Integrative Structure Validation Report ?

October 09, 2025 - 04:44 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	9A3Y | pdb_00009a3y
PDB-Dev ID	PDBDEV_00000219
Structure Title	Predicted model of tissue transglutaminase 2 (TG2) in complex with plasma fibronectin type I modules 7, 8 and 9 (FNI7-9)
Structure Authors	Selcuk, K.; Leitner, A.; Le Blanc, F.
Deposited on	2023-10-04

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 4 dataset(s) were used to build this entry.

| Name	Type	Count
Crosslinking-MS data	Experimental data	1
Experimental model	Starting model	2
De Novo 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 ⍰
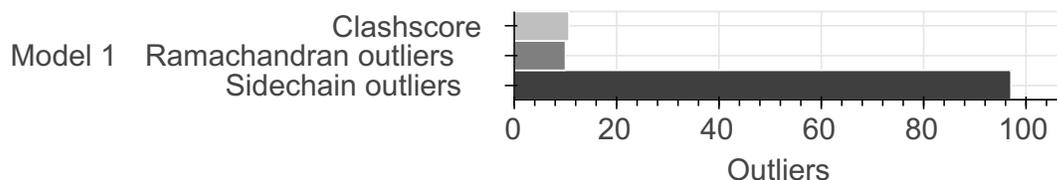
Fit to Data Used for Modeling ⍰
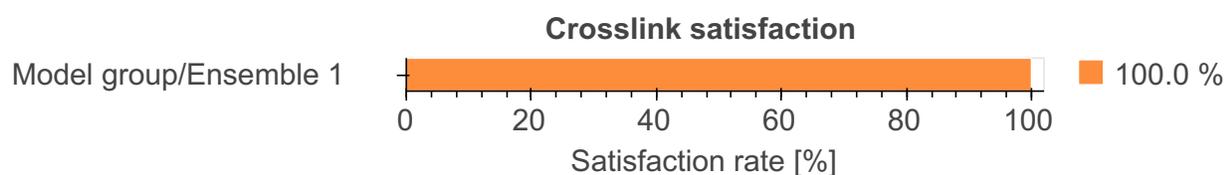
2. Model Details ⍰

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	Protein-glutamine gamma-glutamyltransferase 2	A	687	-	1-687	100.00 / 100.00	Atomic
		2	Fibronectin	B	140	-	1-140	100.00 / 100.00	Atomic

2.3. Datasets used for modeling ⍰

There are 4 unique datasets used to build the models in this entry.

ID	Dataset type	Database name	Data access code
1	Crosslinking-MS data	PRIDE	PXD043976
2	Experimental model	PDB	pdb_00004pyg
3	De Novo model	Not available	Not available

ID	Dataset type	Database name	Data access code
4	Experimental model	PDB	pdb_00003ejh

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	Refinement	Iterative Threading ASSEmbly Refinement	structural refinement of starting models guided by experimental crosslinks	Not available	False	False
2	1	structure prediction	AI prediction of protein 3D structure from its amino acid sequence	structure prediciton of FN type I modules 7-9	Not available	False	False
3	1	satisfied/violated crosslinks identification	Validate measured chemical cross-links on a protein 3D structure	calculation of Euclidean distances between crosslinked residues	Not available	False	False
4	1	calculation of accessible area	calculates the atomic accessible area when a probe is rolled around the Van der Waal's surface of a macromolecule	calculation of accessible area	Not available	False	False
5	1	structure evaluation	composite scoring function which is able to derive both global (i.e. for the entire structure) and local (i.e. per residue) absolute quality estimates on the basis of one single model	Not available	Not available	False	False
6	1	accessible interaction space	Not available	determine the number of complexes consistent with the restraints	Not available	False	False
7	1	docking	data-driven biomolecular docking	Not available	Not available	False	False

There are 7 software packages reported in this entry.

ID	Software name	Software version	Software classification	Software location
1	I-TASSER	Not available	structure refinement	https://zhanggroup.org/I-TASSER/

ID	Software name	Software version	Software classification	Software location
2	AlphaFold2	Not available	structure prediction	https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb
3	Xwalk	Not available	calculation of Euclidean distances	https://www.xwalk.org/home.cgi
4	NACCESS	Not available	solvent accessibility	http://www.bioinf.manchester.ac.uk/naccess/
5	QMEAN	Not available	Estimation of the quality of protein structure models	https://swissmodel.expasy.org/qmean/
6	DisVis	Not available	accessible interaction space	https://wenmr.science.uu.nl/disvis/
7	HADDOCK	Not available	protein-protein docking	https://wenmr.science.uu.nl/haddock2.4/

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.

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	10.68	138

There are 138 clashes. The table below contains the detailed list of all clashes based on a MolProbity analysis. Bad clashes are >= 0.4 Angstrom. The output is limited to 100 rows.

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:11:ASP:HB3	A:42:THR:HB	0.80	1	1
B:4:GLU:HB3	B:16:ILE:HD11	0.80	1	1
A:593:ILE:HD12	A:682:VAL:HG11	0.77	1	1
A:504:VAL:HA	A:533:ASN:HA	0.74	1	1
A:445:TYR:HB2	A:452:GLU:HB3	0.68	1	1
A:87:ASP:HA	A:107:PRO:HB3	0.67	1	1
A:500:ALA:HA	A:536:PRO:HB3	0.66	1	1
A:280:PHE:HB3	A:359:PRO:HG2	0.65	1	1
A:513:THR:HG23	A:555:LEU:HD23	0.65	1	1
A:29:GLU:HB2	B:20:TRP:HB2	0.65	1	1
A:67:PRO:HA	A:73:THR:HB	0.65	1	1
A:237:LEU:HD23	A:272:VAL:HB	0.64	1	1
A:668:PHE:HB3	A:676:VAL:HB	0.64	1	1
A:559:ASN:HB3	A:584:LEU:HB2	0.64	1	1
A:81:ASP:HA	A:92:VAL:HG22	0.63	1	1
A:509:LEU:HD13	A:544:LEU:HD21	0.61	1	1
A:631:GLU:HB3	B:62:ASN:HD21	0.61	1	1
A:491:VAL:HB	A:544:LEU:HB2	0.60	1	1
A:61:VAL:HB	A:115:TYR:HD2	0.59	1	1
A:226:GLY:O	A:230:CYS:HB2	0.59	1	1
A:623:THR:HB	A:635:THR:HG23	0.59	1	1
A:157:GLN:HG2	A:161:LEU:HD12	0.59	1	1
B:7:THR:HG23	B:13:MET:HG2	0.58	1	1
A:251:PRO:HA	A:254:TRP:CD1	0.58	1	1
A:179:PRO:HG3	A:679:PHE:HB3	0.58	1	1
B:115:TYR:HE1	B:120:ARG:HG2	0.58	1	1
A:251:PRO:HA	A:254:TRP:HD1	0.57	1	1
A:343:THR:HA	A:352:GLU:HG3	0.57	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:26:LEU:HD22	A:185:PHE:HA	0.57	1	1
A:267:HIS:HB3	A:270:GLN:HB2	0.57	1	1
A:307:GLN:HB3	A:313:ILE:HG23	0.56	1	1
A:548:TYR:HA	A:551:TYR:CE2	0.56	1	1
A:624:VAL:HG12	A:666:VAL:HG12	0.56	1	1
A:166:PHE:HE1	A:663:LYS:HE3	0.56	1	1
A:68:SER:HB2	A:71:ALA:HB3	0.56	1	1
A:618:GLU:HG3	A:641:PRO:HB3	0.56	1	1
A:315:TYR:HB2	A:330:MET:SD	0.55	1	1
A:631:GLU:HB3	B:62:ASN:ND2	0.55	1	1
A:62:VAL:HB	A:116:ARG:HB3	0.55	1	1
A:120:GLU:HG2	A:129:SER:HB3	0.54	1	1
A:69:GLN:HA	A:74:LYS:O	0.54	1	1
A:239:GLY:HA2	A:275:GLY:O	0.54	1	1
A:611:ASN:OD1	A:644:ALA:HA	0.54	1	1
A:479:VAL:HB	A:489:PHE:HZ	0.53	1	1
A:669:GLU:HG3	A:674:LYS:HG2	0.53	1	1
A:243:ASN:HB2	A:556:THR:HG21	0.53	1	1
A:213:ARG:HA	A:213:ARG:NE	0.53	1	1
A:14:LEU:HA	A:31:LEU:HD22	0.53	1	1
A:262:ARG:CZ	A:632:GLU:HG2	0.52	1	1
A:320:PHE:HA	A:508:LEU:HD13	0.52	1	1
A:311:LEU:HD13	A:395:ALA:HA	0.52	1	1
B:48:SER:HA	B:53:GLN:HB2	0.52	1	1
B:120:ARG:HB2	B:138:LEU:HB2	0.52	1	1
B:69:LYS:HB3	B:79:CYS:SG	0.50	1	1
A:59:PHE:HB2	A:77:PHE:CE2	0.50	1	1
A:430:SER:HB2	A:433:ARG:HG3	0.50	1	1
A:468:LYS:HD3	A:469:GLU:N	0.50	1	1
A:653:ASP:HB3	B:64:ASN:HB2	0.50	1	1
A:499:THR:HG23	A:501:GLU:H	0.49	1	1
A:70:GLU:CD	A:70:GLU:H	0.49	1	1
A:227:MET:O	A:235:GLY:HA2	0.48	1	1
A:114:LEU:H	A:114:LEU:HD23	0.48	1	1
A:447:GLU:HA	A:452:GLU:CD	0.48	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
B:122:GLN:HB2	B:138:LEU:HD11	0.48	1	1
A:240:ARG:HD3	A:274:TYR:HE1	0.48	1	1
A:163:GLN:HA	A:184:GLN:OE1	0.47	1	1
A:184:GLN:HG2	A:289:ARG:HB2	0.47	1	1
B:31:ARG:HB2	B:47:TYR:HE1	0.47	1	1
A:401:VAL:HB	A:418:ARG:O	0.47	1	1
A:492:PHE:HD2	A:541:SER:HB2	0.47	1	1
A:608:SER:HA	A:648:VAL:O	0.47	1	1
A:328:SER:HB3	A:519:ILE:HA	0.47	1	1
A:258:VAL:HB	A:262:ARG:NH2	0.47	1	1
A:311:LEU:HD12	A:460:ASN:ND2	0.47	1	1
A:106:THR:HG22	A:107:PRO:HD2	0.46	1	1
B:5:ILE:HG22	B:13:MET:HB3	0.46	1	1
A:260:ILE:HG23	A:272:VAL:HG21	0.46	1	1
A:569:GLU:HG3	A:572:ILE:O	0.46	1	1
A:347:LEU:HD22	A:387:LYS:HE3	0.46	1	1
A:460:ASN:ND2	A:464:LYS:HE2	0.46	1	1
A:664:LEU:O	A:679:PHE:HA	0.46	1	1
B:115:TYR:CE1	B:120:ARG:HG2	0.45	1	1
A:38:PRO:HA	A:104:LEU:O	0.45	1	1
A:255:ILE:HD11	A:623:THR:HG23	0.45	1	1
A:441:HIS:HB3	A:444:LYS:O	0.45	1	1
A:610:GLN:O	A:612:PRO:HD3	0.45	1	1
A:447:GLU:CD	A:447:GLU:H	0.45	1	1
A:479:VAL:HB	A:489:PHE:CZ	0.45	1	1
A:180:TRP:CH2	A:279:VAL:HG12	0.44	1	1
A:153:GLU:HA	A:156:ARG:HG2	0.44	1	1
A:509:LEU:HB2	A:527:LYS:HB3	0.44	1	1
A:372:GLY:HA2	A:374:VAL:HG23	0.44	1	1
B:107:GLN:O	B:110:ASP:HB2	0.44	1	1
A:507:LEU:HD23	A:567:LEU:HD12	0.44	1	1
A:33:VAL:HG23	A:137:LEU:HD23	0.43	1	1
B:123:CYS:HA	B:134:HIS:O	0.43	1	1
A:298:VAL:CG2	A:439:ILE:HD11	0.43	1	1
A:360:THR:HG22	A:361:PRO:HD2	0.43	1	1

Atom 1	Atom 2	Clash(Å)	Model ID (Worst)	Models (Total)
A:87:ASP:CA	A:107:PRO:HB3	0.43	1	1
A:557:GLU:CD	A:557:GLU:H	0.43	1	1

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	823	746	67	10

There are 10 unique backbone outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	47	GLY	1
A	94	ASP	1
A	108	ALA	1
A	230	CYS	1
A	301	TYR	1
A	431	VAL	1
A	620	CYS	1
A	632	GLU	1
A	640	ASP	1
A	686	PRO	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	719	507	115	97

There are 97 unique sidechain outliers. Detailed list of outliers are tabulated below.

Chain	Res	Type	Models (Total)
A	5	LEU	1
A	26	LEU	1
A	46	GLU	1
A	54	VAL	1
A	58	THR	1
A	70	GLU	1
A	73	THR	1

Chain	Res	Type	Models (Total)
A	89	THR	1
A	93	VAL	1
A	97	ASP	1
A	99	THR	1
A	100	LEU	1
A	101	SER	1
A	102	LEU	1
A	103	GLN	1
A	106	THR	1
A	112	ILE	1
A	154	GLU	1
A	184	GLN	1
A	186	GLU	1
A	196	LEU	1
A	200	ASN	1
A	206	ASN	1
A	209	ARG	1
A	211	CYS	1
A	215	SER	1
A	227	MET	1
A	238	LEU	1
A	250	SER	1
A	264	TRP	1
A	267	HIS	1
A	271	ARG	1
A	299	THR	1
A	307	GLN	1
A	313	ILE	1
A	314	GLU	1
A	324	GLN	1
A	336	CYS	1
A	346	ASP	1
A	347	LEU	1
A	362	GLN	1
A	363	GLU	1

Chain	Res	Type	Models (Total)
A	364	LYS	1
A	368	THR	1
A	386	THR	1
A	396	GLU	1
A	398	ASN	1
A	403	ASP	1
A	405	ILE	1
A	420	LEU	1
A	435	GLU	1
A	442	THR	1
A	450	SER	1
A	463	ASN	1
A	471	THR	1
A	478	ARG	1
A	483	MET	1
A	485	MET	1
A	507	LEU	1
A	508	LEU	1
A	526	THR	1
A	527	LYS	1
A	531	ASN	1
A	532	LEU	1
A	534	LEU	1
A	541	SER	1
A	549	GLU	1
A	552	ARG	1
A	564	ARG	1
A	574	SER	1
A	576	LEU	1
A	577	LEU	1
A	580	ARG	1
A	588	GLU	1
A	602	LYS	1
A	618	GLU	1
A	620	CYS	1

Chain	Res	Type	Models (Total)
A	621	THR	1
A	637	GLU	1
A	640	ASP	1
A	646	GLU	1
A	657	LEU	1
A	664	LEU	1
A	683	ILE	1
B	15	ARG	1
B	18	ASP	1
B	43	THR	1
B	47	TYR	1
B	48	SER	1
B	57	ASP	1
B	63	VAL	1
B	72	GLU	1
B	80	THR	1
B	100	SER	1
B	110	ASP	1
B	132	GLU	1
B	140	THR	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 47 crosslinking restraints combined in 1 restraint groups.

Linker	Residue 1	Atom 1	Residue 2	Atom 2	Restraint type	Distance, Å	Count
DSS	LYS	CA	LYS	CA	upper bound	35.00	27
DSS	LYS	CA	LYS	CA	upper bound	25.00	20

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: upper bound, 35.0 Å

Model Group 1; Self-links: upper bound, 25.0 Å

Model Group 1; Self-links: upper bound, 35.0 Å
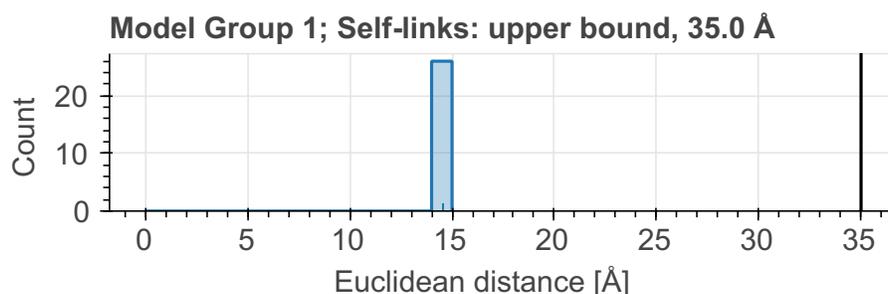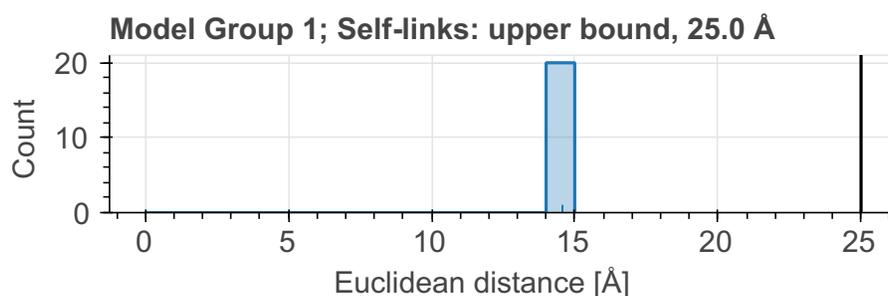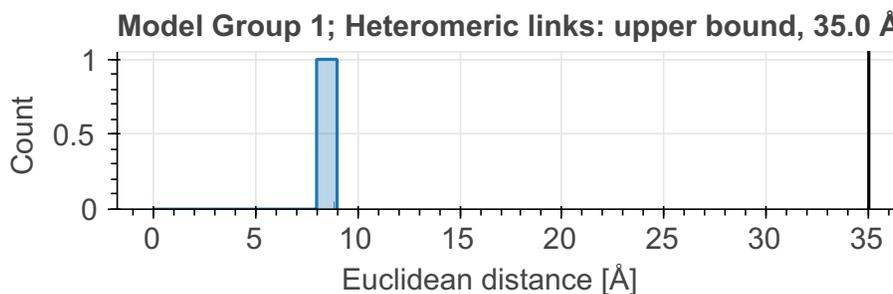
5.2.2. Satisfaction of restraints ?

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=1)
1	1	1	1/1	All	100.00	0.00	1
				Ambiguous entity/ Ambiguous	100.00	0.00	1

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
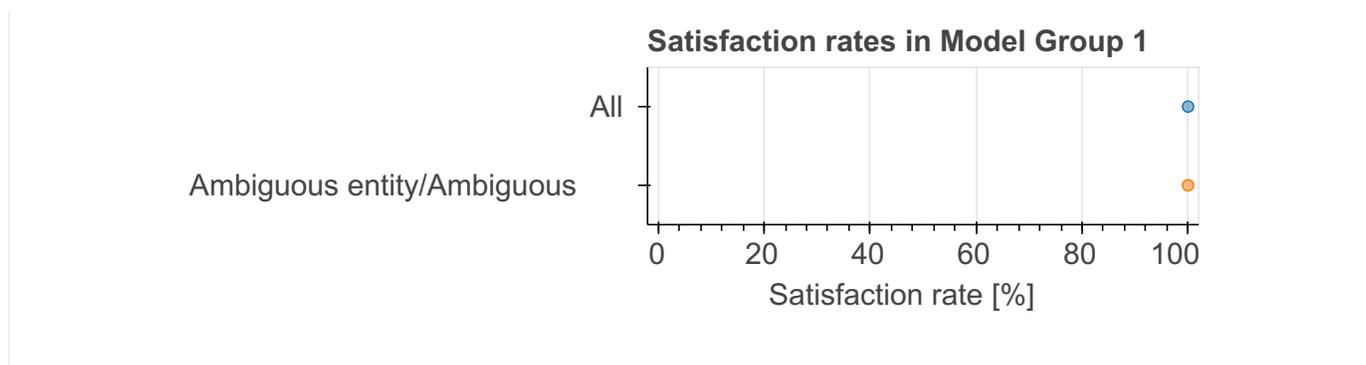
6. Fit to Data Used for Validation Assessment ❓

Validation for this section is under development.