

Integrative Structure Validation Report

September 10, 2024 - 11:10 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	9A5J
PDB-Dev ID	PDBDEV_00000276
Structure Title	Integrative model of DNAK-GUDB 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

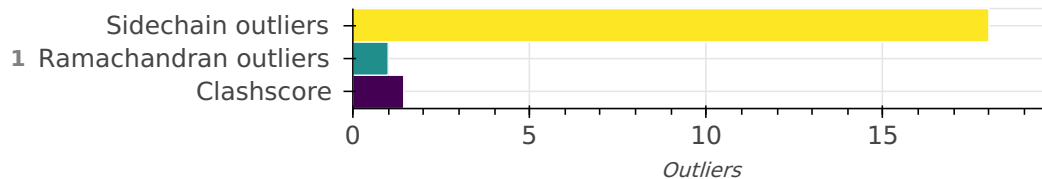
A user guide is available at https://pdb-dev.wwpdb.org/validation_help.html with specific help available everywhere you see the  symbol.

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	DNAK_BACSU	A	A	611
1	2	2	GUDB_BACSU	B	B	427

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	-	1-611
B	-	1-427

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

Standard geometry: bond outliers ?

There are 7985 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--HB1	1.09	0.97	91

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CA--HA	1.09	0.97	950
CG2--HG22	1.09	0.97	218
CD--HD2	1.09	0.97	161
CB--HB2	1.09	0.97	732
CG--HG2	1.09	0.97	314
CG--HG3	1.09	0.97	314
CB--HB3	1.09	0.97	732
CE--HE3	1.09	0.97	107
CB--HB	1.09	0.97	218
OH--HH	0.96	0.84	23
CG2--HG21	1.09	0.97	218
CD2--HD21	1.09	0.97	78
CD1--HD11	1.09	0.97	149
CG1--HG13	1.09	0.97	157
CD1--HD12	1.09	0.97	149
CG2--HG23	1.09	0.97	218
CD--HD3	1.09	0.97	161
CA--HA2	1.09	0.97	88
CD2--HD22	1.09	0.97	78
CE--HE2	1.09	0.97	107
CD2--HD23	1.09	0.97	78
CG1--HG12	1.09	0.97	157
CE--HE1	1.09	0.97	24

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
NZ--HZ1	1.01	0.89	83
OG1--HG1	0.96	0.84	61
CA--HA3	1.09	0.97	88
OG--HG	0.96	0.84	55
NZ--HZ2	1.01	0.89	83
CG1--HG11	1.09	0.97	86
NZ--HZ3	1.01	0.89	83
CD1--HD13	1.09	0.97	149
CG--HG	1.09	0.97	78
N--H3	1.01	0.89	2
N--H2	1.01	0.89	2
N--H1	1.01	0.89	2
SG--HG	1.33	1.20	2
SG--HG	1.34	1.20	3
CE1--HE1	1.08	0.93	65
N--H	1.01	0.86	1002
CD1--HD1	1.08	0.93	56
CD2--HD2	1.08	0.93	65
NH1--HH11	1.01	0.86	44
NH2--HH21	1.01	0.86	44
CE2--HE2	1.08	0.93	51
NE--HE	1.01	0.86	44
ND2--HD22	1.01	0.86	43
ND2--HD21	1.01	0.86	43

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CZ--HZ	1.08	0.93	28
NH1--HH12	1.01	0.86	44
NE2--HE22	1.01	0.86	37
NE2--HE21	1.01	0.86	37
ND1--HD1	1.01	0.86	13
NH2--HH22	1.01	0.86	44
CH2--HH2	1.08	0.93	5
NE2--HE2	1.01	0.86	1
CZ3--HZ3	1.08	0.93	5
CZ2--HZ2	1.08	0.93	5
CE3--HE3	1.08	0.93	5
NE1--HE1	1.01	0.86	5

Standard geometry: angle outliers

There are 50 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	134.63	1
C-N-CA	121.70	134.00	1
O-C-N	123.00	113.16	1
OE1-CD-NE2	122.60	116.58	1
CA-CB-CG2	110.40	120.23	1
OE1-CD-NE2	122.60	116.86	1
CA-C-N	116.20	127.51	1
CA-CB-CG1	110.40	119.87	1
C-N-CA	121.70	131.33	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
OE1-CD-NE2	122.60	117.58	1
CA-CB-CG	113.80	118.68	1
OE1-CD-NE2	122.60	117.76	1
CA-CB-CG	112.60	117.42	1
OE1-CD-NE2	122.60	117.82	1
OE1-CD-NE2	122.60	117.85	1
OE1-CD-NE2	122.60	117.87	1
CA-CB-CG	112.60	117.31	1
C-N-CA	121.70	130.11	1
NH1-CZ-NH2	119.30	113.24	1
CB-CG-CD2	131.20	125.23	1
OE1-CD-NE2	122.60	118.08	1
NE-CZ-NH1	121.50	125.98	1
OE1-CD-NE2	122.60	118.14	1
OE1-CD-NE2	122.60	118.17	3
OE1-CD-NE2	122.60	118.18	1
CA-CB-CG	112.60	117.01	1
C-CA-CB	111.40	119.76	1
OE1-CD-NE2	122.60	118.25	1
CD-NE-CZ	124.40	130.48	1
C-N-CA	121.70	129.30	1
OE1-CD-NE2	122.60	118.40	1
OE1-CD-NE2	122.60	118.41	2

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
OE1-CD-NE2	122.60	118.42	1
OE1-CD-NE2	122.60	118.44	1
CA-CB-CG	113.80	117.90	1
CB-CG-CD2	131.20	125.87	1
OD1-CG-ND2	122.60	118.51	2
OE1-CD-NE2	122.60	118.53	1
CB-CG-CD2	131.20	125.96	1
OE1-CD-NE2	122.60	118.59	1
NE-CZ-NH2	119.20	122.80	1
HZ1-NZ-HZ3	96.98	109.00	1
HH21-NH2-HH22	107.84	120.00	1
C-CA-HA	93.85	109.00	1
N-CA-HA	94.33	110.00	1
C-N-H	107.65	124.30	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	1.44	23

All 23 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	B:188:GLY:HA2	B:199:ARG:CZ	0.718
1	B:89:VAL:HB	B:123:VAL:HG22	0.570
1	A:151:LEU:HD11	A:361:LEU:HD11	0.559

Model ID	Atom-1	Atom-2	Clash overlap (Å)
1	B:188:GLY:CA	B:199:ARG:CZ	0.554
1	B:86:HIS:O	B:89:VAL:HG23	0.547
1	A:421:ASN:HA	A:485:MET:HE3	0.507
1	A:97:LEU:HA	A:100:TYR:CE1	0.494
1	B:85:PHE:CE1	B:122:ILE:HD12	0.478
1	A:513:VAL:HG11	A:543:LYS:HE3	0.476
1	B:414:VAL:HA	B:417:MET:HE2	0.476
1	A:181:LEU:HD23	A:186:PHE:CE1	0.472
1	B:194:GLY:HA2	B:373:TYR:CE1	0.465
1	A:421:ASN:HA	A:485:MET:CE	0.458
1	A:151:LEU:HD11	A:361:LEU:CD1	0.453
1	A:381:LEU:HD22	A:446:PHE:CD2	0.447
1	B:38:VAL:HG22	B:427:ILE:HD12	0.445
1	B:24:VAL:HG13	B:410:TYR:CD2	0.422
1	B:89:VAL:HG22	B:93:GLU:OE2	0.422
1	A:409:HIS:CE1	A:426:ARG:HH21	0.419
1	B:326:GLU:CD	B:406:ARG:HH22	0.419
1	B:257:TYR:CE2	B:283:LYS:HE3	0.412
1	A:53:ALA:HA	A:56:GLN:HE21	0.410
1	B:75:SER:HB3	B:426:TRP:CZ3	0.404

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
----------	----------	---------	---------	----------

Model ID	Analyzed	Favored	Allowed	Outliers
1	1034	1012	21	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	859	814	27	18

Detailed list of outliers are tabulated below.

Model ID	Chain	Residue ID	Residue type
1	A	229	LEU
1	A	302	SER
1	A	330	GLU
1	A	335	VAL
1	A	527	VAL
1	A	572	LEU
1	A	573	TYR
1	A	575	GLU
1	A	593	ASP
1	A	606	ASP
1	A	607	ASP
1	B	7	THR
1	B	10	THR
1	B	89	VAL
1	B	105	SER
1	B	273	ASP

Model ID	Chain	Residue ID	Residue type
1	B	292	GLN
1	B	294	LEU

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 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.