WORLDWIDE PDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 21, 2024 – 06:49 AM EDT

PDB ID	:	2H43
Title	:	Crystal Structure of Human Fragment D Complexed with Ala-His-Arg-Pro-amide
Authors	:	Doolittle, R.F.; Pandi, L.
Deposited on	:	2006-05-23
Resolution	:	2.70 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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.
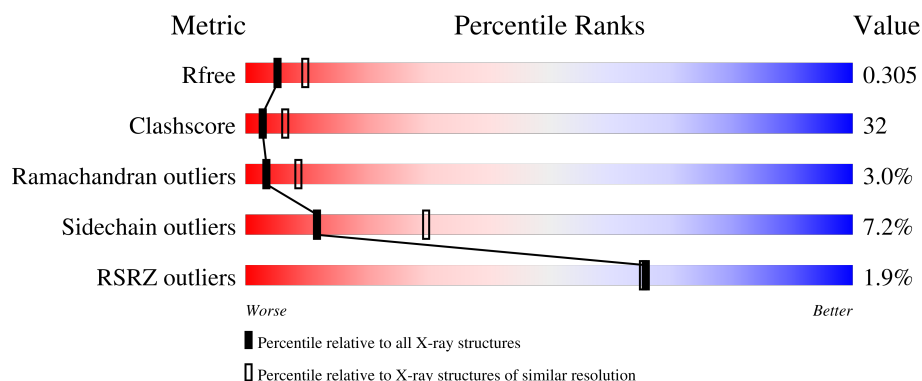
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	3333 (2.70-2.70)
Clashscore	180529	3684 (2.70-2.70)
Ramachandran outliers	177936	3633 (2.70-2.70)
Sidechain outliers	177891	3633 (2.70-2.70)
RSRZ outliers	164620	3333 (2.70-2.70)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower 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% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	87	
1	D	87	
2	B	328	
2	E	328	
3	C	323	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	323	47% 39% 10%
4	I	5	20% 80%
4	J	5	60% 40%
5	G	2	50% 50%
6	H	2	50% 50%

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 10770 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Fibrinogen alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	70	Total	C	N	O	S	0	0	0
			570	354	106	107	3			
1	D	70	Total	C	N	O	S	0	0	0
			570	354	106	107	3			

- Molecule 2 is a protein called Fibrinogen beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	302	Total	C	N	O	S	0	0	0
			2421	1510	428	461	22			
2	E	302	Total	C	N	O	S	0	0	0
			2421	1510	428	461	22			

- Molecule 3 is a protein called Fibrinogen gamma chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	290	Total	C	N	O	S	0	0	0
			2329	1477	394	447	11			
3	F	290	Total	C	N	O	S	0	0	0
			2329	1477	394	447	11			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	408	ALA	-	insertion	UNP P02679
C	409	GLY	-	insertion	UNP P02679
C	410	ASP	-	insertion	UNP P02679
C	411	VAL	-	insertion	UNP P02679
F	408	ALA	-	insertion	UNP P02679
F	409	GLY	-	insertion	UNP P02679
F	410	ASP	-	insertion	UNP P02679
F	411	VAL	-	insertion	UNP P02679

WORLDWIDE PDB
PROTEIN DATA BANK

- Molecule 4 is a protein called GLY-HIS-ARG-PRO-AMIDE peptide ligand.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	I	5	Total 34	C 20	N 10	O 4	0	0	1
4	J	5	Total 34	C 20	N 10	O 4	0	0	1

- Molecule 5 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-a cetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	G	2	Total 28	C 16	N 2	O 10	0	0	0

- Molecule 6 is an oligosaccharide called 2-acetamido-2-deoxy-alpha-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	H	2	Total 28	C 16	N 2	O 10	0	0	0

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

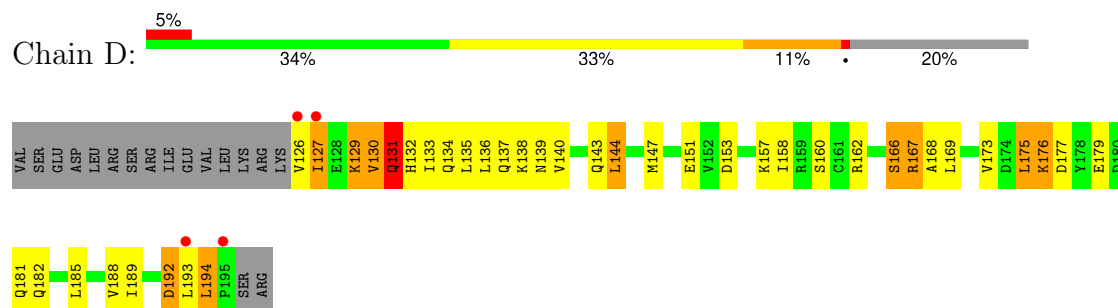| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	2	Total 2	Ca 2	0	0
7	C	1	Total 1	Ca 1	0	0
7	E	2	Total 2	Ca 2	0	0
7	F	1	Total 1	Ca 1	0	0

3 Residue-property plots ⓘ

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). 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.
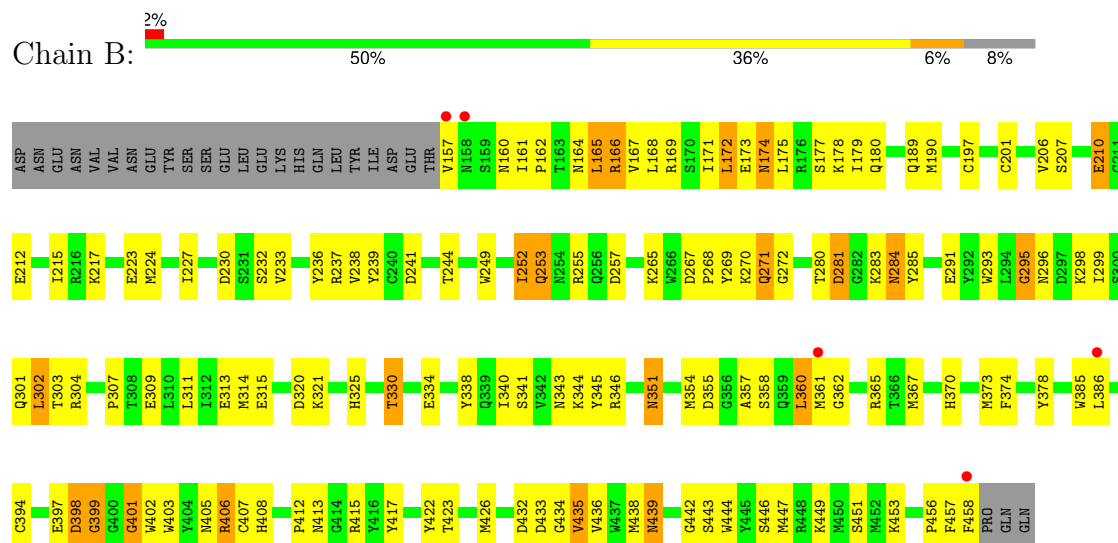
- Molecule 1: Fibrinogen alpha chain

- Molecule 1: Fibrinogen alpha chain
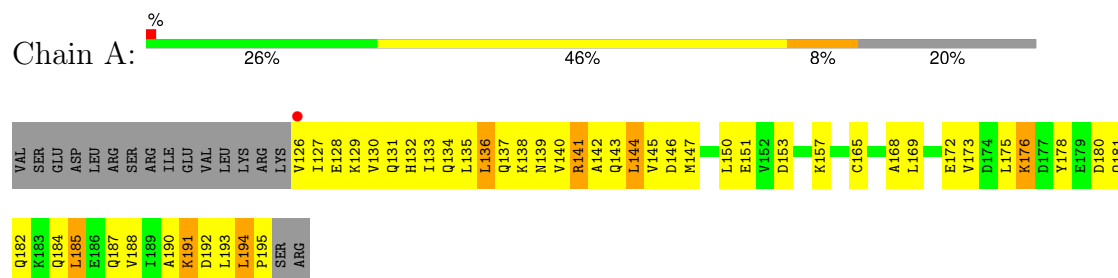
- Molecule 2: Fibrinogen beta chain

• Molecule 2: Fibrinogen beta chain

Chain E:
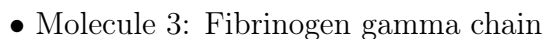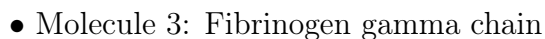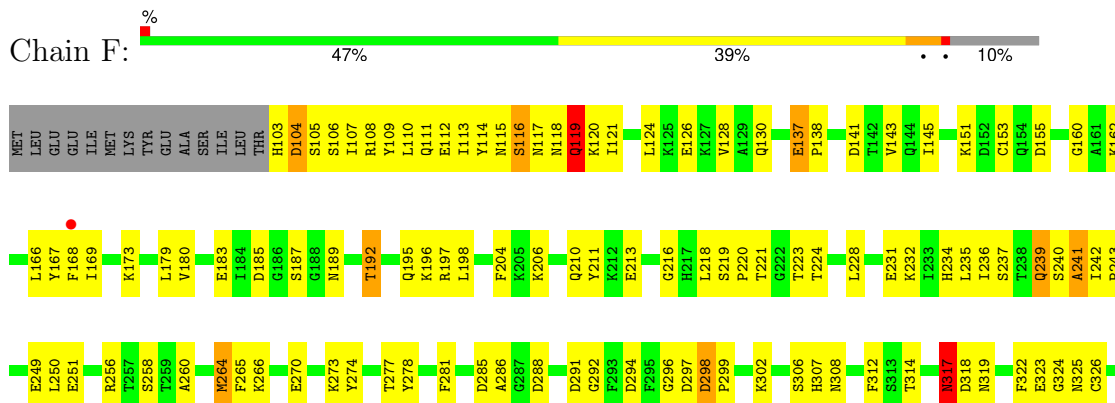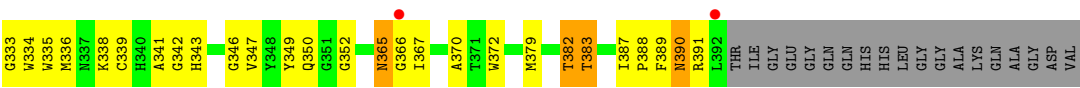
• Molecule 3: Fibrinogen gamma chain

Chain C:

• Molecule 3: Fibrinogen gamma chain

Chain F:

G333 W334 W335 M336 N337 K338 C339 H340 A341 G342 H343 G346 V347 Y348 Y349 Q350 G351 G352 N365 G366 I367 A370 T371 W372 M379 T382 T383 I387 P388 F389 N390 R391 L392 THR ILE GLY GLU GLY GLN GLN HIS HIS LEU GLY ALA LYS GLN ALA GLY ASP VAL

• Molecule 4: GLY-HIS-ARG-PRO-AMIDE peptide ligand

Chain I:

20% 80%

A1 H2 R3 P4 NH25

• Molecule 4: GLY-HIS-ARG-PRO-AMIDE peptide ligand

Chain J:

60% 40%

A1 H2 R3 P4 NH25
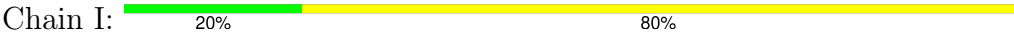
• Molecule 5: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain G:

50% 50%

NAG1 NAG2
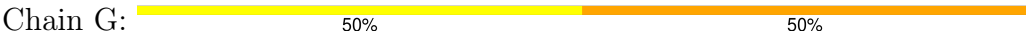
• Molecule 6: 2-acetamido-2-deoxy-alpha-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain H:

50% 50%

NAG1 NDG2
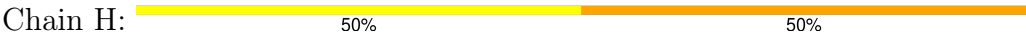
4　Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	105.00Å　　47.90Å　　171.40Å 90.00°　　105.41°　　90.00°	Depositor
Resolution (Å)	30.00　–　2.70 30.00　–　2.70	Depositor EDS
% Data completeness (in resolution range)	(Not available) (30.00-2.70) 91.7 (30.00-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	2.19 (at 2.72Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.251　,　0.307 0.243　,　0.305	Depositor DCC
R_{free} test set	2053 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å^2)	61.2	Xtriage
Anisotropy	0.056	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.25 , 41.5	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.009 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	10770	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality (i)

5.1   Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: CA, NDG, NH2, NAG

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).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#|Z| > 5$	RMSZ	$\#|Z| > 5$
1	A	0.38	0/572	0.66	0/765
1	D	0.39	0/572	0.70	0/765
2	B	0.42	0/2482	0.71	2/3352 (0.1%)
2	E	0.41	0/2482	0.68	2/3352 (0.1%)
3	C	0.41	0/2394	0.62	0/3237
3	F	0.40	0/2394	0.61	0/3237
4	I	0.45	0/34	0.58	0/45
4	J	0.48	0/34	0.59	0/45
All	All	0.41	0/10964	0.66	4/14798 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	401	GLY	N-CA-C	-6.64	96.49	113.10
2	E	399	GLY	N-CA-C	6.50	129.35	113.10
2	B	399	GLY	N-CA-C	6.21	128.64	113.10
2	E	401	GLY	N-CA-C	-5.67	98.93	113.10

There are no chirality outliers.

There are no planarity outliers.

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.

W O R L D W I D E
wwPDB
PROTEIN DATA BANK

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	570	0	602	71	0
1	D	570	0	602	61	0
2	B	2421	0	2288	168	0
2	E	2421	0	2288	168	0
3	C	2329	0	2174	139	0
3	F	2329	0	2174	137	0
4	I	34	0	34	3	0
4	J	34	0	34	5	0
5	G	28	0	25	3	0
6	H	28	0	24	1	0
7	B	2	0	0	0	0
7	C	1	0	0	0	0
7	E	2	0	0	0	0
7	F	1	0	0	0	0
All	All	10770	0	10245	675	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 32.

The worst 5 of 675 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:357:ALA:HB3	2:E:360:LEU:HD21	1.37	1.05
1:D:134:GLN:HE21	1:D:193:LEU:HD13	1.16	1.01
1:D:181:GLN:HE22	2:E:174:ASN:ND2	1.60	1.00
3:C:110:LEU:H	3:C:110:LEU:HD12	1.19	1.00
3:C:153:CYS:HB2	3:C:192:THR:HG22	1.43	0.98

There are no symmetry-related clashes.

5.3　Torsion angles ⓘ

5.3.1　Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	Percentiles	
1	A	68/87 (78%)	52 (76%)	13 (19%)	3 (4%)	2	4
1	D	68/87 (78%)	58 (85%)	4 (6%)	6 (9%)	0	0
2	B	300/328 (92%)	248 (83%)	47 (16%)	5 (2%)	7	20
2	E	300/328 (92%)	249 (83%)	45 (15%)	6 (2%)	6	16
3	C	288/323 (89%)	247 (86%)	31 (11%)	10 (4%)	3	7
3	F	288/323 (89%)	243 (84%)	35 (12%)	10 (4%)	3	7
4	I	3/5 (60%)	3 (100%)	0	0	100	100
4	J	3/5 (60%)	1 (33%)	2 (67%)	0	100	100
All	All	1318/1486 (89%)	1101 (84%)	177 (13%)	40 (3%)	3	9

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	190	ALA
2	B	281	ASP
2	B	295	GLY
1	D	127	ILE
1	D	130	VAL

5.3.2   Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

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	Percentiles	
1	A	65/82 (79%)	58 (89%)	7 (11%)	5	13
1	D	65/82 (79%)	57 (88%)	8 (12%)	4	9
2	B	260/286 (91%)	243 (94%)	17 (6%)	14	34
2	E	260/286 (91%)	244 (94%)	16 (6%)	15	36
3	C	244/269 (91%)	226 (93%)	18 (7%)	11	28
3	F	244/269 (91%)	228 (93%)	16 (7%)	14	33
4	I	3/3 (100%)	3 (100%)	0	100	100
4	J	3/3 (100%)	3 (100%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	1144/1280 (89%)	1062 (93%)	82 (7%)	12	30

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	284	ASN
3	F	183	GLU
2	E	302	LEU
2	E	458	PHE
3	F	317	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 67 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	176	GLN
3	F	189	ASN
3	F	350	GLN
3	C	134	GLN
3	C	123	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates ⓘ

4 monosaccharides are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#|Z| > 2	Counts	RMSZ	#|Z| > 2
5	NAG	G	1	5,2	14,14,15	0.84	0	17,19,21	0.95	0
5	NAG	G	2	5	14,14,15	0.89	1 (7%)	17,19,21	0.74	0
6	NAG	H	1	6,2	14,14,15	0.82	0	17,19,21	1.11	3 (17%)
6	NDG	H	2	6	14,14,15	0.76	1 (7%)	17,19,21	0.71	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	G	1	5,2	-	5/6/23/26	0/1/1/1
5	NAG	G	2	5	-	6/6/23/26	0/1/1/1
6	NAG	H	1	6,2	-	3/6/23/26	0/1/1/1
6	NDG	H	2	6	-	3/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	2	NAG	C1-C2	2.43	1.55	1.52
6	H	2	NDG	C1-C2	2.17	1.55	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	H	1	NAG	C2-N2-C7	-2.57	119.46	122.90
6	H	1	NAG	C4-C3-C2	-2.56	107.26	111.02
6	H	1	NAG	C1-O5-C5	2.04	114.93	112.19

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	G	1	NAG	C1-C2-N2-C7
5	G	1	NAG	C8-C7-N2-C2
5	G	1	NAG	O7-C7-N2-C2
5	G	2	NAG	C8-C7-N2-C2
5	G	2	NAG	O7-C7-N2-C2

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	1	NAG	3	0
6	H	1	NAG	1	0
5	G	2	NAG	3	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

Oligosaccharide Chain G

Bond lengths   Bond angles

Torsions   Rings

Oligosaccharide Chain H

Bond lengths

Bond angles

Torsions

Rings
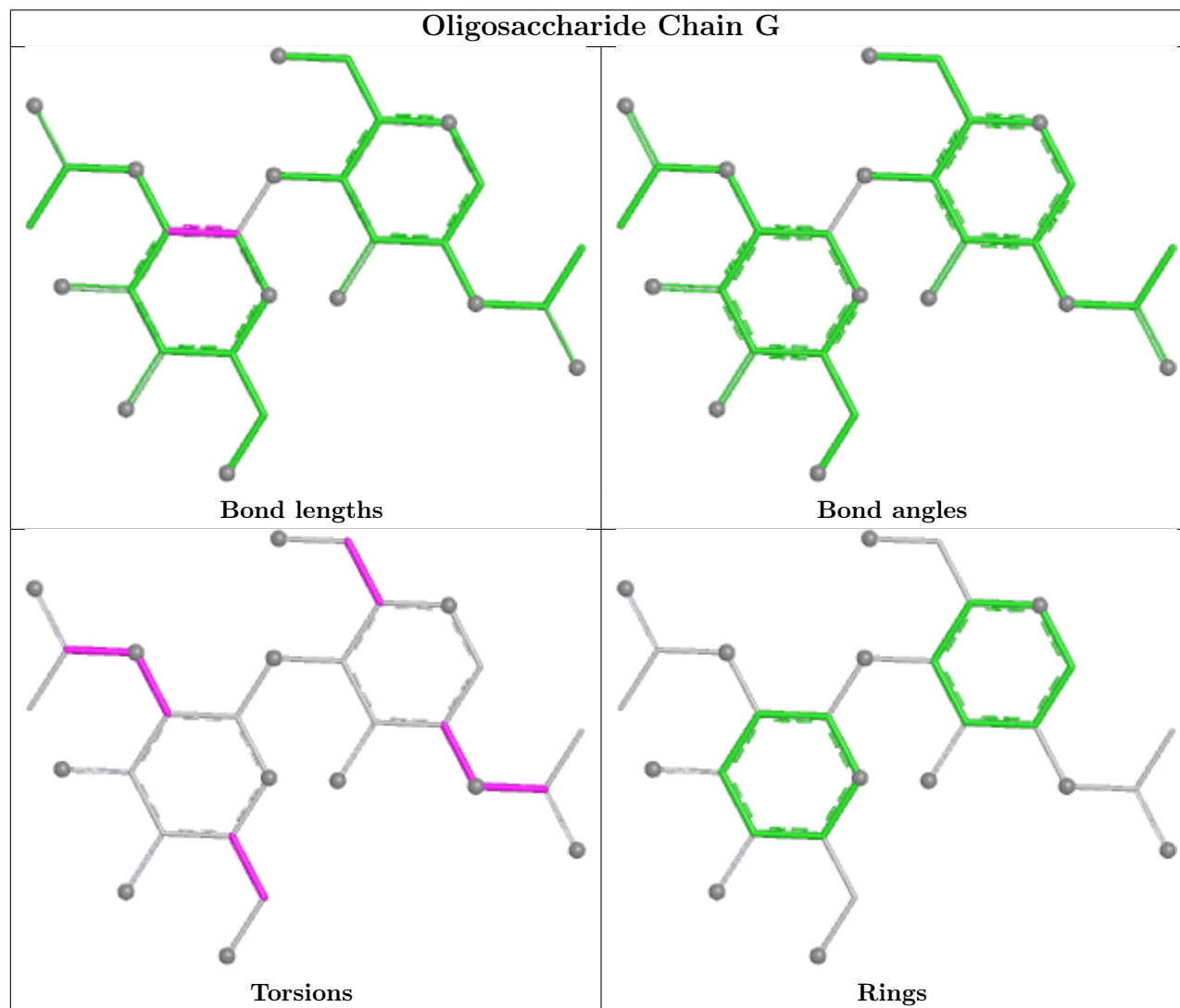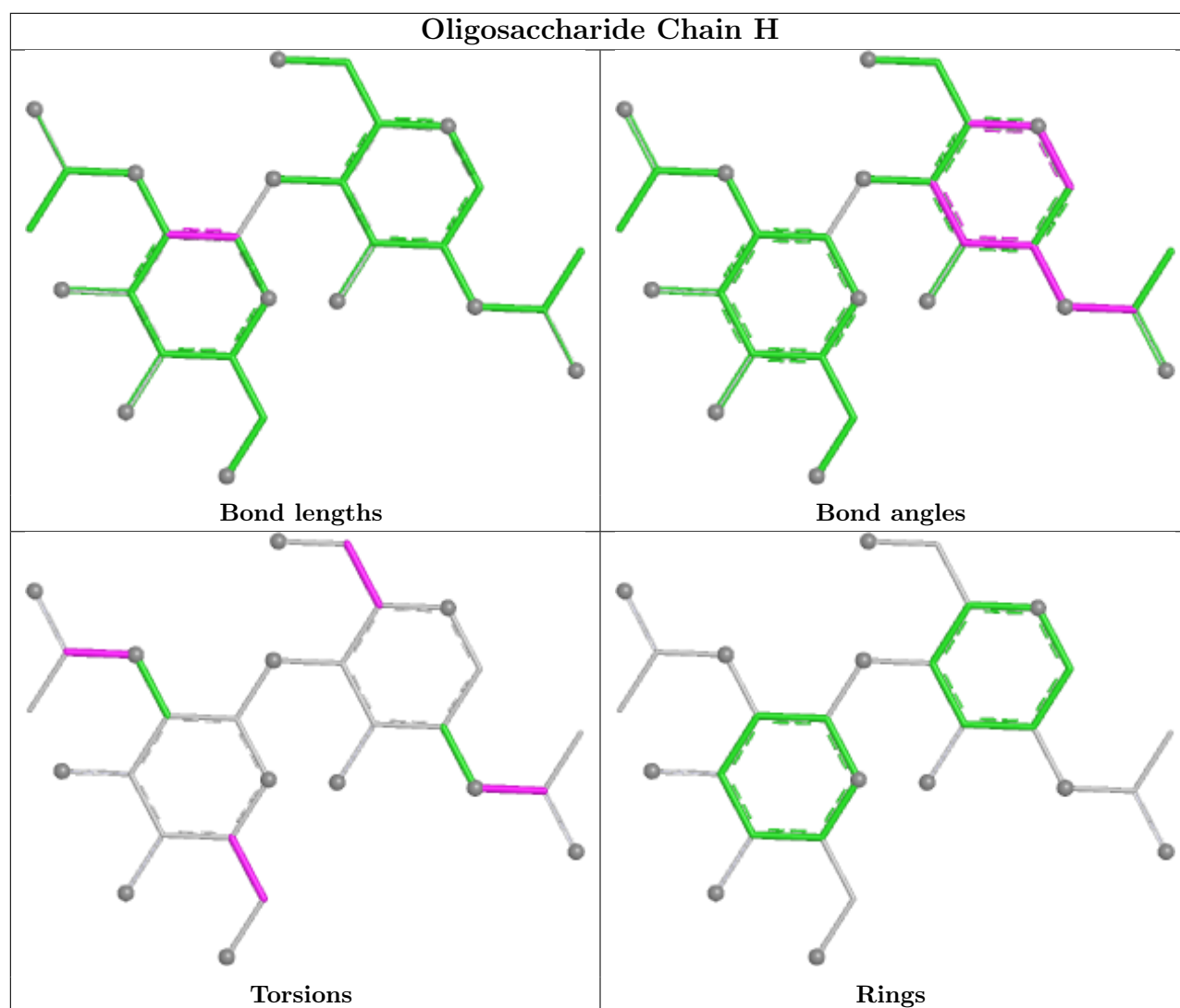
5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	70/87 (80%)	-0.11	1 (1%)	73	73	37, 65, 114, 146	0
1	D	70/87 (80%)	0.08	4 (5%)	30	28	38, 64, 122, 162	0
2	B	302/328 (92%)	-0.16	5 (1%)	69	68	29, 51, 87, 115	0
2	E	302/328 (92%)	0.12	8 (2%)	57	56	30, 61, 91, 111	0
3	C	290/323 (89%)	0.02	4 (1%)	73	73	36, 62, 92, 112	0
3	F	290/323 (89%)	0.28	3 (1%)	79	79	38, 66, 97, 131	0
4	I	4/5 (80%)	0.62	0	100	100	56, 61, 65, 76	0
4	J	4/5 (80%)	-0.06	0	100	100	61, 65, 68, 81	0
All	All	1332/1486 (89%)	0.06	25 (1%)	66	65	29, 61, 95, 162	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	126	VAL	5.6
1	D	127	ILE	5.3
1	A	126	VAL	4.5
2	B	386	LEU	3.8
3	F	392	LEU	3.6

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

6.3 Carbohydrates ⓘ

SUGAR-RSR INFOmissingINFO

6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
7	CA	E	3	1/1	0.68	0.14	111,111,111,111	0
7	CA	B	3	1/1	0.79	0.17	107,107,107,107	0
7	CA	B	2	1/1	0.91	0.15	52,52,52,52	0
7	CA	E	2	1/1	0.94	0.08	58,58,58,58	0
7	CA	F	1	1/1	0.94	0.06	62,62,62,62	0
7	CA	C	1	1/1	0.96	0.09	60,60,60,60	0

6.5 Other polymers ⓘ

There are no such residues in this entry.