

Integrative Structure Validation Report

July 22, 2024 - 05:00 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	9A2J
PDB-Dev ID	PDBDEV_00000168
Structure Title	Model of E. coli RplJ 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

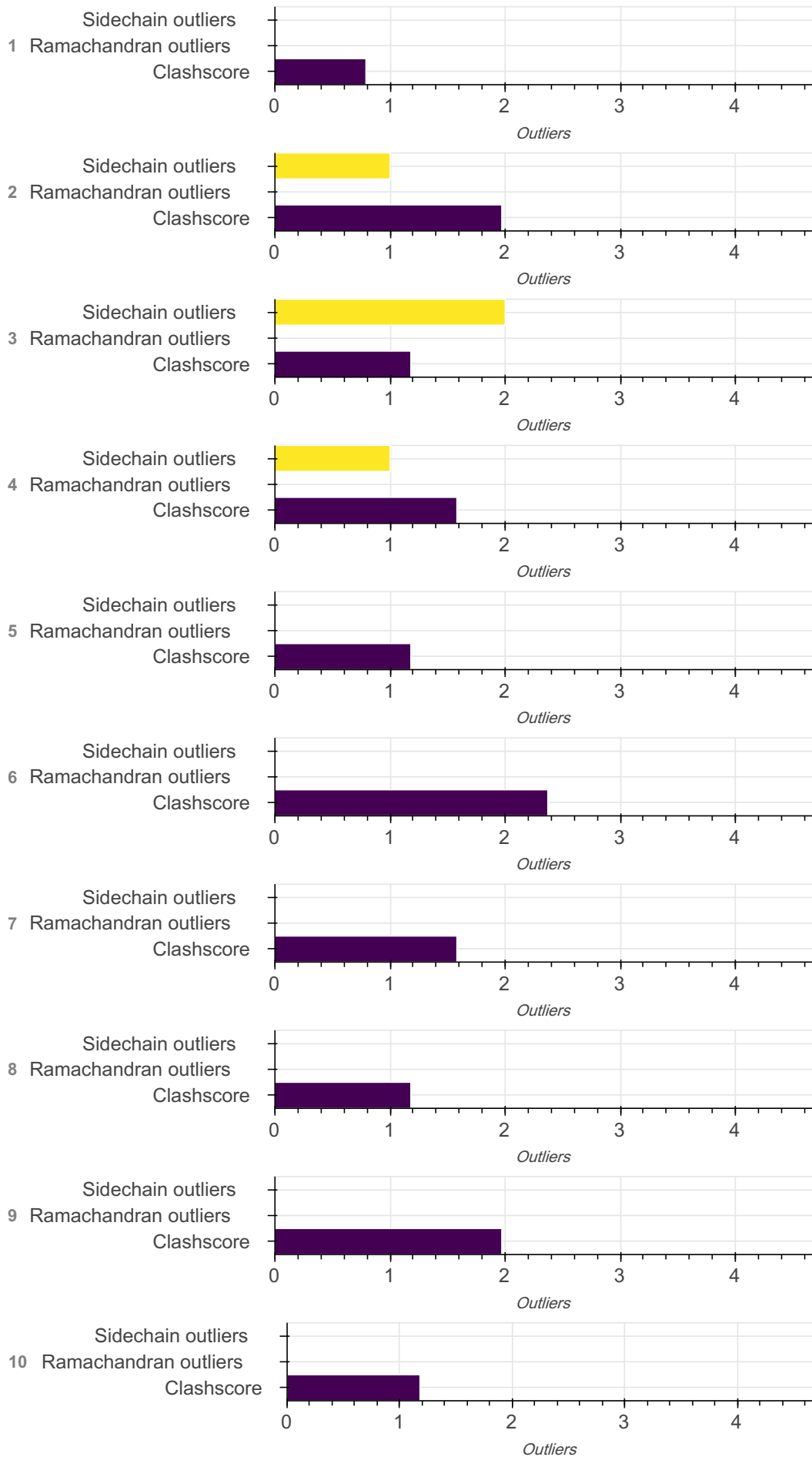
A user guide is available at https://pdb-dev.wwpdb.org/validation_help.html with specific help available everywhere you see the  symbol.

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



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	P0A7J3	A	A	165
2	1	1	P0A7J3	A	A	165
3	1	1	P0A7J3	A	A	165
4	1	1	P0A7J3	A	A	165
5	1	1	P0A7J3	A	A	165
6	1	1	P0A7J3	A	A	165
7	1	1	P0A7J3	A	A	165
8	1	1	P0A7J3	A	A	165
9	1	1	P0A7J3	A	A	165
10	1	1	P0A7J3	A	A	165

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	-	1-165

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

Standard geometry: bond outliers?

There are 12920 bond outliers in this entry. A summary is provided below, and a detailed list of outliers can be found [here](#).

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CB--HB2	1.09	0.97	1270
CA--HA	1.09	0.97	1560
NZ--HZ3	1.01	0.89	120
CG2--HG21	1.09	0.97	290
OH--HH	0.96	0.84	30
CD2--HD22	1.09	0.97	150
CG--HG2	1.09	0.97	530
CG--HG3	1.09	0.97	530
CB--HB1	1.09	0.97	330
CD1--HD11	1.09	0.97	200
CG2--HG23	1.09	0.97	290
CB--HB	1.09	0.97	290
CB--HB3	1.09	0.97	1270
CG1--HG12	1.09	0.97	200
CG--HG	1.09	0.97	150
CD--HD2	1.09	0.97	290
CD2--HD23	1.09	0.97	150
CG2--HG22	1.09	0.97	290
CD1--HD12	1.09	0.97	200
CG1--HG13	1.09	0.97	200
CD--HD3	1.09	0.97	290
CD1--HD13	1.09	0.97	200

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CD2--HD21	1.09	0.97	150
NZ--HZ2	1.01	0.89	120
CE--HE2	1.09	0.97	180
NZ--HZ1	1.01	0.89	120
N--H1	1.01	0.89	10
CE--HE3	1.09	0.97	180
CA--HA3	1.09	0.97	90
CG1--HG11	1.09	0.97	150
CE--HE1	1.09	0.97	60
OG1--HG1	0.96	0.84	90
OG--HG	0.96	0.84	60
CA--HA2	1.09	0.97	90
N--H2	1.01	0.89	10
N--H3	1.01	0.89	10
SG--HG	1.33	1.20	2
SG--HG	1.34	1.20	8
N--H	1.01	0.86	1590
ND2--HD22	1.01	0.86	30
NE--HE	1.01	0.86	120
NH2--HH21	1.01	0.86	120
CE2--HE2	1.08	0.93	90
CD1--HD1	1.08	0.93	90
NE2--HE22	1.01	0.86	30

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CZ--HZ	1.08	0.93	60
CE1--HE1	1.08	0.93	100
CD2--HD2	1.08	0.93	100
ND1--HD1	1.01	0.86	10
NH1--HH12	1.01	0.86	120
NE2--HE21	1.01	0.86	30
NH2--HH22	1.01	0.86	120
ND2--HD21	1.01	0.86	30
NH1--HH11	1.01	0.86	120

Standard geometry: angle outliers

There are 53 angle outliers in this entry. A summary is provided below, and a detailed list of outliers can be found [here](#).

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
NE-CZ-NH2	119.20	123.98	1
CA-CB-CG	112.60	117.43	2
CA-CB-CG	112.60	117.39	1
CB-CG-CD2	131.20	125.06	1
CA-CB-CG	112.60	117.31	1
CA-CB-CG	112.60	117.30	1
CA-CB-CG	112.60	117.25	1
CB-CG-CD2	131.20	125.29	1
OE1-CD-NE2	122.60	118.10	1
CA-CB-CG	112.60	117.09	1
OE1-CD-NE2	122.60	118.13	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
CA-CB-CG	112.60	117.01	1
OE1-CD-NE2	122.60	118.20	1
CG-CD-CE	111.30	101.24	1
OE1-CD-NE2	122.60	118.24	2
OE1-CD-NE2	122.60	118.26	1
NH1-CZ-NH2	119.30	113.71	1
OE1-CD-NE2	122.60	118.31	1
OE1-CD-NE2	122.60	118.33	1
OE1-CD-NE2	122.60	118.37	1
OE1-CD-NE2	122.60	118.38	2
OE1-CD-NE2	122.60	118.39	3
CB-CG-CD2	131.20	125.73	1
OE1-CD-NE2	122.60	118.42	3
OE1-CD-NE2	122.60	118.43	1
OE1-CD-NE2	122.60	118.44	1
NE-CZ-NH2	119.20	122.93	1
OE1-CD-NE2	122.60	118.45	1
CB-CG-CD2	131.20	125.82	1
OE1-CD-NE2	122.60	118.48	1
CA-CB-CG	113.80	117.91	1
CB-CG-CD2	131.20	125.89	1
OE1-CD-NE2	122.60	118.53	2
CB-CG-CD2	131.20	125.92	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
OE1-CD-NE2	122.60	118.54	1
OE1-CD-NE2	122.60	118.55	2
OD1-CG-ND2	122.60	118.57	1
OE1-CD-NE2	122.60	118.58	1
OE1-CD-NE2	122.60	118.60	1
HH11-NH1-HH12	106.33	120.00	1
HH11-NH1-HH12	105.68	120.00	1
HH11-NH1-HH12	104.83	120.00	1
HH11-NH1-HH12	104.78	120.00	1
HH11-NH1-HH12	104.16	120.00	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 the models in this entry.

Model ID	Clash score	Number of clashes
1	0.79	2
2	1.97	5
3	1.18	3
4	1.58	4
5	1.18	3
6	2.37	6
7	1.58	4
8	1.18	3
9	1.97	5

Model ID	Clash score	Number of clashes
10	1.18	3

All 38 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:83:ALA:HB2	A:96:PHE:CZ	0.487
1	A:88:HIS:CD2	A:131:THR:HA	0.460
2	A:132:TYR:CZ	A:136:ILE:HD11	0.516
2	A:42:ARG:HA	A:52:MET:HE3	0.490
2	A:55:VAL:HG23	A:60:LEU:HG	0.476
2	A:29:ASP:HB2	A:109:LYS:HE3	0.463
2	A:41:LEU:O	A:52:MET:HE1	0.413
3	A:132:TYR:CE1	A:136:ILE:HD11	0.486
3	A:83:ALA:HB2	A:96:PHE:CZ	0.453
3	A:140:MET:SD	A:143:MET:CE	0.425
4	A:83:ALA:HB2	A:96:PHE:CZ	0.493
4	A:14:GLU:CD	A:53:ARG:HH12	0.488
4	A:126:LEU:HD12	A:129:LEU:HD12	0.430
4	A:132:TYR:CE1	A:136:ILE:HD11	0.418
5	A:132:TYR:CE1	A:136:ILE:HD11	0.540
5	A:83:ALA:HB2	A:96:PHE:CZ	0.466
5	A:29:ASP:HB2	A:109:LYS:HE3	0.420
6	A:143:MET:HG2	A:151:LEU:HD22	0.529
6	A:29:ASP:HB2	A:109:LYS:HE3	0.509
6	A:45:GLY:HA3	A:52:MET:HE2	0.481

Model ID	Atom-1	Atom-2	Clash overlap (Å)
6	A:143:MET:HG2	A:151:LEU:CD2	0.464
6	A:83:ALA:HB2	A:96:PHE:CZ	0.441
6	A:42:ARG:HA	A:52:MET:HE3	0.411
7	A:88:HIS:CD2	A:131:THR:HA	0.547
7	A:132:TYR:CE1	A:136:ILE:HD11	0.500
7	A:83:ALA:HB2	A:96:PHE:CZ	0.424
7	A:43:LYS:HE2	A:47:GLU:OE1	0.404
8	A:83:ALA:HB2	A:96:PHE:CZ	0.456
8	A:43:LYS:HE2	A:47:GLU:OE1	0.412
8	A:42:ARG:HA	A:52:MET:CE	0.411
9	A:143:MET:HA	A:151:LEU:HD11	0.534
9	A:29:ASP:HB2	A:109:LYS:HE3	0.453
9	A:45:GLY:HA3	A:52:MET:HE2	0.452
9	A:83:ALA:HB2	A:96:PHE:CZ	0.430
9	A:143:MET:HG2	A:151:LEU:HD22	0.405
10	A:83:ALA:HB2	A:96:PHE:CZ	0.503
10	A:29:ASP:HB2	A:109:LYS:HE3	0.458
10	A:132:TYR:CE1	A:136:ILE:HD11	0.407

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	163	162	1	0
2	163	162	1	0

Model ID	Analyzed	Favored	Allowed	Outliers
3	163	162	1	0
4	163	162	1	0
5	163	163	0	0
6	163	162	1	0
7	163	161	2	0
8	163	162	1	0
9	163	163	0	0
10	163	163	0	0

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	123	120	3	0
2	123	119	3	1
3	123	118	3	2
4	123	119	3	1
5	123	120	3	0
6	123	120	3	0
7	123	122	1	0
8	123	121	2	0
9	123	118	5	0
10	123	118	5	0

Detailed list of outliers are tabulated below.

Model ID	Chain	Residue ID	Residue type
----------	-------	------------	--------------

Model ID	Chain	Residue ID	Residue type
2	A	55	VAL
3	A	5	LEU
3	A	154	THR
4	A	154	THR

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 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.