

Integrative Structure Validation Report

July 22, 2024 - 04:05 PM PDT

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

Python-IHM Version 1.3

MolProbity Version 4.5.2

Integrative Modeling Validation Version 1.2

| | |
|-------------------|---|
| 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 |

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

We welcome your comments at pdb-dev@mail.wwpdb.org

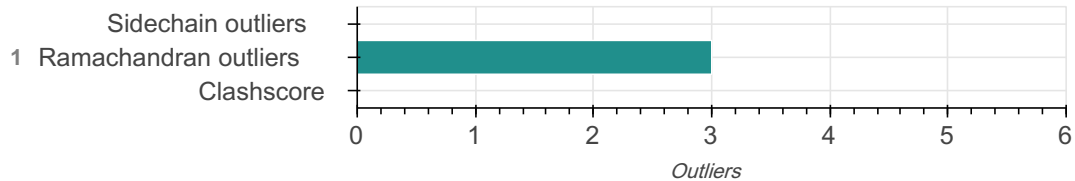
A user guide is available at https://pdb-dev.wwpdb.org/validation_help.html with specific help available everywhere you see the  symbol.

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 assessment for SAS datasets and fit to model assessments for SAS 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 unique models, with 2 subunits in each model. A total of 6 datasets or restraints were used to build this entry. Each model is represented by 0 rigid bodies and 2 flexible or non-rigid units.

Entry composition ?

There is 1 unique type of models in this entry. This model is titled None/Best scoring model.

| Model ID | Subunit number | Subunit ID | Subunit name | Chain ID | Chain ID [auth] | Total residues |
|----------|----------------|------------|--------------|----------|-----------------|----------------|
| 1 | 1 | 1 | HgcA | A | A | 329 |
| 1 | 2 | 2 | HgcB | B | B | 95 |

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 | Other | 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 |

| ID | Dataset type | Database name | Data access code |
|----|--------------------|---------------|------------------|
| 6 | Experimental model | PDB | 6FFV |

Representation ?

This entry has only one representation and includes 0 rigid bodies and 2 flexible units

| Chain ID | Rigid bodies | Non-rigid segments |
|----------|--------------|--------------------|
| A | - | 1-329 |
| B | - | 1-95 |
| C | - | - |
| D | - | - |
| E | - | - |

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 | ab initio modeling | modeling with restraints from coevolution analysis | 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

Model quality

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

Standard geometry: bond outliers

Bond length outliers can not be evaluated for this model

Standard geometry: angle outliers

Bond angle outliers do not exist or can not be evaluated for this model

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 the models in this entry.

| Model ID | Clash score | Number of clashes |
|----------|-------------|-------------------|
| 1 | 0.00 | 0 |

All 0 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

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 |

Detailed list of outliers are tabulated below.

Torsion angles: Protein sidechains

In the following table, sidechain 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 |

Detailed list of outliers are tabulated below.

Fit of model to data used for modeling ?

Fit of model to data used for validation ?

Validation for this section is under development.

Acknowledgements

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 ABI awards (DBI-1756248, DBI-2112966, DBI-2112967, DBI-2112968, and DBI-1756250). The [PDB-Dev team](#) and members of [Sali lab](#) contributed model validation metrics and software packages.

Implementation of validation methods for SAS data and SAS-based models are funded by [RCSB PDB](#) (grant number DBI-1832184). Dr. Stephen Burley, Dr. John Westbrook, and Dr. Jasmine Young from [RCSB PDB](#), Dr. Jill Trehwella, Dr. Dina Schneidman, and members of the [SASBDB](#) repository are acknowledged for their advice and support in implementing SAS validation methods.

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