Integrative Structure Validation Report

July 22, 2024 - 05:29 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	9A3Y
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.

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

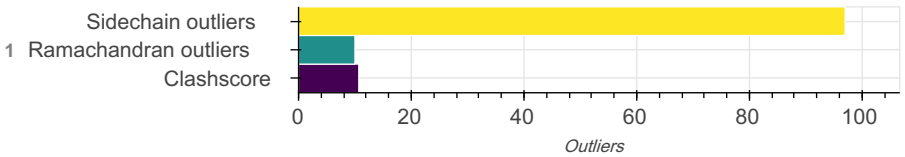
A user guide is available at https://pdb-dev.wwpdb.org/validation_help.html with specific help available everywhere you see the ? symbol.

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 4 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	Protein-glutamine gamma-glutamyltransferase 2	A	A	687
1	2	2	Fibronectin	B	B	140

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	4PYG
3	De Novo model	Not available	Not available
4	Experimental model	PDB	3EJH

Representation

This entry has only one representation and includes 0 rigid bodies and 2 flexible units

Chain ID	Rigid bodies	Non-rigid segments
A	-	1-687
B	-	1-140

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	Iterative Threading ASSEmbly Refinement	Refinement	structural refinement of starting models guided by experimental crosslinks	None	False	False
2	1	AI prediction of protein 3D structure from its amino acid sequence	structure prediction	structure prediciton of FN type I modules 7-9	None	False	False
3	1	Validate measured chemical cross-links on a protein 3D structure	satisfied/violated crosslinks identification	calculation of Euclidean distances between crosslinked residues	None	False	False
4	1	calculates the atomic accessible area when a probe is rolled around the Van der Waal's surface of a macromolecule	calculation of accessible area	calculation of accessible area	None	False	False
5	1	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	structure evaluation	None	None	False	False
6	1	None	accessible interaction space	determine the number of complexes consistent with the restraints	None	False	False
7	1	data-driven biomolecular docking	docking	None	None	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/

Data quality

Crosslinking-MS

Validation for this section is under development.

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.

Standard geometry: bond outliers

There are 1469 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
N--H	0.98	0.86	726
ND2--HD22	0.98	0.86	38
NE2--HE22	0.98	0.86	35
ND2--HD21	0.98	0.86	32
NE2--HE2	0.98	0.86	15
NE2--HE21	0.98	0.86	32

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
ND1--HD1	0.98	0.86	3
NE1--HE1	0.98	0.86	15
OG--HG	0.96	0.84	15
OH--HH	0.96	0.84	9
OG1--HG1	0.96	0.84	14
NE--HE	0.98	0.86	20
SG--HG	1.32	1.20	2
OG--HG	0.97	0.84	24
NE--HE	0.99	0.86	25
NE2--HE22	0.99	0.86	1
N--H	0.99	0.86	65
OH--HH	0.97	0.84	19
OG1--HG1	0.97	0.84	28
ND2--HD21	0.99	0.86	8
NE1--HE1	0.99	0.86	3
SG--HG	1.33	1.20	18
NE2--HE2	0.99	0.86	4
NE2--HE21	0.99	0.86	4
ND2--HD22	0.99	0.86	2
NE--HE	1.00	0.86	3
OG--HG	0.98	0.84	4
OG1--HG1	0.98	0.84	2
NH1--HH11	1.00	0.86	33
OH--HH	0.98	0.84	3
NH2--HH21	1.00	0.86	29
NH1--HH12	1.00	0.86	36
NH2--HH22	1.00	0.86	24
NH2--HH22	1.01	0.86	20

Bond type	Observed distance (Å)	Ideal distance (Å)	Number of outliers
NH1--HH12	1.01	0.86	11
NH2--HH21	1.01	0.86	17
NH1--HH11	1.01	0.86	15
NZ--HZ2	1.04	0.89	31
NZ--HZ3	1.04	0.89	28
NZ--HZ1	1.04	0.89	28
NH2--HH21	1.02	0.86	2
NZ--HZ2	1.05	0.89	5
NZ--HZ3	1.05	0.89	8
NH2--HH22	1.02	0.86	4
NZ--HZ1	1.05	0.89	8
NH1--HH12	1.02	0.86	1

Standard geometry: angle outliers

Bond angle outliers do not exist or can not be evaluated for this model

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	10.68	138

All 138 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	A:11:ASP:HB3	A:42:THR:HB	0.797
1	B:4:GLU:HB3	B:16:ILE:HD11	0.795
1	A:593:ILE:HD12	A:682:VAL:HG11	0.771
1	A:504:VAL:HA	A:533:ASN:HA	0.745
1	A:445:TYR:HB2	A:452:GLU:HB3	0.678
1	A:87:ASP:HA	A:107:PRO:HB3	0.667
1	A:500:ALA:HA	A:536:PRO:HB3	0.663
1	A:280:PHE:HB3	A:359:PRO:HG2	0.650

Model ID	Atom-1	Atom-2	Clash overlap (Å)
1	A:513:THR:HG23	A:555:LEU:HD23	0.648
1	A:29:GLU:HB2	B:20:TRP:HB2	0.645
1	A:67:PRO:HA	A:73:THR:HB	0.645
1	A:237:LEU:HD23	A:272:VAL:HB	0.642
1	A:668:PHE:HB3	A:676:VAL:HB	0.642
1	A:559:ASN:HB3	A:584:LEU:HB2	0.636
1	A:81:ASP:HA	A:92:VAL:HG22	0.628
1	A:509:LEU:HD13	A:544:LEU:HD21	0.614
1	A:631:GLU:HB3	B:62:ASN:HD21	0.606
1	A:491:VAL:HB	A:544:LEU:HB2	0.597
1	A:61:VAL:HB	A:115:TYR:HD2	0.592
1	A:226:GLY:O	A:230:CYS:HB2	0.592
1	A:623:THR:HB	A:635:THR:HG23	0.590
1	A:157:GLN:HG2	A:161:LEU:HD12	0.587
1	B:7:THR:HG23	B:13:MET:HG2	0.585
1	A:251:PRO:HA	A:254:TRP:CD1	0.581
1	A:179:PRO:HG3	A:679:PHE:HB3	0.577
1	B:115:TYR:HE1	B:120:ARG:HG2	0.576
1	A:251:PRO:HA	A:254:TRP:HD1	0.573
1	A:343:THR:HA	A:352:GLU:HG3	0.573
1	A:26:LEU:HD22	A:185:PHE:HA	0.570
1	A:267:HIS:HB3	A:270:GLN:HB2	0.568
1	A:307:GLN:HB3	A:313:ILE:HG23	0.563
1	A:548:TYR:HA	A:551:TYR:CE2	0.562
1	A:624:VAL:HG12	A:666:VAL:HG12	0.561
1	A:166:PHE:HE2	A:663:LYS:HE3	0.558
1	A:68:SER:HB2	A:71:ALA:HB3	0.556
1	A:618:GLU:HG3	A:641:PRO:HB3	0.556

Model ID	Atom-1	Atom-2	Clash overlap (Å)
1	A:315:TYR:HB2	A:330:MET:SD	0.552
1	A:631:GLU:HB3	B:62:ASN:ND2	0.547
1	A:62:VAL:HB	A:116:ARG:HB3	0.546
1	A:120:GLU:HG2	A:129:SER:HB3	0.540
1	A:69:GLN:HA	A:74:LYS:O	0.537
1	A:239:GLY:HA2	A:275:GLY:O	0.536
1	A:611:ASN:OD1	A:644:ALA:HA	0.535
1	A:479:VAL:HB	A:489:PHE:HZ	0.534
1	A:669:GLU:HG3	A:674:LYS:HG2	0.534
1	A:243:ASN:HB2	A:556:THR:HG21	0.531
1	A:213:ARG:HA	A:213:ARG:NE	0.530
1	A:14:LEU:HA	A:31:LEU:HD22	0.526
1	A:262:ARG:CZ	A:632:GLU:HG2	0.519
1	A:320:PHE:HA	A:508:LEU:HD13	0.519
1	A:311:LEU:HD13	A:395:ALA:HA	0.517
1	B:48:SER:HA	B:53:GLN:HB2	0.516
1	B:120:ARG:HB2	B:138:LEU:HB2	0.516
1	B:69:LYS:HB3	B:79:CYS:SG	0.503
1	A:59:PHE:HB2	A:77:PHE:CE1	0.501
1	A:430:SER:HB2	A:433:ARG:HG3	0.498
1	A:468:LYS:HD3	A:469:GLU:N	0.496
1	A:653:ASP:HB3	B:64:ASN:HB2	0.496
1	A:499:THR:HG23	A:501:GLU:H	0.494
1	A:70:GLU:CD	A:70:GLU:H	0.486
1	A:227:MET:O	A:235:GLY:HA2	0.485
1	A:114:LEU:H	A:114:LEU:HD23	0.484
1	A:447:GLU:HA	A:452:GLU:CD	0.483
1	B:122:GLN:HB2	B:138:LEU:HD11	0.483

Model ID	Atom-1	Atom-2	Clash overlap (Å)
1	A:240:ARG:HD3	A:274:TYR:HE1	0.476
1	A:163:GLN:HA	A:184:GLN:OE1	0.475
1	A:184:GLN:HG2	A:289:ARG:HB2	0.475
1	B:31:ARG:HB2	B:47:TYR:HE2	0.474
1	A:401:VAL:HB	A:418:ARG:O	0.471
1	A:492:PHE:HD1	A:541:SER:HB2	0.470
1	A:608:SER:HA	A:648:VAL:O	0.470
1	A:328:SER:HB3	A:519:ILE:HA	0.468
1	A:258:VAL:HB	A:262:ARG:NH2	0.467
1	A:311:LEU:HD12	A:460:ASN:ND2	0.466
1	A:106:THR:HG22	A:107:PRO:HD2	0.462
1	B:5:ILE:HG22	B:13:MET:HB3	0.462
1	A:260:ILE:HG23	A:272:VAL:HG21	0.460
1	A:569:GLU:HG3	A:572:ILE:O	0.460
1	A:347:LEU:HD22	A:387:LYS:HE3	0.459
1	A:460:ASN:ND2	A:464:LYS:HE2	0.458
1	A:664:LEU:O	A:679:PHE:HA	0.455
1	B:115:TYR:CE1	B:120:ARG:HG2	0.452
1	A:38:PRO:HA	A:104:LEU:O	0.451
1	A:255:ILE:HD11	A:623:THR:HG23	0.449
1	A:441:HIS:HB3	A:444:LYS:O	0.449
1	A:610:GLN:O	A:612:PRO:HD3	0.448
1	A:447:GLU:CD	A:447:GLU:H	0.447
1	A:479:VAL:HB	A:489:PHE:CZ	0.445
1	A:180:TRP:CH2	A:279:VAL:HG12	0.444
1	A:153:GLU:HA	A:156:ARG:HG2	0.443
1	A:509:LEU:HB2	A:527:LYS:HB3	0.442

Model ID	Atom-1	Atom-2	Clash overlap (Å)
1	A:372:GLY:HA2	A:374:VAL:HG23	0.440
1	B:107:GLN:O	B:110:ASP:HB2	0.438
1	A:507:LEU:HD23	A:567:LEU:HD12	0.437
1	A:33:VAL:HG23	A:137:LEU:HD23	0.434
1	B:123:CYS:HA	B:134:HIS:O	0.433
1	A:298:VAL:CG2	A:439:ILE:HD11	0.432
1	A:360:THR:HG22	A:361:PRO:HD2	0.432
1	A:87:ASP:CA	A:107:PRO:HB3	0.431
1	A:557:GLU:CD	A:557:GLU:H	0.431
1	A:447:GLU:HA	A:452:GLU:OE1	0.430
1	A:47:GLY:HA2	A:98:CYS:SG	0.428
1	A:365:SER:O	A:366:GLU:HB3	0.428
1	A:240:ARG:HD3	A:274:TYR:CE1	0.427
1	A:311:LEU:HD22	A:395:ALA:CB	0.427
1	A:184:GLN:HG2	A:289:ARG:CB	0.426
1	A:2:ALA:HB1	A:6:VAL:HG21	0.425
1	B:79:CYS:HA	B:91:CYS:HA	0.425
1	A:196:LEU:HG	A:227:MET:SD	0.424
1	A:478:ARG:CZ	A:492:PHE:HB2	0.423
1	A:293:ILE:HA	A:294:PRO:HD3	0.422
1	A:311:LEU:O	A:399:ALA:HA	0.421
1	A:565:ALA:HB3	A:578:ALA:HB3	0.420
1	B:68:HIS:HB3	B:76:MET:HE3	0.420
1	B:92:ASP:HA	B:93:PRO:HD3	0.420
1	A:394:PHE:O	A:398:ASN:HB2	0.419
1	A:379:ILE:HG22	A:394:PHE:HB2	0.418
1	A:636:VAL:HG12	B:66:THR:OG1	0.417

Model ID	Atom-1	Atom-2	Clash overlap (Å)
1	A:62:VAL:HG13	A:67:PRO:HG2	0.416
1	A:262:ARG:NH2	A:632:GLU:HG2	0.416
1	A:9:ARG:HB2	A:44:HIS:ND1	0.415
1	A:167:ILE:HB	A:178:ILE:HG23	0.415
1	B:84:GLN:OE1	B:88:ARG:HD3	0.415
1	B:72:GLU:H	B:72:GLU:HG2	0.415
1	B:114:LYS:HE2	B:123:CYS:SG	0.414
1	A:589:ILE:HA	A:609:LEU:HD13	0.412
1	A:603:LEU:HB3	A:654:LEU:HB2	0.412
1	A:347:LEU:HB3	A:348:GLN:H	0.411
1	B:83:GLY:HA2	B:87:GLY:HA2	0.410
1	A:508:LEU:HG	A:529:LEU:HB3	0.408
1	B:82:PHE:HE2	B:90:LYS:HB3	0.408
1	A:360:THR:CG2	A:361:PRO:HD2	0.406
1	A:485:MET:HA	A:584:LEU:HD23	0.404
1	A:116:ARG:HD2	A:134:HIS:HB3	0.403
1	A:235:GLY:O	A:271:ARG:HB3	0.403
1	A:302:ASN:HA	A:302:ASN:HD22	0.403
1	A:507:LEU:O	A:529:LEU:HA	0.402
1	A:112:ILE:H	A:112:ILE:HD13	0.401

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	823	746	67	10

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

Model ID	Analyzed	Favored	Allowed	Outliers
1	719	507	115	97

Detailed list of outliers are tabulated below.

Model ID	Chain	Residue ID	Residue type
1	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	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

Model ID	Chain	Residue ID	Residue type
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	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

Model ID	Chain	Residue ID	Residue type
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	A	621	THR
1	A	637	GLU
1	A	640	ASP
1	A	646	GLU

Model ID	Chain	Residue ID	Residue type
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

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 labcontributed 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.