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	9A1E	pdb_00009a1e
PDB-Dev ID	PDBDEV_00000086
Structure Title	Integrative structure of the XcpHIJK quaternary complex of a type II secretion system pseudopilin
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 5 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	2

| Name	Type	Count
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 ?

Crosslink satisfaction

28.57 %
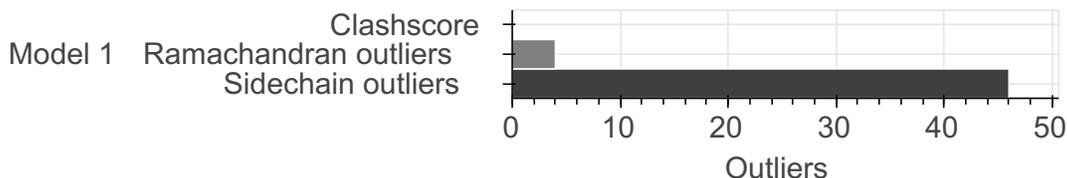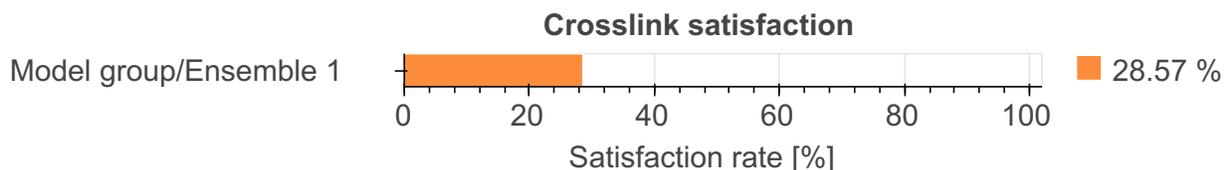
2. Model Details ?

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	XcpH	A	131	1-131	-	100.00 / 100.00	Atomic
		2	XcpI	B	89	1-89	-	100.00 / 100.00	Atomic
		3	XcpJ	C	162	1-162	-	100.00 / 100.00	Atomic
		4	XcpK	D	273	1-273	-	100.00 / 100.00	Atomic

2.3. Datasets used for modeling ?

There are 5 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

2.4. 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	XcpH modeling	Homology modeling	Not available	Not available	False	False
2	1	Docking	HADDOCK	Not available	1000	False	False

There are 3 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	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 no bond length outliers.

Standard geometry: angle outliers ⍰

There are no bond angle outliers.

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	596	568	24	4

There are 4 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
B	68	PRO	1
D	112	ASP	1
D	113	ALA	1
D	115	GLN	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	519	416	57	46

There are 46 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	44	GLU	1
A	90	ASP	1
A	98	SER	1
A	99	SER	1

Chain	Res	Type	Models (Total)
A	103	SER	1
A	112	ARG	1
A	127	ARG	1
B	8	LEU	1
B	17	LEU	1
B	25	THR	1
B	28	SER	1
B	35	GLU	1
B	47	THR	1
B	56	ASP	1
C	32	ASP	1
C	39	SER	1
C	43	ASN	1
C	51	ARG	1
C	63	SER	1
C	89	GLN	1
C	97	GLN	1
C	98	VAL	1
C	100	ASP	1
C	103	THR	1
C	106	SER	1
C	122	THR	1
C	124	GLU	1
D	46	THR	1
D	50	LEU	1
D	51	ASP	1
D	72	LEU	1
D	76	ARG	1
D	92	THR	1
D	115	GLN	1
D	141	SER	1
D	174	LEU	1
D	191	ASP	1
D	204	ARG	1
D	205	GLU	1

Chain	Res	Type	Models (Total)
D	211	ASP	1
D	214	THR	1
D	228	SER	1
D	254	THR	1
D	264	ARG	1
D	268	ARG	1
D	272	GLN	1

5. Fit to Data Used for Modeling Assessment ⍰
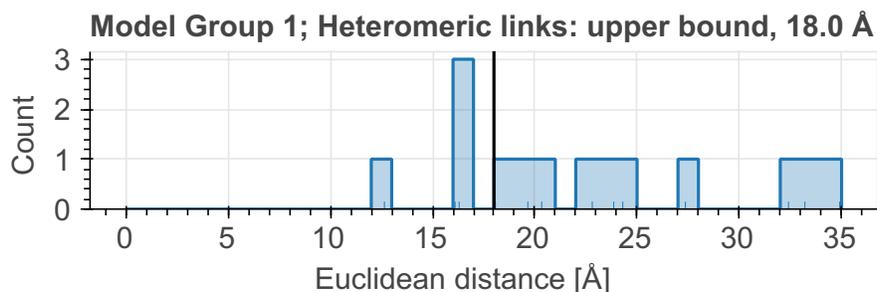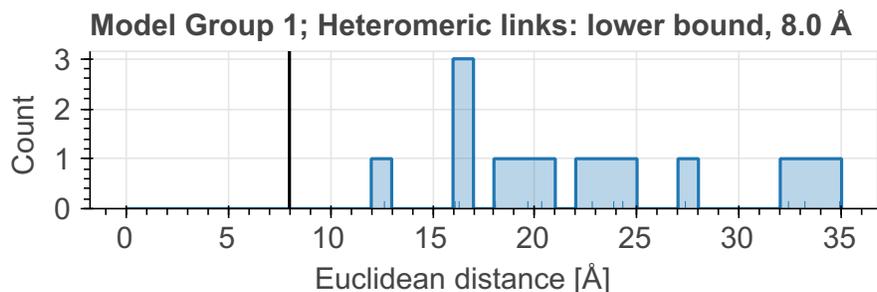
5.2. Crosslinking-MS ⍰

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	ASP	CA	GLU	CA	lower bound	8.00	4
ADH	ASP	CA	GLU	CA	upper bound	18.00	4
ADH	GLU	CA	GLU	CA	lower bound	8.00	7
ADH	GLU	CA	GLU	CA	upper bound	18.00	7
ADH	GLN	CA	GLU	CA	lower bound	8.00	1
ADH	GLN	CA	GLU	CA	upper bound	18.00	1
ADH	ASP	CA	ASP	CA	lower bound	8.00	2
ADH	ASP	CA	ASP	CA	upper bound	18.00	2

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; Heteromeric links: lower bound, 8.0 Å

Model Group 1; Heteromeric 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	28.57	71.43	14
				Heteromeric links/ Intermolecular	28.57	71.43	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
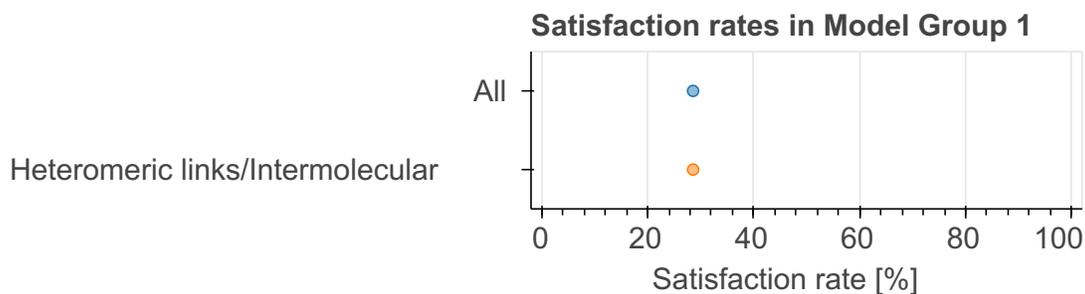
5.4. NMR ?

Validation for this section is under development.

6. Fit to Data Used for Validation Assessment ?

Validation for this section is under development.