

wwPDB EM Validation Summary Report (i)

Dec 6, 2023 - 04:20 pm GMT

PDB ID	:	8AT3
EMDB ID	:	EMD-15632
Title	:	Structure of the augmin holocomplex in open conformation
Authors	:	Zupa, E.; Pfeffer, S.
Deposited on	:	2022-08-22
Resolution	:	33.00 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org* A user guide is available at https://www.wwpdb.org/validation/2017/EMValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

EMDB validation analysis	:	0.0.1.dev 70
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
MapQ	:	FAILED
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $ELECTRON\ MICROSCOPY$

The reported resolution of this entry is 33.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Matria	Whole archive	EM structures
INTEGLIC	$(\# { m Entries})$	$(\# { m Entries})$
Clashscore	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Qualit	y of chain
1	А	286	59%	28% 8% 5%
2	В	597	57%	27% 9% 7%
3	С	353	60%	33% 5% •
4	D	666	56%	32% 6% 5%
5	Е	222	48%	37% 9% •
6	F	978	22% 13% ••	60%
7	G	348	47%	36% <u>9</u> % 5% •
8	Н	367	33% 18%	• 45%



2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 24599 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

• Molecule 1 is a protein called HAUS augmin-like complex subunit 1.

Mol	Chain	Residues		At	AltConf	Trace			
1	А	286	Total 2282	C 1436	N 380	0 453	S 13	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
А	156	ARG	GLN	variant	UNP Q3B8L5

• Molecule 2 is a protein called HAUS augmin-like complex subunit 3.

Mol	Chain	Residues		At	AltConf	Trace			
2	В	597	Total 4771	C 2988	N 817	0 943	S 23	0	0

• Molecule 3 is a protein called HAUS augmin like complex subunit 4 L homeolog.

Mol	Chain	Residues		At	AltConf	Trace			
3	С	353	Total 2885	C 1807	N 508	0 554	S 16	0	0

• Molecule 4 is a protein called HAUS augmin-like complex subunit 5.

Mol	Chain	Residues		Α	AltConf	Trace			
4	D	666	Total 5415	C 3362	N 1000	O 1022	S 31	0	0

• Molecule 5 is a protein called HAUS augmin like complex subunit 2 L homeolog.

Mol	Chain	Residues		At	AltConf	Trace			
5	Е	217	Total 1717	C 1075	N 296	0 334	S 12	0	0

• Molecule 6 is a protein called HAUS augmin like complex subunit 6 L homeolog.



Mol	Chain	Residues		At		AltConf	Trace		
6	F	387	Total 3171	C 2020	N 574	O 558	S 19	0	0

• Molecule 7 is a protein called HAUS augmin like complex subunit 7 S homeolog.

Mol	Chain	Residues		At	AltConf	Trace			
7	G	339	Total 2687	C 1694	N 441	O 533	S 19	0	0

• Molecule 8 is a protein called HAUS augmin-like complex subunit 8.

Mol	Chain	Residues		At	AltConf	Trace			
8	Н	203	Total 1671	C 1048	N 285	0 331	${f S}{7}$	0	0



3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.



 \bullet Molecule 1: HAUS augmin-like complex subunit 1

 \bullet Molecule 3: HAUS augmin like complex subunit 4 L homeolog







• Molecule 5: HAUS augmin like complex subunit 2 L homeolog



IT 9 17 9 1810 1811 1812 1823 1834 1845 1845 1845 1845 1845 1845 1845 1846 1846 1846 1846 1846 1846 1846 1847 1846 1847 1847 1848 1849 1849 1840 1411 1412 1412 1413 1414 1413 1414 1413 1413 1413 1413 1414 1415 1416 1416 1416 1416 1416 1416 1416 1414

• Molecule 6: HAUS augmin like complex subunit 6 L homeolog

Ch	Chain F: 22%						13% ••						60%																											
M1 Q2	S3	SS	R6 P7	H8 T 0	L3 A10	B13	CTN	M16	L23	-	827	628	A34	G35	K36 T27	L38	V39	E73	644		M47	F48	P51	N52 VE3		Y57	F60	1	r04 G65	K66 167	D68	180 1	P81	P82	L83 D84	00 10	NS/ D88	BOJ	100	498 L99
S103	V106	G107	A108 G109	F110 D111		V114	S116	1117 5110	F118 L119	S120	P121	6122 6123	P124		L130	R135		M138	A 146	D147	A148	A156	L157	0,158 c 160	0012	D163	г 10 4 Q165	K166	A16/ L168	A169 B170	171N	K172 L173	A174	R175	0/Th	L179	V181	L182	E185	N186 L187
E191	Y192	R194	K195 A196	Q197	q 202	TOOR BOOR	M206		ROZH	L213	, , , , , ,	47.16	L219	A220	E221	V223	D224	R225 VDD6	1927	S228	D229	K230	12 <mark>34</mark>	TO 37	R238	C239 M240	M241		1244 M245	Q246 M247		E250 M251	E252	K253 5754	E254 V255	D256 VD57	V25/ V258	D259 A260	V261	V262 R263
GLY ASN	ILE	NID	TYR CYS	LEU	GLY	THR NOTE	A276	T277 1070	L2/8 N279	1280	P281	N282 1.283	L284		1287 1288		M292	H293	1.295	Q296	M297	Y301		G304 V305	V306		E OCT	0313		E317	L319	K320 L321		E325	K320	Y329	0331 C331	K332 C333	V334	R335 L336
D337 L338	0339 V340	L341	K344	D 2 E O	H360	1362	2004	D366	1370		R377	F380	W381		K384	L389	G390	K391	P393		K398	GLY LEU	ASN	PRO AT A	TEU	GLU	ASN	PRO	MET	ALA	PHE	SER PHE	ASP	PRO AT A	ALA SER	CLU GLU	VAL	LEU	SER	SER VAL
PHE CYS	HIS	PRO	ALA SER	LEU	GLU	TYR	SIH	TAS	GLU UTS	PRO	LEU	L Y S A SP	PHE	ASN	0 LLN	SIH	SER	GLY	U.F.U	VAL	GLN	ARG	ILE	GLY	THR	VAL	THR	ARG	GLY	ARG	SER	THR SER	SER	LEU	MET	ALA	PRO	ASN	ARG	ARG MET
SER LEU	ASN	ARG	GLU	GLN	PRO	THR	ASN	ASP	AKG	GLY	PHE	GLN ARG	THR	PRO	SER	ALA	VAL	GLN GLN	ARG	ARG	ALA	ASP VAL	SER	TRP I VS	THR	ALA	ALA	SER	LEU	PRO	THR	PRO THR	PRO	T YR I VS	GLN	ASP	LYS	ASN	ALA	GLN
GLN	ALA	GLN	VAL ALA	ASP	ILE	VAL	GLU	SER	ARG	SER	SER	AT 10	ARG	GLY	UET.	TEU	ASP	ASP I EII	LEU LEU	GLY	VAL	LEU	SER	ASP	PHE	LEU	ARG	LYS	ILE	PRO	THR	PRO GLU	ASN	LEU	SER	ASP	ARG	SER	TRP	ARG LYS
ALA ILE	GLN	GLU	GLU SER	LEU	VAL	ALA	PRO	VAL	ALA ALA	PRO	CYS	MET.	SER	THR	ALA GI II	TEU	GLU	SER AT A	HLS	CYS	SER	GLN	ASP	LEU	MET	ALA	PHE	LEU	THR	SER	ALA	SER	GLN	ASN	CYS	SER	THR	ARG	PRO	ASN
THR GLY	GLY	LYS	ALA SER	SER	HIS	SER	VAL	VAL	SIH	GLU	SER	ASP	MET	LYS	SER	ILE	ASP	SER	CI.N	PRO	GLU	ASP LEU	SER	LEU	ILE	LYS	ASN	GLU	ALA	ASP	TEU	ILE ASN	LEU	LEU	ASP	LYS	GLU	SER	ASN	ALA ASP
THR LEU	THR	PRO	VAL GLU	PHE	PHE	SER	GLN	PRO	SER	LYS	LEU	ASP	TAS	THR	DHF	ILE	SER	VAL	VID VID	GLN	GLU	ASN LEU	SER	ALA	THR	THR	SER	TRP	SER	SER	MET	THR	SER	ASP	SER	SER	THR	HIS	ILE	GLN
PHE GLY	ILE	SIH	GLU THR	LEU	GLU	ASN AT A	GLY	ASN	VAL SER	LEU	ASN	THR	TEU	SER		GLY	ASN	GLU	PRO	PHE	GLU	LYS SER	GLU	LEU	ASP	SIH	PHE	THR	ASP	ARG T VS	GLU	HIS SER	SER	ARG	MET	GLU	LYS	MET	ILE	ASN SER
ILE ARG	SER	TYR	GLU ALA	LEU	ARG	THR	THR	SER	THR	ASP	GLU	GLU TVR	HIS	GLY	ASP	ASP	THR	SER	MET	ARG	PHE	THR LYS	SIH	LYS	GLU	SER	LEU	ILE	GLY	SER	VAL	TYR SER	PRO	VAL	CT O	VAL	SER	LEU	LEU	GLU TYR
LEU THR	THR	SER	PRO LYS	ASP	LYS	LEU	LEU	PRO GIN	GLN	ILE	SER	SER	PRO	GLU	TI F	ARG	SER	GLU	GI.N	GLU	ASP	LEU	ASP	VAL	GLU	SER	ASP	PHE	ASN	PRO	SAT	THR PHE	ASP	PHE	SER	SER	LEU	ASP	GLN	LYS SER

THR ASP GLU GLU GLU GLU CLEU ILEU LEU LEU









4 Experimental information (i)

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	11969	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE	Depositor
	CORRECTION	
Microscope	TFS TALOS L120C	Depositor
Voltage (kV)	120	Depositor
Electron dose $(e^-/\text{\AA}^2)$	101.8	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	Not provided	
Image detector	FEI CETA $(4k \ge 4k)$	Depositor



5 Model quality (i)

5.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with |Z| > 5 is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mal	Chain	B	ond lengths	Bond angles				
	Ullalli	RMSZ	# Z > 5	RMSZ	# Z > 5			
1	А	1.47	38/2309~(1.6%)	1.41	48/3102~(1.5%)			
2	В	1.19	41/4836~(0.8%)	1.36	76/6496~(1.2%)			
3	С	0.99	23/2920~(0.8%)	1.18	29/3925~(0.7%)			
4	D	1.10	55/5502~(1.0%)	1.24	61/7397~(0.8%)			
5	Е	1.34	40/1743~(2.3%)	1.08	17/2359~(0.7%)			
6	F	1.22	55/3229~(1.7%)	1.12	34/4333~(0.8%)			
7	G	1.19	48/2736~(1.8%)	1.31	37/3698~(1.0%)			
8	Н	1.12	19/1692~(1.1%)	0.95	6/2278~(0.3%)			
All	All	1.19	319/24967~(1.3%)	1.24	308/33588~(0.9%)			

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	А	0	14
2	В	0	68
3	С	0	11
4	D	0	39
5	Ε	0	11
6	F	0	19
7	G	0	27
8	Н	0	1
All	All	0	190

The worst 5 of 319 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms		Observed(Å)	$\mathrm{Ideal}(\mathrm{\AA})$
2	В	223	GLU	C-N	26.97	1.81	1.33
1	А	277	VAL	C-N	24.03	1.79	1.34
4	D	288	SER	C-N	19.16	1.78	1.34
2	В	591	THR	C-N	18.66	1.66	1.33

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	А	280	PRO	C-N	17.83	1.75	1.34

The worst 5 of 308 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
7	G	130	ILE	O-C-N	-21.95	87.58	122.70
4	D	286	GLU	O-C-N	-21.61	88.13	122.70
3	С	117	THR	O-C-N	-21.59	88.16	122.70
2	В	392	GLU	O-C-N	-21.40	88.46	122.70
3	С	122	PRO	O-C-N	-21.14	80.94	121.10

There are no chirality outliers.

5 of 190 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	А	241	LEU	Mainchain
1	А	242	ALA	Mainchain
1	А	243	PRO	Mainchain
1	А	273	MET	Peptide
1	А	274	GLU	Mainchain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	А	2282	0	2354	224	0
2	В	4771	0	4777	438	0
3	С	2885	0	2950	242	0
4	D	5415	0	5436	545	0
5	Е	1717	0	1709	197	0
6	F	3171	0	3232	260	0
7	G	2687	0	2641	264	0
8	Н	1671	0	1673	218	0
All	All	24599	0	24772	1532	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 31.



Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:E:155:ILE:CG2	6:F:334:VAL:HG21	1.33	1.58
1:A:278:PRO:HB2	1:A:279:GLU:CB	1.32	1.57
2:B:490:GLU:CA	4:D:560:ARG:NH2	1.71	1.51
1:A:281:SER:CB	1:A:283:ARG:HD2	1.41	1.48
1:A:284:ARG:C	1:A:285:LEU:N	1.69	1.45

The worst 5 of 1532 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	\mathbf{P}	ercentiles		
1	А	284/286~(99%)	273 (96%)	4 (1%)	7 (2%)		5	32	
2	В	595/597~(100%)	542 (91%)	18 (3%)	35~(6%)		1	17	
3	С	351/353~(99%)	338 (96%)	7 (2%)	6 (2%)		9	42	
4	D	662/666~(99%)	622 (94%)	19 (3%)	21 (3%)		4	26	
5	Е	213/222~(96%)	203 (95%)	5 (2%)	5 (2%)		6	34	
6	F	383/978~(39%)	365~(95%)	9 (2%)	9 (2%)		6	34	
7	G	335/348~(96%)	311 (93%)	9 (3%)	15 (4%)		2	22	
8	Н	199/367~(54%)	197 (99%)	1 (0%)	1 (0%)		29	69	
All	All	3022/3817~(79%)	2851 (94%)	72 (2%)	99 (3%)		6	26	

5 of 99 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	А	272	MET
1	А	277	VAL
1	А	278	PRO

Continued on next page...



Continued from previous page...

Mol	Chain	Res	Type
1	А	279	GLU
1	А	282	LYS

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Perce	entiles
1	А	260/260~(100%)	248~(95%)	12 (5%)	27	52
2	В	541/541~(100%)	506 (94%)	35~(6%)	17	42
3	С	326/326~(100%)	312~(96%)	14 (4%)	29	53
4	D	605/605~(100%)	570 (94%)	35~(6%)	20	45
5	Ε	190/195~(97%)	176~(93%)	14 (7%)	13	38
6	F	343/882~(39%)	327~(95%)	16 (5%)	26	51
7	G	313/320~(98%)	295~(94%)	18 (6%)	20	45
8	Н	192/328~(58%)	184 (96%)	8 (4%)	30	54
All	All	2770/3457~(80%)	2618 (94%)	152 (6%)	25	47

5 of 152 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	F	173	LEU
7	G	344	SER
6	F	278	LEU
7	G	48	LEU
8	Н	363	ASP

Sometimes side chains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 35 such side chains are listed below:

Mol	Chain	Res	Type
7	G	18	GLN
7	G	92	HIS
8	Н	196	ASN

Continued on next page...



Continued from previous page...

Mol	Chain	\mathbf{Res}	Type
4	D	68	GLN
4	D	49	HIS

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
5	Ε	18
6	F	17
4	D	14
2	В	13
1	А	13
7	G	9
8	Н	7
3	С	6

The worst 5 of 97 chain breaks are listed below:



Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	298:ASN	С	299:GLU	N	4.29
1	В	223:GLU	С	224:GLY	N	1.81
1	А	277:VAL	С	278:PRO	N	1.79
1	D	288:SER	С	289:GLU	N	1.78
1	А	280:PRO	С	281:SER	N	1.75



6 Map visualisation (i)

This section contains visualisations of the EMDB entry EMD-15632. These allow visual inspection of the internal detail of the map and identification of artifacts.

Images derived from a raw map, generated by summing the deposited half-maps, are presented below the corresponding image components of the primary map to allow further visual inspection and comparison with those of the primary map.

6.1 Orthogonal projections (i)

This section was not generated.

6.2 Central slices (i)

This section was not generated.

6.3 Largest variance slices (i)

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color) (i)

This section was not generated.

6.5 Orthogonal surface views (i)

This section was not generated.

6.6 Mask visualisation (i)

This section was not generated. No masks/segmentation were deposited.



7 Map analysis (i)

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution (i)

This section was not generated.

7.2 Volume estimate versus contour level (i)

This section was not generated.

7.3 Rotationally averaged power spectrum (i)

This section was not generated. The rotationally averaged power spectrum had issues being displayed.



8 Fourier-Shell correlation (i)

This section was not generated. No FSC curve or half-maps provided.



9 Map-model fit (i)

This section was not generated.

