Integrative Structure Validation Report ?

October 09, 2025 - 04:39 PM PDT

The following software was used in the production of this report:

IHMValidation Version 3.0
Python-IHM Version 2.5
MolProbity Version 4.5.2

PDB ID	9A1F	pdb_00009a1f
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
Deposited on	2021-05-03

This is a PDB-IHM Structure Validation Report.

We welcome your comments at helpdesk@pdb-ihm.org

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

List of references used to build this report is available here.

1. Overview ?

1.1. Summary ?

This entry consists of 1 model(s). A total of 7 dataset(s) were used to build this entry.

Name	Type	Count
Crosslinking-MS data	Experimental data	1
NMR data	Experimental data	1
Experimental model	Starting model	3

| Name	Type	Count
Integrative model	Starting model	1
Comparative model	Starting model	1

1.2. Overall quality ⍰

This validation report contains model quality assessments for all structures, data quality and fit to model assessments for SAS and crosslinking-MS 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 ⍰
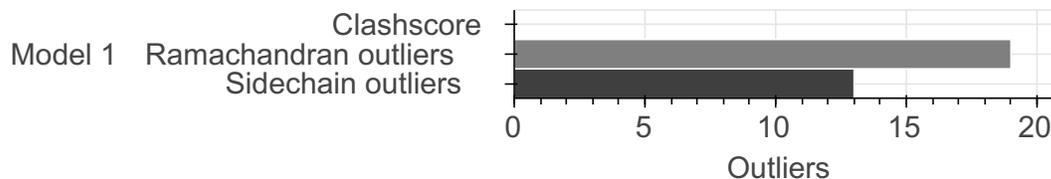
Fit to Data Used for Modeling ⍰
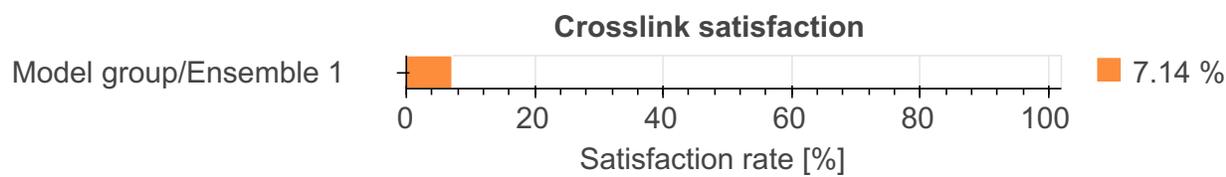
2. Model Details ⍰

2.1. Ensemble information ⍰

This entry consists of 0 distinct ensemble(s).

2.2. Representation ⍰

This entry has 1 representation(s).

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
1	1	1	XcpG	A	134	1-24, 25-134	-	100.00 / 82.09	Atomic
				B					
				C					
				D					
				E					
				F					
				G					
				H					

ID	Model(s)	Entity ID	Molecule name	Chain(s) [auth]	Total residues	Rigid segments	Flexible segments	Model coverage/ Starting model coverage (%)	Scale
				I					
				J					
				K					
				L					
				M					
				N					
				O					
				P					
		2	XcpH	Q	161	1-30, 31-161	-	100.00 / 81.37	Atomic
		3	XcpI	R	120	1-31, 32-120	-	100.00 / 74.17	Atomic
		4	XcpJ	S	199	1-37, 38-199	-	100.00 / 81.41	Atomic
		5	XcpK	T	309	1-36, 37-309	-	100.00 / 88.35	Atomic

2.3. Datasets used for modeling ⑦

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	pdb_00005vtm
2	NMR data	BMRB	50449
3	Crosslinking-MS data	MASSIVE	MSV000086915
4	Comparative model	Not available	Not available
5	Experimental model	PDB	pdb_00002qv8
6	Experimental model	PDB	pdb_00002kep
7	Integrative model	PDB	pdb_00009a1e

2.4. Methodology and software ⑦

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	Docking	HADDOCK	Not available	1000	False	False
2	1	Helix models	Homology modeling	Not available	Not available	False	False

Step number	Protocol ID	Method name	Method type	Method description	Number of computed models	Multi state modeling	Multi scale modeling
3	1	Add helices	Manual modeling	Not available	Not available	False	False
4	1	Modeling Helices	Semi-Manual modeling	Not available	300	False	False
5	1	Create XcpG filament	Manual modeling	Not available	Not available	False	False
6	1	Minimize filament	Minimization	Not available	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

3. Data quality ?

3.2. Crosslinking-MS

At the moment, data validation is only available for crosslinking-MS data deposited as a fully compliant dataset in the PRIDE Crosslinking database. Correspondence between crosslinking-MS and entry entities is established using pyHMMER. Only residue pairs that passed the reported threshold are used for the analysis. The values in the report have to be interpreted in the context of the experiment (i.e. only a minor fraction of in-situ or in-vivo dataset can be used for modeling).

Crosslinking-MS dataset is not available in the PRIDE Crosslinking database.

3.4. NMR ?

Validation for this section is under development.

4. 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.

4.1b. MolProbity Analysis ?

Excluded volume satisfaction for the models in the entry are listed below. The Analysed column shows the number of particle-partice or particle-atom pairs for which excluded volume was analysed.

Standard geometry: bond outliers ?

There are 52 bond length outliers in this entry (0.23% of 22825 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
T	37	VAL	C-N	12.23	1.16	1.33	1	1
P	35	VAL	C-N	11.49	1.49	1.33	1	1
N	35	VAL	C-N	11.45	1.49	1.33	1	1
J	35	VAL	C-N	11.45	1.49	1.33	1	1
E	35	VAL	C-N	11.43	1.49	1.33	1	1
H	35	VAL	C-N	11.42	1.49	1.33	1	1
O	35	VAL	C-N	11.40	1.49	1.33	1	1
F	35	VAL	C-N	11.39	1.49	1.33	1	1
G	35	VAL	C-N	11.39	1.49	1.33	1	1
K	35	VAL	C-N	11.37	1.49	1.33	1	1
C	35	VAL	C-N	11.37	1.49	1.33	1	1
M	35	VAL	C-N	11.37	1.49	1.33	1	1
I	35	VAL	C-N	11.37	1.49	1.33	1	1
L	35	VAL	C-N	11.36	1.49	1.33	1	1
A	35	VAL	C-N	11.36	1.49	1.33	1	1
D	35	VAL	C-N	11.32	1.49	1.33	1	1
B	35	VAL	C-N	11.27	1.49	1.33	1	1
H	36	ALA	C-N	7.61	1.44	1.33	1	1
D	36	ALA	C-N	7.57	1.44	1.33	1	1
O	36	ALA	C-N	7.56	1.43	1.33	1	1
M	36	ALA	C-N	7.56	1.43	1.33	1	1
G	36	ALA	C-N	7.55	1.43	1.33	1	1
S	37	GLN	C-N	7.54	1.22	1.33	1	1
K	36	ALA	C-N	7.54	1.43	1.33	1	1
N	36	ALA	C-N	7.53	1.43	1.33	1	1
F	36	ALA	C-N	7.53	1.43	1.33	1	1
P	36	ALA	C-N	7.53	1.43	1.33	1	1
E	36	ALA	C-N	7.52	1.43	1.33	1	1
B	36	ALA	C-N	7.51	1.43	1.33	1	1
A	36	ALA	C-N	7.51	1.43	1.33	1	1
I	36	ALA	C-N	7.50	1.43	1.33	1	1
C	36	ALA	C-N	7.49	1.43	1.33	1	1
L	36	ALA	C-N	7.47	1.43	1.33	1	1
J	36	ALA	C-N	7.45	1.43	1.33	1	1
O	131	ASN	C-N	4.65	1.26	1.33	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
A	131	ASN	C-N	4.65	1.26	1.33	1	1
L	131	ASN	C-N	4.63	1.26	1.33	1	1
G	131	ASN	C-N	4.62	1.26	1.33	1	1
H	131	ASN	C-N	4.60	1.26	1.33	1	1
K	131	ASN	C-N	4.60	1.26	1.33	1	1
J	131	ASN	C-N	4.60	1.26	1.33	1	1
M	131	ASN	C-N	4.59	1.26	1.33	1	1
I	131	ASN	C-N	4.58	1.26	1.33	1	1
N	131	ASN	C-N	4.57	1.26	1.33	1	1
C	131	ASN	C-N	4.57	1.27	1.33	1	1
E	131	ASN	C-N	4.56	1.27	1.33	1	1
B	131	ASN	C-N	4.56	1.27	1.33	1	1
P	131	ASN	C-N	4.54	1.27	1.33	1	1
F	131	ASN	C-N	4.52	1.27	1.33	1	1
P	28	PRO	N-CD	4.51	1.54	1.47	1	1
D	131	ASN	C-N	4.44	1.27	1.33	1	1
O	28	PRO	N-CD	4.28	1.53	1.47	1	1

Standard geometry: angle outliers ?

There are 85 bond angle outliers in this entry (0.27% of 31052 assessed bonds). A summary is provided below.

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
S	37	GLN	C-N-CA	25.16	76.40	121.70	1	1
S	37	GLN	O-C-N	20.46	155.73	123.00	1	1
R	3	LEU	O-C-N	18.74	93.01	123.00	1	1
T	37	VAL	O-C-N	15.68	97.92	123.00	1	1
S	37	GLN	CA-C-N	14.55	87.10	116.20	1	1
R	3	LEU	CA-C-N	11.49	139.18	116.20	1	1
P	36	ALA	C-N-CA	10.31	103.14	121.70	1	1
N	36	ALA	C-N-CA	10.30	103.17	121.70	1	1
H	36	ALA	C-N-CA	10.29	103.18	121.70	1	1
C	36	ALA	C-N-CA	10.28	103.19	121.70	1	1
O	36	ALA	C-N-CA	10.28	103.19	121.70	1	1
M	36	ALA	C-N-CA	10.28	103.19	121.70	1	1
A	36	ALA	C-N-CA	10.28	103.20	121.70	1	1
F	36	ALA	C-N-CA	10.28	103.20	121.70	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
D	36	ALA	C-N-CA	10.27	103.21	121.70	1	1
G	36	ALA	C-N-CA	10.26	103.22	121.70	1	1
K	36	ALA	C-N-CA	10.26	103.23	121.70	1	1
I	36	ALA	C-N-CA	10.26	103.23	121.70	1	1
B	36	ALA	C-N-CA	10.26	103.23	121.70	1	1
J	36	ALA	C-N-CA	10.26	103.24	121.70	1	1
E	36	ALA	C-N-CA	10.25	103.25	121.70	1	1
L	36	ALA	C-N-CA	10.25	103.25	121.70	1	1
T	37	VAL	CA-C-N	6.03	128.27	116.20	1	1
P	36	ALA	O-C-N	5.62	131.99	123.00	1	1
A	36	ALA	O-C-N	5.58	131.93	123.00	1	1
L	36	ALA	O-C-N	5.57	131.92	123.00	1	1
F	36	ALA	O-C-N	5.57	131.92	123.00	1	1
H	36	ALA	O-C-N	5.57	131.91	123.00	1	1
N	36	ALA	O-C-N	5.56	131.90	123.00	1	1
M	36	ALA	O-C-N	5.56	131.90	123.00	1	1
J	36	ALA	O-C-N	5.56	131.89	123.00	1	1
S	32	GLN	CA-C-N	5.55	127.31	116.20	1	1
B	36	ALA	O-C-N	5.54	131.87	123.00	1	1
D	36	ALA	O-C-N	5.54	131.86	123.00	1	1
G	36	ALA	O-C-N	5.54	131.86	123.00	1	1
O	36	ALA	O-C-N	5.54	131.86	123.00	1	1
C	36	ALA	O-C-N	5.53	131.85	123.00	1	1
I	36	ALA	O-C-N	5.52	131.83	123.00	1	1
S	14	ALA	CA-C-N	5.51	127.23	116.20	1	1
E	36	ALA	O-C-N	5.51	131.82	123.00	1	1
K	36	ALA	O-C-N	5.50	131.80	123.00	1	1
G	36	ALA	CA-C-N	5.49	105.22	116.20	1	1
F	36	ALA	CA-C-N	5.49	105.22	116.20	1	1
M	36	ALA	CA-C-N	5.49	105.23	116.20	1	1
L	36	ALA	CA-C-N	5.48	105.23	116.20	1	1
O	36	ALA	CA-C-N	5.48	105.24	116.20	1	1
A	36	ALA	CA-C-N	5.47	105.25	116.20	1	1
H	36	ALA	CA-C-N	5.47	105.26	116.20	1	1
N	36	ALA	CA-C-N	5.47	105.26	116.20	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
P	36	ALA	CA-C-N	5.47	105.27	116.20	1	1
D	36	ALA	CA-C-N	5.46	105.27	116.20	1	1
E	36	ALA	CA-C-N	5.46	105.28	116.20	1	1
C	36	ALA	CA-C-N	5.46	105.28	116.20	1	1
K	36	ALA	CA-C-N	5.46	105.28	116.20	1	1
I	36	ALA	CA-C-N	5.46	105.29	116.20	1	1
B	36	ALA	CA-C-N	5.44	105.32	116.20	1	1
R	14	ALA	CA-C-N	5.44	127.08	116.20	1	1
J	36	ALA	CA-C-N	5.43	105.34	116.20	1	1
S	32	GLN	O-C-N	5.23	114.62	123.00	1	1
S	21	TYR	CA-C-N	5.01	126.21	116.20	1	1
R	21	LEU	CA-C-N	4.99	126.18	116.20	1	1
S	19	ALA	N-CA-CB	4.97	102.94	110.40	1	1
S	19	ALA	C-CA-CB	4.92	117.88	110.50	1	1
O	29	ASP	CA-CB-CG	4.65	107.95	112.60	1	1
R	13	PHE	CA-C-N	4.39	124.98	116.20	1	1
H	102	ALA	C-CA-CB	4.37	117.05	110.50	1	1
L	102	ALA	C-CA-CB	4.37	117.05	110.50	1	1
F	102	ALA	C-CA-CB	4.37	117.05	110.50	1	1
B	102	ALA	C-CA-CB	4.34	117.01	110.50	1	1
A	102	ALA	C-CA-CB	4.33	117.00	110.50	1	1
C	102	ALA	C-CA-CB	4.33	116.99	110.50	1	1
J	102	ALA	C-CA-CB	4.33	116.99	110.50	1	1
K	102	ALA	C-CA-CB	4.33	116.99	110.50	1	1
D	102	ALA	C-CA-CB	4.32	116.98	110.50	1	1
I	102	ALA	C-CA-CB	4.32	116.98	110.50	1	1
O	102	ALA	C-CA-CB	4.31	116.96	110.50	1	1
N	102	ALA	C-CA-CB	4.31	116.96	110.50	1	1
G	102	ALA	C-CA-CB	4.30	116.96	110.50	1	1
E	102	ALA	C-CA-CB	4.30	116.95	110.50	1	1
P	102	ALA	C-CA-CB	4.30	116.94	110.50	1	1
S	13	PHE	CA-C-N	4.29	124.79	116.20	1	1
M	102	ALA	C-CA-CB	4.29	116.93	110.50	1	1
R	21	LEU	O-C-N	4.07	116.49	123.00	1	1
T	2	ALA	C-N-CA	4.06	114.39	121.70	1	1

Chain	Res	Type	Atoms	|Z|	Observed (Å)	Ideal (Å)	Model ID (Worst)	Models (Total)
R	14	ALA	O-C-N	4.01	116.59	123.00	1	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 atomic models in this entry.

Model ID	Clash score	Number of clashes
1	0.00	0

There are no too-close contacts.

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
1	2842	2675	148	19

There are 19 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
C	102	ALA	1
E	22	PRO	1
E	55	PHE	1
F	22	PRO	1
G	22	PRO	1
I	22	PRO	1
J	22	PRO	1
K	22	PRO	1
K	55	PHE	1
L	22	PRO	1
M	22	PRO	1
M	55	PHE	1
N	55	PHE	1
O	22	PRO	1
O	55	PHE	1
P	22	PRO	1
P	55	PHE	1
Q	78	ARG	1
S	4	LEU	1

Torsion angles : Protein sidechains ?

In the following table, sidechain rotameric outliers are listed. The Analysed column shows the number of residues for which the sidechain conformation was analysed.

Model ID	Analysed	Favored	Allowed	Outliers
1	2357	2299	45	13

There are 13 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
D	92	ASP	1
E	92	ASP	1
F	9	VAL	1
G	9	VAL	1
G	88	LYS	1
G	92	ASP	1
H	92	ASP	1
I	92	ASP	1
J	92	ASP	1
L	92	ASP	1
M	92	ASP	1
O	9	VAL	1
P	92	ASP	1

5. Fit to Data Used for Modeling Assessment ?

5.2. Crosslinking-MS ?

5.2.1. Restraint types ?

This table summarizes information about crosslinker(s) used for data generation, and how crosslinking information was translated into actual modeling restraints. Restraints assigned "by-residue" are interpreted as between CA atoms. Restraints between coarse-grained beads are indicated as "coarse-grained". Restraint group represents a set of crosslinking restraints applied collectively in the modeling.

There are 28 crosslinking restraints combined in 14 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
ADH	LEU	CA	PHE	CA	lower bound	8.00	1
ADH	LEU	CA	PHE	CA	upper bound	18.00	1
ADH	PHE	CA	VAL	CA	lower bound	8.00	2
ADH	PHE	CA	VAL	CA	upper bound	18.00	2
ADH	LEU	CA	MET	CA	lower bound	8.00	1

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
ADH	LEU	CA	MET	CA	upper bound	18.00	1
ADH	MET	CA	VAL	CA	lower bound	8.00	1
ADH	MET	CA	VAL	CA	upper bound	18.00	1
ADH	VAL	CA	VAL	CA	lower bound	8.00	1
ADH	VAL	CA	VAL	CA	upper bound	18.00	1
ADH	LEU	CA	LEU	CA	lower bound	8.00	4
ADH	LEU	CA	LEU	CA	upper bound	18.00	4
ADH	LEU	CA	VAL	CA	lower bound	8.00	4
ADH	LEU	CA	VAL	CA	upper bound	18.00	4

Distograms of individual restraints

Distograms (i.e., histogram plots of distances) provide an overview of distributions of distances between residues for which chemical crosslinks were identified. The shift of the distogram relative to the threshold value may indicate a poor model. Restraints with identical thresholds are grouped into one plot. Only the best distance per restraint per model group/ensemble is plotted. Inter- and intramolecular (including self-links) restraints are also grouped into one plot. Distance for a restraint between coarse-grained beads is calculated as a minimal distance between shells; if beads intersect, the distance will be reported as 0.0. A bead with the highest available resolution for a given residue is used for the assessment.

Model Group 1; Self-links: lower bound, 8.0 Å

Model Group 1; Self-links: upper bound, 18.0 Å
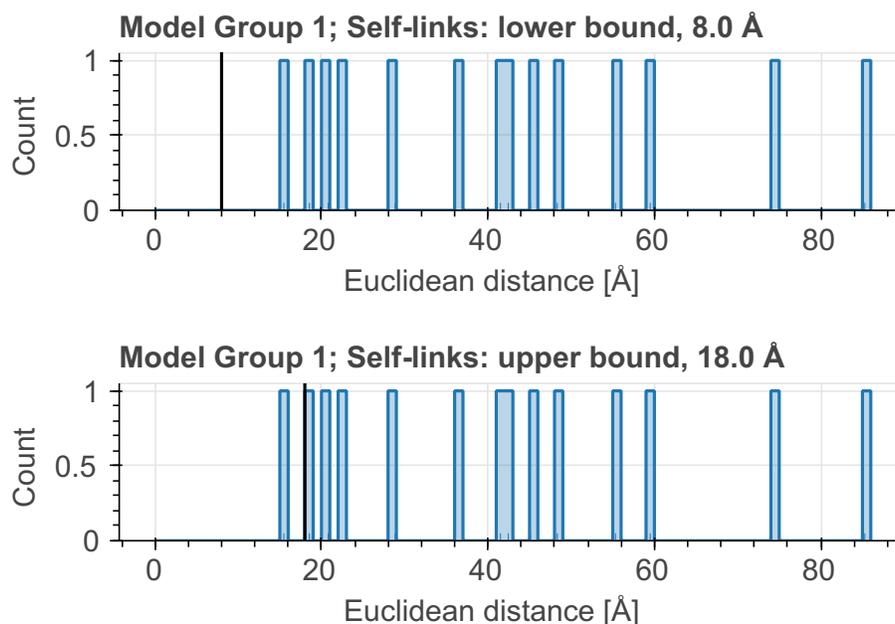
5.2.2. Satisfaction of restraints ?

Satisfaction of restraints is calculated on a restraint group (a set of crosslinking restraints applied collectively in the modeling) level. Satisfaction of a restraint group depends on satisfaction of individual restraints in the group and the conditionality (all/any). A restraint group is considered satisfied, if the condition was met in at least one model of the model group/ensemble. The number of measured restraints can be smaller than the total number of restraint groups if crosslinks involve non-modeled residues. Only deposited models are used for validation right now.

State group	State	Model group	# of Deposited models/Total	Restraint group type	Satisfied (%)	Violated (%)	Count (Total=14)
1	1	1	1/1	All	7.14	92.86	14
				Self-links/ Intermolecular	7.14	92.86	14

Per-model satisfaction rates in ensembles

Every point represents one model in a model group/ensemble. Where possible, boxplots with quartile marks are also plotted.

Satisfaction rates in Model Group 1

5.4. NMR ?

Validation for this section is under development.

6. Fit to Data Used for Validation Assessment ?

Validation for this section is under development.