

Integrative Structure Validation Report

July 22, 2024 - 04:26 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	9A1F
PDB-Dev ID	PDBDEV_00000087
Structure Title	Integrative structure of the XcpGHIJK pseudo pilus filament model of a type II secretion system
Structure Authors	Escobar CA; Douzi B; Ball G; Barbat B; Alphonse S; Quinton L; Voulhoux R; Forest KT

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

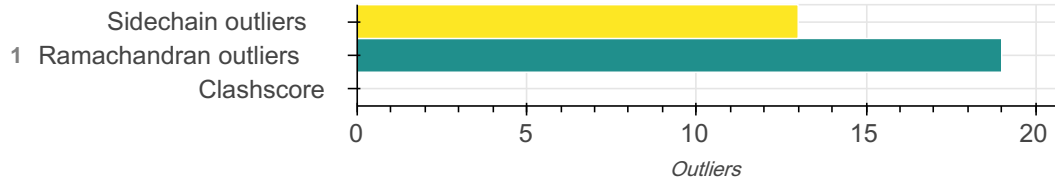
A user guide is available at https://pdb-dev.wwpdb.org/validation_help.html with specific help available everywhere you see the  symbol.

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 20 subunits in each model. A total of 7 datasets or restraints were used to build this entry. Each model is represented by 40 rigid bodies and 0 flexible or non-rigid units.

Entry composition ?

There is 1 unique type of models in this entry. This model is titled XcpGHIJK model/XcpGHIJK.

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Chain ID [auth]	Total residues
1	1	1	XcpG	A	A	134
1	2	1	XcpG	B	B	134
1	3	1	XcpG	C	C	134
1	4	1	XcpG	D	D	134
1	5	1	XcpG	E	E	134
1	6	1	XcpG	F	F	134
1	7	1	XcpG	G	G	134
1	8	1	XcpG	H	H	134
1	9	1	XcpG	I	I	134
1	10	1	XcpG	J	J	134
1	11	1	XcpG	K	K	134

Model ID	Subunit number	Subunit ID	Subunit name	Chain ID	Chain ID [auth]	Total residues
1	12	1	XcpG	L	L	134
1	13	1	XcpG	M	M	134
1	14	1	XcpG	N	N	134
1	15	1	XcpG	O	O	134
1	16	1	XcpG	P	P	134
1	17	2	XcpH	Q	Q	161
1	18	3	XcpI	R	R	120
1	19	4	XcpJ	S	S	199
1	20	5	XcpK	T	T	309

Datasets used for modeling

There are 7 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Experimental model	PDB	5VTM
2	NMR data	BMRB	50449
3	Crosslinking-MS data	MASSIVE	MSV000086915
4	Comparative model	Not available	Not available
5	Experimental model	PDB	2QV8
6	Experimental model	PDB	2KEP
7	Integrative model	PDB-Dev	PDBDEV_00000086

Representation

This entry has only one representation and includes 40 rigid bodies and 0 flexible units

Chain ID	Rigid bodies	Non-rigid segments
A	1-24:None, 25-134	-
B	1-24:None, 25-134	-
C	1-24:None, 25-134	-
D	1-24:None, 25-134	-
E	1-24:None, 25-134	-
F	1-24:None, 25-134	-
G	1-24:None, 25-134	-
H	1-24:None, 25-134	-
I	1-24:None, 25-134	-
J	1-24:None, 25-134	-
K	1-24:None, 25-134	-
L	1-24:None, 25-134	-
M	1-24:None, 25-134	-
N	1-24:None, 25-134	-
O	1-24:None, 25-134	-
P	1-24:None, 25-134	-
Q	1-30:None, 31-161	-
R	1-31:None, 32-120	-
S	1-37:None, 38-199	-
T	1-36:None, 37-309	-

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	HADDOCK	Docking	None	1000	False	False
2	1	Homology modeling	Helix models	None	None	False	False
3	1	Manual modeling	Add helices	None	None	False	False
4	1	Semi-Manual modeling	Modeling Helices	None	300	False	False
5	1	Manual modeling	Create XcpG filament	None	None	False	False
6	1	Minimization	Minimize filament	None	10	False	False

There are 5 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	Pymol	Not available	Visualization-modeling	https://pymol.org/2/
2	Haddock	Not available	Docking	https://alcazar.science.uu.nl/services/HADDOCK2.2/
3	Python	Not available	Programing language	https://www.python.org/
4	pyRosetta	Not available	Modeling	http://www.pyrosetta.org
5	Phyre2	Not available	Homology modeling	http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index

Data quality 

Crosslinking-MS

Validation for this section is under development.

NMR

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 22434 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
CA--HA	1.07	0.97	22
CA--2HA	1.07	0.97	7
CB--1HB	1.07	0.97	2
CA--2HA	1.08	0.97	185
CA--HA	1.08	0.97	1353
CB--1HB	1.08	0.97	190
CD--1HD	1.08	0.97	24
CA--1HA	1.08	0.97	192
CB--2HB	1.08	0.97	192
CE--1HE	1.08	0.97	1
CB--3HB	1.08	0.97	143
CD--2HD	1.08	0.97	24
CG--2HG	1.08	0.97	24
CG--1HG	1.08	0.97	24
N--2H	1.00	0.89	20
N--3H	1.00	0.89	20
CE--2HE	1.08	0.97	1

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
N--1H	1.00	0.89	20
N--H	0.97	0.86	58
CA--HA	1.09	0.97	1182
N--H	0.98	0.86	1291
CB--3HB	1.09	0.97	136
CD--1HD	1.09	0.97	307
CB--1HB	1.09	0.97	1780
CD--2HD	1.09	0.97	307
CE--1HE	1.09	0.97	274
CB--2HB	1.09	0.97	1780
CA--1HA	1.09	0.97	67
CG--2HG	1.09	0.97	611
CD2--2HD2	1.09	0.97	314
CG--1HG	1.09	0.97	611
CG2--1HG2	1.09	0.97	482
NZ--1HZ	1.01	0.89	211
NZ--3HZ	1.01	0.89	211
CG1--2HG1	1.09	0.97	374
OG--HG	0.96	0.84	106
CG2--3HG2	1.09	0.97	482
CE--2HE	1.09	0.97	274
CA--2HA	1.09	0.97	67
CD1--3HD1	1.09	0.97	451

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CG2--2HG2	1.09	0.97	482
CG1--3HG1	1.09	0.97	237
CB--HB	1.09	0.97	482
CG1--1HG1	1.09	0.97	374
OH--HH	0.96	0.84	107
CG--HG	1.09	0.97	314
CD1--2HD1	1.09	0.97	451
CD2--3HD2	1.09	0.97	314
OG1--HG1	0.96	0.84	108
CD1--1HD1	1.09	0.97	451
CD2--1HD2	1.09	0.97	314
CE--3HE	1.09	0.97	63
NZ--2HZ	1.01	0.89	211
NE--HE	0.98	0.86	4
NE2--2HE2	0.98	0.86	1
NE2--1HE2	0.98	0.86	1
N--H	0.99	0.86	75
OH--HH	0.97	0.84	1
NE--HE	0.99	0.86	1
CB--1HB	1.10	0.97	182
CG--1HG	1.10	0.97	182
CB--2HB	1.10	0.97	182
OG--HG	0.97	0.84	2

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
CG--2HG	1.10	0.97	182
CD--1HD	1.10	0.97	182
CA--HA	1.10	0.97	61
CD--2HD	1.10	0.97	182
SG--HG	1.33	1.20	1
ND2--2HD2	1.00	0.86	147
ND2--1HD2	1.00	0.86	147
NE--HE	1.00	0.86	3
NE2--2HE2	1.00	0.86	132
NE2--1HE2	1.00	0.86	132
ND1--HD1	1.00	0.86	3
N--H	1.00	0.86	1
NZ--3HZ	1.04	0.89	1
N--H	1.01	0.86	1227
NH2--2HH2	1.01	0.86	88
CE1--HE1	1.08	0.93	4
NH1--1HH1	1.01	0.86	88
NH1--2HH1	1.01	0.86	88
NH2--1HH2	1.01	0.86	88
CD1--HD1	1.08	0.93	4
CD2--HD2	1.08	0.93	4
NE1--HE1	1.01	0.86	64
CE2--HE2	1.08	0.93	4

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
NE--HE	1.01	0.86	88
NE2--HE2	1.01	0.86	6
NZ--1HZ	1.04	0.89	1
CZ--HZ	1.08	0.93	3
NZ--2HZ	1.04	0.89	1
CD1--HD1	1.09	0.93	240
CZ--HZ	1.09	0.93	69
CE1--HE1	1.09	0.93	185
CE2--HE2	1.09	0.93	176
CE3--HE3	1.09	0.93	64
CD2--HD2	1.09	0.93	185
CH2--HH2	1.09	0.93	64
CZ3--HZ3	1.09	0.93	64
CZ2--HZ2	1.09	0.93	64
NH1--1HH1	1.04	0.86	4
NH2--2HH2	1.04	0.86	4
NH1--2HH1	1.04	0.86	2
NH2--1HH2	1.04	0.86	3
NH2--2HH2	1.05	0.86	2
NH1--1HH1	1.05	0.86	4
NH2--1HH2	1.05	0.86	3
NH1--2HH1	1.05	0.86	5
NH2--2HH2	1.06	0.86	2

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
NH1--2HH1	1.06	0.86	1
NH2--1HH2	1.06	0.86	2

Standard geometry: angle outliers

There are 89 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	76.40	1
O-C-N	123.00	155.73	1
O-C-N	123.00	93.01	1
O-C-N	123.00	97.92	1
CA-C-N	116.20	87.10	1
CA-C-N	116.20	139.18	1
C-N-CA	121.70	103.14	1
C-N-CA	121.70	103.17	1
C-N-CA	121.70	103.18	1
C-N-CA	121.70	103.19	3
C-N-CA	121.70	103.20	2
C-N-CA	121.70	103.21	1
C-N-CA	121.70	103.22	1
C-N-CA	121.70	103.23	3
C-N-CA	121.70	103.24	1
C-N-CA	121.70	103.25	2
CA-C-N	116.20	128.27	1
O-C-N	123.00	131.99	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
O-C-N	123.00	131.93	1
O-C-N	123.00	131.92	2
O-C-N	123.00	131.91	1
O-C-N	123.00	131.90	2
O-C-N	123.00	131.89	1
CA-C-N	116.20	127.31	1
O-C-N	123.00	131.87	1
O-C-N	123.00	131.86	3
O-C-N	123.00	131.85	1
O-C-N	123.00	131.83	1
CA-C-N	116.20	127.23	1
O-C-N	123.00	131.82	1
O-C-N	123.00	131.80	1
CA-C-N	116.20	105.22	2
CA-C-N	116.20	105.23	2
CA-C-N	116.20	105.24	1
CA-C-N	116.20	105.25	1
CA-C-N	116.20	105.26	2
CA-C-N	116.20	105.27	2
CA-C-N	116.20	105.28	3
CA-C-N	116.20	105.29	1
CA-C-N	116.20	105.32	1
CA-C-N	116.20	127.08	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
CA-C-N	116.20	105.34	1
O-C-N	123.00	114.62	1
CA-C-N	116.20	126.21	1
CA-C-N	116.20	126.18	1
N-CA-CB	110.40	102.94	1
C-CA-CB	110.50	117.88	1
CA-CB-CG	112.60	107.95	1
CA-C-N	116.20	124.98	1
C-CA-CB	110.50	117.05	3
C-CA-CB	110.50	117.01	1
C-CA-CB	110.50	117.00	1
C-CA-CB	110.50	116.99	3
C-CA-CB	110.50	116.98	2
C-CA-CB	110.50	116.96	3
C-CA-CB	110.50	116.95	1
C-CA-CB	110.50	116.94	1
CA-C-N	116.20	124.79	1
C-CA-CB	110.50	116.93	1
O-C-N	123.00	116.49	1
C-N-CA	121.70	114.39	1
O-C-N	123.00	116.59	1
C-CA-HA	96.14	109.00	1
CB-CA-HA	122.19	109.00	1

Angle type	Observed angle (°)	Ideal angle (°)	Number of outliers
C-CA-HA	92.45	109.00	1
C-N-H	164.55	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	0.00	0

All 0 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

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	2842	2675	148	19

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	2357	2299	45	13

Detailed list of outliers are tabulated below.

Model ID	Chain	Residue ID	Residue type
1	D	92	OASP
1	E	92	OASP
1	F	9	OVAL
1	G	9	OVAL
1	G	88	OLYS

Model ID	Chain	Residue ID	Residue type
1	G	92	0ASP
1	H	92	0ASP
1	I	92	0ASP
1	J	92	0ASP
1	L	92	0ASP
1	M	92	0ASP
1	O	9	0VAL
1	P	92	0ASP

Fit of model to data used for modeling ?

Crosslinking-MS

Validation for this section is under development.

NMR

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.