Integrative Structure Validation Report

July 22, 2024 - 04:59 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	9A2H
PDB-Dev ID	PDBDEV_00000166
Structure Title	Model of E. coli GlgA by in-cell photo-crosslinking MS and deep learning
Structure Authors	Stahl, K.; Graziadei, A.; Dau, T.; Brock, O.; Rappsilber, J.

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

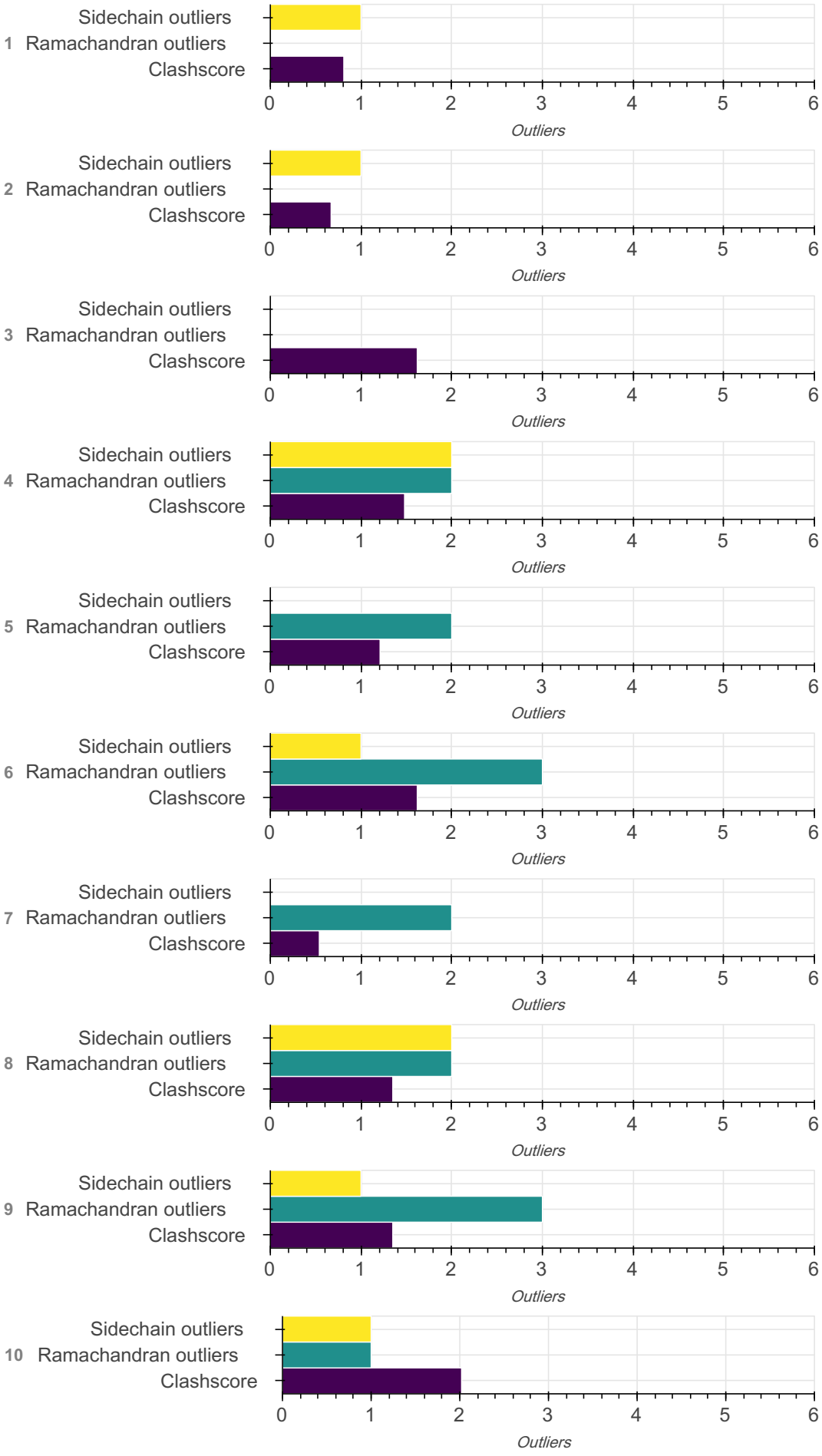
A user guide is available at https://pdb-dev.wwpdb.org/validation_help.html with specific help available everywhere you see the ? symbol.

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


1
2
3
4
5
6
7
8
9
10
Sidechain outliers
Ramachandran outliers
Clashscore
0
1
2
3
4
5
6
Outliers

Ensemble information

This entry consists of 0 distinct ensemble(s).

Summary

This entry consists of 10 unique models, with 1 subunits in each model. A total of 1 datasets or restraints were used to build this entry. Each model is represented by 0 rigid bodies and 1 flexible or non-rigid units.

Entry composition

There are 10 unique types of models in this entry.These models are titled None, None, None, None, None, None, None, None, None, None respectively.

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Chain ID [auth]	Total residues
1	1	1	P0A6U8	A	A	477
2	1	1	P0A6U8	A	A	477
3	1	1	P0A6U8	A	A	477
4	1	1	P0A6U8	A	A	477
5	1	1	P0A6U8	A	A	477
6	1	1	P0A6U8	A	A	477
7	1	1	P0A6U8	A	A	477
8	1	1	P0A6U8	A	A	477
9	1	1	P0A6U8	A	A	477
10	1	1	P0A6U8	A	A	477

Datasets used for modeling

There is 1 unique dataset used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Crosslinking-MS data	jPOSTrepo	JPST001851

Representation

This entry has only one representation and includes 0 rigid bodies and 1 flexible units

Chain ID	Rigid bodies	Non-rigid segments
A	-	1-477

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	AlphaLink with 10 msa subsamples	AlphaLink	None	10	False	False

There is 1 software package reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	AlphaLink	1.0	model building	https://github.com/lhatsk/AlphaLink

Data quality

Crosslinking-MS

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

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	0.81	6
2	0.67	5
3	1.62	12
4	1.48	11
5	1.21	9
6	1.62	12
7	0.54	4
8	1.35	10
9	1.35	10
10	2.02	15

All 94 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Model ID	Atom-1	Atom-2	Clash overlap (Å)
1	A:14:LEU:HD21	A:94:PRO:HG3	0.468
1	A:260:LEU:HD13	A:358:ALA:HA	0.457
1	A:206:ASP:HA	A:240:ARG:HH21	0.448
1	A:95:TYR:CE1	A:139:HIS:NE2	0.444
1	A:215:TYR:CE2	A:219:ILE:CG2	0.436
1	A:13:LEU:HD22	A:50:GLY:HA3	0.431
2	A:11:PHE:CD2	A:12:PRO:HA	0.566

Model ID	Atom-1	Atom-2	Clash overlap (Å)
2	A:16:THR:HA	A:95:TYR:CE2	0.543
2	A:260:LEU:HD13	A:358:ALA:HA	0.467
2	A:115:TRP:CD1	A:150:ARG:HH22	0.428
2	A:13:LEU:HD22	A:50:GLY:HA3	0.414
3	A:14:LEU:HD21	A:94:PRO:HG3	0.692
3	A:380:GLY:HA3	A:383:GLN:HE22	0.662
3	A:260:LEU:HD13	A:358:ALA:HA	0.578
3	A:407:ASP:CG	A:437:ARG:HH22	0.489
3	A:292:VAL:HG13	A:321:GLY:O	0.478
3	A:215:TYR:CE2	A:219:ILE:CG2	0.455
3	A:384:LEU:HD22	A:460:PHE:CG	0.432
3	A:187:HIS:HB3	A:203:TYR:CE2	0.420
3	A:11:PHE:CD2	A:12:PRO:HA	0.408
3	A:245:LEU:HD21	A:404:THR:HG22	0.407
3	A:380:GLY:HA3	A:383:GLN:NE2	0.407
3	A:198:LEU:HD23	A:228:MET:HE1	0.405
4	A:235:ARG:HH12	A:240:ARG:CZ	0.520
4	A:206:ASP:HA	A:240:ARG:HH21	0.517
4	A:14:LEU:HD21	A:94:PRO:HG3	0.478
4	A:235:ARG:HH12	A:240:ARG:NH1	0.449
4	A:246:ASN:ND2	A:379:CYS:SG	0.441
4	A:189:LEU:HG	A:200:ALA:HB2	0.421
4	A:248:VAL:CG2	A:384:LEU:HB3	0.420

Model ID	Atom-1	Atom-2	Clash overlap (Å)
4	A:215:TYR:CE2	A:219:ILE:CG2	0.418
4	A:384:LEU:HD21	A:404:THR:HG21	0.418
4	A:11:PHE:CD2	A:12:PRO:HA	0.417
4	A:421:PHE:CZ	A:438:ALA:HA	0.407
5	A:260:LEU:HD13	A:358:ALA:HA	0.655
5	A:429:TRP:CH2	A:433:ARG:NH2	0.523
5	A:384:LEU:HD21	A:404:THR:HG21	0.510
5	A:215:TYR:CE2	A:219:ILE:CG2	0.440
5	A:301:LEU:HD12	A:335:GLN:HG2	0.440
5	A:215:TYR:CE2	A:219:ILE:HG23	0.413
5	A:397:ARG:HE	A:424:GLU:CD	0.412
5	A:314:LEU:HD13	A:324:LEU:HD22	0.402
5	A:115:TRP:NE1	A:150:ARG:HH22	0.401
6	A:115:TRP:CD1	A:150:ARG:HH22	0.581
6	A:199:LYS:HG3	A:228:MET:HE1	0.533
6	A:270:LEU:HD21	A:456:MET:SD	0.455
6	A:22:VAL:HG21	A:161:HIS:HE1	0.436
6	A:90:ARG:NH2	A:109:ARG:HD2	0.431
6	A:25:ALA:HB1	A:462:TRP:CD1	0.425
6	A:171:ALA:HB2	A:184:PHE:CZ	0.424
6	A:2:GLN:HE21	A:128:TRP:CD1	0.417
6	A:215:TYR:CE2	A:219:ILE:CG2	0.416
6	A:384:LEU:HD21	A:404:THR:HG21	0.413

Model ID	Atom-1	Atom-2	Clash overlap (Å)
6	A:90:ARG:HH22	A:105:ASP:CG	0.407
6	A:245:LEU:HD21	A:404:THR:HG22	0.402
7	A:260:LEU:HD13	A:358:ALA:HA	0.505
7	A:115:TRP:NE1	A:150:ARG:HH22	0.466
7	A:421:PHE:CZ	A:438:ALA:HA	0.428
7	A:2:GLN:HE21	A:128:TRP:CD1	0.416
8	A:260:LEU:HD13	A:358:ALA:HA	0.550
8	A:119:GLU:CD	A:152:ARG:HH22	0.475
8	A:2:GLN:HE21	A:128:TRP:CD1	0.422
8	A:179:LEU:HD23	A:184:PHE:CZ	0.420
8	A:186:ILE:HG21	A:203:TYR:CZ	0.420
8	A:14:LEU:HD21	A:94:PRO:HD3	0.418
8	A:384:LEU:HD21	A:404:THR:HG21	0.413
8	A:118:ALA:HB1	A:152:ARG:HH11	0.409
8	A:210:ALA:HB1	A:215:TYR:CD1	0.409
8	A:421:PHE:CZ	A:438:ALA:HA	0.407
9	A:179:LEU:HD23	A:184:PHE:CE2	0.687
9	A:407:ASP:CG	A:437:ARG:HH22	0.489
9	A:174:MET:SD	A:184:PHE:CE2	0.474
9	A:309:LEU:HD21	A:375:ARG:NH2	0.464
9	A:260:LEU:HD13	A:358:ALA:HA	0.462
9	A:292:VAL:HG13	A:321:GLY:O	0.435
9	A:146:TYR:CD1	A:179:LEU:HD13	0.429

Model ID	Atom-1	Atom-2	Clash overlap (Å)
9	A:380:GLY:CA	A:383:GLN:HE22	0.416
9	A:171:ALA:HA	A:195:ILE:CD1	0.411
9	A:206:ASP:HA	A:240:ARG:HH21	0.409
10	A:379:CYS:SG	A:401:LEU:HD21	0.707
10	A:202:LEU:HD12	A:228:MET:HE3	0.701
10	A:260:LEU:HD13	A:358:ALA:HA	0.592
10	A:199:LYS:HA	A:228:MET:HE1	0.579
10	A:199:LYS:CB	A:228:MET:HE1	0.562
10	A:206:ASP:HA	A:240:ARG:HH21	0.553
10	A:199:LYS:CA	A:228:MET:HE1	0.531
10	A:14:LEU:HD21	A:94:PRO:HG3	0.489
10	A:115:TRP:CD1	A:150:ARG:HH22	0.476
10	A:2:GLN:HE21	A:128:TRP:CD1	0.475
10	A:379:CYS:HB2	A:401:LEU:HG	0.429
10	A:309:LEU:HD21	A:375:ARG:NH2	0.427
10	A:421:PHE:CZ	A:438:ALA:HA	0.413
10	A:161:HIS:O	A:378:PRO:HG2	0.407
10	A:210:ALA:HB1	A:215:TYR:CD1	0.401

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	Analyzed	Favored	Allowed	Outliers
1	475	460	15	0
2	475	460	15	0

Model ID	Analyzed	Favored	Allowed	Outliers
3	475	461	14	0
4	475	458	15	2
5	475	461	12	2
6	475	458	14	3
7	475	458	15	2
8	475	462	11	2
9	475	452	20	3
10	475	457	17	1

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	Analyzed	Favored	Allowed	Outliers
1	382	377	4	1
2	382	375	6	1
3	382	377	5	0
4	382	373	7	2
5	382	374	8	0
6	382	376	5	1
7	382	378	4	0
8	382	374	6	2
9	382	373	8	1
10	382	375	6	1

Detailed list of outliers are tabulated below.

Model ID	Chain	Residue ID	Residue type

Model ID	Chain	Residue ID	Residue type
1	A	463	GLN
2	A	209	THR
4	A	139	HIS
4	A	212	SER
6	A	209	THR
8	A	137	ASP
8	A	271	GLU
9	A	181	TRP
10	A	137	ASP

Fit of model to data used for modeling

Crosslinking-MS

Validation for this section is under development.

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 labcontributed model validation metrics and software packages.

Implementation of validation methods for SAS data and SAS-based models are funded byRCSB PDB (grant number DBI-1832184). Dr. Stephen Burley, Dr. John Westbrook, and Dr. Jasmine Young from RCSB PDB, Dr. Jill Trewhella, 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.