ?

Predicted contacts

Validation for this section is under development.

Model quality ?

For models with atomic structures, MolProbity analysis is performed. For models with coarse-grained or multi-scale structures, excluded volume analysis is performed.

Standard geometry: bond outliers ?

There are 10 bond length outliers in this entry (0.29% of 3448 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	1	B12	C15-C16	6.21	1.45	1.33	1	1
C	1	B12	C41-C8	5.64	1.64	1.52	1	1
C	1	B12	C4-N21	5.17	1.31	1.42	1	1
C	1	B12	C16-N24	5.07	1.31	1.42	1	1
C	1	B12	C2P-C3P	4.90	1.43	1.52	1	1
C	1	B12	C19-N24	4.60	1.51	1.42	1	1
C	1	B12	C1-C2	4.54	1.61	1.52	1	1
A	119	LEU	CB-CG	4.46	1.44	1.53	1	1
C	1	B12	C10-C9	4.14	1.39	1.48	1	1
C	1	B12	C9-N22	4.04	1.34	1.42	1	1

Standard geometry: angle outliers ?

There are 29 bond angle outliers in this entry (0.61% of 4758 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
C	1	B12	C16-N24-CO	8.49	134.98	109.50	1	1
C	1	B12	C18-C60-C61	7.92	133.22	109.47	1	1
C	1	B12	O2-P-O3	7.74	125.83	102.60	1	1
A	234	PHE	CA-CB-CG	7.60	121.40	113.80	1	1
C	1	B12	C4-N21-CO	7.06	130.67	109.50	1	1
B	8	ASP	CA-CB-CG	6.50	119.10	112.60	1	1
C	1	B12	C6-N22-CO	6.21	128.14	109.50	1	1
C	1	B12	C14-N23-CO	6.11	127.82	109.50	1	1
A	187	PHE	CA-CB-CG	6.06	107.74	113.80	1	1
C	1	B12	O2-P-O4	5.95	91.66	109.50	1	1
C	1	B12	C17-C18-C60	5.94	128.83	111.00	1	1
C	1	B12	C3R-C2R-O7R	5.63	126.37	109.47	1	1
A	91	PHE	CA-CB-CG	5.53	119.33	113.80	1	1
C	1	B12	C18-C19-N24	5.49	103.54	120.00	1	1
A	135	PHE	CA-CB-CG	5.00	118.80	113.80	1	1
A	206	PRO	CA-C-N	4.76	124.04	116.90	1	1
A	166	PHE	C-N-CA	4.64	130.05	121.70	1	1

Chain	Res	Type	Atoms	Z	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	80	ILE	C-N-CA	4.63	113.37	121.70	1	1
C	1	B12	C11-N23-CO	4.61	123.33	109.50	1	1
A	66	THR	CA-C-O	4.53	113.10	120.80	1	1
A	233	LEU	C-N-CA	4.53	129.85	121.70	1	1
A	53	ASN	CA-CB-CG	4.45	117.05	112.60	1	1
C	1	B12	C9-N22-CO	4.40	122.70	109.50	1	1
C	1	B12	C13-C12-C47	4.27	123.81	111.00	1	1
A	117	PRO	C-N-CA	4.24	114.07	121.70	1	1
C	1	B12	C19-C1-C2	4.20	123.60	111.00	1	1
C	1	B12	C1-C2-C3	4.18	98.46	111.00	1	1
A	190	ILE	CA-C-N	4.15	123.12	116.90	1	1
C	1	B12	C1-C19-N24	4.14	107.58	120.00	1	1

Too-close contacts ?

The following all-atom clashscore is based on a MolProbity analysis. All-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The table below contains clashscores for all atomic models in this entry.

Model ID	Clash score	Number of clashes
1	0.00	0

There are no too-close contacts.

Torsion angles: Protein backbone ?

In the following table, Ramachandran outliers are listed. The Analysed column shows the number of residues for which the backbone conformation was analysed.

Model ID	Analysed	Favored	Allowed	Outliers
1	420	402	15	3

There are 3 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	118	GLN	1
A	167	THR	1
B	70	GLY	1

Torsion angles : Protein sidechains ?

In the following table, sidechain rotameric outliers are listed. The Analysed column shows the number of residues for which the sidechain conformation was analysed.

Model ID	Analysed	Favored	Allowed	Outliers
1	335	334	1	0

Fit of model to data used for modeling ?

Predicted contacts

Validation for this section is under development.

Fit of model to data used for validation ?

Validation for this section is under development.

Acknowledgments

The development of integrative model validation metrics, implementation of a model validation pipeline, and creation of a validation report for integrative structures are funded by NSF awards to the [PDB-IHM team](#) (DBI-1756248, DBI-2112966, DBI-2112967, DBI-2112968, and DBI-1756250) and awards from NSF, NIH, and DOE to the [RCSB PDB](#) (DBI-2321666, R01GM157729, and DE-SC0019749). The PDB-IHM team and members of the [Sali lab](#) contributed model validation metrics and software packages.

Dr. Jill Trewhella, Dr. Dina Schneidman, and members of the [SASBDB](#) repository are acknowledged for their advice and support in implementing SAS validation methods. Team members from the labs of Dr. Juri Rappsilber, Dr. Alexander Leitner, Dr. Andrea Graziadei, and members of [PRIDE](#) database are acknowledged for their advice and support in implementing crosslinking-MS validation methods. We are grateful to Dr. Shruthi Viswanath for discussions about uncertainty assessment of integrative structural models.

Members of the [wwPDB Integrative/Hybrid Methods Task Force](#) provided recommendations and community support for the project.