

# Integrative Structure Validation Report

July 22, 2024 - 05:16 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	9A3A
PDB-Dev ID	PDBDEV_00000195
Structure Title	Model of E. coli MlaD 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](mailto:pdb-dev@mail.wwpdb.org)*

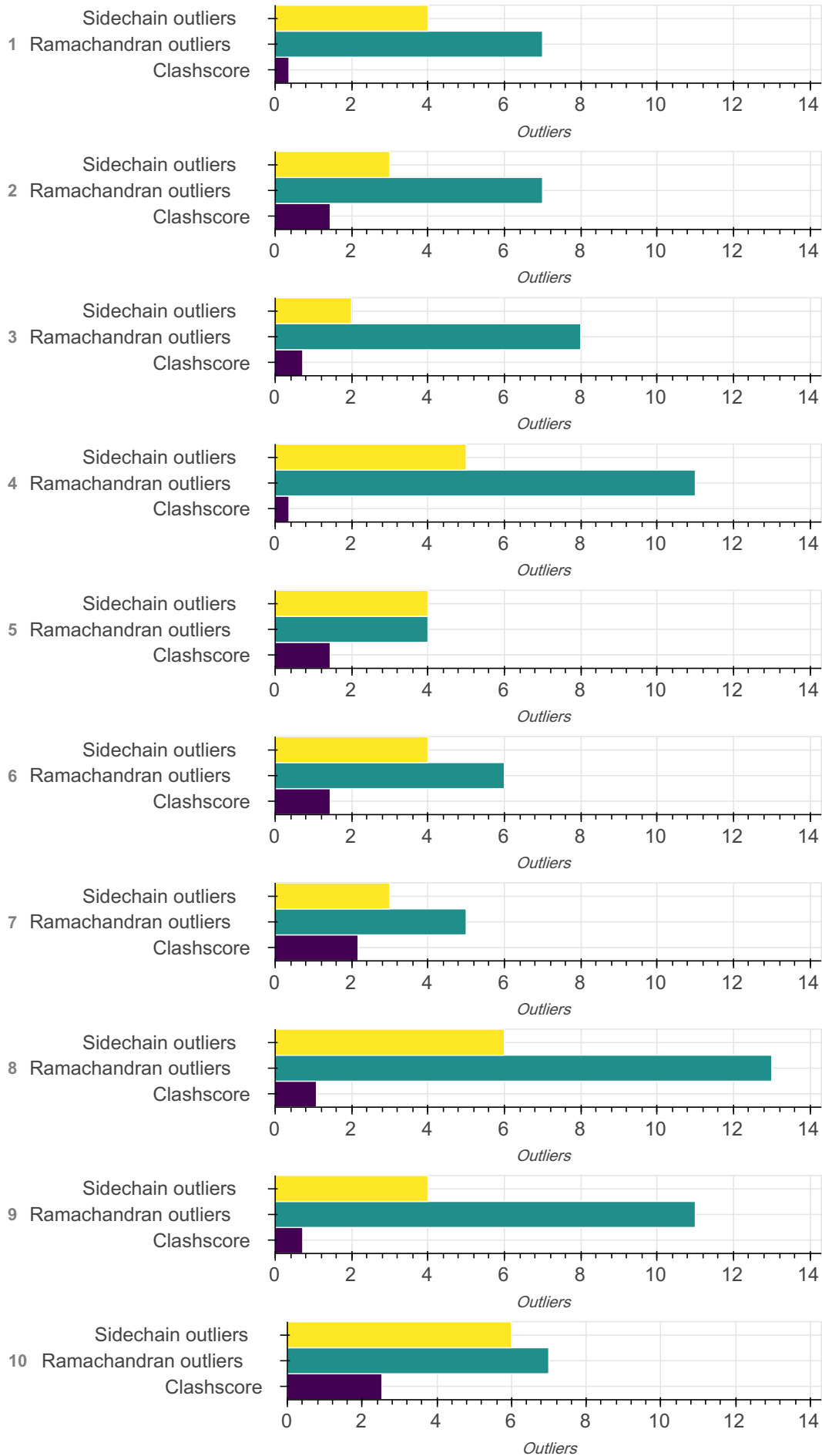
*A user guide is available at [https://pdb-dev.wwpdb.org/validation\\_help.html](https://pdb-dev.wwpdb.org/validation_help.html) with specific help available everywhere you see the  symbol.*

*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	P64604	A	A	183
2	1	1	P64604	A	A	183
3	1	1	P64604	A	A	183
4	1	1	P64604	A	A	183
5	1	1	P64604	A	A	183
6	1	1	P64604	A	A	183
7	1	1	P64604	A	A	183
8	1	1	P64604	A	A	183
9	1	1	P64604	A	A	183
10	1	1	P64604	A	A	183

### 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	-	1-183

### 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	<a href="#">AlphaLink</a>	1.0	model building	<a href="https://github.com/lhatsk/AlphaLink">https://github.com/lhatsk/AlphaLink</a>

### Data quality ?

#### Crosslinking-MS

Validation for this section is under development.

### 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?

There are 13920 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
CD1--HD12	1.09	0.97	330
CB--HB3	1.09	0.97	1240
CG--HG2	1.09	0.97	420
CA--HA3	1.09	0.97	170
CB--HB2	1.09	0.97	1240
CA--HA	1.09	0.97	1660
CD2--HD22	1.09	0.97	210
CG--HG3	1.09	0.97	420
CG2--HG23	1.09	0.97	420
CG--HG	1.09	0.97	210
CG2--HG22	1.09	0.97	420
CG1--HG13	1.09	0.97	240
CA--HA2	1.09	0.97	170
OG1--HG1	0.96	0.84	180
CD--HD3	1.09	0.97	250
OG--HG	0.96	0.84	100
N--H1	1.01	0.89	10
CD2--HD23	1.09	0.97	210
CG2--HG21	1.09	0.97	420
CD2--HD21	1.09	0.97	210
CD1--HD13	1.09	0.97	330
NZ--HZ2	1.01	0.89	100

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CE--HE3	1.09	0.97	120
CD--HD2	1.09	0.97	250
CE--HE2	1.09	0.97	120
CG1--HG11	1.09	0.97	120
CB--HB1	1.09	0.97	160
CB--HB	1.09	0.97	420
CG1--HG12	1.09	0.97	240
CD1--HD11	1.09	0.97	330
NZ--HZ1	1.01	0.89	100
N--H3	1.01	0.89	10
OH--HH	0.96	0.84	60
NZ--HZ3	1.01	0.89	100
CE--HE1	1.09	0.97	20
N--H2	1.01	0.89	10
SG--HG	1.33	1.20	3
SG--HG	1.34	1.20	7
N--H	1.01	0.86	1730
CD2--HD2	1.08	0.93	120
NE2--HE22	1.01	0.86	50
CD1--HD1	1.08	0.93	120
NH1--HH11	1.01	0.86	60
CZ--HZ	1.08	0.93	50
ND2--HD21	1.01	0.86	90

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CH2--HH2	1.08	0.93	10
NH2--HH21	1.01	0.86	60
CE2--HE2	1.08	0.93	110
ND2--HD22	1.01	0.86	90
NH1--HH12	1.01	0.86	60
CE1--HE1	1.08	0.93	120
NH2--HH22	1.01	0.86	60
NE--HE	1.01	0.86	60
NE1--HE1	1.01	0.86	10
CE3--HE3	1.08	0.93	10
ND1--HD1	1.01	0.86	10
CZ2--HZ2	1.08	0.93	10
CZ3--HZ3	1.08	0.93	10
NE2--HE21	1.01	0.86	50

### Standard geometry: angle outliers

There are 102 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
C-N-CA	121.70	135.70	1
C-N-CA	121.70	134.10	1
CA-CB-CG	112.60	119.47	1
C-N-CA	121.70	133.59	1
OE1-CD-NE2	122.60	116.06	1
C-N-CA	121.70	132.56	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
CA-CB-CG	112.60	118.54	1
C-N-CA	121.70	132.39	1
C-N-CA	121.70	132.05	1
CA-CB-CG	112.60	118.34	1
C-N-CA	121.70	131.86	1
OD1-CG-ND2	122.60	117.00	1
C-N-CA	121.70	131.60	1
OE1-CD-NE2	122.60	117.11	1
C-N-CA	121.70	131.55	1
CA-CB-CG	112.60	118.03	1
C-N-CA	121.70	131.44	1
C-N-CA	121.70	131.43	1
C-N-CA	121.70	131.17	1
OE1-CD-NE2	122.60	117.36	1
C-N-CA	121.70	131.07	1
CA-N-CD	112.00	104.72	1
CA-N-CD	112.00	104.83	1
C-N-CA	121.70	130.64	1
CA-CB-CG	112.60	117.52	1
C-N-CA	121.70	130.47	1
C-N-CA	121.70	130.39	2
CA-CB-OG1	109.60	116.83	1
CA-CB-CG	112.60	117.41	1



Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
CA-CB-CG	112.60	117.40	1
C-N-CA	121.70	130.32	1
OD1-CG-ND2	122.60	117.81	1
CD-NE-CZ	124.40	131.10	1
C-N-CA	121.70	130.29	1
OE1-CD-NE2	122.60	117.87	1
CA-CB-OG1	109.60	116.68	1
OE1-CD-NE2	122.60	117.89	1
CA-CB-CG2	110.50	118.49	1
OE1-CD-NE2	122.60	117.91	1
CA-CB-CG2	110.50	118.44	1
CA-CB-OG1	109.60	116.60	1
C-N-CA	121.70	130.07	1
NE-CZ-NH2	119.20	123.38	1
CA-CB-CG	112.60	117.23	1
OE1-CD-NE2	122.60	118.03	1
OE1-CD-NE2	122.60	118.08	1
OE1-CD-NE2	122.60	118.10	1
NE-CZ-NH1	121.50	125.95	1
OD1-CG-ND2	122.60	118.16	1
C-N-CA	121.70	129.63	1
OE1-CD-NE2	122.60	118.20	1
OD1-CG-ND2	122.60	118.20	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
CA-CB-CG	112.60	116.97	1
OD1-CG-ND2	122.60	118.23	1
C-N-CA	121.70	129.53	1
OD1-CG-ND2	122.60	118.25	1
OE1-CD-NE2	122.60	118.26	2
C-N-CA	121.70	129.47	1
OE1-CD-NE2	122.60	118.28	1
NE-CZ-NH1	121.50	125.81	1
OE1-CD-NE2	122.60	118.29	1
C-N-CA	121.70	129.46	1
OE1-CD-NE2	122.60	118.31	1
O-C-N	123.00	116.16	1
OE1-CD-NE2	122.60	118.33	1
OE1-CD-NE2	122.60	118.34	1
C-N-CA	121.70	129.36	1
C-N-CA	121.70	129.35	1
OD1-CG-ND2	122.60	118.35	1
OE1-CD-NE2	122.60	118.37	1
OE1-CD-NE2	122.60	118.38	1
OD1-CG-ND2	122.60	118.38	2
OE1-CD-NE2	122.60	118.40	1
OE1-CD-NE2	122.60	118.42	1
OE1-CD-NE2	122.60	118.44	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
N-CA-C	111.00	122.63	1
C-N-CA	121.70	129.13	1
OE1-CD-NE2	122.60	118.48	2
C-CA-CB	110.10	117.92	1
OD1-CG-ND2	122.60	118.49	2
OE1-CD-NE2	122.60	118.50	1
CD-NE-CZ	124.40	130.09	1
OE1-CD-NE2	122.60	118.54	1
OE1-CD-NE2	122.60	118.56	1
C-N-CA	121.70	128.94	1
OE1-CD-NE2	122.60	118.58	1
CA-CB-CG2	110.40	117.23	1
OE1-CD-NE2	122.60	118.59	1
C-N-CA	121.70	128.90	1
CZ-NH2-HH21	107.38	120.00	1
HH11-NH1-HH12	107.14	120.00	1
C-N-H	109.54	124.30	1
CZ-NH2-HH21	104.69	120.00	1
C-N-H	108.05	124.30	1
C-N-H	105.52	124.30	1
HH21-NH2-HH22	99.70	120.00	1
HH21-NH2-HH22	99.01	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	0.36	1
2	1.44	4
3	0.72	2
4	0.36	1
5	1.44	4
6	1.44	4
7	2.17	6
8	1.08	3
9	0.72	2
10	2.53	7

All 34 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:141:MET:SD	A:146:LEU:HD21	0.411
2	A:35:ARG:C	A:37:GLU:H	0.457
2	A:42:LEU:HD11	A:128:LEU:HD13	0.419
2	A:46:PHE:CE1	A:137:THR:HB	0.418
2	A:116:VAL:CG1	A:128:LEU:HD11	0.405
3	A:35:ARG:HH21	A:39:THR:HG21	0.500
3	A:141:MET:SD	A:146:LEU:HD21	0.441
4	A:93:ILE:HD13	A:114:LEU:HD21	0.569
5	A:37:GLU:CD	A:89:ARG:HH21	0.550
5	A:93:ILE:HD13	A:114:LEU:HD21	0.467

Model ID	Atom-1	Atom-2	Clash overlap (Å)
5	A:35:ARG:C	A:37:GLU:H	0.456
5	A:141:MET:SD	A:146:LEU:HD21	0.433
6	A:35:ARG:HH12	A:85:GLU:CB	0.482
6	A:35:ARG:HH12	A:85:GLU:HB3	0.464
6	A:93:ILE:HD13	A:114:LEU:HD21	0.440
6	A:60:ILE:HD11	A:114:LEU:HD23	0.434
7	A:42:LEU:HD11	A:128:LEU:HD13	0.436
7	A:116:VAL:HG13	A:128:LEU:HD11	0.434
7	A:125:THR:C	A:127:ILE:H	0.422
7	A:119:GLU:HA	A:135:GLN:HE21	0.420
7	A:46:PHE:CD1	A:137:THR:HB	0.409
7	A:107:LEU:CD1	A:143:LEU:HD22	0.401
8	A:93:ILE:HD13	A:114:LEU:HD21	0.529
8	A:30:ASN:CG	A:55:ARG:HH22	0.413
8	A:157:ASP:C	A:159:ASN:H	0.412
9	A:28:ALA:HA	A:31:VAL:HG22	0.496
9	A:35:ARG:HH12	A:85:GLU:CD	0.469
10	A:27:LYS:O	A:31:VAL:HG23	0.592
10	A:116:VAL:CG1	A:128:LEU:HD21	0.576
10	A:142:VAL:HG12	A:144:GLU:OE2	0.557
10	A:124:GLY:O	A:125:THR:HG23	0.484
10	A:33:SER:HB3	A:67:ARG:HH12	0.475
10	A:46:PHE:CE1	A:137:THR:HB	0.417

Model ID	Atom-1	Atom-2	Clash overlap (Å)
10	A:52:LEU:HD13	A:103:THR:HG22	0.408

### 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	181	163	11	7
2	181	158	16	7
3	181	156	17	8
4	181	155	15	11
5	181	167	10	4
6	181	163	12	6
7	181	165	11	5
8	181	149	19	13
9	181	156	14	11
10	181	163	11	7

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	150	139	7	4
2	150	138	9	3
3	150	145	3	2
4	150	138	7	5
5	150	139	7	4

Model ID	Analyzed	Favored	Allowed	Outliers
6	150	138	8	4
7	150	139	8	3
8	150	138	6	6
9	150	140	6	4
10	150	139	5	6

*Detailed list of outliers are tabulated below.*

Model ID	Chain	Residue ID	Residue type
1	A	3	THR
1	A	36	THR
1	A	73	LEU
1	A	172	ASN
2	A	31	VAL
2	A	181	THR
2	A	182	THR
3	A	118	PHE
3	A	181	THR
4	A	3	THR
4	A	18	LEU
4	A	32	THR
4	A	89	ARG
4	A	181	THR
5	A	18	LEU
5	A	125	THR

Model ID	Chain	Residue ID	Residue type
5	A	181	THR
5	A	182	THR
6	A	31	VAL
6	A	36	THR
6	A	181	THR
6	A	182	THR
7	A	3	THR
7	A	123	LEU
7	A	182	THR
8	A	1	MET
8	A	3	THR
8	A	26	LEU
8	A	158	ASP
8	A	172	ASN
8	A	182	THR
9	A	3	THR
9	A	36	THR
9	A	123	LEU
9	A	182	THR
10	A	3	THR
10	A	14	LEU
10	A	18	LEU
10	A	31	VAL



Model ID	Chain	Residue ID	Residue type
10	A	142	VAL
10	A	181	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.*