Integrative Structure Validation Report

September 11, 2024 - 01:18 AM 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	9A7P
PDB-Dev ID	PDBDEV_00000354
Structure Title	Integrative model of SIGA-RPOB by crosslinking MS and deep learning
Structure Authors	Kolja Stahl; Oliver Brock; Juri Rappsilber

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

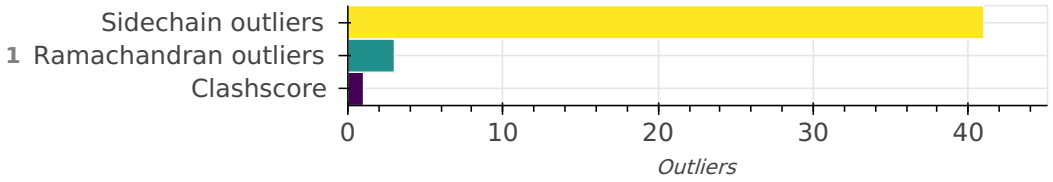
A user guide is available at https://pdb-dev.wwpdb.org/validation_help.html with specific help available everywhere you see the ? symbol.

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 1 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/None.

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Chain ID [auth]	Total residues
1	1	1	SIGA_BACSU	A	A	371
1	2	2	RPOB_BACSU	B	B	1193

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	PRIDE	PXD035508

Representation

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

Chain ID	Rigid bodies	Non-rigid segments

Chain ID	Rigid bodies	Non-rigid segments
A	-	1-371
B	-	1-1193

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	AlphaLink2	AlphaLink2	None	1	False	False

There is 1 software package reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	AlphaLink2	1.0	model building	https://github.com/Rappsilber-Laboratory/AlphaLink2

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

There are 12330 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	1140

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CG--HG	1.09	0.97	140
OG1--HG1	0.96	0.84	77
CG2--HG22	1.09	0.97	303
CD1--HD12	1.09	0.97	243
CA--HA	1.09	0.97	1443
CE--HE1	1.09	0.97	40
CG--HG2	1.09	0.97	527
CG--HG3	1.09	0.97	527
NZ--HZ3	1.01	0.89	85
CD--HD2	1.09	0.97	266
CD--HD3	1.09	0.97	266
CG1--HG11	1.09	0.97	123
CE--HE2	1.09	0.97	125
CB--HB3	1.09	0.97	1140
CG2--HG23	1.09	0.97	303
CD1--HD13	1.09	0.97	243
CD2--HD21	1.09	0.97	140
CB--HB1	1.09	0.97	88
CD2--HD23	1.09	0.97	140
CA--HA2	1.09	0.97	121
CG1--HG13	1.09	0.97	226
CG2--HG21	1.09	0.97	303
CD2--HD22	1.09	0.97	140

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CD1--HD11	1.09	0.97	243
CB--HB	1.09	0.97	303
CE--HE3	1.09	0.97	125
OG--HG	0.96	0.84	78
CA--HA3	1.09	0.97	121
CG1--HG12	1.09	0.97	226
NZ--HZ2	1.01	0.89	85
OH--HH	0.96	0.84	51
NZ--HZ1	1.01	0.89	85
N--H2	1.01	0.89	2
N--H3	1.01	0.89	2
N--H1	1.01	0.89	2
SG--HG	1.34	1.20	4
N--H	1.01	0.86	1495
NE2--HE21	1.01	0.86	58
CD2--HD2	1.08	0.93	120
NH2--HH21	1.01	0.86	114
NE--HE	1.01	0.86	114
ND2--HD21	1.01	0.86	60
NH1--HH12	1.01	0.86	114
ND2--HD22	1.01	0.86	60
NH1--HH11	1.01	0.86	114
ND1--HD1	1.01	0.86	20
NH2--HH22	1.01	0.86	114

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CD1--HD1	1.08	0.93	106
CE1--HE1	1.08	0.93	120
CE2--HE2	1.08	0.93	98
NE2--HE22	1.01	0.86	58
CE3--HE3	1.08	0.93	8
CZ2--HZ2	1.08	0.93	8
CZ--HZ	1.08	0.93	47
NE1--HE1	1.01	0.86	8
CZ3--HZ3	1.08	0.93	8
NE2--HE2	1.01	0.86	2
CH2--HH2	1.08	0.93	8

Standard geometry: angle outliers

There are 56 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
CA-CB-CG	112.60	118.51	1
C-N-CA	121.70	131.68	1
CA-CB-CG	112.60	117.99	1
OE1-CD-NE2	122.60	117.26	1
CA-CB-CG	112.60	117.85	1
N-CA-C-N-OD1-CG-ND2	122.60	117.57	1
OE1-CD-NE2	122.60	117.71	1
NE-CZ-NH1	121.50	126.39	1
OE1-CD-NE2	122.60	117.84	1
OE1-CD-NE2	122.60	117.87	2

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
CA-CB-CA-CB-CG	112.60	117.30	1
CA-CB-CG	112.60	117.27	2
OE1-CD-NE2	122.60	117.94	2
CA-CB-CA-CB-CG	112.60	117.25	1
N-CA-OE1-CD-NE2	122.60	118.01	1
OE1-CD-NE2	122.60	118.02	1
OE1-CD-NE2	122.60	118.04	1
OE1-CD-NE2	122.60	118.06	1
N-CA-OE1-CD-NE2	122.60	118.17	1
CA-CB-CG	112.60	116.99	1
CA-CB-CG	112.60	116.97	1
OE1-CD-NE2	122.60	118.24	1
OE1-CD-NE2	122.60	118.25	1
OE1-CD-NE2	122.60	118.26	1
OE1-CD-NE2	122.60	118.27	1
CA-CB-CG	112.60	116.91	1
OE1-CD-NE2	122.60	118.30	1
CB-CG-CD2	131.20	125.62	1
OE1-CD-NE2	122.60	118.31	1
OE1-CD-NH1-CZ-NH2	119.30	113.77	1
OE1-CD-NE2	122.60	118.35	2
CA-CB-CG	112.60	116.84	1
OD1-CG-ND2	122.60	118.38	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
CB-CG-CD2	131.20	125.78	1
OE1-CD-NE2	122.60	118.46	2
NE-CZ-NH2	119.20	122.91	1
OE1-CD-NE2	122.60	118.48	1
N-CA-CA-CB-CG	112.60	116.69	1
OE1-CD-NE2	122.60	118.52	1
OE1-CD-NE2	122.60	118.54	1
OE1-CD-NE2	122.60	118.55	1
OD1-CG-ND2	122.60	118.56	1
CA-C-OD1-CG-ND2	122.60	118.57	1
OD1-CG-ND2	122.60	118.57	2
OE1-CD-NE2	122.60	118.58	1
C-N-H	111.79	124.30	1
C-N-H	111.77	124.30	1
HZ2-NZ-HZ3	96.33	109.00	1
HH21-NH2-HH22	106.81	120.00	1
HH21-NH2-HH22	96.73	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	1.01	25

All 25 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Model ID	Atom-1	Atom-2	Clash overlap (Å)

Model ID	Atom-1	Atom-2	Clash overlap (Å)
1	B:777:MET:HG3	B:781:LEU:HD12	0.585
1	A:70:THR:HG22	A:177:LYS:HZ1	0.579
1	B:471:MET:HE1	B:483:LYS:HE2	0.551
1	B:179:VAL:HG13	B:220:THR:HA	0.525
1	B:791:HIS:CE1	A:177:LYS:HZ1	0.504
1	A:70:THR:CG2	B:916:LYS:HE2	0.496
1	B:957:LEU:HD11	B:954:ASP:HB3	0.480
1	A:293:HIS:CD2	B:982:TYR:CZ	0.470
1	B:917:ARG:HH11	B:916:LYS:HE3	0.448
1	B:788:THR:HG21	B:374:LYS:HE2	0.440
1	B:781:LEU:HD11	B:118:ILE:HD11	0.433
1	B:978:MET:HE3	A:52:TYR:CD2	0.415
1	B:957:LEU:CD1	B:712:LEU:HD23	0.412
1	B:651:MET:CE	B:976:MET:HE3	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	1560	1528	29	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	Analyzed	Favored	Allowed	Outliers
1	1355	1259	55	41

Detailed list of outliers are tabulated below.

Model ID	Chain	Residue ID	Residue type
1	A	6	THR
1	A	9	THR
1	A	39	MET
1	A	288	THR
1	A	312	ASP
1	A	341	VAL
1	A	370	LEU
1	B	2	THR
1	B	48	MET
1	B	71	LEU
1	B	86	VAL
1	B	159	THR
1	B	161	THR
1	B	174	THR
1	B	189	LYS
1	B	192	VAL
1	B	255	LEU
1	B	319	LEU
1	B	353	THR
1	B	368	TYR
1	B	391	LEU
1	B	440	THR
1	B	578	ASN
1	B	590	ASP

Model ID	Chain	Residue ID	Residue type
1	B	753	LEU
1	B	831	ARG
1	B	838	ASP
1	B	846	VAL
1	B	852	THR
1	B	861	LEU
1	B	942	GLU
1	B	1038	ILE
1	B	1044	VAL
1	B	1048	LEU
1	B	1132	ILE
1	B	1169	LEU
1	B	1180	THR
1	B	1185	VAL
1	B	1186	GLU
1	B	1189	VAL
1	B	1191	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 labcontributed 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 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.